3D stochastic simulation of caves: application to Saint-Sébastien case study (SE, France)

Sophie Viseur, Johan Jouves, Arnaud Fournillon, Bruno Arfib, Yves Guglielmi

Sophie VISEUR[1], Johan JOUVES[1,3], Arnaud FOURNILLON[2], Bruno ARFIB[1] and Yves GUGLIELMI[1]

(1) Aix Marseille University, CNRS, IRD, CEREGE UM34, Sedimentary and Reservoir Systems Department, Marseille, France
(2) Beicip-Franlab, Rueil-Malmaison, France
(3) Cenote, 1, chemin de Valdegour, 30900 Nîmes
viseur@cerege.fr

3D stochastic simulation of caves: application to Saint-Sébastien case study (SE, France)

ABSTRACT: Karst networks are complex structures that are difficult to observe or survey. However, they have a major impact on fluid flows and water resources as well as on petroleum reservoir behaviour. In such applications, geostatistical approaches are generally used to simulate several 3D models of geological structures, on which risk assessment analyses are performed in terms of reserve assessment or flow behaviour. In this paper, such an approach is proposed to simulate karst networks. The approach is based on the classical method Sequential Indicator Simulation (SIS). It relies on the decomposition of the karst network into conduit families, depending on their orientation and dimensions. Conduit shapes are characterized by variograms and are simulated sequentially from the first to the last family. Conditioning points are randomly simulated to assert the connectivity between each conduit family. Moreover, special attention is paid to the integration of geological and hydrogeological knowledge into the karst network simulation. The proposed approach is applied to the case study of Saint-Sébastien (Gréoux-les-Bains, Provence) to illustrate the possibility of integrating analogue and geological information.

KEYWORDS: geostatistics, cave, network.

RÉSUMÉ : SIMULATION STOCHASTIQUE 3D DE GROTTES : APPLICATION À LA GROTTE DE SAINT-SÉBASTIEN (PROVENCE). Les réseaux karstiques sont des structures complexes et difficiles à observer. Cependant, ils ont un impact majeur sur les écoulements et sur les ressources en eau ou hydrocarbures. Ainsi, dans de telles applications, il est souvent nécessaire de modéliser le réseau karstique afin d'estimer les risques en termes d'écoulement ou de réserve, au vu des incertitudes qui peuvent exister. Deux approches existent pour modéliser les structures géologiques : les approches à imitation de processus et celles à imitation de structures [Koltermann & Gorelick, 1996]. Les premières visent à reproduire les processus de genèse des structures pour les modéliser en 3D. Les secondes ont pour objectif de générer l'organisation spatiale et la géométrie des structures géologiques telles qu'elles sont actuellement. Les méthodes à imitation de processus génèrent par définition des modèles cohérents. Elles sont toutefois coûteuses en temps de calcul et ne peuvent que difficilement être conditionnées à des données ponctuelles. Au contraire, les méthodes à imitation de structures, telles que les méthodes géostatistiques, présentent l'avantage de pouvoir être conditionnées à des données ponctuelles ou globales. Ces méthodes sont aussi très efficaces et peuvent ainsi s'intégrer facilement dans une stratégie de simulation stochastique. En effet, lorsque les données disponibles sont insuffisantes pour reproduire un modèle unique fiable, des techniques de simulations stochastiques sont proposées pour générer un ensemble de modèles équiprobables et possibles au vu des données. Ces modèles représentent un échantillon des possibles sur lequel des études d'incertitudes [Scheidt & Caers, 2009] et de risque [Srivastava, 2013] peuvent être réalisées.

Dans ce papier, nous proposons une méthode à imitation de structures stochastique pour reproduire un réseau karstique. Seuls les conduits karstiques humainement observables sont reproduits, des données statistiques pouvant être déterminées à partir des acquisitions spéléologiques (cartes, cheminements, etc.). Ces conduits karstiques sont simulés dans ce qui est nommé « matrice », qui correspond en fait à l'encaissant des conduits, comprenant la roche ainsi que les discontinuités à petite échelle. L'une des principales préoccupations de ce papier est de montrer comment il est possible d'intégrer des connaissances géologiques pour simuler un réseau karstique en utilisant une technique classique de géostatistique, la Simulation Séquentielle par Indicatrice (SIS). La méthode proposée fonctionne de manière similaire aux travaux proposés dans Fournillon et al. [2011]. Le réseau karstique est décomposé en N familles de conduits $C_i, i \in [1;N]$. Chaque famille de conduits C_i est caractérisée par une orientation et des dimensions (longueur, largeur et épaisseur). Un variogramme d'indicatrice γ_i est associé à chaque famille ainsi qu'une valeur de proportion p_i correspondant au ratio du volume des conduits de la famille C_i sur le volume d'étude. La SIS est ensuite utilisée de manière récursive pour simuler les conduits de chaque famille l'une après l'autre. Toutefois, la méthode proposée dans ce papier diffère de Fournillon et al. [2011] dans sa gestion de la connectivité entre les familles de conduits. En effet, dans Fournillon et al. [2011], la connectivité entre conduits n'est pas directement prise en compte, elle est induite par les valeurs de proportions et la zone de simulation imposées. Dans ce papier, la connectivité est assurée en simulant des points conditionnants dans les corps générés de la famille C_i pour contraindre la génération de la famille C_{i+1} et faire en sorte qu'elle recoupe les conduits C_i.

D'autres données conditionnantes sont aussi proposées pour prendre en compte des informations géologiques ou hydrogéologiques. Une approche est ainsi proposée pour déterminer par approximation la proportion globale p_i d'une famille C_i de conduits à partir des données disponibles. De plus, des techniques sont présentées pour prendre en compte des structures géologiques comme chemin préférentiel des réseaux karstiques (i.e. axe de drainage préférentiel) ou pour imposer des contraintes sur les réseaux karstiques à partir des points connus de sortie et d'entrée du karst.

La méthode proposée ainsi que la manière d'intégrer des connaissances géologiques ont été appliquées sur le cas d'étude de la grotte de Saint-Sébastien à Gréoux-les-Bains (Provence). Une étude statistique sur l'orientation et la taille des conduits a été réalisée et a permis de déterminer deux familles de conduits. Différentes données conditionnantes ont aussi été intégrées pour assurer la reproduction d'un réseau cohérent aux informations géologiques connues. Les exemples de simulations sont montrés pour illustrer les capacités de la méthode proposée à reproduire un réseau karstique. Des calculs de volumes ont été réalisés pour comparer les modèles. L'ensemble de ces résultats est présenté et discuté.

MOTS-CLEFS : géostatistique, grotte, réseau.

