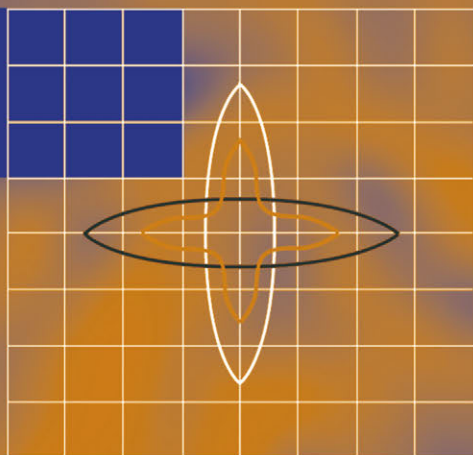


Margaret Armstrong · Alain G. Galli
Gaëlle Le Loc'h · François Geffroy
Rémi Eschard

Plurigaussian Simulations in Geosciences



Springer

Plurigaussian Simulations in Geosciences

Margaret Armstrong • Alain G. Galli
Gaëlle Le Loc'h • François Geffroy • Rémi Eschard

Plurigaussian Simulations in Geosciences

With 77 Figures



Springer

Dr. M. Armstrong
Dr. A.G. Galli
Dr. G. Le Loc'h
Dr. F. Geffroy
Ecole des Mines de Paris
Centre de Géostatistique
35 Rue Saint-Honoré
77305 Fontainebleau
France

Dr. R. Eschard
Institut Français du Pétrole
4 Avenue du Bois Préau
92852 Rueil Malmaison
France

ISBN 978-3-662-12720-9 ISBN 978-3-662-12718-6 (eBook)
DOI 10.1007/978-3-662-12718-6

Cataloging-in-Publication Data applied for

A catalog record for this book is available from Library of Congress.

Bibliographic information published by Die Deutsche Bibliothek.

Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag Berlin Heidelberg GmbH.

Violations are liable for prosecution under the German Copyright Law.

<http://www.springer.de>

© Springer-Verlag Berlin Heidelberg 2003

Originally published by Springer-Verlag Berlin Heidelberg New York in 2003

Softcover reprint of the hardcover 1st edition 2003

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from relevant protective laws and regulations and therefore free for general use.

Production: PRO EDIT GmbH, Heidelberg, Germany

Cover Design: Erich Kirchner, Heidelberg, Germany

Typesetting: Camera-Ready by Authors

Printed on acid-free paper

32/3141/Di

5 4 3 2 1 0

PREFACE

Although many papers have presented different aspects of truncated gaussian and plurigaussian simulations, no single work gives a comprehensive view of the theory and practical applications of this new and interesting geostatistical simulation method. Our aim in writing the book has been to encourage readers to test this method of simulating the geometry of reservoirs and orebodies for themselves. That is why we have included a CDrom containing two small programs called `pluri_demo_simu` and `pluri_demo_vario` to demonstrate how the method works.

Initially this material was developed as the notes for short courses on plurigaussian simulations. The first such course was given in Fontainebleau in September 1998. Others have given at Petrobras training centre in Rio de Janeiro in 1999 and at the IAMG annual conference in Trondheim in 1999.

The original impetus for developing the truncated gaussian method came from Georges Matheron in response to a request from the Institut Français de Pétrole, in particular from Lucien Montadert who realized the importance of being able to simulate the internal architecture of reservoirs. The plurigaussian method was a natural development from the truncated gaussian method. This cooperation between the IFP and the Centre de Géostatistique resulted in the development of the *HERESIM* program as well as practical applications for many oil companies including AGIP, Gaz de France, Petrobras and PDVSA.

Applications of the methodology to mining were initiated during a short course to engineers and geologists of Anglo American Corporation in 1999. Soon afterward a test case was carried out for Rio Tinto Ltd on data from the Rossing uranium mine.

The authors would like to thank these mining and petroleum companies for their support in developing and testing this new simulation method. Special thanks are due to three of our colleagues: Christian Ravenne from the IFP for his geological insights in developing the method, and to Brigitte Doligez also from the IFP and Hélène Beucher from the Centre de Géostatistique who played a key role in implementing the new methodology. This book would never have been published without their help.

TABLE OF CONTENTS

1. Introduction	1
Putting reservoir simulations into context	2
Idea behind truncated and plurigaussian simulations	3
Key steps in a plurigaussian simulation	6
Step 1: Choosing the model type	6
Step 2: Estimating the parameter values	6
Step3: Generating gaussian values at wells/drill-holes	7
Step 4: Simulating values at grid nodes given values at wells	7
Layout of the book	7
 2. Basic properties of indicators	 9
Spatial covariances, variograms and cross-variograms	10
Spatial covariances	10
Variograms and cross-variograms	11
Variogram properties	11
Property 1:	11
Property 2: Two facies	12
Property 3: Covariances for two facies	12
Comments	13
Property 4: Three facies	13
Property 5: Is it possible for two facies to be independent?	15
Property 6: One consequence of the triangle inequality	16
Need for a mathematically consistent method	16
Transition Probabilities	17
First type of transition probability	17
Example of the first type of transition probability	18
Second type of transition probability	20
Example	20
Exercises	22

3. Proportions	27
How to calculate vertical proportion curves	27
Example 1: the Ravenscar sequence	29
Example 2: facies with contrasting anisotropies in a gold deposit	30
Horizontal non-stationarity	31
Choosing the reference level	35
Non-stationarity	37
 4. Thresholds	 43
Basic principle in the truncated gaussian method	43
Defining the thresholds	43
Transitions between facies	43
Link between thresholds and proportions	45
Non-stationary facies	47
Idea behind the plurigaussian method	47
Example 1	47
Example 2	48
Link with the proportions	48
Parameter simplification: use of thresholds	50
Choice of the partition	50
Choice of correlation matrix	52
Calculating thresholds	53
Generalisation to non-stationary case	54
When simulations show “prohibited” transitions	54
 5. Variograms	 57
Truncated gaussian method	57
Experimental variograms and crossvariograms for facies	57
Linking the indicator variograms to the underlying variogram	58
Cross-variograms	59
Truncated plurigaussian	59
Variograms	60
Cross-variograms	62
Generalisation to non-stationary	62
Experimental variograms	62
Remarks	63
Indicator variogram model (truncated gaussian model)	64
Variogram and cross-variogram model (truncated plurigaussian method)	65
Comparing variograms	66
Sill of the indicator model	66

Shape of the indicator model	66
Practical range of the indicator model	67
Anisotropies	68
Variogram fitting	70
Stationary case	70
Non-stationary case	70
Transition probabilities	74

6. Gibbs sampler 77

Part 1: Why we need a two step simulation procedure	77
Simulating $Y(x)$ and $Y(y)$ when there are no constraints	78
Simulating $Y(x)$ and $Y(y)$ when $Y(y)$ belongs to an interval	78
Simulating $Y(x)$ and $Y(y)$ when both belong to intervals	81
Direct simulation using an acceptance/rejection procedure	81
Gibbs sampler	82
Four sample example	82
Exponential variogram	83
Step 1: Initialising the procedure	83
Step 2: The iterative procedure	83
Point No 1	84
Point No 2	84
Point No 3	84
Point No 4	85
Results	85
Alternative updating strategies	86
Blocking factor	87
Experimentally testing convergence	87
Burn-in period	87
Effect of the range on the burn-in period	88
Parallels runs	89
Part 2: Gaussian framework	91
Basic properties of a single gaussian	91
Extension to the plurigaussian case	92
Conditional simulation	93
Implementation of the Gibbs sampler for the model used	94
To summarise	94
Convergence from a theoretical point of view	94
The untruncated case	94
Blocking factor	95
Block size of 1 (order: 1, 2, 3, 4)	96
Block size of 2 (points 1&2 and 3&4)	96
Block size of 2 (points 1&3 and 2&4)	97
Convergence in the truncated gaussian case	97
Markov chains	98
Two integral operators	98

X Table of contents

Stationary distribution	99
Irreducible chain	99
Distance metric and convergence	100
Theorem 1 (Tierney, 1994 and 1996)	100
The Gibbs sampler	102
Truncated gaussian case	103
Definition	104
Theorem 2	104
Comments	105

7. Case studies and practical examples 107

Choosing which simulation method to use	107
Sequence-based pixel methods	107
Object-based methods	108
Nested simulations	109
Building up the reservoir or orebody model	110
Step 1: Defining the lithotypes	110
Step 2: Dividing the reservoir or orebody into units	110
Step 3: Defining the reference level	110
Step 4: Choosing the grid spacing	111
Petroleum applications of the plurigaussian approach	112
Constructing the lithotype rule with geological constraints	112
Simulation of a reservoir with complex facies transitions	112
Simulation of the effects of primary diagenesis in a complex reservoir	115
Simulating progradational patterns	117
Simulation of fractures affecting a specific facies	118
Testing the impact of simulation parameters	119

8. Freeware 121

Introduction	121
Installation	121
Installation instructions for Windows users	121
Installation instructions for Solaris users	122
Installation instructions for Linux users	122
Description of Pluri_Demo_Simu	123
Choosing two of the underlying gaussian images	124
Choosing the rock type rule	125
Exercise A1: Default values for parameters	126
Exercise A2: Inverting the two images	127
Exercise A3: effect of anisotropy	127
Exercise A4: Changing the amount of each facies	128

Exercise A5: Changing the rock type rule – four facies	128
Exercise A6:	128
Exercise A7:	129
Exercise A8:	130
Exercise A9:	130
Variable Proportions	130
Exercise A10: Variable proportions	130
Exercise A11: Variable proportions on two gaussians	131
Exercise A12: working out which rock type rule was used	132
Limitations of the programme	132
Description of Pluri_Demo_Vario	132
Exercise B1: Default values for parameters	133
Exercise B2: Changing the variogram model	134
Exercise B3: Changing the rock type rule	134
Exercise B4: Effect of anisotropy	135
Exercise B5: Adding a nugget effect	136
Exercise B6: Truncated gaussian case	137
Exercise B7: Beatriz's example	138
 9. References	 139
 10. Index	 145

1 INTRODUCTION

This book focuses on two methods for simulating facies and lithotypes: truncated gaussian (Matheron et al, 1987) and plurigaussian simulations (Galli et al, 1994). The first method was developed in the late 80s for simulating the lithotypes found in oil reservoirs because it makes much more sense to simulate the geometry and internal architecture of the reservoir first, then generate suitable values of porosity and permeability once the lithotype is known. These simulations were designed for reservoirs where the lithotypes occur in a sequential order; for example, when sandstone is followed by shaly sandstone then shale. Plurigaussian simulations are a natural extension of these. They were designed to produce a much wider range of patterns and to allow for more complicated types of contacts between facies.

The photo below shows canyons along the San Juan river in Utah where phylloidal algal mounds of Pennsylvanian age are beautifully exposed. This is typical of the sedimentary oil bearing formations that can be modelling using truncated gaussian and plurigaussian methods. This series shows tabular prograding sequences of shallow carbonates and algal mounds. The carbonates can be simulated using the simpler, truncated gaussian method. The upper part of the series consisting of algal mound and intermound facies with the sandstones capping them, is more complicated. The plurigaussian method was used to simulate its structure. The details are given in Chapter 7.



Fig 1.1: Outcrops of a typical carbonate reservoirs along the San Juan River in Utah

Over recent years the mining industry has also started using these two approaches to simulate the facies present in orebodies. For example, the recovery during froth flotation (a mineral processing technique for concentrating sulphide ore) depends on the rock type, rather than the grades. So for both mining deposits and oil reservoirs, the message is clear: first simulate the geometry of the lithotypes or facies then simulate its properties as a function of that type. Before outlining how truncated gaussian and plurigaussian simulations work, we would like to take the time to explain why geostatistical simulations of lithofacies are important in the petroleum industry, since this motivated the development of both methods.

Putting reservoir simulations into context

The aim of reservoir modelling is to construct a gridded model of the reservoir containing its petrophysical properties such as porosity, permeability and capillary pressure, in order to simulate its behaviour during production. As the wells are generally widely spaced and as seismics only provides indirect geological information with a low resolution, the distribution of the geological heterogeneity between wells is uncertain. Geostatistical simulations are a way of quantifying the uncertainty by providing reservoir engineers with representations of the spatial distribution of reservoir heterogeneity. Fluid flow simulations can then be performed on the model to optimise the field's development.

The precise objectives vary depending on the field's stage of development. At the appraisal stage when the aim is to produce a global reservoir model rather than a detailed one, simulations are used to estimate reserves and to quantify their uncertainty. They are also used to define recovery process scenarios. At this stage only a few wells are available and so the seismics play a crucial role.

When the field has started to be developed using primary recovery processes, the aim is to optimise the location of wells. At this stage more wells are available and so the geological model is constructed using the detailed well descriptions together with the seismic interpretation. Detailed reservoir characterisation studies will have been performed, and the petrophysical variables will be available from core and log analysis. Pressure measurements in wells are used to estimate the reservoir connectivity and well test information provides a way of estimating the permeability around the wells. Seismic data can also provide information about the heterogeneity with layers. As we will see later, the degree of detail required in the simulations (i.e. the grid size) depends on the recovery process chosen.

During the final stage of the field's development, the initial reservoir model has to be updated in order to make the recovery process more efficient. A very detailed reservoir model is then needed, for example, to optimise the drilling of wells with a complex geometry. By this time, the well database is much larger; it contains both geological and petrophysical information. Because of its low resolution the seismic information is less important. As the production history of the field is generally available, the observed oil and water flow rates can be compared with those computed from the simulation and used to modify the model iteratively until a good fit is obtained. This process is called history matching.

The complete procedure for reservoir simulation including simulating the lithofacies on a very fine grid and transforming these into reservoir properties, is described by Matheron et al. (1987). The initial geological modelling is a critical step because it ensures the consistency of the simulation from a geological point of view. Another crucial step is upscaling the reservoir properties to a coarse grid to reduce the number of cells because the capacity of most fluid flow simulators is still limited.

Finally, all the information available on the field has to be integrated into the model. Different data have been studied at different scales and have different characteristics. For example, microscopic information such as pore throat types obtained from special core analysis has to be combined with information derived from seismic campaigns having a vertical resolution greater than 20m. Hard data such as well logs based on physical measurements in wells, have to be combined with soft information resulting from geological interpretation of the data. So simulations have to balance hard and soft constraints.

Idea behind truncated and plurigaussian simulations

The basic idea is to start out by simulating one or more gaussian variables (i.e. with a $N(0,1)$ distribution) at every point in the study area and then use the *rock type rule* to convert these values back into lithotypes. Figure 1.2 summarises the procedure when only one gaussian is used. This is called the truncated gaussian approach. The greytone image (left) represents the gaussian value at each point in space. In the image on the right, values below -0.6 have been coloured in dark grey, indicating one facies. Values above 0.5 have been shaded white and intermediate values have been coloured light grey (Figure 1.2b). Three facies have been generated in this way.

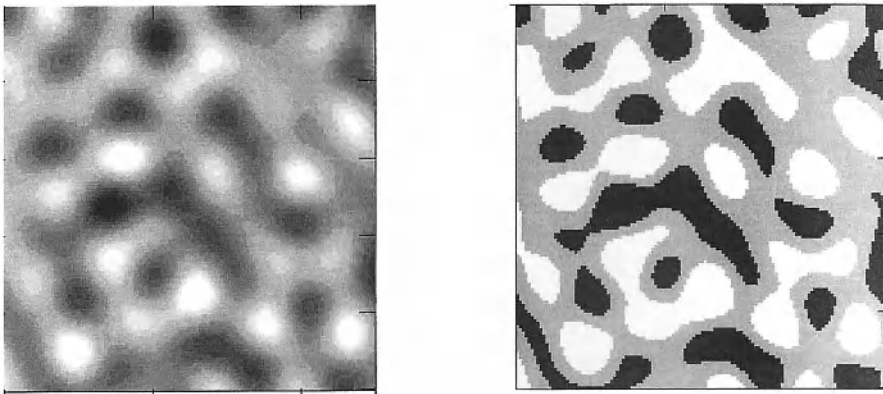


Fig 1.2 (a) Simulated greytone image. Values have a $N(0,1)$ distribution. (b) Same image after being truncated at the cutoffs -0.6 and 0.5 . Values below -0.6 have been shaded dark grey, those between -0.6 and 0.5 are coloured light grey while values above 0.5 are shown in white

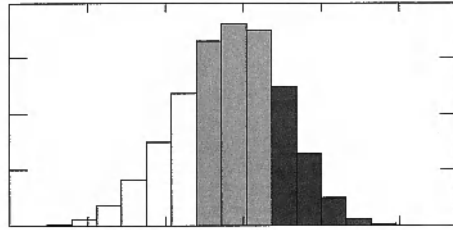


Fig1.3: Histogram of a standard normal distribution (i.e. $N(0,1)$) showing the two cut-offs, -0.6 and 0.5

Looking at Figure 1.2b we see an important feature of the truncated gaussian approach: the grey facies can touch the other two facies, but black and white never touch. Or if they do, it is because the pixel size is too large. If four or five facies were simulated in this way, they would occur in a fixed order, which is defined by the sequence stratigraphy.

Figure 1.3 shows the two thresholds, -0.6 and 0.5, on a standard normal $N(0,1)$ distribution. It is easy to calculate the areas under the three parts of the curve (25%, 45% and 30%) and hence deduce the proportion of space occupied by each facies. In fact there is a 1-1 relationship between the proportions and the thresholds, once the order has been established. In practice we compute the proportions experimentally and use these to deduce the thresholds.

In many cases, the truncated gaussian approach proves to be too restrictive; for example, if there is no natural sequence in the facies or if certain facies can be in contact with more than two facies. So it has been extended to two or more gaussians. Figure 1.4 illustrates this *plurigaussian* procedure for the case of two gaussians. At the top we see two gaussian images (simulations obtained using a gaussian random function). The one on the left has its long range in the NS direction while the other one has its long range in the EW direction. The square shown in the bottom right summarises the *rock type rule*, the rule that is used to assign each point to a facies. Values of the first gaussian, Y_1 , can range from $-\infty$ to $+\infty$. They are plotted along the horizontal axis. Similarly for those of Y_2 along the vertical axis. This square is divided into three regions corresponding to different lithotypes. If $Y_2 < 0$, the rock is coded as dark grey; if $Y_2 > 0$ and $Y_1 < 0$, the rock is classified as being white whereas if $Y_2 < 0$ and $Y_1 > 0$, the rock is light grey. The resulting three facies are shown in the lower left square.

The underlying idea in both truncated gaussian and plurigaussian simulations is to set up one or more simulations of standard normal random functions in the area of interest and to attribute the lithotype or facies depending on the simulated values at each point. This is done by truncating. When only one gaussian is used, the truncation is effectively defined by the values of thresholds. When two or more gaussians are used, the situation is more complex. It is represented graphically via the rock type rule.

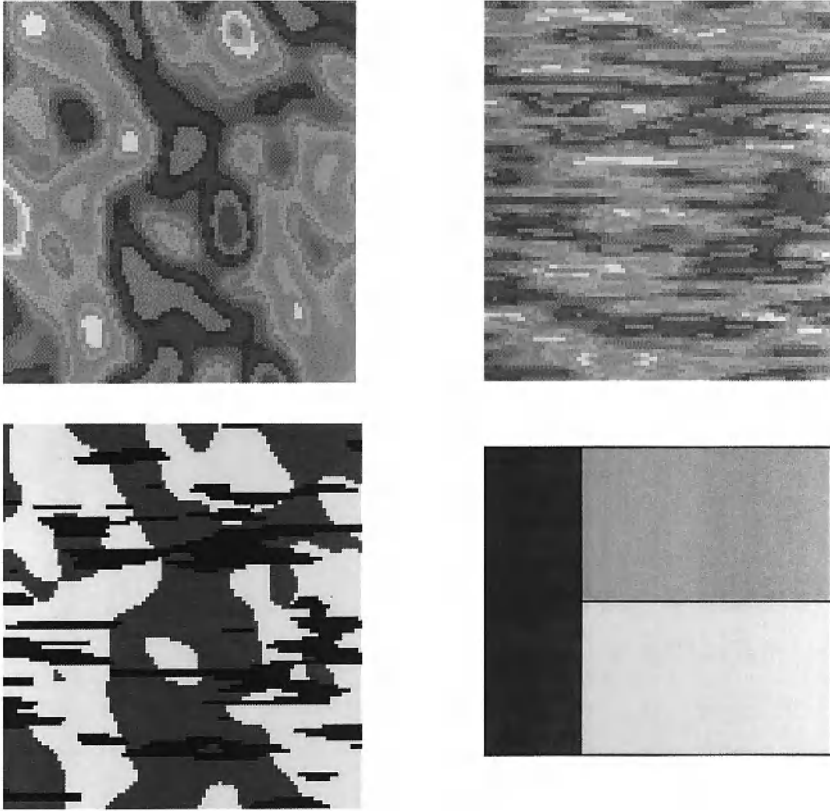


Fig 1.4: How plurigaussian simulations work. The two images at the top are realisations of gaussian (i.e. $N(0,1)$) random functions with anisotropies in the NS and EW directions. The rock-type rule shown bottom right has been used to truncate these two images to obtain the facies or lithotypes shown bottom left. Note the contrast between the anisotropies in the simulated facies

In this book we will almost always use rectangles to divide the rock type rule into lithofacies. We do this because it makes it easier to work out the thresholds from the experimental proportions.

In most plurigaussian applications, the two gaussian random functions are independent of each other but it is also possible to use correlated gaussian RFs. When this is done the facies tend to "wrap around" each other. This makes it more difficult to compute the thresholds from the experimental proportions.

As well as being able to use independent or correlated RFs, more advanced models can be obtained by using derivatives or translated RFs:

- $$Y_2(x) = \frac{\partial}{\partial x_i} Y_1(x)$$

$$\blacksquare \quad Y_2(x) = Y_1(x + b)$$

where x_i represents one of the directions and b is the translation vector. Even more complicated models can be obtained by combining these. The paper by Armstrong and Galli (1999) shows how to generate crevasse splays in anastomosed channels using derivative RFs.

Key steps in a plurigaussian simulation

Having explained the broad principles behind plurigaussian simulations we need to go into more detail. Here we assume that readers are familiar with basic geostatistics (variograms and kriging) and with simulating gaussian random functions. The four main steps in a plurigaussian simulation are choosing the appropriate type of model, estimating the values of its parameters, generating gaussian values corresponding to the lithotypes at sample points and lastly running the conditional simulation, using the gaussian values generated in the previous step.

Step 1: Choosing the model type

Plurigaussian simulations can be divided into several broad families, depending on the types of relations between the facies or lithotypes. For example, in some cases there is a natural sequential order among the facies. In fluvial channel reservoirs, the lithotypes - sandstone, shaly sandstone and shale - generally occur in that order because of depositional conditions. By looking through the catalogue, the reader can choose the type of model that seems best suited to his/her data. When there is a clear sequential order in the lithotypes, a single gaussian usually suffices; otherwise two or more gaussians can be used. In some cases, one of the facies may appear to be a « shifted » version of another facies. Increments or derivatives of gaussians can be used to obtain special effects.

Step 2: Estimating the parameter values

Two key factors control plurigaussian simulations: the thresholds at which the different gaussians are truncated and the variogram model of the underlying gaussian variable. The proportion of each facies, the “rock type” rule and the correlation between the two underlying gaussian random functions, determines the thresholds. Once the thresholds have been worked out, the variograms and cross-variograms of the indicators are calculated experimentally. Knowing the mathematical relation between the indicator variograms and the variograms of the underlying gaussian variable(s), we can find a suitable model for the underlying variogram and estimate the values of its parameters.

Step 3: Generating gaussian values at wells/drill-holes

The samples available at wells (or drill-holes) give us the facies or lithotype for each sample but this does not tell us the corresponding gaussian values. We merely know that they must fall in certain domains which reduce to intervals when a rectangular partition is used. The third step in the procedure consists of generating gaussian values in the appropriate intervals and having the right properties (e.g. variogram model). Obviously these values are not unique. A special statistical method called a Gibbs sampler is used to generate these values.

Step 4: Simulating values at grid nodes given values at wells

Once the gaussian values corresponding to the facies/lithotypes have been generated at wells or drill-holes, the rest of the simulation procedure is quite straightforward. Any algorithm can be used for conditionally simulating the gaussian values. Once the gaussian values at all the grid nodes have been obtained, the last step is to convert these back to facies using the rock type rule.

Layout of the book

This guidebook to plurigaussian simulations is divided into 7 chapters plus an appendix containing a visual catalogue of simulated images. As we will see, indicator variables are deceptively simple. People have the mistaken impression that they can be treated like ordinary variables - like grades, or porosity, or the depth to a horizon. So Chapter 2 reviews their properties and those of indicator variograms. Chapter 3 is on proportions. It shows how to calculate vertical proportion curves which summarise the vertical variability in the lithotypes found in oil reservoirs, and generalises this for the case where there is horizontal non-stationarity as well. Once the proportions have been modelled, the thresholds separating the facies can be calculated. Chapter 4 presents this step. The next step (Chapter 5) is to calculate the experimental indicator variograms and fit models to the variograms of the underlying gaussian variables.

Before introducing the Gibbs sampler, an important theoretical question has to be addressed. We have to show that simulating the gaussian values at target points knowing which facies each sample belongs to is equivalent to simulating those values at target points given the gaussian values generated by the Gibbs sampler at data points. Chapter 6 gives this result and then goes on to explain the Gibbs sampler which is used to generate gaussian values corresponding to the lithotypes/facies at the sample points. Several applications of truncated gaussian and plurigaussian simulations applied to the oil industry are presented in Chapter 7.

Colour illustrations are vital to understanding plurigaussian simulations but they greatly increase the printing costs. You will find a CD in the pocket in the back cover. It contains colour copies of many of the figures plus two

demonstration programs. The first one called "pluri-demo" was designed to demonstrate how plurigaussian simulations function and to allow readers to test the impact of key parameters for themselves. The second program is called "pluri-vario-demo". It shows the relationship between the variogram models of the two stationary random functions (i.e. the two underlying gaussians) and those of the indicators. The appendix at the end of the book contains instructions on how to run the demonstration software together with exercises that the reader can carry out using the programs.

2 BASIC PROPERTIES OF INDICATORS

The aim of this section is to review some of the basic properties of indicators that will be needed later. We let F be a random set (i.e. all the points belonging to one particular facies). Let \bar{F} be its complement (i.e. the points that do not belong to F). The indicator function for the set F takes the value 1 at all the points inside F ; it takes the value 0 elsewhere. This indicator function is denoted by $1_F(x)$. In Figure 2.1, the set F has been shaded. Its complement \bar{F} includes all of the non-shaded area. The set F could be any shape or form, and need not be a single piece. It could be split into several parts.

The first property of indicator functions relies on the fact that inside F its indicator takes the value 1 whereas the indicator of its complement takes the value zero. Outside F the converse is true. That is, the union of F and its complement fills the whole space. Consequently,

$$1 = 1_{F \cup \bar{F}}(x) = 1_F(x) + 1_{\bar{F}}(x)$$

One property of a geological facies that can be measured experimentally is the proportion of space that it occupies. If we let $P_F(x)$ be the probability that point x lies in F , then it is equal to the mean or mathematical expectation of $1_F(x)$. That is,

$$P_F(x) = E[1_F(x)]$$

So $1_F(x)$ is a random function. As it takes the values 0 or 1, its expected value must lie between these two values.

$$0 \leq P_F(x) \leq 1$$

Moreover, because of the linearity of taking expectations,

$$P_F(x) + P_{\bar{F}}(x) = 1$$

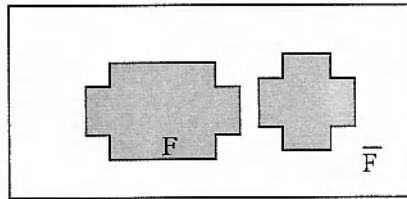


Fig 2.1: Set F (shaded grey) and its complement \bar{F} in white

Now we calculate its variance. Since $[1_F(x)]^2 = [1_F(x)]$,

$$\text{Var}[1_F(x)] = E[1_F(x)] - \{E[1_F(x)]\}^2 = P_F(x)(1 - P_F(x)) \leq 0.25$$

Spatial covariances, variograms and cross-variograms

In this section we focus on the spatial relation between indicators by studying their centred and non-centred covariances, then their variograms and cross-variograms. After defining these concepts we derive their key properties. The first step is to study the relationship between the facies at point x and the one at y , firstly via the (non-centred) spatial covariance $C_F(x, y)$ and then via the variogram $\gamma_F(x, y)$.

Spatial covariances

Let x and $x + h$ be any two points. Their non-centred covariance is defined as:

$$\begin{aligned} C_F(x, x + h) &= E[1_F(x) 1_F(x + h)] = P[1_F(x) = 1 \ \& \ 1_F(x + h) = 1] \\ &= P[x \in F \ \& \ x + h \in F] \end{aligned}$$

It measures the probability that both points lie inside F . Similarly the centred covariance $\sigma_F(x, x + h)$ and the centered cross covariance between facies F_i and F_j , denoted by $\sigma_{F_i F_j}(x, x + h)$ are respectively defined as

$$\begin{aligned} \sigma_F(x, x + h) &= E\{[1_F(x) - P_F(x)][1_F(x + h) - P_F(x + h)]\} \\ &= C_F(x, x + h) - P_F(x)P_F(x + h) \end{aligned}$$

$$\sigma_{F_i F_j}(x, x + h) = E\{[1_{F_i}(x) - P_{F_i}(x)][1_{F_j}(x + h) - P_{F_j}(x + h)]\}$$

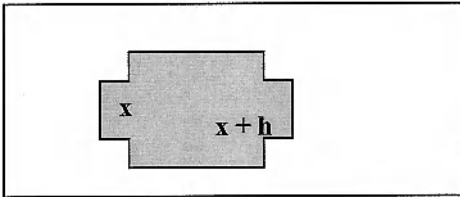


Fig 2.2: Two points, x and $x + h$, lying in facies F

Variograms and cross-variograms

Let $A(x)$ be an arbitrary variable. It could be an ordinary variable $Z(x)$ or an indicator $1_F(x)$. Its variogram is defined as:

$$\begin{aligned}\gamma_A(x, x+h) &= \frac{1}{2} \text{Var} [A(x) - A(x+h)] \\ &= \frac{1}{2} \left\{ E \left([A(x) - A(x+h)]^2 \right) - \left(E[A(x) - A(x+h)] \right)^2 \right\}\end{aligned}$$

Consequently the indicator variogram of the facies F is just

$$\begin{aligned}\gamma_F(x, x+h) &= \frac{1}{2} \text{Var} [1_F(x) - 1_F(x+h)] \\ &= \frac{1}{2} \left\{ E \left([1_F(x) - 1_F(x+h)]^2 \right) - \left(E[1_F(x) - 1_F(x+h)] \right)^2 \right\}\end{aligned}$$

Note that if the indicator is stationary, the second term disappears.

The cross variogram between facies F_i and F_j is defined as:

$$\gamma_{F_i F_j}(x, x+h) = \frac{1}{2} E \left\{ [1_{F_i}(x) - 1_{F_i}(x+h)] [1_{F_j}(x) - 1_{F_j}(x+h)] \right\}$$

When this product is expanded, two of the four terms disappear because 1_{F_i} and 1_{F_j} cannot both take the value 1 at the same point since we have one and only one facies at each point. This gives

$$\gamma_{F_i F_j}(x, x+h) = -\frac{1}{2} \left\{ E[1_{F_i}(x) 1_{F_j}(x+h)] + E[1_{F_j}(x) 1_{F_i}(x+h)] \right\}$$

Variogram properties

Property 1

As the indicators can only take the values 0 or 1, the values of the variogram must satisfy the inequality

$$\gamma_F(x, x+h) = \frac{1}{2} \text{Var}[1_F(x) - 1_F(x+h)] \leq 0.5$$

One consequence of this is that indicator variograms must be bounded. *So a power function could never be an appropriate model for an indicator variogram.*

Property 2: Variograms for two facies (stationary case)

$$\gamma_F(h) = \gamma_{\bar{F}}(h)$$

Proof. By definition the variograms of F and \bar{F} are

$$\gamma_F(h) = \frac{1}{2} E[1_F(x+h) - 1_F(x)]^2 \quad \text{and} \quad \gamma_{\bar{F}}(h) = \frac{1}{2} E[1_{\bar{F}}(x+h) - 1_{\bar{F}}(x)]^2$$

The first result follows because

$$1_F(x+h) - 1_F(x) = -(1_{\bar{F}}(x+h) - 1_{\bar{F}}(x))$$

Property 3: Covariances for two facies

If F is a stationary random set, its variogram depends on the vector between the two points, but not on their positions, and similarly for the covariances. The next property follows from this.

$$\sigma_F(h) = \sigma_{\bar{F}}(h) = -\sigma_{F\bar{F}}(h) = -\sigma_{\bar{F}F}(h)$$

Proof. By definition, we have

$$\sigma_F(h) = E[1_F(x+h) - P_F(x+h)][1_F(x) - P_F(x)]$$

Because of the stationarity, $P_F(x+h) = P_F(x) = P_F$ and similarly for its complement. So

$$\sigma_F(h) = E[1_F(x+h) - P_F][1_F(x) - P_F]$$

$$\sigma_{\bar{F}}(h) = E[1_{\bar{F}}(x+h) - P_{\bar{F}}][1_{\bar{F}}(x) - P_{\bar{F}}]$$

$$\sigma_{F\bar{F}}(h) = E[1_F(x) - P_F][1_{\bar{F}}(x+h) - P_{\bar{F}}]$$

$$\sigma_{\bar{F}F}(h) = E[1_{\bar{F}}(x) - P_{\bar{F}}][1_F(x+h) - P_F]$$

Moreover

$$1_F(x+h) - P_F = 1 - 1_{\bar{F}}(x+h) - P_F = -[1_{\bar{F}}(x+h) - P_{\bar{F}}]$$

$$1_F(x) - P_F = -[1_{\bar{F}}(x) - P_{\bar{F}}]$$

Substituting these relations into the covariance formulas gives the second result.

Comments

There are three practical consequences of these results. Firstly two complementary random sets are forcibly correlated. They cannot be independent of each other. Secondly their cross covariance is the same as the direct covariance up to a change of sign. Thirdly, the indicator covariance is not a connectivity index. A connectivity index tells us whether it is possible to find a path joining any two points in the facies and lying completely inside that facies, whereas the indicator covariance merely tells us whether two end points are in the facies. It says nothing about paths joining the two points.

A facies F and its complement rarely have the same connectivity and yet their indicator covariances are identical. Figure 2.3 illustrates this point. The background is divided into two separate parts but the shaded facies is a single piece.

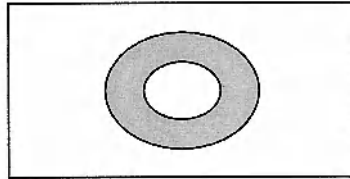


Fig 2.3: The facies F and its complement do not have the connectivity but their indicator covariances are identical

Property 4: Three facies

Instead of two facies we now consider three facies labelled A , B and C , which together fill the whole space under study. See for example Figure 2.4. At present we assume that the three random sets are stationary.

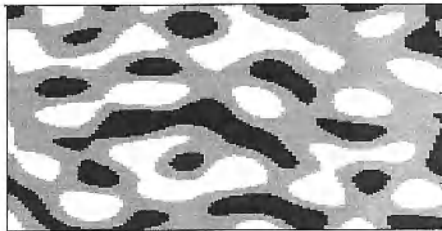


Fig 2.4: Three facies A , B and C shown in white, light grey and dark grey

14 Basic Properties

For these three facies

$$\sigma_A(h) = -\sigma_{AB}(h) - \sigma_{AC}(h)$$

$$\sigma_B(h) = -\sigma_{BA}(h) - \sigma_{BC}(h)$$

$$\sigma_C(h) = -\sigma_{CA}(h) - \sigma_{CB}(h)$$

$$\sigma_{AB}(h) - \sigma_{BA}(h) = \sigma_{BC}(h) - \sigma_{CB}(h) = \sigma_{CA}(h) - \sigma_{AC}(h)$$

Proof. To prove this result, we note that any point x lies in one and only one of the three sets; that is

$$I_A(x) + I_B(x) + I_C(x) = 1$$

Consequently

$$P_A + P_B + P_C = 1$$

Hence

$$I_A(x) - P_A = -[I_B(x) - P_B + I_C(x) - P_C]$$

We now substitute this into the second term of

$$\sigma_A(h) = E[I_A(x+h) - P_A][I_A(x) - P_A]$$

This gives

$$\begin{aligned}\sigma_A(h) &= -E[I_A(x+h) - P_A][I_B(x) - P_B + I_C(x) - P_C] \\ &= -\sigma_{AB}(h) - \sigma_{AC}(h)\end{aligned}$$

This proves the first three parts of property #3. Now to get the last part, we start from the definitions :

$$\sigma_{AB}(h) = E[I_A(x) - P_A][I_B(x+h) - P_B]$$

$$\sigma_{BA}(h) = E[I_B(x) - P_B][I_A(x+h) - P_A]$$

As before, we substitute the relation linking I_A to $I_B + I_C$ into these two, giving

$$\sigma_{AB}(h) = -E[I_B(x) - P_B + I_C(x) - P_C][I_B(x+h) - P_B]$$

$$= -\sigma_B(h) - \sigma_{CB}(h)$$

$$\sigma_{BA}(h) = -E[1_B(x) - P_B] [1_B(x+h) - P_B + 1_C(x+h) - P_C]$$

$$= -\sigma_B(h) - \sigma_{BC}(h)$$

Subtracting gives the result :

$$\sigma_{AB}(h) - \sigma_{BA}(h) = \sigma_{BC}(h) - \sigma_{CB}(h)$$

Property 5: Is it possible for two facies to be independent?

As A, B and C are a partition of the space, they cannot all be uncorrelated. But we can ask whether it is possible for two facies, A and B (say), to be uncorrelated. We are now going to show that this is not possible. If there were more than three facies, we could simply regroup all the other facies (except A and B) into a single facies C. So it suffices to prove the result for three arbitrary facies, A, B and C. Suppose that facies A and B are uncorrelated. Then

$$\sigma_{AB}(h) = E[1_A(x) - P_A] E[1_B(x+h) - P_B] = 0$$

Similarly for $\sigma_{BA}(h)$. So

$$0 = \sigma_{AB}(h) - \sigma_{BA}(h)$$

hence

$$\sigma_{BC}(h) = \sigma_{CB}(h) \text{ and } \sigma_{CA}(h) = \sigma_{AC}(h)$$

Consequently

$$\sigma_A(h) = 0 - \sigma_{AC}(h)$$

$$\sigma_B(h) = 0 - \sigma_{BC}(h)$$

$$\sigma_C(h) = -\sigma_{CA}(h) - \sigma_{CB}(h)$$

From this it is easy to show that

$$\sigma_C(h) = \sigma_A(h) + \sigma_B(h)$$

The next step is to substitute $h = 0$ into this equation. Since $\sigma(0)$ is equal to the variance, we get

16 Basic Properties

$$P_C(1 - P_C) = P_A(1 - P_A) + P_B(1 - P_B)$$

$$\text{and because } P_C = 1 - (P_A - P_B)$$

$$P_A P_B = 0$$

This means that one of the two facies A or B is not present, which contradicts the initial hypothesis. This proves that facies cannot be independent.

