



HAL
open science

Simulation stochastique des réseaux karstiques

Yves Frantz

► **To cite this version:**

Yves Frantz. Simulation stochastique des réseaux karstiques. Géomorphologie. Université de Lorraine, 2021. Français. NNT : 2021LORR0215 . tel-03652808

HAL Id: tel-03652808

<https://hal.univ-lorraine.fr/tel-03652808>

Submitted on 27 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



Simulation stochastique des réseaux karstiques

THÈSE

présentée et soutenue publiquement le 13 Octobre 2021

pour l'obtention du grade de

Docteur de l'Université de Lorraine

Spécialité Géosciences

par

Yves FRANTZ

Composition du jury:

<i>Présidente du jury:</i>	Prof. Valérie PLAGNES	Sorbonne Université, France
<i>Rapporteurs:</i>	Prof. Anne JOHANNET	IMT Mines Alès, France
	Prof. Nicolas MASSEI	Université de Rouen Normandie, France
<i>Examineurs:</i>	Prof. Philippe RENARD	Université de Neuchâtel, Suisse
	Dr. Sophie VISEUR	Aix-Marseille Université, France
<i>Directrice de thèse:</i>	Dr. Pauline COLLON	Université de Lorraine, France

GeoRessources - UMR 7359

Université de Lorraine – CNRS
École Nationale Supérieure de Géologie
2 rue du Doyen Marcel Roubault - BP 10162
54505 Vandœuvre-lès-Nancy, FRANCE

Résumé

Difficiles d'accès, les réseaux karstiques ne sont généralement que partiellement décrits. Les techniques d'exploration classiques produisent par ailleurs des données éparses et sujettes à des incertitudes concernant la position des conduits et leurs dimensions, paramètres essentiels aux simulations d'écoulement. Les simulations stochastiques permettent de mieux gérer et évaluer ces incertitudes en proposant plusieurs représentations équiprobables de réseaux karstiques. Le simulateur idéal devrait permettre de construire des réseaux de drains karstiques tridimensionnels, honorant les observations de terrain (indices de karstification), les connaissances apportées par les essais de traçage, les informations recueillies sur des parties accessibles du réseau et celles apportées par l'étude d'autres réseaux. C'est dans ce contexte que s'inscrit cette thèse qui propose 3 principales contributions.

La première contribution est l'analyse statistique d'une base de données de 49 réseaux karstiques. Elle se focalise sur l'étude de la géométrie des conduits, à travers l'analyse de deux métriques : le rayon équivalent et le ratio largeur-hauteur. Aucune loi statistique générique qui décrirait la géométrie des conduits n'a pu être mise en évidence. Néanmoins, la variabilité spatiale des propriétés géométriques à différentes échelles a été caractérisée, notamment grâce au développement de variogrammes 1D-curvilinéaires. L'hypothèse répandue d'une organisation hiérarchique des géométries de conduits a également été analysée et rejetée.

La deuxième contribution est le développement de deux méthodes de simulation stochastique de propriétés le long de réseaux karstiques s'appuyant sur les résultats de l'analyse statistique. La première méthode se focalise sur la reproduction de la variabilité des propriétés à l'échelle du réseau, tandis que la deuxième se concentre sur la reproduction de la variabilité au sein et entre les branches du réseau. Elles sont toutes deux basées sur les méthodes de Simulations Gaussiennes Séquentielles (SGS), ici adaptées à des objets 1D-curvilinéaires.

La troisième contribution est le prototype d'une méthode ayant pour but de simuler stochastiquement des réseaux karstiques discrets sous forme de graphes (qualifiés de *squelettes des réseaux*). Il est espéré qu'une fois entièrement développée, elle permette de simuler différents types de réseaux et de prendre directement en compte les données de terrain ainsi que les informations géologiques. Elle consiste en trois grandes étapes : i) la génération d'un nuage de points, ii) le calcul de leur connectivité et iii) leur connexion pour former le squelette du réseau karstique.

Ces contributions ouvrent de nouvelles perspectives quant à la simulation de réseaux de drains karstiques utilisables pour la simulation d'écoulement (e.g., SWMM, Epanet, Modflow-CFP), ce qui devrait permettre une meilleure caractérisation des flux en leur sein.

Mots-clés karst, réseaux karstiques, analyse statistique, géométrie de conduits, simulation stochastique, graphe

Abstract

Despite intensive explorations by speleologists, karstic networks remain only partially described as many conduits are not accessible to humans. The classical exploration techniques produce sparse data subject to uncertainties concerning the conduit position and their dimensions, which are essential parameters for flow simulations. Stochastic simulations make it possible to better handle and assess these uncertainties by offering several equally probable karstic network representations. The ideal simulator should allow for the construction of tridimensional karstic drain networks, respecting the field observations (karstification markers), the knowledge brought by tracer tests, the information collected in the accessible parts of the network and those obtained by the study of other networks. In this context, this PhD thesis offers 3 main contributions.

The first contribution is the statistical analysis of a database of 49 karstic networks. It focuses on the study of conduit geometry, through the analysis of two metrics : the equivalent radius and the width-height ratio. No generic statistical law describing the network geometry was found. Nonetheless, the spatial variability of the geometrical properties at different scales was characterized, mostly through the development of 1D-curvilinear variograms. The widespread hypothesis of a hierarchical organization of the conduit geometries has also been analysed and rejected.

The second contribution is the development of two methods allowing stochastic simulations of properties along karstic networks and based on the results of the statistical analysis. The first method focuses on the reproduction of the property variability at the network scale, while the second one focuses on the reproduction of the variability within and between the network branches. Both are based on the Sequential Gaussian Simulation methods and are adapted to 1D-curvilinear objects.

The third contribution is the prototype of a method aiming to stochastically simulate discrete karstic networks as graphs (known as *network skeletons*). We hope that once completed, it would allow the simulation of different network types, while taking directly into account field data and geological information. It is divided in three main steps : i) the generation of a point cloud, ii) the computation of the point connectivity and iii) their connection to create the skeleton of the karstic network.

These contributions open new prospects regarding the simulation of karstic drain networks usable for flow simulation (e.g., SWMM, Epanet, Modflow-CFP), which should allow a better characterization of the associated flows.

Keywords karst, karstic networks, statistical analysis, conduit geometry, stochastic simulation, graph

Remerciements

Merci.

J'hésitais à faire la blague mais j'ai fini par céder (les manuscrits ayant déjà été imprimés, ça ne devrait pas gâcher de papier). N'étant pas doué pour ce genre de remerciements, je vais éviter de trop en faire. Je tiens tout d'abord à dire un grand merci à Pauline, ma directrice de thèse, sans qui rien n'aurait commencé, ni abouti. Merci pour ton encadrement, la route aura été un peu plus longue que prévu mais le jeu en valait la chandelle.

Je tiens ensuite à remercier les membres du jury pour leurs remarques pertinentes (et pour m'avoir accordé le titre de docteur), plus particulièrement Sophie et Philippe qui m'ont également prodigué de nombreux conseils éclairés au cours de cette thèse.

Merci à toutes les personnes de l'équipe pour leur aide, pour leur présence et pour avoir supporté mon humour. La liste est longue et j'espère n'oublier personne. Il y a bien sûr les permanents de l'équipe RING: Pauline, Guillaume, Paul, François, Christophe, Mustapha et Marie (sans oublier Fatima, Sophie, Nageoi et Christine)... mais aussi les thésards et post-docs: Benjamin, Jonathan, Antoine, Gabriel, Marion, Gautier, Pierre, Modeste, Julien, Margaux, Nicolas 1, Nicolas 2, Anais, Melchior, Corentin, Paul, Zoé, Capucine, Enrico, Fabrice et Jérémie et... je crois que que je n'ai oublié personne. Je tiens bien sûr à également remercier les sponsors du consortium RING-Gocad pour avoir financé cette thèse et le personnel de l'ASGA pour avoir assuré son bon fonctionnement (surtout Armelle et Maryse). Merci aux membres du CEREGE m'ayant réservé un chaleureux accueil à Marseille et aux spéléologues ayant fourni les données utilisées lors de cette thèse. N'oublions pas non plus les élèves de l'ENSG dont j'ai pu encadrer les travaux: Gaspard, Zakaria et Pierre.

Merci à ceux qui ont lu, relu et corrigé ce long (très long) manuscrit, surtout Pauline et ma mère mais également mon père, Corentin, Marie et Killian. Merci à ceux présents sur place lors de ma soutenance, Mr et Mme Zingo, Ludovic, Raphael, les 3A... mais aussi à ceux ayant assisté à distance sans forcément comprendre grand-chose (non il n'y a pas d'aspirateur dans le réseau de l'Aspirateur). Merci aux autres potes, je suis trop feignant pour citer tout le monde mais les concernés se reconnaîtront. Enfin, merci à ma famille pour m'avoir soutenu au cours de ces quelques 25 années ayant précédé l'obtention de mon doctorat.

Résumé : Merci à toutes les personnes m'ayant soutenu.

PS: Je dédie cette thèse à feu mon chien Xerus et feu mon chat Voyou.

Table des matières

Résumé	i
Abstract	iii
Remerciements	v
Introduction	1
1 État de l’art	7
1.1 Introduction	8
1.2 Les réseaux karstiques	9
1.2.1 La géomorphologie karstique	9
1.2.2 La formation des réseaux karstiques	14
1.2.3 Classification via motifs morphologiques	28
1.2.4 L’exploration des réseaux	30
1.3 Modélisation et simulation des systèmes karstiques	35
1.3.1 Approches fonctionnelles	36
1.3.2 Approches structurelles	39
1.4 Conclusion	47
2 Inférence statistique	49
2.1 Introduction	50
2.2 Données	51
2.2.1 Base de données utilisée	51
2.2.2 Métriques utilisées	53
2.2.3 Représentation des réseaux via des graphes	55
2.3 Méthodes analytiques	56
2.3.1 Analyses univariées	57
2.3.2 Analyses multivariées	61
2.4 Résultats	64
2.4.1 Analyses univariées	65
2.4.2 Analyses multivariées	71
2.5 Discussion	78
2.6 Conclusion	83

3	Simulation des propriétés géométriques des conduits	85
3.1	Introduction	86
3.2	Méthodes de simulation	86
3.3	Résultats	88
3.3.1	SGS 1D-curvilinéaire	91
3.3.2	SGS 1D-curvilinéaire <i>branchwise</i>	98
3.4	Discussion	105
3.5	Conclusion	107
4	Simulation du squelette des réseaux	109
4.1	Introduction	110
4.2	Concept de la méthode de simulation des squelettes	111
4.3	État actuel de la méthode	116
4.3.1	Calcul de la connectivité	116
4.3.2	Connexion des nœuds	119
4.4	Résultats	119
4.5	Discussion	123
4.6	Conclusion	129
	Conclusions générales	131
	Annexes	137
A	Réseaux utilisés lors de l'étude statistique et résultats	138
B	Analyse statistique des motifs morphologiques	143
C	Analyse du degré des nœuds	145
D	Résultats de simulations de propriétés géométriques	147
	Bibliographie	154

Introduction

Motivations

Les réseaux karstiques

Les grottes ont toujours exercé une certaine fascination sur les Hommes. Lien direct avec le monde souterrain dans de nombreuses mythologies, elles continuent à être le cadre de nombreuses fictions modernes telles le célèbre roman *Voyage au centre de la Terre* (Verne, 1864) ou encore l'*Outreterre* du jeu de rôle *Donjons et Dragons* (e.g., Heinsoo et Collins, 2010). Il s'agit de lieux singuliers, abritant des écosystèmes uniques et précaires, et offrant pour seule vision des murs rocheux et des concrétions appelées *spéléothèmes*, telles que les stalactites et stalagmites. À une époque où les satellites permettent d'obtenir des images précises du monde qui nous entoure, elles restent un des rares territoires peu explorés à la surface de la Terre. Bien que leurs entrées aient pu servir d'habitat à certains de nos ancêtres, l'exploration des grottes à proprement parler, aussi connue sous le nom de spéléologie, ne s'est réellement développée qu'au 19^{me} siècle, sous l'influence de pionniers tels Adolf Schmidl (1802-1863) et Édouard-Alfred Martel (1859-1938). L'évolution de cette discipline a permis l'exploration de nombre d'entre elles et la compréhension des phénomènes leur étant liés.

Les grottes ne sont en soi qu'une partie des grands ensembles géomorphologiques nommés karsts. Les karsts sont des structures géomorphologiques associées à des formes de surface et subsurface caractéristiques, notamment : i) un paysage très fortement altéré en surface, ii) un réseau hydrographique souterrain, iii) de nombreux vides souterrains. Ces vides, qui incluent les grottes, sont qualifiés de réseaux karstiques. Les modes d'expression des karsts en surface peuvent être très variés, allant des paysages montagneux alpins aux célèbres cénotes des mers tropicales.

La formation des karsts est intimement liée à l'action érosive de l'eau. Ils peuvent contenir de grandes réserves hydrauliques et la présence des réseaux fait que les écoulements souterrains diffèrent de ceux observés dans les autres types de nappes souterraines. Pour cette raison, en plus d'être définis en tant que structures géomorphologiques, ils peuvent également l'être en tant qu'aquifères.

Ressources et enjeux

On estime que les karsts couvrent entre 7% et 12% des terres émergées et qu'ils sont source d'eau potable pour près du quart de la population humaine mondiale (Ford et Williams, 2007; Hartmann et al., 2014). En France, les aquifères karstiques présentent un intérêt majeur puisqu'ils contribuent à environ 40% de l'alimentation en eau potable

(Jouves, 2018). Les écoulements sont souvent beaucoup plus rapides dans les réseaux karstiques que dans les milieux poreux, rendant ces systèmes plus vulnérables aux pollutions. Une bonne gestion des aquifères karstiques est donc cruciale pour le bien-être des populations humaines.

D'un point de vue industriel, les principaux enjeux concernent les ressources minérales (e.g., gisements *Mississippi Valley-type* (Jébrak et Marcoux, 2008)) ou pétrolières que peuvent contenir les réseaux paléokarstiques (réseaux karstiques non actifs d'un point de vue hydrogéologique et généralement enfouis). On estime en effet que les réservoirs carbonatés contiennent environ 60% des réserves de pétrole mondiales et 40% des réserves de gaz (Schlumberger, 2007; Jouves, 2018). Les réservoirs carbonatés peuvent être associés à un réseau paléokarstique qui agit comme un drain et contracte fortement les écoulements fluides. Une bonne caractérisation de ces derniers est donc nécessaire pour optimiser l'exploitation du pétrole et du gaz.

La caractérisation des réseaux karstiques est également importante en géotechnique car la présence de vides de grande taille et d'eau peut fortement impacter la stabilité des infrastructures. Les karsts présentent également un intérêt dans d'autres domaines tels que la climatologie, l'archéologie ou encore la paléontologie.

Simulation des réseaux

La modélisation numérique des karsts présente un intérêt certain pour l'exploitation des ressources associées. Les informations qu'apportent les modèles peuvent en effet améliorer la compréhension de la structure des karsts et de leur fonctionnement. On distingue différents types de modèles selon leurs objectifs. Les méthodes *fonctionnelles* ont pour but de simuler le fonctionnement hydraulique de l'aquifère tandis que les méthodes *structurelles* cherchent à simuler explicitement des réseaux discrets (Jeannin, 1996; Fournillon, 2012; Jouves, 2018).

Les méthodes *fonctionnelles* sont particulièrement importantes pour la gestion des ressources en eau présentes dans les aquifères karstiques. Elles permettent de reproduire le comportement dynamique de l'aquifère au moyen de modélisations d'écoulement (e.g., Teutsch, 1993; Scanlon et al., 2003; Mazzilli et al., 2017). Cela permet par exemple de prédire les changements de l'aquifère en fonction du pompage et du climat, ou d'évaluer l'impact d'éventuelles pollutions.

Les méthodes *structurelles* se concentrent sur la simulation des réseaux eux-mêmes afin de mieux caractériser les incertitudes qui leur sont associées. Plusieurs méthodes visant à simuler des réseaux karstiques discrets ont été développées au cours de la dernière décennie (e.g., Pardo-Igúzquiza et al., 2012; Borghi et al., 2012; Collon-Drouaillet et al., 2012; Fournillon et al., 2012; Viseur et al., 2014; Hendrick et Renard, 2016a; Mourlanette et al., 2016). La majorité d'entre elles se basent sur des processus stochastiques afin d'obtenir plusieurs réalisations équiprobables, ce qui permet de mieux appréhender les incertitudes associées. Certaines approches cherchent à reproduire les processus physiques et chimiques contrôlant la formation des réseaux (e.g., Kaufmann et Braun, 2000; Birk et al., 2003; Dreybrodt et al., 2005). D'autres se focalisent directement sur la reproduction de la structure finale en se servant de techniques basées sur les statistiques spatiales (e.g., Fournillon et al., 2012; Viseur et al., 2014). Des simplifications et approximations du processus de formation, combinées ou non à des

statistiques spatiales, ont également permis de proposer des approches de modélisation originales (e.g., Jaquet et al., 2004; Labourdette et al., 2007; Pardo-Igúzquiza et al., 2012; Borghi et al., 2012; Collon-Drouaillet et al., 2012; Hendrick et Renard, 2016a). Certains logiciels permettent d’effectuer des simulations d’écoulement sur ces réseaux discrets (e.g., SWMM, Epanet, Modflow-CFP).

Limites actuelles et objectifs de la thèse

Malgré de nombreuses études, les spéléologues ne peuvent recueillir des données que dans les parties accessibles des réseaux karstiques, ce qui mène à des explorations partielles. Les essais de traçage permettent d’obtenir des informations intéressantes sur les écoulements fluides en leur sein mais ne permettent pas de déterminer des topologies (relations de voisinage entre les conduits) et dimensions précises. L’acquisition de données dans les réseaux paléokarstiques est encore plus complexe et force à utiliser d’autres techniques (e.g., méthodes sismiques, étude pétrophysique des déblais de forage) ayant leurs propres limitations. Ainsi, les techniques d’exploration classiques produisent des données éparses conduisant à de grandes incertitudes concernant la position des conduits et leurs dimensions. Ces paramètres étant essentiels aux simulations d’écoulement, c’est dans ce contexte que les méthodes *structurelles* sont principalement utilisées.

Les différentes méthodes développées au cours de la dernière décennie permettent en effet de mieux caractériser ces incertitudes mais elles présentent certaines limitations. La majorité d’entre elles ne permettent pas d’intégrer toutes les informations géologiques et observations de terrain connues de l’utilisateur (e.g., essais de traçage, indices de karstification), ne permettent généralement pas d’imposer un type d’organisation topologique et géométrique au réseau de drains et se basent sur l’utilisation de grilles, ce qui peut mener à la génération de conduits en *marches d’escalier* et à des temps de calculs importants. Les dimensions des conduits simulés ne se basent de plus habituellement pas sur des observations de terrain.

Dans ce contexte, l’objectif de la thèse était de développer une nouvelle méthode de simulation stochastique de réseaux karstiques discrets prenant en compte différentes informations contraignant la position des conduits et leurs dimensions. Ce travail s’est articulé autour de 3 principaux axes de recherche : i) l’analyse statistique de réseaux connus pour identifier d’éventuelles règles statistiques génériques, ii) le développement de méthodes capables de simuler les dimensions des conduits et iii) le développement d’une méthode permettant de simuler le squelette des réseaux sous forme de graphes.

Plan détaillé et résumé des contributions

Le présent manuscrit de thèse se compose de quatre chapitres et de quatre annexes.

Le premier chapitre consiste en un état de l’art des connaissances générales sur les karsts et des différentes méthodes de simulation existantes.

Le second chapitre présente les résultats de l’analyse statistique d’une base de données forte de 49 réseaux karstiques différents. Elle se focalise principalement sur l’étude de la géométrie des conduits, c’est à dire leurs dimensions, à travers l’analyse de deux métriques : le rayon équivalent et le ratio largeur-hauteur. Les réseaux sont analysés individuellement puis comparés afin d’identifier des règles statistiques génériques. La

variabilité spatiale de ces propriétés géométriques est caractérisée, notamment grâce au développement de variogrammes 1D-curvilinéaires. L'existence d'une potentielle organisation hiérarchique des conduits est également évaluée à travers l'étude de l'organisation des propriétés géométriques étudiées.

Le troisième chapitre propose deux méthodes de simulation de propriétés le long de réseaux karstiques. Elles se basent sur les résultats présentés dans le Chapitre 2. La première méthode se focalise sur la reproduction de la variabilité des propriétés à l'échelle du réseau, tandis que la deuxième se concentre sur la reproduction de la variabilité au sein et entre les branches du réseau. Leur but est de permettre de simuler des dimensions de conduits semblables aux observations de terrain. Ces méthodes sont appliquées à quatre cas synthétiques et un cas réel.

Le quatrième et dernier chapitre présente un prototype de méthode dont le but est de simuler stochastiquement des réseaux karstiques discrets sous forme de graphes. Il est espéré qu'une fois entièrement développée, cette méthode permette de simuler différents types de réseaux tout en prenant directement en compte les données de terrain et les informations géologiques. Bien qu'elle ne soit encore qu'aux premiers stades de son développement, ses principes généraux, son état actuel et des exemples d'application sont présentés et commentés.

L'annexe A présente les résultats détaillés de l'analyse statistique des 49 réseaux étudiés, tandis que l'annexe B se focalise sur les résultats des différents motifs morphologiques. L'annexe C présente les résultats détaillés de l'analyse du degré des nœuds des réseaux. Enfin, l'annexe D correspond à certains résultats, jugés moins essentiels, des méthodes de simulation présentées dans le Chapitre 3.

Ainsi, les contributions majeures de cette thèse sont les suivantes :

- L'analyse statistique d'une base de données comprenant 49 réseaux karstiques différents et le développement d'outils dédiés à cet effet.
- Le développement de deux méthodes de simulation permettant de générer la géométrie (les dimensions) des conduits.
- Les premières étapes du développement d'une méthode de simulation de squelette de réseau karstique.

Contexte opérationnel de la thèse

Cette thèse a été effectuée entre 2018 et 2021, au sein de l'équipe RING¹ (*Research for Integrative Numerical Geology*) du laboratoire Georessources² de l'Université de Lorraine. Les recherches effectuées au sein du laboratoire Georessources portent sur toutes les étapes de détection, d'acquisition et de traitement des ressources géologiques, principalement les ressources minières et pétrolières. Dans ce contexte, les travaux de l'équipe RING se concentrent principalement sur la modélisation du sous-sol et le traitement des problèmes associés, afin d'optimiser l'exploitation desdites ressources.

¹<https://www.ring-team.org>

²<http://georessources.univ-lorraine.fr/>

Cette thèse a été financée par le consortium RING-Gocad, géré par l'ASGA (Association Scientifique pour la Géologie et ses Applications) dont j'ai été l'employé pendant un CDD de 3 ans. J'ai passé la majorité de ces 3 années au sein des locaux de l'équipe RING, situés dans l'ENSG (École Nationale Supérieure de Géologie). Mes travaux s'inscrivent dans la continuité d'autres travaux réalisés par des membres de l'équipe RING, notamment ceux de Pauline Collon et de Guillaume Rongier. Ce travail a bénéficié des commentaires éclairés de Sophie Viseur et Philippe Renard, membres du comité de suivi, et se base sur des données partagées par plusieurs spéléologues (principalement B. Arfib, J. Bodin, J. Botazzi, J-P. Cassou, J-L. Galera, S. Jaillet, P-Y. Jeannin, J. Jouvès, T. Kincaid, E. Pardo-Iguzquiza, C. Vuilleumier et les membres de *the Quintana Roo Speleological Survey*). Une partie de ce travail a été effectuée dans le plugin RingKarst du logiciel SKUA-Gocad. Il s'agit d'un logiciel de géomodélisation, actuellement exploité par la société Emerson³.

Une partie des mes travaux ont été publiés dans la revue *Geomorphology* (Frantz et al., 2021) et ont pu être présentés lors d'une conférence internationale : l'IAMG 2019 (Frantz et al., 2019). Ils ont également été présentés lors de trois colloques organisés par l'équipe RING et principalement destinés aux membres du consortium RING-Gocad (RING Meetings 2018, 2019 et 2020).

Publications associées

Publications dans des journaux à comité de lecture

- Y. Frantz, P. Collon, P. Renard et S. Viseur (2021). Analysis and stochastic simulation of geometrical properties of conduits in karstic networks. *Geomorphology*, 377. doi : 10.1016/j.geomorph.2020.107480.

Publications dans des actes de conférences internationales

- Y. Frantz, P. Collon, P. Renard et S. Viseur (2019). Statistical analysis and simulation of conduit dimensions within karstic systems. *Annual Conference IAMG 2019*.

Publications pour le RING Meeting

- Y. Frantz, P. Collon et P. Renard (2018). Statistical analysis of conduit dimension distributions within karstic systems. *RING Meeting 2018*.
- Y. Frantz, P. Collon, P. Renard et S. Viseur (2019). Stochastic simulations of conduit dimensions inside karstic networks. *RING Meeting 2019*.
- Y. Frantz et P. Collon (2020). Stochastic simulations of discrete karstic networks. *RING Meeting 2020*.

³<https://www.emerson.com/fr-fr>

Chapitre 1

État de l'art

Sommaire

1.1	Introduction	8
1.2	Les réseaux karstiques	9
1.2.1	La géomorphologie karstique	9
1.2.2	La formation des réseaux karstiques	14
1.2.3	Classification via motifs morphologiques	28
1.2.4	L'exploration des réseaux	30
1.3	Modélisation et simulation des systèmes karstiques	35
1.3.1	Approches fonctionnelles	36
1.3.2	Approches structurelles	39
1.4	Conclusion	47

Résumé Ce chapitre correspond à un état de l'art sur les karsts et leur modélisation. Je me focalise dans un premier temps sur les karsts en tant qu'objets géologiques. L'accent est mis sur les processus de formation et sur l'organisation interne des réseaux karstiques. Les différentes méthodes permettant l'acquisition de données sur les réseaux sont également présentées. Je décris dans un second temps les deux grandes familles de méthodes de modélisation des systèmes karstiques : les approches *fonctionnelles* visant à simuler leur fonctionnement hydrologique et les approches *structurelles* visant à générer des réseaux de conduits discrets.

1.1 Introduction

Le nom *karst* provient du mot allemand *Karst*, équivalent du mot slovène *Kras* et du mot italien *Carso*. Tous trois sont probablement dérivés du mot latinisé *Carsus*, le mot pré-indo-européen original ayant pour base *Karus-* dont la racine *kar* est identique à celle du mot *karra/gara*, signifiant pierre (Gams, 1993; Kranjc, 2011). Il désignait initialement un plateau carbonaté des Alpes dinariques, s'étendant du Nord-est de l'Italie au Nord-Ouest de la Croatie et dont la majorité se situe en Slovénie.

L'emploi du mot *karst* s'est démocratisé pour désigner un type de paysage dont les roches sont très fortement affectées par des processus de dissolution. Cette définition peut cependant être discutée au vu des nombreux aspects géologiques et hydrologiques qui leur sont liés. En effet, suivant le domaine, le karst sera tantôt considéré comme une structure géomorphologique et tantôt comme un aquifère. Les karsts sont principalement associés à des roches solubles, notamment des carbonates ou des évaporites, voire à des roches à ciment carbonaté. Dans le cas de structures similaires formées par dissolution de roches non carbonatées, le terme *parakarst* est plutôt employé. Il arrive que des structures ressemblent aux karsts sans avoir été formées via des processus de dissolution. Ces dernières sont alors désignées par le nom de *pseudokarsts*. Les pseudokarsts peuvent avoir des origines diverses comme le refroidissement de tubes de lave, l'écoulement d'eau dans des glaciers ou l'action de la marée sur des falaises (e.g., Halliday, 2007, Holler, 2019).

On estime que les karsts couvrent entre 7% et 12% des terres émergées. 21.6% de la surface terrestre européenne serait constituée de roches carbonatées mais seulement 13.8% affleuraient (Chen et al., 2017) (Figure 1.1). En France, les formations carbonatées concernent environ 35% des roches affleurantes et 49.6% du territoire en comptant les formations de subsurface (Chen et al., 2017). Bien que le degré soit variable, toutes ces formations sont potentiellement associées à des phénomènes de karstification. Les aquifères karstiques sont source d'eau potable pour près du quart de la population humaine mondiale (Ford et Williams, 2007; Hartmann et al., 2014) et contribuent à environ 40% de l'alimentation en eau potable française (Jouves, 2018).

Ce chapitre dresse un état de l'art sur les systèmes karstiques et se divise en deux parties. La première consiste en une description générale des karsts, des processus associés et des méthodes d'analyse existantes. Elle commence par une présentation de la géomorphologie et des réseaux karstiques, plus généralement désignés par le terme

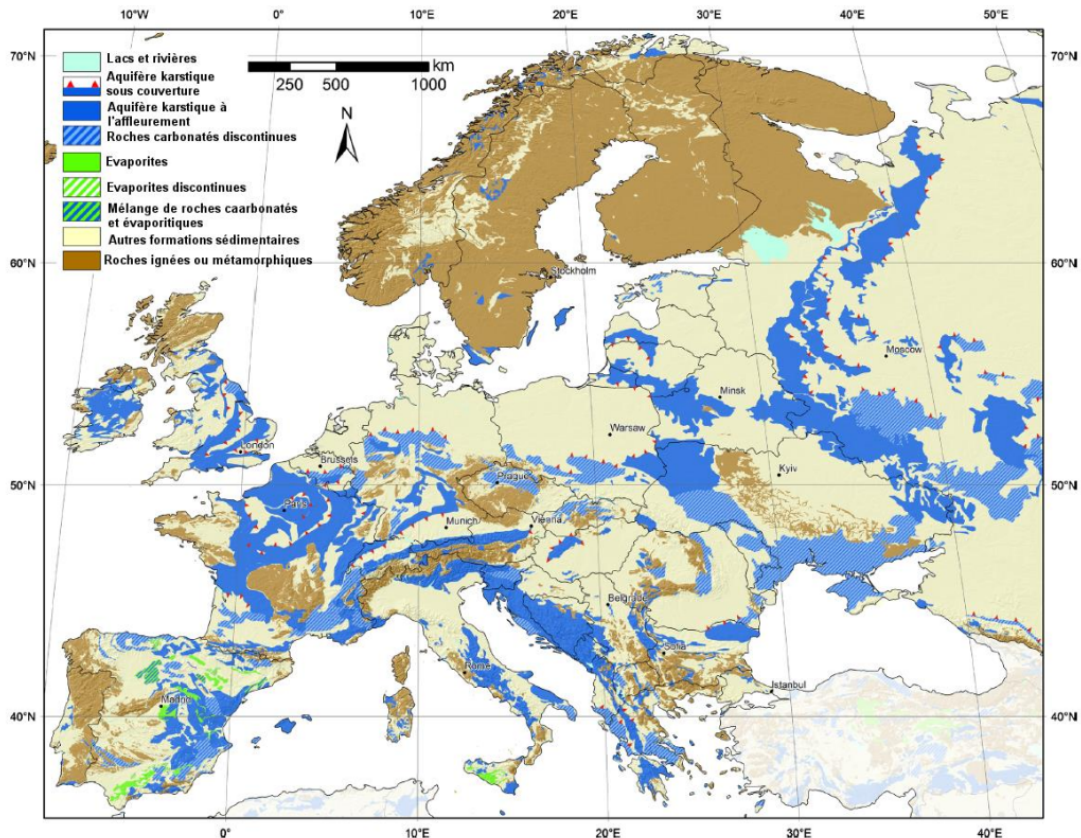


Figure 1.1 – Carte des formations carbonatées karstifiables d'Europe (Lorette, 2019, modifié d'après Chen et al., 2017).

vague de *grottes*. S'en suit une présentation de la spéléogénèse, terme faisant référence aux phénomènes directement impliqués dans la formation des réseaux. Sont d'abord présentés les processus de karstification, les généralités sur la spéléogénèse et l'impact de l'origine des eaux. Une classification des réseaux basée sur leur organisation est ensuite exposée. Enfin, différentes techniques permettant l'acquisition de données sur les réseaux sont décrites. La deuxième partie se focalise sur les méthodes de simulation des systèmes karstiques. On distinguera les approches *fonctionnelles* visant à simuler leur fonctionnement hydrologique des approches *structurelles* visant à générer des réseaux de conduits discrets.

1.2 Les réseaux karstiques

1.2.1 La géomorphologie karstique

Les paysages karstiques sont souvent facilement reconnaissables par certains marqueurs de surface et de subsurface caractéristiques (Figure 1.2). La partie surfacique des karsts est appelée *exokarst*, en opposition à la partie souterraine, incluant notamment les réseaux, appelée *endokarst*. Les principaux éléments indiquant la présence d'une zone karstifiée sont :

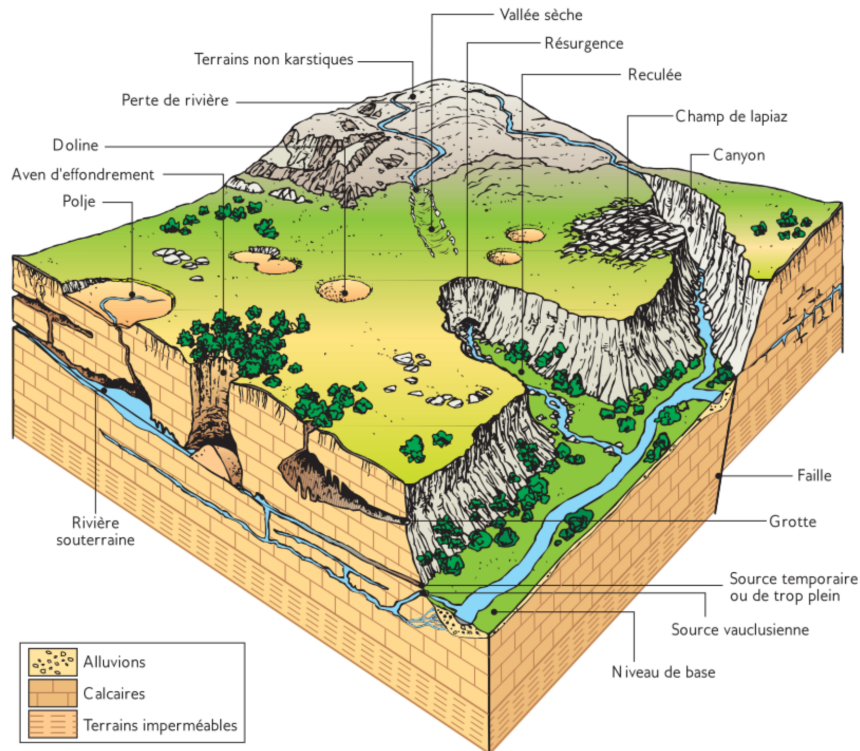


Figure 1.2 – Bloc-diagramme idéalisé d'un paysage karstique (Bakalowicz, 1999).

- Un paysage très fortement altéré.
- La présence d'un réseau hydrographique souterrain.
- De nombreux vides souterrains, parfois qualifiés de grottes.

Bien que la majorité des paysages terrestres soient façonnés par l'action érosive de l'eau, les paysages karstiques en sont peut-être l'exemple le plus frappant. Ils sont souvent associés à des marqueurs à grande échelle comme les vallées mais également à des formes d'altération caractéristiques à plus petite échelle, issues de la dissolution des roches calcaires en surface mais aussi en subsurface (Figure 1.3).

Les plus notables de ces formes d'altération de surface sont les lapiaz et les dolines, principaux points d'infiltration de l'eau dans la roche sous-jacente. Les lapiaz, connus sous d'autres noms comme *lapiés* ou *karren*, sont des formes d'altération de surface généralement semblables à des fissures ou rigoles empruntées par l'eau de pluie, et pouvant être de formes très diverses (Lundberg, 2013). Leurs tailles vont de l'ordre du centimètre à la dizaine de mètres pour les grandes, la majorité étant de taille inférieure à 2 mètres (Ford et Williams, 2007; Lundberg, 2013). Bien qu'elles puissent également être observables sur des roches beaucoup moins solubles, elles restent le principal indicateur surfacique de la présence de karsts lorsque le paysage est peu marqué. Les lapiaz sont souvent connectés, formant de grands ensembles appelés champs de lapiaz (Figure 1.3A,B).

Les dolines (*sinkhole*) sont des dépressions plus ou moins grandes formées par une dissolution de la roche en profondeur, parfois couplée à un affaissement de surface (Figure 1.3B,C). Le diamètre des dolines peut aller d'un mètre à presque un kilomètre et

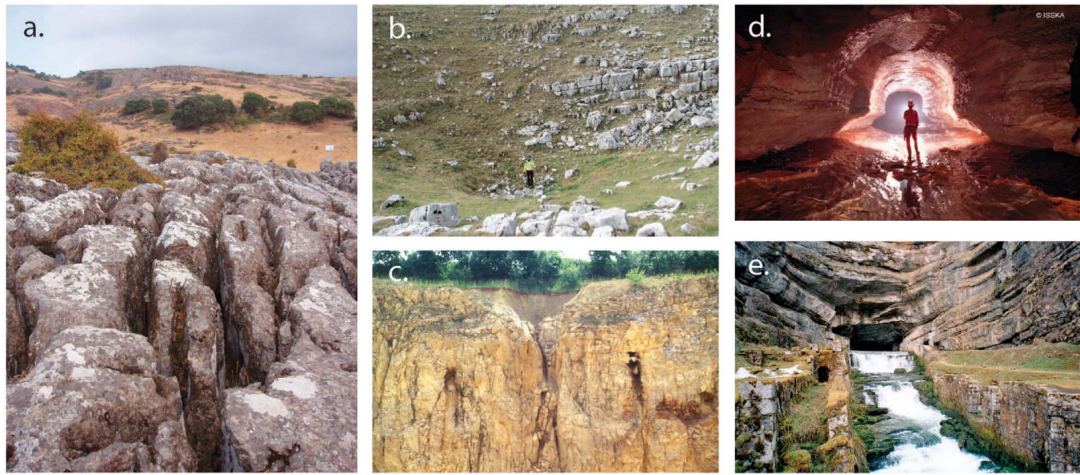


Figure 1.3 – Marqueurs de surface et de subsurface typiques des karsts. Compilation d’Hartmann et al., 2014. A : Champ de lapiaz au Nord de Málaga, Espagne (photographie par Andreas Hartmann). B : Champ de lapiaz et doline dans le Jura Suisse (photographie par Nico Goldscheider). C : Coupe verticale à travers la zone épikarstique et doline observée dans une carrière du Jura souabe, Allemagne (photographie par Nico Goldscheider). D : Grotte en Ardèche, France (photographie par Remy Wenger, SISKKA). E : Source de la Loue, France (photographie par Nico Goldscheider).

leur profondeur de quelques mètres à plusieurs centaines, ce qui engendre des pentes d’inclinaison variable et des formes générales diverses et variées (Ford et Williams, 2007). Leurs tailles et leurs origines sont certes variables (Delle Rose et al., 2004; Waltham et Fookes, 2005) mais elles sont toujours liées à une, voire plusieurs évacuations d’eau souterraine. Les principaux processus de formation des dolines sont une dissolution progressive de la roche présente en surface et en profondeur (dolines de dissolution), la formation d’une cavité en profondeur associée à l’effondrement de sa couverture (dolines d’effondrement), ou l’évacuation de sédiments de surface par l’eau dans le réseau hydrographique souterrain (dolines de subsidence) (Waltham et Fookes, 2005). Les dolines d’effondrement (et parfois les dolines de subsidence) peuvent mener à la formation de grand gouffres (ou abîmes) très visibles dans le paysage.

Les poljés correspondent à des dépressions à fond plat de grande dimension (parfois plusieurs centaines de kilomètres carrés) et au moins partiellement closes, associées à un drain du réseau souterrain, via une perte appelée ponor. De nombreux poljés sont inondés lors des périodes humides et asséchés lors des périodes sèches. En effet, le niveau piézométrique (altitude de la nappe phréatique) étant habituellement plus haut en période humide qu’en période sèche, ce dernier peut dépasser l’altitude du fond du poljé, menant à son inondation. Ce phénomène joue lui-même un rôle important pour la formation des poljés. En période sèche, une bonne partie de l’eau de surface tend à être évacuée dans le réseau souterrain à travers le ponor. Cette alternance mène à l’existence de lacs temporaires, présents uniquement en période humide.

Les karsts sont associés à de nombreux vides souterrains tridimensionnels interconnectés, qualifiés de réseaux karstiques (Figure 1.3D). Ils sont communément appelés grottes lorsque les réseaux sont horizontaux, gouffres, abîmes ou avens quand les réseaux sont très verticaux. La notion de réseau karstique peut s’étendre à différentes échelles,

allant de la porosité des roches à des vides plurimétriques. Ces réseaux sont habituellement formés par différents processus de dissolution (cf Section 1.2.2) et sont le lieu de flux hydrauliques souterrains. Les vides les plus importants, également appelés conduits karstiques, sont les principaux drains du réseau hydraulique souterrain. L'eau s'écoulant dans les réseaux en sort généralement par des exutoires, appelés sources, pouvant être associés à des débits très importants (Figure 1.3E). Ainsi, les aquifères karstiques se distinguent des aquifères traditionnels sur plusieurs points : i) la majeure partie des écoulements a lieu dans de grands vides interconnectés (les réseaux karstiques), ii) les écoulements sont de fait hétérogènes sur l'ensemble de l'aquifère, iii) les vitesses d'écoulement dans les conduits peuvent être très importantes (pouvant atteindre parfois plusieurs centaines de mètres par heure), iv) les débits sur l'ensemble du réseau sont très variables dans le temps, pouvant aller d'une valeur nulle sur une période de plusieurs mois à plusieurs dizaines de m^3/s à la source (Bakalowicz, 2005). L'eau traverse également la roche encaissante, appelée matrice, grâce à sa perméabilité. Les flux matriciels sont assez similaires à ceux intervenant dans une nappe phréatique classique et sont beaucoup plus lents et moins importants que ceux passant par les conduits, représentant rarement plus d'1% du total des flux (e.g., Worthington et al., 2000). La matrice joue néanmoins un rôle très important vis-à-vis du stockage de l'eau, représentant plus de 96% du stockage dans les aquifères étudiés par Worthington et al. [2000]. Bien que la porosité matricielle fasse techniquement partie du réseau hydraulique, nous désignerons par réseau karstique l'ensemble des conduits servant de principaux drains pour l'eau souterraine, c'est à dire au moins pluricentimétriques. L'étendue des réseaux peut être très variable, allant d'à peine quelques dizaines de mètres de longueur à plus de 650 kilomètres comme c'est le cas pour la grotte *Mammoth Cave* (Kentucky, USA), qui est actuellement le plus étendu des réseaux connus (Palmer, 2016). Quand ces réseaux ne sont plus submergés, des phénomènes de précipitation des matières dissoutes par l'eau, lors de son passage à travers la roche encaissante (notamment des carbonates), mènent à la formation de concrétions nommées spéléothèmes (e.g., stalactites, stalagmites).

Les paléokarsts correspondent à des karsts n'étant plus associés à une activité hydraulique et n'évoluant plus. Les plus anciens exemples connus sont situés dans des roches archéennes du bouclier canadien et auraient été formés il y a plus de 2.5 milliards d'années. Ils sont habituellement remplis par des sédiments plus récents (généralement d'origine sédimentaire mais parfois également volcanique (e.g., Osborne, 2017)) et, à terme, finissent enfouis en même temps que les couches géologiques dans lesquelles ils sont présents. Il arrive cependant que des paléokarsts soient exhumés et retrouvent une activité hydraulique. Les paléokarsts constituent fréquemment des réservoirs de ressources telles que du pétrole, du gaz ou divers minerais.

L'expression des karsts en surface est variable suivant la zone géographique. Elle dépend, en effet, très fortement du temps de karstification mais également du climat. Les morphologies les plus classiques correspondent aux karsts des régions tempérées. D'autres morphologies sont observables dans les régions aux climats plus extrêmes, comme les zones froides et polaires, les zones arides et surtout les zones tropicales (Salomon et Pulina, 2005). Les karsts des régions polaires et montagneuses situés sous la glace et ceux des régions froides situés sous le permafrost sont souvent difficilement

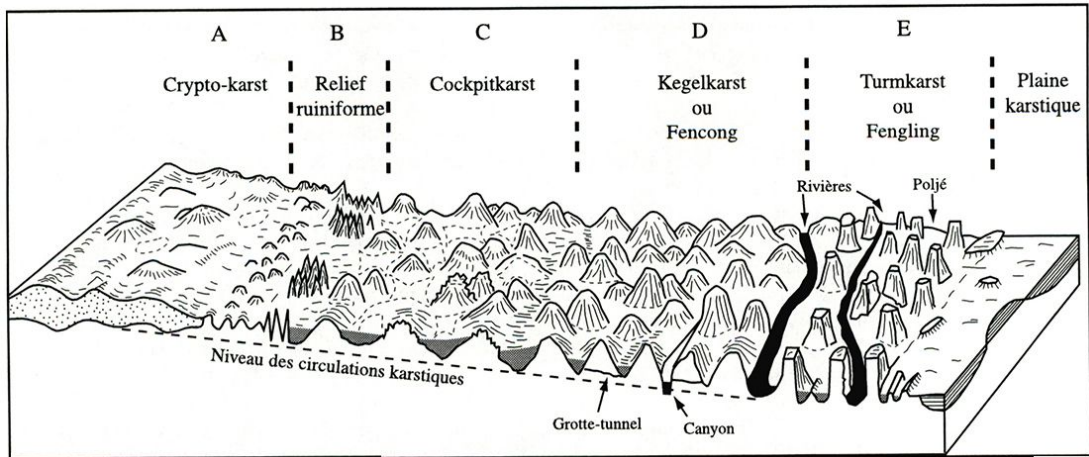


Figure 1.4 – Schéma théorique de l'évolution des différents types de karsts tropicaux chinois (Salomon, 2000).

observables en surface (Salomon et Pulina, 2005). La fonte partielle de la glace en période estivale est fréquente dans les zones subpolaires mais rare dans les zones polaires. Les karsts présents dans les régions subpolaires sont généralement issus de la circulation d'eaux provenant de cette fonte partielle, tandis que ceux des régions polaires sont communément formés par la circulation d'eaux thermales liées à l'activité volcanique (Salomon et Pulina, 2005). Dans les zones tempérées froides, c'est principalement la fonte du pergélisol qui apporte l'eau nécessaire aux processus de karstification. Les glaciers présents en altitude dans des zones plus tempérées peuvent être à l'origine de phénomènes de karstification similaires (e.g., Maire, 1978; Bini et al., 1998; Skoglund et al., 2010), la dissolution de la roche s'effectuant en parallèle de processus d'érosion typiques des glaciers comme l'abrasion par les sédiments charriés ou l'excavation glaciaire (*plucking*) (Herman et Braun, 2008; Beaud et al., 2014). Les eaux glaciaires étant très agressives et le niveau piézométrique pouvant être situé à une altitude relativement basse, il en résulte des réseaux très développés verticalement dans les parties non saturées (Palmer, 1991). Les karsts des milieux arides et semi-arides sont également peu observables en surface et, bien que différentes études aient été réalisées, ils restent moins caractérisés que les karsts des milieux tempérés. Dans ces milieux, les précipitations sont souvent rares mais brutales. L'eau s'écoule rapidement et ce contact rapide avec la roche limite le développement du karst dans les zones proches de la surface. Les karsts tropicaux sont associés à des paysages très caractéristiques (Figure 1.4). Ils sont soumis à des processus de dissolution beaucoup plus importants et longs qu'en climat tempéré, la majorité de la dissolution intervenant dans les zones superficielles (Palmer, 1991). Il en résulte des reliefs très ruiniformes dans lesquelles de hautes barres calcaires escarpées peuvent alterner avec des zones beaucoup plus basses et planes. Les principales conditions de formation des karsts tropicaux seraient des températures élevées (moyenne annuelle supérieure à 17-18 °C), des précipitations annuelles importantes (supérieures à 1000 millimètres) et une grande activité biologique (Salomon et Pulina, 2005). L'absence de cryoclastie est également une des raisons permettant la conservation de structures très hautes, contrairement aux zones tempérées.

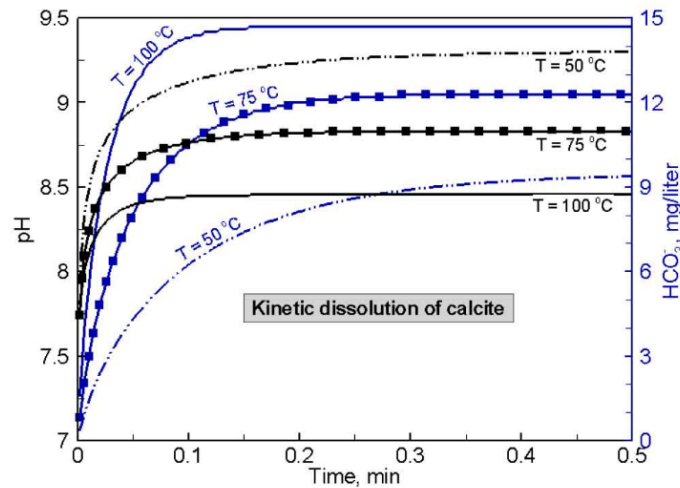


Figure 1.5 – Cinétique de dissolution de la calcite à différentes températures (Khaledialidusti et Kleppe, 2017).

1.2.2 La formation des réseaux karstiques

Tandis que la karstification désigne l'ensemble des processus de formation des karsts à l'échelle du massif, le terme de spéléogénèse fait plutôt référence à la formation des réseaux de conduits associés. Ces deux phénomènes sont intrinsèquement liés puisque le deuxième est la conséquence directe du premier. Les processus de dissolution impliqués lors de la spéléogénèse sont de fait les mêmes que pour la karstification. Seuls les processus associés aux roches carbonatées seront décrits par la suite.

1.2.2.a Les processus de karstification

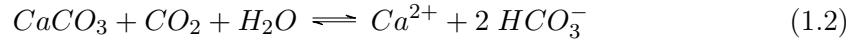
La karstification est un procédé complexe, impliquant de nombreux phénomènes. Elle combine altération chimique et altération mécanique, mais l'altération chimique est prépondérante. Le phénomène le plus important correspond à la dissolution chimique des roches carbonatées, généralement couplée à une évacuation mécanique des produits de dissolution de façon synchrone ou asynchrone. Le principal processus d'altération impliqué lors de la karstification est la dissolution du calcaire par l'eau. Son équation est la suivante (Buhmann et Dreybrodt, 1985; Ford et Williams, 2007) :



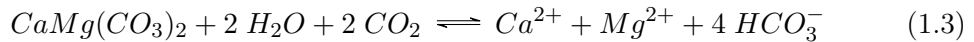
La température a un double impact sur la dissolution de la calcite par l'eau (Palmer, 1991; Khaledialidusti et Kleppe, 2017) : plus la température est élevée, plus la cinétique de dissolution de la calcite est importante (Figure 1.5). Cependant, contrairement à la majorité des solides, la concentration en calcite dissoute à l'équilibre est inversement proportionnelle à la température. Cela signifie que, bien que la calcite soit dissoute plus rapidement à haute température, l'état d'équilibre est atteint pour une concentration finale plus faible en Ca^{2+} et HCO_3^- .

La cinétique de cette réaction reste relativement lente. Les carbonates sont en effet normalement peu solubles dans l'eau. Cette solubilité va néanmoins fortement augmenter en fonction de la teneur en CO_2 , passant de 60 mg.L^{-1} pour une pression

partielle $P_{CO_2} = 10^{-3} \text{ bar}$ à 400 mg.L^{-1} pour $P_{CO_2} = 10^{-1} \text{ bar}$ dans le cas d'une eau à 25 degrés Celsius et à une pression de 1 bar (Ford et Williams, 2007). La dissolution du calcaire s'opère donc généralement par une eau enrichie en CO_2 , d'après l'équation suivante (Frisia et Borsato, 2010; Ford et Williams, 2007) :

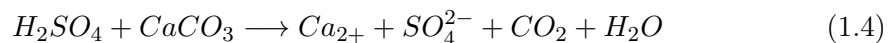


Pour que l'altération puisse progresser de manière efficace, il faut que le flux d'eau de subsurface soit suffisamment important pour évacuer l'eau saturée et ainsi garder de l'eau non saturée en contact avec la roche. Une eau saturée ne peut en effet plus dissoudre de calcaire. Des processus similaires s'appliquent pour d'autres roches carbonatées, notamment les dolomites (Ford et Williams, 2007) :



L'eau météorique est initialement associée à une phase gazeuse de CO_2 atmosphérique mais sa pression partielle augmente lors de son trajet dans le sol, amplifiant l'agressivité de l'eau. Dans la partie superficielle des réseaux, P_{CO_2} est souvent égale à celle du massif, résultant en un équilibre entre l'air, l'eau et la roche subissant le processus de dissolution (Houillon, 2016). Le CO_2 des sols provient de plusieurs sources, principalement de la respiration racinaire et de l'activité microbienne décomposant la matière organique (Salomon et Pulina, 2005; Houillon, 2016). Nonobstant, il est estimé qu'environ 90% du CO_2 fabriqué dans le sol s'échappe dans l'atmosphère et qu'uniquement 10% contribue à l'agressivité des eaux d'infiltration (Palmer, 2007). Les concentrations en CO_2 du sol varient fortement selon les régions, principalement en lien avec le climat, et sont une des raisons des phénomènes de karstifications plus intenses dans les zones superficielles des régions tropicales (Brook et al., 1983). En fonction de l'évolution de différents paramètres, notamment la concentration de l'eau en carbonates, sa température et la pression partielle en CO_2 , des phénomènes de reprécipitation des carbonates peuvent survenir.

La dissolution des roches carbonatées n'est cependant pas exclusive aux eaux riches en CO_2 , différents acides pouvant contribuer grandement aux processus de dissolution (Audra et Palmer, 2013). Pour la majorité des réseaux de surface, il s'agira principalement d'acides organiques issus de l'activité biologique dans les sols, notamment dans le cas des karsts tropicaux. Leur rôle exact est cependant relativement mal connu. Dans d'autres cas, il s'agit d'acide sulfurique, voire un mélange d'acide sulfurique et d'eaux riches en CO_2 ou d'eaux ayant des propriétés différentes (e.g., température, concentration en différents éléments), qui seront présentées en Section 1.2.2.c. L'équation bilan du processus de dissolution du calcaire par l'acide sulfurique est la suivante (Ford et Williams, 2007; Galdenzi et al., 2008) :



Il existe trois principaux processus de karstification aboutissant à des structures de drainage spécifiques : la corrosion sur roche nue, la crypto-altération et la fantômisiation (Quinif, 1999; Jouvès, 2018). La corrosion sur roche nue est le processus le plus courant. Elle correspond à l'altération chimique de la roche calcaire par dissolution et

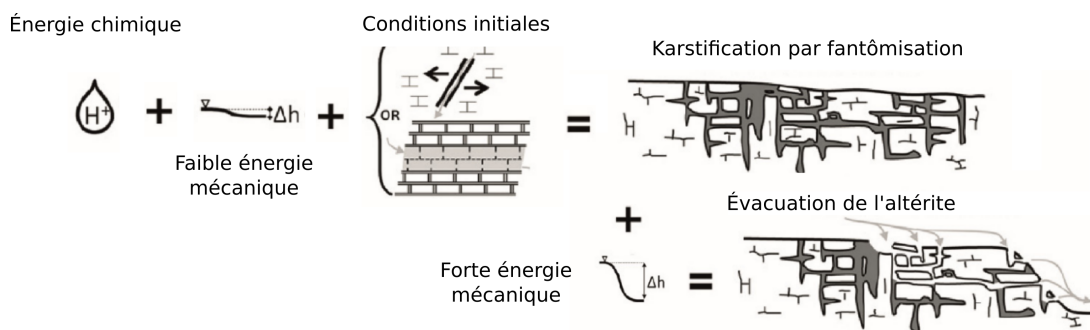


Figure 1.6 – Schéma du processus de fantômisiation. Dans une première phase, les minéraux solubles sont évacués chimiquement. Dans une deuxième phase, les minéraux restants (le fantôme de roche) sont évacués mécaniquement (modifié d'après Jouves, 2018 et Dubois et al., 2014).

à l'évacuation mécanique de la matière dissoute et des résidus. Le contact de la roche et de l'eau aboutit aux formes sur roches nues en surface et dans l'endokarst (Jouves, 2018). Les autres principaux processus de karstification des roches carbonatées sont la crypto-altération et la fantômisiation. La crypto-altération correspond à l'altération de la roche sous une couche perméable non karstifiable (généralement des sables) ayant permis la formation d'un aquifère captif (le toit de la couche calcaire n'affleure pas). Les dépressions formées par dissolution (crypto-lapiaz, crypto-dolines) se retrouvent progressivement comblées par la couche toit (Bonte, 1962; Vergari et Quinif, 1997; Jouves, 2018). Elle mène à la formation de karsts sous couverture. La fantômisiation consiste initialement en une altération chimique des minéraux situés sous le niveau piézométrique (dans la zone saturée en eau) (Figure 1.6). Les minéraux solubles sont emportés par l'eau tandis que les autres restent sur place et forment le *fantôme* de roche. Dans un second temps, les minéraux résiduels sont évacués mécaniquement, généralement lors de l'augmentation d'énergie potentielle liée à une baisse du niveau piézométrique (Quinif et al., 2014; Jouves, 2018). Cette seconde phase peut former des cavités en seulement quelques années (e.g., Bruxelles et al., 2009) et les réseaux formés via fantômisiation peuvent se retrouver superposés à un autre réseau, qui aura apporté l'eau nécessaire à cette altération mécanique.

