
Probability aggregation methods and multiple-point statistics for 3D modeling of aquifer heterogeneity from 2D training images

thesis presented at the
FACULTY OF SCIENCES
INSTITUTE OF HYDROGEOLOGY AND GEOTHERMICS
(CHYN)

UNIVERSITY OF NEUCHÂTEL

for the degree of
DOCTOR OF NATURAL SCIENCES

presented by
ALESSANDRO COMUNIAN

accepted on the recommendation of

Renard PHILIPPE, *supervisor*
Julien STRAUBHAAR, *co-supervisor*
Ghislain DE MARSILY, *examiner*
Mauro GIUDICI, *examiner*
Denis ALLARD, *examiner*
Guillaume CAUMON, *examiner*

DEFENDED ON JANUARY 14, 2011

UNIVERSITY OF NEUCHÂTEL
2011

FACULTE DES SCIENCES
Secrétariat-Décanat de la faculté
Rue Emile-Argand 11
CH-2000 Neuchâtel

IMPRIMATUR POUR LA THESE

**Probability aggregation methods and multiple-point
statistics for 3D modeling of aquifer heterogeneity
from 2D training images**

Alessandro Comunian

UNIVERSITE DE NEUCHATEL

FACULTE DES SCIENCES

La Faculté des sciences de l'Université de Neuchâtel,
sur le rapport des membres du jury

MM. P. Renard (directeur de thèse), J. Straubhaar, G. de Marsily (Paris),
M. Giudici (Milan), D. Allard (Avignon) et G. Caumon (Nancy)

autorise l'impression de la présente thèse.

Neuchâtel, le 8 février 2011

Le doyen :
P. Kropf



*A Circe, che con le sue graffiate e i suoi morsi ha
saputo tenere compagnia a Linda in questi anni.*



Abstract

Multiple-point statistics (MPS) is a rising method for the characterization of heterogeneity. Its strength and its Achilles' heel lie in the training image, which is the conceptual model of geological heterogeneity on which MPS simulations are based. Indeed, on one side the use of the training image allows great flexibility when for example for bi-dimensional (2D) simulation a training image can be provided by a photo-mosaic of an outcrop or by a sketch drawn by a geologist. On the other side, in three-dimensions (3D) a training image is rarely available.

When the information provided by a 3D image is not accessible, then one must somehow use probabilistic information which comes from lower dimension sources, like for example 2D training images. If different 2D sources of information are available, one possibility is to aggregate the corresponding probability information. This problem is very general and several methods exist. In this thesis, the linear opinion pool, the Beta-transformed linear pool (BLP), the logarithmic pool, the disjunction and conjunction of probabilities, the Bordley's formula, the tau model, the nu model and the Markovian-type categorical prediction (MCP) are tested on different case studies to evaluate their main properties and their reliability. Two main categories of methods are distinguished: those based on the sum (convex) and those based on the multiplication (non-convex). When the weighting factors can be determined from some training data, the best reliabilities are obtained with the BLP and the Bordley's formula. Instead, when training data are not accessible, reasonably reliable results can be obtained with the Bordley's formula and with the MCP.

One convex method (linear pooling formula) and one non-convex method (Bordley's formula) are then tested for the aggregation of information coming from 2D training images. For the tests, one 3D image of a micro-computed tomography of a sandstone and one 3D realization of a fluvio-glacial environment are used as references. Two dimensional slices of the reference 3D images are used as training images for providing the information to be aggregated with the methods cited above, but also for the simulation with two novel method proposed here. One of this methods is based on merging the lists obtained for each 2D training images using the impala algorithm. The other is based on sequential 2D simulations conditioned by the data computed during the previous simulation steps (method `s2Dcd`). With this last method it is possible to obtain , without the use of a 3D training image, 3D simulations which can be considered close to the reference images according to most comparison criteria considered (visual comparisons, facies proportions, connectivity, equivalent conductivity). Moreover, while the results obtained with the method `s2Dcd` are close to the results obtained with a MPS simulation which make use of a 3D training image, the CPU time required by `s2Dcd` is from two to four orders of magnitude smaller than with

a traditional 3D simulation. This computational efficiency is a step forward for the introduction of MPS in frameworks which require a great number of realizations in a reasonably restricted amount of time, like for example Monte Carlo methods or stochastic inverse problems.

Other techniques exist to deal with the simulation in 3D when a 3D training image is not available. One of these techniques, developed in this thesis' framework, is applied for the simulation of the image of the fluvio-glacial aquifer analog used as reference for the tests depicted above. It is a hierarchical technique: six parallel outcrops mapped during a quarry excavation serve to recognize geological features at different scales and on different depositional layers. Once the complexity of the observed heterogeneity is simplified, object based techniques are used to simulate 3D training images containing simple shapes. Maps of the orientation of the main geological structures are created by interpolating orientations derived from morphological analysis. All this information is then included in a MPS simulation framework. The geological heterogeneity reproduced using this technique is more realistic than the one which can be for example obtained by approaches based on transition probability and Markov chains, and can provide an high-resolution benchmark for fluid flow and transport problems.

In summary, this thesis demonstrates that it is possible to apply MPS methods obtaining credible results from a geological point of view even in absence of a 3D training image.

Contents

1	Introduction	1
1.1	Motivation	3
1.2	Thesis's structure	4
	Bibliography	6
2	Probability aggregation methods in Earth sciences	9
2.1	Introduction	11
2.2	Background	12
2.3	Some important properties	13
2.3.1	Convexity	14
2.3.2	Unanimity and dictatorship	15
2.3.3	Response to extreme values	15
2.3.4	Independence preservation	15
2.3.5	Marginalization	16
2.3.6	External Bayesianity	17
2.3.7	Reliability and sharpness	17
2.4	Some mathematical aggregation methods	18
2.4.1	Methods based on the sum	18
2.4.2	Multiplicative methods	20
2.5	Weights determination	24
2.6	Case studies	25
2.6.1	Two experts and a dice	25
2.6.2	Probability densities	27
2.6.3	Trivariate event	27
2.6.4	Simulation studies	29
2.7	Results and discussion	42
2.8	Conclusions	45
	Bibliography	47
3	3D MPS simulation using 2D training images	51
3.1	Introduction	53
3.2	Background and terminology	55
3.3	Methodology	56

3.3.1	Probability aggregation	56
3.3.2	Sequential 2D simulations with conditioning data . . .	61
3.3.3	Lists merging	62
3.4	Comparison criteria	64
3.5	Micro-computed tomography image	67
3.5.1	Preliminary bi-dimensional calibration	68
3.5.2	Choose the weights for the Bordley's formula	68
3.5.3	Methods comparison	68
3.6	Reproduction of sedimentary structures	70
3.7	The Herten aquifer analog	71
3.8	Results and discussion	72
3.8.1	Micro CT image and paBor weights	72
3.8.2	Methods comparison on the micro CT image	76
3.8.3	Brussels sand sedimentary structures	77
3.8.4	The Herten aquifer analog	77
3.9	Check compatibility among training images	83
3.10	Conclusion	84
	Bibliography	87
4	Geostat. modeling of a 3D high resolution fluvio-glacial aquifer analog	91
4.1	Introduction	93
4.2	The Herten dataset	96
4.3	Motivations for a hierarchical approach	96
4.4	The hierarchical simulation framework	99
4.4.1	Subdivide the domain into large scale regions	100
4.4.2	Small scale: MP statistics	103
4.5	Results and Discussion	112
4.6	Conclusions	117
4.7	Acknowledgments	118
	Bibliography	118
5	Conclusions and outlook	125
	Bibliography	130
A	<i>whypda</i>	131
A.1	Introduction	133
A.2	Main <i>whypda</i> features	135
A.3	Content organization, methods and technologies	135
A.4	Data sources	139
A.5	Data quality	140
A.6	Consultation example	141
A.7	Contribution example	141
A.8	Discussion	143

A.9 Conclusion	144
Bibliography	146
B Some statistics about the visits to <i>wwhypda</i>	149
C Three-dimensional high resolution fluvio-glacial aquifer analog: Part 1: field study	153
D Comparison of three geostatistical methods for hydrofacies simulation: a test on mixed bedload-suspended load alluvial sediments	183
E Preliminary bi-dimensional calibration, micro CT image	215
F Simulation of 3D channels with the method <i>s2Dcd</i>	223
Bibliography	225

keywords

multiple-point statistics,
training image,
geological heterogeneity,
aquifer analog,
probability aggregation methods,
hydrogeological parameters,
aquifer characterization.

Chapter 1

Introduction

1.1 Motivation

Geological heterogeneity plays a key role in quantitative hydrogeology and reservoir production (7), in particular for solute and transport problems (6; 17).

Stochastic methods allow to deal with heterogeneity (6), and among them the multiple-point statistics method (MPS, 8; 16; 9) allows to overcome a big part of the issues which, according to Renard (14), limit the application of stochastic methods and the interest of practitioners. Indeed, MPS allows to produce simulation which are realistic from a geological point of view and to integrate diverse secondary information (4; 15). Moreover, recent advances in its implementation algorithms (15) have reduced the computational effort required for the simulation.

However, the fundamental requirement of MPS, the training image, is not accessible in some situations. This is for example frequent for three-dimensional problems. The main motivation of this thesis is to look for alternative tools to allow applying MPS techniques even when a 3D training image is not available.

One possibility to face this issue is to combine the information coming from one or several bi-dimensional (2D) training images and/or from covariance functions (12; 13; 3). These authors employ two different types of methods to aggregate the information expressed in probabilistic terms. But many other methods exist and the selection of the best method for a given application is a challenge. Analyzing the properties and the characteristics of various aggregation methods can guide the choice of the method, and chapter 2 faces this topic. This chapter also deals with the issue of the determination of the parameters when the aggregation methods are parametric and depends on some weighting factors, like for example the Bordley's formula (or τ model, 2; 10).

Note that the two problems reported over here have a broad range of potential application. For example, aggregation methods are used to account for secondary information in MPS simulations (4). Again, they are of interest in all the situations where various sources of information have to be aggregated into a unique term, thus in domains like for example management sciences, meteorology, and risk assessment.

The aggregation of probability is not always the best solution to overcome the lack of 3D training images in MPS. For example, in the case studies presented in chapter 3, the results obtained with a method based on the aggregation of probability are worse than the ones obtained using a different technique based on sequential 2D simulation constrained by conditioning data.

When a number of 2D training images is used for collecting the information which then are used to simulate and depict the same geological heterogeneity in 3D, it is crucial to check if they are compatible among them.

Chapter 3 contains some suggestions about how to check this compatibility.

Another possibility to simulate in 3D with MPS when a 3D training image is missing is to generate this last using process based, object based or computer assisted design techniques (11). However, it can be difficult to apply these techniques for the simulation of complex geological environments. A challenging example of a highly heterogeneous medium, in a fluvio-glacial environment, is the Herten aquifer analog (1). In this example, one single technique is not sufficient to deal with the simulation of such a complex heterogeneity. There is a need for new simulation frameworks which tackle the problem by mixing different techniques, and in chapter 4 a novel hierarchical method is proposed. This chapter deals with another important issue which affects all the geostatistical methods, MPS included – the non stationarity.

As claimed by the scientific community (5; 6), the use of stochastic techniques would be facilitated by a catalog of statistical properties of the rock types. The description of a project which attempts to provide such a tool can be found in appendix A.

1.2 Thesis's structure

Although the main chapters of this thesis, as pointed out in the previous section, are all tied up by the same *fil rouge*, they are self-contained and can be read independently.

Chapter 2 contains an overview of probability aggregation methods of interest for the Earth sciences. The main properties of the considered methods are described and tested on diverse case studies. For the methods which depends on parameters and weighting factors, some solutions to identify the parameters are proposed and tested.

A paper based on this chapter by A.Comunian, P.Renard and D.Allard “*Probability aggregation method in Earth sciences*” is in preparation for Mathematical Geosciences.

Chapter 3 In this chapter, two methods studied in chapter 2 are used for the aggregation of the information coming from bi-dimensional (2D) training images in order to perform three-dimensional MPS. Moreover, two novel simulation methods are proposed to face the same problem. One is based on sequential 2D simulations along different directions and constrained to the data simulated during a previous 2D simulation step. The other one is based on merging of the compatible patterns from different training images. This last method, even if it does not provide satisfactory simulations, provide a tool to check the compatibility among the 2D training images which are used to depict a 3D geological heterogeneity.

A paper based on this chapter by A.Comunian, P.Renard and J. Straubhaar “*3D multiple points statistics simulation using 2D training images*” is in preparation for Mathematical Geosciences.

Chapter 4 illustrates a novel hierarchical methodology applied to the simulation at high resolution of a fluvio-glacial aquifer analog, the Herten aquifer analog. From the information coming from six 2D parallel outcrops, mapped during a quarry excavation, a hierarchical structure is deduced and the simulation domain is divided into smaller domains. On these smaller domains, the features of the observed geological shapes is mimicked with MPS using simple 3D training images simulated with an object based technique. In the same domains, the non stationarity in the orientation of the geological bodies is tackled by constraining the MPS simulation by a map of orientation obtained from a morphological analysis of the outcrops and an interpolation of those orientation data.

This chapter was submitted to Journal of Hydrology by: A.Comunian, P.Renard, J.Straubhaar and P.Bayer “*Three-dimensional high resolution fluvio-glacial aquifer analog: Part 2: geostatistical modeling.*”.

Chapter 5 contains concluding remarks and perspectives.

Appendix A depicts *wwhypda* (the World Wide Hydrogeological Parameters Database), a project which attempts to collect in a collaborative database, accessible on line, all the measurements of the Hydrogeological properties of rock types published in papers, technical reports and other sources which report measurements made around the world.

This appendix was published on Hydrogeology Journal by: A. Comunian and P. Renard “*Introducing *wwhypda*: a world-wide collaborative Hydrogeological parameters database.*”.

Appendix B contains a brief statistical overview of the visits to the web interface of the database *wwhypda*, depicted in appendix A.

Appendix C contains the detailed description of the dataset used to perform the geostatistical simulation on the Herten aquifer analog proposed in chapter 4.

This appendix was submitted to Journal of Hydrology by: P. Bayer, P. Huggenberger, P.Renard and A.Comunian “*Three-dimensional high resolution fluvio-glacial aquifer analog: Part 1: field study.*”.

Appendix D describes a real case study where one of the methods described in chapter 2 is tested and compared with other methods on the simulation of mixed bedload-suspended load alluvial sediments.

This appendix was submitted to Hydrogeology Journal by: D. Dell'Arciprete, R. Bersezio, F. Felletti, M. Giudici, A. Comunian, P. Renard

“Comparison of three geostatistical methods for hydrofacies simulation: a test on mixed bedload - suspended load alluvial sediments.”.

Appendix E contains a description of the methodology used to define some parameters (search template size and multi-grid level) for the simulation of the micro-computed tomography image studied in chapter 3.

Appendix F contains some tests on simulations of 3D channels performed with the methods illustrated in chapter 3.

Bibliography

- [1] Bayer, P., Huggenberger, P., Renard, P., and Comunian, A. (2010). Three-dimensional high resolution fluvio-glacial aquifer analog: Part 1: field study. Submitted to Journal of Hydrology.
- [2] Bordley, R. F. (1982). A multiplicative formula for aggregating probability assessments. *Management Science*.
- [3] Caers, J. (2006). A general algorithm for building 3D spatial laws from lower dimensional structural information. Technical report, Stanford Center for Reservoir Forecasting.
- [4] Chugunova, T. L. and Hu, L. Y. (2008). Multiple-point simulations constrained by continuous auxiliary data. *Mathematical Geosciences*, 40(2):133–146.
- [5] Dagan, G. (2002). An overview of stochastic modeling of groundwater flow and transport: From theory to applications. *Eos Trans. AGU*, 83(53):621.
- [6] de Marsily, G., Delay, F., Goncalves, J., Renard, P., Teles, V., and Violette, S. (2005). Dealing with spatial heterogeneity. *Hydrogeology Journal*, 13(1):161–183.
- [7] Eaton, T. T. (2006). On the importance of geological heterogeneity for flow simulation. *Sedimentary Geology*, 184(3-4):187–201.
- [8] Guardiano, F. B. and Srivastava, R. M. (1993). Multivariate geostatistics: Beyond bivariate moments. In Soares, A., editor, *Geostatistics: Troia '92*, volume 1, pages 133–144. Kluwer, Dordrecht, The Netherlands.
- [9] Hu, L. Y. and Chugunova, T. L. (2008). Multiple-point geostatistics for modeling subsurface heterogeneity: A comprehensive review. *Water Resources Research*, 44(W11413):14.

-
- [10] Journel, A. G. (2002). Combining knowledge from diverse sources: An alternative to traditional data independence hypotheses. *Mathematical Geology*.
 - [11] Koltermann, C. E. and Gorelick, S. M. (1996). Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches. *Water Resources Research*, 32(9):2617–2658.
 - [12] Okabe, H. and Blunt, M. J. (2004). Prediction of permeability for porous media reconstructed using multiple-point statistics. *Physical Review*.
 - [13] Okabe, H. and Blunt, M. J. (2007). Pore space reconstruction of vuggy carbonates using microtomography and multiple-point statistics. *Water Resources Research*.
 - [14] Renard, P. (2007). Stochastic hydrogeology: What professionals really need? *Ground Water*, 45(5):531–541.
 - [15] Straubhaar, J., Renard, P., Mariethoz, G., Froidevaux, R., and Besson, O. (2010). An improved parallel multiple-point algorithm using a list approach. *Mathematical Geosciences*, accepted.
 - [16] Strebelle, S. (2002). Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, 34:1–21.
 - [17] Sudicky, E. A., Illman, W. A., Goltz, I. K., Adams, J. J., and McLaren, R. G. (2010). Heterogeneity in hydraulic conductivity and its role on the macroscale transport of a solute plume: From measurements to a practical application of stochastic flow and transport theory. *Water Resour. Res.*, 46(1):W01508.

Chapter 2

Probability aggregation methods in Earth sciences

Abstract

The need of combining in a probabilistic framework different sources of information is frequent in Earth sciences. This can occur for example when modeling a reservoir using direct geological observations, geophysics, remote sensing, etc. From each source of information it is often possible to extract the probability of occurrence of a certain event (*i.e.* the presence of a certain lithofacies at a certain location). Then, one needs to select a method for aggregating the different probability distributions into a single one before being able to simulate the reservoir.

Ten methods have been identified as candidates to solve this problem: linear opinion pool, Beta-transformed linear pool, logarithmic pool, disjunction and conjunction of probabilities, Bordley's formula, tau model, nu model and Markovian-type categorical prediction. They are classified according to their main properties: the convexity, the response to the impossible event or to the certain event, the external Bayesianity. The methods are compared and their properties are tested on five case studies. When optimized weighting factors can be computed, the best results are obtained with the Beta-transformed linear pool and the Bordley's formula. When the optimized weights cannot be computed, and when the hypothesis of conditional independence among sources of information is justified, good results can be obtained with the Markovian-type categorical prediction and the Bordley's formula with unitary weights.

Besides, a method can be selected according to its properties. For example, in the case of a sequential simulation framework, the external Bayesianity may be required and then the logarithmic pool or the Bordley's formula can be used.

2.1 Introduction

Aggregating all the information available about a problem is a crucial step towards solving it. But how to do that? How to aggregate the opinions of n experts about the risk of a nuclear power plant accident, about the risk of an economical crisis, or about the diagnosis of health disease? How to combine the geological interpretation provided by a seismic survey and an electrical tomography survey for the same area? How to merge the results of different weather forecasting models?

In the Earth Sciences, this problem is frequent and many authors approached it in different ways. Tarantola and Valette (33) developed the concept of conjunction and disjunction of probabilities for solving inverse problems. Benediktsson and Swain (3) adopted consensus theoretic classification methods to aggregate geographical data like satellite images coming from different sources. Journel (17) proposed the *tau model* in a very broad perspective.

This model was subsequently used for example by Strebelle et al. (31) to map lithofacies using seismic information and multiple-point statistics, or by Caers (7) to combine the probability assessments derived from different 2D geostatistical models to simulate 3D geological structures. Okabe and Blunt (24; 25) used the linear probability combination method to simulate 3D porous media from 2D multiple-point statistics extracted from microscope images of a rock sample. Ranjan and Gneiting (27) combined weather forecasts coming from different models with the beta-transformed linear opinion pool (BLP). Mariethoz et al. (21) used the probability conjunction method to develop a co-located co-simulation algorithm allowing to model any complex probability relationship between the primary and secondary variable. In the context of risk analysis, Clemen and Winkler (11; 12) and Genest and Zidek (16) provide detailed reviews about probability aggregation methods and their properties.

The diversity of approaches that one can find in the literature may be surprising, but this is because aggregating probabilities is usually an ill posed problem: there is often in practice a lack of information and models to describe accurately the interactions between the sources of information. In that framework, we are left with making assumptions and select a method without being able to check the accuracy of the estimations. In other words, there is neither a single method nor a single set of parameters (as several methods are parametric) that can aggregate probabilities accurately under all possible circumstances. Instead, the selection of the most suitable aggregation method depends on the specific problem which is addressed; a clear understanding of the properties characterizing each aggregation method is therefore an important step.

In this paper, we provide an overview of most of the available techniques to aggregate probabilities and discuss their properties in the perspective of Earth sciences applications. In section 2.3, we define the main mathematical properties of the aggregation methods. We then review the aggregation methods of interest for the Earth Sciences (section 2.4). Most of the methods are parametric, and therefore section 2.5 contains a brief overview of the standard approaches used to determine the parameters. Through a series of examples, section 2.6 illustrates the different behaviors of the methods under simple and more complex situations. Section 2.7 discusses the results and provides guidelines for the selection of a suitable aggregation method.

2.2 Background

The problem of aggregating information coming from n different sources S_i , $i \in (1, \dots, n)$ involves a broad range of disciplines, therefore the information can be shared using different parameters and different scales. Here we consider only frameworks where the information are provided as probability

mass functions, probability density functions, probability odds or logistic-type ratios. Moreover, consideration about the elicitation processes and the quantification of the degree of belief of each S_i are not considered.

The probabilistic framework considered here is akin to the one proposed by Journel (17): to assess the probability of an event E knowing the probability of E conditioned to some data events D_i , $i = 1, \dots, n$.

For example, the event E can be a spatially localized lithofacies value, while the data D_i can represent the information provided about the spatial distribution of the lithofacies by a seismic survey, a core sample, an electrical tomography or a training image.

In real-world problems $P(E|D_1, \dots, D_n)$ is rarely accessible, hence the challenge is to find a good approximation of it, that is to determine a global aggregated probability term P_G (a *pooling operator*) such that:

$$P(E|D_1, \dots, D_n) \approx P_G(P_0(E), P(E|D_1), \dots, P(E|D_n)) \quad (2.1)$$

Here $P_0(E)$ can be considered as a prior knowledge about the event E independent from the information provided by some fictitious data D_0 , that is $P(E|D_0) = P_0(E)$. The term $P_0(E)$ defined in this way can be included with more generality in the aggregation formulas, like one of the source of information which are aggregated.

In the following we use in general the letter P and the term “probability” as a shorthand for probability density function (pdf), probability mass function or probability measure. When it is required to specify that we deal with a probability density function then the lowercase p is used.

Some probability aggregation methods are formulated in terms of *odd ratios* or in terms of *logistic-type ratios*; odds are the inverse of logistic-type ratios. To denote the odds we adopt the letter O ; the relation between odds and probabilities of an event E is the following:

$$O(E) = \frac{P(E)}{1 - P(E)}. \quad (2.2)$$

In the following the more concise notation P_i will be used to denote the terms $P(E|D_i)$; for example, the right-hand side of the equation (2.1) will be often rewritten as:

$$P_G(P_0, P_1, \dots, P_n)(E). \quad (2.3)$$

2.3 Some important properties

In the following sub-sections, we define and discuss the main properties which can be taken into account to characterize an aggregation method. Note that axiomatic approaches use some of the following properties as the starting point to derive a pooling formula which have to satisfy one or more of them. The prior term as defined in the previous section can be included in great

generality in all the aggregation formula at the same level as the other aggregated terms. Still, the prior holds a particular position in some formulas. The recapitulative table at the end of the manuscript (tab. 2.4) also contains a reference to this detail.

2.3.1 Convexity

If the combination method provides values of P_G such that:

$$P_G \in [\min(P_0, \dots, P_n), \max(P_0, \dots, P_n)] \quad (2.4)$$

then the method is *convex*. The notion of convexity can also be extended in the case of probability densities; one possible definition of convexity in the case where $n + 1$ probability densities $p_i(x)$ with $i = 0, \dots, n$ are combined into a global term $p_G(x)$ is:

$$\forall x \in \mathbb{R} \quad : \quad \min[p_0(x), \dots, p_n(x)] \leq p_G(x) \leq \max[p_0(x), \dots, p_n(x)] \quad (2.5)$$

To illustrate the concept of convexity, consider the throw of a dice and the probability related to the face five. One source of information is “*the face is odd*”, so it associates to the face five a probability of $1/3$. The other source of information tells “*the face is greater than or equal to 4*”; again we have for the face five a probability of $1/3$. Combining these two probabilities with our mental (*non-convex*) combination method, the global probability for the face five is of 100%. Using a convex combination method (like for example the linear combination method, which will be explained in the following sections) we would obtain a probability of $1/3$, a result which is realistic but does not corresponds to the full use of the available information. This example is discussed in more detail further.

In Earth Sciences, in some situations a convex combination method can be a limitation. Consider for example the case of a multiple-point statistics simulation of a binary variable which can be “sand” or “shale”. Suppose that scanning the training image we find a probability of 0.8 to have “sand” in a given point, and suppose that we get the same result from a secondary data map; then it is reasonable to assign to the probability of occurrence of the facies “sand” a probability greater than 0.8.

However enforcing (or weakening) some probability assessment due to the non-convexity of a combination method have sense only if the sources of information are *distinct* (*differently informed*). Using again the example of the rolling of a dice and the information “*the face is odd*” and “*the face is not even*” and the probability of the face to be five. In this case computing an aggregated probability greater than $1/3$ has no sense, because the probabilities provided by the two sources are extracted from the same information, therefore *aggregating* information no longer makes sense. Here it is easy to realize that the pooling procedure is no longer an information aggregation,

but in other cases it can be difficult. Consider for example the case where the two sources of information are the results of a seismic survey and of an electric tomography survey on the same geological domain Ω . The surveys are based on different physical processes, but a fraction of the information extracted from the same geological domain is always deduced from the same informing features. In aggregating the information coming from these two different sources, how to distinguish the pieces of information which allows to perform a real information aggregation from the redundant ones? To find and answer to this question is a challenging question.

2.3.2 Unanimity and dictatorship

A method preserves unanimity when the aggregation of identical probabilities P^* provides the same value $P_G(P^*, \dots, P^*)(E) = P^*(E)$. In particular this property entails the *zero preservation property* (or the *null set property*): $P_G(0, \dots, 0) = 0$. Unanimity is strictly related to convexity: each convex method have the unanimity property.

A method is *dictatorial* (16) when the probability P_i provided by the i -th source of information is always taken as the group assessment, that is $P_G(P_1, \dots, P_i, \dots, P_n)(E) = P_i(E)$. For example, the linear opinion pool (see following sections) is dictatorial when $w_i = 1$ for some i and the other $w_{j \neq i}$ are null.

2.3.3 Response to extreme values

While situations where the dictatorship holds are restricted to some pathological cases, it is important to check the behavior of an aggregation method in situations where some sources of information provide a probability equal to zero (impossible event) or to one (certain event).

A property of interest for soil sciences is the *0/1 forcing property* (1). A method which has this property produces an aggregated probability term $P_G(E) = 0$ (resp. $= 1$) when $P_i(E) = 0$ (resp. $= 1$) for some i , and can be convenient for example in the case where catenary patterns of depositional sequences have to be reproduced.

2.3.4 Independence preservation

A method preserves the independence if whenever we choose two events A and B for which $P_i(A \cap B) = P_i(A) P_i(B)$ is valid for every $i = 1, \dots, n$ then:

$$P_G(P_1, \dots, P_n)(A \cap B) = P_G(P_1, \dots, P_n)(A) P_G(P_1, \dots, P_n)(B) \quad (2.6)$$

holds.

Authors like Lehrer and Wagner (20); Genest (13); Wagner (34); Genest and Wagner (15) faced the challenge of finding a pooling formula which preserve independence, but nobody found a pooling formula of interest. Their

opinion can be summed up quoting Genest and Zidek (16): “*independence preservation is not a reasonable requirement to impose on consensus-finding procedures*”.

2.3.5 Marginalization

Consider a vectorial random variable Z associated to the event E . For each component k of Z we can define the *marginalization operator* M_k such as:

$$M_k(P(Z)) = P_k(Z) \quad (2.7)$$

where $P_k(Z)$ is the marginal probability related to the component k of Z . A pooling procedure has the *marginalization property* if, for each component k , the operator M_k commutes with the pooling operator:

$$P_G(M_k(P_1), \dots, M_k(P_n)) = M_k(P_G(P_1, \dots, P_n)). \quad (2.8)$$

Let us illustrate this property in the case of a vectorial random variable with two continuous components, $Z = \{X, Y\}$, n sources of information which provide n probability density functions $p_1, \dots, p_n$ and a given pooling method which provides the global term p_G . The marginal probability density functions for the component X and the n sources considering the joint $p_{i_{X,Y}}(x, y)$ are:

$$p_{i_X}(x) = \int_y p_{i_{X,Y}}(x, y) dy \quad \text{with } i = 1, \dots, n. \quad (2.9)$$

The global marginal probability density function computed using the pooling method with (2.9) is:

$$p_{G_X}(x) = p_{G_X}(p_{1_X}(x), \dots, p_{n_X}(x)) \quad (2.10)$$

and the global joint probability density function is:

$$p_{G_{X,Y}}(x, y) = p_{G_{X,Y}}(p_{1_{X,Y}}(x, y), \dots, p_{n_{X,Y}}(x, y)). \quad (2.11)$$

Then, according to (2.10), the pooling method has the marginalization property if:

$$p_{G_X}(x) = \int_y p_{G_{X,Y}}(x, y) dy \quad (2.12)$$

and if the same relation is valid for the other component Y .

The same considerations are valid for random variables with more components and can be easily reduced to the discrete case.

In Earth Sciences, joint distributions are often used. In many situations, the marginal distributions need to be considered and therefore the marginalization property should be required. However we will see later that only one aggregation method (the linear pooling) can ensure this property.

2.3.6 External Bayesianity

The *external Bayesianity* property (14) is related to the behavior of a pooling method when additional information become available. If we identify the process of revising and updating a probability term with a *likelihood operator* L , then a probability combination method is external Bayesian if L commutes with the method, viz:

$$P_G(LP_1, \dots, LP_n)(E) = LP_G(P_1, \dots, P_n)(E). \quad (2.13)$$

In words, this means that if additional information about the event E are learned, then to include them in each aggregated term (and then aggregate) or include them directly in the aggregated term P_G has the same effect.

External Bayesianity is a suitable property in the Earth Sciences. This property is useful for example in all the sequential simulation algorithms, where the information added at each simulation step is used to update the probability distributions used to perform the subsequent simulation step. Moreover, often additional information can be added and combined with the available one in different phases of the simulation. The flexibility provided by an aggregation method which has the external Bayesianity property is then desirable.

2.3.7 Reliability and sharpness

Reliability (or calibration) and sharpness are two aspects which can be used to evaluate the quality of probability forecasts. For example, these two aspects are frequently used to evaluate weather forecasts (23).

According to Murphy (23), a (probability) forecast has a good quality if there is a good correspondence between the forecasts and the matching observations.

Reliability (or calibration) is related to correspondence between a forecast value and the mean of the observations related to it, averaged over all forecasts. Consider for example the probability forecast of 0.1 for a snow event and the mean of the observations related to this forecast. The agreement between 0.1 and the mean of the observations gives an idea of the reliability for the forecast, a concept which can be extended considering an average on all the forecast values and illustrated using reliability plots, which will be depicted in the following sections.

Another criteria to evaluate the quality of a probability forecast is the sharpness. For example, a rain forecast has a good sharpness if it often close to the extreme values zero and one. On the other hand, a forecast close to the central value 0.5 has a low sharpness.

Reliability and sharpness can be used to check the quality of probability aggregation formulas; this is explained in detail in the section devoted to the simulation case studies, where reliability plots, histograms and score rules are used to evaluate visually and numerically these two properties.

2.4 Some mathematical aggregation methods

Clemen and Winkler (11) propose a classification of the probability aggregation methods into mathematical combination methods and behavioral approaches. Mathematical aggregation methods are functions or models which aggregate probabilities estimations P_i coming from different sources and combine them into a global term P_G . Behavioral approaches are instead based on the interaction among experts; the aggregation process concludes with an agreement about a common probability term. Here we focus on mathematical aggregation methods.

Note that in the context of behavioral approaches the word “interaction” has a meaning strictly related to the fact that the experts are human beings who can exchange advice and discuss their assessments; it is different from the *non-interaction* referred by Bordley (4) or from the *independence* quoted in the section 2.3.4 and in the rest of this work.

Mathematical methods can in turn be divided into methods derived from axiomatic approaches and methods derived from Bayesian considerations. Another dichotomy for the mathematical methods is possible. Some methods combine the information summing its terms (and eventually apply some transformation to the sum), with an approach which is akin to the *disjunction* of probability distributions (33; 32). Other methods are multiplicative, that is the probabilities are aggregated using a multiplication; this aggregation approach is close to the *conjunction* of probability distributions (33; 32). This last dichotomic criterion defines also two groups within which the aggregation methods have many common properties.

The next subsections, following and extending the work of Genest and Zidek (16), Clemen and Winkler (11), Clemen and Winkler (12), provide a summary of some of the most important aggregation methods in Earth sciences.

2.4.1 Methods based on the sum

A weighted sum is a natural way of combining the probability terms $P_1, \dots, P_n$ into a global probability term:

$$P_G(E) = \sum_{i=1}^n w_i P_i(E). \quad (2.14)$$

In the equation (2.14) the w_i are positive weighting factors; in order to have a meaningful global probability, their sum must be one. The formula (2.14), attributed to Laplace by Bacharach (2), is the *linear opinion pool*. In equation (2.14), one can also account for the term P_0 (then the index i in the sum starts from 0); in this case the formula represents a *generalized linear opinion pool*.

The linear opinion pool is simple to understand and to implement, satisfies

unanimity and it is the only aggregation method which satisfies the marginalization property (11). It is applied in many fields with good empirical results; it is probably the most common used consensus rule. However Ranjan and Gneiting (27) demonstrate that the linear opinion pool is intrinsically sub-optimal: the result of any aggregation of individually calibrated forecasts is uncalibrated and lacks sharpness.

Moreover the linear opinion pool is convex (as it can be easily deduced from the simple case studies considered in section 2.6 and from equation (2.14) with $\sum w_i = 1$) and this can be a limitation.

This combination method is dictatorial when a certain w_i is equal to one. It does not have the 0/1 forcing property.

Laddaga (19) pointed out that the linear opinion pools cannot satisfy the independence preservation property unless they are dictatorial, but as explained in section 2.3 many authors argue that this is not a suitable property.

If we provide an equal weight to every probability term P_i the method reduces to an arithmetic average; in this case the method coincides with the *disjunction* of probabilities (33; 32)

The linear opinion pool can be formulated with equation (2.14) also in the case of probability density functions. In this case, the method has some good properties: the aggregation's result is a probability distribution, and the aggregation process conserves for example the absolute continuity of the data sources.

The combined probability density results is, in most cases, multi-modal. From the interpreter's point of view, this can be a problem, because it is difficult to evaluate a preferred value (16). At the same time, a multi-modality can reflect the real nature of the considered phenomenon. Therefore in some cases this ambiguous situation can be preferred to a well defined peak far from any of the aggregated information. We will explore these issues in the following sections.

One of the main drawback of the linear opinion pool is that it does not satisfy the external Bayesianity property which is desirable, especially in situations where there are frequent additions of information before the achievement of a final global probability estimation.

Beta-transformed linear pool

A transformation can be applied to a linear pool in order to improve its performances. Ranjan and Gneiting (27) propose to aggregate probability forecasts of a binary event using a beta-transformed linear pool (BLP) of the form:

$$P_G = H_{\alpha,\beta} \left(\sum_{i=1}^n w_i P_i \right). \quad (2.15)$$

In equation (2.15), the weights respect the same constraints as a linear pool, and the function $H_{\alpha,\beta}$ is the cumulative density function of a beta distribution with shape parameters $\alpha > 0$ and $\beta > 0$, that is:

$$H_{\alpha,\beta}(x) = B(\alpha, \beta)^{-1} \int_0^x t^{\alpha-1} (1-t)^{\beta-1} dt \quad (2.16)$$

with $x \in [0, 1]$ and $B(\alpha, \beta) = \int_0^1 t^{\alpha-1} (1-t)^{\beta-1} dt$.

In their work, Ranjan and Gneiting (27) show, on simulations and on real case studies, that the BLP outperforms the linear pool. By adding the two parameters α and β , this aggregation method enlarges the possibilities offered by the linear pool. For example, using the BLP it is possible to obtain non convex aggregation of probabilities. Moreover the BLP includes the linear pool when α and β are one.

The marginalization property is lost, due to the non-linear transformation; also, the convexity of the method is a function of the parameters α and β . The responses of the BLP with respect to the extreme values zero and one are the same of the linear pool. The beta transformation can be also applied to probability densities. Note that in order to apply this method to multivariate events a normalization factor must be added.

2.4.2 Multiplicative methods

A weighted sum of the logarithms of the probability (or odd) terms can be considered in place of their simple sum, obtaining methods based on the multiplication. With this approach applied to probabilities one can obtain the *logarithmic pooling formula*.

$$P_G(E) \propto \prod_{i=0}^n P(E|D_i)^{w_i}. \quad (2.17)$$

The variability of the weights can be restricted in order to obtain that their sum S_w is one, that is $\sum_{i=1}^n w_i = 1$. In this case the formula preserves the unanimity and the external Bayesianity, overcoming with this last property one of the lacks of the linear pool. However, in general the weights can be selected such that the normalization factor in equation (2.17) is finite.

All the multiplicative methods respect the 0/1 forcing property, the unanimity and in general they are not convex. These properties can be explained in Boolean terms. In fact, while the methods based on the sum aggregate the terms using the *or* operation (disjunction), the multiplicative methods use instead the *and* operation.

The multiplicative methods do not have the marginalization property and they do not preserve the independence.

In terms of probability density function, for example for some $p_0, \dots, p_n$ on $\mathbb{R}$, the logarithmic formula is:

$$p_G(p_0, \dots, p_n) = \frac{\prod_{i=0}^n p_i(x)^{w_i}}{\int_{\mathbb{R}} \prod_{i=0}^n p_i(x)^{w_i} dx} \quad (2.18)$$

with $w_i \geq 0$ for each i and $\sum_{i=0}^n w_i = 1$ (or more generally, w_i are selected such that the denominator of equation (2.18) is finite).

In the case of probability densities the logarithmic opinion pools provide a distribution which is close to unimodal (see section 2.6.2), that is a distribution that is in general easier to interpret than a distribution obtained with a linear pool.

When it is not possible to determine the weights w_i , one approach consists in setting $w_i = 1$ for each i . This corresponds to the *conjunction of probabilities* (33; 32).

Another solution is to set all the weights w_i equal to $1/(n+1)$; in this case the logarithmic pool reduces to a geometric mean of the probabilities. Other possibilities are explored in the following.

Bordley's formula

A multiplicative formula can be used to aggregate probabilities expressed in terms of odds (O_i). Bordley (4) derived, for binary events, a formula of this kind using an axiomatic approach, obtaining:

$$O_G = \prod_{i=0}^n O_i^{w_i} \quad (2.19)$$

where the condition $w_0 = 1 - \sum_{i=1}^n w_i$ is imposed, while the other weights w_i can vary in $[0, \infty)$.

The Bordley's formula is close to a logarithmic opinion formula, but its different formulation entails some advantages: for example the formula (3.5) includes the need of the logarithmic pools to have the sum of the weights equal one (a requirement for the external Bayesianity) into a constraint on the value of w_0 . In this way more flexibility is left to the choice of the other weights w_i for $i \geq 1$.

The prior term P_0 has a special position inside the formula, and its influence on the aggregated result P_G can be tuned changing the value of the sum of the other weighting factors S_w . If $S_w = 1$ then the contribution of P_0 is neglected. An brief overview of the dependence of the aggregated results on S_w is provided by the following example (4). Suppose that $P_i = P^*$ for each $i = 1, \dots, n$, then the following cases can be distinguished: (1) if $S_w = 1$, then $P_G = P^*$; (2) if $S_w > 1$, then P_G is always further from P_0 than P^* ; (3) if $S_w < 1$, then P_G is always closer from P_0 than P^* . Other examples of this features can be found in the following sections.

The Bordley's formula can also be rewritten in this form:

$$\frac{O_G}{O_0} = \prod_{i=1}^n \left(\frac{O_i}{O_0} \right)^{w_i} \quad (2.20)$$

or in term of probabilities:

$$P_G = \frac{P_0 \prod_{i=1}^n (P_i/P_0)^{w_i}}{P_0 \prod_{i=1}^n (P_i/P_0)^{w_i} + (1 - P_0) \prod_{i=1}^n [(1 - P_i)/(1 - P_0)]^{w_i}}. \quad (2.21)$$

The expressions (2.21) (resp. (2.20)) enlighten the fact that the terms O_i (resp. P_i) for $i \geq 1$ are normalized by the prior term O_0 (P_0).

Another possible representation of the Bordley's formula is the following:

$$\ln \left(\frac{O_G}{O_0} \right) = \sum_{i=1}^n w_i \ln \left(\frac{O_i}{O_0} \right) \quad (2.22)$$

where $\ln(O_i/O_0)$ could be interpreted as the “net information” provided by the source of information i .

Even if the determination of the weighting factors represents a challenge, the tuning of the weights provides flexibility. It is probably for this reason that the Bordley's aggregation formula is widely used in the Earth Sciences - see for example the works of Benediktsson and Swain (3), Journal (17), Caers (7), Polyakova and Journal (26) and Krishnan (18). Furthermore the Bordley's formula can be expressed in terms of odd densities and has the 0/1 forcing property.

Tau model

Journal (17) derived the Bordley's aggregation formula using a different approach, that is applying the permanence of ratios paradigm and considering *logistic-type ratios* in place of odd ratios. Moreover, he focused on the aggregation of conditional probability statements about a data event E when the data events D_i are observed, that is $P(E|D_i)$ with $i = 1, \dots, n$. Journal named this aggregation formula the *tau model*.

Krishnan (18) proposed an analytical expression for the tau weights. Unfortunately, the expression depends on the available data and on the sequence whereby the data are considered. Even if the dependence on the data sequence can be attenuated computing the mean of the weights over all the possible sequences, it is difficult to compute the weights when few data are available.

In this model, setting $w_i = 1$ for all i corresponds to consider the hypothesis of conditional data independence. When justified, this hypothesis entails a great simplification in the model and shows some similarities with other models, *i.e.* with the Markovian type categorical predictor, depicted in the following.

The nu model

Another different way of expressing the Bordley's model was proposed by Polyakova and Journal (26) and named the *nu model*. In analogy to the analytical expressions for the weighting factors for the τ model proposed by Krishnan (18), Polyakova and Journal (26) propose analytical expressions for the weights of the ν model. The expressions share the same drawbacks of the expression proposed by Krishnan (18): in real case studies the determination of the weights requires a huge amount of data, which is rarely accessible. In the original formulation of the ν model, the weights can be regrouped obtaining a simplified expression for the model with only one weighting factor ν_0 :

$$\frac{O_G}{O_0} = \frac{1}{\nu_0} \prod_{i=1}^n \frac{O_i}{O_0} \quad (2.23)$$

To the advice of Polyakova and Journal (26) the nu representation “allows a clear and useful separation” of the terms related to the individual data information and the data interaction. The condensed formula (2.23) with $\nu_0 = 1$ is considered to provide a less constrictive modelisation than the hypothesis of data conditional independence. The similarities with the formula (2.23) and the Bordley's formula (or τ model) with weights equal one are strong. Polyakova and Journal (26) suggest that the condensed nu model is preferable in cases when it is possible to compute explicitly the weights with their analytical expression or estimate them from training data; instead, when the weights are considered constant or independent of the data values, then the tau model should be preferred. The simulation case studies depicted in the following sections show two situations where optimizing only one ν weight it is possible to obtain reliable results and one case where the Bordley's formula with two parameters provide a better reliability.

Markovian-type categorical prediction

Allard et al. (1) proposed, for the prediction of spatial categorical variables, a multiplicative formula which does not require to determine the weights w_i . The formula, called *Markovian-type categorical prediction* (MCP), is a very good approximation of the Bayesian maximum entropy principle (8), with the advantage of being computationally efficient. In our framework the MCP formula can be rewritten as:

$$P_G(E = k) = \frac{P_0(E = k)^{1-n} \prod_{i=1}^n P(D_i, E = k)}{\sum_{k=1}^K P_0(E = k)^{1-n} \prod_{i=1}^n P(D_i, E = k)}. \quad (2.24)$$

In the equation (2.24) the event E is a categorical variable which can assume the finite values $1, \dots, K$.

The MCP formula represents an interesting aggregation approach when the accessible data set does not allow to determine reliable values of the weighting

parameters. Strong similarities can be observed among the Bordley's formula (tau model) and the nu model under the hypothesis of conditional data independence ($w_i = 1, \nu_0 = 1$) and the MCP formula, that is among the equation (2.21) with $w_i = 1$, the equation (2.23) with $\nu_0 = 1$ and the equation (2.24). However, the trivariate case study illustrated in the following sections shows that the methods can provide diverse results (fig. 2.3,2.4,2.5,2.6).

2.5 Weights determination

The majority of the aggregation methods presented above are parametric. How to determine the weighting factors w_i (and shape parameters α and β) when analytical expressions like the one proposed by Krishnan (18) for the tau model and by Polyakova and Journel (26) for the nu model cannot be used?

Winkler (35) suggests four ways of assessing the weights for the linear pool, which could in general be applied also to the other methods: (1) equal weights; (2) weights proportional to a ranking based on expert's advice; (3) weights proportional to a self-rating (each source of information provides a rank for itself); (4) weights based on some comparison of previously assessed distributions with actual outcomes.

The first approach was applied in a large part of the tests described in section 2.6. In many situations, indeed, there is no element which allows to prefer one source of information to another.

Looking for a ranking criteria for the sources of information, many possibilities exists. One consists in ranking according to the *information content* (29) of the sources.

Bordley (4) points out that in its pooling formula, beside the weights themselves, an important role is played by the sum of the weights $S_w = \sum_i w_i$. The three main configurations considered by Bordley (4) ($0 < S_w < 1$, $S_w = 1$ and $S_w > 1$) are treated in the examples of section 2.6. These configurations represent three diverse ways of incorporating the prior term in the pooling formula; they are tested on case studies in section 2.6.

When a large amount of training data is available it is possible to estimate the optimum weights (and shape factors) using an optimization method. One possibility is to optimize the parameters looking for the maximum of the log likelihood of the method adopted. This approach is applied by Ranjan and Gneiting (27) and in section 2.6 to compute the optimum weights for the different methods applied.

Another approach, which can be applied using the tau or the nu model and the explicit expressions for the weights, consists in computing the mean value of the weights for all the possible data sequences (18). However this approach requires a large amount of training data and as a consequence is rarely applicable on real-world problems.

2.6 Case studies

This section illustrates and compares the features of the probability aggregation methods described in section 2.4. The parametric methods are checked using different parameter sets, in order to show how the parameterization can control the aggregated result.

2.6.1 Two experts and a dice

This case was introduced in section 2.3. A dice is thrown and we are interested in the event E “the result is five”. Without any additional information, the prior distribution of this random process is the uniform distribution P_0 illustrated in fig. 2.1 (a). To estimate the probability $P(E)$ we use the information coming from two data sets D_1 and D_2 . The information provided by the data set D_1 tell if the face is odd or not; in our case for example the face is odd and the probability distribution provided is $P_1 = P(E|D_1)$ (fig. 2.1 (b)). The data source D_2 can only distinguish if the face is ≥ 4 or not; the distribution $P_2 = P(E|D_2)$ is provided (fig. 2.1 (c)). P_1 and P_2 are aggregated using the pooling formulas depicted in section 2.4: the disjunction of probabilities and the linear pool with equal weights (fig. 2.1 (d)); the conjunction of probabilities, the logarithmic pool with equal weights and the Bordley’s method with the sum of weights which can be greater, equal or less than one (fig. 2.1 (e)); the generalized opinion pool with equal weights (fig. 2.1 (f)); the BLP with equal weights and equal shape parameters less than one (fig. 2.1 (g)) and greater than one (fig. 2.1 (h)).

The results illustrated in figure 2.1 allow to distinguish between two types of responses with respect to the extreme value zero (impossible event). For the conjunction of probabilities, the Bordley’s method and the logarithmic pool, when a component of the combined probability mass function is zero, then the aggregated result is zero (2.1 (e)); for these methods null probability statements are vetoes. This is a common feature of the multiplicative methods. The other methods behave differently: the results provided are a mean of the aggregated probability terms. For the generalized linear pool this mean includes the prior term P_0 (fig. 2.1 (f)).

Moreover figure 2.1 illustrates the convexity of the disjunction of probabilities and of the linear methods. The conjunction of probabilities, the Bordley’s method and the logarithmic pool are non convex, while the BLP can be both convex ($\alpha, \beta \leq 1$) and non convex ($\alpha, \beta > 1$). This differentiation of features concerning the convexity still reflects the differences between methods based on the multiplication and methods based on the sum, except for the BLP, which has a bivalent behavior.

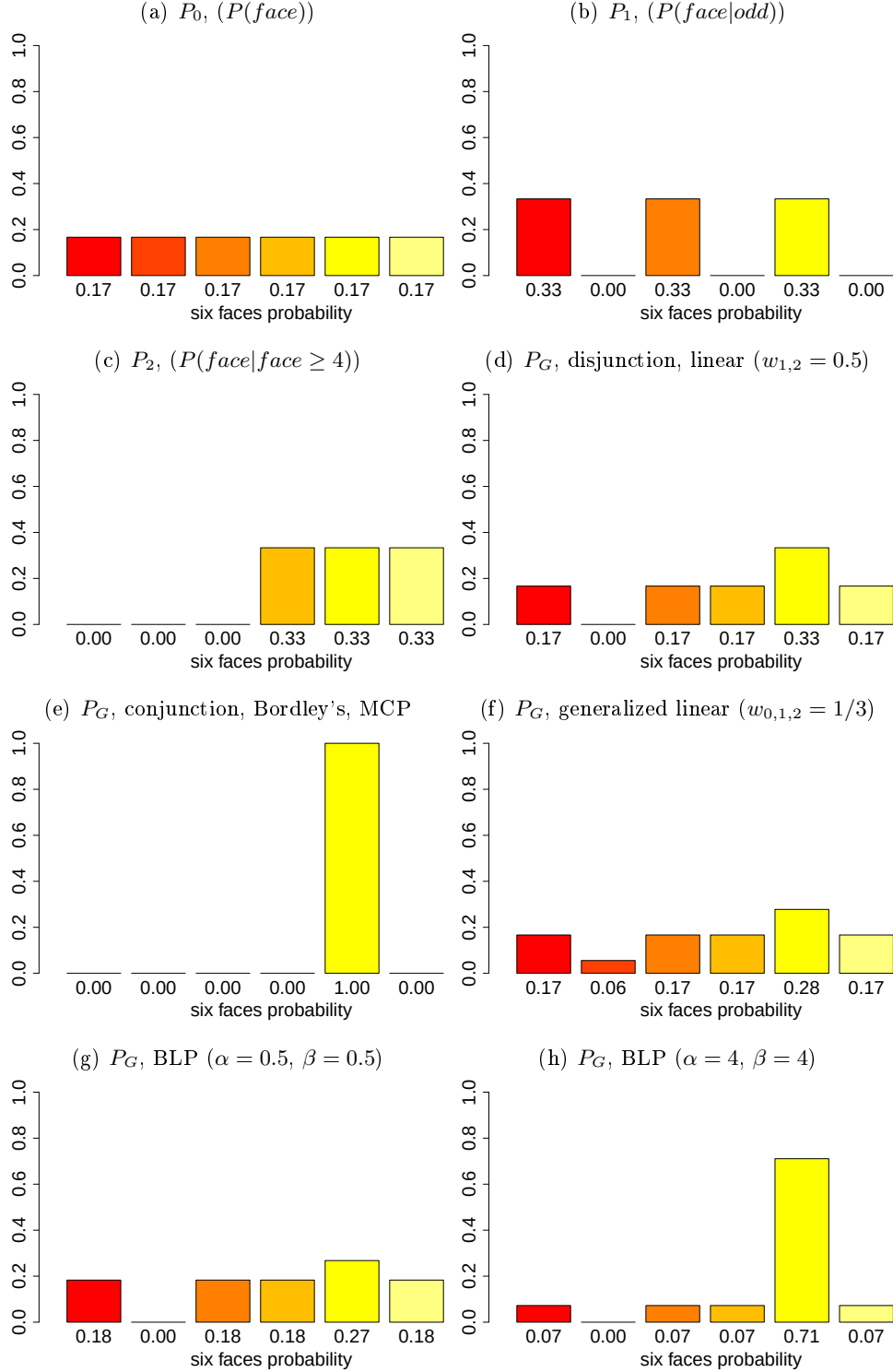


Figure 2.1: Obtaining a five when throwing a dice: (a) prior distribution P_0 for the six faces of a dice; (b) information provided by a source which can check only if the face is odd or not; (c) same as in (b) but when the source distinguish only if a face is ≥ 4 or not; (d) results of aggregation using the disjunction of probabilities and the linear method with equal weights; (e) aggregation with the conjunction, Bordley's and logarithmic method; (f) the generalized linear pool with equally weighted terms; (g) the BLP with shape parameters less than one and (h) more than one.

2.6.2 Probability densities

The information provided by different sources is often expressed in terms of probability density functions. In this example, we consider a continuous random variable x with prior law $\mathcal{N}(\mu_0, 4)$ (fig. 2.2 (a)), and two different sources of information which provide the density functions $\mathcal{N}(\mu_1, 1)$ and $\mathcal{N}(\mu_2, 1)$ respectively (fig. 2.2 (b), (c)). These density functions are combined using the following methods and the following weights sets: linear pool method (fig. 2.2 (d)); generalized linear pool method (fig. 2.2 (e)); BLP with shape parameters greater and less than one (fig. 2.2 (f)); conjunction of probabilities (fig. 2.2 (g)); logarithmic method (fig. 2.2 (h)); Bordley's method with weights with sum greater, equal or less than one (fig. 2.2 (i)). It is to be noted that all the methods except the linear pools require a normalization factor, which is computed by integrating the combined term on the interval $(-\infty, +\infty)$.

The aggregation methods can be divided into two main groups: (1) methods which provide an aggregated multi-modal distribution (fig. 2.2 (d), (e) and (f)) and (2) methods which provide an aggregated uni-modal distribution (fig. 2.2 (g), (h) and (i)). The first group contain the linear pooling formulas, its transformations (Beta) and the disjunction of probabilities. The second group the conjunction of probabilities, the logarithmic method and the Bordley's one.

The modes for the methods of the second group which take into account for the prior term are shifted with respect to the mean value $(\mu_1 + \mu_2)/2$ (fig. 2.2 (g) and (i)). This shift is in the direction of the mode of the prior distribution μ_0 when the sum of the weights in the Bordley's formula S_w is less than one; the shift is in a direction opposite to μ_0 for the conjunction of probabilities or when $S_w > 1$.

Figure 2.2 (d) illustrates the concept of convexity expressed, for the continuous case, by equation (2.5).

2.6.3 Trivariate event

Figures 2.3, 2.4, 2.5 and 2.6 illustrate the features of the aggregation methods described in section 2.4 when two probability statements P_1 and P_2 about a trivariate event with prior distribution P_0 are aggregated.

Four different configurations of P_0 , P_1 and P_2 are considered: first (fig. 2.3), a case where P_1 and P_2 both express a high probability for one outcome (in this case the first one, represented in the plots of fig. 2.3 as the first red bar), while the term P_0 is neutral; second (fig. 2.4), a case analogous to the first one but when for one outcome the probability is low; third (fig. 2.5), a case when only the prior term expresses a low probability for one outcome, while P_1 and P_2 are neutral; fourth (fig. 2.6) a case when both P_0 and P_1 express a high probability for one outcome, but P_2 remains neutral.

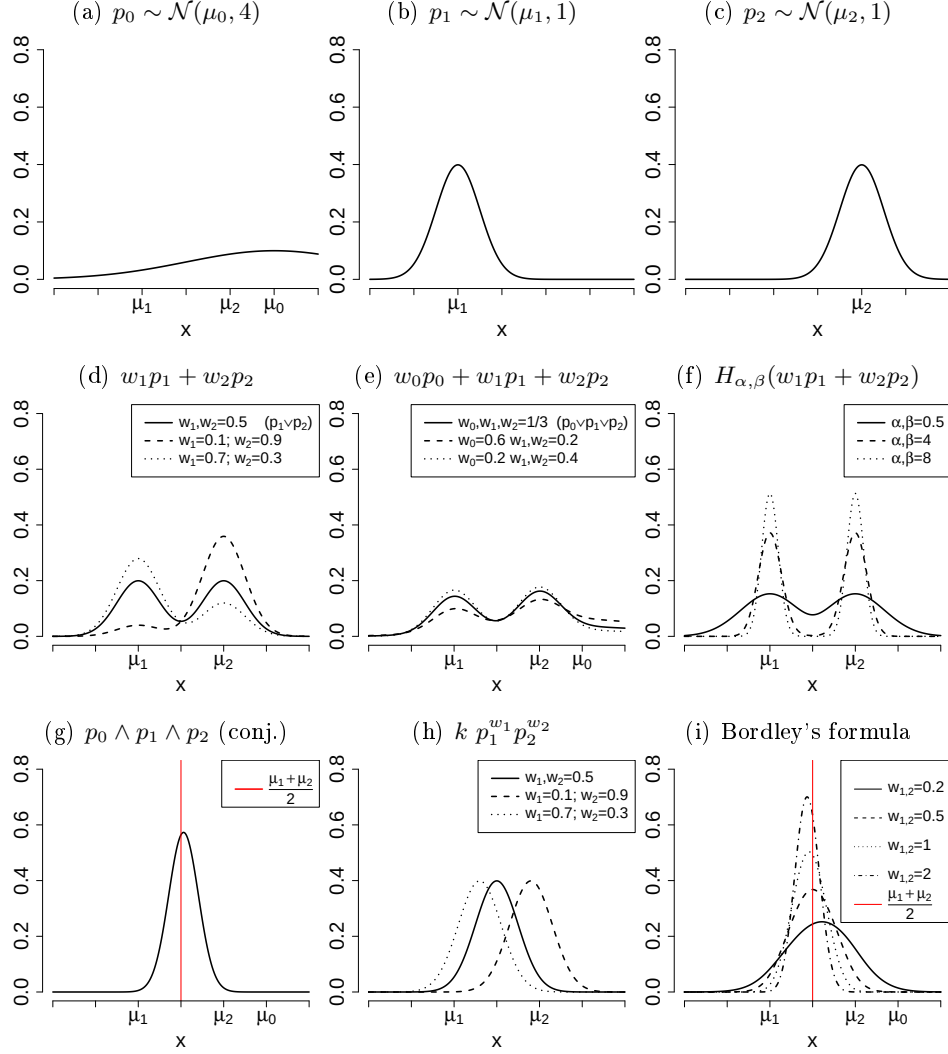


Figure 2.2: Application of different pooling methods to aggregate the probability densities (b) $p_1(x)$ and (c) $p_2(x)$, for a random variable x which have the prior density (a) $p_0(x)$: (d) linear (disjunction in the case $w_{1,2} = 0.5$); (e) generalized linear; (f) BLP; (g) conjunction; (h) logarithmic; (i) Bordley's formula.

Again, for this case study we can observe the convexity of the linear methods and of the BLP for $\alpha, \beta < 1$. Moreover, in the first three cases considered, the probabilities aggregated with linear pool, the logarithmic pool and the Bordley's formula for $S_w = 1$ are always close. In fact, when $S_w = 1$ the term P_0 in the Bordley's formula is filtered and the formula becomes closer to the one of a logarithmic pooling where the term P_0 is not considered. The similarities with the linear pool can be explained by the fact that the combined terms P_1 and P_2 are very close, and accordingly their arithmetic mean is very close to their geometric mean.