Property 6: One consequence of the triangle inequality

Matheron (1987) proved that all indicator variograms must satisfy the triangle inequality for any two distances, h_1 and h_2 :

$$\gamma(h_1 + h_2) \leq \gamma(h_1) + \gamma(h_2)$$

This tells us about the behaviour of indicator variograms near the origin. Suppose that the behaviour was a power function near $h = 0$; say $\gamma(h) \approx |h|^\alpha$. What values are possible for the power α ? Let $h_1 = h = h_2$. Then

$$\gamma(h_1 + h_2) \approx |2h|^\alpha \text{ and } \gamma(h_1) + \gamma(h_2) \approx 2|h|^\alpha$$

Consequently α must be less than or equal to 1. *This means that the gaussian variogram should not be used as a model for an indicator variogram.* Two examples of the types of inconsistencies that can arise when unsuitable variogram models are used for indicators, are given in exercises 2.3 and 2.4 at the end of the chapter. The point in including these examples is to show just how difficult it is to give a general characterisation for the variogram of an indicator random function, or in other words, of a random set. Matheron (1987, 1989 and 1993) produced several papers on this subject. The third one gives results concerning groups of 3, 4 and 5 points and then makes a conjecture to generalise these. These results are relevant in "multi-point" geostatistics.

Need for a mathematically consistent method

In the previous section we saw that although indicators are very simple themselves, their properties are far from intuitive. For example, most people believe that it would be possible for some of the facies or lithotypes to be independent of each other, until they have seen the proof to the contrary. Nor do they realise that power functions and the gaussian model are not suitable models for indicator variograms. Although we have proved that these two models are not acceptable, this does not mean that the others can be used. It merely means that no counter examples have yet found to them. Perhaps someone will one tomorrow.

This raises a troubling question: which variogram models can be used for indicators. Or putting it more generally, Bochner's theorem provides a general characterisation for positive definite functions (i.e. for the covariances of gaussian random functions, and hence their variograms). Is there an equivalent theorem for random sets and indicator variograms? To the best of our knowledge there is not.

So we are faced with two options: make ad hoc choices about the indicator variograms with the risk of ending up with mathematical inconsistencies, or construct a mathematically consistent model. We prefer the second option. The way that we have chosen to do this is, as we saw in Chapter 1, by truncating simulations of gaussian random functions in a suitable way. The advantages of using gaussian random functions are clear: firstly, the normal (gaussian) distribution has many nice statistical properties; secondly, there are many ways of simulating gaussian random functions. The good properties, from a theoretical point of view, of the normal distribution and gaussian random functions include

- In simple kriging the error is orthogonal to the estimate, so it is independent of it. This allows us to condition simulations by kriging.
- Any positive definite function can be used as the covariance model of a gaussian random function. This is not true for other distributions, either discrete or continuous. We have already seen this for indicators which are discrete variables. In a similar vein, Armstrong (1992) shows that the spherical variogram is not always compatible with a lognormal distribution.

Transition probabilities

From experimental data, it is easy to compute the proportion of samples belonging to a particular facies. Although it is a little more complex, we can also calculate the conditional probability of going from one facies to another, or of staying in the same facies for points a certain distance apart. We are now going to express these transition probabilities in terms of the indicator covariance and the proportions.

First type of transition probability

The first type of transition probability is just the probability of being in facies F_j at point $x + h$ knowing that point x is in facies F_i :

$$P(x + h \in F_j \mid x \in F_i) = P[1_{F_j}(x + h) = 1 \mid 1_{F_i}(x) = 1]$$

$$= \frac{P[1_{F_j}(x + h) = 1 \ \& \ 1_{F_i}(x) = 1]}{P[1_{F_i}(x) = 1]}$$

$$\text{So } P(x+h \in F_j \mid x \in F_i) = \frac{E[1_{F_j}(x+h) 1_{F_i}(x)]}{E[1_{F_i}(x)]}$$

$$= \frac{C_{ij}(x, x+h)}{p_{F_i}(x)}$$

where $C_{ij}(x, x+h)$ is the noncentred crosscovariance between $1_{F_i}(x)$ and $1_{F_j}(x+h)$, and $p_{F_i}(x)$ is the proportion of facies F_i at point x .

This is the probability of going from facies F_i to facies F_j , not being sure that we have a transition between x and $x+h$. This formula is also correct if $j = i$. If the facies are stationary, then this probability contains essentially the same information as the indicator variogram.

Example of the first type of transition probability

To illustrate how to calculate this type of transition probability, we have selected four neighbouring drill-holes extending from one chronostratigraphic marker (N°11) down to the following one (N°10). Only two facies are present, A and B. Observations of the facies were made at 2m intervals down the holes (Table 1).

Table1: Facies observed at 2 m spacing going from marker N°11 down to marker N°10

Drill-hole N° 92	Drill-hole N°58	Drill-hole N°39	Drill-hole N° 84
A	-	-	-
B	-	-	-
A	A	-	-
B	A	-	-
A	A	B	-
A	A	A	-
A	A	B	-
B	A	B	-
B	A	A	-
B	B	A	-
B	A	A	-
B	A	B	B
B	A	B	B
B	B	B	A
B	B	B	A
B	B	B	B
B	A	B	A
B	A	A	B
B	B	B	A
A	A	B	A
B	A	B	B
A	B	B	B
B	B	B	B

As the drill-hole sections are of different lengths and as Marker N°10 is considered as the reference level (see chapter 3 for more information on reference levels), the sections are given from this level upward rather than from marker N°11 downward, with blanks at the top. Our objective is to calculate the transition probabilities. By definition, these depend on the two points x and $x+h$. If the data are stationary horizontally (which is often the case), we can average along the rows. Similarly, if they are stationary vertically we can average in that direction but vertical stationarity is rare in sedimentary rock-types. Here as we only have four drill-holes in this example, we will assume vertical and horizontal stationarity for simplicity.

Table 2a shows the numbers of each type of transition (i.e. $A \rightarrow A$, $A \rightarrow B$, $B \rightarrow A$ and $B \rightarrow B$) for a distance of 2m upward while Table 2b shows the number of transitions that would be expected, assuming a random distribution of facies. As 33 of the 75 observations are facies A, the $\text{Pr}(A) = 0.44$ and $\text{Pr}(B) = 0.56$. Consequently since there are 71 transitions each 2m long, the expected number of transitions $A \rightarrow B$ or $B \rightarrow A$ is just $0.44 \times 0.56 \times 71$. In fact, the observed number of transitions $A \rightarrow A$ and $B \rightarrow B$ is higher than would be expected if the facies were arranged randomly, which was expected because of the sequential structure of sedimentary facies at this scale. Similarly, Table 3a shows the numbers of each type of transition for a distance of 10m upward while Table 3b shows the corresponding number of transitions that would be expected, assuming independence.

Table 2(a) Observed number of transitions for a distance of 2m upward (out of a total of 71), (b) expected number of transitions assuming independence.

	Up to A	Up to B
From A	16	13
From B	15	27

(a)

	Up to A	Up to B
From A	13.7	17.5
From B	17.5	22.3

(b)

Table 3(a) Observed number of transitions for a distance of 10m upward (out of a total of 55), (b) expected number of transitions assuming independence.

	Up to A	Up to B
From A	11	8
From B	14	22

(a)

	Up to A	Up to B
From A	10.7	13.5
From B	13.5	17.3

(b)

One difference between the observed values in Tables 2 and 3 is that Table 2 is almost symmetric around the diagonal but Table 3 is definitely not. The lower line representing sections starting with facies B contains 36 out of 55 observations (i.e. 65% rather than 56%). This indicates that facies B tends to be lower in the stratigraphic sequence than facies A; that is, this is evidence of non-stationarity. So averaging along the vertical direction is not a valid operation in this case. This problem will be encountered again when calculating variograms.

Second type of transition probability

Here we consider the case where there are three or more facies. We are interested in the probability of being in facies F_j at point $x+h$ knowing that we are in facies F_i at point x but not at $x+h$. This probability is:

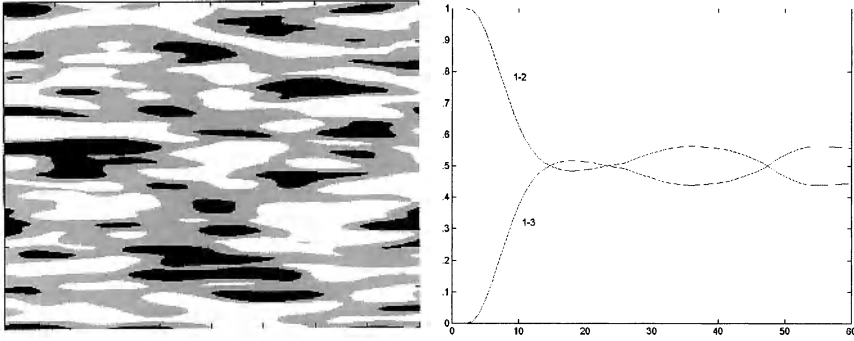
$$\begin{aligned}
 & P(x+h \in F_j \mid x \in F_i \ \& \ x+h \notin F_i) \\
 &= P[1_{F_j}(x+h) = 1 \mid 1_{F_i}(x) = 1 \ \& \ 1_{F_i}(x+h) = 0] \\
 &= \frac{P[1_{F_j}(x+h) = 1 \ \& \ 1_{F_i}(x) = 1 \ \& \ 1_{F_i}(x+h) = 0]}{P[1_{F_i}(x) = 1 \ \& \ 1_{F_i}(x+h) = 0]} \\
 &= \frac{E[1_{F_j}(x+h) \cdot 1_{F_i}(x) [1 - 1_{F_i}(x+h)]]}{E\{1_{F_i}(x) [1 - 1_{F_i}(x+h)]\}} \\
 &= \frac{E[1_{F_j}(x+h) \cdot 1_{F_i}(x)]}{E\{1_{F_i}(x) [1 - 1_{F_i}(x+h)]\}} \\
 &= \frac{C_{ij}(x, x+h)}{P_{F_i}(x) - C_{ii}(x, x+h)}
 \end{aligned}$$

where $C_{ii}(x, x+h)$ is the non centred covariance between $1_{F_i}(x)$ and $1_{F_i}(x+h)$. This is the probability of going from facies F_i to facies F_j , when we know that the facies at $x+h$ is not the same as at x . Clearly this probability cannot be computed if $j = i$.

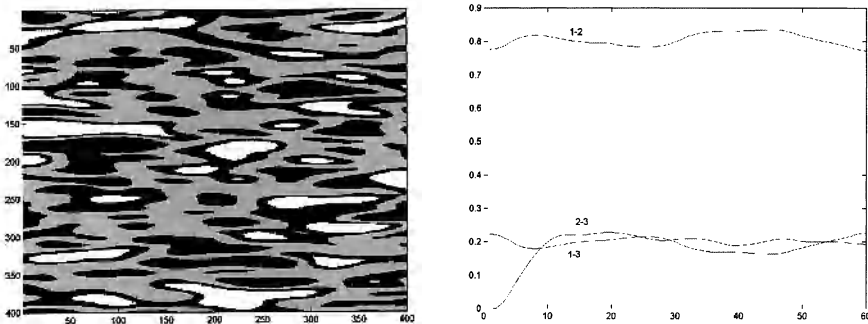
Example

The second type of transition probabilities were calculated in the vertical direction for the image shown in Figure 2.5, which was obtained by truncating a gaussian with an anisotropic gaussian variogram. The EW range was 50 units compared to 10 units in the NS direction. Figure 2.6 shows the probability that the second point ($x+h$) lies in facies 2 (or facies 3) given that the first point (x) is in facies 1 and that the second one is not in facies 1. We see that for short distance up to about 4 units the only possible transition is from facies 1 to 2. As the distance increases the probability that the second point lies in facies 3 rises steadily to about 0.35.

As a contrast, the image shown in Figure 2.7 was obtained by truncating the same gaussian random function using a different rock type rule. In the previous case, the rock type rule consisted of three intervals in the order: black, then grey and then white. Now there are four intervals: black, grey, black and white.



Figs 2.5 and 2.6: Simulation of a gaussian random function with a gaussian variogram with an EW range of 50 units and a NS range of 10 units (left), and on the right the probability of a point at position $x + h$ being in facies 2 (or facies 3, respectively) given that the first point at position x lies in facies 1 and that the second one does not, plotted as a function of the distance between the two points



Figs 2.7 and 2.8: Simulation of a gaussian random function with a gaussian variogram with an EW range of 50 units and a NS range of 10 units (left), and on the right the probability of a point at position $x + h$ being in facies 2 (or facies 3, respectively) given that the first point at position x lies in facies 1 and that the second one does not, plotted as a function of the distance between the two points

Figure 2.8 shows the corresponding transition probability. The essential difference between these two is that in the first case the black never touches the white whereas it does in the second case.

Exercises

The first two exercises in this chapter are designed to familiarise readers with indicator random functions. The next two illustrate the types of inconsistencies that arise with unsuitable models for indicator covariances. The last exercise gives some rather interesting results on multi-point statistics.

Exercise 2.1

- (a) Figure 2.9 shows twelve regularly spaced points in a 2D zone containing two facies (coloured grey and white). Let F denote the white facies. Write the values of the indicators of F and \bar{F} in Table 2.4.

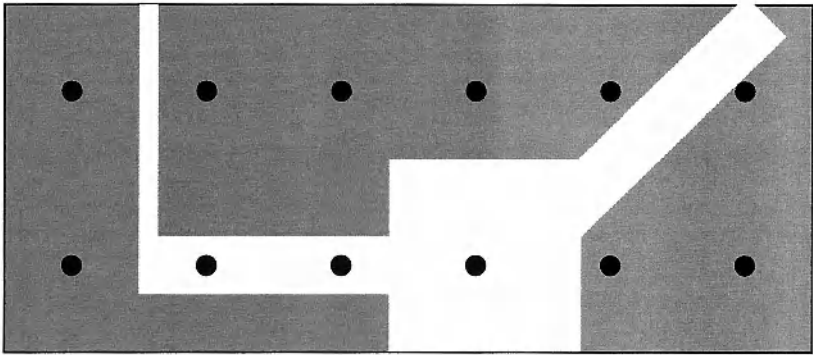


Fig 2.9: Twelve samples are available in an area contains a facies F that is coloured in white. Its complement \bar{F} is shaded grey

Table 2.4: Values of the indicator for the facies F and its complement

Point N°	1	2	3	4	5	6
$1_F(x)$						
$1_{\bar{F}}(x)$						
$1_F(x) + 1_{\bar{F}}(x)$						

Point N°	7	8	9	10	11	12
$1_F(x)$						
$1_{\bar{F}}(x)$						
$1_F(x) + 1_{\bar{F}}(x)$						

- (b) Calculate the experimental mean and variance of the indicator for F and of its complement. Check that the means lie between 0 and 1, and that the variances are no greater than 0.25. Under what circumstances would the variances equal 0.25.
- (c) Calculate the experimental variogram for facies F, and then for its complement for up to 3 lags in the EW direction. Plot these. Why are they the same? What shape do they have near the origin?
- (d) The next step is to calculate the centered spatial covariances for F and its complement, then their cross covariances. Check that the direct covariances are identical and that the cross covariances are the negative of these.

Exercise 2.2

Assume that a facies F is stationary with probability $P_F = 0.6$ in the area under study. Show that the variance of the indicator $1_F(x)$ equals 0.24. Suppose that its variogram can be modelled as an exponential with a scale parameter equal “a”.

- (a) What is the value of its sill?
- (b) Write down the equation for the variogram of its complement.
- (c) Write down the equations for the centred covariance of the indicator for F and for its complement.
- (d) Plot the non-centred covariances for these two indicators.

Exercise 2.3 Inappropriate models for indicator variograms

This example taken from Matheron (1987) and Armstrong (1992) highlights the types of inconsistencies that can arise when inappropriate models are chosen as indicator variograms. In this case we are going to show again that the gaussian variogram is unsuitable as a model for the variogram of the indicators.

Consider two points x_1 and x_3 . Let x_2 be their midpoint. As each of the three indicators can take two values, 0 and 1, there are 8 possible combinations for the values of the three indicators. Let ω denote the probability that all three indicators take the value 1. That is,

$$\omega = \Pr[1_F(x_1) = 1_F(x_2) = 1_F(x_3) = 1]$$

As usual the non-centred covariance is:

$$C_{ij} = C_F(x_i, x_j) = E[1_F(x_i) 1_F(x_j)] = \Pr[1_F(x_i) = 1_F(x_j) = 1]$$

- (a) Show that

$$\Pr[1_F(x_1) = 1_F(x_2) = 1 \text{ \& } 1_F(x_3) = 0] = C_{12} - \omega$$

24 Basic Properties

$$\Pr[1_F(x_2) = 1_F(x_3) = 1 \& 1_F(x_1) = 0] = C_{23} - \omega$$

$$\Pr[1_F(x_1) = 1_F(x_3) = 1 \& 1_F(x_2) = 0] = C_{13} - \omega$$

Hint: split the event $[1_F(x_1) = 1_F(x_2) = 1]$ into two mutually exclusive events according to the value of the third indicator e.g.

$$\begin{aligned} \Pr[1_F(x_1) = 1_F(x_2) = 1] &= \Pr[1_F(x_1) = 1_F(x_2) = 1_F(x_3) = 1] \\ &+ \Pr[1_F(x_1) = 1_F(x_2) = 1 \& 1_F(x_3) = 0] \end{aligned}$$

(b) Similarly show that

$$\Pr[1_F(x_1) = 1 \& 1_F(x_2) = 0 \& 1_F(x_3) = 0] = C_{11} - C_{12} - C_{13} + \omega$$

$$\Pr[1_F(x_1) = 0 \& 1_F(x_2) = 1 \& 1_F(x_3) = 0] = C_{22} - C_{12} - C_{23} + \omega$$

$$\Pr[1_F(x_1) = 0 \& 1_F(x_2) = 0 \& 1_F(x_3) = 1] = C_{33} - C_{13} - C_{23} + \omega$$

Hint: consider the event $[1_F(x_1) = 1 \& 1_F(x_2) = 1_F(x_3) = 0]$

(c) The contradiction becomes apparent if the three points are set 0.15 apart (compared to a scale parameter of 1) and if the probability that $1_F(x_i) = 1$ is 0.1.

$$C_{11} = C_{22} = C_{33} = 0.9$$

Calculate the value of the gaussian variogram for a distance of 0.15 and show that the following values are obtained for the noncentred covariance:

$$C_{12} = C_{23} = 0.1 \times 0.9 \times 0.494 = 0.0396 \text{ and } C_{13} = 0.1 \times 0.9 \times 0.478 = 0.0430$$

Show that $0 \leq 0.043 - \omega$ since $\Pr[1_F(x_1) = 1_F(x_3) = 1 \& 1_F(x_2) = 0] \geq 0$

In the same way, use the fact that

$$\Pr[1_F(x_1) = 0 \& 1_F(x_2) = 1 \& 1_F(x_3) = 0] = C_{22} - C_{13} + C_{23} - \omega \geq 0$$

to prove that $\omega \geq 0.821$.

Since it is impossible for $\omega \leq 0.043$ and $\omega \geq 0.821$, the gaussian variogram model is incompatible with indicator data.

Exercise 2.4: Another unsuitable variogram model

At first one might be tempted to think that the gaussian variogram is not allowable because it is infinitely differentiable or because it is quadratic near the origin. To show that this is not the case, Matheron produced a construction for a variogram model

$$f(x) = \left[a^2 e^{-ax} - b(2a - b) e^{-bx} \right]$$

where $a < b < 2a$ and $\omega^2 = 2ab - b^2$.

This is licit for indicators in 1-D provided that $\omega < b$. However he also proved that if $\omega > b$, the triangle inequality is not respected. To see this, let $b = 0.1$, $\omega = 1$ and $h = 0.2$.

Exercise 2.5: Multi-point Statistics

Covariances and variograms are examples of two point statistics. This exercise gives some theoretical results on multi-point statistics. Part (a) shows that once one of the three point covariances is known, all the others can be expressed in terms of it and of lower order statistics.

(a) Let

$$C_{FF\bar{F}}(x, y, z) = E(l_F(x)l_F(y)l_{\bar{F}}(z))$$

Show that :

$$C_{FF\bar{F}}(x, y, z) = -C_{FFF}(x, y, z) + C_{FF}(x, y)$$

$$C_{FF\bar{F}}(x, y, z) = C_{FFF}(x, y, z) - C_{FF}(x, y) - C_{FF}(x, z) + P_F(x)$$

$$\begin{aligned} C_{\bar{F}\bar{F}\bar{F}}(x, y, z) = & -C_{FFF}(x, y, z) + C_{FF}(x, y) + C_{FF}(x, z) + C_{FF}(y, z) + 1 \\ & - P_F(x) - P_F(y) - P_F(z) \end{aligned}$$

By permuting the order of the coordinates show that

$$C_{FF\bar{F}}(x, y, z) = -C_{FFF}(x, y, z) + C_{FF}(x, z)$$

Deduce similar expressions for the remaining three point statistics. (Exercise 2.3 is based on these.) We can now extend this result to higher order multi-point statistics. First we need to define the notation. Let

$$C_{FF}(x_1, x_2) = C_{F^2} \quad \text{and} \quad C_{FFF}(x_1, x_2, x_3) = C_{F^3}$$

More generally, let the covariance for the case where the first $(p-k)$ points belong to the facies F and the remaining k points to its complement be

$$C_{\underbrace{F, \dots, F}_{(p-k)}, \underbrace{\bar{F}, \dots, \bar{F}}_k}(x_1, \dots, x_p) = C_{F^{p-k} \bar{F}^k}(x_1, \dots, x_p)$$

Show that for any two integers k and p

$$C_{F^{p-k}, \bar{F}^k}(x_1, x_2, \dots, x_p) = (-1)^k C_{F^p}(x_1, x_2, \dots, x_p) + \Phi[(x_1, x_2, \dots, x_p)]$$

The function Φ involves only lower order covariances $C_{F^{p-i}}$ and proportions. This can be deduced by expanding the appropriate product. For example, to find $C_{FF\bar{F}}(x, y, z)$, we expand

$$F(x)[1 - F(y)][1 - F(z)] = F(x)F(y)F(z) - F(x)F(y) - F(x)F(z) + F(x)$$

This gives

$$C_{FF\bar{F}}(x, y, z) = C_{FFF}(x, y, z) - C_{FF}(x, y) - C_{FF}(x, z) + P_F(x)$$

By permuting the coordinates we can show that any n -point statistics can be expressed using the basic n -point covariance $C_{F^p}(x_1, x_2, \dots, x_p)$ and terms involving lower order covariances.

Exercise 2.6: Erosion model

Consider two models with two facies where A and D are the main facies and \bar{A}, \bar{D} their complements. Then form a model with three facies, A , B and C , by letting A erode D and its complement.

- Write the indicators of A , B , C using the indicators of A , and D .
- Use the previous results to compute the indicator covariances of A , B , C for the case when the facies A is independent of D .

Exercise 2.7: Porous medium

Consider a porous media. Let A be the random set describing the pores. The porosity of support V is the mean value of the indicator A on the volume V centred at the point x . Compute the analytical form of

- The covariance of the porosity of support V
- The covariance of another indicator F with the porosity of support V .

3 PROPORTIONS

The simplest quantity that can be measured from an image showing different facies or lithotypes is the percentage of space taken up by each facies. So the first step in a plurigaussian study is to calculate the proportions from the experimental data. In the stationary case, we only need calculate the percentage of all the data that belong to each facies to get the proportions. However in most practical cases, the geology is more complicated. For example, petroleum reservoirs are not stationary in the vertical direction because of cyclic changes during their deposition. Vertical proportion curves were designed to quantify these changes. We will show how to compute and interpret these curves. In more complicated cases the proportions vary laterally as well. In that case we build up a 3D matrix of proportion curves. Although plurigaussian simulations were first developed for the oil industry, they are now used in mining as well. Proportion curves and 3D proportion matrices are also used there to estimate the proportions of each rock type.

The first section in this chapter describes vertical proportion curves for the simplest case where the reservoir or deposit is stationary horizontally. These curves summarise the vertical variability in the proportions. After that we treat the more general case of non-stationarity where we use 3D proportion matrices. One key factor when calculating vertical proportion curves (VPC, for short) is the choice of the reference level. An example is presented to illustrate the impact of inappropriate choices on VPC.

How to calculate vertical proportion curves

Vertical proportion curves first proposed by Beucher et al (1987) are a simple tool for quantifying the evolution in the amount of each facies or lithotype present as a function of depth. They are computed along lines parallel to the chosen reference level (generally a chrono- stratigraphic marker). The results are presented as a graph showing the proportion of each facies at each level. We illustrate this procedure using a simple example.

Figure 3.1 shows five fictive wells each containing five core sections of equal length. Three lithotypes which we call sandstone, shaly-sandstone and shale, have been recorded. Looking at this figure, we see that there is more shale at the tops of the wells than at the bottoms. In fact, all five core sections in the top row are shale (i.e. 100% shale). In the second row there is 80% shale and 20% shaly-sandstone. At the bottom there is 40% shaly-sandstone and 60% sandstone, but no shale.

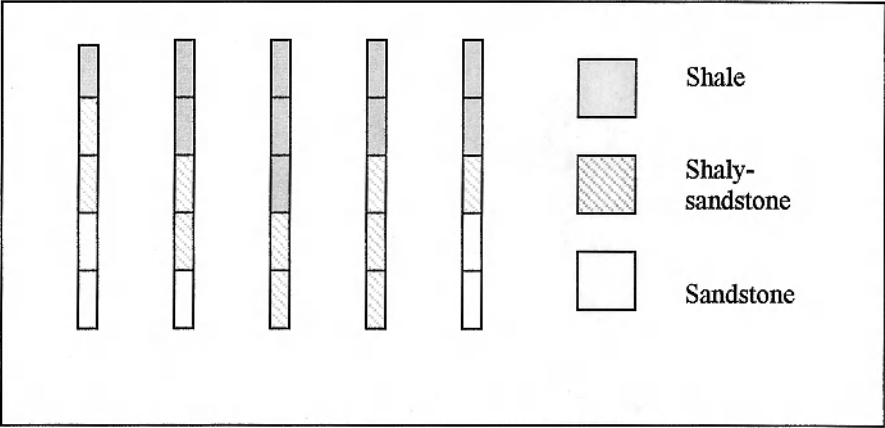


Fig 3.1: Five drill-holes each containing five core sections of equal length

These proportions are presented graphically in Figure 3.2. The order that the facies are arranged in is important. It must reflect the evolution seen in the geology. In this case the interpretation is simple: the grain size at the top of the wells is much smaller than at the bottom. This indicates sedimentological evolution from a high energy medium at the bottom to a low energy one at the top.

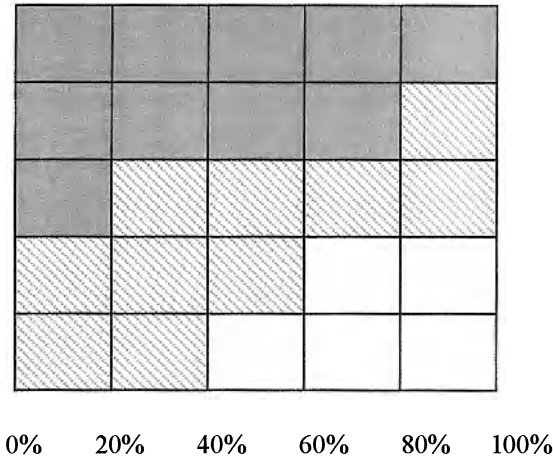


Fig 3.2: Vertical Proportion Curve showing the proportion of each facies per vertical level

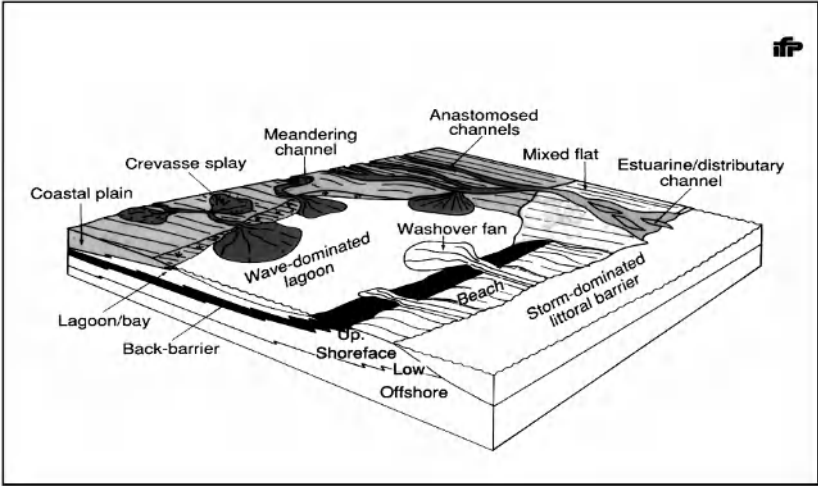


Fig 3.3: Conceptual model of the deposition

Example 1: the Ravenscar sequence

The Ravenscar sequence which outcrops in cliff faces near Scarborough in North Yorkshire, consists of about 200m of siliclastic sediments dating from the Middle Jurassic era. The Cleveland basin to which it belongs is a progradation of a deltaic system with marine influences. Figure 3.3 (taken from Beucher, Doligez and Yarus, 2000) shows a conceptual model of it.



Fig 3.4: Vertical proportion curve for the Saltwick unit of the Ravenscar sequence

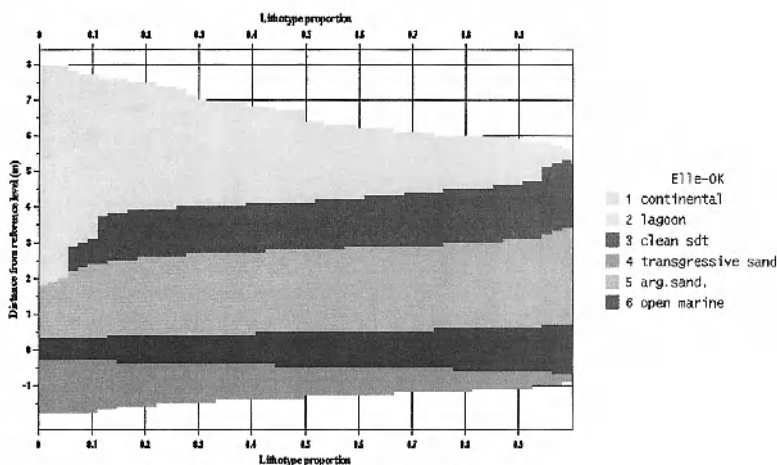


Fig 3.5: Vertical proportion curve for the Ellerbeck unit of the Ravenscar sequence

The Ravenscar sequence consists of seven units. From the top down, these are Scarborough, Gristhorpe, Millepore, Sycharham, Cloughton, Ellerbeck and Saltwick. Each had a distinctly different depositional environment. For example, the Saltwick unit was laid down in a continental environment whereas the Ellerbeck unit is characterised by steady changes in the sea level. Figures 3.4 and 3.5 (from Beucher, Doligez and Yarus, 2000) show the vertical proportion curves for these two units. The three main channel producing episodes are clearly visible in the VPC for the Saltwick unit. In contrast to this, the VPC for the Ellerbeck unit shows regular changes from an open marine environment (marine mudstones and marls) at the bottom of the unit to continental facies at the top. Intermediate lithounits include argillaceous sandstones and transgressive sandstone corresponding to the shoreface and clean sandstone from the foreshore and upper shoreface. The top of the unit is characterised by a lagoon environment with a floodplain, continental mudstone and argillaceous sands.

Example 2: facies with contrasting anisotropies in a gold deposit

The Lupin mine in Northwest Canada is a good example of a stratiform banded iron formation (BIF) hosted gold deposit. Some of the gold is uniformly disseminated in thin, laterally continuous units of sulphur-rich BIF while the rest is contained in steeply inclined quartz veins that have overprinted the BIF. See Kerswill et al (1996) for more information. The two gold-bearing lithofacies and the barren background are present in the schematic vertical cross-section of the orebody shown in Figure 3-6 from Roth et al (1998).

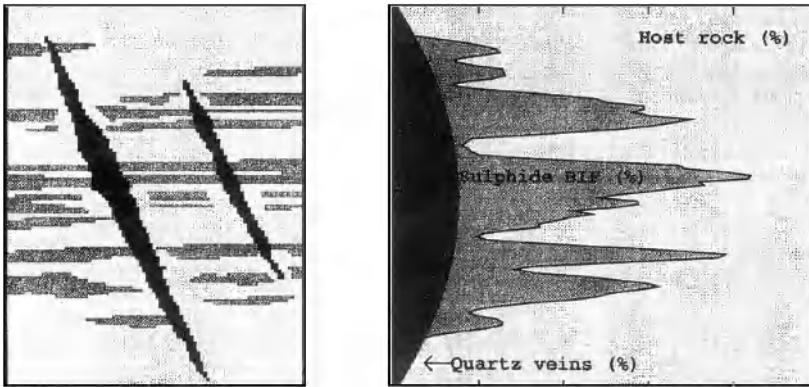


Fig 3.6 (left): Schematic vertical cross-section of the Lupin orebody showing two mineralised lithofacies with contrasting anisotropies on a waste background; **Fig 3.7** (right): the vertical proportion curve corresponding to Fig 3.6

Figure 3.7 shows the experimental VPC computed from Figure 3.6, with the proportion of quartz vein in dark grey and that of sulphide BIF in light grey. The proportion of quartz vein changes gradually reaching a maximum near the middle of the cross-section. In contrast to this, the proportion of sulphide BIF is erratic. Before this VPC can be used to calculate the thresholds we have to decide whether these peaks are representative of genuine geological features in the deposit or whether they are due to statistical fluctuations. Looking back at Figure 3.4, we see that same problem arose with the VPC for the Saltwick unit of the Ravenscar sequence. If the peaks are a real feature they should be modelled. Otherwise we should smooth them. Alternatively, the reference level has to be changed. This decision can only be made after discussion between the geologist or the engineer and the geostatistician. In both of these cases it was decided that the peaks represented real features of the deposit/reservoir that should be reproduced in the simulations. So the thresholds were set to generate more BIF (or more channels, as the case may be) at the corresponding depths.

Horizontal nonstationarity

The vertical proportion curves presented in the previous section were computed assuming that the deposit/reservoir was at least horizontally stationary. In many cases, this is not true. The proportions vary laterally from one area to another. The Millpore and Gristhorpe units of the Ravenscar sequence in Yorkshire are a typical example of this. The depositional environment varied from marine and littoral deposits. It is characterised by marine mudstone, wash-over argillaceous sandstones through to deltaic environment with a floodplain and fluvial and

deltaic channels. Beucher, Doligez and Yarus (2000) constructed a 3D matrix of proportion curves to model the evolution in the proportions both vertically and laterally. Their first step was to group wells locally and compute a vertical proportion curve for each set. The percentage of each of eight lithotypes was then kriged at the nodes of a regular grid (Fig 3.7a) and the kriged estimates were recombined to give a vertical proportion curve. As the proportions were kriged individually, they do not add up to 100% at each grid node. So the values were rescaled to guarantee this. Figure 3.7b shows nine cells from the 3D proportion matrix.

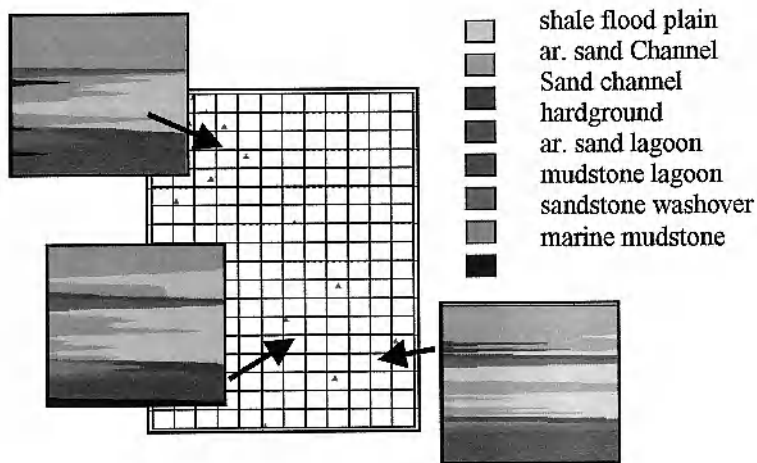


Fig 3.7a: Kriging VPC onto a regular grid

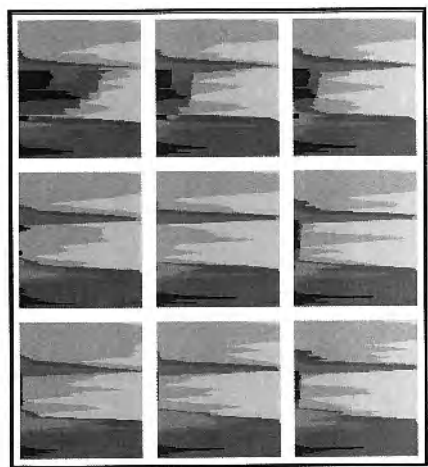


Fig3.7b: Nine cells from the 3D proportion matrix computed by Beucher, Doligez and Yarus (2000)

Having to rescale the proportions to make them sum to 100% is not entirely satisfactory. One could envisage using cokriging, but this would result in difficulties setting up the structural model (i.e. the variograms and the cross-covariograms). A more fundamental reason for using a simple technique like kriging is that it is then easier to incorporate additional information, for example from seismic data. Moulière (1996, 1998) considered a case where 2D seismics gave the cumulative thickness of one lithotype (shaly sandstone) for each unit in the formation. So it was important to incorporate this constraint into the proportion curves. She proposed two methods for doing this: one based on cokriging and the other using kriging. She showed that the simulations obtained after taking account of the available seismic information were much more realistic than those based on data from the limited numbers of wells that were available.

Lateral non-stationarity was also encountered by Doligez, Beucher, Fonnesu et al (1994) in their study of the Cajigar 2 succession. As the proportions changed from one area to another, they constructed a matrix of lithofacies proportions on a 300m x 300m grid. Each horizontal proportion curve was obtained by averaging the percentages of each lithofacies in the three levels above and below, as well as the level itself. This has the advantage of maintaining the vertical correlations between successive levels, as well as smoothing the results. Whereas in vertical proportion curves, the proportions are plotted horizontally as a function of the vertical height, the horizontal proportion curves show the proportions on the Y-axis with horizontal coordinate along the X-axis.

Non-stationarity can also arise when studying mining deposits. Compared with petroleum data, far more drill holes are usually available but there is often no chronostratigraphic marker to help interpret the geology. Betzhold and Roth (2000) studied the Mantos Blancos copper orebody which is located 45km northeast of the city of Antofagasta in the north of Chile. Their objective in simulating the orebody was to improve the ore homogenisation procedure by providing mine-planning engineers with more accurate images of the key mineralogical units. The rocks at Mantos Blancos come from a volcanic sequence consisting of andesitic flows and flow breccia at the top, then flow-breccia and flows of porphyritic dacite, and at the bottom, flows of augen (quartz-eye) dacite.

The sequence dips to the southwest at 10 - 20° in the mine area. The mineralisation forms an irregular blanket ranging from 100m to 200m thick. As different mineralogical ensembles with similar average grades do not increase the grade variability, they can effectively be grouped together and treated as one new unit. This led Betzhold and Roth to define three ore type classifications:

- high grade ore comprising chalcocite-bornite ore
- low grade ore comprising chalcopyrite and chalcopyrite-pyrite ores
- waste rocks that are not sent to the flotation plant. These include the oxide ore (atacamite-malachite), pyrite and the barren rock that is generally found beneath the orebody.