1.2.2.b Généralités de la spéléogénèse

La spéléogénèse est la partie du processus de karstification conduisant à la formation des réseaux karstiques. Elle a lieu sur de longues périodes temporelles, allant parfois jusqu'à plusieurs dizaines voire centaines de millions d'années. L'eau s'infiltré initialement dans les roches à travers leur porosité et surtout le long de fissures. Elle suit habituellement le gradient hydraulique de son point d'entrée jusqu'au niveau piézométrique, les écoulements se faisant principalement sous l'action de la gravité. Lors de son passage, l'eau dissout la roche avec laquelle elle est en contact, menant au développement d'une porosité interconnectée sous forme de conduits qui s'élargissent au fil du temps. La majorité de la dissolution initiale se faisant dans la zone de surface, l'eau s'infiltrant dans les zones les plus profondes est presque entièrement saturée en carbonates dissous, limitant la poursuite de la dissolution. L'accès aux zones profondes est facilité par le

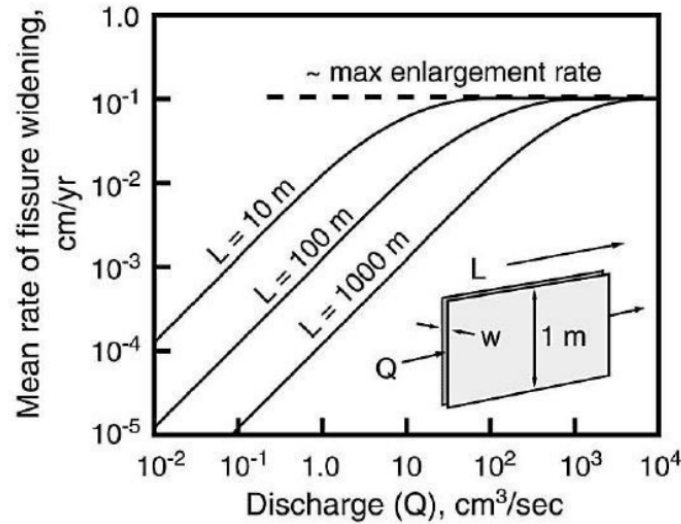


Figure 1.7 – Taux d’élargissement moyen d’une fissure dans du calcaire, en fonction du débit (Q) et de la distance d’écoulement (L). Q correspond au débit par mètre de hauteur de la fissure (plus longue dimension de la section transversale de la fissure). Ces valeurs sont obtenues pour une température de $10 \text{ }^\circ\text{C}$ et une pression partielle en CO_2 de 0.01 atmosphère (Palmer, 2003).

développement du réseau, ce qui réduit le temps de trajet de l’eau et donc sa saturation en carbonates. Quand l’eau atteint une *sortie*, le flux s’accélère grandement, le temps pour atteindre ce stade étant appelé temps de percée (*breakthrough time*). Les conduits de grande taille offrent une grande surface de contact et sont associés à des flux hydrauliques plus importants. Cette augmentation des flux et de la surface soluble couplée à une eau moins chargée en carbonates augmentent la vitesse de dissolution (Figure 1.7), menant à une rétroaction positive. Ainsi, les plus gros conduits correspondent aux chemins préférentiels de l’eau dans le réseau et s’élargissent plus rapidement que les autres. La vitesse de dissolution des parois croît jusqu’à atteindre 0.01-0.1 centimètres par an dans les conduits les plus importants. Cela correspond à la limite de la cinétique chimique, cette valeur ne changeant plus même lorsque le débit augmente (Palmer, 1991). D’après Palmer [1991], les conduits moins importants sont plus ou moins abandonnés, ou du moins voient leur vitesse de croissance stagner, et contribuent beaucoup moins au fonctionnement du réseau. Ce processus d’élargissement le long de chemins préférentiels mène à une espèce d’auto-organisation propre aux réseaux karstiques. Les conduits sont organisés pour faciliter les flux selon la direction du gradient hydraulique et les réseaux évoluent conjointement à la dissolution provoquée par ces flux (Huntoon, 1995; Worthington et Ford, 2009). Le cas de la spéléogénèse par fantômisation diffère de celle décrite précédemment car elle a lieu en deux phases. La première consistant uniquement en la dissolution des minéraux solubles sans altération mécanique, les conduits n’apparaissent réellement que lors de la seconde phase, lorsque l’énergie mécanique est suffisante pour évacuer les minéraux résiduels.

Bien que les réseaux puissent se développer uniquement grâce à la perméabilité des roches, certains facteurs géologiques peuvent grandement impacter ce processus. Palmer [1991] estime, à partir des réseaux utilisés lors de son étude, que la formation

des conduits est contrôlée à 57% par les plans de stratification, à 42% par les fractures et à 1% par la porosité, même si les proportions exactes restent discutables. Les failles et fractures facilitant le transit de l'eau dans la roche, les zones très fracturées sont soumises à des processus de karstification accélérés (Romanov et al., 2002). Les plus larges ouvertures et les points d'infiltration initiaux influencent plus la cinétique de formation du réseau que le gradient hydraulique le plus important et le chemin d'écoulement le plus court, surtout dans les premières phases (Dreybrodt, 1990; Audra et Palmer, 2013). De manière générale, les conduits se forment préférentiellement au niveau d'horizons appelés horizons d'inception (Lowe, 1992; Knez, 1998; Filipponi et al., 2010). Ils sont définis dans les roches carbonatées comme étant "n'importe quel élément lithostratigraphique d'une séquence favorisant, de manière active ou passive, le commencement localisé d'un processus de dissolution en raison d'un changement physique, lithologique ou chimique vis-à-vis du faciès carbonaté prédominant dans cette séquence" (Lowe, 1992). Ils sont généralement présents aux limites entre différentes couches géologiques, au niveau de plans ou de joints de stratifications et sont également trouvables dans d'autres types de roches, telles des grès ou des évaporites (Lowe et Gunn, 1997). On distingue trois principaux types d'horizons d'inception (Filipponi, Marco et Jeannin, 2008; Filipponi et al., 2010) (Figure 1.8) : i) le type 1 où le développement initial des conduits s'initie au sein de l'horizon car sa perméabilité est légèrement supérieure à celles des deux couches en contact, ii) le type 2 où le développement initial des conduits s'initie juste au-dessus d'un horizon dont la perméabilité est légèrement inférieure à celle de la couche supérieure et iii) le type 3 où les conduits se forment le long des joints de stratification, à cause d'irrégularités telles des stries ou des brèches. Ces zones facilitent généralement le passage de l'eau lors des premières phases de formation du réseau mais au final, seule une minorité des plans de stratification correspondrait à des horizons d'inception. Filipponi et al. [2009] estiment que probablement moins de 10% d'entre eux seraient des horizons d'inception mais qu'ils guideraient plus de 70% des conduits.

Le climat influence directement la formation des réseaux, car il influe sur les précipitations, les températures et les teneurs en CO_2 et autres acides des sols intervenant lors du processus de karstification. La température a également une grande importance sur le bilan hydrique, via les pourcentages d'évaporation et d'évapotranspiration, sur le chimisme de l'eau (principalement pour la teneur en CO_2 dissous), ainsi que sur l'absence de gel. Leur rôle est moins caractérisé, mais les nombreux types de microorganismes présents dans les réseaux peuvent également avoir un grand impact lors de la spéléogénèse (Northup et Lavoie, 2001; Engel et al., 2004; Boston et al., 2009).

1.2.2.c Impact de l'origine de l'eau

Suivant l'origine et la source d'agressivité des eaux impliquées dans le processus de karstification, on distingue trois principaux types de karsts (Bakalowicz, 1996; Husson, 2014) :

- Les karsts épigènes, formés via dissolution par une eau météorique dont l'acidité provient principalement du CO_2 biogénique, et par écoulement gravitaire entre la zone de recharge et les sorties de l'aquifère karstique (sources).

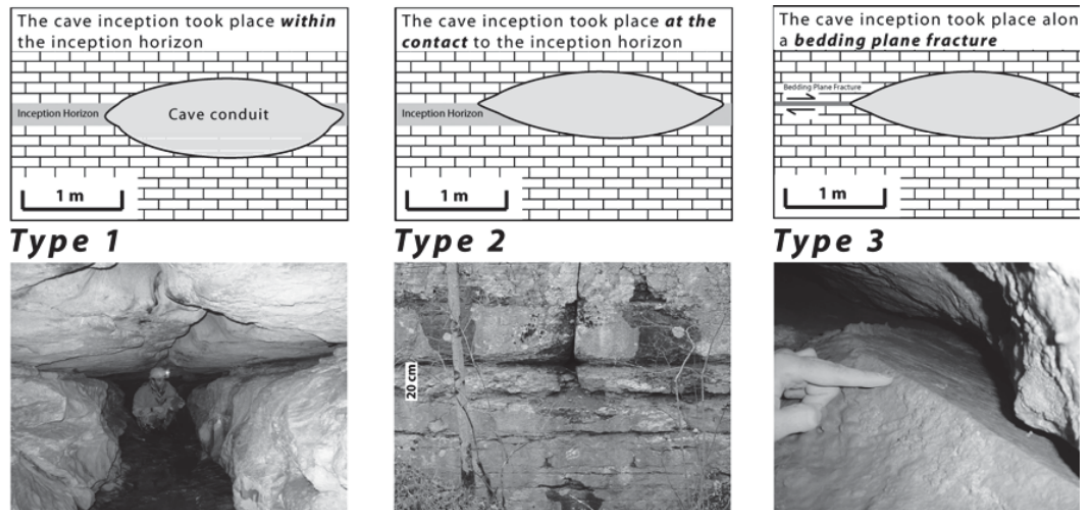


Figure 1.8 – Les trois principaux types d’horizons d’inception : ceux où le développement initial des conduits s’initie au sein de l’horizon, ceux où le développement initial des conduits s’initie au contact de l’horizon et ceux se formant le long des joints de stratification. Les images du haut correspondent aux schémas des différents types tandis que les images du bas sont des photographies de terrain (Filipponi, Marco et Jeannin, 2008).

- Les karsts hypogènes, formés via dissolution par remontée d’eaux connées ou hydrothermales ayant acquis leur acidité en profondeur (majoritairement par du CO_2 et de l’ H_2S).
- Les karsts littoraux, formés via dissolution par le mélange eaux douces/eaux salées et abrasion mécanique par les vagues, les courants et les variations du niveau marin (Fleury, 2005). Les karsts littoraux étant moins étudiés que les karsts épigènes et hypogènes et ne présentant pas d’intérêt particulier pour cette thèse, ils ne seront pas présentés de manière plus détaillée.

Plusieurs types de karstification peuvent être combinés au sein d’un même système, notamment à travers plusieurs phases de spéléogénèse successives (e.g., Figure 1.9) (e.g., Laurent et al., 2021).

Réseaux épigènes

Les réseaux épigènes représentent environ 90% des réseaux connus. Il s’agit de réseaux formés par des eaux d’origine météorique et dont le développement dépend beaucoup du contexte hydrogéologique. Le mode de formation des conduits est fortement influencé par leur position au sein du réseau, vis-à-vis du niveau piézométrique, et par le type de recharge hydraulique (Figure 1.10).

L’eau présente dans un système peut avoir différentes origines suivant le type de dénudation de la partie surfacique du karst (Ford et Williams, 2007) :

- Le système est dit autochtone si la zone de surface est constituée uniquement de roches carbonatées et que l’eau provient uniquement des précipitations s’y infiltrant (Figure 1.11A).

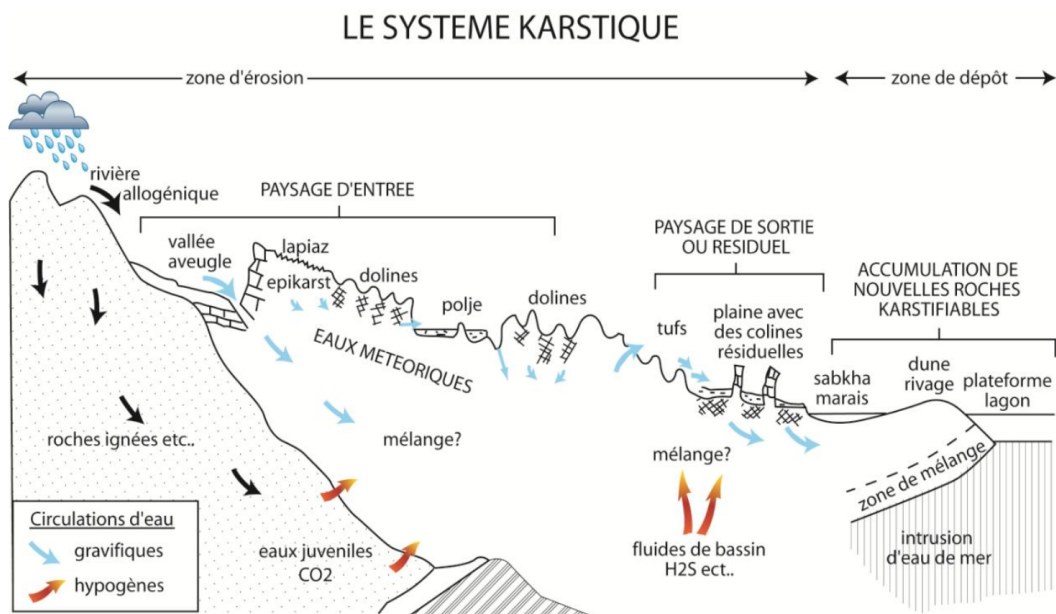


Figure 1.9 – Schéma des principaux phénomènes interagissant au sein d'un système karstique (Husson, 2014, modifié d'après Ford et Williams, 2007).

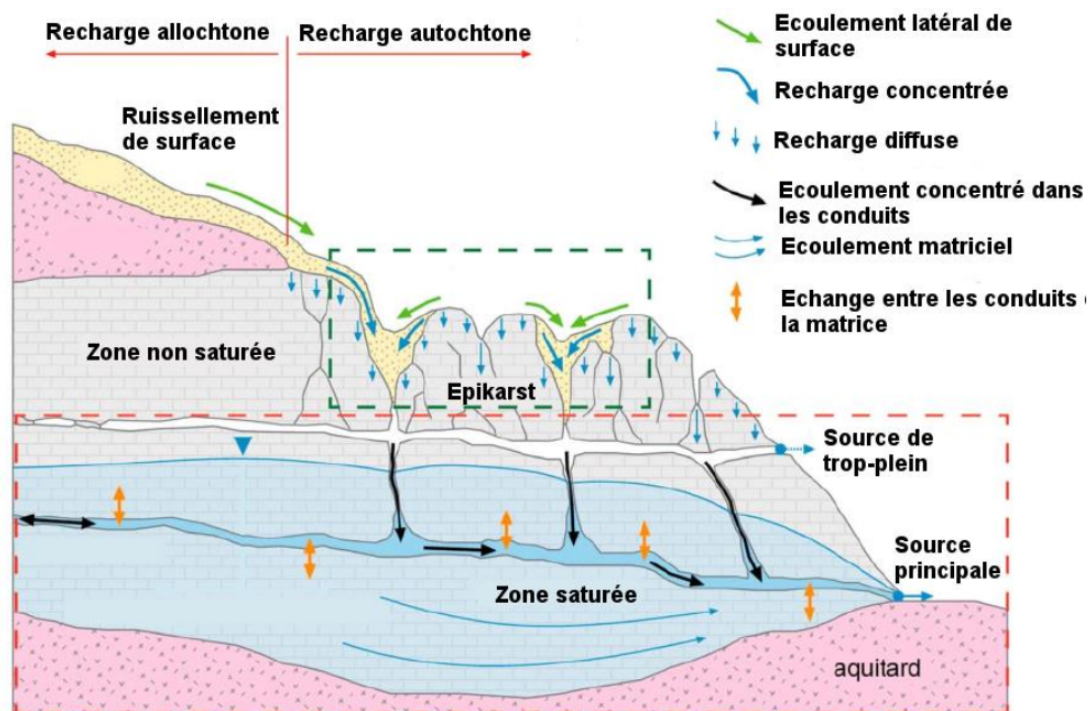


Figure 1.10 – Schéma des écoulements dans les réseaux karstiques en fonction du type de recharge et de la position relative au niveau piézométrique. Les pointillés verts et rouge correspondent respectivement aux compartiments superficiels et profonds de l'aquifère karstique (Lorette, 2019, modifié de Hartmann et al., 2014).

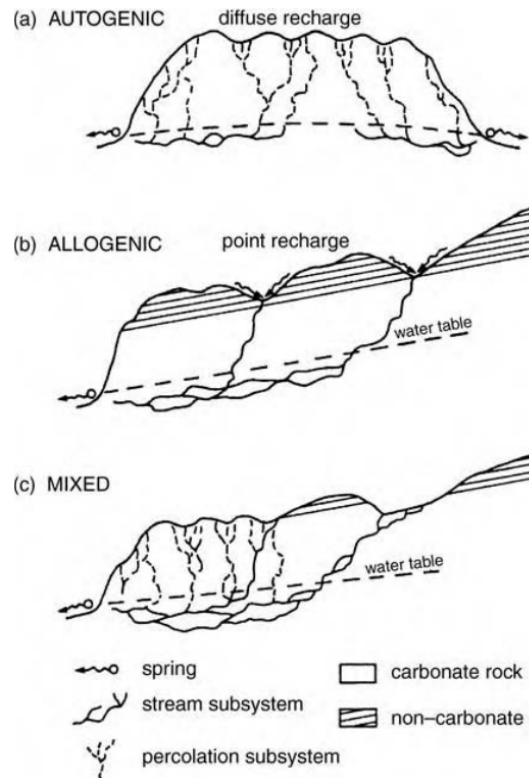


Figure 1.11 – Trois types de dénudation des systèmes karstiques. A : Dénudation autochtone. B : Dénudation allochtone. C : Dénudation mixte (Ford et Williams, 2007).

- Le système est dit allochtone si l'eau provient d'une zone de drainage constituée de roches non karstifiées (Figure 1.11B).
- La majorité des systèmes karstiques est cependant généralement un mélange de parties autochtones et allochtones (Figure 1.11C).

Suivant le couvert ainsi que d'autres paramètres, il y a deux principaux types de recharge hydraulique pour les réseaux (Figure 1.12) :

- La recharge est dite irrégulière quand la roche est directement à nu et que les entrées du réseau sont exposées, ce qui correspond généralement à des systèmes autochtones. L'eau s'infiltré directement dans le réseau et l'altitude du niveau piézométrique dépend beaucoup des conditions météorologiques, provoquant la formation de conduits d'altitude variant fortement localement (Figure 1.12 A).
- La recharge est dite régulière lorsqu'il y a une couverture sur la roche, notamment un sol, mais cela ne correspond pas forcément à un système allochtone. L'eau s'infiltré assez lentement mais continuellement tout au long de l'année, ce qui mène à de faibles variations du niveau piézométrique sur de courtes périodes temporelles. Les conduits ont donc plutôt tendance à suivre le niveau piézométrique dans la partie basse du réseau, sans présenter d'importantes variations d'altitude (Figure 1.12 B).

D'une région du monde à une autre, l'organisation interne des réseaux épigènes reste relativement similaire. Les structures dépendent principalement de la position

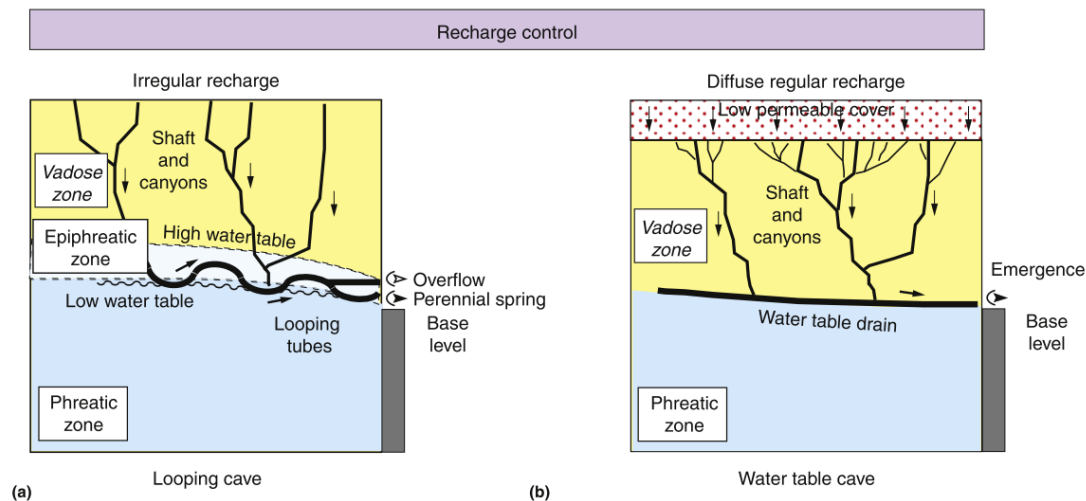


Figure 1.12 – A : Exemple de recharge irrégulière conduisant à des changements d'altitude du niveau piézométrique. Les conduits ont tendance à former des boucles verticales calées sur les hétérogénéités de la roche dans la partie épiphréatique. B : Exemple de recharge régulière, l'altitude du niveau piézométrique reste constante. Les conduits ont tendance à se former le long du niveau piézométrique (Audra et Palmer, 2013).

relative des conduits par rapport au niveau piézométrique et à ses changements durant leur formation. On peut diviser les karsts épigènes en quatre grandes parties, chacune associée à des morphologies de conduits différentes :

- L'épikarst correspond à la zone superficielle de la partie souterraine (Figure 1.13). Suivant les définitions, le sol peut être inclus ou non dans l'épikarst s'il est présent au-dessus des roches karstifiées (Klimchouk, 2004). L'épikarst est la principale zone d'infiltration d'eau dans le réseau et sert de manière générale comme interface entre la roche et l'atmosphère ainsi que la biosphère. Son épaisseur va habituellement de l'ordre du mètre à la quinzaine de mètres (Klimchouk, 2004). L'épikarst est très hétérogène et la porosité, ainsi que la perméabilité de la roche, baissent avec la profondeur, les zones profondes étant moins altérées par les infiltrations d'eau. Ces différences peuvent conduire à l'existence de zones temporairement saturées en eau, qualifiées de nappes perchées (Robineau, 2019). L'eau stockée dans cette zone peut soit s'infiltrer dans le réseau, soit s'écouler le long des discontinuités. La composante horizontale des écoulements dans cette zone peut ainsi être importante (Mangin, 1975).
- La zone vadose correspond à la partie située sous l'épikarst, au-dessus du niveau piézométrique et n'étant pas saturée en eau. Passé l'épikarst, l'eau transite dans la zone vadose jusqu'à atteindre le niveau piézométrique. Les écoulements peuvent être de nature très variable : il peut s'agir d'écoulements lents à travers la matrice ou d'écoulements rapides dans des conduits généralement sub-verticaux. Cela dépend principalement de l'avancement de la karstification. Plus cette dernière est avancée, plus les écoulements en conduits sont favorisés, dépassant à terme les 90% (e.g., Worthington et al., 2000). Si les temps de transit sont suffisam-

ment importants, il est possible d'atteindre un équilibre entre le CO_2 dissous et celui présent dans l'air. La pression partielle en CO_2 de l'eau va ainsi évoluer conjointement à celle de l'air, notamment lors des changements des conditions météorologiques (Peyraube et al., 2012; Peyraube et al., 2013; Houillon, 2016; Sivelles, 2019).

- La zone phréatique correspond à la zone saturée en eau située sous le niveau piézométrique. La majeure partie de l'eau est stockée dans la matrice et est associée à des écoulements lents suivant le gradient hydraulique. La majorité des écoulements est cependant rapide et a lieu dans les conduits. Les conduits formés dans cette zone sont souvent sub-horizontaux et relativement circulaires.
- La zone épiphréatique est située à l'interface de la partie vadose et de la partie phréatique. Suivant le type de recharge, les conditions météorologiques et d'autres facteurs, la quantité d'eau s'infiltrant dans le réseau varie, ce qui impacte l'altitude du niveau piézométrique. Cela a pour conséquence une zone parfois saturée, parfois non, conduisant à d'importantes variations locales de l'altitude des conduits. Les vitesses d'écoulements associées à ce mode de formation sont généralement assez élevées. Dans certains cas, lorsque les écoulements sont rapides, la zone épiphréatique peut atteindre jusqu'à plusieurs centaines de mètres de hauteur (e.g., Häuselmann et al., 2003).

L'environnement hydrogéologique contraint ces différentes zones, notamment au niveau des chemins d'écoulement, et cela influe sur les réseaux formés. Il est généralement assez facile d'identifier si une partie d'un réseau a été formée en zone vadose ou phréatique (Palmer, 1991). Les réseaux sont souvent auto-organisés, leur morphologie dépendant des flux hydraulique : les conduits ont souvent tendance à suivre la direction du gradient hydraulique dans leur zone de formation (Huntoon, 1995; Worthington et Ford, 2009).

En géomorphologie, le niveau de base correspond à la surface en-dessous de laquelle il n'y a plus d'érosion. L'érosion s'opérant à l'altitude du niveau de base est purement horizontale, contrairement à celle s'effectuant au-dessus qui présente habituellement une grande composante verticale. Il est possible de définir des niveaux de base locaux mais le niveau de base final de tout système hydrologique est toujours l'altitude de la mer. Le niveau piézométrique désignant l'altitude locale de la nappe phréatique, il correspond plus à un *niveau local actuel* alors que le niveau de base serait le *niveau à atteindre*. Les réseaux se formant sur de longues périodes temporelles, l'altitude du niveau de base reste rarement constante, ce qui fait varier l'altitude du niveau piézométrique et influe directement sur celle des conduits, ces variations pouvant être importantes (Figure 1.14). Elle peut en effet changer pour différentes raisons : creusement ou remplissage d'une vallée, changement du niveau de la mer, causes tectoniques... Lors des baisses du niveau de base, les conduits initialement présents dans la zone phréatique se retrouvent localisés dans la zone vadose. Les conduits peuvent être abandonnés au profit de nouveaux conduits descendants. Le gradient gravitaire étant plus important, l'eau a également tendance à creuser le fond des conduits préexistants de sorte que leur section finale soit en forme de trou de serrure (Audra et Palmer, 2013). A l'inverse, une augmentation du niveau base se traduit par un passage du milieu vadose au milieu

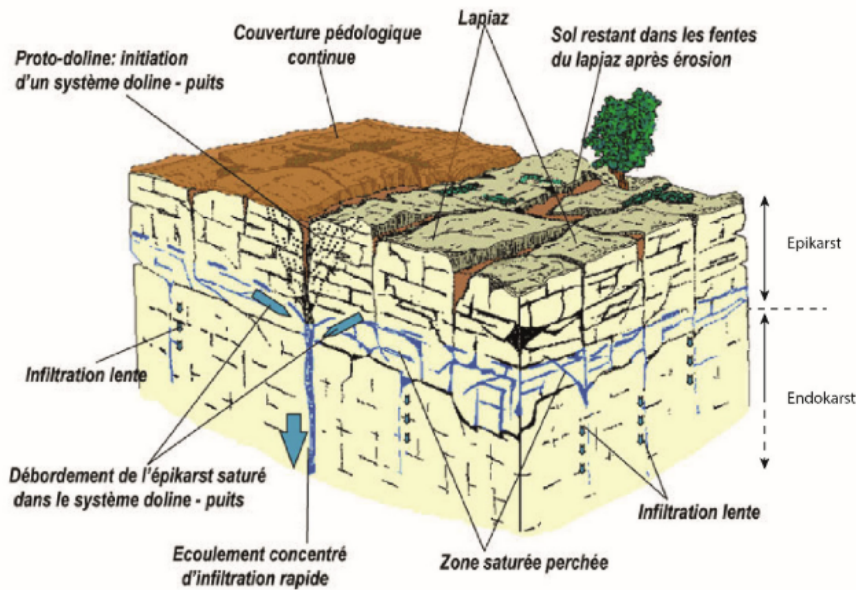


Figure 1.13 – Schéma de la structure de l'épikarst (Jouves, 2018, modifié depuis Bakalowicz, 2013 et Mangin, 1975).

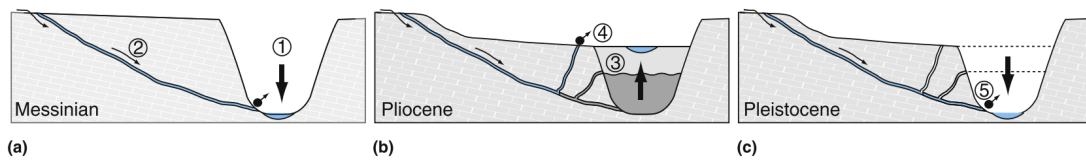


Figure 1.14 – Modèle de spéléogénèse *per ascensum*. A : Creusement du canyon par l'eau (1), le conduit rejoint le niveau de base (2). B : Comblement du canyon lors de la montée du niveau de base en deux étapes, menant chacune à la formation d'un conduit ascendant (3 et 4). C : Exhumation du canyon lors de la baisse du niveau de base. Les conduits 3 et 4 sont abandonnés (Audra et Palmer, 2013).

phréatique, ce qui peut provoquer la formation de nouveaux chemins ascendants via une spéléogénèse *per ascensum* (Farrant et Smart, 2011; Audra et Palmer, 2013; Audra et Palmer, 2015) (Figure 1.14). Ce phénomène mène à la formation de sources dites vauclusiennes, qui correspondent à des résurgences d'eau souterraine. L'alternance de ces montées et de ces baisses du niveau piézométrique peut conduire à des morphologies de réseau relativement complexes.

Réseaux hypogènes

Les réseaux hypogènes représentent moins de 10% des réseaux connus, mais comme ils ne sont pas toujours accessibles via la surface, leur proportion réelle est probablement plus importante (Palmer, 1991). Bien que ces réseaux soient connus depuis longtemps (Martel, 1935), tout comme le rôle de l'acide sulfurique (Socquet, 1803), ils furent longtemps considérés comme des cas *exotiques* comparés aux réseaux épigènes (Audra et al., 2009a). Les modèles de formation, ainsi que la caractérisation des principaux attributs des réseaux hypogènes sont cependant relativement récents (Klimchouk, 2007; Audra et al., 2009a).

Contrairement aux réseaux épigènes, les processus de spéléogenèse associés à la formation des réseaux hypogènes sont liés à l'action de fluides d'origine profonde. Ces fluides remontent vers la surface, générant des conduits de bas en haut. On parle alors de spéléogenèse *per ascensum* (à ne pas confondre avec la spéléogenèse *per ascensum* épigène présentée précédemment). Les principaux fluides impliqués sont les suivants :

- Des eaux connées, qui sont des eaux initialement présentes dans les couches géologiques lors de leur enfouissement, qui se sont trouvées *piégées* à l'intérieur et qui ont été libérées par compaction/déshydratation des roches (Jouves, 2018).
- Des eaux d'origine météorique infiltrées en profondeur et dont les longs temps de transit ainsi que le passage à travers différentes couches géologiques leur ont permis de dissoudre des minéraux sur leur passage (souvent présents dans des roches peu solubles) et d'acquérir une forte agressivité.
- Des mélanges de plusieurs types d'eau, ce qui conduit généralement à une forte augmentation de leur agressivité.

Ces eaux ont généralement des concentrations en sulfates, sulfures, métaux mais aussi en CO_2 plus importantes que les eaux météoriques. La dissolution des minéraux sulfatés et sulfurés, comme la pyrite, aboutit souvent à la formation d'acide sulfurique, qui est un des principaux agents chimiques impliqués dans le processus de formation des réseaux hypogènes.

Les eaux d'origine profondes seaturent généralement vite en carbonates, empêchant de poursuivre la dissolution de manière efficace mais cependant, trois phénomènes majeurs, principalement provoqués par le mélange d'eaux, leur permettent de renouveler leur agressivité (Palmer, 1991) :

- L'agressivité de l'eau peut être renouvelée par le mélange d'eaux ayant des compositions chimiques différentes, notamment par rapport à leurs concentrations en gaz dissous, principalement le sulfure d'hydrogène (H_2S) mais aussi le dioxyde de carbone (CO_2). La Figure 1.15 montre les valeurs de saturation de la calcite et de la dolomite en fonction de la concentration en CO_2 ou en H_2S à l'équilibre pour une température de 25 degrés Celsius (Palmer, 1991). Le mélange de deux solutions saturées peut conduire à l'obtention d'une solution sous-saturée vis-à-vis de la calcite ou de la dolomite.
- L'oxydation des sulfures, surtout celle du sulfure d'hydrogène (H_2S) en acide sulfurique (H_2SO_4) se produit généralement lorsque des fluides profonds se mélangent aux eaux météoriques, au niveau de la nappe. L'agressivité de l'eau augmente beaucoup, l'acide sulfurique pouvant dissoudre des quantités de calcite plus importantes que le sulfure d'hydrogène. Si la concentration en sulfates est assez importante, les parois peuvent être remplacées par du gypse (e.g., Temovski et al., 2013; Laurent et al., 2021).
- Un refroidissement des fluides survient généralement lors de leur remontée et de leur trajet dans les roches carbonatées. Comme la solubilité de la calcite dans l'eau est inversement proportionnelle à la température, cette baisse de température augmente l'agressivité de l'eau. Cela produit souvent des cavités relativement

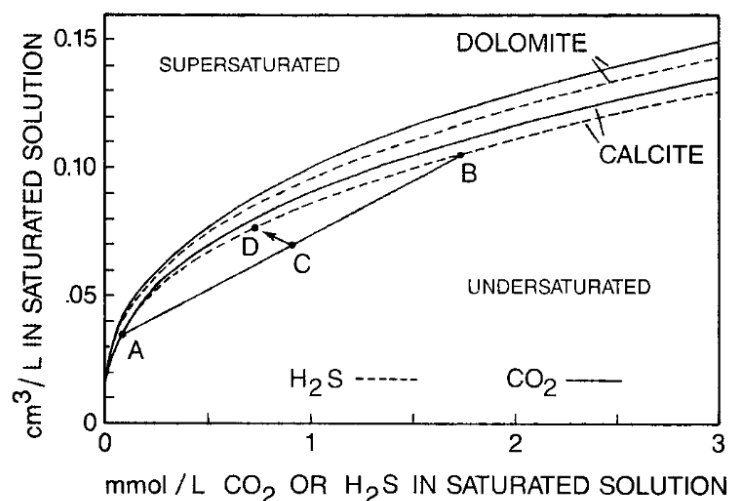


Figure 1.15 – Concentration de saturation de la calcite et de la dolomite en fonction de la concentration en CO_2 ou en H_2S à 25 degrés Celsius, en volume dissout par litre. Le mélange de deux solutions saturées (comme A et B) produit une solution sous-saturée (C). La dissolution suit alors la ligne C-D (Palmer, 1991).

petites (Palmer, 1991). Le mélange d'eaux de températures différentes peut, là aussi, renouveler leur agressivité.

En plus de ces processus physicochimiques, certaines études semblent indiquer une importante participation de certains microorganismes à la spéléogénèse hypogène, notamment dans la transformation de l' H_2S en acide sulfurique (Engel et al., 2004).

Les réseaux hypogènes sont souvent reconnaissables par leur manque de lien avec la surface et entre leurs différents niveaux. Les zones de discontinuité les plus favorables tendent à s'élargir au même rythme, conduisant, dans la majeure partie des cas, à l'absence de chemin préférentiels. Il est possible de reconnaître les spéléogénèses hypogènes à plusieurs niveaux. A petite échelle, on observe de nombreux éléments caractéristiques produits par corrosion et déposition (Audra et al., 2009b). Les processus impliqués dans la spéléogénèse hypogène dépendent principalement de la position de la zone considérée par rapport au niveau piézométrique (Figure 1.16) :

- La zone phréatique se traduit par la présence de chemins ascendants, principalement formés via corrosion lors de la remontée des eaux profondes. Les flux ayant lieu alternativement le long de plans de stratification et de fractures, cela mène à des profils semblables à des escaliers (Audra et al., 2009c). La majorité de la dissolution s'effectue généralement le long du chemin emprunté lors de la remontée des fluides, tandis que des zones de dissolution plus petites sont observables autour de ce chemin. De nombreux marqueurs des flux ascendants sont également observables, comme des trainées de bulles (Audra et al., 2009a; Audra, 2017). Lorsque le flux ascendant rencontre un horizon de perméabilité plus faible, les conduits se développent directement sous celui-ci et suivent son pendage (Audra, 2017). Dans des conditions très favorables à la dissolution, il arrive que des conduits de très grande taille se forment lors de cette remontée, leur agrandissement se faisant souvent conjointement à des effondrements de surface (Audra, 2017).

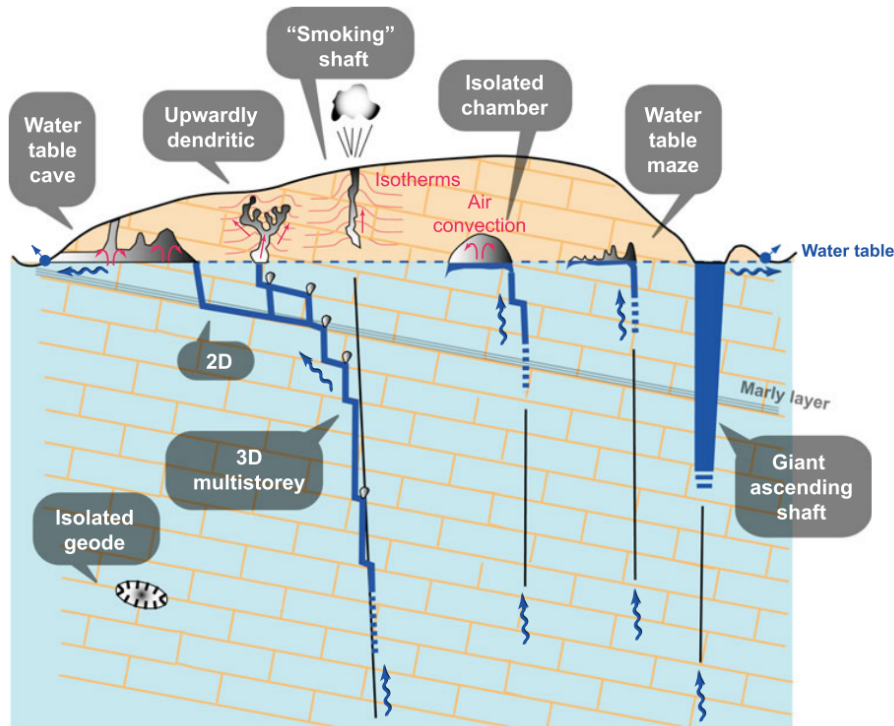


Figure 1.16 – Schéma conceptuel des principaux types de réseaux hypogènes, en fonction des structures géologiques et du type de flux (Audra, 2017, modifié d’après Audra et al., 2007 et Audra et al., 2009c).

- L’altitude du niveau piézométrique se traduit souvent par des encoches de dissolution. Ses variations sont donc relativement facilement observables (Audra, 2017). Lorsque le niveau de l’eau ne subit que peu de variations sur de longues périodes de temps, l’importante dissolution à sa hauteur forme de grandes cavités très planes.
- Au-dessus du niveau piézométrique, la dissolution est principalement associée à des phénomènes de condensation-corrosion (Sarbu et Lascu, 1997; Audra et al., 2009a; Audra, 2017). Les eaux sous-jacentes s’évaporent partiellement dans la partie non submergée. Elles peuvent dégazer du CO_2 et de l’ H_2S , créant une atmosphère très humide. Ces gaz chauds montent vers les parties supérieures des réseaux par convection. L’eau se condense sur les parois à cause de la différence de température. Plus cette dernière est importante, plus la condensation-corrosion est intense. Ainsi, les parties des réseaux les plus impactées par ce phénomène sont généralement celles proches de la surface et celles relativement bien isolées du reste du réseau. L’eau présente sur les parois peut ensuite atteindre un équilibre avec le CO_2 et l’ H_2S présents dans l’atmosphère. Elle devient alors très agressive, conduisant à la corrosion des parois, principalement si les temps de contact sont suffisamment longs, et à la formation de coupoles caractéristiques.

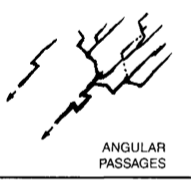
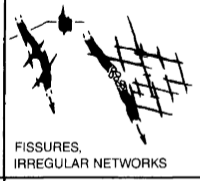
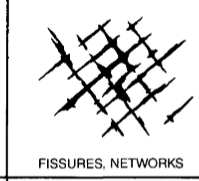
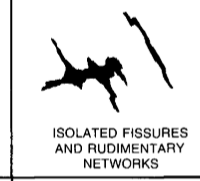
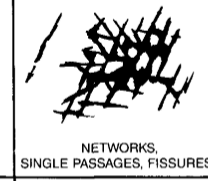
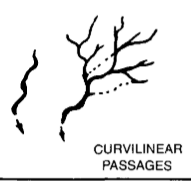
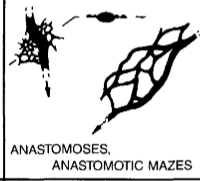
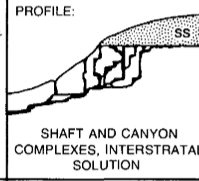
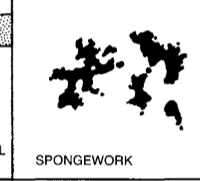
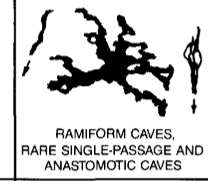
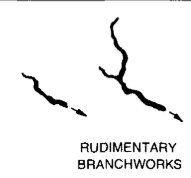

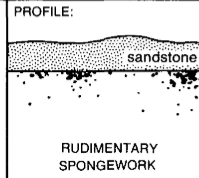
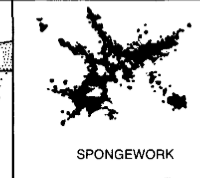
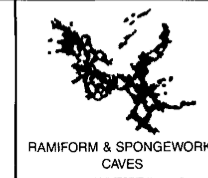
		TYPE OF RECHARGE				
		VIA KARST DEPRESSIONS		DIFFUSE		HYPOGENIC
		SINKHOLES (LIMITED DISCHARGE FLUCTUATION)	SINKING STREAMS (GREAT DISCHARGE FLUCTUATION)	THROUGH SANDSTONE	INTO POROUS SOLUBLE ROCK	DISSOLUTION BY ACIDS OF DEEP-SEATED SOURCE OR BY COOLING OF THERMAL WATER
		BRANCHWORKS (USUALLY SEVERAL LEVELS) & SINGLE PASSAGES	SINGLE PASSAGES AND CRUDE BRANCHWORKS, USUALLY WITH THE FOLLOWING FEATURES SUPERIMPOSED:	MOST CAVES ENLARGED FURTHER BY RECHARGE FROM OTHER SOURCES	MOST CAVES FORMED BY MIXING AT DEPTH	
DOMINANT TYPE OF POROSITY	FRACTURES	 ANGULAR PASSAGES	 FISSURES, IRREGULAR NETWORKS	 FISSURES, NETWORKS	 ISOLATED FISSURES AND RUDIMENTARY NETWORKS	 NETWORKS, SINGLE PASSAGES, FISSURES
	BEDDING PARTINGS	 CURVILINEAR PASSAGES	 ANASTOMOSES, ANASTOMOTIC MAZES	PROFILE:  SHAFT AND CANYON COMPLEXES, INTERSTRATAL SOLUTION	 SPONGEWORK	 RAMIFORM CAVES, RARE SINGLE-PASSAGE AND ANASTOMOTIC CAVES
	INTERGRANULAR	 RUDIMENTARY BRANCHWORKS	 SPONGEWORK	PROFILE:  RUDIMENTARY SPONGEWORK	 SPONGEWORK	 RAMIFORM & SPONGEWORK CAVES

Figure 1.17 – Mise en relation de différents motifs morphologiques avec le type de recharge et le principal type de porosité lors du processus de spéléogénèse. Les réseaux sont en vue du dessus sauf quand précisé (Palmer, 1991).

1.2.3 Classification via motifs morphologiques

Comme vu précédemment, le type de spéléogénèse influe très fortement sur la morphologie des réseaux. En fonction du type de recharge, de la position par rapport au niveau piézométrique, de son évolution au cours du temps et d'autres paramètres (e.g., perméabilité des roches, lithologies), différents modes de spéléogénèse ont été identifiés. Ces modes de spéléogénèse influencent la morphologie des réseaux, que cela soit au niveau de leur topologie ou de leur géométrie. La topologie est une branche des mathématiques étudiant les propriétés des ensembles invariantes par transformations continues, notamment les relations de voisinage. Dans le cas des réseaux karstiques, elle peut être définie comme la façon dont les conduits sont connectés. La géométrie est quant à elle définie comme les dimensions de la section locale des conduits en chaque point du réseau. Leurs variations permettent de diviser et classifier les différents types de morphologies sous différents noms, également appelés motifs. Palmer [1991] proposa la première classification de ces morphologies (Figure 1.17). D'autres études ont permis d'améliorer la caractérisation de ces différents motifs (e.g., Audra et Palmer, 2013; Jouvès et al., 2017). Les réseaux étant généralement associés à différentes phases de spéléogénèse, que cela soit de façon synchrone ou asynchrone, plusieurs motifs morphologiques peuvent être présents au sein d'un même réseau (Figure 1.18) (Audra et Palmer, 2013; Jouvès et al., 2017). Nous présenterons ici les plus communs.

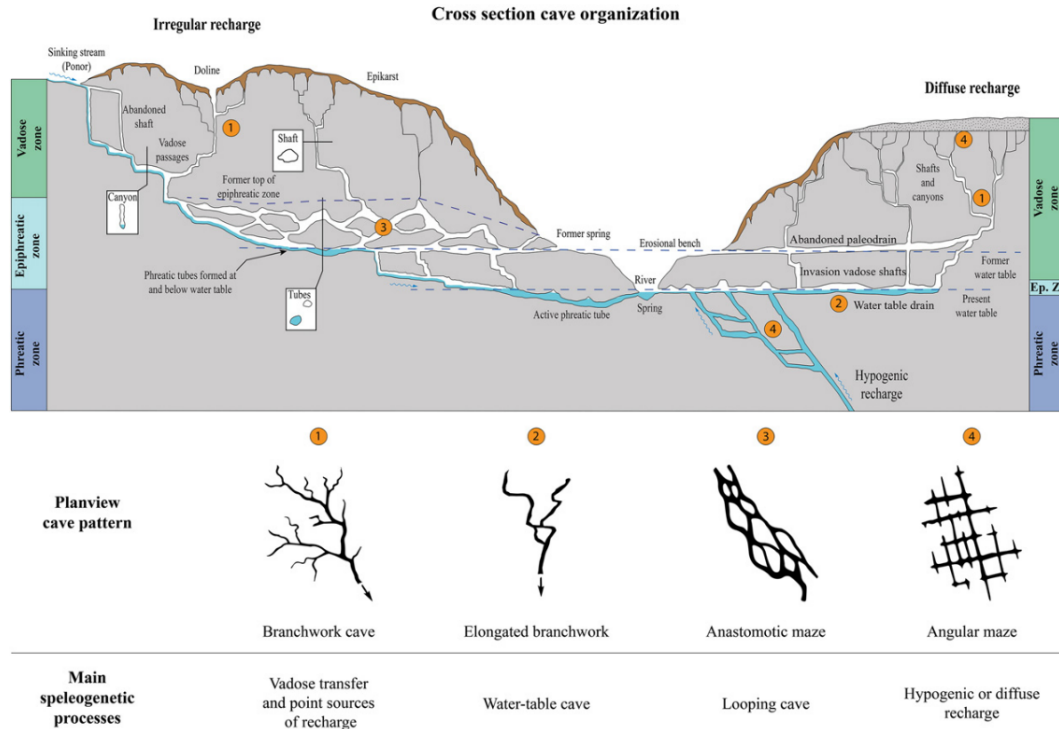


Figure 1.18 – Coupe transversale idéalisée d'un système karstique avec zonation spatiale verticale du karst et vue en plan des motifs associés (Jouves et al., 2017).

Les motifs *Vadose Branchwork* ou arborescents se développent dans les zones vadoses du réseau, principalement quand la recharge associée est irrégulière. Ils correspondent à de nombreux conduits sub-verticaux, provenant généralement d'écoulements gravitaires issus de différents points de recharge (notamment des dolines), qui convergent progressivement, les conduits en aval étant souvent théorisés comme étant plus grands que ceux en amont (Palmer, 1991).

Les motifs *Water Table Cave* ou motifs arborescents allongés s'observent principalement à l'emplacement d'un niveau piézométrique présent ou passé. Comme le niveau d'eau a tendance à rester constant à cause de la recharge régulière, cela forme de grands réseaux de conduits sub-horizontaux. Les faibles variations du niveau piézométrique limitent fortement le développement de la zone épiphréatique, conduisant à de très faibles variations d'altitude des conduits. Les phénomènes de dissolution s'effectuant au niveau piézométrique ou en-dessous ont tendance à former des conduits aux sections très circulaires (Audra et Palmer, 2013).

Les motifs *Looping Cave* ou motifs en labyrinthe anastomosé sont formés dans les parties épiphréatiques des réseaux, principalement le long de plans de stratifications ou de fractures, quand la recharge est irrégulière. À cause des fluctuations du niveau piézométrique, les conduits forment de nombreuses boucles verticales ayant généralement une faible pente en amont et une forte pente en aval (Jouves, 2018).

Les motifs *Angular Maze* ou motifs en labyrinthe angulaire sont habituellement associés à des réseaux de fractures au sein des roches karstifiées. Ils s'observent principalement dans des réseaux épigènes associés à une recharge diffuse, dans des réseaux formés par fantômisations ou dans des réseaux hypogènes (Audra et Palmer, 2013). Les

conduits sont organisés en grands réseaux interconnectés angulairement selon la fracturation naturelle du massif, formant une espèce de maillage. Un contact direct entre un glacier et une roche carbonatée peut également être à l'origine de ce genre de motif (Skoglund et al., 2010).

Les motifs *Spongework* ou en éponge sont associés à une importante porosité intergranulaire (Palmer, 1991). Ils se forment par la croissance et la coalescence des pores intergranulaires ainsi que d'autres petites ouvertures, lorsqu'il n'y a pas de fractures majeures ou autres horizons d'inception. Cela mène à des écoulements hydrauliques chaotiques, provoquant la formation de réseaux complexes. Ces réseaux sont généralement associés à une origine hypogène ou un mélange d'eau douce et d'eau salée (Audra et Palmer, 2013).

1.2.4 L'exploration des réseaux

Une bonne caractérisation des réseaux est importante pour optimiser l'exploitation des ressources associées ainsi que pour la gestion de leur vulnérabilité. Les observations de surface peuvent apporter des informations intéressantes. L'observation du paysage permet en effet de repérer les traces indiquant l'existence de phénomènes de karstification, ainsi que les entrées et sorties hydrologiques des réseaux. Des photographies aériennes et satellites peuvent également être utilisées lors de certaines études (Kresic, 1995; Tam et al., 2004). L'observation du paysage ne permet cependant pas d'obtenir des informations précises sur les réseaux souterrains. De nombreuses méthodes ont de fait été développées pour permettre leur étude.

1.2.4.a Mesures directes par des spéléologues

Les réseaux karstiques sont des objets géologiques particuliers qui, contrairement à beaucoup d'autres, nécessitent habituellement une exploration souterraine par des humains pour pouvoir être caractérisés de manière précise (Kambesis, 2007). L'acquisition de données par les spéléologues lors de l'exploration des réseaux est complexe. Dans le but de cartographier certains réseaux, ils peuvent mesurer la largeur et la hauteur des conduits en différents points appelés stations (Figure 1.19), en même temps que la distance, l'orientation et le pendage entre stations voisines. Il arrive que les spéléologues prennent note d'informations tectoniques telles que les failles mais c'est rarement le cas des informations pétrographiques et stratigraphiques (e.g., Knez, 1998). Il est possible d'établir la position des points en coordonnées euclidiennes et de reconstruire les réseaux en 3 dimensions à partir des mesures effectuées. Les dimensions des conduits sont généralement estimées visuellement et ne sont pas toujours représentatives de leur morphologie, même si les récentes technologies laser, notamment le LIDAR, laissent espérer des données plus précises à l'avenir (e.g., Sadier, 2013; Schiller et Renard, 2016). Les spéléologues ne peuvent cependant faire l'acquisition de données que dans les parties accessibles des réseaux, ce qui empêche de les explorer complètement. Seuls les conduits à taille humaine sont réellement explorables et un rétrécissement très localisé au sein d'un conduit peut empêcher la caractérisation d'une grande zone d'un réseau. Bien que cela puisse être très dangereux, il est pour autant possible de prendre des mesures dans des conduits submergés. L'acquisition des données par les spéléologues ainsi que leur traitement seront présentés plus en détail dans le Chapitre 2.

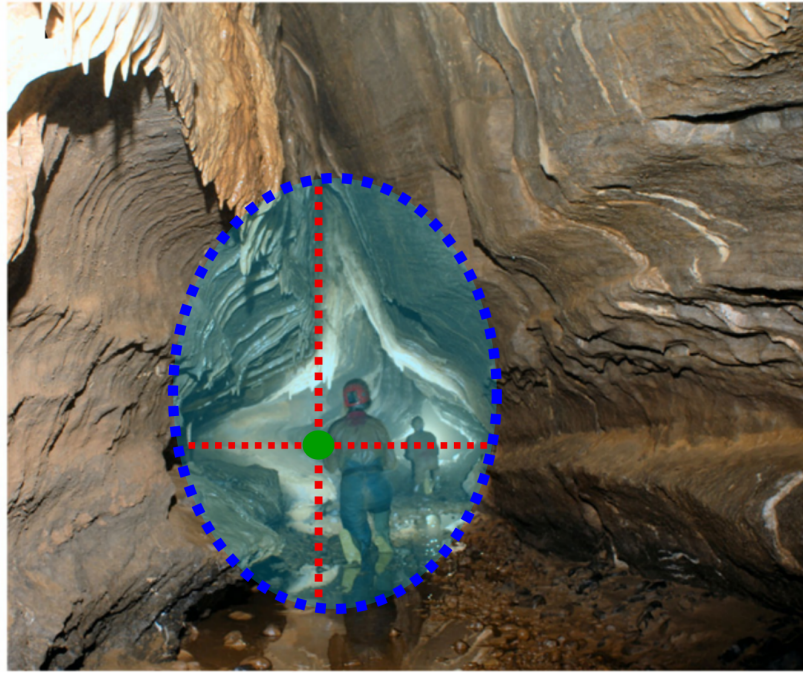


Figure 1.19 – Exemple de mesure de la largeur et de la hauteur d'un conduit (rouge) au niveau d'une station (vert). L'ellipse bleue correspond à la section approximative obtenue de cette manière mais elle ne respecte pas parfaitement la morphologie du conduit et n'est pas centrée sur le conduit (modifié d'après Farrant et Smart, 2011, photographie par Peter Hann).

1.2.4.b Étude du fonctionnement hydrologique, essais de traçage et analyse de l'eau

L'étude du fonctionnement hydrologique des karsts est très importante pour la bonne gestion des ressources en eau présente en leur sein. Les écoulements étant directement dépendants de leur connectivité et de leur morphologie, leur étude permet également d'améliorer la connaissance des réseaux. Certaines méthodes, notamment de modélisation inverse, permettent d'obtenir des informations sur les réseaux (e.g., connectivité, rayon des conduits) en se servant de données hydrologiques connues comme les débits (e.g., Pardo-Iguzquiza et al., 2011; Borghi et al., 2016; Tinetti et al., 2019). Les informations obtenues sur la géométrie des conduits sont cependant peu précises et dépendent des nombreuses hypothèses sous-jacentes aux modèles.

Les essais de traçage consistent à marquer de l'eau avec un traceur, notamment de la fluorescéine ($C_{20}H_{12}O_5$), pour suivre ses déplacements. Les essais de traçages peuvent être utilisés pour des eaux de surface, telles les rivières, mais également pour des eaux souterraines. Le traceur est injecté en un point de manière ponctuelle ou continue, et sa concentration est suivie à plusieurs endroits. Les essais de traçage permettent notamment de mettre en avant l'existence d'une connexion hydraulique entre deux points précis et de caractériser le temps de trajet de l'eau entre ces deux points (e.g., Dewaide, 2018; Tinetti et al., 2019). L'analyse des temps de transfert et des quantités de traceur recueillies peut mettre en évidence l'existence de connexions complexes.

L'analyse hydrologique des réseaux ne se limite pas aux seuls essais de traçage. L'analyse des débits aux entrées et sorties connues du réseau, couplée à l'analyse des précipitations permet également de caractériser la réponse hydrologique (e.g., temps, variations du niveau d'eau) et de calculer le budget hydrologique du réseau. L'analyse des hydrogrammes, via l'analyse des courbes de récession lors des crues permet par exemple d'obtenir des informations sur le fonctionnement de la zone vadose du réseau et d'évaluer la quantité stockée dans la zone phréatique (Mangin, 1975).