Indeed, the difference among the aggregated probabilities obtained with the three methods increases in the case of fig. 2.6, where P_1 and P_2 are different. Also, this last figure shows that in general the results obtained with the Bordley's formula and the logarithmic pool, even if close, can be different (see fig. 2.6(g) and 2.6(j)).

This case study highlights a counterintuitive behavior of the MCP and of the Bordley's formula for $S_w > 1$. Figure 2.5 shows that the trend provided by the prior term P_0 is completely reversed by the aggregation process in cases (h),(k),(l) and (m). The same can be observed in the fourth example when comparing figure 2.6(a) with 2.6 (h),(l) and (m). This behavior is analogue to the shift of the modes observed for the case study with probability densities (fig. 2.2 (i)).

Even though counterintuitive, the Bordley's formula for $S_w > 1$ and the conjunction of probabilities can represent real situations. For example, one of this situation is described in Bordley (4) (example 2 of the case II). Another situation where the setting $S_w > 1$ in the Bordley's formula is adequate is provided in the following sections.

The main lesson learned from this case study is that the choice of the weights with $S_w > 1$ in the Bordley's formula should be handled with care (an analogous observation was made by Krishnan (18)). Note nevertheless that the assumption $S_w > 1$ (for example $w_i = 1$ for each $i = 1, \dots, n$, a weak assumption of independence) is used in many applications of the Bordley's formula (*i.e.* (30; 9; 10)), and is included in the MCP method.

2.6.4 Simulation studies

Two simulation studies on two analytical models are used to compare the features of the different aggregation methods. The first model is a truncated Gaussian model, with correlation between the sources of information which are aggregated. In the second simulation study the aggregated information are independent.

For both cases, the probabilistic framework is close to the one depicted by Ranjan and Gneiting (27). The main procedure used to conduct the simulation studies is the following. The joint distribution of the random vector (X, Y) is considered, where X and Y are two random variables with a given

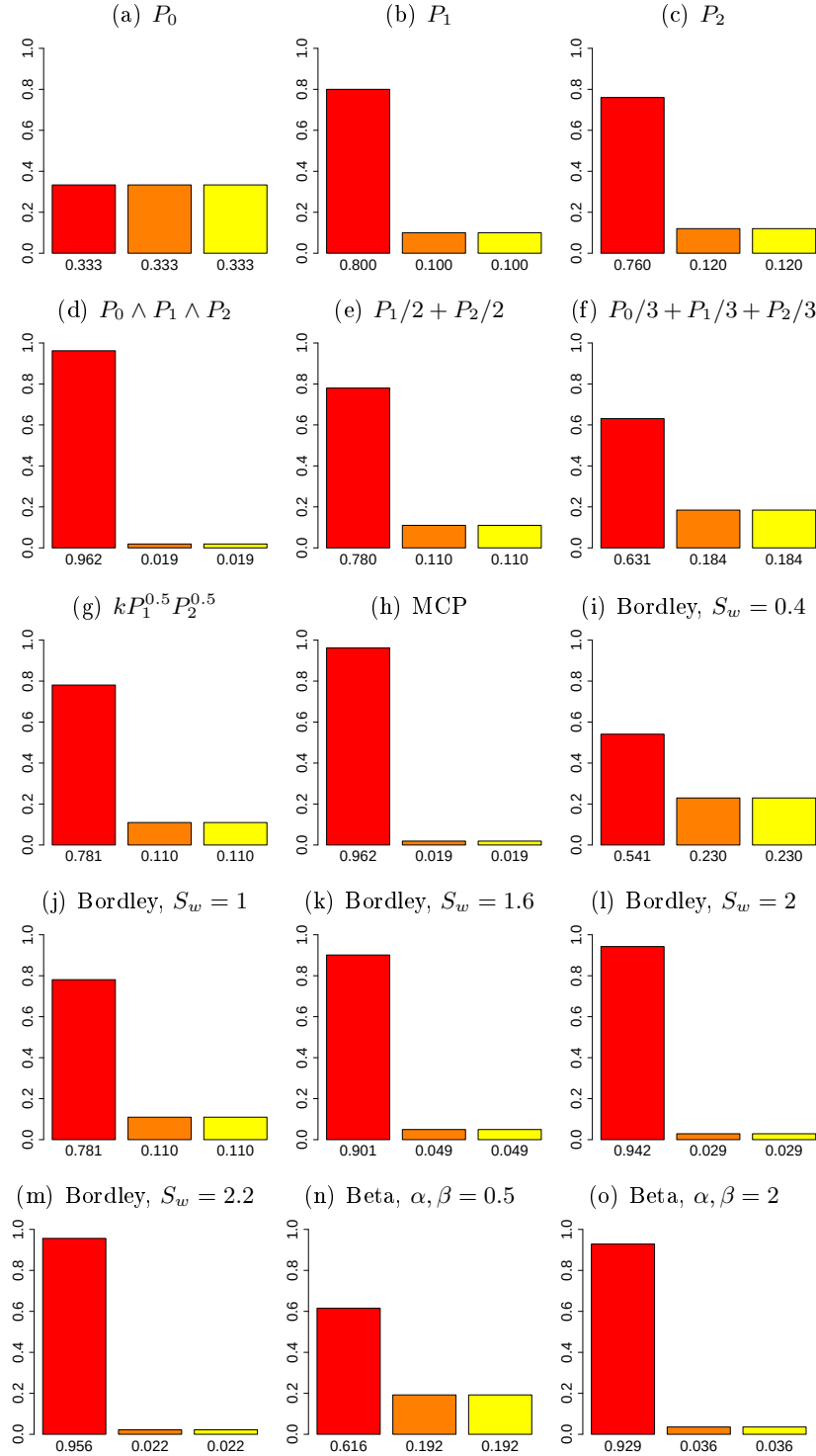


Figure 2.3: Combining the probabilities for a trivariate event. The three bars contained in each sub-figure represent each one the probability of one of the components, which is also reported at the base of each bar. In this case, two sources of information P_1 and P_2 which express a high probability for one outcome and a prior term which is neutral are combined: (a) prior term P_0 ; (b), (c) probability statements P_1 and P_2 ; (d) P_1 and P_2 combined with the conjunction of probabilities; (e) linear pool; (f) generalized linear; (g) logarithmic; (h) MCP; (i), (j), (k), (l) and (m) Bordley's method with $S_w = 0.4, 1.0, 1.6, 2.0$ and 2.2 ; (n), (o) BLP with shape factors $<$ and $>$ 1 respectively.

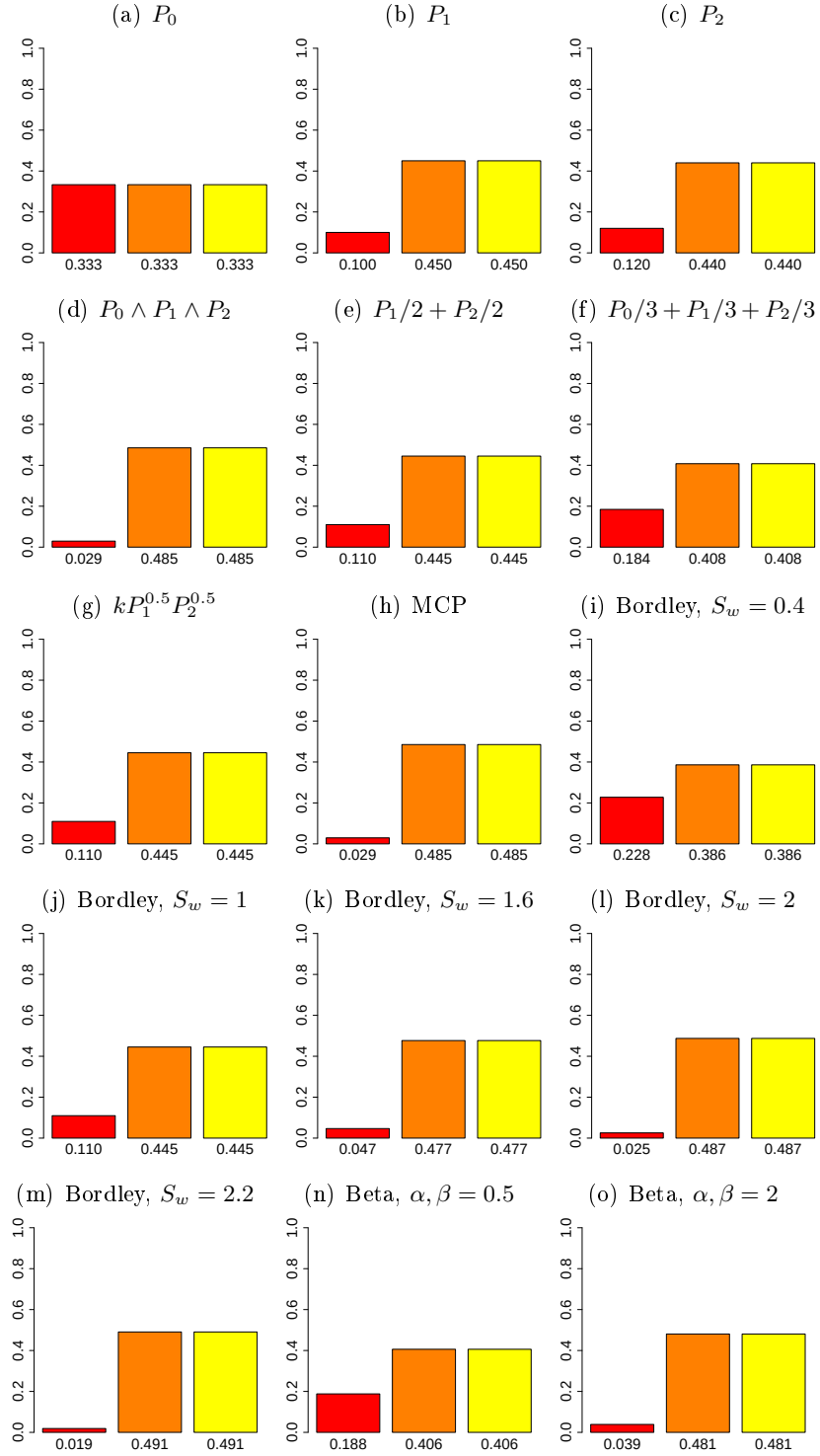


Figure 2.4: Same as fig. 2.3 but when the two sources of information P_1 and P_2 express a low probability for one outcome.

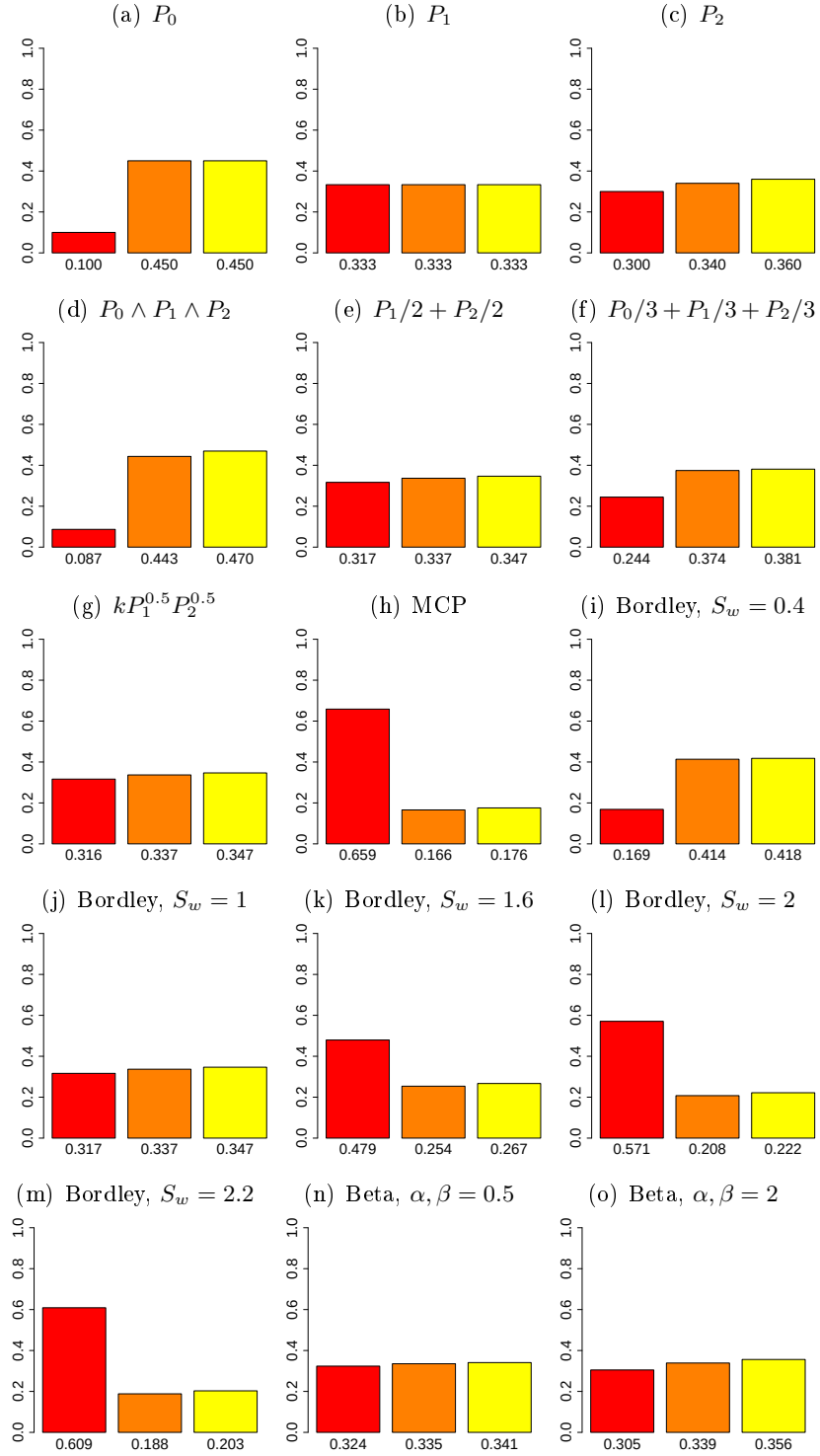


Figure 2.5: Same as fig. 2.3 but when the two sources of information P_1 and P_2 are neutral and the prior term express a low probability for one outcome.

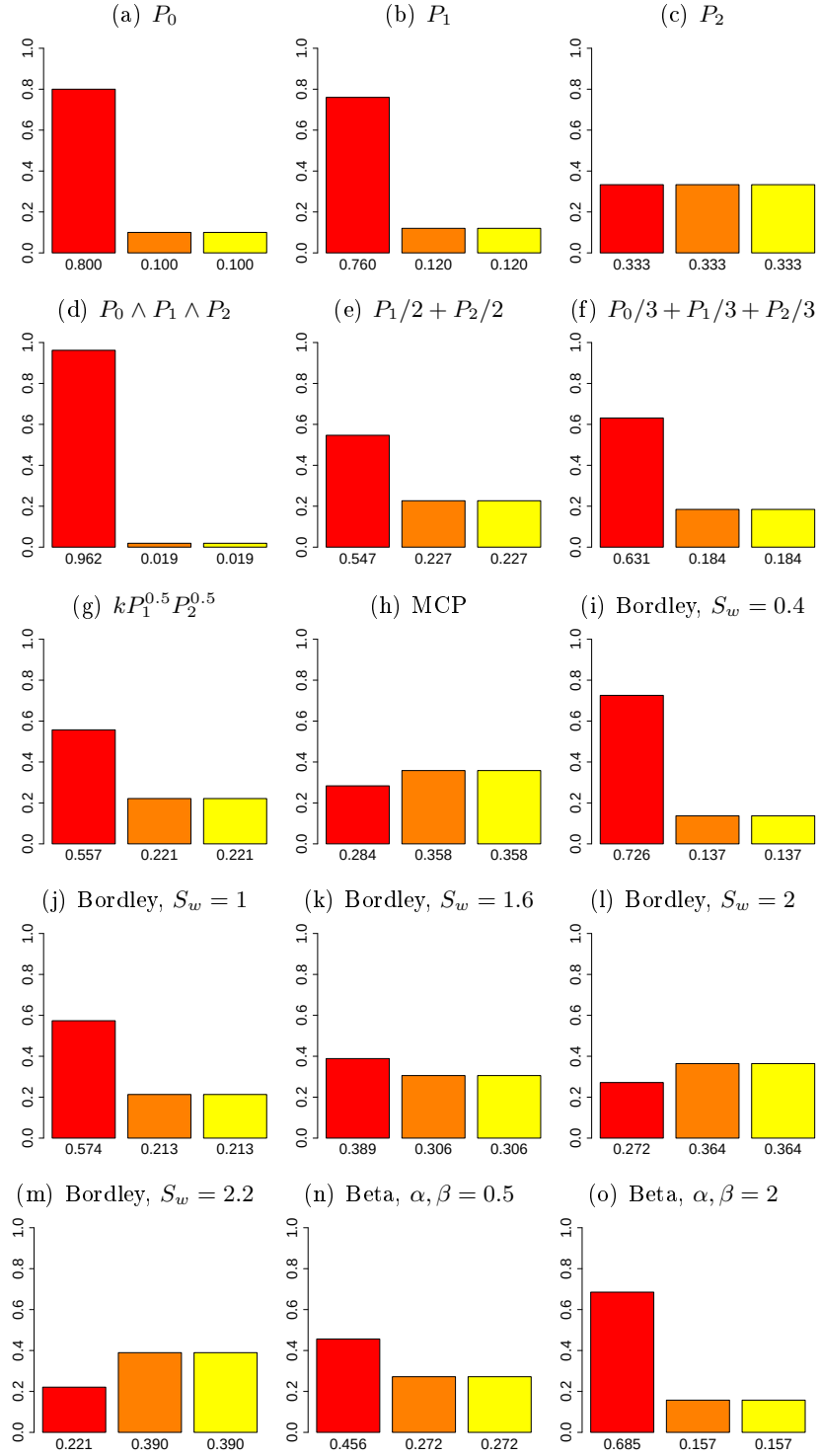


Figure 2.6: Same as fig. 2.3 but when one source of information and the prior term express a low probability for one outcome and the other source of information is neutral.

distribution law. E is a Bernoulli random variable with probability of success $P(E = 1|X, Y)$. The laws of X and Y are selected such that it is possible to compute analytically an expression for $P_1 = P(E|X)$ and $P_2 = P(E|Y)$, and the same for $P(E|X, Y)$.

Then, the distribution laws of X and Y are used to compute a number N of independent samples of the two variables. For each pair of the simulated (X, Y) , the terms $P_1 = P(E|X)$, $P_2 = P(E|Y)$ and $P(E = 1|X, Y)$ are calculated. The terms P_1 and P_2 are considered as a sample of information, which are then combined into a sample of aggregated information P_G . A sample of this aggregated information is computed for each one of the aggregation methods considered.

The terms $P(E = 1|X, Y)$ are used to create a sample of *verifications* for the event E , that is a serie of N values like for example $\{0, 0, 1, 0, \dots, 1, 0, 1\}$. At the end of the simulation procedure we have a set of N verifications values, which can represent for example the observations made about a rain/no-rain event, and a set of N probabilities of occurrence of such events (*forecast probabilities*), obtained aggregating the information coming from the simulated P_1 and P_2 .

With the verifications and the forecast probabilities calculated at the previous step it is possible to trace reliability plots and to compute some scoring rules.

Reliability plots (or calibration curves 28) are tools to evaluate the capability of a forecast in predicting correctly an event. One way to trace reliability plots (6) is to divide the forecasts into bins containing the same number of forecasts. For each bin a central value of probability can be selected and considered in abscissa. Using the bins created for the forecasts the “observed” frequency of the verifications can be computed, and used for the ordinate. The curves plotted in this way provide a visual overview of how well the predicted probabilities of an event correspond to its observed frequencies.

The quality of a forecast can be evaluated quantitatively using the Brier score (BS, 5). This quantity and its components (reliability (REL), resolution (RES) and uncertainty (UNC), (22)) are computed in order to support the illustration provided by the reliability plots (tables 2.1, 2.2 and 2.3). The mean Brier score provides an estimation of the magnitude of the probability forecast errors, the reliability quantifies the average agreement between the forecasted and the observed values; the smaller these parameters, the better the quality of the forecast. Resolution is related to the ability of providing different results for the forecasts of different events. Uncertainty represents an index of variability of the observations.

The histograms included in each reliability plot illustrate the frequency distribution of the “forecast” probabilities; they can be used to check visually the sharpness of the forecast.

Having access to the complete analytical expressions of the aggregated terms and of the aggregation results, we could compare the aggregation meth-

ods using these. Instead, hereinafter we compare the performances of the different aggregation methods using simulations and reliability plots. The reasons are the following: first, reliability plots (which are explained in detail later on) allow to check in an easy way the reliability (or calibration) of a pooling formula. Reliability, a concept which in general is applied to a probability forecast, is an important parameter related to the quality of a forecast, because it represents the agreement between the forecast values and the observed values. Moreover, some scoring rules can be applied to the simulated results in order to obtain a quantitative estimation of the reliability of a method.

Second, it is difficult to visualize and compare multidimensional models. For example, the truncated Gaussian model considered hereinafter depends on X , Y and Z . One possibility to visualize the results of an aggregation process in two dimensions is to fix the values of some variables; however this provides an incomplete representation of the results. Instead, a reliability plot provides a global overview of the results, whatever dimension they have.

Finally, simulation case studies provide a guideline for real case studies, where the analytical expressions of the involved parameters are seldom known.

Truncated Gaussian model

The first analytical model considered is the truncated Gaussian model described in Chugunova and Hu (9). The model consists of a Gaussian vector (X, Y, Z) where the three components are obtained with a linear combination of four independent standard Gaussian random variables U , S_1 , S_2 and S_3 , that is:

$$\begin{aligned} X &= \sqrt{\alpha} U + \sqrt{1-\alpha} S_1 \\ Y &= \sqrt{\beta} U + \sqrt{1-\beta} S_2 \\ Z &= \sqrt{\gamma} U + \sqrt{1-\gamma} S_3 \end{aligned} \quad (2.25)$$

with $\alpha, \beta, \gamma \in (0, 1)$.

The Gaussian vector (X, Y, Z) has the following covariance matrix:

$$\begin{bmatrix} 1 & \sqrt{\alpha\beta} & \sqrt{\beta\gamma} \\ \sqrt{\alpha\beta} & 1 & \sqrt{\alpha\gamma} \\ \sqrt{\beta\gamma} & \sqrt{\alpha\gamma} & 1 \end{bmatrix} = \begin{bmatrix} 1 & \rho_{XY} & \rho_{XZ} \\ \rho_{XY} & 1 & \rho_{YZ} \\ \rho_{XZ} & \rho_{YZ} & 1 \end{bmatrix}. \quad (2.26)$$

For the events $A(x) = \{X|X < x\}$, $B(y) = \{Y|Y < y\}$ and $C(z) = \{Z|Z < z\}$ it is possible to compute the analytical expression of $P(A)$, $P(B)$, $P(C)$ and of the joint probabilities $P(A, B)$, $P(A, C)$, $P(B, C)$ and $P(A, B, C)$. Using the definition of conditional probability we can get the analytical expression for the terms $P(A|B, C)$, $P(A|B)$ and $P(A|C)$.

Within this analytical framework, it is possible to compute, according to the

simulation procedure described above, 10,000 samples of the terms $P(A|B, C)$, $P_0 = P(A)$, $P_1 = P(A|B)$ and $P_2 = P(A|C)$ using the simulated samples of U , S_1, S_2 and S_3 . The reliability diagrams of figures 2.7 and 2.8 are plotted using the term $P(A|B, C)$ to produce the “verifications”. The verifications are simulated with a binomial law using, as probability of success, the joint probability $P(A|B, C)$, while the corresponding forecast probabilities are computed aggregating P_1 and P_2 with the methods described in section 2.4.

Figure 2.7 represents the reliability plots computed for this analytical framework when $\rho_{XY} = 0.5$, $\rho_{XZ} = 0.5$ and $\rho_{YZ} = 0.3$. The plots are drawn for the following forecasts: (a) the term $P(A|B, C)$; (b) linear pool with equal weights; (c) optimally weighted linear pool; (d) optimally weighted BLP; (e) generalized linear pool with equal weights; (f) optimally weighted generalized linear pool; (g) conjunction of probabilities; (h) Bordley’s method with sum of the weights less than one; (i) Bordley’s method with sum of the weights equal one; (j) Bordley’s method with sum of the weights greater than one ($w_{1,2} = 1$); (k) Bordley’s method with sum of the weights greater than one ($w_{1,2} = 2$); (l) optimally weighted Bordley’s method; (m) optimally weighted ν model; (n) optimally weighted logarithmic method; (o) MCP.

The optimal weights are computed using the maximum likelihood method for log likelihood functions (see for example Ranjan and Gneiting (27) for the log likelihood function for the BLP). Figure 2.8 represents the same terms as figure 2.7 but computed for a case where the correlation between the components Y and Z is higher, that is for $\rho_{YZ} = 0.9$. For the two cases the Brier score and its components are computed in order to provide a more quantitative comparison among the methods (table 2.1 and 2.2).

In the first case ($\rho_{YZ} = 0.3$, fig. 2.7) the best results are obtained with the optimized BLP, the optimized Bordley’s formula and the optimized nu model (fig. 2.7 (d),(l) and (m)). The low Brier and reliability scores confirm the observation which can be made from the figures (table 2.2). It is interesting to note that in this case also the Bordley’s formula with $w_{1,2} = 1$ and the MCP method, although not optimized, can provide reliable estimations (fig. 2.7 (j) and (o)). Their reliability is higher (that is the parameter REL is closer to zero) than the one obtained with the optimized linear pools and the optimized logarithmic pool (fig. 2.7 (c),(f) and (n)).

The situation is slightly different in the case where $\rho_{YZ} = 0.9$ (fig. 2.8). With this stronger correlation the assumptions of conditional independence of the MCP model and of the Bordley’s formula with $w_{1,2} = 1$ deteriorate the reliability obtained with these two methods. The more reliable methods are the linear pool, the optimized ones, and the optimized BLP, Bordley’s formula, nu formula and logarithmic formula. In this case two non optimized methods with a high reliability are the linear pool and the Bordley’s formula for $w_{1,2} = 0.5$.

The nu model (fig. 2.7 and 2.8 (m)), with only one optimized parameter,

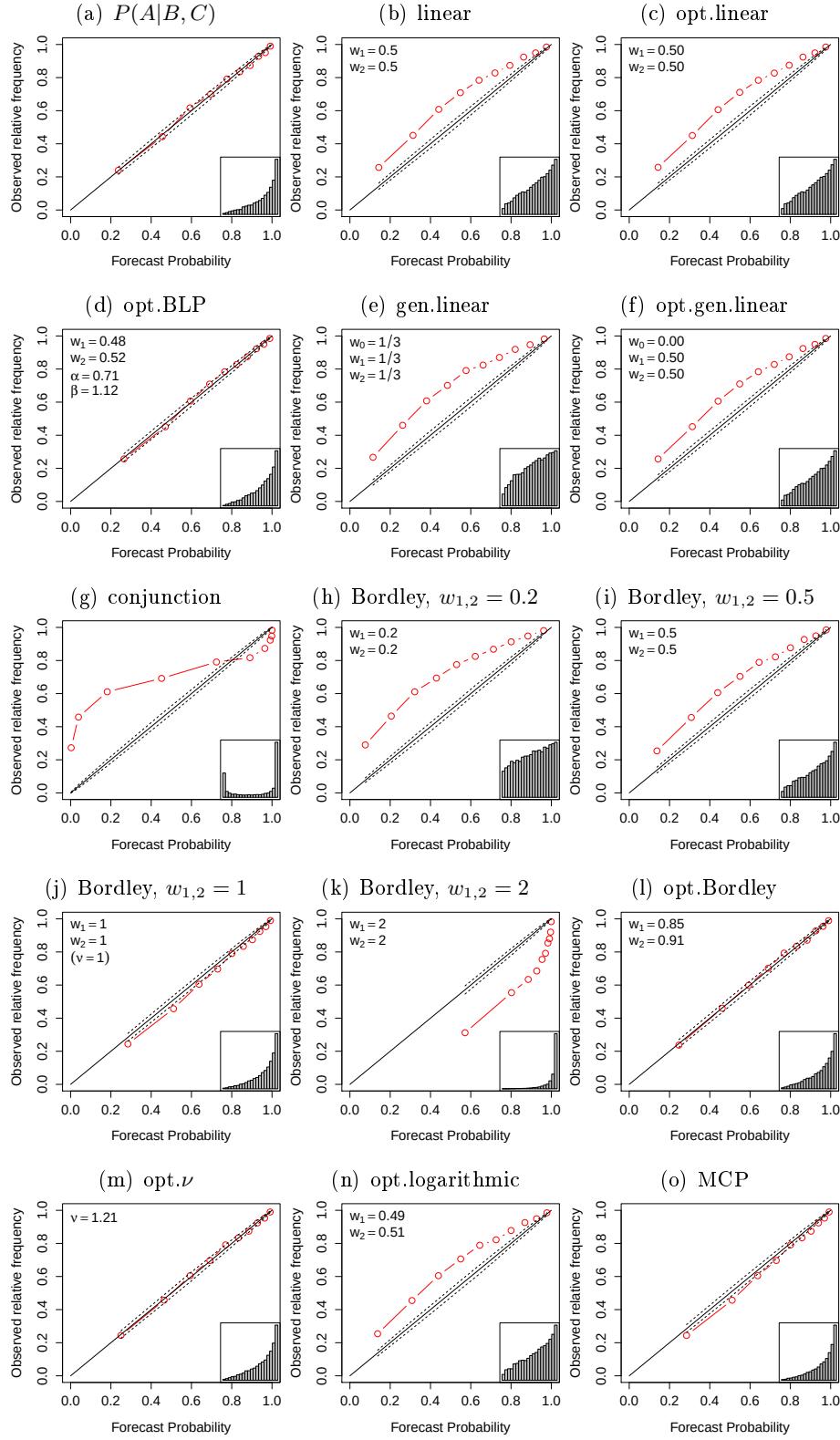


Figure 2.7: Reliability diagrams for the simulation test on a truncated Gaussian model defined by the vector (2.25) when $\rho_{YZ} = 0.3$.

	BS	REL	RES
$P(A B, C)$	0.1478	0.0002	0.0463
P_0	0.2261	0.0682	0.0359
P_1	0.1667	0.0136	0.0408
P_2	0.1663	0.0138	0.0415
linear, $w_{1,2} = 0.5$	0.1629	0.0129	0.0439
linear, optimal weights	0.1628	0.0128	0.0439
gen. linear, $w_{1,2,3} = 1/3$	0.1780	0.0263	0.0421
gen. linear, opt. weights	0.1629	0.0129	0.0439
beta, opt. weights	0.1501	0.0001	0.0439
log., opt. weights	0.1624	0.0127	0.0442
Bordley, $w_{1,2} = 0.2$	0.1953	0.0408	0.0394
Bordley, $w_{1,2} = 0.5$	0.1625	0.0127	0.0441
Bordley, $w_{1,2} = 1$ ($nu = 1$)	0.1494	0.0009	0.0453
Bordley, $w_{1,2} = 2$	0.1998	0.0358	0.0298
Bordley, opt.weights	0.1478	0.0001	0.0462
opt. ν	0.1487	0.0001	0.0453
conjunction	0.2037	0.0516	0.0418
MCP	0.1494	0.0009	0.0453

Table 2.1: Brier score, reliability term, resolution for the truncated Gaussian simulation case study when $\rho_{YZ} = 0.3$. The uncertainty term in each case is 0.2393.

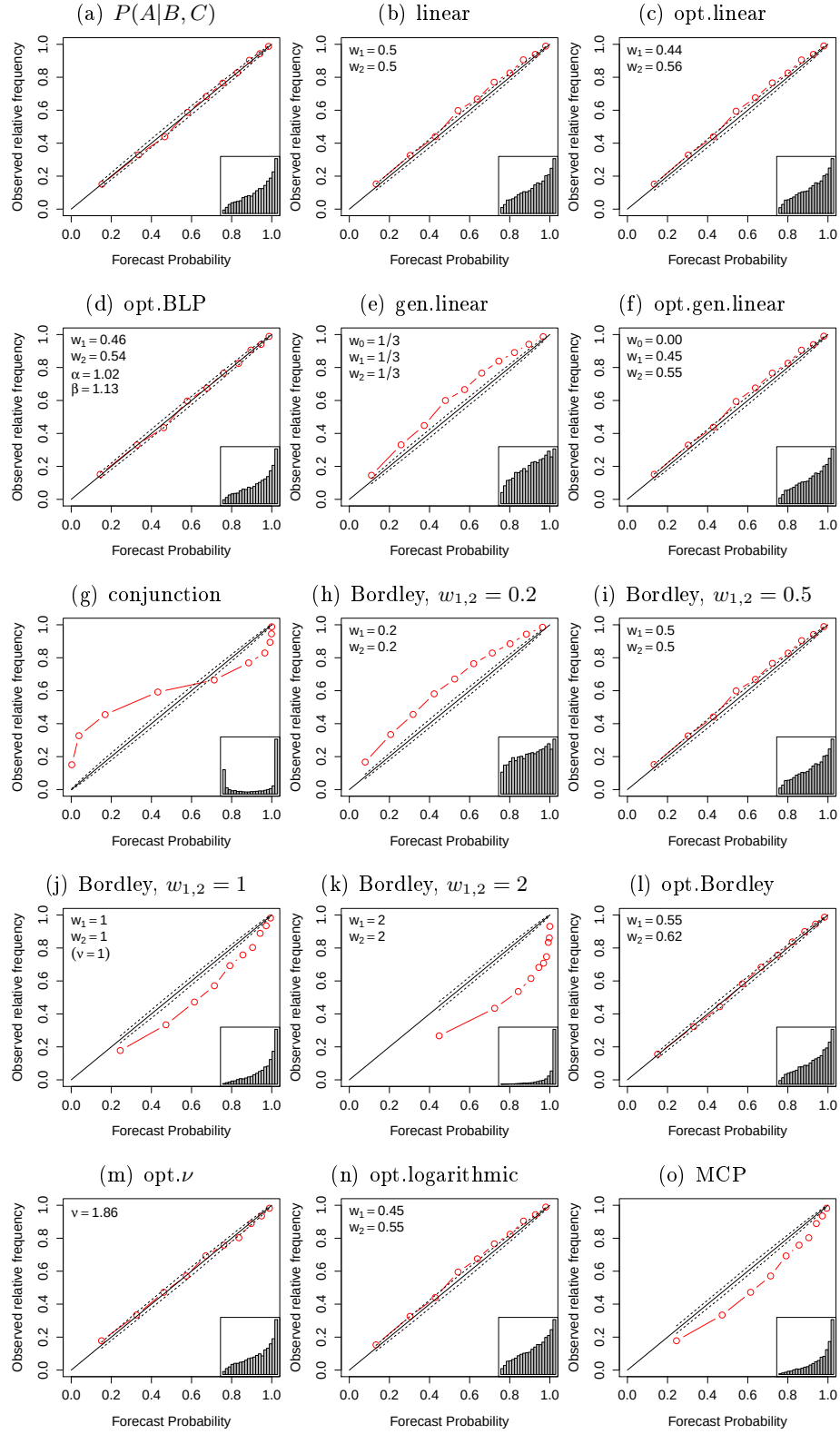


Figure 2.8: Reliability diagrams for the simulation test on a truncated Gaussian model defined by the vector (2.25) when $\rho_{YZ} = 0.9$.

	BS	REL	RES
$P(A B, C)$	0.1638	0.0001	0.0603
P_0	0.2023	0.0309	0.0525
P_1	0.1651	0.0009	0.0598
P_2	0.1653	0.0008	0.0595
linear, $w_{1,2} = 0.5$	0.1649	0.0010	0.0600
linear, optimal weights	0.1649	0.0009	0.0599
gen. linear, $w_{1,2,3} = 1/3$	0.1703	0.0062	0.0598
gen. linear, opt. weights	0.1649	0.0009	0.0599
beta, opt. weights	0.1640	0.0001	0.0600
log., opt. weights	0.1648	0.0009	0.0600
Bordley, $w_{1,2} = 0.2$	0.1805	0.0134	0.0568
Bordley, $w_{1,2} = 0.5$	0.1648	0.0009	0.0600
Bordley, $w_{1,2} = 1$ ($nu = 1$)	0.1795	0.0098	0.0542
Bordley, $w_{1,2} = 2$	0.2467	0.0532	0.0305
Bordley, opt.weights	0.1639	0.0001	0.0601
nu model, opt.nu	0.1700	0.0003	0.0542
conjunction	0.1904	0.0259	0.0594
MCP	0.1795	0.0098	0.0542

Table 2.2: Same situation of tab. 2.1 but when $\rho_{YZ} = 0.9$. The uncertainty term in each case is 0.2386.

provides results which are comparable with the result obtained by the other formulas optimized on two or three parameters. Thus, in terms of optimization costs, in this case the nu model should be preferred.

If the correlation varies from low to high values ($\rho_{YZ} = 0.3$ and $\rho_{YZ} = 0.9$), the optimized methods which conserve a good reliability in both cases are the multiplicative ones, together with the BLP.

Note that the Bordley's formula with weights $w_i = 1$ coincide with the nu model with $\nu = 1$; evidently the performances in these models deteriorate when the correlation ρ_{YZ} increases.

Independent sources of information

Another simulation study is performed using the analytical setting adopted by Ranjan and Gneiting (27). In this setting the variables considered are $a_1 \sim \mathcal{N}(0, 1)$ and $a_2 \sim \mathcal{N}(0, 2)$, and the joint probability term is a function of a_1 and a_2 :

$$P = \Phi(a_1, a_2) \quad (2.27)$$

where Φ is the standard normal cumulative distribution function. For the terms $P_1 = P(Y = 1|a_1)$ and $P_2 = P(Y = 1|a_2)$ Ranjan and Gneiting (27) compute the analytical expressions:

$$P_1 = \Phi\left(\frac{a_1}{\sqrt{3}}\right) \quad \text{and} \quad P_2 = \Phi\left(\frac{a_2}{\sqrt{2}}\right). \quad (2.28)$$

With the same procedure as depicted for the truncated Gaussian framework it is possible to simulate 10,000 samples of a_1 and a_2 , P_1 and P_2 and of the joint distribution P . The samples of P are then used as success probabilities to simulate the verifications Y . Here as prior term we consider a constant value. Figure 2.9 contains the reliability plots related to this case study for: (a) verifications simulated using the joint probability P ; (b) linear pool with equal weights; (c) optimally weighted linear pool; (d) optimally weighted BLP; (e) generalized linear pool; (f) optimized generalized linear pool; (g) conjunction of probability; (h,i,j,k) Bordley's formula, $w_{1,2} = 0.2$, $w_{1,2} = 0.2$, $w_{1,2} = 0.2$, $w_{1,2} = 0.2$; (l) optimized Bordley's formula; (m) optimized ν model; (n) optimally weighted logarithmic pool; (o) MCP. In this case the best reliability is obtained with the optimized BLP and with the optimized Bordley's formula (2.9 (d) and (l)); the other aggregation formulas, both optimized or not, are less reliable. This case study is an example of situation where, with the Bordley's formula, the suitable weighting factors have a sum $S_w > 1$. The results obtained for the previous case study (fig. 2.5 and 2.6), which at a first glance look counterintuitive, are found now to provide more reliable results.

The inadequacy of some formulas with fixed weights is evident (fig. 2.9 (b),(e), (h), (i)); however, with some of them (conjunction, Bordley's formula

with $w_{1,2} = 1, 2$ and MCP) the reliability obtained is not so distant from the limits defined by the bootstrap technique of Bröcker and Smith (6). It is interesting to note that this is true, in the case of Bordley's formula, also for high values of $w_{1,2}$. Though, here the prior term can be considered neutral as the one considered for the trivariate case of fig. 2.3 and 2.4.

The results obtained in this case with the conjunction, the Bordley's formula for $w_{1,2} = 1$ and the MCP coincide due to the choice of the prior term.

	BS	REL	RES
P	0.1401	0.0000	0.1100
P_0	0.2500	0.0002	0.0002
P_1	0.2266	0.0004	0.0238
P_2	0.1864	0.0001	0.0638
linear, $w_{1,2} = 0.5$	0.1783	0.0363	0.1080
linear, optimal weights	0.1765	0.0072	0.0806
gen. linear, $w_{1,2,3} = 1/3$	0.1975	0.0555	0.1080
gen. linear, opt. weights	0.1765	0.0072	0.0806
beta, opt. weights	0.1419	0.0000	0.1081
log., opt. weights	0.1689	0.0155	0.0966
Bordley, $w_{1,2} = 0.2$	0.2067	0.0649	0.1082
Bordley, $w_{1,2} = 0.5$	0.1688	0.0270	0.1082
Bordley, $w_{1,2} = 1$ ($nu = 1$)	0.1458	0.0040	0.1082
Bordley, $w_{1,2} = 2$	0.1435	0.0017	0.1082
Bordley, opt. weights	0.1402	0.0000	0.1099
nu model, opt. nu	0.1458	0.0040	0.1082
conjunction	0.1458	0.0040	0.1082
MCP	0.1458	0.0040	0.1082

Table 2.3: Brier score, reliability term and resolution for the simulation case study proposed by Ranjan and Gneiting (27). The uncertainty term is always 0.2500. The term P represents the joint probability.

2.7 Results and discussion

Table 2.4 recapitulates the previous sections about the aggregation methods and their properties. It provides a quick overview which allows to select a method according to the required properties. Table 2.4 shows that there is not a method that can be considered the best in each situation because first none of them has all the suited properties, and second because the properties themselves can be required in some cases but inadequate in others.

Again, table 2.4 illustrates the dichotomization between methods which aggregate the terms using the sum (linear methods) and methods based on

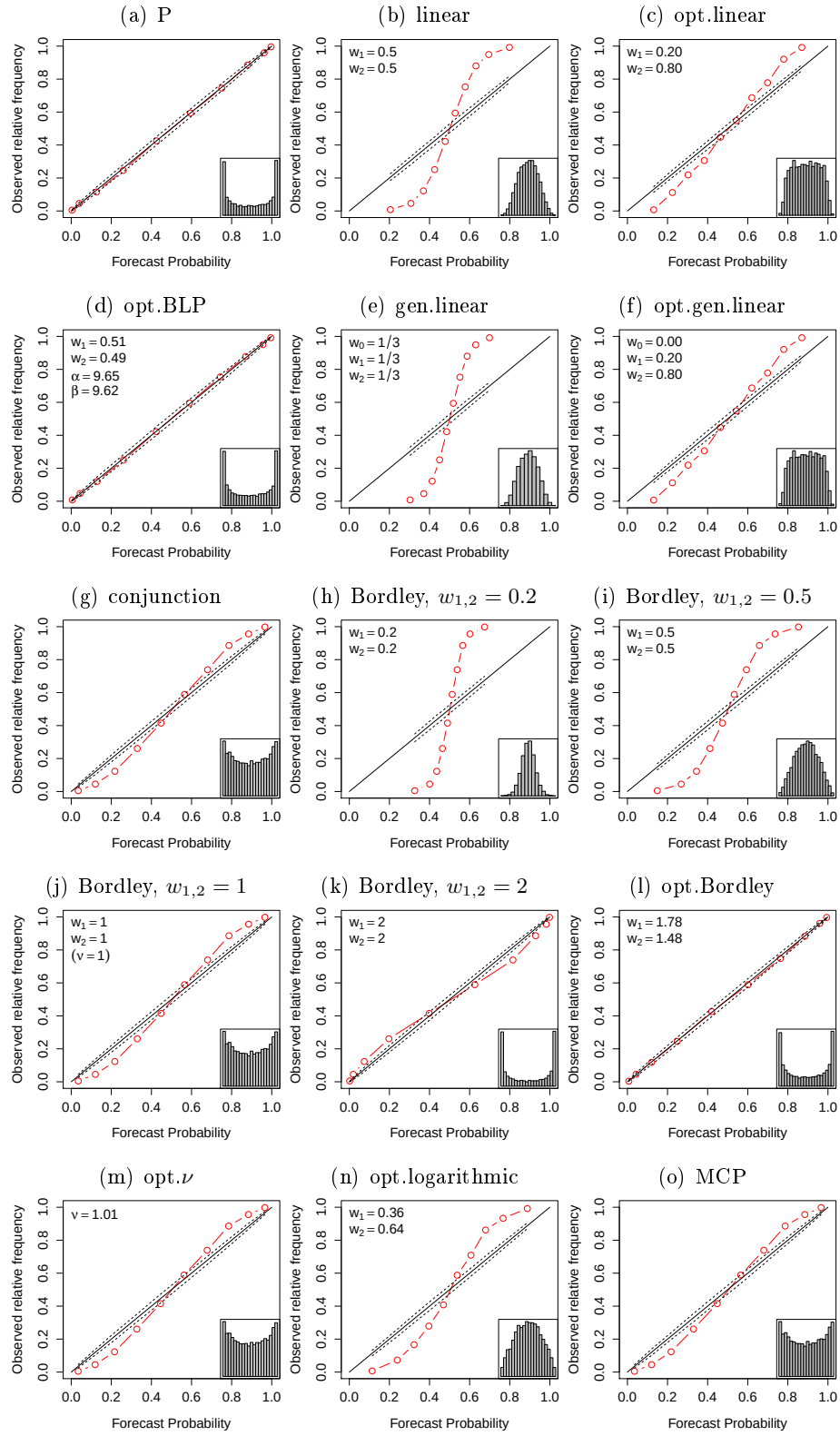


Figure 2.9: Simulation test using the analytical model proposed by Ranjan and Gneiting (27).

multiplication. The BLP, even if based on the transformation of a linear methods, has properties which lie in between the two categories.

Let us then consider each property in detail on the light of the simple case studies examined in section 2.6.

As discussed previously, the non convexity of a method can be a reasonable requirement when sources of information are distinct and probabilities should be reinforced when combining these information. The only convex methods are the linear one and the BLP for $\alpha, \beta \leq 1$. This last method with $\alpha, \beta > 1$ can be selected when the non convexity of the multiplicative methods is required, but the 0/1 forcing property is not.

However, this last property can be desirable in many frameworks, for example in frameworks related to Geology, where the occurrence of catenary patterns must be respected; and only the multiplicative methods have it.

The expression for the prior term proposed in section 2.2 allows to include it in the pooling formulas as if it were one of the terms to be aggregated, hence all the aggregation formulas proposed here can take into account for the prior information. However, some formulas (Bordley's, ν model, MCP) reserve a particular position to this term. This must be considered when using an aggregation formula, because the weight assigned to the prior term can considerably influence the aggregated result (see *i.e.* fig. 2.5 and fig. 2.6). Using the Bordley's formula, for example, it is possible to manage the influence of the prior term acting on the value of the sum of the weights S_w . When $S_w = 1$ the prior term is filtered, and using $S_w < 1$ or $S_w > 1$ its influence on the aggregated results can have opposite effects (fig. 2.5, fig. 2.6).

The external Bayesianity is an important property; in a sequential simulation framework, the methods which have this property (Bordley's formula, logarithmic pool formula and MCP) should be preferred.

Another important point to consider in the selection of an aggregation method is the number of parameters to be determined. When there are no information which can help in the inference of a specific set of weights, an aggregation method which does not require weights or one with equally fixed weights can provide reliable results. In the three case studies considered, when the sources of information are independent (fig. 2.9) or poorly correlated (fig. 2.7, $\rho_{YZ} = 0.3$), with the Bordley's formula with $w_{1,2} = 1$ and the MCP it is possible to obtain reliable results without computing optimized weights. When the correlation is higher (fig. 2.8, $\rho_{YZ} = 0.9$), the results obtained with an equally weighted linear pool and the Bordley's formula with $S_w = 1$ are more reliable than the ones obtained with the Bordley's formula with $w_{1,2} = 1$ and the MCP.

When a number of training data is accessible and the weighting factors can be determined with a technique like for example maximizing the log likelihood, then the most reliable aggregated results are obtained with the BLP and the optimized Bordley's formula. In the truncated Gaussian simulation case study, with the *nu* formula it is possible to obtain reliable results us-

ing one parameter only (fig. 2.7 and 2.8). In the other simulation case study (fig. 2.9) using two and more optimized weights (Bordley's formula and BLP) the reliability is improved.

2.8 Conclusions

Selecting a method to aggregate informations coming from different sources is a challenge. Knowing the properties of the commonly used methods in Earth Science can help in the selection of a method well suited for a given application. In this work, the main properties of the diverse aggregation methods are described, recapitulated (table 2.4) and illustrated using case studies.

The methods can be dichotomized into two main groups: methods which aggregate the terms using the sum and methods based on the multiplication. A method based on the sum should be used when the result must be convex, and the 0/1 forcing property is not required. Instead, a multiplicative method should be used when the 0/1 forcing property is required. If the external Bayesianity is required, then methods like the Bordley's formula and the logarithmic method with sum of the weights equal one can be used. The BLP lies in between the two groups, and can be for example used (with shape parameters $\alpha, \beta > 1$) when the 0/1 forcing property and the convexity are not required. Thus, when the aggregation method depends on some parameters, these last can be used to tune and adapt the features of the method to some requirements. However, the choice of these parameters is a delicate point. For example, the Bordley's formula with sum of the weights greater than one can lead to counterintuitive results (fig. 2.5 and 2.6).

The simulation case studies show that the most reliable methods in case where optimized weights can be computed are the BLP and the Bordley's formula. With the ν model it is possible to obtain reliable results with only one optimized parameter, but this is not true in all situations (fig. 2.9(m)). When training data are not accessible, then equally weighted pooling formulas or formulas which does not require weighting factors can be applied. On two of the three simulation cases, when there is poor or no dependence among the sources of information, discretely reliable results can be obtained with the MCP and the Bordley's formula with weights equal one. In one case study, where the correlation between two sources of information is higher, then the reliability of these two methods deteriorates.

Even if not exhaustive, the case studies included in this work provide useful guidelines for the selection of a pooling formula and illustrate the main properties, the strengths and weaknesses of the most commonly used probability aggregation methods in Earth Sciences.

	linear	BLP	logarithmic	conjunction	Bordley	ν	MCP
convexity	<i>yes</i>	*	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>
0/1 forcing prop.	<i>no</i>	<i>no</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
role of the prior	<i>normal</i>	<i>normal</i>	<i>normal</i>	<i>normal</i>	<i>special</i>	<i>special</i>	<i>special</i>
marginalization	<i>yes</i>	<i>no</i> *	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>	<i>no</i>
ext. Bayesianity	<i>no</i>	<i>no</i>	<i>no</i> *	<i>no</i>	<i>yes</i>	<i>yes</i>	<i>yes</i>
# of parameters	$n+1$	$n+1+2$	$n+1$	0	$n+1$	1	0

Table 2.4: Recapitulation of the properties of methods for aggregating n sources of information and a prior term. The symbol * means that the property depends on the parameters (w_i , α and β); the symbol * on superscript it denotes that the property is valid in general, but may change for some parameter sets (*i.e.* $w_i = 1$ and $w_j = 0$ for each $j \neq i$). Note that the linear case includes the disjunction of probabilities (when $w_i = 1/n$ for $i = 1, \dots, n$), and the nu model, even if it is a special case of the Bordley's formula, it is considered separately in order to show that its formulation reduce the number of parameters to be determined.

Bibliography

- [1] Allard, D., D'Or, D., and Froidevaux, R. (2010). Estimating and simulating spatial categorical data using an efficient maximum entropy approach. Accepted for publication in the *European Journal of Soil Science*.
- [2] Bacharach, M. (1979). Normal bayesian dialogues. *Journal of the American Statistical Association*.
- [3] Benediktsson, J. and Swain, P. (1992). Consensus Theoretic Classification Methods. *IEEE Transactions on Systems Man and Cybernetics*, 22(4):688–704.
- [4] Bordley, R. F. (1982). A multiplicative formula for aggregating probability assessments. *Management Science*.
- [5] Brier, G. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, 78:1–3.
- [6] Bröcker, J. and Smith, L. A. (2007). Increasing the reliability of reliability diagrams. *Weather and Forecasting*, 22(3):651–661.
- [7] Caers, J. (2006). A general algorithm for building 3D spatial laws from lower dimensional structural information. Technical report, Stanford Center for Reservoir Forecasting.
- [8] Christakos, G. (1990). A bayesian/maximum-entropy view to the spatial estimation problem. *Mathematical Geology*, 22(7):763–777.
- [9] Chugunova, T. and Hu, L. (2008a). An assessment of the tau model for integrating auxiliary information. In Ortiz, J. M. and Emery, X., editors, *Geostats 2008 - VIII International Geostatistics Congress*. Mining Engineering Department, University of Chile.
- [10] Chugunova, T. L. and Hu, L. Y. (2008b). Multiple-point simulations constrained by continuous auxiliary data. *Mathematical Geosciences*, 40(2):133–146.
- [11] Clemen, R. T. and Winkler, R. L. (1999). Combining probability distributions from experts in risk analysis. *Risk Analysis*.
- [12] Clemen, R. T. and Winkler, R. L. (2007). Aggregating probability distributions. In *Advances in Decision Analysis*. Edwards, W and Miles, R.F. and von Winterfeldt, D.
- [13] Genest, C. (1984). Pooling operators with the marginalization property. *The Canadian Journal of Statistics*.

- [14] Genest, C., McConway, K. J., , and Schervish, M. J. (1986). Characterization of externally bayesian pooling operators. *The Annals of Statistics*.
- [15] Genest, C. and Wagner, C. G. (1987). Further evidence against independence preservation in expert judgement syntheses. *Aequationes Mathematicae*, 32(1):74–86.
- [16] Genest, C. and Zidek, J. V. (1986). Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*.
- [17] Journel, A. G. (2002). Combining knowledge from diverse sources: An alternative to traditional data independence hypotheses. *Mathematical Geology*.
- [18] Krishnan, S. (2008). The tau model for data redundancy and information combination in earth sciences: Theory and application. *Mathematical Geosciences*.
- [19] Laddaga, R. (1977). Lehrer and the consensus proposal. *Synthese*.
- [20] Lehrer, K. and Wagner, C. (1983). Probability amalgamation and the independence issue: A reply to laddaga. *Synthese*.
- [21] Mariethoz, G., Renard, P., and Froidevaux, R. (2009). Integrating collocated auxiliary parameters in geostatistical simulations using joint probability distributions and probability aggregation. *Water Resources Research*, 45.
- [22] Murphy, A. H. (1973). A new vector partition of the probability score. *Journal of applied meteorology*, 12:595–600.
- [23] Murphy, A. H. (1993). What is a good forecast? an essay on the nature of goodness in weather forecasting. *Weather and Forecasting*, 8:281–293.
- [24] Okabe, H. and Blunt, M. J. (2004). Prediction of permeability for porous media reconstructed using multiple-point statistics. *Physical Review*.
- [25] Okabe, H. and Blunt, M. J. (2007). Pore space reconstruction of vuggy carbonates using microtomography and multiple-point statistics. *Water Resources Research*.
- [26] Polyakova, E. I. and Journel, A. G. (2007). The nu expression for probabilistic data integration. *Mathematical Geology*.
- [27] Ranjan, R. and Gneiting, T. (2008). Combining probability forecast. Technical Report 543, Department of Statistics, University of Washington.
- [28] Sanders, F. (1963). On subjective probability forecasting. *Journal of applied Meteorology*, 2:191–201.

-
- [29] Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656.
 - [30] Strebelle, S., Payrazyan, K., and Caers, J. (2002). Modeling of a deep-water turbidite reservoir conditional to seismic data using multiple-point geostatistics. In *SPE 77425*. Society of Petroleum Engineers Inc.
 - [31] Strebelle, S., Payrazyan, K., and Caers, J. (2003). Modeling of a deepwater turbidite reservoir conditional to seismic data using principal component analysis and multiple-point geostatistics. *SPE Journal*, 8(3):227–235.
 - [32] Tarantola, A. (2005). *Inverse Problem Theory*. SIAM.
 - [33] Tarantola, A. and Valette, B. (1982). Inverse problems = quest for information. *Journal of Geophysics*, 50:159–170.
 - [34] Wagner, C. (1984). Aggregating subjective probabilities: some limitative theorems. *Notre Dame J. Formal Logic*, 25(3):233–240.
 - [35] Winkler, R. L. (1968). The consensus of subjective probability distributions. *Management Science*, 15(2):B61–B75.

Chapter 3

3D multiple points statistics
simulation using 2D training
images

Abstract

One of the main issues for the application of multiple-point statistics (MPS) for the simulation of three-dimensional (3D) blocks is the lack of a suitable 3D training image.

In this work, we compare three methods to overcome this issue using the information coming from bi-dimensional (2D) training images. One method is based on the aggregation of probabilities, and it is tested using two different aggregation formula, the linear opinion pool and the Bordley's formula. The other methods are novel approaches. One relies on merging the lists obtained using the *impala* algorithm from diverse 2D training images, creating a list of compatible data events which is then used for the MPS simulation. The other method (**s2Dcd**) is based on sequential simulations of 2D slices constrained by the conditioning data computed at the previous simulation steps.

These three methods are tested on the reproduction of two 3D images which are used as reference: a micro-computed tomography of a sandstone and a simulation of a fluvio-glacial aquifer analog obtained with a hierarchical geostatistical framework. Another benchmark considered here is a real case study, where two training images of sedimentary structures observed in Brussels Sand are considered.

The tests performed with the proposed methods show that it is possible to obtain 3D MPS simulations with at least two 2D training images. The simulations obtained with these methods, in particular those obtained with the **s2Dcd** method, are close to the references from a visual point of view, in terms of facies proportions, in terms of connectivity and in terms of equivalent conductivity. When a 3D reference is not available, the comparisons made on the conductivity tensors of the training images show a good agreement with the simulations obtained with the **s2Dcd** method.

The CPU time required to simulate with the method **s2Dcd** is from two to four orders of magnitude smaller than the one required by a MPS simulation performed using a 3D training image, while the results obtained are comparable using the majority of the comparison criteria considered in this work. This computational efficiency and the possibility to use MPS for 3D simulation without the need of a 3D training image facilitates the inclusion of MPS in Monte Carlo, uncertainty evaluation and stochastic inverse problems frameworks.

3.1 Introduction

Multiple-point statistics (MPS, 11; 31) is a geostatistical simulation technique which allows to reproduce complex geological patterns and to provide more realistic results than standard geostatistical techniques. Recent imple-

mentations of the MPS technique allow to include efficiently auxiliary information (7; 28) and to solve the computational issues related to the simulation of large grids in terms of RAM and CPU time (28). Still an important barrier remains for applying the MPS simulation technique: the lack of training images, especially for three-dimensional (3D) situations. A suitable training image (TI), i.e. a conceptual statistical model of the geology which has to be simulated, is the fundamental requirement for MPS simulation. Finding a 3D TI is often challenging. This is claimed for example in the work of Huysmans and Dassargues (13), where two 2D training images are deduced from two quarry walls in the Brussels Sands formation (Belgium), or in the work of Le Coz et al. (19), where only one horizontal training image is available from detailed aerial/satellite images of the Komadougou Yobé alluvium (Niger).

The lack of a 3D TI can be faced by using different approaches. One possibility is to build the 3D TI using object-based or process-based techniques (for a review of the methods see Koltermann and Gorelick (17) and de Marsily et al. (10)). However, in many cases it can be difficult to apply these techniques. For example, with an object-based technique it is challenging to reproduce all the kind of geological shapes which can be represented by an outcrop mapped at high resolution, or to account for non stationarity. To overcome this difficulty, one can for example mix different techniques and a hierarchical approach (9), but the simulation framework can become complex. Similarly, the use of process-based techniques to generate training data sets can lead to complex algorithms (22).

Another possibility, which is explored in this work, is to use the statistical information obtained from several 2D training images.

Along this line of research, Okabe and Blunt (23; 24) aggregated with a linear pooling formula the information coming from a 2D thin section of a carbonate rock in order to reproduce the micro-pores at the sub-micron scale. In this way they improved the estimation of the porosity of a sample of carbonate rock where the 3D porosity at macro scale was investigated using micro-computed tomography imaging.

The statistical information coming from variograms and 2D training images was used by Caers (5) to simulate geological structures at different scales. Caers (5) aggregates the statistical information using the tau model (15) (or the Bordley's formula, (4)), and obtains encouraging results.

This paper is an attempt to investigate in more details the problems studied by Okabe and Blunt (23; 24) and Caers (5) and to propose new approaches to face the *interesting research topic* (12) of 3D multiple-point simulation using 2D training images.