Over the whole field, there is almost 10% of high-grade ore and about 45% of both low grade and non-copper sulphide material but these proportions are not constant in space. For example, there is much more low grade material near the top of the orebody. The proportions have to be modelled as a function of the point considered in the field.

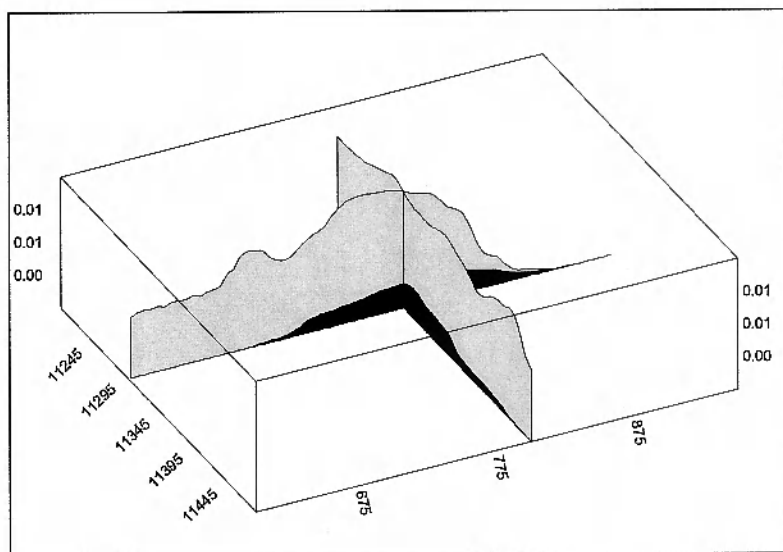
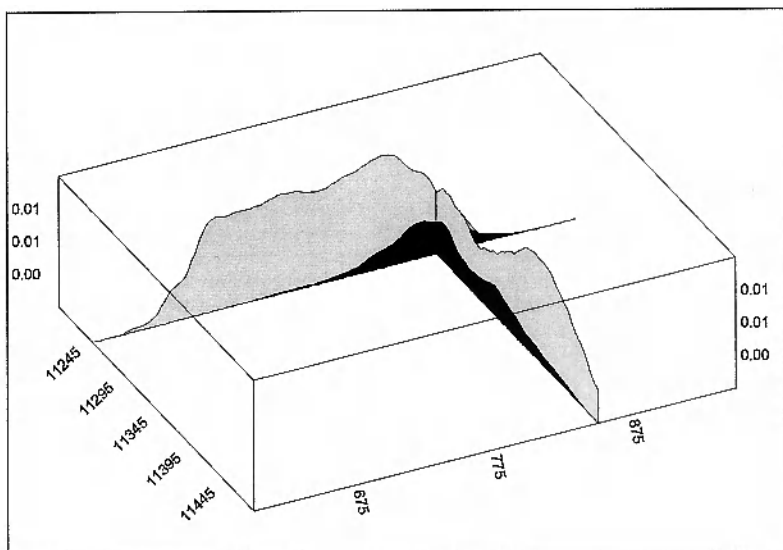


Fig 3.8: Two sets of proportion curves, an HPC and a VPC, calculated using just over 900 samples. The horizontal axes are the east-west direction (on the left) and the vertical direction (on the right). In the top diagram, the HPC was calculated close to the surface, whereas the VPC came from the Western side of the field. In the lower diagram both proportion curves were computed closer to the centre of the field. The proportions are shown as a coloured surface on the vertical axis with high grade ore in black, low grade ore in middle grey and poor material in light grey

Proportion curves were computed by a simple moving average procedure that was refined and complemented manually to incorporate knowledge about the geological characteristics of the site. Figure 3.8 shows two sets of proportion curves from a 2D test zone containing just over 900 samples. The horizontal axes are the east-west direction (on the left) and the vertical direction (on the right). In the top diagram, the horizontal proportion curve was calculated close to the surface, and the vertical proportion curve on the Western side of the field. In the lower diagram the proportion curves were computed closer to the centre of the field. The proportions of the three ore types are shown as a coloured surface on the vertical axis: high grade ore in black, low ore in middle grey and poor grade in light grey. The sum of the facies proportions is always equal to 100%. It is interesting to see how much the proportions of low and poor facies can vary over relatively small distances. In contrast the proportion of high grade ore seems more regular.

Choosing the reference level

The shape of the vertical proportion curve and the resulting simulations depend on the choice of the reference level. For oil reservoirs this is a specific geological marker which is used to restore the geometry of the reservoir at the time of deposition. This level must have been horizontal during sedimentation, and, should, if possible, correspond to a time line. That is, it should be a chronostratigraphic marker not an erosional unconformity. The reservoir is then flattened using this as the reference level.

Common choices for the reference level are the top of the unit or the bottom. Alternatively, a proportional grid could have been used to take account of differential subsidence. Figure 3.9 illustrates two of these three possibilities. Different reference levels can be used for each reservoir unit, or one serve for several of them.

To illustrate how important this choice is, Beucher (1998, personal communication) tested the impact of three possible reference levels on the vertical proportion curve and then the simulations. The Ellerbeck unit of the Ravenscar sequence which was described earlier in the chapter, was used as an example. The correct level is a flooding surface inside the Ellerbeck unit near the bottom. The two incorrect levels she considered were the base of the Ravenscar sequence and its top (denoted by R1 and R2 respectively). The base is a chronostratigraphic marker, the top is an unconformity. Figure 3.10 illustrates these three possibilities.

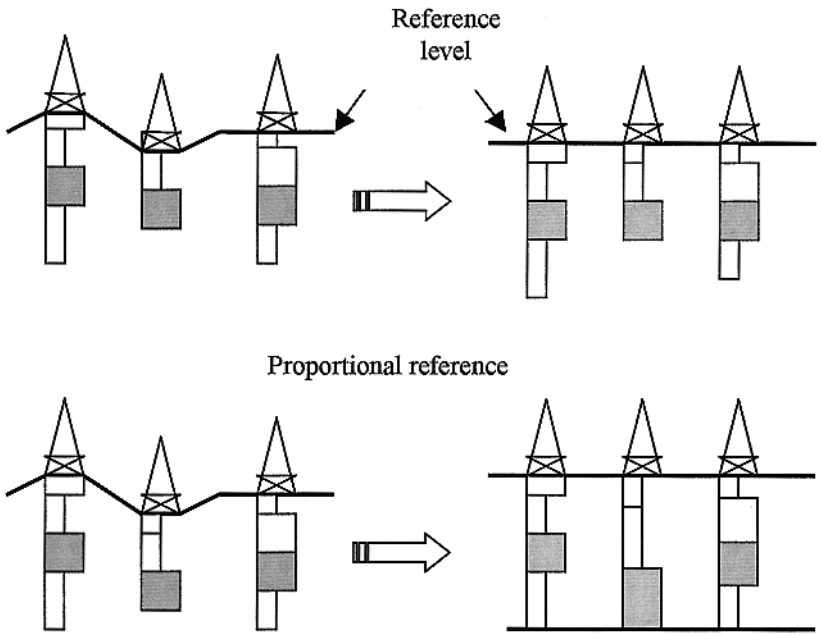


Fig 3.9: Two ways of flattening a reservoir, (a) using the top as reference level, (b) proportionately to account for differential subsidence

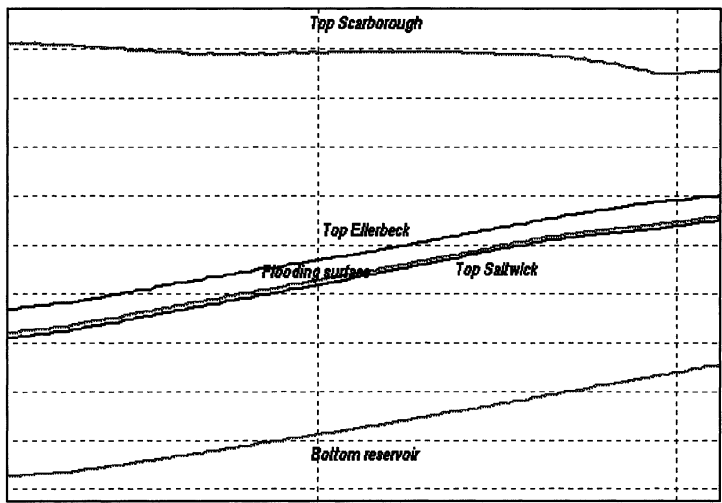


Fig 3.10: Three possible choices for the reference level; the flooding surface inside the Ellerbeck formation (the correct choice), the bottom of the reservoir (R1) and the top (R2)

Beucher computed the vertical proportion curves for the three candidate reference levels. Figure 3-11a shows the VPC for the best choice (denoted by OK) shows a steady gradation in the depositional environment. Starting from the bottom these are

- mudstone coming from an open marine environment,
- argillaceous sandstone from the low to mid shore,
- transgressive sandstone,
- clean sandstone from the foreshore,
- a mixture of argillaceous sandstone and mudstone from a lagoon environment, and finally
- continental mudstone from a flood plain environment at the top.

The second VPC (Fig 3.11b) corresponds to reference level R1 which is not the correct one. But its irregularities could have been due to cycles within the depositional cycles. The third VPC corresponding to reference level R2 is obviously incorrect (Fig 3.11c). Figure 3.12 shows one simulation of the unit obtained using each of the VPC. The correct one reproduces the type of continuity that would be expected at this scale, in this type of depositional environment. The lithotypes in the simulation corresponding to R1 are much less continuous and would lead to quite different fluid flows. The third simulation for R2 suggests no connectivity from one part of the unit to another.

It is not always obvious which level is the most appropriate for the reference. When in doubt, seismic data can sometimes provide guidance because it can give an indication of the cumulative thickness of sandstone. This could be compared to the cumulative thickness found in simulations. Fig 3.13 shows the cumulative sandstone thickness for the three simulations. The differences are quite marked. Volpi, Galli and Ravenne (1997) and Ravenne et al (2002) discuss vertical proportion curves in detail. The definition of a level has a marked impact on the result of the simulation. Onlap or toplap configurations can be produced depending on the choice of a reference level with regards to the unit geometry.

Non-stationarity

From a geological point of a view, a non-stationarity is characterised by a significant lateral change in the lithotype distribution in a reservoir unit, within the study area. It often shows up in the horizontal proportion curves as significant variations in the lithotype proportions in a given direction.

As geological phenomena always have some non-stationary aspects, depending on the scale of investigation, the choice between stationary or non-stationary models is subjective. It implies firstly a classification of the heterogeneity and then setting up a hierarchy. The distribution of the relevant heterogeneity will be then analysed to determine if it is stationary or not.

Non-stationarity concerns both object and sequence-based models. In object based models, it implies lateral variations of the object frequency. In sequence - based models, non stationarity can be handled by computing several proportion

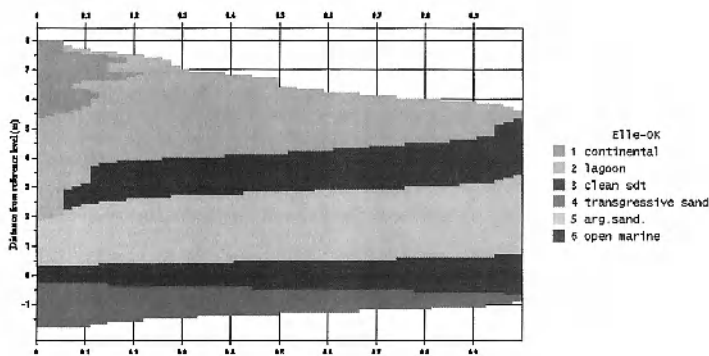


Fig 3.11a: Vertical proportion curve corresponding to the correct reference level

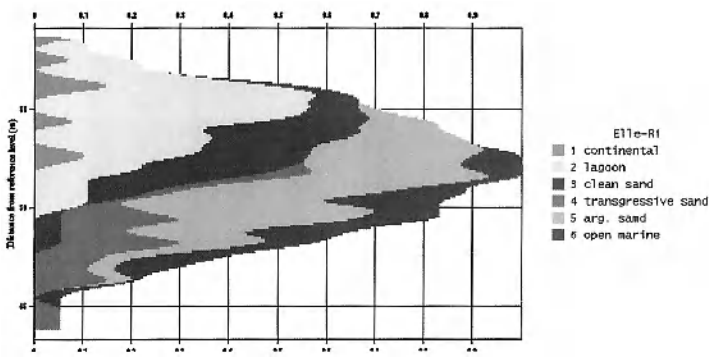


Fig 3.11b: Vertical proportion curve corresponding to the reference level R1

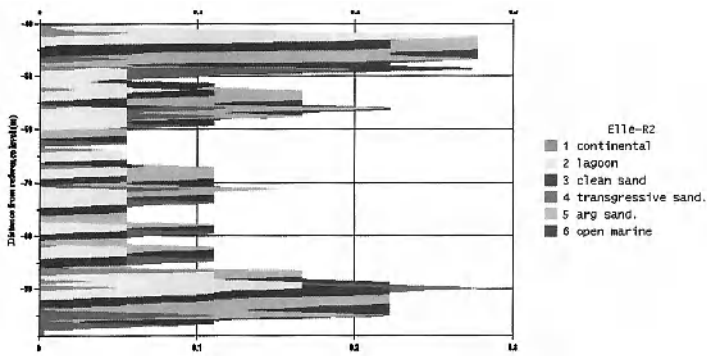


Fig 3.11c: Vertical proportion curve corresponding to the reference level R2

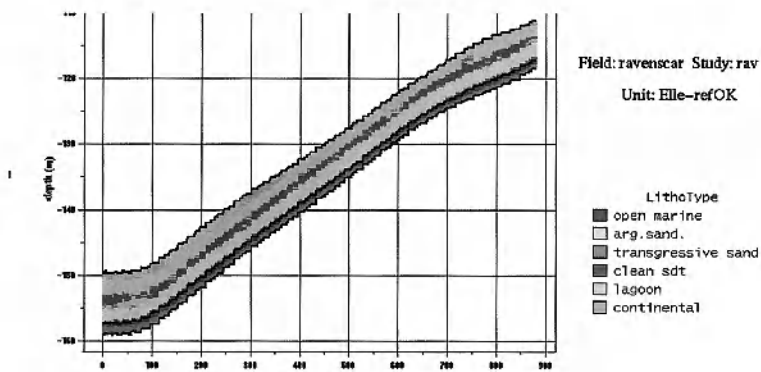


Fig 3.12a: One simulation of the unit obtained using the correct reference level

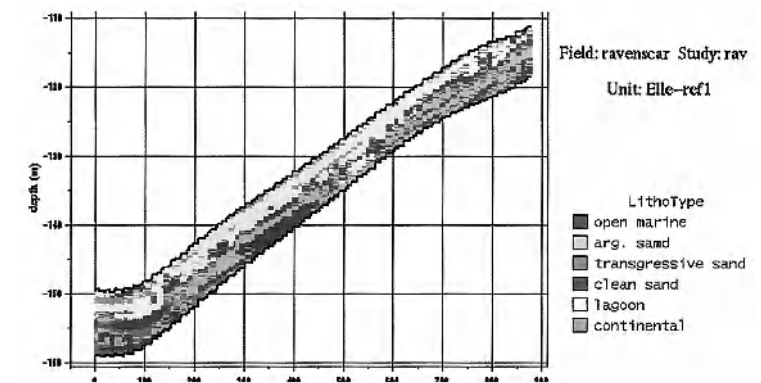


Fig 3.12b: One simulation of the unit obtained using the reference level R1

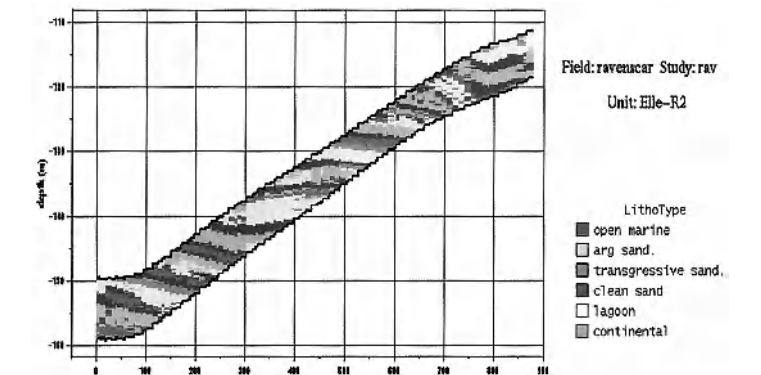


Fig 3.12c: Simulation of the unit obtained using the reference level R2

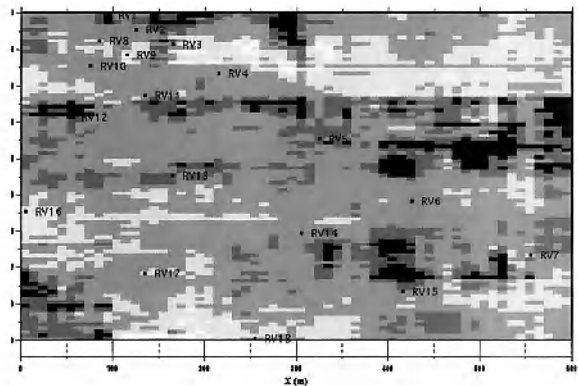


Fig 3.13a: Accumulated sand thickness obtained using the correct reference level

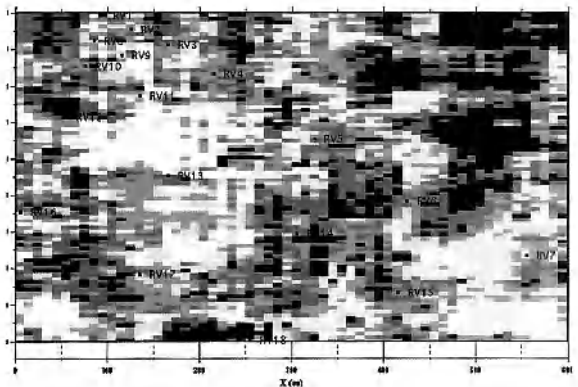


Fig 3.13b: Accumulated sand thickness obtained using reference level R1

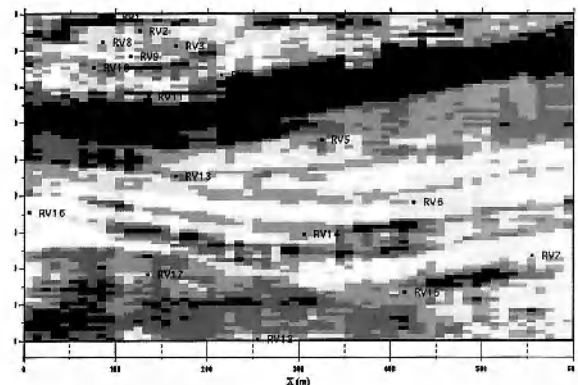


Fig 3.13c: Accumulated sand thickness obtained using reference level R2

curves for lithotypes in different areas of the reservoir. This leads to the construction of a proportion matrix, in which vertical proportion curves are computed in each cell of a grid (Beucher et al, 1993; Doligez et al, 1999). This approach is very flexible and gives realistical lateral variation of lithofacies. Furthermore, it is easy to integrate soft constraints such as an external drift into the computation of the proportion matrix. Geological information derived from seismic campaigns can then be integrated in the simulations (Mouliere et al., 1997; Fournier and Derain, 1997; Johann et al., 1996; Doligez et al., 1999).

4. TRUNCATION AND THRESHOLDS

Basic principle in the truncated gaussian method

In truncated gaussian simulations, the lithofacies are not simulated directly: a stationary gaussian random function is simulated first, and is then transformed into the lithofacies variable by truncation. For example, if we want to simulate two lithofacies, F_1 and F_2 , a very intuitive way to transform the simulated gaussian variable into facies values is to say “if the numerical value of the simulated gaussian is lower than the number t_1 , we obtain the first facies F_1 ; otherwise, we obtain the second facies F_2 ”. The value t_1 is called a threshold.

Defining the thresholds

Let x be any point in the simulated domain, let $1_{F_1}(x)$ and $1_{F_2}(x)$ be the indicators of the lithofacies F_1 and F_2 and let $Z(x)$ be the simulated gaussian function at point x . The transformation used is described mathematically by:

$$1_{F_1}(x) = 1 \Leftrightarrow -\infty \leq Z(x) < t_1$$

When we want to simulate a lithofacies variable which can take more than two values, we have to define more than one threshold ($N-1$ thresholds for N possible lithofacies values). The i^{th} facies F_i is defined by:

$$x \in F_i \Leftrightarrow 1_{F_i}(x) = 1 \Leftrightarrow t_{i-1} \leq Z(x) < t_i$$

The thresholds are in increasing order

$$t_1 \leq t_2 \leq \dots \leq t_{i-1} \leq t_i \leq t_{i+1} \leq \dots \leq t_{N-1}.$$

And we can obviously have one and only one facies value at each point.

Transitions between facies

Provided that the nugget effect in the variogram is zero, the simulated gaussian function is continuous. We are going to study the simulated values as we go from point x to point y .

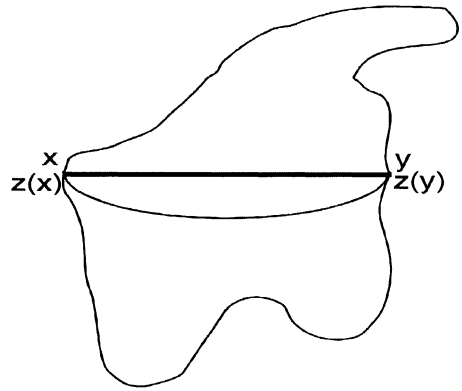


Fig 4.1: Examples of possible paths

For any intermediate value between $z(x)$ and $z(y)$, there must be at least 1 point where z takes that value. This is true for every path going from x to y , no matter how complicated it is (Figure 4.1).

This has important consequences. Suppose that the points x and y belong to different facies, respectively F_i and F_j , where $i < j$. Then we will find all the intermediate facies F_k ($i < k < j$) on all paths joining the points x and y (provided that the thresholds do not vary in space). Graphically, this corresponds to a characteristic pattern of concentric shapes (Figures 4.2 and 4.3). In both cases, if we go from a point in the white facies to one in the dark grey, we have to cross the light grey. So the contacts that are possible using the truncated gaussian model are defined by the ordering of the thresholds, i.e. by the ordering of the facies.

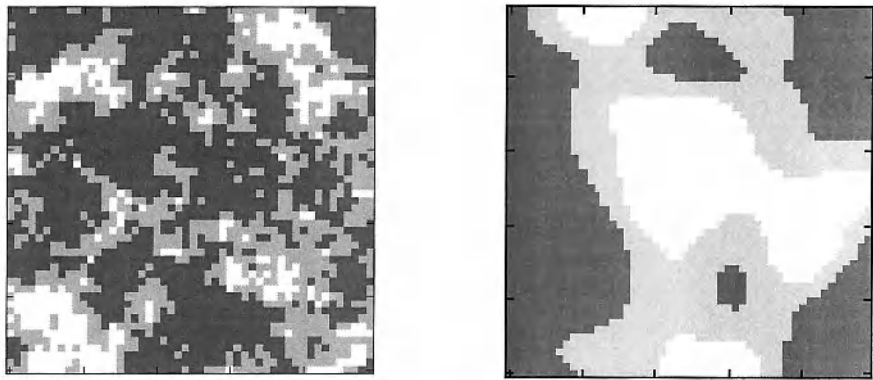


Fig 4.2 and 4.3: Characteristic patterns of concentric circles generated using an exponential variogram (left) and a gaussian variogram (right)

Figure 4.4 shows three simulations obtained with different choices of facies ordering. In the top figure, the light grey always occurs between white and dark grey. In the middle figure, the dark grey forms the buffer between the other two colours whereas in the bottom figure it is the white. This relationship is summarised in the rocktype rules to the right of each figure. The discretisation of the space can attenuate this effect, especially when the properties of the gaussian variable allow sharp (but nevertheless continuous) variations (as is the case in Figure 4.2). Variation of the thresholds in space can also attenuate this effect, especially when some facies disappear locally.

Link between thresholds and proportions

Earlier we saw that the proportion of a particular facies F_i at point x is the probability of having this facies F_i at that point x . This can be written as:

$$P_{Fi} = P(\text{facies at point } x = F_i) = E[1_{Fi}(x)]$$

As we have

$$1_{Fi}(x) = 1 \Leftrightarrow t_{i-1} \leq Z(x) < t_i,$$

we can also write:

$$\begin{aligned} p_{Fi}(x) &= P(t_{i-1} \leq Z(x) < t_i) = P(-\infty < Z(x) < t_i) - P(-\infty < Z(x) < t_{i-1}) \\ p_{Fi}(x) &= G(t_i) - G(t_{i-1}) \end{aligned}$$

where $G(t)$ is the cumulative distribution function for the standard normal distribution $N(0,1)$. (The choice of the normal is entirely conventional). As the proportions of each facies are known experimentally, we just invert this relationship to deduce the thresholds:

$$\begin{aligned} t_1 &= G^{-1}[p_{F1}(x)] \\ t_2 &= G^{-1}[p_{F1}(x) + p_{F2}(x)] \\ t_i &= G^{-1}[p_{F1}(x) + p_{F2}(x) + \dots + p_{Fi}(x)] \end{aligned}$$

Once the facies ordering has been chosen, we have a one to one relationship between the thresholds and the proportions. This shows a limitation of the truncated gaussian method: as the model orders the facies, this ordering must be realistic from a geological point of view. (See Figure 4.4)

Remark: Once the facies sequence is defined, it is of no importance whether they are numbered in an increasing order or a decreasing one. By replacing Z with $-Z$, both cases are equivalent.

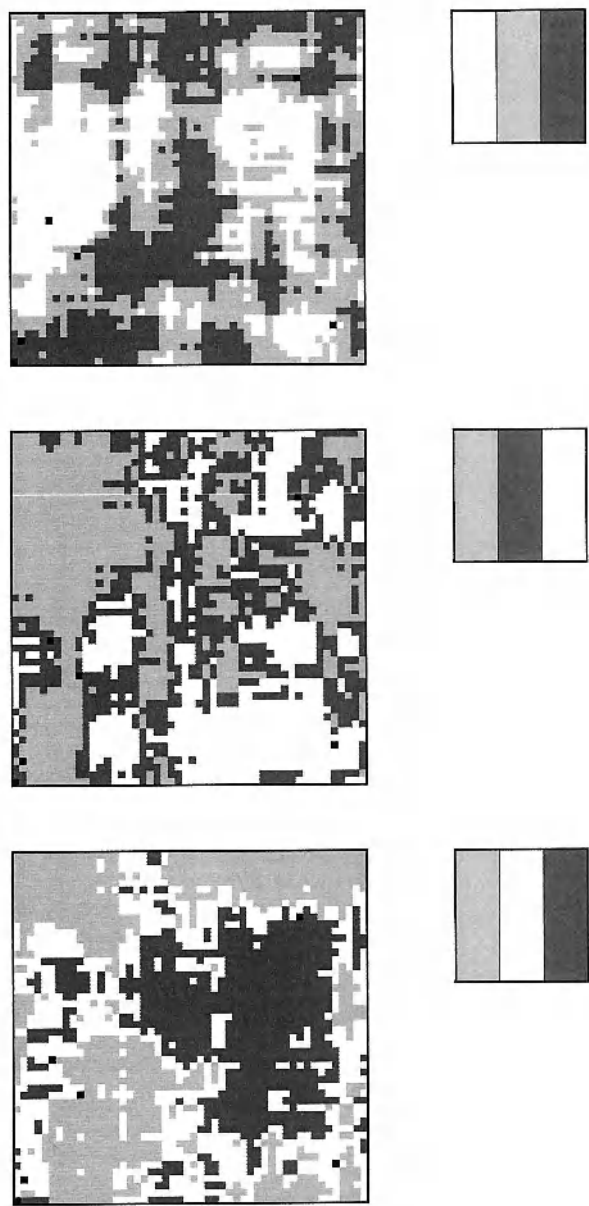


Fig 4.4: Three truncated gaussian simulations with the facies ordered in different ways. At the top, the light grey always occurs between white and dark grey. In the middle, the dark grey forms the buffer between the other two colours whereas at the bottom it is the white. This relationship is summarised in the rocktype rules to the right of each figure

Non stationary facies

If the facies are stationary, the thresholds are constant in space; otherwise they vary. The example of how to calculate vertical proportion curves given in chapter 3 is a typical case of vertical non stationarity. Table 4.1 lists the proportions of each facies in each level for that example and Table 4.2 gives the corresponding thresholds.

Table 4.1: Proportions of each facies shown in Figure 3.1

Level (numbered from bottom to top)	Shale (F1)	Shaly sandstones (F2)	Sandstones (F3)
5	100%	0%	0%
4	80%	20%	0%
3	20%	80%	0%
2	0%	60%	40%
1	0%	40%	60%

Table 4.2: Thresholds corresponding to proportions shown in Table 4.1

Level	t_0	t_1	t_2	t_3
5	$-\infty$	∞	∞	∞
4	$-\infty$	0.84	∞	∞
3	$-\infty$	- 0.84	∞	∞
2	$-\infty$	$-\infty$	0.25	∞
1	$-\infty$	$-\infty$	- 0.25	∞

The first and last thresholds are always equal to $-\infty$ and ∞ respectively, and they are not taken into account in the $N - 1$ thresholds (if we consider them, there are $N+1$ thresholds).

In non stationary cases, facies sometimes disappear locally. For example, if the i^{th} facies is not present, then $t_{i-1} = t_i$ as occurred in Table 4.2.

Idea behind the plurigaussian method

With the plurigaussian method, we simulate *several* gaussian functions instead of one. A transformation is made on the whole set of gaussian functions to obtain the facies value. We will illustrate how this is done via a few examples.

Example 1:

To start with, we limit ourselves to two gaussian functions $Z_1(x)$ and $Z_2(x)$, to simplify the graphical display. Figure 4.5 shows the 2D gaussian space and the position of two simulated points in this space. As the partition chosen here contains only two subsets, we will have two lithofacies (white and grey). At the first point x_1 the simulated values are $Z_1(x_1)$ and $Z_2(x_1)$, which we will shorten to z_{1-1} and z_{2-1} . This pair of values belongs to the subset labelled F_1 . This means that the white facies will be attributed to x_1 . At point x_2 the simulated values are z_{1-2} and z_{2-2} . This pair of values belongs to the subset labelled F_2 , which means that the grey facies will be attributed to x_2 . (N.B. the gaussian functions take all values between $-\infty$ and $+\infty$. We represent them in a rectangle, to simplify the display.)

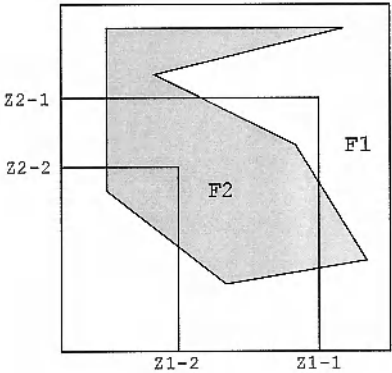


Fig 4.5: Rocktype rule for 2 facies

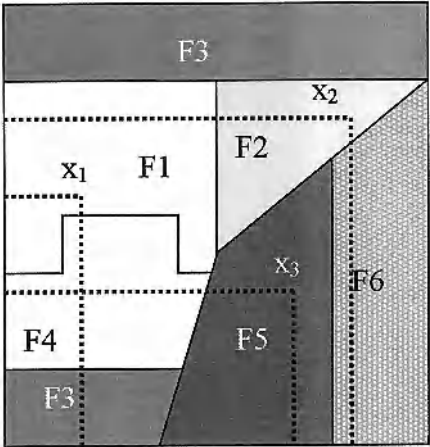


Fig 4.6: Rocktype rule with 6 facies. The three points lie in facies N° 1, 2 and 5 respectively

Example 2:

In this example, we have also two gaussian functions, but there are six lithofacies. Figure 4.6 shows the 2D gaussian space with its partition, and the position of three simulated points in this space. Here we attribute the facies F_1 to point x_1 , F_2 to point x_2 and F_5 to point x_3 .

Remarks:

1. In these figures, we see that the limits of the subsets can be rather complicated, so we can no longer speak of thresholds.
2. If we simulate N gaussian functions, they define a theoretical space with N dimensions. At any point x , the values of the simulated gaussian functions, z_1, z_2, \dots, z_N , define the co-ordinates of the simulated value in this new space. The facies are labels attached to parts of this space. If we want to be able to give a facies to each simulated point, these parts must define a partition of the gaussian space. If we have N gaussian functions Z_j , and if we let D_i be the subset of the gaussian space which is labelled as facies F_i , we can write:

$$1_{F_i}(x) = 1 \Leftrightarrow (Z_1(x), Z_2(x), \dots, Z_N(x)) \in D_i$$

3. The gaussian functions need not to be independent. They can be correlated.

Link with the proportions

As for the truncated gaussian, the proportion of facies F_i at point x is just the probability of having this facies F_i at that point. It can be written as :

$$p_{F_i}(x) = P(\text{facies at point } x = F_i) = E[1_{F_i}(x)]$$

As we have $1_{F_i}(x) = 1 \Leftrightarrow (Z_1(x), Z_2(x), \dots, Z_N(x)) \in D_i$ we also have:

$$p_{F_i}(x) = P\left\{\left[Z_1(x), \dots, Z_N(x)\right] \in D_i\right\} = \int_{D_i} g_{\Sigma}(z_1, \dots, z_N) dz_1 \dots dz_N$$

where $g_{\Sigma}(z_1, z_2, \dots, z_N)$ is the N -variate gaussian density function with mean 0 and variance 1, and Σ is its correlation matrix. Computing p_{F_i} is quite easy when we know Σ and D_i , but determining Σ and D_i even when we know all the p_{F_i} is impossible in the general case because there are an infinity of solutions. To solve this problem, we have to impose some constraints on the parameters.

Parameter simplification: use of thresholds

To overcome this problem, we choose to partition the gaussian space into rectangles (if we have 2 gaussian functions) or in rectangular parallelepipeds (with more than 2 gaussian functions). Figure 4.7 shows an example of one such partition with 2 gaussian functions. These rectangular boxes are defined by their projections on the gaussian co-ordinate axes. The end points of these projections can be considered as thresholds (see Figure 4.8). Each rectangular box is defined by 2^N thresholds (some of them can be infinite). When we know $2^N - 1$ of them, it is possible to invert the equation

$$p_{Fi}(x) = \int_{D_i} g(z_1, z_2, \dots, z_N) dz_1 dz_2 \dots dz_N$$

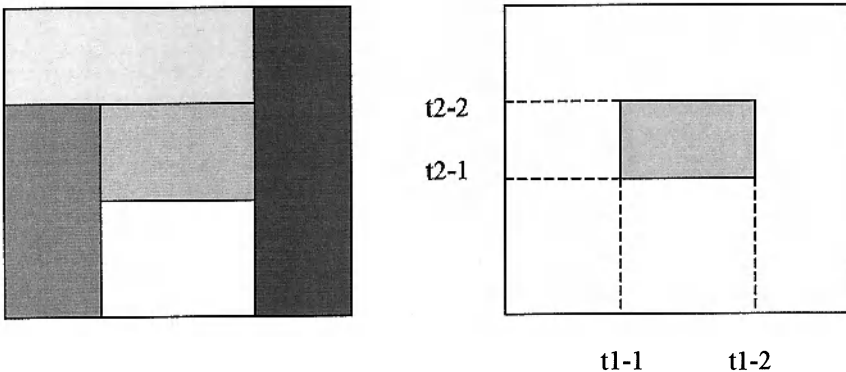
numerically to find the last unknown threshold. If we have N gaussian functions and M facies, we have $M \times 2^N$ thresholds. As we do not allow “empty space”, our rectangular boxes constitute a partition of this space. This considerably decreases the number of independent thresholds. For example, if we include the infinite thresholds, we have:

- M+1 thresholds with 1 gaussian function (i.e. truncated gaussian method),
- M+3 thresholds with 2 gaussian functions,
- M+5 thresholds with 3 gaussian functions,
- M+2N-1 thresholds with N gaussian functions,

which means $M-1$ finite independent thresholds.

Choice of the partition

Even after having decided to use a rectangular partition, there are many possible layouts. For example with two gaussian functions and four lithofacies, we have ten



Figs 4.7 and 4.8: Partition in rectangles (left) and thresholds (right)

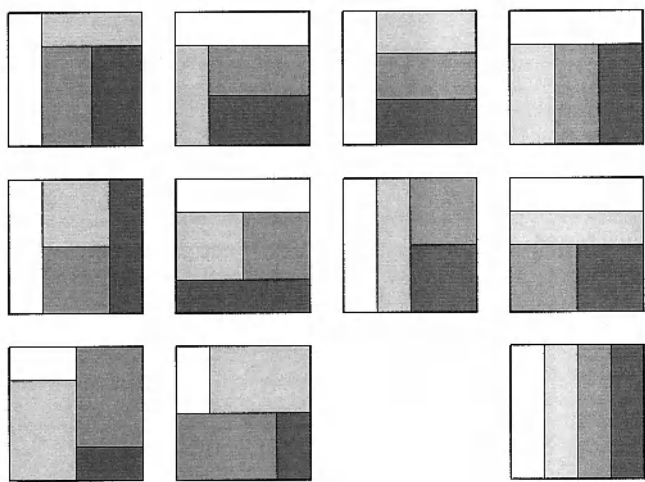


Fig 4.9: Rectangle rocktype rules for 4 facies

possibilities, plus one which is equivalent to the case with only one gaussian function. Figure 4.9 shows these partitions. With 6 facies, we have approximately 140 possibilities. This shows that we will not be able to test all the possible partitions. We have to make a choice, and then find a criterion to help us do this. Figure 4.10 shows examples of simulations corresponding to two cases from Figure 4.9.

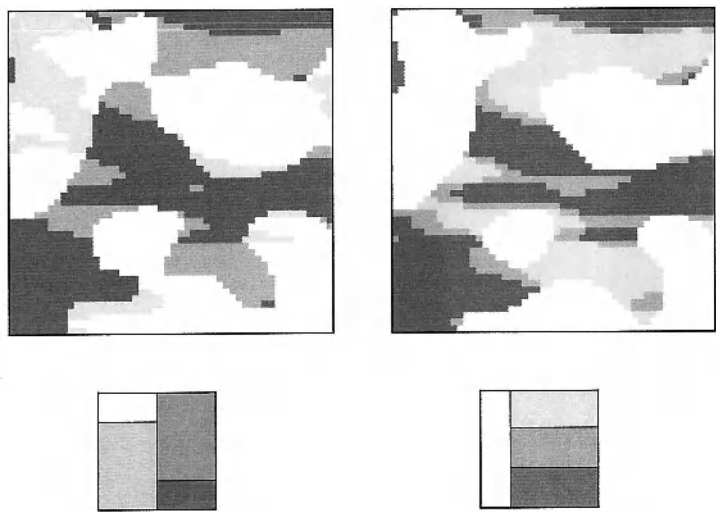


Fig 4.10: Examples of simulations for two of the rocktype rules shown in Fig 4.9;

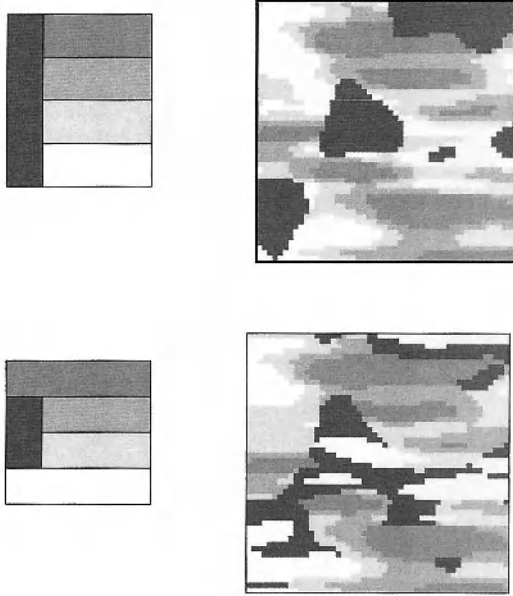


Fig 4.11: Examples of simulations with the same contacts, but with different rocktype rules

We can see that the facies that touch each other in the rocktype rule are also in contact in the simulation. This is a general rule when the proportions are constant, and when the simulated field is large enough to be statistically representative.