Le couplage d'hydrogrammes (variation des quantités d'eau) et de chémogrammes (variation de la composition de l'eau) en sortie de réseaux avec des diagrammes climatiques permet de réaliser des suivis poussés de la qualité de l'eau sur de longues périodes. D'autres paramètres de l'eau peuvent également être analysés, comme son pH, sa température, sa turbidité ou encore sa conductivité électrique. Les avancées technologiques ont permis la création de stations de mesures in situ permettant d'analyser en continu toutes ces propriétés et de les communiquer en temps réel (e.g., Lorette, 2019).

1.2.4.c Méthodes géophysiques

Les cavités souterraines ayant d'importants contrastes de propriétés avec la roche encaissante, cela permet d'envisager l'utilisation de méthodes géophysiques pour les détecter. En plus d'être utilisables pour la caractérisation de réseaux de surface, les méthodes géophysiques sont particulièrement utiles dans le cas où les réseaux sont inaccessibles (e.g., paléokarsts).

Les cavités remplies d'air ont une forte résistivité électrique, une faible densité et les ondes élastiques les traversant ont une faible vitesse. Comparées à elles, celles remplies d'eau ou de sédiments sont associées à une résistivité électrique plus faible de deux ou trois ordres de magnitude (Vyzhva et al., 2020). De manière générale, les milieux très hétérogènes présentant de nombreuses fissures et vides sont caractérisés par des valeurs de densité et de vitesse des ondes élastiques relativement faibles.

L'utilisation des méthodes géophysiques dépend de nombreux critères, notamment : i) le but de l'étude (prospection, exploration, suivi) et son échelle, ii) la taille, la profondeur et la morphologie des cavités, iii) la nature de la roche encaissante, ses hétérogénéités et les propriétés géologiques de la région (plateau ou zone plissée), iv) la présence d'eau, d'air, d'un mélange des deux ou d'autre chose dans la matrice ainsi que dans les cavités et v) la composition de l'eau remplissant les cavités (Martínez-Moreno et al., 2014; Vyzhva et al., 2020).

Parmi les nombreuses méthodes géophysiques, voici certaines des plus importantes :

- Les méthodes de profilage électrique sont principalement basées sur le calcul de la résistivité électrique et sont surtout utiles pour la subsurface. Elles consistent à faire passer un courant électrique dans le sol, entre plusieurs électrodes, et à analyser les différences mesurées à la surface en différents points (les capteurs n'étant pas forcément localisés au niveau des électrodes). La présence d'air dans les cavités et fissures a tendance à augmenter la résistivité électrique tandis que l'eau l'abaisse. L'utilisation de ces méthodes peut permettre d'identifier la limite entre roche karstifiée et roche non karstifiée en profondeur, sur des zones étendues. Comme l'acquisition s'effectue entre plusieurs points, il est souvent nécessaire de

définir une grille de mesure pour pouvoir couvrir des surfaces conséquentes (Ford et Williams, 2007).

- Les méthodes gravimétriques se basent sur l'étude des variations du champ gravitaire pour déterminer la densité des différentes roches du sous-sol et ainsi les différencier. Les vides associés aux réseaux provoquent une baisse des valeurs mesurées, liée à la masse manquante. Les variations étant infimes dans le cas des karsts, les mesures nécessitent des instruments de haute précision et une grande méticulosité sur le terrain (McGrath et al., 2002; Ford et Williams, 2007). Ces mesures sont généralement réalisées le long d'une grille régulière et requièrent de longues périodes d'acquisition.
- Les méthodes sismiques consistent à analyser la propagation des ondes sismiques dans le sous-sol, leur comportement variant lorsqu'elles traversent une discontinuité ou changent de milieu. Elles sont principalement utilisées pour identifier et *cartographier* les différentes couches géologiques du sous-sol mais peuvent également être appliquées à des terrains karstiques (e.g., Sumanovac et Weisser, 2001; Valois et al., 2010; Valois et al., 2011; Cassidy et al., 2014; Tian et al., 2017). Les ondes sismiques analysées sont générées artificiellement via percussion et explosions et les ondes réfléchies et réfractées sont généralement captées par un réseau de géophones. Ces méthodes permettent d'analyser des roches situées à plusieurs centaines de mètres de profondeur mais les hétérogénéités karstiques complexifient l'interprétation des mesures (Schmelzbach et al., 2015). Ces méthodes permettent néanmoins d'identifier des grandes structures associées aux terrains karstiques, comme des karsts en tours ou des canyons (e.g., Tian et al., 2017), voire des structures plus petites comme des dolines (e.g., Valois et al., 2011). La résolution des résultats étant plurimétrique à forte profondeur, cela rend difficile l'identification de conduits métriques (Figure 1.20). Ces méthodes ne permettent donc pas de cartographier des réseaux entiers (Francesconi et al., 2018).
- Les méthodes électromagnétiques consistent à envoyer des ondes électromagnétiques dans le sol (Martínez-Moreno et al., 2014). Le fonctionnement de ces méthodes est très semblable à celui des méthodes sismiques, la principale différence étant que l'on analyse des différences de permittivité diélectrique au lieu de changements d'impédance acoustique.
- Les méthodes magnétiques, comme la magnétométrie, sont principalement utiles lorsque les cavités sont remplies de minéralisations magnétiques, comme des minéralisation d'oxydes de fer (Mochales et al., 2008; Martínez-Moreno et al., 2014; Vyzhva et al., 2020).

Ces méthodes étant complémentaires, elles sont habituellement utilisées en parallèle et la combinaison de leurs résultats permet de réduire les incertitudes associées (e.g., Sumanovac et Weisser, 2001; McGrath et al., 2002; Martínez-Moreno et al., 2014). Les méthodes géophysiques peuvent ainsi avoir une bonne précision et être efficaces pour identifier les cavités et la présence de fluides, principalement en subsurface. Elles ont également l'avantage d'être non destructives et souvent peu coûteuses (Vyzhva et al.,

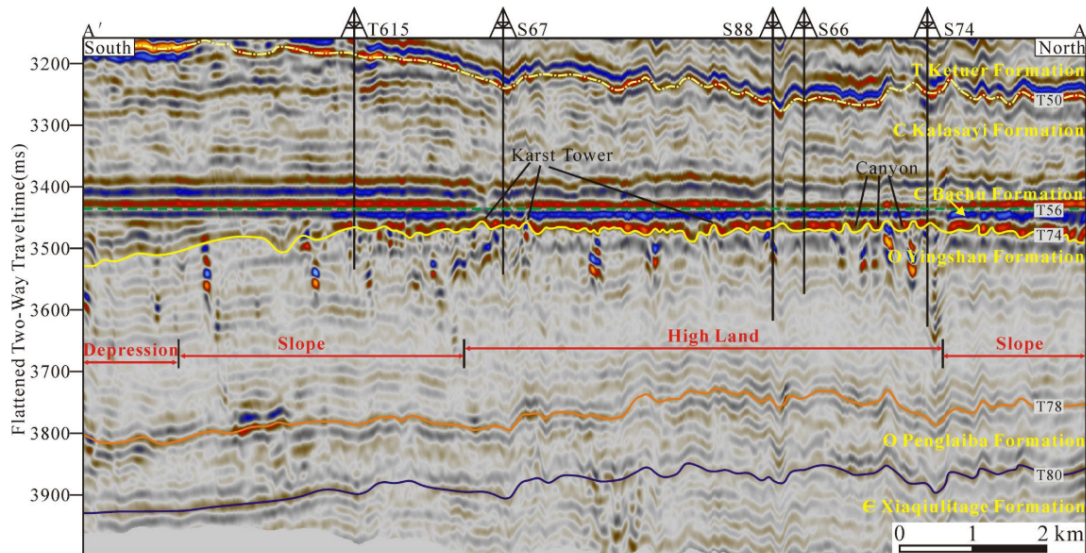


Figure 1.20 – Profil sismique réalisé dans la formation de Bachu, Chine. Des morphologies de karsts tropicaux (tours et canyons) sont mis en évidence mais les conduits des réseaux ne sont pas identifiables (Tian et al., 2017).

2020). Elles restent cependant très limitées quant à la taille des vides identifiables et à leur profondeur d'investigation.

1.2.4.d Utilisation des puits et forages

Bien que de nombreux réseaux soient présents en subsurface, beaucoup de systèmes paléokarstiques sont enfouis profondément sous terre et ne sont pas explorables humainement. Leur profondeur limite également les informations accessibles par l'utilisation de méthodes géophysiques. Tandis que les systèmes de surfaces ont généralement un rôle d'aquifère, les paléokarsts peuvent constituer des réservoirs potentiels de pétrole et de gaz. Dans ces contextes, le forage de puits est à la fois une fin et un moyen. Une fin car les puits permettent l'extraction des ressources et un moyen car ils offrent des informations permettant de caractériser le sous-sol et d'optimiser l'extraction. Les conduits karstiques fonctionnant comme des drains, ils contrôlent très fortement les écoulements fluides. Lorsqu'un conduit est traversé lors d'un forage, le flux dans le puits correspondant peut être très conséquent, similairement à une source (Ford et Williams, 2007). Les conséquences sont importantes car cela peut augmenter les volumes de pétroles extraits mais également accélérer la venue d'eau et précipiter la fin de l'exploitation du puits dans un contexte pétrolier. Il est cependant assez rare que les forages intersectent réellement des conduits importants car le pourcentage volumique qu'ils occupent est assez faible, généralement moins de 1% du volume d'étude et très rarement plus de 2.5% (Worthington, 1999). La majorité des forages intersectent des petits vides n'étant que peu connectés aux drains majeurs, ce qui engendre des flux relativement peu importants au sein du puits correspondant, comparativement aux conduits avoisinants, mais plus importants que ceux de puits présents dans des zones peu karstifiées.

L'étude pétrophysique des déblais de forage (*well cuttings*) permet de caractériser les propriétés des roches traversées, notamment leur nature et leur perméabilité, à travers

une grande diversité de techniques (Freeze et Cherry, 1979; Chapellier, 1992; Ford et Williams, 2007). Les résultats sont présentés sous forme de diagraphies, indiquant les caractéristiques des roches rencontrées sur toute la profondeur du forage. Ces analyses permettent ainsi d’obtenir des données concernant i) la localisation des zones les plus perméables de la formation karstifiée, ii) la conductivité hydraulique d’un petit volume de roche autour du puits, iii) le stockage de cette partie de l’aquifère (Ford et Williams, 2007). Les fortes hétérogénéités au sein des roches karstifiées rendent cependant peu précises les extrapolations réalisées à partir de ces données locales.

Qu’il s’agisse de paléokarsts enfouis profondément ou de réseaux actifs peu profonds, les puits permettent également de procéder à des études complémentaires à celle du comportement naturel des réseaux présentée en Section 1.2.4.b. En effet, qu’il s’agisse d’extraire du pétrole ou de caractériser le fonctionnement d’un aquifère, la présence d’un ou de plusieurs puits permet de procéder à des essais de recharge, de pompage, de traçage, d’injection et de récupération. Ces tests consistent à injecter ou extraire de l’eau ou du pétrole au niveau d’un puits et à observer les variations qui en découlent sur de courtes ou de longues périodes. Ils donnent des informations sur la conductivité hydraulique des roches autour du puits, ainsi que sur le stockage. Si plusieurs puits sont présents autour du puits de test, cela permet d’observer les évolutions de la hauteur de l’eau ou de l’huile sur une zone plus étendue. L’injection d’un traceur et sa récupération dans les puits avoisinants permet d’avoir plus de précision quant aux liens entre les différentes zones.

De manière générale, les forages fournissent des données de haute résolution quant aux roches et vides rencontrés, contrairement aux méthodes géophysiques, mais leur couverture tridimensionnelle est très limitée (Francesconi et al., 2018). Il s’agit cependant souvent de la seule source d’information hydrogéologique sur les systèmes enfouis. Il convient également de noter que la présence d’un puits ou la connexion de plusieurs aquifères peut augmenter les processus de karstification (Vyzhva et al., 2020).

1.3 Modélisation et simulation des systèmes karstiques

La modélisation numérique des karsts permet d’améliorer leur compréhension et d’optimiser l’exploitation des ressources associées. Les modèles apportent des informations sur la structure des karsts et sur leur fonctionnement. Suivant leur cadre d’utilisation, on distingue deux principales familles de méthodes de modélisation : les approches *fonctionnelles* et les approches *structurelles* (Jeannin, 1996; Fournillon, 2012; Jouvès, 2018). Les approches *fonctionnelles* ont pour principal but de modéliser le fonctionnement hydrologique des aquifères karstiques. Les approches *structurelles* cherchent à modéliser explicitement des réseaux karstiques discrets, ce qui permet de mieux appréhender les incertitudes inhérentes aux méthodes d’acquisition présentées en Section 1.2.4. On distinguera également les approches *déterministes* et les approches *stochastiques*. A travers la résolution d’équations, les approches déterministes fournissent un résultat unique considéré comme *optimal*. Inversement, les approches stochastiques incluent une part d’aléatoire, de manière à produire plusieurs réalisations équiprobables à partir d’un ensemble de paramètres d’entrée. Les méthodes de modélisation sont généralement déterministes mais l’utilisation de méthodes stochastiques est très fréquente en

géosciences car elles permettent une bonne gestion des incertitudes. Une classification des différentes méthodes de modélisation des karsts est proposée par Jouves [2018].

1.3.1 Approches fonctionnelles

L'analyse des comportements hydrologiques des aquifères karstiques permet à terme de les représenter sous forme de modèles numériques dynamiques. Ces modèles sont très importants pour la gestion des ressources en eau présentes dans les aquifères, par exemple pour prédire les changements de l'aquifère en fonction du pompage et du climat ou de limiter l'impact d'éventuelles pollutions. Ainsi, les approches fonctionnelles se focalisent sur le comportement temporel de l'aquifère, à travers sa dynamique, plus que sur son comportement spatial.

Bien que les écoulements s'effectuant dans les milieux fissurés soient assez peu caractérisés, ceux s'effectuant dans les milieux poreux ou dans les réseaux karstiques sont relativement bien compris (Gale, 1984). Les flux matriciels sont laminaires tandis que les flux au sein des conduits sont généralement turbulents. Ces écoulements turbulents peuvent être problématiques s'ils ne sont pas pris en compte correctement. De nombreux modèles hydrologiques se basent en effet sur la loi de Darcy qui n'est utilisable que pour des écoulements laminaires (Scanlon et al., 2003).

Plusieurs types d'approches fonctionnelles existent, mais elles peuvent être divisées en deux catégories. D'une part, les approches dites globales, ou pluie-débit, permettent de caractériser la réponse hydraulique des réseaux sans prendre en compte les hétérogénéités spatiales des aquifères. D'autre part, les approches discrétisées ou distribuées qui prennent en compte la distribution spatiale de l'aquifère, notamment celle du réseau correspondant, à travers différents paramètres physiques. Très récemment, Jeannin et al. [2021] ont appliqué plusieurs méthodes à un même cas d'étude afin d'en faciliter la comparaison. Pour autant, il reste difficile de définir une méthode *idéale*, tant les résultats dépendent de nombreux facteurs, notamment des données disponibles et des objectifs poursuivis.

Approches globales Les approches globales permettent de modéliser le flux en sortie d'un réseau sans prendre en compte sa morphologie, omettant toute information spatiale dans les équations associées. Ainsi, il est généralement possible de modéliser le fonctionnement hydrologique de l'aquifère via des fonctions de transfert linéaires, liant recharge et décharge, ne prenant habituellement pas les processus physiques en compte (e.g., Mangin, 1975; Dreiss, 1989; Long et Derickson, 1999; Scanlon et al., 2003). On parle alors d'approches linéaires. Les entrées du système correspondent généralement aux informations pluviométriques à l'échelle du bassin versant et les sorties aux débits mesurés aux sources. Les résultats modélisés sont généralement proches des débits réels mesurés aux sorties du réseau (e.g., Barrett et Charbeneau, 1996; Zhang et Bai, 1996) mais restent limités à ceux-ci (Peterson et Wicks, 2006). Certaines approches plus complexes visent à modéliser le fonctionnement de l'aquifère sous forme de plusieurs réservoirs interconnectés (e.g., Fleury et al., 2007; Baudement et al., 2017; Mazzilli et al., 2017). Le principe est similaire, mais les fonctions de transfert associées ne sont pas toujours linéaires. Les fonctions de transfert sont habituellement de la forme

suivante (Hartmann et al., 2014) :

$$Q = \frac{dS}{dt} = a.S^b \quad (1.5)$$

avec Q le débit du réservoir, S le volume stocké en son sein, a et b des paramètres propres au système devant être calibrés. Des approches se basant sur l'apprentissage automatique (*machine learning*), ayant pour but de lier les valeurs en entrée et en sortie grâce l'utilisation de réseaux de neurones, commencent également à se développer (e.g., Hu et al., 2008; Kong A Siou et al., 2011; Kong-A-Siou et al., 2014). L'utilisation des approches globales requiert cependant l'enregistrement en continu, et sur de longues durées, des informations pluviométriques à l'échelle du bassin versant ainsi que des débits aux sources, afin de pouvoir calibrer les paramètres et de tester les modèles rigoureusement. Ces méthodes sont efficaces pour modéliser le comportement hydrologique du système à l'échelle globale, plus particulièrement sa réponse aux précipitations. Bien que plusieurs méthodes puissent prendre en compte certaines différences spatiales des caractéristiques de l'aquifère (e.g., Bittner et al., 2018), ce n'est pas le cas de la majorité d'entre elles, ce qui limite leur intérêt à l'échelle locale (Peterson et Wicks, 2006).

Approches distribuées Pour obtenir des informations plus précises sur les flux souterrains, on aura plutôt tendance à utiliser des approches distribuées dans lesquelles différents paramètres sont définis spatialement. Les systèmes sont discrétisés sur des grilles 2D ou 3D dont les cellules sont associées à différentes propriétés hydrauliques. Suivant le nombre de paramètres utilisés, les temps de calibration peuvent devenir très importants (Jeannin et al., 2021). Les écoulements dans les cellules correspondant à la matrice sont habituellement basés sur l'équation de la diffusivité (Hartmann et al., 2014) :

$$S_s \frac{\partial H}{\partial t} = \nabla(K\nabla H) \quad (1.6)$$

avec S_s le coefficient d'emmagasinement spécifique, K la conductivité hydraulique, H la charge hydraulique, ∇ l'opérateur de divergence et t le temps. Les écoulements dans les conduits sont quant à eux généralement définis par l'équation de Darcy-Weisbach (Young et al., 2010; Reimann et al., 2011; Hartmann et al., 2014) :

$$\frac{\partial H}{\partial x} = -\mu \frac{q^2}{2gd} \quad (1.7)$$

avec μ le coefficient de friction, d le diamètre du conduit, q la vitesse moyenne d'écoulement, x la distance du chemin d'écoulement et g la constante gravitationnelle.

L'approche simple continuum, également désignée par le nom d'approche par milieu poreux équivalent (*Equivalent Porous Medium* : EPM), englobe toutes les hétérogénéités de l'aquifère à l'échelle locale au sein d'une seule valeur par cellule pour chaque propriété hydrodynamique. Cela correspond à un équivalent des valeurs de la matrice, des fractures et des conduits pour chaque propriété (Figure 1.21B). Cette approche est efficace pour simuler les flux moyens à l'échelle de la région mais manque de précision pour les systèmes très karstifiés, associés à la présence de nombreux conduits. Les écoulements

rapides ayant lieu dans ces derniers ne sont en effet que très peu pris en compte.

L'approche double continuum permet de modéliser la dualité des écoulements karstiques. Un continuum correspond à la matrice (faible conductivité, grand stockage, flux diffus) et l'autre aux conduits (forte conductivité, faible stockage) (Scanlon et al., 2003) (Figure 1.21C). Les deux continnum sont connectés entre eux et les flux dans chacun d'entre eux sont considérés comme étant laminaires (Scanlon et al., 2003), bien qu'ils soient en réalité turbulents dans les conduits. Les échanges d'eau dépendent principalement de la différence de pression entre les deux ensembles (Teutsch, 1993) et peuvent être décrits via des fonctions linéaires, à travers un coefficient d'échange (e.g., Maréchal et al., 2008; Kordilla et al., 2012). Il faut cependant réussir à estimer les propriétés hydrodynamiques de la matrice et des conduits pour que cette approche soit efficace. Elle est similaire à l'approche double porosité utilisée dans d'autres domaines, notamment l'industrie pétrolière (Warren et Root, 1963).

L'approche multi-continuum est similaire à l'approche double-continuum, exception faite qu'elle intègre également des flux de surface sous forme d'un troisième continuum (Bresinsky et al., 2019; MedWater). Ce continuum interagit avec le continuum matriciel, permettant de prendre en compte les différentes infiltrations d'eau plus complexes associées aux systèmes karstiques. Ce terme multi-continuum est cependant peu utilisé dans la littérature et certains auteurs s'en servent pour désigner l'approche double-continuum (e.g., Liedl et al., 2003).

L'approche mixte discrète sur continuum (Combined Discrete-Continuum approach) utilise des réseaux de conduits discrets inclus dans un continuum correspondant à la matrice, ce qui permet d'obtenir des modèles assez précis spatialement (e.g., Kiraly, 1984; Kiraly, 1998; Macquarrie et Sudicky, 1996) (Figure 1.21D). Elle est notamment utilisée dans MODFLOW-CFP (Kuniansky et al., 2008; Reimann et Hill, 2009). Elle nécessite de connaître les propriétés de la matrice et d'autres paramètres, notamment le coefficient d'échange entre la matrice et les conduits qui peut grandement impacter les résultats (e.g., Dufoyer et al., 2019), mais surtout la localisation et la géométrie des conduits, rendant compliqué son utilisation en l'absence de données. Les réseaux utilisés étant généralement définis avant la simulation, leurs topologies et géométries sont rarement ajustées pendant le processus de calibration (Jeannin et al., 2021).

Les flux peuvent également être simulés directement sur des réseaux de conduits discrets, souvent représentés par des tuyaux rectilinéaires, sans forcément prendre en compte les écoulements dans la matrice (e.g., SWMM, Epanet, Modflow-CFP) (e.g., Thrailkill, 1974; Jeannin, 2001; Campbell et Sullivan, 2002; Peterson et Wicks, 2006; Chen et Goldscheider, 2014; Jeannin et al., 2015; Gill et al., 2020). Ils sont très utilisés pour les écoulements dans les réseaux urbains d'adduction d'eau, notamment pour la gestion des eaux de tempêtes et des débordements associés (e.g., Rossman, 2015). Ces modèles ont l'avantage de permettre le suivi des écoulements en n'importe quel point des réseaux. Ils peuvent être efficaces pour mettre en relation les débits en entrée et en sortie mais, pour lier précipitations et débits en sortie, il est nécessaire d'utiliser un modèle de recharge en complément (Jeannin et al., 2021).

La principale limite des approches utilisant des réseaux de conduits discrets concerne les réseaux eux-mêmes. Un léger manque de précision peut fortement affecter les résultats des simulations d'écoulement (e.g., Peterson et Wicks, 2006). L'utilisation de

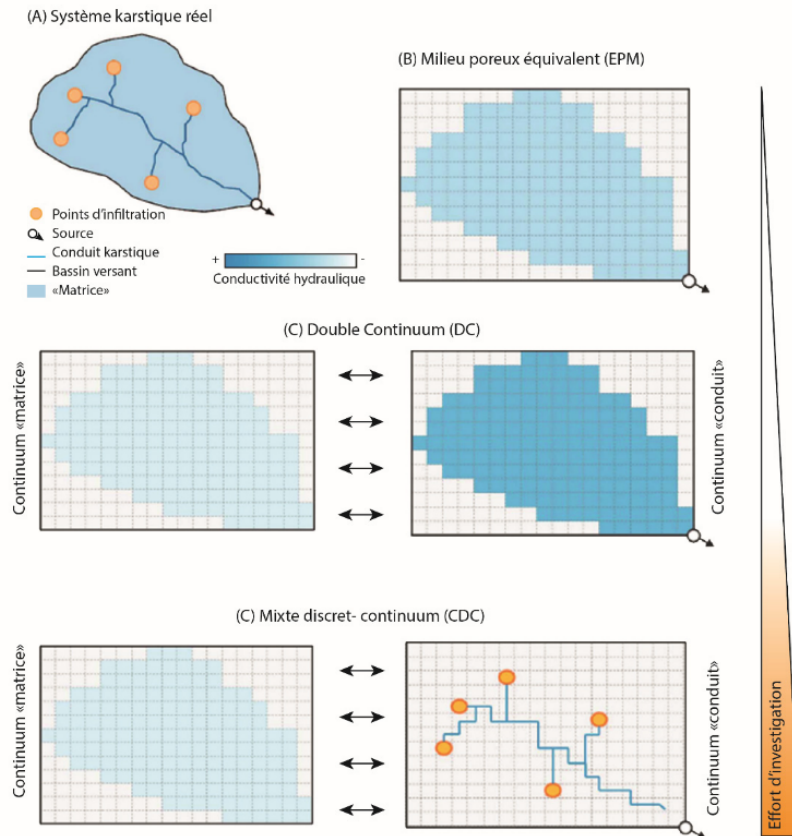


Figure 1.21 – Schéma de différentes approches de modélisation distribuées développées pour les aquifères karstiques. Système karstique réel (A) modélisé via différentes approches. B : Approche milieux poreux équivalent (EPM). C : Approche double continuum (DC). D : Approche mixte discrète sur continuum (CDC) (Jouves, 2018, modifié d'après Hartmann et al., 2014 et Ghasemizadeh et al., 2012).

modèles de réseaux proches des observations de terrain est ainsi recommandée. Pour autant, l'obtention d'informations détaillées est complexe (Section 1.2.4). Dans le cas où les données ne sont pas suffisantes, il est nécessaire de se servir de réseaux simulés. Les approches structurelles sont un moyen d'obtenir de tels réseaux discrets, pouvant servir de support aux simulations.

1.3.2 Approches structurelles

Les approches structurelles visent à simuler des réseaux karstiques entiers ou éventuellement des parties de réseaux dont l'existence est connue mais n'a pas pu être caractérisée. Contrairement aux approches fonctionnelles, elles se focalisent sur l'état *statique* (ou *spatial*) du karst, plus que sur son comportement *dynamique* (ou *temporel*) (Jouves, 2018). Certaines méthodes permettent la prise en compte d'informations connues sur le réseau à simuler ou sur des réseaux analogues, ce qui permet de mieux appréhender les incertitudes associées. L'utilisation de méthodes stochastiques permet de tenir compte au mieux de ces informations. Cette sous-partie se focalise sur deux aspects de la simulation des réseaux : i) La simulation du squelette des réseaux, ce qui comprend leur position dans l'espace et leur topologie et ii) la simulation de leur géométrie.

1.3.2.a Simulation du squelette des réseaux karstiques

Il existe de nombreuses méthodes visant à simuler des réseaux karstiques discrets (e.g., Pardo-Igúzquiza et al., 2012; Borghi et al., 2012; Collon-Drouaillet et al., 2012; Fournillon et al., 2012; Viseur et al., 2014; Hendrick et Renard, 2016a). Elles peuvent se baser sur des concepts très différents les uns des autres et le choix d’une méthode dépend principalement du cas d’étude, de l’échelle de la simulation et des données disponibles. Les approches structurelles sont parfois divisées en deux grandes catégories, suivant qu’elles se basent sur des approches de modélisation basées-processus ou basées-structures (Koltermann et Gorelick, 1996; Jouvès, 2018).

Approches basées-processus Les approches basées-processus cherchent à reproduire les processus physiques et chimiques contrôlant la formation des réseaux (e.g., Kaufmann et Braun, 2000; Birk et al., 2003; Dreybrodt et al., 2005). La reproduction du processus de spéléogénèse se base sur la résolution d’équations physiques et chimiques, décrivant notamment la dissolution des roches carbonatées. Elles permettent de simuler l’évolution des réseaux au cours du temps et présentent ainsi un grand intérêt théorique pour la compréhension de la spéléogénèse (e.g., Dreybrodt et al., 2005; Gabrovšek et al., 2014). Les réseaux simulés sont ainsi *cohérents* d’un point de vue géologique et physique (Jouvès, 2018). Ces approches présentent néanmoins certaines limites, rendant difficile leur utilisation sur des cas réels : i) elles nécessitent la connaissance de nombreux paramètres physico-chimiques, environnementaux et géologiques ainsi que leur évolution sur de grandes périodes temporelles, ii) elles sont déterministes et les temps de calcul associés sont longs, limitant la prise en compte des incertitudes et iii) elles intègrent difficilement les positions de conduits connues (*hard data*) (Jouvès, 2018).

Approches basées-structures Les approches basées-structures cherchent à reproduire la structure des réseaux sans prendre en compte les processus spéléogénétiques. La majorité d’entre elles sont stochastiques, ce qui permet une bonne prise en compte des incertitudes. Beaucoup de méthodes se basent sur l’utilisation de statistiques spatiales (e.g., Fournillon et al., 2012; Viseur et al., 2014) et peuvent être qualifiées de méthodes *géostatistiques*. Elles se servent de statistiques calculées sur la structure spatiale de réseaux ou portions de réseaux connus pour les reproduire lors de simulations. Des simplifications et approximations du processus de formation des réseaux, combinées ou non à des statistiques spatiales, ont également permis de proposer des méthodes de modélisation originales (e.g., Jaquet et al., 2004; Labourdette et al., 2007; Pardo-Igúzquiza et al., 2012; Borghi et al., 2012; Collon-Drouaillet et al., 2012; Hendrick et Renard, 2016a). Les méthodes géostatistiques se divisent en deux grandes catégories : les méthodes basées-objets et les méthodes basées-pixel.

Méthodes basées-objets Les méthodes basées-objets décrivent les objets à simuler à travers différents paramètres, notamment géométriques, et les simulent de manière discrète dans un volume souhaité. Les paramètres peuvent être ajustés par l’utilisateur pour obtenir le résultat voulu. Ces méthodes sont notamment utilisées pour la simulation de chenaux (e.g., Viseur, 2004; Rongier et al., 2017; Parquer, 2018) mais restent en-

core peu utilisées pour la simulation de réseaux karstiques (Anquez et al., 2015; Mourlanette et al., 2016). Elles ont l'avantage de ne pas être basées sur des grilles, ce qui facilite la simulation d'objets curvilinéaires et réduit les temps de calcul. Elles sont néanmoins limitées car : i) l'obtention de modèles cohérents requiert souvent l'utilisation d'un grand nombre de paramètres, ii) elles intègrent difficilement les positions de conduits connues (*hard data*) et iii) elles sont peu utilisées pour des objets tridimensionnels tels que les réseaux karstiques.

Méthodes basées-pixel Les méthodes basées-pixel simulent la valeur d'une variable sur chaque voxel d'une grille correspondant à la zone d'intérêt. Elles visent à reproduire sur la grille des statistiques pouvant soit être obtenues sur des réseaux connus, soit être définies directement par l'utilisateur. Dans le cas présent, la variable à simuler correspond généralement à la présence ou non d'un conduit sur le voxel considéré. Il s'agit d'une variable discrète pouvant être assimilée à un *faciès* et ayant deux valeurs possibles : soit le voxel fait partie d'un conduit, soit cela n'est pas le cas. Ces méthodes peuvent être divisées en deux catégories, suivant qu'elles se basent sur l'utilisation de variogrammes (statistiques deux-points) ou d'images d'entraînement (statistiques multipoints) (Figure 1.22).

Les variogrammes sont des fonctions mathématiques décrivant la variabilité spatiale d'une variable aléatoire donnée dans une direction (Chapitre 2). Il existe plusieurs méthodes permettant de reproduire des variogrammes de variables discrètes. Parmi les plus utilisées se trouvent les méthodes de simulation séquentielle d'indicatrices (SIS) (e.g., Goovaerts, 1997; Deutsch et Journel, 1998). Il s'agit de méthodes séquentielles, ce qui signifie que les valeurs des voxels sont simulées l'une après l'autre, dans un ordre aléatoire, et qu'elles conditionnent la simulation des suivantes. Pour chaque voxel, une loi de probabilité (probabilité que le voxel soit associé à un conduit) est déterminée à partir des résultats d'un krigeage effectué en se basant sur les valeurs des voxels voisins et des paramètres d'entrée (variogramme, proportion de chaque faciès), puis un tirage aléatoire a lieu. Elles ont l'avantage de respecter les données conditionnantes, les variogrammes et les proportions de chaque faciès donnés en entrée et ont pu être adaptées au cas des réseaux karstiques (e.g., Fournillon et al., 2012; Viseur et al., 2014). Néanmoins, ces méthodes utilisant à la fois des grilles et des variogrammes, il est relativement compliqué de simuler des objets curvilinéaires.

Les méthodes multipoints utilisent une image d'entraînement similaire aux structures devant être simulées. Elles sont habituellement séquentielles et basées sur des grilles. L'algorithme analyse la configuration des cellules voisines du voxel devant être simulé et recherche des configurations identiques dans l'image d'entraînement. Suivant la méthode, le voxel se verra assigner soit une valeur tirée aléatoirement parmi celles des voxels au centre d'une configuration similaire dans l'image d'entraînement (e.g., Mariethoz et al., 2010), soit une valeur de probabilité (e.g., Straubhaar et al., 2011). Des méthodes de MPS non maillées, permettant d'utiliser des nuages de points au lieu de grilles, ont également été proposées pour permettre la simulation de réseaux karstiques (Erzeybek et al., 2012; Ghobadi, 2016). Malheureusement ces approches sont largement limitées par l'utilisation d'une grille qui doit avoir une grande résolution et leur difficulté à reproduire des réseaux connectés.

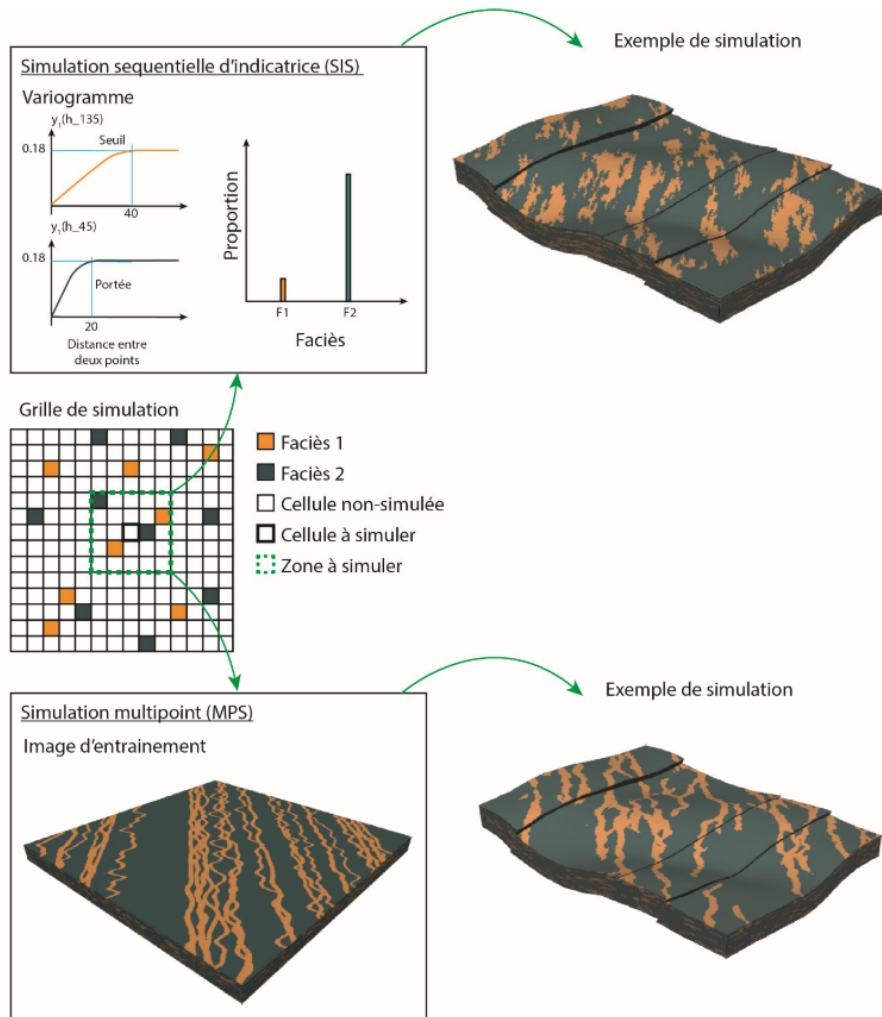


Figure 1.22 – Comparaison de deux méthodes de simulation basées-pixel : la simulation séquentielle d'indicateurs (SIS) utilisant des variogrammes et la simulation multipoints (MPS) utilisant des images d'entraînement. Les valeurs simulées proviennent d'un modèle issu de la configuration connue des cellules dans la zone autour du voxel à simuler (Jouves, 2018, modifié d'après Rongier et al., 2016).

Méthodes notables Parmi les méthodes existantes, nous considérons les méthodes suivantes comme étant les plus notables au vu de l'objectif de cette thèse :

- Henrion et al. [2008] et Collon-Drouaillet et al. [2012] proposent de discrétiser l'ensemble matrice-réseau de fractures en un unique réseau de tubes (*pipe network*). Le chemin le plus court entre une entrée et une sortie est ensuite calculé dans ce réseau grâce à l'algorithme A^* (Figure 1.23A). Dans Collon-Drouaillet et al. [2012], de nouveaux points sont ajoutés de manière récursive, pour reconstruire un réseau arborescent dont la topologie respecte des critères d'Horton-Strahler (Horton, 1945). Les nouveaux points d'entrée sont raccordés au reste du réseau karstique grâce à l'algorithme A^* . Cette méthode a l'avantage de respecter les données conditionnantes comme les positions de conduits connus, les réseaux de fractures, les horizons d'inception et les anciens niveaux piézométriques, mais ses applications directes sont principalement limitées aux réseaux arborescents.

- Borghi et al. [2012] se servent d'un modèle géologique et d'un réseau de fractures discrètes, habituellement simulé stochastiquement, pour générer un champ de vitesses, les hétérogénéités étant associées à des vitesses plus grandes que la matrice. Les entrées et sorties du réseau sont déterminées ou simulées stochastiquement à partir des informations topographiques et hydrogéologiques. Un algorithme de *fast-marching* est alors utilisé pour calculer le chemin le plus court entre les sorties et entrées, chemin correspondant au réseau simulé (Figure 1.23B'). Cet algorithme peut être utilisé de manière itérative pour simuler plusieurs phases de karstification, les conduits simulés étant utilisés pour mettre à jour le champ de vitesses. Il peut respecter les données relatives aux positions des conduits connus mais l'utilisation d'un réseau de fractures et d'une grille aboutit à des conduits en *marches d'escalier*.
- Hendrick et Renard [2016a] génèrent un amas de percolation et définissent le réseau comme l'ensemble des liens connectés traversés par un flux supérieur à un certain seuil défini par rapport au flux maximal (Figure 1.23C). Cette méthode est cependant principalement limitée au cas de réseaux bidimensionnels et s'appuie là encore sur une grille. Les conduits sont donc discrétisés en segments orthogonaux les uns aux autres.
- Fournillon et al. [2012] et Viseur et al. [2014] divisent les parties connues des réseaux en différentes familles de conduits, caractérisent leurs paramètres géométriques et topologiques et utilisent ces statistiques dans des simulations séquentielles d'indicatrices pour générer des conduits dans une grille, de manière récursive (Figure 1.23D). Pour autant, ces méthodes utilisant à la fois des grilles et des variogrammes, cela rend difficile la simulation de conduits curvilinéaires.
- Pardo-Igúzquiza et al. [2012] génèrent un ensemble de conduits basés sur la géométrie de conduits types (*templates*) (Figure 1.23E). Les conduits générés de la sorte sont ensuite déplacés aléatoirement. Lorsqu'un conduit en rencontre un autre, ils sont agrégés et c'est le nouvel agrégat qui continue le déplacement aléatoire. Ce processus d'agglomération continue aussi longtemps que voulu ou jusqu'à ce qu'un unique ensemble soit formé (Figure 1.23E'). Cette méthode intègre cependant difficilement les informations de terrain, notamment la position des conduits connus.
- Une méthode basée-objet, utilisant les systèmes de Lindenmayer, a été développée au sein de l'équipe RING sans connaître de publication officielle (Anquez et al., 2015; Mourlanette et al., 2016). Les systèmes de Lindenmayer (Lindenmayer, 1968; Prusinkiewicz et Lindenmayer, 1990) correspondent à une grammaire formelle, ce qui revient à réécrire un objet complexe en le séparant en plusieurs morceaux simples et en appliquant à chacun d'entre eux un ensemble de règles. Une fois traduits en informations graphiques, cela aboutit à la création de segments connectant les données brutes avec des nœuds nouvellement créés. Cette méthode est actuellement principalement adaptée au cas des réseaux arborescents et anastomosés.

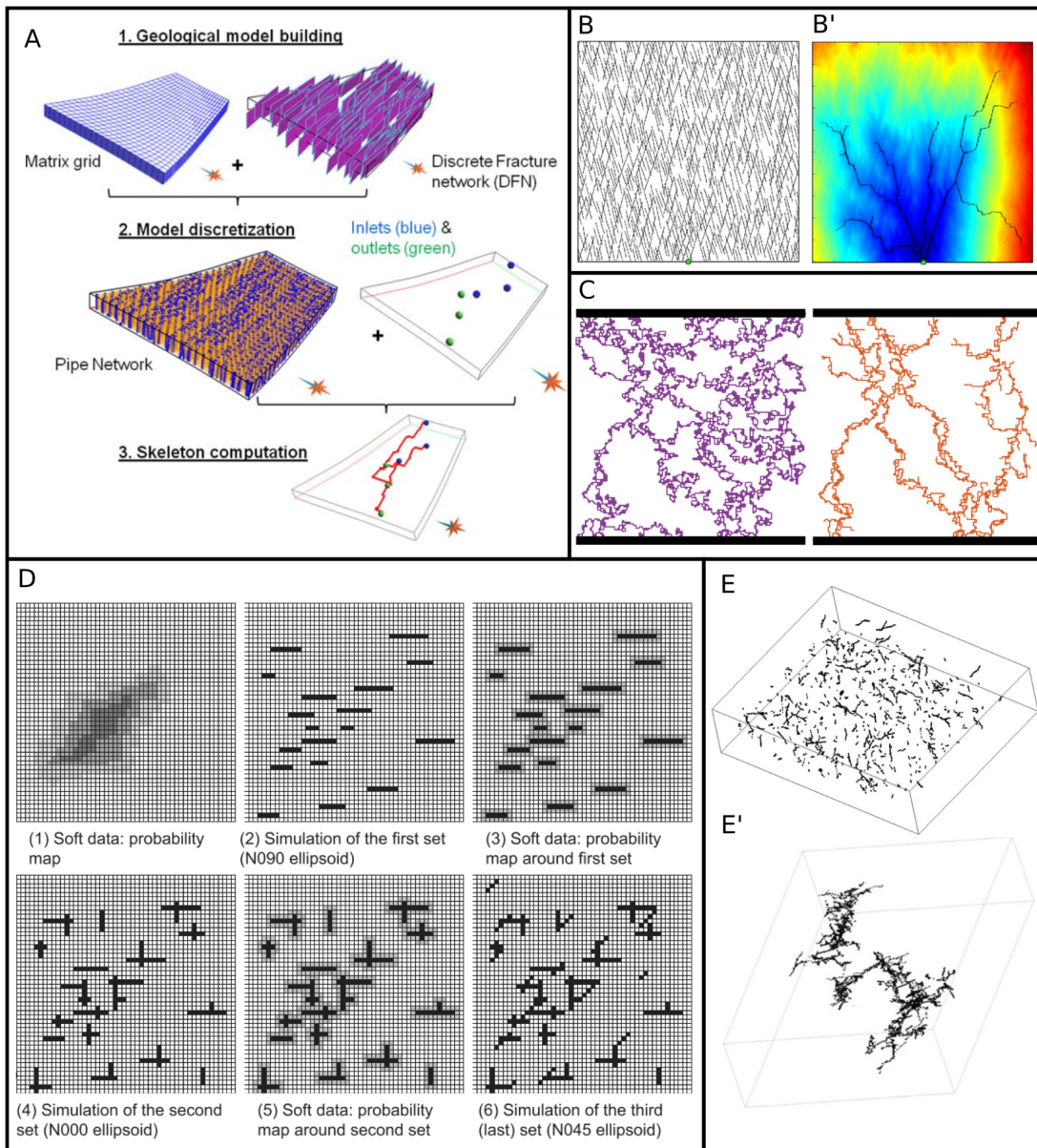


Figure 1.23 – Exemples de méthodes permettant la simulation de squelettes de réseaux karstiques. A : Génération d'un squelette de réseau à partir d'une grille et d'un réseau de fractures discrètes (modifié d'après Collon-Drouaillet et al., 2012). B : Exemple d'un réseau de fractures discrètes. B' : Réseau obtenu après plusieurs itérations de l'algorithme de *fast-marching* (modifié d'après Borghi et al., 2012). La couleur correspond au temps nécessaire pour atteindre cette cellule de la grille depuis la source (vert), les petits temps étant en bleu et les grands temps en rouge. C : Exemple de deux réseaux obtenus à partir du même amas de percolation mais pour des seuils différents (plus élevé à droite) (modifié d'après Hendrick et Renard, 2016a). D : Simulation stochastique de plusieurs familles de conduits (modifié d'après Fournillon et al., 2012). E : Ensemble de conduits générés aléatoirement. E' : Exemple de réseau obtenu après leur agrégation (modifié d'après Pardo-Igúzquiza et al., 2012).

Bilan Exception faite des méthodes basées-processus, la majorité des méthodes de simulation de réseaux discrets ont l'avantage d'être stochastiques. Elles présentent néanmoins certaines limitations. Elles se basent pour beaucoup sur l'utilisation de grilles, ce qui limite la simulation de conduits curvilinéaires. Les squelettes simulés sont en effet alignés sur les axes de la grille, provoquant des problèmes *d'aliasing* et ces *marches d'escalier* peuvent influencer le résultat des simulations d'écoulement. Utiliser des résolutions de grille fines permet de réduire l'impact de ce problème mais les réseaux simulés étant tridimensionnels, cela peut induire d'importants temps de calcul. Beaucoup de méthodes ne simulent pas de morphologie de réseau particulière (e.g., Pardo-Igúzquiza et al., 2012) ou sont spécifiques à un seul motif morphologique (e.g., Borghi et al., 2012; Collon-Drouaillet et al., 2012). Pour autant, les processus spéléogénétiques ayant habituellement lieu en plusieurs phases, les réseaux réels combinent souvent plusieurs types de motifs (e.g., réseau arborescent vertical dans la partie vadose, réseau anastomosé dans la partie épiphréatique) (Jouves et al., 2017). Les méthodes actuelles ne sont donc pas adaptées à la simulation de réseaux entiers, combinant plusieurs motifs distincts. Ces méthodes prennent en compte différentes données, propres à chacune d'entre elles. Certaines informations sont nécessaires à leur utilisation et inversement certaines informations ne peuvent pas être prises en compte. Bien que certaines méthodes permettent de se servir de données acquises dans des cas analogues (e.g., Pardo-Igúzquiza et al., 2012; Fournillon et al., 2012; Viseur et al., 2014), d'autres sont inutilisables si certaines informations sur le système devant être modélisé sont manquantes, notamment les méthodes basées-processus (e.g., Kaufmann et Braun, 2000). Dans l'idéal, il faudrait que la majorité des informations connues de l'utilisateur puissent contraindre la simulation sans être obligatoirement requises, qu'il s'agisse d'observations de terrain (e.g., position des conduits connus) ou d'informations géologiques et hydrogéologiques (e.g., horizons d'inception, failles, altitude du niveau piézométrique). Le prototype d'une méthode visant à s'affranchir de ces limites sera présenté dans le Chapitre 4.

1.3.2.b Simulation de la géométrie des réseaux karstiques

Les sections des conduits sont souvent de formes complexes et les acquisitions de terrain consistent généralement en la mesure de la plus grande hauteur et de la plus grande largeur sur cette section (Section 1.2.4.a) (Figure 1.19). L'approximation de ces sections sous forme d'une ellipse permet, dans une moindre mesure, de s'approcher de la forme réelle des conduits. La géométrie des conduits est un paramètre important lors des simulations d'écoulement. Peterson et Wicks [2006] montrent en effet qu'une différence de diamètre des conduits de l'ordre de 10% peut changer les écoulements de façon significative. Pour autant, la géométrie réelle des conduits karstiques est rarement prise en compte lors de leur modélisation.

Lorsque des réseaux discrets sont simulés en se servant d'une grille, ils sont généralement associés à ses voxels (e.g., Fournillon et al., 2012; Viseur et al., 2014). Les dimensions des conduits sont alors très dépendantes de la résolution de ladite grille, ce qui peut conduire à de grosses approximations.

Dans le cadre de simulations d'écoulement sur des réseaux discrets, de nombreux auteurs assimilent les conduits à des tubes de section circulaire et utilisent des valeurs de rayon constantes au sein de chacun d'entre eux, les choix étant souvent effectués de

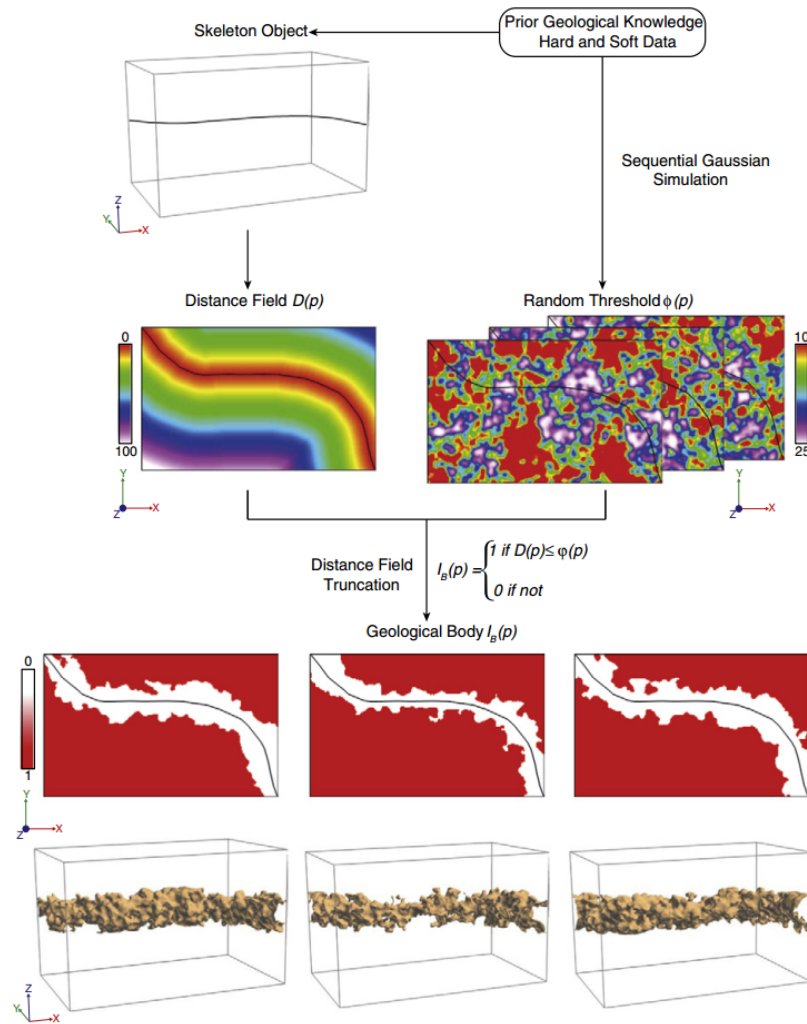


Figure 1.24 – Méthode ODSIM appliquée à la génération d'un conduit karstique à partir de son squelette. Le champ de distance $D(p)$ est tronqué par le seuil aléatoire $\varphi(p)$, ce qui permet d'obtenir une propriété d'indicatrice $I_B(p)$ correspondant à la présence du conduit (Rongier et al., 2014, modifié d'après Henrion et al., 2010).

manière très arbitraire. Ils peuvent spécifier une unique valeur de rayon pour tous les conduits du réseau (e.g., Ronayne, 2013), des valeurs différentes pour chaque conduit (e.g., Chen et Goldscheider, 2014) ou une loi de puissance dépendant de la hiérarchie des conduits dans le réseau : des fortes valeurs pour les conduits recevant la plus forte recharge, des faibles valeurs pour les conduits recevant une recharge plus faible et un élargissement à chaque confluence (e.g., Vuilleumier et al., 2013; Borghi et al., 2016). Souvent, le rayon et les paramètres contrôlant sa distribution sont calibrés directement durant les simulations d'écoulement pour reproduire les débits observés en différents endroits (e.g., Vuilleumier et al., 2013; Jeannin et al., 2015).

Certaines méthodes simulent des géométries de conduits proches des observations de terrain en plus de simuler le squelette du réseau (e.g., Collon-Drouaillet et al., 2012; Pardo-Igúzquiza et al., 2012). La méthode ODSIM (Henrion et al., 2010; Rongier et al., 2014; Schiller et Renard, 2016) utilisée par Collon-Drouaillet et al. [2012] permet notamment de générer les dimensions des conduits de manière stochastique. Elle consiste à

calculer un champ de distance autour du squelette du réseau généré préalablement puis à le tronquer via un seuil généré par SGS (Simulation Gaussienne Séquentielle) (Figure 1.24). Les conduits ainsi générés ont des géométries très variées, tant géométriquement que spatialement. Cependant, cette méthode calcule le champ de distance et effectue la SGS sur une grille. La résolution de cette dernière contrôle donc la capacité à reproduire une géométrie fine et détaillée et induit des temps de calcul qui peuvent être prohibitifs à l'échelle d'un système karstique entier.

1.4 Conclusion

Cette étude bibliographique présente les karsts en tant qu'objets géologiques, les processus intervenant lors de leur formation, ainsi que l'organisation des réseaux associés. Les techniques permettant l'acquisition de données sur les réseaux karstiques et leurs limites sont également décrites. Elles produisent des données éparses conduisant à de grandes incertitudes concernant la position des conduits et leurs dimensions. Différentes méthodes permettant d'améliorer la compréhension de la structure des karsts et de leur fonctionnement sont présentées. Celles ayant pour but de simuler le fonctionnement hydraulique de l'aquifère sont qualifiées de méthodes fonctionnelles. Elles offrent des résultats intéressants à l'échelle du système entier mais manquent souvent de précision à l'échelle locale. L'utilisation de réseaux de conduits discrets permettrait d'améliorer la compréhension des flux au sein des réseaux. Les méthodes structurelles permettent de générer de tels modèles de réseaux discrets en prenant en compte les incertitudes liées aux données, mais elles présentent également certaines limites. Elles sont généralement spécifiques à une morphologie de réseau, ne peuvent pas intégrer toutes les données connues de l'utilisateur et impliquent la manipulation de grandes grilles, ce qui engendre des problèmes d'*aliasing* et de longs temps de calcul. Concernant la modélisation de la géométrie des sections de conduits karstiques, on dispose d'un côté d'approches globales très *grossières*, assimilant chaque conduit à un tube de dimensions fixes, et de l'autre d'approches très locales, produisant des géométries, très, voire peut-être trop détaillées.

Les questions que l'on peut donc légitimement se poser sont les suivantes :

- Existe-t-il des lois statistiques génériques décrivant la géométrie des sections des conduits, ainsi que sa variabilité spatiale, suivant différents critères?
- Quelles méthodes de simulation permettraient de rendre compte de ces informations statistiques à une échelle acceptable pour appréhender le réseau karstique dans son ensemble?
- Peut-on définir une méthode capable de simuler stochastiquement des réseaux non alignés sur une grille, permettant l'intégration de données de terrain ainsi que d'informations géologiques et utilisable pour différentes morphologies de réseaux, tout en offrant des temps de calculs raisonnables?

Chapitre 2

Inférence statistique

Sommaire

2.1	Introduction	50
2.2	Données	51
2.2.1	Base de données utilisée	51
2.2.2	Métriques utilisées	53
2.2.3	Représentation des réseaux via des graphes	55
2.3	Méthodes analytiques	56
2.3.1	Analyses univariées	57
2.3.2	Analyses multivariées	61
2.4	Résultats	64
2.4.1	Analyses univariées	65
2.4.2	Analyses multivariées	71
2.5	Discussion	78
2.6	Conclusion	83

Résumé Ce chapitre présente l'étude statistique de la géométrie des conduits de 49 réseaux karstiques à travers l'analyse de deux métriques : le rayon équivalent et le ratio largeur-hauteur. Les différents réseaux sont analysés de manière individuelle puis comparés les uns aux autres afin de vérifier si des règles statistiques universelles pourraient en être extraites. La variabilité spatiale de ces mêmes propriétés est également quantifiée. Enfin, la manière dont les dimensions des conduits sont organisées au sein du réseau est étudiée, notamment pour vérifier l'existence d'une organisation hiérarchique.

La majorité des résultats présentés dans ce chapitre a été publiée dans l'article suivant : Frantz et al. [2021].