In section 3.2, the terminology used to describe the multiple-point statistics is introduced; section 3.3 contains a description of two novel methods and the description of the method based on the aggregation of probabilities. The three methods are then tested and compared using different benchmarks;

when a 3D training image is accessible, a direct comparison of the results with a reference using the criteria depicted in section 3.4 is made. The simulation methods are tested and compared to the simulation of a micro computed tomography image (section 3.5), a sedimentary environment observed in the Brussels sands (section 3.6) and a fluvio-glacial aquifer analog (section 3.7). The last two sections contain results and discussions.

The approaches presented here are illustrated using a 3D simulation as target and 2D training images as sources of information. However, they can be extended to a more generic framework where the target simulation has dimension m , and the information is collected from sources of dimension $n < m$.

3.2 Background and terminology

This section reports the definitions and the notation required to describe the methodologies illustrated in the following. The notation and the MPS implementation algorithm considered here are the one developed by Straubhaar et al. (28) in the paper which describes the *impala* algorithm.

The MPS method is a sequential simulation technique. All the nodes are simulated in a sequential order, and the same algorithm is applied repeatedly while accounting from the previously simulated nodes (30).

At a given iteration, one has to consider the current node to be simulated. Its location is denoted u (a vector of coordinates). A *search template* is then defined. It consists in N lag vectors $h_1, \dots, h_N$ in a 2D or 3D space. The search template τ of size N referred to a node location u is defined as a set of locations centered on u :

$$\tau(u) = \{u + h_1, \dots, u + h_N\}. \quad (3.1)$$

A *data template* is then associated to a search template. If s is a function which associates the value of a facies code to a spatial location, then a data template d of size N at the location u is the vector of the facies codes at all the neighboring nodes of u :

$$d(u) = \{s(u + h_1), \dots, s(u + h_N)\}. \quad (3.2)$$

Once defined a search template, and before starting the sequential simulation algorithm, the first step of a multiple-point simulation is the construction of a catalog made of all the data templates contained in the training image. The catalog contains information about frequency of the occurrence of a facies code at the central node of a data template.

All the approaches presented hereinafter are based on the *impala* algorithm and on the use of lists. Even if the following methodology could be

extended to other data storage paradigms (*i.e* the *three* of *snesim* algorithm, 31), its implementation is easier with lists. Moreover, the use of the lists allows improved computational performances on parallel machines.

A list $\mathcal{L}$ is a collection of list elements L , where each element is composed by two vectors (c, d) . The vector $d = (s_1, \dots, s_N)$ contains one of the facies configurations which are observed in the training image scanned with the search template τ . Given a training image composed of M different facies codes, the vector $c = (c_1, \dots, c_M)$ contains the counters for the occurrences of each facies at the center of the search template (see Straubhaar et al. (28) for more details).

3.3 Methodology

In order to face the lack of a complete 3D training image when a 3D simulation is required, we propose and compare three methods. The first is based on probability aggregation methods; the second is based on 2D simulation constrained by the conditioning data computed at a previous 2D simulation step; the third is based on the mixing of compatible data events of the lists extracted from different 2D training images.

Note that in the following we consider to have access to at least two different training images oriented along perpendicular directions. When only one training image is accessible, we assume that this same image can be used to depict the required features along another direction too.

3.3.1 Probability aggregation

The first method is based on the aggregation of probability, that was already explored by Okabe and Blunt (23; 24) and Caers (5). In this paper, we apply this concept using two aggregation methods, we test it more extensively than the above authors using synthetic and real case studies and we check how the results obtained depend on the weighting factors.

The principle of the method, illustrated in fig. 3.2, is the following. Since a complete 3D multiple-point conditional probability distribution function (cpdf) is not accessible, an approximation is obtained by aggregating the probabilities which are computed from several sources of information having a lower dimensionality.

Consider for example a case where N bi-dimensional training images $TI_1, \dots, TI_N$ are accessible. These images can be for example three training images having their normal vectors oriented along the coordinates of a 3D Cartesian coordinate system (fig. 3.1(b), 3.1(d) and 3.1(e)). The first steps of a MPS simulation performed with this approach proceed for each training image separately like a standard simulation. The training images $TI_1, \dots, TI_N$ are scanned separately using corresponding search templates $\tau_1, \dots, \tau_N$. For example, the search templates can have the shapes illustrated

in fig. 3.3(a), (b) and (c). The node locations h_i which compose each data template are selected in order to lie in the plane defined by the corresponding training image. During the scan, one catalog (a *list* in case of the *impala* algorithm) for each training image is created. Then, at a point u of the simulation domain, the conditional probabilities densities $\mathbb{P}(s(u) = k|d_1(u)), \dots, \mathbb{P}(s(u) = k|d_N(u))$ of finding the facies k are computed for each training image and for each facies. Here the terms $d_1, \dots, d_N$ are the data events found in the simulation grid using the corresponding search templates $\tau_1, \dots, \tau_N$. In the following, the cpdf obtained from the different training images will be denoted for the sake of brevity $\mathbb{P}_1, \dots, \mathbb{P}_N$.

At this step, the different cpdf are aggregated into a global probability term $\mathbb{P}_G$, which is an approximation of the conditional probability

$$\mathbb{P}_G(u) \approx \mathbb{P}(s(u) = k|d_1(u), \dots, d_N(u)). \quad (3.3)$$

$\mathbb{P}_G$ is used to draw a given facies at the point u , then the simulation proceeds sequentially using the same approach.

How to aggregate the different cpdf coming from the different training images? Comunian et al. (8) propose a critical overview of the methods which can be used to aggregate probability terms. Here two methods which represents two different categories of aggregation methods are tested: the *linear pooling formula* and the *Bordley's formula* (4) (or *tau model*, (15)). For the sake of brevity, in the following we will often refer to the two methods using the acronyms **paLin** (probability aggregation using a **L**inear pooling formula) and **paBor** (probability aggregation using the **B**ordley's pooling formula).

Linear pooling formula

The first aggregation method considered is the linear pooling formula; it aggregates the probability terms using a weighted sum. The method is convex and it does not have the external Bayesianity property, a property which is suitable in the case of sequential simulations (see Comunian et al. (8) for details). However, the method is appealing because of its flexibility and its simplicity. It is widely used (for example it was used by Okabe and Blunt (23; 24)). With this method, the aggregated probability is given by:

$$\mathbb{P}_G(u) = \sum_{i=1}^N w_i \mathbb{P}_i \quad \text{with } w_1, \dots, w_N \in \mathbf{R}^+ \quad (3.4)$$

In order to obtain a meaningful probability term $\mathbb{P}_G$, the weights w_i are selected so that their sum S_w equals one.

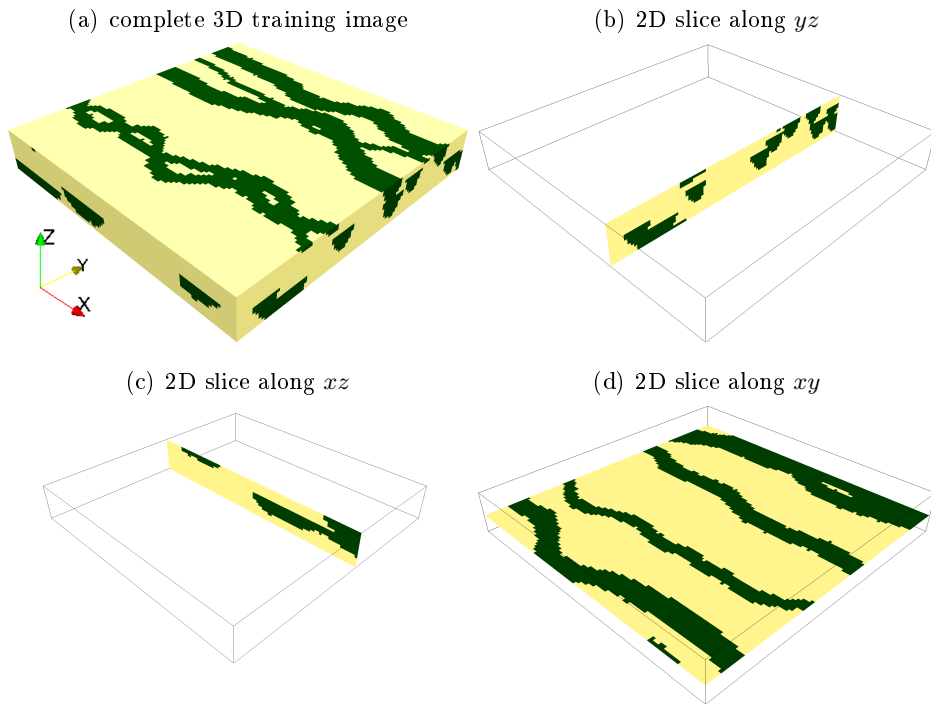


Figure 3.1: (a) A complete 3D training image containing channel structures and (b), (c), (d) its slices along different orientation planes. These slices can be used as 2D training images in the proposed approaches.

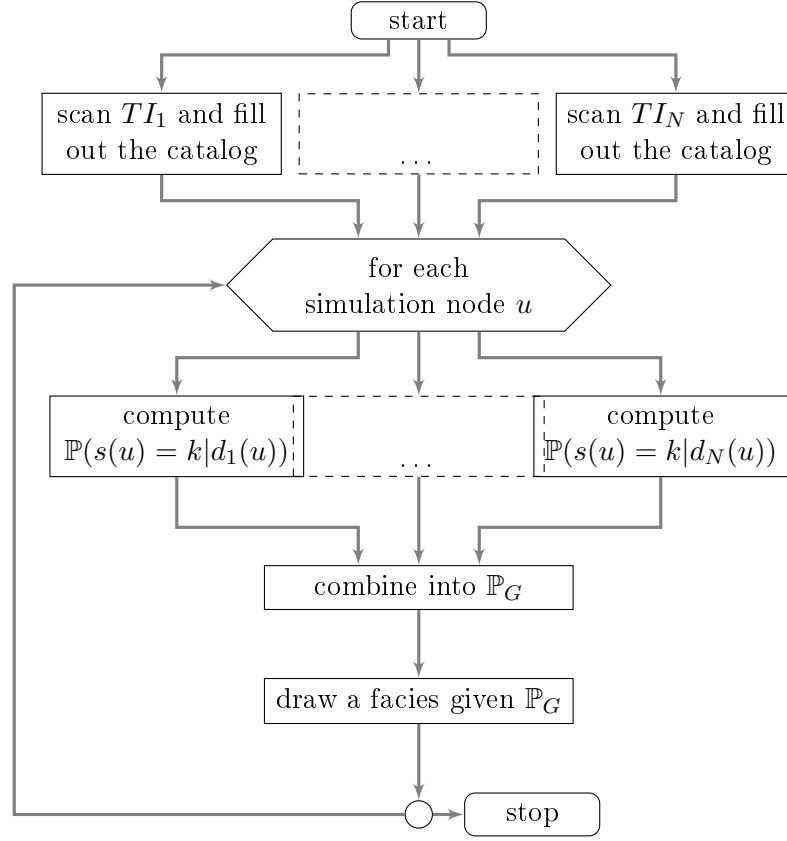


Figure 3.2: The procedure to perform 3D MPS simulations aggregating the probabilities from 2D training images. The dashed blocks represent operations which are performed for each considered training image.

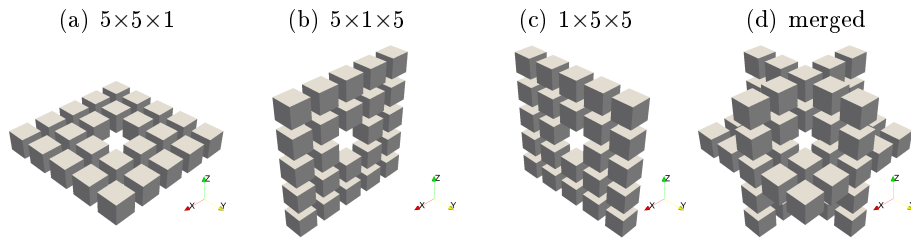


Figure 3.3: Examples of bi-dimensional search templates used to scan training images oriented along different direction ((a),(b) and (c)) and a search template obtained by merging the first three search templates ((d)).

Bordley's formula

The other aggregation method considered is the Bordley's formula, which is expressed in terms of odds:

$$\mathbb{O}_G(u) \propto \mathbb{O}_0^{1-S_w} \prod_{i=1}^N \mathbb{O}_i^{w_i} \quad \text{and} \quad S_w = \sum_{i=1}^N w_i. \quad (3.5)$$

The odds $\mathbb{O}_i$ are related to the corresponding probability terms $\mathbb{P}_i$ by the following relation:

$$\mathbb{O}_i = \frac{\mathbb{P}_i}{1 - \mathbb{P}_i} \quad \text{with} \quad i \in \{G, 0, 1, \dots, N\}. \quad (3.6)$$

In equation (3.5), the term $\mathbb{O}_0$ is the odd of the prior probability term $\mathbb{P}_0$. This last term represents a prior information that the Bordley's formula allows to integrate in the aggregation process; it can for example be used to incorporate the target facies proportions derived from a previously performed survey. The weights $w_1, \dots, w_N \in \mathbf{R}^+$ can be selected such that their sum S_w is greater, lower or equal to one. In the last case the contribution of $\mathbb{P}_0$ is ignored. More details about the other cases can be found in Bordley (4) and Comunian et al. (8).

The Bordley's formula has properties which are desirable in a geostatistical simulation framework. One of these properties is the *0/1 forcing property* (2): if one probability term provides a null value (impossible event), then the aggregated result will be null; the same is valid for the value one (certain event). Other properties of the Bordley's formula are the external Bayesianity and the non convexity.

Weights selection

In order to apply the aggregation methods presented before, the weighting factors must be defined. Comunian et al. (8) review the methods which can be used to determine these weights and apply some methods to the solution of synthetic case studies. Here, three weighting schemes are proposed and tested.

The first weighting scheme considers the weights equal and fixed during the entire simulation process. If there are no elements which allow to prefer one training image rather than another, this is a reasonable choice which can be easily applied for the linear pooling formula: the aggregation then reduces to the arithmetic mean of the cpdf extracted from each training image. However, this can be more complicated for the Bordley's formula, where there is not the restriction $S_w = 1$. For this formula, one of the most common approaches is to consider unitary weights, which corresponds to the assumption of conditional independence among sources of information (15; 18). Different weight choices are explored in the following test cases. In

the case of the Bordley's formula, important differences are expected when $S_w > 1$ or when $S_w < 1$; these two situation correspond to negative or positive values of the exponent of the term $\mathbb{O}_0$.

The second weighting scheme considers that the weights should vary during the simulation process. It is known that the weights should not be constant (18), because they may depend on the data configuration. However, the complete cpdf (eq. (3.3)) being not accessible, one is left with having to provide a reasonable approximation. A possible scheme, when the multi-grids are employed, is to assign a different weight for each multi-grid level, giving a greater confidence (with larger weights) to the information which comes from the training images for the simulation of the first multi-grid levels; then, when the main structures are formed and for the simulation of the final multi-grid levels, give more weight to the prior terms in order to try to get closer to the target facies proportions. With this weighting scheme, we performed some tests which are not reported here because they did not shown significant improvements in the results obtained.

The last weighting scheme proposed here is novel: weights vary at each step of the simulation process. They are computed at each point u of the simulation path, and are proportional to the size of the data events which are used to compute the cpdf from the different training images. The weights are selected proportional to the size of the data event, but of course their sum must respect the condition imposed in the case of the linear pool ($S_w = 1$) or for the Bordley's formula (S_w should be equal to the sum of the "default" values of the weights).

3.3.2 Sequential 2D simulations with conditioning data

The second approach proposed here is based on sequential 2D multiple-point simulations and on the use of conditioning data (in brief approach **s2Dcd**, sequential 2D conditioning data) . The procedure, depicted in figure 3.4, is the following. At the first step, a sequence for the sequential simulation of the 2D surfaces is defined. Then, for each simulation step i , a 2D MPS simulation is performed along a surface S_i using the training image that describes the heterogeneity along the direction of S_i ; the conditioning data used are the points contained in $S_i \cap (\bigcup_{j < i} S_j)$ and the points contained in $S_i \cap D$, where D is the set of hard conditioning data provided (if available) at the beginning of the simulation.

These steps are repeated until all the domain is simulated.

The definition of the sequence of the simulation surfaces is crucial for this approach, that relies on the information provided by the conditioning data, and must therefore try to include, at each simulation step, as many conditioning data as possible. At the same time, the geometrical relations among the dimensions of the simulation domain must be considered in the choice of the simulation sequence. If the dimensions of the simulation do-

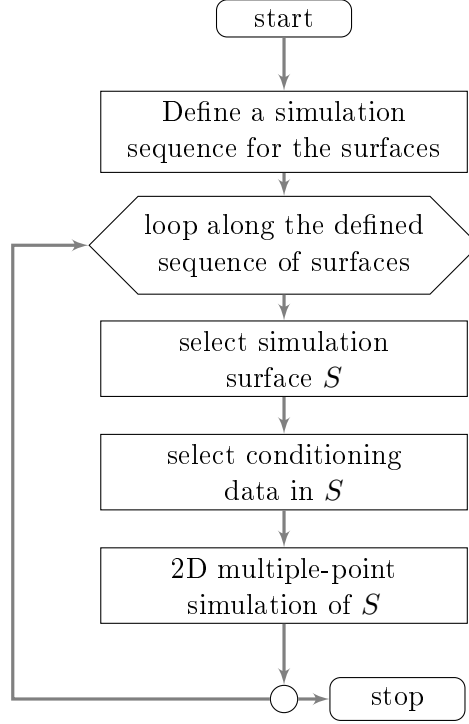


Figure 3.4: The procedure to simulate a 3D volume using sequential bi-dimensional MPS simulations and conditioning data (approach **s2Dcd**).

main are comparable, then a possible simulation order for the surfaces S_i is illustrated in fig. 3.5(a). It refers to a square domain where there is no conditioning data at the first simulation step and where only two training images (one perpendicular to the axis x and the other one perpendicular to the axis y) are available. When conditioning data are available or when the dimensions of the simulation domain differ significantly, the order of simulation of the 2D surfaces should be customized according to the location of the conditioning data and to the shape of the simulation domain. As a general rule, the simulation order of the 2D surfaces should be selected in order to gradually fill up the simulated domain and to include as many conditioning data as possible. However, different choices can be done according to the correlation length of the structures contained in the training images and/or to location of the conditioning data. An example of simulation sequence for a simulation domain with $x \gg y$ is illustrated in fig. 3.5(b).

3.3.3 Lists merging

Another method to simulate in 3D using 2D training images, illustrated in fig. 3.6, relies on the use of the lists. Nevertheless, the method described

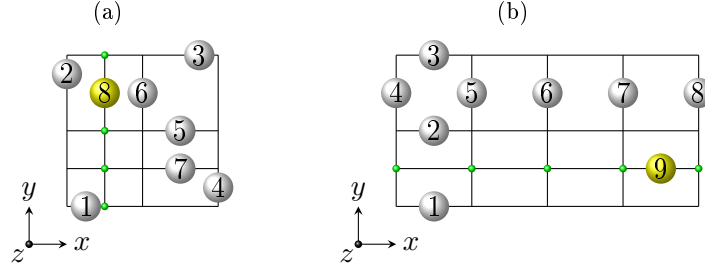


Figure 3.5: The top view on the plane xy of two possible sequences of bi-dimensional surfaces S_i for the **s2Dcd** method when (a) the dimensions x and y of the domain are comparable or (b) when $x \gg y$. Only the first eight(a)/nine(b) surfaces are shown. The green dots represent the vertical columns of conditioning data considered for the simulation of the surfaces S_8 (a) and S_9 (b).

here for the case of the *impala* algorithm can be adapted to other MPS implementations.

Let us illustrate this method with an example (fig. 3.7). Consider two 2D binary training images T_{xz} and T_{xy} oriented along the planes xz and xy and two corresponding search templates τ_{xy} and τ_{xz} of size 3×3 pixels. For each training image, the corresponding lists $\mathcal{L}_{xz}$ and $\mathcal{L}_{xy}$ are computed using the *impala* algorithm. Each element of the lists, for example $L_m \in \mathcal{L}_{xz}$, is composed by a vector which contains the configuration of a data event d_m and a vector c_m , which contains counters for the occurrences of each facies for the given d_m . Figure 3.7 illustrates some situations where the data events d_m and d_n , belonging to the lists elements $L_m, L_n \in \mathcal{L}_{xz}$, can be combined with data events d_i, d_j and d_k , belonging to the elements of the list $\mathcal{L}_{xy}$. For example, the data events d_m and d_i are compatible, and a list element of the merged list $\mathcal{L}_{merged}$ can be created. This last list element contains the data event $d_{m,i}$, obtained by merging d_m and d_i . The compatibility between data events of the list $\mathcal{L}_{xz}$ and $\mathcal{L}_{xy}$ is checked by looking at the facies at the positions α and β . With the same procedure, the compatibility between all the data events of the two lists is verified and, if it is the case, an element of $\mathcal{L}_{merged}$ is created. Of course, it can happen that a data event of one list (for example $d_n \in \mathcal{L}_{xz}$) is compatible with more than one data event of the other list (for example $d_j, d_k \in \mathcal{L}_{xy}$). In this case, one element of $\mathcal{L}_{merged}$ is created for each compatible combination. Once the compatibility among the data events belonging to the elements of the two lists is verified and a merged data event is created, another step is required to fill up the element of $\mathcal{L}_{merged}$ – the computation of the vector of counters. These counters can be computed as a function of the counters of the two combined list elements. Once again, this

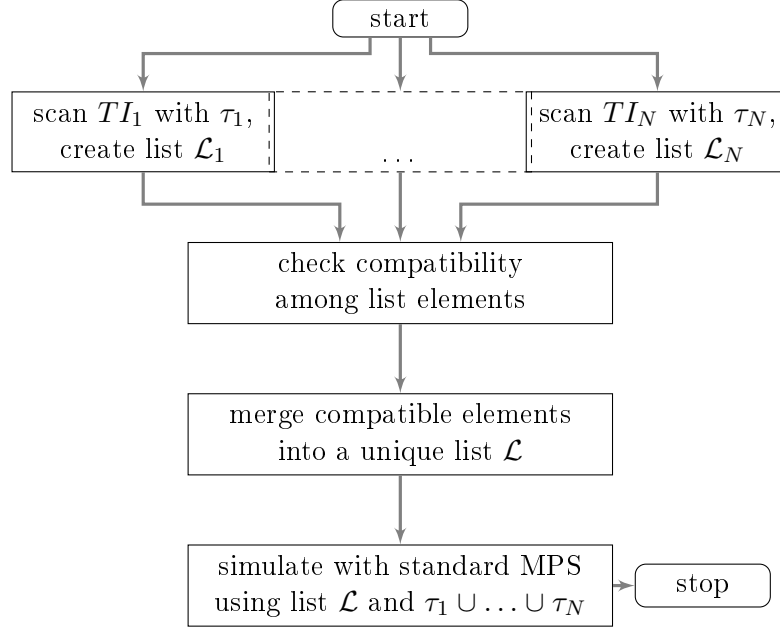


Figure 3.6: The procedure to perform 3D multiple-point simulation aggregating the probabilities from 2D training images. The dashed blocks represent operations which are performed for each considered training image.

problem is ill posed because there is no way to be sure that the combination is correct, unless the 3D information is available. Here four approaches are proposed and tested. For the first one, the counters are computed using the minimum value of the counters available for the compatible data events in each training image. For the other three approaches the mean value, the maximum value or the sum of the counters is used.

The merged list computed in this way can be used to simulate the 3D domain. At the same time, the size of the merged list can be used as an indicator of the compatibility of the training images considered. Indeed, in some situation it can be difficult to determine if the models provided by two 2D training images are coherent and can depict different but compatible aspect of the same 3D features which must be reproduced. This method offers a quantitative approach to deal with this issue.

3.4 Comparison criteria

For the test considered when a 3D reference TI is available, direct comparisons among this 3D training image, the simulation obtained with the methods proposed here and the standard MPS simulation methods which

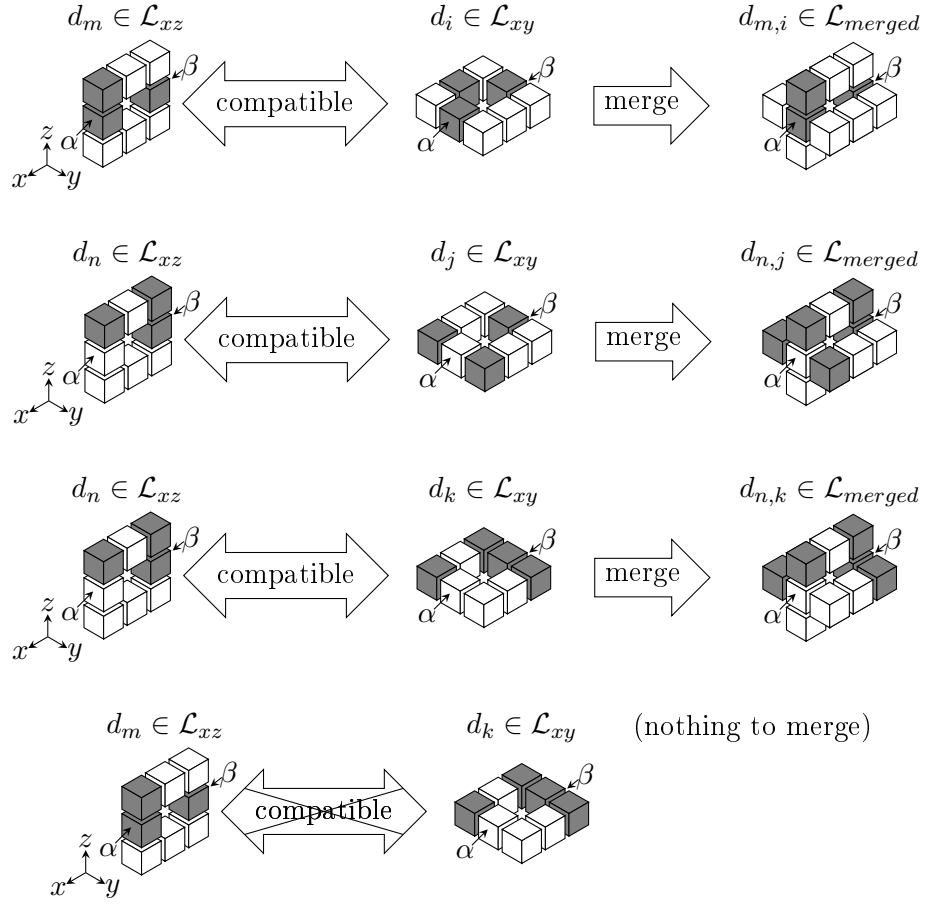


Figure 3.7: The data events of the couples d_m and d_i , d_n and d_j , d_n and d_k belonging to the lists $\mathcal{L}_{xz}$ and $\mathcal{L}_{xy}$ are compatible. Therefore, they can be merged into the data events $d_{m,i}$, $d_{n,j}$ and $d_{n,k}$. The data events d_m and d_k are not compatible (different facies in the position α), and they cannot be merged.

make use of this 3D training image are performed. When a 3D reference is not available, the features of the results are compared with the ones of the considered training images. The comparisons between the simulations obtained with the proposed methods and the references are based on visual comparisons, the reproduction of facies proportions, the computing time and the following criteria and parameters:

number of geobodies the number of connected components (clusters) composed of the same facies code. Comparing this value computed for the reference image and the one obtained for the simulations allows to check the continuity of the structures which are simulated. However, as we show in the following, this parameter should be considered as indicative only.

Moreover, while the interpretation of this parameter is intuitive for the simulations of binary images, it is less intuitive for complex structures and when the number of simulated facies is greater than two.

geobodies area (volume) the area (volume) of the connected components of the same facies are computed counting the number of pixels (voxels) contained in each component. For each simulation, the statistical distribution of the areas (volumes) is compared with one of the reference image.

connectivity functions Following the definition of connectivity proposed by Strauffer and Amnon (29) and Allard (1) for a binary medium Ω , two point x and y are connected when there exists at least one path completely contained in the permeable component $\Omega_{|perm.}$ of the medium which joins the two points; in this case the notation $x \leftrightarrow y$ is used. With this notation, a connectivity function between two points at a distance h can be defined as:

$$g(h) = \mathbb{P}(x \leftrightarrow x + h \mid x \in \Omega_{|perm.}). \quad (3.7)$$

In this work we compare the connectivity functions computed for the reference image and for the simulations. In the cases considered the images can be considered isotropic and only the connectivity along the direction x is computed.

Equivalent hydraulic conductivity tensor The comparisons among the morphological features of the geobodies can become difficult when the number of facies which is simulated is greater than two. In this case, we support the comparisons using another criterion, based on the computation of the equivalent hydraulic conductivity tensor $\mathbf{K}$. $\mathbf{K}$ is computed using typical values of hydraulic conductivity of the geological facies simulated, performing stationary flux simulations with linear boundary conditions along the three different axis directions

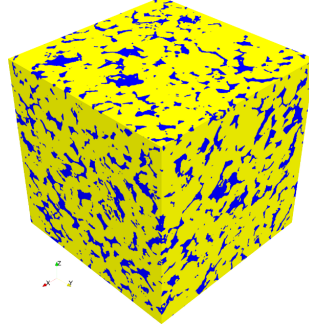


Figure 3.8: The micro CT of a sample of Berea sandstone used as 3D reference. In the image, of size $400 \times 400 \times 400$ voxels, the blue represents the vugs and the yellow the matrix.

(27; 26; 25). To compare the equivalent permeability of the reference $\mathbf{K}_{ref}$ with the one of the simulation $\mathbf{K}_{sim}$, a relative error ε based on the Frobenius norm $\| \cdot \|$ is computed (25):

$$\varepsilon = \frac{\|\mathbf{K}_{ref} - \mathbf{K}_{sim}\|}{\|\mathbf{K}_{ref}\|}. \quad (3.8)$$

When the computational effort allows it, all the parameters above are computed for a minimum of 10 simulations to a maximum of 100 simulations.

3.5 Micro-computed tomography image

The first realistic benchmark used to test the proposed methods is the result of a micro-computed tomography (micro CT) on a sample of Berea sandstone (23; 24). It is a 3D pore-space binary image discretized in $400 \times 400 \times 400$ voxels (3.8). It was selected for a number of reasons. Firstly, because the 3D results of the simulation methods proposed can be directly compared with the actual 3D reference. Secondly, because it is sufficiently stationary to be considered as a training image. Finally, it shows that the methods proposed here can be applied to a range of spatial scales, pore-scale included.

The data set extracted from the reference image is composed of: a 3D training image of dimension $200 \times 200 \times 200$, smaller than the original one in order to allow a faster simulation for the methods which require a full 3D training image; a subset of the original image of dimension $100 \times 100 \times 100$, that is of the same size as the 3D simulations obtained with the methods under examination, which can be used for direct comparisons with the simulations; three bi-dimensional training images obtained as slices 400×400 of the reference image along the planes xy , xz and yz .

3.5.1 Preliminary bi-dimensional calibration

General guidelines about how to set up the parameters of a multiple-point simulation exist (20). However, these parameters strongly depend on the considered training image and should be adapted to the specific case. To look directly on 3D simulations with a trial and error procedure for a suitable parameters set requires higher computation and interpretation efforts than on 2D simulations. For this reason, the procedure used to select the number of multi-grid levels and the size of the search template was performed only on 2D cases. This preliminary step is performed for all the case studies considered in this work.

The results, in the case of the micro CT image, are illustrated in appendix E.

3.5.2 Choose the weights for the Bordley's formula

The application of the two proposed aggregation formulae requires the definition of weighting factors. For the case of the micro CT image, there is no element which allow to prefer one of the three 2D training images extracted as sections along the planes xy , xz and yz to another. The 3D image can be considered isotropic. As a consequence, the same weight is associated to each considered training image. While in the case of the linear pooling formula this reduces the aggregation process to a simple arithmetic mean of the probability terms, in the case of the Bordley's formula more freedom is left on the choice of the weights, and on their sum S_w in particular. In our case, having access to the properties of the target 3D results, we performed a trial and error procedure in order to determine which set of equal weights w_i provides the result closer to the reference image.

The **paBor** method is applied for eleven different values of the weights $w_{1,2}$ ranging from 0.1 to 2.0; two training images parallel to the xz and to the xy are used. For each parameters set, 100 realizations are obtained.

3.5.3 Methods comparison

The number of multi-grid levels and the size of the data template used for the following tests are defined by the trial and error procedure discussed in section 3.5.1. The training images used in the following are illustrated in fig. 3.9.

For the method based on probability aggregation, both the linear formula and the Bordley's formula are tested.

The first one is tested with two training images parallel to the planes xy and xz and with weights $w_{1,2} = 0.5$ (test (a) tab. 3.1) or proportional to the size of the data event and such that $S_w = 1$ (test (b) tab. 3.1).

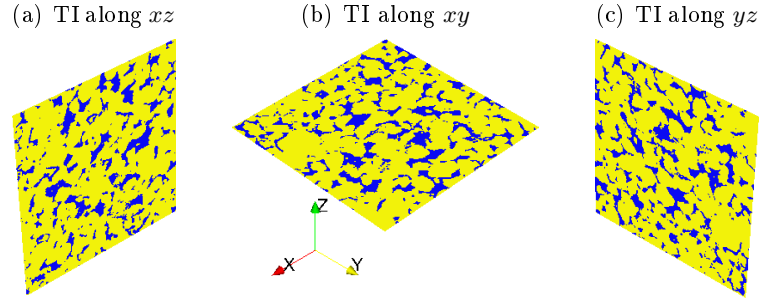


Figure 3.9: Training images used along the different directions for the tests on the micro CT image.

The Bordley's aggregation formula is tested using two training images parallel to the planes xy and xz (test (c) and (d) tab. 3.1) and three training images parallel to the planes xy , xz and yz (test (e) and (f) tab. 3.1). The weights used are $w_{1,2} = 1.0$, $w_{1,2,3} = 1.0$ and weights proportional to the size of the data event with $S_w = \sum_i w_i$ and $w_i = 1$ for each i ; the considered prior probability terms are the porosity and $1 - \text{porosity}$.

The approach **s2Dcd** is tested using two training images parallel to the planes xy and xz (test (g) tab. 3.1) or three training images parallel to the planes xy , xz and yz (test (h) tab. 3.1).

The results obtained with the method based on the merging of lists are not shown here because their quality is worse than the one of the results obtained with the other methods. However, an application to check the compatibility among training images based on this method is illustrated in section 3.9.

The availability of a 3D training image allows to perform MPS simulations which use it in order to compare the results with the ones obtained with the methods which use 2D training images only. Here the simulations with a 3D training image (tests (i), (j) and (k) tab. 3.1) are performed using diverse data templates: (k) a complete $7 \times 7 \times 7$ template; (i) a template obtained by merging two bi-dimensional templates of dimensions $7 \times 7 \times 1$ and $7 \times 1 \times 7$; (j) a template obtained by merging three data templates $7 \times 7 \times 1$, $7 \times 1 \times 7$ and $1 \times 7 \times 7$ (see for example fig. 3.3). Incomplete 3D search templates are considered in tests (i) and (j) in order to compare the result of the standard MPS technique and the ones tested in this work using search templates of comparable size.

The results obtained with the set up described in this section are illustrated in tab. 3.1, fig. 3.16 and fig. 3.17 and discussed in section 3.8.2.

test	method	method details	time [sec]	# TI	# geob.	porosity [%]
(a)	paLin	$w_{1,2} = 0.5$	1299	2	2916	21.01
(b)	paLin	$w_{1,2} \propto d$	1333	2	2384	20.38
(c)	paBor	$w_{1,2} = 1$	1152	2	592	20.74
(d)	paBor	$w_{1,2} \propto d$	1172	2	519	25.72
(e)	paBor	$w_{1,2,3} = 1$	1689	3	269	23.44
(f)	paBor	$w_{1,2,3} \propto d$	1999	3	227	29.01
(g)	s2Dcd	2 TI	459	2	226	22.10
(h)	s2Dcd	3 TI	434	3	212	19.58
(i)	3D	$\tau_{xy} \cup \tau_{xz}$	72500	1	775	20.86
(j)	3D	$\tau_{xy} \cup \tau_{xz} \cup \tau_{yz}$	131234	1	395	21.33
(k)	3D	$7 \times 7 \times 7$	493414	1	79	22.40
ref					44	20.13

Table 3.1: Comparison of different methods for the simulation of the micro CT image of a Berea sandstone. The cases where the weights are selected at each simulation step proportional to the data event size are indicated with “ $\propto d$ ”.

3.6 Reproduction of sedimentary structures

The methods **paBor** and **s2Dcd** are tested using two training images which represents some typical sedimentary structures mapped in the Brussels Sands by Huysmans and Dassargues (13). In the training images, one can distinguish a matrix of clay-rich bottom-sets and distinct mud drapes (fig.3.10). For this case study, the results are compared visually and from the point of view of the equivalent conductivity. An equivalent conductivity tensor is computed in 2D for one of the training images (the one oriented along the plane xz , fig. 3.10 (a)). Another equivalent conductivity tensor is computed in 3D for each one of the ten simulations performed for the method **paBor** and for the ten simulations with the method **s2Dcd**.

To compute the relative error ε on K we use as K_{ref} the 2×2 tensor computed for the training image, and as K_{sim} the four elements of the 3×3 tensors computed for the simulations which correspond to the four elements of K_{ref} . In order to observe a remarkable anisotropy in the equivalent conductivity tensor, for the flow simulations a contrast between the two facies of 100 is used in place of the real values of conductivity measured in the field (14).

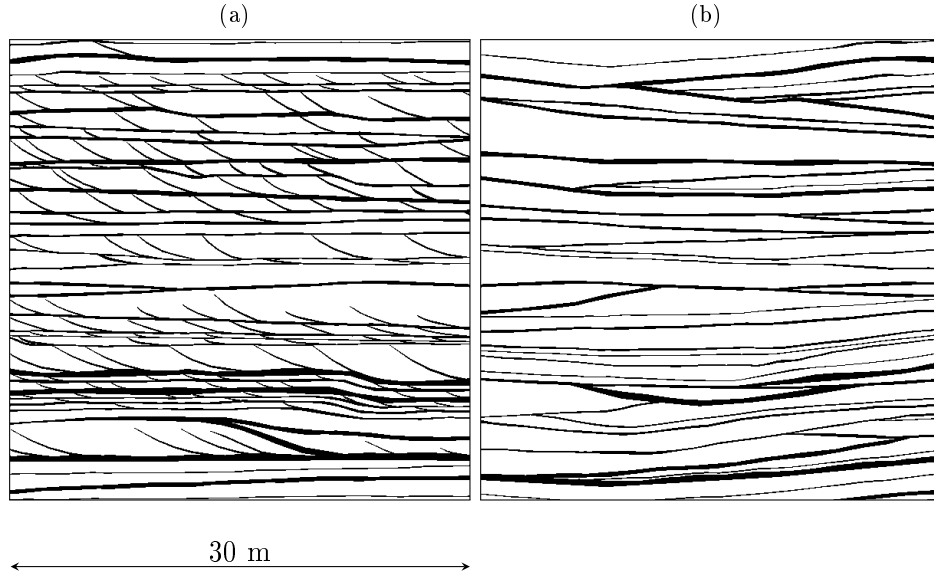


Figure 3.10: The training images considered for the Brussels Sands sedimentary environment (a) along the plane xz and (b) along the plane yz . Clay-rich bottom-set are in white, while mud drapes are in black.

3.7 The Herten aquifer analog

The **s2Dcd** and **paBor** methods are tested on the reproduction of the hydrofacies distribution of a high-resolution fluvio-glacial aquifer analog also. The Herten aquifer analog is described and studied in detail by Bayer et al. (3). The part of the data set used here consists of six parallel outcrops (oriented along the plane xz in the reference system adopted here) where ten hydrofacies were recognized and mapped during the excavation of a quarry. A recent study (9) allowed to obtain a detailed and realistic 3D reconstruction of the entire volume with a hierarchical geostatistical simulation approach; the results of that study are used here as reference.

The data set is used as conditioning data and training images as well. In our test, we used one of the profiles as a 2D training image for both the plane xz and the plane yz . This is a strong hypothesis which can be debatable, but for the application of the proposed methods we decided to rely only on field data, and none was available along the perpendicular direction.

For this data set, we test the method based on the aggregation of probabilities using the Bordley's formula and weighting factors $w_{1,2} = 1$ (fig. 3.19(c)). The method **s2Dcd** is tested with and without the use of the coordinate z as auxiliary variable (fig. 3.19(d) and 3.19(e)).

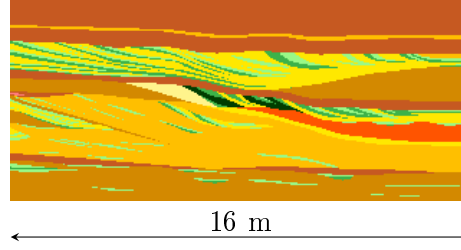


Figure 3.11: One of the outcrops of the Herten aquifer analog dataset. It is used as training image in both xz and yz direction for the simulation with the methods **paBor** and **s2Dcd**.

3.8 Results and discussion

3.8.1 Micro CT image and paBor weights

The tests performed with the method **paBor** on the micro CT image using different weighting factors suggest that the best set of weights that can be used in this case is $w_{1,2} = 1$. While the 3D representations (fig. 3.12), and the connectivity curves (fig. 3.13) show that results close to the reference can be obtained with $w_{1,2} = 0.8, 0.9$ or 1.0 , the vugs proportions (fig. 3.15) show that the best weights set in this case is $w_{1,2} = 1$.

Note that, as expected, the orientations along which the shapes contained in the reference images are reproduced in a better way are the orientations along which a training image is used (fig. 3.12). The information about the volume of the geobodies (fig. 3.14) did not provide a significant insight for the selection of the most appropriate set of weights.

The number of geobodies follows an opposite trend than the other criteria. The best results are obtained with the set of weights $w_{1,2} = 2$: a mean number on the 100 simulation of 318 geobodies, while 608 geobodies are observed in average for $w_{1,2} = 1$, with a target number in the reference image of 44 geobodies. This is related to the overestimation of the pore space obtained with $w_{1,2} = 2$ (fig. 3.15). Thus, the results which are closer to the reference image according to the majority of the comparison criteria are obtained with $w_{1,2} = 1$.

The issue raised by the comparison between the results for $w_{1,2} = 1$ and $w_{1,2} = 2$ shows that the best parameters set cannot be selected considering only one parameter (i.e. the number of geobodies); instead, a number of parameters which entails different aspects and descriptions of the results must be considered.

For the following comparisons with other methods, the **paBor** method is used with weights $w_i = 1$ for each i .

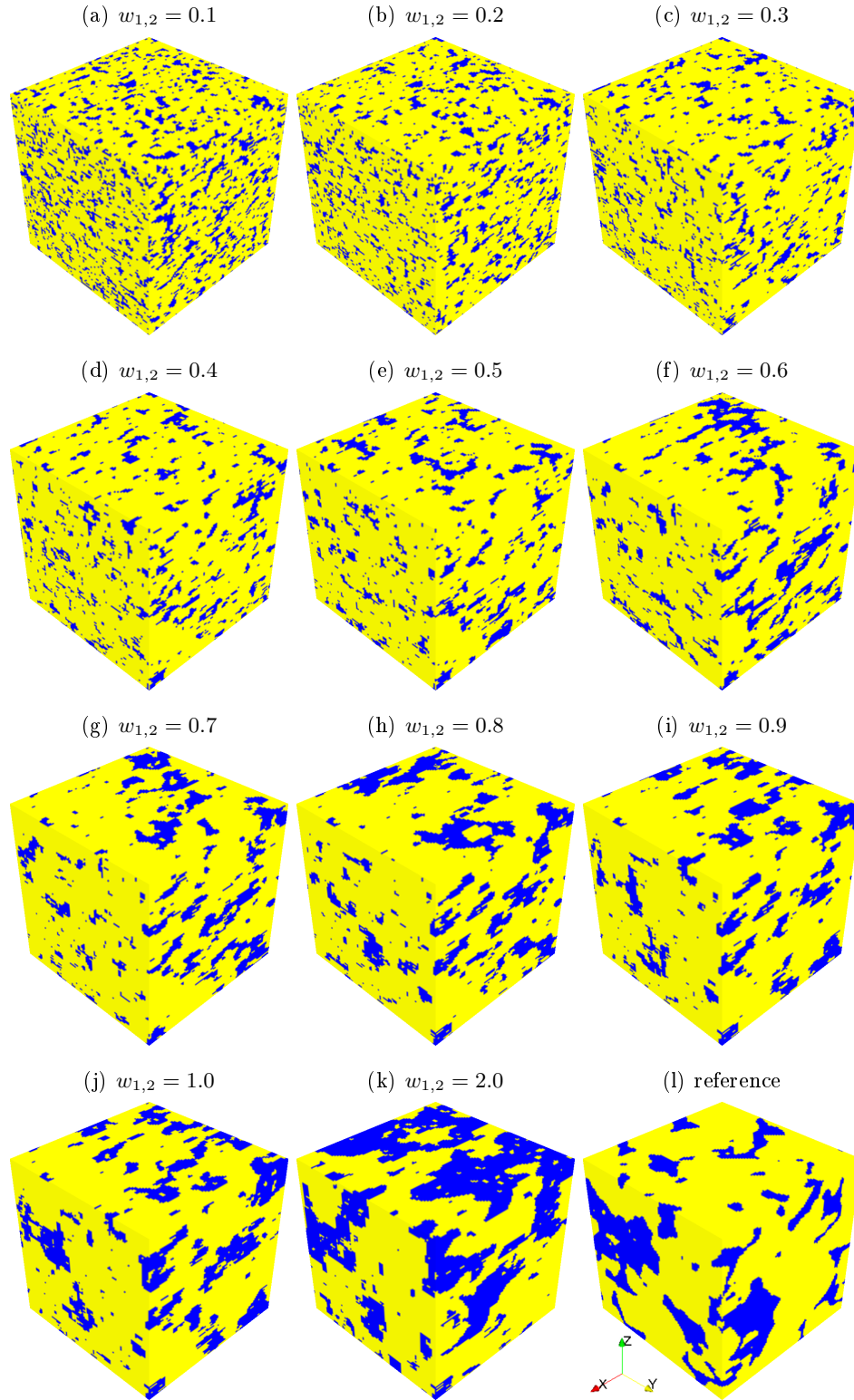


Figure 3.12: Simulations of the micro CT image obtained with the **paBor** method with different values of $w_{1,2}$. The two training images used are parallel to the planes xy and xz .

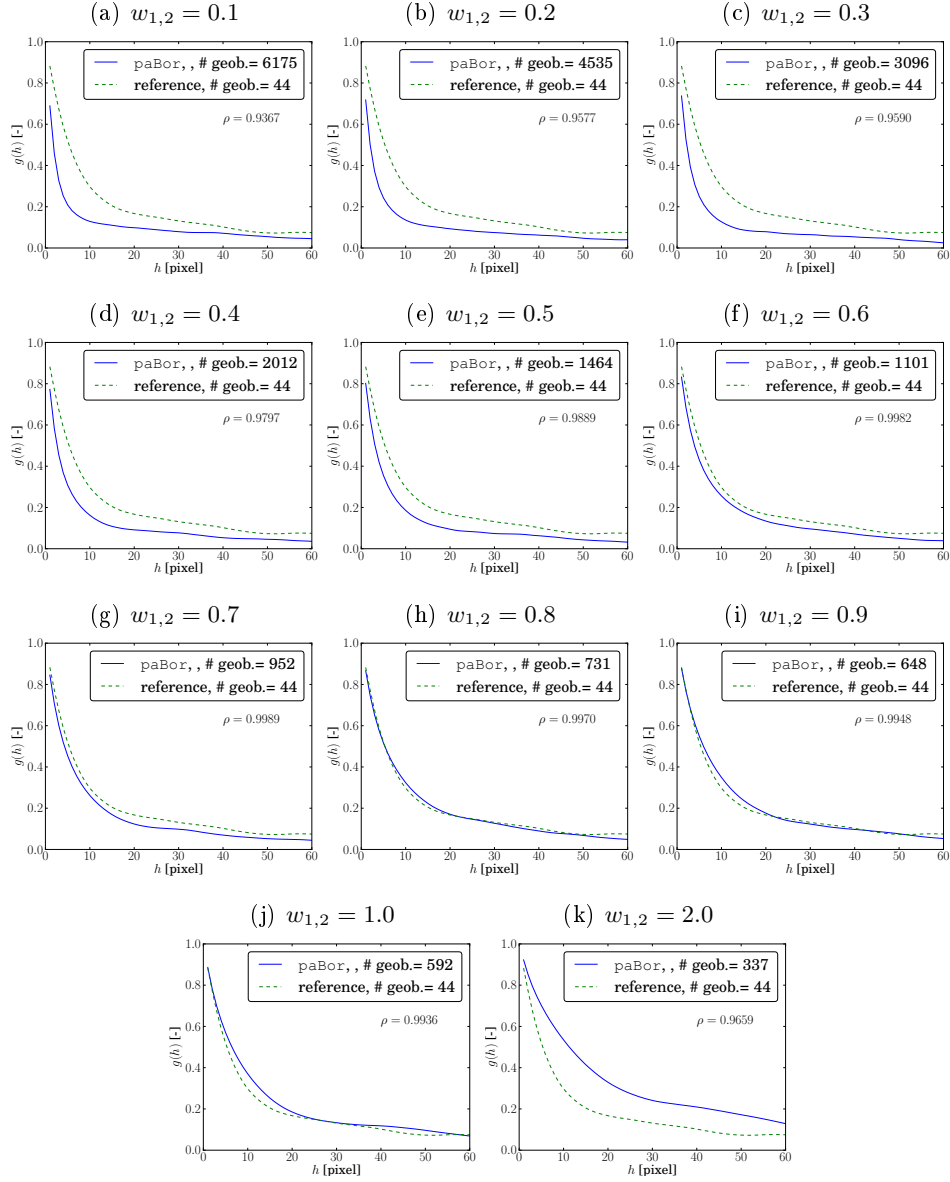


Figure 3.13: Connectivity functions computed along the coordinate x for the 3D simulations obtained with the **paBor** method and different values of $w_{1,2}$. The dashed line represents the reference; the correlation coefficient ρ between the reference and the computed curve is reported.

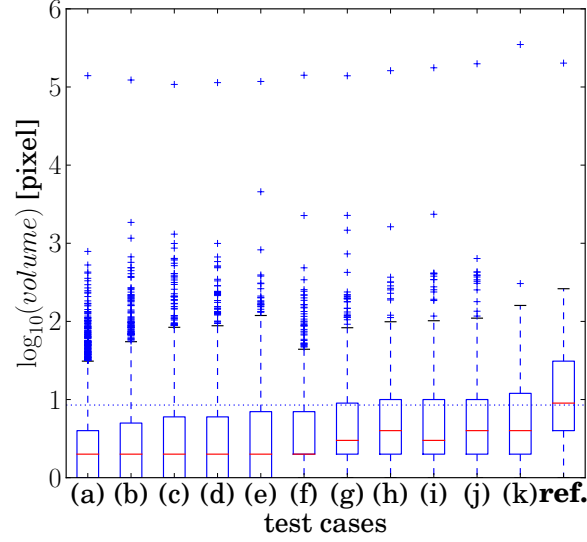


Figure 3.14: For one of the 100 simulations obtained for each value of $w_{1,2}$, the $\log_{10}$ of the volume of the vugs of the micro CT image. The simulation are obtained using two training images and the Bordley's aggregation formula, the weight value varies from 0.1 (a) to 2.0 (k). The dashed line represents the median of the volume of the geobodies contained in the reference image.

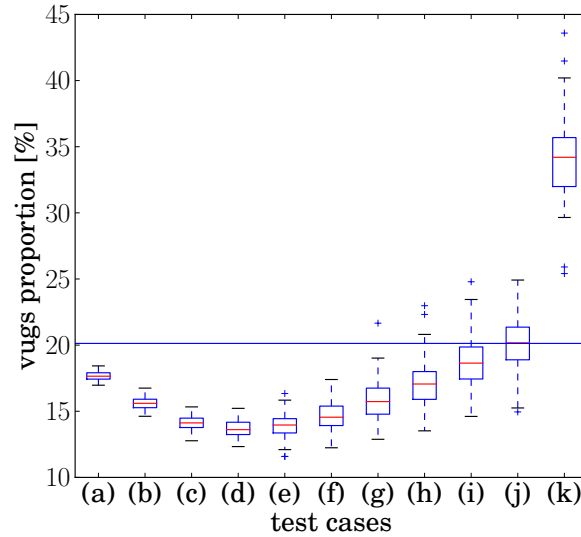


Figure 3.15: Proportions of vugs for micro CT image for different values of the weights $w_{1,2}$ and the method **paBor**. The horizontal line represents the porosity of the reference image.

3.8.2 Methods comparison on the micro CT image

The method which better reproduces the features of the reference training image is the **s2Dcd**. The visual inspection of the simulations (fig. 3.16), the connectivity functions (fig. 3.17) and the data of table 3.1 provide a strong support for this. Note that the number of geobodies obtained with this method is closer to the one of the reference training image than the one obtained by the standard MPS simulation approach but with search templates reduced to a merge of two or three 2D search templates oriented along different orientations (tab. 3.1, methods (i) and (j)). Moreover, the computational effort of the method **s2Dcd** is more than two orders of magnitude smaller than for methods (i) and (j). In general, this is true for all the aggregation methods tested on this case study: the simulation time required by the methods **paLin**, **paBor** and **s2Dcd** is from two to three orders of magnitude smaller than the one required by a standard multiple-point simulation with a complete 3D training image.

The comparisons among the linear aggregation method (**paLin**, tests (a) and (b)) and the Bordley's formula (**paBor**, test (c),(d),(e),(f)) show that with the second method it is possible to obtain results which are more satisfactory in terms of visual aspect (fig. 3.16), connectivity function (fig. 3.17) and number of geobodies (tab. 3.1). Therefore in this situation the Bordley's aggregation formula is more suitable than the linear formula.

Of course, the results obtained using three training images are more realistic than the ones obtained using only two training images (fig. 3.16, tab. 3.1). Moreover, in the case of the **s2Dcd**, using three training images does not entails a greater global computational effort as it is for the methods based on the aggregation of probability. Indeed, even if more simulation steps are required for the **s2Dcd** method when three training images are used, the simulation time decreases faster than in the case where only two training images are used, because of the faster increase of the number of conditioning data (fig. E.5). This can result in total computing times which are smaller when three training images are used (tab. 3.1).

Selecting the weights as proportional to the size of the data event did not improved noticeably the results. When this option is used (tab. 3.1 tests (b), (d) and (f)), the number of geobodies is slightly closer to the target number of geobodies, but at the same time the proportions get farther from the target proportions. The observation concerning the set of weights $w_{1,2} = 2$ made on the previous section support this hypothesis.

Some tests are performed using the method based on lists merging for compatible data events. They are not reported because the quality of the simulations was not comparable with the simulations obtained with the other methods. Another drawback of this method is that when there is a high compatibility among the lists which are merged, the size of the merged list

becomes unacceptable.

This problem can be faced in different ways. One possibility is to increase the size of the search templates; this entails a reduction of the number of the compatible data events. Another possibility is to develop some criteria to reduce the size of the lists in such a way that the final result is not sensibly altered and the information content is preserved.

An additional issue related to this method is about how to compute the occurrences counters in the merged list as a function of the counters of the departure lists.

Despite its drawbacks, this method reveals its usefulness when it is required to check the compatibility among 2D training images which are selected to represent different aspect of a 3D training image. Appendix 3.9 is devoted to this last aspect.

3.8.3 Brussels sand sedimentary structures

With the method **s2Dcd** it is possible to obtain credible 3D simulations of the sedimentary structures depicted along two different sections by the training images of fig. 3.10. These simulations contain less noise than the ones obtained with the method **paBor** (fig. 3.18). Note that the slices of fig. 3.18 are not selected looking for the best or the worst situation, but trying to depict an equilibrated outline.

The better simulation quality that is observed for the method **s2Dcd** in fig. 3.18 is confirmed by the relative error ε (fig. 3.19). This last is computed comparing the equivalent conductivity tensor of the training image of fig. 3.10(a) and the concerning components of the tensors computed for the 3D simulations. The cluster of simulations performed with the method **s2Dcd** has a lower relative error than the corresponding cluster of simulations obtained with the **paBor** method.

3.8.4 The Herten aquifer analog

The results obtained for the Herten aquifer analog are compared visually (fig. 3.20). The ones obtained with the **paBor** method contain a noisy component which is less evident in the simulations obtained with the other methods. With the method **s2Dcd** the results obtained considering or not the coordinate z as auxiliary variable (fig. 3.20 (e) and (f) respectively) are comparable; they allow to obtain features which are very close to the ones of the reference image (fig. 3.20 (b)).

To compare the features of the simulated image in this case study using the same criteria as used for the micro CT images can be tedious. For example, the connectivity function can be computed for each of the ten hydrofacies used in the simulations, and the same for the number of geobodies and the proportions, but these results cannot be compared as easily as for

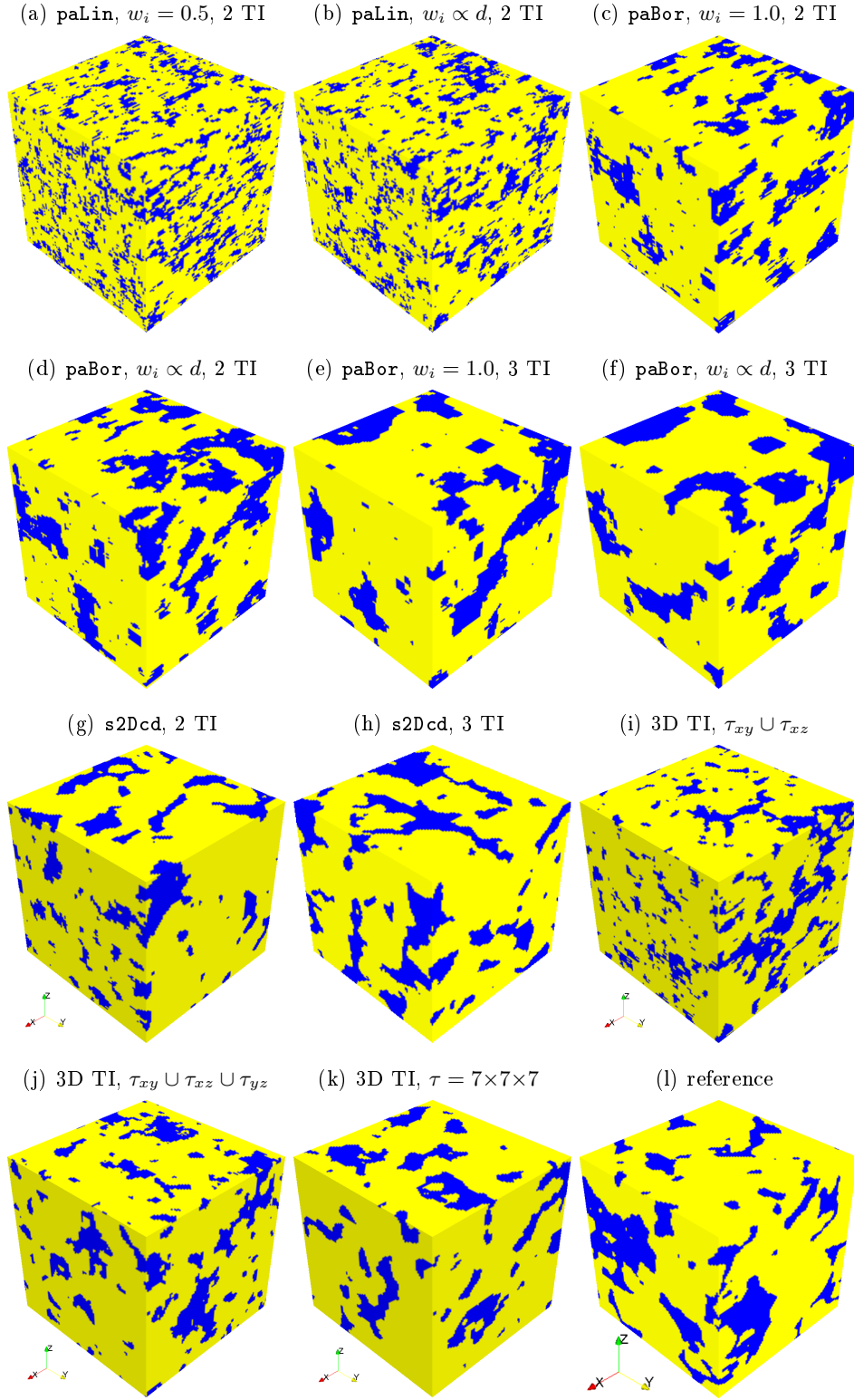


Figure 3.16: Simulations obtained with different methods for the micro CT image. When two 2D TI are used, they are oriented along the planes xy and xz . The methods which adopt weights proportional to the size of the data event and with $S_w = \sum_i w_i$ with $w_i = 1$ are denoted by “ $\propto d$ ”. The figure (l) is considered as reference and it is used as TI for the simulations (i), (j) and (k).