This rule is usually sufficient to choose the partition if there are only a few facies, but when the number increases, we can find several partitions which give the same contact possibilities (Figure 4.11). In that case, the shape of these contacts can also help us. We have to decide which gaussian function will guide the shape of which contact.

When the proportions vary, some facies can disappear in some parts of the simulated field. In that case, the rocktype rule also varies, and the forbidden contact will locally disappear.

Choice of the correlation matrix

With two gaussian functions, the correlation matrix which can be written as

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

is consistent provided that $-1 < \rho < 1$. The values $\rho = \pm 1$ are also authorised, but in that case it is better to work with only one gaussian function.

The covariance matrix must always be positive definite. With more than two gaussians, it would be virtually impossible to get a positive definite matrix just by

picking values for the terms. It is better to give a consistent model of relationships between the gaussians, and then deduce a suitable covariance matrix from this. We will not go into more detail here.

With two gaussian functions, once the rocktype rule is chosen, we still need to know the M-1 thresholds and the correlation, ρ , which gives M unknowns. So we need M equations linking the unknowns to the proportion values. We have M equations but they are not independent because the partition automatically ensures that the proportions sum to 1. As a consequence, we have more variables than independent equations. The number of solutions is infinite. Worse, these solutions are not equivalent: they give rise to simulations that are quite different.

For example Figure 4.12 shows two simulations obtained with the same partition but different correlations, 0 and 0.6. The shape of the black facies is exactly the same in both cases because the first gaussian function is the same, but those of the three other facies change. This is easiest to see on the white facies. In the right hand simulation where the correlation is 0.6, the white tends to wrap around the black, especially in the bottom left corner. In chapter 7, we present a case study where correlation will be used to drape one facies over algal bioherms. The example given above shows that increasing the correlation coefficient introduces a border effect, which looks like the ordering effect we have with only one gaussian function. This is the reason why we suggest choosing the value of the correlation coefficient arbitrarily, depending on whether we want to have a strong border effect or not. Note that the correlation coefficient is a property of the gaussian functions: it will remain constant over the whole domain.

Calculating the thresholds

When we have chosen the correlation coefficient, we have M-1 independent equations of the form:

$$p_{Fi}(x) = \int_{D_i} g(z_1, z_2, \dots, z_N) dz_1 dz_2 \dots dz_N$$

where the D_i are rectangles $[t_{i-1}, t_i] \times [s_{i-1}, s_i]$.

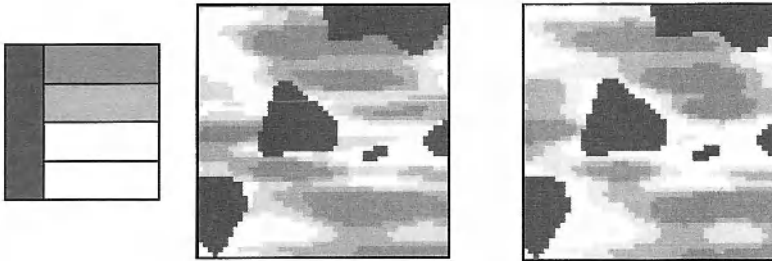


Fig 4.12: Simulations with the same lithotype rule and different correlation coefficients, $\rho=0$ and $\rho=0.6$

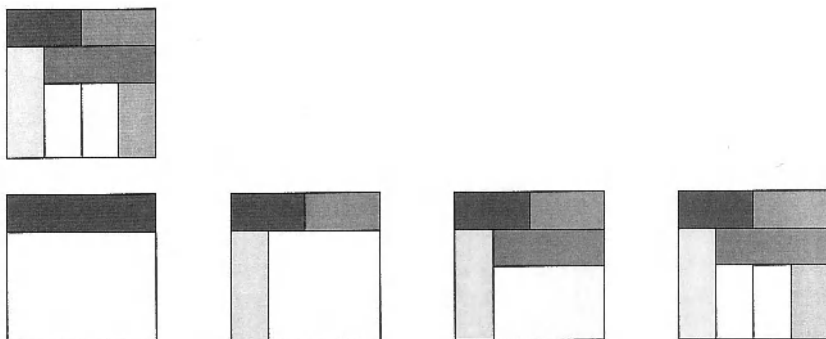


Fig 4.13: Successive groupings to obtain the thresholds

We have already seen that there are $M-1$ independent thresholds. This equation system cannot be solved analytically, but iterative methods can give us the solution. The trial and error method, testing successively all the thresholds, would be too slow, because it does not automatically ensure the consistency of the partition in rectangles.

In general, it is better to perform a global optimisation, minimising for example the global square error on the proportions. But in many cases, it is possible to group the facies and work successively on one gaussian function, then the other one (we may have to iterate this procedure). Figure 4.13 shows an example of this. Here the trial and error method is very quick, because the partition is automatically consistent. The partition we want to obtain is shown on the top line. The second line shows the order in which the thresholds are evaluated, starting with the top block.

Generalisation to non stationary case

As for the truncated gaussian, non-stationarity is obtained via varying proportions. This results in varying thresholds for both gaussian functions. The rocktype rule must be given including all the facies, even if some of them can disappear locally.

When simulations show “prohibited” contacts

Sometimes simulations show contacts which do not exist in the rock type rule diagram. In the stationary case, there are two possible reasons for this: one (or more) of the gaussian functions is discontinuous, or the discretisation is too coarse to show the continuity. This problem occurs much more often in the non-stationary case. We will demonstrate this in the most common case of vertical non stationarity for the truncated gaussian, but the same applies for horizontal non stationarity and in the plurigaussian case.

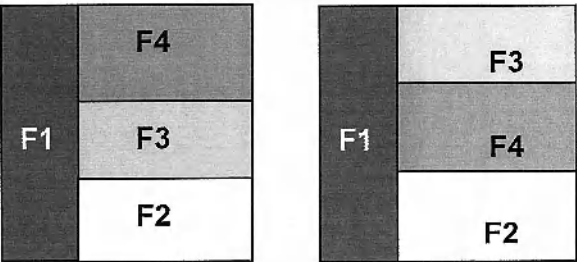


Fig 4.14: Equivalent rocktype rules if F3 and F4 do not appear on the same level.

When one facies disappears, the two facies which should be separated by it come into direct contact. This is obvious when the local rock type rule is shown, but can be forgotten when only the global rocktype rule is given. One consequence is that the relative position of two facies which never appear at the same level is of no importance (see Figure 4.14). What must be taken into account is their relationship with the other facies.

Two facies which are not in contact in the rock type rule can sometimes touch if the proportions (and hence the thresholds) vary sharply between consecutive levels. This is easy to see on the example below, where there are three facies: shale (F1), shaly sandstone (F2) and sandstone (F3). On level 1, the proportions are respectively 60%, 20% and 20%, which gives the thresholds $t_1 = 0.25$ and $t_2 = 0.84$. On level 2, they are 20%, 20% and 60%, and so the thresholds are $t_1 = -0.84$, $t_2 = -0.25$. The gaussian values between -0.25 and 0.25 represent 20% of the histogram. When the value simulated on level 1 is between -0.25 and 0.25 (which corresponds to the facies shale), there is a high probability that on level 2 the simulated gaussian value is located within the same interval. (This of course depends on the variogram model). If that happens the sandstone will be directly in contact with the shale vertically below it.

Table 4.3: Schematic representation of the rocktype rules on two consecutive levels, respecting the ordering in the threshold values

Level 2	Shale	Shaly Sandstones	Sandstones	
Level 1	Shale		Shaly Sandstones	Sandstones

5 VARIOGRAMS AND STRUCTURAL ANALYSIS

Truncated Gaussian Method

This chapter describes how to calculate experimental variograms for the facies and how to fit models to them. As the codes chosen to represent the facies are purely conventional and have no physical meaning (except for their ordering), the only representative numbers we can use to compute experimental variograms are the facies indicators. (The relationship between the facies indicators and the underlying gaussian values is not one-to-one, so we cannot back transform the indicator values to get the values of the gaussians). There is a theoretical relation linking the variograms of the underlying gaussians and those of the indicators. Rather than invert it, we use an indirect iterative procedure to fit a variogram model. But first, we show how to calculate the experimental variograms for the facies indicators.

Experimental variograms and crossvariograms for facies

In Chapter 2, the definition for the indicator variogram of the facies F was given:

$$\begin{aligned}\gamma_F(x, x+h) &= \frac{1}{2} \text{Var}[l_F(x) - l_F(x+h)] \\ &= \frac{1}{2} \left\{ E\left([l_F(x) - l_F(x+h)]^2\right) - (E[l_F(x) - l_F(x+h)])^2 \right\}\end{aligned}$$

If the facies indicators are second order stationary or intrinsic with constant mean, the second term is zero and the formula simplifies to:

$$\gamma_F(x, x+h) = \frac{1}{2} E\left([l_F(x) - l_F(x+h)]^2\right)$$

So the formula for computing the experimental variogram for facies F is:

$$\gamma_F^*(x, x+h) = \frac{1}{2N} \sum_{|x_\alpha - x_\beta| = h} [l_F(x_\alpha) - l_F(x_\beta)]^2$$

where the data points x_α and x_β are separated by the vector h (possibly with a tolerance). The summation is carried out over all the data pairs separated by this distance, and N is the number of pairs. In the stationary case, this is an unbiased

estimator of variogram. In the non-stationary case, it is not but we nevertheless use it because of the difficulties in getting reliable estimates of the mean difference between the two indicators. More information will be given later on how to compute the experimental variogram in the non-stationary case.

In exactly the same way, the cross variogram for facies F_i and F_j is defined as:

$$\gamma_{F_i F_j}(x, x+h) = \frac{1}{2} E \left([1_{F_i}(x) - 1_{F_i}(x+h)] [1_{F_j}(x) - 1_{F_j}(x+h)] \right)$$

Consequently the formula for the corresponding experimental crossvariogram is:

$$\gamma_{F_i F_j}^*(x, x+h) = \frac{1}{2N} \sum_{|x_\alpha - x_\beta| = h} [1_{F_i}(x_\alpha) - 1_{F_i}(x_\beta)] [1_{F_j}(x_\alpha) - 1_{F_j}(x_\beta)]$$

Linking the indicator variograms to the underlying variograms

Firstly we consider the truncated gaussian case. If t_{i-1} and t_i are the thresholds for facies F_i then

$$\begin{aligned} \gamma_{F_i}(x, x+h) = & \frac{1}{2} \{ P[t_{i-1} \leq Z(x) < t_i] + P[t_{i-1} \leq Z(x+h) < t_i] \} \\ & - P[t_{i-1} \leq Z(x) < t_i \text{ and } t_{i-1} \leq Z(x+h) < t_i] \end{aligned}$$

If the correlation function of the gaussian function Z is $\rho(h)$, the indicator variogram is given by:

$$\begin{aligned} \gamma_{F_i}(x, x+h) &= p_{F_i} - \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} g_{\rho(h)}(u, v) \, du \, dv \\ &= p_{F_i} - \frac{1}{2\pi\sqrt{1-\rho(h)^2}} \int_{t_{i-1}}^{t_i} \int_{t_{i-1}}^{t_i} \exp\left(-\frac{u^2 + v^2 - 2\rho(h)uv}{2(1-\rho(h)^2)}\right) du \, dv \end{aligned}$$

where p_{F_i} is the proportion of facies F_i and $g_{\rho(h)}$ is the bigaussian density function. The correlation between the two gaussian variables u and v is $\rho(h)$. Once we have chosen the variogram for the gaussian function, we know $\rho(h)$ and can compute the variogram for each facies using the previous formula.

Cross-variograms

In Chapter 2, we showed that the cross-variogram can be written as

$$\gamma_{F_i F_j}(x, x+h) = -\frac{1}{2} \left\{ E \left[1_{F_i}(x) 1_{F_j}(x+h) \right] + E \left[1_{F_j}(x) 1_{F_i}(x+h) \right] \right\}$$

Consequently, for the stationary truncated gaussian case this becomes:

$$\gamma_{F_i F_j}(x, x+h) = - \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} g_{\rho(h)}(u, v) \, du \, dv$$

Truncated Plurigaussian Method

Variograms

The experimental variograms and cross-variograms are computed in exactly the same way as for the truncated gaussian method. We will give the equations linking the indicator variogram models to those of the gaussians for the case with only two gaussian functions. They can easily be extended to the general case. As for the standard truncated gaussian method, the variogram can be written:

$$\begin{aligned} \gamma[1_{F_i}(x), 1_{F_i}(x+h)] \\ = \frac{1}{2} \left\{ P[1_{F_i}(x) = 1] + P[1_{F_i}(x+h) = 1] - 2P[1_{F_i}(x) = 1 \text{ and } 1_{F_i}(x+h) = 1] \right\} \end{aligned}$$

In the stationary case, we have

$$1_{F_i}(x) = 1 \Leftrightarrow [Z_1(x), Z_2(x)] \in D_i$$

where D_i is the domain in the gaussian space used to define the facies F_i . Consequently:

$$\begin{aligned} \gamma[1_{F_i}(x), 1_{F_i}(x+h)] = \frac{1}{2} \left\{ P[(Z_1(x), Z_2(x)) \in D_i] + P[(Z_1(x+h), Z_2(x+h)) \in D_i] \right. \\ \left. - 2P[(Z_1(x), Z_2(x)) \in D_i \text{ and } (Z_1(x+h), Z_2(x+h)) \in D_i] \right\} \end{aligned}$$

From this we obtain:

$$\gamma[l_{Fi}(x), l_{Fi}(x+h)] = p_{Fi} - \iint_{D_i} \iint_{D_i} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2$$

Here Σ is the covariance matrix for the four gaussian variables $Z_1(x)$, $Z_2(x)$, $Z_1(x+h)$ and $Z_2(x+h)$ and g_{Σ} is the quadrivariate gaussian density. Let ρ be the correlation coefficient between Z_1 and Z_2 . Similarly let $\rho_{Z1}(h)$ and $\rho_{Z2}(h)$ be the covariance functions of the two gaussians Z_1 and Z_2 . With this information we can compute 12 of the 16 terms in the covariance matrix:

$$\Sigma = \begin{pmatrix} 1 & \rho & \rho_{Z1}(h) & ? \\ \rho & 1 & ? & \rho_{Z2}(h) \\ \rho_{Z1}(h) & ? & 1 & \rho \\ ? & \rho_{Z2}(h) & \rho & 1 \end{pmatrix}$$

The other four values indicated by question marks depend on the coregionalization model between Z_1 and Z_2 . A wide choice of coregionalization models is available. For example, we can use the standard linear model where Z_1 and Z_2 are linear combinations of independent factors. If $Y_1(x)$ and $Y_2(x)$ are independent $N(0,1)$ gaussian functions having the covariances $\rho_{Y1}(h)$ and $\rho_{Y2}(h)$, we can define $Z_1(x)$ and $Z_2(x)$ as

$$Z_1(x) = Y_1(x) \quad \text{and} \quad Z_2(x) = \rho Y_1(x) + \sqrt{1-\rho^2} Y_2(x)$$

where

$$\rho_{Y1}(h) = \rho_{Z1}(h) \quad \text{and} \quad \rho_{Y2}(h) = \rho^2 \rho_{Z1}(h) + (1-\rho^2) \rho_{Z2}(h)$$

Then the covariance matrix Σ is

$$\Sigma = \begin{pmatrix} 1 & \rho & \rho_{Z1}(h) & \rho \rho_{Z1}(h) \\ \rho & 1 & \rho \rho_{Z1}(h) & \rho_{Z2}(h) \\ \rho_{Z1}(h) & \rho \rho_{Z1}(h) & 1 & \rho \\ \rho \rho_{Z1}(h) & \rho_{Z2}(h) & \rho & 1 \end{pmatrix}$$

Alternatively Z_2 could be the derivative of Z_1 (in a given direction) or it could equal the first gaussian shifted by a vector "a". That is,

$$Z_2(x) = \frac{\partial Z_1(x)}{\partial x_i} \quad \text{or} \quad Z_2(x) = Z_1(x - a)$$

Combinations of these are also possible. For example, one model that covers quite a wide range of cases is:

$$\begin{cases} Z_1(x) = Y_1(x) \\ Z_2(x) = \lambda_1 Y_1(x+a) + \lambda_2 Y_2(x) \end{cases}$$

If we want a specific correlation ρ between Z_1 and Z_2 , then:

$$\lambda_1 = \frac{\rho}{\rho_{Z1}(a)}, \quad \lambda_2 = \pm \sqrt{1 - \frac{\rho^2}{\rho_{Z1}(a)^2}}$$

In that case, the covariance matrix Σ is:

$$\Sigma = \begin{pmatrix} 1 & \rho & \rho_{Z1}(h) & \frac{\rho \rho_{Z1}(h+a)}{\rho_{Z1}(a)} \\ \rho & 1 & \frac{\rho \rho_{Z1}(h-a)}{\rho_{Z1}(a)} & \rho_{Z2}(h) \\ \rho_{Z1}(h) & \frac{\rho \rho_{Z1}(h-a)}{\rho_{Z1}(a)} & 1 & \rho \\ \frac{\rho \rho_{Z1}(h+a)}{\rho_{Z1}(a)} & \rho_{Z2}(h) & \rho & 1 \end{pmatrix}$$

where

$$\rho_{Z2}(h) = \frac{\rho^2}{\rho_{Z1}^2(a)} \rho_{Z1}(h) + \left[1 - \frac{\rho^2}{\rho_{Z1}^2(a)} \right] \rho_{Y2}(h)$$

As before, $\rho_{Y2}(h)$ is the covariance function of the second gaussian Y_2 .

When the domain D_i is a rectangle defined by the thresholds t_{i1} and t_{i2} for the gaussian function Z_1 , and s_{i1} and s_{i2} for the gaussian function Z_2 , the equation becomes:

$$\gamma[1_{Fi}(x), 1_{Fi}(x+h)] = p_{Fi} - \int_{t_{i1}}^{t_{i2}} \int_{s_{i1}}^{s_{i2}} \int_{t_{i1}}^{t_{i2}} \int_{s_{i1}}^{s_{i2}} g_{\Sigma}(u_1, u_2, v_1, v_2) \, du_1 \, du_2 \, dv_1 \, dv_2$$

Knowing the proportions, the thresholds and the covariance matrix of Z_1 and Z_2 , it is easy to compute (and plot) the indicator variograms.

Cross-variograms

We saw in Chapter 2 that the cross variogram model can be written:

$$\gamma_{F_i F_j}(x, x+h) = -\frac{1}{2} \left\{ E \left[1_{F_i}(x) 1_{F_j}(x+h) \right] + E \left[1_{F_j}(x) 1_{F_i}(x+h) \right] \right\}$$

In the stationary case, this gives:

$$\begin{aligned} \gamma \left[1_{F_i}(x), 1_{F_j}(x+h) \right] &= \\ -\frac{1}{2} &\left\{ \iint_{D_i} \iint_{D_j} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2 + \iint_{D_i} \iint_{D_j} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2 \right\} \\ &= -\iint_{D_i} \iint_{D_j} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2 \end{aligned}$$

Generalisation to the non stationary case

Experimental variograms

Non-stationarity makes it more difficult to compute the experimental variograms. To understand this, we come back to the formula (for the stationary and intrinsic cases):

$$\begin{aligned} \gamma \left[1_{F_i}(x), 1_{F_i}(x+h) \right] &= \frac{1}{2} \text{Var} \left[1_{F_i}(x) - 1_{F_i}(x+h) \right] \\ &= \frac{1}{2} E \left(\left[1_{F_i}(x) - 1_{F_i}(x+h) \right]^2 \right) \end{aligned}$$

This expectation should be interpreted as the average of all possible values of $[1_{F_i}(x) - 1_{F_i}(x+h)]^2$ over all realisations of the random function. As we have only one realisation, we replace the average over all the realisations by the average over all the data pairs (separated by the distance h). In general we are allowed to do that if the random function is stationary and ergodic. However in our model the underlying gaussians are always stationary, nonstationarity is introduced by allowing the thresholds to vary. This is why it is meaningful to use the noncentered variogram:

$$\gamma_{\text{exp}} \left(1_{F_i}(x_{\alpha}), 1_{F_i}(x_{\beta}) \right) = \frac{1}{2N} \sum_{|x_{\alpha} - x_{\beta}|=h} \left[1_{F_i}(x_{\alpha}) - 1_{F_i}(x_{\beta}) \right]^2$$

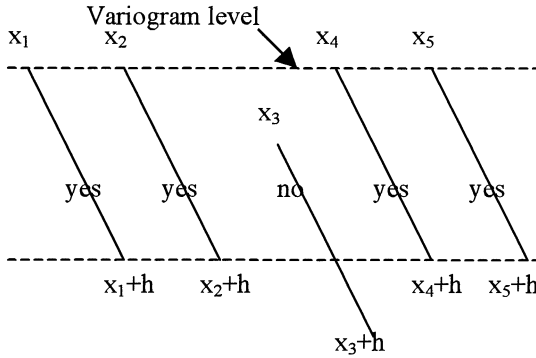


Fig 5.1: Variogram level

Although the proportions vary along the vertical direction (i.e. vertical after flattening), they can often be considered as constant in the horizontal plane. We call this vertical non-stationarity. In this case, the expectations

$$E[I_{Fi}(x) - I_{Fi}(x+h)] \quad \text{and} \quad E([I_{Fi}(x) - I_{Fi}(x+h)]^2)$$

are constant when the point x moves horizontally. So we can estimate the expectation by an average in space, provided that we average pairs where the first point x stays in the same horizontal plane (Figure 5.1). In this way, we obtain an experimental variogram for each horizontal level, which is called the variogram level.

Remarks

In this case, the indicators cannot be isotropic, and so we have to compute variograms in different directions. In theory we should fit all these variograms simultaneously (using the same parameters for the gaussian function), but this has two disadvantages:

- in practice, we usually have hundreds of levels, and it takes too much time to check the fit on all these levels;
- if the data come from only a few wells, the statistical fluctuations in the variograms computed level by level are too important, and we cannot fit these variograms at all.

These are the reasons why we average all the variogram levels rather than fitting a model to them all directly. It should be kept in mind that this average is no longer a variogram but a mean variogram. It is still possible to fit it, so we have to find a model for it.

Another point to note is that when all the wells cross all the levels vertically, there is no difference between this average level by level and the average in the whole space. Otherwise, these averages are different.

For the cross variograms, the basic formula

$$\gamma_{\text{exp}}(l_{F_i}(x), l_{F_j}(x+h)) = \frac{1}{2N} \sum_{|x_\alpha - x_\beta| = h} [l_{F_i}(x_\alpha) - l_{F_i}(x_\beta)] [l_{F_j}(x_\alpha) - l_{F_j}(x_\beta)]$$

is exact for the stationary case, but should be centred for the non-stationary case. For the reasons given above, we have chosen not to centre it because we use the relation with the noncentered covariance. Depending on the type of (non-) stationarity, the average will be calculated over the whole space, or level by level, as for the simple variograms. When the cross variograms are computed level by level, we average them again to obtain a mean cross variogram.

Indicator variogram model (truncated gaussian model)

We now present the variogram model for the non-stationary case. Earlier in the chapter we saw that for the stationary case, the variogram for the indicator function $l_{F_i}(x)$ is:

$$\gamma_{F_i}(x, x+h) = \frac{1}{2} \{P[t_{i-1} \leq Z(x) < t_i] + P[t_{i-1} \leq Z(x+h) < t_i]\} \\ - P[t_{i-1} \leq Z(x) < t_i \text{ and } t_{i-1} \leq Z(x+h) < t_i]$$

As the thresholds now vary with location, the equivalent formula is:

$$\gamma[l_{F_i}(x), l_{F_i}(x+h)] = \\ \frac{1}{2} \{P[t_{i-1}(x) \leq Z(x) < t_i(x)] + P[t_{i-1}(x+h) \leq Z(x+h) < t_i(x+h)] \\ - 2P[t_{i-1}(x) \leq Z(x) < t_i(x) \text{ and } t_{i-1}(x+h) \leq Z(x+h) < t_i(x+h)]\}$$

where t_{i-1} and t_i are the thresholds for the facies F_i . This can be written as:

$$\gamma[l_{F_i}(x), l_{F_i}(x+h)] = \frac{1}{2} \left\{ p_{F_i}(x) + p_{F_i}(x+h) - \int_{t_{i-1}(x)}^{t_i(x)} \int_{t_{i-1}(x+h)}^{t_i(x+h)} g_{\rho(h)}(u, v) \, du \, dv \right\} \\ = \frac{1}{2} \left\{ p_{F_i}(x) + p_{F_i}(x+h) - \frac{1}{\pi \sqrt{1-\rho(h)^2}} \int_{t_{i-1}(x)}^{t_i(x)} \int_{t_{i-1}(x+h)}^{t_i(x+h)} \exp \left(-\frac{u^2 + v^2 - 2\rho(h)uv}{2(1-\rho(h)^2)} \right) du \, dv \right\}$$

where $p_{F_i}(x)$ and $p_{F_i}(x+h)$ are the proportions of facies F_i at points x and $x+h$, and $g_{\rho(h)}$ is the same bigaussian density function as in the stationary case. The cross variogram model is given by:

$$\gamma[1_{F_i}(x), 1_{F_j}(x+h)] = -\frac{1}{2} \left\{ \int_{t_{i-1}(x)}^{t_i(x)} \int_{t_{j-1}(x+h)}^{t_j(x+h)} g_{\rho(h)}(u, v) du dv + \int_{t_{i-1}(x+h)}^{t_i(x+h)} \int_{t_{j-1}(x)}^{t_j(x)} g_{\rho(h)}(u, v) du dv \right\}$$

Variogram and cross-variogram model (truncated plurigaussian method)

The variogram equation is now:

$$\begin{aligned} & \gamma[1_{F_i}(x), 1_{F_i}(x+h)] \\ &= \frac{1}{2} \left\{ P[(Z_1(x), Z_2(x)) \in D_i(x)] + P[(Z_1(x+h), Z_2(x+h)) \in D_i(x+h)] \right. \\ & \quad \left. - 2P[(Z_1(x), Z_2(x)) \in D_i(x) \text{ and } (Z_1(x+h), Z_2(x+h)) \in D_i(x+h)] \right\} \end{aligned}$$

That is,

$$\gamma[1_{F_i}(x), 1_{F_i}(x+h)] = \frac{1}{2} \left[p_{F_i}(x) + p_{F_i}(x+h) - 2 \iint_{D_i(x)} \iint_{D_i(x+h)} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2 \right]$$

where Σ is the same covariance matrix as in the stationary case.

When the domain $D_i(x)$ is a rectangle defined by the thresholds $t_{i1}(x)$ and $t_{i2}(x)$ for the gaussian function Z_1 , and $s_{i1}(x)$ and $s_{i2}(x)$ for the gaussian function Z_2 , the equation becomes:

$$\gamma(1_{F_i}(x), 1_{F_i}(x+h)) = \frac{1}{2} \left[p_{F_i}(x) + p_{F_i}(x+h) - 2 \int_{t_{i1}(x)}^{t_{i2}(x)} \int_{s_{i1}(x)}^{s_{i2}(x)} \int_{t_{i1}(x+h)}^{t_{i2}(x+h)} \int_{s_{i1}(x+h)}^{s_{i2}(x+h)} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2 \right]$$

The cross variogram model is given by:

$$\begin{aligned} \gamma[(1_{F_i}(x), 1_{F_j}(x+h))] &= \frac{1}{2} \left\{ \iint_{D_i(x)} \iint_{D_j(x+h)} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2 \right. \\ & \quad \left. + \iint_{D_i(x+h)} \iint_{D_j(x)} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2 \right\} \end{aligned}$$

Comparing variogram models for indicators and gaussian functions

Only the stationary case will be discussed here in detail. In the non stationary case, the proportions vary and this has various effects on the individual variograms. The impact on the mean variogram is much more complex as can be seen from the examples given in the section on “variogram fitting”.

Sill of the indicator model

In the stationary case, the sill of the indicator variogram for facies F_i is equal to $p_{Fi}(1 - p_{Fi})$. As there is no link between the sill of the indicator variograms and that of the gaussians, we can choose any sill for the latter and the standard normal $N(0,1)$ is the simplest choice. In the case of non stationarity, we can no longer speak of a sill. The variogram shape for long distances can be rather complicated; it need not stabilize. This often happens for vertical variograms. This long distance shape is completely controlled by the proportions.

Shape of the indicator model

We are now going to see which features of the variogram of the gaussian functions reappear in the indicator variogram. Figure 5.2 shows a gaussian variogram with a range of 50, and the corresponding indicator variogram for the truncated gaussian model. We already know (see Chapter 2) that the indicator variogram cannot have a zero derivative at the origin. We see here that the indicator variogram is linear near the origin. The curvature near the origin which is characteristic of the gaussian variogram has disappeared.

Now look at Figure 5.3 which shows an exponential variogram with the same practical range of 50, and the corresponding indicator variogram (same proportion and same rock-type rule as in Figure 5.2). In this case, the overall shape of the indicator variogram for the exponential is closer to that of the exponential than for the gaussian variogram in Figure 5.2. In general, there is less difference between the shapes of the indicator variograms than between those of the underlying gaussians. So it is not possible to determine the type of model or its parameters from the indicator variogram. In top of this as will be seen in the next section, changes in the proportions have are marked impact on their shape. So the choice between a gaussian variogram and an exponential is made by studying the consequences of this choice on the resulting simulations, especially the continuity and regularity of each facies, rather than the shape of the variogram.

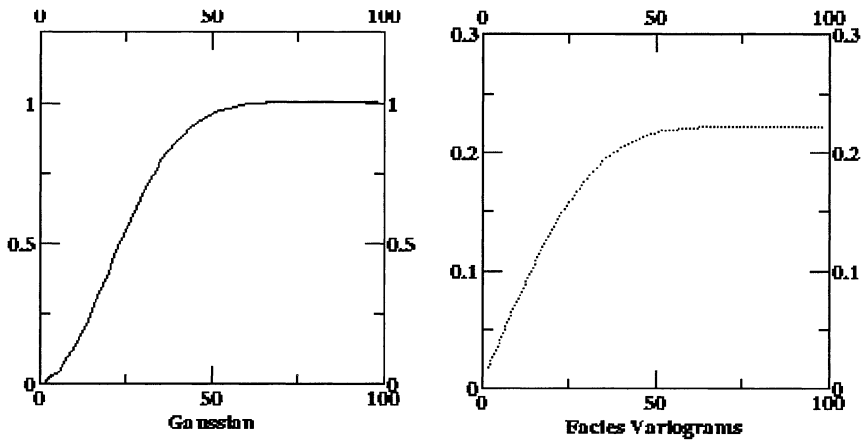


Fig 5.2: Gaussian variogram (above) and the corresponding indicator variogram

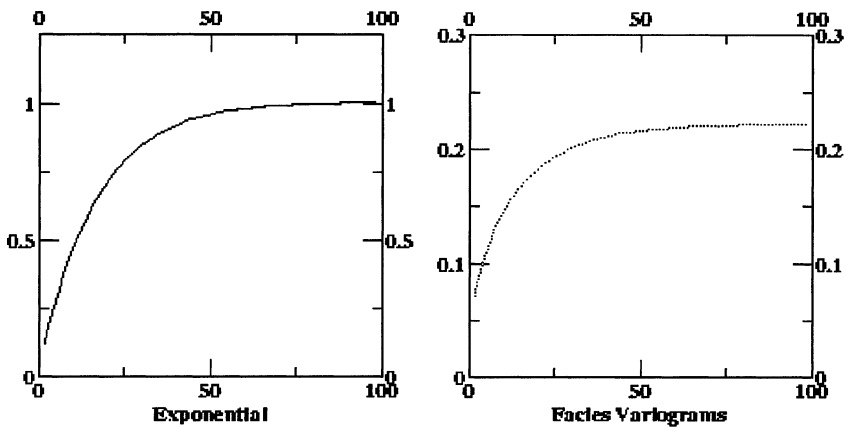


Fig 5.3: Exponential variogram (above) and the corresponding indicator variogram

Practical range of the indicator model

The previous figures correspond to the facies F_1 in the rocktype rule given in Figure 5.4. Figure 5.5 shows the indicator variogram for facies F_2 using the same exponential variogram and the same proportion (33%) as in Figure 5.3. These figures clearly show that the practical range of the indicator variogram depends on the position of the corresponding facies in the rocktype rule as well as the practical range of the gaussian function variogram.

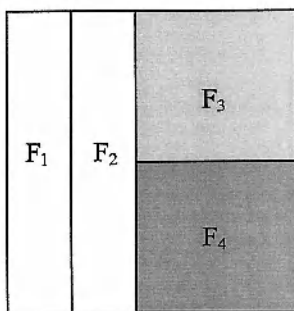


Fig 5.4: Rock-type rule

Figures 5.6 and 5.7 show what happens to this indicator variogram when the proportion is increased from 33% to 50% or decreased to 15%, respectively. As expected in the first case, the sill increases (a proportion of 50% gives the maximum sill in the stationary case). We see that the practical range has increased too. In the latter case, the practical range decreases as does the sill. This shows that in the same way that we cannot choose the variogram model for the gaussian function just by looking at the shape of the indicator variogram, we cannot deduce its range directly from that of the indicator variogram.

Anisotropies

If the gaussian function model has a geometric anisotropy, the isovalue lines in the variogram map are ellipses. Figure 5.8 shows the isovalue line for the indicator variogram corresponding to 90% of the sill for the rocktype rule depicted in Figure 5.4. The two gaussian functions have exponential variograms with a geometric anisotropy with ranges of 25 and 100. The long axis of continuity is oriented north-south for the first gaussian function Z_1 , and east-west for the other gaussian function Z_2 . These are independent.

For facies F_1 and F_2 , the isoline curves are still ellipses, as the truncation is carried out on only one gaussian function. The shape is more complicated for facies F_3 and F_4 (as they have the same proportions and are symmetrical in the rocktype rule, their variograms are the same). The result looks like a combination of the two ellipses in perpendicular directions.

Figure 5.9 was obtained in the same way as Figure 5.8 except that the anisotropy axes have been rotated for Z_2 . The long axis is now north-east/south-west. We immediately see that the isolines for facies F_1 and F_2 have not changed, as the truncation for these facies is made only on Z_1 which has not changed. However, the isolines for facies F_3 and F_4 have changed markedly. Here the ellipses with both sets of main anisotropy axes are superposed.

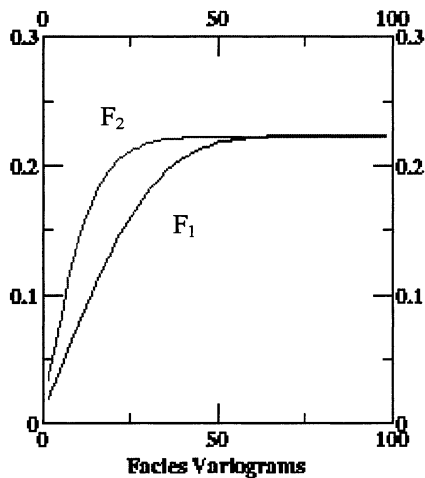
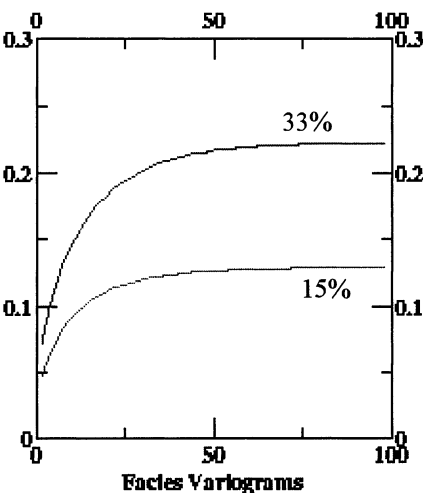
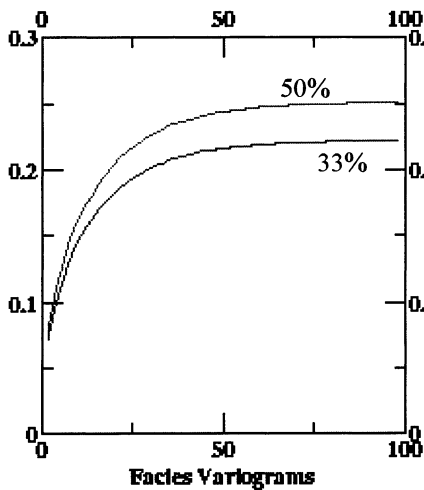


Fig 5.5 : Comparison of the indicator variograms for facies F_1 and F_2



Figs5.6 (left) and 5.7 (right): Comparison of the indicator variograms for facies F_1 and F_2

Variogram fitting

As the principle behind variogram fitting is the same in Truncated Gaussian case and in Plurigaussian case, both cases will be presented together.

Stationary case

We now use the theoretical relation between the variograms of the underlying gaussian variables and the facies indicators to fit models to them. We first choose a model for the gaussian function, to calculate the equivalent indicator model and to plot this. The key point in the variogram fitting is to have good quality experimental variograms in order to compare them with the model. The examples shown in Figures 5.8 and 5.9 show what is required (at least 4 directions are required to define the variogram model in a given plane). If the data do not allow us to compute all these experimentally, we need additional information, for example from the geologist.

In the plurigaussian case, if some facies are defined by truncating on one gaussian function only, it is better to fit them first. Once the values of the parameters for this gaussian have been fitted, then the facies using the second gaussian function can be fitted.

Figure 5.10 shows an example of fitting a vertical variogram in the stationary case. Note that this vertical stationarity is not common. In this particular case, experimental vertical proportions show a slight vertical non stationarity, which can be interpreted as statistical fluctuations as there is no clear trend.

Non stationary case

The variogram fitting is done the same way as in the stationary case. We have to average the theoretical variograms in the same way as the experimental variograms. In the non stationary case, the long distance variations are completely controlled by the proportions. Figure 5.11 shows the same experimental variogram as in Figure 5.10, but the slight vertical non stationarity has been taken into account during the fitting process.

Note that the long distances are better fitted here (see black circle). Even in the stationary case, it can be easier to fit the variogram with a non stationary approach taking into account the experimental fluctuations of the proportions. The stationary proportions will be used for the simulation, together with the fitted variogram model. If the case is clearly non stationary, we can choose to smooth the proportions for the simulation in order to reduce their experimental fluctuations, but the fitting must be done with the raw proportions.