2.1 Introduction

Les systèmes karstiques sont associés à des réseaux complexes qui ne sont que partiellement explorés. De nombreux réseaux enfouis ne peuvent être caractérisés de manière précise, seules les grandes dolines étant généralement observables. D'importantes incertitudes sont de fait présentes lors de la modélisation des réseaux. Les simulations requièrent de nombreuses données pour leur conditionnement mais également pour vérifier la cohérence des résultats. L'étude statistique de réseaux connus permet, par analogie, de caractériser partiellement ces réseaux enfouis, fournissant des données utilisables lors des simulations.

En géomorphologie, l'analyse statistique d'objets curvilinéaires est ancienne puisque Horton propose dès 1945 des applications aux systèmes fluviaux (Horton, 1945). Différentes études se sont ensuite focalisées sur les réseaux karstiques à partir des années 1960 (e.g., Curl, 1966; Howard, 1971). Les métriques telles que l'extension des réseaux, la longueur des conduits, le pendage, l'orientation, la largeur ou la hauteur ont été étudiées par plusieurs auteurs (e.g., Jeannin, 1996; Frumkin et Fischhendler, 2005; Fournillon et al., 2012; Collon et al., 2017), mais seules de rares études ont porté sur l'analyse des dimensions des conduits (Pardo-Iguzquiza et al., 2011; Jouvès et al., 2017).

Dans ce travail nous disposons des données de 49 réseaux karstiques. Les dimensions de leurs conduits sont étudiées à travers deux métriques : le rayon équivalent et le ratio largeur-hauteur. Afin d'identifier des lois ou des comportements de ces métriques, qui seraient observables dans une majorité de réseaux, nous les comparons entre eux et effectuons également des analyses individuelles. La définition de variogrammes 1D-curvilinéaires et le calcul de l'assortativité des dimensions des conduits permet de démontrer l'existence d'une corrélation spatiale au sein des réseaux, corrélation encore plus importante au sein de chaque conduit. Deux hypothèses sont parfois utilisées en modélisation : i) les conduits sont organisés hiérarchiquement au sein des réseaux et ii) leur taille dépend de leur centralité ou de la distance à l'entrée ou la sortie la plus proche. Nous proposons de les analyser, notamment grâce à un algorithme de calcul de ladite hiérarchie au sein des réseaux karstiques. La mise en relation des propriétés

géométriques avec d'autres paramètres, notamment l'altitude, le pendage et l'orientation des conduits, est également présentée.

La première partie de ce chapitre présente la base de données utilisée, les métriques analysées et la représentation des réseaux par des graphes. Dans la seconde partie, je décris les méthodes utilisées pour analyser les métriques étudiées, avec en premier les méthodes d'analyse univariées : la comparaison des différents réseaux, l'analyse de la variabilité spatiale à travers l'assortativité et de nouvelles méthodes variographiques, puis l'analyse des distributions individuelles. Je décris ensuite les méthodes d'analyse multivariées : l'impact de la hiérarchie des conduits au sein des réseaux à travers l'étude de la centralité des nœuds, de la distance aux entrées et sorties les plus proches et d'une nouvelle méthode de hiérarchisation, et enfin, l'impact de l'altitude, de l'orientation et du pendage des conduits. Je présente ensuite les résultats de cette étude et explique la pertinence et les limitations des méthodes employées.

2.2 Données

2.2.1 Base de données utilisée

Cette étude statistique a été réalisée sur 49 réseaux karstiques différents. Les données ont été partagées par différents spéléologues dans le cadre de deux études antérieures (Collon et al., 2017; Jouvès et al., 2017), fusionnées au cours de cette thèse, et complétées par quelques nouveaux réseaux. Le nom et la localisation des réseaux sont donnés en Annexe A.

Les réseaux initialement utilisés par Jouvès et al. [2017] sont divisés en plusieurs groupes correspondant à différents motifs morphologiques et donc à des processus spéléogénétiques différents. Différentes parties d'un même réseau peuvent correspondre à des motifs différents mais la majorité des réseaux étudiés ne sont associés qu'à un seul motif. Tous les résultats présentés par la suite ont, sauf mention du contraire, été obtenus sur les réseaux entiers, regroupant leurs différentes parties.

Les données ont été acquises dans les réseaux par des spéléologues. Ces mesures sont réalisées en différents points appelés stations, les spéléologues mesurant la distance, l'orientation et le pendage entre les stations voisines (Figures 2.1; 2.2). Avec ces informations, il est possible d'établir la position des points en coordonnées cartésiennes et ainsi de reconstruire les réseaux en 3 dimensions. L'échantillonnage est généralement effectué lors de grands changements de direction des conduits, menant à des mesures ponctuelles et irrégulières. Les distances moyennes d'échantillonnage sont malgré tout relativement similaires d'un réseau à l'autre et nous considérons que ces approximations se compensent globalement, étant donné la manière dont les données sont interprétées. Suivant la façon dont il est effectué, l'échantillonnage peut conduire à des erreurs de représentation du réseau (Figure 2.3). Bien que cela ne soit pas systématique, les spéléologues mesurent souvent la hauteur et la largeur de la section du conduit à chaque station, voire parfois les 4 dimensions principales pour être plus précis : droite, gauche, bas et haut. Ainsi, cette description ne prend pas en compte la géométrie réelle des sections de conduits mais seulement leurs dimensions principales. Ces dimensions sont généralement estimées visuellement même si les récentes technologies laser, notamment le LIDAR, laissent espérer des données plus précises à l'avenir (e.g., Sadier, 2013;

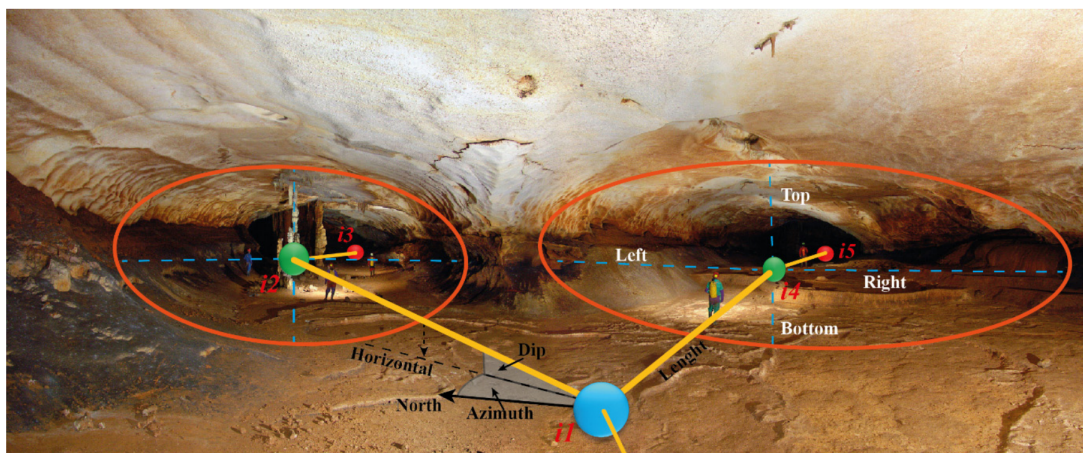


Figure 2.1 – Exemple de données numériques recueillies dans un réseau. Photographie : V. Ferrer. Grotte de Saint Marcel, France. Les sphères de couleur correspondent aux stations, tandis que les lignes jaunes représentent les visées entre elles. Ces visées sont associées à des mesures de distance, d’azimut et de pendage. Les mesures des hauteurs et largeurs à chaque station permettent d’obtenir une section approximative, représentée en orange, bien qu’elle ne tienne pas parfaitement compte des morphologies réelles (Jouves et al., 2017).

Schiller et Renard, 2016).

Les données fournies par les spéléologues ne sont cependant pas utilisables telles quelles et doivent être uniformisées en termes de format pour être analysées. Deux formats, *.xtb* ou *.tro*, compatibles respectivement avec les logiciels GHTopo¹ et Visual Topo² nous ont été fournis par les spéléologues. Les fichiers au format *.xtb* sont ouverts dans GHTopo pour être convertis au format *.tro*. Les données au format *.tro* sont toutes chargées dans Visual Topo, puis sont vérifiées individuellement. En effet, pour certains réseaux, les mesures d’orientation et de pendage ne suivent pas les mêmes conventions, ce qui nécessite des changements manuels. La position des points en coordonnées cartésiennes est ensuite calculée avec ce même logiciel. Les résultats sont alors exportés sous forme d’un fichier *.dxf*, représentant le squelette du réseau, et sous la forme d’un tableau Excel, lui-même converti en fichier *.txt*, pour conserver les informations concernant leur géométrie. Le fichier *.dxf* est importé dans le logiciel SKUA-Gocad³, ce qui permet d’obtenir le squelette du réseau en question. Nous créons ensuite des points de données en important le fichier *.txt*, puis les informations géométriques associées à ces points sont transférées sur le squelette pour obtenir une polyline avec des propriétés de dimensions associées. Si les valeurs de rayon équivalent et de ratio largeur-hauteur ne sont pas directement présentes, elles sont calculées à partir des largeurs et des hauteurs via l’usage d’un script. Gocad permet d’observer les réseaux en 3D de manière plus précise que Visual Topo. Les réseaux sont alors nettoyés semi-automatiquement pour corriger les erreurs de mesures (Collon et al., 2017) : par exemple, les stations présentes en double sont retirées ou d’autres sont reliées (Figure 2.3A,B,D). Les réseaux, associés aux valeurs de rayon équivalent et ratio WH, sont ensuite exportés en format *.pl*. Ces

¹<https://ghtopo.blog4ever.com>

²<http://vtopo.free.fr/>

³<https://www.pdgm.com/products/skua-gocad>

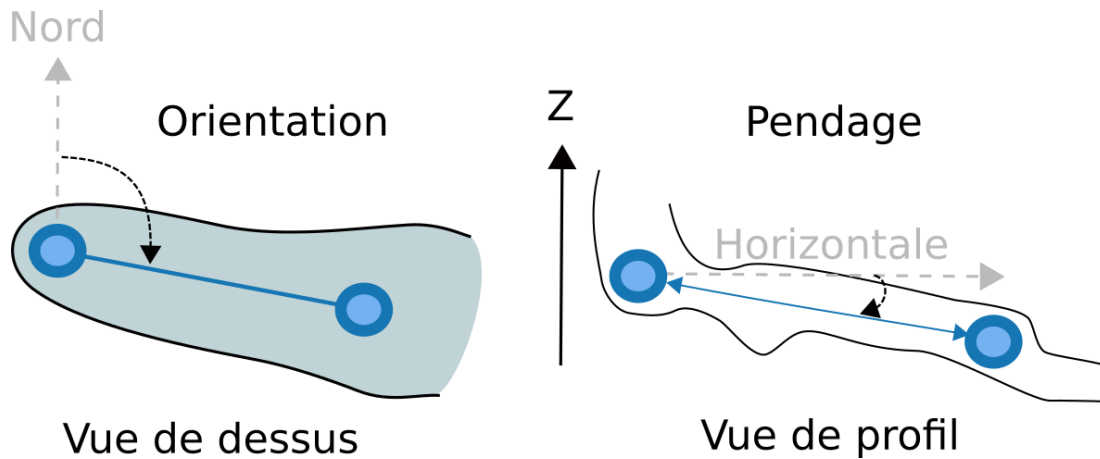


Figure 2.2 – Exemple de mesure de terrain d’orientation et de pendage (modifiée d’après Collon et al., 2017).

fichiers sont convertis via un script python en deux fichiers *.dat*, l’un correspondant aux liens entre les nœuds, l’autre aux coordonnées des nœuds et aux valeurs des propriétés associées. Ces fichiers sont enfin importés dans Matlab et convertis en graphes, les propriétés géométriques étant stockées sur les nœuds. Les réseaux sont alors utilisables pour l’étude statistique présentée par la suite.

Des bases de données initiales, seuls les réseaux comportant les informations géométriques d’intérêt ont été conservés ici. La conversion des données et surtout le processus de nettoyage ayant été repris depuis les données brutes, il existe ponctuellement des différences avec les réseaux présentés dans Collon et al. [2017] (notamment certaines parties sans informations ou parties isolées qui ont pu être supprimées).

L’extension des réseaux varie beaucoup. Le plus grand, Sieben Hengste (Suisse), s’étend sur plus de 80 kilomètres et est associé à 15340 points de données, tandis que le plus petit, Baume Galinière (France), comporte moins de 200 mètres de conduits et n’est associé qu’à 50 points de données. La majorité des réseaux étudiés sont relativement petits, la longueur médiane étant de 2135 mètres, et l’échantillonnage médian de 269 points. La distance d’échantillonnage moyenne est d’un point tous les 7.5 mètres. Il convient de noter que l’échantillonnage n’est pas homogène et que des parties entières de certains réseaux ne sont pas associées à des informations géométriques. L’échantillonnage est de plus couplé à de grandes incertitudes, la distance entre les stations étant généralement arrondie au décimètre tandis que le pendage et l’orientation sont arrondis au degré près. Les dimensions mesurées sont souvent arrondies au décimètre lorsqu’elles sont petites mais l’arrondi peut se faire au multiple de 0.5m ou au mètre le plus proche lorsqu’elles sont plus importantes. Ces informations sont présentées plus en détail en Annexe A.

2.2.2 Métriques utilisées

Les métriques telles que la longueur des conduits, le pendage, l’orientation, la largeur ou la hauteur ont déjà été étudiées par de nombreux auteurs (e.g., Jeannin, 1996; Loucks, 1999; Frumkin et Fischhendler, 2005; Fournillon et al., 2012; Collon et al., 2017). D’autres études ont eu pour but de caractériser l’étendue des réseaux dans leur entièreté,

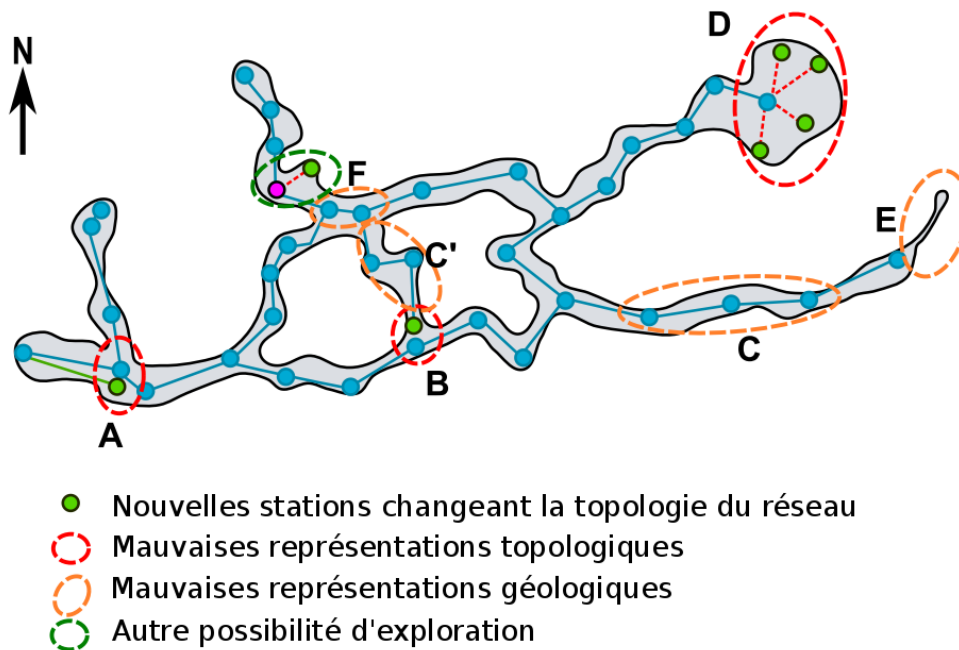


Figure 2.3 – Exemple d’erreurs liées au processus d’acquisition de données. Les cas A et B sont détectés et corrigés automatiquement lors du traitement des données. A : Station et ligne de visée additionnelles résultant d’un retour en arrière. Le nœud vert et le segment n’auraient pas dû être pris en compte. B : Un trou apparaît entre le premier et le dernier nœud du cycle, résultant en un cycle non fermé. C et C’ : Impact de la localisation de l’échantillonnage sur la sinuosité apparente du conduit exploré. Un échantillonnage régulier et un positionnement systématique des stations au milieu des conduits permettraient de le limiter. D : Configuration en *Rayons de Soleil* résultant d’un choix spécifique d’échantillonnage d’une grande salle. E : Certaines parties du réseau n’ont pas été explorées à cause d’un manque d’accessibilité. F : L’addition d’une nouvelle station peut changer la topologie locale du réseau (modifiée d’après Collon et al., 2017).

sans se focaliser sur les dimensions des conduits (e.g., Frumkin et Fischhendler, 2005; Piccini, 2011; Hendrick et Renard, 2016b). Plus récemment, Collon et al. [2017] ont cherché à caractériser la topologie des réseaux via des métriques adaptées. Néanmoins, seules de rares études sont directement centrées sur l’analyse des dimensions des conduits (Pardo-Iguzquiza et al., 2011; Jouvès et al., 2017). Dans notre optique de simulation de réseaux *réalistes*, nous avons choisi de nous focaliser principalement sur l’étude de ces dimensions.

La taille des conduits est souvent définie par le rayon ou le diamètre de leur section. Plusieurs auteurs utilisent des valeurs de rayon constantes pour réaliser des simulations d’écoulement sur des réseaux de conduits discrets. Ils peuvent utiliser une unique valeur de rayon pour tous les conduits (e.g, Ronayne, 2013), des valeurs différentes pour chaque conduit (e.g., Chen et Goldscheider, 2014), voire des valeurs dépendant d’une loi de puissance liée à la position hiérarchique du conduit dans le réseau : de fortes valeurs dans les conduits supposés recevoir les flux hydrauliques les plus importants, de faibles valeurs dans ceux recevant des flux plus faibles et un élargissement à chaque

confluence (Vuilleumier et al., 2013; Borghi et al., 2016). Le rayon ou les paramètres contrôlant sa distribution spatiale sont parfois calibrés directement durant les simulations d'écoulement afin de reproduire les débits observés en différents points du réseau (e.g., en ses sorties). Dans ces travaux, les valeurs de rayon sont généralement définies comme constantes dans chaque conduit (e.g., Vuilleumier et al., 2013; Jeannin et al., 2015). La variabilité du rayon au sein du réseau, mais surtout au sein d'un conduit est de fait rarement prise en compte. Dans la réalité, les conduits des réseaux karstiques peuvent présenter une grande diversité de tailles et de formes.

Plutôt qu'utiliser directement les valeurs de hauteur et de largeur présentes dans la base de données, nous avons choisi de travailler avec le rayon équivalent du conduit (Re_i) car il s'agit d'un paramètre important lors des simulations d'écoulement. Il est défini à une station i à partir de sa largeur W_i :

$$W_i = R_i + L_i \quad (2.1)$$

et de sa hauteur H_i :

$$H_i = U_i + D_i \quad (2.2)$$

R_i , L_i , U_i et D_i étant respectivement les distances vers la droite, la gauche, le haut et le bas depuis la station. La forme réelle des conduits étant trop complexe pour être représentée simplement, il a été choisi de se baser sur une approximation elliptique (Figure 2.4). Chaque section de conduit est assimilée à une ellipse ayant la même hauteur et la même largeur et le rayon équivalent est défini comme le rayon d'un cercle de même aire que cette ellipse :

$$Re_i = \sqrt{\frac{W_i H_i}{4}} \quad (2.3)$$

Cette approximation permet de prendre directement en compte les largeurs et hauteurs des conduits tout en respectant leur forme usuelle. Pour quantifier l'anisotropie de forme, nous utilisons le ratio largeur-hauteur, désigné également par WH pour *Width-Height ratio* (e.g., Pardo-Iguzquiza et al., 2011; Jouves et al., 2017) (Figure 2.5) :

$$WH_i = \frac{W_i}{H_i} \quad (2.4)$$

2.2.3 Représentation des réseaux via des graphes

Dans le cadre de l'analyse statistique, nous représentons les réseaux sous forme de graphes (Collon et al., 2017). Ils correspondent aux squelettes des réseaux et plus particulièrement aux conduits observés et mesurés par les spéléologues. Chaque nœud coïncide avec une station de mesure et porte les informations de rayon équivalent et de ratio largeur-hauteur (Figure 2.6). Les nœuds sont directement reliés par des arêtes (*edges*) portant comme information les longueurs, pendages et orientations des conduits. Comme la direction des écoulements au sein des réseaux n'est pas connue, nous utilisons des graphes non-orientés. Il s'agit de graphes utilisant des arêtes bidirectionnelles, signifiant que, s'il y a un lien direct entre un nœud A et un nœud B, il y a également un

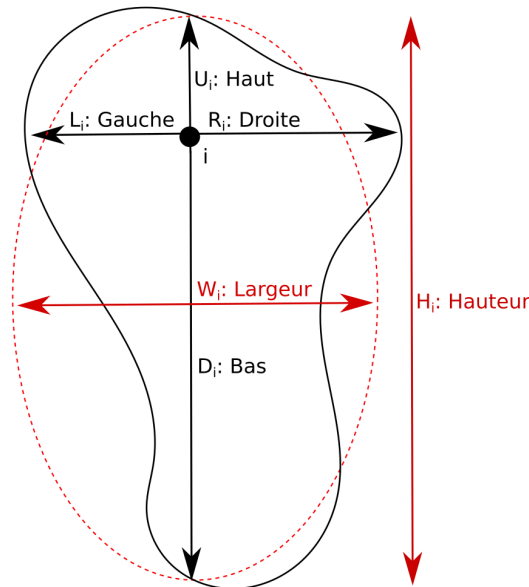


Figure 2.4 – Mesure des dimensions du conduit à une station i du réseau. L'approximation elliptique est montrée en rouge.

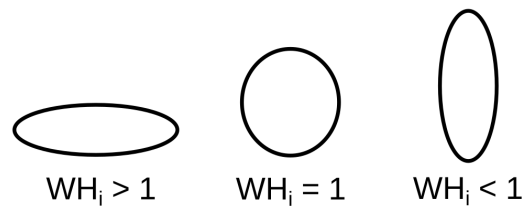


Figure 2.5 – Schématisation du ratio largeur-hauteur.

lien direct entre le nœud B et le nœud A. Inversement, dans le cas de graphes orientés, même si un lien direct existe entre un nœud A et un nœud B, il n'y aura pas forcément de lien entre le nœud B et le nœud A.

On appelle degré des nœuds le nombre de voisins qu'a chacun d'entre eux. Ainsi, si un nœud est relié directement à 3 autres nœuds, son degré sera égal à 3. Les extrémités sont définies comme les nœuds de degré 1, tandis que les intersections sont les nœuds de degré 3 ou plus. Il est possible de diviser les réseaux en plusieurs parties appelées branches (Figure 2.7). Une branche est définie comme un ensemble d'arêtes adjacentes connectant deux nœuds, de degré différent de 2, de manière similaire à Collon et al. [2017] et Jouvès et al. [2017]. Elles correspondent aux parties situées entre deux nœuds qui sont des intersections de conduits et/ou des extrémités du réseau incluant les intersections et les extrémités elles-mêmes.

2.3 Méthodes analytiques

Sont présentées ici les méthodes d'analyse statistique utilisées dans le cadre de cette étude. Elles sont divisées en deux principales catégories : les analyses univariées, portant

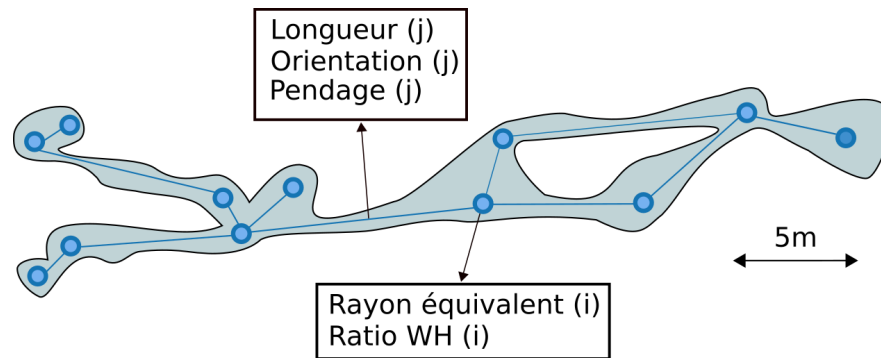


Figure 2.6 – Exemple de graphe correspondant à un réseau karstique. Il s’agit de nœuds reliés via des arêtes, ce qui correspond à une représentation simplifiée du réseau. Les nœuds et les arêtes peuvent porter des informations : dans le cas présent, les nœuds portent les informations de rayon équivalent et de ratio largeur-hauteur, tandis que les arêtes sont couplées aux longueurs, orientations et pendages des segments correspondant.

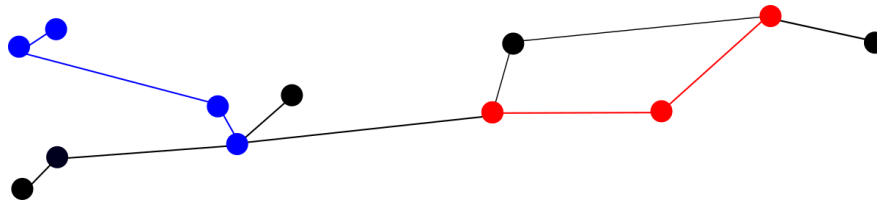


Figure 2.7 – Exemple de branches dans un réseau. Parmi les sept branches de ce réseau, une est représentée en bleu et une autre en rouge.

uniquement sur les métriques étudiées et les analyses multivariées mettant en relation les métriques étudiées avec d’autres paramètres.

2.3.1 Analyses univariées

2.3.1.a Comparaison des différents réseaux

Les distributions des métriques étudiées ont été comparées dans les différents réseaux afin de mettre en avant de potentielles similarités. Ces analyses statistiques auraient pu prendre en compte la longueur des segments associés aux différents points de mesure pour pouvoir les pondérer, mais nous avons choisi de ne pas procéder ainsi. En effet, les mesures sont très ponctuelles et rien n’indique qu’elles s’appliquent au segment amont ou aval. De plus, la distance moyenne entre deux stations voisines reste relativement similaire d’un réseau à un autre. Nous avons également effectué une analyse statistique selon les différents motifs morphologiques identifiés par Jouvès et al. [2017].

En plus de ces analyses quantitatives, nous avons aussi procédé à des tests statistiques, également appelés tests d’hypothèse. Différents tests ont été utilisés pour analyser les similarités entre les réseaux, à la fois pour le rayon équivalent et pour le ratio largeur-hauteur. Bien que leurs fonctionnements diffèrent sur de nombreux points, la démarche reste identique. Les tests calculent, via différents processus, une valeur appelée valeur- p (p -value ou encore *probabilité critique*). Cette valeur est ensuite comparée

avec une valeur α choisie par l'utilisateur, α étant le risque de première espèce, aussi appelé signification statistique. α correspond au risque de rejeter l'hypothèse testée H_0 alors qu'elle est vraie. Si la valeur-p est inférieure au α choisi, alors l'hypothèse H_0 est rejetée en faveur de l'hypothèse alternative H_1 . Dans le cas contraire, l'hypothèse H_0 est dite non-rejetée (ou *acceptée*), ce qui ne signifie pas pour autant qu'elle soit validée. Les valeurs de α sont généralement de 0.01, 0.05, voire 0.1, correspondant respectivement à une chance de rejet de 1, 5 et 10 % de l'hypothèse H_0 alors qu'elle est vraie.

Nous utilisons les tests suivants avec différentes valeurs d' α :

- Le test des rangs signés de Wilcoxon, pour vérifier si les distributions de tous les réseaux pourraient être des échantillons d'une distribution ayant une médiane spécifique choisie par l'utilisateur (Wilcoxon, 1945; Hollander et Wolfe, 1999; Gibbons et Chakraborti, 2003). Ici nous utilisons la médiane des médianes de tous les réseaux.
- Le test de la somme des rangs de Wilcoxon (qui est équivalent au test U de Mann-Whitney), pour comparer les médianes des réseaux deux à deux afin de vérifier si deux distributions pourraient être de possibles échantillons de distributions ayant la même médiane (Wilcoxon, 1945; Hollander et Wolfe, 1999; Gibbons et Chakraborti, 2003).
- Le test de Kolmogorov-Smirnov à double échantillon, afin de vérifier si chaque paire de réseaux est potentiellement issue d'une même distribution (Massey, 1951; Miller, 1956; Marsaglia et al., 2003).
- Le test de Kruskal-Wallis, pour vérifier si, pour chaque paire de réseaux, ceux-ci sont potentiellement des échantillons issus de la même distribution (Kruskal et Wallis, 1952; Gibbons et Chakraborti, 2003). Il s'agit d'une extension du test de la somme des rangs de Wilcoxon.

2.3.1.b Quantification de la variabilité spatiale

Il est fréquent d'observer des changements de dimension brusques au sein d'un conduit karstique, mais également de remarquer une certaine constance sur des étendues limitées. Pour vérifier si les données sont spatialement corrélées, nous avons calculé l'assortativité des dimensions des conduits et développé des méthodes variographiques adaptées aux réseaux karstiques.

Assortativité

L'assortativité est définie comme la tendance qu'ont les nœuds d'un réseau à se connecter à des nœuds de valeur similaire. Elle est généralement calculée sur le degré des nœuds (e.g., Newman, 2002; Collon et al., 2017) mais peut être utilisée pour étudier la corrélation d'autres paramètres (e.g., Noldus et Mieghem, 2015). Pour estimer la corrélation entre les nœuds voisins dans les graphes correspondant aux réseaux, nous avons calculé les valeurs d'assortativité des métriques étudiées. Nous nous sommes servis du coefficient de corrélation de Pearson appliqué à toutes les arêtes du graphe

(Newman, 2002) :

$$r = \frac{\frac{1}{N_e} \sum_{e=1}^{N_e} Z(i_e)Z(j_e) - [\frac{1}{N_e} \sum_{e=1}^{N_e} \frac{1}{2}(Z(i_e) + Z(j_e))]^2}{\frac{1}{N_e} \sum_{e=1}^{N_e} \frac{1}{2}(Z(i_e)^2 + Z(j_e)^2) - [\frac{1}{N_e} \sum_{e=1}^{N_e} \frac{1}{2}(Z(i_e) + Z(j_e))]^2} \quad (2.5)$$

où N_e est le nombre d'arêtes dont les deux nœuds sont associés à une valeur, $Z(i_e)$ la valeur à la première extrémité de l'arête e et $Z(j_e)$ la valeur à la deuxième extrémité de l'arête e . Seules les arêtes dont les deux extrémités sont associées à des valeurs sont prises en compte lors de ce calcul.

Une valeur de r positive signifie que le réseau est assortatif : Les nœuds de grande valeur sont généralement reliés à des nœuds de grande valeur. Une valeur de r nulle signifie que le réseau n'est pas assortatif. Une valeur de r négative signifie que le réseau est disassortatif : les nœuds de grande valeur sont généralement reliés à des nœuds de faible valeur. Les valeurs de r sont comprises entre -1 et 1 . Plus r est proche de 1 , plus le réseau est assortatif, et inversement lorsqu'il est proche de -1 .

Bien qu'elle soit intéressante pour les relations entre voisins directs, l'assortativité reste très limitée quant à la distance considérée. La variographie, elle, permet l'analyse de la variabilité spatiale sur de grandes distances.

Variogrammes

Les variogrammes sont des fonctions mathématiques décrivant la variabilité spatiale d'une variable aléatoire donnée. Ils peuvent être omnidirectionnels ou calculés dans une direction spécifique. Le variogramme expérimental d'une fonction aléatoire Z est calculé de la manière suivante :

$$\gamma_e(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad (2.6)$$

où h est une distance et $N(h)$ le nombre de paires de nœuds distants de h .

Les variogrammes peuvent être calculés dans un espace à une, deux ou trois dimensions. Ils sont généralement calculés dans un espace tridimensionnel cartésien XYZ mais il est souvent nécessaire de les calculer dans un espace paramétrique UVW permettant de quantifier la corrélation spatiale le long de structures géologiques (Mallet, 2002). En se servant de variogrammes calculés dans un système de coordonnées cartésiennes, Pardo-Iguzquiza et al. [2011] ont montré que les largeurs, hauteurs et formes des conduits sont corrélées spatialement avec une longueur de corrélation d'approximativement 40 mètres. Nous avons choisi d'approfondir cette analyse en caractérisant la variabilité spatiale le long des réseaux, qui selon nous, ne doivent pas être simplement assimilés à des espaces uni- ou tridimensionnels. Nous utilisons ainsi des variogrammes 1D-curvilinéaires. Des approches relativement similaires ont déjà été proposées dans le cas des rivières (Ver Hoef et al., 2006; de Fouquet, 2006). Les variogrammes utilisés durant cette thèse sont basés sur la plus courte distance curvilinéaire entre chaque paire de nœuds et sont donc indépendants d'une quelconque direction spécifique. L'algorithme de Dijkstra (Dijkstra, 1959) est utilisé pour le calcul des distances. Les variogrammes obtenus de la sorte sont nommés variogrammes globaux.

Visuellement, la variabilité spatiale des métriques étudiées semble plus grande à

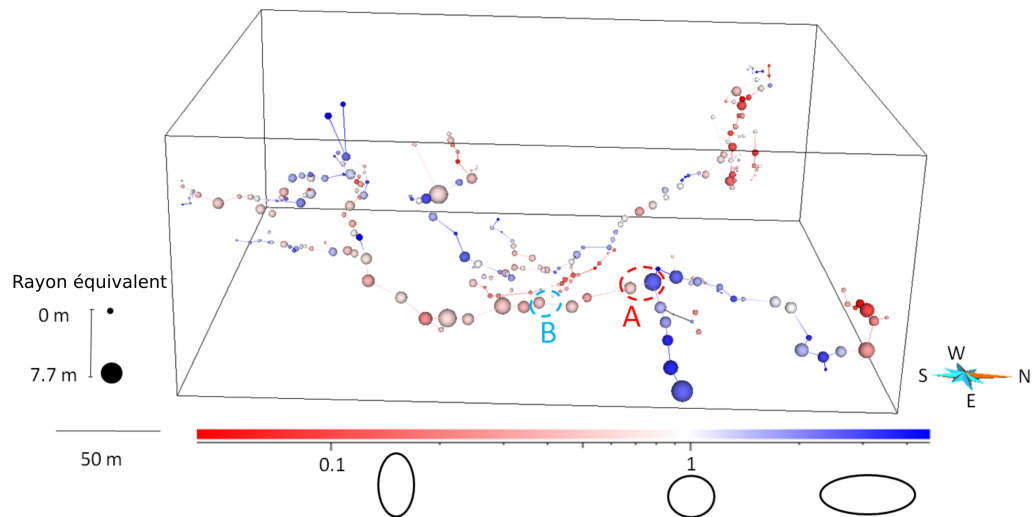


Figure 2.8 – Représentation du réseau de l'Aspirateur (269 nœuds). La taille des nœuds est proportionnelle à leur rayon équivalent tandis que leur couleur représente leur ratio largeur-hauteur. Des différences notables entre des nœuds voisins peuvent être observées au niveau des intersections. A : Différence de ratio largeur-hauteur. B : Différence de rayon équivalent.

l'intérieur des branches elles-mêmes. De plus, des différences importantes peuvent être observées à certaines des intersections entre les différents conduits (Figure 2.8). Pour analyser ce comportement, nous définissons un variogramme intrabranché, qui utilise uniquement les paires de nœuds localisés dans une même branche (Figure 2.9C). De la même manière, nous définissons un variogramme interbranché, utilisant les paires de nœuds situés dans des branches différentes (Figure 2.9D).

2.3.1.c Distributions statistiques

Pour mettre en évidence des comportements spécifiques, les distributions des métriques sont comparées à des lois de distribution spécifiques, notamment les lois log-normale ou de Pareto. Différents tests ont été employés à cet effet. Ces tests sont conçus pour des données non corrélées et ont ainsi été effectués après l'analyse variographique sus-citée, pour pouvoir les appliquer à des sous-échantillons dont les nœuds sont suffisamment distants afin de minimiser l'impact de la corrélation spatiale.

- Le test de validité de l'ajustement du χ^2 (Pearson, 1900; Gibbons et Chakraborti, 2003) et le test de Kolmogorov-Smirnov à simple échantillon (Massey, 1951; Miller, 1956; Marsaglia et al., 2003) avec $\alpha = 5\%$ permettent de vérifier si les réseaux suivent individuellement une loi de Pareto.
- Le test de validité de l'ajustement du χ^2 (Pearson, 1900; Gibbons et Chakraborti, 2003), le test de Kolmogorov-Smirnov à simple échantillon (Massey, 1951; Miller, 1956; Marsaglia et al., 2003), le test de Lilliefors (Lilliefors, 1967; Lilliefors, 1969; Conover, 1999), le test d'Anderson-Darling (Anderson et Darling, 1952; Anderson et Darling, 1954) et le test de Jarque-Bera (Jarque et Bera, 1987) avec $\alpha = 5\%$ permettent de vérifier si le logarithme décimal des métriques suit une loi normale.

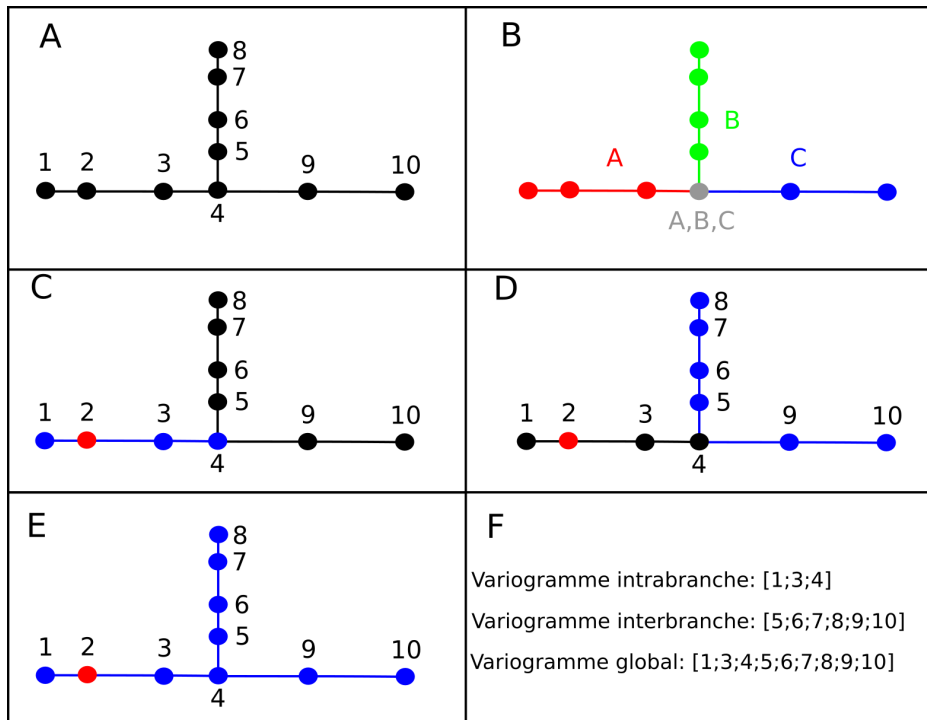


Figure 2.9 – A : Exemple de réseau. B : Les trois branches de ce réseau, chacune représentée par une couleur différente. Le nœud d’intersection (en gris) appartient aux trois branches. C : Nœuds (bleus) appariés avec le nœud 2 (rouge) durant le calcul du variogramme intrabranche. D : Nœuds (bleus) appariés avec le nœud 2 (rouge) durant le calcul du variogramme interbranche. E : Nœuds (bleus) appariés avec le nœud 2 (rouge) durant le calcul du variogramme global. F : Liste des nœuds appariés avec le nœud 2 en C, D et E.

2.3.2 Analyses multivariées

Les analyses multivariées permettent de mettre en évidence d’éventuels liens entre plusieurs métriques.

2.3.2.a Hiérarchisation des réseaux

Dans le but d’affecter des valeurs de rayon équivalent dans un réseau en s’appuyant sur un nombre minimal de paramètres, certains travaux sont partis du principe que ce dernier peut être exprimé sous forme d’une loi de puissance dépendant du niveau hiérarchique du conduit correspondant à l’intérieur du réseau (Vuilleumier et al., 2013; Borghi et al., 2016). Cette hypothèse se base sur l’idée que la taille des conduits est proportionnelle au flux les traversant et qu’en général, un conduit localisé en aval sera plus large qu’un conduit localisé en amont (Figure 2.10A, B). Dans ces articles, le niveau hiérarchique est estimé en prenant en compte la taille de la zone drainée pour chaque entrée du réseau et se propage en aval. Nous vérifions ici cette supposition que la taille des conduits varie ainsi le long du réseau en étudiant des données réelles. N’ayant cependant pas accès aux informations concernant les flux traversant les conduits, 3 moyens de mettre en avant une potentielle hiérarchisation sont proposés.

Centralité des nœuds

On peut supposer que les zones les plus centrales des réseaux sont associées à des flux plus importants et que cela influence d'une quelconque manière les dimensions des conduits. De nombreux moyens de définir la centralité des nœuds d'un graphe existant, nous nous sommes limités aux 4 suivants :

- La centralité de degré, qui correspond au degré des nœuds du graphe. Elle est complétée par l'étude du ratio intersections/extrémités.
- La centralité de proximité, définie comme l'inverse de la somme des longueurs des chemins les plus courts entre le nœud considéré et tous les autres nœuds du réseau.
- La centralité d'intermédiarité, correspondant au nombre de chemins les plus courts entre chaque paire de nœuds passant par ce nœud en particulier.
- La centralité de vecteur propre, partant du principe que les connexions aux nœuds de forte valeur ont plus d'impact sur la valeur du nœud considéré que celles aux nœuds de faible valeur. Les nœuds de forte valeur sont ainsi connectés à d'autres nœuds de forte valeur.

Ces différentes définitions de la centralité offrent des résultats assez différents. Nous explorons alors leurs éventuelles relations avec les métriques étudiées.

Distance aux entrées et sorties les plus proches

La notion de flux convergents au sein des réseaux nous incite à penser que les entrées, associées à des flux plus faibles, ont des tailles plus petites que les sorties, associées à des flux plus importants. Nous proposons donc d'étudier la relation entre les dimensions des conduits et la distance des nœuds considérés par rapport à l'entrée ou la sortie la plus proche. Pour simplifier l'algorithme, tous les nœuds de degré 1 (nœuds d'extrémité) sont considérés comme des entrées ou des sorties en fonction de la différence d'altitude entre leur position et celle du nœud d'intersection le plus proche au sein du graphe. Si le nœud d'extrémité a une altitude supérieure à celle du nœud d'intersection, il est considéré comme une entrée puisque l'eau descendra le long du réseau sous l'effet de la gravité, et ce, quelle que soit sa position dans le réseau. Autrement, l'extrémité est considérée comme une sortie. Pour respecter l'organisation des réseaux, nous avons utilisé la distance curvilinéaire le long du réseau au lieu d'une distance euclidienne.

Nouvelle méthode de hiérarchisation

Pour permettre une analyse applicable à tous les réseaux, il est nécessaire de définir la hiérarchie de manière simple et reproductible. Pour cette raison, nous avons développé une méthode de hiérarchisation capable de prendre en compte la séparation des flux ainsi que les boucles observées dans les réseaux considérés et qui ne requiert aucune information quant aux surfaces drainées par les entrées (Figure 2.10C, D). Les valeurs hiérarchiques se propagent depuis les entrées situées dans la partie supérieure du réseau jusqu'aux sorties situées plus bas, similairement à un fluide s'écoulant dans le réseau sous l'action de la gravité (en accord avec le mode de formation épigénique de la majorité de réseaux).

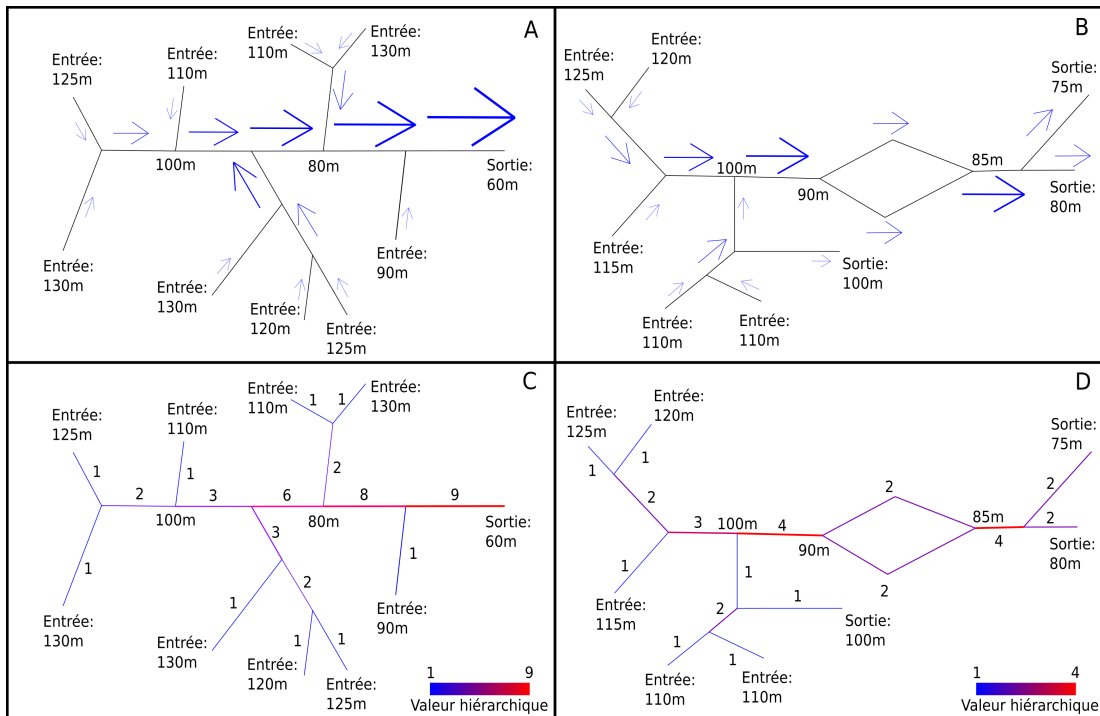


Figure 2.10 – A : Exemple d'un réseau arborescent pouvant correspondre à un réseau fluvial ou karstique. Les flèches bleues montrent la direction du flux et leurs tailles sont proportionnelles aux valeurs de flux théoriques. B : Exemple de réseau complexe, avec une boucle et trois sorties, pouvant correspondre à un réseau karstique. C-D : Méthode de hiérarchisation proposée (vue du dessus), appliquée respectivement aux réseaux A et B.

Quand deux conduits de valeur hiérarchique n_A et n_B se rejoignent, le conduit résultant obtient une valeur de $n_A + n_B$. Quand un conduit de valeur n se divise en deux, les deux conduits ainsi formés sont associés à une valeur de $n/2$, ce qui prend en compte la séparation du flux (Figure 2.10B). Si le *flux* est bloqué dans un nœud de basse altitude (dans un siphon), nous le forçons à remonter pour atteindre des nœuds n'ayant pas encore été atteints. Cette remontée est cependant très compliquée à définir et n'est donc pas implémentée entièrement. Le présent algorithme n'est donc adapté que pour des petits réseaux, présentant peu de siphons. Une fois que les valeurs hiérarchiques des conduits sont calculées, il est possible de chercher une corrélation avec leurs dimensions.

2.3.2.b Autres relations

La zonation verticale des karsts pouvant avoir un impact potentiel sur la géométrie des conduits, nous avons cherché à étudier le lien entre les altitudes des nœuds et les dimensions des conduits.

Comme les réseaux se forment préférentiellement le long de réseaux de fractures préexistantes, les pendages et orientations des conduits sont également comparés avec leurs propriétés géométriques. Il est en effet possible de s'attendre à différentes géométries de conduits en fonction des familles de fractures associées. Les calculs ne sont pas directement effectués sur les conduits mais sur des paires de nœuds voisins. L'orientation

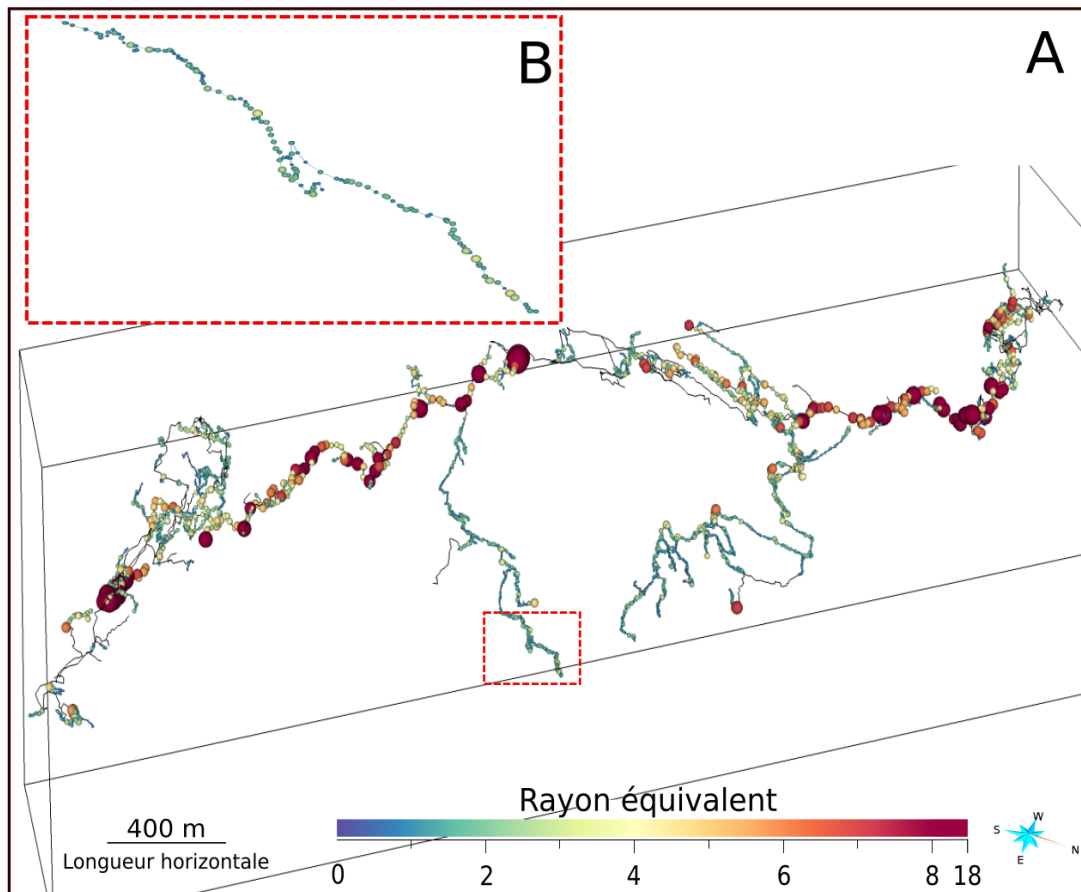


Figure 2.11 – Rayon équivalent du réseau d'Arrestelia dans le réseau entier (A) et dans une petite partie (B). La taille des nœuds est directement proportionnelle à leur valeur.

et le pendage sont directement calculés sur les arêtes, qui peuvent être de longueur variable, tandis que les rayons équivalents et ratios largeur-hauteur considérés sont la moyenne des valeurs des deux nœuds considérés. Si un des deux nœuds n'a pas de valeur associée, le segment n'est pas pris en compte.

2.4 Résultats

De nombreux résultats individuels sont présentés sur le réseau d'Arrestelia⁴, dont les données ont été traitées pour ne garder que les principales parties du réseau dans lesquelles les dimensions des conduits sont connues. Ainsi, après prétraitement, le réseau étudié fait 58 kilomètres de long et est représenté par 6000 nœuds bien que seulement 4283 d'entre eux soient associés à une valeur de rayon équivalent et de ratio largeur-hauteur (Figure 2.11).

⁴Pyrénées-Atlantiques, France, données fournies par J.P. Cassou

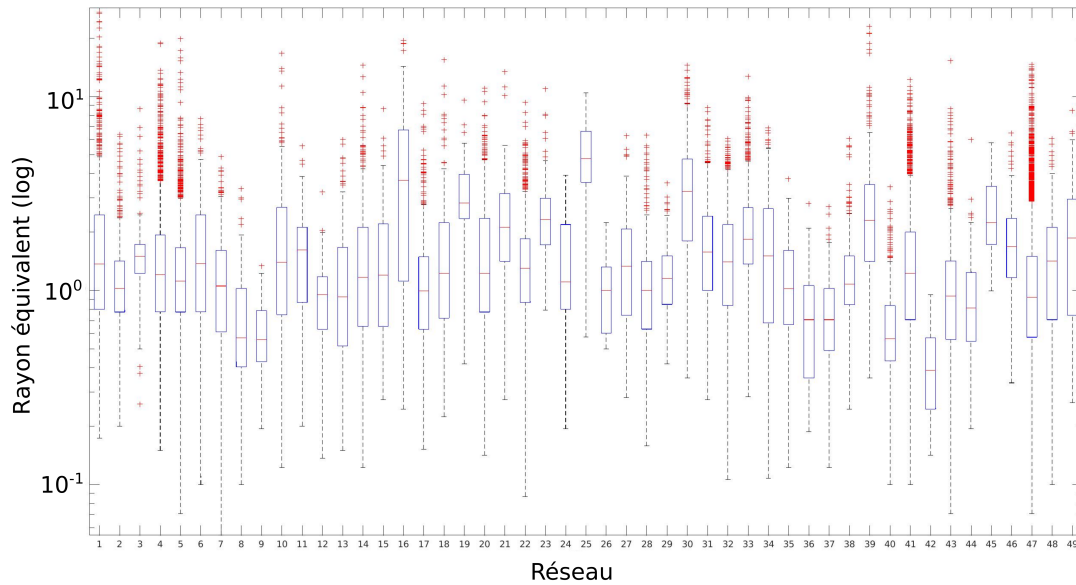


Figure 2.12 – Diagramme en boîte du rayon équivalent dans les différents réseaux. Nous utilisons une échelle logarithmique pour l’axe vertical. Le numéro des réseaux est le même que celui indiqué en Annexe A. Les boîtes correspondent à l’intervalle entre le premier et le troisième quartile, tandis que le trait rouge correspond à la valeur médiane. La zone entre les deux extrémités des moustaches représente environ 99% des valeurs, les autres étant des *valeurs aberrantes*.

2.4.1 Analyses univariées

2.4.1.a Comparaison des différents réseaux

Nous avons dans un premier temps comparé les distributions de tous les réseaux. La majorité d’entre eux tendent à avoir des ordres de magnitude similaires pour le rayon équivalent (Figure 2.12) et le ratio largeur-hauteur (Figure 2.13) mais des différences notables restent observables d’un réseau à l’autre. L’inspection visuelle de tous les histogrammes révèle une forte asymétrie des distributions pour tous les réseaux, ainsi que des formes de distribution relativement similaires lors de l’utilisation d’une échelle logarithmique.

La moyenne du rayon équivalent pondérée par le nombre de points dont les informations géométriques sont connues dans chaque réseau est égale à 1.52 mètres, tandis que celle du ratio largeur-hauteur est égale à 1.10. La médiane des médianes du rayon équivalent des différents réseaux est égale à 1.22 mètres, tandis que celle du ratio largeur-hauteur est égale à 1.00. Beaucoup de réseaux de grande taille sont cependant associés à des valeurs médianes de ratio largeur-hauteur inférieures, voire très inférieures à 1. Les résultats statistiques sont détaillés dans l’Annexe A. Un résumé est présenté dans le Tableau 2.1.

Les tests statistiques sont effectués sur des échantillons complets, dans le cadre d’une étude aveugle. Les tests des rangs signés de Wilcoxon sont rejetés dans presque tous les cas, à la fois pour le rayon équivalent et le ratio largeur-hauteur, exception faite des réseaux avec des médianes extrêmement proches de la médiane des médianes. Les

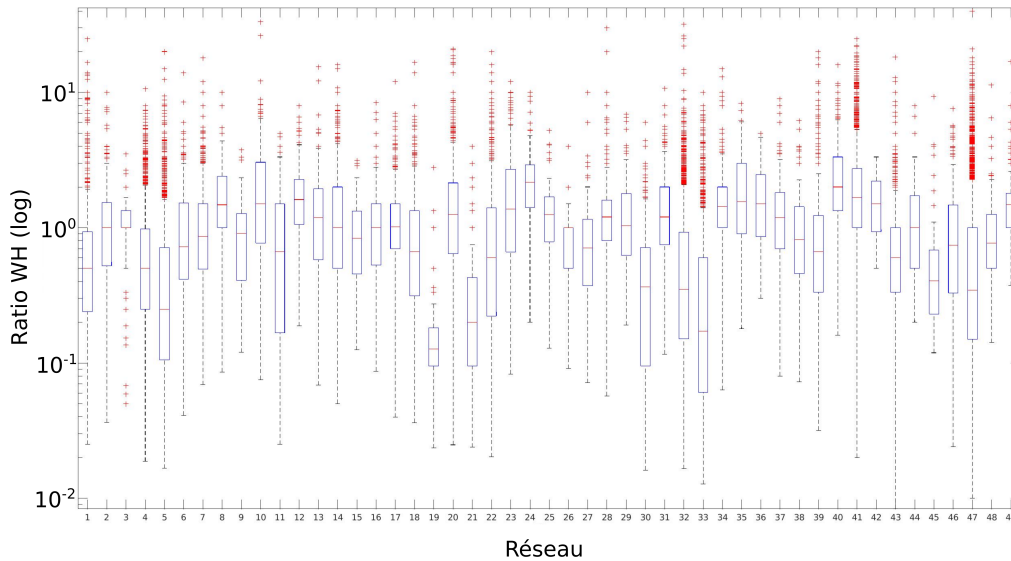


Figure 2.13 – Diagramme en boîte du ratio largeur-hauteur dans les différents réseaux. Nous utilisons une échelle logarithmique pour l’axe vertical.

résultats des autres tests sont résumés dans le Tableau 2.2 pour le rayon équivalent et le Tableau 2.3 pour le ratio largeur-hauteur.