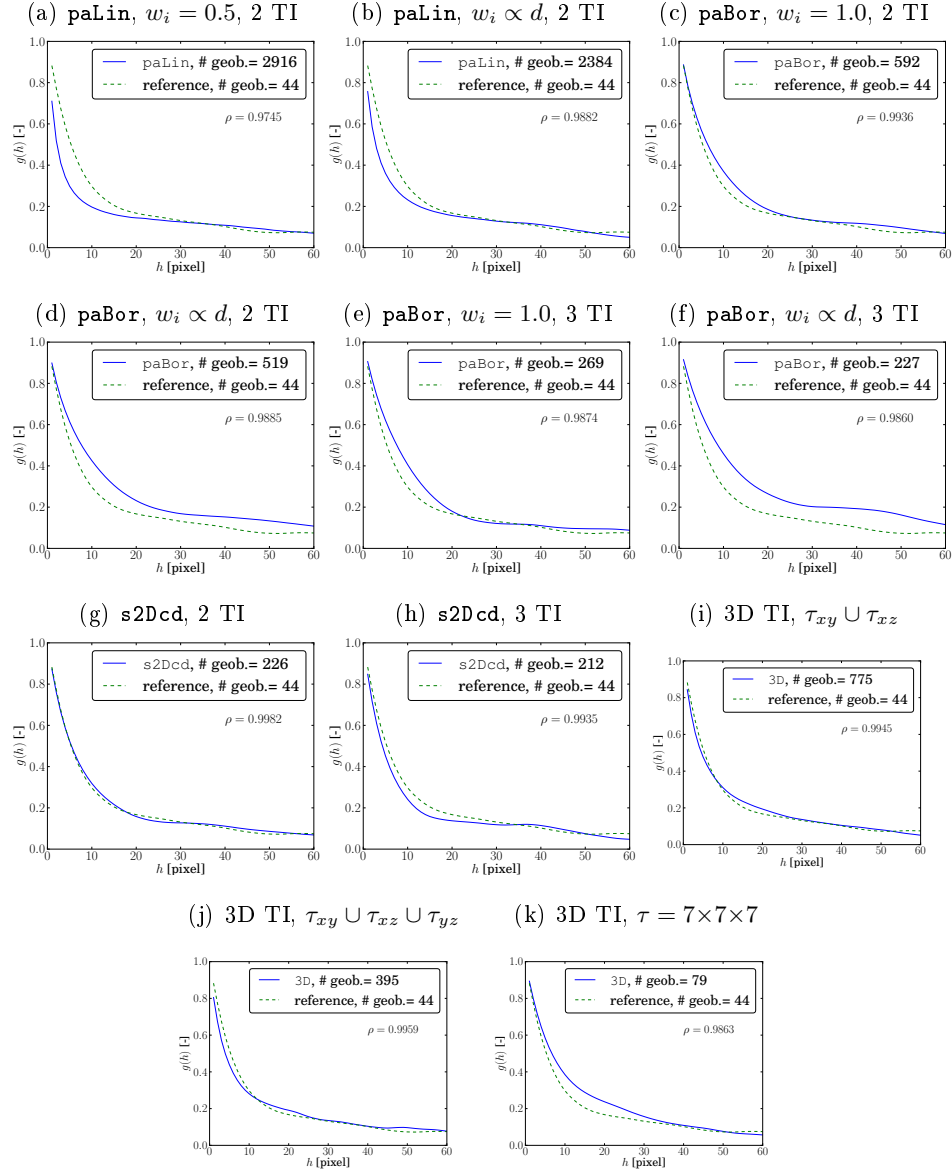


Figure 3.17: Connectivity functions along the x axis for the 3D tests performed for the micro CT image. The dashed curve is the function computed on the reference training image. The correlation coefficient between the curves and the number of geobodies is reported.

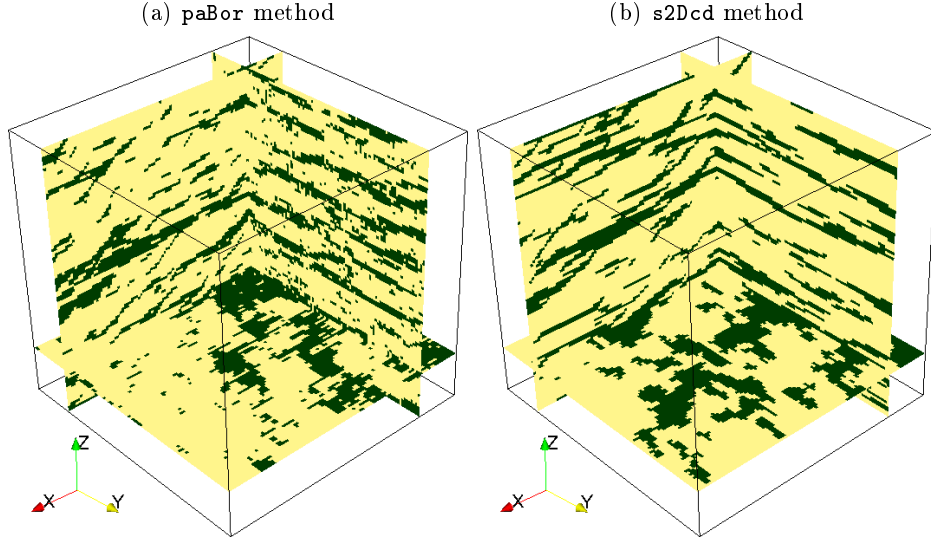


Figure 3.18: Simulations of the Brussels Sands environment (training images in fig. 3.10) obtained using (a) the **paBor** method and (b) the **s2Dcd** method.

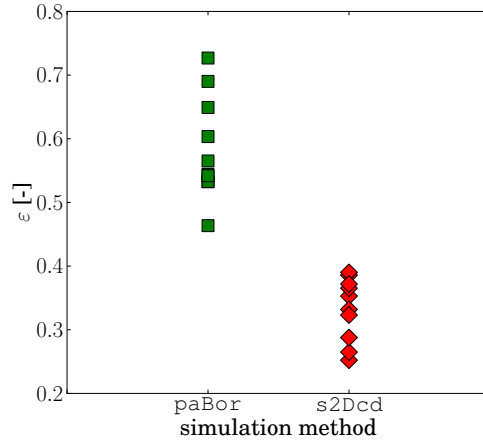


Figure 3.19: The relative error ε on $\mathbf{K}$ computed for 10 simulations performed with the **paBor** method and 10 simulations with the **s2Dcd** method. The reference conductivity tensor is computed on the training image of fig. 3.10(a).

data template	multi-grid levels	method	time [hours]
$7 \times 1 \times 7, 1 \times 7 \times 7$	3	paBor	2.1
$7 \times 1 \times 7, 1 \times 7 \times 7$	3	s2Dcd	3.5
$3 \times 1 \times 3, 1 \times 3 \times 3$	6	s2Dcd	0.2
$3 \times 3 \times 3$	6	3D	22.9
$7 \times 7 \times 7$	3	3D	>216.0

Table 3.2: Computation times for the different methods used to simulate the Herten aquifer analog.

a binary image. For this reason, another criterion is used. The comparisons are made by computing for each simulation an equivalent conductivity tensor $\mathbf{K}$ and its distance in terms of Frobenius norm to the equivalent conductivity of the reference image. This distance ε , normalized by the value of the Frobenius norm of the reference, provides a global parameter which allows a quantitative comparison of the characteristics of the simulation results in term of hydraulic parameters (fig. 3.21).

The distance from the reference image allows to distinguish a cluster of simulations performed with the **paBor** method, while the results obtained with the method **s2Dcd** with and without the auxiliary variable z have the same distance from the reference, five times smaller than the distance obtained with the **paBor** method. The distance of the methods **s2Dcd** is comparable with the one of the simulation method TProgs (transition probability/Markov chain, TP/MC, 6; 21). The method which allows to obtain the conductivity tensor closer to the reference is the one which uses it as 3D training image.

Concerning the computation times, the results are reported in table 3.2. The simulation times required for simulations with the methods **paBor** and **s2Dcd** are comparable (2.1 and 3.5 hours respectively, with two 2D search templates 7×7), while the time required for a simulation which uses a 3D training image is one order of magnitude greater using a search template $3 \times 3 \times 3$. If a template of size $7 \times 7 \times 7$ is used for the simulation with the 3D training image, then the computing times are more that two orders of magnitude greater than the times required by the **paBor** and **s2Dcd** methods. Note that the computation times reported here and in table 3.2 correspond to serial implementations of the codes for the MPS simulation. Of course, the parallel implementations of the MPS would have provided smaller computational times (28), but some of the simulation tools developed for this study are serial only, for the moment. Therefore, the simulations are performed using serial implementations, in order to have a fair comparison.

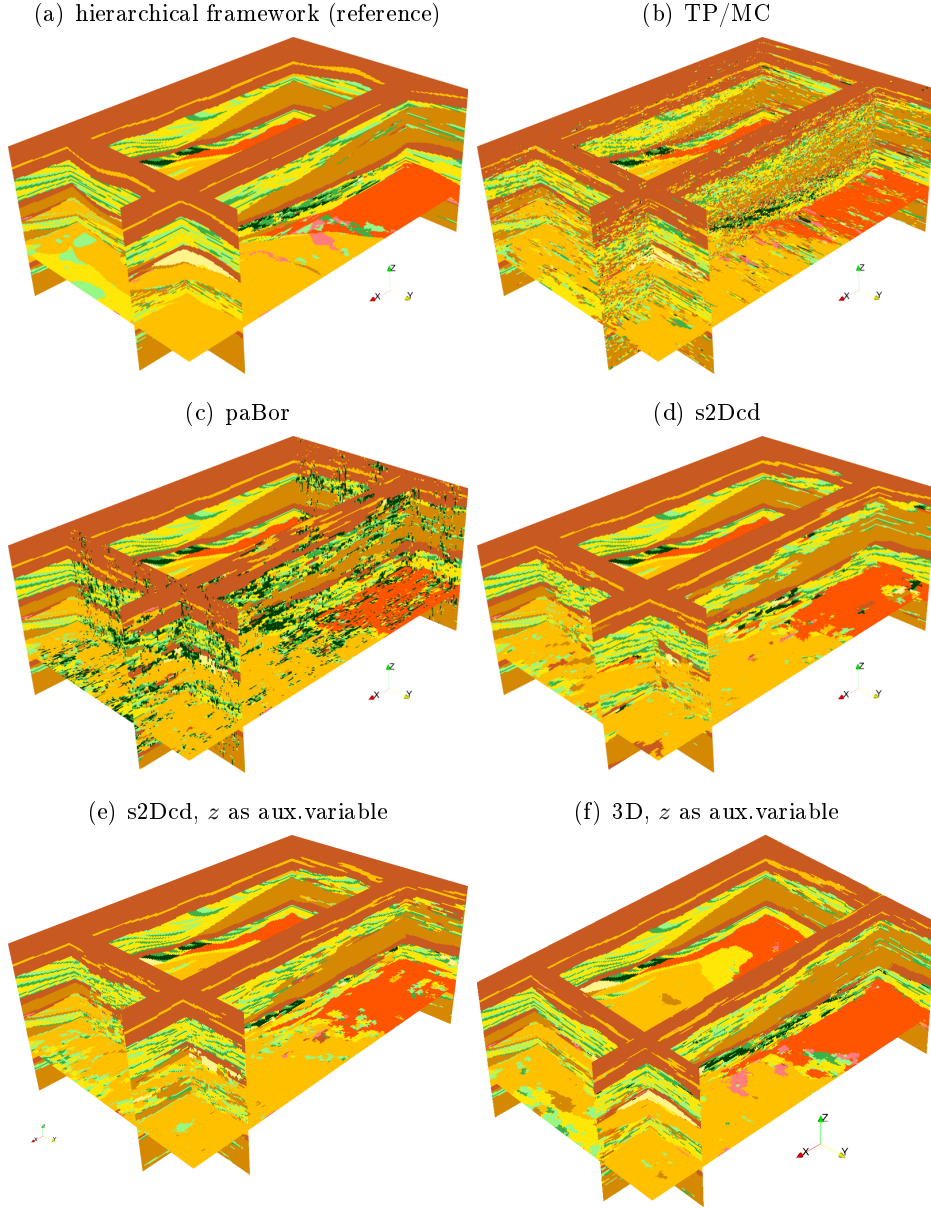


Figure 3.20: The simulations obtained with the different methods for the Herten aquifer analog: (a) hierarchical technique (considered as reference, 9); (b) Transition Probability/Markov Chain approach (21); (c) **paBor** approach; (d) **s2Dcd** with no auxiliary variables; (e) **s2Dcd** with the coordinate z as auxiliary variable; (f) with the use of (a) as 3D training image. The background section along xz contains one of the six sections used as conditioning data, while the other section is placed in between two parallel sections of conditioning data.

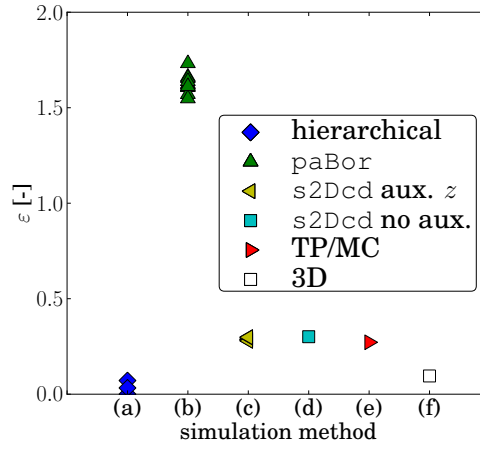


Figure 3.21: The relative error on K computed using as reference one realization of the Herten aquifer analog obtained with a hierarchical technique (9) for (a) two other realizations obtained with the hierarchical approach; (b) ten realization obtained with **paBor**; (c) ten realization obtained with **s2Dcd** and z as auxiliary variable; (d) one realization obtained with **s2Dcd** and no auxiliary variable; (e) one realization obtained with a TP/MC approach (21); (f) one realization obtained using the reference image as TI.

3.9 Check compatibility among training images

In the examples treated in this chapter, the 2D training images that are used are compatible with each other because they are either from the same 3D image, or observed in the field from a single geological environment. However, in practice the user is left with providing two or more independent training images, and it may happen that they are incompatible. Here we discuss how the technique to merge the lists proposed in section 3.3.3 can be used to test their compatibility.

Here the tests are performed on two training images. One training image is the training image containing channels (fig. 3.22(h)), introduced by Strebelle (31), considered as oriented along the xz plane. This image is associated to the list $\mathcal{L}_1$. The other training image, oriented along the yz plane, is chosen among the different images obtained by eroding or dilating the image used along the plane xz (fig. 3.22, list $\mathcal{L}_2$).

We then define an indicator of compatibility between the training images as the ratio between the size of the merged list $\mathcal{L}_{merged}$ and the size that the list would have in case all the data events were compatible, that is $\#\mathcal{L}_1 \times \#\mathcal{L}_2$ (fig. 3.23). The size of $\mathcal{L}_{merged}$ alone cannot provide a clear indication, because it depends on both the sizes of $\mathcal{L}_2$ and $\mathcal{L}_1$, and on the number of combined compatible data events.

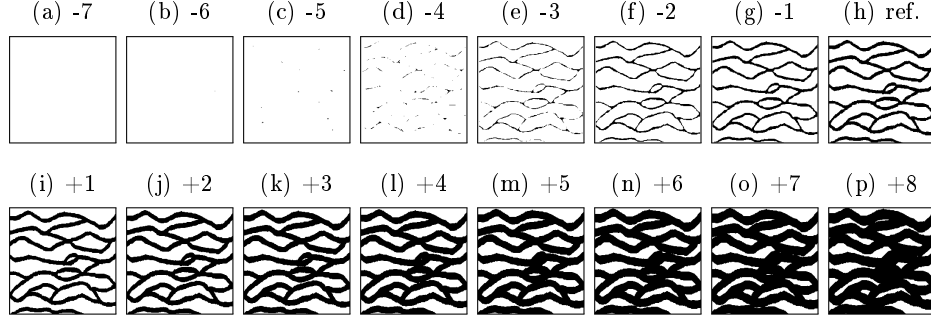


Figure 3.22: The reference training image used along the plane $xz(h)$. The images (a)-(g) and (i)-(p) are respectively obtained eroding (-) or dilating (+) the reference training image (h) by a number of successive iterations (shown in the subcaptions).

Figure 3.23 shows, for different multi-grid levels, the ratio

$$\#\mathcal{L}_{merged}/(\#\mathcal{L}_1 \times \#\mathcal{L}_2). \quad (3.9)$$

For each multi-grid level, a maximum of this ratio is observed in correspondence to the abscissa 0, that is for the value of erosion/dilation which corresponds to the departure image, when the training images used along the two directions are the same.

Even if this feature should be confirmed by other case studies, the results shown here indicate that this criterion could be used as comparison criterion to check the compatibility of 2D training images to be used for a joint simulation of a 3D block.

3.10 Conclusion

This work demonstrates that even without the use of a 3D training image, it is possible to obtain 3D simulations which reproduce the main characteristics of simulations obtained with a standard multiple-point statistics (standard in the sense that it uses a complete 3D training image).

One of the novel methods proposed here (**s2Dcd**), produces simulations which are closer to the reference 3D image than methods based on the aggregation of probability, according to most of the evaluation criteria considered here. The other novel method proposed, based on the merging of lists of compatible data events, does not allow to obtain realistic simulation results. However, it is based on a principle which is useful and easy to use for determining when two (or more) 2D training images can be considered compatible for the joint description of 3D geological structures.

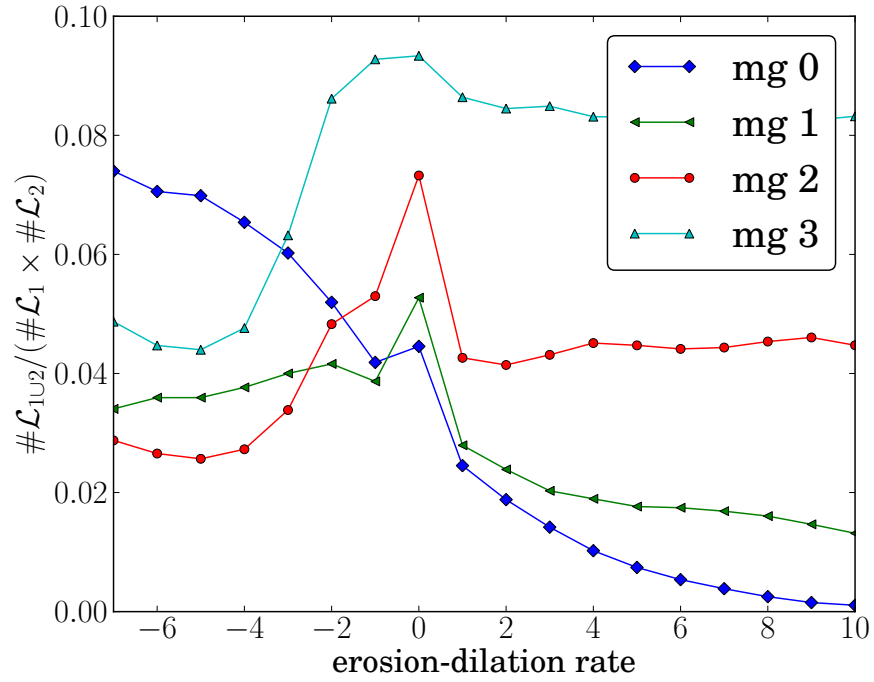


Figure 3.23: The ratio between the size of the merged list $\mathcal{L}_{1\cup 2}$ and the size of the merged list under the hypothesis that all the elements of the two merged lists $\mathcal{L}_1$ and $\mathcal{L}_2$ are compatible. The list $\mathcal{L}_1$ is computed from the training image of fig. 3.22(h), while the list $\mathcal{L}_2$ is obtained from the same image eroded or dilated (fig. 3.22). The size of the merged lists are computed for four different multi-grid levels.

An important aspect of the methods proposed here is their computational efficiency. Even if, of course, these methods cannot reach the same accuracy in the reproduction of complex geological patterns of MPS technique with a 3D training image, the computation time is from two to four orders of magnitude smaller than the methods using a full 3D image. Therefore, with these methods, it is possible to approach problems which are prohibitive for the standard MPS simulation in terms of CPU time required. For example, with the same computational effort, with these method it is possible to select search templates bigger than the ones allowed by the standard MPS . This computational efficiency allows to include MPS in Monte Carlo and uncertainty evaluation simulation frameworks, or in inverse problems which require a high number of realizations performed within a reasonable computation time.

Moreover, with the method **s2Dcd** it is possible to face simulation problems which would normally require huge RAM resources, because the simulation process is reduced to a sequence of 2D MPS simulations. This allowed for example to simulate domains of about 200 millions of voxels, using two 2D training images of sand lenses from a clay till outcrop in Denmark mapped by Kessler et al. (16).

Still, the **s2Dcd** method has some limitations. Even if the overall quality of the results presented here is good, some artifacts can be observed. For example, looking the 2D slices along their simulation sequence, a gradual deterioration of their quality can be noticed. While at the beginning the 2D simulations can be easily constrained by the available conditioning data, the constraints can become too strong when many conditioning data have been simulated. The strength of the constraints is due to the fact that the 2D simulations are obtained ignoring lateral correlation (the only information used are the conditioning data on the same simulation surface), and therefore some incoherent features can arise during the simulation procedure. These drawbacks can be reduced selecting a suitable simulation sequence of the surfaces, but further investigations are required in order to minimize them.

Bibliography

- [1] Allard, D. (1992). On the connectivity of two random set models: the truncated gaussian and the boolean. In *Geostatistics Troia '92*. Kluwer, Dordrecht, The Netherlands.
- [2] Allard, D., D'Or, D., and Froidevaux, R. (2010). Estimating and simulating spatial categorical data using an efficient maximum entropy approach. Accepted for publication in the European Journal of Soil Science.
- [3] Bayer, P., Huggenberger, P., Renard, P., and Comunian, A. (2010). Three-dimensional high resolution fluvio-glacial aquifer analog: Part 1: field study. Submitted to Journal of Hydrology.
- [4] Bordley, R. F. (1982). A multiplicative formula for aggregating probability assessments. *Management Science*.
- [5] Caers, J. (2006). A general algorithm for building 3D spatial laws from lower dimensional structural information. Technical report, Stanford Center for Reservoir Forecasting.
- [6] Carle, S. (1996). *A transition probability-based approach to geostatistical characterization of hydrostratigraphic architecture*. PhD thesis, Department of Land, Air and Water Resources, University of California Davis.
- [7] Chugunova, T. L. and Hu, L. Y. (2008). Multiple-point simulations constrained by continuous auxiliary data. *Mathematical Geosciences*, 40(2):133–146.
- [8] Comunian, A., Renard, P., and Allard, D. (2010a). Probability aggregation methods in earth sciences. in preparation for the submission to Mathematical Geosciences.
- [9] Comunian, A., Renard, P., Straubhaar, J., and Bayer, P. (2010b). Three-dimensional high resolution fluvio-glacial aquifer analog: Part 2: geostatistical modeling. (submitted to Journal of Hydrology).
- [10] de Marsily, G., Delay, F., Goncalves, J., Renard, P., Teles, V., and Violette, S. (2005). Dealing with spatial heterogeneity. *Hydrogeology Journal*, 13(1):161–183.
- [11] Guardiano, F. B. and Srivastava, R. M. (1993). Multivariate geostatistics: Beyond bivariate moments. In Soares, A., editor, *Geostatistics: Troia '92*, volume 1, pages 133–144. Kluwer, Dordrecht, The Netherlands.
- [12] Hu, L. Y. and Chugunova, T. L. (2008). Multiple-point geostatistics for modeling subsurface heterogeneity: A comprehensive review. *Water Resources Research*, 44(W11413):14.

- [13] Huysmans, M. and Dassargues, A. (2010). Application of multiple-point geostatistics on modelling groundwater flow and transport in a cross-bedded aquifer. In Atkinson, P. M. M. and Lloyd, C. D. D., editors, *geoENV VII, Geostatistics for Environmental Applications*, volume 16 of *Quantitative Geology and Geostatistics*, pages 139–150. Springer Netherlands.
- [14] Huysmans, M., Peeters, L., Moermans, G., and Dassargues, A. (2008). Relating small-scale sedimentary structures and permeability in a cross-bedded aquifer. *Journal of Hydrology*, 361(1-2):41 – 51.
- [15] Journel, A. G. (2002). Combining knowledge from diverse sources: An alternative to traditional data independence hypotheses. *Mathematical Geology*.
- [16] Kessler, T., Klint, K. E. S., Renard, P., Nilsson, B., and Bjerg, P. L. (2010). Geostatistical description of geological heterogeneity in clayey till as input for improved characterization of contaminated sites. In *GQ10: Groundwater Quality Management in a Rapidly Changing World. (Proc. 7th International Groundwater Quality Conference, Zurich, 13-18 June 2010)*.
- [17] Koltermann, C. E. and Gorelick, S. M. (1996). Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches. *Water Resources Research*, 32(9):2617–2658.
- [18] Krishnan, S. (2008). The tau model for data redundancy and information combination in earth sciences: Theory and application. *Mathematical Geosciences*.
- [19] Le Coz, M., Genthon, P., and Adler, P. M. (2010). Multiple-point statistics for modeling facies heterogeneities in a porous medium: the Komadugu-Yobe alluvium, Lake Chad basin. submitted to *Mathematical Geosciences*.
- [20] Liu, Y. (2006). Using the Snesim program for multiple-point statistical simulation. *Computers & Geosciences*, 32(10):1544–1563.
- [21] Maji, R. and Sudicky, E. (2008). Influence of mass transfer characteristics for DNAPL source depletion and contaminant flux in a highly characterized glaciofluvial aquifer. *Journal of Contaminant Hydrology*, 102(1-2):105 – 119.
- [22] Michael, H. A., Li, H., Boucher, A., Sun, T., Caers, J., and Gorelick, S. M. (2010). Combining geologic-process models and geostatistics for conditional simulation of 3-D subsurface heterogeneity. *Water Resour. Res.*, 46(5):W05527.

-
- [23] Okabe, H. and Blunt, M. J. (2004). Prediction of permeability for porous media reconstructed using multiple-point statistics. *Physical Review*.
- [24] Okabe, H. and Blunt, M. J. (2007). Pore space reconstruction of vuggy carbonates using microtomography and multiple-point statistics. *Water Resources Research*.
- [25] Renard, P. and Ababou, R. (2010). Equivalent permeability tensor of heterogeneous and fractured media: upscaling methods and properties (review and analyses). soon published...
- [26] Renard, P., Genty, A., and Stauffer, F. (2001). Laboratory determination of the full permeability tensor. *J. Geophys. Res.*, 106(B11):26443–26452.
- [27] Rubin, Y. and Gómez-Hernández, J. J. (1990). A Stochastic Approach to the Problem of Upscaling of Conductivity in Disordered Media: Theory and Unconditional Numerical Simulations. *Water Resources Research*, 26:691–701.
- [28] Straubhaar, J., Renard, P., Mariethoz, G., Froidevaux, R., and Besson, O. (2010). An improved parallel multiple-point algorithm using a list approach. Mathematical Geosciences, accepted.
- [29] Strauffer, D. and Amnon, A. (1994). *Introduction to percolation theory*. Taylor and Francis.
- [30] Strebelle, S. (2000). *Sequential simulation drawing structure from training images*. PhD thesis, Stanford University, Stanford, CA, USA.
- [31] Strebelle, S. (2002). Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, 34:1–21.

Chapter 4

Geostatistical modeling of a three-dimensional high resolution fluvio-glacial aquifer analog

Abstract

Reconstructing the three-dimensional (3D) structure of a highly heterogeneous sedimentary deposit is a challenge, even when a huge data set is accessible. Multiple-point (MP) statistics is a geostatistical method that can be used to face this problem, but it requires a suitable training image and is based on the stationarity assumptions.

For the Herten aquifer analog (see companion paper, 2) several vertical outcrop sections provide a large quantity of data. These data can be used both as conditioning data and to deduce the spatial variability of the hydrofacies distribution. Still, the Herten aquifer analog is a challenging test for MP statistics because (1) a complete 3D training image is not available and (2) the hydrofacies distribution observed in the sections presents multiple non stationarities.

To overcome these shortcomings, we propose a hierarchical simulation method. The method subdivides the domain into small-scale regions. In these regions it is easy to generate 3D training images with a boolean technique; the training images can then be used for the MP statistics simulation of hydrofacies. Again, in these small-scale regions, the orientations of the geological structures is extracted using a morphological technique and then interpolated and mapped into the simulation domain. The orientation maps are included in the MP statistic simulation framework and allow to deal with the non stationarity of the orientation of the geological structures.

The reconstructed 3D aquifer analog reproduces realistic geological structures. These structures respect the spatial distribution of the hydrofacies observed in the outcrop sections. Moreover, the realism of the results from the point of view of the equivalent hydraulic conductivity and of breakthrough concentration curves is confirmed by a comparison made on the sections of the data set and an on five 2D slices extracted from the simulated 3D analog.

The reconstructed 3D aquifer analog is provided as supplementary to this paper as a unique solution that can be used as a research benchmark for a variety of applications in hydrology and hydrogeology. Moreover, the training images and the variogram models can be reused, for example, for the simulation of another gravel sediment deposit that stems from the same deposition environment but where few data are available.

4.1 Introduction

To deal with heterogeneity, there is the need for proficient modeling tools and detailed data sets to test and constrain them.

Among the existing data sets, geological analogs play a central role; they represent very detailed descriptions that can be used to infer accurate statis-

tical measures, allowing us to feed the models. Moreover, geological analogs provide benchmarks to test the applicability and quality of simulation techniques. Analog allow detailed observation at the surface of the geological structures that are expected to occur in deeper reservoirs. In this paper, we focus on alluvial sedimentary deposits that are extremely important in central Europe because they constitute the majority of aquifers used for water supply. These aquifers are often closely related to surface water and located in densely populated areas. Analog studies and heterogeneity modeling in these types of environments have been presented by many authors (1; 20; 18; 16; 15; 3; 4; 34; 19). The Herten site (2), located in the Rhine river valley in Germany close to the city of Basel (Switzerland), is selected for the present study.

We are particularly interested in this question: how can detailed 2D geological sections provided from analog studies be used to constrain 3D MP statistics models of such fluvio-glacial environments? The works of Koltermann and Gorelick (22) and de Marsily et al. (10) are interesting reviews of the techniques that are available to deal with heterogeneity. Among these techniques, MP statistics (12; 32; 17), allow to constrain the model using conceptual geological knowledge related to the expected shape and patterns of the geological facies. This conceptual model of the geological heterogeneity is provided by the user as a Training Image (TI). For example, the TI can represent meanders if the user knows that the geological environment that he wants to model is made of meanders.

For the user of MP statistics, the main question that he has to answer is how to get the TI? A classical proposal was to use geological analogs. However, only a few attempts to use real geological analogs have been made so far (25; 19). In these works, a difficulty that was rapidly identified is that analog observations are often highly non stationary (because of the physical nature of the geological processes that lead to the heterogeneity). If one uses such analog observations as training images, the simulations are not satisfactory (*e.g.* Boucher (5)). Indeed, one of the core principles of MP statistics algorithms is to assume that there is a repetition of the patterns in the training image allowing to build the MP probability distributions (ergodicity). If the training image is not stationary, all the different patterns are mixed in the inferred probability distributions and the simulated result is not what one would expect. Different techniques have been developed to overcome this limitation by providing additional information (*i.e.* orientation maps, results of geophysical surveys, probability maps) to avoid mixing the patterns (5; 9; 11; 31).

Note that the problem of non-stationarity of the data used to infer the statistics is not specific to MP statistics. The same problem affects geostatistical techniques based on lower order statistics. However, when only two-point statistical moments have to be reproduced, the simulated structures are not expected to resemble the original ones, and the problem is then less ap-

parent; indeed, it is known that two-point moments do not contain enough information to constrain the shape of the structures (see e.g. Strebel (32)).

Another issue when trying to use analog data to simulate 3D heterogeneity with MP statistics is that one needs in principle a 3D training image to simulate the 3D domain. In general, analog observations can only provide 2D training images. It is therefore necessary to develop specific methods that facilitate the use of MP statistics in that situation. Again, if we look at standard geostatistical procedures, the practice in those cases would be to assume some kind of invariance of the probability laws by rotation (8). The covariances or variograms would be inferred along the available data in the directions included in the geological section, and the variability in the perpendicular directions would be assumed to follow the same laws or to be simply rescaled to account for some anisotropy but without changing the type of distributions. Here similar assumptions could be made, but the current implementations of MP statistics do not yet offer these possibilities. When no other technique is available, the user therefore has to build himself some 3D training image that is – as far as possible – compatible with the 2D observations. One can do it for example by using object based techniques as proposed by Caers (6) and Maharaja (26), or by using 3D geometrical modeling techniques. If the geological structures are highly complex and intricate, this can be very tedious.

To overcome the problem of non stationarity and the problem of insufficient access to a 3D training image, the approach followed in this paper consists of subdividing the domain in smaller regions. In these smaller regions the heterogeneity is simplified and one can use basic but reasonable 3D training images constructed with object based techniques. The non stationarity related to the presence of different architectural geological elements is treated by using regions and the non stationarity of orientations of these elements is modeled separately by interpolating their orientations. In deltaic reservoirs, a similar approach has already proven to be very efficient (21).

The main object of the paper is to develop a methodology to reconstruct in 3D and in a realistic manner the geological heterogeneity which is contained in the six sections which compose the available data set. Moreover, the methodology must deal with the non stationarity observed in the data set and with the lack of 3D training images which impede the direct application of the MP statistics simulation technique.

The paper is structured as follows. First we present the Herten analog data set. Then we illustrate the limits of current modeling approaches on this data set in order to motivate the need for the proposed approach. The following sections are devoted to the step by step description of the approach. The paper ends with a presentation of the results and a discussion. All the MP statistic simulations are obtained with the algorithm IMPALA (Improved Multiple-Point Algorithm using a List Approach, Straubhaar et al. (31)).

4.2 The Herten dataset

The Institute for Geoscience of the University of Tübingen has studied in detail the Herten site, which has been intensively characterized. The data set contains six parallel geological sections having a dimension of 16 by 7 meters (figure 4.1). In the following we will refer to the six sections as $S_1, \dots, S_6$. These data were obtained by mapping the geological facies in successive outcrop sections exposed during the pit exploitation.

The distance between each section is 2 m. In the reference system adopted in this manuscript these sections are parallel to the plane xz .

Ten different hydrofacies types are mapped within the sections with a discretization of 5 cm along both x and z direction. For each hydrofacies, hydraulic conductivity and effective porosity was measured on a number of samples ranging from 3 to 15. For an exhaustive description of the mapping procedure, the hydrofacies coding conventions, as well as the available sedimentary, hydrogeological and geophysical data sets the reader is referred to part 1 of the companion paper (2).

In the present paper, we will use the hydrofacies data from the Herten data set and build a 3D stochastic model within a novel hierarchical MP statistics framework. The hydrofacies from the outcrops (fig. 4.1) are used both as hard conditioning data (268 000 points) to constrain the simulation and more generally as a source of information to develop the whole methodology. The discretization is identical to that used by Maji et al. (28). The simulation domain Ω is divided into 9 million ($320 \times 201 \times 140$) cubic cells having a side of 5 cm.

This discretization is selected for three reasons: (1) to reflect the very high resolution of the data; (2) to allow direct comparison with related previous work on the same site; (3) to provide a high resolution 3D description of the heterogeneity of fluvio-glacial deposits that could then be used in the future to investigate how the heterogeneity can control physical processes in these systems.

4.3 Motivations for a hierarchical approach

Before entering into the description of the proposed hierarchical approach, it is important to visualize and understand the impact of some limits of available approaches for modeling the 3D heterogeneity of the Herten site.

Several studies have used the Herten dataset (see companion paper, 2). Among them, the most interesting for our purposes are the works of Maji (27) and Maji et al. (28). These authors applied a transition-probability-based Markov chain approach (TP/MC, 7), using the software TProgs, to reconstruct the complete 3D volume and to use it for analyzing the impact of heterogeneity on DNAPL transport.

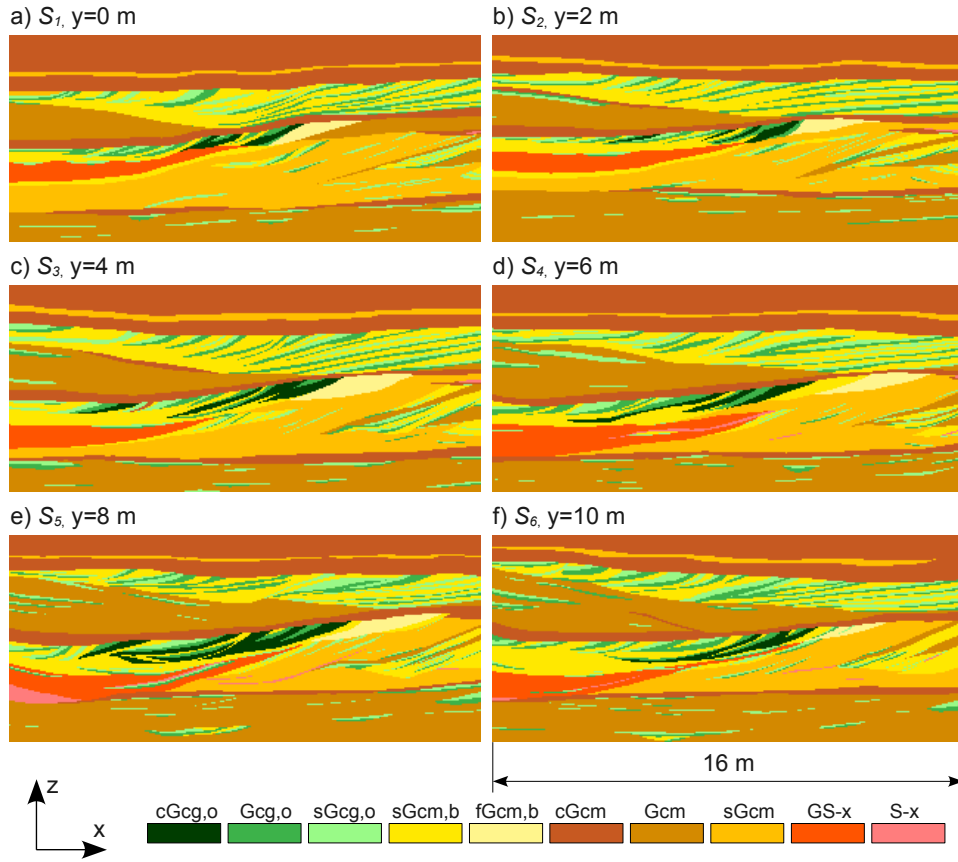


Figure 4.1: Rasterized six sections S_j (16 by 7 m) of the Herten case study. Figure 4.7(a) provides an *ensemble* view. The hydrofacies codes are explained in detail in the companion paper (2).

The TP/MC approach accounts for the probability of transition from one hydrofacies to another with the use of a transition probability matrix. This matrix is composed of transition probability curves which are adjusted from the data contained in the available sections; the 3D simulation is performed using all the sections as conditional data.

Two slices within a 3D simulation obtained with the TP/MC approach are displayed in figures 4.2(c) and (d). Figure 4.2(c) shows a slice parallel and close (0.25 m) to the conditioning data of the first section S_1 (fig.4.2(a)); Figure 4.2(d) is placed in between S_1 and S_2 , at a distance of one meter from both. These two slices show that the simulation becomes very noisy when the simulated nodes get further away from the conditioning data. More precisely, thin continuous layers, especially in the upper part of the domain, become clearly disconnected, and may have substantial influence on the hydraulic behavior of the whole heterogeneous domain. Even if the transition

probability curves shown by Maji (27) agree rather well with the experimental data, it is possible that better results could be obtained with an improved parameterization within TProgs. Actually, we believe that the main reasons for the problems with the TP/MC simulations are (1) the high non stationarity of the data set and (2) the poor information available for the computation of the probability transition matrix along the direction normal to the outcrops.

As a remedy the TP/MC approach could be applied, for example, in a sequence stratigraphic framework (33). However, this would be still constrained by the use of statistical models fitted on experimental curves that are computed considering only two points at a time. To overcome this last limitation, we investigate MP statistics.

The analog sections are only 2D, thus we cannot use them directly as training images to simulate a 3D domain with MP statistics. However, one of these sections (*i.e.* S_1) can be used as training image to perform 2D unconditional MP statistics simulations. Figure 4.2(e) is a simulation obtained without taking into account the non-stationarity of the training image. As shown by Chugunova and Hu (9) and Boucher (5), all the patterns present in the training image are mixed up in the simulation and the architectural elements contained in the section are not reproduced.

To overcome this, one can use auxiliary variables to describe the trends in the training image and in the simulation domain (9; 31). Briefly, an auxiliary variable can be represented by a continuous field superimposed to the training image and which is known in the simulation domain. This field is used to constrain MP statistics: the patterns which are observed in the training image only for a given value of the auxiliary variable field are reproduced in the simulation domain only for the corresponding values of the auxiliary variables (with a given tolerance). Here, for example, the training image (fig. 4.2(a)) contains an evident vertical trend along the z . The results obtained with the MP statistics using z as auxiliary variable (fig. 4.2(f)) are remarkably improved if compared with the ones obtained without the use of the auxiliary variable (fig. 4.2(e)). Probably, one could go a step further by using a second auxiliary variable along the x coordinate (horizontal) to improve the results. In addition, note that the two 2D MP statistics simulations are obtained without using conditioning data.

Even if the simulation of fig. 4.2(f) is rather good, we still note that the thin inter-bedded layer (with hydrofacies code *sGcm*, see companion paper, 2) on top of the section is duplicated. This observation reflects the fact that, in general, MP statistics is suitable for reproducing repetitive patterns, but it is not suited to modeling single discrete features that should rather be modeled with other methods.

Finally, the application of MP statistics on the whole 3D domain remains a problem because there is still a need for a 3D training image or for an improved MP statistics methods allowing 3D volumes to be simulated from

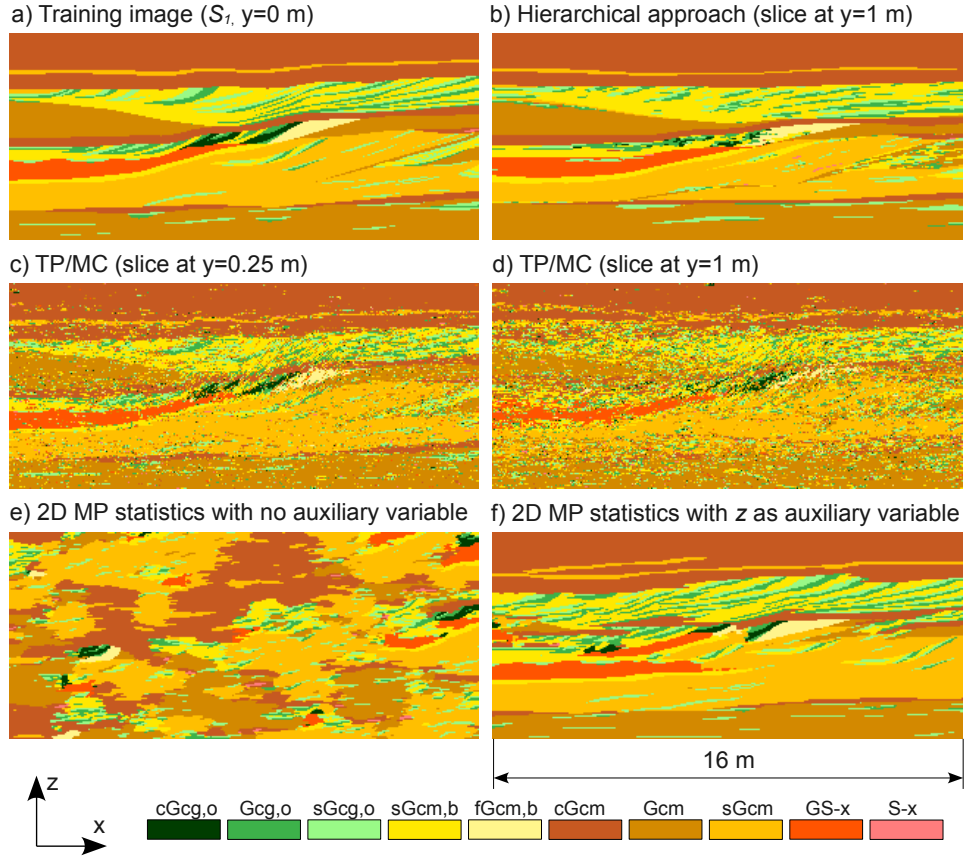


Figure 4.2: A comparison between the simulations obtained with TP/MC, MP statistics with and without auxiliary variables and the proposed hierarchical approach. Fig. b),c) and d) are slices parallel to the plane xz of 3D simulations obtained using the six S_j as conditioning data. Fig. e) and f) are obtained with the training image a) and no conditioning data. The hydrofacies codes are explained in detail in the companion paper (2).

2D training images. An appropriate procedure to overcome this problem is presented subsequently.

Therefore we propose to dichotomize the problem and to model the main layers and the small-scale heterogeneity separately. A first encouraging result of this method is shown in fig. 4.2(b) precisely at the same location as the one shown in fig. 4.2(d) for the TP/MC based approach.

4.4 The hierarchical simulation framework

Figure 4.1 exhibits large-scale structures composed of sub-horizontal layers with small-scale heterogeneities within them. It becomes hence natural to

adopt a hierarchical approach in order to simulate the two structures separately.

Figure 4.3 describes the major steps of the proposed framework, which are explained in the following.

The first major step is to recognize the main layered large-scale structures inside the sections, and extend the large-scale structures to the 3D simulation domain with this procedure:

1. define the hierarchy among structures and identify in each section separate regions R_i
2. infer variogram and interpolate with ordinary Kriging, in the whole 3D domain, the separation surfaces between two adjacent regions
3. create the large-scale layered structure, applying given erosion rules on the interpolated surfaces.

At the end of this step, the simulation domain is divided into a number of homogeneous and heterogeneous sections.

The second major step concerns the heterogeneous regions only. It consists in applying the MP statistics method to simulate, at small scale, the fine heterogeneous structures:

1. exploratory and morphological analysis of the shapes contained in each heterogeneous region, in order to get the geometrical parameters suited for the application of object based simulation methods
2. simulate suitable 3D training images with an object based technique
3. interpolate in all the R_i the orientation of the small-scale heterogeneities, in order to create a map to be used as input for the MP statistics simulation technique
4. simulate internal heterogeneity within each region with the MP statistic using the conditioning data provided by the outcrops and the orientation fields
5. assemble the results.

The heterogeneous regions simulated in this second step are then included in the 3D domain.

4.4.1 Subdivide the domain into large scale regions

The first step of the hierarchical framework consists of partitioning the whole domain Ω in a number of regions R_i such that $\Omega = \cup_i R_i$ and $R_i \cap R_j = \emptyset$ when $i \neq j$.

The regions R_i are first identified manually inside the sections S_j ($R_i|_{S_1, \dots, S_6}$,

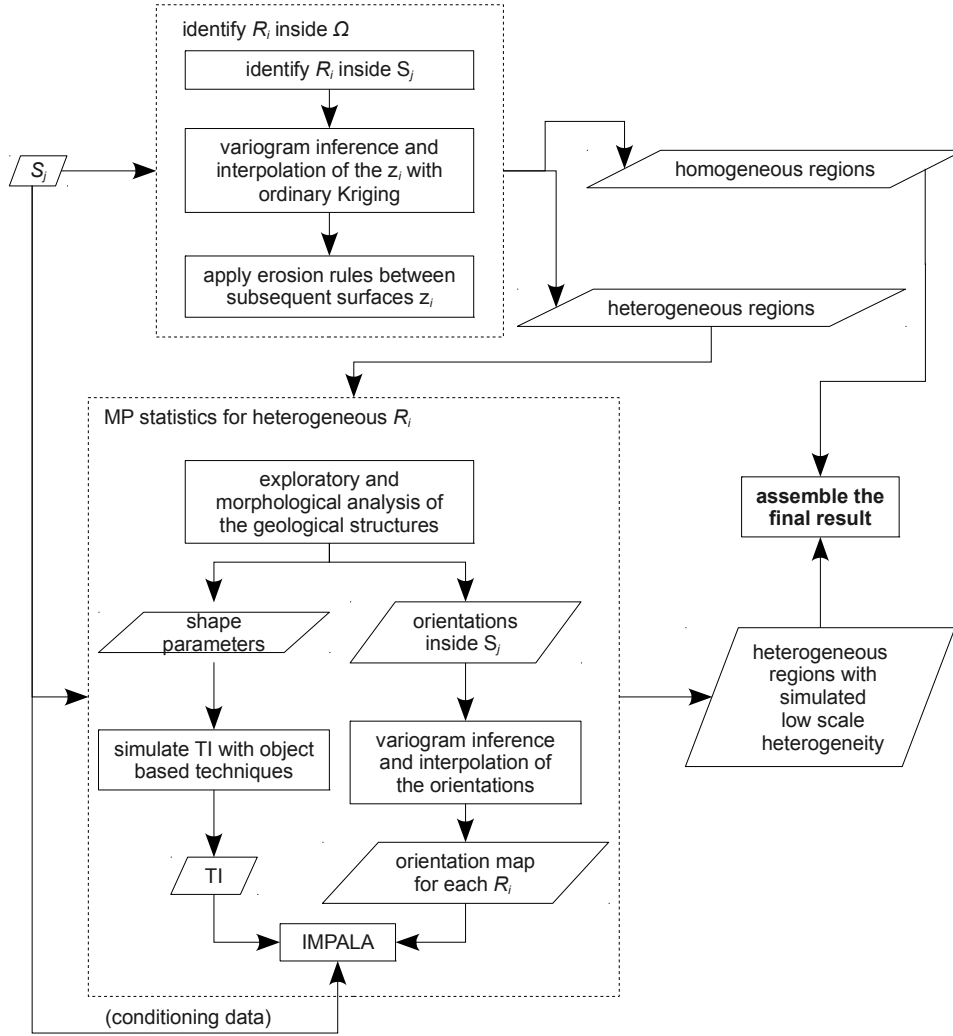


Figure 4.3: A schematic representation of the proposed hierarchical framework.

see fig. 4.4). The genetic units defined in the companion paper (2) are used as guidelines for the definition of the regions, which are identified following one of the two following rules. The first rule is used when the distribution of the hydrofacies presents a small-scale heterogeneity (*i.e.* of the order of decimeters, see lower part of fig. 4.5(a)). With this rule, the cells are clustered in order to form a unique region; this is the case of R_{19} in fig. 4.5(b). The regions defined with this rule are heterogeneous. The second rule is used when homogeneous large-scale layers or architectural elements are observed. This is for example the case of the genetic unit IV (see companion paper, 2), where three homogeneous regions R_{21} , R_{22} and R_{23} are identified (fig. 4.5(a) and (b)).

A third rule was added in order to facilitate the interpolation procedure: the identified regions R_i do not have to overlap. This is illustrated in fig. 4.5 (c) and (d), where following only the two rules presented before the regions R_9 and R_{11} would have been considered as a whole; the upper limit z of this last region would not have been uniquely defined by its x and y coordinates, creating a problem for the subsequent interpolation procedure.

Adding this last rule, the relation among the upper limits z of the identified regions is the following

$$z_{i|S_1, \dots, S_6} \geq z_{i-1|S_1, \dots, S_6} \quad \text{for } i = 2, \dots, n_R - 1. \quad (4.1)$$

In the equation (4.1) n_R stands for the number of identified R_i ; the vertical z axis is pointing upwards and the index increases from lower z to higher z . Eleven homogeneous regions and twelve heterogeneous regions $R_{i|S_1, \dots, S_6}$ are identified (in total, $n_R = 23$) following the rules illustrated before. The results of this identification procedure are depicted in figure 4.4.

A systematic shift of the heights $z_{i|S_2}$, especially for the lower regions (R_i with $i \leq 7$) was identified by preliminary interpolations (see table 4.1). Due to the impossibility of investigating the reason for this systematic shift in depth, we decided to subtract its mean value (0.3 m) from all the $z_{i|S_2}$ (for more details see table 4.1). Another possibility would be to ignore the z values provided by the second section during the interpolation process, for example for all $z_{i|S_2}$ with $i \leq 7$. However, we preferred to apply the shift to all the $z_{i|S_2}$ in order to keep all the data for the interpolation.

Structural analysis is performed on the corrected heights, for each region, in order to fit a variogram model and interpolate the z_i with ordinary Kriging on Ω . The variogram model adopted for z_1 is depicted in figure 4.6 (a).

To avoid overlapping among the interpolated surfaces some erosion rules are applied according to equation (4.1). The z_i interpolated are corrected taking as reference the lower surface, starting from $i = 1$, with the following

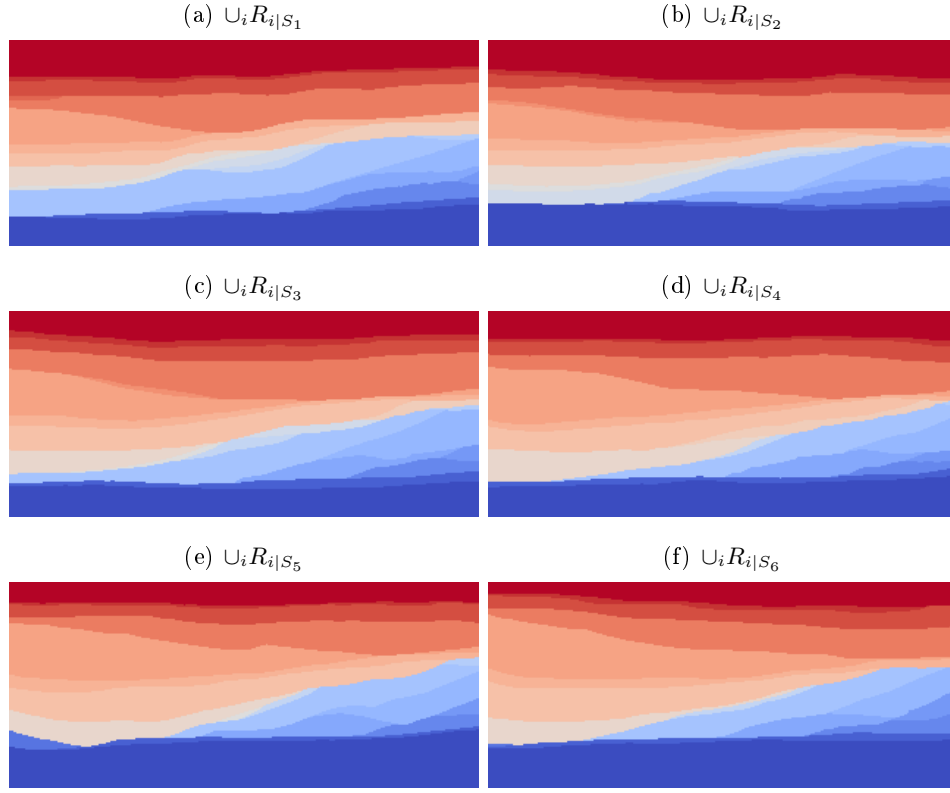


Figure 4.4: The different colors represent the regions R_i identified inside each section S_j from the geological sections displayed in figure 4.1. For a color-region correspondence see fig.4.7(c).

rule:

$$z_i = \begin{cases} z_i & \text{if } z_i \geq z_{i-1} \\ z_{i-1} & \text{if } z_i < z_{i-1} \end{cases} \quad \text{for } i = 2, \dots, n_R - 1 \quad (4.2)$$

Figure 4.7 (b) represents the layered structure obtained in this first reconstruction step.

4.4.2 Small scale: MP statistics

While eleven regions of the layered structure obtained by interpolation can be considered homogeneous, the other twelve regions R_i with $i = 1, 4, 5, 6, 7, 8, 9, 10, 13, 15, 17, 19$ have a heterogeneous internal structure. Their structure is simulated using MP statistics with an individual training image I_i for each region.

The following sections describe the procedure adopted to generate the training images required to apply MP statistics. Also, the steps required to take into account the non stationarity of the orientations of the geological struc-

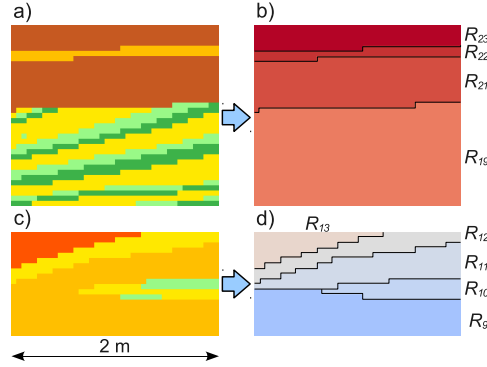


Figure 4.5: Procedure adopted to identify the regions R_i inside subsets of the section S_1 . Lower part of (a) and (b): R_{19} is created aggregating neighboring cells which contain small-scale architectural elements. Upper part of (a) and (b): R_{21} , R_{22} , and R_{23} are created aggregating homogeneous layers. (c) and (d): two regions (in this case R_9 and R_{11}) are created in order to facilitate the interpolation of their upper limit z .

tures in MP simulations are reported. The following procedure is applied to each heterogeneous region.

Simulate the missing training images

At the scale of the regions R_i one can recognize geological structures inside the six 2D sections $R_{i|S_1, \dots, 6}$ and deduce some 2D training images. However, 3D training images are required in order to apply MP statistics methods. The missing 3D training images are generated using object-based techniques, relying on the information provided by the 2D sections. We suppose that the 3D shapes can be represented as lenses, as in the case study described by Jussel et al. (20). The objects are generated at random locations with a shape defined by:

$$\frac{(x - x_O)^2 + (y - y_O)^2}{r_e^2} + \frac{(z - z_O - [z - z_O]\lambda r_p)^2}{r_p^2} = 1, \quad \lambda \in [0, 1]. \quad (4.3)$$

Expression (4.3) represents an ellipsoid flattened by a factor λ , centered in the point (x_O, y_O, z_O) with equatorial radii r_e and polar radius r_p . The parameters r_e and r_p are varied with uniform laws $\mathcal{U}([r_{e \min}, r_{e \max}])$ and $\mathcal{U}([r_{p \min}, r_{p \max}])$ for each simulated object; the simulation is stopped when a target proportion is reached.

The values of the parameters $r_{e \min}$, $r_{e \max}$, $r_{p \min}$, $r_{p \max}$ and λ adopted for the boolean simulations are reported in table 4.2. These values are estimated manually by trial and error from the outcrops instead of using an automated procedure which would have been required to account for the

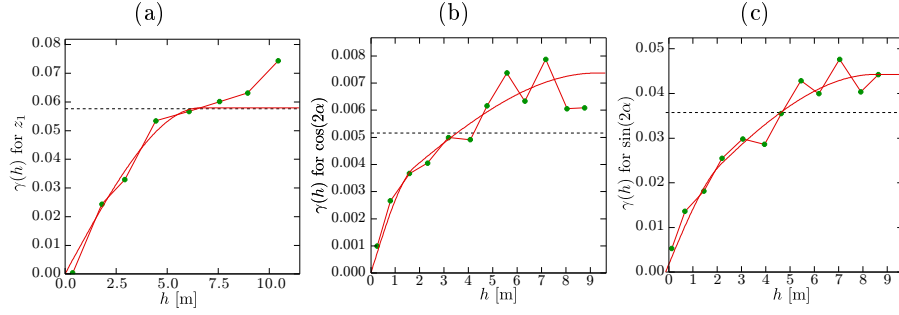


Figure 4.6: The variograms adopted to interpolate (a) the top of region R_1 ; (b) $\cos(2\alpha)$ and (c) $\sin(2\alpha)$ for R_{19} . Here α is the orientation of the objects in the plane xz ($\alpha = 0$ on the x axis and positive counterclockwise).