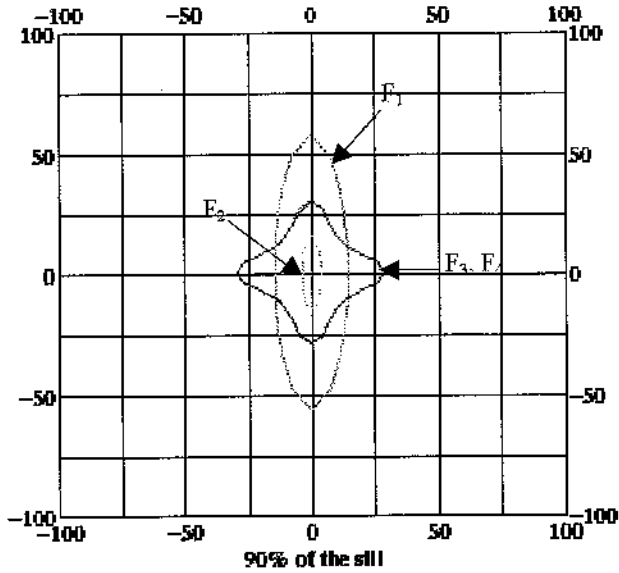


Fig 5.8: For the facies of Figure 5.4, isolines representing $\gamma(h)=90\%$ of the sill. Long anisotropy axes of the gaussian random functions are perpendicular

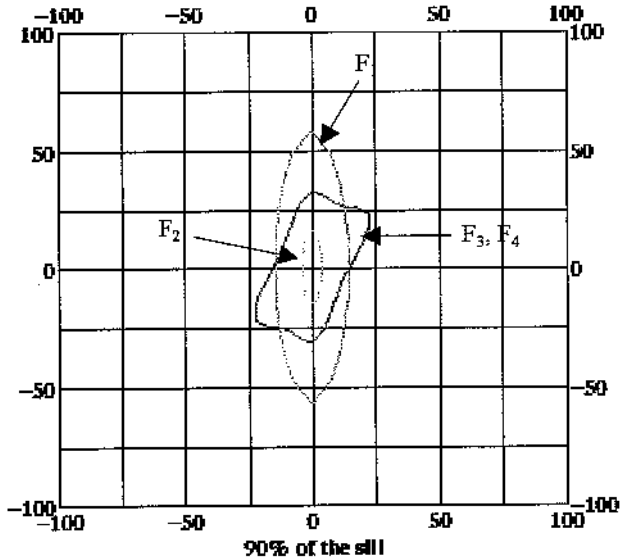


Fig 5.9: For the facies of Figure 5.4, isolines representing $\gamma(h)=90\%$ of the sill. Long anisotropy axes of the gaussian random functions make a 45° angle

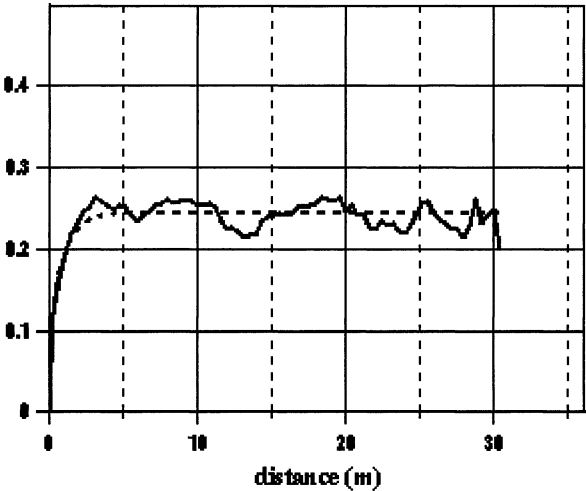


Fig 5.10: Example of stationary variogram fitting

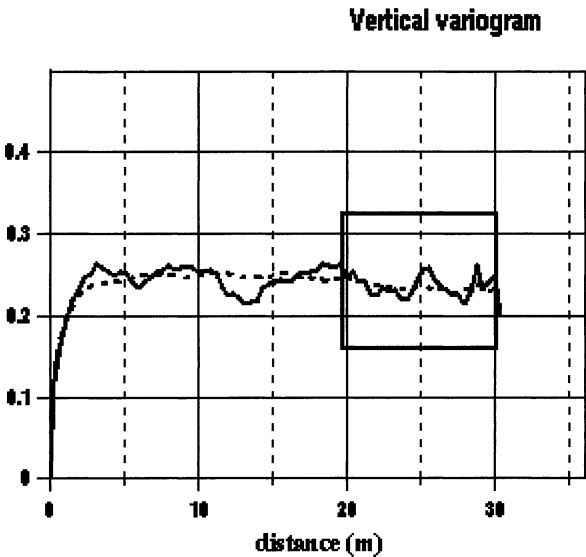


Fig 5.11: Example of use of the raw varying proportions in a stationary case

Figure 5.12 shows a variogram fitting in a non stationary case. Its shape is completely irregular, due to the marked variability in the proportions. Figure 5.13 shows the variogram model obtained using the same ranges but computed using smoothed proportions.

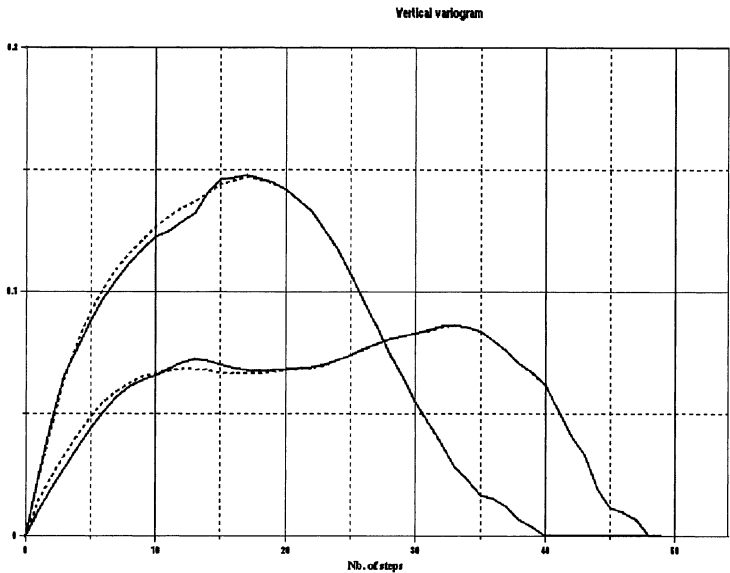


Fig 5.12: Non stationary variogram fitting using raw proportions

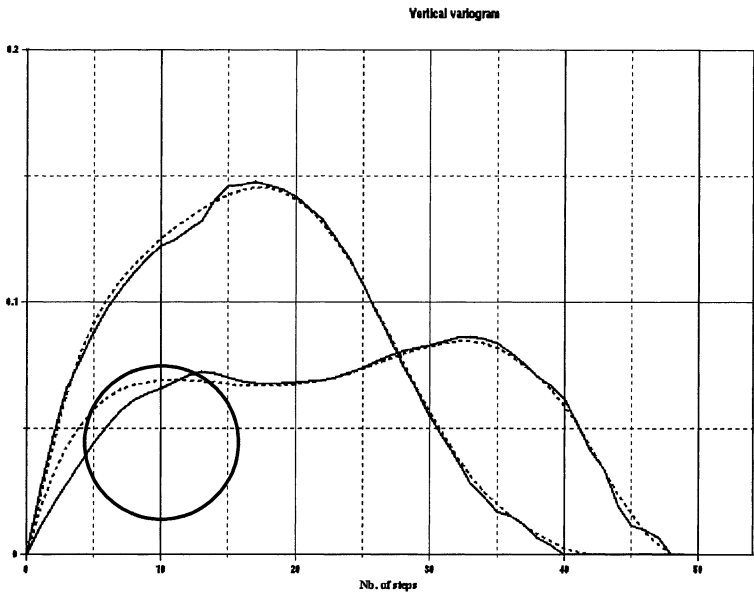


Fig 5.13: Non stationary variogram computed using smoothed proportions.

Smoothing has produced quite marked changes particularly near the origin. In some cases it has been impossible to get a good fit after modifying the proportions

Transition probabilities

We saw earlier that transition probabilities are simple to compute when we know the non-centred covariances of the indicator functions :

$$P[x + h \in F_j \mid x \in F_i] = \frac{C_{ij}(x, x + h)}{p_{Fi}(x)}$$

and

$$P[x + h \in F_j \mid x \in F_i \text{ and } x + h \notin F_i] = \frac{C_{ij}(x, x + h)}{p_{Fi}(x) - C_{ii}(x, x + h)}$$

These non centred covariances are easy to compute for the plurigaussian method:

$$C_{ij}(x, x + h) = \iint_{D_i(x)} \iint_{D_j(x+h)} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2$$

and

$$C_{ii}(x, x + h) = \iint_{D_i(x)} \iint_{D_i(x+h)} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2$$

This gives the following transition probabilities :

$$P[x + h \in F_j \mid x \in F_i] = \frac{\iint_{D_i(x)} \iint_{D_j(x+h)} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2}{p_{Fi}(x)}$$

and

$$\begin{aligned} & P[x + h \in F_j \mid x \in F_i \text{ and } x + h \notin F_i] \\ &= \frac{\iint_{D_i(x)} \iint_{D_j(x+h)} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2}{p_{Fi}(x) - \iint_{D_i(x)} \iint_{D_i(x+h)} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2} \end{aligned}$$

We see that these probabilities provide the same information as the non centred covariances. They differ from variograms because they are not symmetrical in general. For example, these probabilities are not symmetrical if the proportions vary in the direction of the vector h , even if the gaussian functions have symmetrical covariances:

$$\iint_{D_i(x)} \iint_{D_j(x+h)} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2$$

$$\neq \iint_{D_j(x)} \iint_{D_i(x+h)} g_{\Sigma}(u_1, u_2, v_1, v_2) du_1 du_2 dv_1 dv_2$$

These probabilities can also help to infer the value of the correlation ρ between the gaussian functions, especially the second one:

$$P[x+h \in F_j \mid x \in F_i \text{ and } x+h \notin F_i] = \frac{C_{ij}(x, x+h)}{p_{F_i}(x) - C_{ii}(x, x+h)}$$

Finally, they are easy to compute experimentally in the stationary case provided that there are enough data points to be able to compute reliable statistics. For example, we can only compute the second one if we have enough points with the facies F_i at level x , and enough points which are not in facies F_i at level $x+h$.

6 GIBBS SAMPLER

Our ultimate objective is to simulate a gaussian random function with a specified covariance structure, given the observed lithotypes (facies) at sample points. As the lithotypes are known at these points, the corresponding gaussian variables must lie in certain intervals but their values are not known. So we need to simulate a gaussian random function subject to interval constraints. A two step procedure is used to do this. Firstly a Gibbs sampler is used to generate gaussian values at sample points that have the right covariance and belong to the right intervals. Once we have this set of point values, any method for conditionally simulating gaussian random functions can be used; for example, turning bands together with a conditioning kriging, sequential gaussian simulations, LU decomposition, etc. See Chiles and Delfiner (1999), Lantuéjoul (2002) or Deutch and Journel (1992). As these techniques are well known, we will not dwell on them here.

One difficulty in writing this chapter is that readers come from widely differing mathematical backgrounds. Those who are more interested in the practical applications would prefer a user-friendly approach whereas the mathematically inclined want proofs of the results, particularly on the convergence of the Gibbs sampler. To cater for both points of view, we have divided this chapter into two parts: the first half gives a “maths-lite” presentation, the second half provides the theory.

The first half starts out by explaining why a two step procedure is required, then gives examples to illustrate how a Gibbs sampler works and what is meant by convergence. The second part begins by reviewing the key properties of gaussian distributions that will be required later in the proofs. The next section proves that the two step procedure described in the first part of the chapter is effectively equivalent to simulating subject to interval constraints. Then the kernel operator for the Gibbs sampler is determined. As the Gibbs sampler involves an iterative procedure, it is important to prove that this converges. The last part of the chapter is devoted to proving this.

Part 1: Why we need a two step simulation procedure

The aim of this section is to highlight the difficulties of simulating gaussian random functions subject to interval constraints. To do this we consider three simple cases where there are only two points:

1. with no constraints,
2. with interval constraints on one variable and
3. with interval constraints on both variables.

In the first case, the conditional distribution of $Z(x)$ given $Z(y)$ turns out to be a gaussian distribution but this is no longer true in the other two cases. The conditional distributions are merely proportional to gaussians.

Simulating $Z(x)$ and $Z(y)$ when there are no constraints

Consider two gaussian variables $Z(x)$ and $Z(y)$ with a correlation coefficient, ρ . In order to simulate $Z(x)$ given $Z(y)$ we need to know its conditional distribution which can be deduced from the joint distribution of the two variables:

$$g(u, v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{(u^2 + v^2 - 2\rho uv)}{2(1-\rho^2)}\right\}$$

where u and v represent $z(x)$ and $z(y)$ respectively. This can be rewritten as

$$g(u, v) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(u - \rho v)^2}{2\sigma^2}\right\} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{v^2}{2}\right\}$$

where $\sigma^2 = 1 - \rho^2$. The second term is just the marginal distribution of $Z(y)$. The first is the conditional distribution that we are looking for. It is clearly a gaussian distribution with mean, ρv , and variance σ^2 . The equation can be written as

$$g(u, v) = g_v(u)g(v)$$

This is equivalent to the well-known decomposition:

$$Z(x) = \rho Z(y) + \sigma R(x)$$

where $R(x)$ is a $N(0,1)$ residual that is independent of $Z(y)$. If we estimate $Z(x)$ given $z(y)$, the simple kriging weight equals ρ , and the SK variance is $\sigma^2 = 1 - \rho^2$.

To simulate pairs of values of $Z(x)$ and $Z(y)$, we first draw two independent $N(0,1)$ values for $Z(y)$ and $R(x)$, then we substitute them into the decomposition formula to get $Z(x)$. Alternatively we could say that we draw one realisation v of a $N(0,1)$ variable for $Z(y)$, followed a $N(\rho v, \sigma^2)$ variable for $Z(x)$. In that case, we use the marginal distribution of $Z(y)$ to draw a realisation of it, then the conditional distribution of $Z(x)$ given $Z(y)$ to draw the other value directly. This is only possible because the form of the conditional distribution is so simple.

Simulating $Z(x)$ and $Z(y)$ when $Z(y)$ belongs to an interval

In this case $Z(y)$ is known to lie in a specified interval, I , and we want to simulate the pair of variables, $Z(x)$ and $Z(y)$, given that $Z(y)$ lies in that interval. The joint density of the two variables is now

$$h(u, v) = k g(u, v) l_I(v)$$

where $1_I(v)$ is the indicator function for the interval I and k is the normation factor required to ensure that the integral of $h(u,v)$ sums to 1. So the joint density is

$$h(u, v) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{(u - \rho v)^2}{\sigma^2}\right\} \frac{k}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{v^2}{\sigma^2}\right\} 1_I(v)$$

This can be written as

$$h(u, v) = h_v(u) h(v)$$

It is clear that $h(v)$ is the marginal distribution of a gaussian variable $Z(y)$ restricted to the interval I . To show that it is the marginal distribution, we just have to prove that

$$\int_{\mathfrak{R}} h_v(u) du = 1$$

Integrating $h(u,v)$ with respect to u gives

$$\int_{\mathfrak{R}} h(u, v) du = \int_{\mathfrak{R}} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(u - \rho v)^2}{2\sigma^2}\right\} du \frac{k}{\sqrt{2\pi}} \exp\left\{-\frac{v^2}{2}\right\} 1_I(v)$$

As the first term on the right hand side is just the integral of a $N(\rho v, \sigma^2)$ variable, it equals 1, which gives us the required result. In order to simulate these we use the same interpretation as before: first simulate $Z(y)$ in the interval I then simulate $Z(x)$ given that $Z(y) \in I$. The first simulation is just a truncated gaussian; the second one corresponds to simulating an independent $N(\rho v, \sigma^2)$ variable. That is, we are still using the classical decomposition.

But there is a fundamental change in the marginal distribution of $Z(x)$. To see this, we integrate the joint density with respect to v .

$$\begin{aligned} \int_{\mathfrak{R}} h(u, v) dv &= \int_{\mathfrak{R}} k g(u, v) 1_I(v) dv \\ &= \int_{\mathfrak{R}} \frac{1}{\sigma\sqrt{2\pi}} 1_I(v) \exp\left\{-\frac{(v - \rho u)^2}{2\sigma^2}\right\} dv \times \frac{k}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} \end{aligned}$$

If we let $t = v - \rho u$, then $v \in I \Leftrightarrow t \in I - \rho u$ and the first term on the right becomes

$$\int_{\mathfrak{R}} \frac{1}{\sigma\sqrt{2\pi}} 1_{I - \rho u}(t) \exp\left\{-\frac{t^2}{2\sigma^2}\right\} dt = E[1_{I - \rho u}]$$

$$\text{As } E[1_{I - \rho u}] = P_{\sigma}[I - \rho u]$$

$$\int_{\mathfrak{R}} h(u, v) dv = k \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{u^2}{2}\right\} \times P_{\sigma}(I - \rho u)$$

This shows that the marginal density of $Z(x)$ given that $Z(y) \in I$, is no longer gaussian. It is proportional to a gaussian density but is multiplied by the probability that the first variable lies in the interval, $I - \rho u$. To illustrate the impact of this change, we have plotted this distribution for two intervals: $[-0.5, 0.5]$ and $[2, 3]$, and for two different correlation factors, -0.5 and 0.8 . Figures 6.1 and 6.2 present the resulting curves.

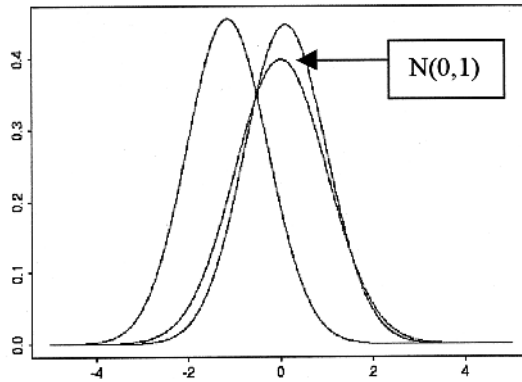


Fig 6.1: Probability densities functions for the marginal distributions for the two intervals, for the case where $\rho = -0.5$ together with the $N(0,1)$ density for comparison purposes

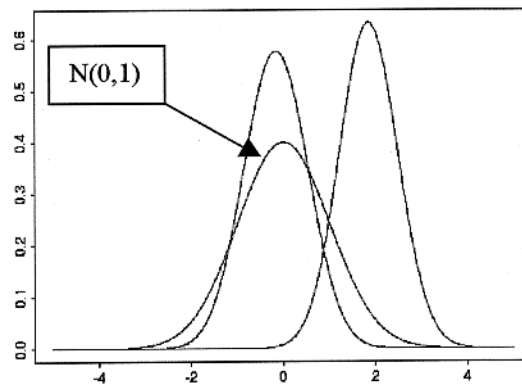


Fig 6.2: Probability densities functions for the marginal distributions for the two intervals, for the case where $\rho = 0.8$ together with the $N(0,1)$ density for comparison purposes

Simulating $Z(x)$ and $Z(y)$ when both belong to intervals

Now suppose that $Y(x)$ belongs to I_1 and $Y(y)$ belongs to I_2 . Their joint density is

$$h(u, v) = k g(u, v) 1_{I_1}(u) 1_{I_2}(v)$$

where k is the appropriate normation factor.

$$h(u, v) = \frac{k}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(u-\rho v)^2}{2\sigma^2}\right\} 1_{I_1}(u) \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{v^2}{2}\right\} 1_{I_2}(v)$$

As expected, the marginal distributions are no longer gaussian or even truncated gaussian. Integrating with respect to u gives

$$\int_{\mathfrak{R}} h(u, v) du \propto \frac{1}{\sqrt{2\pi}} P_{\sigma_2}(I_1 - \rho \ v) \exp\left\{\frac{-v^2}{2}\right\} 1_{I_2}(v) \quad (1)$$

Similarly integrating with respect to v gives

$$\int_{\mathfrak{R}} h(u, v) dv \propto \frac{1}{\sqrt{2\pi}} P_{\sigma_1}(I_2 - \rho \ u) \exp\left\{\frac{-u^2}{2}\right\} 1_{I_1}(u) \quad (2)$$

Because of the increasing difficult in simulating these distributions directly as the number of interval constraints increases, we rapidly reach the point where direct simulation is no longer practicable and we have to resort to an indirect approach such as the Gibbs sampler.

These examples suggest the idea of using a two step procedure to conditionally simulate gaussian values when some of them are constrained to lie in specified intervals. The two steps are:

1. Generating gaussian values at data points, in the prescribed intervals and with the right covariance structure, and then
2. Using any algorithm for conditionally simulating gaussian random functions given the values generated in step (1).

We have given an ad hoc justification of why a two step procedure is needed. A formal proof is given in the second half of the chapter.

Direct simulation using an acceptance/rejection procedure

Having seen that the first step is to generate a set of gaussian values at sample points, the next question is how should this be done. We might be tempted to try an acceptance/rejection procedure, by generating gaussian values with the right

covariance structure and rejecting those lying outside the specified intervals. If only ten or twenty samples were available this could be done directly using an LU decomposition (i.e. a Cholesky decomposition of the covariance matrix). The LU decomposition could be carried out for much large matrices (up to about 500 x 500). The limiting factor in the procedure is the rate of rejection. For example, suppose that ten samples were available and that there were only two facies each present 50% of the time. Then if there was no spatial correlation (pure nugget effect), the probability of getting all ten values in the right intervals would be 1 in 2^{10} , that is, about 1 in 1,000. This procedure becomes prohibitively slow as the number of samples increases. As there are usually hundreds or thousands of data in mining and petroleum applications, another approach is needed. This is why we have to resort to more complicated methods.

Gibbs sampler

Statisticians routinely use iterative methods based on Markov chain Monte Carlo simulations (MCMC, for short) for sampling complicated distributions and for estimating parameter values. The best known are the Hastings-Metropolis algorithm and the Gibbs sampler. The latter is a particular case of the Hastings-Metropolis method. See Meyn and Tweedie (1993), Cowles and Carlin (1996) and Robert (1996) for information on these methods. Freulon (1992) and Freulon and de Fouquet (1993) adapted the Gibbs sampler to truncated gaussian simulations. To introduce this technique, we present an example to show how this method works and to illustrate the concept of convergence.

Four sample example

Suppose that there are only two lithotypes F_1 and F_2 , and that they are present in equal proportions. So it is natural to use a zero threshold to separate them. Negative gaussian values correspond to F_1 ; positive ones, to F_2 . Figure 6.3 shows a simplified well or drill hole containing four samples, with the top two belonging to F_1 and the other two belonging to F_2 . The gaussian values assigned to the top samples must be negative, the others have to be positive.

The procedure relies on the standard decomposition of the gaussian random function into its simple kriging estimate and an orthogonal residual.

$$Z(x) = Z_{SK}(x) + \sigma_{SK}R(x)$$

where the index SK denotes the simple kriging estimate or its variance and $R(x)$ is a $N(0,1)$ residual. Note that for gaussian random variables, the SK estimate equals the conditional expectation.



Fig 6.3: Simplified well or drill hole containing 4 samples

Exponential variogram

Suppose that the four samples are 1m apart and that the underlying variogram for the gaussians is an exponential with a unit sill and a scale parameter $a = 2\text{m}$ (i.e. a practical range of 6m). The SK weights for estimating the top point using the other three points as data are:

$$\lambda_2 = 0.61, \lambda_3 = 0, \lambda_4 = 0 \text{ and } \sigma_{\text{SK}}^2 = 0.63 \Rightarrow \sigma_{\text{SK}} = 0.79$$

(Points are numbered from the top down). By symmetry the weights are the same but in the reverse order when kriging the fourth point from the other three.

The weights for the second point (or similarly the third one) are:

$$\lambda_1 = 0.44, \lambda_3 = 0.44, \lambda_4 = 0 \text{ and } \sigma_{\text{SK}}^2 = 0.46 \Rightarrow \sigma_{\text{SK}} = 0.68$$

As the configuration does not change from one iteration to the next the weights remain the same.

Step 1: Initialising the procedure

The first step consists of choosing gaussian values that belong to the appropriate intervals. Here we select $(-1, -1, +1, +1)$.

Step 2 : Iterative procedure

Simple kriging is applied to the points in turn. For example, the value of the top point is kriged using the other three points as input data. Then we move down to the second point and krige it using the initial values for the points below it and the new updated value for the top point. After completing the second point we move down to the third one which is kriged using the updated values for the points above it and the old value for the point below it. Similarly for the fourth point.

When all the points have been updated by kriging, one iteration has been completed.

Point No 1

The kriged estimate for $Z(x_1)$, abbreviated to $Z(1)$, based on the initial values (i.e. $-1, +1, +1$) for the other three points is:

$$Z_{sk}(1) = -1 \times 0.61 + 1 \times 0 + 1 \times 0 = -0.61$$

And the corresponding residual must satisfy

$$R(1) \leq -(-0.61) / 0.79 = 0.77$$

Suppose for argument's sake that we draw a value of 0.52 (from a $N(0,1)$ distribution). Then the updated value would be

$$Z(1) = -0.61 + 0.79 \times 0.52 = -0.20$$

Point No 2

The kriged estimate for $Z(2)$ based on the initial values for $Z(3)$ and $Z(4)$, and the updated value of $Z(1)$ is:

$$Z_{sk}(2) = -0.20 \times 0.44 + 1 \times 0.44 + 1 \times 0 = 0.42$$

The corresponding residual must satisfy

$$R(2) \leq 0.42 / 0.68 = 0.62$$

If we draw a value of -0.75 , then the updated value of $Z(2)$ is

$$Z(2) = 0.42 + 0.68 \times -0.75 = -0.09$$

Point No 3

Following the same procedure, the kriged estimate for $Z(3)$ and the inequality to be satisfied by its residual are

$$Z_{sk}(3) = -0.20 \times 0 - 0.09 \times 0.44 + 1 \times 0.44 = +0.40$$

$$R(3) \geq 0.40 / 0.68 = 0.59$$

If we draw a value of 0.28, then the updated value of $Z(3)$ is 0.21.

Point No 4

In the same way, the kriged estimate for $Z(4)$ and the inequality to be satisfied by its residual are

$$Z_{sk}(4) = 0 \times (0.20) + 0 \times (-0.09) + 0.61 \times 0.21 = 0.13$$

$$R(4) \geq -0.13 / 0.79 = -0.16$$

Drawing a value of 0.63 would give an updated value of 0.63 for $Z(4)$.

Results

Table 6.1 summarises the intermediate results during first iteration. This updating procedure is repeated iteratively, in general for several hundred or several thousand iterations. Table 6.2 shows the results of the first five iterations.

Table 6.1: Successive steps in the first iteration of this Gibbs sampler

		Initial				
	F_1	-1	-0.20	-0.20	-0.20	-0.20
	F_1	-1	-1	-0.09	-0.09	-0.09
	F_2	+1	+1	+1	+0.21	+0.21
	F_1	-1	-1	-1	-1	+0.63

Table 6.2: Results of first five iterations of the Gibbs sampler

Point No	Initial	No 1	No2	No3	N° 4	N° 5
Point 1	-1	-0.20	-0.37	-1.34	-0.34	-0.13
Point 2	-1	-0.09	-0.35	-0.24	-0.31	-0.20
Point 3	+1	+0.21	+0.15	+0.05	+0.19	+0.20
Point 4	-1	+0.63	+0.86	+0.10	+0.26	+0.32

Alternative updating strategies

In the previous example, individual points were sequentially updated. A variant of this consists of sequentially updating from the top down, then from the bottom up on the next iteration.

Blocking factor

It is also possible to update blocks of points simultaneously. For example, the four points could be grouped into two blocks each consisting of two points. There are three possible groupings of this type:

Points 1 & 2	and	Points 3 & 4
Points 1 & 3	and	Points 2 & 4
Points 1 & 4	and	Points 2 & 3

In the first case, the values of the top two points are updated using the values of the other two as the conditioning data (i.e. using kriging) and simulated in the right interval with the right correlations, and vice versa for the other pair. We will illustrate this procedure later in the chapter and will show that suitably chosen blocking strategies can significantly improve the speed of convergence.

Experimentally testing convergence

Having seen how the procedure works, several questions need to be answered. Firstly, does the algorithm converge? If so, after how many iterations? What factors affect the speed of convergence? How should we choose the initial values?

Burn-in period

In this section we illustrate the difference between the initial burn-in period and the subsequent stationary part of the Markov chain. To do this we continue the previous example but the variogram is changed to a gaussian model with a practical range of 3. So the correlation between adjoining samples is 0.95. Five hundred iterations of the corresponding Gibbs sampler were run starting from a very extreme set of initial values (+5,+5,-5,-5). This choice lengthens the initial burn-in period, making it visually much more obvious.

Figure 6.4 shows the output for each component as a function of the number of iterations. The values of the first component (top left) decrease steadily from the initial value of +5 until they are below 1.0. The curve seems to stabilise after approximately 100 iterations so the burn-in period must be at least this long. Similarly for the third and fourth components. But it appears to be much shorter for the other component (bottom left), about 20 iterations. This shows that the burn-in period need not be the same for all components in a Gibbs sampler. In MCMC theory it is well-known that different states can have different rates of convergence; see Meyn and Tweedie (1993, p362-363).

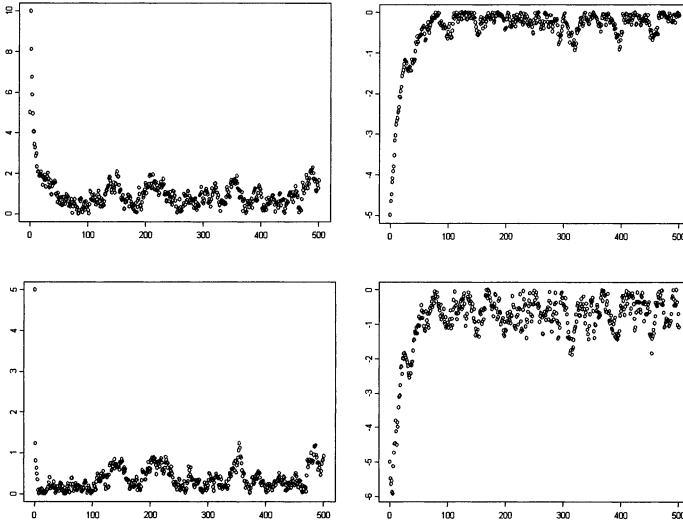


Fig 6.4: Output of the Gibbs sampler for all 4 components for 500 iterations, starting from initial values of $(+5, +5, -5, -5)$. The second component (lower left) seems to stabilise after about 20 iterations components whereas the other three are much slower. They take at least 100 iterations to reach their stationary distribution

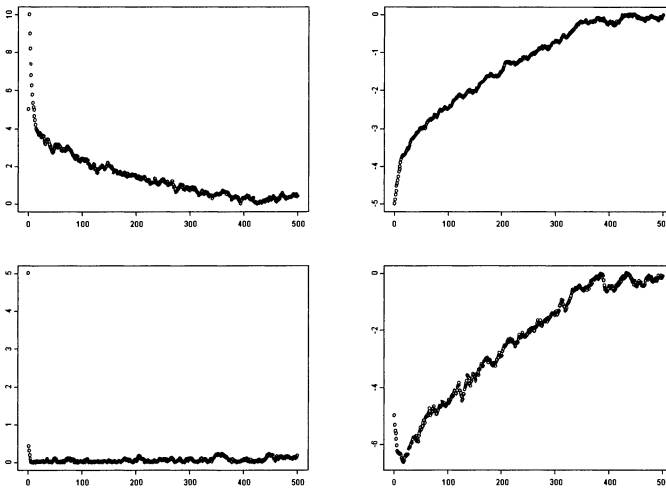


Fig 6.5: Output of the Gibbs sampler for the first of 4 components for 500 iterations, starting from initial values of $(+5, +5, -5, -5)$. Compared to Fig 6.4, the correlation between adjoining points has been increased from 0.95 to 0.98. Note the increase in the burn-in period

The implications of not necessarily having the same burn-in period for all components are important in practice. When there are only four components it is possible to check the convergence of all of them but if there were 1000 samples it would be virtually impossible to inspect the output of the Gibbs sampler for all 1000 components. We could only check a few of them visually and we might have the bad luck to choose those with shorter burn-ins. Inspecting the results for selected components gives us an idea of the burn-in period but it is not foolproof.

Effect of the range on the burn-in period

Several factors including the range of the variogram and the number of components have a marked effect on the length of the burn-in period. To illustrate the effect of the range, we repeated the previous example using a practical range of 5 instead of 3. This increases the correlation between adjoining samples from 0.95 to 0.98. Figure 6.5 shows the output.

Whereas the burn-in period for the first component was about 100 beforehand, it is now closer to 500. Conversely decreasing the range would decrease the burn-in period. Looking at Figure 6.5 we also notice how smooth the curves are compared to the corresponding ones in Figure 6.4. The strong serial correlation between successive values makes it more difficult to determine whether the Markov chain has converged.

These two examples show that it is not simple to judge whether a Gibbs sampler has reached its stationary distribution just by studying the output from a single run (even a very long one). It would be better to run a large number of samplers in parallel and study their output after 1, 5, 10, ... , 50 iterations and so on. Ideally we should compare the experimental distribution of the output with the stationary distribution. How could this be tested experimentally?

A multivariate normal distribution is fully specified when the means and the covariance matrix are known. By extension, a truncated gaussian is fully defined when the truncation thresholds are known together with the means and the covariance matrix for the full distribution. Having said that, it is clear that after truncation the means are not going to be the same as before. See Eqns (1) and (2). Nor are the variances or the correlations. In the case under study, the four components come from a quadrivariate normal distribution with the covariance matrix given in Table 6.3. Because of the high correlation between adjoining samples, its density is an elongated cigar shape. Figure 6.6 shows a diagram representing the bivariate densities of X_1 and X_2 , and X_2 and X_3 respectively.

Table 6.3: Quadrivariate covariance matrix for the 4-sample case

$$\begin{bmatrix} 1 & 0.95 & 0.80 & 0.61 \\ 0.95 & 1 & 0.95 & 0.80 \\ 0.80 & 0.95 & 1 & 0.95 \\ 0.61 & 0.80 & 0.95 & 1 \end{bmatrix}$$

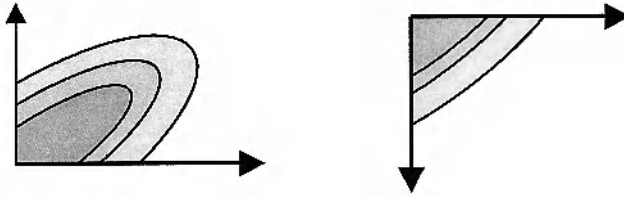


Fig 6.6: Schematic representation of the bivariate densities of X_1 and X_2 (left) and of X_2 and X_3 (right)

The impact of different types of truncations is evident from these. In one case, we are dealing with the elongated part of the ellipse whereas in the other, it is merely a triangular "corner". Intuition can often be misleading when trying to guess the properties of truncated gaussian distributions. For example people expect the marginal distributions in this example to be "half gaussians". The marginal distributions given in Figure 6.7 show just how wrong intuition can be and also confirms what was shown in the two variable case given at the beginning of the chapter.

Parallel runs

Up till now we have illustrated the difference between the burn-in period and the stationary part by focussing on individual components. An alternative is to run many Gibbs samplers in parallel and study the results after a certain number of iterations. Figure 6.8 plots the first and second components for 50 parallel runs. Figure 6.8 (a) shows their locations after a single iteration; both started out from an initial value of +5. Figures 6.8 (b), (c) and (d) give the output after 5, 10 and 50 iterations. As expected, the centre of the cloud moves downwards and disperses outward from this. Initially the distribution is far from the target cloud but as the number of iterations increases, it steadily tends toward it.

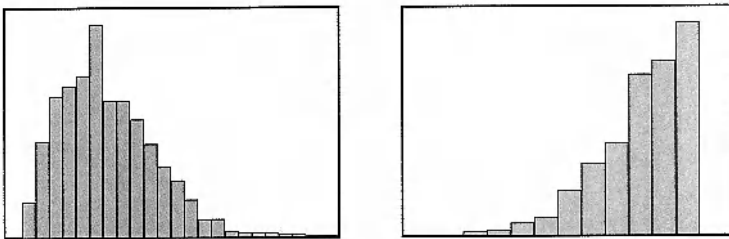


Fig 6.7: The marginal distributions of X_1 and of X_3 , from the 4 component example

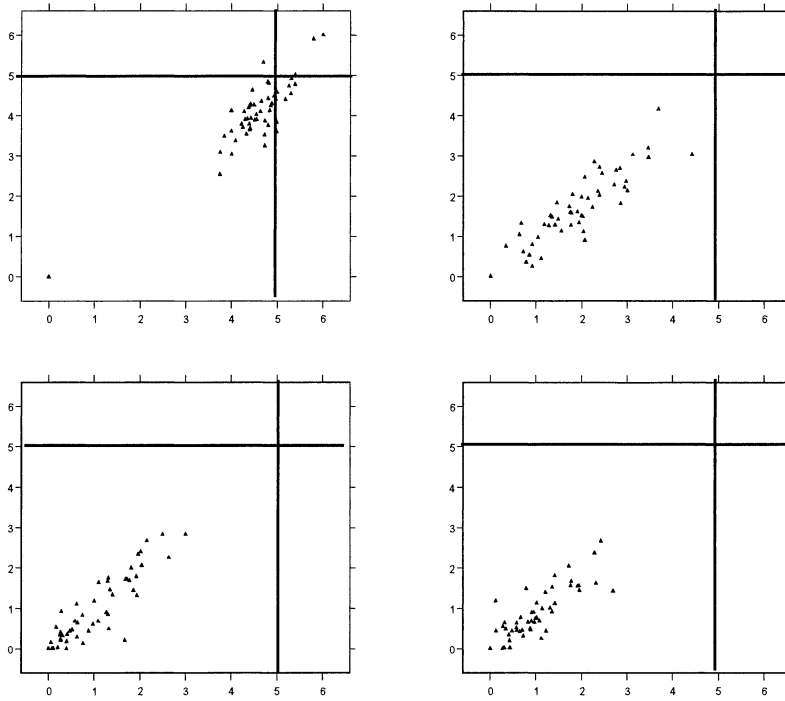


Fig 6.8: The values of the first and second components after 1 iteration (top left), then 5 (top right), 10 (bottom left) and 50 iterations (bottom right) starting out from initial values of +5 (as indicated by the crosshairs)

Part 2: Gaussian framework

Basic properties of a single gaussian

Let Z be an arbitrary gaussian vector of dimension $p+n$. Let w_1, \dots, w_{p+n} be its $p+n$ components. Assume its covariance matrix Σ has an inverse Σ^{-1} . Its density is:

$$g(w_1, w_2, \dots, w_{p+n}) = \frac{1}{(2\pi)^{(p+n)/2} |\Sigma|} \exp \left\{ -\frac{1}{2} (w - m)^t \Sigma^{-1} (w - m) \right\}$$

where w and m are column vectors with $m = E(Z)$ and where $|\Sigma|$ is the determinant of the covariance matrix.