Comme presque tous les tests sont rejetés, et au vu des différences observables entre les réseaux (Figures 2.12 et 2.13), nous supposons qu’il n’y a pas de distribution générique, à la fois pour le rayon équivalent et pour le ratio largeur-hauteur. Il serait donc préférable d’utiliser, si possible, des données acquises spécifiquement sur le système étudié pour définir une distribution de référence du réseau à modéliser.

Un aperçu des résultats obtenus sur les motifs morphologiques est présenté en Annexe B. Comme certains réseaux étudiés par Jouvès et al. [2017] n’avaient pas les informations géométriques nécessaires, le nombre de réseaux associés à chaque motif peut être légèrement différent. Au vu du faible nombre de réseaux associés à chaque motif, ces résultats ne doivent pas être surinterprétés.

2.4.1.b Quantification de la variabilité spatiale

Assortativité

Le calcul des coefficients de corrélation de Pearson du rayon équivalent et du ratio WH dans les différents réseaux montre que les réseaux sont assortatifs vis-à-vis de ces deux métriques (Tableau 2.4 et Annexe A). Les valeurs obtenues pour le rayon équivalent démontrent une assortativité assez importante, avec une valeur médiane de 0.69. Celles du ratio largeur-hauteur sont dans l’ensemble plus faibles, avec une valeur médiane de 0.45 mais témoignent malgré tout d’une assortativité non négligeable. Toutes les valeurs obtenues sont positives et, pour le ratio largeur-hauteur, seuls trois réseaux présentent une valeur de coefficient de corrélation de Pearson inférieure à 0.2.

Ces résultats démontrent une certaine continuité des métriques à faible distance. Ils ne sont cependant obtenus que sur des voisins directs et sont de fait très fortement impactés par les distances d’échantillonnage.

	Nombre de nœuds	Nombre de nœuds sans valeur	Rayon moyen	Rayon médian	Écart type rayon
Maximum	15340	1717	5.02	4.80	4.37
Minimum	45	0	0.42	0.39	0.21
Moyenne	1050	102	1.76	1.44	1.31
Moyenne pondérée	/	333	1.52	1.20	1.24
Médiane	269	16	1.62	1.22	1.14
Écart Type	2378	265	0.94	0.82	0.81

	Ratio WH moyen	Ratio WH médian	Écart Type ratio WH	Longueur Totale (m)	Distance moyenne entre les nœuds (m)
Maximum	2.77	2.16	3.39	82238	12.2
Minimum	0.25	0.13	0.46	189	3.2
Moyenne	1.32	0.97	1.34	7757	7.5
Moyenne pondérée	1.10	0.69	1.34	40834	7.3
Médiane	1.21	1.00	1.29	2135	7.7
Écart type	0.54	0.48	0.55	15370	2.5

Tableau 2.1 – Résumé de l'analyse statistique des données brutes des 49 réseaux différents. La moyenne pondérée l'est par rapport au nombre de points dont les informations géométriques sont connues dans chaque réseau.

α	5%	10%	20%
Somme des rangs de Wilcoxon	0.84	0.86	0.90
Kruskal-Wallis	0.84	0.87	0.90
Kolmogorov-Smirnov à double échantillon	0.91	0.93	0.95

Tableau 2.2 – Proportion de tests de similarité des réseaux rejetés pour le rayon équivalent en fonction du risque de première espèce.

α	5%	10%	20%
Somme des rangs de Wilcoxon	0.84	0.86	0.89
Kruskal-Wallis	0.84	0.86	0.89
Kolmogorov-Smirnov à double échantillon	0.89	0.92	0.94

Tableau 2.3 – Proportion de tests de similarité des réseaux rejetés pour le ratio largeur-hauteur en fonction du risque de première espèce.

	Coefficient de corrélation de Pearson du rayon équivalent	Coefficient de corrélation de Pearson du ratio WH
Maximum	0.91	0.86
Minimum	0.28	0
Moyenne	0.67	0.46
Moyenne pondérée	0.72	0.52
Médiane	0.69	0.45
Écart Type	0.15	0.18

Tableau 2.4 – Résumé de l'analyse statistique de l'assortativité des 49 réseaux différents.

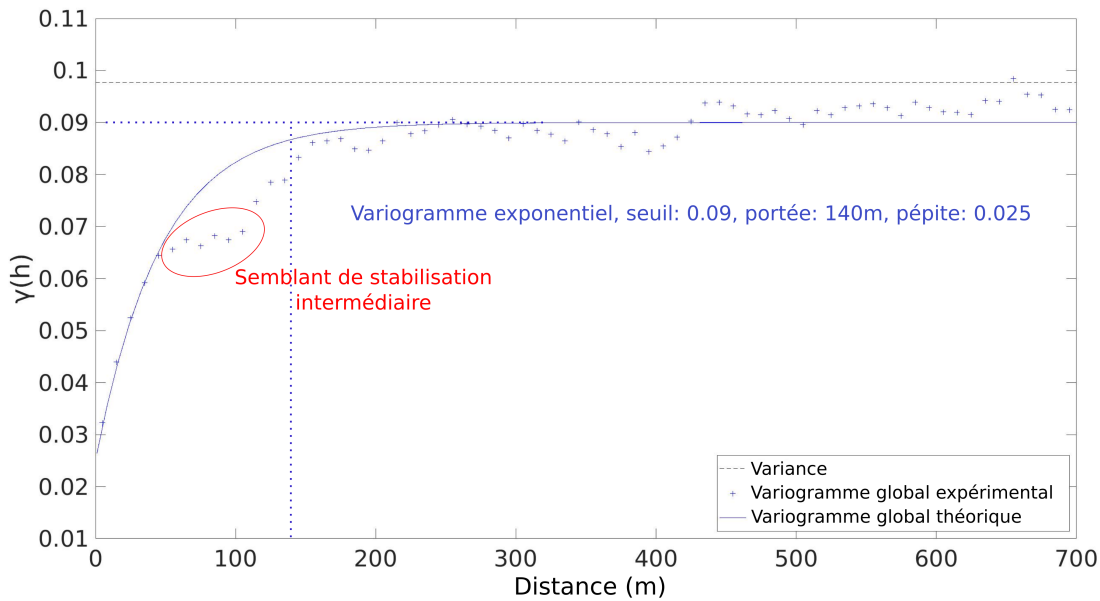


Figure 2.14 – Variogramme global expérimental et variogramme global théorique du logarithme décimal du rayon équivalent du réseau d’Arrestelia (4283 points de données). La distance affichée est limitée à 700 mètres mais la distance maximale est proche de 2000 mètres.

Variogrammes

L’étude variographique a permis de bien caractériser la variabilité spatiale sur l’ensemble des réseaux. Quand les réseaux ont suffisamment de nœuds (plus de 100), la majorité des variogrammes expérimentaux du logarithme décimal du rayon équivalent ont un comportement proche de celui montré en Figure 2.14. Dans le cas contraire, les variogrammes sont bien souvent trop erratiques pour pouvoir être interprétés. Les portées et seuils des variogrammes diffèrent d’un réseau à l’autre mais les valeurs de γ sont généralement basses à faibles distance, démontrant une bonne corrélation spatiale entre les nœuds proches. Il y a cependant un effet pépité, inévitable au vu de l’échantillonnage habituel (un point de données tous les 7.5 mètres en moyenne dans les réseaux étudiés). Après une augmentation rapide, les valeurs des variogrammes expérimentaux oscillent autour de la valeur du seuil, ce qui signifie que les métriques deviennent spatialement indépendantes. Passée une certaine distance, les valeurs deviennent généralement erratiques, à cause du faible nombre de points de données. Dans la majorité des cas, des modèles de variogrammes théoriques exponentiels représentent correctement les variogrammes expérimentaux. Nous avons également observé une stabilisation intermédiaire de plusieurs variogrammes. Cela pourrait être modélisé via des variogrammes imbriqués, combinaisons de plusieurs modèles de variogramme. Cela peut être interprété comme deux échelles de variabilité, ce qui est probablement une conséquence directe de la complexité des réseaux. Néanmoins, pour ne pas trop complexifier le problème, nous avons décidé de n’utiliser que des variogrammes théoriques simples. Les portées des variogrammes globaux des plus grands réseaux (plus de 1000 points) sont présentés dans le Tableau 2.5.

Tandis que l’on observe une faible variabilité spatiale du rayon et du ratio largeur-hauteur dans les réseaux, la variabilité des métriques est encore plus basse à l’intérieur

Nom	Nombre de points ayant une valeur	Portée rayon équivalent (m)	Portée ratio WH (m)
Abisso Chimera	999	50	50
Arphidia Robinet	1425	500	100
Arrestelia	4283	150	100
Autran	1514	100	100
Charentais Heche	1741	100	200
Foussoubie	2406	200	100
Malaval	2077	100	250
Mazo Chico	1847	200	500
Saint Marcel	3801	200	400
Sakany	1677	50	50
Sieben Hengste	15144	500	500

Tableau 2.5 – Résumé de l’analyse statistique des variogrammes globaux des 11 plus grands réseaux. Les portées indiquées sont arrondies au multiple de 50 le plus proche pour simplifier l’analyse.

des branches : le variogramme intrabranche a des valeurs plus faibles que le variogramme global (Figure 2.15). Inversement, deux nœuds localisés dans des branches différentes ont une probabilité non négligeable d’avoir des valeurs plus éloignées que deux nœuds situés dans la même branche et ce, même à faible distance. Les valeurs des variogrammes interbranches sont plus grandes que celles des variogrammes globaux à faible distance. Cela met en évidence l’importante variabilité spatiale existant entre les différentes branches. Comme elles ont, de par leur définition, des tailles très limitées, les valeurs des variogrammes intrabranches deviennent rapidement erratiques à cause du manque de paires de données. Finalement, le variogramme interbranche commence à se superposer au variogramme global passé une certaine distance, dépendant plus ou moins de la taille du réseau.

Ces résultats mettent en évidence la corrélation spatiale des dimensions des conduits au sein des réseaux et de leurs branches et constituent une information à intégrer à la simulation de ces paramètres.

2.4.1.c Distributions statistiques

Nous avons également vérifié si les distributions des métriques suivaient une loi de distribution spécifique. A l’échelle logarithmique, les histogrammes obtenus pour la majorité des réseaux de grande taille semblent proches d’une distribution normale (Figure 2.16), à la fois pour le rayon et pour le ratio largeur-hauteur. Les conduits étroits tendent à être sous-échantillonnés et il est relativement compliqué de réaliser des mesures dans des conduits verticaux. Cela mène à une sous-estimation des petites valeurs, à la fois pour le rayon équivalent et pour le ratio largeur-hauteur. Cela contribue à l’aspect log-normal des distributions et mène à une représentation légèrement trompeuse à l’échelle linéaire. Ainsi, après avoir tronqué les distributions au-dessus des plus petites valeurs dans le but de limiter ce biais d’échantillonnage, elles semblent suivre une loi de Pareto, ou du moins une loi de puissance, à l’échelle linéaire (Figure 2.17).

L’analyse variographique ayant démontré l’existence d’une corrélation spatiale à courte distance, les tests statistiques d’ajustement à une loi ne peuvent être réalisés sur les données brutes. Afin de s’affranchir de cette corrélation, un sous-échantillonnage des

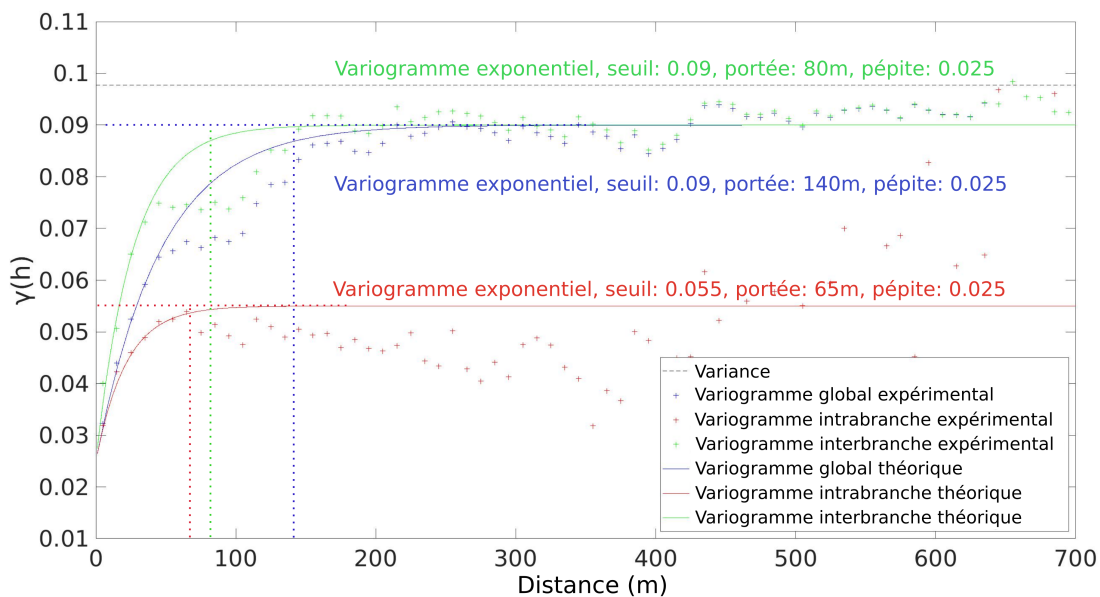


Figure 2.15 – Les 3 variogrammes expérimentaux du rayon équivalent du logarithme décimal du réseau d’Arrestelia (4283 points de données), avec les variogrammes théoriques correspondant. Après 350 mètres, les valeurs du variogramme intrabranche deviennent trop erratiques pour être considérées comme valides.

	Rayon équivalent	Ratio WH
Validité de l’ajustement du χ^2	0.27	0.27
Kolmogorov-Smirnov à simple échantillon	0.55	0.18

Tableau 2.6 – Proportion de tests rejetés lors de la vérification de l’ajustement de la distribution à une loi de Pareto sur les 11 sous-échantillons non corrélés, pour $\alpha = 5\%$.

réseaux a donc été effectué de façon à ne conserver que les points séparés par une distance supérieure à la portée du variogramme correspondant. Seuls les 11 réseaux de plus de 1000 points de données ont été jugés assez grands pour permettre une analyse statistique pertinente après sous-échantillonnage, les portées de leurs variogrammes globaux étant présentées dans le Tableau 2.5. Le sous-échantillonnage est effectué à la fois pour le rayon équivalent et le ratio largeur-hauteur car les portées de leurs variogrammes sont habituellement différentes. Les sous-échantillons non corrélés obtenus de cette manière ont un nombre moyen de 53 points (les portées des variogrammes pouvant être importantes).

Les résultats indiquent que l’hypothèse de l’existence d’une loi de Pareto (Tableau 2.6) ou log-normale (Tableau 2.7) n’est généralement pas rejetée, à la fois pour le rayon équivalent et le ratio largeur-hauteur. Il faut néanmoins noter l’incompatibilité entre ces deux lois, le faible nombre de réseaux sous-échantillonnés et le faible nombre de valeurs associées. Une loi de puissance pourrait également être représentative de la distribution du rayon équivalent mais le manque de faibles valeurs dans nos données ne permet pas de le prouver.

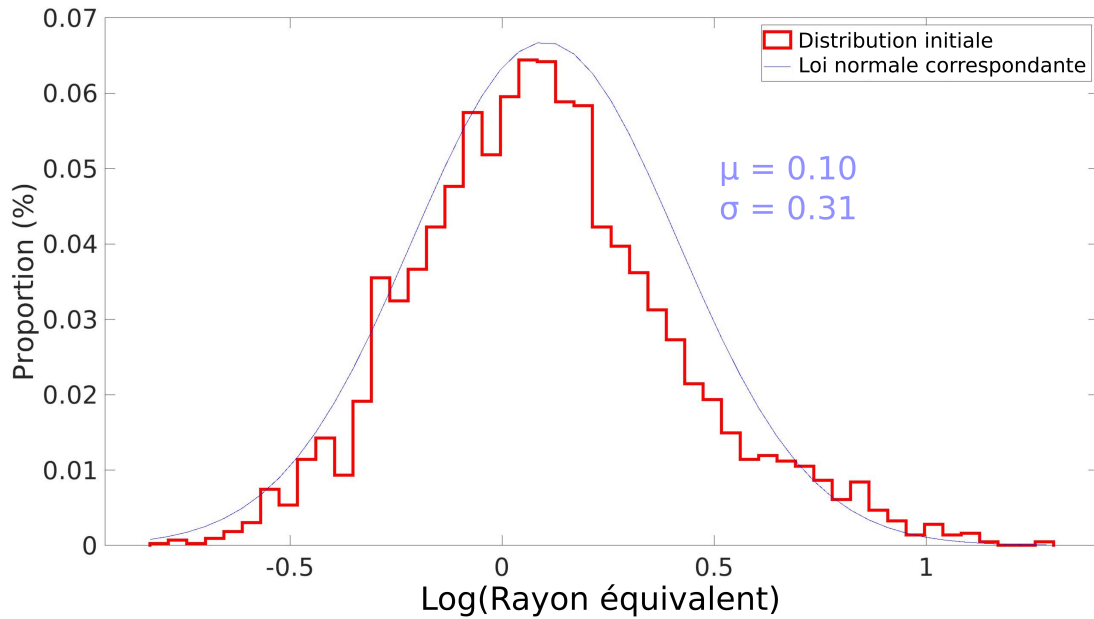


Figure 2.16 – Histogramme du logarithme décimal du rayon équivalent du réseau d'Arrestelia (4283 points de données) et sa loi normale ajustée.

	Rayon équivalent	Ratio WH
Validité de l'ajustement du χ^2	0.09	0.36
Kolmogorov-Smirnov à simple échantillon	1	0.82
Lilliefors	0	0.27
Anderson-Darling	0.09	0.36
Jarque-Bera	0.09	0.36

Tableau 2.7 – Proportion de tests rejetés lors de la vérification de l'ajustement de la distribution à une loi log-normale sur les 11 sous-échantillons non corrélés, pour $\alpha = 5\%$.

2.4.2 Analyses multivariées

2.4.2.a Hiérarchisation des réseaux

Centralité des nœuds

Les valeurs de centralité diffèrent fortement d'une définition à une autre (Figure 2.18) et la mise en relation avec le rayon équivalent et le ratio largeur-hauteur en est fortement impactée (Figure 2.19). L'étude des diagrammes en boîte semble indiquer que seule la centralité de degré peut potentiellement avoir un réel impact sur le rayon équivalent. Les valeurs médianes du rayon équivalent semblent augmenter avec le degré des nœuds dans certains réseaux, le réseau d'Arrestelia étant un des exemples les plus marqués. Les valeurs médianes semblent rester indépendantes des centralités d'intermédierité et de proximité. Les résultats obtenus sur la centralité de vecteur propre sont quant à eux très erratiques, montrant également un manque de lien. Le ratio largeur-hauteur apparaît comme relativement indépendant des quatre types de centralité étudiés. Au vu de ces résultats généraux, seule la centralité de degré et son lien avec le rayon équivalent seront présentés plus en détail.

La distribution du degré des nœuds est relativement similaire d'un réseau à l'autre.

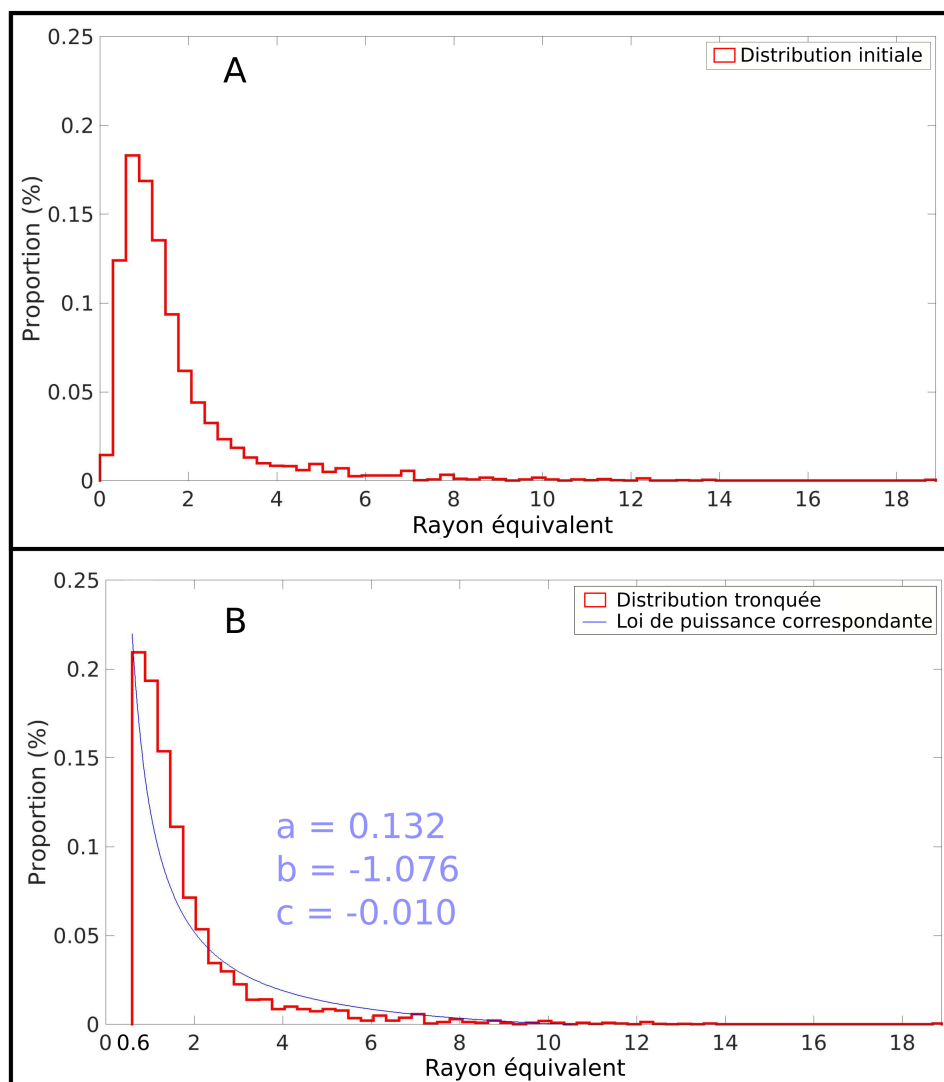


Figure 2.17 – A : Histogramme du rayon équivalent du réseau d’Arrestelia (4283 nœuds de données). B : Histogramme du rayon équivalent du réseau d’Arrestelia sans les valeurs inférieures à 0.6m (3573 nœuds de données). Une loi de puissance est ajustée à cette distribution tronquée : $f(x) = 0.132 * x^{-1.076} - 0.010$.

La grande majorité des nœuds est de degré 2, tandis que les nœuds d’extrémité (degré 1) et d’intersection (degré 3 ou plus) sont en proportions relativement similaires, bien que la proportion des nœuds d’intersection soit généralement plus importante dans les grands réseaux. Cette proportion est quantifiée à travers le ratio intersections/extrémités, qui aura son importance lors du Chapitre 4. Un résumé des résultats est présenté dans le Tableau 2.8. Les résultats détaillés sont présentés en Annexe C.

Dans beaucoup de réseaux, il semble y avoir une légère augmentation de la médiane et du troisième quartile avec le degré des nœuds pour le rayon équivalent, augmentation particulièrement marquée pour le réseau d’Arrestelia (Figure 2.19). Les plus grandes valeurs sont observables pour les nœuds de degrés 2, plus que pour les nœuds d’ordre supérieur (bien que les médianes associées soient plus grandes). Cela résulte en partie de la grande proportion de nœuds de degré 2 dans les réseaux par rapport aux nœuds

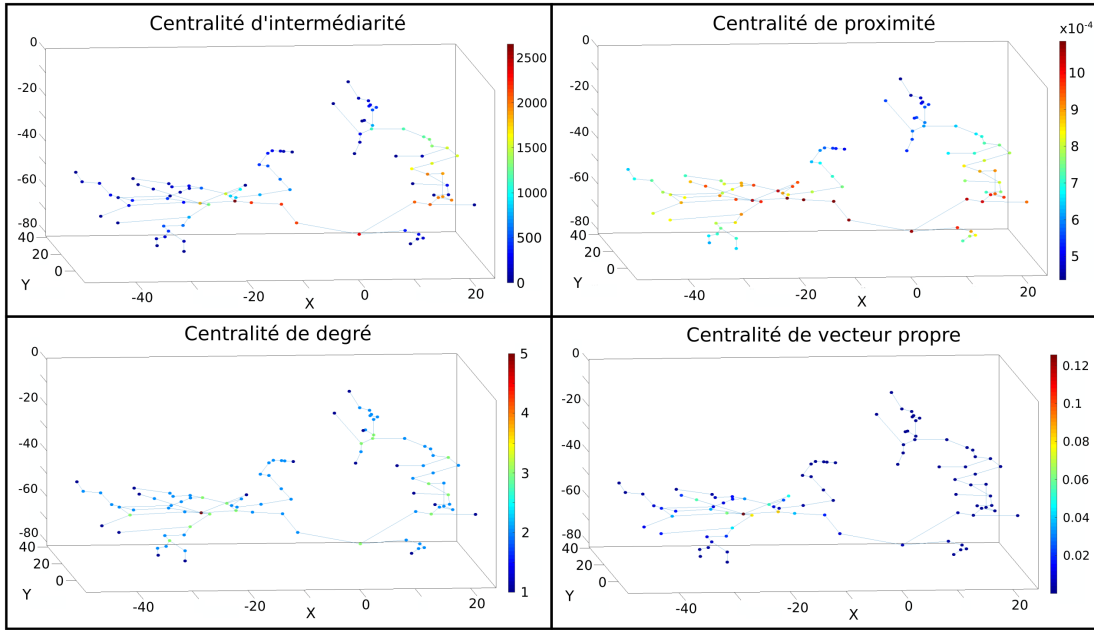


Figure 2.18 – Comparaison des valeurs des 4 centralités retenues au sein du réseau Everest (94 nœuds). La petite taille de ce réseau permet de bien illustrer les différences entre ces dernières.

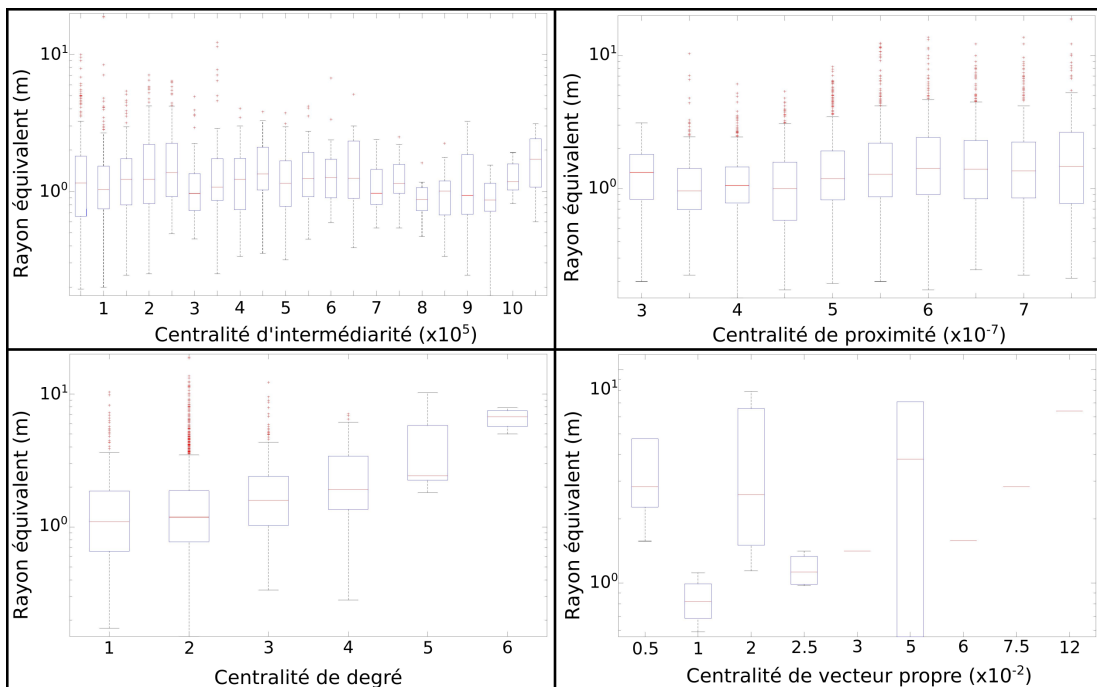


Figure 2.19 – Diagrammes en boîte du rayon équivalent en fonction des valeurs des différentes centralités des nœuds sur le réseau d'Arrestelia (4283 nœuds sont associés à une valeur de rayon).

	Nombre de nœuds	Proportion de nœuds de degré 2	Nombre d'extrémités	Nombre d'intersections	Ratio int/ext
Maximum	15340	0.98	843	1160	1.53
Minimum	45	0.4	2	0	0
Moyenne	1050	0.82	71.7	77.6	0.89
Moyenne pondérée	/	0.86	377.3	466	1.10
Médiane	269	0.87	23	28	0.93
Écart type	2378	0.13	140.9	177	0.26

Tableau 2.8 – Résumé de l'analyse statistique du degré des nœuds des 49 réseaux étudiés

d'intersection. Les nœuds de degré 1 ont généralement des valeurs plus basses, et ce de façon marquée, car ils correspondent souvent aux parties où les conduits deviennent trop étroits pour pouvoir être explorés par les spéléologues. Les grandes valeurs médianes des nœuds de degré supérieur à 2 peuvent également être liées à la difficulté d'effectuer des mesures aux nœuds d'intersection. Il y a pour certains réseaux une légère diminution du ratio largeur-hauteur conjointement à l'augmentation du degré des nœuds mais ce comportement est beaucoup moins systématique, empêchant de tirer de réelles conclusions à ce sujet.

Distance aux entrées et sorties les plus proches

Aucune relation directe entre le rayon équivalent des conduits et la distance à la sortie la plus proche n'a pu être mise en évidence. La Figure 2.20 montre le résultat obtenu pour le réseau d'Arrestelia. Les boîtes utilisées ont une largeur égale à 50 mètres mais les mêmes résultats sont également observables pour des largeurs moins grandes. Aucune relation n'a également été trouvée pour le ratio largeur-hauteur et il en va de même avec la distance aux entrées les plus proches.

Nouvelle méthode de hiérarchisation

Considérons à présent les résultats obtenus avec notre approche de modélisation. La Figure 2.21A montre un exemple de valeurs hiérarchiques calculées sur le réseau Everest. Ce réseau est en effet un des réseaux les plus simples présents dans notre base de données, avec suffisamment peu de siphons pour permettre une hiérarchisation cohérente et ce, bien qu'aucune information de flux ne soit disponible. La Figure 2.21B montre le diagramme en boîte du rayon équivalent pour différents niveaux hiérarchiques. Nous observons une légère augmentation du rayon pour les plus grandes valeurs hiérarchiques mais pas d'augmentation régulière. Les grandes valeurs ne concernent de plus que quelques rares points. Enfin, une simple analyse visuelle des réseaux (par exemple via la Figure 2.8) montre que le rayon varie d'une manière complexe dans l'espace car des grandes valeurs sont observables partout dans le réseau. Elles ne semblent pas augmenter de manière hiérarchique.

2.4.2.b Autres relations

Relation avec l'altitude des nœuds

Considérons maintenant les éventuelles relations entre l'altitude des nœuds et les propriétés géométriques associées. La Figure 2.22A montre la distribution des altitudes

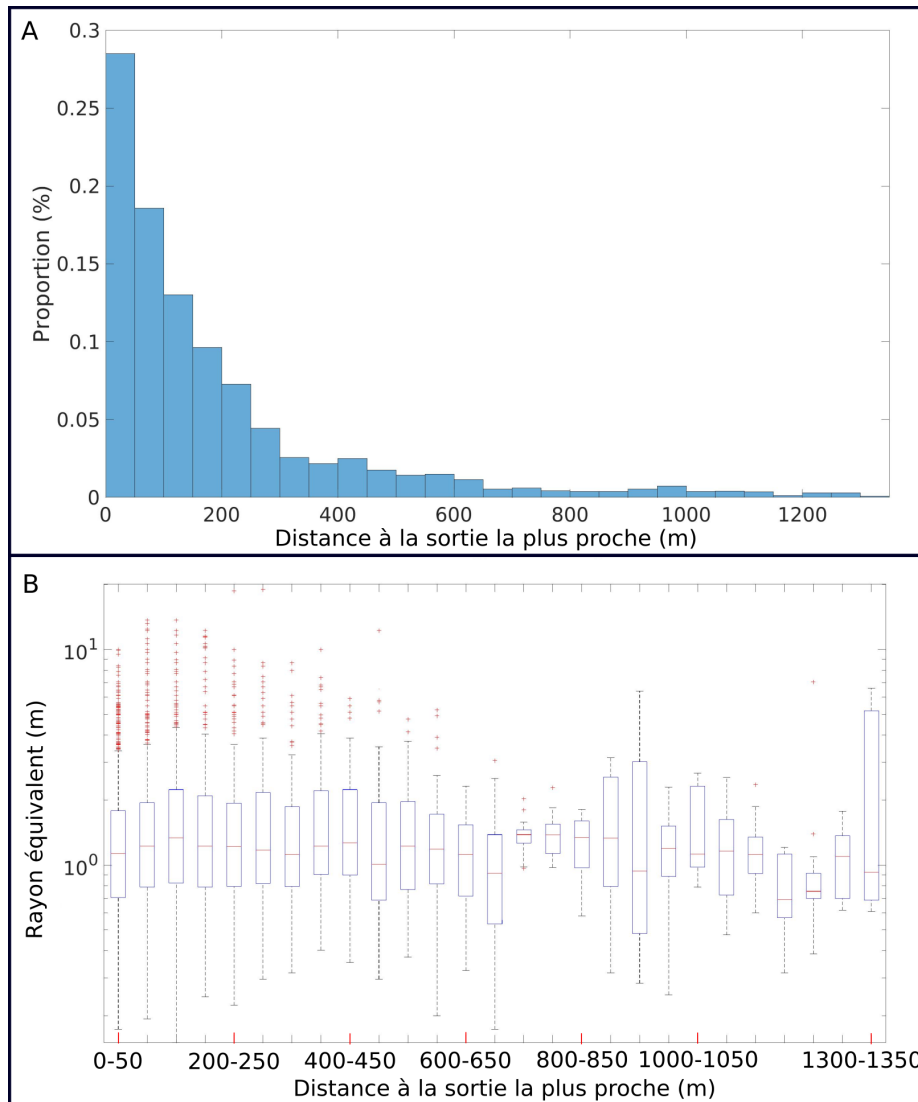


Figure 2.20 – A : Histogramme de la distance à la sortie la plus proche au sein du réseau d’Arrestelia (4283 nœuds). B : Diagramme en boîte du rayon équivalent en fonction de la distance à la sortie la plus proche au sein du réseau d’Arrestelia.

au sein du réseau d’Arrestelia, tandis que la Figure 2.22B montre le diagramme en boîte du rayon équivalent pour différentes altitudes. Cet exemple est représentatif des résultats que nous obtenons pour la majorité des réseaux karstiques étudiés : aucune relation claire avec les propriétés géométriques des conduits n’est observable dans les différents réseaux. Néanmoins, certains d’entre eux ont de plus faibles valeurs de rayon équivalent associées aux nœuds de haute altitude (e.g., 2.22B). C’est probablement parce que ces zones de haute altitude correspondent aux parties vadoses des réseaux, dans lesquelles les conduits tendent à être plus étroits (Jouves et al., 2017). Cela dépend cependant très fortement des cas et ne devrait pas être généralisé pour des réseaux inexplorés. De plus, des discussions avec des spéléologues nous ont mis en garde quant à la diversité des définitions de hauteur et largeur dans les conduits sub-verticaux, qui sont en général moins échantillonnés.

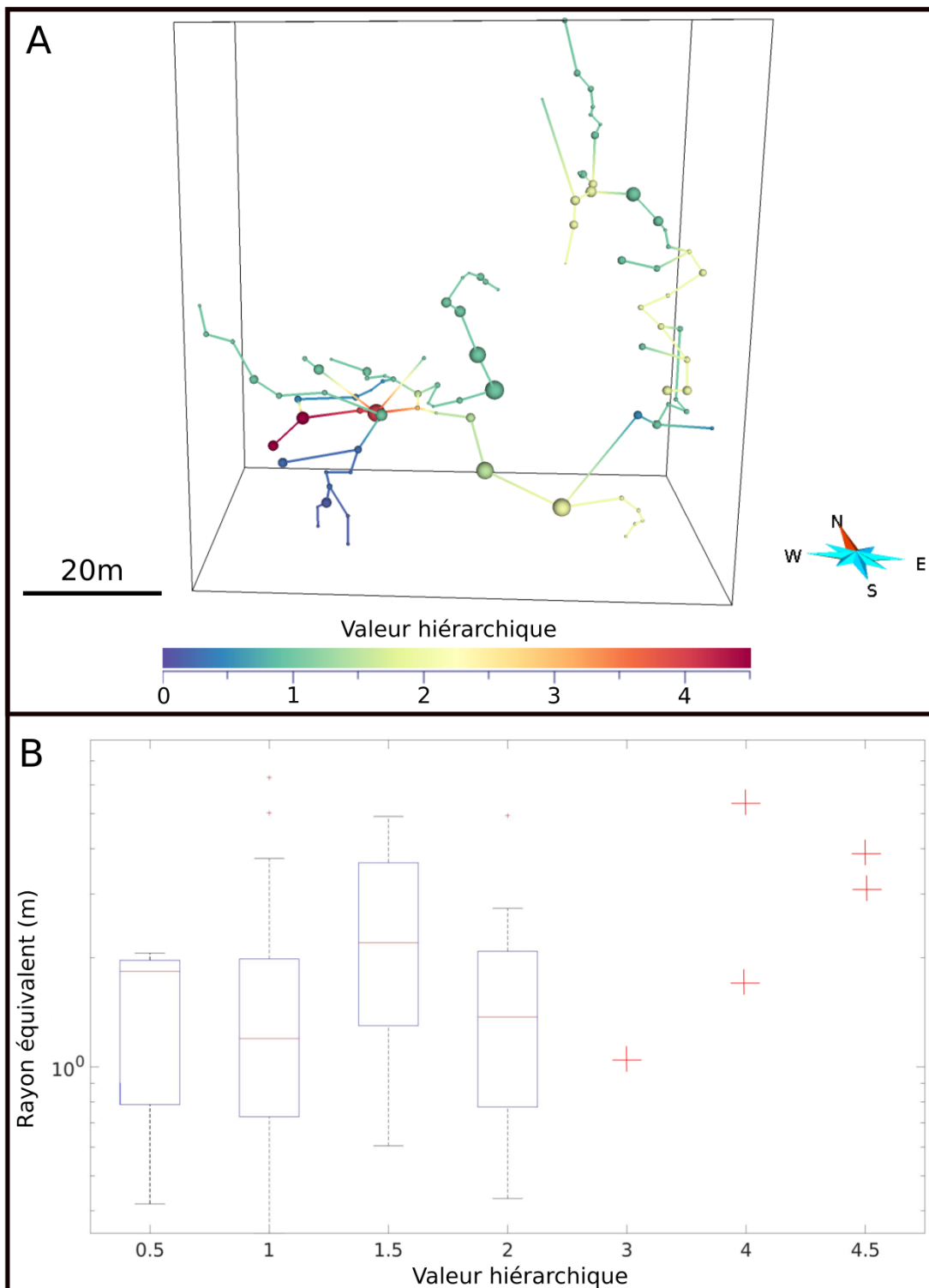


Figure 2.21 – A : Hiérarchisation des conduits au sein du réseau Everest (94 nœuds). La taille des nœuds est directement proportionnelle à la valeur de rayon équivalent, les valeurs allant de 0.28 à 6.29 mètres. B : Diagramme en boîte du rayon équivalent en fonction de la valeur hiérarchique des nœuds (92 nœuds ont une valeur de rayon).

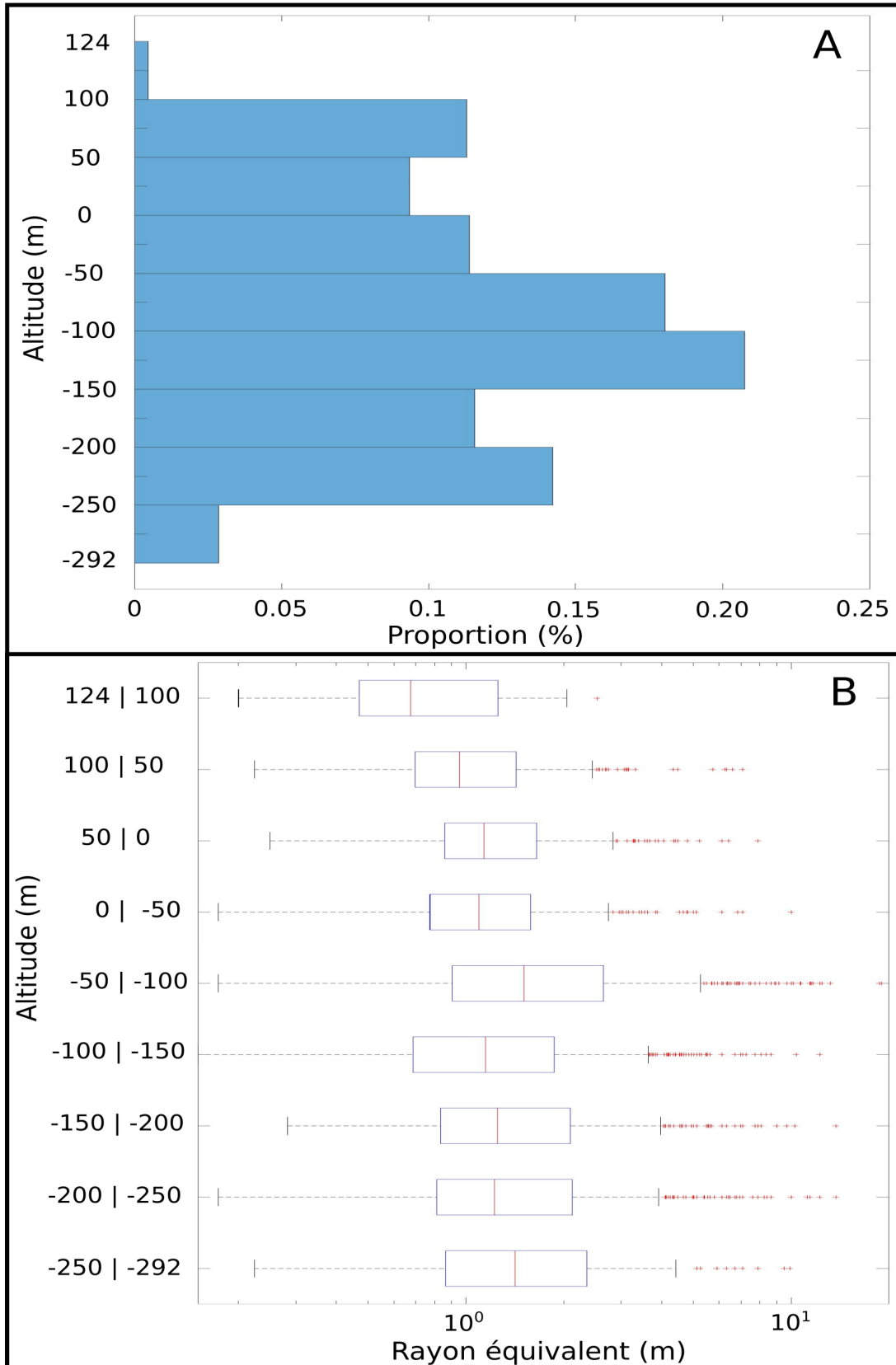


Figure 2.22 – A : Histogramme des altitudes des nœuds du réseau d'Arrestelia (4283 nœuds, seuls ceux associés à une valeur de rayon équivalent ont été pris en compte). B : Diagramme en boîte du rayon équivalent en fonction des valeurs d'altitude.

Relation avec les pendages et orientations des conduits

Bien qu'une relation entre les dimensions et le pendage ainsi que l'orientation des conduits soit observable dans certains réseaux (e.g., Figure 2.23), elle semble généralement négligeable dans la majorité d'entre eux, comme le réseau d'Arrestelia (Figure 2.24). En conséquence, aucune relation générique ne peut être généralisée pour l'ensemble des réseaux. Il semblerait néanmoins que de telles relations soient identifiables dans certains cas particuliers, probablement lorsque des fractures préexistantes ont été fortement impliquées lors du processus spéléogénétique. Ces relations doivent néanmoins être démontrées spécifiquement dans le contexte étudié car dans de nombreux cas, bien qu'un développement dans certaines directions préférentielles soit visible, aucun lien réel n'est observable (e.g., Figure 2.24).

2.5 Discussion

Cette analyse statistique a été effectuée sur une base de données forte de 49 réseaux karstiques. Aucune étude similaire n'a, à notre connaissance, été réalisée sur autant de réseaux. Exception faite de différences mineures, principalement liées au nettoyage des réseaux, les ratios largeur-hauteur moyens sont similaires à ceux présentés par Jouvès et al. [2017] et Pardo-Igúzquiza et al. [2011].

Cette analyse n'a mis en évidence aucune loi statistique simple et universelle, indépendante des conditions locales, qui permettrait de décrire les distributions du rayon équivalent et du ratio largeur-hauteur. Des tests statistiques ont été effectués pour pouvoir étayer nos observations, mais l'existence d'une distribution identique dans tous les réseaux est rejetée. Pourtant, beaucoup de ces réseaux sont localisés dans les mêmes régions et certains font probablement partie d'un même système, ce qui induit un certain biais.

Les tests effectués sur des sous-échantillons non corrélés ne rejettent pas l'hypothèse que les distributions des grands réseaux suivent une loi log-normale ou une loi de Pareto. Ces tests ayant été effectués sur seulement 11 sous-échantillons, eux-mêmes relativement petits (nombre moyen de 53 valeurs), il est délicat d'extrapoler ces résultats. On peut noter qu'il n'y a pas de raison évidente à ce que le rayon équivalent suive une loi log-normale ou une loi de Pareto. En effet, les données ont été directement recueillies par des spéléologues, principalement dans des conduits de taille humaine. Cela signifie que les conduits de dimension moindre ne sont pas pris en compte dans ces données. Au vu du comportement fractal observé dans certains réseaux (e.g., Hendrick et Renard, 2016b) et qui pourrait être généralisé dans la majorité des cas (Pardo-Igúzquiza et al., 2019), la distribution réelle des rayons équivalents des conduits serait probablement plus proche d'une loi de puissance. Par ailleurs, il est compliqué de juger le ratio largeur-hauteur puisqu'il dépend fortement de l'échantillonnage, ce dernier pouvant être variable dans les zones très verticales. Dans les gouffres et conduits à fort pendage, les hauteurs et largeurs sont généralement mesurées dans les directions perpendiculaires au déplacement des spéléologues ; la hauteur mesurée ne correspond ainsi pas à la hauteur réelle du conduit (qui est en fait la longueur mesurée). Lorsque les mesures sont effectuées au fond de canyons et que la progression des spéléologues s'effectue majoritairement à l'horizontal, les valeurs de hauteur et de largeur correspondent normalement aux vérita-

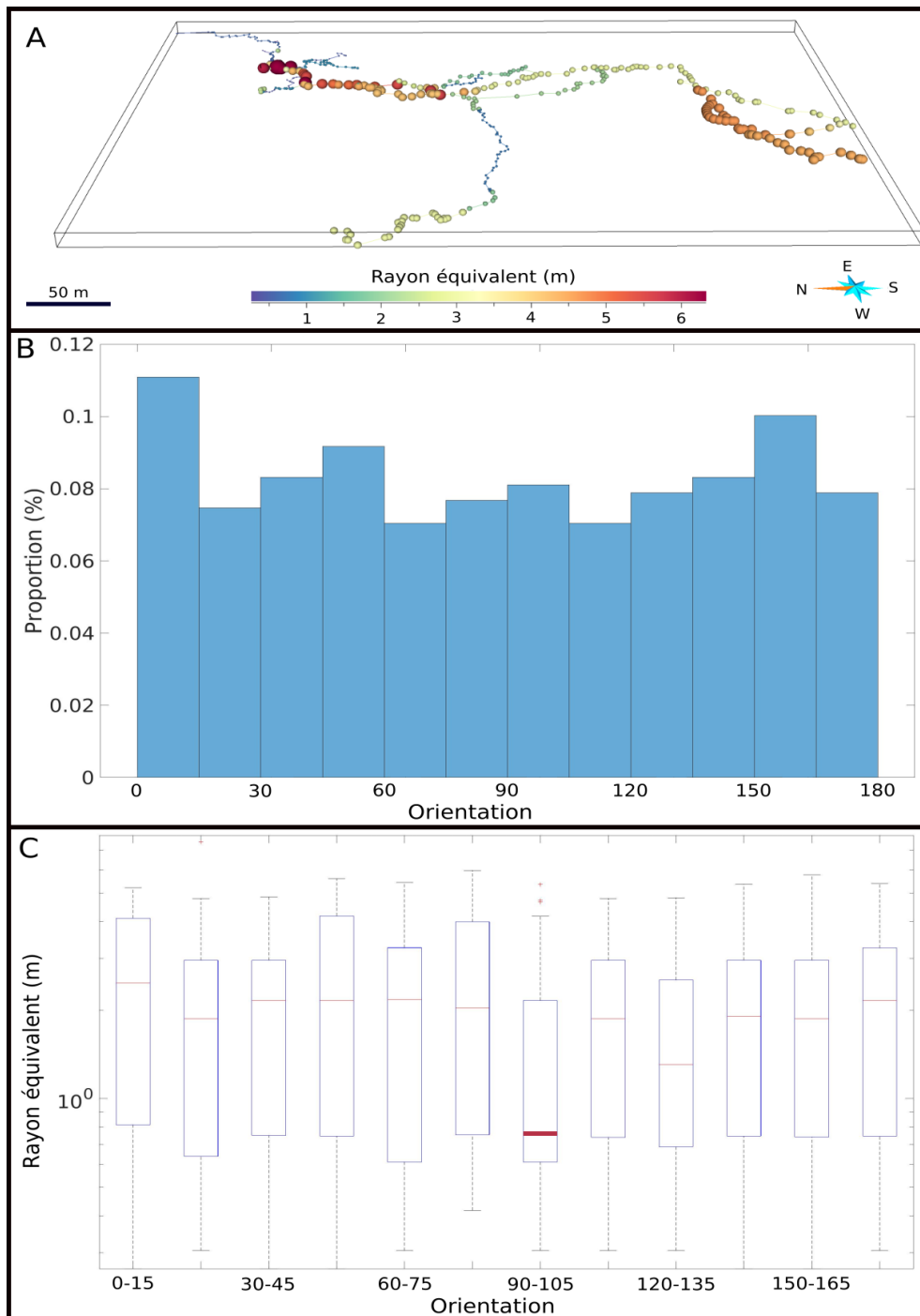


Figure 2.23 – A : Réseau de Wakulla (474 nœuds). La taille des nœuds est directement proportionnelle aux rayons équivalents associés. Il semble y avoir une évidente différence de comportement entre les conduits orientés Nord-Sud et ceux orientés Est-Ouest. B : Histogramme des orientations des segments de conduits du réseau de Wakulla (469 segments, seuls ceux associés à une valeur de rayon équivalent sont pris en compte). C : Diagramme en boîte du rayon équivalent en fonction des valeurs d'orientation. La médiane pour la boîte 90-105 est beaucoup plus basse que les autres (le trait est épaissi pour assurer sa bonne visibilité).

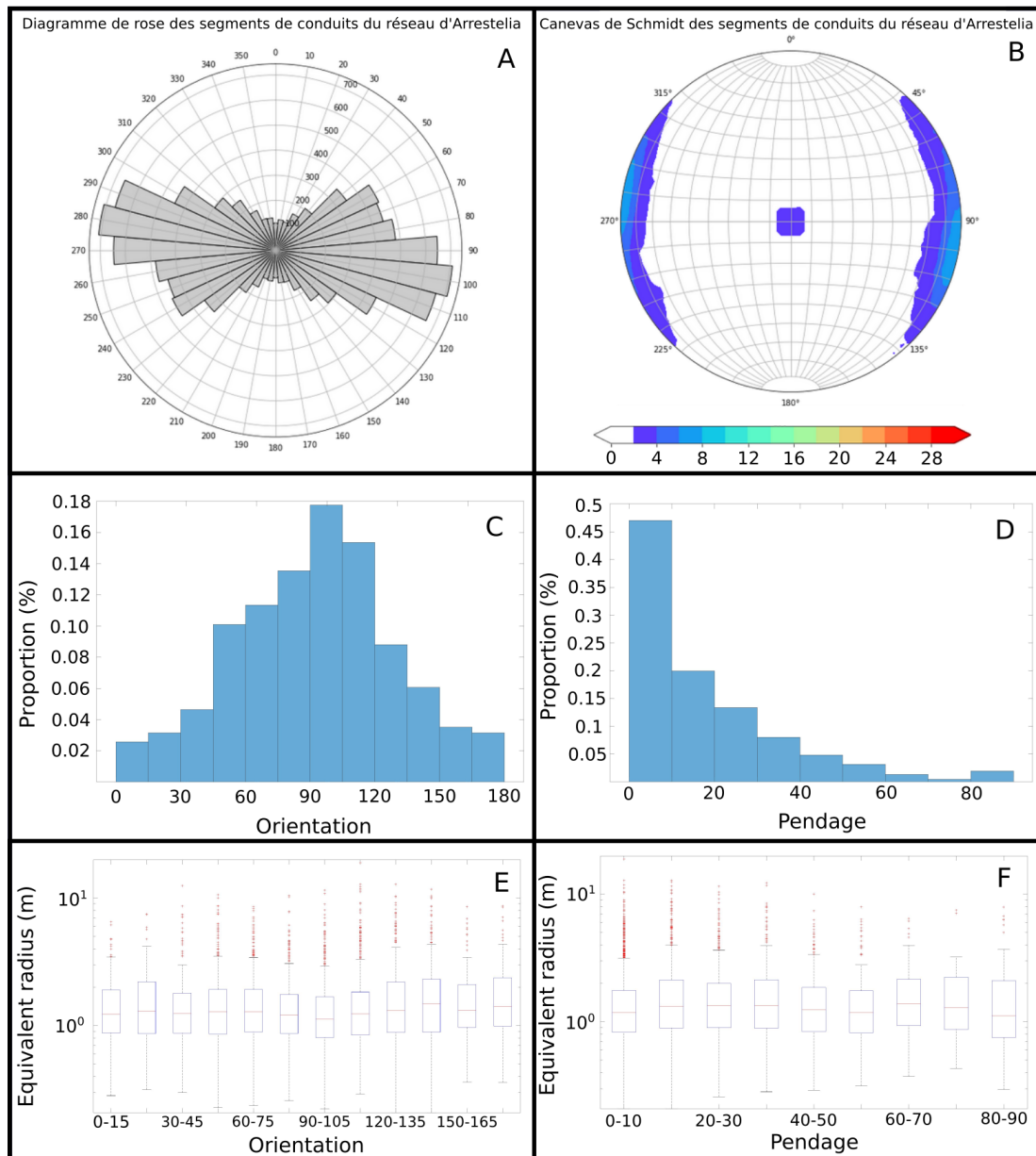


Figure 2.24 – A : Diagramme de rose des segments de conduit du réseau d'Arrestelia (6012 segments). B : Canevas de Schmidt des segments de conduit du réseau d'Arrestelia (6012 segments). C : Histogramme des orientations des segments de conduit du réseau d'Arrestelia (4009 segments, seuls ceux associés à une valeur de rayon équivalent et n'étant pas verticaux sont pris en compte). D : Histogramme des pentages des segments de conduit du réseau d'Arrestelia (4088 segments, seuls ceux associés à une valeur de rayon équivalent sont pris en compte). E : Diagramme en boîte des rayons équivalents en fonction des valeurs d'orientation. F : Diagramme en boîte des rayons équivalents en fonction des valeurs de pentage.

bles hauteurs et largeurs. Les grandes zones planes et basses de plafonds sont difficiles à échantillonner à cause des difficultés d'accès, ce qui pourrait être une des causes de la prépondérance des faibles valeurs de ratio largeur-hauteur dans les grands réseaux.