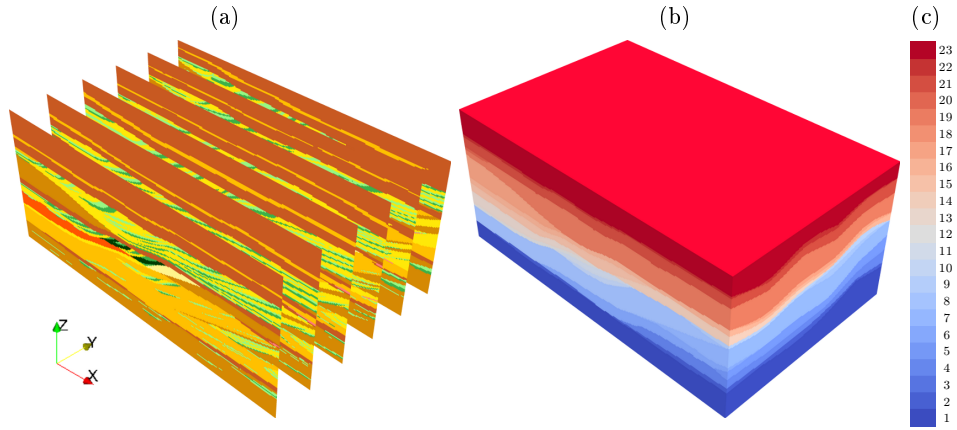


Figure 4.7: (a) The six sections S_j of the Herten data set and (b) the layered structure obtained after the interpolation of the upper limit of the 23 regions R_i .

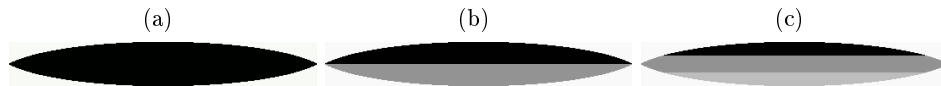


Figure 4.8: The slices of the three kinds of flattened ellipsoids, used for the boolean simulation of the TIs, composed by (a) one, (b) two and (c) three facies.

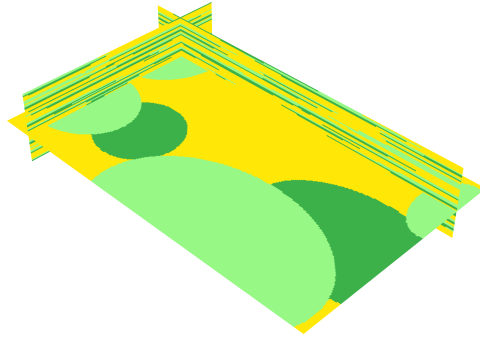


Figure 4.9: The 3D training image used to simulate with MP statistics the region R_{19} .

limited size of the outcrops, for object truncation due to erosion processes and for the non stationarity in orientation of the objects.

The facies proportions used as threshold to stop the object simulation procedure are extracted from the data contained in the outcrops. However, inferred 3D object properties from 2D outcrops introduces a bias: for this reason, even if for each heterogeneous region R_i we try to simulate a training image with proportions close to those of the outcrops $R_{i|S_{1,\dots,6}}$, we also accept training images simulated with proportions having the same order of magnitude of the proportions extracted from the sections. For each heterogeneous region R_i , the proportions extracted from the outcrops $R_{i|S_{1,\dots,6}}$, the proportions obtained for the training image I_i , and the proportions obtained from two realizations of conditional MP simulations are in agreement (fig. 4.10).

The boolean simulation code allows the generation of three different types of flattened ellipsoid having a composition of one, two or three facies (fig. 4.8); in fact, we suppose that the shape of the small-scale geological structures which can be elicited from the outcrops can be approximated with these types of ellipsoids. Consequently, the training images for the 12 heterogeneous regions are generated with a combination of the three types, in order to mirror the compositions observed in the sections S_j .

Objects orientation

The non stationarity inherent to the layered structure of the site at large scale can be taken into account by subdividing the domain Ω into smaller regions R_i . However, at the scale of the heterogeneous regions R_i , there is another source of non stationarity: the one related to the orientation of the heterogeneities within the region. With the MP statistics simulation algorithm IMPALA (31) it is possible to account as well for this type of non stationarity; inside a simulation domain, different zones can be defined where individual simulation parameters can be applied, like a dilatation or

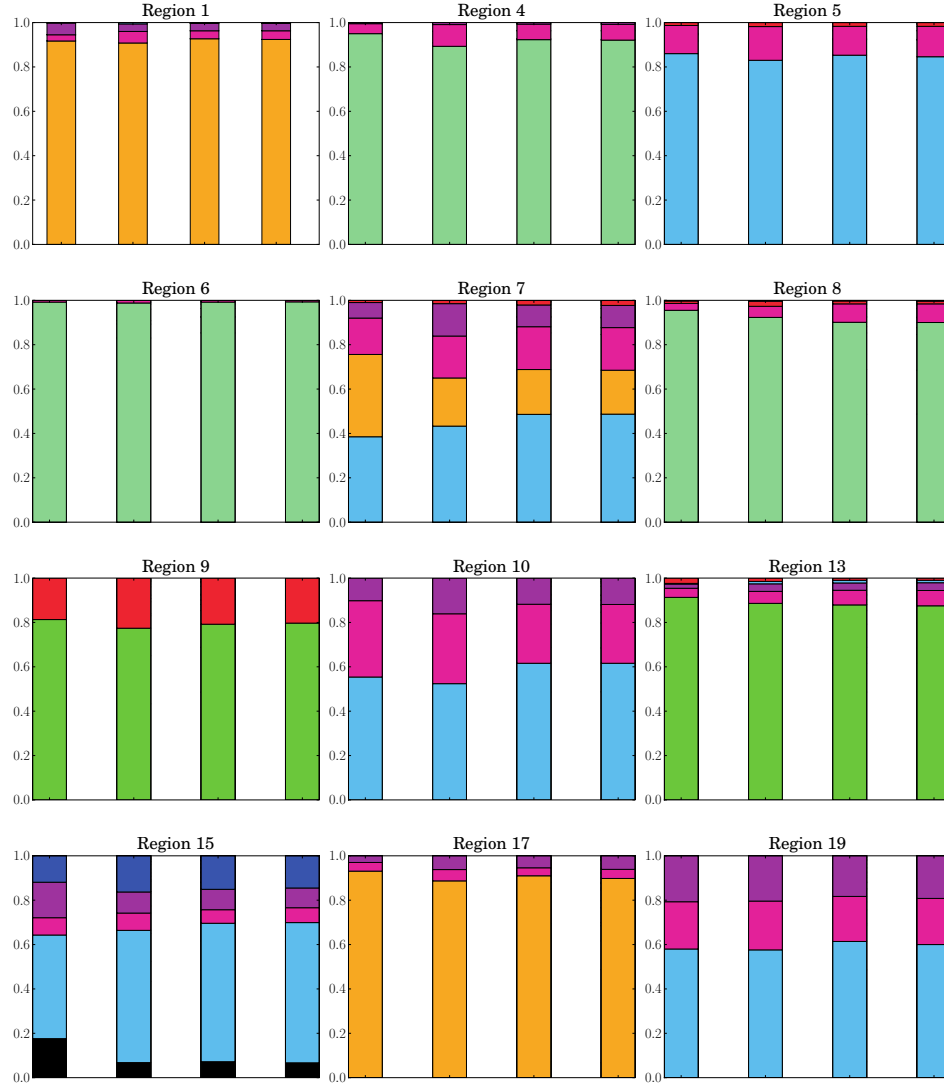


Figure 4.10: The hydrofacies proportions computed for each heterogeneous region R_i . For each region, the first bar on the left represents the proportions computed on the conditioning data, the second the proportions on the training image I_i used for the simulation, the third and the fourth the proportions of two realizations of a conditional MP simulation.

a rotation of the training image. In order to perform MP simulation in each heterogeneous R_i , IMPALA requires a suitable training image and an orientation map for the objects. This section is devoted to this last map.

In the following examples, and for testing the sets of control variables of the MP simulation, we focus on region R_{19} ; in fact this region groups a great number of features whose reproduction are a challenge in the 3D simulations, like, for example, extended continuous layers.

Figure 4.11 depicts the main steps performed in order to detect the orientations of the objects. The steps are implemented in Matlab and the figure reports as an example the case of region R_{19} inside section S_1 . Because we only have access to sections along the outcrop plane xz , we only determine orientations on this plane.

For each heterogeneous region $R_{i|S_1, \dots, S_6}$ the orientations are computed with the following procedure:

- Distinguish the matrix and the objects. For example the hydrofacies Gcg, o and $sGcg, o$ can be considered as forming unique objects inside a matrix of $sGcm, b$ (fig. 4.11(a),(b)).
- Apply a morphological operation in order to reduce the objects into thin shapes (fig. 4.11(c)).
- Label each reduced object using the connectivity rule described by (14) and considering eight neighboring points (fig. 4.11(d)).
- Fit a polynomial (red lines, fig. 4.11(d)) for each labeled object with a least squares method. Use the derivative of the polynomial to get the orientation value α in each point of the object skeleton (dark arrows, fig. 4.11(e)).

Orientation interpolation

To define the 3D orientation maps required by the MP statistics simulations the punctual values of orientations computed along the skeleton of the small-scale architectural elements have to be interpolated.

Interpolating local directional data is a challenge (23; 13). The main problem is that orientation data are circular when they are represented as angles. They only vary between 0° and 360° , and the value 0° is closer to 370° than the value 20° , for example. A change of variable must be used to ensure that the interpolation is made properly.

In geostatistics, several approaches have been designed for the interpolation of orientation data. Lajaunie et al. (23) and Chilès and Delfiner (8), chapter 5.5 observe that orientation data can be represented as the gradient of a scalar field; starting from this consideration, they propose using general co-Kriging equations to interpolate the scalar field and map the orientations.

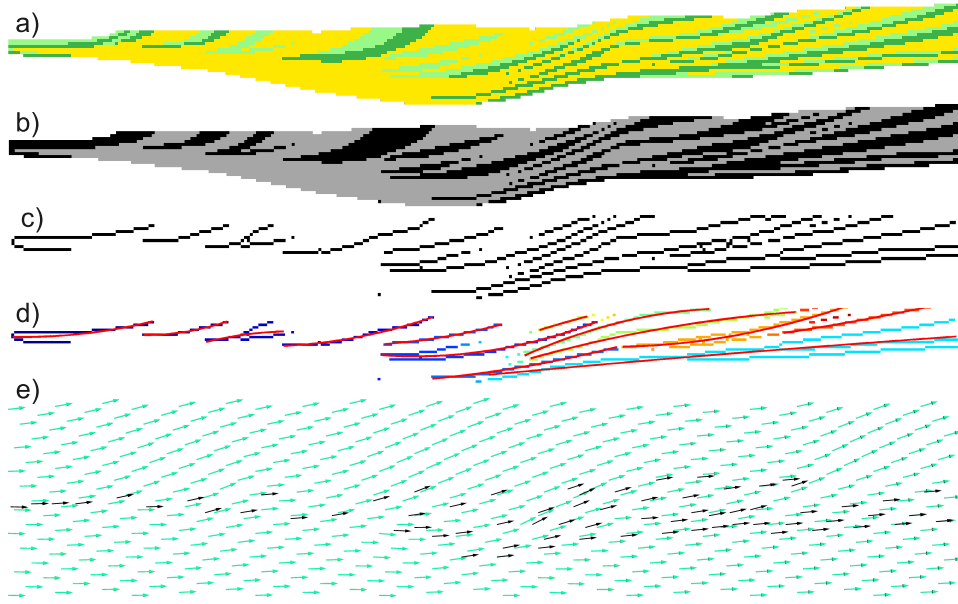


Figure 4.11: Part of the steps performed to compute an orientation map for region R_{19} and the section S_1 : (a) $R_{19} \cap S_1$ (b) the objects inside the matrix are identified; (c) objects are reduced into their skeleton; (d) the skeletons of distinct objects are labeled (with distinct colors), and polynomials of order two are fitted on the skeletons (red lines); (e) the orientations are deduced from the derivatives of the polynomials (dark arrows). The bright arrows in (e) are a slice of the 3D interpolation of the orientations.

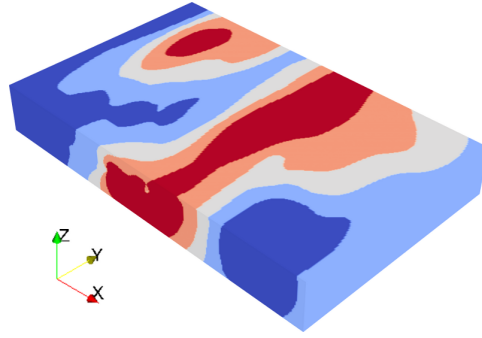


Figure 4.12: The five orientation zones defined to take into account for the non stationarity in the orientation of the geological structures for region R_{15} .

(13) provide a review of the main methods and propose also a simple solution to the problem. Instead of interpolating directly the orientation angle α , they recommend interpolating $\cos(2\alpha)$ and $\sin(2\alpha)$; then the values of α on the whole interpolation domain are computed with $\arctan(\sin(2\alpha)/\cos(2\alpha))/2$. In this work, due to the small range of variation of the orientation of the object, we apply this last approach.

Structural analysis is performed on $\cos(2\alpha)$ and $\sin(2\alpha)$ for each heterogeneous region R_i . Here we report as examples the variogram models adopted to interpolate with ordinary Kriging $\cos(2\alpha)$ and $\sin(2\alpha)$ for R_{19} (fig. 4.6(b) and 4.6(d)). The bright arrows of fig.4.11(e) represent a slice of the orientations interpolated for R_{19} .

For each heterogeneous region R_i , the interpolated orientations are grouped in five orientation classes, in order to define five zones in which to apply the MP statistic method with the same training image but with different orientations. The classification of the orientations is done according to the quantiles: a quantile step of 10% is computed, and while the 0%, 20%, 40%, 60%, 80% and 100% quantiles are adopted as bounds for the five classification intervals, the 10%, 30%, 50%, 70%, and 90% quantiles are used as the orientations to characterize each zone. Table 4.3 contains all the values adopted for the previous classifications. Figure 4.12 represents the five orientation zones used to apply MP statistics to region R_{15} .

A test performed on R_{19} using ten orientation classes (quantile step of 5%) did not improve the simulation result; therefore, the classification using five classes is considered sufficient.

MP statistics simulation

Selecting a suitable set of control variables of the MP statistics simulation is a crucial point. Here we test different sets with a “trial and error” procedure,

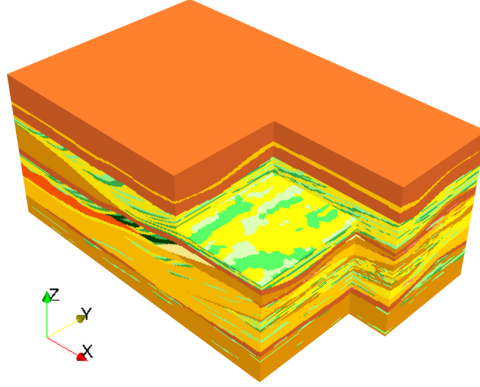


Figure 4.13: One realization obtained with the hierarchical simulation framework taking into account the conditioning data.

following the guidelines provided by Liu (24) and guided by our experience. A large part of the tests are performed only on region R_{19} .

An important question is how to evaluate and compare the results obtained with different sets of variables. In this context the results of the simulations are evaluated for their ability to reproduce, first, the observed hydrofacies proportions, and, second, the continuity of the structures, which can be deduced from visual inspection of the sections S_j .

For a detailed explanation of the variables tested and their influence on the simulated results, refer to Liu (24) and Straubhaar et al. (31); here after we report only a brief overview of the variables considered in the tests.

The first variables inspected are the number of multi-grid levels and the size of the search neighborhood. With a big search neighborhood it is possible to capture large-scale features of the training image, but an increased neighborhood requires more RAM and more CPU time. Increasing the multi-grid level allows the reproduction of large-scale features with a smaller search neighborhood. Here we test neighborhood of size $3 \times 3 \times 3$, $5 \times 5 \times 5$ and $7 \times 7 \times 7$ and multi-grid levels from 1 to 6.

Another control variable of interest for the MP statistics is the simulation path. Depending on the disposition of the conditioning data and on the shapes which have to be reproduced, this variable can influence the simulation result. The simulation paths tested in this case study are the following: a random path; a path performed varying the coordinates of the simulation grid in the order x , y and z (“unilateral” path); a path in which the grid nodes closer to the conditioning data are simulated first.

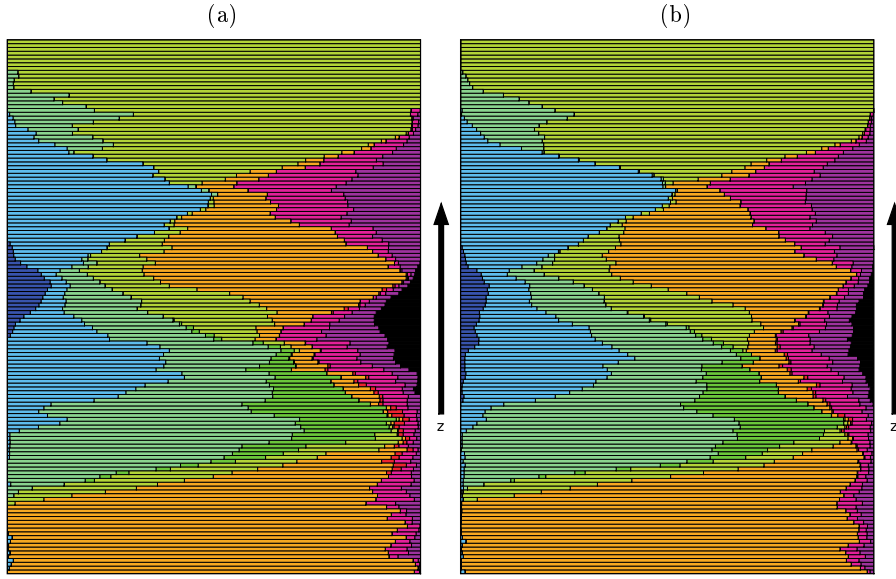


Figure 4.14: The hydrofacies proportions computed on horizontal slices (parallel to xy plane) of 3D simulations. Slices are computed for each discretization step along the axis z (140 slices in total). Proportions on slices are computed (a) on the conditioning data and (b) on the 3D simulation of figure 4.13.

4.5 Results and Discussion

Figure 4.13 illustrates one realization obtained with the hierarchical simulation framework. Different realizations are obtained with the same set of control variables, while using a different seed for the random MP statistics simulations. The control variables used are 5 multi-grid levels along the directions x and y and 1 multi-grid level along z , a random simulation path and a search neighborhood of size $5 \times 5 \times 5$. Using this set of variables the realizations obtained reproduce reasonably well the complex structure of the medium given by the six vertical sections of figure 4.1. Thin layers are now continuous over most of the domain. Internal heterogeneity within a region is constrained within the region. Moreover, the variation of the hydrofacies proportion computed along the vertical axis z for the simulated domain and the conditional data are in good agreement (fig. 4.14). These results are closer to the observation and much less noisy than previous simulations made using transition probability techniques (fig. 4.2 (c) and (d)).

Two numerical flow simulation tests are performed in order to corroborate the results of the visual comparisons and of the comparisons on the proportions.

The first tests are performed under stationary conditions, with two different

configuration of linear boundary conditions along the x and along the z axes. Using the results of these simulations and the values of hydraulic conductivity provided in the companion paper (2), it is possible to compute the equivalent hydraulic conductivity tensor $\mathbf{K}$ (30; 29). $\mathbf{K}$ is computed for the six sections of the data set, for five 2D slices in the 3D simulation obtained with the TP/MC approach and for five 2D slices in the 3D simulation obtained with the hierarchical approach. These slices are parallel to the planes of the sections and in between of them, that is for $y = 1, 3, 5$ and 7 m. K_{xx} and K_{xz} are the components of $\mathbf{K}$ which provide indications about the continuity of the lithological layers observed in the sections along the direction x . The values of K_{xx} and K_{xz} of the slices extracted from a simulation obtained with the hierarchical approach are closer to the ones of the reference sections than the values of K_{xx} and K_{xz} computed with the TP/MC approach (fig. 4.15 (a) and (b)). Concerning the component K_{zz} , taking into account its range of variation we can consider comparable the values obtained for the sections and for the slices extracted from the 3D simulations obtained with the two simulation methods (fig. 4.15 (c)).

The second tests are based on transport simulations performed on the same sections and on the same slices where $\mathbf{K}$ is computed. The solute transport problem is simulated imposing an horizontal flux along the direction x with permeameter boundary conditions, and with a fixed concentration solute boundary condition at $x = 0$ m (fig. 4.16). The evolution of the average concentration is observed on the outflow boundary (fig. 4.17). Figure 4.17 shows that the contaminant breakthrough curves for the slices in the simulation obtained with the hierarchical approach (fig. 4.17 (b)) are closer to the curves related to the six surfaces of the data set (fig. 4.17 (a)) than the curves for the slices in the simulation obtained with the TP/MC approach (fig. 4.17 (c)). The same considerations can be applied to the result shown in fig. 4.16, where the relative concentration at two different time steps in the section S_1 and in two 2D slices parallel to S_1 at $y = 1$ m extracted from the 3D simulation obtained with the hierarchical approach and the TP/MC approach are compared.

The hierarchical approach described here allows the use of MP simulation tools with simple training images generated using well known boolean simulation techniques. This is an important advantage because the accessibility of a suitable training image, especially in 3D cases, is often a problem. At the same time, while the small-scale intricacy is reproduced using MP methods, the large-scale structures can be reproduced using standard geostatistical methods like for example ordinary Kriging.

With the simplification introduced by the hierarchical framework it is possible to take into account the different aspects of the non stationarity which characterizes complex sites (like *i.e.* the Herten case study). One aspect of the non stationarity is tackled by splitting the domain in sub-horizontal regions R_i . Another aspect of the non stationarity is related to

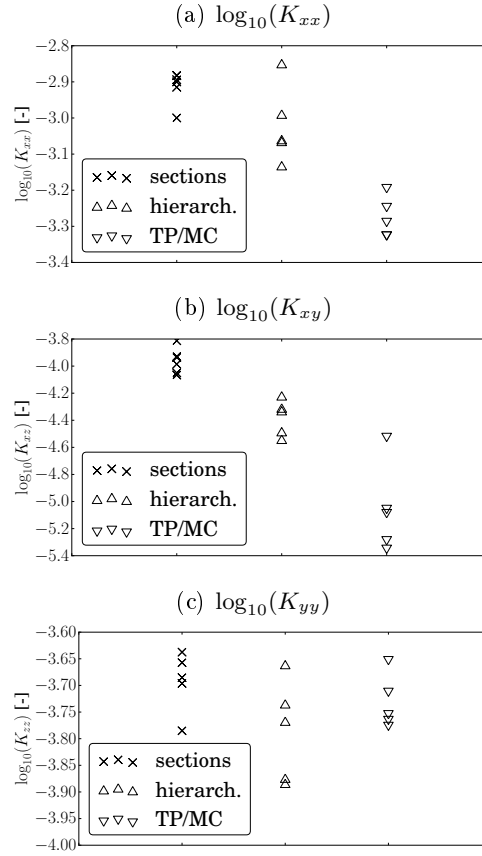


Figure 4.15: The $\log_{10}$ of the components of the equivalent conductivity tensor $\mathbf{K}$. (a) K_{xx} , (b) K_{xy} and (c) K_{yy} are computed on the six sections of the data set, on five 2D slices in the 3D simulation obtained with the hierarchical approach and on five 2D slices obtained with the TP/MC approach.

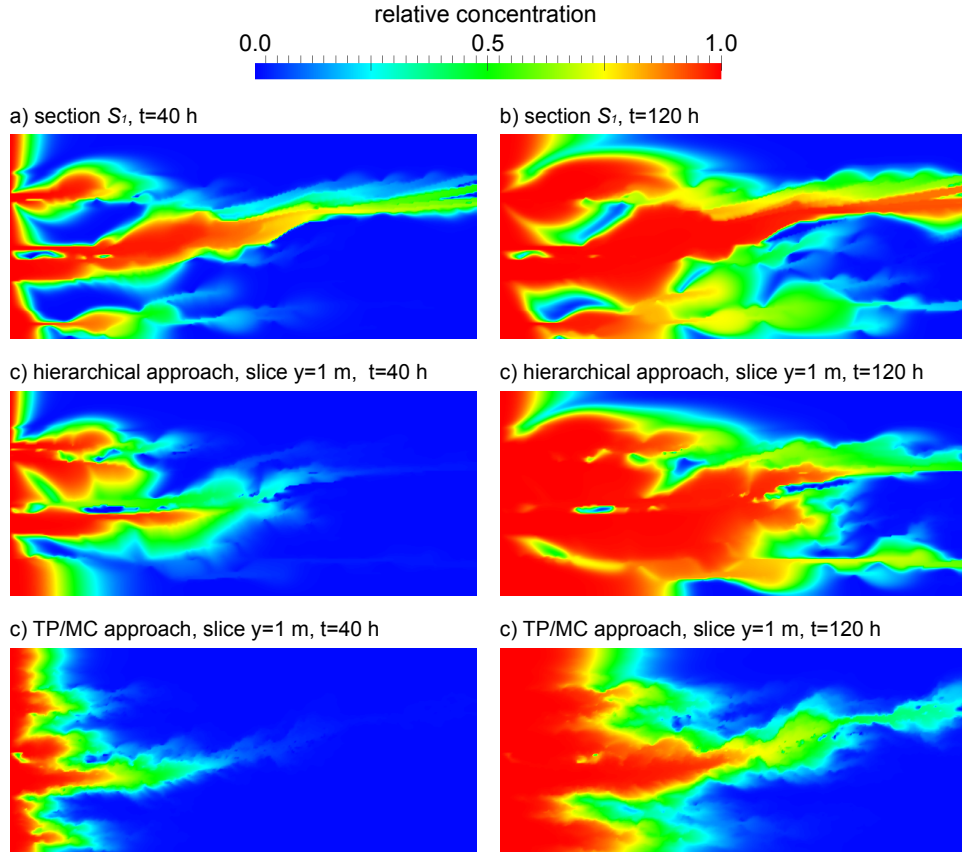


Figure 4.16: The relative concentration after 40 and 120 hours into (a) the section S_1 , (b) a slice on the 3D simulation obtained with the hierarchical approach, parallel to S_1 and in between of S_1 and S_2 ($y = 1$), (c) a slice on the 3D simulation obtained with the TP/MC approach in the same position as the slice (b).

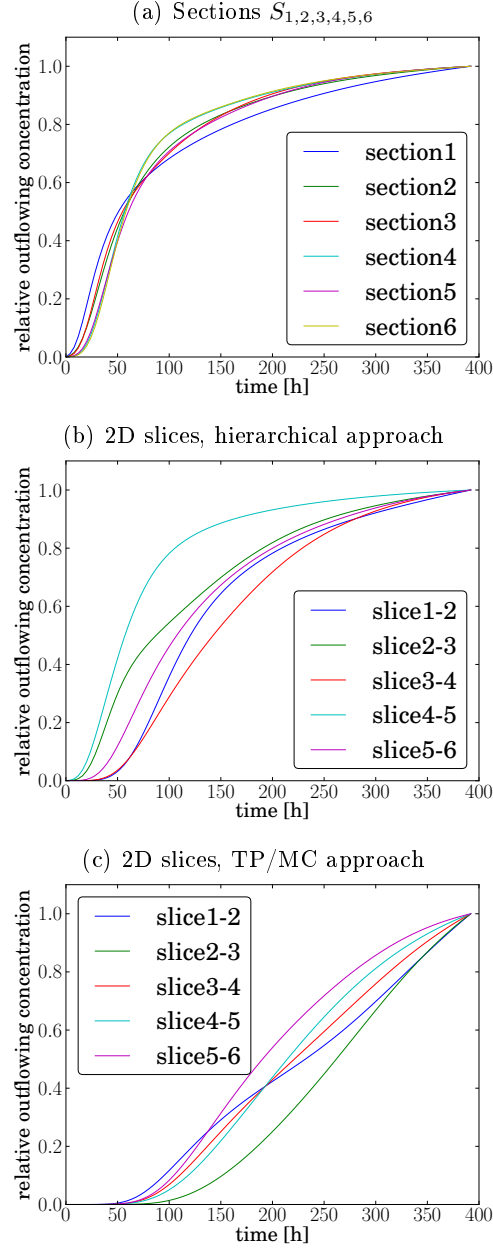


Figure 4.17: Contaminant breakthrough curves for (a) the section S_1 , (b) a slice on the 3D simulation obtained with the hierarchical approach, parallel to S_1 and in between of S_1 and S_2 , (c) a slice on the 3D simulation obtained with the TP/MC approach in the same position as the slice (b).

the orientation of the geological structures contained in the heterogeneous regions R_i . The hierarchical approach allows one to apply MP statistics algorithms which can handle this last aspect. Again, within this hierarchical framework a simple procedure is developed to detect and interpolate spatial orientation of the geological structures.

Another problem faced in this study is the choice of the best MP statistics simulation parameters. A number of tests were performed with different sets of parameters for the region R_{19} . They show that selecting a big search neighborhood and a high number of multi-grid levels does not always coincide with a better simulation result: often it is possible to obtain good results at moderate computational effort burden with a search neighborhood with a radius of three grid cells and five multi-grid levels. Moreover, since the continuity of the geological structures is mainly horizontal, along the direction z one multi-grid level provided satisfactory results.

In regards to the simulation path, this study, in which conditioning data are not sparse, finds that a simulation path based on the distance from the simulated point and the conditioning data is not ideal; the main structures are simulated starting from the vertical conditional sections, and this results in an artificial lack of continuity between the two planes. The best results are obtained with a random simulation path.

4.6 Conclusions

With this work, we have shown that hybrid or hierarchical models allow their users to circumvent most difficulties (non stationarity and lack of a 3D training image) encountered when dealing with multi-scale data such as the Herten site. With the advent of MP statistics, there was a temptation to think that all types of geological structures could be modeled with a single tool provided that an adequate training image is available. It is clear that progress has been made in that direction, but we also see that it is often a better strategy to combine different techniques.

One should note that splitting the domain into regions is something that is done routinely when building 3D geological models. Here the difference is that we apply the same principle at a small scale and within a single geological formation to describe different scales of internal structures within the formation.

The 3D model of the Herten site is provided as supplementary and can be used in the future in several ways.

Indeed, the 3D heterogeneity of the model is certainly close to the one which can be observed in this type of fluvio-glacial aquifers and at that scale. Therefore, the model could be used to investigate physical processes that are controlled by this heterogeneity (see e.g. 28). A whole range of numerical experiments could be done.

Another possibility is to use some statistical information that is derived from the sections (for example the number of regions per unit thickness, the variograms of the boundaries between the regions, the type of orientation structures, etc.) to extend the dimension of the 3D model. Indeed, it is possible to simulate successively an ensemble of random surfaces following the statistical parameters and the erosion rules inferred on the dataset to build an unconditional partition of a large domain. If borehole data are available, one could also simulate these surfaces conditional to the borehole data. The surfaces can be used to define separate regions where one could simulate orientation fields based on the orientation observed in the data set. The orientation fields can then be used to simulate pertinent geological structures with MP statistics (with the training images derived as depicted before).

The last point that we want to stress is that analog data sets almost always provide only 2D information. This fact clearly emphasizes the need to extend the MP statistics framework in order to allow simulating 3D domains from 2D training images and a set of reasonable assumptions in the future. The presented procedure demonstrates a successful first step in this direction.

4.7 Acknowledgments

Funding for this research was provided by Swiss National Science foundation (Grants PP002-106557 and PP002-124979) and the Swiss Confederation's Innovation Promotion Agency (CTI project No. 8836.1 PFES-ES).

Bibliography

- [1] Anderson, M. P. (1989). Hydrogeologic facies models to delineate large-scale spatial trends in glacial and glaciofluvial sediments. *Geological Society of America Bulletin*, 101(4):501–511.
- [2] Bayer, P., Huggenberger, P., Renard, P., and Comunian, A. (2010). Three-dimensional high resolution fluvio-glacial aquifer analog: Part 1: field study. Submitted to Journal of Hydrology.
- [3] Bersezio, R., Bini, A., and Giudici, M. (1999). Effects of sedimentary heterogeneity on groundwater flow in a Quaternary pro-glacial delta environment: joining facies analysis and numerical modelling. *Sedimentary Geology*, 129(3-4):327 – 344.
- [4] Bersezio, R., Giudici, M., and Mele, M. (2007). Combining sedimentological and geophysical data for high-resolution 3-D mapping of fluvial architectural elements in the Quaternary Po plain (Italy). *Sedimentary Geology*, 202(1-2):230 – 248.

-
- [5] Boucher, A. (2009). Considering complex training images with search tree partitioning. *Computers & Geosciences*, 35(6):1151–1158.
- [6] Caers, J. (2005). *Petroleum Geostatistics*. Society of Petroleum Engineers.
- [7] Carle, S. (1996). *A transition probability-based approach to geostatistical characterization of hydrostratigraphic architecture*. PhD thesis, Department of Land, Air and Water Resources, University of California Davis.
- [8] Chilès, J. and Delfiner, P. (1999). *Geostatistics: modeling spatial uncertainty*. Wiley, New York.
- [9] Chugunova, T. L. and Hu, L. Y. (2008). Multiple-point simulations constrained by continuous auxiliary data. *Mathematical Geosciences*, 40(2):133–146.
- [10] de Marsily, G., Delay, F., Goncalves, J., Renard, P., Teles, V., and Violette, S. (2005). Dealing with spatial heterogeneity. *Hydrogeology Journal*, 13(1):161–183.
- [11] de Vries, L. M., Carrera, J., Falivene, O., Gratacos, O., and Luit, J. S. (2009). Application of multiple point geostatistics to non-stationary images. *Mathematical Geosciences*, 41(1):29–42.
- [12] Guardiano, F. B. and Srivastava, R. M. (1993). Multivariate geostatistics: Beyond bivariate moments. In Soares, A., editor, *Geostatistics: Troia '92*, volume 1, pages 133–144. Kluwer, Dordrecht, The Netherlands.
- [13] Gumiaux, C., Gapais, D., and Brun, J. (2003). Geostatistics applied to best-fit interpolation of orientation data. *Tectonophysics*, 376(3-4):241–259.
- [14] Haralick, R. M. and Shapiro, L. G. (1992). *Computer and Robot Vision*, volume I. Addison-Wesley.
- [15] Heinz, J., Kleinedam, S., Teutsch, G., and Aigner, T. (2003). Heterogeneity patterns of Quaternary glaciofluvial gravel bodies (SW-Germany): application to hydrogeology. *Sedimentary Geology*, 158(1-2):1 – 23.
- [16] Hornung, J. and Aigner, T. (1999). Reservoir and aquifer characterization of fluvial architectural elements: Stubensandstein, Upper Triassic, southwest Germany. *Sedimentary Geology*, 129(3-4):215 – 280.
- [17] Hu, L. Y. and Chugunova, T. L. (2008). Multiple-point geostatistics for modeling subsurface heterogeneity: A comprehensive review. *Water Resources Research*, 44(W11413):14.

- [18] Huggenberger, P. and Aigner, T. (1999). Introduction to the special issue on aquifer-sedimentology: problems, perspectives and modern approaches. *Sedimentary Geology*, 129(3-4):179 – 186.
- [19] Huysmans, M., Peeters, L., Moermans, G., and Dassargues, A. (2008). Relating small-scale sedimentary structures and permeability in a cross-bedded aquifer. *Journal of Hydrology*, 361(1-2):41 – 51.
- [20] Jussel, P., Stauffer, F., and Dracos, T. (1994). Transport modeling in heterogeneous aquifers: 1. Statistical description and numerical generation of gravel deposits. *Water Resources Research*, 30(6):1803–1817.
- [21] Khan, D., Strebelle, S., Hanggoro, D., Willis, B., and Pyrcz, M. (2008). Non-stationary multiple point simulation to model complex deltaic deposits for flow simulation. In Ortiz, J. and Emery, X., editors, *Geostatistics Santiago 2008*, pages 217–226. Springer, Netherland.
- [22] Koltermann, C. E. and Gorelick, S. M. (1996). Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches. *Water Resources Research*, 32(9):2617–2658.
- [23] Lajaunie, C., Courrioux, G., and Manuel, L. (1997). Foliation fields and 3D cartography in geology: Principles of a method based on potential interpolation. *Mathematical Geology*, 29(4):571–584.
- [24] Liu, Y. (2006). Using the Snesim program for multiple-point statistical simulation. *Computers & Geosciences*, 32(10):1544–1563.
- [25] Maharaja, A. (2005). Stochastic modeling of the Rhine-Meuse delta using mutiple-point geostatistics. In Leuangthong, O. and Deutsch, C. V., editors, *Geostatistics Banff 2004*, volume 14 of *Mathematics and Statistics*, pages 653–659. Springer Netherlands.
- [26] Maharaja, A. (2008). Tgenerator: Object-based training image generator. *Computers & Geosciences*, 34(12):1753–1761.
- [27] Maji, R. (2005). *Conditional Stochastic Modelling of DNAPL Migration and Dissolution in a High-resolution Aquifer Analog*. PhD thesis, University of Waterloo.
- [28] Maji, R., Sudicky, E., Panday, S., and Teutsch, G. (2006). Transition probability/markov chain analyses of DNAPL source zones and plumes. *Ground Water*, 44(6):853–863.
- [29] Renard, P., Genty, A., and Stauffer, F. (2001). Laboratory determination of the full permeability tensor. *J. Geophys. Res.*, 106(B11):26443–26452.

-
- [30] Rubin, Y. and Gómez-Hernández, J. J. (1990). A Stochastic Approach to the Problem of Upscaling of Conductivity in Disordered Media: Theory and Unconditional Numerical Simulations. *Water Resources Research*, 26:691–701.
- [31] Straubhaar, J., Renard, P., Mariethoz, G., Froidevaux, R., and Besson, O. (2010). An improved parallel multiple-point algorithm using a list approach. *Mathematical Geosciences*, accepted.
- [32] Strebelle, S. (2000). *Sequential simulation drawing structure from training images*. PhD thesis, Stanford University, Stanford, CA, USA.
- [33] Weissmann, G. and Fogg, G. (1999). Multi-scale alluvial fan heterogeneity modeled with transition probability geostatistics in a sequence stratigraphic framework. *Journal of Hydrology*, 226(1-2):48–65.
- [34] Zappa, G., Bersezio, R., Felletti, F., and Giudici, M. (2006). Modeling heterogeneity of gravel-sand, braided stream, alluvial aquifers at the facies scale. *Journal of Hydrology*, 325(1-4):134 – 153.

[illegible]

Table 4.2: The geometrical parameters used to simulate with an object based method the training images required to apply the MP simulation method in the heterogeneous regions R_i . For all the objects $\lambda = 0.5$.

region	$r_{e\ min}$ [m]	$r_{e\ max}$ [m]	$r_{p\ min}$ [m]	$r_{p\ max}$ [m]
R_1	0.55	3.50	0.05	0.35
R_4	0.75	3.50	0.05	0.20
R_5	0.75	3.50	0.05	0.20
R_6	0.75	1.50	0.05	0.10
R_7	2.00	5.00	0.05	0.30
R_8	1.00	4.00	0.05	0.20
R_9	1.50	3.50	0.05	0.15
R_{10}	1.50	3.50	0.05	0.20
R_{13}	1.50	3.50	0.05	0.20
R_{15}	1.50	3.50	0.20	0.75
R_{17}	1.50	5.00	0.10	0.20
R_{19}	3.00	7.50	0.10	0.30

Table 4.3: The orientation angles used to define the five rotation zones (quantiles 20%, 40%,60% and 80%) and the five corresponding rotations (quantiles 10%,30%,50%,70% and 90%) associated to each region.

region	 min [°]	zone 0 Q10 [°]	) Q20 [°]	zone 1 Q30 [°]	) Q40 [°]	zone 2 Q50 [°]	) Q60 [°]	zone 3 Q70 [°]	) Q80 [°]	zone 4 Q90 [°]	 max [°]
R_1	-17.37	-1.27	-0.25	0.33	0.79	1.20	1.57	2.11	3.00	4.50	24.18
R_4	-15.93	-9.27	-3.12	0.00	2.19	5.48	6.53	7.59	8.38	9.35	14.94
R_5	-28.99	-16.91	-15.27	-12.19	-8.85	-4.89	0.00	0.00	5.07	10.08	16.51
R_6	-2.55	0.00	0.00	0.00	0.00	2.55	5.36	8.71	10.64	11.80	13.23
R_7	-36.87	7.88	9.37	10.20	12.56	14.23	15.62	17.18	19.41	22.28	31.88
R_8	-10.78	4.51	7.29	10.37	13.42	16.19	17.35	19.06	20.48	22.78	30.51
R_9	-6.26	-2.47	-2.33	-2.11	-0.91	0.00	0.00	1.10	3.78	7.53	18.05
R_{10}	-0.95	10.82	13.89	15.11	18.50	19.21	20.64	22.87	24.71	26.79	32.75
R_{13}	-18.55	-1.70	0.00	1.53	3.38	5.18	7.44	10.07	13.45	17.23	23.04
R_{15}	-44.54	-0.20	2.43	4.31	5.98	7.81	9.53	11.48	13.81	17.24	44.48
R_{17}	-17.69	-11.26	-9.21	-7.69	-6.50	-6.04	-5.66	-5.22	-4.33	-2.44	4.26
R_{19}	-12.56	-1.96	0.58	2.31	3.69	4.97	6.29	7.72	9.77	13.87	38.00

Chapter 5

Conclusions and outlook

Probability aggregation methods in Earth sciences

In the first part of this thesis, we investigate the properties and the features of a number of probability aggregation methods. The main motivation for that work is the multiple-point statistics (MPS) simulation of 3D blocks using 2D training images. However, this research topic has a much broader range of potential applications. For example, it can be directly used for the integration of secondary information in MPS (2; 1). More generally, it is of interest in all the scientific fields where probabilistic information coming from diverse sources must be aggregated.

The methods analyzed can be classified into two main groups. One relies on the sum for the aggregation. The corresponding methods are convex and do not have the 0/1 forcing property. The other group of methods relies on the multiplication of probabilities or odd ratios. These methods are not convex but have the 0/1 forcing properties.

When a training data set is accessible for the determination of the weighting factors, the more reliable results are obtained with the Bordley's formula and the Beta transformed linear opinion pool. When a training data set is not available, acceptable results are obtained with the Markovian type categorical predictor method and the Bordley's formula with weights equal to one.

Through case studies, we illustrate the properties of the different methods under simple to understand conditions and on two simulation studies where there is a different correlation between the sources of information. To corroborate the conclusions made on this study, other tests on different case studies could be made. For example, it would be interesting to consider a number of multiple-point statistics realizations as training data set, where a maximum likelihood principle could be used to determine the optimal weights, instead of the trial and error procedure illustrated in chapter 3. Some work could be required for the conceptualization of this aspect.

3D multiple points statistics simulation using 2D training images

In the second chapter we show that multiple-point statistics simulation can be performed in 3D even without the use of a 3D training image. The best results are obtained with a method which performs sequential 2D simulations conditioned to the data simulated during the previous simulation steps (**s2Dcd** method). The **s2Dcd** method allows to reproduce different features of the 3D images used as a reference. Of course, the information content of two or three 2D training images is not comparable with the one of a 3D training image. However, the results obtained are encouraging, especially from a computational point of view. Indeed, with the method **s2Dcd** the simulation is from two to four order of magnitude faster than using a full 3D training

image. This computational efficiency could allow to include multiple-point statistics simulation in contexts which at the moment are prohibitive for MPS with a 3D training image.

A drawback of the **s2Dcd** is that the simulation presents some artifacts. These artifacts are due to the fact that the method does not rely on a true multiple-point multi directional correlation among the facies simulated, but on the conditioning data only. Some research could deepen this topic. An idea to reduce these defeats it to mix the results obtained with the method for the first simulation steps with methods which can take into account spatial correlation along different directions. These last methods could be for example the **paBor** method or the Direct Sampling (5). The need of using the synergy among different techniques is elicited in another part of this work, summarized in the following section.

Geostatistical modeling of a three-dimensional high resolution fluvio-glacial aquifer analog

In the last chapter , we suggest that using a single modeling technique for the reconstruction of a complex geometry such as the one of the Herten aquifer analog is not sufficient. This is shown in this work where ordinary Kriging, MPS, morphological techniques and boolean simulation are used together to face the 3D simulation of this aquifer analog. The results obtained with this hierarchical simulation framework are more realistic than the ones obtained for example by transition probability/Markov chain approaches.

This idea is very general, and should be pursued by developing tools and software facilitating the integration of various modeling techniques in a single platform.

From a hydrogeological point of view, the high resolution simulations that have been obtained should be used in the future to conduct a wide range of numerical experiments. For example, using these 3D blocks the effect of heterogeneity on the interactions between ground water and surface water could be investigated in great detail.

The world wide hydrogeological parameters database

To conclude, I would like to emphasize a part of the work that was done in the early stages of the thesis and which is described in detail in appendix A: a world wide hydrogeological parameters database (<http://wwhypda.org>). I believe that more resources should be devoted to such a project. The on-line database **wwhypda** is a first attempt to provide an useful catalog of hydrogeological properties. Since it is on-line (April, 2008), more that 560 unique visitors from 60 different countries visited the database (see appendix B for more details), but for the moment the contributions from the hydrogeologist

community were few. One hypothesis which could explain this fact is that probably the critical mass of users is too small to contain an important part of users which not only consult the database, but also contribute actively by adding some measurements. For example, in the case of collaborative encyclopedia Wikipedia, 80% of the revisions are submitted by less than 7% of the user (4). This could explain the small amount of contributions submitted to wwwhyda.

Another possible reason is the web interface itself, which was designed to upload large amounts of data at once. This requires that the user downloads a spreadsheet and fills it before being able to contribute.

A more user friendly interface could facilitate the entry of data by a larger number of users. Along this line of thought, a new interface is under development (<http://dev.wwhypda.org>).

Some actions could improve and extend the usage of the database.

Advertisement and promotion: it is important to devote some resources to improve the success of wwwhyda, even if the database was already presented in a number of international conferences and by some publications (6; 3). Some actions which could improve its spreading could be for example:

- Include the use of the database in classes and for educational meetings.
- Continue the promotion in conferences and international seminars.
- Try to include in measurements protocols or in systematic measurement campaigns a work-flow and data structures in synergy with the database.
- Propose, in collaborations with international journals, the possibility for authors to include their data set in a wwwhyda related format.

Standard Markup Languages The relation with the geo-localization of the data-set could be emphasized in order to include the database in web-based GIS tools, like for example Google Earth. Another important effort should be oriented to the relationship with Standard Markup Languages like the Geographic Markup Language (GML) or the Hydrology Markup Language (HYDROML).

Other actions could be devoted to the improvement of the quality of the data and the quantity of data (for example including data set of other similar projects). Moreover, the database would be more appealing if it could include different types of statistical information, like for example variograms of training images.

Bibliography

- [1] Chugunova, T. and Hu, L. (2008a). An assessment of the tau model for integrating auxiliary information. In Ortiz, J. M. and Emery, X., editors, *Geostats 2008 - VIII International Geostatistics Congress*. Mining Engineering Department, University of Chile.
- [2] Chugunova, T. L. and Hu, L. Y. (2008b). Multiple-point simulations constrained by continuous auxiliary data. *Mathematical Geosciences*, 40(2):133–146.
- [3] Comunian, A. and Renard, P. (2009). Introducing *wwhypda* : a world-wide collaborative hydrogeological parameters database. *Hydrogeology Journal*, 17:481–489.
- [4] Javanmardi, S., Ganjisaffar, Y., Lopes, C., and Baldi, P. (2009). User contribution and trust in wikipedia. In *Proceedings of the 5th International Conference on Collaborative computing: Networking, Applications and Worksharing, Washington DC, USA*.
- [5] Mariethoz, G., Renard, P., and Straubhaar, J. (2010). The direct sampling method to perform multiple-point geostatistical simulations. *Water Resour. Res.*, 46(11):W11536–.
- [6] Renard, P. (2007). Stochastic hydrogeology: What professionals really need? *Ground Water*, 45(5):531–541.

Appendix A

Introducing *wwhypda*: a
world-wide collaborative
hydrogeological parameters
database

Abstract

Since the seminal publication of Henry Darcy's work in the 1850s, a very large number of rock property values (such as hydraulic conductivity, permeability, compressibility, porosity, etc.) has been measured and published. These data are, however, dispersed and difficult to access. To overcome this problem and to facilitate site characterization (especially stochastic), a worldwide hydrogeological parameter database (*wwhypda*) is proposed. It is an open and collaborative catalog allowing users to store and retrieve measurements. The catalog is accessible through a web interface (<http://wwhypda.org>). Presently, it provides individual values and probability density functions of the properties as a function of lithology, scale of observation, location, and geological environment.

A.1 Introduction

Since the first experiments of Henry Darcy, published in 1856, and for more than a century, experiments have been designed and conducted at different scales and all around the world to characterize not only hydraulic conductivity or intrinsic permeability, but also porosity, specific storage or compressibility of earth materials. Large-scale equivalent properties have been obtained by calibrating regional aquifer models. All these experiments and their interpretations have been published in articles, books or technical reports and in different fields such as Hydrogeology, petroleum engineering or civil engineering.

In addition, many national geological surveys are centralizing geological and groundwater data within national databases or geographical information systems (GIS). These data sets are available upon request (3; 19), or via Web interfaces—for example, the Canadian geospatial data infrastructure (CGDI), the ADES groundwater national portal in France (ade), or the Ground Water Atlas of the United States by the United States Geological Survey (usg). Similarly, GIS databases are often developed to facilitate data management within specific consulting or research projects (5; 18). From a research perspective, there are also initiatives to manage site specific datasets in a standardized format in order to ensure their long term availability and to facilitate the collaboration among various research teams (11). In this last case, the aim is to improve the efficiency and the quality of the research, rather than to initiate a public service such as more general databases provided by geological surveys.

Parallel to the acquisition of datasets of parameter values, the analysis and modelling of their spatial distribution have evolved significantly thanks to the development of geostatistical theoretical frameworks (23; 8; 16; 7; 26; 12). Therefore, in addition to the physical parameter values themselves,

corresponding statistical parameters which describe their distributions and their spatial continuities (covariance functions) have also been collected from a large number of test sites worldwide (see for example Rubin (26) for a review).

Despite these efforts, in most sites where a hydrogeological study is to be conducted, there is still a limited amount of data available to characterize the site. This is especially true and poses a problem when one considers the first stages of a new study. Anderson et al. (4) estimated that, in most cases, less than 1% of the total volume of an aquifer is directly sampled. Even at the most extensively studied sites, the volume of the aquifer that is directly sampled remains small. In their detailed study of a braided alluvial system, Anderson et al. (4) estimated this sampled volume to be around 20-40% of the total volume if they include the geophysical measurements.

In addition, in most practical situations it is necessary to cope with uncertainties because of the lack of data. This is why the use stochastic techniques as one component of a general framework to manage hydrogeological projects is widely recognized among researchers (15), even if the application of stochastic techniques in daily practice is still limited. The reasons for such limitations are diverse and include many practical and theoretical aspects (9; 32; 25). However, several authors have suggested that one step to facilitate the use of stochastic techniques would be to provide a catalog of statistical properties of the different parameters and of the structures of heterogeneity that are typical of different geological environments (9; 12).

Such a catalog would be useful both in a stochastic and deterministic framework. It could answer questions such as: What is the order of magnitude of intrinsic permeabilities in a fine sandstone? What is the most probable value for the porosity of pillow lavas in an ophiolitic environment? Or what should be a safe value for the variance of the hydraulic conductivity in fractured sandstone to design a pump-and-treat system? It could provide parameter distributions to use as priors or plausibility terms for inverse problem solving (29; 2)). It could also provide training images describing the typical spatial arrangement of different rock types in a given geological environment that could be used as input for multiple-point simulation algorithm (27; 20)).

The aim of this article is to describe a first attempt at providing such a “world-wide hydrogeological parameters database” (*wwhypda*). It is based on open source technology to allow a wide distribution and free access via Internet. The idea is not to offer a substitute for local investigation but a well-documented complement of information when there is a lack of data. In addition, an important concept is that such a catalog needs to be supported by a broad community and should therefore be based on a collaborative web site in which users can provide data. This article presents the main features and the structure of *wwhypda*.

A.2 Main *wwhypda* features

wwhypda is designed to store values of the most important properties of earth materials, and to relate them either to a specific material or to a hydrogeological environment. These values can then be post-processed automatically to extract statistical moments or full probability density functions (Fig. A.1a). To allow such type of queries, the database contains an extensible catalog of typical rock types and typical geological environments.

Because the aim of *wwhypda* is to provide essentially statistical distributions of rock properties that can be borrowed from one site to be used in another site, the exact location of the measurement is not crucial. Therefore, *wwhypda* is not based on a GIS, even if optional information about the measurement site location can be stored. For the same reason, time dependent data such as time series of piezometric heads are not considered in the present work.

In terms of its usage, *wwhypda* is available via a Web interface. Anybody can consult the database, download it, or add data, under the terms of an open licence which ensures that *wwhypda* is and will be freely available. More precisely users are free to consult, copy, distribute, display, and make use of the database, provided that a proper reference is made to the *wwhypda* project (detailed license terms are provided online).

Finally, a key feature of *wwhypda* is that it provides an interface which allows users to contribute. The only requirement is a registration step, which allows user contributions to be recognized and tracked. The counterpart of this open policy is that even if a validation procedure is used to control the quality of the data, it is not possible to exclude erroneous measurements, which will become part of the database. This is why the database is provided without any warranty.

A.3 Content organization, methods and technologies

The entity-relationship schema (6) was adopted to design the conceptual model of the database. Figure A.2 is a simplified graphical representation of *wwhypda*'s model, built with a unified modeling language diagram (UML). It shows the entities which compose the database, their relationships and their main attributes.

The basic entity composing the database is an ensemble of values of measurements of hydrogeological parameters (Fig.A.3). Only geometry-independent parameters are considered. Information about a single measurement is stored in the entity **Measure** whose main attributes are:

value the measured value of the parameter.

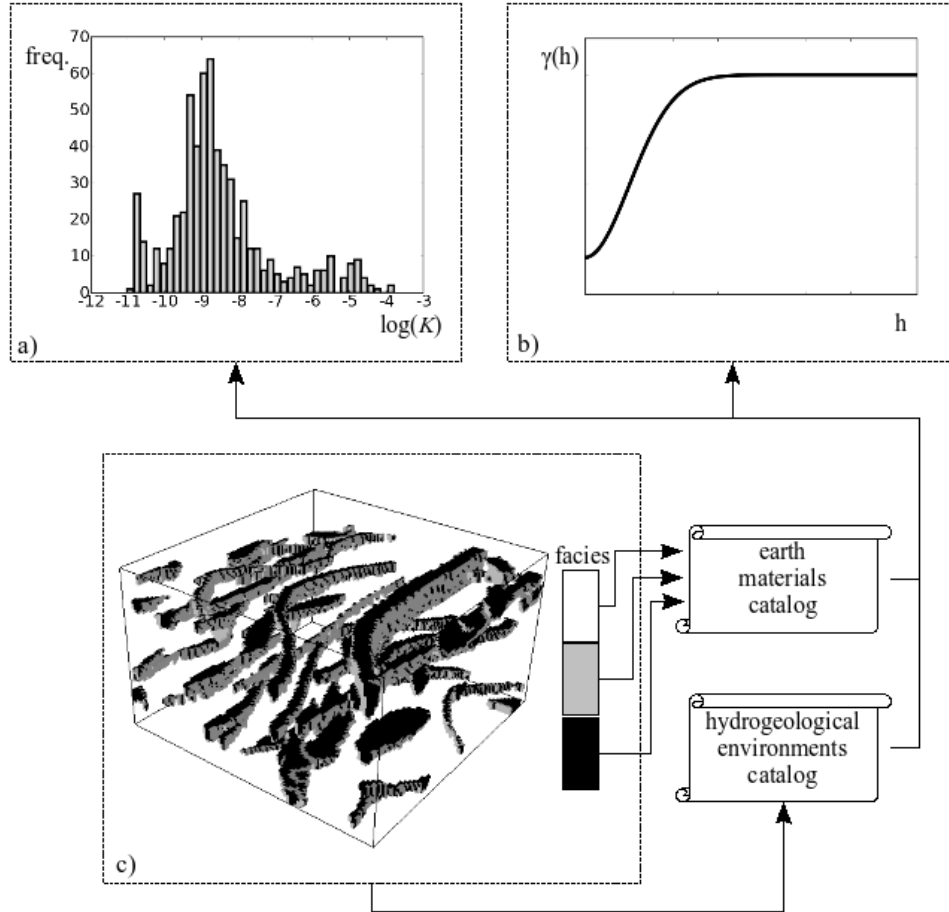


Figure A.1: Typical examples of *wwhypda* content and output: **a** histogram of the log-hydraulic conductivity (K) for one facies. **b** Variogram (γ) describing the correlation of the properties versus the distance (h). **c** Three-dimensional training image of the internal architecture of a fluvial plain obtained from a fluvsim simulation (13).

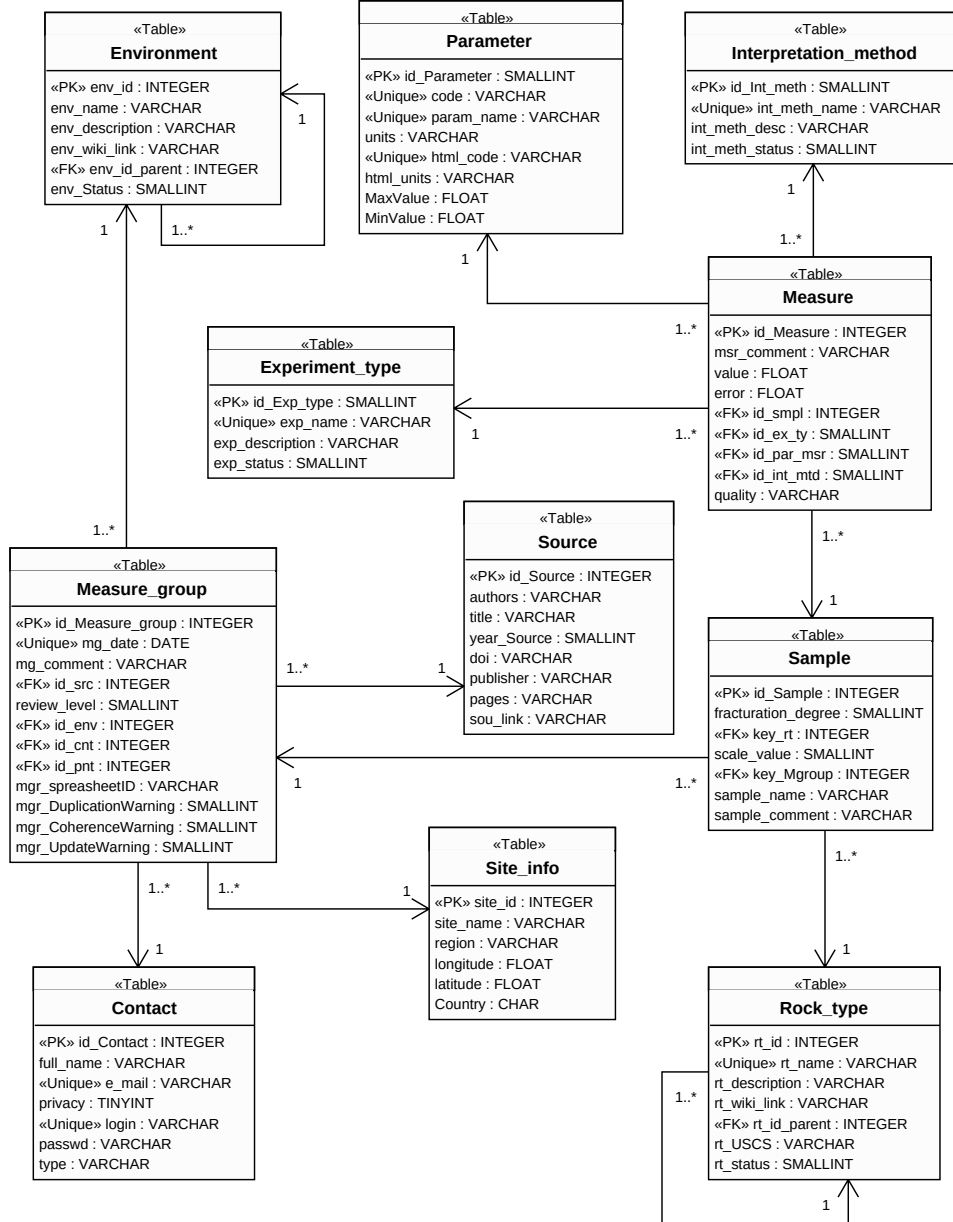


Figure A.2: Simplified conceptual schema of *wwhypda* using an *UML diagram* (the 1 and the 1..* represent the cardinality of the relationships between tables). The actual schema contains more than 28 *SQL* tables

rock type name and LINK	n_e effective porosity	K hydraulic conductivity	α_L longitudinal dispersivity	k permeability	n porosity	a rock compressibility	S_s specific storage	S_y specific yield	α_T transversal dispersivity	all param.
Basalt	0	23	0	0	0	0	0	0	0	23
Chalk	0	1097	0	0	0	0	0	9	0	1106
Clay	0	17	16	0	0	0	0	0	0	33
Clay, kaolinite	0	17	0	0	0	0	0	0	0	17
Clay, sandy	0	0	5	0	0	0	0	0	0	5
Coal	0	225	0	0	0	0	0	0	0	225
Coal, spoil	0	40	0	0	0	0	0	0	0	40
Dolomite	0	5	0	0	0	0	0	0	0	5
Gravel	0	11	3	1	20	0	0	0	1	36
Gravel, clayey	0	0	0	1	20	0	0	0	0	21
Gravel, sandy	0	0	3	0	0	0	0	0	1	4
Lignite	0	14	0	0	0	0	0	0	0	14
Limestone	0	599	0	0	530	0	1	0	0	1130
Loam	0	0	93	0	0	0	0	0	0	93
Marl	0	24	0	0	24	0	0	0	0	48

Figure A.3: Part of the table showing rock type versus parameter, allowing for rapid search of the database.

error if the value is obtained as mean value of many measurements, this parameter will represent the associated uncertainty, otherwise it should be used to store an estimate of the uncertainty related to the measurement technique and interpretation.

quality this information reports an evaluation of the data-quality given by the user. It can be set to low, medium or high.