I. Introduction

Karst networks have a major influence on fluid flow at the reservoir scale but at this scale, they are hardly continuously observable. Therefore, it is required to model the non-observable parts for reservoir characterization applications. Two 3D modelling approaches exist for reproducing geological features [Koltermann & Gorelick, 1996): the process-based and the structure-based. On one hand, process-based approaches numerically reproduce the speleogenesis processes to generate the shape of the karst networks. Based on the speleogenesis processes, the obtained 3D models are geologically sound. However, several difficulties are encountered when applying such approaches: 1) they imply knowledge of a wide parameter set that is difficult to estimate from observed data; 2) because they are time-consuming, they are generally determinist and not integrate uncertainties; 3) they are hard to condition to available punctual observed data. On the other hand, structure-based approaches reproduce the geometry and the spatial distribution of targeted objects (e.g., channel geobodies, and fracture or karst networks). These approaches are dedicated to be conditioned to local (e.g., local values of rock measurements, and field observation) or global data (i.e., trends). They are also very efficient and can be integrated into a stochastic procedure. Indeed, because few data are available, the use of a single interpolated model is not sufficient because many uncertainties exist in the geometry and spatial distribution of the objects. In this under-constrained problem, stochastic procedures aim at generating several equiprobable and plausible models for the geological structures. These models are used as support for uncertainty analyses [Scheidt & Caers, 2009] and risk assessment studies [Srivastava, 2013). In the present applications, stochastic simulations aim to reproduce field observations of karst networks. These approaches commonly require input parameters estimated from observed analogues. However, because the generated 3D models are only based on geometrical inputs of karst morphology (dimensions, orientations, etc.) and on their spatial distribution (proportion of karst conduits, etc.), their geological consistency is generally not insured. Classical geostatistical approaches [Fournillon et al., 2011; Mariethoz, 2009] and algorithms dedicated to karst network simulation [Borghi, 2013] have been proposed to reproduce the non-observed karst parts. The consistency is generally reproduced by integrating trends and secondary data into the simulation. The main difficulty in karst modelling is to reproduce accurate network connectivity (e.g., number of network cycles) while maintaining the geological and hydrogeological consistency of the models.

In this paper, it is proposed to show how it is possible to integrate geological knowledge into a classical geostatistical algorithm, SIS (Sequential Indicator Simulation), so that it reproduces a cave morphology from parameters obtained by an analogue study. The proposed technique is roughly similar to the approach proposed in [Fournillon et *al.*, 2011] because it relies on SIS and hierarchical simulations of karst conduits but it differs in the way in which karst connectivity is accounted for. Therefore, the SIS technique and the method proposed by [Fournillon et *al.*, 2011] are brie y described in the first section. Then, we explain how the different kinds of data (proportions, connectivity, geological features, etc.) can be integrated into the simulations so that the generated cave morphology is consistent with the known karstogenesis processes. In the second section, it is shown how this methodology can be applied to a real case study. The results of this case study are then discussed to assess the applicability of the proposed methodology.

II. Proposed approach

The proposed approach is termed as HSIS (Hierarchical Sequential Indicator Simulation) within this paper. The main strategy is to decompose caves into "basic elements" and to successively reproduce the spatial organisation of these elements via SIS to obtain the whole cave morphology.

A. Sequential Indicator Simulation

The SIS technique [Goovaerts, 1997; Journel et *al.*, 1998] corresponds to a geostatistical simulation approach commonly used for reproducing 3D models of categorical variables (e.g., facies, classes) from available conditioning data. The spatial repartition of categorical variables is simulated in a 3D grid composed of $nu \times nv \times nw$ cells (figure 1). In each grid cell, a facies value is simulated and the set of cells characterized by the same facies value forms the expected related 3D geobodies. The characterisation of the 3D geobody geometry relies on the estimation of a 2-point statistical parameter, referred to as the indicator variogram [Goovaerts, 1997; Journel et *al.*, 1998]. The indicator experimental variograms are computed in several directions from the available dataset. The variogram of one facies calculates the frequency of change from this facies to another. Variogram models are generally characterized by two kinds of parameters: the sill and the ranges. In stationary cases, the sill corresponds to the sample variance [Gringarten & Deutsch, 2001]. In the case of indicator variables, the variance is by definition $p(1 - p)$, where p is the proportion of the simulated facies or class. From the computed variograms, the main averaged dimensions and orientations of the geobodies are estimated by defining the variogram ranges and anisotropy direc-

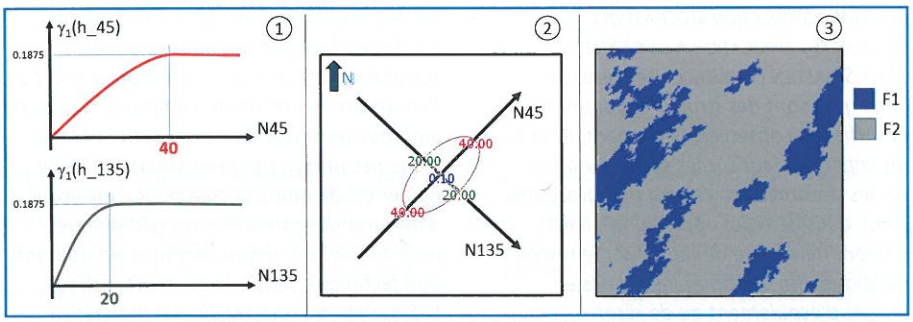

Figure 1: Examples of SIS simulation: 1) Variogram models of F1 facies in the two areal directions. Variogram ranges are shown as well as sills. Sills are equal to the variance of the F1 facies indicator $\sigma^2 = p_1(1 - p_1)$, where p_1 is the proportion of F_1 ; 2) Variogram ellipsoid of the facies; 3) Simulation of the two facies, whose proportions are 0.25 for F1 and 0.75 for F2.

Exemples de simulation SIS : 1) Modèles de variogramme pour le faciès F1 dans deux directions horizontales. Les portées des variogrammes sont montrées ainsi que le palier. Ce palier est égal à la variance de l'indicatrice du faciès F1, c'est-à-dire $\sigma^2 = p_1(1-p_1)$, avec p_1 la proportion de faciès F1 ; 2) Ellipsoïde du variogramme du faciès F1 ; 3) Simulation des deux faciès, avec les contraintes de proportion suivantes 0,25 pour F1 et 0,75 pour F2.