Bien que nous n'ayons pas non plus jugé nécessaire de présenter ces résultats dans le présent manuscrit, nous n'avons pas observé de corrélation directe entre le rayon équivalent et le ratio largeur-hauteur. Il est cependant relativement rare d'observer de grandes valeurs de rayon associées à de grandes valeurs de ratio largeur-hauteur, alors qu'elles sont plus fréquemment observées pour de très faibles valeurs de ratio largeur-hauteur. La majorité des grandes valeurs de rayon équivalent sont cependant associées à des valeurs de ratio largeur-hauteur relativement proches de 1.

Les résultats de l'analyse des motifs morphologiques (Annexe B) offrent de nouvelles informations mais ils ne doivent pas être généralisés à cause de leur manque de robustesse. En effet, nous avons trop peu de réseaux associés à chaque motif. De plus, la majorité de ces parties sont petites, surtout comparées à un réseau entier. Par exemple, la célèbre grotte de Lechuguilla est une des plus longues grottes du monde (plus de 200 km de conduits connus) et a été formée via des processus hypogéniques (e.g., Duchene et al., 2017). Les 5 réseaux hypogènes présentés dans cette étude ont une longueur médiane de 314 mètres et ne sont donc pas représentatifs d'une telle cavité. La division des réseaux selon les motifs se base sur des informations de terrain mais également sur leur topologie et leur géométrie, ce qui peut induire un biais lors de leur analyse statistique. Il est logique que les conduits localisés dans les parties des réseaux considérées comme vadoses soient plus verticaux et qu'ils aient ainsi des ratios largeur-hauteurs plus faibles. Ainsi, bien que ces résultats soient intéressants, il faut veiller à ne pas les surinterpréter.

A cause de la complexité des réseaux, les mesures peuvent manquer de précision, ce qui impacte directement les résultats de cette étude statistique. L'approximation elliptique est également une potentielle source d'erreur impactant les résultats puisqu'elle lisse la forme réelle des conduits. Néanmoins, une approximation circulaire ne prendrait pas en compte la différence entre les hauteurs et largeurs et une approximation rectangulaire serait moins respectueuse de la forme habituelle des conduits. La cumulation de la distance dans deux directions pour travailler avec les largeurs et hauteurs induit malgré tout une perte d'information car elle modifie le centrage des conduits, contrairement à l'utilisation des distances vers la droite, la gauche, le haut et le bas mais seules les informations de largeur et de hauteur sont disponibles dans de nombreux cas. La représentation des réseaux par des graphes est également une grande approximation de la réalité. Les propriétés géométriques sont uniquement portées par les nœuds, ce qui empêche de connaître précisément les valeurs entre ces derniers. Leur nettoyage (comme par exemple la suppression des nœuds dupliqués) implique l'émission de certaines hypothèses, comme le fait de considérer que deux nœuds voisins sont en fait une seule position mesurée deux fois, ce qui est également une source d'erreurs potentielle. La conversion des données a également nécessité certains changements manuels impactant les réseaux finaux, plusieurs parties de petite taille, non connectées au reste du réseau ou sur lesquelles les propriétés géométriques ne sont pas connues, n'ayant pas été prises en compte. Notre définition des branches étant purement topologique, ces dernières ne sont pas parfaitement représentatives des conduits eux-mêmes. En dépit de ces erreurs, les simulateurs d'écoulement (e.g., SWMM, Epanet, Modflow-CFP) se basent sur une

telle approximation de la géométrie des conduits. Ainsi, ce traitement est en accord avec notre but, qui est de fournir un moyen de calculer le rayon équivalent et non la géométrie exacte des conduits.

La décroissance de la valeur-p conjointement à l'augmentation de la taille de l'échantillon testé est un problème connu (e.g., Berkson, 1938; Lin et al., 2013) mais peu caractérisé. Cela signifie que les tests effectués sur de grands échantillons ont tendance à être rejetés même si H_0 est vraie. Il a en effet pu être constaté que la majorité des tests effectués lors de cette étude (certains n'étant pas présentés ici) ont été rejetés lorsqu'ils concernaient de grands réseaux, et ce même avec de faibles valeurs α , tandis que ce n'était généralement pas le cas lorsqu'ils étaient effectués sur de petits réseaux. Bien que les tailles de nos échantillons soient relativement limitées comparées à d'autres études, les résultats obtenus sont assez discutables. Le fait d'avoir trop de données contraint aussi, dans certains tests, les distributions originales dont pourraient être issus les échantillons, de manière similaire à du surajustement. Je suspecte également que l'effet inverse se produise lors des tests effectués sur de petits échantillons, c'est à dire que les tests effectués sur les petits réseaux, incluant ceux réalisés sur les réseaux sous-échantillonnés, seraient potentiellement moins rejetés à cause du faible nombre de nœuds.

Les valeurs d'assortativité du rayon équivalent et du ratio largeur-hauteur indiquent que les nœuds voisins ont généralement des valeurs similaires, témoignant d'une corrélation spatiale à courte distance au sein des réseaux. Les variogrammes 1D-curvilinéaires mettent en évidence l'existence d'une corrélation spatiale pour le rayon équivalent et le ratio largeur-hauteur au sein des réseaux, et ce, sur de grandes distances. Cette corrélation est généralement plus importante à l'intérieur même des branches. Les portées des variogrammes diffèrent d'un réseau à un autre mais elles semblent cependant plus grandes que celles calculées par Pardo-Iguzquiza et al. [2011] sur le réseau de Sakany, qui utilisait des distances euclidiennes. Ne pas prendre en compte la curvilinéarité de la structure des réseaux pourrait en effet mener à une sous-estimation de la corrélation spatiale des dimensions des conduits. Par ailleurs, les variogrammes intrabranches ont une portée limitée par la définition même des branches. Une branche étant définie entre deux nœuds d'intersection ou d'extrémité, chaque confluence définit une nouvelle branche, or la définition d'un conduit pourrait être également perçue comme un vide relativement linéaire. Si un long conduit est rejoint en différents endroits par d'autres conduits, seules de petites parties de ce conduit seront individuellement prises en compte comme étant des branches.

Parmi les 4 types de centralité utilisés, seule la centralité de degré semble réellement avoir un lien avec les métriques étudiées, surtout pour le rayon équivalent. Cependant, bien que les valeurs médianes de rayon équivalent semblent souvent augmenter conjointement avec le degré des nœuds, les plus grandes valeurs sont beaucoup plus nombreuses pour les nœuds de degré 2 car la proportion de ces derniers est beaucoup plus importante. Les faibles valeurs des nœuds de degré 1 sont en partie expliquées par le fait qu'ils correspondent souvent aux endroits où la taille des conduits devient trop petite pour poursuivre leur exploration. Les difficultés de mesure aux extrémités et intersections peuvent également avoir un impact.

Aucun lien direct n'a pu être observé entre la distance à l'entrée ou la sortie la plus

proche et les propriétés géométriques des conduits lors de cette étude. Cela remet également en question l'idée que les conduits karstiques croîtraient vers l'aval des réseaux, avec des valeurs maximales aux sources.

L'algorithme de hiérarchisation a été développé dans l'optique d'obtenir des valeurs équivalentes à un flux théorique traversant le réseau ; il gère la présence de boucles et l'existence de multiples sorties dans la majorité des réseaux. Il est actuellement seulement utilisable pour de petits réseaux, principalement à cause des difficultés liées aux siphons, mais aucun lien n'a été trouvé entre les dimensions des conduits et les valeurs hiérarchiques calculées. Il peut, de plus, être compliqué de définir où sont les entrées et sorties du réseau. Les principales entrées et sorties peuvent être connues pour les réseaux actifs mais il n'est pas facile de les déterminer dans le cas où cette information n'est pas disponible, comme pour les réseaux paléokarstiques. Notre méthode permettant de définir les entrées et sorties du réseau n'est certes ni efficace, ni réaliste, mais il n'existe a priori aucune manière évidente de les identifier quand ces informations ne sont pas disponibles. Qui plus est, les flux matriciels ne sont pas non plus pris en compte, bien que leur impact ne soit en réalité pas négligeable. Borghi et al. [2016] associe chaque conduit à une valeur proportionnelle à celle de la surface drainée théorique correspondante. Les surfaces drainées sont calculées à partir de la topographie et sont reliées à chaque entrée du réseau. Ainsi, si nous avions eu les informations topologiques nécessaires, il aurait été intéressant d'utiliser des valeurs calculées ainsi au lieu d'assigner une valeur de 1 à chaque entrée potentielle. Il est néanmoins compliqué de définir une méthode de hiérarchisation efficace car le flux d'eau dans les réseaux change en fonction de différents paramètres. Le manque habituel d'informations relatives aux potentielles entrées et sorties et le fait que les réseaux ne soient généralement pas complètement cartographiés sont également des problèmes majeurs. En dépit de ces difficultés, l'inspection visuelle des données montre qu'il n'existe aucune hiérarchie évidente des distributions des dimensions des conduits dans les réseaux.

Pour résumer cette partie du travail, nous considérons que cette étude amène à rejeter la supposition habituelle d'une dépendance entre les dimensions des conduits et leur organisation hiérarchique, la distance aux entrées et sorties, et la centralité des nœuds dans les réseaux. Cette hypothèse n'est pas confirmée par les données de terrain et est probablement beaucoup plus complexe en réalité.

Aucune relation générique n'a été observée entre l'altitude des nœuds, l'orientation ou le pendage des conduits et les propriétés géométriques. Néanmoins, il semble y avoir de tels liens dans certains réseaux. Pour cette raison, nous encourageons une analyse au cas par cas.

2.6 Conclusion

Cette analyse du rayon équivalent et du ratio largeur-hauteur dans 49 réseaux différents a permis de mieux caractériser les dimensions des conduits des réseaux karstiques. Bien qu'aucune loi valide dans tous les réseaux n'ait été trouvée, les résultats peuvent servir de référence pour de futures études. Ces résultats sont également un important progrès vers la simulation de dimensions de conduits réalistes. Ils peuvent être appliqués dans différents contextes tels que les simulations d'écoulement dans des réseaux de conduits

discrets (e.g., SWMM, Epanet, Modflow-CFP). Néanmoins, cette analyse se base sur des relevés effectués de manière irrégulière, à des distances plurimétriques et uniquement dans les parties accessibles des réseaux. Elle n'a, de plus, été effectuée *que* sur 49 réseaux, la majorité d'entre eux étant situés dans des localités françaises proches les unes des autres et faisant parfois partie du même système. L'émergence de nouveaux outils d'exploration et d'acquisition de données permettront très probablement de compléter cette étude préliminaire en l'élargissant à un plus grand nombre de réseaux karstiques à travers le monde.

Chapitre 3

Simulation des propriétés géométriques des conduits

Sommaire

3.1	Introduction	86
3.2	Méthodes de simulation	86
3.3	Résultats	88
3.3.1	SGS 1D-curvilinéaire	91
3.3.2	SGS 1D-curvilinéaire <i>branchwise</i>	98
3.4	Discussion	105
3.5	Conclusion	107

Résumé Ce chapitre présente deux méthodes de simulation de propriétés le long de réseaux karstiques. Elles se basent sur les résultats présentés dans le Chapitre 2. La première méthode se focalise sur la reproduction du variogramme global, tandis que la deuxième se concentre sur la reproduction des variogrammes intrabranche et interbranche. Les résultats obtenus sur 4 cas synthétiques et sur un cas réel sont présentés.

Ces méthodes et une partie des résultats ont été publiées dans l'article suivant : Frantz et al. [2021].

3.1 Introduction

Comme vu précédemment, certains logiciels permettent d'effectuer des simulations d'écoulement sur des réseaux de conduits discrets décrits par le rayon de leur section (e.g., SWMM, Epanet, Modflow-CFP). Peterson et Wicks [2006] ont démontré que de petits changements de géométrie peuvent mener à des changements statistiquement significatifs. Être capable de simuler les dimensions des conduits de façon *réaliste* pourrait grandement améliorer la caractérisation des écoulements fluides en leur sein.

Lors de simulations d'écoulement sur des réseaux de conduits discrets, le rayon peut être défini de plusieurs manières (e.g., Ronayne, 2013; Vuilleumier et al., 2013; Chen et Goldscheider, 2014; Borghi et al., 2016). Les valeurs de rayon sont souvent constantes au sein de chaque conduit (e.g., Vuilleumier et al., 2013; Jeannin et al., 2015) et calibrées au cours de la simulation. La variabilité du rayon au sein du réseau et de ses conduits a pu être caractérisée dans le Chapitre 2 à travers les variogrammes globaux, intrabranches et interbranches, mais aucune méthode ne la prend en compte.

Ainsi, au vu de l'importance de la géométrie des conduits, nous avons développé deux méthodes de simulation permettant de générer des valeurs de rayon équivalent et de ratio largeur-hauteur le long des réseaux karstiques tout en respectant leur variabilité spatiale. Ces méthodes visent à remplir les parties des réseaux qui ne sont pas associées à des informations géométriques. Elles peuvent à la fois être appliquées à des réseaux connus ou à des squelettes de réseaux simulés. Nous nous sommes basés sur les résultats majeurs présentés dans le Chapitre 2 pour leur conception, et plus particulièrement sur les variogrammes. Dans la première partie de ce chapitre, je présente les concepts théoriques derrière les deux méthodes de simulation développées durant cette thèse. Dans la seconde partie, je présente les résultats obtenus sur 4 cas synthétiques et sur le réseau d'Arrestelia.

3.2 Méthodes de simulation

Deux méthodes basées sur les Simulations Gaussiennes Séquentielles (SGS) ont été développées au cours de cette thèse. Elles ont pour but de générer les propriétés géométriques, principalement le rayon équivalent et le ratio largeur-hauteur, le long

des graphes représentant les réseaux. Ces simulations peuvent porter sur des parties de réseaux dépourvues d'information ou sur des réseaux issus eux-mêmes de simulations. Les méthodes développées respectent les données conditionnantes ainsi que les informations géostatistiques spécifiées par l'utilisateur (distributions et variogrammes). La stochasticité des processus permet de gérer les incertitudes associées (e.g., Chilès et Delfiner, 2009).

Les Simulations Gaussiennes Séquentielles (SGS) sont des méthodes permettant la simulation de variables continues principalement utilisées dans le cadre des géostatistiques. Contrairement aux estimateurs tels les krigeages, elles ne génèrent pas un seul modèle *optimal* mais plusieurs modèles stochastiques respectant les histogrammes et variogrammes d'entrée. Leur utilisation suppose une distribution d'entrée gaussienne. Si ce n'est pas le cas, il est nécessaire d'avoir recours à une anamorphose gaussienne de la propriété simulée pour obtenir une distribution suivant une loi normale (Goovaerts, 1997; Deutsch et Journel, 1998). Les paramètres des variogrammes doivent alors être estimés à partir de la propriété transformée au lieu de la propriété initiale. Finalement, après obtention des résultats, une étape de transformation inverse est effectuée pour obtenir les vraies valeurs simulées.

Les SGS impliquent la réalisation d'un krigeage sur chaque nœud du réseau. Le krigeage est un estimateur linéaire qui estime la valeur d'un nœud à partir des valeurs de ses voisins. Comme la moyenne de la distribution est connue (et égale à 0 si la distribution a été transformée), un krigeage simple est utilisé. Pour chaque réalisation, un chemin aléatoire permettant de visiter tous les nœuds à simuler (nœuds non-associés à des valeurs) est généré. L'algorithme détermine le voisinage de chaque nœud sans valeur. Ce dernier correspond à l'ensemble des nœuds associés à une valeur et étant situés plus proches du nœud considéré qu'une distance définie par l'utilisateur. Si le voisinage est vide, une valeur est échantillonnée à partir de la distribution d'entrée. Dans le cas contraire, un krigeage simple est effectué et permet l'obtention d'une valeur estimée Z_0^* et de sa variance d'estimation σ_k^2 . Une valeur aléatoire est alors obtenue via un échantillonnage par la méthode de Monte-Carlo dans la distribution gaussienne $N(Z_0^*, \sigma_k^2)$. Le résultat est ensuite ajouté aux nœuds conditionnants pris en compte lors de la simulation des nœuds suivants.

Les principales différences entre nos méthodes et les méthodes de SGS classiques sont : i) l'utilisation de la distance curvilinéaire le long du réseau au lieu d'une distance euclidienne et ii) la définition du voisinage. En plus de la distance et de l'existence d'une valeur associée, la présence des voisins dans la même branche ou dans une branche différente est également prise en compte. Ces méthodes utilisent une matrice de distance calculée sur toutes les paires de nœuds via l'algorithme de Johnson (Johnson, 1977).

Deux variantes de la méthode de simulation ont été développées durant cette thèse :

- La méthode de **SGS 1D-curvilinéaire** qui respecte la variabilité spatiale calculée à l'échelle du réseau (voir Algorithme 1). Toutes les valeurs sont ainsi simulées en se servant du variogramme global.
- La méthode de **SGS 1D-curvilinéaire *branchwise*** (par branche) qui prend principalement en compte la différence de variabilité spatiale au sein des branches et essaye de plus de reproduire la variabilité à l'échelle du réseau (voir Algorithme 2). Elle commence par la simulation d'un certain nombre de nœuds dans chaque

Algorithme 1 : SGS 1D-curvilinéaire

Entrée: Squelette du réseau, distribution, variogramme 1D-curvilinéaire, données conditionnantes (si simulation conditionnelle)

Sortie: Valeurs de la propriété simulée

Si distribution non gaussienne **alors**
 Anamorphose gaussienne de la distribution et des données conditionnantes

Fin si

Calcul de la plus courte distance curvilinéaire entre toutes les paires de nœuds en utilisant l'algorithme de Johnson

Pour chaque réalisation **faire**
 Génération d'un chemin aléatoire pour visiter tous les nœuds sans valeur

Pour chaque nœud sur le chemin **faire**
 Recherche de son voisinage
Si voisinage vide **alors**
 Tirage d'une valeur aléatoire dans la distribution d'entrée

Sinon
 Krigeage simple avec le variogramme global sur le nœud pour obtenir la valeur attendue Z_0^* et sa variance d'estimation σ_k^2
 Échantillonnage de la distribution normale $N(Z_0^*, \sigma_k^2)$ pour obtenir la valeur du nœud

Fin si
 Ajout du résultat aux nœuds conditionnants

Fin pour

Si une anamorphose gaussienne a été effectuée **alors**
 Transformation inverse des résultats

Fin si

Fin pour

branche en se servant du variogramme interbranche (Figure 3.1B-D). Ce nombre dépend de deux paramètres définis par l'utilisateur : une proportion *prop* et un nombre maximal de nœuds par branche *nbnoeudsmax*. Lors de cette étape, les données conditionnantes réduisent d'autant le nombre de nœuds à simuler. Les nœuds restants et localisés dans les branches sont simulés en se servant du variogramme intrabranché lors d'une seconde étape (Figure 3.1E-J). Le voisinage utilisé est cette fois uniquement composé de nœuds situés dans la même branche que le nœud simulé. Enfin, les nœuds d'intersection sont simulés avec le variogramme global dans une troisième et dernière étape (Figure 3.1K).

Ces deux méthodes sont implémentées en C++ dans le plugin RingKarst du logiciel SKUA-Gocad ¹.

3.3 Résultats

Afin de démontrer le bon fonctionnement des méthodes de simulation, je présente dans cette partie les résultats obtenus sur plusieurs cas synthétiques et sur un cas réel. Seuls des résultats obtenus pour la simulation du rayon équivalent sont présentés mais la méthode de simulation du ratio largeur-hauteur ou d'autres propriétés est en tout point identique.

Le réseau d'Arrestelia est ici utilisé comme base pour illustrer et tester les méthodes de simulation proposées sur un cas réel. La distance moyenne entre chaque station de

¹<https://www.pdgm.com/products/skua-gocad>

Algorithme 2 : SGS 1D-curvilinéaire *branchwise*

Entrée: Squelette du réseau, distribution, variogramme 1D-curvilinéaire intrabranche, variogramme 1D-curvilinéaire interbranche, variogramme 1D-curvilinéaire global, données conditionnantes (si simulation conditionnelle), proportion de nœuds simulés par branche avec le variogramme interbranche *prop*, nombre maximal de nœuds simulés par branche avec le variogramme interbranche *nbnoeudsmx*

Sortie: Valeurs de la propriété simulée

Si distribution non gaussienne **alors**
 Anamorphose gaussienne de la distribution et des données conditionnantes

Fin si

Calcul de la plus courte distance curvilinéaire entre toutes les paires de nœuds en utilisant l'algorithme de Johnson

Pour chaque réalisation **faire**
 Conditionnement interbranche
 Pour chaque branche **faire**
 Calcul du nombre de nœuds devant être simulés dans cette branche à partir du nombre de nœuds dans la branche, nœuds d'intersection exceptés, ainsi que de la proportion *prop* et du nombre maximal de nœuds *nbnoeudsmx* définis en entrée
 Génération d'un chemin aléatoire pour visiter le même nombre de nœuds sans valeur situés dans la branche
 Pour chaque nœud sur le chemin **faire**
 Recherche de son voisinage
 Si voisinage vide **alors**
 Tirage d'une valeur aléatoire dans la distribution d'entrée
 Sinon
 Krigeage simple avec le variogramme interbranche sur le nœud pour obtenir la valeur attendue Z_0^* et sa variance d'estimation σ_k^2
 Échantillonnage de la distribution normale $N(Z_0^*, \sigma_k^2)$ pour obtenir la valeur du nœud
 Fin si
 Ajout du résultat aux nœuds conditionnants
 Fin pour
 Fin pour
 Simulation intrabranche
 Génération d'un chemin aléatoire pour visiter tous les nœuds sans valeur, nœuds d'intersection exceptés
 Pour chaque nœud sur le chemin **faire**
 Recherche de son voisinage dans la même branche
 Si voisinage vide **alors**
 Tirage d'une valeur aléatoire dans la distribution d'entrée
 Sinon
 Krigeage simple avec le variogramme intrabranche sur le nœud pour obtenir la valeur attendue Z_0^* et sa variance d'estimation σ_k^2
 Échantillonnage de la distribution normale $N(Z_0^*, \sigma_k^2)$ pour obtenir la valeur du nœud
 Fin si
 Ajout du résultat aux nœuds conditionnants
 Fin pour
 Simulation des intersections
 Pour chaque nœud d'intersection sans valeur **faire**
 Recherche de son voisinage
 Si voisinage vide **alors**
 Tirage d'une valeur aléatoire dans la distribution d'entrée
 Sinon
 Krigeage simple avec le variogramme global sur le nœud pour obtenir la valeur attendue Z_0^* et sa variance d'estimation σ_k^2
 Échantillonnage de la distribution normale $N(Z_0^*, \sigma_k^2)$ pour obtenir la valeur du nœud
 Fin si
 Ajout du résultat aux nœuds conditionnants
 Fin pour
 Si une anamorphose gaussienne a été effectuée **alors**
 Transformation inverse des résultats
 Fin si
Fin pour

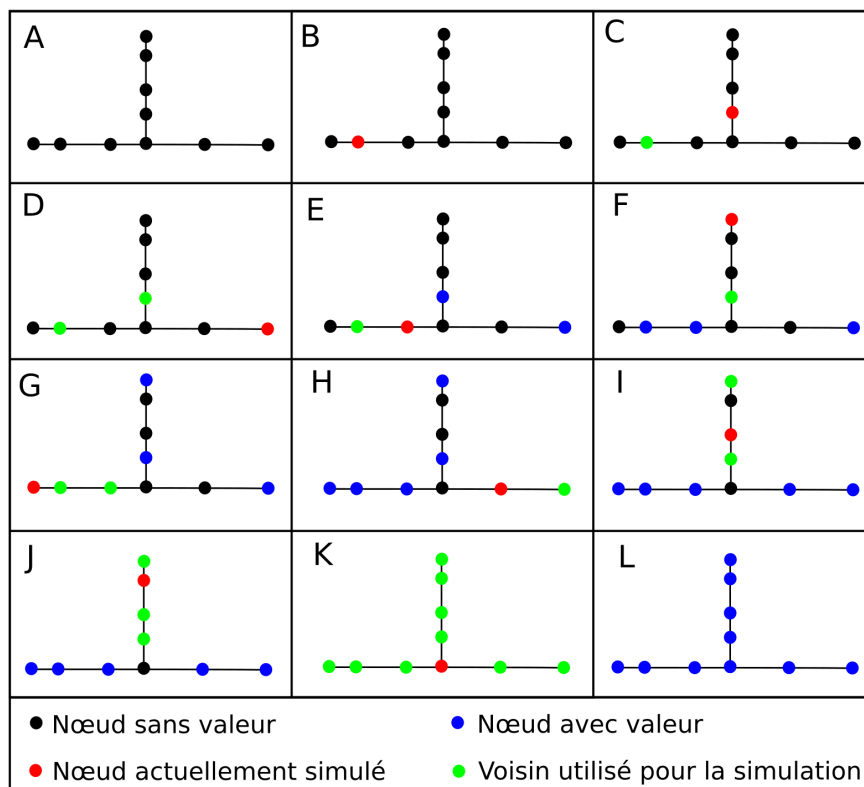


Figure 3.1 – Exemple d’une simulation non-conditionnelle avec la méthode de SGS 1D-curvilinéaire *branchwise*. L’ordre des nœuds simulés est tiré aléatoirement à chaque étape. A : Réseau initial. B-D : Conditionnement interbranche. Simulation d’un nœud par branche en se servant du variogramme interbranche. E-J : Simulation intrabranché. Simulation de tous les nœuds sans valeur, nœuds d’intersection exceptés en se servant du variogramme intrabranché. K : Simulation des intersections. Simulation du nœud d’intersection en se servant du variogramme global. L : Résultat final.

mesure est de 9.7 mètres mais comme l’échantillonnage n’est pas régulier, nous avons densifié le réseau pour que la distance maximale entre 2 nœuds voisins soit de 5 mètres. La valeur de 5 mètres est choisie afin de disposer d’assez de nœuds pour effectuer une analyse robuste des résultats des simulations, tout en gardant un temps de calcul raisonnable. Le réseau utilisé pour ces simulations a ainsi 14610 nœuds, dont 4283 portant des informations géométriques. Les informations statistiques de ce réseau sont présentées plus en détail dans le Chapitre 2.

Les simulations présentées par la suite, à la fois sur les cas synthétiques et sur le cas réel, ont été effectuées en prenant comme paramètres d’entrée la distribution du logarithme décimal du rayon équivalent du réseau d’Arrestelia et les paramètres des variogrammes théoriques de sa transformée par anamorphose gaussienne, la distribution n’étant pas gaussienne. Nous avons choisi d’effectuer ces simulations sur le logarithme décimal du rayon équivalent au lieu du rayon lui-même car ses variations sont moins importantes. Une taille de voisinage maximale de 16 nœuds a été choisie pour limiter le temps de calcul.

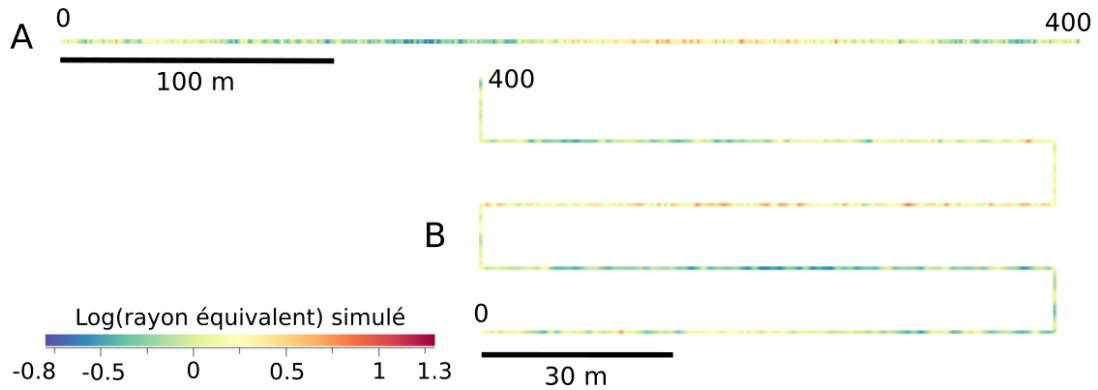


Figure 3.2 – Comparaison de résultats obtenus par simulation non-conditionnelle via la méthode de SGS 1D-curvilinéaire sur les cas synthétiques A et B. Pour des paramètres et une graine aléatoire identiques, les résultats sont parfaitement équivalents, exception faite des coordonnées cartésiennes des nœuds. La longueur totale et la densité de nœuds des deux cas sont égales mais les échelles utilisées sont cependant différentes pour permettre leur affichage sur la même figure. Les repères 0 et 400 font référence aux mêmes nœuds sur les 2 courbes.

3.3.1 SGS 1D-curvilinéaire

Cas synthétiques

Afin de s’assurer de son bon fonctionnement, cette méthode a été testée sur de nombreux cas synthétiques au fur et à mesure de son développement. Nous présentons ici les résultats obtenus sur deux cas synthétique simples. Le cas synthétique A correspond à une ligne droite tandis que le cas synthétique B correspond à une ligne brisée, les deux ayant une longueur de 400 mètres. Cette longueur représente environ le triple de la portée du variogramme devant être reproduit. Nous avons choisi une densité d’un nœud tous les mètres, les résultats présentés correspondant ainsi à 401 nœuds. Le choix de cette densité permet d’obtenir des résultats assez robustes pour notre étude et pouvant être visualisés, bien qu’il soit relativement compliqué de distinguer les nœuds individuellement.

Les résultats obtenus lors de simulations non-conditionnelles sur ces deux cas synthétiques sont parfaitement identiques (Figure 3.2). En effet, l’utilisation de la même graine aléatoire (*seed*) fournit exactement les mêmes résultats, bien que les coordonnées des nœuds soient très différentes. Cela prouve à la fois la reproductibilité des résultats et le fait que seule la distance curvilinéaire soit prise en compte.

Les histogrammes et variogrammes des résultats sont également parfaitement identiques sur les deux cas testés et sont présentés en Figure 3.3 et Figure 3.4. Les distributions des résultats sont très proches de la distribution d’entrée. Ce cas consistant en unique branche, les variogrammes globaux et intrabranche sont confondus et il n’y a pas de variogramme interbranche. Les résultats pouvant être assez différents d’une réalisation à une autre, il en va de même pour les variogrammes calculés sur lesdites réalisations. Pour cette raison, nous présentons un variogramme moyenné. Les variogrammes obtenus sont relativement proches du variogramme théorique, bien que les

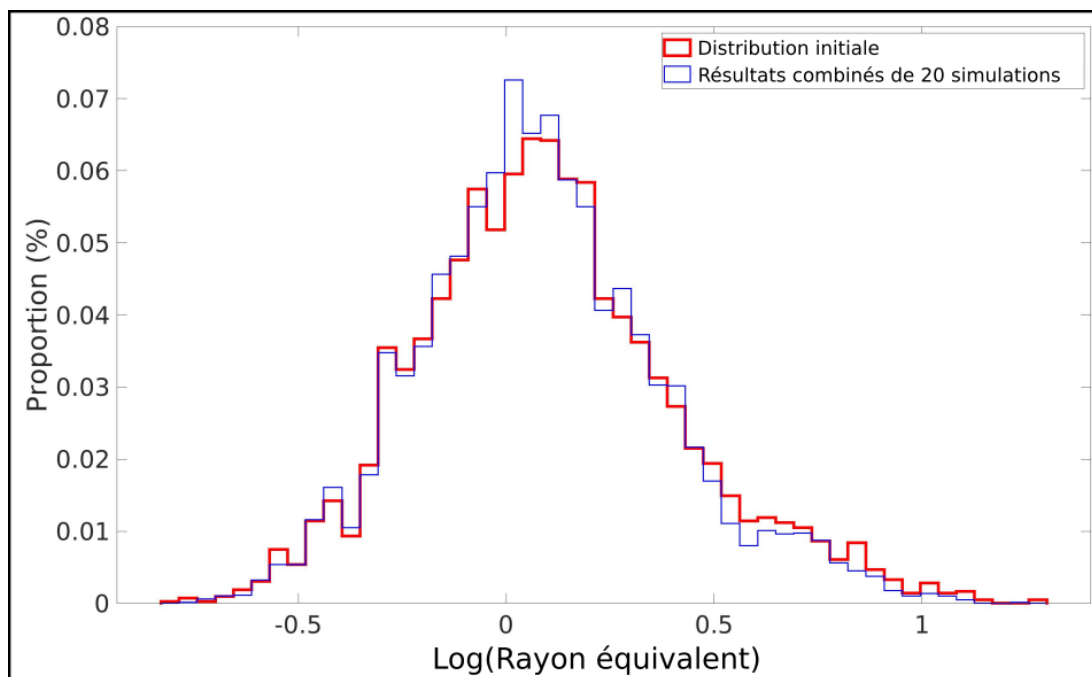


Figure 3.3 – Histogramme de la somme des résultats de 20 simulations non-conditionnelles effectuées sur le cas synthétique A avec la méthode de SGS 1D-curvilinéaire et distribution initiale du logarithme décimal du rayon équivalent utilisée comme paramètre d’entrée. La densité utilisée est d’un nœud tous les 1m pour 401 nœuds au total. Les résultats obtenus sur le cas synthétique B sont parfaitement identiques.

valeurs à faible distance, et surtout l’effet pépite, soient plus élevées. Les valeurs du variogramme expérimental moyenné commencent à s’éloigner du seuil du variogramme théorique passé 140 mètres (ce qui correspond à la portée du variogramme théorique). Il a pour cette raison été choisi de limiter la distance maximale affichée à 200 mètres, ce qui correspond à la moitié de la longueur de la branche, le nombre de paires de nœuds devenant de plus en plus faible au-delà.

La simulation conditionnelle a été testée sur le cas synthétique A en définissant aux nœuds situés à 50, 150, 250 et 350m du nœud d’origine des valeurs de -0.5 , 0 , 0.5 et 1 . Les données conditionnantes étant définies arbitrairement afin de pouvoir analyser visuellement la répartition spatiale des résultats, les histogrammes et variogrammes obtenus diffèrent fortement des paramètres théoriques et leur étude n’est pas pertinente. Certains résultats de simulations conditionnelles sont présentés en Figure 3.5. On observe bien une continuité des valeurs et les tendances imposées par les données conditionnantes sont généralement respectées, ce qui, combiné avec le respect de la distribution et du variogramme d’entrée des simulations non-conditionnelles, semble indiquer le bon fonctionnement de la méthode.

Réseau d’Arrestelia

20 simulations non-conditionnelles du logarithme décimal du rayon équivalent ont été effectuées sur l’ensemble du réseau d’Arrestelia (e.g., Figure 3.6). Comme précédemment

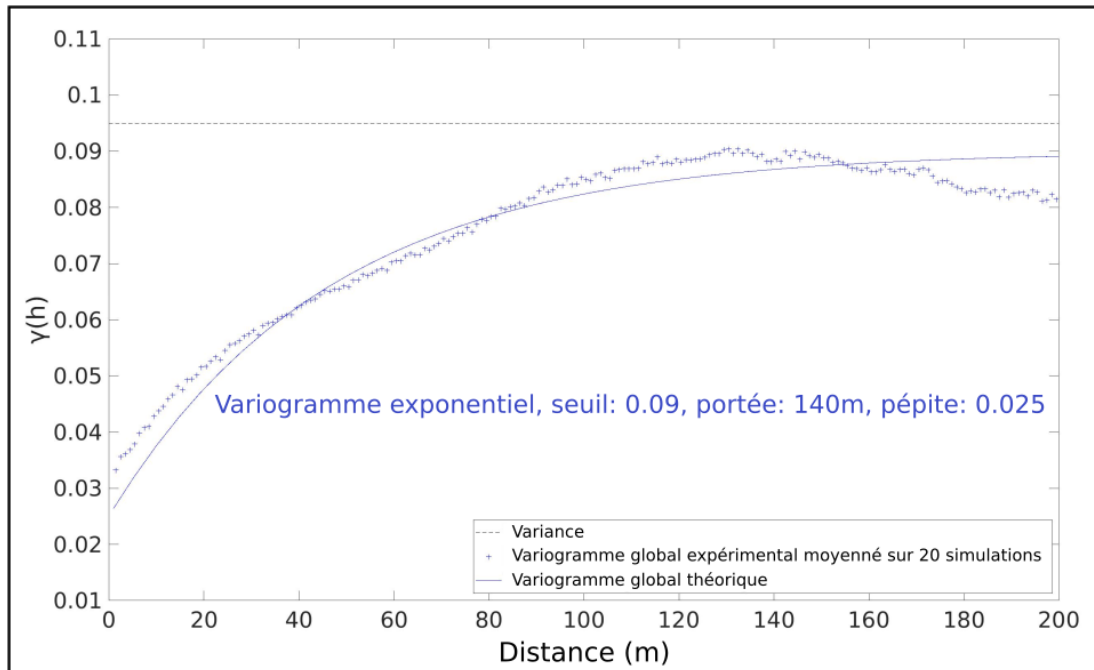


Figure 3.4 – Variogramme expérimental moyenné des résultats de 20 simulations non-conditionnelles effectuées sur le cas synthétique A avec la méthode de SGS 1D-curvilinéaire et variogramme théorique du logarithme décimal du rayon équivalent. La densité utilisée est d'un nœud tous les 1m pour 401 nœuds au total. Les résultats obtenus sur le cas synthétique B sont parfaitement identiques.

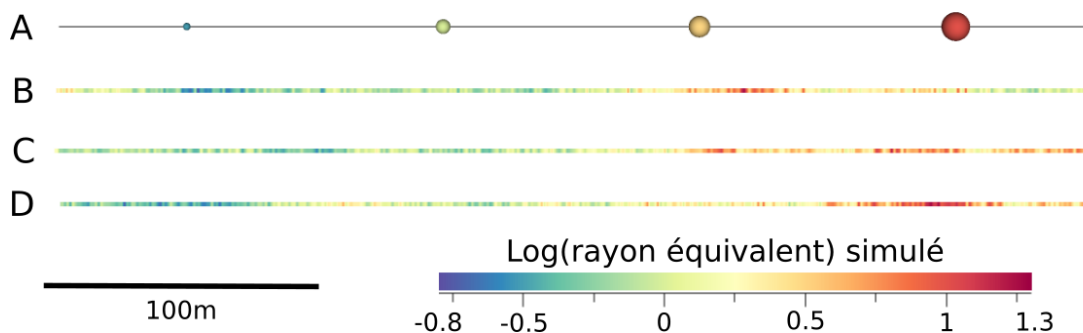


Figure 3.5 – A : Données conditionnantes pour le cas synthétique A, la taille des sphères est proportionnelle à leur valeur. B-D : 3 exemples de résultats obtenus par simulation conditionnelle via la méthode de SGS 1D-curvilinéaire sur le cas synthétique A. Tous ces résultats sont obtenus pour une densité d'un nœud tous les mètres, conduisant à un total de 401 nœuds.

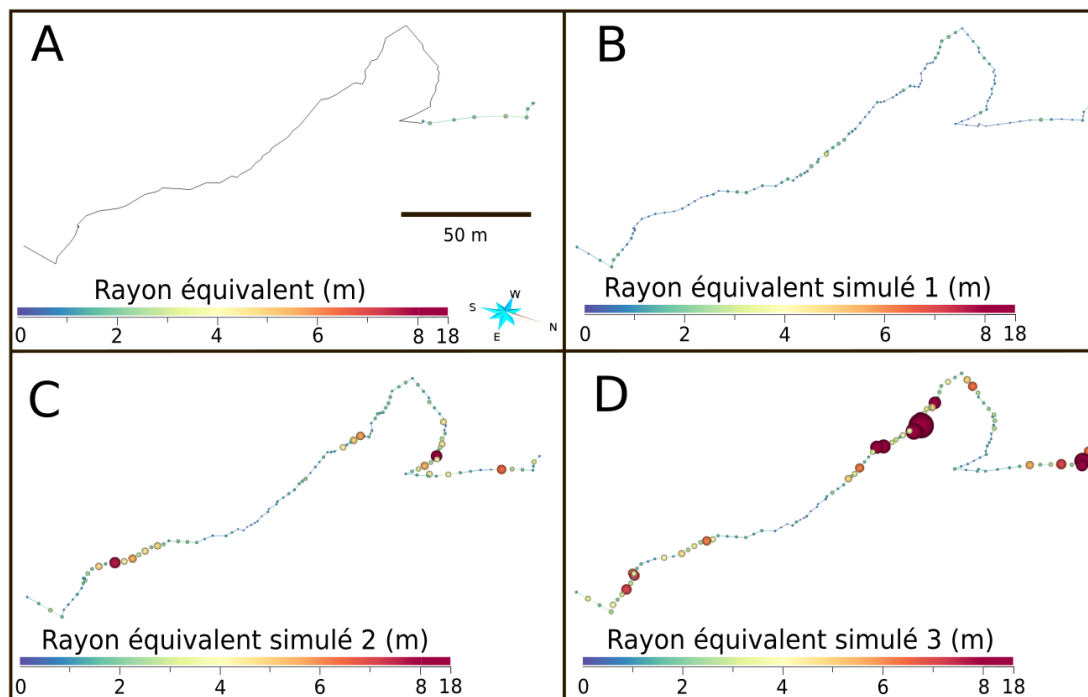


Figure 3.6 – A : Partie du réseau d’Arrestelia. B-D : 3 exemples de résultats obtenus sur cette partie par simulation non-conditionnelle via la méthode de SGS 1D-curvilinéaire sur le réseau entier (aucune donnée conditionnante n’est utilisée). Le rayon équivalent affiché a été calculé à partir des valeurs logarithmiques simulées.

nous nous sommes servis de la distribution du logarithme décimal du rayon équivalent du réseau d’Arrestelia, ainsi que du variogramme global théorique de sa transformée comme paramètres d’entrée.

Les distributions des résultats sont relativement proches de la distribution initiale mais sont légèrement plus centrées autour de la moyenne (Figure 3.7). Les moyennes et les médianes sont presque identiques aux valeurs de la distribution d’entrée : la moyenne des résultats est bien de 0.101, tandis que la médiane des résultats est de 0.082 au lieu de 0.081. Les variances sont un peu plus basses, avec une valeur moyenne de 0.090 au lieu de 0.098.

Les variogrammes globaux expérimentaux calculés sur les différentes réalisations sont proches du variogramme global théorique donné en entrée (Figure 3.8). Les variogrammes interbranches et intrabranches n’étant pas utilisés comme paramètres d’entrée, seul le comportement global est réellement reproduit lors de la simulation. Les variogrammes interbranches expérimentaux sont relativement similaires au variogramme interbranche théorique à cause de leur proximité avec les variogrammes globaux, mais aucun comportement intrabranche n’est présent.

20 simulations conditionnelles ont également été effectuées (e.g., Figure 3.9). Les données conditionnantes ont été créées en choisissant aléatoirement des nœuds distants de plus de 100 mètres l’un de l’autre, 203 au total. Un échantillonnage de Monte-Carlo de 203 valeurs aléatoires a ensuite été effectué dans la distribution initiale et elles ont été assignées aléatoirement aux nœuds, afin de pouvoir être utilisées comme données conditionnantes. Les données conditionnantes ainsi obtenues ont une valeur moyenne

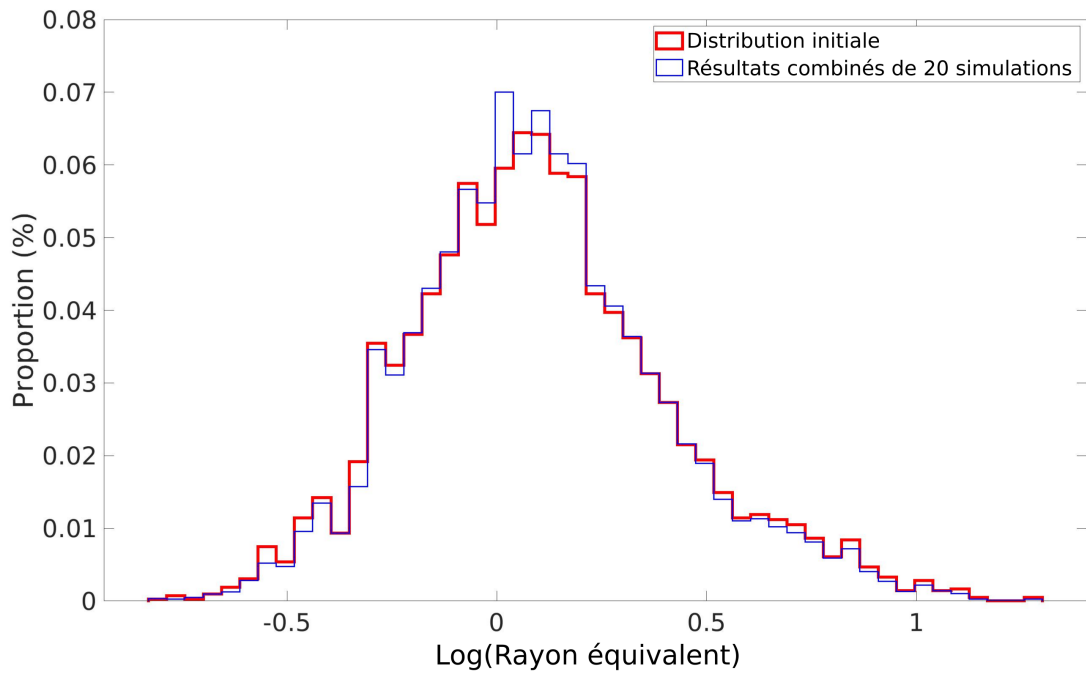


Figure 3.7 – Histogramme de la somme des résultats de 20 simulations non-conditionnelles effectuées sur le réseau d'Arrestelia avec la méthode de SGS 1D-curvilinearé et distribution initiale du logarithme décimal du rayon équivalent.

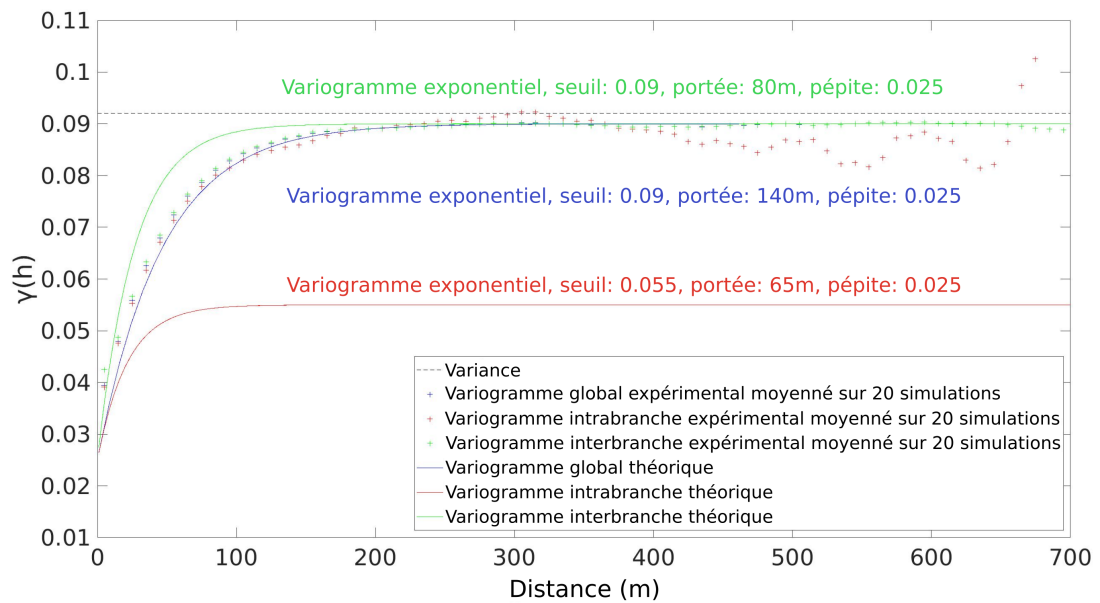


Figure 3.8 – Variogrammes expérimentaux moyennés des résultats de 20 simulations non-conditionnelles effectuées sur le réseau d'Arrestelia avec la méthode de SGS 1D-curvilinearé et variogrammes théoriques du logarithme décimal du rayon équivalent. Les variogrammes intrabranche et interbranche théoriques n'ont pas été utilisés lors des simulations mais sont donnés à titre indicatif.

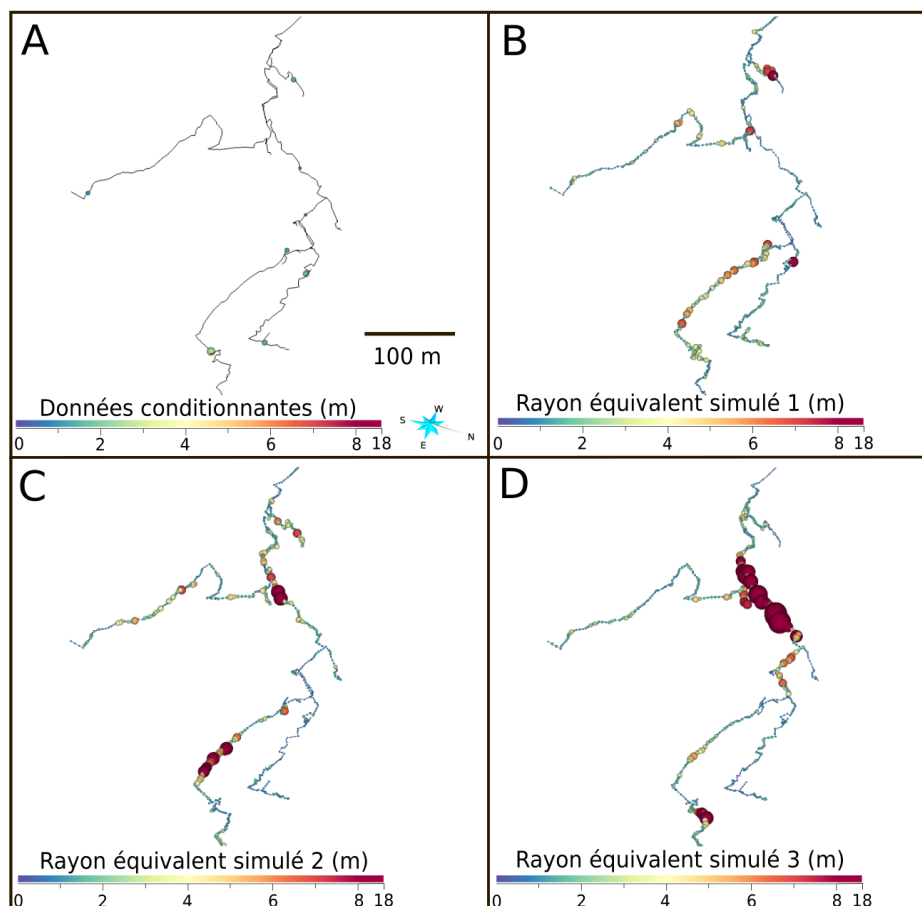


Figure 3.9 – A : Partie du réseau d'Arrestelia et données conditionnantes (les nœuds correspondant sont légèrement agrandis par soucis de visibilité). B-D : 3 exemples de résultats obtenus sur cette partie par simulation conditionnelle via la méthode de SGS 1D-curvilinéaire sur le réseau entier (usage de données conditionnantes). Le rayon équivalent affiché a été calculé à partir des valeurs logarithmiques simulées. Les données conditionnantes étant très éloignées, nous avons préféré présenter une zone relativement grande, rendant difficile l'observation du respect des données conditionnantes sur les résultats.

de 0.076, une médiane de 0.098 et une variance de 0.079. La distribution d'entrée, les variogrammes et la taille maximale du voisinage utilisés pour les simulations sont les mêmes que ceux utilisés pour les simulations non-conditionnelles.

Les valeurs statistiques des résultats sont légèrement différentes de celles des simulations non-conditionnelles, avec une moyenne de 0.094, une médiane de 0.076 et une variance de 0.089. Les histogrammes et variogrammes (Figures 3.10; 3.11) sont légèrement différents de ceux des simulations non-conditionnelles, mais ces écarts sont minimes et peuvent être en partie dus aux données conditionnantes qui ont des valeurs statistiques différentes de la distribution initiale.

Un résumés des résultats est visible dans le Tableau 3.1.

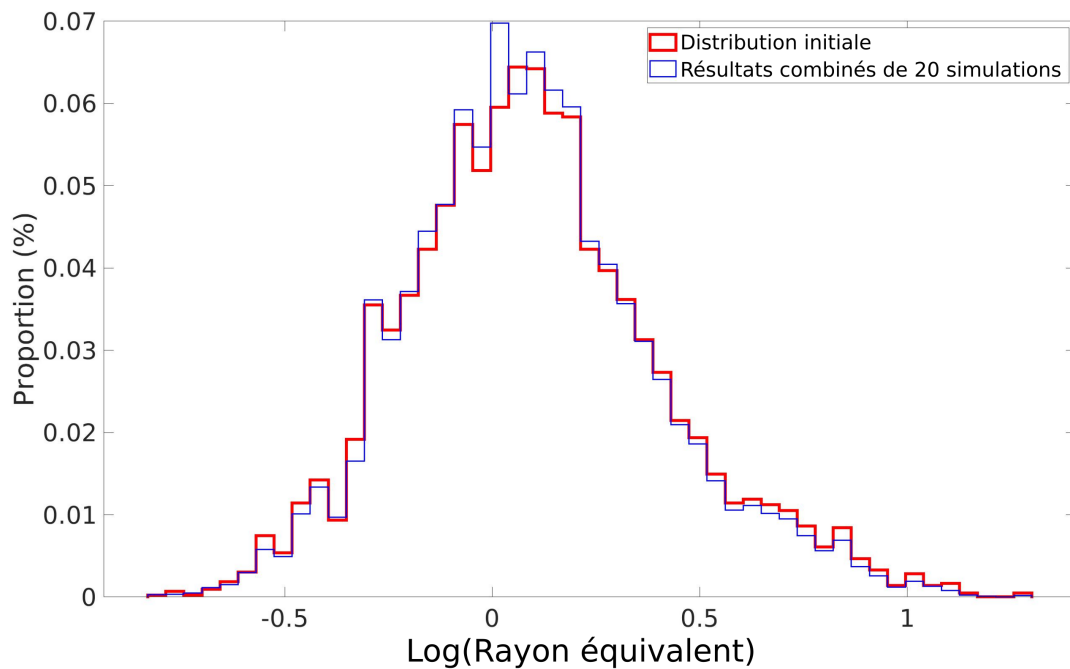


Figure 3.10 – Histogramme de la somme des résultats de 20 simulations conditionnelles effectuées sur le réseau d'Arrestelia avec la méthode de SGS 1D-curveilinéaire et distribution initiale du logarithme décimal du rayon équivalent.

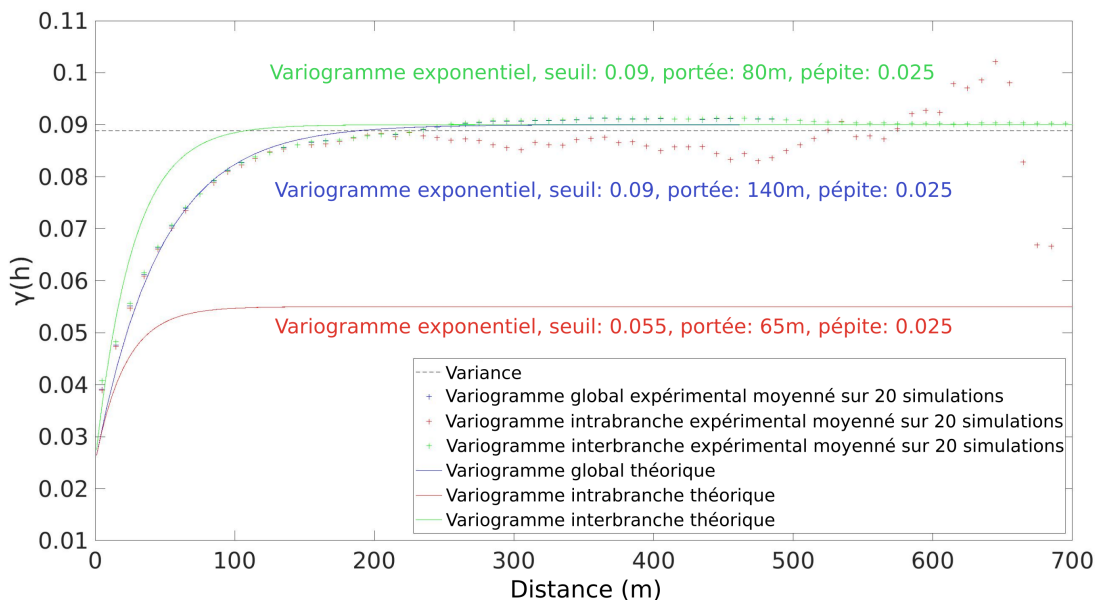


Figure 3.11 – Variogrammes expérimentaux moyennés des résultats de 20 simulations conditionnelles effectuées sur le réseau d'Arrestelia avec la méthode de SGS 1D-curveilinéaire et variogrammes théoriques du logarithme décimal du rayon équivalent. Les variogrammes intrabranche et interbranche théoriques n'ont pas été utilisés lors des simulations mais sont donnés à titre indicatif.