We split the vector Z into two parts X and Y , of dimension p and n respectively, which correspond to the p data points and the n target points to be simulated. Partitioning the covariance matrix in the same way we obtain:

$$\Sigma = \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{XY}^t & \Sigma_{YY} \end{pmatrix}$$

where Σ_{XX} and Σ_{YY} are the covariance matrices of X and Y , respectively, and Σ_{XY} is the cross covariance matrix between X and Y . Let Λ be the vector defined by

$$\Lambda = \Sigma_{XX}^{-1} \Sigma_{XY}$$

It is easy to see that its i^{th} column consists of the weights for the minimum variance linear estimate of the i^{th} element of Y using the vector X as information. If the data are regionalised variables, the weights are just the simple kriging weights when estimating the $(p + i)^{\text{th}}$ component of Z using the first p values as data. Let R be the covariance matrix of the residuals when estimating Y from X ; that is, $R = \text{cov}(Y - \Lambda^t X)$. It is well known (see for example Galli & Gao (2001)) that the inverse of the covariance matrix Σ^{-1} can be written as

$$\Sigma^{-1} = \begin{bmatrix} \Sigma_{XX}^{-1} + \Lambda R^{-1} \Lambda^t & -\Lambda R^{-1} \\ -R^{-1} \Lambda^t & R^{-1} \end{bmatrix}$$

If we let u and v be the vectors containing the first p components of w and the remaining n components then

$$(u^t \ v^t) \Sigma^{-1} \begin{pmatrix} u \\ v \end{pmatrix} = (v - \Lambda^t u)^t R^{-1} (v - \Lambda^t u) + u^t \Sigma_{XX}^{-1} u$$

Now we use this decomposition to express the gaussian density given earlier as the product:

$$g(u, v) = k \ g_u(v) \ g(u)$$

where the normation constant k is just

$$k = \frac{|\Sigma_{xx}| |R|}{|\Sigma|}$$

$$g(u) = \frac{1}{(2\pi)^{p/2} |\Sigma_{xx}|} \exp \left\{ -\frac{1}{2} u^t \Sigma_{xx}^{-1} u \right\}$$

and

$$g_u(v) = \frac{1}{(2\pi)^{n/2} |R|} \exp \left\{ -\frac{1}{2} (v - \Lambda^t u)^t R^{-1} (v - \Lambda^t u) \right\}$$

Both are normal distributions; the first has zero mean and covariance Σ_{xx} , the second has mean $\Lambda^t u$ and covariance R . As their integrals are equal to 1, the normation factor k turns out 1, which was not obvious initially. The first one $g(u)$ is the marginal distribution of the vector X . The second is the conditional density of Y given X . It is still gaussian but its mean now equals the minimum variance estimate of Y given X , and its covariance matrix is R , the covariance matrix of the residuals.

Extension to the plurigaussian case

Although the plurigaussian algorithm can be described in a more general framework (Galli, Beucher et al, 1994) we will consider the most commonly used case where there are two gaussian random functions Z_1 and Z_2 with zero mean and unit variance and where the lithotype rule is defined using rectangles. That is, if a data point x belongs to the facies F_k then the pair $(Z_1(x), Z_2(x))$ has to lie in the rectangle R_k . This is equivalent to saying that $Z_1(x)$ belongs to the interval $I_{1,k}$ and $Z_2(x)$ belongs to the interval $I_{2,k}$.

As before we want to simulate at n points (for example, on a grid) given the facies information at p data points (i.e. given that the two gaussians lie in certain intervals at the p points). Let Z be the vector of length $2(n+p)$ consisting of the p data pairs followed by the n target pairs:

$$Z^t = \underbrace{(Z_1(x_1), Z_2(x_1)) \cdots (Z_1(x_p), Z_2(x_p))}_{p \text{ sample points}}, \underbrace{(Z_1(x_{p+1}), Z_2(x_{p+1})) \cdots (Z_1(x_{p+n}), Z_2(x_{p+n}))}_{n \text{ target points}}$$

The joint density of the target values and the data can be written as:

$$h(u, v) = k g(u) g_u(v) \varphi(u) \quad (3)$$

where $g(u)$ is the $2p$ variate gaussian density of the data. The next term is the $2n$ variate gaussian density of the points to be simulated centred by the simple cokrigings. Its covariance is the covariance of the residuals. Finally, the last term is the product of the indicators:

$$\varphi(u) = \prod_{k=1}^p 1_{R_k}(u)$$

Because of the way we chose to write Z each indicator depends only on u_{2k+1} and u_{2k+2} .

The goal of the simulation is to sample from $h(u, v)$. Integrating with respect to v we find that the product $kg(u)\varphi(u)$ is the marginal density corresponding to the data points. However as in the examples with two points at the beginning of the chapter, the distribution of the points to be simulated is not gaussian. Instead their density $h(v)$ is (up to a factor) the product of the gaussian density and the probability that the data points lie in the domain D translated by the cokriging of the data points using the values of the points to be simulated. That is,

$$h(v) = k' g(v) P_R(D - \tilde{\Lambda}'v)$$

The domain D is the product of the rectangles R_k (i.e. hyper-rectangle in \mathfrak{R}^p). The matrix $\tilde{\Lambda}$ contains the weights for cokriging X given Y . In fact, it is the matrix of the weights when estimating Z_1 and Z_2 at the data points using the values of both variables at the points to be simulated. So if Z_1 and Z_2 are independent, the weights are just the SK weights, as before. Alternatively if they are correlated, the weights are the simple co-kriging weights. So $\tilde{\Lambda}$ satisfies the equation:

$$\Sigma_{YY}\tilde{\Lambda} = \Sigma_{XY}$$

Finally the index R in the gaussian probability P_R denotes the covariance matrix of the cokriging residuals.

Conditional simulation

Equation (3) shows that in order to sample from $h(u, v)$, we have to work in two steps.

1. At data points draw a gaussian vector $(Z_1(x_1), Z_2(x_1), \dots, Z_1(x_p), Z_2(x_p))$ satisfying the constraints (with density $g(v)\varphi(v)$).
2. Once the gaussian values satisfying the constraints are known, we have only to simulate a gaussian vector having the (co) kriging with previous values as its mean and the covariance matrix of the residuals. That is, it is a classical conditional gaussian simulation.

The difficulty lies in step 1. As we cannot sample from the required density in a single step, we have to resort to a Gibbs sampler.

Implementation of the Gibbs sampler for the model used

It is quite simple to implement the two step procedure when the two gaussians Z_1 and Z_2 are independent. When they are correlated it suffices to consider two independent gaussians Y_1 and Y_2 as was shown earlier. This considerably simplifies the two steps of the simulations. Here we assume that the Gibbs sampler updates points one after another (i.e. with a blocking factor of 2). This means that at each step we have to draw a point according to the conditional density knowing the values at all points except one. So its conditional density is a noncentered truncated plurigaussian of dimension 2 and the mean is the cokriging. If the gaussian values $(Z_1(x_k), Z_2(x_k))$ belong to the rectangle, $R_k = I_k^1 \times I_k^2$, then

$$Y_1(x_k) \in I_k^1 \quad \text{and} \quad Y_2(x_k) \in \frac{I_k^2 - \rho Y_1(x_k)}{\sqrt{1 - \rho^2}}$$

For simplicity let the second interval be $I_k^2(Y_1, \rho)$. So now we have only to simulate two independent gaussians, firstly Y_1 in the interval I^1 and then Y_2 in the interval $I^2(Y_1, \rho)$. And the means are just the two simple krigings of Y_1 and Y_2 . The same decomposition of Z_1 and Z_2 into two independent gaussians is also used in the second step which is a bivariate conditional simulation. As expected the conditioning by cokriging reduces to two simple krigings in Y_1 and Y_2 . At the end Y_1 and Y_2 are recombined to give Z_1 and Z_2 .

To summarise:

The simulation method is quite straightforward: first simulate two independent gaussian random functions Y_1 and Y_2 with values in the appropriate intervals at the data points, then use these values as conditioning data to simulate two independent gaussian random functions at target points.

Convergence from a theoretical point of view

The untruncated case

The Gibbs sampler is known to converge in the untruncated gaussian case. See Barone and Frigessi (1990), Amit and Grenader (1991), Roberts and Sahu (1997) or Galli and Gao (2001). In fact, the Gibbs Sampler corresponds to Gauss-Seidel iterations on the inverse of the covariance matrix. This can be proved by expressing the inverse of the covariance matrix C^{-1} as $D(I - L - U)$ where D is the

block diagonal part of C^{-1} , L and U are lower and upper block triangular matrices and I is the identity matrix. Then the iteration matrix of the Gibbs sampler can be written as $A = (I - L)^{-1} D^{-1} U$. So the rate of convergence depends on its spectral radius $\rho(A)$. A simple proof based on relating these iterations to linear fix points iterations is given in Galli and Gao (2001).

The iteration is:

$$Y^{(n+1)} = AY^{(n)} + (I - L)^{-1} R^{(n+1)}$$

where $R^{(n+1)} = L_R V^{(n+1)}$. Here L_R is the lower triangular part of the Cholesky decomposition of D^{-1} and $V^{(n+1)}$ is a vector containing independent $N(0,1)$ components. Consequently the covariance matrix of $R^{(n+1)}$ is D^{-1} . The equation can be written in a more compact form.

$$Y^{(n+1)} = AY^{(n)} + U^{(n+1)}$$

where the covariance matrix of $U^{(n+1)}$ is

$$\Sigma = (I - L)^{-1} D^{-1} (I - L)^{-t}$$

Blocking factor

Galli and Gao (2001) also demonstrated the importance of the block size used in the Gibbs sampler (i.e. in the matrices D , L and U) for geostatistical applications. The most natural choice for D is simply the diagonal terms of C^{-1} . This corresponds to blocks of size 1×1 . Other choices are possible and these can significantly increase the speed of convergence.

To illustrate its impact, we compare the spectral radius for the 1×1 blocking size with that for the 2×2 case for the 4 sample example considered earlier in the chapter. In that case the covariance matrix is of the form:

$$C = \begin{bmatrix} 1 & a & a^2 & a^3 \\ a & 1 & a & a^2 \\ a^2 & a & 1 & a \\ a^3 & a^2 & a & 1 \end{bmatrix}$$

and so

$$C^{-1} = \frac{1}{(a^2 - 1)} \begin{bmatrix} 1 & a & 0 & 0 \\ a & -(a^2 + 1) & a & 0 \\ 0 & a & -(a^2 + 1) & a \\ 0 & 0 & a & 1 \end{bmatrix}$$

In this case, it is not difficult to calculate the various matrices and hence find the spectral radii for the two block sizes and also if the points are arranged in a different order.

Block size of 1 (order: 1, 2, 3, 4)

$$D = \frac{1}{(a^2 - 1)} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -(a^2 + 1) & 0 & 0 \\ 0 & 0 & -(a^2 + 1) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A = (I - L)^{-1}U = \begin{bmatrix} 0 & b & 0 & 0 \\ 0 & b^2 & b & 0 \\ 0 & b^3 & b^2 & b \\ 0 & b^4 & b^3 & b^2 \end{bmatrix} \quad \text{where } b = \frac{a}{(a^2 + 1)}$$

and hence the spectral radius $\rho(A)$ is the largest root of the equation:

$$\lambda^2 [\lambda^2 - 3b^2\lambda + b^4] = 0$$

That is, it is

$$\rho(A) = b^2 \frac{3 + \sqrt{5}}{2} < 1 \quad \text{since} \quad \max(b^2) = \max\left(\frac{a^2}{(a^2 + 1)^2}\right) = \frac{1}{4}$$

Block size of 2 (points 1&2 and 3&4)

$$D = \frac{1}{(a^2 - 1)} \begin{bmatrix} 1 & a & 0 & 0 \\ a & -(a^2 + 1) & 0 & 0 \\ 0 & 0 & -(a^2 + 1) & a \\ 0 & 0 & a & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 0 & a^2 & 0 \\ 0 & 0 & a & 0 \\ 0 & 0 & a^2 & 0 \\ 0 & 0 & a^3 & 0 \end{bmatrix}$$

and hence the spectral radius $\rho(A)$ is the largest root of the equation:

$$\lambda^3 [a^2 - \lambda] = 0 \quad \Rightarrow \rho(A) = \sqrt{a}$$

Block size of 2 (points 1&3 and 2&4)

$$D = \frac{1}{(a^2 - 1)} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -(a^2 + 1) & 0 & 0 \\ 0 & 0 & -(a^2 + 1) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 0 & b & 0 \\ 0 & 0 & b & b \\ 0 & 0 & 2b^2 & b^2 \\ 0 & 0 & b^2 & b^2 \end{bmatrix}$$

and hence the spectral radius $\rho(A)$ is the largest root of the equation:

$$\lambda^2 [\lambda^2 - 3b^2\lambda + b^4] = 0$$

So it is the same as in the first case:

$$\rho(A) = b^2 \frac{3 + \sqrt{5}}{2} < 1$$

Convergence in the truncated gaussian case

In his thesis Freulon (1992) considered two iterative ways of simulating truncated gaussian random functions. The first was based on work in the field of statistical mechanics by Metropolis (1953) while the second is the Gibbs sampler (see Geman and Geman, 1984). The procedure given above follows the second of these. Freulon (pp58-65) proved that both procedures converge to a truncated gaussian distribution, for the total variation distance, but his demonstration of their ergodicity only works for the case where the constraint is a compact (i.e. a closed and bounded interval). Secondly it does not provide any way to compute the rate of convergence because it is the minimum of $k(x,y)$ on $D \times D$ which is difficult to evaluate because of the denominator. In order to obtain more general results, we need to work in a more general framework, and so we introduce Markov chains.

Markov chains

Here we give an intuitive introduction to Markov chains. For more precise and more complete descriptions see Numelin (1984), Meyn and Tweedie (1993) or Gikhman and Skorokhod (1969). A Markov chain is a chain $X^{(0)}, X^{(1)}, \dots, X^{(n)} \dots$ with the property that for n greater than 1:

$$E(X^{(n)} | X^{(n-1)}, \dots, X^{(1)}, X^{(0)}) = E(X^{(n)} | X^{(n-1)})$$

So a chain is fully determined as soon as we know the transition probabilities from one state to another. Here the possible states are not countable so they are characterised by an integral kernel $k_n(x, y)$. If the kernel does not depend on the position in the chain, ie $k_n(x, y) = k(x, y)$ the chain is said to be homogeneous. We will only consider this case.

Two integral operators

We can define two integral operators from this kernel.

(1) If $f(x)$ is a density then the operator $g(y) = Kf(y)$ as defined below is a density:

$$g(y) = \int_S k(x, y) f(x) dx$$

Furthermore, if f is the density of the $(n-1)^{\text{st}}$ iterate then $g(y)$ is the density of the n^{th} one.

(2) Let $h(x)$ be the conditional expectation of $t(X^{(n)})$ given $X^{(n-1)}$. Then

$$h(x) = E(t(X^{(n)}) | X^{(n-1)} = x) = \int_S k(x, y) t(y) dy$$

The two operators can be expressed in a more symmetrical way by letting

$$\tilde{k}(y, x) = k(x, y)$$

In that case

$$K_F t(y) = \int_S \tilde{k}(x, y) t(x) dx$$

Let $K^{(n)}$ be the iterated operator, that is,

$$K^{(n)}(f) = K(K^{(n-1)}(f)).$$

Similarly let $K_F^{(n)}$ be

$$K_F^{(n)}(t) = K_F(K_F^{(n-1)}(t))$$

Finally let $k^{(n)}(x,y)$ be the corresponding kernel. The operator $K^{(n)}$ expresses the relationship between the density of $X^{(p)}$ and that of $X^{(p+n)}$. In contrast $K_F^{(n)}$ gives the conditional expectation of a function of $X^{(p+n)}$ given $X^{(p)}$

$$E(t(X^{(n+p)}) | X^{(p)})$$

Note that for a homogeneous chain the joint distribution of the iterates $X^{(n+p)}$ and $X^{(n)}$ is:

$$g^{(n+p)}(x,y) = k^{(n)}(x,y)f^{(p)}(x) = g_y^{(n)}(x)f^{(n+p)}(y)$$

Here $f^{(p)}$ and $f^{(n+p)}$ stand for the marginal densities obtained by running the chain. The last equality holds because of Bayes Theorem.

If we let $k^*(y,x)$ be the transition kernel corresponding to reverse passage from y to x , then $g_y^{(n)}(x)$ corresponds to $k^{*(n)}(y,x)$. So we get:

$$g^{(n+p)}(x,y) = k^{(n)}(x,y)f^{(p)}(x) = k^{*(n)}(y,x)f^{(n+p)}(y) \quad (4)$$

Stationary distribution

We say that a chain has a stationary distribution if the influence of the initial point $X^{(0)}$ disappears after some time. More formally a chain has a stationary distribution if $P(X^{(n)} \in A | X^{(0)})$ tends toward $\pi(A)$ for any measurable set A , whatever the initial state $X^{(0)}$. In our case this is equivalent to $k^{(n)}(x,y) \rightarrow \pi(y)$. The chain is said to have an invariant distribution π if there exists a distribution $\pi(x)$ such that

$$\pi(y) = \int_S k(x,y)\pi(x)dx.$$

Clearly a stationary distribution must be an invariant distribution. Assuming that $\pi(x)$ is an invariant distribution, and applying (4) with $p = 1$, $n = 1$, and $f(x) = \pi(x)$ we get:

$$g^{(2)}(x,y) = k(x,y)\pi(x) = k^*(y,x)\pi(y) \quad (5)$$

Irreducible chain

A Markov chain is said to be ψ irreducible if for each x in the support of a measure ψ , and for each measurable set A such that $\psi(A) > 0$, there exists an integer $p \geq 1$ such that

$$P(X^{(p)} \in A | X^{(0)} = x) > 0$$

A chain is said to be periodic if there exists an integer $d \geq 2$ and a sequence of nonempty disjoint sets D_1, \dots, D_{d-1} $i = 1, \dots, (d-1)$ such that

$$\forall i = 1, \dots, d-1 \text{ and } \forall x \in D_i, K(x, D_j) = 1 \text{ where } j = (i+1) \text{ modulo } d$$

A chain is aperiodic if it is not periodic. Note that if a chain is irreducible with $p=1$, then it is aperiodic.

Distance metric and convergence

To define convergence we first have to define a distance criterion. One of the most commonly used distances is total variation (TV). The fact that a measure μ_n converges toward a measure μ for this distance, implies that for any bounded measurable function f we have convergence of the expectation. That is,

$$\|\mu_n - \mu\|_{TV} \rightarrow 0 \text{ implies that } \int (\mu_n - \mu)(dx)f(x) \rightarrow 0$$

for any measurable bounded function f . This can be written as

$$E_n(f) \rightarrow E(f)$$

We have the following convergence result for Markov chains.

Theorem 1 (Tierney, 1994 and 1996)

If a chain is an irreducible aperiodic markov chain with transition kernel k and an invariant distribution π then:

$$\|k_F^{(n)}(x, \cdot) - \pi(\cdot)\|_{TV} \rightarrow 0$$

Although this result is interesting, its usefulness is limited because of its generality. Secondly it does not give the rate of convergence. Stronger results can be obtained in L^2 type spaces, provided one makes some additional assumptions about the compactness of the integral operators defined previously. Here we will restrict ourselves to the case where the state space S is compact. For the truncated gaussian case this excludes unbounded intervals. From a practical point of view, it means adding two additional facies with infinitesimal proportions which are not seen in the data in order to bound the interval by a large value.

For that case the previous result provides a method to estimate the speed of convergence in practice, at least for reasonably small spaces. In the case of truncated gaussian this means a reasonable number of constraints. The idea is to estimate the transition operator while running the chain as was proposed by Lantuejoul (2002). But here the chain has a continuous state space, so it is not possible to consider all possible transitions.

Given a finite partition of S , that is, a family of disjoint sets Δ_i such that:

$$S = \bigcup_{i=1}^n \Delta_i$$

we consider the transition matrix:

$$P_{i,j}^{\Delta} = \frac{1}{|\Delta_i|} \int_{\Delta_i} \int_{\Delta_j} k(x, y) dx dy$$

This can be estimated reasonably by running the chain for a long enough time. Then we can compute the eigenvalues of $P_{i,j}^{\Delta}$. So the question is now: How close are these eigenvalues to those of K ? Our assumptions are

1. The state space S is compact.
2. The kernel k is continuous
3. The assumptions of theorem 1 hold.

The first and second assumptions ensure that K will map $C(S)$ to $C(S)$. This is easy to verify because for continuous functions on a compact, there is no problem inverting the limit and the integral. Then using a result in Yosida (1978, p 277) we can show that this new operator K considered from $C(S)$ to $C(S)$ is compact. The convergence result of theorem 1 ensures that the eigenvalues of K are lower than or equal to 1, and that it has one and only one eigenvalue 1, which is associated to the eigenvector π . This is easy to show because the set of bounded measurable functions on S includes continuous functions. Consequently

$$\sup_{f \in C[S], \|f\| \leq 1} \left| \int_S (k^{(n)}(x, y) - \pi(y)) f(y) dy \right| \leq \|k^{(n)} - \pi\|_{TV}$$

If there were a distribution $g \neq \pi$ such that $K_F(g) = cg$ with $c \geq 1$, then $K_F^{(n)}(g)$ would equal $c^n g$ and the integral

$$\int_S (c^n g(y) - \pi(y)) dy$$

would not converge to 0. This, in turn, implies that $\|k^{(n)} - \pi\|_{TV}$ would not converge. Finally if we define

$$P_{\Delta}(f)(t) = \frac{1}{|\Delta_i|} \int_{\Delta_i} f(u) du \quad \text{for } t \text{ in } \Delta_i$$

We see that when

$$|\Delta| \rightarrow 0, \quad P_{\Delta} f(t) \rightarrow f(t)$$

And finally it can be shown that this corresponds to a Galerkin approximation. See Chatelin (1983, pp170-174). That guarantees the convergence of P^Δ toward k as $|\Delta|$ tends toward 0. Interestingly, the rate of convergence depends on the derivability of k (see exercice 4.39 p 191 in Chatelin (1983)). Furthermore the convergence of eigenvalues is also ensured and bounds are known.

The Gibbs Sampler

Let $\pi(x)$ be the density of the vector (X_1, \dots, X_n) , and let $\pi(x_i | x_{i-}, x_{i+})$ be the conditional density of X_i given (X_{i-}, X_{i+}) where

$$X_{i-} = (X_1, \dots, X_{i-1}) \quad \text{and} \quad X_{i+} = (X_{i+1}, \dots, X_n)$$

By convention

$$X_{0-} = \emptyset \quad \text{and} \quad X_{n+} = \emptyset$$

The Gibbs sampler modifies all the coordinates of the vector, one after another, using the newly simulated one and the former ones.

Note that X_i may be either a single component or a block of components. In the latter case we should speak of the block Gibbs sampler, but for simplicity we will call it a Gibbs sampler even if we are using a blocking factor.

The transition kernel for one iteration of the Gibbs sampler from x to y is:

$$k(x, y) = \prod_{i=1}^n \pi(y_i | y_{i-}, x_{i+})$$

The positivity condition is said to hold if the support of π is the product of the supports of its marginal distributions. This turns out to be true in our case provided the constraints are products of unions of intervals.

Using this positivity condition we can establish the Hammersley Clifford decomposition (Besag, 1974). It is easy to show that

$$\begin{aligned} \pi(x) &= \pi(x_1 | x_{1-}, x_{1+}) \pi(x_{1-}, x_{1+}) \\ &= \frac{\pi(x_1 | x_{1-}, x_{1+})}{\pi(y_1 | x_{1-}, x_{1+})} \pi(y_1 | x_{1-}, x_{1+}) \end{aligned}$$

By repeating this operation on $\pi(y_1 | x_{1-}, x_{1+})$ we obtain:

$$\pi(x) = \frac{\prod_{i=1}^n \pi(x_i | y_{i-}, x_{i+})}{\prod_{i=1}^n \pi(y_i | y_{i-}, x_{i+})} \pi(y) \quad (\text{Hammersley Clifford})$$

The denominator is $k(x,y)$ and the numerator is the kernel $k^*(y,x)$ corresponding to one iteration of the Gibbs sampler from y to x in the reverse order. So we can rewrite the Hammerseley Clifford decomposition in the following way:

$$\pi(x)k(x,y) = \pi(y)k^*(y,x) \quad (6)$$

Integrating this gives

$$\begin{aligned} \int \pi(x)k(x,y)dx &= \int \pi(y)k^*(y,x)dx \\ &= \pi(y) \int k^*(y,x)dx \\ &= \pi(y) \end{aligned}$$

Consequently π is an invariant distribution of the Markov chain with transition kernel $k(x,y)$ and also of the Markov chain with transition kernel $k^*(y,x)$. Equation (6) could be obtained directly from equation (5) without the positivity assumption but with the assumption that π is the stationary distribution of the chain.

Truncated Gaussian case

In the truncated gaussian case

$$\pi(x) = \kappa g(x)l_D(x)$$

where κ is the normalisation constant and $l_D(x)$ is the indicator of constraints. We will only consider constraints which are products of unions of intervals. So the positivity condition holds because the conditional distributions $\pi(x_i | x_{i-}, x_{i+})$ are truncated gaussians whose domain is the union of intervals corresponding to the constraint for the i^{th} component. Let the inverse of the covariance matrix of X be

$$C^{-1} = D(I - L - U)$$

where D is the block diagonal part of the inverse of C , and U and L are the upper and lower triangular parts.

Using the results for the non truncated gaussian case, it is easy to show that the transition kernel is:

$$k(x,y) = k(x)e^{-\frac{1}{2}(y-Ax)^T \Sigma (y-Ax)} l_D(y)l_D(x)$$

where $k(x)$ is a normation factor whose value depends on x and

$$A = (I - L)^{-1}U \quad \text{and} \quad \Sigma = (I - L)^T D (I - L)$$

Similarly,

$$k^*(y, x) = k^*(y) e^{-\frac{1}{2}(x - \tilde{A}y)^t \tilde{\Sigma}(x - \tilde{A}y)} l_D(y) l_D(x)$$

where

$$\tilde{A} = (I - U)^{-1} L \quad \text{and} \quad \tilde{\Sigma} = (I - U)^t D (I - U)$$

It is then quite obvious that the conditions for theorem 1 hold so the Gibbs Sampler converges toward our candidate distribution π , for the total variation distance.

Then the kernels are continuous, in the support of π , and if D is compact the two operators K and K_F corresponding to the truncated gaussian case define compacts operators from $C(D)$ to $C(D)$. So we can apply the approximation method described earlier to estimate the eigenvalues, and to estimate the speed of convergence experimentally. It is also interesting to note that provided the constraints are the product of compact intervals, the minorisation condition holds.

Definition

We say that the kernel $k(x, y)$ satisfies the minorisation condition if there exists a density ψ , and a parameter $0 < \rho < 1$ such that

$$\forall x \in D \quad \forall y \in D \quad k(x, y) \geq \rho \psi(y)$$

Theorem 2

If the constraints D are compact the kernel $k(x, y)$ of the transition operator satisfies the minorisation condition.

Proof

$$k(x, y) = k(x) e^{-\frac{1}{2}(y - Ax)^t \Sigma(y - Ax)} l_D(y) l_D(x)$$

As the proof depends on the properties of convex sets, we start out by reviewing some basic facts about them.

A point v in a convex is extremal if it cannot be written as a convex combination of other points in the convex. In a polyhedral convex such as $y - A(D)$, the extremal points are just the vertices of the polyhedra.

Because D is compact and $k(x)$ is continuous, it is bounded below by a parameter c which is greater than 0 because $k(x) \approx P(D - Ax)$, and its minimum is reached at at least one point x_0 and $P(D - Ax_0) > 0$.

In order to minimise the exponential term we have to maximise the exponent. Let $u = y - Ax$. This is equivalent to maximising

$$\|u\|_K^2 = u^t \Sigma u$$

subject to the constraints

$$u \in y - A(D)$$

As the function is continuous and as the set $y - A(D)$ is compact, the minimum is attained. Because $\Sigma \geq 0$, the function is strictly convex.

It is well known that the maximum of a convex function on a convex is reached at extremal points of the convex. See for example Rockafellar (1970). Now each extremal point of $y - A(D)$ is of the form $y - A(x_i)$ where x_i is an extremum in D . There are 2^n extremal points x_i in D , they are the vertices of the hyperrectangle. Let T_i be the points y whose corresponding extremal point is x_i . By definition D is the union of the T_i and we can write:

$$k(x, y) > c \sum_i 1_{T_i}(y) e^{-\frac{1}{2}(y - Ax_i)^t \Sigma (y - Ax_i)}$$

If we let $\tilde{f}(y)$ denote

$$\tilde{f}(y) = c \sum_i 1_{T_i}(y) e^{-\frac{1}{2}(y - Ax_i)^t \Sigma (y - Ax_i)}$$

we find that

$$1 = \int_D k(x, y) dy > \int_D \tilde{f}(y) dy > 0$$

If we let ρ be the integral of \tilde{f} over D we can define the density $f(y) = \tilde{f}(y)/\rho$. We find that:

$$\forall y \in D, \forall x \in D \quad k(x, y) > \rho f(y) \quad \text{with } \rho < 1$$

This minorisation condition ensures the convergence of the Gibbs sampler in L^1 . See for example, Robert and Polson (1994).

Comments

The Gibbs sampler is a very powerful simulation method. It is not difficult to establish its convergence in the truncated gaussian case. However it is far more difficult to find its rate of convergence experimentally, even in the case of compact constraints which considerably simplifies the problem. For this case we proposed a finite rank operator which approximates the transition operator (which is compact).

7 CASE STUDIES AND PRACTICAL EXAMPLES

Choosing which simulation method to use

Simulation methods fall into two broad classes: pixel methods and object-based methods. Before presenting any practical examples, it is important to decide which method would be the most appropriate for the problem under study. So we first review the pros and cons of two classes of models, sequence-based pixel methods such as the truncated gaussian and plurigaussian simulations, and object-based methods, in particular boolean simulations. As this book focuses on plurigaussian simulations, we do not intend to present boolean simulations here. Interested readers can consult Matheron (1968, 1975), Lantuéjoul (1997a, 1997b and 2002) Chilès and Delfiner (1999), and Molchanov (1997). Sequence-based pixel methods and boolean simulations can also be combined to get the best of both approaches. These are called nested simulations.

Sequence-based pixel models

As their name suggests, sequence-based models are well suited to simulating geological configurations where lithotypes are organised in sedimentary sequences. Within a sedimentary sequence, lithotypes are correlated both vertically and horizontally. According to Walther's law, vertical successions of lithotypes also imply a lateral correlation between the lithotypes. Another advantage of pixel simulations is that they realistically reproduce the facies interfingering. Sedimentary sequences occur at different scales, and their organisation is a function both of auto and allocyclic processes. For example, fluvial channel sequences are characterised by a particular succession of lithotypes from base to top as the channels filled. The progradation of a shoreface which is mostly controlled by relative sea-level variations also shows a typical vertical succession of facies from offshore to foreshore. Over the past ten years, tremendous progress has been made in sequence stratigraphy and this has led to a much better understanding of the facies architecture of depositional sequences of different orders, in a wide range of depositional environments.

Even though the relative volume of facies and their geometric organisation (the facies partitioning) changes according to sequence organisation and although the sequences are usually non-stationary, proportion curves have

proved to be a simple but highly effective tool for imaging sequences and for quantifying the associated non-stationarity. As they measure the relative proportion of lithofacies, both vertically and horizontally, they provide a good way of describing the spatial distribution of the lithofacies. They can guide the geologist in his interpretation and can help to define the reservoir layering (Eschard et al., 1999).

Vertical proportion curves give information on the facies partitioning within the sequences and indicate the stratigraphic levels where the main vertical permeability barriers can be expected, which is important in reservoir characterisation.

Computing horizontal proportion curves also helps in understanding the spatial distribution of the lithofacies and the geometry of the reservoir. Their representativity depends on the well spacing and distribution in the area to be simulated. When calculated in different directions across the field, they can show whether there are any major lateral facies variations in the study area. Furthermore, they can aid in deciding whether there is non-stationarity. When the relative proportions of facies vary significantly laterally within a reservoir unit, then non-stationarity is present.

In sequence-based pixel models, the correlation between facies is quantified by the vertical and horizontal variograms. In standard geostatistical studies, the range of the variogram has a direct geological significance; here the relationship is indirect. It is a function of the size of the heterogeneity, but also of its spatial frequency. The type of variogram used in the simulations also has a strong impact on the way the simulated facies are organised. Exponential variograms, for example, are appropriate when facies transition is progressive. On the contrary, gaussian variograms generate more "rounded" shapes in simulations.

Many depositional systems can be simulated well by sequence-based pixel approaches. They can handle carbonate depositional systems, which are often characterised by a cyclic organisation of facies and progressive facies interfingerings. In clastic depositional environments, the approach is also suitable for deltaic or shoreface sediments which are organised in transgressive-regressive cycles. Internal heterogeneity within fluvial sand sheets can also be simulated in this way. These types of reservoirs often show an internal sequential organisation which can be reproduced easily by sequence-based models.

One difficulty in the method is computing reliable horizontal variograms from wells because the well spacing is generally greater than the mean heterogeneity size. Different variogram ranges can then be tested, and the one which provides the best fit with production data is chosen (Eschard et al., 1998). Alternatively, data bases on analogous reservoirs or outcrops can be used to give an idea of the range.

Object-based models

Object-based models were designed to simulate geological objects which have a well-defined geometry. These generally correspond to sedimentary bodies isolated in a non-reservoir matrix (e.g. fluvial channels in floodplain mudstones).

Similarly, shale breaks with a specific geometry within a massive sandstone unit can also be generated as objects. Different types of objects can be simulated together. Attraction or repulsion functions can also be applied to promote or prevent connections between different types of objects. Objects also can be superposed with erosion or preservation rules. The vertical frequency of the objects can be computed at different levels from the well data.

Generally speaking, object-based simulations require a great deal of *a priori* geological information. First, the geologist has to determine and classify the different types of objects from the well data. Then, he has to impose a simplified shape on them. This requires a good knowledge of the geological setting, and also the possibility of consulting reservoir data bases containing the geometry (width versus thickness, etc) and shapes (sinuosity, etc) for each depositional environment. This information generally comes from outcrop studies, from observations of comparable modern depositional environments or from well-documented subsurface fields. It is generally summarised in the form of cross plots and frequency diagrams in which the variability of the measured parameters is shown. This variability is often high for a given sedimentary object because the shapes and the amalgamation rate of the objects depend on their location within a sequence (Eschard et al., 1999). During periods when there is little space available for sedimentation, channels deeply incise the floodplain mudstones, and the amalgamation rate of fluvial channels is high. When this space is large, the floodplain aggrades, and channels become less incised in floodplain mudstones. However, although this type of information is critical for controlling the object size, it is rarely available in reservoir data bases.

Siliclastic reservoirs can often be simulated by object based methods. The most classical objects that are simulated with this approach are fluvial channels and crevasse splays in a matrix composed of floodplain mudstones. Channels can either be straight or sinuous, isolated in the floodplain or amalgamated. Crevasse splays are connected to the channels.

Nested simulations

Nested simulations combine object and sequence-based pixel models. Typically, the objects are simulated first, then filled with lithotypes using a sequence-based simulation. This approach is generally used to simulate large geological objects which present internal heterogeneities. For example, the heterogeneity in fluvial and estuarine channels corresponds to mudstone plugs deposited within the channels. The object, the channel, is simulated first, then filled by heterogeneous material using a sequence-based algorithm.

Building up the reservoir or orebody model

The first step in a plurigaussian simulation of an orebody or a reservoir is to calculate the proportions of each facies. But even before doing this, we have to make a series of preliminary choices. We have to define the lithotypes to be studied, to divide the reservoir or orebody into units, to choose the references levels to be used for flattening each unit and finally to choose the parameters of the grid to be simulated.

Step 1: Defining the lithotypes

The lithotypes constituting the orebody or reservoir have to be defined from the available data: core samples, well logs and seismic data. They have to honour the geological information, and where applicable, the petrophysical information. In general, the geologists define more facies than can reasonably be simulated. So these have to be grouped into lithotypes.

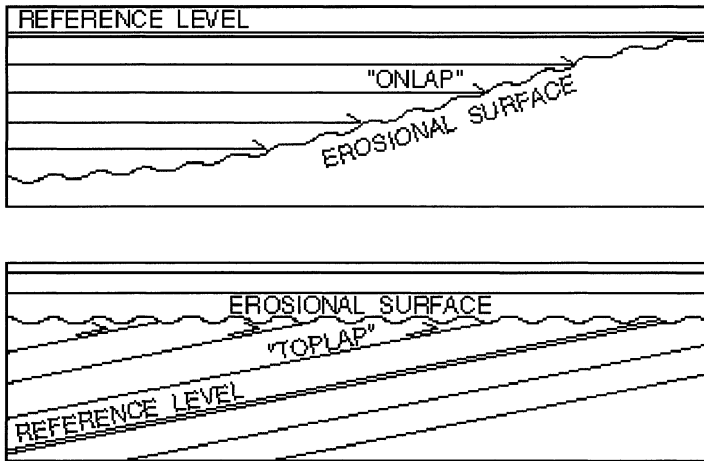
The definition adopted is crucial, firstly in the data integration process and later in the simulations themselves. For example, as well as reproducing the geology of the reservoir realistically, reservoir simulations must reproduce the key reservoir properties (porosity, relative permeability, capillary pressure, etc) from a fluid flow point of view. This implies reconciling the geological and petrophysical descriptions of the reservoir when the lithotypes are first defined. Furthermore, the lithotype data base must be homogeneous in all the wells.

Step 2: Dividing reservoir or orebody into units

Reservoirs generally consist of several stacked units. Their geology may differ from one unit to another. So each unit must be simulated independently with different parameters. Variogram ranges and anisotropies are generally different for each of the units. Sometimes the geostatistical techniques used can vary from one unit to another, depending on their geology. It is not possible to correlate lithotypes across unit boundaries, and moreover the simulation grid follows the geometry of the units. This is why the definition of the reservoir units is a such critical step in reservoir modelling.

Step 3: Defining the reference level

The reference level for the simulation is a specific geological layer which is used to restore the geometry of the orebody or reservoir at the time of deposition. This level must have been deposited horizontally during sedimentation and should, if



Figs 7.1 and 7.2: Onlap (above) or toplap (below) configurations are produced depending on the choice of a reference level relative to the unit geometry

possible, correspond to a time line. The reservoir is then flattened using this as the reference level. Different reference levels can be used for each reservoir unit, or one can serve for several of them.

As for the definition of the lithotypes, the choice of the reference level has a marked impact on the result of the simulation. For example, onlap or toplap configurations can be produced depending on the choice of a reference level with regards to the unit geometry (Figures 7.1 and 7.2).

Step 4: Choosing the grid spacing

The grid size is in theory directly dependent on the size of the heterogeneities to be reproduced in the simulation. The detail needed in a simulation also depends on the recovery process used to produce the reservoir. Gas is less sensitive than oil to small scale reservoir heterogeneities. Fields developed with horizontal wells generally require very detailed reservoir models. From a geological point of view, it is tempting to simulate as much detail as possible, but in practice, the number of cells is limited by computer capacity. Because of the limitations on fluid flow simulators, the fine grid simulated has to be upscaled to a coarser one. Within each cell of the latter grid, the upscaling algorithm has to summarise all the values in the fine grid in a single value. This is quite simple for porosity where it means taking the average, but it is more difficult for permeability. Only approximate solutions exist (unless we run a fluid flow simulator for each cell of the coarse grid, and that would be prohibitively time consuming). Consequently many different upscaling techniques exist and so differences may appear in the simulation depending on which upscaling method is applied, especially when the number of cells is very large.