Every measure is related to a description of the experimental procedure (such as a pumping test, a slug test, etc.) as well as the interpretation method that was used to analyze the data. This information is stored in lists that can be extended when required.

Measurements of different properties can be made on the same sample. The database structure requires the user to link each measurement with a sample. This allows, for example, the user to study the correlation between parameters. The entity **Sample** has the following main attributes:

scale_value the majority of the parameters included in the database have statistical distributions that are scale dependent (22; 17; 24). The scale of observation is described by four classes: micro (for lengths $< 1m$), macro ($1m < \text{lengths} < 10m$), meso ($10m < \text{lengths} < 100m$) and mega (for lengths $> 100m$).

fracturation_degree this is a characteristic which strongly affects the values of the measured properties.

To minimize redundancy, part of the information related to the sample and to the measure is stored in the entity **Measure_group**. It contains the information that is shared by a group of measurements such as:

review_level this attribute depicts the level of reviews undergone by a group of measurements. Values among high (such as a paper published in a peer reviewed journal), low (such as an internal report or a MSc thesis), medium (such as a PhD thesis or a reviewed technical report) and not reviewed are allowed.

The entity **Measure_group** also contains a link to the entity **Contact** in order to identify the contributor.

Every sample is related to an earth material (rock type) or to a hydro-geological environment. To facilitate this task, the entities **Rock_type** and **Environment** contain typical sets of earth materials and environments. In practice, these two database components represent two catalogs organized and presented as tree structures; contributors are able to add their own **Rock_type** or their own **Environment** as branches of the proposed tree classification. The definitions of the rock types are specified by a brief description and, when possible, are followed by a link to the free encyclopedia Wikipedia, which provides a more detailed description. Finally, geographic information is stored in the entity **Site_info** (**site_name**, **region**, **longitude** and **latitude**), in order to locate a measure group.

Open source technology is used to implement the conceptual schema in a relational database with Sun Microsystems' MySQL, and to develop the web interface with the content management system Plone, based on the content management framework Zope.

To conclude this section, the structure of *wwhypda* is designed to be flexible: internal catalogs of earth materials, environments and measurement methods are open to users' contributions. Other database characteristics such as the parameters list, are also modifiable if required.

A.4 Data sources

wwhypda should progressively assimilate data from diverse sources. On one hand, some efforts have already been made by various authors to collect data for specific rock types or parameters, and those existing data sets can be integrated rather rapidly. For example, Vanderborcht and Vereecken (31) contributed by providing all dispersivity values they collected for their review article. Their database was directly imported into *wwhypda*. The second important data sources are the values published in scientific journals. This information is currently difficult to access because it is dispersed

in many articles and many journals. Progressively, it should be added to *wwhypda* to allow easy access. Finally, technical reports constitute another very large source of information. For example, NAGRA (the Swiss agency in charge of nuclear waste management) has founded extensive research and published reports in which they provide permeability and porosity measurements (e.g. Keller et al. (21)). Many of such type of reports are public and can be integrated in the database. Another incredible amount of measurements has been collected by oil companies. However, these data are usually confidential and are not accessible. At the moment, *wwhypda* contains more than 20,000 measurements of properties of about 25 earth materials. This is still very little and needs to be extended with contributions from the whole Hydrogeology community.

A.5 Data quality

The evaluation of data quality is of extreme importance. In *wwhypda*, a first screening consists of rejecting data that are not acceptable from a physical point of view (porosity must be between 0 and 1, permeability must be positive, etc). Once these trivial checks are satisfied, the quality of the data is described with three criteria stored as three different attributes (**quality**, **review_level** and **error**) linked either to the group of measures or to the measure itself.

The person entering the data has to provide subjective and, if possible, quantitative information about the data quality. Three attributes are used: the **review_level**, the **error**, and the **quality**. The **review_level** is attached with the source of the measurement. It indicates whether the source has been reviewed thoroughly such as in an international journal or not (the four states described previously are available: not reviewed, low, medium and high). The **error** is an optional attribute (because often it is not estimated) corresponding to the error range or uncertainty related to the measurement itself (30). Finally, the **quality** attribute (which can take only the values high, medium or low) allows the user to provide a subjective judgment based on his/her knowledge of the field conditions during the data acquisition, as well as the type of interpretations that were used for the estimation of the parameter values.

In addition, it is planned that regular scan of the database will be made by the administrator in order to manually check the plausibility of the values. A tag will be used to identify measurements which have been controlled and will allow users to select only those measurements. An interface to let users correct themselves or indicate dubious measurements has still to be developed. At the moment, it is recommended that users send a message to the administrator when potentially incorrect information has been identified.

A.6 Consultation example

It is expected that most users will be more interested in consulting the database than contributing to it. Therefore the design of the web interface has been made to allow a very rapid and simple access to the data. Different search tools are available. They are freely available without any registration. The fastest is an interactive table that shows the number of measurements currently available for each rock type and for each parameter A.3. On this table, one can see very rapidly if *wwhypda* contains the measurements that one is looking for. Suppose for example, that the user is conducting a project on an alluvial aquifer and would like to check the typical values of hydraulic conductivity of a gravely sand. For that purpose, he or she can check the table and will find that for the rock type *Sand, gravely* there are 458 hydraulic conductivity measurements available. Clicking on the number brings one to a search interface in which the main attributes were automatically set. In that window, one can specify additional search criteria and type of results that are to be obtained. For example, the user can define that he or she needs to get basic statistics for those values. He or she can also include criteria such as the scale of the measurement, or the type of experimental method used to obtain the parameter value, or a certain quality level. The result of the search is then displayed in a page like the one showed in A.4. The user in this case has asked to access the raw data, including (when available) a link to the original publication. He or she can, as well, download all the raw data corresponding to the result of the query as a csv file that can be used for further statistical treatment.

A.7 Contribution example

An important feature of *wwhypda* is that it allows for contribution to the database. The procedure for entering property values is accomplished essentially via a spreadsheet (Fig. A.5) and can be summarized as follows:

- The user downloads the contribution spreadsheet from the web site either in OpenOffice Calc or MS Excel format.
- The user fills the spreadsheet off line. The web site and the spreadsheet itself contain information about how to do this. The spreadsheet (Fig. A.5) contains drop-down lists from the internal catalogues such as earth materials list or type of parameters. The spreadsheet contains generic information about a set of values (such as the site description or the source of the data), data related to the samples (rock type, degree of fracturation), and then property values with their attributes. In this way, different parameters can be linked together and relations between parameters can be studied.

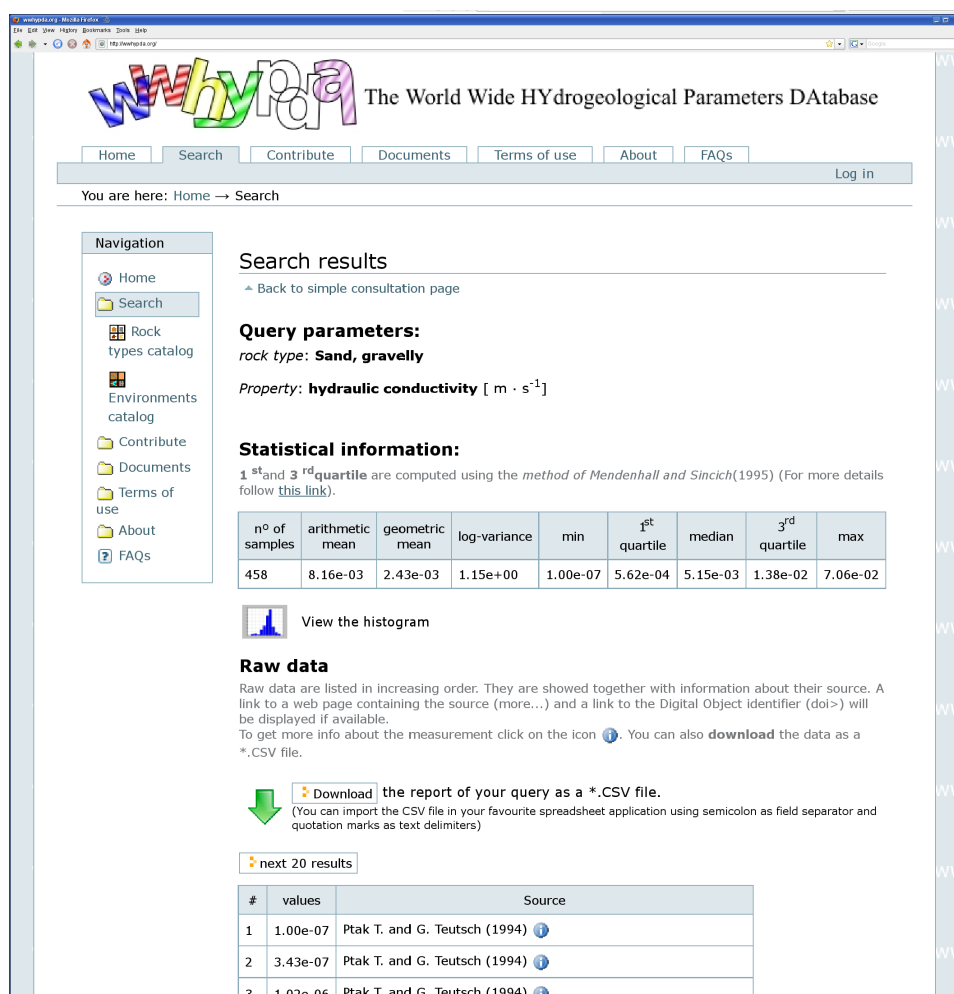


Figure A.4: Screen shot of the results page

- Finally, the user has to upload the spreadsheet on the web site. This last action is available only for registered users. The spreadsheet is then checked by an automated procedure; if all checks are successful, the information is inserted into *wwhypda* and available for consultation.

A.8 Discussion

Building a database that combines standards and flexibility requires some restrictions and simplifications to be adopted. The first difficulty is the choice of a classification system for the rock types and the hydrogeological environments. There are two problems. The first is that geologists, hydrogeologists, civil engineers, and environmental engineers are not using a unique and common classification. To make matters even worse, the classifications often differ from one country to another: fine sand may not have the same granulometry in France and in Great Britain. Which one should be adopted? The second problem is technological: How to represent the classification in a simple and efficient way? On this topic, Davenport et al. (10) and Struik et al. (28) proposed interesting solutions in order to facilitate digital geological mapping. However, their solution was too sophisticated to be implemented within the framework of this project.

Among the classifications, the challenge is to remain simple and flexible. A classification such as the one proposed by the British Geological Survey (BGS) is, for example, too detailed for *wwhypda* needs at the moment (1). But even if the classification needs to be simple, it should allow coding an accurate geological description when it is available. Thus, it was decided to store the classification in a tree. The base elements correspond to the most common families of rocks (such as basalt, chalk, etc.). The sub-elements correspond to different levels of refinement in the classification (for example one branch of the tree is Clay, Organic clay, High plasticity organic clay). Each element or sub-element contains a brief definition. When available, a symbolic description made with the Unified Soil Classification System (AST) is provided. Such a tree allows querying the database at different levels, depending on the amount of data available. For example, it may not be possible to find in the database a sufficient number of measurements for the specific storage for ferruginous meta-sandstone, but there may be enough data at a broader level to characterize the meta-sandstone or sandstone. Furthermore, when a rock-type is not available, the user can propose to add it in the tree structure.

This approach seems to be an affordable compromise at the present state of the project. However it also has limitations just like any other classification system. For example, rocks that are a mixture of different base classes have to be derived only from one parent. An equalmixture of sand and clay for example will have to be linked either to the general clay class or sand class,

but not to both. Additional characterization criteria such as the depth of sampling, are not accounted for in the database for the moment.

At present, the main focus of the work is related to data acquisition and interface development for the hydrogeological properties. However, the database is designed to store spatial statistics too such as variograms (Fig. A.1b) and typical two-dimensional and three-dimensional images of lithofacies distributions (Fig. A.1d), which could be used as training images for multiple-point statistics simulations (27). These images are digitally coded representations of facies architectures for a given geological environment. The coding corresponds to rock types which are also available in the database. In this way, it is possible to obtain statistical information related to the geometry from the digital image and to retrieve statistical information related to the physical parameters themselves within the different architectural elements. The combination of these two aspects of the database could allow, for example, simulating lithofacies distributions for a given site by combining the local borehole observations with the statistical parameters extracted from *wwhypda*. In a second step, the grid cells corresponding to the different lithologies could be populated by parameters values. Further work is now required to develop the web interface and obtain training images showing the internal structure of a variety of typical hydrogeological environments that could be stored and distributed in *wwhypda*.

Finally, it is important to stress that the greatest challenge lies not in the technical aspects of the database, but in the community involvement to enrich the database. Within the project, the database is fed with an initial data set that will progressively increase; however, this is not sufficient. *wwhypda* can really become a useful worldwide reference if and only if it receives contributions from a very large community. The recent development of projects such as *Wikipedia* and other intensive user collaborations on the Internet indicates that community involvement is feasible.

A.9 Conclusion

wwhypda is certainly a utopia; however, 150 years after the publication of the seminal work of Henry Darcy, the authors think that the time has now come to follow the spirit of people like Diderot and D'Alembert. This is why the hydrogeological community is invited to participate in the creation of a world-wide catalog of hydrogeological properties and hydrogeological environment architectures freely accessible through the Internet. Quoting Diderot (14) *“This is a work that cannot be completed except by a society of men of letters and skilled workmen, each working separately on his own part, but all bound together solely by their zeal for the best interests of the human race and a feeling of mutual good will”*.

14					
15					
16		site_name	Horkheimel Insel site		
17		region	North of Stuttgart, Neckar Valley		
18		country_name	Germany		
19		longitude			
20		latitude			
21					
22					
23					Ge
24		env_name	Fluvial environment		
25		review_level	high (such as a paper published in peer reviewed journal)		
26					
27					
28					Source of information
29		authors	Ptak T. and G.Teutsch		
30		title	A comparison of investigation methods for the prediction of flow and t		
31		source	Transport and Reactive Processes in Aquifers, Dracos & Stauffer, Balk		
32		year_Source	1994		
33		pages	157-164		
34		doi			
35		source_link			
36					
37					
38					Measurement "Sample" (the bc
39			A name to identify the sample in the Measurements table.	the name of the rock type	the scale to which the measurements are made
40		#	smpI_name	rock_type	scale
41		1	sample01	Sand, gravelly	micro (lengths < 1 m)
42		2	sample02	Sand, gravelly	micro (lengths < 1 m)
43		3	sample03	Sand, gravelly	micro (lengths < 1 m)
44		4	sample04	Sand, gravelly	micro (lengths < 1 m)

Figure A.5: Screen shot of a part of the contribution spreadsheet.

Bibliography

- [ade] Ades groundwater national portal <http://www.adeseaufrance.fr>.
- [AST] *Classification of soils for engineering purposes: annual book of ASTM standards*. American Society for Testing and Materials.
- [usg] Ground water atlas of the unites states
<http://pubs.usgs.gov/ha/ha730/gwa.html>. *Unites States Geological Survey, Reston, VA*.
- [1] (2008). The bgs rock classification scheme
<http://www.bgs.ac.uk/bgsrscs/home.html>.
- [2] Alcolea, A., Carrera, J., and Medina, A. (2006). Pilot points method incorporating prior information for solving the groundwater flow inverse problem. *Advances in Water Resources*, 29(11):1678 – 1689.
- [3] Allen, D., Brewerton, L., Coleby, L., Gibbs, B., Lewis, M., Mac Donald, A., Wagstaff, S., and Williams, A. (1997). The physical properties of major aquifers in england and wales. Technical report, British Geological Survey, Keyworth, UK.
- [4] Anderson, M. P., Aiken, J. S., Webb, E. K., and Mickelson, D. M. (1999). Sedimentology and hydrogeology of two braided stream deposits. *Sedimentary Geology*, 129(3-4):187 – 199.
- [5] Barazzuoli, P., Bouzelboudjen, M., Cucini, S., Kiraly, L., Menicori, P., and Salleolini, M. (1999). Olocenic alluvial aquifer of the River Cornia coastal plain (southern Tuscany, Italy): database design for groundwater management. *Environmental Geology*, 39(2):123–143.
- [CGDI] CGDI. The canadian geospatial infrastructure
<http://www.geoconnections.org>.
- [6] Chen, P. P. (1976). The entity-relationship model—toward a unified view of data. *ACM Trans. Database Syst.*, 1(1):9–36.
- [7] Chilès, J. and Delfiner, P. (1999). *Geostatistics: modeling spatial uncertainty*. Wiley, New York.
- [8] Dagan, G. (1989). *Flow and transport in porous formations*. Springer-Verlag.
- [9] Dagan, G. (2002). An overview of stochastic modeling of groundwater flow and transport: From theory to applications. *Eos Trans. AGU*, 83(53):621.

- [10] Davenport, P., Boisvert, E., Quat, M., Okulitch, A., Brodaric, B., Colman-Sadd, S., Nolan, L., Struik, B., MacIntyre, D., Tzeng, P., Scott, D., Gilbert, C., Abbott, G., Bassan, A., Journeay, M., Francis, J., and Houlihan, T. (2002). A scalable, digital map database of bedrock geology for Canada: a progress report. Technical Report 02-370, US Geological Survey Open-File.
- [11] de Dreuzy, J.-R., Bodin, J., Le Grand, H., Davy, P., Boulanger, D., Battais, A., Bour, O., Gouze, P., and Porel, G. (2006). General database for ground water site information. *Ground Water*, 44(5):743–748.
- [12] de Marsily, G., Delay, F., Goncalves, J., Renard, P., Teles, V., and Violette, S. (2005). Dealing with spatial heterogeneity. *Hydrogeology Journal*, 13(1):161–183.
- [13] Deutsch, C. and Wang, L. (1996). Hierarchical object-based stochastic modeling of fluvial reservoirs. *Mathematical Geology*, 28(7):857–880.
- [14] Diderot, D. (1755). *Encyclopédie [Encyclopedia]*. Hackett, Indianapolis, IN, USA. Extracted from English edition: Diderot D (2001) Rameau's nephew and other works. (translated by Barzun J, Bowen RH).
- [15] Freeze, A. R., Massmann, J., Smith, L., Sperling, T., and James, B. (1990). Hydrogeological decision analysis: 1. a framework. *Ground Water*, 28(5).
- [16] Gelhar, L. (1993). *Stochastic Subsurface Hydrology*. Prentice-Hall inc., Englewood, Cliffs, NJ 07632 (USA).
- [17] Gelhar, L., Welty, C., and Rehfeldt, K. (1992). A critical-review of data on field-scale dispersion in aquifers. *Water Resources Research*, 28(7):1955–1974.
- [18] Gogu, R., Carabin, G., Hallet, V., Peters, V., and Dassargues, A. (2001). GIS-based hydrogeological databases and groundwater modelling. *Hydrogeology Journal*, 9(6):555–569.
- [19] Jones, H., Morris, B., Cheney, C., Brewerton, L., Merrin, P., Lewis, M., MacDonald, A., Coleby, L., Talbot, J., McKenzie, A., Bird, M., Cunningham, J., and Robinson, V. (2000). The physical properties of minor aquifers in England and Wales. Technical report, British Geological Survey.
- [20] Journel, A. and Zhang, T. (2006). The necessity of a multiple-point prior model. *Mathematical Geology*, 38(5):591–610.
- [21] Keller, B., Bläsi, H., Plat, N., Mozley, P., and Matter (1990). Sedimentaere architektur der distalen unteren süßwassermolasse und ihre beziehung zur diagenese und den petrophysikalischen. eigenschaften am

- beispiel der bohrungen langenthal. [sedimentary architecture of distal lower freshwater molasse in connection with its diagenesis and petrophysical properties, exemplified by the lagenthal boreholes]. Technical report, NAGRA.
- [22] Kiraly, L. (1976). Rapport sur l'état actuel des connaissances dans le domaine des caractères physiques des roches karstiques. *Hydrogeology of karstic terrains*, B(3):52–67.
- [23] Matheron, G. (1962). *Traité de géostatistique appliquée [Treatise of applied geostatistics]*. Technip, Paris.
- [24] Neuman, S. and Di Federico, V. (2003). Multifaceted nature of hydrogeologic scaling and its interpretation. *Reviews of Geophysics*, 41(3).
- [25] Renard, P. (2007). Stochastic hydrogeology: What professionals really need? *Ground Water*, 45(5):531–541.
- [26] Rubin, Y. (2003). *Applied stochastic hydrogeology*. Oxford University Press, New York.
- [27] Strebelle, S. (2002). Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, 34:1–21.
- [28] Struik, L., Quat, M., Davenport, P., and AV, O. (2002). A preliminary scheme for multihierarchical rock classification for use with thematic computer-based query systems. Technical report, Geological Survey of Canada, Ottawa.
- [29] Tarantola, A. (2005). *Inverse Problem Theory*. SIAM.
- [30] Taylor, B. and Kuyatt, E. (1994). *Guidelines for evaluating and expressing the uncertainty of NIST measurement results*. NIST. United States Department of Commerce Technology Administration, National Institute of Standards and Technology, Gaithersburg, MD, USA.
- [31] Vanderborght, J. and Vereecken, H. (2007). Review of Dispersivities for Transport Modeling in Soils. *Vadose Zone J*, 6(1):29–52.
- [32] Zhang, Y. and Zhang, D. (2004). Forum: The state of stochastic hydrology. *Stochastic Environmental Research and Risk Assessment*, 18(4):265.

Appendix B

Some statistics about the visits
to *wwhypda*

The Google Analytics web service (<http://www.google.com/analytics/>) provides statistics which can be useful to evaluate the success of a web site. Table B.1 reports some of them, while fig.B.1 shows from where the wwhypda was visited. The data shown was collected from 01/04/2008 to 25/11/2010.

visits	1,139
absolute unique visitors	564
page views	12,334
average page views	10.83
time on site	00:07:15
bounce rate	29.85%
new visits	48.38%

Table B.1: Some statistics about the web visits to wwhypda (source: Google Analytics): the total number of visits, the number of absolute unique visitors, the total number of viewed pages, the average number of pages viewed during a visit, the average duration of a visit, the percentage of single page visits, the percentage of visits by people who had never visited the site before.

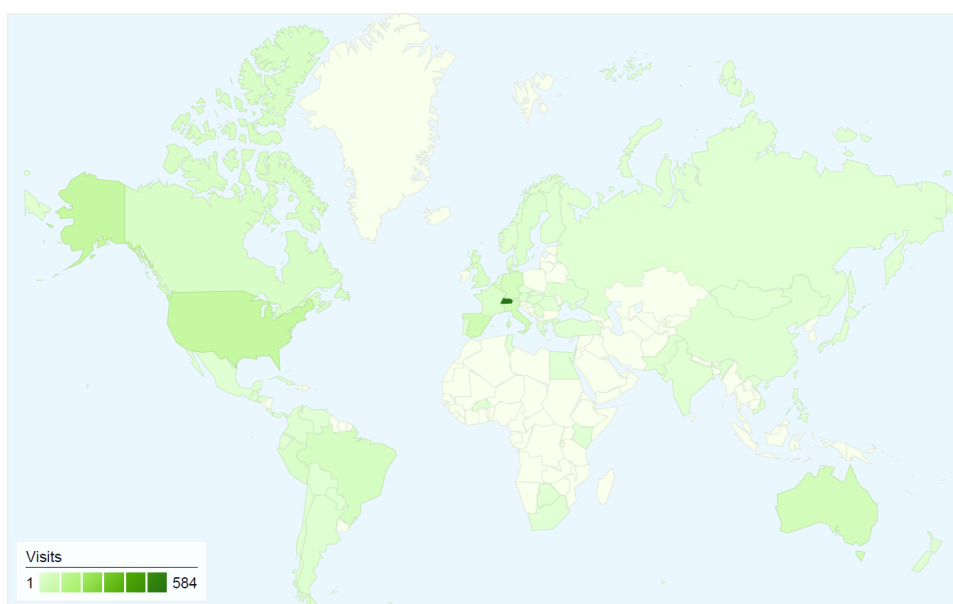


Figure B.1: wwhypda has been visited by people from 60 different countries.

Appendix C

Three-dimensional high
resolution fluvio-glacial aquifer
analog: Part 1: field study

Three-dimensional high resolution fluvio-glacial aquifer analog: Part 1: field study

P. Bayer^{a,*}, P. Huggenberger^b, P. Renard^c, A. Comunian^c

^a *University of Tübingen, Sigwartstrasse 10, 72076 Tübingen, Germany*

Present address: ETH Zürich, Engineering Geology, Sonneggstrasse 5, 8092 Zürich, Switzerland

^b *University of Basel, Geological Institute, Bernoullistrasse 32, 4056 Basel, Switzerland*

^c *University of Neuchâtel, Centre of Hydrogeology and Geothermics (CHYN), Rue Emile-Argand 11
CP 158, 2009 Neuchâtel; Switzerland*

Abstract

In recent years, analog studies have contributed to an improved understanding of those geological structures that control flow and transport in natural aquifers. Special attention has been drawn to unconsolidated sedimentary deposits, which host many of the important shallow aquifers in Central Europe. Among these are relatively young fluvio-glacial and fluvial sediments in the Rhine basin. Gravel pits found in excavation show excellent sections of the sedimentary sequence and thus offer direct insight into the structural and textural composition of the subsoil. A major problem is that usually analogs are obtained from single outcrop walls, which only deliver a cross-sectional profile but hardly capture the true three-dimensional conditions. This paper describes an approach to also inspect the third dimension: by mapping during the ongoing excavation it is possible to obtain a three-dimensional representation of a local field of observation within a short period of time. A detailed description of this “Herten case” is presented and the findings from sedimentological, hydrogeological and geophysical analyses are compared. A medium-scale (16m x 7m x 10m) gravel body is studied that delivered a high-resolution data set of lithofacies, hydrofacies and ground penetrating radar (GPR) profiles. In Part 2 of the companion paper (Comunian et al., 2010), a procedure is presented that captures the multiple scales to geostatistically reproduce the structure of such gravel sediments.

Keywords: aquifer analog, unconsolidated sediment, digital mapping, ground penetrating radar (GPR), hydrofacies, heterogeneity

**Corresponding author*

Email addresses: bayer@erdw.ethz.ch (P. Bayer), peter.huggenberger@unibas.ch (P. Huggenberger), alessandro.comunian@unine.ch (A. Comunian), philippe.renard@unine.ch (P. Renard)

1 Introduction

Sedimentary aquifers in river valleys are often built up by non-uniform sequences of layers, cross beddings and deposits that are gradually changing. Such features can hardly be captured by simple interpolation between point measurements or standard geostatistical methods (Anderson 1989, Jussel et al. 1994, Weissmann and Fogg 1999, Eaton 2006, Chen et al. 2010). Even advanced noninvasive techniques such as modern tomographical geophysical examination can only approximate subsurface structures (e.g. Becht et al. 2006, Dietrich et al. 2008, Brauchler et al. 2010) and often delivers non-unique results. Site investigation thus is ideally complemented by process-based geological analysis and characterization of deposition environments, as is common also in reservoir engineering. Due to the complexity of the braided river systems, there is still a lack of combinations of sedimentological, hydrogeological and geophysical analyses of different sites that portray the character and the bandwidth of heterogeneity scales in this type of deposits. Controversies have arisen mainly as a result of the difficulty to find simple relationships between surface morphology to be observed in modern braided river systems and the geometry of architectural elements of ancient braided river deposits in vertical profiles (Huggenberger and Regli 2003). This is mainly due to differences in preservation potential of architectural elements, which is governed by morphologic elevation in the floodplain, lateral and longitudinal mobility of depositional elements and aggradation rates. The interplay of these factors results in a large number possibilities for different characters of aquifer heterogeneities. As a consequence it would be important to find ways to catch the essence of such a deposit with respect to heterogeneity characterization at different scales, which is a main topic of the field investigation methods presented in this study.

Sedimentological analysis and interpretation can offer a systematic framework for reconstruction of architectural elements and their spatial continuities (Scheibe and Freyberg 1995, Fraser and Davies 1998, Heinz et al. 2003, Ezzy et al. 2006). Utilizing this knowledge is fundamental to achieve a geologically plausible description of heterogeneity, to delineate ranges of high-conductivity zones and for consistent re-interpretation of geophysical data (Webb and Anderson 1996, Stanford and

Ashley 1998, Whittaker and Teutsch 1999, Asprion and Aigner 1999, Hyndman and Tronicke 2005, Goutaland et al. 2008, Iversen et al. 2008). Inclusion of such information in hydrogeological model calibration prevents non-realistic realizations of spatial hydraulic parameter distribution and exploits complementary knowledge to reduce ill-posedness of (geostatistical) inversion problems (Chen and Rubin 2003, Harp et al. 2008, Maier et al. 2009, Dafflon et al. 2010). Even in case of significant data scarcity, a conceivable portrayal of the subsurface, which would hardly be possible based on only a few measurements, is feasible.

For bringing together sedimentological and hydrogeological characterization, aquifer analog studies have become popular, particularly in the last decade: vertical outcrops of sedimentary bodies are inspected in detail to obtain a model imitate of the real reservoirs that is as exact as possible (e.g. Smith 1989, Anderson et al. 1999, Huggenberger and Aigner 1999, Hornung and Aigner 1999, Willis and White 2000, Heinz et al. 2003, Bersezio et al. 2007, Huysmanns et al. 2008). Such accuracy means a considerable workload, because sedimentary facies may vary on the centimeter scale. Facies types have to be categorized with respect to sedimentary and hydrogeological criteria. Lithofacies reflect well defined deposition events or environments, whereas hydrofacies stand for quasi-homogeneous subunits that can be characterized by single representative hydraulic and/or hydrochemical parameters (Anderson 1989, Huggenberger and Aigner 1999, Klingbeil et al. 1999, Ouellon et al. 2008). These parameters have to be determined by appropriate field and lab methods.

The present work reports a comprehensive three dimensional aquifer analog study on an unconsolidated braided river deposit at the scale of tens of meters. This scale is of importance, for instance at the scale of smaller to medium scale contaminant sources of spills and proximal plumes downstream. The previously unpublished mapping (Bayer 2000) delivered a detailed dataset that may be used in realistic 3D hydrogeological modeling studies. A gravel pit found in excavation showed excellent sections of the sedimentary sequence and thus offered direct insight into the structural and textural composition of the subsoil. Study of lithofacies and hydrofacies distribution is supplemented by ground penetrating radar (GPR) measurements that were conducted before mining.

Meanwhile, several studies have been using individual sections of the obtained data set for demonstration of geophysical forward (Kowalski et al. 2001) or inverse (Hu et al. 2009) modeling, for hydraulic or transport modeling (Maier et al. 2005, Werth et al. 2005, Maji et al. 2006), and for sedimentological analysis (Heinz et al. 2003). In parallel, many authors have started to study the impact of complex geological structures on flow and transport processes (e.g. Sánchez-Vila et al. 1996, Zinn and Harvey 2003, Kerrou et al. 2008). Many of these studies are only based on conceptual model of heterogeneity and therefore there is still a need to investigate those processes on real geological structures. Most often such data sets are unavailable in 3D, and the corresponding studies may be biased by the choice of the heterogeneity models. On the other hand modeling of heterogeneity (de Marsily et al. 2004) is often a step required to compensate the lack of direct observations of the underground. Once again, real exhaustive data sets are missing to test the quality of various methods. To address this growing interest, a detailed description of the sedimentary block is given in this paper and the updated full hydrofacies data set is provided as downloadable supplement. The companion paper (Comunian et al. 2010) presents how such a data set can be used to obtain a full 3D model using classical and multiple point geostatistical techniques.

2 Architecture and facies types

2.1 Sedimentological setting and concept

The study site is located in the upper Rhine valley in Southern Germany close to the Swiss border near Basel and 500 m west of the town of Herten (see Figure 1, Kieswerk Rheinfelden, R3403877, H5268837; 276m a.s.l.). The gravel deposits originate from glacial outwash 50 km downstream of the front of the Rhine-glacier. An accumulation phase at the glacier front reached its maximum 18'000 a bp and down river, in the Basel region about 12000bp (see Kock et al. 2009). Several degradation phases followed and formed series of river terraces. The sedimentary structures outcropping as nearly vertical sections in various gravel pits clearly document the braided character of the majority of the deposit. Exceptions are remnants of coarse poorly sorted gravel documenting events of high flows

with an erosional episode followed by deposition of poorly sorted coarse gravel. Overbank deposits are quasi absent, indication of a large lateral channel mobility of the active channel belt or a lateral extent of the active channel belt including the entire valley floor, at least for the aggrading phase. Among others, Siegenthaler and Huggenberger (1993), Beres et al. (1995), Heinz and Aigner (2003b) and Heinz et al. (2003) intensively studied the pro-glacial environment along the river Rhine. They interpreted their observations from multiple outcrop studies with respect to depositional conditions and presented regional facies models that can be distinguished according to discharge types of the former glacier outwash plain.

The Herten gravel body is categorized as part of the main discharge area, also including high-energetic depositions of coarse sediments. In the following, the genetic classification and sedimentological interpretation will only be tangent to the general presentation and the reader is referred to Heinz and Aigner (2003b) and Kostic et al. (2005) for more details.

Lithofacies and sedimentary structures have been carefully mapped, from the maximum aggradation level, in a series of parallel outcrops, oriented in east western direction more or less parallel to the former River Rhine. To perform a genetic analysis on such deposits, several scales need to be taken into consideration. On the small scale, separate lithofacies types with comparable structural and textural characteristics and which stem from similar transport and depositional processes are distinguished. On the larger scale, architectural elements can be differentiated according to characteristic element boundaries (generally made up of erosional surfaces and in some cases, but in the Rhine gravel rarely, of former geomorphological surfaces). These architectural elements include cut-and-fill elements and accretionary elements as well as transitional categories between these two types. The former is the result of a process in which an erosive phase creates scours. The scours are then filled, as a consequence due to flow separation between channel base and scour, with well sorted among open framework and bimodal gravel sediments (Heinz et al. 2003, Huggenberger 2003). The gouging is usually characterized by internal trough-shaped cross-beddings that can be composed of alternating sequences. Here, lithofacies varies on the decimeter-scale. These cut-and-fill elements

(e.g. scour pools) are frequent in the main discharge area of the Rhine valley and also typical for the central part of the strata studied at the Herten site.

Accretionary elements (e.g. gravel sheets) are interpreted as accumulations in the active channel belt and dominate distant discharge areas. Usually, unarticulated horizontal layer units of matrix-dominated gravel grow over a continuously flat base. In contrast to the cut-and-fill elements, they appear homogeneous and can show significant lateral extension to hundreds of meters.

2.2 Lithofacies

The classification into small-dimensional sedimentological units is, for the most part, based on the classification used by Keller (1996) and Heinz and Aigner (2003b). According to this system, it is possible to differentiate between typical lithofacies types that can be characterized by the spectrum of grain sizes; the sorting, level of roundness, and the composition of the particles; as well as the structure and the layering.

A code ($[i_1]I_1i_2i_3[,i_4]$) is used, which for each facies combines a unique sequence of indices that represent abbreviations for typical features. Index I_1 denotes the main grain size, and optionally i_1 may be considered to describe subordinate components or matrix. For specification of I_1 , we find gravel (G) and sand (S) dominated facies types, as well a mixture of both (GS). Index i_1 highlights the appearance of cobbles (c), sand (s) or fines (f). The texture (c: clast-supported, m: matrix-supported) is indicated by the next code letter, i_2 . This may be not further be specified in well-sorted facies and accordingly represented by $i_2 = '-'$. The following index i_3 reflects stratification: x: stratified, m: massive (no bedding), g: graded. Additional information may be added, separated by a comma as i_4 . This includes alternation (a), open framework (o) and bimodal structure (b).

At the Herten quarry site, four lithofacies types are distinguished::

poorly sorted gravel (*Gcm*): gravel with a broad grain size distribution, cobbles and also sand as secondary components; the layered appearance is due to changing sand content, but imbrication of

particles indicates horizontal- to cross-bedding; this litho- and hydrofacies originates from bedloads of high-energetic flood events (Klingbeil et al. 1999, Heinz et al. 2003b); it can be of significant lateral extension up to hundreds of meters and as such form accretionary elements (Figure 2a).

alternating gravel (*Gcg,a*): sequence of matrix-filled (bimodal) and matrix-free (open framework) gravel; lower, bimodal unit with matrix and made of sand or silt; horizontal to cross-bedding; typical deposit of cut-and-fill elements, which can appear in different variants, partly eroded and incomplete (Figure 2b).

well sorted gravel and sand (*GS-x*): mixture of gravel and sand, matrix- or component-based; often cross-bedded, sheet-bound grain size change; interpreted as typical bedload deposit in high-energetic fluvial systems (Heinz et al. 2003b) (Figure 2a).

pure sand (*S-x*): very well sorted sand with occasional gravel debris, horizontal- to cross-bedding; occurs rarely at Herten site and typically represents low-energetic bedload deposit (Figure 2c).

2.3 Hydrofacies

The lithofacies classification is oriented toward physical attributes that allow direct subdivision of outcrop walls into the sedimentary categories. Mosaics can be generated that are clustered in well-defined sub-units. Such lithofacies types are distinguished from a sedimentological perspective and cannot always be transferred to hydrofacies classes. For example, for genetic reasons it is common to combine typical sequences of alternating gravel beddings (*Gcg,a*) into one lithofacies cluster. A basal matrix-supported gravel unit is overlain by sequences of clast-supported, matrix-free gravels (open framework) (Figure 2b). The lower part shows a substantially lower hydraulic conductivity than the upper highly permeable open-framework. Consequently, it is reasonable to further subdivide such lithofacies into subclasses of different hydrofacies. Vice versa, different lithofacies types with similar hydraulic properties may be combined into one hydrofacies.

In contrast to lithofacies, hydrofacies types can quantitatively be described by the specific hydraulic parameters. Thus classification based on hydrofacies means that types with the same or comparable permeability and porosity could be merged. Heinz et al. (2003) present the results from laboratory measurements (using permeameter columns) that are complemented by empirical estimation of hydraulic conductivity based on the grain size distribution. The measurements with multiple samples show a considerable variability, in particular for the matrix-supported facies types. Therefore, former estimations presented by Huggenberger et al. (1988), Jussel et al. (1994), Kleineidam (1998), Klingbeil et al. (1999) and Bayer (2000) can differ compared with the findings from the more comprehensive analyses as presented by Heinz et al. (2003). Table 1 lists the specifications taken from Bayer (2000), Heinz et al. (2003), and Kostic et al. (2005), whereas the latest are considered the more realistic and precise data. Partially empirical values as given in the original work (Bayer 2000) are replaced by later experimental conductivity and porosity estimates. In a comprehensive overview, Zappa et al. (2006) compiled the hydraulic data from different studies on alluvial sediments. In addition to the hydraulic analysis, Kleineidam (1998) and Kleineidam et al. (1999) also examined the chemical properties of the sand and gravel grains.

As presented in Table 1, hydrofacies follow the lithofacies classification. However, the poorly sorted gravel, *Gcm*, as well as the alternating gravel, *Gcg,a*, are subdivided into separate hydrofacies types (see also Figure 2a,b), which results in a total of ten categories. For the massive gravel, sandy and cobble-rich variants are distinguished, which can be characterized by individual conductivity and porosity values. As mentioned above, the alternating sequence can be further differentiated according to the gradation. Apparently, though considered as single sedimentological unit, the *Gcg,a* spans a wide range of hydraulic conductivity values over nearly seven orders of magnitude. Of major hydraulic importance appear the highly conductive, matrix free compartments of typically thin open framework deposits (*Gcg,o*).

The lithofacies-based sub-classification of hydrofacies types is standardized (e.g., Heinz et al. 2003), but could be simplified. For example, nearly the same conductivities are measured for *sGcm,b* and

sGcm; also Gcm and cGcm appear to be similar. Therefore depending on the specific hydrological objectives, it appears desirable to combine those hydrofacies types listed in Table 1 into major groups. For example, an option is to cluster and visualize them based on a log scale as will be done in the remainder of this study.

2.4 GPR and radar facies

Before excavation, the sedimentary block was examined by a ground-penetrating radar (GPR) survey. GPR is a fast, inexpensive and established geophysical technique for non-invasive collection of stratigraphical data (e.g. Huggenberger, 1993; Jol, 2009). It utilizes high-frequency electromagnetic waves that are sent in the ground, where they propagate (for gravels mostly in the order of 8-11 cm/ns), are refracted, reflected or attenuated. Changes in residual water content, clay fraction and grain size distribution determine the reflection coefficients of the different materials. The two-way travel-times of the returning signals are recorded. The vertical resolution of sedimentary structures is in the order of a quarter wave length, for gravel at the 10 cm scale. However, for a given wave length, resolution decreases with depth (i.e. increasing size of footprint with time, Daniels, 2009). The strong decrease in amplitudes with increasing depth of reflections is mainly controlled by attenuation of the energies of the electromagnetic waves due to conductive electrical properties of the different sediments. Penetration depth in the Rhine gravel depends on electrical properties of different lithofacies, which are mainly dependent on clay content and chemical composition of the pore water. For gravel it varies from a few meters to > 20 m for areas with very low apparent electrical conductivities.

During field measurement, GPR is commonly traced along a surface line to obtain a vertical 2D visualization of shallow underground structures. Multiple closely-spaced GPR profiles can be combined into a 3D representation (e.g. Beres et al. 1995). Such structural information has to be backed up by borehole or outcrop measurements to correlate the litho-, hydrofacies and/or architectural elements to the reflections. Typically, reflection amplitudes can be correlated to the

contrasts between different hydrofacies types. They can be used to identify bounding surfaces and the geometry of internal structures such as foresets of different lithologic zones and bodies. However, no unambiguous evidence of single strata properties can be given (Heinz and Aigner, 2003a). Hence, GPR profiles are ideally explained by comparison to a well-defined set of default reflection patterns and their associated interpretation. Such radar facies types are not unique and have to be defined for different depositional environments.

Asprion and co-workers (Asprion and Aigner, 1997, 1999) have examined in detail the characteristic reflection patterns in fluvio-glacial sediments as can be found in the Rhine valley. In line with their work, the GPR profiles obtained for the Herten site will be examined with respect to specific patterns that can be correlated to typical geometric structures, genetic units or architectural elements as identified in the outcrop sections. Accordingly, the measurement and signal filtering procedures as given by Asprion and Aigner (1997) were adopted, and a GSSI (Geophysical Survey Systems Inc., North Salem, NK, USA) radar system (model SIR 10A) with a pair of 300 MHz antenna was applied. Separation of transmitter and receiver was held constant at 1.4 m. Asprion and Aigner (1999) provide a more complete description of the field measurement procedure. In a related study, Kowalsky et al. (2001) selected section 4 to construct models with various water saturation distributions for GPR forward modeling.

3 Mapping procedure

For the purpose of this mapping, a 16m x 7m x 10m gravel body excavated in the summer of 1999 in the Herten gravel pit was chosen. To reduce the strong attenuation of the GPR-signal, the top-soil containing electrically high loss material has been removed. Then, perpendicular to the desired direction of excavation, GPR measurements were conducted. Finally, unconsolidated rock until a depth of ca. 9 m was dug out. The excavation extended laterally and did not reach the groundwater table. During the ongoing mining it was possible to obtain a three-dimensional image of a local field of observation within a short period of time. The sedimentary body was documented by parallel and

equidistant (2 m) vertical sections, which follow the gradual excavation (Figure 3). From the geophysical measurements, GPR maps could be assigned to each of the cross sections.

Sketches of geological structures were drawn, and small- and large-scale digital photographs were taken from all six cross-sections in order to fully capture 2D heterogeneity on the cm to dm-level. Wide-angle photographs (6 cm x 17 cm, Linhof Technorama 617 S camera with tripod and Schneider centre filter) of the entire outcrop width were scanned and served as basemaps for the digitization procedure. Bedding surfaces and facies boundaries could then be traced to produce digital maps (GIS shape format) that accurately depict the facies distribution on each face. As a result, six parallel portrayals of litho-, hydro- and radar facies were obtained. In this manner, the basis for the production of an interpolated full 3D-model was created (Maji et al 2005, Comunian et al. 2010).

As an example, Figure 4 depicts inner section 4. Already at the low resolution of the outcrop photo shown here (Figure 4a), major structures can be delineated. This is due to the variability of material, water content and stability specific to different facies types. Furthermore, in matrix-rich compartments, drying of the exposure wall surface from the sunny weather was less effective, a fact that increased the visual contrasts.

4 Results

4.1 Architectural elements

The digitization procedure and facies-based distinction delivered six very similar facies mosaics (Figure 5, no. 1 to 6 in the direction of excavation) with typical, recurring patterns. In all sections, we can distinguish six major genetic units that are dominated by specific architectural elements. These units can be distinguished as follows (Figure 4b):

The lowermost unit consists of sand-rich, poorly sorted gravel intercalated with patches and discontinuous beds of openwork gravels (10-20 cm in thickness) which indicate subhorizontal stratification.

A heterogeneous thick unit with local small sections through alternating sequences, with variably cross-bedded poorly sorted gravel dipping to the South and a massive well sorted gravel body in the Southern (left) half.

A trough-shaped, relatively thin and continuous unit with mainly alternating sequences, dipping in a Southern direction.

A wedge-shaped sand-rich, poorly sorted gravel deposit with discontinuous thin beds of openwork gravels that indicate cross stratification gently dipping to the North.

A laterally extended unit with heterogeneous sequences of inclined, concave and trough-shaped bimodal gravel and open frameworks; dipping to the South, similar to unit III.

Horizontally stratified, thick continuous gravel sheets in the uppermost ca. 1.5 m of the outcrop.

The geometry of the bounding surfaces and the dip of the tangential and parallel foresets are indicators that the sections are oriented almost parallel to the mean ancient flow direction. Genetic unit III and V can be interpreted as typical scour-pool fillings that originate from pools moving to the South with sedimentation of alternating sequences downstream. The bimodal couples found in the units II, III and V portray the existence of flow separation zone developing at the transition from the channel into the scour. These scour pool fillings represent architectural elements with erosion as well as deposition character. The size of lower unit III slightly increases in direction of excavation with greatest extension in profiles 5 and 6. This could also indicate a trough shape in this direction. Unit IV represents a gravel-sheet, which is genetically related to the scours (e.g. Ashmore 1982, Marti 2002). Alternating sequences sporadically appear also in unit V. These can be interpreted as partly eroded relics of former more extensive scour pool fills. Particularly in the Northern part, a wedge-shaped zone in the cross section can be found in all profiles that could be the beginning of a further big scale scour pool fill that dominates beyond the mapping boundaries. While the rather homogeneous units I and V cannot be easily categorized, unit VI shows typical features of erosive gravel sheets. Such gravel sheets are typical architectural elements, in particular in the more distant

glacier outwash area. Therefore, its appearance on top of the profiles could indicate a later phase of outwash deposition during further retreat of the glacier. The transition between unit V and VI is also described at different localities in NE-Switzerland, e.g., Hüntwangen pit (Siegenthaler and Huggenberger 1993).

4.2 Litho- and hydrofacies distribution

The hydrofacies show the highest variation in the alternating sequences (lithofacies Gcg,a) that dominate unit III and V. These alternating sequences account for about 25% of the profiles. The sandy bimodal gravel ($sGcg,b$) clearly dominates the base of the sequences. The finer $fGcg,b$ common to Gcg,a sequences only appears rarely and concentrated locally as trough-shaped deposits in the northern part of unit III (in profiles 1-5). The highly permeable coarse framework gravels ($cGcg,o$) also occur exclusively in this central unit. They are most pronounced in profile 5. The total portion of open framework is about 10% on the average and slightly increases towards the (western) direction of excavation (Figure 6).

The preeminent lithofacies (ca. 70%) is the matrix-supported gravel (Gcm) with the hydrofacies variants $sGcm$, $cGcm$, Gcm . In contrast to the alternating sequences, they commonly show larger extensions and more continuity from one section to the next. For example, top unit VI is built up by a continuous $cGcm - sGcm - cGcm$ strata in all profiles. Such continuous $cGcm$ gravel sheets also represent the top of unit I and the basis of unit IV. Compared with the homogeneous unit VI, the matrix-supported gravel in the other units is frequently interbedded by local trough-shaped or thin-layered framework.

Unit II is dominated by the sandy matrix-supported gravel. Therefore, as depicted in the log-based conductivity profiles (Figure 5), it is an extensive zone of low conductivity. Only in this unit, a substantial $GS-x$ wedge can be found which increases its size towards the direction of excavation. It represents a continuous zone of relatively high conductivity. Overall, well sorted lithofacies $GS-x$ and $S-x$ are rare, and the latter represents less than 1% in the profiles.

4.3 Radar profiles

Figure 5 compares the derived conductivity maps to the GPR profiles, and shows that several typical structures and bedforms are well reproduced by GPR. Radar permittivity and reflection is controlled especially by the residual water content contrasts. Thus hydrofacies boundaries are ideally reproduced if they are significant and continuous such as at the interfaces between some genetic units and within alternating sequences. This is true in particular between and within the upper units. The horizontal layering of the homogeneous gravel sheet that makes up top unit VI can be directly correlated from the outcrop facies maps to the continuous horizontal high-amplitude reflectors in the upper part of the radar profiles (see top arrow GPR profile of section 3). The following lower unit V is separated clearly from unit VI by commonly short, dipping reflectors that delineate the internal cross-beddings of the bundles of alternating gravel. Interestingly, the amplitudes of these reflections increase towards the bottom, which could reflect drainage in the upper part and ponding of residual water in the lower. A basal strong reflection of unit V originates from the contrast to the underlying massive matrix-supported gravel of unit IV.

The alternating sequences generate considerable small-scale contrasts, apparently delineating the substantial hydraulic contrasts between clast-supported and matrix-supported facies types. Especially on the Southern part, reflection dips of unit V follow the relatively horizontal layering, whereas the inclined reflections of the Northern part indicate the cross-beddings and more concave, inclined alternating sequences. Nevertheless, resolution is limited, and even though the radar profiles reproduce characteristic features well, individual beds are rarely resolved. Overall, the reflectors help to reconstruct the sequences, although the reflection volume limits depiction of small scale features. Radar facies determined based on these GPR maps thus would have to be based on the genetic units rather than single facies variability.

Deeper units are more difficult to correlate with the GPR profiles. While the wedge-shaped form of unit IV can be identified also in the reflections, it is hardly possible to deduce from radar information alone that it is underlain by a thin high-conductivity unit V. This is certainly due to the resolution that

technically decreases with the profile depth, and due the small width of unit III that cannot be resolved by the given wavelength. Again, internal lithofacies variations can hardly be interpreted by the radar profiles alone. As an exception, the major structural trends in the basal units, cross-beddings and trough-shaped bedforms in unit II and more horizontal layers in unit I, can be seen in the reflector patterns. This enables the tracing of the boundary between the units in the radar profiles, where inclined reflectors downlap on basal horizontal reflectors (as accentuated by the arrows on the GPR profile for section 2, Figure 5).

5 Conclusions

A high resolution portrayal of a braided river deposit block was developed, which is considered one of the best characterized gravel bodies of this size. Combination of sedimentary and hydrogeological facies analysis enabled a reliable structural analysis on decimeter and meter scale. The gravel body was investigated by recording full 2D cross sections during excavation. The sections were digitized based on outcrop drawings and photographs. Facies mosaics are attained, for each of which a sequence of six characteristic genetic sedimentary units can be distinguished. The facies mosaics were compared to GPR profiles. The latter had been obtained from a survey before the excavation. In particular the best resolved upper part of the profiles delivers characteristic reflection patterns: continuous high amplitudes correlate with layered gravel sheets and inclined small-scale reflections delineate alternating gravel sequences as deposited in scour pools. These patterns would be suitable for framing radar facies types. For the lower units in the gravel body, interpretation of the GPR profiles would be ambiguous without the facies maps.

The gravel body is highly heterogeneous, in particular in the alternating sequences of units III and V, and at other parts where highly conductive open frameworks are embedded in matrix supported gravel beds. Continuity in the third dimension, towards the direction of excavation, can be assessed by sequential comparison of neighboring facies profiles. It is revealed that the major genetic units recur in each section, with slight variability and some trends in the unit boundaries. Among the

individual hydrofacies zones, only larger layers and wedges show continuity. Those with high variability in the 2D section such as again the alternating sequences cannot easily be traced in the third dimension. Meanwhile, Maji et al. (2005) generated a transition probability based interpolation between the six sections. In Part II of the companion paper (Comunian et al., 2010), a purpose built two-step procedure based on kriging and multiple points statistics is introduced to ultimately derive a realistic full 3D reproduction of the gravel body.

This presentation is intended to offer more insight into the underlying “real time” mapping procedure as well as into the types of sedimentological, hydrogeological and geophysical data available. It is also meant to stimulate further use in future applications. The data set, in 2D sections or reconstructed in 3D, is certainly perfect grounds for testing and validating innovative hydrogeological modeling concepts. The efficiency of both forward and inverse modeling techniques can be examined on a highly realistic and fully digitalized field case.

Acknowledgements

The diligent assistance of Margaret Hass during the preparation of the manuscript is highly acknowledged. Special thanks go to Jürgen Heinz, Georg Teutsch, Thomas Aigner and Gerhard Lörcher for their support during mapping, digitization and facies interpretation.

6 References

- Anderson, M.P., 1989. Hydrogeologic facies models to delineate large-scale spatial trends in glacial and glaciofluvial sediments. *Geol. Soc. of Am. Bull.* 101(4), 501–511.
- Anderson, M.P., Aiken J.S, Webb E.K., Mickelson, D.M., 1999. Sedimentology and hydrogeology of two braided stream deposits. *Sediment. Geol.* 129(3-4), 187-199.
- Ashmore, P.E., 1982. Laboratory modeling of gravel braided stream morphology. *Earth Surf. Process.* 7, 2201-2225.

- Asprion, U., Aigner, T., 1997. Aquifer architecture analysis using ground penetrating radar: Triassic and Quaternary examples (S Germany). *Environ. Geol.* 28, 1-10.
- Asprion, U., Aigner, T., 1999. Towards realistic aquifer models: a three-dimensional georadar case study of quaternary gravel deltas (Singen Basin, SW Germany). *Sediment. Geol.* 129(3-4), 281-297.
- Bayer, P., 2000. Aquifer-Analog-Studie in grobklastischen 'braided river' Ablagerungen: Sedimentäre/-hydrogeologische Wandkartierung und Kalibrierung von Georadarmessungen – Diplomkartierung. University of Tübingen (unpublished).
- Becht, A., Bürger, C., Kostic, B., Appel, E., Dietrich, P., 2007. High-resolution aquifer characterization using seismic cross-hole tomography: an evaluation experiment in a gravel delta. *J. Hydrol.* 336(1-2), 171-185.
- Beres, M., Green, A.G., Huggenberger, P., Horstmeyer, H., 1995. Detailed 3-D georadar studies of glaciofluvial sediments. *Geol.* 12, 1087-1091.
- Bersezio, R., Giudici, M., Mele, M., 2007. Combining sedimentological and geophysical data for high-resolution 3-D mapping of fluvial architectural elements in the Quaternary Po plain. *Sediment. Geol.* 202(1-2), 230-248.
- Brauchler, R., Hu, R., Vogt, T., Halbouni, D., Heinrichs, T., Ptak, T., Sauter, M., 2010. Cross-well slug interference tests: An effective characterization method for resolving aquifer heterogeneity. *J. Hydrol.*, doi:10.1016/j.jhydrol.2010.01.004.
- Chen, J., Rubin, Y., 2003. An effective Bayesian model for lithofacies estimation using geophysical data. *Water Resour. Res.* 39(5), 1118, doi:10.1029/2002WR001666.
- Chen, X., Song, J., Wang, W., 2010. Spatial variability of specific yield and vertical hydraulic conductivity in a highly permeable alluvial aquifer. *J. Hydrol.* 388, 3-4, 379-388.
- Comunian, A., Straubhaar, J., Bayer, P., Renard, P., 2010. Three-dimensional high resolution fluvio-glacial aquifer analog: Part 2: geostatistical modelling. *J. Hydrol.*, submitted.
- Dafflon, B., Irving, J., Holliger, K., 2010. Calibration of high-resolution geophysical data with tracer test measurements to improve hydrological predictions. *Advances in Water Resources*, 33(1), 55-68.
- Daniels, D.J., 2009. Antennas, in: Jol, H.M. (Ed.), *Ground Penetrating Radar: Theory and Applications*, pp.104-139.

- de Marsily, G., Delay, F., Gonçalves, J., Renard, P., Teles, V., Violette, S., 2005. Dealing with spatial heterogeneity. *Hydrogeology J.* 13, 161-183.
- Dietrich, P., Butler Jr, J.J., Faiß, K., 2008. A rapid method for hydraulic profiling in unconsolidated formations. *Ground Water*, 46(2), 323-328.
- Eaton T., 2006. On the importance of geological heterogeneity for flow simulation. *Sediment. Geol.* 184(3-4), 187-201.
- Ezzy, T., Cox, M., O'Rourke, A., Huftile, G., 2006. Groundwater flow modelling within a coastal alluvial plain setting using a high resolution hydrofacies approach. *Hydrogeol J.* 14(5), 675-688.
- Fraser, G.S., Davis, J.M., 1998. *Hydrogeologic Models of Sedimentary Aquifers*, SEPM Concepts in Hydrogeology and Environmental Geology vol. 1. Soc. for Sediment. Geol., Tulsa, OK, 180 pp.
- Goutaland, D., Winiarski, T., Dubé, J.S., Bièvre, G., Buoncristiani, J.F., Chouteau, M., Giroux, B., 2008. Hydrostratigraphic characterization of glaciofluvial deposits underlying an infiltration basin using ground penetrating radar. *Vadose Zone J.* 7, 194-207.
- Harp, D.R., Dai, Z., Wolfsberg, A.V., Vrugt, J.A., Robinson, B.A., Vesselinov, V.V., 2008. Aquifer structure identification using stochastic inversion. *Geophysical Research Letters*, 35, L08404, doi:10.1029/2008GL033585.
- Heinz, J., Aigner, T., 2003a. Three-dimensional GPR analysis of various Quaternary gravel-bed braided river deposits (southwestern Germany), in: Bristow, C.S., Jol, H.M. (Eds.) *Ground Penetrating Radar in Sediments*, Spec. Publ. Geol. Soc., London, 211, 99–110.
- Heinz, J., Aigner, T., 2003b. Hierarchical dynamic stratigraphy in various Quaternary gravel deposits, Rhine glacier area (SW Germany): Implications for hydrostratigraphy. *Int. J. Earth Sci.*, 92, 923-938.
- Heinz, J., Kleinedam, S., Teutsch, G., Aigner, T., 2003. Heterogeneity patterns of Quaternary glaciofluvial gravel bodies (SW-Germany): Application to hydrogeology. *Sediment. Geol.* 158(1-2), 1-23.
- Hornung, J., Aigner, T., 1999. Reservoir and Aquifer Characterization of Fluvial Architectural Elements: Stubensandstein, Upper Triassic, Southwest Germany. *Sediment. Geol.* 129(3-4), 215-280.
- Hu, R., Brauchler, R., Herold, M., Bayer, P., Sauter, M., 2009. A hydraulic tomography approach coupling travel time inversion with steady shape analysis based on aquifer analogue study in coarsely clastic fluvial glacial deposits. EGU General Assembly, Vienna.