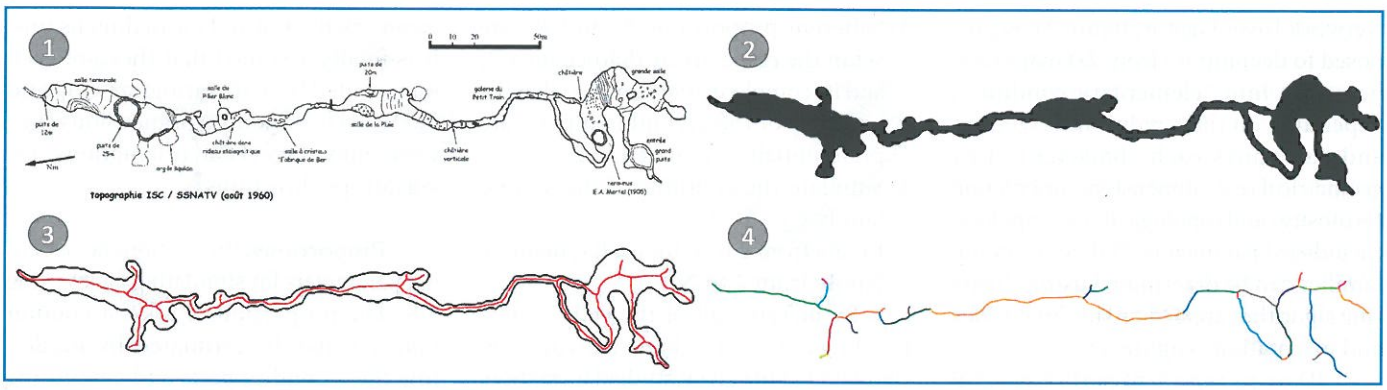

Figure 2: KNIT decomposition: 1) plan view map of the reference cave; 2) filling of the cave area; 3) extraction of the cave skeleton; 4) automated decomposition based on a curvature threshold.
Décomposition via KNIT : 1) Représentation en plan de la cavité de référence ; 2) Remplissage de l'aire de la cavité ; 3) Extraction du squelette ; 4) Décomposition automatique en fonction d'un seuil appliqué sur la courbure du squelette.

tion, respectively. Guo & Deutsch [2010], Ma & Jones [2001], Jones & Ma [2001]. show the relationships between parameters of indicator variograms and the modelled geobody dimensions and orientations. Thus, if data are not available, variogram models can be established from mean dimensions and orientations of the geobodies to be reproduced.

The main advantage of SIS is its flexibility because several conditioning data (data points, trends, etc.) can be easily taken into account and any kind of categorical variables can be simulated in 3D, regardless of what they represent and their nature. In this paper, the simulated facies corresponds to karst conduits and the background is the matrix, which includes rock and small heterogeneities (e.g., fractures). Because SIS uses 2-point statistics (variograms), it is only able to reproduce straight shapes, roughly representing ellipsoids that can be amalgamated. Certain authors have proposed variant SIS algorithms that allow curvilinear shapes to be reproduced using variable azimuth maps for defining local variogram directions [Deutsch, 2002; Pyrcz & Deutsch, 2014]. However, difficulties are met when reproducing highly continuous curvilinear geobodies, such as channels. Object-based [Viseur et *al.*, 2004, 1999] and multiple-point techniques [Mariethoz, 2009 ; Caers & Zhang, 2004] are alternative approaches to generate 3D models of complex object stacking.

B. Cave decomposition and simulation

The proposed approach relies on the technique presented in [Fournillon et *al.*, 2011]. The main idea is to decompose the cave network into elementary conduits. These elementary conduits correspond to cave branches that share the same geometrical characteristics: main direction and dimensions (height and width). The whole network is then cut into several elementary conduits that can be clustered into conduit families. To reproduce the cave morphologies, the simulator generates the conduits from the first family to the last one. The order of simulation may be random or performed according to a given chronological or process-based order. In Abelard et *al.* [2011], an automated approach termed KNIT (Karst

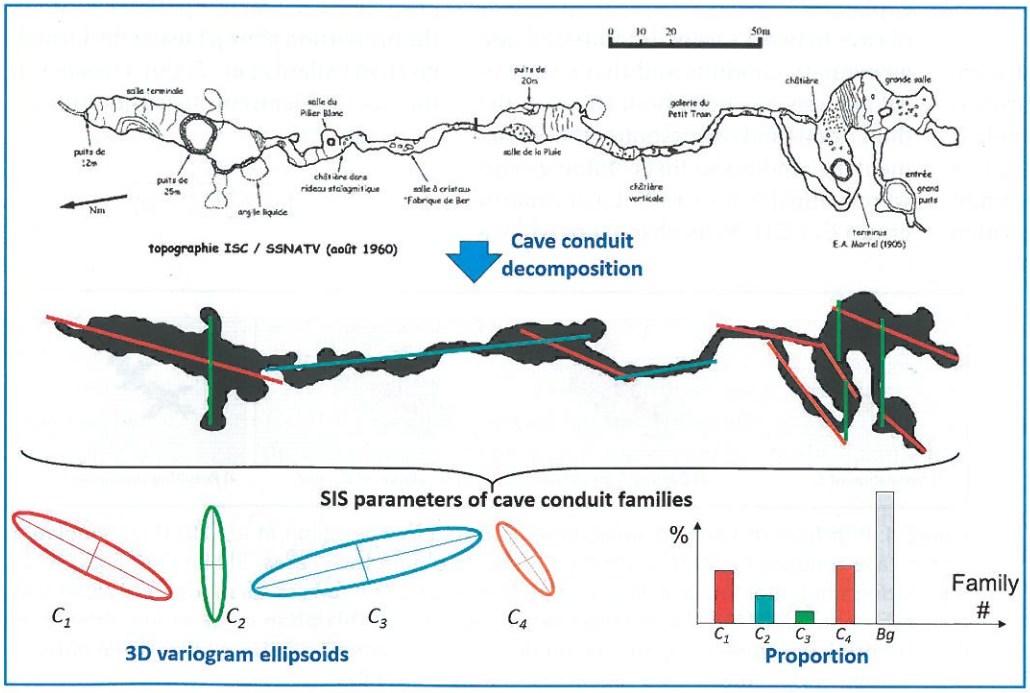

Figure 3: Principles of karst decomposition for SIS simulation, from top to bottom: 1) 2D map of cave; 2) decomposition of cave skeleton; 3) input parameters for SIS simulation, including 3D variogram ellipsoids and proportions. $C_i, i \in \{1;4\}$ are the cave conduit families and Bg is the background, i.e. matrix.
Principes de la décomposition de la cavité pour la simulation SIS, de haut en bas : 1) Carte 2D de la grotte ; 2) Décomposition du squelette de la cavité ; 3) Paramètres d'entrée pour la SIS, comprenant les ellipsoïdes 3D de variogramme ainsi que les proportions pour chaque famille de conduits $C_i, i \in \{1;4\}$ ainsi que pour la matrice.