	Moyenne	Médiane	Variance
Distribution d'entrée	0.101	0.081	0.098
Données conditionnantes	0.076	0.098	0.079
Résultats des simulations non-conditionnelles	0.101	0.082	0.090
Résultats des simulations conditionnelles	0.094	0.076	0.089

Tableau 3.1 – Comparaison des valeurs statistiques de la distribution d'entrée et des moyennes calculées sur les résultats des simulations non-conditionnelles et conditionnelles, respectivement 20 de chaque, effectuées sur le réseau d'Arrestelia avec la méthode de SGS 1D-curvilinéaire.

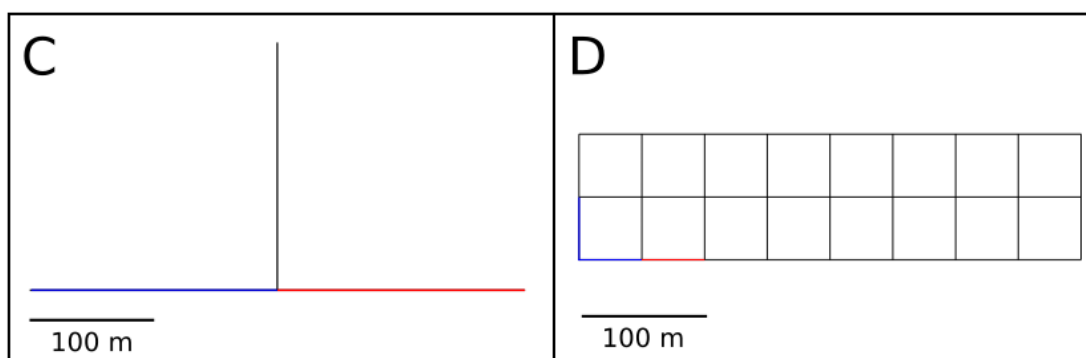


Figure 3.12 – Cas synthétiques C et D. Sur chaque cas, une branche est colorée en rouge, une autre en bleu et les autres en noires.

3.3.2 SGS 1D-curvilinéaire *branchwise*

Cas synthétiques

La méthode de SGS 1D-curvilinéaire *branchwise* a été testée sur deux autres cas synthétique (Figure 3.12), l'objectif étant ici de reproduire les comportements des 3 variogrammes théoriques. Le cas synthétique C correspond à l'intersection de trois segments, chacun long de 200 mètres. Le cas synthétique D correspond à une grille de 400 mètres de long et 100 mètres de large, le côté des cellules faisant 50 mètres de long. Ils sont densifiés à hauteur d'un nœud tous les mètres, pour un total de 601 nœuds pour le cas C et 2087 nœuds pour le cas D. La longueur des branches du cas C représente le triple de la portée du variogramme intrabranche, tandis que les longueurs de la majorité des branches du cas D lui sont légèrement inférieures. En plus des paramètres utilisés auparavant, les variogrammes interbranches et intrabranche calculés sur la transformée du logarithme décimal du rayon équivalent servent également de paramètres d'entrée. La proportion *prop* et le nombre maximal de nœuds simulés *nbnoeudsmax* par branche ont été choisis de façon à simuler 10% des nœuds de chaque branche lors de l'étape de conditionnement interbranche (10% et 20 nœuds pour le cas C et 10% et 5 nœuds pour le cas D).

Dans ces deux cas synthétiques, les distributions des résultats sont plus centrées autour de la moyenne que la distribution d'entrée (Figure 3.13). L'utilisation du variogramme intrabranche conduit en effet à une variance plus basse. A cause de cette

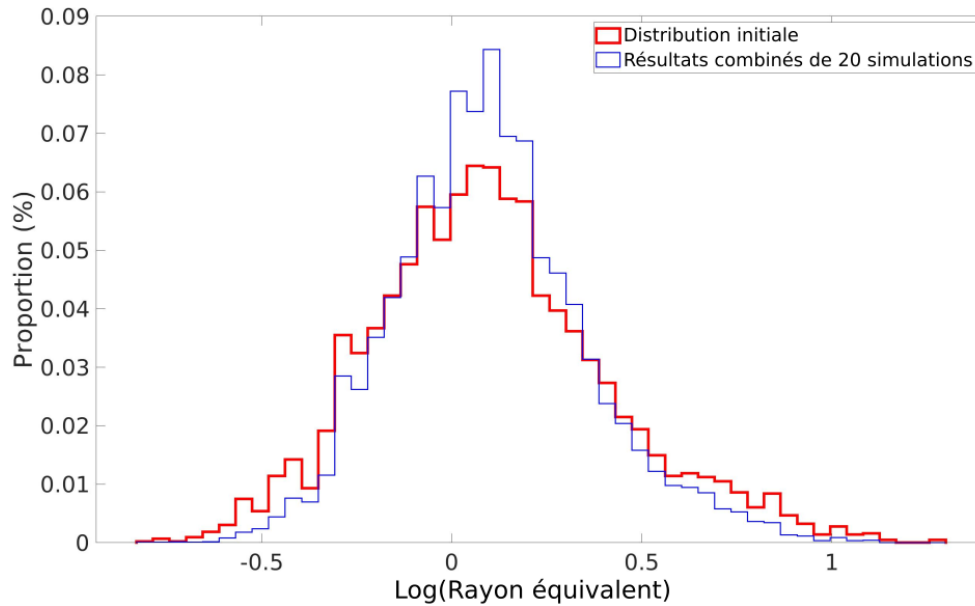


Figure 3.13 – Histogramme de la somme des résultats de 20 simulations non-conditionnelles effectuées sur le cas synthétique C avec la méthode de SGS 1D-curvilinéaire *branchwise* et distribution initiale du logarithme décimal du rayon équivalent utilisée comme paramètre d’entrée. La densité utilisée est d’un nœud tous les 1m pour 601 nœuds au total.

variance plus faible, les moyennes et médianes des résultats sont sensiblement plus proches de celles de la distribution d’entrée que celles obtenues avec la méthode précédente. Ces cas étant composés de plusieurs branches différentes, les variogrammes globaux et intrabranches expérimentaux sont bien distinguables et il est possible de calculer les variogrammes interbranches. La Figure 3.14 montre que les valeurs des variogrammes globaux et intrabranches sont très proches mais surtout que les valeurs des variogrammes interbranches sont comparativement bien plus grandes à faible distance. À très faible distance, les variogrammes globaux sont confondus avec les variogrammes intrabranches, tandis qu’à grande distance, ils se confondent avec les variogrammes interbranches. Pour le cas synthétique C, le faible nombre de nœuds, le faible nombre de branches et leur très grande taille conduit néanmoins à des variogrammes relativement erratiques, résultant en des valeurs de variogrammes intrabranches plus grandes que les valeurs des variogrammes interbranches par endroit.

Le cas synthétique D a principalement été créé pour permettre une bonne visualisation des résultats. Mise à part le fait que la majorité des branches fassent 50 mètres (les branches aux coins de la grille faisant 100 mètres), ce qui limite la distance maximale de calcul, les variogrammes intrabranches expérimentaux sont semblables à ceux du cas C tout en étant moins erratiques. Les simulations conditionnelles permettent bien d’imposer des tendances différentes sur chacune des branches comme cela peut-être vu en Figure 3.15C. Une meilleure discontinuité entre les branches au niveau du nœud d’intersection que lors de l’utilisation de la méthode de simulation précédente semble observable, bien qu’il soit difficile de réellement la mettre en évidence. La rupture n’est de plus pas supposée être radicale grâce à la prise en compte du variogramme interbranche.

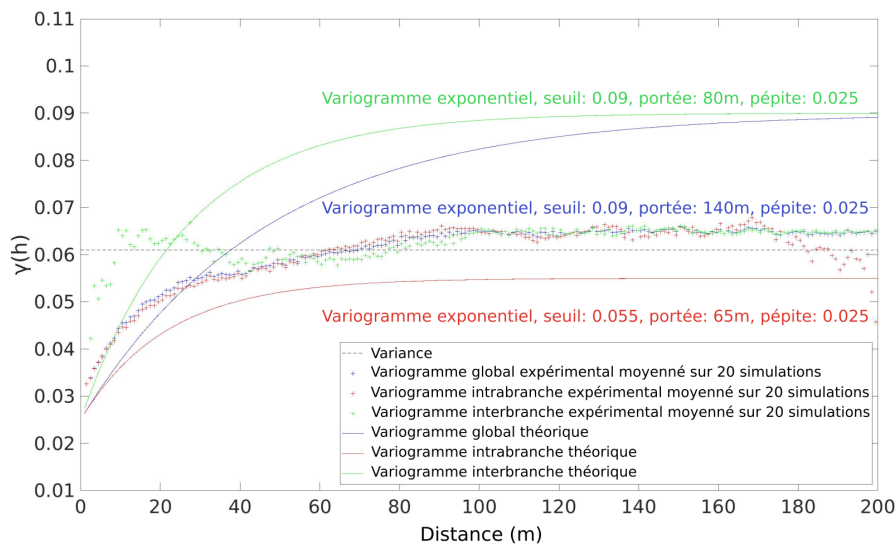


Figure 3.14 – Variogrammes expérimentaux moyenné des résultats de 20 simulations non-conditionnelles effectuées sur le cas synthétique C (601 nœuds) avec la méthode de SGS 1D-curvilinéaire *branchwise* et variogramme théorique du logarithme décimal du rayon équivalent.

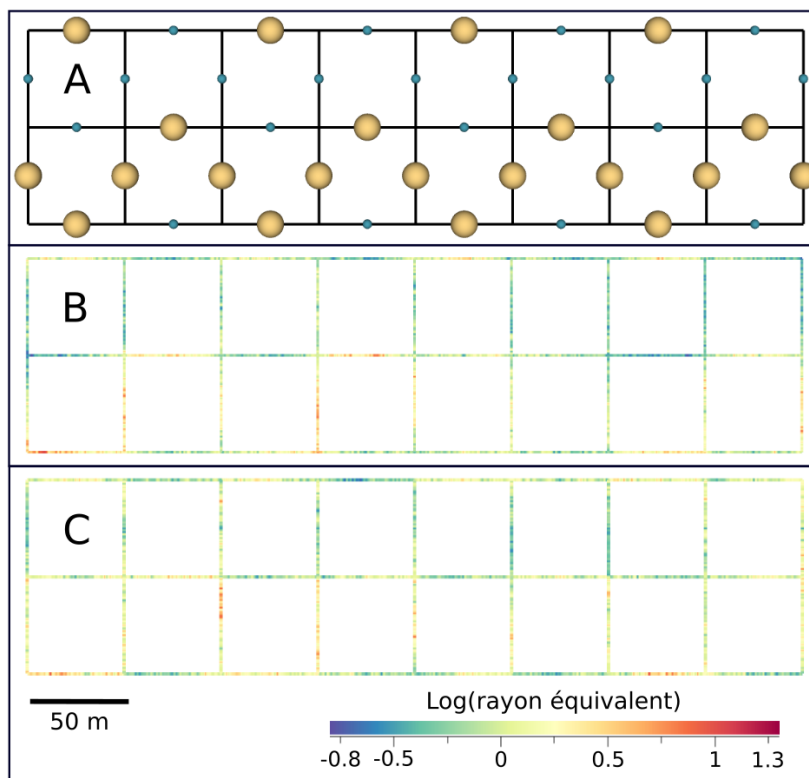


Figure 3.15 – A : Données conditionnantes pour le cas synthétique D, la taille des sphères est proportionnelle à leur valeur. B : Exemple de résultat obtenu par simulation conditionnelle via la méthode de SGS 1D-curvilinéaire. C : Exemple de résultat obtenu par simulation conditionnelle via la méthode de SGS 1D-curvilinéaire *branchwise*. Tous ces résultats sont obtenus pour une densité d'un nœud tous les mètres, conduisant à un total de 2087 nœuds. Les différences de valeur au niveau des points d'intersection semblent plus distinctes en C, bien qu'il soit difficile de le confirmer.

Réseau d'Arrestelia

De nombreuses simulations non-conditionnelles du logarithme décimal du rayon équivalent ont été effectuées sur le réseau d'Arrestelia via l'utilisation de la méthode de SGS 1D-curvilinéaire *branchwise* (e.g., Figure 3.16). La distribution d'entrée, les variogrammes et la taille du voisinage sont les mêmes que ceux utilisés pour la première méthode. Après avoir testé différentes valeurs de proportion *prop* et de nombre de nœuds *nbnoeudsmax* simulés via l'utilisation du variogramme interbranche dans chaque branche pendant l'étape de préconditionnement, nous avons observé que les résultats reproduisant le mieux la double variabilité étaient obtenus avec un nombre maximal de 15 nœuds par branche et une proportion de 100% comme paramètre pour l'étape de conditionnement interbranche. Cela signifie que si une branche a plus de 15 nœuds sans données, 15 d'entre eux sont simulés avec le variogramme interbranche, tandis que les autres sont simulés avec le variogramme intrabranché (cf l'Algorithme 2). Dans le cas contraire, ils sont tous simulés avec le variogramme interbranche. 20 simulations ont été effectuées avec ces paramètres.

Les distributions des résultats sont proches de la distribution initiale mais moins qu'avec la méthode précédente (Figure 3.17). Les moyennes et médianes restent très proches de celles de la distribution d'entrée : la moyenne des résultats est de 0.097 au lieu de 0.101, tandis que leur médiane est de 0.080 au lieu de 0.081. Les distributions sont cependant plus centrées autour de la moyenne à cause d'une importante baisse de la variance, qui est de 0.075 au lieu de 0.098.

Comme le variogramme intrabranché contribue plus que dans la méthode précédente et que son seuil est plus bas que le seuil du variogramme global, il est naturel d'observer une variance résultante plus faible. Ce phénomène peut être observé dans les variogrammes calculés (Figure 3.18). Le variogramme intrabranché théorique tend en effet à être plus respecté (même si les variogrammes expérimentaux sont loin d'être parfaitement calés dessus) mais les valeurs des variogrammes globaux et interbranches sont significativement plus basses que celles de leurs variogrammes théoriques. Les différences entre les variogrammes intrabranches et les variogrammes globaux et interbranches sont clairement visibles, ce qui prouve que la variabilité simulée à l'échelle des branches et celle à l'échelle du réseau sont bien distinctes.

Nous avons également effectué des simulations conditionnelles de la même manière que précédemment et les résultats sont également proches de ceux des simulations non-conditionnelles. Les valeurs statistiques des résultats des simulations conditionnelles et non-conditionnelles sont présentées dans le Tableau 3.2.

D'autres simulations effectuées avec différentes proportions et nombres maximaux de nœuds simulés par branche avec le variogramme interbranche ont également été effectuées. Les valeurs statistiques des résultats sont présentées dans le Tableau 3.3 et certains histogrammes et variogrammes sont comparés en Annexe D.

De manière générale, plus le nombre de nœuds simulés dans chaque branche en se basant sur le variogramme interbranche est élevé, plus les variogrammes expérimentaux des résultats s'approchent des variogrammes globaux et interbranches théoriques. Quand seule la proportion est utilisée, le variogramme intrabranché est très proche des deux autres, témoignant d'une quasi-absence de comportement spécifique au sein des

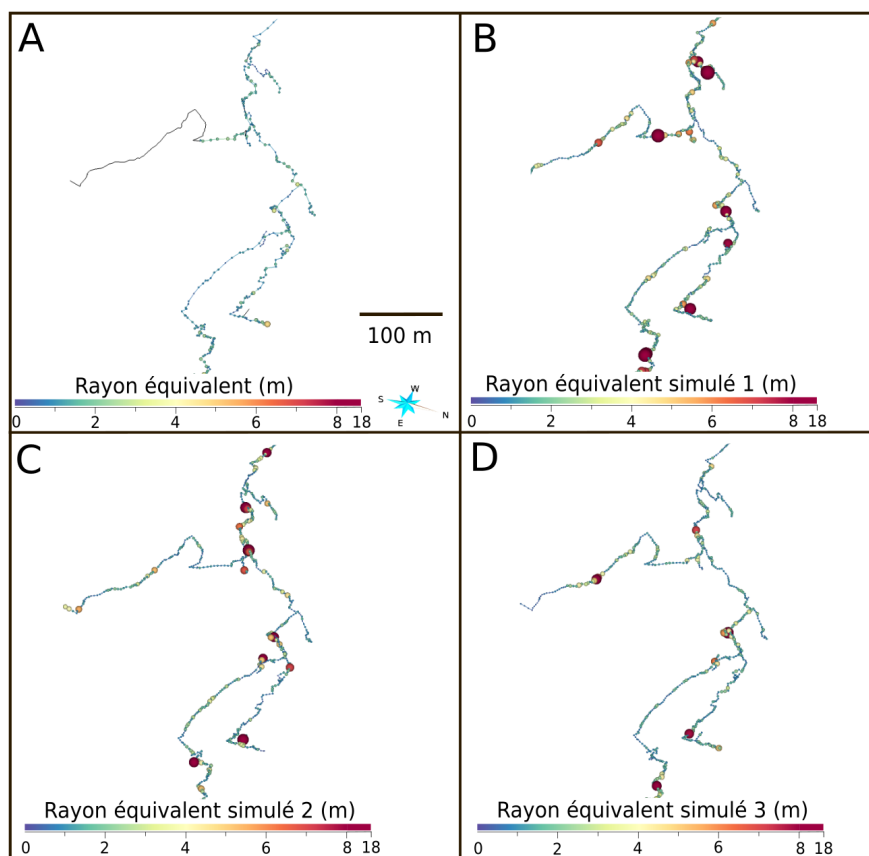


Figure 3.16 – A : Partie du réseau d’Arrestelia initial. B-D : 3 exemples de résultats obtenus sur cette partie par simulation non-conditionnelle via la méthode de SGS 1D-curvilinéaire *branchwise* sur le réseau entier (aucune donnée conditionnante n’est utilisée). Le rayon équivalent affiché a été calculé à partir des valeurs logarithmiques simulées.

branches (Figure D.1). A contrario, quand seul le nombre maximal de nœuds est utilisé, les valeurs des variogrammes intrabranches expérimentaux sont clairement inférieures à celles des variogrammes interbranches et globaux, témoignant de l’existence d’un double comportement, à l’échelle du réseau et à l’échelle des branches (Figure D.2). Lorsque ces deux paramètres sont utilisés, le double comportement est mieux observable que dans le cas où seule la proportion est utilisée mais moins que dans le cas de l’utilisation du seul nombre maximal de nœuds (Figure D.3).

Concernant l’étude des valeurs statistiques, la variance augmente avec la proportion et le nombre de nœuds maximal, comme cela était attendu. Les moyennes et médianes sont également croissantes mais une légère baisse est cependant observable lors du passage de 5 à 7 nœuds maximaux. Cela est probablement causé par la taille des branches du réseau mais la raison exacte est difficile à déterminer. L’étude des histogrammes n’apporte pas d’informations particulières. Moins le nombre de nœuds simulés via le variogramme interbranche est important, plus ils sont centrés autour des valeurs des médianes et moyennes. Cela mène à un pic plus important au centre et des proportions plus faibles sur les côtés (Figure D.4). Inversement, plus le nombre de nœuds est élevé, plus les histogrammes ont tendance à s’approcher de la distribution initiale.

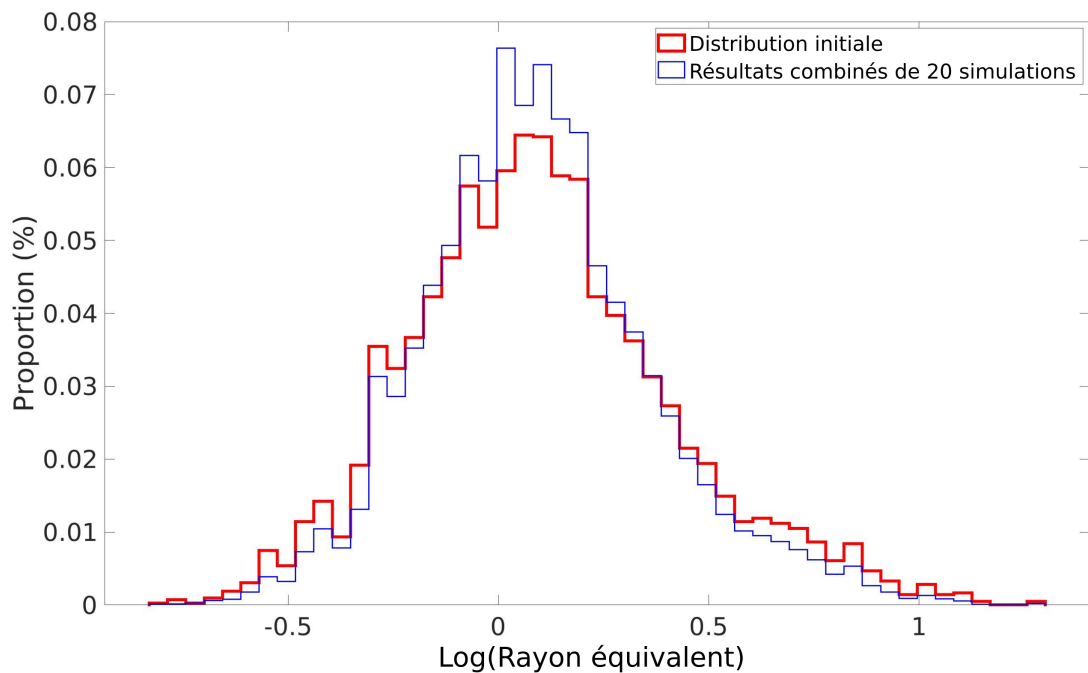


Figure 3.17 – Histogramme de la somme des résultats de 20 simulations non-conditionnelles effectuées sur le réseau d'Arrestelia avec la méthode de SGS 1D-curvilinéaire *branchwise* et distribution initiale du logarithme décimal du rayon équivalent.

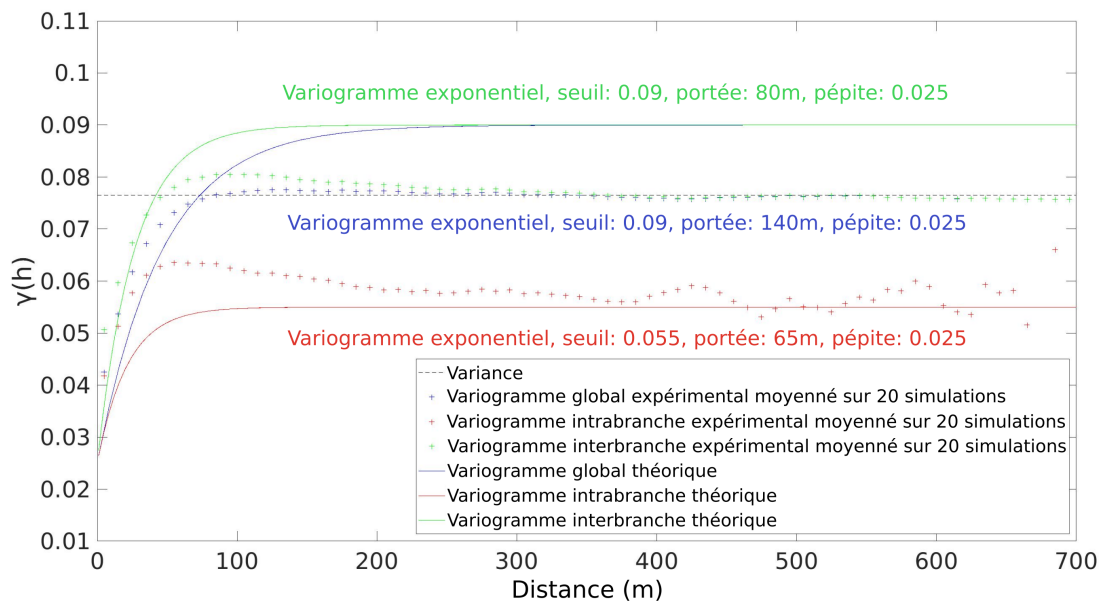


Figure 3.18 – Variogrammes expérimentaux moyennés des résultats de 20 simulations non-conditionnelles effectuées sur le réseau d'Arrestelia avec la méthode de SGS 1D-curvilinéaire *branchwise* et variogrammes théoriques du logarithme décimal du rayon équivalent.

	Moyenne	Médiane	Variance
Distribution d'entrée	0.101	0.081	0.098
Données conditionnantes	0.076	0.098	0.079
Résultats des simulations non-conditionnelles	0.097	0.080	0.075
Résultats des simulations conditionnelles	0.094	0.079	0.075

Tableau 3.2 – Comparaison des valeurs statistiques de la distribution d'entrée et des moyennes calculées sur les résultats des simulations non-conditionnelles et conditionnelles, respectivement 20 de chaque, effectuées sur le réseau d'Arrestelia avec la méthode de SGS 1D-curvilinéaire *branchwise*.

Proportion	Nombre maximal de nœuds	Moyenne	Mediane	Variance
Distribution d'entrée		0.101	0.081	0.098
/	1	0.090	0.079	0.054
/	3	0.092	0.079	0.060
/	5	0.095	0.082	0.064
/	7	0.091	0.077	0.067
/	9	0.093	0.078	0.070
/	15	0.097	0.080	0.075
/	30	0.099	0.082	0.082
1	/	0.091	0.080	0.054
5	/	0.091	0.079	0.055
10	/	0.091	0.079	0.057
20	/	0.092	0.078	0.063
50	/	0.097	0.081	0.076
10	7	0.091	0.077	0.057
10	15	0.092	0.079	0.058
20	7	0.093	0.082	0.060
20	15	0.093	0.079	0.062

Tableau 3.3 – Comparaison des valeurs statistiques de la distribution d'entrée et des moyennes calculées sur les résultats des simulations non-conditionnelles effectuées avec la méthode de SGS 1D-curvilinéaire *branchwise*, pour différentes proportions et nombres de nœuds maximaux simulés dans chaque branche avec le variogramme interbranche (20 simulations par ensemble de paramètres).

3.4 Discussion

Deux approches sont proposées pour simuler des propriétés le long des conduits de réseaux karstiques complexes, tout en reproduisant les informations statistiques d'entrée et les données conditionnantes. La démonstration de l'existence d'une faible variabilité spatiale des propriétés géométriques des conduits à courte distance dans le Chapitre 2 est le facteur clé ayant pu permettre leur développement. Ces deux approches honorent la distribution d'entrée. La méthode de SGS 1D-curvilinéaire offre des résultats intéressants quant à la reproduction du variogramme global mais ne permet pas de rendre compte des différences de comportement observées au sein des branches lors de l'analyse statistique présentée dans le Chapitre 2. La méthode de SGS 1D-curvilinéaire *branchwise*, qui s'appuie sur ces différences de comportement, prend en compte les trois variogrammes, bien que le variogramme global n'intervienne que très peu. Les résultats reproduisent correctement le comportement du variogramme intrabranche théorique mais cela induit des valeurs plus faibles pour les variogrammes globaux et interbranches, ainsi qu'une distribution plus centrée autour de la moyenne. Le comportement à l'échelle du réseau et le comportement à l'échelle des conduits peuvent cependant être simulés de façon distincte.

Il est actuellement compliqué de trouver les bons paramètres à utiliser afin de pouvoir reproduire cette double variabilité. Les résultats obtenus sur le réseau d'Arrestelia semblent indiquer que la meilleure piste serait d'utiliser uniquement le nombre maximal de nœuds simulés via le variogramme interbranche, sans prendre en compte la proportion. Cela présente néanmoins quelques limitations conceptuelles car dans ce cas, contrairement aux grandes branches, les petites sont entièrement simulées via le variogramme interbranche. Ainsi, la variabilité intrabranche n'est respectée que sur les grandes branches, tandis que la variabilité interbranche est majoritairement respectée dans les petites branches. Bien que les variogrammes obtenus de cette manière soient relativement satisfaisants, il est dur de juger le succès des simulations. Dans l'état actuel de la méthode *branchwise*, la prise en compte du variogramme interbranche théorique empêchera toujours de reproduire parfaitement le comportement intrabranche et ce même sur une unique branche (e.g., Figure D.5).

Le caractère unidimensionnel des simulations peut conduire à un effet écran plus important que lors de l'utilisation de SGS bi- ou tridimensionnelles. Il a par exemple été observé lors de simulations sur des réseaux réels non densifiés que les voisins les plus proches du nœud simulé pouvaient avoir un poids extrêmement important comparativement aux voisins plus éloignés. Ainsi, les poids cumulés des deux voisins les plus proches d'un nœud pouvaient dépasser 80 voire 90% du poids total des voisins.

Les méthodes présentées utilisent actuellement un krigeage simple, ce choix ayant été motivé par une bonne reproduction des informations statistiques connues. L'utilisation d'un krigeage ordinaire est également envisagée car elle permettrait potentiellement d'obtenir une meilleure cohérence locale, au risque de moins bien respecter les données d'entrée (Emery, 2004), mais cela n'a pas encore pu être testé.

Les difficultés à reproduire les variogrammes intrabranches et interbranches simultanément sont directement liées à l'implémentation actuelle de notre méthode de SGS 1D-curvilinéaire *branchwise*. Si l'on considère leur reproduction stricte comme une priorité, un travail supplémentaire sur cet algorithme est nécessaire. D'autres méthodes,

également basées sur les SGS, ont pu être développées et testées durant cette thèse mais leurs résultats n'étaient pas suffisamment bons pour les garder. L'adaptation d'autres méthodes de simulation utilisées en géostatistique comme les méthodes de simulations spectrales, de moyennes mobiles ou de *turning bands* au cas présent pourrait s'avérer intéressant. Elles sont en effet connues pour permettre une bonne reproduction des variogrammes d'entrée, bien que les méthodes de SGS semblent particulièrement bien adaptées pour la reproduction de variogrammes ayant une faible portée et un effet pépite (Pinto et Deutsch, 2020). Le développement de méthodes basées sur les Simulations Multipoints (MPS) est également envisagé, même si elles ne permettent pas de reproduire les variogrammes.

Les résultats obtenus sont meilleurs sur le réseau d'Arrestelia que sur les cas synthétiques car les paramètres d'entrée ont directement été évalués sur ce dernier et son nombre de nœuds plus important induit une meilleure robustesse. La différence de topologie, de nombre de nœuds et de densité de nœuds impacte forcément les simulations. Le faible nombre de nœuds des cas synthétiques conduit notamment à des variogrammes relativement erratiques passé une certaine distance. Pour autant, effectuer des simulations sur des cas beaucoup plus densifiés que la distance entre les points sur lesquels ont été recueillies les données n'est pas pertinent à cause de la différence d'échelle. L'effet pépite des variogrammes calculés sur des points distants de 5 mètres sera généralement plus important que celui devant être utilisé pour des points distants de 0.1 mètres. Les tests montrent effectivement qu'une densité de nœuds plus élevée ne produit pas forcément des résultats plus satisfaisants (e.g., Figure D.6). La faible taille des cas synthétiques comparée à la portée des variogrammes conduit également à un faible nombre de tirage de nœuds dans la distribution initiale, ce qui est une potentielle cause de différence.

La méthode d'analyse des résultats elle-même pourrait être améliorée. Les résultats peuvent en effet beaucoup varier d'une réalisation à une autre. Les histogrammes et variogrammes présentés dans cette thèse sont des moyennes effectuées sur 20 réalisations différentes. L'utilisation de barres d'erreur ou de boîtes similaires à celles des diagrammes en boîte pourrait permettre de mieux quantifier ces différences sur les variogrammes expérimentaux.

Les principaux avantages des deux méthodes de simulation présentées sont les suivants :

- Elles ne sont pas basées sur une grille régulière ou cartésienne. Cela permet des temps de calculs plus courts et donne à l'utilisateur la possibilité de définir librement la résolution des réseaux. Néanmoins, les temps de calcul peuvent quand même être importants pour les grands réseaux (environ 30 secondes par simulation non-conditionnelle pour simuler environ 15000 nœuds sur un ordinateur portable standard).
- Elles permettent de respecter la variabilité spatiale des propriétés simulées au sein de chaque conduit au lieu d'utiliser une seule valeur pour le conduit entier.
- Elles utilisent un squelette défini par l'utilisateur comme support et il n'y a pas de dépendance entre sa création et la génération des propriétés. Cela donne à

l'utilisateur le choix de travailler sur des réseaux déjà explorés ou sur des réseaux simulés.

- Elles sont stochastiques, ce qui permet d'avoir un bon aperçu des incertitudes concernant la géométrie des conduits.
- Ces méthodes peuvent être appliquées à toute propriété qui présenterait une corrélation spatiale le long d'un réseau curvilinéaire ou d'un graphe.

3.5 Conclusion

Les méthodes de simulation présentées permettent de reproduire, ne serait-ce que partiellement, la complexité des réseaux karstiques grâce à l'usage des géostatistiques. Ces algorithmes permettent la génération d'un rayon équivalent et d'un ratio largeur-hauteur proches des observations réelles présentées dans le Chapitre 2. Tandis que les distributions d'entrée sont respectées, les variogrammes d'entrée ne sont pour le moment pas parfaitement reproduits suivant la méthode de simulation. La méthode de SGS 1D-curvilinéaire permet de reproduire le variogramme global mais le comportement sur les branches n'est pas respecté. D'un autre côté, la méthode de SGS 1D-curvilinéaire *branchwise* permet une meilleure reproduction du variogramme intrabranche mais les variogrammes globaux et interbranches ont des valeurs plus faibles qu'elles ne devraient l'être.

Ces résultats sont malgré tout un important progrès vers la simulation de réseaux ayant des géométries de conduits *réalistes*. Ces méthodes pourront ainsi être utilisées dans différents contextes, notamment pour les simulations d'écoulement dans des réseaux de conduits discrets (e.g., SWMM, Epanet, Modflow-CFP).

Chapitre 4

Simulation du squelette des réseaux

Sommaire

4.1	Introduction	110
4.2	Concept de la méthode de simulation des squelettes	111
4.3	État actuel de la méthode	116
4.3.1	Calcul de la connectivité	116
4.3.2	Connexion des nœuds	119
4.4	Résultats	119
4.5	Discussion	123
4.6	Conclusion	129

Résumé Ce chapitre présente les fondements d’une nouvelle méthode visant à simuler stochastiquement des réseaux karstiques discrets sous forme de graphes. Elle est conçue de sorte à pouvoir prendre directement en compte les données de terrain (e.g., entrées et sorties connues des réseaux) et utiliser des informations géologiques et hydrogéologiques comme données conditionnantes. La connectivité du réseau est imposée via un ratio entre le nombre de nœuds d’intersections et d’extrémités. Contraindre ainsi la connectivité devrait permettre de pouvoir à terme simuler différentes morphologies de réseaux (e.g., arborescent, anastomosé). La première étape de cette méthode consiste à définir une grille de probabilité dans la zone étudiée et à tirer à l’intérieur un nombre de points choisi par l’utilisateur. La probabilité associée à chaque voxel dépend directement des données conditionnantes et des paramètres d’entrée. Le nombre de connexions attendues pour chaque point, lorsqu’il sera nœud du graphe, est ensuite calculé en respectant le ratio intersections/extrémités et d’autres paramètres donnés en entrée. Enfin, les connexions entre les nœuds sont générées automatiquement. Cette méthode n’en est encore qu’aux premiers stades de son développement mais ses principes généraux, son état actuel et des exemples d’application sont présentés et commentés.

4.1 Introduction

Comme vu dans les chapitres précédents, il est possible d’effectuer des simulations d’écoulement sur des réseaux de conduits discrets en se servant de différents logiciels (e.g., SWMM, Epanet, Modflow-CFP). Des méthodes permettant de simuler les propriétés géométriques des conduits, de façon à reproduire les observations de terrain, ont été présentées dans le Chapitre 3. Pour autant, ces propriétés sont attachées à des réseaux de drains qui, s’ils ne sont pas connus, doivent être simulés. Il s’agit de l’objet du présent chapitre.

Comme évoqué au Chapitre 1, différentes approches ont été proposées ces vingt dernières années pour simuler les squelettes de réseaux karstiques (e.g., Jaquet et al., 2004; Pardo-Igúzquiza et al., 2012; Borghi et al., 2012; Collon-Drouaillet et al., 2012; Fournillon et al., 2012; Viseur et al., 2014; Hendrick et Renard, 2016a). Bien que chacune d’entre elles propose une solution pertinente, certaines limites nous ont motivés à imaginer une méthode alternative.

La principale contrainte que nous nous sommes fixée est celle d’une méthode unique, utilisable quelle que soit la topologie éventuellement prédéfinie du réseau. Le simulateur idéal devrait permettre de générer indifféremment des réseaux à dominante arborescente ou anastomosée, avec une influence ou non de la fracturation sur l’orientation des conduits. Ces *motifs morphologiques* devraient même pouvoir varier localement dans le domaine d’étude. Il serait ainsi possible de reproduire la diversité observée au sein même des réseaux réels et reflétant la spéléogenèse polyphasée des systèmes naturels (Jouves et al., 2017). D’autres conditions nous semblent également importantes pour la simula-

tion de réseaux *réalistes*. Une méthode efficace devrait être conçue de sorte à générer des réseaux tridimensionnels, qui se déploient dans tout l'espace et pas uniquement sur certains plans reliés par des drains purement verticaux. Par ailleurs, les méthodes actuelles reposent habituellement sur une discrétisation structurée de l'espace. La simulation de réseaux de grande taille implique alors la manipulation de larges grilles, ce qui augmente considérablement les temps de calcul lorsque l'on travaille en 3D et engendre des problèmes d'*aliasing* : les squelettes résultants sont alignés sur les axes de la grille. Ces *marches d'escalier*, visibles lorsque l'on zoome sur les réseaux, sont connues pour influencer sur les résultats de simulation d'écoulement. Une motivation au développement d'une nouvelle méthode de simulation est donc de pouvoir générer des réseaux de conduits discrets qui ne soient pas contraints à passer par les arêtes d'une grille. Bien qu'elles soient souvent partielles, les informations connues de l'utilisateur doivent pouvoir contraindre le processus de simulation : i) localisation des entrées et sorties du système, ii) portions explorées et cartographiées du réseau, iii) indices de karstification visibles en surface (e.g., dolines), iv) horizons d'inception et fracturation du massif, v) informations hydrogéologiques. La gestion des incertitudes implique d'avoir une approche stochastique, ce qui requiert des temps de calculs raisonnables. L'approche présentée dans ce chapitre est un premier pas dans cette direction.

La première partie décrit les trois grandes étapes de cette méthode : i) la génération d'un nuage de points, ii) le calcul de leur connectivité et iii) leur connexion. Je présente ensuite l'implémentation ayant pu être réalisée dans la deuxième partie et les premiers résultats dans la troisième partie. La pertinence et les limites actuelles de la méthode, ainsi que plusieurs perspectives sont discutées dans la quatrième partie.

4.2 Concept de la méthode de simulation des squelettes

La méthode de simulation du squelette des réseaux karstiques développée ici se focalise sur les aspects suivants :

- Simulation de réseaux tridimensionnels, non alignés sur une grille
- Intégration des données de terrain (localisation de conduits connus) et des informations géologiques et hydrogéologiques.
- Contrôle de la topologie des réseaux par l'utilisateur.
- Stochasticité.
- Temps de calcul réduits, de sorte à ce que la simulation de cent réseaux de plusieurs centaines de nœuds prenne moins d'une heure.

La méthode proposée vise à simuler le squelette des réseaux sous forme de graphes tels que définis dans le Chapitre 2. Les dimensions des conduits pourront être simulées ultérieurement grâce aux méthodes présentées dans le Chapitre 3.

Je propose de diviser la simulation des squelettes en trois grandes étapes : i) la génération d'un nuage de points, ii) le calcul de leur connectivité et iii) leur connexion pour former le squelette (Figure 4.1).

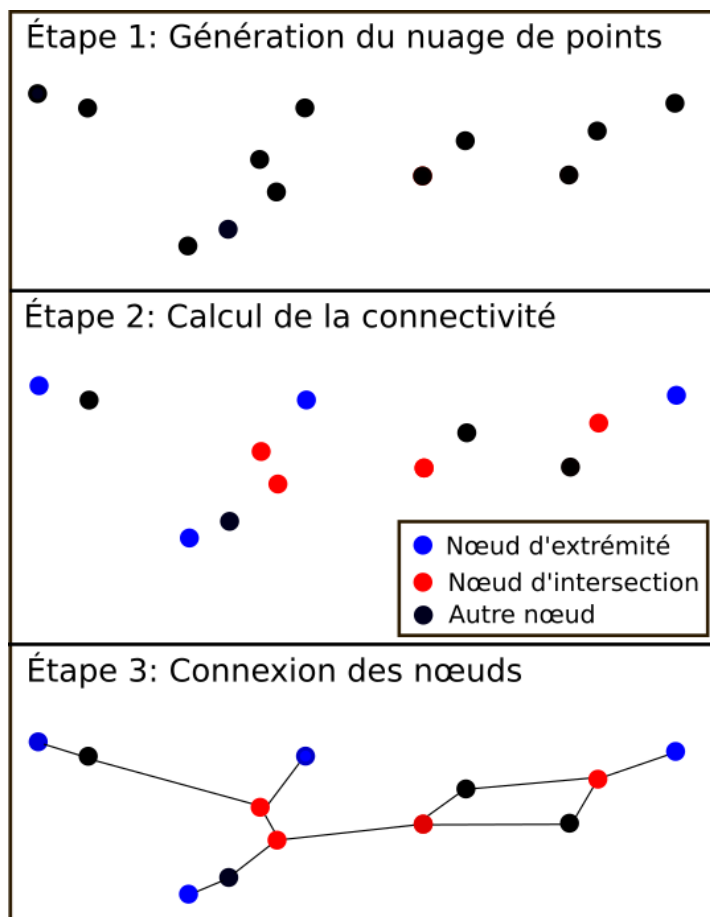


Figure 4.1 – Étapes de la méthode de simulation. A : Exemple de génération d'un nuage de points. B : Exemple de connectivité calculée pour un nombre d'extrémités égal à 4 et un ratio intersections/extrémités égal à 1. La connectivité est assimilable au degré théorique qu'aura le point en tant que nœud du graphe. C : Exemple de graphe obtenu après connexion des nœuds. Ce graphe correspond au squelette du réseau.

La première étape n'est pas implémentée mais elle consisterait à générer aléatoirement un nombre voulu de *points* dans une grille de probabilité. Comme les points générés dans cette étape ne seront pas nécessairement intégrés au réseau simulé, on nommera *points* les vertex générés lors de la première étape et *nœuds* ceux faisant partie du graphe final. Cette étape commencera par la création de la grille de probabilité à partir de diverses informations (Figure 4.2). Chaque voxel sera ainsi associé à une probabilité de présence de points. Le calcul des probabilités se décomposerait suivant deux principaux aspects qui seraient combinés : i) l'intégration des données et ii) le remplissage automatique de la grille :

- L'intégration des données comprend à la fois l'intégration des données de terrain (e.g., position des conduits connus) et l'intégration des informations géologiques et hydrogéologiques (e.g., horizons d'inception, failles, altitude du niveau piézométrique). Les conduits connus devraient être directement intégrés comme partie du réseau et leur présence dans un voxel de la grille devrait également augmenter les probabilités associées au voxel en question et à ses voisins. En effet, les conduits sont selon toute vraisemblance reliés à d'autres conduits proches et il

faudra être en mesure de les connecter au réseau généré. Les données géologiques devraient moins impacter la grille de probabilité que les données de terrain mais nous souhaiterions laisser la possibilité à l'utilisateur de choisir l'importance de leur influence. Par exemple, si de nombreux conduits sont observés à l'interface entre 2 couches géologiques, il est envisageable de la considérer comme un horizon d'inception et de lui assigner une valeur d'influence de 10. En comparaison, une faille sur laquelle aucun conduit n'a pu être observé pourrait se voir assigner une valeur de 2. Il est plus probable que des conduits se développent le long de cette dernière que dans la matrice mais la probabilité reste à priori faible. Si plusieurs structures sont présentes dans le même voxel, les probabilités associées s'additionneront.

- Le remplissage automatique de la grille devra générer les probabilités des voxels n'étant pas associés à des données à partir de paramètres choisis par l'utilisateur. Ainsi, nous souhaiterions être en mesure de générer des grilles entières même sans avoir accès à des données. Les probabilités calculées de façon automatique s'apparenteraient à un bruit de fond pouvant se superposer aux probabilités calculées via l'intégration des données.

Après création de la grille, autant de voxels que de points voulus sont tirés aléatoirement, en accord avec les probabilités associées, le même voxel pouvant être tiré plusieurs fois (Figure 4.3B). La position des points est ensuite tirée aléatoirement au sein de chacun de ces voxels (Figure 4.3C).

Ce tirage aléatoire des coordonnées des points au sein des voxels limite la génération d'amas de points et de conduits en *marches d'escalier* entre les voxels même pour des résolutions grossières, ce qui permet d'utiliser des grilles n'étant pas nécessairement très fines. Le choix de la résolution impactera très fortement la position des points générés et les temps de calcul. Une résolution trop grossière conduirait à des points très regroupés au sein des voxels les plus favorables et potentiellement éloignés des zones favorisant la karstification présentes en leur sein (e.g., horizons d'inception). Une résolution trop fine provoquerait quant à elle des temps de calcul conséquents.

La grille englobant la surface topographique, certains points pourraient être générés hors de la roche. Suivant ses attentes, nous souhaiterions permettre à l'utilisateur de choisir entre trois options de traitement de ces points avant d'effectuer la simulation. Chaque point généré en dehors de la surface topographique serait supprimé puis : i) un nouveau point serait tiré dans le même voxel, ii) un nouveau point serait généré dans la grille, iii) aucun point ne serait généré. La première option permettrait de respecter au mieux la grille de probabilité mais pourrait provoquer la génération de points très proches les uns des autres dans les voxels en question. La seconde limiterait la création d'amas de points dans les zones proches de la surface topographique mais réduirait les chances de générer des points dans ces voxels. Enfin la troisième respecterait mieux le tirage initial mais aboutirait en un nombre de points plus faible que voulu par l'utilisateur.

La seconde étape consiste à calculer la connectivité des points (Figure 4.1,B). Ils peuvent être définis comme des extrémités, des intersections ou des nœuds de liaison. La connectivité calculée sur chaque point sert à contraindre de manière souple le degré du nœud correspondant dans le graphe final. Les extrémités ont un degré égal à 1, les

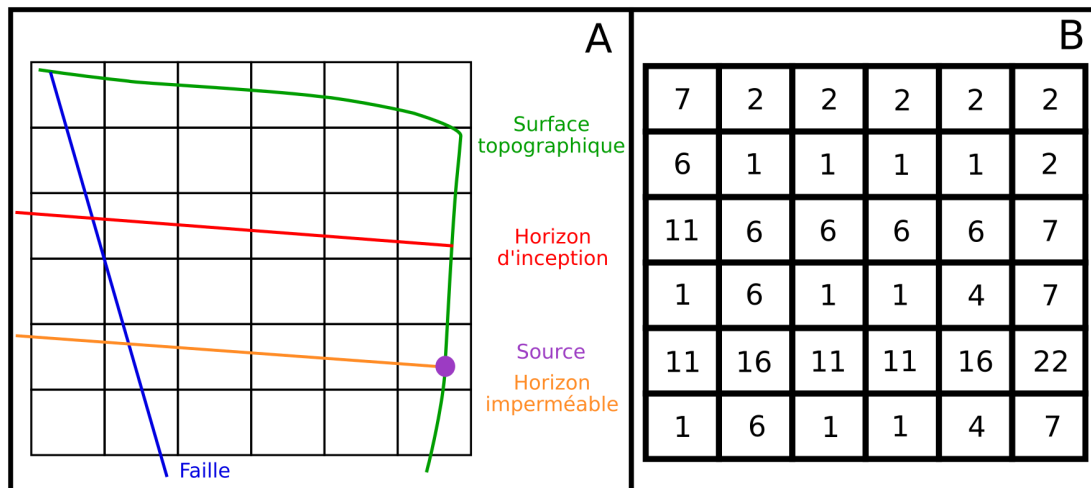


Figure 4.2 – A : Exemple de grille superposée à des informations géologiques connues. B : Exemple de grille de probabilité créée à partir des informations de la grille A. Les probabilités associées à chaque information sont cumulatives. La source correspondant à un conduit connu, le point est directement intégré comme nœud du réseau. La probabilité de la cellule correspondante augmente de 10, celle de ses voisines directes de 5 et celles de ces voisines indirectes (situés en diagonale) augmente de 3. La probabilité des cellules directement en contact avec la surface topographique augmente de 1. La probabilité de celles recoupées par la faille augmente de 5, de même pour celles recoupées par l'horizon d'inception. Celles situées au niveau de l'horizon imperméable voient leur probabilité augmenter de 10. Toutes les cellules se voient par défaut assigner une probabilité de 1 via le remplissage automatique. Ces valeurs sont purement arbitraires et sont données ici en guise d'exemple.

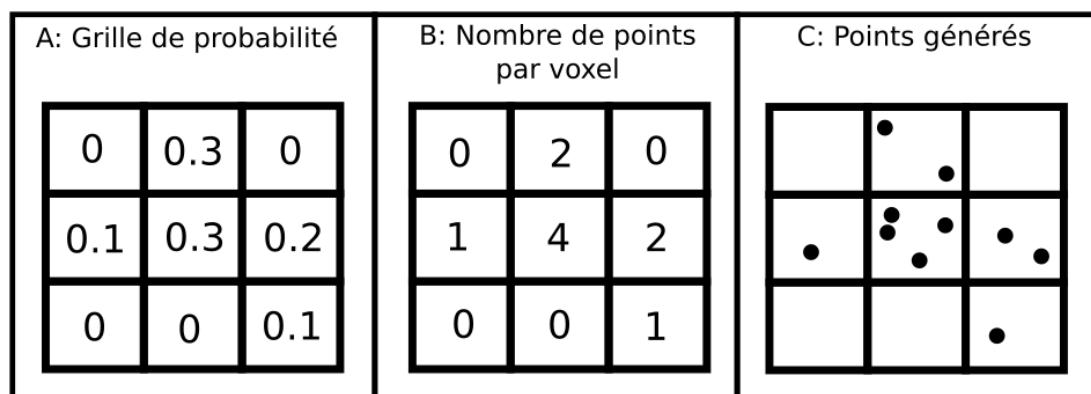


Figure 4.3 – Exemple 2D du tirage aléatoire de 10 points. A : Probabilité de présence de point. B : Exemple de tirage aléatoire des voxels. Le nombre correspond au nombre de tirages de chaque voxel. C : La position finale des points est tirée aléatoirement au sein des voxels correspondant.

intersections un degré égal à 3 ou plus et les autres nœuds un degré égal à 2. Le nombre d'extrémités Ext et le ratio intersections/extrémités R_{ie} , qui permettent de calculer le nombre d'intersections Int , sont définis par l'utilisateur, bien qu'ils ne puissent pas toujours être respectés. Le degré des nœuds et le ratio intersections/extrémités ont été étudiés lors de l'étude statistique et les résultats sont présentés dans le Chapitre 2 et

l'Annexe C, ce qui permet de se baser sur des exemples analogues. En comparaison, l'utilisation des proportions d'extrémités et d'intersections par rapport au nombre total de points n'a pas été jugée pertinente car celles-ci dépendent très fortement de la densité des mesures.

Le ratio intersections/extrémités permet de contraindre la topologie des réseaux simulés (Figure 4.4). Bien que l'étude réalisée sur les motifs morphologiques ne soit pas robuste, des faibles valeurs de ratio intersections/extrémités correspondront principalement à des réseaux arborescents tandis que les grandes valeurs seront plus représentatives des réseaux anastomosés. Il est en effet possible de démontrer que la valeur du ratio intersections/extrémités d'un réseau composé d'une seule composante connexe et d'aucun cycle est nécessairement comprise dans l'intervalle $[0;1[$. Le nombre de cycles c d'un graphe non-orienté constitué de e arêtes, n nœuds et p composantes connexes se calcule de la manière suivante :

$$c = e - n + p \quad (4.1)$$

Dans le cas d'un réseau présentant une seule composante connexe et aucun cycle, on obtient la relation suivante :

$$n = e + 1 \quad (4.2)$$

Dans ce contexte, l'ajout d'un nœud correspond à l'ajout d'une unique arête. Pour simplifier, nous raisonnons sur un graphe réduit, c'est à dire sans prendre en compte les nœuds de degré 2. Rajouter une extrémité au graphe revient soit à la connecter à une intersection déjà existante, soit à la connecter à une intersection nouvellement créée et le rajout d'une intersection ne peut être réalisé que conjointement au rajout d'une extrémité. Il est possible de démontrer par récurrence la relation suivante entre le nombre d'intersections Int et le nombre d'extrémités Ext de ce réseau purement arborescent :

$$Int \leq Ext - 2 \quad (4.3)$$

Partons de U_0 , un cas composé d'une unique branche reliant deux extrémités : $Ext_0 = 2$ et $Int_0 = 0$ donc $Int_0 = Ext_0 - 2$. Nous appelons n le nombre d'extrémités ajoutées à cette branche initiale. U_1 correspond à trois extrémités reliées entre elles via une intersection : $Ext_1 = 2 + 1 = 3$ et $Int_1 = 0 + 1 = 1$ donc $Int_1 = Ext_1 - 2$. A partir de $n = 2$, la nouvelle extrémité sera soit liée à une nouvelle intersection, soit à une intersection existant déjà, autrement dit : $Ext_{n+1} = 2 + n + 1$ et $1 \leq Int_{n+1} \leq n + 1$, ce qui mène bien à $Int_{n+1} \leq Ext_{n+1} - 2$. Cela donne une indication sur le ratio intersections/extrémités :

$$R_{ie} = \frac{Int}{Ext} \leq \frac{Ext - 2}{Ext} = 1 - \frac{2}{Ext} < 1 \quad (4.4)$$

Ainsi, un réseau sans cycle composé d'une seule composante connexe aura nécessairement une valeur de ratio intersections/extrémités comprise dans l'intervalle $[0;1[$. Cela signifie également qu'un réseau pour lequel $R_{ie} \geq 1$ contient au moins un cycle. Néanmoins, un réseau pour lequel $R_{ie} < 1$ peut contenir un ou plusieurs cycles si plusieurs

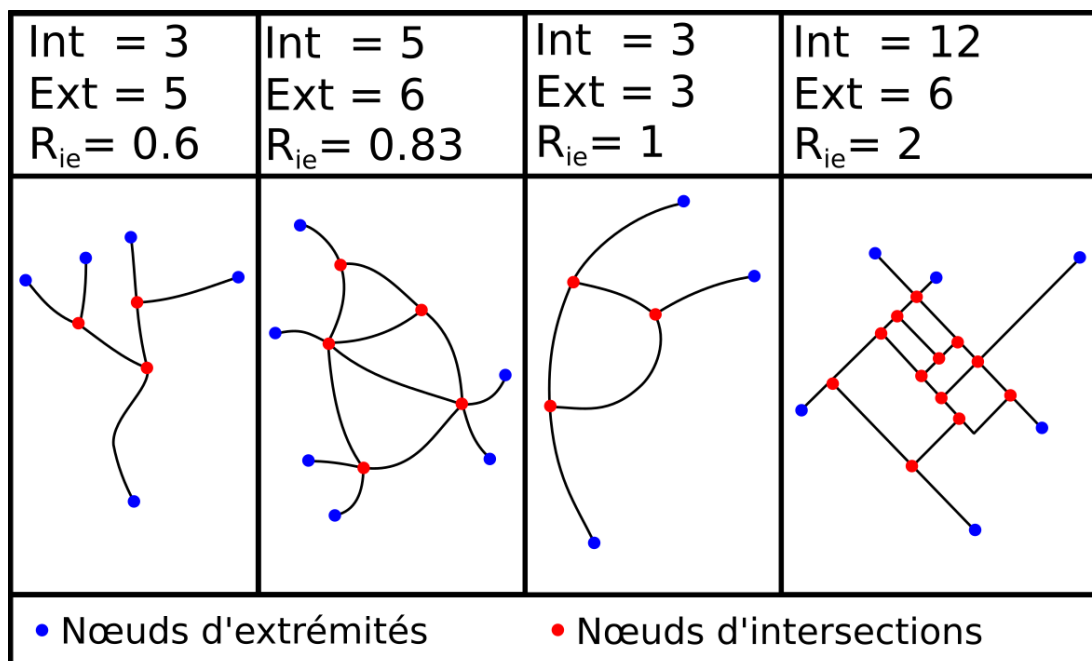


Figure 4.4 – Exemple de topologies pour 4 ratios intersections/extrémités différents.

de ses intersections ont un degré supérieur 3 (Figure 4.4). La définition du ratio par l'utilisateur, couplée à un contrôle du degré maximal des intersections devrait donc permettre à l'utilisateur de contraindre la topologie du réseau.

La troisième étape correspond à la connexion des nœuds. Ils sont reliés ensemble tout en respectant les connectivités calculées lors de l'étape précédente (Figure 4.1 C). L'approche actuelle consiste à partir d'une intersection et à relier les nœuds de proche en proche en favorisant les nœuds situés à faible distance les uns et des autres. Cette approche basée sur la proximité des nœuds permet de créer des liaisons relativement cohérentes même sans que des informations géologiques ne soient disponibles. Différents paramètres peuvent être définis par l'utilisateur pour obtenir un résultat proche de ce qu'il souhaite, notamment la distance maximale entre deux nœuds reliés D_{connex} .