- Huggenberger, P., Siegenthaler, C., Stauffer, F. 1988. Grundwasserströmung in Schottern; Einfluss von Ablagerungsformen auf die Verteilung der Grundwasserfließgeschwindigkeit. *Wasserwirtschaft*, 78/5, 202-212.
- Huggenberger, P., 1993. Radar facies: recognition of facies patterns and heterogeneities within Pleistocene Rhine gravels, NE Switzerland, in: Best, J.L., Bristow, C.S., (Eds.), *Braided Rivers*. Geol. Soc. London, Spec. Publ. 75, 163–176.
- Huggenberger, P., Aigner T., 1999. Introduction to the special issue an Aquifer-sedimentology: problems, perspectives and modern approaches. *Sediment. Geol.*, 129(3-4), 179-186.
- Huggenberger, P., Regli, C., 2006. A sedimentological model to characterize braided river deposits for hydrogeological applications, in: Sambrook-Smith, G., Best, J., Bristow, C., Petts, G. (Eds.) *Braided Rivers, Process, Deposits, Ecology and Management*, ISA Special Publication 36, 51-74.
- Huysmans, M., Peeters, L., Moermans, G., Dassargues, A., 2008. Relating small-scale sedimentary structures and permeability in a cross-bedded aquifer. *J. Hydrol.*, 361(1-2), 41-51.
- Hyndman, D., Tronicke, J., 2005. Hydrogeophysical studies at the local scale: the saturated zone, in: *Hydrogeophysics*, vol 50. Water Science and Technology Library, Springer, NY, pp 391-412.
- Iversen, B.V., van der Keur, P., Vosgerau, H., 2008. Hydrogeological Relationships of Sandy Deposits: Modeling of Two-Dimensional Unsaturated Water and Pesticide Transport, 37(5), 1909-1917.
- Jol, H.M., 2009. Ground penetrating radar: theory and applications. Elsevier Lt., 544 p.
- Jussel, P., Stauffer, F., Dracos, T., 1994. Transport modeling in heterogeneous aquifers, 1. Statistical description and numerical generation of gravel deposits. *Water Resour. Res.*, 30(6), 1803-1817.
- Keller, B., 1996. Lithofazies-Codes für die Klassifikation von Lockergesteinen. *Mitt. Schweiz. Gesell. Boden Felsmechanik*, 132, 5-12.
- Kerrou, J., Renard, P., Hendricks-Franssen, H.J., Lunati, I. 2008. Issues in characterizing heterogeneity and connectivity in non-multiGaussian media. *Adv. Water Res.*, 31, 147-159.
- Kleineidam, S. 1998. Der Einfluß von Sedimentologie und Sedimentpetrographie auf den Transport gelöster organischer Schadstoffe im Grundwasser [Impact of sedimentology and sediment petrography on the fate and transport of organic compounds in groundwater]. PhD Thesis, Tübinger Geowissenschaftliche Arbeiten, 41, Germany.

- Kleineidam, S., Rügner, H., Grathwohl, P., 1999. Influence of petrographic composition/organic matter distribution of fluvial aquifer sediments on the sorption of hydrophobic contaminants. *Sediment. Geol.*, 129(3-4), 311-325.
- Klingbeil, R., Kleineidam, S., Asprion, U., Aigner, T., Teutsch, G., 1999. Relating lithofacies to hydrofacies: outcrop-based hydrogeological characterisation of quaternary gravel deposits. *Sediment, Geol.*, 129(3-4), 299–310.
- Kock, S., Huggenberger, P., Preusser, F., Rentzel, P., Wetzel, A. 2009. Formation and evolution of the lower Terrace of the Rhine River in the area of Basel, *Swiss J. Geosci.* 102, 307-321.
- Kostic, B., Becht, A., Aigner, T., 2005. 3-D sedimentary architecture of a Quaternary gravel delta (SW Germany): implications for hydrostratigraphy. *Sediment. Geol.* 181(3-4), 143-171.
- Kowalsky, M.B., Dietrich, P., Teutsch, G., Rubin, Y. 2001, Forward modeling of ground-penetrating radar data using digitized outcrop images and multiple scenarios of water saturation. *Water Resour. Res.*, 37(6), 1615-1625.
- Maier, U., Becht, A., Kostic, B., Bürger, C., Bayer, P., Teutsch, G., Dietrich, P., 2005. Characterization of Quaternary gravel aquifers and their implementation in hydrogeological models, in *Proceedings of GQ2004. 4th International Groundwater Quality Conference, Waterloo, Canada, July 2004, IAHS Publ*, 297:159-168.
- Maier, U., DeBiase, C., Baeder-Bederski, O., Bayer, P., 2009. Calibration of hydraulic parameters for large-scale vertical flow constructed wetlands. *J. Hydrol.*, doi: 10.1016/j.jhydrol.2009.02.032.
- Maji, R., 2005. Conditional stochastic modelling of DNAPL migration and dissolution in a high-resolution aquifer analog. PhD Thesis, Dept of Earth Sciences, University of Waterloo.
- Maji, R., Sudicky, E.A., Panday, S., Teutsch, G., 2006. Transition Probability/Markov Chain Analysis of DNAPL Source-zones and Plumes. *Ground Water* 44(6), 853-863. doi: 10.1111/j.1745-6584.2005.00194.x.
- Marti, C., 2002. Morphodynamics of widenings in steep rivers. In: Bousmar, D., Zech, Y., (Eds.), *River Flow*, Vol. 2, pp. 865–873.
- Ouellon, T., Lefebvre, R., Marcotte, D., Boutin, A., Blais, V., Parent, M., 2008. Hydraulic conductivity heterogeneity of a local deltaic aquifer system from the kriged 3D distribution of hydrofacies from borehole logs, Valcartier, Canada., *J. Hydrology*, 351, 71-86.

- Regli, C., Huggenberger, P., Rauber, M., 2002 Interpretation of drill core and georadar data of coarse gravel deposits. *J. Hydrol.*, 255(1-4), 234-252.
- Sánchez-Vila, X., Carrera, J., Girardi, J.P., 1996. Scale effects in transmissivity. *J. Hydrol.*, 183, 1-22.
- Scheibe, T., Freyberg, D.L., 1995. The use of sedimentological information for geometric simulation of natural porous media structure. *Water Resour. Res.*, 3259-3270.
- Smith, S.A., 1989. Sedimentation in a meandering gravel-bed river: the River Tywi, South Wales. *Geol J.* 24, 193–204.
- Stanford, S.D., Ashley, G.M., 1998. Using three-dimensional geologic models to map glacial aquifer systems: an example from New Jersey, in: Fraser, G.S., Davis, J.M. (Eds.) *Hydrogeologic models of sedimentary aquifers*, vol 1: Concepts in hydrogeology and environmental geology, no 1 SEPM, Oklahoma, USA, pp 69-84.
- Webb, E.K., Anderson, M.P. 1996. Simulating of preferential flow in three-dimensional, heterogeneous conductivity fields with realistic internal architecture. *Water Resour. Res.*, 32(3), 533-545.
- Weissmann, G.S., Fogg, G.E., 1999. Multi-scale alluvial fan heterogeneity modeled with transition probability geostatistics in a sequence stratigraphic framework. *J. Hydrol.*, 226(1– 2), 48-65.
- Werth, C.J., Cirpka, O.A., Grathwohl, P., 2006. Enhanced mixing and reaction through flow focusing in heterogeneous porous media. *Water Resour. Res.*, 42, W12414.
- Whittaker, J., Teutsch, G., 1999. The simulation of subsurface characterization methods applied to a natural aquifer analogue. *Adv. Water Resour. Res.*, 22(8), 819-829.
- Willis, B.J., White, C.D., 2000. Quantitative outcrop data for flow simulation. *J. Sediment Res.* 70(4), 788-802.
- Zappa, G., Bersezio, R., Felletti, F., Giudici, M., 2006. Modeling heterogeneity of gravelsand, braided stream, alluvial aquifers at the facies scale. *J. Hydrol.* 325, 134-153.
- Zinn, B., Harvey, C.F., 2003. When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields. *Water Resour. Res.* 39, 1051, doi 10.1029/2001WR001146.

Table

Table 1. Litho- and hydrofacies types at Herten site with hydraulic properties.

lithofacies code	description	hydrofacies code	further description	hydraulic conductivity, K [m/s]	porosity	reference
Gcm	poorly sorted, matrix supported gravel	Gcm	normal	$2.5 \cdot 10^{-4}$	0.17	(2)
		cGcm	cobble-rich	$2.3 \cdot 10^{-4}$	0.15	(2)
		sGcm	sand-rich	$6.1 \cdot 10^{-5}$	0.13	(2)
Gcg,a	alternating gravel		matrix-free, clast supported open framework coarse-fine pebbles	$2.6 \cdot 10^{-2}$	0.26	(3)
		cGcg,o	cobbles-coarse pebbles openwork	$1.3 \cdot 10^{-1}$	0.26	(3)
		sGcg,o	granules/sand open framework	$9.5 \cdot 10^{-2}$	0.23	(2)
		sGcm,b	bimodal basal subunit with sand matrix	$4.3 \cdot 10^{-5}$	0.22	(1)
		fGcm,b	bimodal basal subunit with silt/clay matrix	$6.0 \cdot 10^{-7}$	0.2	(1)
GS-x	well sorted gravel (and coarse sand)	GS-x		$2.3 \cdot 10^{-3}$	0.27	(2)
S-x	pure, well-sorted sand	S-x		$1.4 \cdot 10^{-4}$	0.36	(2)

(1) Bayer 2000, (2) Heinz et al. 2003, (3) Kostic et al. 2005

Figures

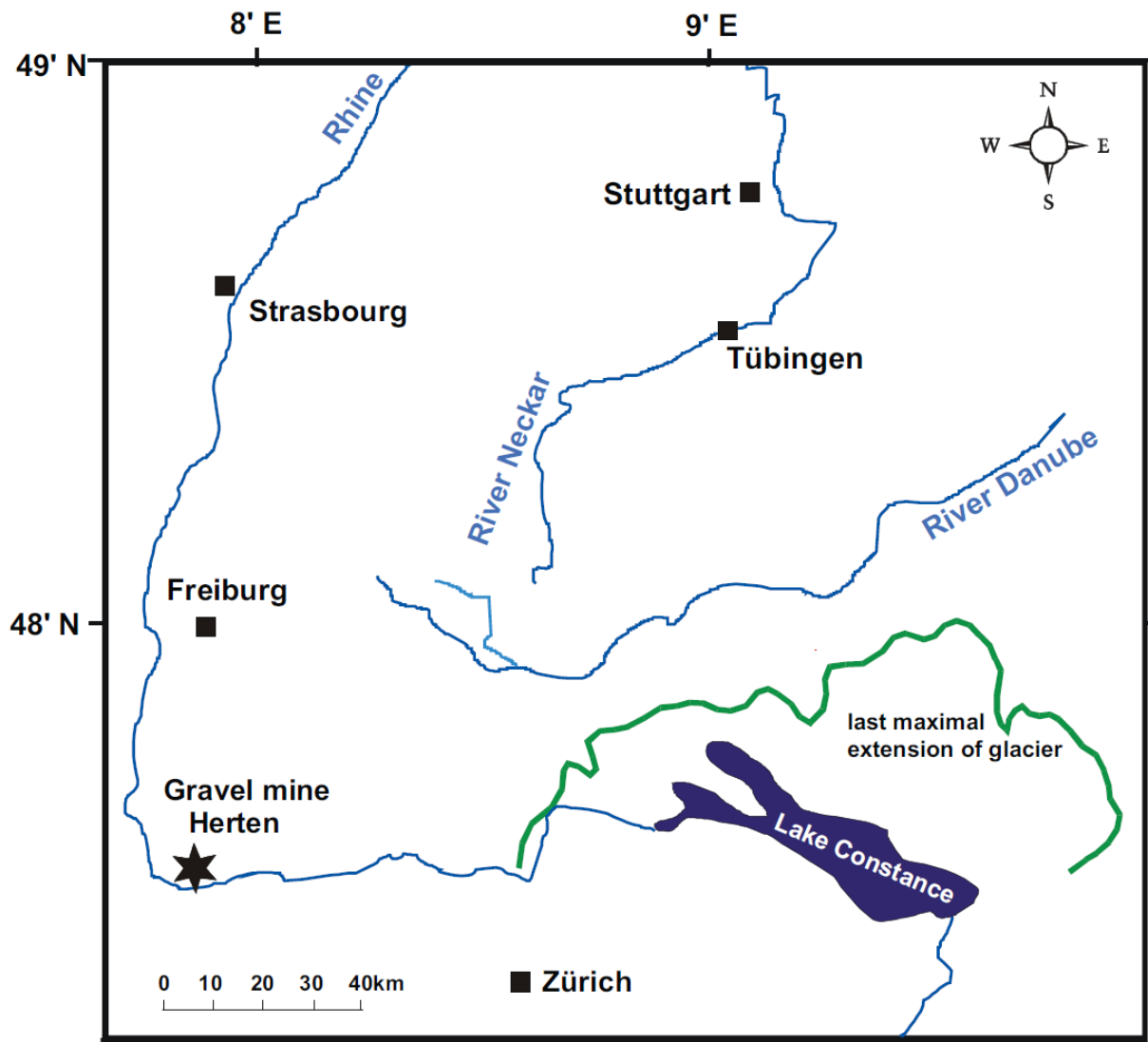


Fig. 1: Map of Herten gravel pit location in SW Germany at the SW bend of the River Rhine.

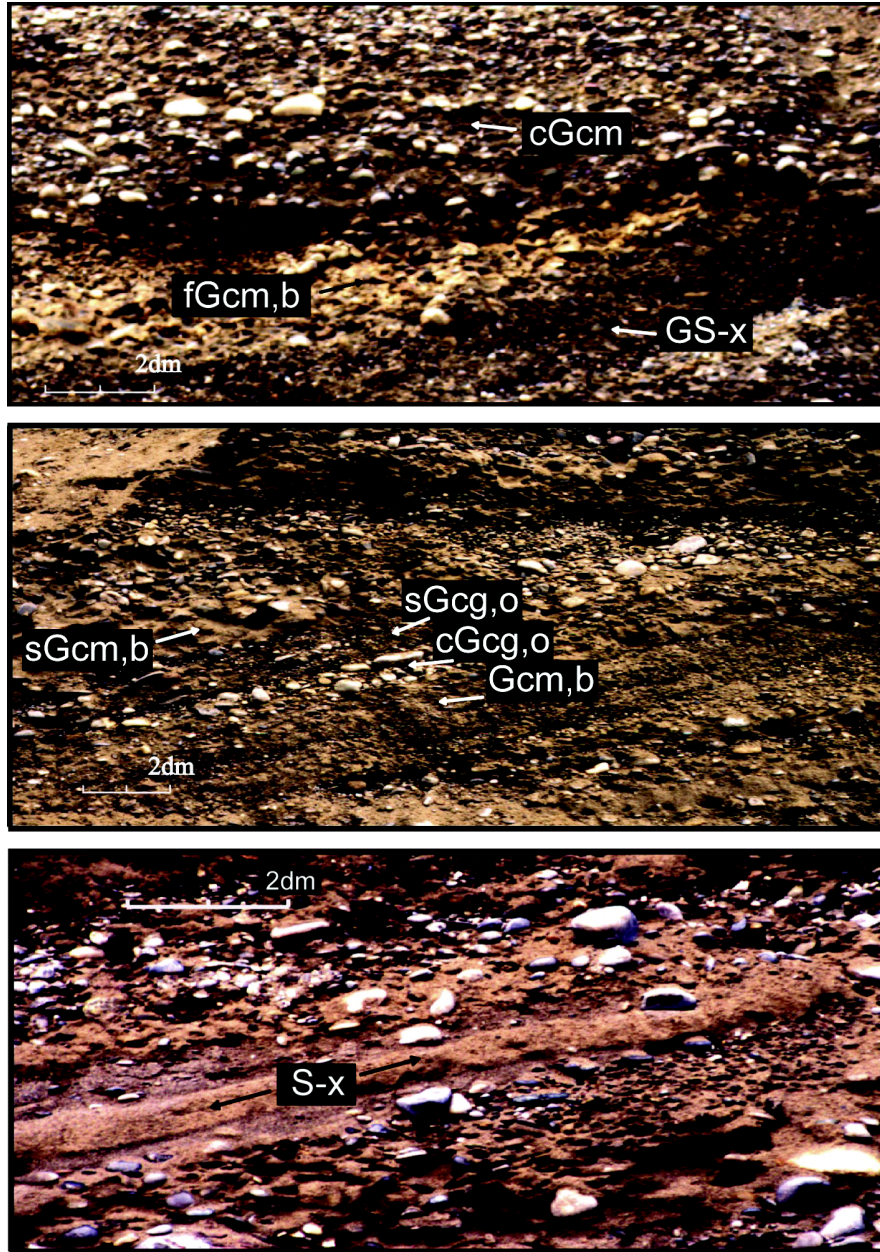


Fig. 2: Photos of typical lithofacies and hydrofacies at study site: a) poorly and well sorted gravel, b) alternating gravel, c) well sorted sand.

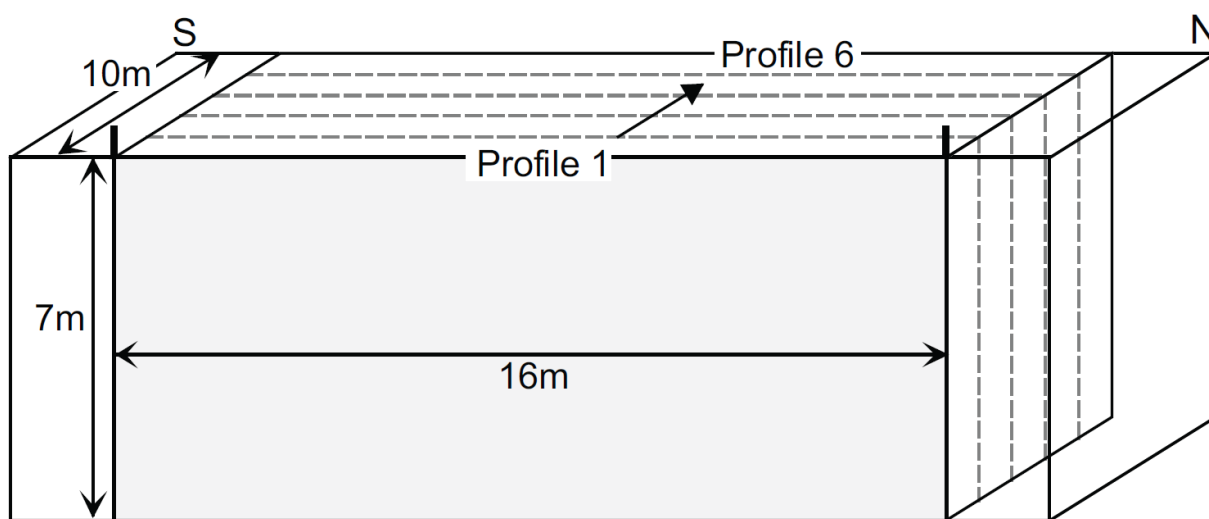


Fig. 3: Orientation of profiles through gravel body (uniform distance between profiles of 2m). During excavation, a pair of pillars on top was used to mark the position of each new section.

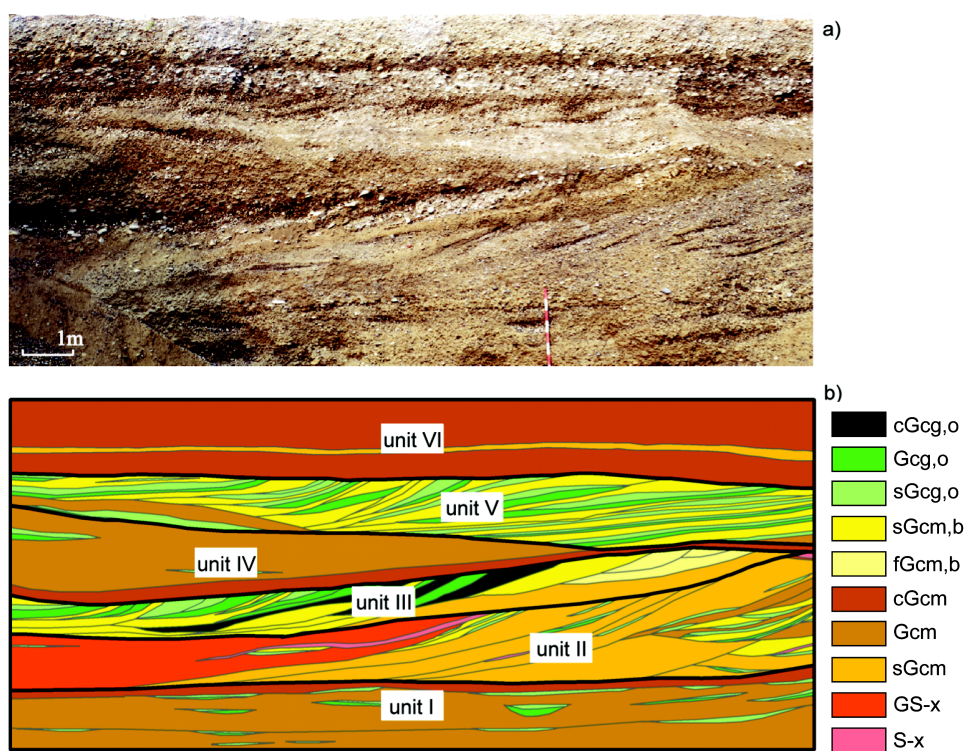


Fig. 4: Profile 4 – a) outcrop photograph, b) litho and hydrofacies distribution after digitization with distinction of main genetic units I - VI.

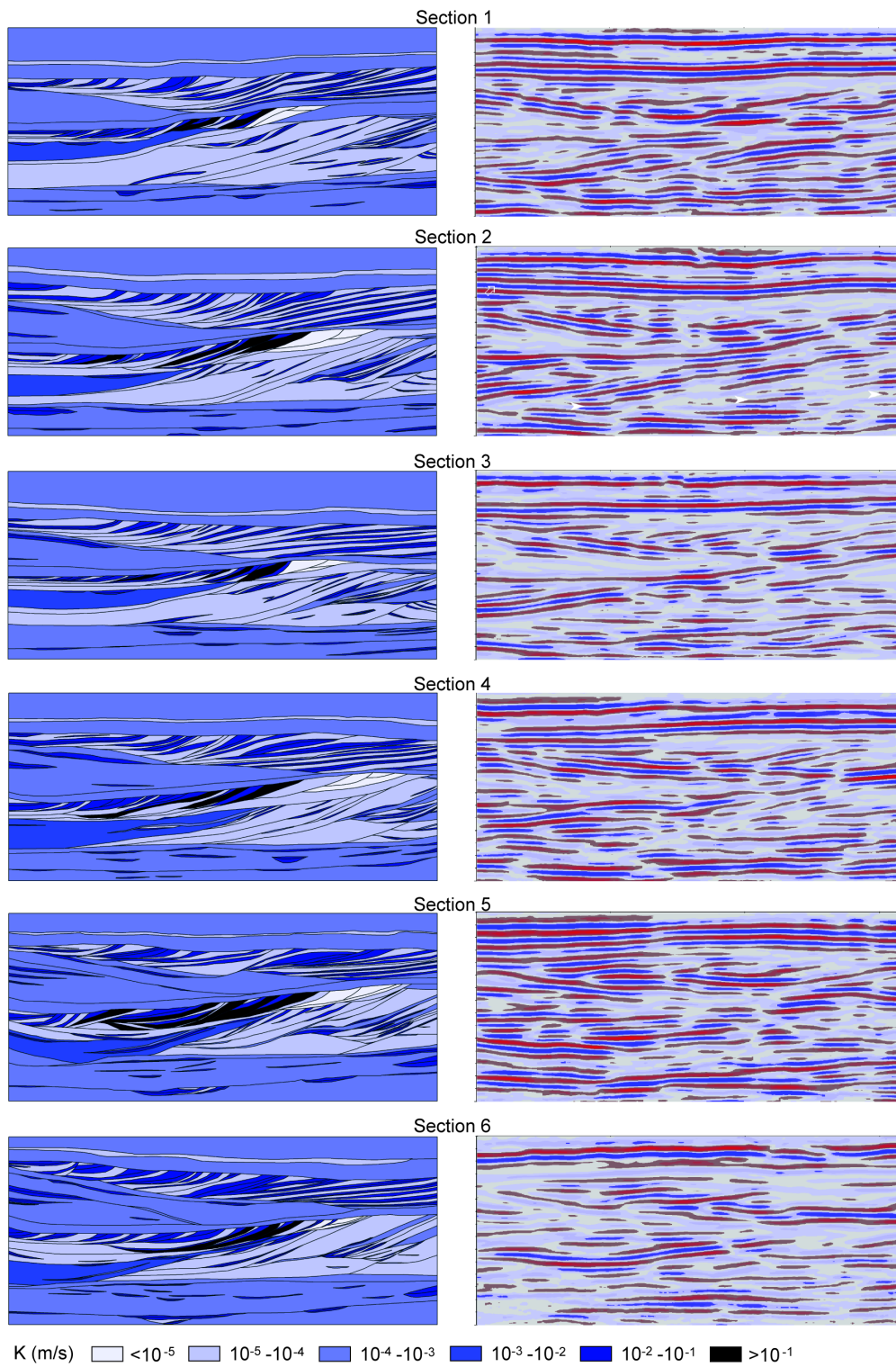


Fig. 5: Hydrofacies based conductivity map (left) and corresponding GPR profiles of sections 1-6.

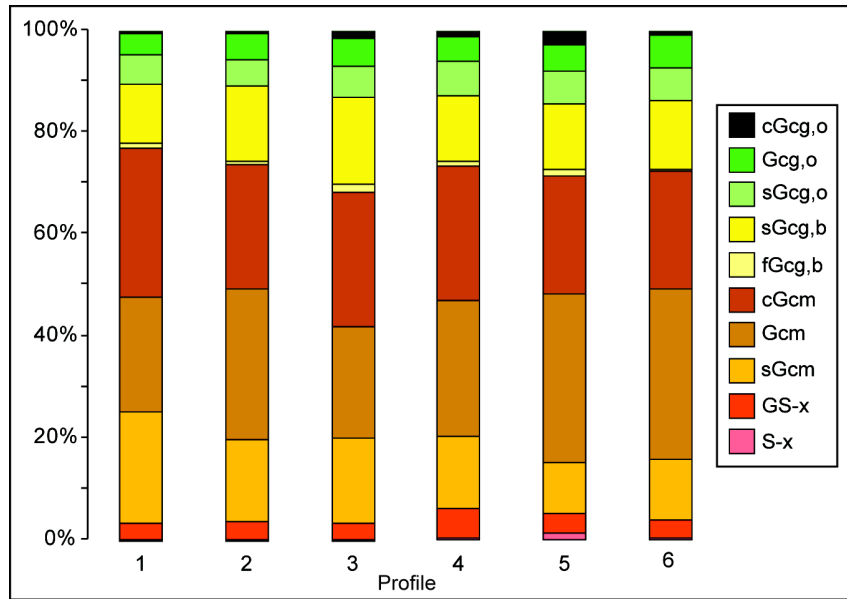


Fig. 6: Areal percentage of hydrofacies classes in the six profiles.

Appendix D

Comparison of three
geostatistical methods for
hydrofacies simulation: a test
on mixed bedload-suspended
load alluvial sediments

Diana dell’Arciprete, Riccardo Bersezio, Fabrizio Felletti, Mauro Giudici, Alessandro Comunian,
Philippe Renard

**Comparison of three geostatistical methods for hydrofacies simulation:
a test on mixed bedload - suspended load alluvial sediments**

Diana dell’Arciprete, Fabrizio Felletti, Mauro Giudici
Università degli Studi di Milano, Dipartimento di Scienze della Terra “A. Desio”,
via Mangiagalli 34, 20133 I-Milano.

Riccardo Bersezio
Università degli Studi di Milano, Dipartimento di Scienze della Terra “A. Desio”,
CNR – IDPA,
via Mangiagalli 34, 20133 I-Milano.

Alessandro Comunian, Philippe Renard
Centre d'Hydrogéologie, Université de Neuchatel,
Rue Emile Argand, 11, 2007 Neuchatel, CH.

Corresponding author:
Mauro Giudici mauro.giudici@unimi.it

Abstract

The transport properties of porous sediments are strongly influenced by the heterogeneity at fine scales, which can be reconstructed by geostatistical simulations. In order to improve the assessment of the properties of three different geostatistical simulation methods (Sequential Indicator Simulation, SISIM; Transition Probabilities Geostatistics, T-ProGS; Multiple Point Simulation, MPS) a comparison test at different scales was performed for a well-exposed aquifer analogue. In the analysed volume (approximately 30000 m³) four operative hydrofacies have been recognised: very fine sand and silt, sand, gravelly sand and open framework gravel.

Several equiprobable realizations were computed with SISIM, MPS and T-ProGS for a test volume of approximately 400 m³ and for the entire volume and the different output were compared with visual inspection and connectivity analysis of the high or low permeable structures.

The comparison of the different simulations shows that the geological model is best reproduced when the simulations are realised separately for each highest rank depositional element and subsequently merged. Moreover, the three methods yield different images of the volume: in particular MPS is efficient in mapping the geometries of the most represented hydrofacies, whereas SISIM and T-ProGS can account for the distribution of the less represented facies.

Keywords

Alluvial sediments, Aquifer characterization, Geostatistics, Hydrofacies, Heterogeneity

1 Introduction

Progress towards improved characterization and modeling of porous groundwater reservoirs requires the integration of several different methods and involves the use of multiple data-sets, consisting of both descriptive “soft” geological data and “hard” sedimentological and hydrological parameters. Hard data are usually available in correspondence of wells and are often distributed in space by stratigraphic correlations and geostatistical interpolations. This paper focuses on the reconstruction of hydrofacies distribution in a mixed bedload-suspended load alluvial aquifer analogue, obtained by different geostatistical techniques and calibrated to direct knowledge of the lithofacies/hydrofacies

distribution and sedimentary architecture. Different geostatistical simulation methods are available to obtain such a goal (Deutsch 2002).

In this study, lithofacies distribution was simulated by three pixel-based simulation methods which are commonly used for water and oil reservoir modeling: Sequential Indicator Simulation (SISIM; Goovaerts 1997; Deutsch and Journel 1992), Transition Probability Geostatistics (T-ProGS; Carle and Fogg 1996; Ritzi 2000) and Multiple Point Simulation (MPS; Strebelle 2002, Liu et al. 2005). Pixel-based algorithms for facies modeling are fast and allow direct hard data conditioning, avoiding iterative time-consuming approaches. These techniques permit to simulate the different facies in the form of coded indicator-type variables, where each value corresponds to a given facies.

Traditional geostatistical methods, such as SISIM and T-ProGS, are based on random function models, chosen according to the geometrical affinity of their outcomes with the geological architecture under investigation. The model parameters (e.g., facies proportion, object density, variogram, etc.) are inferred from available quantitative data, and the model realizations are generated and conditioned to these data. This approach is effective in practice provided that the structure of the random function model captures the geometrical features. Although this approach has been successfully applied to many cases, it is not always possible to build a random function model which represents some complex geological architectures, for instance, in a meandering channel system. MPS was proposed by Guardiano and Srivastava (1993) for modeling subsurface heterogeneity. The first efficient implementation of the method was developed by Strebelle (2002) and it started to be an active research topic in recent years. Unlike traditional geostatistical simulations based on analytical random function models, MPS does not require such an explicit definition of a random function. It directly utilizes empirical multivariate distributions inferred from training images. Therefore MPS is applicable to any geological environment, provided that (i) a training image representative of the geological heterogeneity is available and (ii) the essential features of this training image can be characterized by statistics defined on a configuration involving few points (Hu and Chugnova 2008).

An important feature when characterizing aquifers and reservoirs with respect to their flow properties is the connectivity of the high or low permeable structures (Alabert et al. 1992; Zinn and Harvey 2003). Failing to capture subsurface connectivity may bias the forecasts of any underground fluid flow and transport modeling project (Gómez-Hernández and Wen 1998; Knudby and Carrera 2005).

This concept has been seldom treated systematically and only few authors have used connectivity in order to compare results of different flow and transport models (Alabert et al. 1992; Zinn and Harvey 2003; Knudby and Carrera 2005). Recently some authors are trying to use the connectivity function as constraints for simulations (Renard and Caers 2008).

The goal of this work is to compare the performances of SISIM, T-ProGS and MPS when applied to complex alluvial sediments, like point-bar/channel systems and at different scales. The test was conducted on an aquifer analogue which was exposed in a sand-gravel quarry dug in the Lambro Valley (Northern Italy, just south of the Milan conurbation), aiming to improve the understanding of the strengths and the weaknesses of SISIM, T-ProGS and MPS at different scales and with a varying ratio between the numbers of the conditioning data and of the simulated grid cells.

In the next section the case study and the methods are briefly summarized.

After this, a description of the data set used for the geostatistical simulations is provided, before the presentation of the principal results and the relative discussion.

2 Case study and methods

The case study is a Holocene point bar–channel system of the Lambro river, a left tributary of the Po river that flows south of Milan encased in a terraced valley. In this study the historical sediments from an intermediate terrace between the present day river course and the LGM (last glacial maximum) alluvial plain are examined. The aquifer analogue was exposed in a sand-gravel open-air quarry, the so called “Ca’ de Geri quarry site”, located in Sant’Angelo Lodigiano, south of Milan (Figure 1).

This area was chosen because a large exposure of sediments belonging to two different phases of river’s evolution could be studied. The sediments were progressively removed during quarrying, so it was possible to study the internal architecture of the point-bar and channel system, by the analysis of different outcrops. In the same area, before the removal of the sandy-gravel material, geophysical surveys were conducted in order to support the reconstruction of the external geometry of the sedimentary bodies (Mele Degree Thesis 2004; Bersezio et al. 2007).

The analyzed volume is approximately 30000 m³ (47 m×75 m×8.6 m), covering an horizontal surface of about 3500 m².

The quarry site exposed three superimposed depositional units formed by sands, gravels and subordinate silt and clay, which could be attributed to an historical age, as it was proved by the findings of Roman to Middle Age and Renaissance Age artifacts (bricks, tiles, ceramics), imbricated within dunes and bars (dell'Arciprete Degree Thesis 2005; Bersezio et al. 2007). Two units correspond to the exposed parts of two composite point bars and channels with minor channel fills on top. They were named respectively unit A (the lower, with Roman-Middle Age findings) and unit B (the upper, with Renaissance Age findings). Unit A shows the lateral transition from a composite point bar to main channel fill, unit B is mostly represented by a composite point bar, with chute channel scour and fills on top. A younger channel (unit C, bounded by the erosion surface β and partly anthropogenic) eroded part of unit B. Together with units A and B, it is cut by the modern and present-day courses of the Lambro river.

The simulation methods require conditioning data, which, for this case study, are taken from the vertical facies maps of five almost orthogonal quarry faces. For modeling purposes we adopted a classification based on four hydrofacies that we defined after permeametric analyses on samples: least permeable hydrofacies (F+fS - very fine sand and silt-clay respectively from topmost channel-fill, silt/clay plugs, drapes and balls), poorly permeable hydrofacies (S - sand from point-bar and channel fill bedforms), intermediate permeable hydrofacies (SG – GS - sandy gravel and gravelly sand from point bars) and most permeable hydrofacies (G - open framework gravels from the lower parts of the lateral accreted units).

The geostatistical simulations were performed on three different domains (Figure 2):

- 1) a *test volume* (11.4 m×11.4 m×2.85 m) discretized with small (0.2 m×0.2 m×0.05 m) cells or fine grid. This test volume was chosen in an area where many conditioning data, belonging to two orthogonal quarry faces, were available covering part of the three units A, B and C (Figure 2);
- 2) the *entire volume* (46 m×74 m×8.4 m) discretized with large (0.4 m×0.4 m×0.1 m) cells or coarse grid;
- 3) the entire volume discretized with *fine grid* (0.2 m×0.2 m×0.05 m).

In order to evaluate efficiency and pitfalls of the different techniques, at different operative scales, the entire volume was reconstructed in two different ways: (a) simulating the undivided entire volume; (b)

simulating separately the units, then merging the simulations through the α boundary which separates Units A and B, that was obtained after kriging of the data points from logging and GPR survey.

The size of the cells was chosen in order to capture the heterogeneities at the hydrofacies scale. This size permits to elaborate the simulations allocating a reasonable computation time, using a PC or a workstation.

A short description of some properties of the simulation methods is now given.

Sequential indicator simulation (SISIM)

SISIM is based on the indicator variograms for the different facies and has been applied at different scales in a variety of depositional settings such as alluvial (Journel et al. 1998; Seifert and Jensen 1999; Zappa et al. 2006; Felletti et al. 2006; Falivene et al. 2007), deltaic (Cabello et al. 2007), aeolian (Sweet et al. 1996), and deep-marine “turbiditic” settings (Journel and Gómez-Hernández 1993; Falivene et al. 2007).

There are legitimate criticisms against SISIM: i) the models can appear very patchy and unstructured; ii) indicator variograms use two-point statistical measures only; iii) SISIM often leads to uncontrolled and geologically unrealistic transitions between the simulated categories; iv) the cross correlation between multiple categories is not explicitly controlled. Despite these criticisms, there are many good reasons to consider SISIM: i) the required statistical parameters are easy to infer from limited data sets; ii) the simulation results are reasonable in settings where large-scale curvilinear features are absent; iii) the algorithm is robust and provides a straightforward way to transfer uncertainty in categories through the resulting numerical models.

Semivariogram computation and SISIM were performed using GSLib (Deutsch and Journel 1992).

Transitional probabilities geostatistics (T-ProGS)

T-ProGS is based on the probabilities of transitions from one facies to another and on the modeling of such probabilities with Markov chain (Fogg et al. 1998; Carle et al. 1998). It has been applied to model facies distribution in braided rivers (Felletti et al. 2006) and in alluvial fans (Fogg et al. 1998; Carle et al. 1998; Weissmann et al. 1999; Weissmann and Fogg 1999; Fleckenstein et al. 2006).

All the steps for simulations, including visualization, are implemented with the T-ProGS library of FORTRAN programs. Transition probability matrices are generated using GAMEAS (Carle 1999), a modification of the GSLIB (Deutsch and Journel 1998) code GAMV for estimating spatial statistics

from irregularly spaced data. Transition probabilities in the horizontal plane are assumed isotropic, separate transition probabilities are determined for the vertical direction. Markov chain models are generated using MCMOD (Carle 1999). Finally, simulation is performed with TSIM (Carle 1999).

Multiple point simulation (MPS)

MPS has been used to reconstruct turbiditic reservoirs using 3D training images and conditioning data from boreholes and geophysical prospecting (Strebelle et al. 2003); also 3D reconstructions starting from 2D training images at the pore scale have been proposed by Okabe and Blunt (2005). MPS is a recent technique (Caers 2001; Strebelle 2002; Liu et al. 2004, Straubhaar et al 2010). At each grid node, the conditional PDF of the facies is obtained by considering hard data, previously simulated nodes and training images. The training image provides a conceptual description of the geological heterogeneity. It is analyzed within a template of fixed dimension; all the possible patterns embedded in that template are counted. This allows to evaluate the conditional probability to find a given facies at the node to be simulated knowing the conditioning data and the previously simulated nodes.

Multiple-point geostatistics should represent an improvement with respect to the classical variogram-based geostatistical methods because it characterizes the spatial structure by considering more than two data points (spatial configurations of several points are searched for in this method, whereas variograms account for correlation between pairs of points only), enabling the reproduction of complex patterns. However, one difficulty with this approach is to provide an adequate training image.

The training images represent conceptual images of the heterogeneity to be reproduced; they depict the patterns of geological heterogeneities deemed relevant to the application under study. The training images do not carry any locally accurate information on the real structure, they merely reflect a prior geological/structural concept. Thus, a training image can be an unconditional realization generated by an object-based algorithm, or a simulated realization of an analogous field, or simply a geologist's sketch processed with CAD algorithms and properly digitized (Strebelle 2002). In this study the discretized vertical facies maps were used as training images. This poses a problem because these images include non stationarity and are only two dimensional while MPS requires a 3D training image to generate a 3D simulation (see discussion in Comunian et al 2010).

Different techniques are currently under development to overcome that difficulty. Here, MPS was performed with a customized version of the code IMPALA (Straubhaar et al. 2010). More precisely, the simulations use two perpendicular training images. The conditional probability density distributions computed from these two images in two different directions are combined following the same principle as Okabe and Blunt (2005) but instead of using a linear combination of the individual pdfs, we use a multiplicative formula proposed by Bordley (1982). In that formula, a weight of 0.5 was given to each pdf.

To compare the results obtained by the different geostatistical techniques, connectivity indicators have been used. Several definitions of connectivity can be given (Vassena et al. 2009). In this study, two indicators of facies connectivity are computed: total and intrinsic connectivity. They are based on the probability that pairs of connected points belong to a subset, characterized by a given property, e.g. texture or hydraulic conductivity, as proposed in Vassena et al. (2009). In particular the total connectivity of a facies is the probability that two, non coincident, points of the domain belonging to the given facies are connected; the intrinsic conductivity is the probability that two points are connected, conditioned on the fact the both belong to the given facies. These indicators share some properties of the connectivity functions proposed by Allard and Group (1993), Allard (1994) and Western et al. (2001), but they do not depend upon the separation lag between the points.

3 Dataset for the geostatistical simulation

The dataset used to compute the geostatistical simulations consists of:

- 1) *Facies proportions*, computed on the vertical facies maps at different hierarchic levels (Figure 3). In particular different facies proportions are used to simulate either Units A and B separately or the whole volume. Note that in the entire volume, the facies F is more abundant than in units A and B, because unit C is rich of fine sediments.
- 2) *Vertical facies maps*, discretized with both small (0.2 m×0.2 m×0.05 m) and large (0.4 m×0.4 m×0.1 m) cells.
- 3) *Variograms* for SISIM simulation, computed along the three space coordinates, for units A and B separately and for the whole volume, with both small and large cells. For the simulations of the test

volume different variograms are computed, using data from the portions of the facies maps inside the test volume (Figure 4 a).

4) *Transition probabilities* for T-ProGS simulation, computed along the three space coordinates, for units A and B separately and for the whole volume, with both small and large cells (Figure 4 b).

5) The almost perpendicular facies maps were used as 2D training images for MPS simulations. The training images consist of two geological sections derived from the field observations. They are displayed on Figure 4 c. A template of 7x7 pixels was used in each direction, as well as 4 levels of multi-grid. The non stationarity displayed in the training images and in the simulation domain has not been treated explicitly.

Variogram models and transition probabilities are fitted to the experimental variograms and transitional probability curves.

Fifty simulations for each method (SISIM, T-ProGS and MPS) were performed on the test volume, which consists of 185193 cells; a portion of discretized facies maps inside the test volume (19380 cells) was used as conditioning data. On the entire volume (coarse grid), 10 simulations were performed for each simulation method. For this volume, which consists of 1787100 cells, all the discretized facies maps (i.e. 30116 cells) were used as conditioning data. On the entire volume (fine grid) one simulation for each method was performed in order to contain computing times for facies modeling. In a first attempt, the undivided volume (15140816 cells), was simulated using all the discretized facies maps (120868 cells) as conditioning data. A second simulation attempt was computed on the separate units A and B (with respectively 78435 and 28372 cells as conditioning data), then merging the simulations by the kriged α boundary between A and B.

For the goals of this work it is important to validate the results, i.e., to check whether the simulation results reproduced the statistical parameters of the real distribution of the hydrofacies; however, even more important is the assessment of similarities and differences among the results of the three simulation methods. The comparisons between the geological model and the simulations and among different simulation methods are performed in four different ways:

- 1) visual inspection of the simulated volumes;
- 2) 3D representation of the probability of occurrence of each facies;

- 3) comparative image analysis of the vertical facies maps and of sections cut through the simulated volumes at increasing separation from the conditioning data;
- 4) connectivity analysis.

4 Discussion of the results

Test volume

In Figure 5 one realization randomly selected out of the 50 simulations obtained for the test volume using the three simulation methods is shown.

The mean and the standard deviation of facies proportions for the 50 simulations are reported in Table 1. The facies proportions in the simulated volumes are different from the proportions used as input parameter, because of the averaging effect between the conditions imposed by different parameters (conditioning data, proportions, variograms, etc). In the case of MPS the training images represent a stronger constrain than facies proportions.

Total and Intrinsic connectivity indicators computed in the *test volume* are displayed in the graph of Figure 6. From the computation of the connectivity indicators, it can be observed that: i) for MPS the total connectivity for facies S is higher than for facies SG, instead for SISIM the total connectivity for S facies is lower than for SG facies; ii) for MPS the intrinsic connectivity of S facies is closed to the intrinsic connectivity of SG facies; iii) the intrinsic connectivity for G facies is less than 0.1 for SISIM simulations.

From the 50 equiprobable simulations of the *test volume*, the probability of occurrence for each facies was computed (Figure 7). At this scale the following results can be highlighted.

- 1) It was almost impossible to find a Markov Chain model fitting the transition probabilities statistics computed on the conditioning faces of the test volume, therefore T-ProGS simulations are unsatisfactory for all facies at this scale (Figure 5). At this scale, in fact, there are few repetitions of the hydrofacies bodies (that mimic the sedimentary architectural elements), therefore transition probabilities statistics are inadequate.
- 2) MPS reproduces laterally persistent hydrofacies bodies for the S, SG and G facies, that are the most abundant in the conditioning faces used to derive the training images; on the other hand the same simulations cannot take into account the geometry and distribution of the F hydrofacies

bodies, that are present only in one of the orthogonal training images. As a consequence, the MPS technique reproduces well connected volumes of the high permeability hydrofacies (SG and G), that could represent the preferential flow paths in the considered region of the aquifer analogue (Figure 6).

3) SISIM reproduces S, SG and G hydrofacies bodies that are less persistent than those generated by MPS; on the other hand all the simulated hydrofacies have a good and realistic spatial continuity (Figure 7).

Entire volume, coarse grid

Figure 8 shows randomly selected examples of the realizations obtained on the *entire volume*, with the *coarse grid*, for each simulation method.

The mean and the standard deviations of facies proportions for the 10 simulations are reported in Table 2. Notice that the proportions of F and G facies for the simulations with MPS are significantly lower than the input data, because this method reproduces the less abundant facies only partially.

Total and Intrinsic connectivity indicators are displayed in the graphs of Figures 9 and 10.

From the 10 equiprobable simulations of the *entire volume* simulated with the *coarse grid*, the probability of occurrence for each facies was computed (Figure 11).

At this scale the following results can be highlighted.

- 1) T-ProGS simulations generate a background of S hydrofacies voxels, within the other hydrofacies are sparse through the whole volume. On the contrary MPS simulations yield well connected volumes of SG hydrofacies bodies that alternate with the S hydrofacies bodies, in agreement with the conditioning data-set (Figure 12).
- 2) MPS simulations can take into account the least represented hydrofacies (F and G, Figure 8) only partially; nonetheless they can reproduce realistic shape and size of the most abundant S and SG hydrofacies. The most represented hydrofacies are reproduced by MPS simulations in the whole volume as bodies whose shape and size are very similar to those appearing in the training images (Figure 12).

3) T-Progs simulations yield a more connected pattern of the low permeable F hydrofacies, than the high permeable G hydrofacies. Both S and G hydrofacies are loosely connected in SISIM realizations (Figures 9 and 10).

Entire volume, fine grid

Two simulations of the *entire volume* with the *fine grid* were obtained with each method, simulating either the whole undivided volume or the units A and B separately (unit C was not considered here, because of the very poor available observation). The results are shown in Figure 13. In Table 3 facies proportions derived from the simulations are reported.

Total and Intrinsic connectivity indicators were computed in the *entire volume* with the *fine grid*, for moving blocks ($57 \times 57 \times 57$ cells, corresponding to $11.4 \text{ m} \times 11.4 \text{ m} \times 2.85 \text{ m}$); the results are displayed in Figure 14.

On the *entire volume* simulated with the *fine grid*, the following results can be highlighted.

- 1) Visual inspection and comparison with the facies maps and field observations show that SISIM and T-ProGS yielded unrealistic results for the undivided volume of Units A and B. The realizations obtained with the separate simulation of the two units were by far more realistic (Figure 13). On the contrary MPS yields more realistic simulations for the undivided volume than for the merge of the separate simulations of A and B. MPS takes in account the differences between the two units that are evident in the training images of the entire volume, and works better over large volumes than on small ones.
- 2) Image analysis shows that all the techniques underestimate spatial continuity and size of the low-rank geological elements (bodies composed of one hydrofacies), in the case of the simulations realized separately for A and B.
- 3) The connectivity analysis performed on the volume simulated with SISIM for Units A and B separately was considered. In this case, the S hydrofacies is more abundant in Unit A than in Unit B. Nonetheless it looks to be well connected among the whole volume (high Total Connectivity in unit A, high Intrinsic Connectivity in both Units A and B). On the contrary, the SG hydrofacies is more abundant in Unit B than in Unit A where it shows a low connection degree. In the Entire

Volume some areas of well connected F and G hydrofacies have been observed. However, the Total Connectivity is low (Figure 14).

4) The distribution of hydrofacies G (open framework gravels along the lower part of the inclined bed-sets of the composite bars) and F (meter-sized lenses of very fine sand and mud at the top of minor channel elements and decimeter-size mud clasts at their base) is not reproduced by T-ProGS simulations, that yield a scattered pattern of small clusters, sparse in a “matrix” of voxels occupied by facies S and SG. MPS yields better results for background hydrofacies than for the least abundant ones (Figure 13). Simulations by SISIM reproduced rather efficiently the size, shape, distribution and orientation (sloping features of lateral and frontal accreted elements) of these low-hierarchy elements (dell’Arciprete et al. 2010, dell’Arciprete PhD Thesis 2010).

5) The geological model shows non-stationary transition from GS and G hydrofacies association to S and less abundant F hydrofacies moving towards the western and southern part of the volume, where the composite point-bar to channel-fill transitions occur. This trend is only partially reproduced by simulations. Visual inspection of the simulated volumes reveals periodical repetitions of the most permeable hydrofacies G, at a separation distance that is a multiple of the variogram range for SISIM and of the minimum of transition probability for T-ProGS. MPS yields a sort of repetition of S and SG bodies, imitating their shape in the training images. In summary, all the simulation approaches used do not account for the non-stationary architecture of composite bars and channels, loosing their real spatial trends and number of connections.

6) SISIM and T-ProGS do not reproduce the elements of the architectural complexity, like minor channels, erosion bases, etc. This problem affects many pixel-oriented methods of simulation and, in our case, it seems to arise from the fact that the semivariogram and correlation matrix are a bivariate isotropic measure (two-point autocorrelation), and therefore any non-linear correlation structure (e.g., curved surfaces) cannot be reproduced. MPS, on the contrary, can reproduce the shape of the curved structures, but fails to reproduce their internal features at this scale. Moreover, vertical tendencies at the scale of the bed-sets and bed-set groups (in the range from 2 to 4 m), which are evident in the cross-variogram and in the off-diagonal vertical transition-probability plots of the facies maps, are partially lost in the 3D simulation. The representation of such non-stationary

periodicities is still an open issue and cannot be resolved using “classical stationary” semivariogram or Markov chain models (Felletti et al. 2006).

A critical parameter in conditional simulations is the ratio between the number of conditioning data and the number of simulated cells. For the test volume this ratio is equal to 0.105; in this case, MPS and SISIM produce satisfactory simulations, whereas the differences between the simulations performed over the entire volume with the coarse grid (ratio equal to 0.017) and the fine grid (ratio equal to 0.008) show that the use of small cells does not significantly improve the results. On the contrary within MPS the shape and the size of S and SG hydrofacies bodies are partly lost. Differently T-ProGS needs a significant number of object repetitions and gives unsatisfactory results for the test volume and for the Unit B simulate separately.

Simulations of the undivided volume are less realistic than simulations obtained separately for Units A and B, because these units are characterized by very different statistical properties (frequency and correlation of hydrofacies). In order to obtain realistic simulations it is necessary that statistical properties do not vary significantly throughout the studied domain (Falivene et al. 2007).

5 Conclusions

Three different pixel-oriented methods were applied to simulate hydrofacies distribution in a point-bar – channel aquifer analogue: SISIM, based on the facies correlation length derived from the variograms; T-ProGS, based on the transition probabilities, that take into account the juxtaposition of the different facies; MPS, based on the training images. Note however, that the MPS approach used here is not conventional since it is based on the use of 2D training images and a recombination of partial probability density functions, while the traditional approach uses a full 3D training image. Consequently the conclusions drawn on the MPS method may not be general. Still, this work enlightens the difficulties related to the selection of a training image when using MPS in real field applications. From their different characteristics the methods were expected to properly reproduce different features of the sedimentary heterogeneity: the length of the facies bodies for SISIM, the transition between hydrofacies (a proxy of their connectivity) for T-ProGS and the shape of the facies bodies for MPS. The results of the different simulations show the different features of the analyzed volume effectively reproduced by each method. MPS can take in account the shape and size of the most abundant facies

bodies, that must occur at least in two orthogonal training images, whereas SISIM and T-ProGS can reproduce efficiently the distribution of the less abundant facies.

To obtain realistic realizations with T-ProGS some repetitions of the analyzed facies bodies along the directions of the Cartesian axes are necessary. Otherwise transition probabilities statistics cannot be estimated with a sufficient accuracy, as it occurred in the simulations of the test volume.

The simulation methods, specifically SISIM and T-ProGS, yield more realistic results if Units A and B are simulated separately, because pixel-oriented simulations work better if the statistic properties of the simulated bodies are homogeneous (Falivene et al. 2007).

The geological model shows a well-defined trend along the NE direction, with a transition from prevailing gravels to fine-grained sediments (fine sands). The simulation methods used in this study take into account this non-stationary behavior only indirectly, since it is introduced through the data of the conditioning faces. The methods reproduce stationary volumes in such a way that the characteristics of the random repetitions of the facies bodies depend on the simulation method and on the parameters that characterize the underlying geostatistical model. In fact, the separation between facies bodies varies over multiples of the variograms ranges (for SISIM) or of the minimum transition probability (for T-ProGS). In the MPS simulations the shapes of these bodies are similar to those appearing in the training images. The increasing abundance of the gravelly hydrofacies G in the upper part of the volume is better reproduced by MPS simulations, whereas it can be reproduced by SISIM and T-ProGS only if Units A and B are simulated separately.

For an efficient pixel oriented conditional simulation a significant number of conditioning data is needed. The simulations of the *entire volume* with *coarse grid*, for example, are more realistic than the simulations of the *entire volume* with *fine grid*, even if the database has a better accuracy than the discretization grids. This fact is an effect of the ratio between the number of the conditioning data and the number of simulated cells, which is much greater for the coarse grid than for the small grid (1.7% vs. 0.8%). With the coarse grid simulations, however, the features of the low hierarchic level are lost.

Acknowledgements

This work was financially supported by the Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR) and the University of Milan through the research project of national interest "Integrated

geophysical, geological, petrographical and modeling study of alluvial aquifer complexes characteristic of the Po plain subsurface: relationships between scale of hydrostratigraphic reconstruction and flow models” (PRIN 2007).

References

- Alabert FG, Aquitaine E, Modot V (1992) Stochastic Models of Reservoir Heterogeneity: Impact on Connectivity and Average Permeabilities. Society of Petroleum Engineers, 24893.
- Allard D (1994) Simulating a geological lithofacies with respect to connectivity information using the truncated Gaussian model. In: Armstrong M, Dowd PA (eds) Geostatistical Simulations: Proceedings of the Geostatistical Simulation Workshop, Fontainebleau, France, 27-28 May 1993. Kluwer, Norwell, MA.
- Allard D, Group H (1993) On the connectivity of two random set models: the truncated Gaussian and the Boolean. In: Soares A (ed) Geostatistics Tróia '92. Kluwer, Dordrecht, The Netherlands.
- Bersezio R, Giudici M, Mele M (2007) Combining sedimentological and geophysical data for high-resolution 3-D mapping of fluvial architectural elements in the Quaternary Po plain (Italy). *Sedimentary Geology* 202: 230–248.
- Bordley R (1982) A multiplicative formula for aggregating probability assessments. *Management Science* 28: 1137-1148
- Cabello P, Cuevas JL, Ramos E (2007) 3D modelling of grain size distribution in Quaternary deltaic deposits (Llobregat Delta, NE Spain). *Geologica Acta* 5: 231-244.
- Caers J (2001) Geostatistical reservoir modelling using statistical pattern recognition. *Journal of Petroleum Science and Engineering* 29: 177–188.
- Carle SF (1999) T-PROGS: Transition Probability Geostatistical Software version 2.1. University of California, Davis.
- Carle SF, Fogg GE (1996): Transition probability-based indicator geostatistics. *Mathematical Geology* 28: 453-477.
- Carle SF, Labolle EM, Weissmann GS, Van Brocklin D, Fogg GE (1998) Conditional Simulation of Hydrofacies Architecture: A Transition Probability/Markov Approach. In: Fraser GS, Davis JM (eds.) *Hydrogeologic models of sedimentary aquifers*. SEPM Special Publication, Concepts in

- Hydrogeology and Environmental Geology: 147-170.
- Comunian A, Renard P, Straubhaar J, Bayer, P (2010) Three-dimensional high resolution fluvio-glacial aquifer analog: Part 2: geostatistical modeling. *Journal of Hydrology*, submitted
- dell'Arciprete D, Felletti F, Bersezio R (2010) Simulation of fine-scale heterogeneity of meandering river aquifer analogues: comparing different approaches. In: Atkinson PM, Lloyd CD (eds.) *geoENV VII – Geostatistics for Environmental Applications: Quantitative Geology and Geostatistics*, Springer 16: 127-137.
- Deutsch CV, Journel, A (1992) *GSLIB: Geostatistical Software Library*. Oxford Univ. Press.
- Deutsch CV (2002) *Geostatistical Reservoir Modeling*. Oxford University Press, New York.
- Deutsch CV, Journel AG (1998) *GSLIB: Geostatistical Software Library: and User's Guide*, second ed. Oxford University Press, New York, NY.
- Falivene O, Cabrera L, Muñoz JA, Arbués P, Fernández O, Sáez A (2007) Statistical grid-based facies reconstruction and modelling for sedimentary bodies. Alluvial-palustrine and turbiditic examples. *Geologica Acta* 5: 199-230.
- Felletti F, Bersezio R, Giudici M (2006) Geostatistical simulation and numerical upscaling, to model groundwater flow in a sandygravel, braided river, aquifer analogue. *Journal of Sedimentary Researchs* 76: 215–1229.
- Fleckenstein JH, Niswonger RG, Fogg GE (2006) River-Aquifer Interactions, Geologic Heterogeneity, and Low-Flow Management. *Ground Water* 44(6): 837–852.
- Fogg GE, Noyes CD, Carle SF (1998) Geologically-based model of heterogeneous hydraulic conductivity in an alluvial setting. *Hydrogeol. Journal* 6: 131–43.
- Giudici M, Vassena C (2008) Spectral Analysis of the Balance Equation of Ground Water Hydrology. *Transp Porous Med* 72: 171–178.
- Gómez-Hernández JJ, Wen XH (1998) To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. *Adv Water Resour* 21: 47–61.
- Goovaerts P (1997) *Geostatistics for Natural Resources Evaluation*. Oxford University Press, Oxford.
- Guardiano F, Srivastava RM (1993) Multivariate geostatistics: Beyond bivariate moments. In Soares A (ed) *Geostatistics-Troia*. Kluwer Academic, Dordrecht: 133–144.
- Hu LY, Chuginova T (2008) Multiple-point geostatistics for modeling subsurface heterogeneity: A

- comprehensive review. *Water Resources Research* 44, W11413.
- Journel AG, Gómez-Hernández JJ (1993) Stochastic Imaging of the Wilmington Clastic Sequence. *Society of Petroleum Engineers Formation Evaluation* March: 33-40.
- Journel AG, Gundersen R, Gringarten E, Yao T (1998) Stochastic modelling of a fluvial reservoir: a comparative review of algorithms. *of Petroleum Science and Engineering* 21: 95-121.
- Knudby C, Carrera J (2005) On the relationship between indicators of geostatistical, flow and transport connectivity. *Adv Water Resour* 28: 405–421.
- Liu Y, Harding A, Abriel W, Strebelle S (2004) Multiple-point simulation integrating wells, 3D seismic data and geology. *American Association of Petroleum Geologists Bulletin* 88: 905–922.
- Liu Y, Harding A, Gilbert R, Journel A (2005) A Workflow for Multiple-point Geostatistical Simulation. In *Geostatistic Banff 2004*: 245-254.
- Okabe H, Blunt MJ (2005) Pore space reconstruction using multiple-point statistics. *Journal of Petroleum Science and Engineering* 46: 121– 137.
- Renard Ph, Caers J (2008) Conditioning facies simulations with connectivity data. In: *Proceedings of the VIII Geostatistical Congress*. Santiago, Chili, December 2008.
- Ritzi RW (2000) Behaviour of indicator variogram and transition probabilities in relation to the variance in lengths of hydrofacies. *Water Resources Research* 36: 3375-3381.
- Seifert D, Jensen JL (1999) Using Sequential Indicator Simulation as a tool in reservoir description: Issues and Uncertainties. *Geology* 31: 527-550.
- Straubhaar J, Renard P, Mariethoz G, Froidevaux R, Besson O (2010) An improved parallel multiple-point algorithm. *Mathematical Geosciences*, in press.
- Strebelle S (2002) Conditional simulation of complex geological structures using multiple-point statistics. *Math. Geol.* 34: 1–21.
- Strebelle S, Payrazyan K, Caers J (2002) Modeling a Deepwater Turbidite Reservoir conditional to seismic data using multiple-point geostatistics. *Society of Petroleum Engineers*, 77425.
- Strebelle S, Payrazyan K, Caers J (2003) Modeling of a deepwater turbidite reservoir conditional to seismic data using principal component analysis and multiple-point geostatistics. *SPE Journal* 8: 227–235.
- Sweet ML, Blewden CJ, Carter AM, Mills CA (1996) Modeling heterogeneity in a low-permeability gas

reservoir using geostatistical techniques, Hyde Field, southern North Sea. AAPG Bulletin 80: 1719-1735.