Petroleum applications of the plurigaussian approach

The geostatistical methods presented above have their limitations when applied in complex geological settings. In some cases, the reservoir architecture presents complex facies transitions which cannot be simulated with mono-gaussian techniques. Similarly, boolean models are only suitable when only a few types of sedimentary bodies are present and when their shapes are well defined. The main advantage of the plurigaussian approach is its ability to handle complex facies relationships, both vertically and laterally, in a pixel simulation. For example, it is possible to impose different anisotropies onto the gaussian functions. Furthermore, different types of variograms each with its own range and anisotropy, can be used for each gaussian function. The combination of all these parameters makes the approach very flexible.

Constructing the lithotype rule with geological constraints.

As the theory of the lithotype rule construction has already been described, we will illustrate it with several examples to show how the rules can be chosen *a priori* depending on the geological context. As usual, lithotypes are represented by different colours and the surface occupied by each one in the matrix is a function of the relative proportion of facies computed in wells. Complex facies transitions can then be reproduced by changing the relative position and surface of the lithotype in the matrix.

If only one gaussian function is used, the lithotype rule only shows superposed bands, with each band representing a lithofacies. When colours are in contact in the lithotype rule, the corresponding lithotypes will be connected in the simulation. Consequently the ordering of the facies in the lithotype rule must then respect a sedimentary sequence in order to correctly reproduce facies inter-fingerings in the simulation. One of the most common sequences encountered corresponds to the progradation of a shoreface (top of Figure 7.3), in which offshore, shoreface, foreshore and coastal plain deposits are organised in a prograding sequence. Its lithotype rule is also shown in Figure 7.3.

Simulation of reservoir with complex facies transitions

The lithotype rule also makes it possible to simulate complex geological settings, especially in carbonate depositional systems. For example, it is difficult to simulate the irregular shape and complex internal architecture of algal or reefal bioherms with classical object or sequence based models (Van Buchem et al, 2000). In this case, the lithotype rule is used together with the vertical proportion curve as a kind of facies substitution diagram to reproduce the sequential organisation of the facies.

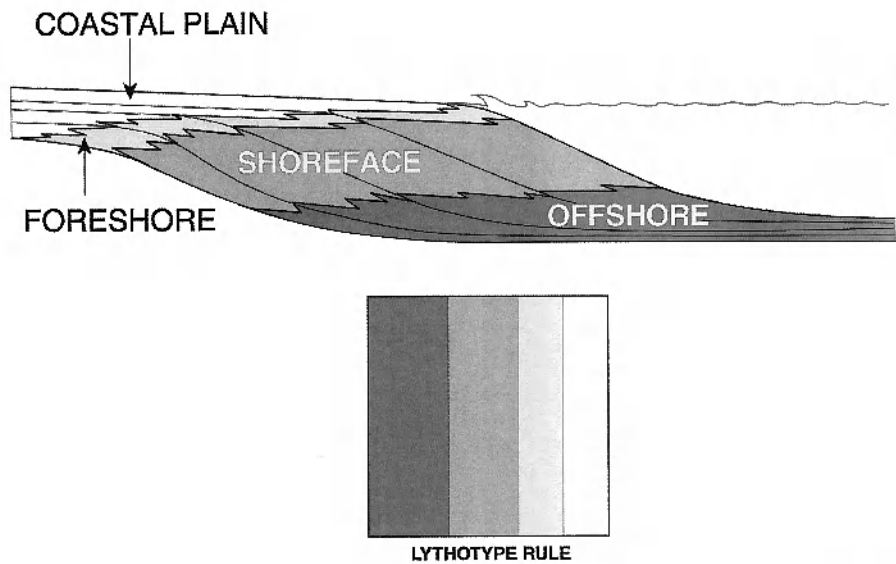


Fig 7.3: Lithotype rule with only one gaussian function used to simulate a prograding sequence. Facies ordering is respected in the rule

Figure 7.4 shows algal mounds in the Paradox basin, of Pennsylvanian age, which were studied in outcrops along the San Juan River (Grammer et al., 2000, and Van Buchem et al, 2000). The algal bioherms correspond to bioconstructions with irregular rounded shapes which were constructed in a mixed carbonate and siliciclastic open shelf. The carbonate shelf first prograded (facies 1 to 3), and then when a certain water depth was reached, incipient mounds (facies 4) started to develop in the shelf setting.

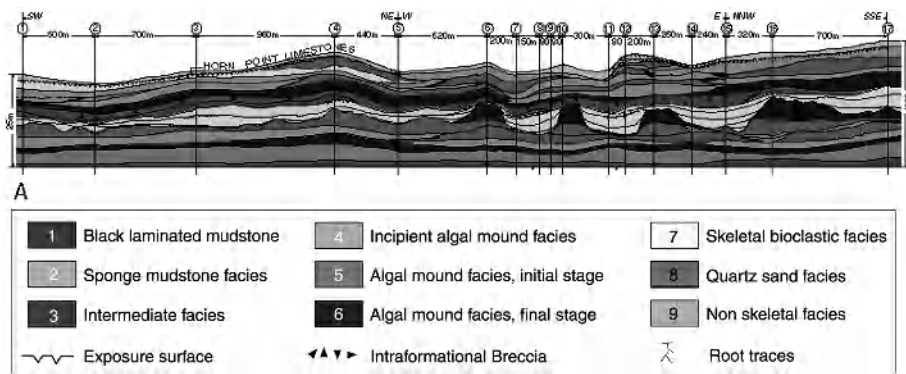


Fig 7.4: Geological model of Pennsylvanian algal mounds in outcrops (from Galli et al., in press)



Fig 7.5: Lithotype rules for the upper (on the right) and lower units of Paradox basin

After this initial stage, the mounds themselves started to grow, and different stages of construction have been observed (facies 5: initial stage, facies 6: final stage). At the end, the mounds have an irregular rounded shape with a height of more than 20m compared to the surrounding shelf. When the algal mound growth stopped, the inter-mounds troughs were progressively filled by in situ shelf sediments and material falling from the mounds. These deposits now form flanking beds in outcrops (facies 7), with the beds onlapping the mound relief. Finally, clastic deposits (facies 8) were deposited in the shelf setting during subsequent relative sea-level drops. On top of the sequence, karstification often restricted marine carbonates (facies 9) during emersion.

The lithotype rule corresponding to this complex geological setting is shown in Figure 7.5. Firstly, the series was divided in two units: the lower unit corresponds to the platform progradation, whereas the algal mounds are located in the upper unit. The lithotype rule of the lower unit is simple as only one gaussian function is used for its simulation. Its vertical proportion curve also reflects the progradational pattern of the shelf.

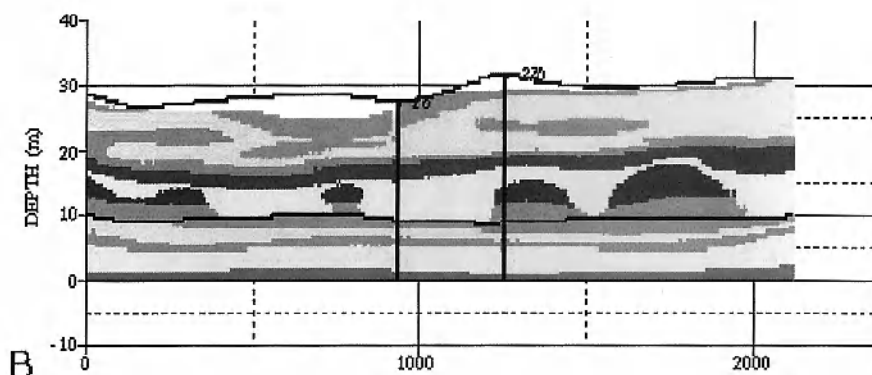


Fig 7.6: Cross section in a simulation of the Pennsylvanian algal mounds using the pluri-gaussian approach. The mound geometry is realistically reproduced (compare with the geological model in fig. 7.4)

In the upper unit, the lithotype rule is more complex because of the spatial organisation of the mounds and intermound facies. The intermound facies (facies 7) can be in contact with the two mounds facies (facies 5 and 6) and with the sandstones capping both the mounds and the inter-mounds deposits (facies 8). The simulation in Figure 7.6 reproduces the geometry of the mounds in this shallow marine platform setting, very realistically both in section and in plan view.

Simulation of the effects of primary diagenesis in complex reservoir

Plurigaussian simulations are particularly suitable for simulating interdependent phenomena which interacted during sedimentation or just afterwards. Moreover, the two phenomena can have different anisotropies, and can be correlated or not. The effects of the primary diagenesis, which affect the sediments just after burial, can be simulated with this approach.

The Paradox basin algal mounds described earlier are a good example of the effects of primary diagenesis. During relative sea-level drops, the topmost part of the mounds was altered when the mound top is emerged. This alteration is a function of the initial carbonate texture and porosity. The flanking beds which were deposited after this episode, were not affected. A new facies corresponding to this alteration of the reef was added into the lithotype rule (facies 9). It is in contact with both the flanking beds laterally, the mound core below and the sandstones above. The simulation in Figure 7.7 reproduces the altered facies capping the mounds.

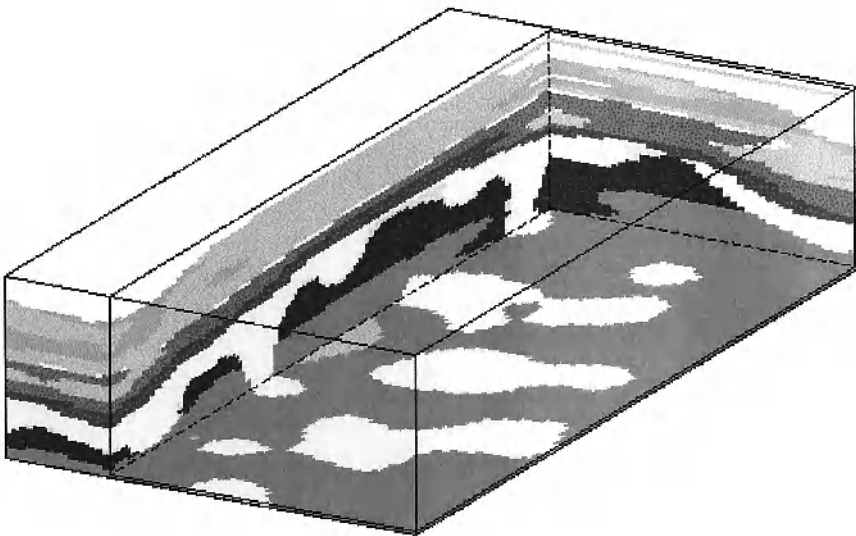


Fig 7.7: Simulation showing the altered facies capping the mounds. The altered facies corresponds to the light grey colour on top of the dark algal mound

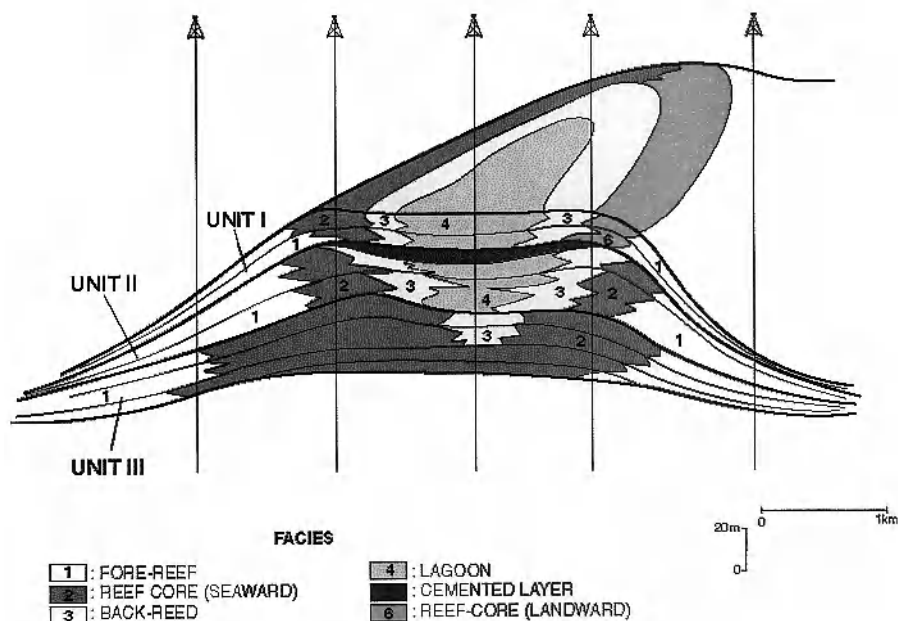


Fig 7.8: Synthetic model of a reef complex from which a pseudo-well data-base was compiled for simulation purposes

A synthetic case study of a reef reservoir was carried out in order to test the ability of the plurigaussian method to simulate this type of reservoir. It is based on the Miocene reefs observed in subsurface by Grötsch and Mercadier (1999). In this synthetic case (Figure 7.8), the reef growth occurs in three stages. Each stage is characterised by a specific organisation of the facies. The basal unit (unit III) corresponds to the initial stage of the reef growth when no lagoon had developed inside the reef. The middle unit (unit II) corresponds to the growth stage of the reef. Back-reef and lagunal deposits were well developed inside the reef. The topmost layers of unit II were cemented during a relative sea-level drop when the topmost part of the reef emerged. Unit I is the final stage of the reef growth. According to Grötsch and Mercadier (1999), cementation can specifically affect the core reef exposed to the ocean but not the lee-side where porosity is preserved.

Six facies were differentiated for the simulation: fore reef (facies 1), core reef, basin-ward (facies 2), back-reef (facies 3), lagoon (facies 4), cemented facies (facies 5), and core-reef landward (facies 6). Facies 5 corresponds to a cemented horizon which affected facies 2, 3 and 4 in unit II. The cemented horizon constitutes a significant permeability barrier within the reservoir, and so it is important to reproduce it in the simulation. Three litho-stratigraphic units were then used for the reservoir simulation, each with a specific facies distribution. The simulation was processed in a proportional grid within a unit, in order to get correlation lines parallel to the top and bottom of each unit.

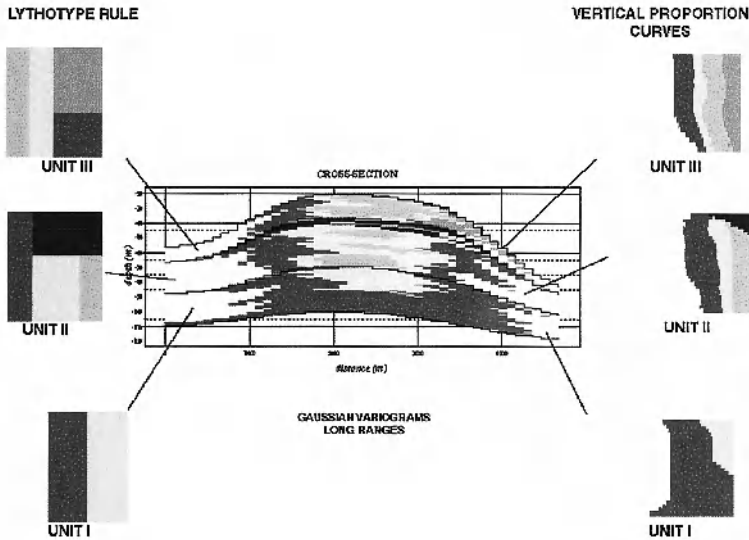


Fig 7.9: Vertical section extracted from a plurigaussian simulation of the reefal complex which was divided into three units. These are shown with their lithotype rules (left) and their vertical proportion curves (right).

In unit III, the truncated gaussian algorithm was used with a simple lithotype rule (Figure 7.9). In unit II, the plurigaussian algorithm was applied in order to respect the complex relationship between the facies. The cemented facies (facies 5) can be in contact with facies 2, 3 and 4 but facies 5 cannot be in contact with facies 1. The vertical proportion curve also constrains facies 5 only to the topmost part of the unit II. In unit I, the porous core reef (facies 6) is only present in the lee side oriented towards the continent. The lithotype rule prevents contact between facies 6 and 4. Gaussian variograms were used for the two gaussian functions in order to respect the rounded shape of the reef facies belts. The continuity of the facies distribution was also ensured by using long variogram ranges of the 6000m in the East-West direction and 15000m in the North-South direction. This anisotropy imposes an ellipsoidal shape on the simulated reef. The results of the simulation are shown in Figure 7.9 in a vertical section across the reef, and in Figure 7.10 in plan view. The complex internal organisation of the reef has been reproduced well in the simulation.

Simulating progradational patterns

In a progradational sequence, the facies are organised in a logical order from the deepest facies at the base of sequence to the shallowest on top. Laterally, the geometry of a progradation is marked by oblique clinoforms each corresponding to a depositional surface.

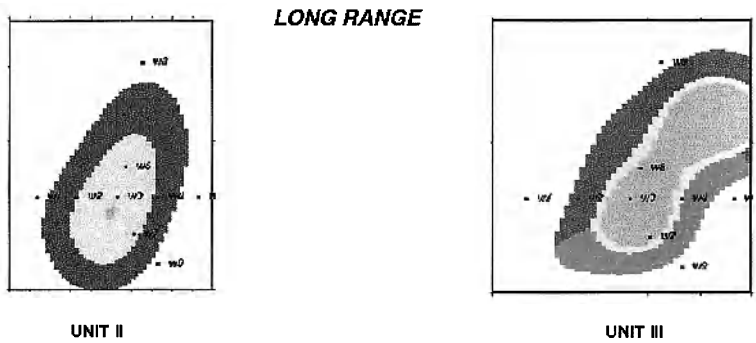


Fig 7.10: Plan view of the reef simulation, using a gaussian variogram with a long range in Unit II (a) and III (b). See Fig 7.9 for unit names

The facies belts are then organised obliquely with their dip corresponding to the direction of progradation (Figure 7.11). In it lithofacies 1 is separated from lithofacies 4 by lithofacies 2 and 3. The first gaussian function controls the thresholds of lithofacies 1 and 4, and the second, lithofacies 2 and 3. An anisotropy was added to simulate the dip of facies 3 within facies 4.

Simulation of fractures affecting a specific facies.

Fracturing which affects some reservoirs after burial can also be simulated with a stochastic approach (Cacas et al., 2001). Fracture networks are simulated as objects in a reservoir matrix. In many cases, there is a relationship between the fracture distribution and the initial lithotypes because the mechanical properties of the latter depend on their lithology and textures. When stressed during burial and tectonic deformation, fracturing affects the lithotypes differentially. Some of them end up intensively fractured, whereas others do not. So it is important to correlate the fracture distribution with the lithotype.

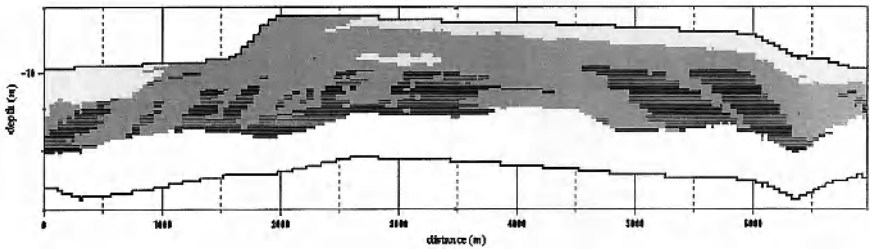


Fig 7.11: Simulation of a progradation

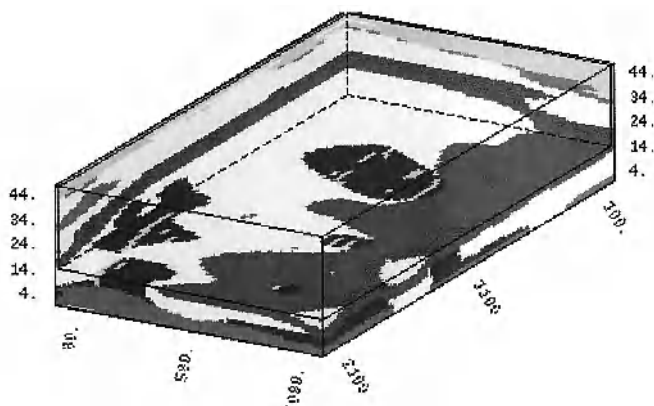


Fig 7.12: Simulation of fractures only affecting one facies within the Pennsylvanian algal mounds

We simulated algal mounds with fractures in them. We consider that only the mound facies (facies 5 and 6) have been fractured significantly. The distribution of the facies was simulated first with the plurigaussian method. Then, fractures were only simulated in the mound facies. Figure 7.12 presents the results.

Testing the impact of simulation parameters.

Different tests were carried out to evaluate the sensitivity of the simulations to variations in input parameters. The synthetic reef case study presented in Figure 7.8 was used as a reference to evaluate the effect of changing the type of variogram model and its range. The influence of the range was tested on unit II. In the reference simulation (Figures 7.9 and 7.10), the horizontal ranges used were 6000m and 15000m in the X and Y directions respectively. In Figure 7.13, smaller ranges were used (X = 1500m, Y=1300m).



Fig 7.13: Testing the variogram range influence on the reefal complex described in Figure 7.8. The figure shows a plan view of the reef using variograms with short ranges. Compare with Figure 7.10 in which the same plan was simulated with a long range

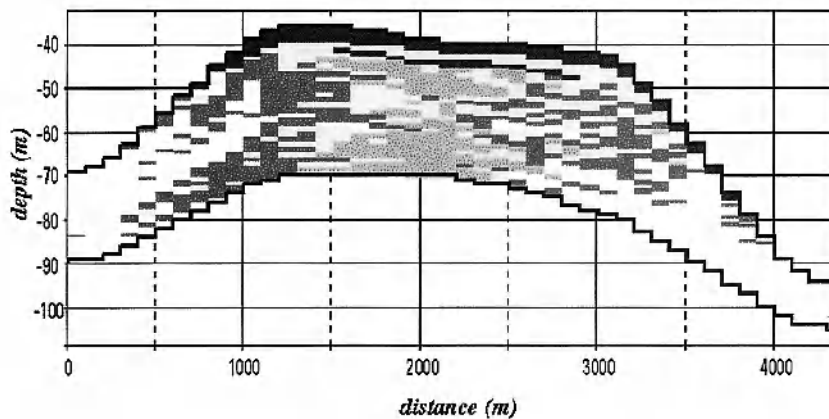


Fig 7.14: Testing the variogram model. The simulation represents a cross section of the reef complex described in Figure 7.8. An exponential variogram model was used here. Compare with Figure 7.9 in which the same reef complex was simulated with a gaussian variogram

Comparing the two simulations shows that the range reduction has dramatic consequences on the shape and distribution of the facies belts. Small isolated reefs are produced by the smaller ranges. The sensitivity of the simulations to the variogram model was then tested. In the reference simulation, two gaussian variograms with ranges of $X = 6000\text{m}$, $Y=15000\text{m}$, $Z= 5\text{m}$ were used for each gaussian function. Figure 7.14 shows the results produced using exponential variograms with similar ranges to the reference simulation. The consequences of this change are again very important. The facies are more scattered in the simulation using an exponential variogram models than in the one using gaussian variograms. So the type of variogram model can have important consequences when computing facies connectivity within a reservoir.

8 FREeware

Introduction

You will find a CD containing two free programmes called **pluri_demo_simu** and **pluri_demo_vario** plus a set of colour images, in the pocket in the back cover of this book. These two programmes were designed to allow users to discover plurigaussian simulations by “playing” with them. The first one makes it possible to visualise simple plurigaussian simulations and to see how changes in the input parameters affect the resulting images. The second one is more technical. It illustrates the relationship between the indicator variogram and those of the underlying gaussian random functions (RFs for short) that are truncated to give the indicators. For each programme we provide a set of graded exercises showing how to run the programme and how the key parameters affect the simulated images or the indicator variograms, as the case may be.

Installation

The package is written in java for the user interface and in C/C++ for the computing algorithm. It runs on INTEL compatible platforms under win95, win98, winnt4.0, win2000 or linux (tested on RedHat distributions) and on SPARC platforms under Solaris 7/8.

Installation instructions for Windows users

On the CD, the directory *Windows* contains 2 files:

- **Jre117Bi-win32.exe** : the java runtime environment, necessary to run the demo programs
- **PluriDemo.zip** : the archive containing the demo programs

To install:

1. Download the two files into a temporary directory.
2. Install the java runtime environment by running the file **Jre117Bi-win32.exe**

3. Unzip the archive file *PluriDemo.zip* in the destination directory (for example in **C:\Program Files**)

Installation instructions for Solaris users

On the CD, the directory *Sun-Solaris* contains 1 file:

- **PluriDemo.tar.Z** : the archive containing the demo programs

To install:

1. Download the file in a temporary directory (for example */tmp*).
2. Check you can start the command *java* or *appletviewer*, if not, ask your system manager.
3. Uncompress the archive file *PluriDemo.tar.Z*:
uncompress PluriDemo.tar.Z
4. Extract the files in the destination directory (for example in: */usr/local*)

```
cd INSTALL_DIR (for example: /usr/local)  
tar xvf TEMP_DIR/PluriDemo.tar
```

5. Edit the 2 scripts *pluri_demo_simu* and *pluri_demo_vario* in the directory **INSTALL_DIR/PluriDemo/sparc-solaris** and set the environment variables as appropriate for your installation.
6. Modify your *PATH* environment variable (for *csh/tcsh* users: ***setenv PATH INSTALL_DIR/PluriDemo/sparc-solaris:\$PATH***).

Installation instructions for Linux users

On the CD, the directory *Linux-x86* contains 2 files:

- **jre118_v3-glibc-2.1.3.tar.bz2** : the java runtime environment, necessary to run the demo programs
- **PluriDemo.tar.gz** : the archive containing the demo programs

To install:

1. Download the 2 files into a temporary directory (for example */tmp*).
2. Install the package containing the java runtime environment.

Choose an installation directory `INSTALL_DIR` (for example `/usr/local`)

```
cd INSTALL_DIR
```

```
tar -Ixf TEMP_DIR/jre118_v3-glibc-2.1.3.tar.bz2
```

3. Extract the PluriDemo files in the destination directory (for example in: */usr/local*)

```
cd INSTALL_DIR
```

```
tar zxvf TEMP_DIR/PluriDemo.tar.gz
```

4. Edit the 2 scripts *pluri_demo_simu* and *pluri_demo_vario* in the directory **INSTALL_DIR/PluriDemo/sparc-solaris** and set the environment variables as appropriate for your installation.
5. Modify your *PATH* environment variable (for *csh/tcsh* users: **setenv PATH INSTALL_DIR/PluriDemo/linux:\$PATH**).

Description of Pluri_Demo_Simu

The user has to select two underlying gaussian images, out of 12 preprogrammed images. The variograms used to generate these images consist of 3 exponential models, 3 gaussian models, 3 cardinal sine models and 3 sphericals. Next the user must choose the “rock type rule” which defines the number of facies and the relationship between them. Thirty-six rock type rules are currently available. Lastly the user specifies the proportions of each facies. In the stationary case, these would be the same at all points in the area under study but as oil reservoirs are usually nonstationary in the vertical direction, it is instructive see what happens if the proportions can be varied. So this program allows the user to vary the proportions linearly from the top of the image to the bottom. Now you can start using the programme.

To start the programme in a **Windows** environment double-click onto the file:

C:\Program Files\Pluridemo\windows\pluri_s.bat

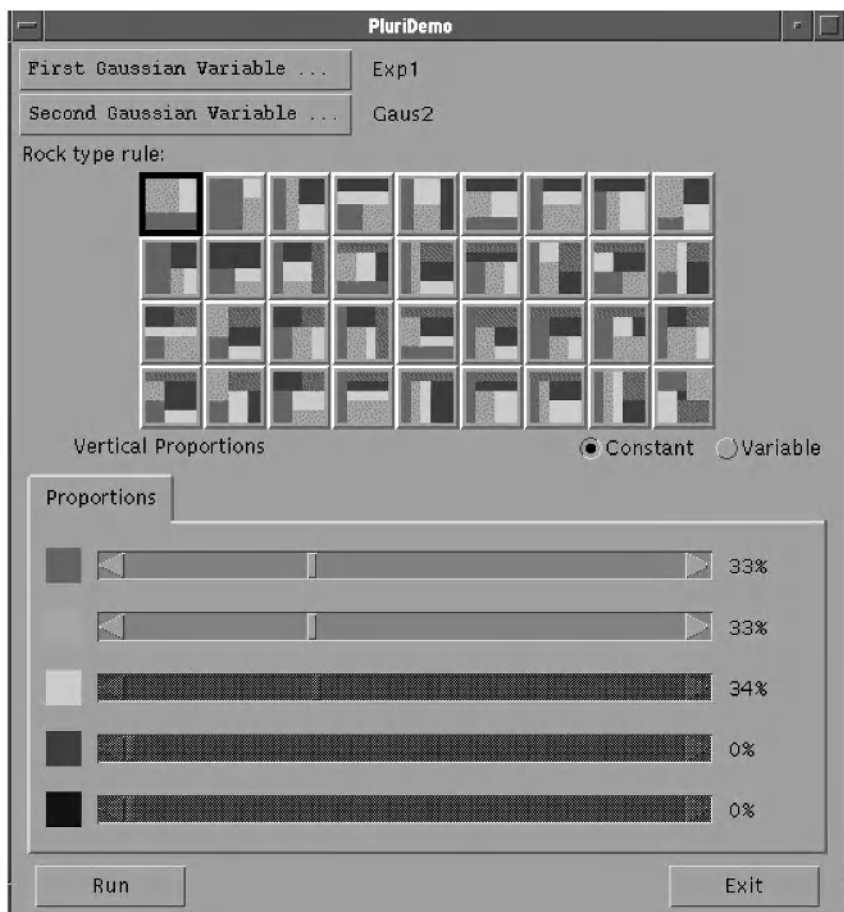


Fig 1: Screen that comes onto the screen when `pluri_demo_simu` is started

To start the programme in a solaris or linux environment type in the command:

`pluri_demo_simu &`

Choosing two of the underlying gaussian images

The program contains 12 preprogrammed images that were obtained by simulating standard gaussian random functions with different variograms (3 exponentials, 3 gaussians, 3 cardinal sines and 3 sphericals). If you click on the rectangle

First gaussian variable.

a new panel will pop up showing the 12 simulated images. Although the three images in each row have the same type of variogram model, they have different ranges and anisotropies. Table A1 lists the parameters of these 12 models. Compare the images in the top row (exponentials) and the bottom one (sphericals) with those in the other two rows. The images generated using the exponentials and the sphericals have a patchy appearance compared to the smoother images given by the gaussian and cardinal sine models. This is due to the linearity of the exponential and spherical models near the origin compared to the quadratic shape of the other two. To change the simulated image being used, just click on the image you want. Its name should appear beside the corresponding rectangle. For the first example we will continue with the default values for parameters; that is

First gaussian variable

Exp1

Second gaussian variable

Gaus2

Table A1: Parameters of the variogram models used to generate the 12 images. (N.B. Angles are measures counter-clockwise starting from east)

	Variogram Type	Long Range	Short Range	Orientation
Exp1	Exponential	35	20	170°
Exp2	Exponential	105	25	40°
Exp3	Exponential	60	6	95°
Gaus1	Gaussian	50	7	0°
Gaus2	Gaussian	40	30	120°
Gaus3	Gaussian	40	18	60°
Sine1	Cardinal sine	70	55	20°
Sine2	Cardinal sine	150	20	140°
Sine3	Cardinal sine	90	35	85°
Sph1	Spherical	80	15	10°
Sph2	Spherical	60	30	160°
Sph3	Spherical	60	50	35°

Choosing the rock-type rule

The thirty-six possible rock-type rules are displayed in 4 rows in the middle of the pluri-demo panel. This collection is by no means exhaustive. We chose to limit the programme to the rules that generate three, four or five facies, not more. The two in the top left corner contain only three facies; the next ones contain four facies and so on. Secondly, only rectangular partitions have been used because it simplifies the calculation of the thresholds once the proportions are given.

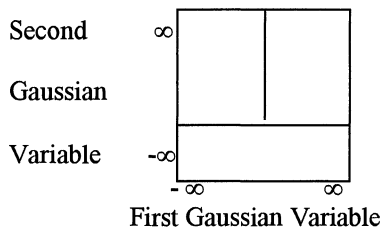


Fig 2: Diagram representing a rock rule

Each square icon is a symbolic representation of the thresholds separating facies. The horizontal axis representing the first gaussian variable goes from $-\infty$ on the left to $+\infty$ on the right; and similarly for the second gaussian variable on the vertical axis. Figure 2 shows an enlarged version of the first rule shown in the top left corner.

Now to explain how to interpret the diagram. As the square is divided into three regions coloured orange, yellow and green, this rule generates 3 facies. Any point in the top left area would have a negative value for the first gaussian whereas its value for the second gaussian would have to be greater than a certain threshold. Similarly any point in the top right area would have a positive value for the first gaussian whereas its value for the second gaussian would be greater than that threshold. Finally any point in the bottom region could have any value for the first gaussian but its value for the second gaussian would be less than the threshold.

In this case we are using two thresholds, one for each gaussian. Their values determine what proportion of each facies there is. Decreasing the value of the threshold on the second gaussian would decrease the amount of the green facies but would increase the amounts of yellow and orange. In the default case the facies are programmed to have equal proportions (here 33%). We will start with this as our first example.

Exercise A1 : Default values for parameters

Set the parameters to their default values :

First gaussian variable

Exp1

Second gaussian variable

Gaus2

Rock type rule



Constant proportions

33%, 33% & 33% for facies throughout the area.

Now click on the run button. The panel that appears on the screen (Fig A1) is divided into four quadrants. The resulting plurigaussian image appears in the bottom left quadrant. The two images that were used in constructing it are shown in the top two quadrants. The bottom right square shows the rock type rule that was used to generate it. The green foreground in the plurigaussian image corresponds to the lowest values of the second gaussian. The contour lines around the green foreground are smooth because its variogram is gaussian. The contour separating the orange from the yellow is broken up because the corresponding variogram is exponential.


Exercise A2 : Inverting the two images

What do you think would happen if we reversed the images that were used for the first and second gaussian variables? (See Fig A2). That is, if we keep the same rock type rule and the same proportions but put:

First gaussian variable	Gaus2
Second gaussian variable	Exp1

Exercise A3 : Effect of anisotropy

To see how anisotropy affects the results, set the first gaussian to Gaus1 which has a strong east-west anisotropy and reset the second gaussian to Gaus2. Keep all the other parameters the same. That is,

First gaussian variable	Gaus1
Second gaussian variable	Gaus2
Rock type rule	
Constant proportions	33% for each facies throughout the area.

The background in the resulting image (Fig A3) now has a marked east-west anisotropy

Exercise A4 : Changing the amount of each facies

Keeping the same underlying gaussian images, we now increase the proportion of green by sliding the first bar from 33% to 60%. Note the second bar automatically drops, so as to maintain a total of 100%. Now slide the second bar up to 25%. The third one moves down to 15% to compensate. Now click on run, giving Fig A4. There is much more green now (60% instead of 33%). Although it is less obvious, there is now more orange than yellow.

Exercise A5 : Changing the rock-type rule - four facies

To change the number of facies we have to change the rock type rule. For example, the top left one has three facies but the third one from the left in the top row has four. Change to this rock type rule, keeping the same underlying gaussian variables and accepting the default proportions (25% each).

First gaussian variable

Gaus3

Second gaussian variable

Gaus2

Rock type rule



Constant proportions

25% for each facies
throughout the area.

Figure A5 shows the new image. Does the green colour ever touch the red or the yellow? Looking back to the rock type rule, you can see that the colours do not touch there either. The green is always surrounded by a halo of orange. Work out which colours can touch? (Note: if the pixel size used to print the images is too large, colours which should not touch appear to do so. With smaller pixels the intermediate colours would show up.)

The next big step in learning about plurigaussian simulations is to allow the threshold to vary. But before we do this, we give the reader some exercises to try for himself/herself.

Exercise A6

Set the parameters to:

First gaussian variable

Gaus2

Second gaussian variable

Exp1

Rock type rule



Constant proportions

33% for each facies
throughout the area.

Compared to Exercise 1, we have reversed the first and second gaussian variables and the rock type rule has been rotated through 90°. The resulting image (Fig A6) is exactly the same. At first glance it seems as if these two rock rules are equivalent (provided that the rock rules are reversed), but this is only true if the proportions are constant throughout the area. As we will see in Exercise 5, this is no longer true when the proportions are variable. This is why we have included both rules.

Exercise A7

Set the parameters to:

First gaussian variable

Gaus3

Second gaussian variable

Gaus2

Rock type rule



Constant proportions

20% for each facies
throughout the area.

Before clicking on run, try to imagine what sort of image will be produced. How many facies will there be? Which ones will touch each other? What will the directions of anisotropy be? Now click on run and see whether you were right (Fig A7).

There is an important difference between this rock type rule and the other 35 rules that are available in Pluri-demo. In all the other rock type rules including the one shown below there is at least one line of demarcation going right from $-\infty$ to $+\infty$ (either horizontally or vertically).

Rock type rule



We could call this a primary line of demarcation. The other line is a secondary demarcation. The (one or more) facies delimited by a primary demarcation can be

simulated first. Those delimited by a secondary demarcation can be simulated in what remains, and so on for lower order demarcation lines.

It would be possible to simulate the three facies shown in the rule above by successive applications of standard simulation packages. In the first gaussian variable, all points where the value is less than the threshold would belong to the green facies; those with larger values could be either orange or yellow. The split between these two colours is determined by the value of the second gaussian. This is not true of the 5 facies rule given above. This type of image can only be generated using the values of both gaussians simultaneously.

Exercise A8

Figure A8 shows a plurigaussian image that was generated using constant and equal proportions and the following rock type rule. Can you work out which of the 12 underlying simulated images were used for the first and second gaussian variables?



Exercise A9

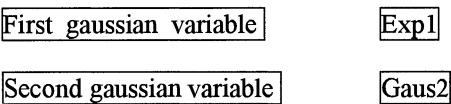
Figure A9 shows a plurigaussian image that was generated using the Gaus3 image for the first gaussian variable and the Sph2 one for the second variable together with constant and equal proportions. Can you work out which rock type rule was used? (Figure A9b gives the answer).

Variable proportions

Having seen how plurigaussian simulations work in the stationary case, we now treat the more realistic case of vertical nonstationarity.

Exercise A10 : Variable proportions

The simplest way to demonstrate how the variable proportions work is by using the same basic data as in Exercise 1:



Rock type rule



However we will vary the proportions from the top to the bottom. Instead of the button for constant proportions, click on **Variable** . Now we have to define the proportions at the top and the bottom of the image. The default values for the proportions are 33% each. To accentuate the effect of changing the proportions set the proportion of green to 0% at the top and 100% at the bottom. Any remaining proportion is divided equally between the yellow and the orange. The program automatically varies the proportions linearly from the top to the bottom. So halfway down, there will be 50% green, 25% yellow and 25% orange. As the program has to calculate the thresholds for each vertical level so that the proportions vary linearly, it takes more time to react than for constant proportions. Now click on run. Figure A10 shows the results.