Network InvesTigator, figure 2), is proposed to decompose from 2D maps cave networks into "elementary conduits", depending on the conduit intersections and curvatures each characterized by geometrical (e.g., dimensions, orientation tortuosity) and topological (i.e., topological indices) parameters. Different conduit families can be determined using clustering algorithm applied on the dimensions and orientations (figure 3).

2D maps represent partial views of cave networks because the cave geometry and description are projected onto a 2D horizontal plane. It is also possible to obtain using a 3D topographic survey the complex path, with conduit width and height, of accessible cave parts. Thus, from this data set, decomposition into elementary conduits could be applied and allows effective orientation parameters to be obtained. Then, it is assumed in the following that statistical analysis of analogue caves has been performed for the dimensions and orientations of elementary conduits for each defined family.

From these statistical analyses, a 3D variogram model (figure 3), corresponding to a 3D ellipsoid, is determined and approximates the conduit shape. Then, a series of variogram models can be modelled so that they represent the different karst conduit geometries. The simulated variables $C_i, i \in [1;N]$ are binary random variables and defined as follows :

$$\begin{cases} C_i(x,y,z) = 1 \text{ if } (x,y,z) \text{ belongs to a karst conduit} \\ C_i(x,y,z) = 0 \text{ else (i.e. matrix)} \end{cases} \quad (1)$$

Let p_i^k be the proportion of karst conduit of the i^{th} family; then, the proportion of background against this family is by definition: $p_i^g = 1 - p_i^k$. In [Fournillon et *al.*, 2011], the cave conduits were simulated using this approach. The algorithm may be summarized as follows (figure 4) :

1) The first conduit family C_1 is simulated in grid area Z_0, corresponding to the whole 3D grid. Each simulated karst conduit corresponds to a set of connected grid cells.
2) Using a morphological operator (dilatation), a zone Z_1 is created around each simulated karst conduit. The extension of this area is equal to the length of the second family C_2 to simulate.
3) Different proportions p_2^k and $p_2'^k$ are set in the respectively defined area Z_1 and its complementary zone so that the conduits of the second family C_2 are preferentially simulated in Z_1.
4) Simulate the conduits of the second family C_2.
5) Iterate from 2 to 4, for each remaining family from 2 to N.

By merging all of the cells corresponding to C_1 to C_N, karst networks are simulated in the given studied formation. In Fournillon et *al.*, 2011], it is shown how a proportion map can be used to simulate karst networks in areas where they are preferentially formed.

Two main remarks may be addressed when considering this algorithm. Firstly, the proportions p_i^k and p'^i_k (step 3) were used to promote connectivity between karst conduit families but their values are hard to determine from analogue data. Moreover, the connectivity is not directly accounted for but is a consequence of the proportions p_{i+1}^k and the size of the Z_i region. Secondly, the different proportions of the karst conduits do not account for intersections between conduits. Adjustments are required to avoid underestimated simulated proportions.

C. Data conditioning

In this section, it is shown how the proportions and connectivity are defined in the proposed approach, compared to the similar approach in [Fournillon et *al.*, 2011]. How geological and hydrogeological knowledge can be integrated is also explained. It is assumed that analogues of cave networks were decomposed into elementary conduits and that a statistical analysis has been performed on the dimensions and orientations of these elementary conduits so that conduit groups were defined from 1 to *N*. Each conduit group $C_i, i \in \{1;N\}$ is characterized by a mean orientation and mean dimensions. It is finally assumed that they are each represented by a 3D variogram ellipsoid modelled from conduit dimensions (i.e., variogram ranges) and orientations (i.e. anisotropy directions).

Proportions: Proportions are essential parameters for simulating facies using SIS. The proportion of a karst conduit family i may be estimated by modelling these conduit parts and computing the proportion as the ratio between the modelled conduit volumes and a given study volume. However, it is often hard to model these 3D volumes because 3D cave data are rare. Another strategy is to draw a parallel between SIS and Boolean models to estimate proportions. Let N_i be the estimated number of karst conduits C_i in a given volume. Because elementary conduits C_i are approximated by a 3D variogram ellipsoid, N_i can be considered as the density θ_i of a Poisson point process simulating ellipsoid objects, with mean volume V_i. In stationary conditions and in the case of Poisson point process, it is shown in [Allard et *al.* [2006]) that the relationship between density and proportion is the following:

$$\theta_i = -\frac{1}{V_i}.ln(1 - p_i^k) \quad (2)$$

Moreover, intersections may occur between C_1 and C_2 and lead to overlap between C_1 and C_2. The overlapping parts correspond to simulated parts of C_2 that are not accounted for in the final proportion. Then, it is possible to adjust the proportion p_i^k by p'^k_i using the formula given in [Allard et *al.*, 2006], considering the case of hierarchical erosion rules:

$$p_i'^k = \frac{p_i^k}{1 - \sum_{j=1}^{j=i-1} p_j^k} \quad (3)$$

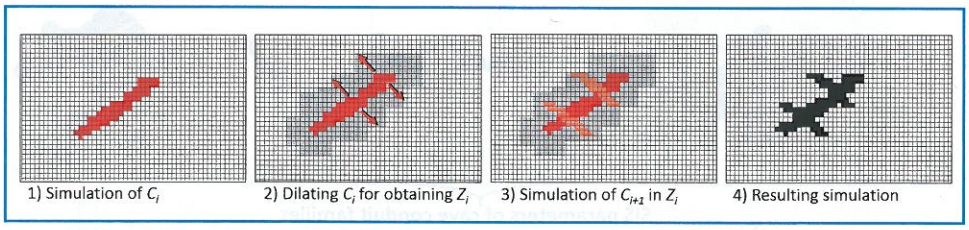

Figure 4: Principles of karst SIS simulation, proposed in (Fournillon et *al.*, 2011): 1) Simulation of the cave conduits C_i ; 2) Dilating the C_i zone to obtain the Z_i area; 3) Simulating C_{i+1} in Z_i ; 4)Resulting simulation of cave. *Principes de la simulation de formes karstiques, proposée dans Fournillon et al., [2011] : 1) Simulation des conduits C_i ; 2) Dilatation de la région obtenue de C_i pour définir une zone Z_i ; 3) Simulation de C_{i+1} dans la zone Z_i ; 4) Simulation finale obtenue pour un réseau karstique à deux familles de conduits C_1 et C_2.*