Vassena C, Cattaneo L, Giudici M (2009) Assessment of the role of facies heterogeneity at the fine scale by numerical transport experiments and connectivity indicators. Hydrogeology Journal, doi:10.1007/s10040-009-0523-2.

Weissmann GS, Carle SF, Fogg GE (1999) Three-dimensional hydrofacies modeling based on soil surveys and transition probability geostatistics. Water Resources Research 35: 1761-1770.

Weissmann GS, Fogg GE (1999) Multi-scale alluvial fan heterogeneity modeled with transition probability geostatistics in a sequence stratigraphic framework. Journal of Hydrology 226: 48-65.

Western A, Blöchl G, Grayson RB (2001) Toward capturing hydrologically significant connectivity in spatial patterns. Water Resour Res 37: 83–97.

Zappa G, Bersezio R, Felletti F, Giudici M (2006) Modeling aquifer heterogeneity at the facies scale in gravel-sand braided stream deposits. Journal of Hydrology 325: 134-153.

Zinn B, Harvey CF (2003) When good statistical models of aquifer heterogeneity go bad: a comparison of flow, dispersion, and mass transfer in connected and multivariate Gaussian hydraulic conductivity fields. Water Resour Res 39: 1051.

Tables

Test Volume Facies Proportions				
	Input parameters	SISIM	T-ProGS	MPS
F	18.0 %	6.6% ±1.3%	17.9%±0.2%	0.5%±0.2%
S	29.0 %	30.6%±3.7%	29.2%±0.2%	48.4%±4.8%
SG	43.0 %	49.4%±3.5%	42.9%±0.3%	39.5%±3.9%
G	10.0 %	13.4%±1.0%	10.0%±0.1%	11.6%±1.8%

Table 1: Input and output facies proportions in the Test Volume. The output proportions for the three simulation methods are expressed as (mean ± standard deviation) over the ensemble of 50 realizations.

Entire Volume Coarse Grid Facies Proportions				
	Input parameters	SISIM	T-ProGS	MPS
F	6.0 %	6.8% ±1.2%	6.8%±1.0%	2.6%±0.2%
S	61.0 %	52.7%±3.1%	59.8%±1.1%	67.4%±1.4%
SG	28.0 %	32.9%±2.5%	28.1%±1.0%	27.7%±1.4%
G	5.0 %	7.6%±1.1%	5.3%±0.3%	2.3%±0.2%

Table 2: Input and output facies proportions in the Entire Volume, coarse grid. The output proportions for the three simulation methods are expressed as (mean ± standard deviation) over the ensemble of 10 realizations.

Entire Volume Fine Grid Facies Proportions (undivided volume)					
	Input parameters		SISIM	T-ProGS	MPS
F	6.0 %		7.1%	6.8%	2.8%
S	61.0 %		58.7%	59.3%	49.5%
SG	28.0 %		28.8%	28.8%	43.5%
G	5.0 %		5.4%	5.1%	4.2%
Entire Volume Fine Grid Facies Proportions (unit A + unit B)					
	Input parameters		SISIM	T-ProGS	MPS
	unit A	unit B			
F	5.0%	0.0%	3.5%	3.8%	3.1%
S	80.0%	22.0%	56.8%	57.4%	50.1%
SG	13.0%	67.0%	33.1%	32.8%	44.8%
G	11.0%	2.0%	6.6%	6.0%	2.0%

Table 3: Facies proportions of the entire volume simulated SISIM, T-ProGS and MPS.

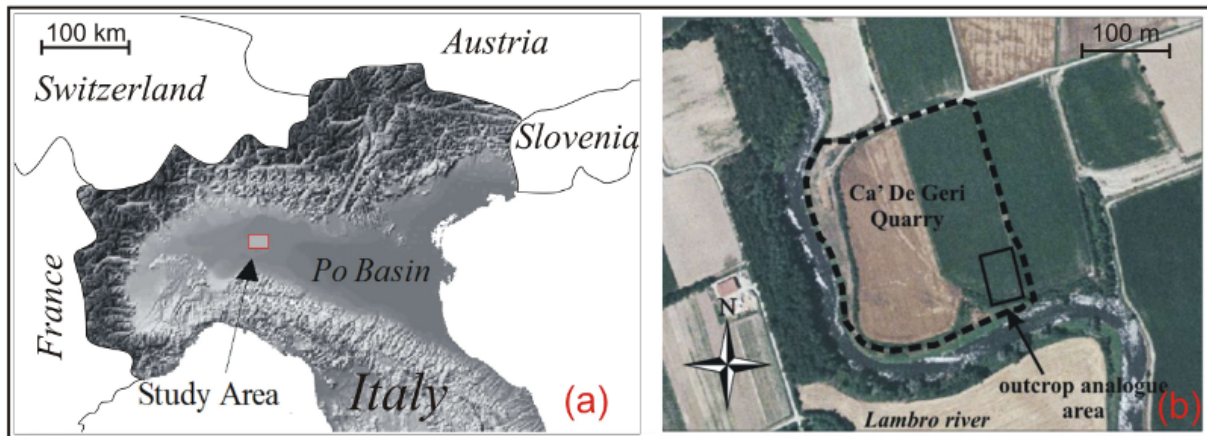


Fig. 1: Location map of the case study.

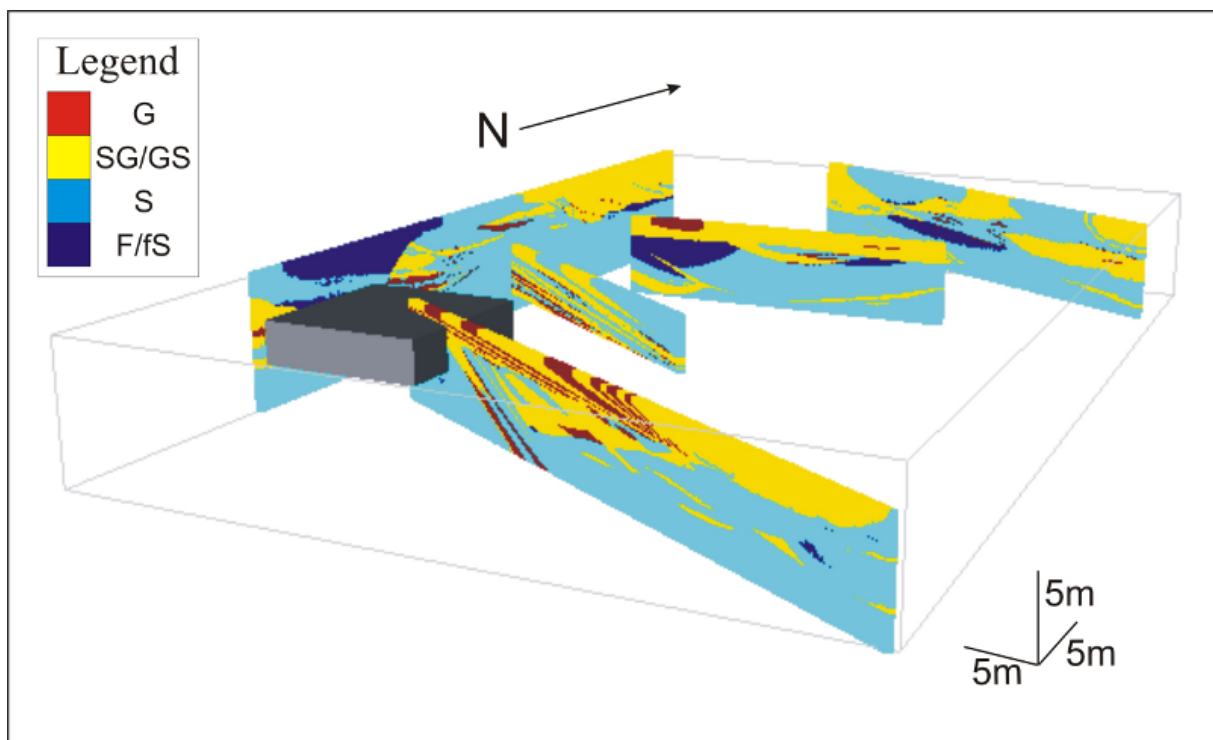


Fig. 2: Domain of the geostatistical simulation. Discretization of the facies maps after hydrofacies classification and location of the test volume are shown.

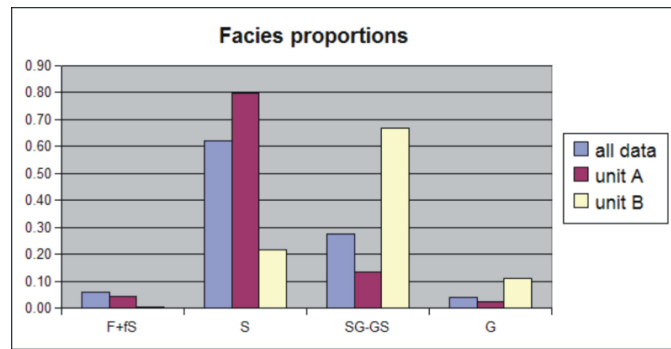


Fig. 3: Facies proportions computed in the conditioning faces for the undivided volume and for Units A and B separately.

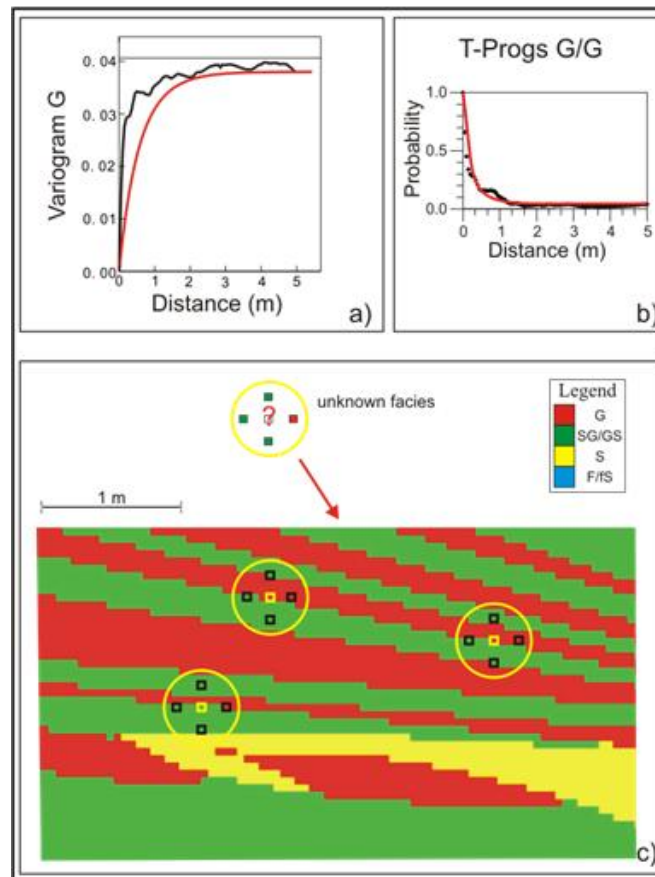


Fig. 4: Example of semivariograms (a) and transitional probabilities (b) for the hydrofacies G (gravel) along the vertical direction; the experimental curves are drawn in black and the fitted models in red. (c) Example of a training image (portion of face 1).

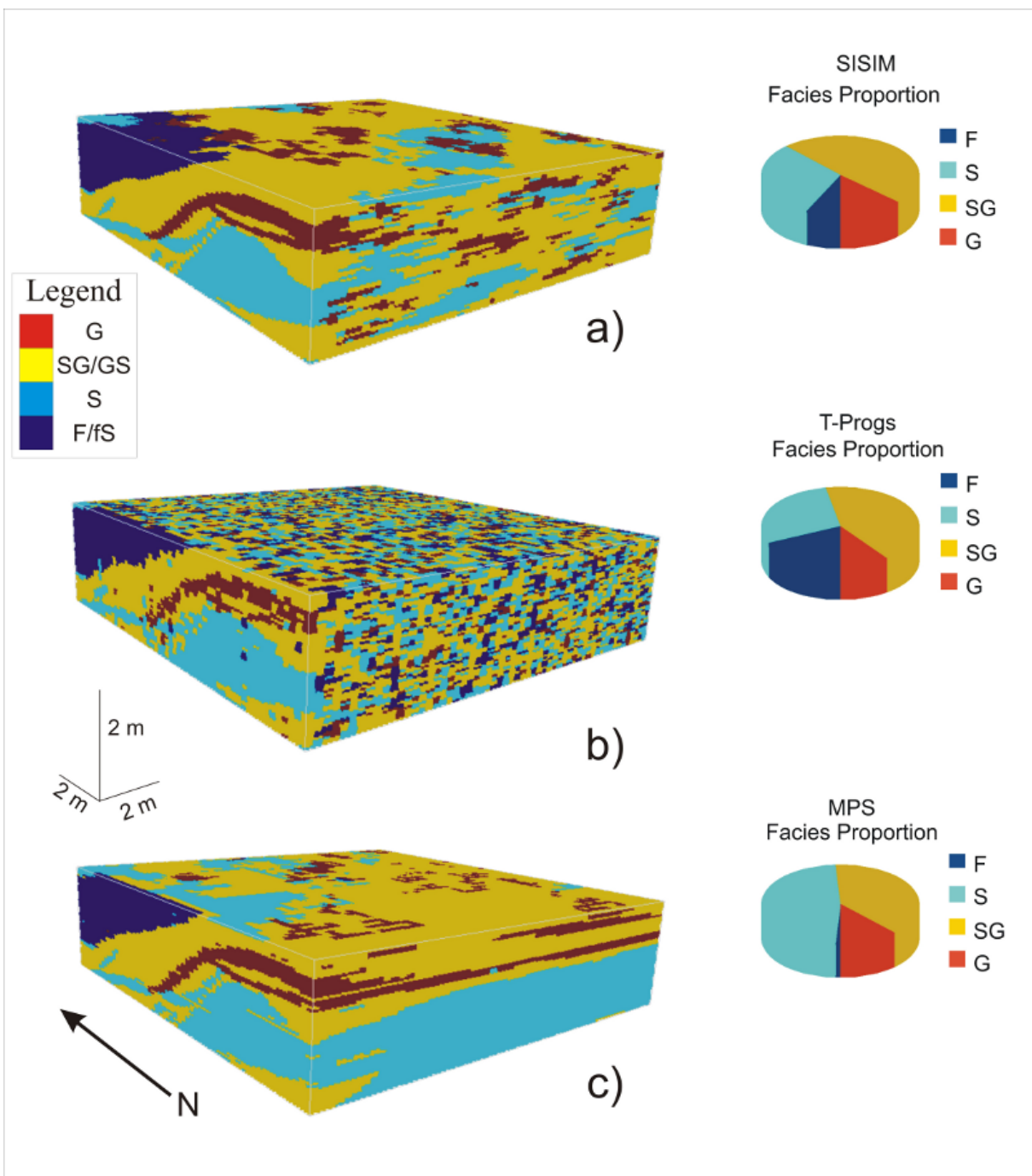


Fig. 5: Randomly selected simulations computed in the test volume with (a) SISIM (b) T-ProGS and (c) MPS. The facies proportions in the simulated volumes are presented in the pie charts.

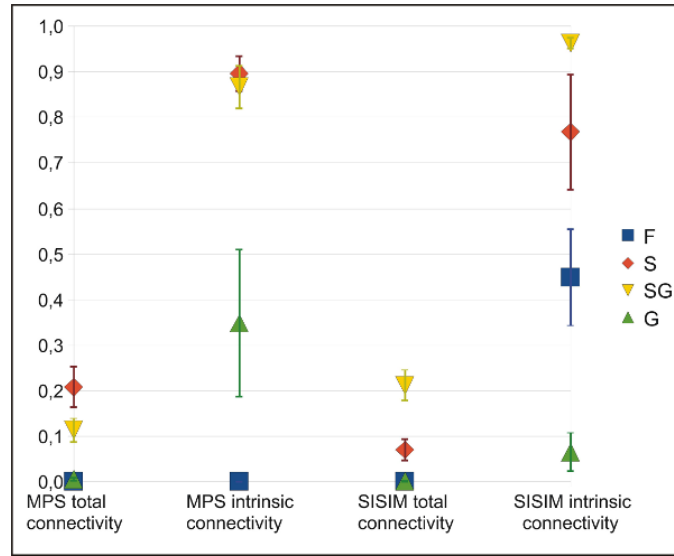


Fig. 6: Total and Intrinsic connectivity (mean $\pm$ st.deviation) computed on the test volume for the 50 simulations realized with SISIM and MPS.

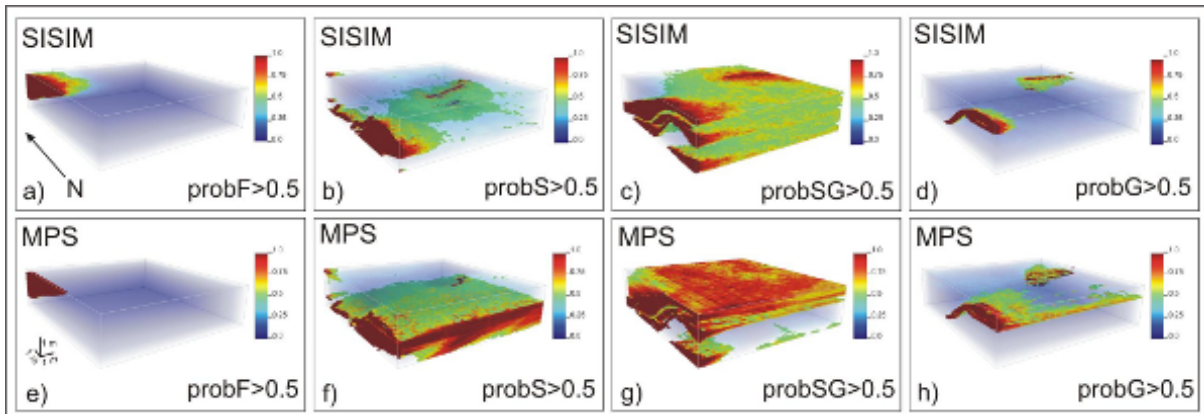


Fig. 7: Probability of finding each facies for the test volume with SISIM - a) facies F, b) facies S, c) facies SG, d) facies G - and with MPS - e) facies F, f) facies S, g) facies SG, h) facies G. The drawn volumes corresponds to the regions where the probability is greater than 0.5.

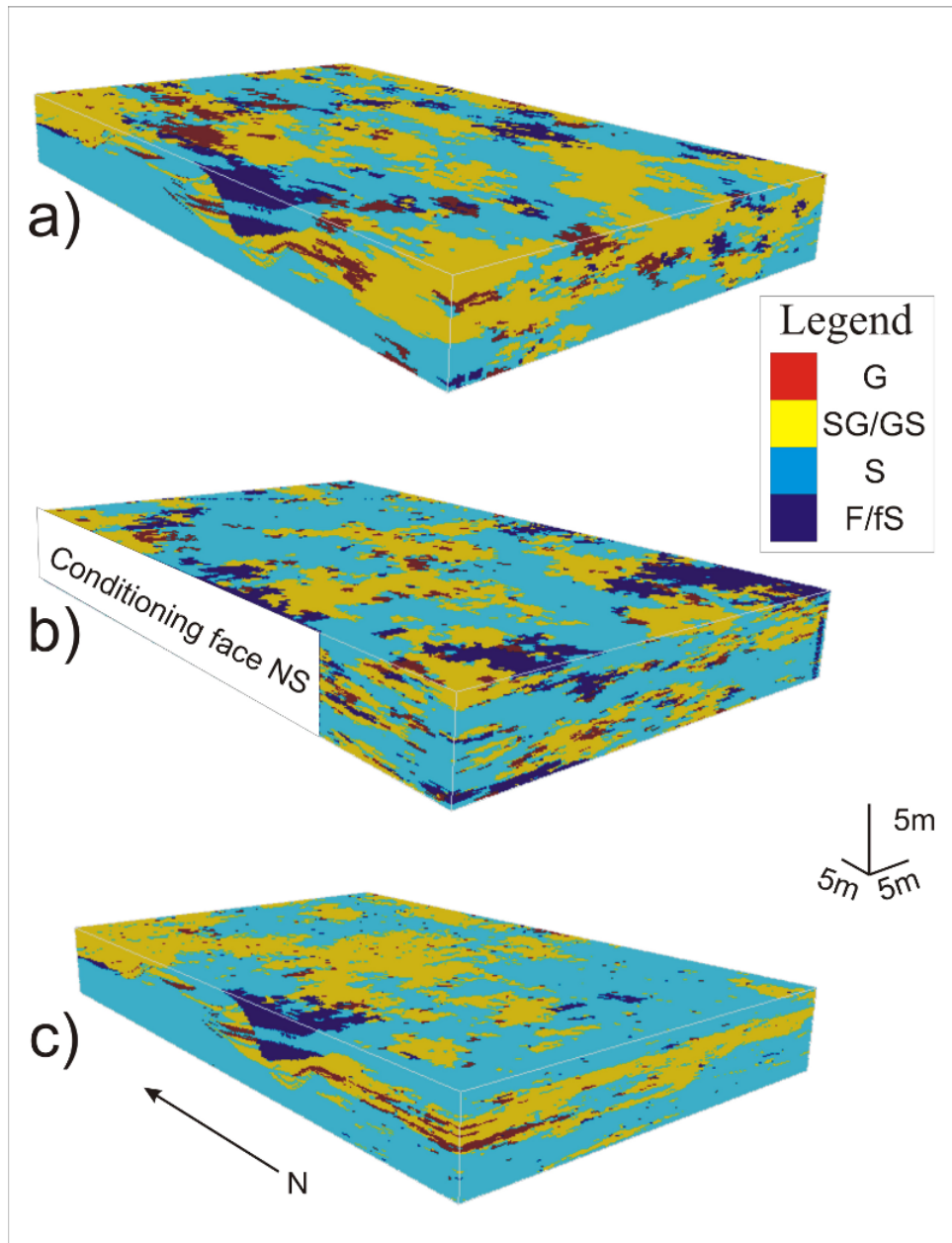


Fig. 8: Equiprobable simulation number 1 of 10, computed in the entire volume (0.4 m by 0.4 m by 0.1 m cells) with (a) SISIM (b) T-ProGS and (c) MPS.

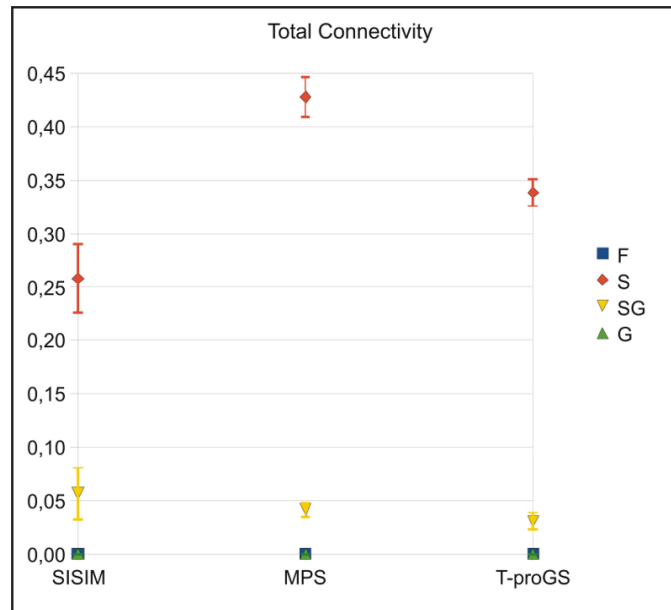


Fig 9: Total Connectivity computed in the Entire Volume simulated with SISIM, MPS and T-Progs with Coarse Grid. The y axis is in logarithmic scale.

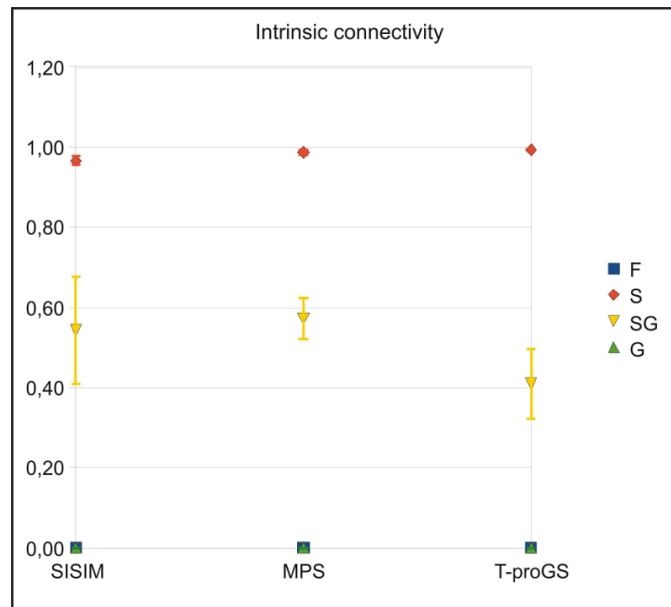


Fig. 10: Intrinsic Connectivity computed in the Entire Volume simulated with SISIM, MPS and T-Progs with Coarse Grid. The y axis is in logarithmic scale.

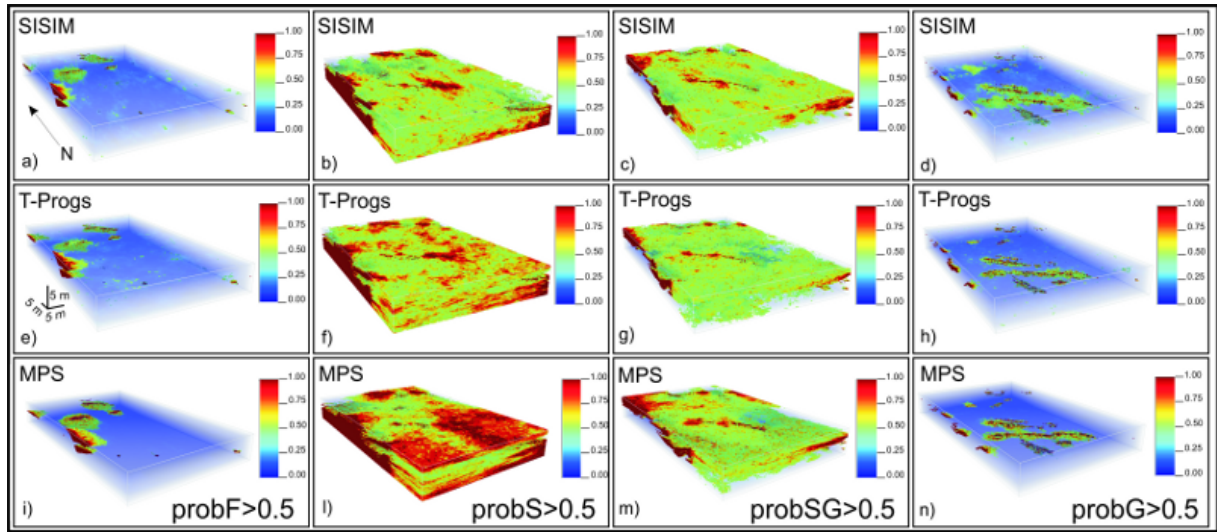


Fig. 11: Probability of finding each facies for the entire volume (coarse grid) with SIS - a) facies F, b) facies S, c) facies SG, d) facies G - with T-ProGS - e) facies F, f) facies S, g) facies SG, h) facies G - and with MPS - i) facies F, l) facies S, m) facies SG, n) facies G. The drawn volumes correspond to the regions where the probability is greater than 0.5.

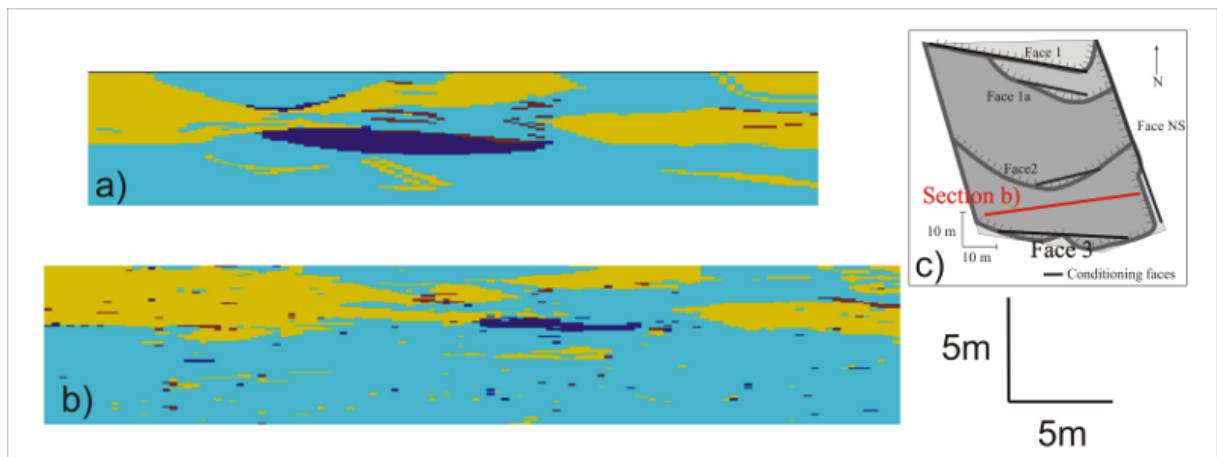


Fig. 12: Training image face 3 discretized with coarse grid and b) section cut into the MPS simulated volume (coarse grid, realization 4). In c) the positions of the face 3 and the section are reported. From these pictures the similarity between the shape of the S and SG hydrofacies bodies in the simulated volume and in the training image are evident.

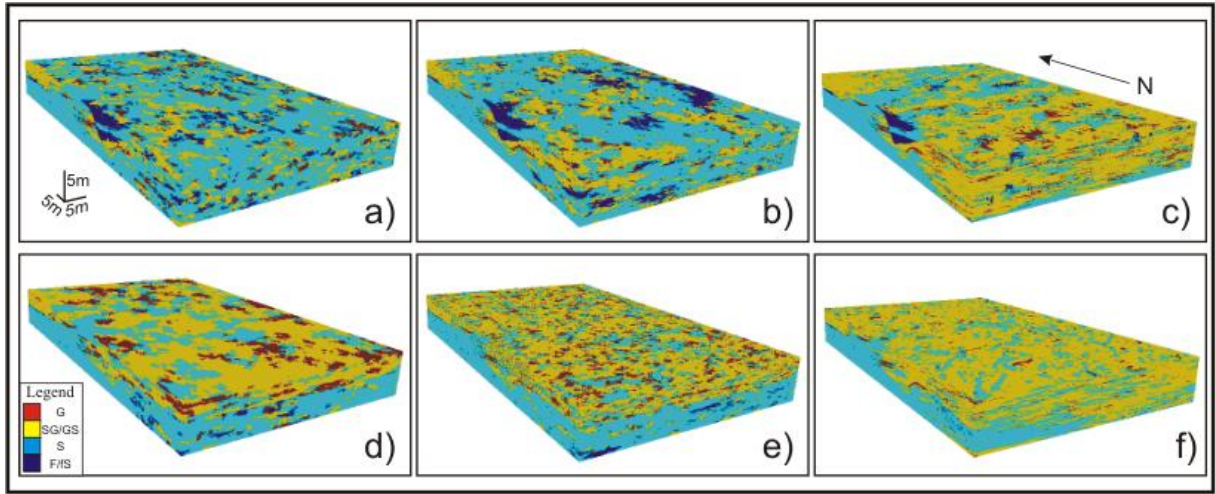


Fig. 13: Simulations of the undivided entire volume (0.2 m by 0.2 m by 0.05 m cells) computed with (a) SISIM (b) T-ProGS and (c) MPS. Simulations of units A and B computed separately using (d) SISIM (e) T-ProGS and (f) MPS.

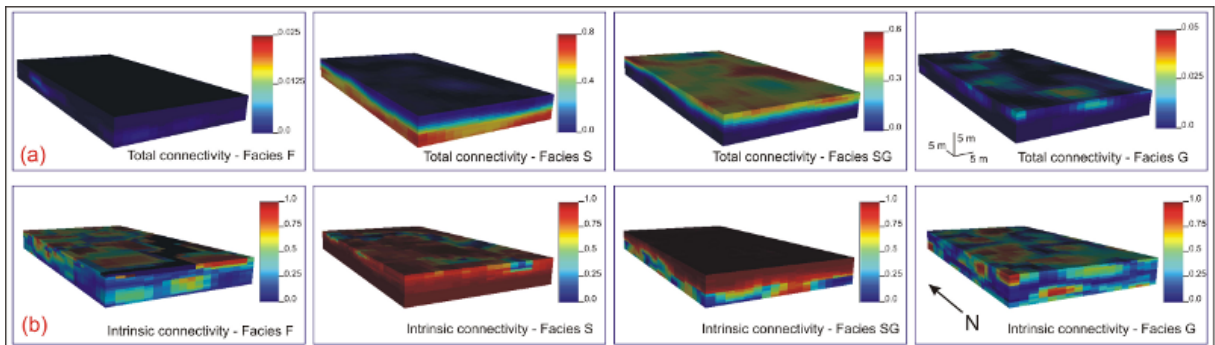


Fig. 14: (a) Total and (b) Intrinsic connectivity computed in the SISIM simulated volume for moving blocks of 57 x 57 x 57 cells.

Appendix E

Preliminary bi-dimensional calibration, micro CT image

This appendix contains the details about the preliminary study performed in order to select in 2D a suitable set of parameters (multi-grid level and size of the search template) to be used for the 3D simulations on the micro-computed tomography image considered in chapter 3.

For each one of the ten combinations of different multi-grid levels and search templates listed in table E.1, 100 realizations of size 400×400 are obtained and compared. A 2D section of the micro-computed tomography of dimensions 400×400 is used as training image and as reference (fig. E.1 and 3.9(b)).

The results are compared in terms of number of geobodies (tab.E.1), facies proportions (tab.E.1, fig.E.2(c)), connectivity functions (fig.E.2), size of the geobodies along the coordinate x (fig.E.2(b)), area of the geobodies (fig.E.2(a)), Euler number (fig.E.2(d)), and visually (fig.E.1).

The results closer to the target training image are obtained with the parameter sets (c),(d) and (h), and the parameters set used for the simulation in 3D is the set (d).

test case	multi-grids number	template size [px]	simulation time [sec]	geobodies number	vugs proportion [%]
(a)	1	7×7	67	533	18.06
(b)	2	7×7	92	266	18.51
(c)	3	7×7	107	226	21.31
(d)	4	7×7	100	210	20.66
(e)	5	7×7	99	203	20.58
(f)	6	7×7	101	185	21.86
(g)	5	3×3	< 1	297	19.32
(h)	5	5×5	12	204	19.75
(i)	5	9×9	367	198	21.33
(j)	6	variable	4	305	21.24
ref.				202	20.13

Table E.1: The parameters used to perform the MPS simulation for the micro CT image in the 2D case. For the case study (j) the dimensions of the search template varies for each multi-grid level: the templates are square with a side of, 7, 7, 5, 5, 3 and 3 pixels from the first to the last multi-grid level. *WARNING: The parameters are computed only on the first of 100 simulations. Maybe for these parameters I could use the mean value...*

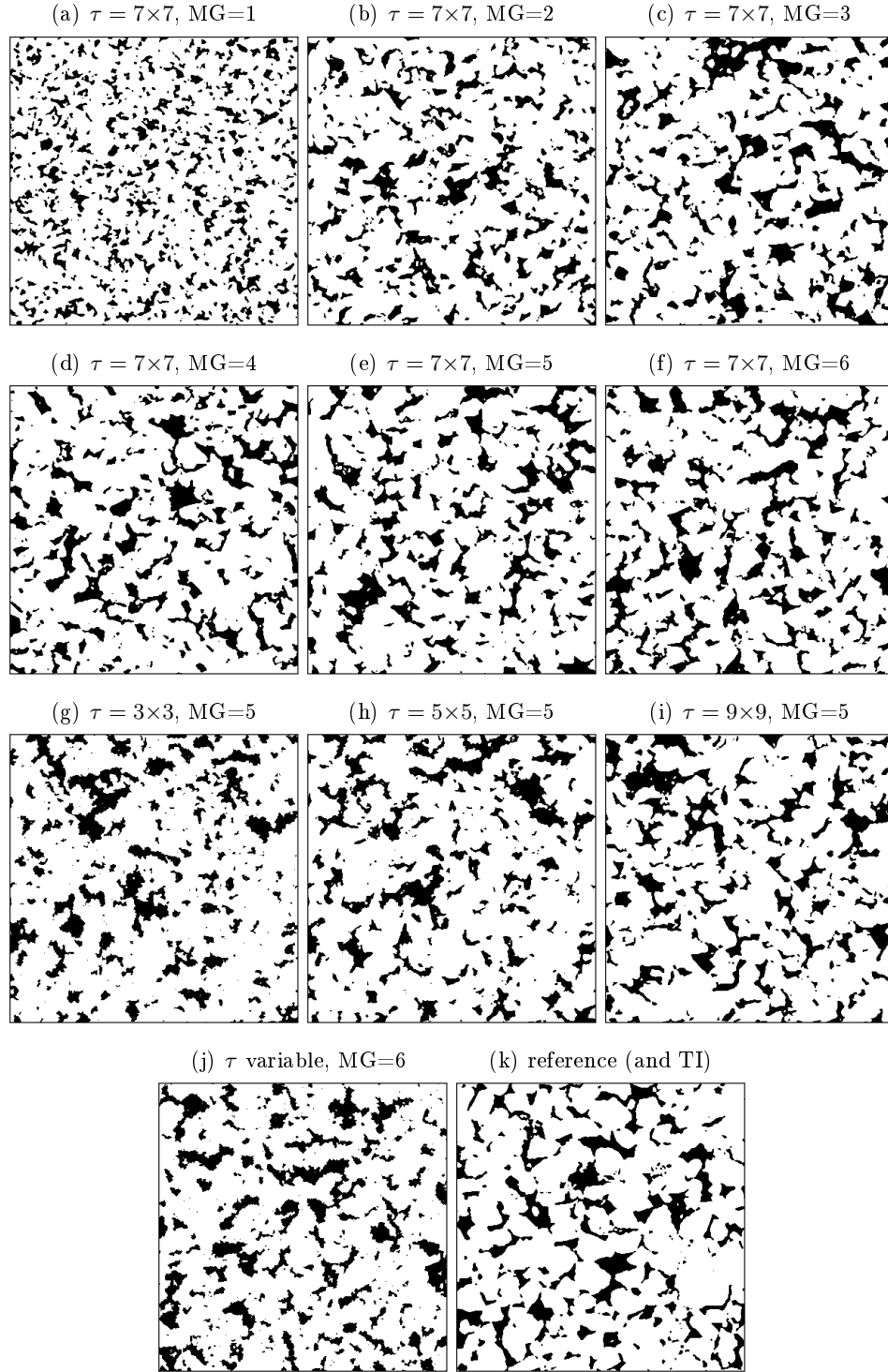


Figure E.1: Result of 2D multiple-point simulations performed with the parameters of tab. E.1 for the micro CT image with different search templates (τ) and multi-grid levels (MG). Image (k) is used as reference and as training image.

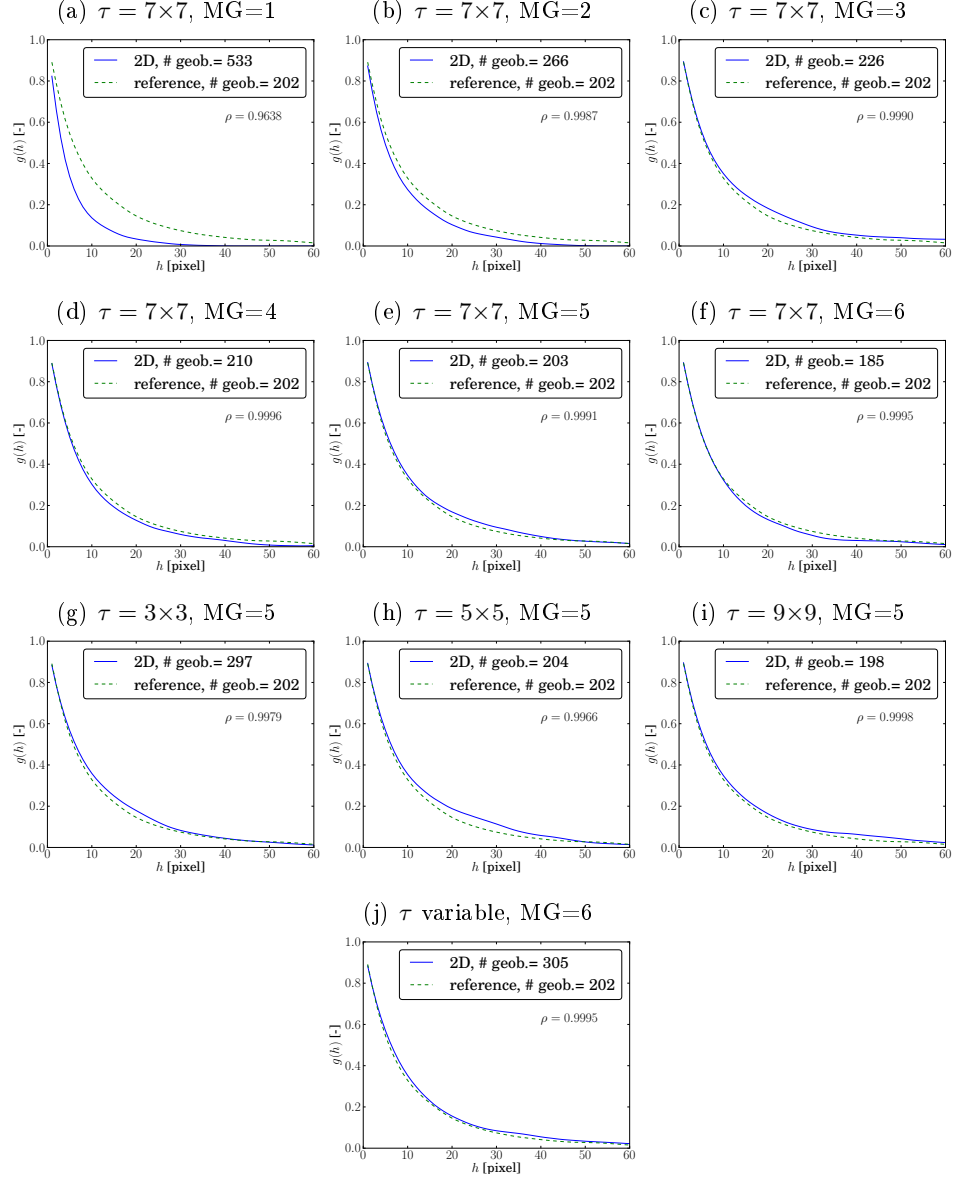


Figure E.2: Connectivity functions along the x axis for the 2D tests performed for the micro CT image. The dashed curve is the function computed on the reference training image. The correlation coefficient between the curves and the number of geobodies are reported.

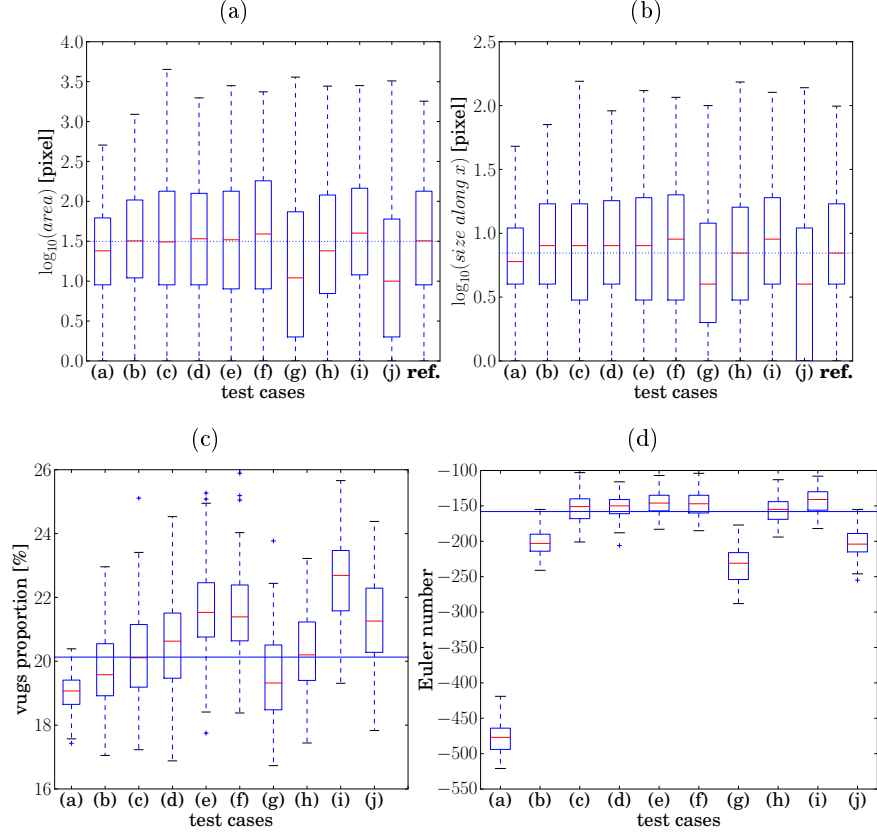


Figure E.3: For the 2D simulations of the micro CT image for each parameters set of table E.1, (a) $\log_{10}$ of the area of the geobodies (the dashed line is the median computed for the reference training image); (b) size of the geobodies along the x axis (the same trend can be observed along the y); (c) mean of vugs proportions for 100 simulations for each parameters set (the horizontal line is the value for the reference image); (d) the Euler number (the horizontal line represents the Euler number for the reference).

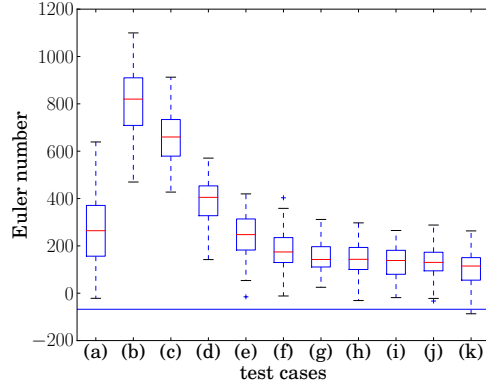


Figure E.4: The Euler number computed for different values of the weighting factors $w_{1,2}$ for the simulations with the **paBor** method.

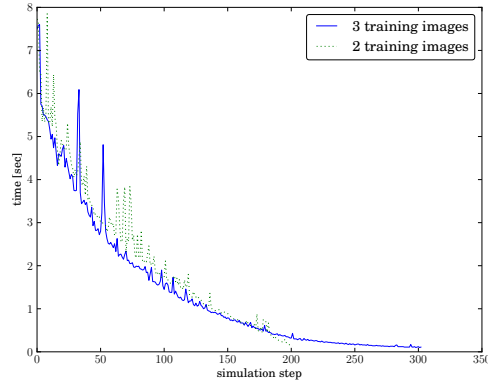


Figure E.5: The CPU time required for the simulation in 3D of the micro CT image for each simulation step for the method **s2Dcd**. Here the case with two and three training images are compared. The total computing time with three training image is smaller than the one with two training images.

Appendix F

Simulation of 3D channels with the method s2Dcd

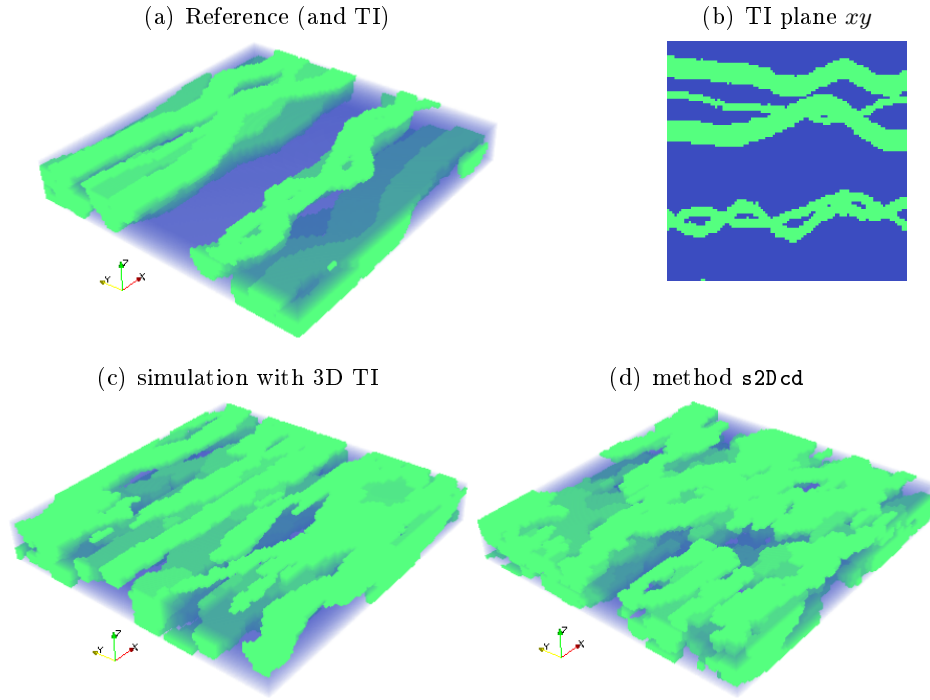


Figure F.1: Channels simulation: (a) 3D training image; (b) one of the 2D sections used as training images; (c) simulation obtained using the 3D training image; (d) simulation with the **s2Dcd** method.

This appendix contains the results of simulations performed in order to reproduce with the method **s2Dcd**, illustrated in chapter 3, 3D training images containing channel structures.

Some tests are performed in order to test the proposed methods for the simulation of channel bodies and to compare these results with the results of previous works (1).

A first test is performed using two 2D training images extracted as slices of a 3D training image containing channels (fig.F.1).

Another test is performed with the same training images used by Caers (1) (fig.F.2(a) and (b))

Bibliography

- [1] Caers, J. (2006). A general algorithm for building 3D spatial laws from lower dimensional structural information. Technical report, Stanford Center for Reservoir Forecasting.
- [2] Deutsch, C. and Journel, A. (1997). *GSLIB: Geostatistical Software Li-*

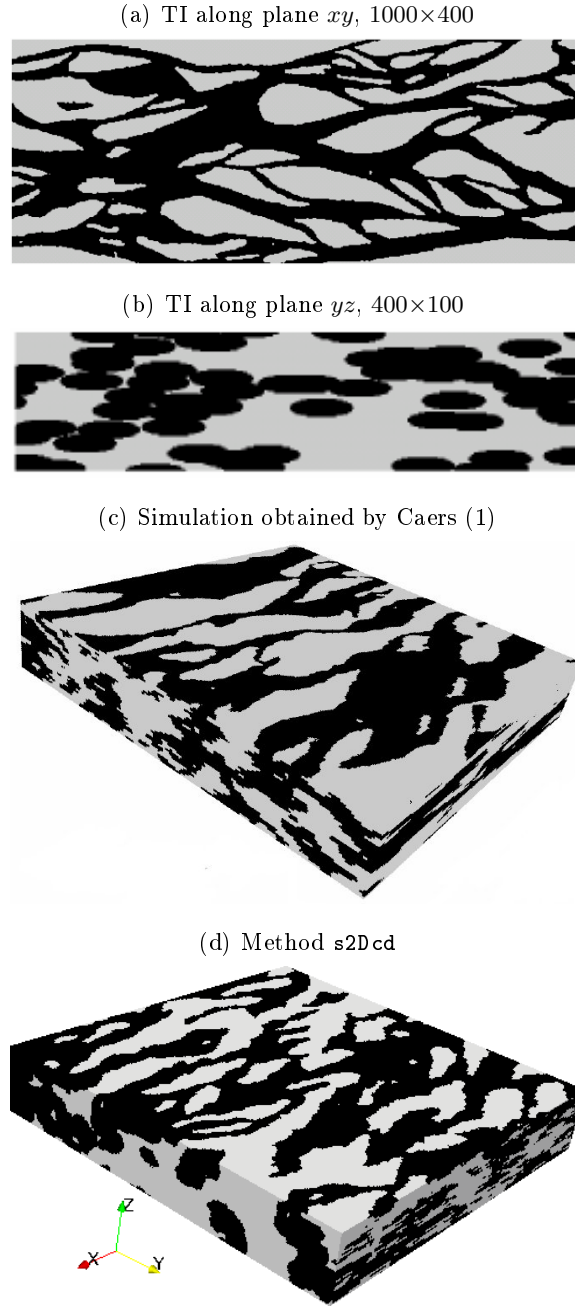


Figure F.2: (a) Bi-dimensional TI of a braided system in New Zealand; (b) bi-dimensional TI obtained with an object based method (b, ellipsim 2); (c) simulation obtained by Caers (1); (d) simulation obtained with the $s2Dcd$ method. The images (c) and (d) are scaled 2 times along the z coordinate.

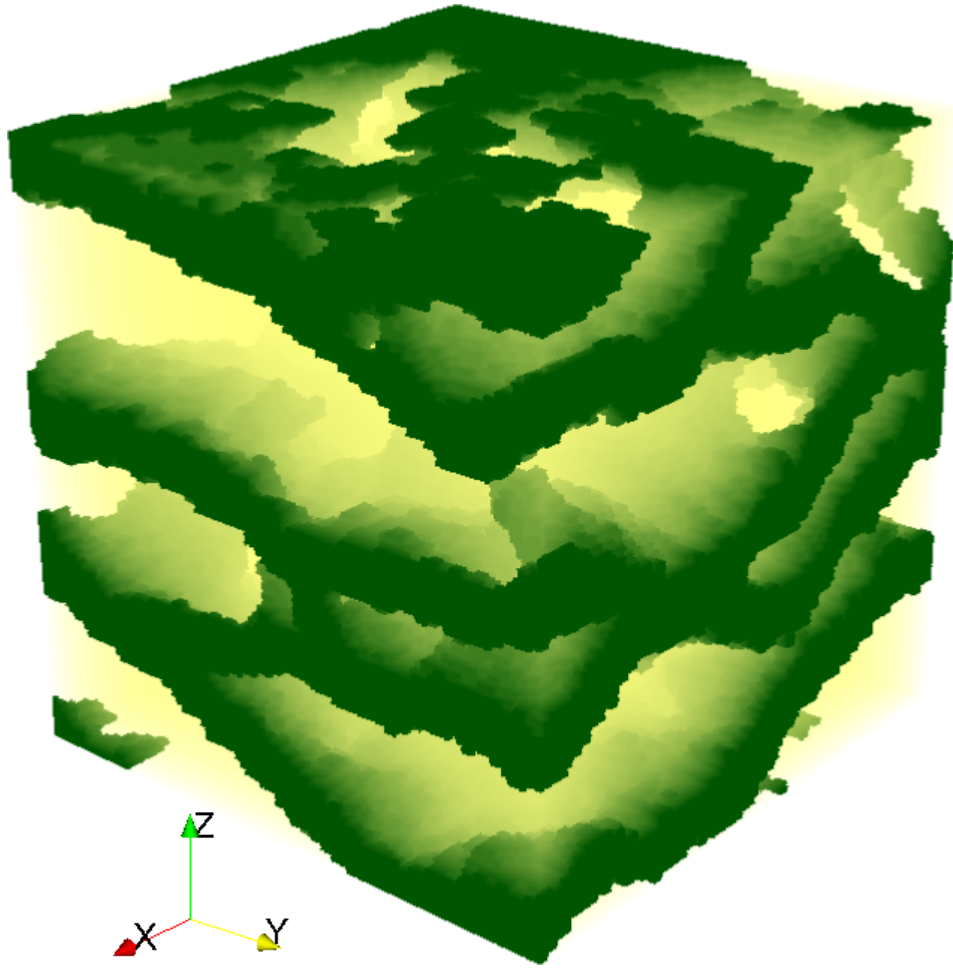


Figure F.3: A simulation $100 \times 100 \times 100$ performed with the `s2Dcd` method using along the planes xz and yz the 2D training image with channels by Strebel (3).

brary and User's Guide. Applied geostatistics series. Oxford University Press, second edition.

- [3] Strebelle, S. (2002). Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, 34:1–21.

Ringraziamenti

Grazie alla mia famiglia, Dario, Laura, Karin e Thomas (ma anche nonna Guglielmina, Itala, Lia), per avermi sempre sostenuto. A Linda, che nonostante la lontananza spaziale ha sempre collaborato per tenerci vicini col cuore; grazie anche alla sua famiglia. Grazie ad Alberto Fusi, per la sua amicizia, e in collaborazione con Valeria per la nipota! Agli “Aros”, in particolare Andrea Milanese e Marco Gargano. Al Prof. Mauro Giudici, per avermi sempre accompagnato ed indicato molte vie, in questo cammino di ricerca. Ad Andrea Borghi, che nei momenti di distrazione ha saputo sempre rallegrarmi, per il comune amore per “il pinguino”, per Benny Lava e i motivetti che rimbalzano per la mente tutto il giorno. A Marco Bonamore, per avermi trasmesso un po’ della sua passione per l’arrampicata e per il suo codice per la semplificazione di immagini. A Fabio Oriani, per χ_E e per aver risvegliato turbe musicali. A Giona Preisig, Diana dell’Arciprete, Luca Guglielmetti e a tutti gli amici italofofi... grazie!

Remerciements

Merci beaucoup à Philippe Renard, pour son enthousiasme, son amitié, sa patience, son aide... merci! Merci beaucoup à Julien Straubhaar, pour les mêmes raisons que Philippe, mais aussi pour ses scripts, sources continues d'inspiration. Merci aussi aux amies des anciennes et nouveaux pauses, Jaouher Kerrou, Antoine Baillieux, Rob de Roji, Damian Glenz, Ellen Milnes, Michiel Pronk, Nicolas Coppo, Fabien Cornaton, Grégoire Mariethoz, Pascale et Romain Ducommun, Francois Gainon, Xenia Boutsiadou, Fairouz Slama, Vivian Gremaud, Andrés Alcolea et tous les amis du CHYN! Merci *al Maestro* Marco Borel, pour avoir partagé avec moi son amour pour la musique et la flute traversière. J'espère bientôt de reprendre avec régularité à jouer! A Paul Bailleux, pour les moments musicaux et non que on a partagé. A Ludmilla et Corinne, pour la musique des lutins. Et merci à Manu aussi, pour sa mission de "concierge des esprits"!

In realtà un lavoro simile non termina mai.

Lo si deve dichiarare concluso quando,
a seconda del tempo e delle circostanze,
si è fatto il possibile.

Johann Wolfgang von Goethe Italienische Reise (1789)