Compared to Exercise 1, the images in the top two squares remain the same. The most obvious changes occur in the bottom right square which now shows three diagrams instead of one. The long diagram on the left shows the proportions varying linearly from 50% yellow, 50% orange and 0% green at the top, to 100% green, 0% yellow and 0% orange at the bottom. The little square at the top right shows the rock type rule at the top (with no green) while the lower square shows it at the bottom of the image (all green).

Looking at the simulated image in the lower left quadrant, there is, as expected, no green at the top and correspondingly, no yellow and no orange at the bottom. The proportion of green increases steadily going down the image.

Exercise A11: Variable proportions on two gaussian variables

Replace the exponential in the first gaussian variable by one with an anisotropic gaussian variogram and see how the resulting image changes (Fig A11).

First gaussian variable

Gaus3

Second gaussian variable

Gaus2

Rock type rule



Top proportion

0% green, 50% yellow, 50% orange

Bottom proportion

100% green, 0% yellow, 0% orange

The green foreground (controlled by Gaus2) is the same as before but the edges between the orange and the yellow are smooth rather than being “furry”.

Exercise A12 : Working out which rock type rule was used

Figures A12 and A13 show plurigaussian images that were created using Gaus3 and Gaus2 as the two underlying gaussian variables. By looking at the number of facies and which ones are in the foreground and the background, deduce which rock type rule has been used and also what the top and bottom proportions were. The answers are given in Figures A14 and A15.

Limitations of the program

As this program was designed to demonstrate how plurigaussian simulations work, it does not have the full capacity of commercial software. This paragraph outlines its main limitations:

- ☐ The program is limited to two independent gaussian variables. No allowance has been made for correlations between the variables.
- ☐ Twelve simulated images were pre-programmed. The user cannot generate more images with the same variogram models or with others.
- ☐ Only 36 rock type rules are available. To simplify the calculation of thresholds the partition is rectangular. Then the number of facies was limited to 3, 4 or 5. Even with these restrictions the catalogue of rock type rules is by no means exhaustive. For example each colour only appears once.
- ☐ The program allows for nonstationarity in the vertical direction (but no other) and then only a linear change in the proportions from top to bottom. In practice the cyclic changes in the geological environment over time mean that the variations are far more complex.
- ☐ Lastly, the program gives only nonconditional simulations. No conditioning data can be included. But the conditioning step is possible in practical cases.

Description of Pluri_demo_vario

The second demonstration programme, **Pluri_demo_vario**, was designed, as its name suggests, to show what the indicator variograms corresponding to different types of underlying random functions look like. Exercises B1-B7 are designed to show you how the programme works. The next exercise illustrates the variogram for a truncated gaussian (i.e. only one gaussian RF). The other exercises involve plurigaussians (i.e. two or more RFs).

The programme asks the user to specify the types of variogram for the two underlying gaussian RFs together with their anisotropies and their ranges (or their practical ranges). Then it calculates and plots the corresponding indicator variograms and cross-variograms. The first screen is similar to that for **Pluri-demo-simu**. The same thirty-six rock-type rules are available, the proportions of each facies can be varied by moving the slide bar. One important difference between the two programmes is that this one allows the user to choose the

anisotropy for each underlying RF. The user specifies the main direction of the anisotropy (an angle between 0° and 180° measured counter-clockwise starting from east) and the ratio. Another difference is that the user can include a nugget effect.

To start the programme in a **Windows** environment double-click onto the file:

C:\Program Files\Pluridemo\windows\pluri_v.bat

To start the programme in a solaris or linux environment type in the command:

pluri_demo_vario &

Exercise B1: Default values for parameters

In the first exercise we use the default values for all the parameters. So the first RF has an isotropic exponential variogram whereas the second one has an isotropic gaussian variogram. Both have ranges of 50m. The proportions for the facies are 50% green, 33% orange and 17% yellow. Calculate the simple variograms by clicking on the "compute" button (Fig B1a). In order to see the cross-variograms rather than the simple variograms, click on the box marked "variograms", then slide to "cross-variograms" (Fig B1b). Note that in contrast to the simple variograms, the values are negative. Can you explain why? Click on the box "file" to return to the main panel.

Now we change the range of the exponential to 20m instead of 50m by sliding the bar. At this point the variogram models under study are:

Model 1: Exponential,	range = 20m,	anisotropy ratio = 100%
	direction = 0° ,	% nugget = 0%

Model 2: Gaussian,	range = 50m,	anisotropy ratio = 100%
	direction = 0° ,	% nugget = 0%

Calculate the corresponding simple variograms (Fig B1c). The green variogram remains the same as before but the range of the yellow one shortens.

Exercise B2: Changing the variogram model

Change the first model from an exponential to a gaussian RF. So now we have two gaussian RFs with ranges of 20m and 50m respectively:

Model 1: Gaussian,	range = 20m, direction = 0°,	anisotropy ratio = 100% % nugget = 0%
Model 2: Gaussian,	range = 50m, direction = 0°,	anisotropy ratio = 100% % nugget = 0%

Calculate their simple variograms (Fig B2a). Note that although both gaussian models are quadratic near the origin, the indicator variograms are not. They are always linear.

Now we are going to change the proportions of each colour. Slide the proportions bar to give 60% green, 20% yellow and 20% orange. Plot the simple variograms (Fig B2b) and the cross-variograms for Facies 1 (Fig B2c). Note how the sill of the green variogram increases compared to Figure B1a. Secondly the yellow and orange variograms are the same (or they would be if their proportions were exactly the same) and so one covers up the other.

Now look at the cross-variograms. The one shown in yellow is actually the cross-variogram between Facies 1 (yellow) versus Facies 3 (green). In order to get the other cross-variograms (e.g. for Facies 2), click on the box "facies", then onto Facies 2 (Fig B2d). Whereas the green one decreases monotonically, the yellow decreases at first then rises up again.

Exercise B3: Changing the rock-type rule

Now change the rock-type rule from three facies to the one with four shown below, with 30% green, 30% orange, 20% yellow and 20% red. (Both RFs still have gaussian variograms with ranges of 20m and 50m respectively).



Calculate the simple variograms (Figure B3a) and the cross-variograms (Figure B3b). In the first of these, the green variogram is not visible. The orange covers the green because their proportions are identical and their positions in the rock-type rule are symmetrical. The red variogram is different from the yellow one even though both are present 20% of the time. This is because they are not symmetrical.

The two underlying variograms are now:

Model 1: Gaussian,	range = 20m, direction = 0°,	anisotropy ratio = 100% % nugget = 0%
---------------------------	---------------------------------	--

Model 2: Gaussian,	range = 50m,	anisotropy ratio = 100%
	direction = 0°,	% nugget = 0%

Exercise B4: Effect of anisotropy

The anisotropy of the variogram models has a marked effect on the indicator variograms. Using the same rock-type rule and the same RFs as in exercise B3 (with gaussian variograms with ranges of 20m and 50m), click on the button "Show Anisotropies". This produces a 2D diagram showing the distance at which each variogram reaches 90% of its sill value as a function of direction (Fig B4a). In this case we see three circles coloured yellow, orange and red. The outer one in red has a radius of about 40m whereas the orange and yellow ones have radii of about 25m and 20m respectively. What do you think happened to the green circle?

Now change the anisotropy ratio for the longer range gaussian (i.e. with a range of 50m) from 100% to 50%. Plot the anisotropy diagram (Fig B4b). The circles have changed to ellipses. The red range in the NS direction has dropped from about 40m to about 22m. Up till now the variograms have all been isotropic so there was no point in changing the "Computation Direction" but now it becomes interesting to compare the different directional variograms. Fig B3a shows the simple variograms in the east direction (0°). Slide the computation direction bar to 90° and recalculate the simple variograms (Fig B4c). The sills are, of course, the same but the ranges are much shorter.

This time we increase the range of the first model from 20m to 50m, change its anisotropy ratio from 100% to 50% and change the main anisotropy direction from 0° to 90°. So its long range will be NS whereas the long range of the second structure is EW. The variograms and the computation angle are now:

Model 1: Gaussian,	range = 50m,	anisotropy ratio = 50%
	direction = 90°,	% nugget = 0%

Model 2: Gaussian,	range = 50m,	anisotropy ratio = 50%
	direction = 0°,	% nugget = 0%

Computation direction = 90°

Plot the anisotropy diagram (Fig B4d). There is a dramatic change in the shapes, especially for the orange facies. Now let's increase the range of the first structure from 50m to 100m while holding the anisotropy ratios constant at 50% (Fig B4e). We can produce even more dramatic effects by reducing both the anisotropy ratios to 20% (Fig B4f). Looking back at the rock-type rule, we see that

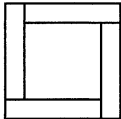
the red and yellow facies depend on cut-offs on only one RF and their variogram diagrams stay elliptical. It is the orange facies which depends on cut-offs on both RFs that produces the really unusual diagrams.

Model 1: Gaussian,	range = 100m,	anisotropy ratio = 20%
	direction = 90°,	% nugget = 0%

Model 2: Gaussian,	range = 50m,	anisotropy ratio = 20%
	direction = 0°,	% nugget = 0%

Computation direction = 90°

Now change to the five facies rule shown below with 12% green, 12% orange, 12% yellow, 52% red and 12% blue. Set both the variogram ranges to 100m with anisotropy ratios of 20%. The first structure is oriented at 90° (i.e. NS) whereas the second one is oriented EW. The anisotropy diagram is particularly interesting. (Fig B4g). As expected, the variograms for blue and green are identical and symmetrical with the opposite pair (yellow and orange). The cross shaped variogram for the red facies is particularly unusual.



Exercise B5: Adding a nugget effect

In order to see the impact of a nugget effect on the indicator variograms, we use the first rock-type rule (top left corner) with the default proportions (50% green, 33% orange and 17% yellow). Set the variogram models to gaussians with ranges of 100m and anisotropy ratios of 20%. As before the first one is oriented NS whereas the second is EW. Starting out with no nugget effect, calculate the variograms in the EW direction i.e. 0° (Fig B5a). Successively increase the amount of nugget effect present in the second structure to 25%, 50% and finally 100%. Figs B5b to B5d illustrate the evolution. Note the change in shape (particularly for the green facies) as the percentage of nugget effect increases.

Model 1: Gaussian,	range = 100m,	anisotropy ratio = 20%
	direction = 90°,	% nugget = 0%

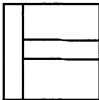
Model 2: Gaussian,	range = 100m,	anisotropy ratio = 20%
	direction = 0°,	% nugget = 0, 25, 50, 100%

Computation direction = 0°

Now change the direction for the variogram calculation to NS (90°) and recalculate the four variograms. (Fig B5e to B5h).

Exercise B6: Truncated gaussian case

First set the rock-type rule to:



Then change the proportions of each colour so that there is no green, 50% red, 30% yellow and 20% orange. As there is no green, the first gaussian plays no effective role and so it comes back to a truncated gaussian rather than a plurigaussian. Sketch the corresponding rock-type rule.

Now set both variograms to gaussians with ranges of 50m (with no anisotropy and no nugget effect). Plot the simple variograms (Fig B6a). Note the relative positions of their sills. These, of course, depend on the proportions of each facies. Now plot the cross-variograms for the red, orange and yellow facies. Fig B6b gives the one corresponding to facies 2 (orange).

Model 1: Gaussian,	range = 50m, direction = 0°,	anisotropy ratio = 100% % nugget = 0%
---------------------------	---------------------------------	--

Model 2: Gaussian,	range = 50m, direction = 0°,	anisotropy ratio = 100% % nugget = 0%
---------------------------	---------------------------------	--

Computation direction = 90°

Now change both variogram models to cardinal sines and plot the corresponding variograms (Fig B6c) and the cross-variograms for Facies 2 (Fig B6d). The hole effect in the parent variogram is passed on to the indicator variograms.

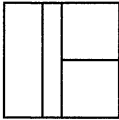
Model 1: Card sine,	range = 50m, direction = 0°,	anisotropy ratio = 100% % nugget = 0%
----------------------------	---------------------------------	--

Model 2: Card sine,	range = 50m, direction = 0°,	anisotropy ratio = 100% % nugget = 0%
----------------------------	---------------------------------	--

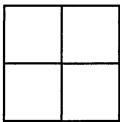
Computation direction = 90°

Exercise B7: Beatriz's example

Set the rock-type rule to the following with 25% per facies.



Set both variograms to gaussian models with scale factors of 25m and 75m respectively (no nugget effect and no anisotropy. Calculate the indicator variograms (Fig B7a) and the cross variograms for Facies 3 (Fig B7b). Without changing the other parameters, change the rock-type rule to the following:



Calculate the simple indicator variograms and the cross variogram for Facies 3. Fig 7c & d). Compare them to those obtained earlier.

9. REFERENCES

- Amit. Y., and Grenader. U. 1991 Comparing sweep strategies for Stochastic Relaxation. *J. Multivariate Anal.* 37 197-222
- Armstrong M. 1992. Positive definiteness is not enough. *Math. Geol.* (1). 135-143.
- Armstrong M. and Galli A.. 1999 Derivative based plurigaussian simulations, in *LAGM 99, Trondheim*, ed S.J. Lippard, A. Naess and R. Sinding-Larsen, pp591-596
- Azpirixaga I., Hernandez E. and Correa A. 1997. Uso del modelaje geostatístico gaussiano truncado en la definición de secuencias estratigráficas en yacimiento C6 del área VLA-8, bloque I, Lago de Maracaibo, In *Memorias del I Congreso Latinoamericano de Sedimentología*, Soc. Venezolana de Geol. Tomo 1, 59-65, Nov. 1997.
- Barone P. and Frigessi A 1990 Improving Stochastic relaxation for gaussian random fields, *Probability in the Engineering & Informational Sciences*, v4, p369-389
- Besag J. 1974. Spatial interaction and the statistical analysis of lattice systems (with discussion), *J. Royal Statist. Soc B* 36 p 192-236
- Betzhold J. and Roth C. 2000. Characterising the mineralogical variability of a Chilean copper deposit using plurigaussian simulations, to appear in *J South African IMM*, March/April issue.
- Beucher H., Galli A., Le Loc'h G. and Ravenne C. 1993. Including a regional trend in reservoir modelling using the truncated Gaussian method. In Soares ed., *Geostat Troia '92*. Vol. 1. Dordrecht: Kluwer. 555-566.
- Beucher H. and Galli A. 1994. La géostatistique dans le domaine pétrolier. In *Proc of the 4th Tunisian Petroleum Exploration Conference*, Tunis pp379-393
- Beucher H., Geffroy F. and Doligez B. 1997. Introducing more geology in reservoir stochastic modelling using truncated Gaussian approach. In *Memorias del I Congreso Latinoamericano de Sedimentología*, Soc. Venezolana de Geol. Tomo 1, 89-95, Nov. 1997.
- Cacas M.C., Daniel J.M. and Letouzey, J., 2001, Nested geological modelling of naturally fractured reservoirs, *Petroleum Geosciences*, Vol. 7, pp. S43-S52
- Chatelin F. 1983 *Spectral Approximation of Linear Operators*. Academic Press
- Chautru J.M., Ravenne C., Lemouzy P. and Galli A. 1992. Heresim 3D: integrating data and disciplines for reservoir characterization. In *Annales du 4eme EC Symposium, Berlin*. 11p.
- Chilès J.P. and Delfiner P. 1999 *Geostatistics: Modeling spatial Uncertainty*, John Wiley and Son, New York, 695pp
- Cowles M.K. and Carlin B.P. 1996. Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review, *Journal of the American Statistical Association* 91, 883-904
- Deutch C. and Journel A.G. 1992. *GSLIB Geostatistical software library and user's guide*, Oxford University Press

- Doligez B., Ravenne C., Galli A. and Lemouzy, P. 1992. Une methodologie pour une etude integree des reservoirs: des donnees de puits aux simulations d'ecoulement en utilisant des outils géostatistiques. *Pét. Tech.* 372, 43-48.
- Doligez B., Beucher H., Fonnesu F., Tebaldi E. and Baruffini L.. 1994. Reservoir modeling: The reference model: Cajigar. In Helbig ed., *Modeling the Earth for Oil Exploration*. Pergamon.81-91.
- Doligez B., Beucher H., Fonnesu F., Tebaldi E. and Baruffini L.. 1994. Reservoir modeling: Limited data sets: Cajigar. In Helbig ed., *Modeling the Earth for Oil Exploration*. Pergamon. 108-110.
- Doligez B., Beucher H., Geffroy, F. and Eschard R. 1999 Integrated reservoir characterisation: improvement in heterogeneous stochastic reservoir modeling by integration of additional external constraints, in Schatzinger R. and Jordans J. eds, *Reservoir Characterization - Recent Advances*, AAPG Memoir N° 71, pp 333-342
- Doligez B., Granjeon D., Joseph P., Eschard R. and Beucher H., 1999 How can stratigraphic information help constrain geostatistical reservoir simulations? In *Numerical experiments in stratigraphic and sedimentologic computer simulations*, SPEM Special Publications N° 62, 1999 pp239-244
- Doligez B., Eschard R., Joseph P. and Beucher H., 1999. Une méthodologie intégrée pour l'estimation des volumes en place dans les réservoirs et la quantification de leurs incertitudes, *Pétrole et technique*, N° 422, Sept-Oct 1999 pp 50-54
- Dowd P.A. 1997. Structural controls in the geostatistical simulation of mineral deposits. In Baafi et al. eds., *Geostatistics Wollongong '96*. Vol 2. Dordrecht: Kluwer. 647-657.
- Cacas M.C., Daniel J.M. & Letouzey, J., in press at Petroleum Geosciences, Nested geological modelling of naturally fractured reservoirs
- Eschard R., Doligez B., Rahon D., Ravenne C. and Le Loc'h G. 1991. A new approach for reservoirs description and simulation using geostatistical methods. In *Advances in Reservoir Technology, Characterization, Modelling and Management, Edinburgh*. 13p.
- Eschard R., Desaubliaux G., Houel, P. and Beucher H. 1992. Geometry of an incised valley-fill complex (Southbay, Middle Jurassic, Yorkshire). In 7th IFP Research Conf. in *Exploration Production, Subsurface Characterization from Outcrop Observations, Scarborough, Yorkshire*.
- Eschard R., Lemouzy, P., Bacchiana, C., Désaubliaux, G., Parpant, J. and Smart B.. 1999. Combining sequence stratigraphy, geostatistical simulations and production data for modeling a fluvial reservoir in the Chaunoy field (Triassic, France), AAPG Bulletin, Vol 82, pp545-568
- Eschard R., Doligez B. and Beucher H. 2002 Using quantitative outcrop databases as a guide for geological reservoir modelling, *Geostatistics Rio 2000*, (eds) Armstrong et al, Kluwer Academic Publishers, Dordrechts, Holland, pp7-18.
- de Fouquet C., Beucher H., Galli A. and Ravenne, C.. 1989. Conditional simulation of random sets: application to an argillaceous sandstone reservoir. In Armstrong ed., *Geostatistics*. Vol.2. Dordrecht: Kluwer. 517-530.
- Freulon X. and de Fouquet C. 1993. Conditioning a Gaussian model with inequalities. In Soares ed., *Geostat Troia '92*. Vol. 1. Dordrecht: Kluwer. 201-212.
- Freulon X., de Fouquet C. and Rivoirard J. 1990. Simulation of the geometry and grades of a uranium deposit using a geological variable. In *Proc. XXII International Symposium APCOM*. Vol. 2. 649-659.
- Freulon X. 1992. *Conditionnement du modèle gaussien par des inégalités ou des randomisées*, Thèse de Docteur en Géostatistique, ENSMP. 168p.

- Freulon X. 1994. Conditional simulation of a Gaussian random vector with nonlinear and/or noisy observations. In Armstrong et al. eds., *Geostatistical Simulations*. Dordrecht: Kluwer. 57-71.
- Fournier, F. and Derain, J.F.. 1997. A statistical methodology for deriving reservoir properties from seismic data, *Geophysics*, Vol 60, N° 5, pp1437-1450
- Galli A., Guérillot D. and Ravenne C. 1990. Combining geology, geostatistics and multiphase fluid flow for 3D reservoir In Guérillot et al. eds., *2nd European Conference on the Mathematics of Oil Recovery*. Paris: Technip. 11-19.
- Galli A., Beucher H., Le Loc'h G. and Doligez B. 1994. The pros and cons of the truncated gaussian method. In Armstrong et al. eds., *Geostatistical Simulations*. Dordrecht: Kluwer. 217-233.
- Galli A., Ravenne C., Richard V. and Guérillot D. 1995. Constraining detailed geostatistical reservoir models with seismic and dynamic data. In *Seamless Seismic to Simulation and Back*, Proc. EAPG/SPE Workshop, Glasgow. 6 p.
- Galli A. and Beucher H. 1997. Stochastic models for reservoir characterization: a user-friendly review. In *Fifth Latin American and Caribbean Petroleum Engineering Conference and Exhibition, Rio de Janeiro, Brazil, 30 August-3 September 1997*. SPE 38999. 11 p.
- Geman S. and Geman D. 1984 Stochastic Relaxation, Gibbs distribution and the bayesian restoration of images, *I. E. E. E. Transactions: Pattern analysis and machine intelligence* 6:721-741
- Gikhman I.I. and Skorokhod A.V. 1996 Introduction to the theory of random processes. Dover Publication
- Gossa H., Gaaya M.H., Yahmadi B., Volpi B., Doligez B., Hu L.Y. and Galli A. 1993. A geostatistical approach aimed at reservoir characterization: a case study. In *2èmes Journées Tunisiennes de Géologie Appliquée, Sfax, Tunisie*. 10p.
- Grammer G.M., Eberli G.B., Van Buchem F.S.P., Stevenson G.M. and Homewood P., 1996 "Application of High-Resolution Sequence Stratigraphy to Evaluate Lateral Variability in Outcrop and Subsurface- Desert Creek and Ismay Intervals, Paradox Basin". In M.W.Longman and M.D.Sonnenfeld, eds., *Paleozoic Systems of the Rocky Mountain Region, Rocky Mountain Section, SEPM*, p.235-266.
- Grötsch J. and Mercadier C., 1999 Integrated 3-D Reservoir modeling based on 3-D seismic : The Tertiary Malampaya and Camago buidups, offshore palawan, Philippines, A.A.P.G. Bulletin, V.83, N°11, p. 1703-1728.
- Guérillot D., Rudkiewicz J.L., Ravenne C., Renard D. and Galli A. 1989. An integrated model for computer aided reservoir description: from outcrop study to fluid flow simulations. In *Proc. 5th European Symposium on Improved Oil Recovery, Budapest*. 651-660.
- Guérillot D., Beucher H., Galli A., Lemouzy P. and Morelon I. 1990. Multiphase fluid flow for 3D geostatistical reservoir description. In *Proc. OMMBKE 90, 21st Petroleum Itinerary Conference and Exhibition, Siofok, Hungary*. 252-261.
- Guérillot D., Lemouzy P., Galli A. and Ravenne C., 1990. 3D fluid flow behaviour in porous media characterized by geostatistical methods. In *SPE Latin American Petroleum Engineering Conference and Exhibition, Rio de Janeiro*.SPE 21081. 8p.
- Jacod J. and Joathon P. 1971 Use of random genetic models in the study of sedimentary processes, *Math. Geol.* vol. 3, no. 3, pp. 265-279.
- Johann, P., Fournier, F., Souza, O., Eschard, R. and Beucher, H. 1996 3D stochastic modeling constrained by well and seismic data on a turbidite field, SPE 36501, SPE Annual Technical Conference and Exhibition, denver, CO October 6-9

- Kerswill, J.A., Henderson, J.R., and Henderson, M.N. 1996 Distribution of gold and sulphides at Lupin, Northwest Territories – a discussion, *Econ Geol*, 91 (5), pp 957-963
- Lantuéjoul C. 1997a Iterative algorithms for conditional simulations In Baafi et al. eds., *Geostatistics Wollongong '96*. Vol 1. Dordrecht: Kluwer. 27-40
- Lantuéjoul C. 1997b Conditional simulation of object-based models, In *Advances in Theory and Applications of Random Sets*, D.Jeuin ed, World Scientific, 271-288
- Lantuéjoul C. 2002. *Geostatistical Simulation: Models and Algorithms*, Springer-Verlag Berlin, 250pp
- Le Loc'h G. 1989. An efficient strategy for combining the permeabilities: practical application on a simulated reservoir. In Armstrong ed., *Geostatistics* Vol. 2. Dordrecht: Kluwer. 557-568.
- Le Loc'h G., Beucher H., Galli A. and Doligez B. 1994. Improvement in the truncated gaussian method: combining several Gaussian functions. In *Proc. ECMOR IV, 4th European Conference on the Mathematics of Oil Recovery, Roros, Norway*. 13p.
- Le Loc'h G and Galli A., 1997 Truncated plurigaussian method: theoretical points of view, In E.Y. Baafi et al eds *Geostatistics Wollongong '96*, Vol 1. Kluwer, Dordrecht, pp211-222
- Lemouzy P.M., Eschard R. and Beucher H. 1991. An integrated approach for evaluation of production scenarios. *66th Annual Technical Conference and Exhibition of the SPE*. 209-219. SPE 22906.
- Mathieu Y., Verdier F., Houel P., Delmas J. and Beucher H. 1993. Reservoir heterogeneity in fluvatile Keuper facies: a subsurface and outcrop study. In Eschard R. and Doligez B. (eds) *Subsurface reservoir characterization from outcrop observations*, Paris: Editions Technip, pp 145-160
- Matheron G. 1968 *Schéma booléen séquentiel de partition aléatoire* Internal report N-83, CMM, Fontainebleau.
- Matheron G. 1975 *Random Sets and Integral Geometry*, Wiley, New York.
- Matheron G., Beucher H., de Fouquet C., Galli A., Guérillot D. and Ravenne C. 1987. Conditional simulation of the geometry of fluvio-deltaic reservoirs. *SPE 1987 Annual Technical Conference and Exhibition*, Dallas, Texas. 591-599. SPE 16753.
- Matheron G., Beucher H., de Fouquet C., Galli A. and Ravenne C. 1988. Simulation conditionnelle à trois facies dans une falaise de la formation du Brent. *Sci. Terre, Inf.* 28, 213-249.
- Matheron G. 1987. Suffit-il, pour une covariance, d'être de type positif? *Sci. Terre, Inf. Géologique* 26, 51-66.
- Matheron G. 1989. The internal consistency of models in geostatistics. In Armstrong ed., *Geostatistics*. Vol. 1. Dordrecht: Kluwer. 21-38.
- Matheron G. 1993. Une conjecture sur la covariance d'un ensemble aléatoire. *Cahiers de Géostatistique* 3, 107-113.
- Mathieu Y., Verdier F. and Beucher H.. 1993. Reservoir heterogeneity in fluvatile Keuper facies: a subsurface and outcrop study. In Eschard et al. eds., *Subsurface reservoir characterization from outcrop observations*. Paris: Technip. 145-160.
- Metropolis N., Rosenbluth A.W., Teller A.H. et al. 1953 Equations of state calculations by fast computing machines, *J. Chem. Phys.*, 21:1087-1091
- Meyn S.P. and Tweedie R.L., (1993) *Markov Chains and Stochastic Stability*, Springer Verlag, Berlin 550pp
- Molchanov I. 1997 *Statistics of the Boolean Model for Practitioners and Mathematicians*, John Wiley & Sons, Chichester, 160pp

- Morelon I.F., Doligez B., Guérillot D., Rahon D. and Touffait Y. 1991. An application of a 3D geostatistical imaging to reservoir fluid flow simulations. In Proc. Sixth SPE Petroleum Computer Conference. 223-231. SPE 22312.
- Moulière D., Beucher H., Hu L-Y. et al 1997 integration of seismic derived information in reservoir stochastic modelling using the truncated gaussian approach, in Baafi E. et al eds *Geostatistics Wollongong '96*, Vol 1., Kluwer pp374-385
- Moulière D. 1998 Intégration d'information sismiques pour la simulation de réservoirs, Thèse de Docteur de l'Ecole des Mines de Paris, pp182
- Numelin E. 1984 General irreducible Markov chains and no negative operators. Cambridge University Press
- Ravenne C., Eschard R., Galli A., Mathieu Y., Montadert L. and Rudkiewicz J.L. 1987. Heterogeneities and geometry of sedimentary bodies in a fluvio-deltaic reservoir. SPE 1987 Annual Technical Conference and Exhibition, Dallas, Texas. Vol. X. 115-122. SPE 16752.
- Ravenne C. and Beucher H. 1988. Recent developments in description of sedimentary bodies in a fluvio-deltaic reservoir and their 3D conditional simulations. In Proc. 63rd Annual Technical Conf. and Exhibition of SPE. 463-476. SPE 18310.
- Ravenne C., Galli A., Beucher H., Eschard R. and Guérillot D. 1991 Outcrop studies and geostatistical modelling of a middle Jurassic Brent Analogue, In Imarisio G et al (eds) *The European Oil and Gas conference, A multidisciplinary approach in exploration and production R&D*, Chapman & Trotman Ltd, London pp 497-520
- Ravenne C., Galli A. and Preat A. 1994. Reservoir modeling: The reference model: Poulseur/CBA. In Helbig ed., *Modeling the Earth for Oil Exploration*. Pergamon. 91-96.
- Ravenne C., Galli A., Doligez, B., Beucher H. and Eschard R., 2002 Quantification of facies relationships via proportion curves, *Geostatistics Rio 2000*, (eds) Armstrong et al, Kluwer Academic Publishers, Dordrechts, Holland, pp19-40.
- Remacre A. and Simon V., 1998 Conditional plurigaussian simulation: a case-study, In *Cahiers de Géostatistique*, Ecole des Mines de Paris, 6 17-28
- Rockafellar R.T., 1970 *Convex Analysis*. Princeton Mathematical Series
- Robert C.P. 1996 *Méthodes de Monte Carlo par Chaines de Markov*, Economica, Paris, 340pp
- Roberts G.O., Polson N.G. 1994 On the geometric convergence of the Gibbs sampler, J. R. Statist. Soc. B 56, N° 2, pp 377-384
- Roberts, G.O. and Sahu S.K. 1997. Updating scheme, correlation structure, blocking and parametrisation for Gibbs sampler, J R. Statist Soc B, v 59, N° 2, p291-317
- Roth C., Armstrong M., Galli A. and Le Loc'h G. 1998. Using plurigaussian simulations to reproduce lithofacies with contrasting anisotropies. In APCOM '98, London, IMM, pp. 201-213.
- Rudkiewicz J.L., Guérillot D. and Galli A. 1990. An integrated software for stochastic modeling of reservoir lithology and properties with an example from the Yorkshire middle Jurassic, In *North Sea Oil, pp399-406 Gas Reservoirs II*, The Norwegian Institute of Technology, Graham and Trotman
- Seigneurin A., Muller D., Galli A. and Ravenne C. 1993. Optimization of the well-spacing with a geostatistical model Tunu field - Mahakam area. In Proc. Indonesian Petroleum Association, 22nd Annual Convention, 17 p.
- Skvortsova T., Armstrong M., Beucher H., Forkes J., Thwaites A. & Turner R. 2000. Applying plurigaussian simulations to a granite-hosted orebody. In *Geostats 2000 Cape Town*, (eds) Kleingeld W and Krige D.G., Proceedings of the 6th international Geostatistics Congress, held in Cape Town 10-14 April 2000, pp904-911.

- Skvortsova T., Armstrong M., Beucher H., Forkes J., Thwaites A. & Turner R. 2002 Simulating the geometry of a granite-hosted uranium orebody, *Geostatistics Rio 2000*, (eds) Armstrong et al, Kluwer Academic Publishers, Dordrechts, Holland, pp85-100.
- Tierney L. 1994 Markov chains for exploring posterior distributions," with discussion, *Annals of Statistics*, 22, 1701-1762.
- Tierney L. 1996 Introduction to general state-space Markov chain theory. In *Markov Chain Monte Carlo in Practice*. Eds W.R. Gilks, S. Richardson, D.J. Spiegelhalter, Chapman & Hall
- Touffait Y., Beucher H. and Guérillot D. 1990. A 3D integrated structure for computer aided reservoir characterization. In *Proc. Fifth SPE Petroleum Computer Conference*. 171-180. SPE 20348.
- Van Buchem F.S., Doligez B., Eschard R., Lerat O., Grammer M. and Ravenne C., 2000, Stratigraphic Architecture and Stochastic Simulation of a Mixed silico-clastic/carbonated platform (Pennsylvanian, paradox basin, USA, in P. Homewood and G.P. Eberli editors, Genetic stratigraphy, Bulletin de Recherche Exploration-Production memoir 28, p. 109-128.
- Volpi B., Galli A. and Ravenne C. 1997. Vertical proportion curves: a qualitative and quantitative tool for reservoir characterization. In *Memorias del I Congreso Latinoamericano de Sedimentologia*, Soc. Venezolana de Geol. Tomo 2, 351-358, Nov. 1997.
- Yosida K. 1978 *Functional Analysis* Fifth Edition, Springer Verlag

INDEX

A

aggrading	110
algal mounds	2, 113 - 116, 120
analogue	109
anastomosed channels	26
andesitic flows	31
anisotropy	6, 21, 27, 69 - 70, 72, 111, 113, 118, 124, 128 - 139
aperiodic	101

B

back barrier	26
banded iron formation	27
blocking factor	87, 95 - 98, 103
Bochner's theorem	17
boolean	108, 113
burn-in period	87 - 91

C

capillary pressure	3, 111
carbonates	2, 109, 114
chalco-bornite	31
chalco-pyrite	31
channel	26, 27, 29, 108 - 110
chronostratigraphic marker	19, 24, 32
coastal plain	113, 114
cokriging	30, 93, 94
compact	98, 101 - 106
conditional expectation	83 - 86, 99, 100
conditional simulation	78, 82, 93, 95
connectivity	14, 34, 121
constraints	
hard	4
interval	78 - 82
soft	4, 38
contacts	
forbidden	55
convergence	78, 87 - 91, 95 - 106
copper	30 - 31
correlation	7, 14, 50, 53, 54, 59, 61, 66, 133

- covariance 78, 83, 89
 - centred 11, 78
 - noncentred 11, 21, 25, 75
 - symmetrical 75
 - 2 facies 12, 16
 - 3 facies 13, 14
- crevasse splay 7, 26, 110
- D**
- decomposition
 - Cholesky 83, 96
 - LU 78, 83
- deltaic reservoir 26, 28, 29, 109
- depositional cycles 7, 24, 27, 34, 108 - 109
- diagenesis 116 - 118
- E**
- eigenvalues 102, 105
- Ellerbeck 27 - 28, 33 - 37
- ergodicity 63, 98
- erosion 27, 32, 110
- estuarine channel 26, 110
- external drift 38
- F**
- flanking beds 115 - 116
- flooding surface 27, 109
- fluid flow 3, 4, 34, 112
- fluvial reservoir 28, 108, 109
- foreshore 26, 34, 113, 114
- fracture network 119 - 120
- G**
- gas 112
- Gauss-Seidel 95 - 96
- Gibbs sampler 8, 78 - 106
- gold 27 - 28
- H**
- Hammersley Clifford 103 - 106
- Hastings Metropolis 83
- history matching 3
- I**
- installation
 - linux 123, 125, 133
 - solaris 123, 125, 133
 - windows 122 - 123, 133
- irreducible 100 - 101
- J**
- Jurassic age 26

K

karstification 115

kernel operator 78, 99 - 106

L

lagoon 26, 27, 29, 34

lithotype rule 46 - 56, 68 - 69, 92, 113 - 121, 124 - 139

Lupin orebody 27 - 28

M

Mantos Blancos mine 30 - 31

Markov chain 83, 87 - 91, 98 - 106

meandering channels 26

mining 3, 27 - 31, 83

minorisation condition 105 - 106

Monte Carlo 83, 87

mudstone 27, 28, 29, 34, 109, 110, 114

multi-point geostatistics 17, 23, 26

N

nested simulations 108, 110

nonstationarity 8, 20, 24, 30, 34 - 38, 48, 55, 59, 63 - 65, 71 - 74, 125

O

object-based 34, 108 - 110, 113, 119 - 120

offshore 113

onlap 34, 112, 115

outcrop study 109, 110, 114

P

Paradox Basin 114 - 118

partition 49 - 53, 92, 102, 126

Pennsylvanian age 2, 114, 115

periodic 101

permeability 3, 111, 112, 117

petrophysical data 3, 111

pixel-based 108, 113

pluri_demo_simu 9, 122, 133 - 139

pluri_demo_vario 9, 122 - 133

porosity 3, 8, 111, 112

porous medium 27

production tests 3

prograding 113, 115, 118 - 119

proportion curve 108

horizontal 30- 32, 109

vertical 8, 25, 28, 31 - 34, 38, 109, 113, 115, 118

proportion matrix 38

Q

quartz veins 27 - 28

R

random function RF

correlated	6
indicator	10 - 27
stationary	8
"shifted"	7, 61
translated	7

Ravenscar 26 - 29, 32

recovery process 30 - 31

reef 113, 117, 118, 120

reference level 32 - 37, 111 - 112

regressive-transgressive 109

reservoir characterisation 109

rock type rule 3, 5 - 7, 46 - 56, 68- 69, 92, 113 - 121, 124 - 139

S

Saltwick 26 - 29, 33

sandstone 2, 7, 24, 25, 27, 29, 30, 56, 110

San Juan 2, 114

Scarborough formation 27 - 28, 32 - 37

sea level changes 108, 115 - 118

seismic data 3, 30, 34, 111

sequence stratigraphy 5, 108

shale 2, 7, 24, 25, 29, 56, 110

shoreface 26, 27, 109, 113

siliclastic reservoir 110, 114

simple kriging SK 83 - 86, 92 - 95

sinuosity 110

spectral radius 97 - 98

stationarity 20, 24, 44, 55, 58 - 60, 65 - 71, 76, 87, 88, 100, 104, 124

horizontal 24

vertical 64

subsidence 32 - 33

sulphide 3, 30

T

thickness 30, 34, 110

thresholds 5, 6, 44 - 56, 59, 62 - 66, 126 - 139

toplap 34, 112

total variation distance 98, 101

transgressive 27, 34, 109

transition probability 18 - 22 , 75 - 76, 99

1st type 18 - 20

2nd type 21 - 22

triangle inequality 17

U

unconformity 32

updating strategy 86 - 87

upscaling 4, 112

V

- variograms
 - cardinal sine 12 - 17, 58 - 76, 87, 111
 - 124 - 126, 138
 - experimental 58- 60, 71
 - exponential 24, 45, 67, 68, 69, 84, 109, 121, 124 - 129, 131 - 133
 - fitting 64, 67 - 74
 - gaussian 17, 21 - 26, 45, 67, 68, 87, 109, 118, 121, 124 - 139
 - horizontal 10
 - indicator 7, 8, 58 - 76, 122, 133
 - levels 64 - 65
 - nugget effect 44, 133 - 139
 - power function 12, 17
 - range 68 - 69, 89, 109, 111, 113 , 118, 120, 136
 - spherical 124 - 126, 131
 - vertical 109
 - 2 facies 12, 16
 - 3 facies 14
- vertical proportion 8, 25, 28, 31 - 34, 38, 108 , 109, 113, 115, 118
- curves

W

- Walther's law 108
- washover fan 26, 28, 29
- waste 27 - 28
- well-logs 111

Y

- Yorkshire cliffs 26 - 28, 32 - 37