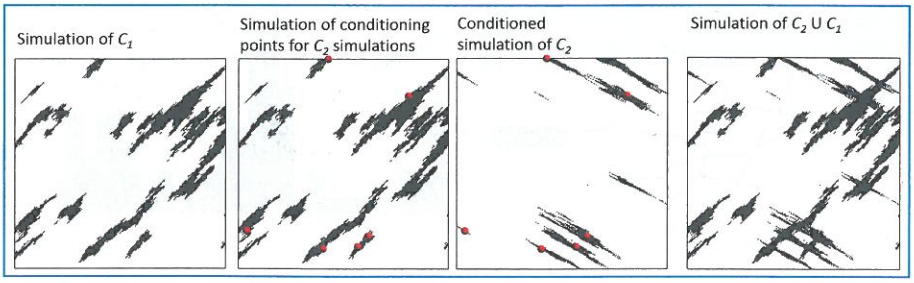

Figure 5: Proposed algorithm for accounting for a given connectivity degree: 1) Simulation of C_1 ; 2) Simulation in C_1 of conditioning points for C_2 simulation; 3) Conditioned simulation of C_2 ; 4) Resulting simulation corresponding to the union between C_1 and C_2.
Algorithme proposé pour prendre en compte un certain degré de connectivité : 1) Simulation de C_1 ; 2) Simulation dans C_1 de points conditionnant pour la simulation de C_2 ; 3) Simulation conditionné de C_2 ; 4) Simulation finale obtenue par l'union entre C_1 et C_2.

It is important to notice that these formulas only give an estimation of the proportion to target, because the SIS algorithm is different than Poisson point processes.

Connectivity: As mentioned previously, connectivity between cave branch families was insured by using preferential areas of simulation around already simulated cave branches in the approach proposed by [Fournillon et *al.*, 2011]. In the proposed approach, connectivity is not reproduced using these areas but using simulated conditioning points. Consider a cave composed by two main conduit directions, respectively, C_1 and C_2. These two families are each characterized by a variogram model γ_1 and γ_2. Let N_{1-2} be the number of intersections between C_1 and C_2 conduits. The cave conduits corresponding to the first family are first simulated in the 3D grid. Let N_1 be the number of grid cells belonging to C_1. To account for a given connectivity between C_1 and C_2, it is proposed to generate and use, on average, N_{1-2} simulated conditioning points among the N_1 cells for simulating the C_2 geometries (figure 5).

To achieve this, a critical probability value p_c is computed as follows:

$$p_c = \frac{N_{1-2}}{N_1} \qquad (4)$$

Then, a random variable is simulated with a Monte-Carlo technique applied on a uniform probability law $U(0;1)$ for each cell belonging to C_1. If the simulated value is below pc, the cell becomes a conditioning point to C_2 simulation. By passing through each cell of C_1, the average number of simulated conditioning points will be N_{1-2} over a set of simulations. Conditioning the simulation of C_2 by these points ensures that C_2 will intersect C_1 N_{1-2} times on average.

Observations: It is possible to observe inlet and outlet points from a cave network. These observations mean that karst conduits pass through these points. These points can be set as conditioning points. However, because in the proposed approach, the cave network is decomposed into several conduit families, only part of these data points condition the simulation of a specific conduit family. Therefore, from the entire dataset, it is proposed to stochastically generate aseries of conditioning points CP_k for each conduit family k. Let consider n conduit families, each characterized by a proportion p_k. These proportions $P_k = \{p_k\}_{k\in\{1;n\}}$ define a probability density function for the conduit family occurrence. Considering that these proportions are stationary, the proposed assumption is the following: the probability that an outlet or an inlet point belongs to a conduit family is equal to its proportion pk. Therefore, a Monte Carlo sampling technique is applied on the set of data points: a random path is chosen among the points and for each point, its membership to a set of given CP_k is simulated by Monte Carlo. Thus, a proportion of p_k data points serve on average to condition the conduit family k. This stochastic sampling can account for the proportion trend (e.g., areal and vertical proportion trends), for non stationary cases.

Control of geological structures: several studies [Filipponi et *al.*, 2009] show that karst preferentially occurs along fractures and stratigraphic discontinuities. Therefore, it would be applicable to simulate karst conduits preferentially near these structures. To achieve this, it is proposed to compute at each grid cell the distance map to these structures. The distance map corresponds to the computation of the shortest distance between each point of a given space to a given structure (e.g., surfaces, lines). Because these structures increase karst conduit occurrence, the probability of karst occurrence should be higher because the computed distance is small. Consider the distance d computed at each point (x,y,z) of a given space. The following equation is proposed to define a consistent probability field that accounts for this assumption:

$$\begin{cases} p_{ks}(x,y,z) &= (\exp(\frac{-d(x,y,z)^4}{\sigma^2})) \\ p_{mks}(x,y,z) &= 1-(\exp(\frac{-d(x,y,z)^2}{\sigma^2})) \end{cases} \qquad (5)$$

The σ parameter controls how fast the impact of the given structure on karst occurrence decreases.

Hydrogeological control: Two kinds of hydrogeological controls are addressed in this paper: the estimation of the karst conduit direction based on the main apparent flow and on a map of cave conduit proportions depending on observed outlet and inlet points. First, in some cases, the orientation of the main conduits is aligned along particular directions. If these directions can be estimated from observed data, it is then possible to tune the direction of the simulation cave conduits using a variable azimuth approach [Boisvert, 2011; Biver et *al.*, 2012; Pyrcz & Deutsch, 2014]. Second, certain conduits stem from the main flows between outlet and inlet points. Their occurrence is preferentially located within the area defined by these points. Thus, to account for this, it is suggested in this paper to use the convex hull of the observed points as follows: within the convex hull the cave conduit proportion is highest; outside the proportions decrease depending on distance using equations similar to 5.

Combining data: If several probability field conditionings exist, it is possible to combine them. Let us consider the two following assumptions:
- Hypothesis A: karst conduits occur preferentially near a given structure.
- Hypothesis B: karst conduits occur preferentially in a given area.

Then, considering the two hypotheses consists in considering the event $A \cap B$ and if A and B are independent, then: $P(A \cap B) = P(A) * P(B)$.

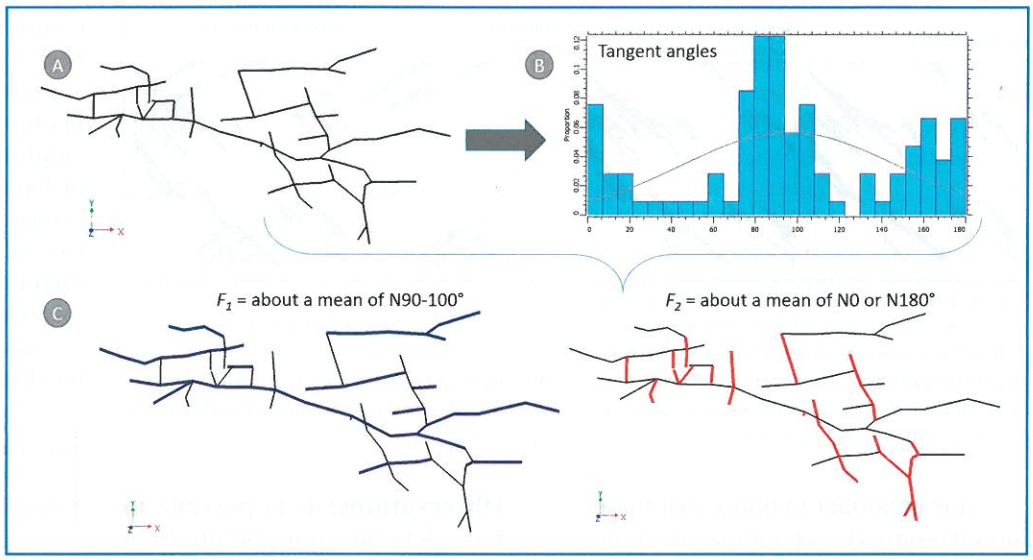

Figure 6: Branch decomposition: two families were determined from statistical analysis of the skeleton directions. A) the cave network of Saint-Sébastien; B) histogram of tangent angles; C) visualisation of the two branch families, *F1* is the main family, which is centred around a mean orientation of *N90–110* and *F2* is the second family centred around a mean orientation of 160.
Décomposition en familles : deux familles ont été déterminées par analyse des directions du squelette de la grotte de Saint-Sébastien. A) Réseau de la grotte St-Sébastien ; B) Histogramme des angles des tangentes au squelette ; C) Visualisation des deux familles de branches définies, F1 correspond à la famille principale orientée autour de N90 – 110 et F2 à l'autre famille centrée autour de N160.

III. Applications and discussions

This application aims to show how it is possible to integrate data in real conditions. As previously mentioned, it is assumed that a statistical analysis is performed on analogue data before applying the proposed method. The statistical analysis allows information about the orientations and sizes of cave conduits. The proposed approach is applied to the case study of the Saint-Sébastien cave (Gréoux-les-Bains, Provence). After describing the geological backgrounds, the statistical analysis of the conduits is shown as well as the used conditioning data and their justifications. Then, the resulting simulations are shown and discussed.

A. Geological settings and data set

Saint-Sébastien Cave (Gréoux-les-Bains, Alpes de Haute Provence, SE France) is a hypogenic 2D maze [Audra et *al.*, 2010] with a total length of 319m. It is a planar maze, which develops within a slightly tilted (25°*SE*) Hauterivian limestone confined below an aquitard of marls. Rising water came from the deeper passage to reach the current entrances. The even recharges of deep water enlarged the entire fissure network at the same rate producing a maze pattern reflecting the tectonic heritage.

In addition to the geological information, the dataset available for this study consists of a 3D linear representation of the cave path (figure 7-A). At acquisition path points, information about the local left and right width of the conduits as well as its bottom and top height is available, except in certain locations (due to acquisition difficulties). The acquisition points are not necessarily located at the middle of the cave conduit. To obtain comparable measurements along the cave network, it was necessary to relocate the path points to the middle of the conduits width and height.

It is important to notice that biases exist in such data because conduit dimensions are measured at the acquisition points only. They do not

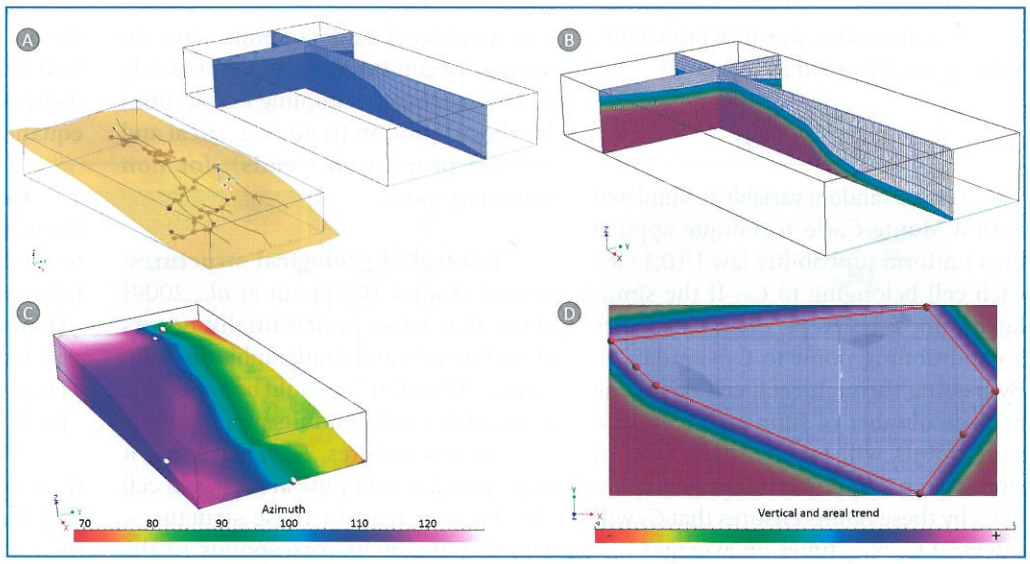

Figure 7: Data conditioning: A) 3D model of the basal surface of marl formation and the 3D grid with gridding conditioned to this surface; B) Probability for conduit occurrence near the marl base surface; C) Variable azimuth for the *F1* family; D) Probability field for conduit occurrence within an area.
Données conditionnantes : A) modèle 3D de la surface de base de la formation marneuse ainsi que la grille 3D dont le maillage est conditionné à cette surface ; B) Probabilité de l'occurrence d'un conduit karstique près de la surface de base des marnes ; C) Champs 3D de l'azimut de la famille F1, utilisé comme contrainte dans la SIS à azimut variable ; D) Champs 3D de probabilité pour la présence d'un conduit karstique dans la zone comprise entre et autour des points d'entrée et de sortie du karst.

correspond to the minimum or maximum values. Because the shape of the conduit cross-sections can be complex, the width and height values acquired at these points do not necessarily reflect the true main cross-section dimensions. Moreover, the acquired cave data only represent a sub-network of the entire network and that can be reached by humans.

B. Data conditioning

A 3D geomodel was built considering the top of the karst network as the base of marls. A surface was built conditioned to top points of the karst networks using DSI technology in Gocad [Mallet, 2002]. Then, a 3D grid was constructed so that the cell layering is parallel to this surface (figure 7-A). The grid resolution is around 40 cm in area and 25 *cm* in height.

A statistical analysis was performed on the tangent direction on the corrected cave path. This analysis allows two cave conduit families to be distinguished (figure 6): a first family C_1 composed of roughly E-W conduit directions (*N*90 – *N*100) and the second family C_2 of conduits along N-S directions (around *N*0 or *N*180). The mean width and height of the individualized conduit families were estimated to determine the variogram ranges of each family. Therefore, using the anisotropy direction (orientation) and these dimensions, two 3D variogram ellipsoids were defined. The C_1 family conduits are oriented in the main flow direction between the outlet and inlet locations. This leads to a fan shaped spatial repartition of these conduits. These orientations follow the main straight directions between these points. Thus, an orientation map (figure 7) has been interpolated, conditioned to these boundary directions. This map was used as a trend for the variogram anisotropy direction of the C_1 family. This corresponds to a variable azimuth conditioning of the SIS. By definition, the sill of the variograms is equal to $p_k \times (1 - p_k)$ with $k \in \{1,2\}$ and p_k the proportion of the conduit family C_k.

The proportions of the C_1 and C_2 conduits were estimated using the formula 2 and 3. The number of conduits per volume of interest was estimated for each family. The elementary conduit volume was based on the 3D variogram ellipsoid ranges. Two trends of the proportions were integrated in the conditioning:

- Flow direction impact: the convex hull of the observed inlet and outlet points was calculated. Within this zone, the proportion is maximal and outside, the proportion decreases depending on the distance using the equation 5. The σ was chosen as the mean conduit dimensions of C_2 (10 *m*) because they are located almost perpendicular to the main flow direction.
- Distance to the marl base surface: distance to this surface was calculated in each cell of the 3D grid. The proportion decreases with this distance using the equation 5. The σ was chosen as the maximal height of the conduits (2 *m*).

Family	r1	r2	r3	azimuth	sill	proportion
C_2	80	1	1	Trend (cf.figure 7-C)	0.0475	0.05
C_1	10	1	1	160	0.024375	0.025

Table 1: Parameters used for simulating conduits of families C_1 and C_2. Distances are in meters and angles are degrees. *Paramètres utilisés pour simuler les familles C_1 et C_2 de conduits. Les distances sont en mètres et les angles en degrés.*

The hard data conditioning are of two categories. Firstly, the inlet and outlet points serve as support for conditioning. The number of these points is randomly separated into two groups: one conditioning the C_1 conduits and one the C_2. The random choice is performed on the data points through Monte-Carlo sampling considering the proportion of conduits. Several simulations can be generated from the same random choice as shown in the next section. Secondly, the C_2 family was conditioned to points belonging to the C_1 simulation to ensure the connectivity between C_1 and C_2. These points are randomly simulated using the technique described in section 2.3. The number of intersections between C_1 and C_2 was estimated as 22 from the dataset. The conditioning data are shown in figure 7 and table 1.

C. Results and discussions

Using the conditioning dataset previously presented, simulations of cave networks were generated. Three of the simulations are shown in the figure 8. Per simulation, several unconnected caves were generated, the ones shown in figure 8 are the largest ones among the simulations. Moreover, only simulations for which a cave geobody connects all inlet and outlet points are kept because they are all connected in the Saint-Sébastien cave network.

The simulated numbers of conditioning points for the C_2 family simulation are shown in the table 2 as well as the volume of the simulated caves. This volume corresponds to the raster volume of the cave, which corresponds to a set of grid cells. To be comparable, the volume of the 3D reference model of the Saint-Sébastien cave was also computed as a raster volume. The simulated numbers

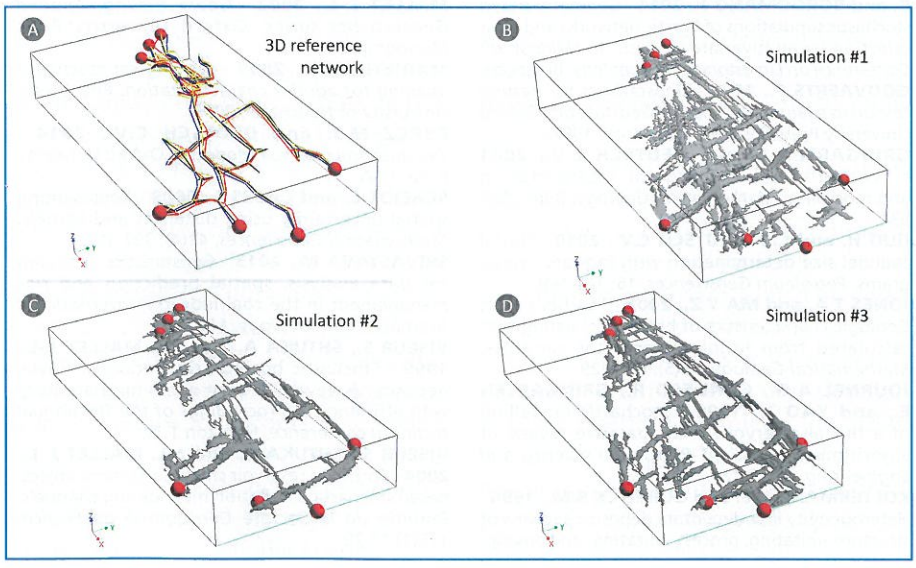

Figure 8: Simulations: A) 3D survey of the reference network; B) Simulation #1; C) Simulation #2; D) Simulation #3. *Simulations : A) Topographie 3D du réseau de référence ; B) Simulation n°1 ; C) Simulation n°2 ; D) Simulation n°3.*

of conditioning data varies around 22. The first simulation is the most connected volume with 27 C_2 conditioning points. This leads to a cave with a large volume because it connects several branch conduits. More simulations must be studied to see the impact of the generated conditioning points on the final cave volumes. Simulation #2 looks visually more similar to the real Saint-Sébastien cave network and its volume is also the closer to the true volume.

This example shows that it is possible to add conditioning data when simulating cave networks.

These data reflect the geological and speleogenetic knowledge of the cave being simulating. However, as previously mentioned, only straight shapes can be reproduced using SIS. Therefore, only straight elementary cave conduits can be simulated. Variable azimuth conditioning can be used to locally modify the conduit directions, as shown in this example. However, the ability to generate curvilinear shapes remains difficult with this technique. Multiple-point statistics are better methods to perform this task.

Finally, efforts must be devoted to the connectivity conditioning. Conditioning points are simulated to assert a certain degree of connectivity between conduits but due to the hazard of the simulation additional intersections between conduits of different families could occur. This leads to over-simulation of average number of conduit intersections. One strategy could be to virtually set as matrix facies the cells of the already simulated conduits, except around conditioning points, while simulating the current family conduits. Moreover, it may be interesting to promote connections, hence conditioning points, at the extremity of a C_i conduit for simulating C_{i+1} conduits. This allows particular speleogenetic processes to be mimicked.

Simulation #	Conditioning data	Volumes
Simulation #1	27	402.741
Simulation #2	19	259.17
Simulation #3	16	314.552
Data 3D model	-	247.015

Table 2: Simulated values for the three simulations shown in figure 8. The conditioning data correspond to the simulated conditioning point in C_1 that condition the simulation of C_2.
Volumes are in m³. *Valeurs simulées pour 3 exemples de simulations montrés dans la figure 8. Les points conditionnants correspondent à ceux simulés dans les conduits de la famille C_1 et utilisés pour conditionner la simulation de la famille C_2. Les volumes sont en m³.*

IV. Conclusion

A method was proposed to simulate the karst network from analogue and geological information. The method relies on a common technique, known as SIS. The karst network is decomposed into elementary conduits that can be approximated by a variogram ellipsoid. Several conduit families are then determined from a given analogue karst database. Each conduit family is successively simulated and conditioned to points that assert connectivity between each conduit family and other secondary data. The final simulation aims to reproduce non-observed karst networks to perform risk assessments. This technique was successfully applied to the case study of Saint-Sébastien. This technique allows simulation of rectilinear karst conduit networks, such as maze patterns. Efforts must be devoted to enhancing the conditioning of the connectivity between families, and more generally of the network topology, as well as reproducing conduits that could be curvilinear.

Acknowledgements

The authors would like to thank ParadigmGeo and the ASGA for their support in providing Gocad software and plugins. The authors would also like to thank Philippe Audra for providing the cave data and for allowing us to use the data for this study. Johan Jouves's PhD scholarship is cofunded by Cenote Company, Véolia, and Agence de l'Eau Rhône Méditerranée et Corse.

Bibliographie

ABELARD S., FOURNILLON A., VISEUR S., ARFIB B. and BORGOMANO J., 2011 - Geomorphological 2d characterization of caves using automatic extraction of topological and geometrical parameters. In *H2Karst: 9th Conference on Limestone Hydrogeology.*

ALLARD D., FROIDEVAUX R. and ALTHOUGH BIVER P., 2006 - Conditional simulation of multitype non stationary markov object models respecting speci ed proportions. *Mathematical Geology,* 38(8): 959 986.

AUDRA P., D'ANTONI-NOBECOURT J.-C. and BIGOT J.-Y., 2010 - Hypogenic caves in France. Speleogenesis and morphology of the cave system. *Bull. Soc. Geol. Fr.,* 4: 327 335.

BIVER P., PIVOT F. and HENRION V., 2012 - Estimation of most likely lithology map in the context of truncated gaussian techniques. In *Proceedings of the Ninth International Geostatistics Congress.*

BOISVERT J.-B., 2011 - *Generating locally varying azimuth fields.* Technical report, CCG Annual Report 13, Ed- monton, AB.

BORGHI A., 2013 - *3D stochastic modeling of karst aquifers using a pseudo-genetic methodology.* PhD thesis, University of Neuchâtel.

CAERS J. and ZHANG T., 2004 - Multiple-point geostatistics: a quantitative vehicle for integrating geologic analogs into multiple reservoir models. In Grammer et *al,* editor, *AAPG Memoir Integration of outcrop and modern analogs in reservoir modeling,* p. 384 394.

DEUTSCH C.V., 2002 - *Geostatistical Reservoir Modeling.* Oxford University Press.

FILIPPONI M., JEANNIN P.-Y., and TACHER L., 2009 - Evidence of inception horizons in karst conduit networks. *Geomorphology,* 106: 86 99.

FOURNILLON A., ABELARD S., VISEUR S., ARFIB B. and BORGOMANO J., 2011 - Geometry-based stochastic simulations of karstic networks and their selection by multivariate analysis. In *H2karst:9th Conference on Limestone Hydrogeology,* Besançon.

GOOVAERTS P., 1997 - *Geostatistics for natural resources evaluation. Applied Geostatistics.* Oxford University Press, Elsevier, New York, 1997.

GRINGARTEN E. and DEUTSCH C.V., 2001 - Teacher's aide. Variogram interpretation and modeling. *Mathematical Geology,* 33(4): 507 534.

GUO H. and C.V. DEUTSCH C.V., 2010 - Fluvial channel size determination with indicator variograms. *Petroleum Geosciences,* 16: 161 169.

JONES T.A. and MA Y.Z., 2001 - Teacher s aide: Geologic characteristics of hole-effect variograms calculated from lithology-indicator variables. *Mathematical Geology,* 33(5): 615 629.

JOURNEL A.G., GUNDESO R., GRINGARTEN E., and YAO T., 1998 - Stochastic modelling of a fluvial reservoir: a comparative review of algorithms. *Journal of Petroleum Science and Engineering,* 21:95 121.

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: 2617-2658.

MA Y.Z. and JONES T., 2001 - Teacher's aide: Modeling hole-effect variograms of lithology-indicator variables. *Mathematical Geology,* 33(5):631-648.

MALLET J.-L., 2002 - *Geomodelling. Applied Geostatistics series.* Oxford University Press Monograph.

MARIETHOZ G., 2009 - *Geological stochastic imaging for aquifer characterization.* PhD thesis, University of Neuchâtel, 2009.

PYRCZ M.J. and DEUTSCH C.V., 2014 - *Geostatistical reservoir Modeling.* Oxford University Press, USA.

SCHEIDT C. and CAERS J., 2009 - Representing spatial uncertainty using distances and kernels. *Mathematical Geosciences,* 41(4): 397 419.

SRIVASTAVA M., 2013 - Geostatistics: A toolkit for data analysis, spatial prediction and risk management in the coal industry. *International Journal of Coal Geology,* 112: 2 13.

VISEUR S., SHTUKA A. and J.-L. MALLET J.-L., 1999 - Stochastic boolean simulation of fluvial deposits: A new approach combining accuracy with efficiency. In *Proceedings of the SPE annual technical conference,* Houston T. X.

VISEUR S., SHTUKA A. and J.-L. MALLET J.-L., 2004 - Turbidite reservoir characterization: object-based stochastic simulation meandering channels. *Bulletin de la Société Géologique de France,* 175(1):11 20.