4.3 État actuel de la méthode

A l'issue de cette thèse, seule une version préliminaire de la deuxième et de la troisième étape a pu être implémentée. La première étape est actuellement limitée à la génération de points dans une grille de probabilité préexistante. Nous nous focalisons donc ici sur le calcul de la connectivité (Section 4.3.1) et la connexion des nœuds (Section 4.3.2). La méthode dans son présent état a été implémentée en C++ dans le plugin RingKarst du logiciel SKUA-Gocad ¹.

4.3.1 Calcul de la connectivité

J'ai choisi de contrôler la connectivité du réseau par l'utilisation de deux paramètres indépendants : i) le nombre voulu d'extrémités Ext et ii) le ratio intersections/extrémités

¹<https://www.pdgm.com/products/skua-gocad>

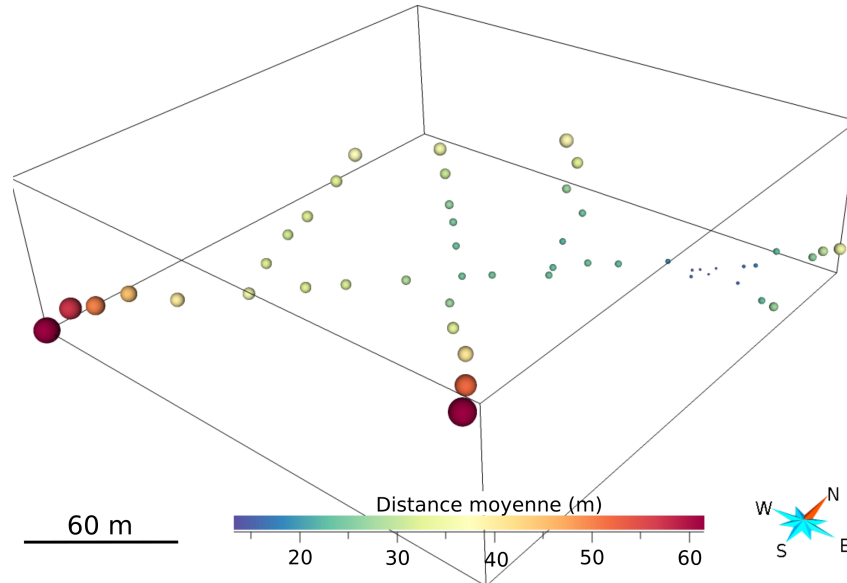


Figure 4.5 – Distance moyenne entre un point et ses 10 voisins les plus proches. Les points les plus distants sont les moins *centraux*.

R_{ie} . Cette étape prend plusieurs paramètres en entrée : Ext , R_{ie} , le nombre de points utilisés lors du calcul de distance moyenne N_{dm} , le nombre théorique de voisins pour chaque point N_{vth} , la distance maximale entre un point et ses voisins D_v , et une valeur de tolérance angulaire θ_{tol} . On appelle voisins d'un point les N_v points les plus proches de lui, tout en lui étant distant de moins de D_v . Un point a théoriquement N_{vth} voisins mais il peut en avoir moins si peu de points se trouvent à une distance inférieure à D_v . L'algorithme commence par le calcul du nombre attendu d'intersections en fonction de Ext et R_{ie} . La distance moyenne entre chaque point et les N_{dm} points les plus proches est ensuite calculée (Figure 4.5), N_{dm} n'étant pas nécessairement égal à N_{vth} . Cela permet de voir quels sont les points les moins *centraux* au sein nuage de points (il ne s'agit cependant pas du calcul d'une centralité théorique, comme cela a pu être fait dans le Chapitre 2).

Les extrémités sont déterminées avant les intersections. Elles sont actuellement définies comme les points les plus éloignés des autres et qui ne sont pas localisés entre plusieurs groupes de points. L'algorithme itère sur tous les points, à commencer par les moins *centraux*, jusqu'à atteindre le nombre d'extrémités voulu. Pour chaque point, l'algorithme vérifie s'il est potentiellement situé dans un conduit à dominante verticale ou horizontale. Si la différence d'altitude entre le point actuel et son voisin le plus éloigné (c'est à dire le N_v -ième point le plus éloigné de lui) est supérieure à leur distance horizontale, le conduit potentiel est considéré comme étant à dominante verticale. Autrement, il est considéré comme étant à dominante horizontale. Si le conduit attendu est subvertical, l'algorithme compare l'altitude du point et de ses voisins. Si ceux-ci ont tous une altitude soit plus basse soit plus élevée que le point actuel (signifiant que le point actuel est le plus bas ou haut de tout son voisinage), il est considéré comme étant une extrémité : sa connectivité est imposée à 1. Si le conduit attendu est subhorizontal, les orientations entre le point et ses voisins sont calculées et chaque orientation est

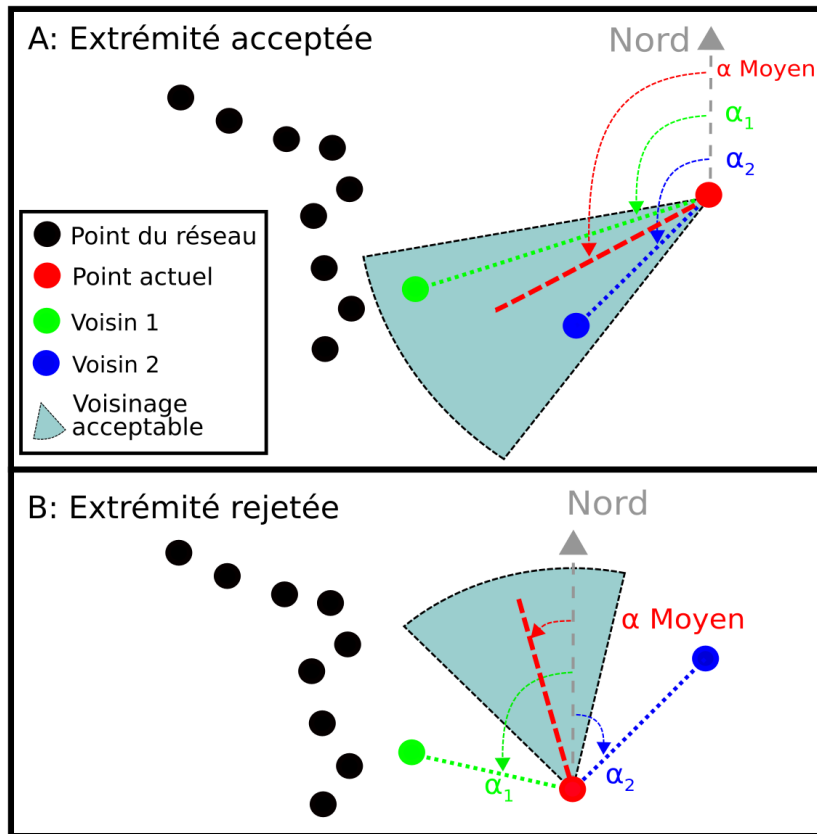


Figure 4.6 – Exemple de détermination d’une extrémité pour un conduit à dominante horizontale. Tous les points font partie du réseau mais ceux sur lesquels sont effectués les calculs ont volontairement été placés à l’écart des autres par souci de visibilité. Pour simplifier l’exemple, le voisinage d’un point est limité aux deux points les plus proches. L’orientation moyenne entre le point actuel et ses voisins est calculée. Si les orientations entre le point actuel et ses deux voisins sont proches de l’orientation moyenne, le point actuel est considéré comme étant une extrémité. A : Les orientations entre le point actuel et ses deux voisins sont proches. Le point actuel est bien une extrémité. B : Les orientations entre le point actuel et ses deux voisins sont très différentes. Le point actuel n’est pas une extrémité.

ensuite comparée à l’orientation moyenne. Si la différence de ne serait-ce qu’une seule orientation et de la valeur moyenne est plus grande que le seuil de tolérance θ_{tol} spécifié par l’utilisateur, le point actuel n’est pas considéré comme une extrémité (Figure 4.6). L’algorithme s’arrête une fois tous les points parcourus et ce, même si le nombre d’extrémités voulu n’a pas été atteint.

Les intersections sont actuellement définies comme les points ayant la plus courte distance moyenne à leurs voisins (ce sont les points les plus *centraux*). L’algorithme itère sur tous les points n’étant pas des extrémités, à commencer par les plus *centraux* et les définit comme étant des intersections jusqu’à atteindre le nombre souhaité. Afin de ne pas avoir des intersections trop proches les unes des autres, les voisins d’une intersection ne peuvent néanmoins pas être des intersections eux-mêmes. Tous les autres points se voient assigner une valeur de connectivité égale à 2.

4.3.2 Connexion des nœuds

Un prototype de la troisième étape est également implémenté. Il prend en entrée un nuage de points associés à des valeurs de connectivité. Une distance maximale entre 2 nœuds reliés D_{connex} peut être définie par l'utilisateur afin d'éviter de créer des liens entre des nœuds trop éloignés. L'algorithme utilisé est semblable à un algorithme de parcours en profondeur (e.g., Tarjan [1972]) mais sert ici à construire le graphe et non à le parcourir (Figure 4.7). Il lie les nœuds deux à deux en commençant depuis un nœud choisi aléatoirement parmi les extrémités. Un lien est créé entre ce nœud et son voisin le plus proche n'ayant pas atteint un degré correspondant à sa connectivité théorique. Cela continue de proche en proche (Figure 4.7B). Quand une intersection est atteinte pour la première fois, elle est stockée dans une liste d'attente créée vide au début de cette étape (Figure 4.7C). En effet, après la création d'un lien avec un autre nœud, son degré n'est que de 2 alors qu'il devrait au moins être égal à 3. Quand une extrémité est atteinte ou qu'aucun nœud suffisamment proche du nœud actif n'est mobilisable pour créer un autre lien, le dernier nœud d'intersection stocké devient le nouveau nœud actif et l'algorithme reprend à partir de ce dernier (Figure 4.7D, E). Quand aucun nouveau lien ne peut être créé et que la liste d'attente est vide, si une extrémité n'a pas encore été atteinte, l'algorithme repart depuis cette dernière. Si toutes les extrémités ont été atteintes, cette partie de l'algorithme s'arrête (Figure 4.7F).

Une option permet à l'utilisateur d'améliorer ou non la connexité du graphe, en augmentant les chances de connecter les points isolés et en créant des connexions directes entre les nœuds d'intersection. Si certains nœuds de degré 2 n'ont pas été atteints, ils sont reliés de la même manière que précédemment à leurs voisins les plus proches, jusqu'à atteindre un nœud d'intersection. À ce stade, le degré des nœuds d'intersections peut devenir supérieur à 3 mais nous avons choisi de ne pas dépasser un degré maximal de 5. Si certains nœuds d'intersections n'ont pas atteint un degré minimal de 3, ils sont reliés à d'autres nœuds d'intersection. La connexion entre les nœuds d'intersections ayant un degré inférieur à 5 est également forcée pour augmenter les chances de créer un réseau connexe et de former des cycles. Si un nœud n'a pas été atteint, il est supprimé automatiquement. L'algorithme a par ailleurs été pensé de manière à éviter de créer des cycles composés uniquement de nœuds de degré 2.

4.4 Résultats

La génération du nuage de points (étape 1) n'étant pas encore implémentée, je ne présente ici que des résultats obtenus sur des nuages de points prédéfinis, afin d'évaluer les algorithmes pour le calcul de la connectivité (étape 2) et la connexion des nœuds (étape 3). Les paramètres d'entrée des différents exemples sont données dans le Tableau 4.1.

Exemple A L'étape de connexion des nœuds est d'abord testée de manière indépendante sur un groupe de 50 points ayant déjà des valeurs de connectivité théorique (Figure 4.8). La distance entre les points les plus proches est d'environ 10 mètres. La connectivité du point central a dès le départ été fixée à 4 bien que, dans l'algorithme développé pour la deuxième étape, les points d'intersections aient une connectivité initiale fixée à

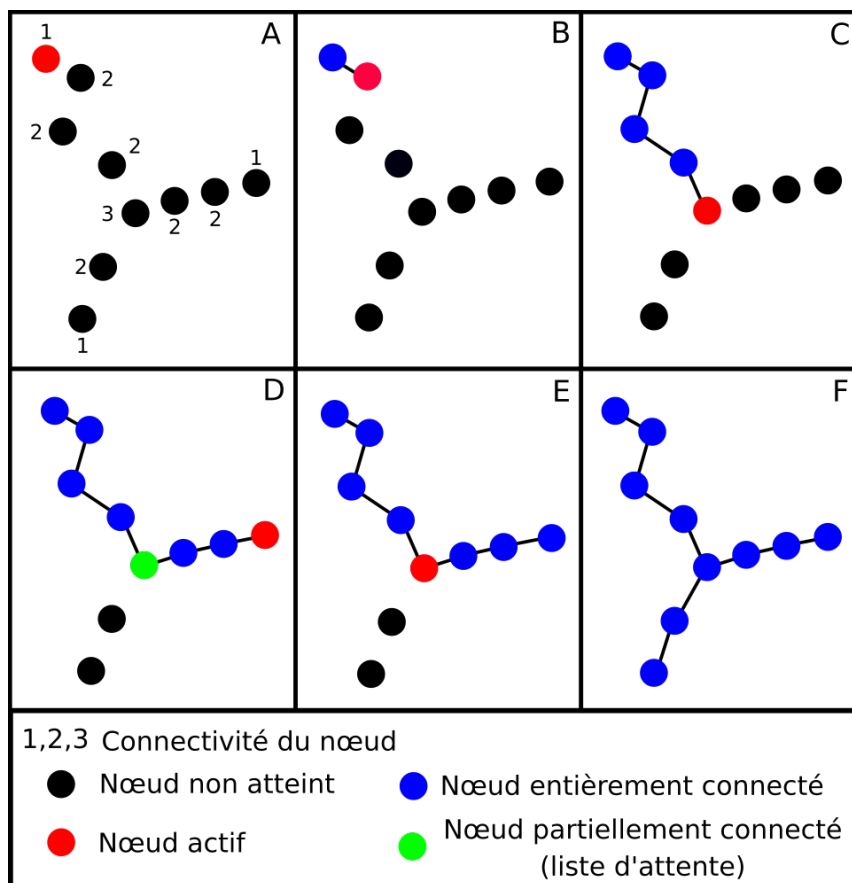


Figure 4.7 – Exemple de création d'un squelette. A : Configuration initiale du réseau. La connectivité des points, qui est équivalente au degré attendu du nœud au sein du graphe, est indiquée. L'extrémité située en haut à gauche est choisie aléatoirement comme nœud de départ. Une liste d'attente pour les nœuds d'intersections est créée vide. B : Le nœud actif est relié à son voisin le plus proche. Ce voisin devient le nouveau nœud actif. C : L'algorithme continue jusqu'à ce qu'une intersection soit atteinte. Le nœud d'intersection est stocké dans la liste d'attente. D : L'algorithme continue jusqu'à ce qu'une extrémité soit atteinte. E : Le dernier nœud dans la liste d'attente devient le nœud actif. F : L'algorithme continue jusqu'à ce qu'une extrémité soit atteinte. Il n'y a plus de nœuds dans la liste d'attente et toutes les extrémités ont été atteintes. Tous les nœuds sont connectés et tous les nœuds d'intersections ont un degré minimal de 3 donc l'algorithme s'arrête.

3. Cette valeur a été choisie de façon à obtenir un résultat adapté à la configuration des points. Différentes réalisations sont obtenues, en fonction de l'extrémité de départ sélectionnée par l'algorithme. Les endroits pour lesquels la variabilité est la plus grande peuvent parfois laisser émerger un manque local de cohérence géologique : des liaisons peuvent s'y superposer (e.g., Figure 4.8B), mais la connectivité théorique des points elle-même est bien respectée.

Exemple	Étape de calcul de la connectivité						Étape de connexion	
	Ext	R_{ie}	N_{dm}	N_{vth}	D_v	θ_{tol}	D_{connex}	Amélioration connexité
A	Connectivité imposée						50	Non
B	7	0.6	10	10	50	45	50	Non
C	Connectivité imposée						50	Oui et non
D	35	1	10	10	50	45	50	Oui et non
E	30	1	10	10	250	45	250	Oui

Tableau 4.1 – Paramètres d’entrée des différents exemples. La colonne Amélioration connexité désigne une option que l’on peut activer ou non.

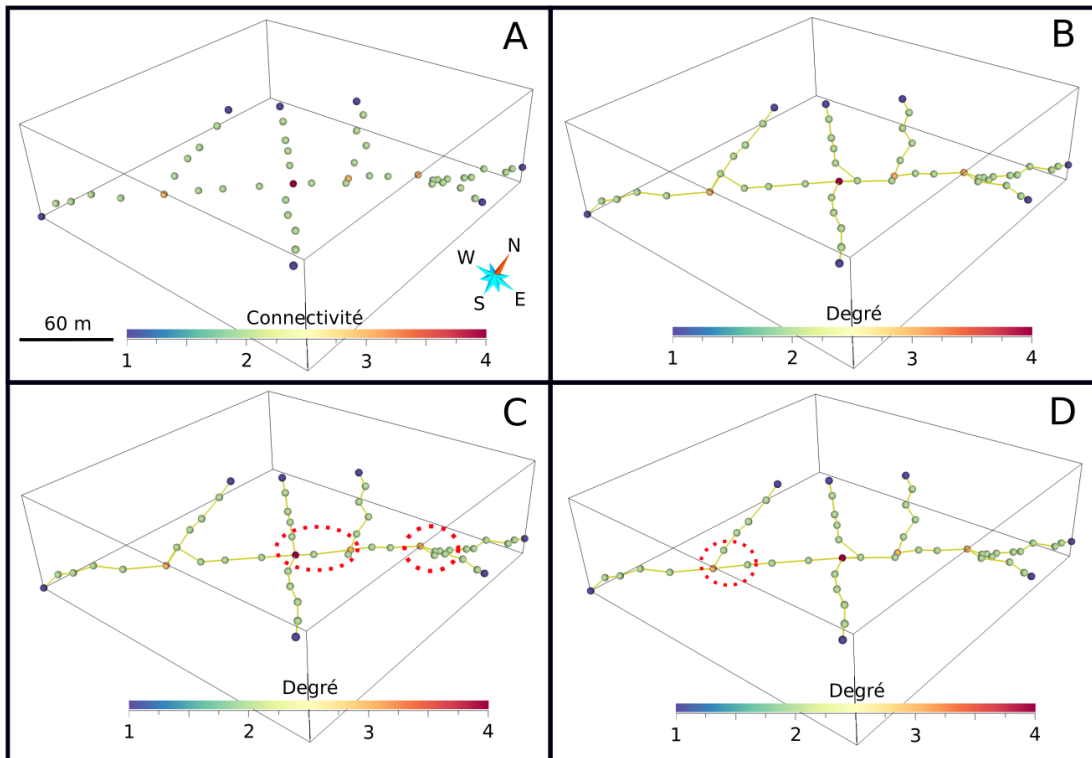


Figure 4.8 – A : Nuage de points initial avec leurs connectivités. B, C, D : Exemples de résultats obtenus en appliquant l’algorithme de connexion au nuage de points et en utilisant la connectivité présentée en A. La propriété affichée correspond au degré final des nœuds et non à la connectivité théorique. Les ellipses rouges indiquent les principales différences entre les réalisations. Pour des raisons arbitraires, nous utilisons B comme référence sur cet exemple.

Exemple B Nous nous sommes ensuite servi du même nuage de points pour tester l’étape de calcul de connectivité (Figure 4.9A). Afin d’essayer d’obtenir le même nombre d’intersections et d’extrémités que précédemment (Figure 4.8A), nous avons utilisé un nombre d’extrémités de 7 et un ratio intersections/extrémités égal à 0.6 comme paramètres d’entrée. Cela devait théoriquement mener au calcul de 7 extrémités et 4 intersections. Cette étape étant actuellement déterministe, nous nous attendions à ce que les connectivités calculées soient similaires à celles de la Figure 4.8A.

Les valeurs de connectivité calculées sont relativement proches de ce qui était at-

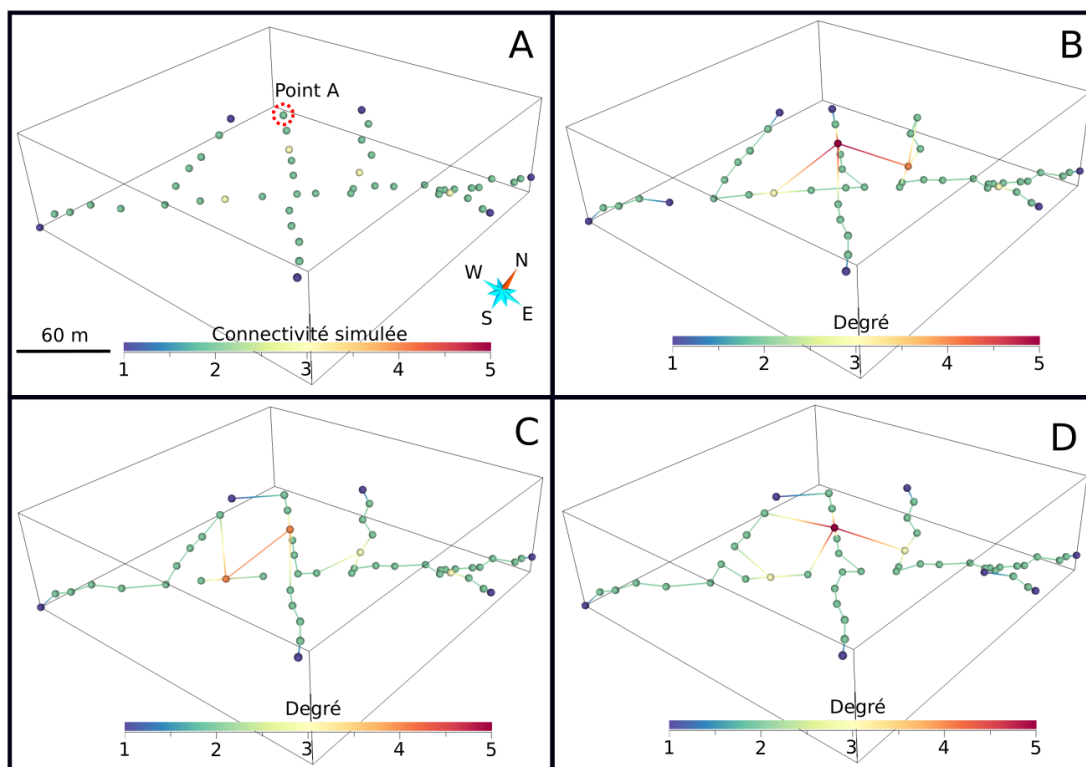


Figure 4.9 – A : Nuage de points avec les valeurs de connectivité calculées durant la deuxième étape. Le point A devrait avoir une valeur de connectivité de 1 mais elle est égale à 2 dans les résultats. B, C, D : Exemples de résultats obtenus durant l'utilisation de la troisième étape, en se servant des valeurs de connectivité calculées durant la deuxième étape (montrées en A). La propriété affichée correspond au degré final des nœuds et non à la connectivité pré-calculée.

tendu, exception faite de la fin du conduit vertical du milieu (point A) qui n'est pas considérée comme une extrémité. En effet, à cause de la présence de deux autres conduits autour, ce conduit est considéré comme horizontal et étant donnée la présence de points à sa gauche et à sa droite, le point A n'est pas défini comme une extrémité. En ce qui concerne les intersections, elles correspondent bien aux points les plus *centraux* (Figure 4.5) n'étant pas dans le voisinage d'autres intersections, même si ce ne sont pas les points *idéaux* au vu du nuage de points initial (elles ne sont pas à la base des différents conduits).

L'étape de connexion est ensuite testée sur ces résultats (Figure 4.9B, C, D). Les résultats obtenus sont cohérents avec l'algorithme proposé, bien que l'on puisse constater que les réseaux générés sont assez différents de ce qui a pu être montré en Figure 4.8. Ces changements sont principalement dus à la position des intersections calculées lors de la deuxième étape.

Exemple C Nous avons ensuite testé la troisième étape sur un réseau réel en nous servant du vrai degré des nœuds comme valeur de connectivité (Figure 4.10). Le réseau utilisé est celui de l'Aspirateur ². Il n'est pas densifié, ainsi seuls 269 nœuds associés à

²Gard, France, données fournies par J.L. Galéra

leur valeur de degré dans le réseau initial servent de base à ce test. La forme globale du réseau est plutôt bien reproduite mais de nombreuses différences locales sont observables. Les connexions changent en fonction de l'extrémité de départ mais les réseaux obtenus ne sont généralement pas connexes. L'algorithme améliorant la connectivité du graphe permet de rajouter quelques nœuds au résultat final mais ce n'est cependant pas suffisant pour permettre d'obtenir un réseau connexe, le nombre de composantes connexes variant selon les réalisations.

Exemple D Nous avons également testé l'enchaînement de la deuxième et de la troisième étape sur le réseau de l'Aspirateur. Nous avons choisi un nombre d'extrémités égal à 35 et un ratio intersections/extrémités égal à 1 comme paramètres d'entrée, ceci afin d'obtenir une connectivité semblable à celle du réseau réel (Figure 4.11). Après calcul de la connectivité, environ 15 points sont considérés comme des extrémités et une vingtaine comme des intersections. Le faible nombre d'extrémités est explicable par le fait que dans le réseau initial, nombre d'entre elles sont situées entre plusieurs groupes de points, ce qui est un critère excluant pour notre méthode. Le nombre d'intersections est également plus bas car avec notre définition, un point ayant un point d'intersection dans son voisinage ne peut lui-même pas être un point d'intersection. Ceci laisse entrevoir certaines limites actuelles de notre méthode. Le résultat final respecte relativement bien la forme du réseau dans sa globalité mais des changements locaux sont également visibles (Figure 4.11B). La gestion des points isolés et l'augmentation de la connectivité permettent de rajouter de nombreux nœuds au résultat final (Figure 4.11C) mais là aussi, les réseaux obtenus ne sont pas complètement connexes.

Exemple E Nous avons finalement effectué des tests sur des nuages de 300 points générés aléatoirement dans un cube faisant 1000 mètres de côté (Figure 4.12). Nous avons choisi un nombre d'extrémités égal à 30 et un ratio intersections/extrémités égal à 1. La distance D_{connex} est très importante (250 mètres), de manière à compenser la très faible densité de points. Bien qu'il soit difficile de les évaluer visuellement à cause de la disposition des points, les résultats correspondent bien à des graphes interconnectés. Les nœuds connectés sont généralement proches, menant à des connexions pouvant être qualifiées de *logiques*. Nous observons de plus que, conformément à nos définitions actuelles, les intersections sont principalement situées au centre du cube, tandis que les extrémités ont tendance à être situées sur ses bords. Néanmoins, à cause de ces mêmes définitions, le nombre d'intersections et d'extrémités générées est inférieur à 30 sur toutes les réalisations analysées.

4.5 Discussion

Bien qu'encore à l'état de prototype, la méthode présentée dans ce chapitre offre des premiers résultats intéressants. J'ai envisagé et testé de nombreuses variations dans ses différentes étapes. L'implémentation présentée ici est celle qui a donné les résultats les plus satisfaisants. Bien qu'imparfaits, ils montrent un certain potentiel. L'intégration des données brutes et des informations géologiques et hydrogéologiques n'est pas encore possible mais plusieurs critères jugés importants sont déjà, au moins partiellement,

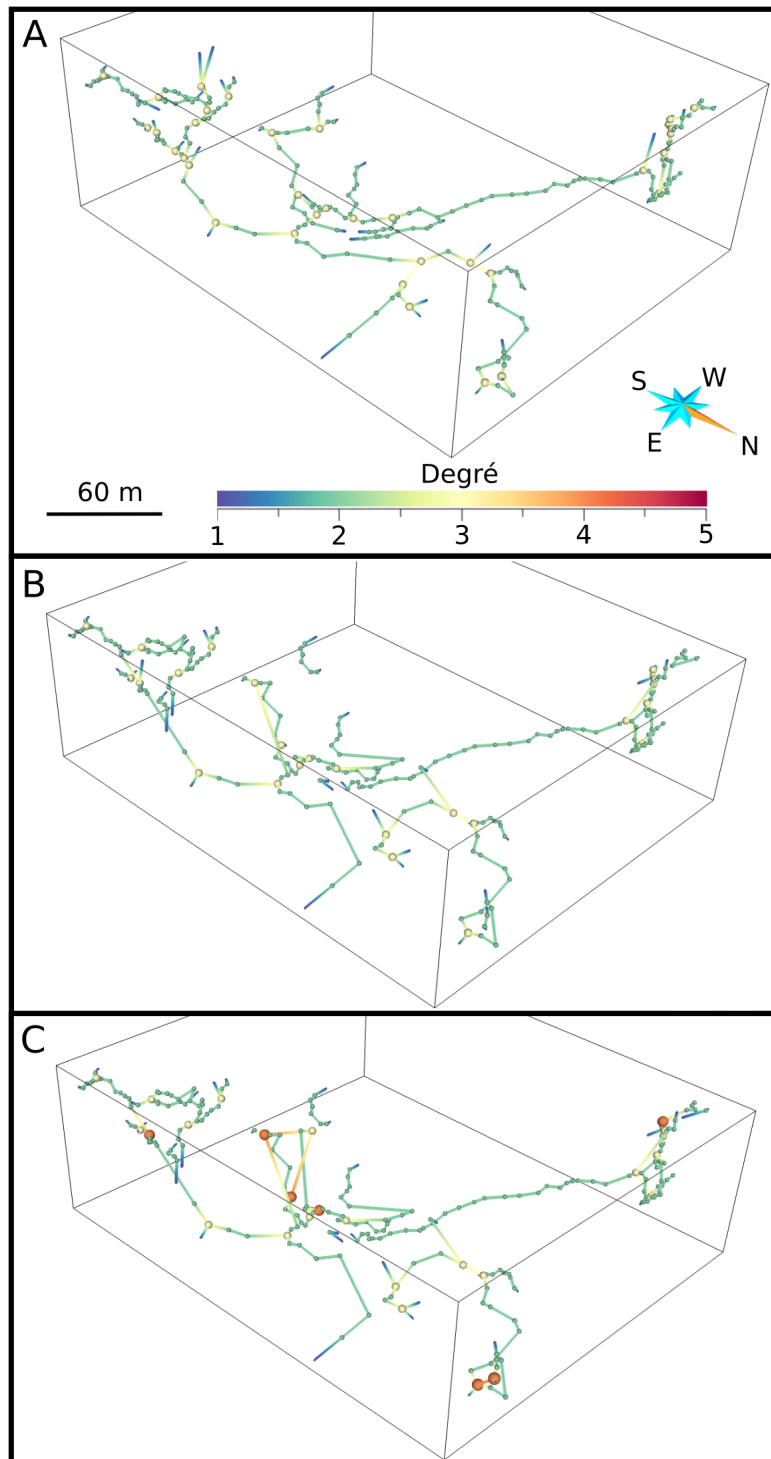


Figure 4.10 – A : Véritable réseau de l’Aspirateur (France, Gard). La taille des nœuds est proportionnelle à leur degré. B : Résultat obtenu en appliquant la connexion automatique aux points du réseau de l’Aspirateur sans utiliser la gestion des points isolés et l’augmentation de la connectivité. Le degré des nœuds est utilisé comme connectivité théorique et la distance maximale entre deux nœuds connectés est de 50 mètres. La propriété présentée correspond au degré réel des nœuds du résultat, celui-ci pouvant différer du degré original. Le graphe n’est pas connexe. C : Résultat obtenu en appliquant cette même étape au réseau de l’Aspirateur mais en gérant les points isolés et en augmentant la connectivité. Bien que certains nœuds aient pu être rajoutés au graphe, ce dernier n’est toujours pas connexe.

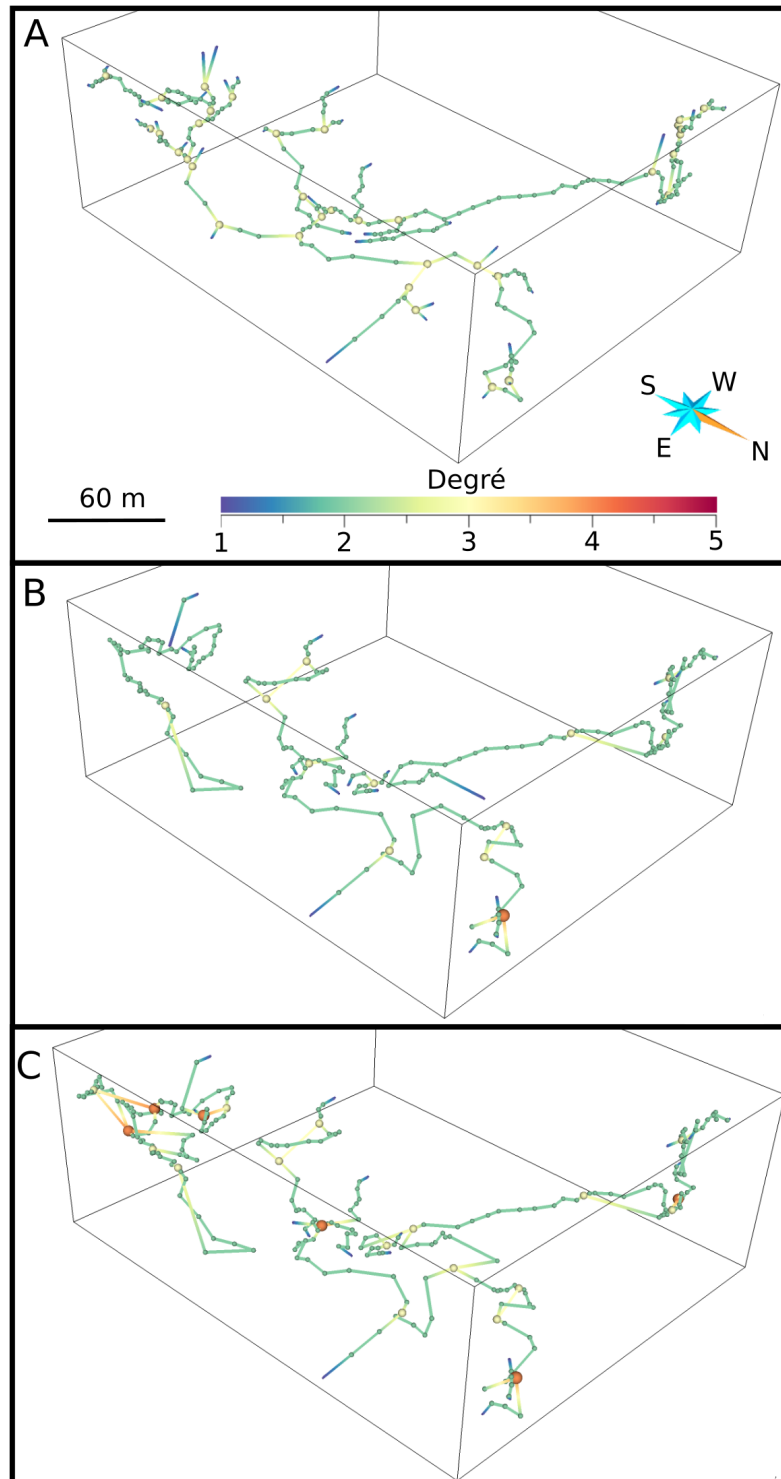


Figure 4.11 – A : Véritable réseau de l’Aspirateur (France, Gard). La taille des nœuds est proportionnelle à leur degré. B : Résultat obtenu en appliquant le calcul de la connectivité et la connexion automatique aux points de réseau de l’Aspirateur sans utiliser la gestion des points isolés et l’augmentation de la connectivité : le graphe n’est pas connexe. C : Résultat obtenu en appliquant ces mêmes étapes au réseau de l’Aspirateur mais en gérant les points isolés et en augmentant la connectivité. Bien que certains nœuds aient pu être rajoutés au graphe, ce dernier n’est toujours pas connexe.

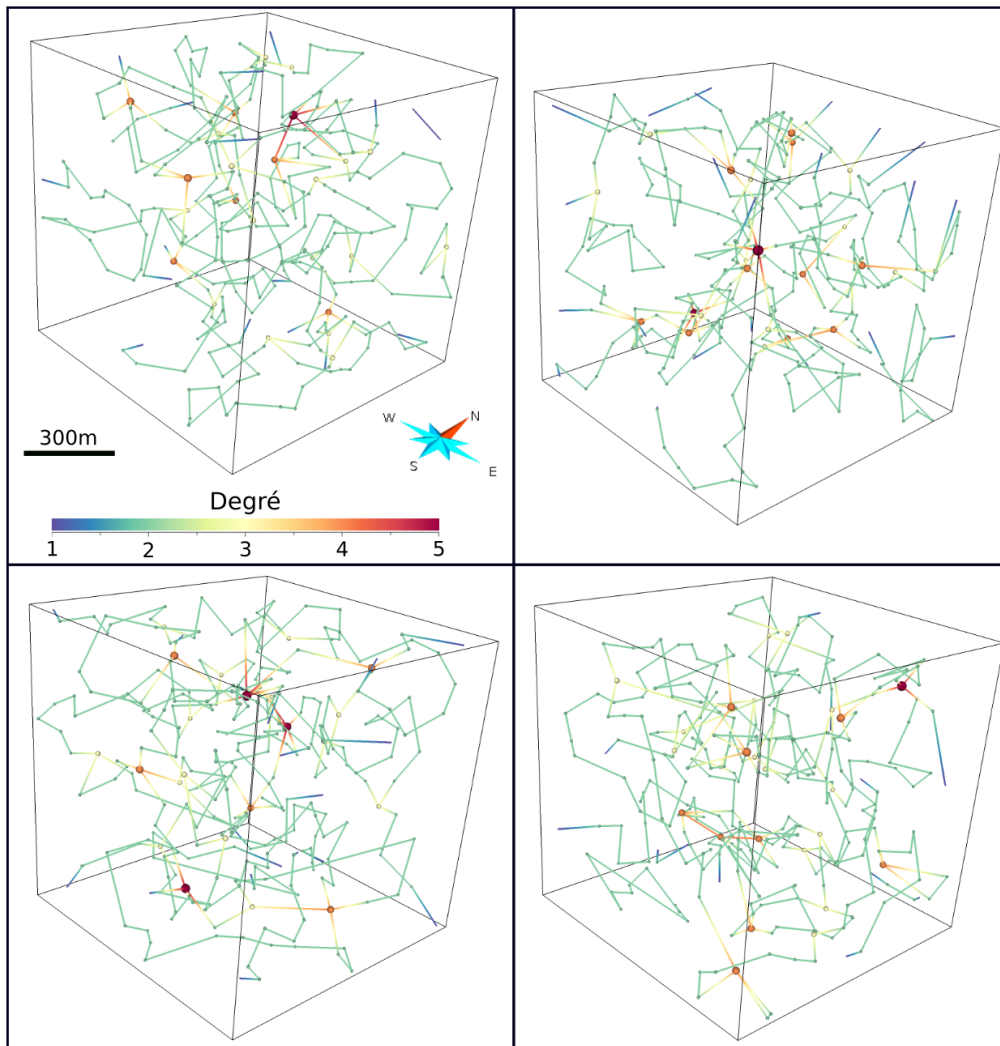


Figure 4.12 – Résultats obtenus en appliquant la deuxième et de la troisième étape sur 4 nuages de points générés aléatoirement. La taille des nœuds est proportionnelle à leur degré.

respectés.

Le principal but de cette méthode est de permettre la simulation de réseaux de conduits discrets tridimensionnels. Bien que l'étape de génération des points ne soit pas encore implémentée, il est possible de générer de tels réseaux à partir de nuages de points prédéfinis. Même s'il est prévu que la première étape utilise une grille de probabilité, les points devraient être générés aléatoirement au sein des voxels, conduisant à des réseaux non alignés sur une grille.

Les squelettes des réseaux sont bien générés de façon stochastique. Il est attendu que cette stochasticité offre un bon aperçu des incertitudes liées à la complexité des réseaux. Actuellement, la stochasticité n'est induite que par l'ordre de parcours des extrémités n'ayant pas encore été connectées au graphe lors de l'étape de connexion. En effet, comme toutes les extrémités ne sont pas forcément atteintes depuis l'extrémité de départ (la connexité du réseau n'étant pas assurée), l'algorithme recommence à partir d'une d'entre elles choisie aléatoirement. Une fois implémentée, la génération des

points à partir de la grille de probabilité devrait grandement accroître la stochasticité des simulations. Intégrer une part d'aléatoire lors du calcul de la grille de probabilité permettrait de générer des nuages de points encore plus différents d'une réalisation à une autre. L'étape de calcul de la connectivité est actuellement purement déterministe. Assigner une probabilité aux différents points lors de la détermination des intersections et des extrémités, cela en fonction du respect des critères considérés et de leur position au sein du nuage de points, puis tirer aléatoirement le nombre voulu d'intersections et d'extrémités la rendrait stochastique. Il devrait également être possible de procéder similairement pour l'étape de connexion, en assignant aux voisins potentiels du point actif des valeurs de probabilité dépendant de leur distance et éventuellement d'autres paramètres (orientations, différences d'altitude).

Cette méthode n'est pas basée sur une grille, exception faite du calcul de probabilité de la première étape. Les temps de calcul de la deuxième et de la troisième étape sont de fait assez raisonnables et ce sans que le code n'ait été optimisé : ils sont de l'ordre de la seconde pour les exemples précédents mais bien sûr, plus le nombre de points est important, plus la simulation est longue. Pour autant, en fonction de la résolution de la grille de probabilité utilisée et de la complexité de l'algorithme de remplissage automatique, les temps de calculs de la première étape pourraient s'avérer importants.

L'algorithme proposé devrait permettre à l'utilisateur de contrôler la topologie des réseaux simulés, notamment grâce au choix d'un nombre d'extrémités et d'un ratio intersections/extrémités. L'utilisation de ce ratio est justifiée car comme démontré précédemment, une valeur supérieure à 1 impliquera nécessairement la présence d'au moins un cycle dans le réseau, bien que l'inverse ne soit pas vrai. Il a de plus pu être caractérisé dans le Chapitre 2 et l'Annexe C. Une fois pleinement implémentée, cette méthode devrait permettre la simulation de plusieurs morphologies de réseaux (e.g., arborescent, anastomosé). J'imagine qu'une simulation séquentielle ou régionalisée permettrait de combiner ces motifs dans un même réseau. Pour autant, bien que la topologie des réseaux soit contrôlable, si les points générés ne sont pas cohérents avec le type souhaité, les résultats obtenus peuvent s'en éloigner. Les résultats obtenus sur l'exemple E pour différents ratios ne correspondaient nullement à des réseaux arborescents ou anastomosés car les points ont été générés de façon purement aléatoire. Il serait ainsi souhaitable de pouvoir prendre en compte le type de réseau voulu dès l'étape de simulation des points, bien que cela risque d'être complexe. Dans l'implémentation actuelle de cette méthode, le nombre d'extrémités et le ratio intersections/extrémités des résultats peuvent de plus légèrement différer des paramètres d'entrée pour deux principales raisons : i) le respect des paramètres dépend beaucoup de la configuration des points, ii) il n'est pas possible d'obtenir un graphe connexe pour certaines proportions alors que c'est ce que nous souhaitons. Il est possible que le bon nombre d'intersections et d'extrémités soit déterminé lors du calcul de la connectivité des nœuds mais que celles-ci ne soient pas respectées lors de l'étape de connexion. Certains nœuds ayant une valeur de connectivité théorique de 2, voire de 3 ou plus, ne sont connectés qu'une seule fois et finissent en nœuds de degré 1. Les points isolés ne sont parfois pas connectés au reste du réseau et sont automatiquement supprimés, ce qui induit également un nombre de nœuds différent de celui choisi par l'utilisateur. Dans l'état actuel de l'implémentation, cela s'explique par le fait que tous les nœuds assez proches

du nœud actif ont déjà atteint leur valeur de connectivité maximale et ne peuvent donc être mobilisés pour effectuer de nouvelles liaisons. Bien qu'une option améliorant la connexité du graphe soit disponible, elle ne remédie que partiellement à ce problème car la connexion dépend grandement de la distance entre les points. Il n'est pas toujours possible de créer un graphe connexe à partir de tous les ensembles de valeurs de connectivité possibles mais la prise en compte d'autres paramètres, comme c'est déjà ici le cas de la distance D_{connex} , ne fait que renforcer ce problème. Cependant, bien que le non-respect des paramètres d'entrée puisse être problématique, le fait que certains groupes de nœuds ne soient pas intégrés au réseau final n'est un problème que pour les tests effectués sur des nuages de points déjà définis. Lorsque les points sont générés de manière stochastique, comme cela devrait être le cas avec la première étape, il peut être souhaitable que les groupes trop éloignés du reste des points soient supprimés. Il conviendrait donc probablement de réévaluer cet aspect une fois la chaîne de simulation totalement implémentée.

Dans l'état actuel de l'algorithme, les réseaux générés lors de la troisième étape ne sont généralement pas complètement cohérents d'un point de vue géologique car nous nous servons de la distance comme seul critère pour connecter les nœuds. Cette approche basée sur la proximité ne permet pas de parfaitement reproduire l'organisation habituelle des réseaux au niveau local. La distance maximale entre deux nœuds connectés doit être adaptée au cadre d'étude. Une distance trop courte empêche de correctement lier les points, résultant en de petits groupes très peu connexes, tandis qu'une distance trop grande génère des liaisons aberrantes. Certains critères comme les orientations préférentielles des conduits (liées à l'orientation des fractures du domaine d'étude) ou les différences d'altitude entre les nœuds ne sont, de plus, pas encore pris en compte. On remarque également que certaines connexions peuvent se superposer, notamment lorsqu'une intersection est connectée à un nœud et à un autre nœud situé directement derrière lui. Il n'est de plus pas prouvé qu'une fois implémentée, la première étape permette de générer des nuages de points semblables aux observations de terrain. La détermination des intersections et extrémités lors de la deuxième étape est également limitée, certaines configurations de points étant incompatibles avec nos définitions actuelles. S'inspirer de l'utilisation des graphes de proximité et de leurs dérivées comme cela a pu être le cas dans d'autres domaines d'application (e.g., Gabriel et Sokal, 1969; Jaromczyk et Toussaint, 1992; Galin et al., 2011) pourrait s'avérer intéressant pour générer des réseaux semblables aux observations de terrain.

Les données de terrain et les informations géologiques et hydrogéologiques ne sont actuellement pas prises en compte. Elles devraient directement intervenir lors du calcul de la grille de probabilité de la première étape. Néanmoins, pour des résolutions de grille grossières, il est probable que les points générés s'éloignent des surfaces d'inception. L'orientation préférentielle des conduits est parfois connue, que cela soit grâce à la connaissance de conduits réels ou de la fracturation locale, mais il n'est pas prévu de la prendre en compte dès la génération des points. Favoriser la création de conduits respectant cette orientation préférentielle lors de l'étape de connexion me semble plus adapté. Il devrait être possible de prendre en compte d'autres informations statistiques lors de l'étape de génération des points, notamment en contraignant les valeurs des voxels de la grille de probabilité avec des lois de probabilité connues, plus particulièrement par des

lois de probabilité selon les 3 dimensions choisies par l'utilisateur (l'histogramme des altitudes ayant été étudié dans le Chapitre 2). Les essais de traçage peuvent apporter des informations quant à l'existence de conduits reliant différentes parties d'un réseau, mais n'ayant pas pu être explorés. Il risque d'être compliqué de se servir de ces informations au sein des différentes étapes de la méthode. Pour autant, il devrait être possible de permettre à l'utilisateur de sélectionner plusieurs nœuds du réseau devant être reliés et de générer des conduits entre ces derniers. La simulation de parties entières de réseaux devrait être possible de cette manière. La définition des conduits devant être reliés ou de la zone à simuler nécessiterait néanmoins l'avis de l'utilisateur et ne pourrait donc pas être automatisée.

Pour remédier à certains problèmes, une étape de post-traitement après la génération du réseau est envisageable. Le but de cette étape serait de modifier légèrement les réseaux, notamment la position des nœuds, pour qu'ils soient plus conformes aux attentes de l'utilisateur. Les graphes étant une structure de données relativement peu complexe, il devrait être possible de les modifier après leur création. Ce post-traitement pourrait avoir lieu directement après la simulation du réseau ou après simulation de ses propriétés géométriques. Il devrait notamment être possible d'éviter que des conduits, dont la géométrie a été simulée, ne se recoupent. En effet, les méthodes de simulation des propriétés présentées dans le Chapitre 3 ne prenant pas en compte les coordonnées cartésiennes des nœuds, certains conduits peuvent potentiellement se recouper une fois leur géométrie simulée. Séparer les conduits correspondant en modifiant légèrement les coordonnées de leurs nœuds permettrait d'obtenir des conduits bien distincts, ce qui serait plus convenable lors des simulations d'écoulement. Il est également envisageable de fusionner les conduits se recoupant excessivement en un unique conduit de grande taille. Les liaisons générées entre des points éloignés pouvant être de taille conséquente, elles peuvent mener à des conduits rectilignes sur plusieurs dizaines de mètres. Une densification du graphe dans son entièreté suivie d'un léger déplacement des nœuds nouvellement créés pourrait être intéressante pour limiter la rectilinéarité desdits conduits.

4.6 Conclusion

Ce chapitre explore une nouvelle approche pour simuler des squelettes de réseaux karstiques tridimensionnels de façon stochastique. Elle reproduit, au moins partiellement, la complexité des réseaux en respectant les paramètres de connectivité.

La base de la première étape est terminée mais la création de la grille de probabilité, incluant l'intégration des données de terrain, constitue un énorme chantier qui n'a pas pu être entamé dans le cadre de cette thèse. La deuxième et la troisième étapes sont bien implémentées mais sont encore améliorables sur de nombreux points. Les problèmes rencontrés pour ces deux étapes sont complexes et dépassent le simple cadre géologique. Explorer la piste des graphes de proximité et de leurs dérivées (e.g., Gabriel et Sokal, 1969; Jaromczyk et Toussaint, 1992; Galin et al., 2011) pourrait permettre de contourner certains problèmes actuels.

En dépit des imperfections des résultats obtenus, des possibilités intéressantes sont mises en évidence quant à la simulation de réseaux de conduits *réalistes*. Une fois com-

plète, cette méthode, combinée aux méthodes présentées dans le Chapitre 3, pourrait être utilisée pour diverses applications comme les simulations d'écoulement au sein de réseaux de conduits discrets (e.g., SWMM, MODFLOW-CFP, Epanet) ou la simulation de parties non-explorées mais dont l'existence est soupçonnée dans des réseaux connus.

Conclusions générales

Récapitulatif des contributions

Le travail effectué au cours de cette thèse visait à développer une nouvelle méthode permettant de simuler des réseaux karstiques discrets de façon stochastique et prenant en compte différentes informations contraignant la position des conduits et leur géométrie. Ce travail a résulté en 3 contributions majeures :

- L'analyse statistique d'une base de données comprenant 49 réseaux karstiques différents et le développement d'outils dédiés à cet effet.
- Le développement de deux méthodes de simulation permettant de simuler la géométrie, c'est à dire les dimensions, des conduits.
- Le prototype d'une méthode pour simuler le squelette des réseaux de drains karstiques.

Les deux premières contributions ont été publiées dans Frantz et al. [2021].

Inférence statistique

Dans le Chapitre 2, j'ai présenté les résultats de l'analyse statistique réalisée sur une base de données forte de 49 réseaux connus. Cette analyse s'est principalement focalisée sur les propriétés géométriques des conduits et, à notre connaissance, il s'agit de la première étude de cette envergure. Les réseaux ont été analysés individuellement puis comparés. Nous n'avons pu mettre en évidence aucune loi statistique simple, universelle et indépendante des conditions locales qui aurait permis de décrire les distributions du rayon équivalent et du ratio largeur-hauteur. Nous avons également analysé les géométries de conduits de différents motifs morphologiques, mais le faible nombre de réseaux disponibles pour chacun d'entre eux, ainsi que leur petite taille, la rendent non robuste. L'autre intérêt majeur de cette analyse statistique a été la mise en évidence d'une faible variabilité spatiale des propriétés géométriques à courte distance, notamment au sein des branches du réseau, grâce à l'implémentation et à l'utilisation de trois variogrammes 1D-curvilinéaires. Finalement, cette étude nous a conduits à rejeter l'hypothèse répandue d'une organisation hiérarchique des géométries des conduits.

Les résultats de cette étude ont servi de base aux méthodes de simulation présentées dans les Chapitres 3 et 4.

Simulation des propriétés géométriques des conduits

Deux méthodes permettant de simuler des propriétés le long des conduits de réseaux karstiques complexes ont été présentées dans le Chapitre 3. Elles reproduisent bien les informations statistiques d'entrée (distribution et variogrammes) ainsi que les données conditionnantes, et permettent de générer des valeurs de rayon équivalent et de ratio largeur-hauteur statistiquement proches des observations de terrain présentées dans le Chapitre 2. Basées sur les méthodes de Simulations Gaussiennes Séquentielles (SGS), elles sont stochastiques, ce qui permet potentiellement de propager les incertitudes concernant la géométrie des conduits jusqu'aux futures simulations d'écoulement. La caractérisation de la variabilité spatiale le long des réseaux grâce aux variogrammes 1D-curvilinéaires présentés dans le Chapitre 2 est le facteur clé ayant permis leur développement. La méthode de SGS 1D-curvilinéaire permet de reproduire la variabilité à l'échelle du réseau. La méthode de SGS 1D-curvilinéaire *branchwise* permet, quant à elle, de reproduire la variabilité à l'échelle des branches tout en reproduisant partiellement la variabilité à l'échelle du réseau. Nous n'avons cependant pas réussi à reproduire ces deux comportements simultanément et de façon distincte.

Les principaux avantages de ces deux méthodes de simulation de propriétés sont les suivants :

- Elles ne sont pas basées sur une grille régulière ou cartésienne.
- Elles permettent de respecter la variabilité spatiale des propriétés simulées au sein de chaque conduit plutôt qu'utiliser une seule valeur pour le conduit entier.
- Elles utilisent un squelette (de type graphe) défini par l'utilisateur comme support et il n'y a pas de dépendance entre sa création et la génération des propriétés.
- Elles sont stochastiques.
- Elles peuvent être appliquées à toute propriété qui présenterait une corrélation spatiale le long d'un réseau curvilinéaire ou d'un graphe.

L'efficacité de ces deux méthodes a pu être démontrée sur plusieurs cas synthétiques et sur un cas réel, le réseau d'Arrestelia (présenté en détail dans le Chapitre 2).

Simulation du squelette des réseaux

Bien qu'elle ne soit encore qu'à l'état de prototype, une méthode ayant pour but de générer des réseaux karstiques discrets sous forme de graphes (*squelettes*) a été présentée dans le Chapitre 4. Cette méthode se décompose en trois étapes distinctes : i) la génération de points à partir d'une grille de probabilité, ii) le calcul de leur connectivité théorique et iii) leur connexion pour former le squelette.

La première étape n'est pas encore implémentée mais elle consisterait à générer aléatoirement un nombre points voulu dans une grille de probabilité. La grille devrait pouvoir prendre en compte les données de terrain, les informations géologiques et d'autres informations données par l'utilisateur. La deuxième étape consiste à déterminer le rôle théorique des points au sein du graphe final. Il leur est assigné un rôle

d'extrémité, d'intersection ou de nœud de liaison, principalement en fonction de leur position au sein du nuage de points et des proportions souhaitées par l'utilisateur. Enfin, la troisième étape consiste à relier les nœuds en respectant la connectivité déterminée à l'étape précédente ainsi que d'autres paramètres (notamment la distance maximale entre deux nœuds reliés) pour obtenir un réseau final cohérent. L'état actuel de la deuxième et de la troisième étape est présenté à travers leur application à différents cas synthétiques.

Dans son état final, les points forts de cette méthode seraient les suivants :

- Elle permettrait la simulation de réseaux tridimensionnels non alignés sur une grille.
- Elle serait stochastique.
- Les temps de calcul devraient être raisonnables.
- La topologie des réseaux simulés pourrait être contrainte par l'utilisateur.
- Elle devrait permettre l'intégration de données de terrain (localisation de conduits connus) et d'informations géologiques et hydrogéologiques.

Perspectives

Inférence statistique

La base de données utilisée lors de l'analyse statistique présentée dans cette thèse comporte 49 réseaux karstiques différents. Aucune étude similaire n'a, à notre connaissance, été effectuée sur un tel nombre de réseaux, mais pour autant sa robustesse reste limitée, surtout pour l'analyse des différents motifs morphologiques. Beaucoup de réseaux utilisés ne comportent qu'un faible nombre de points de données. Nombre d'entre eux sont également situés dans des localités françaises proches les unes des autres et faisant parfois partie du même système ce qui conduit à un manque de diversité. Étendre cette étude à un nombre de réseaux bien plus important augmenterait significativement sa robustesse. Les données disponibles peuvent également manquer de précision. Leur acquisition ayant pour but principal la cartographie des réseaux, cela résulte en des échantillonnages irréguliers. Les dimensions des conduits sont, de plus, estimées visuellement et approximées. Procéder à des échantillonnages beaucoup plus réguliers et à des mesures plus précises améliorerait grandement la robustesse de cette étude. Les nouvelles techniques telles que les récentes technologies LIDAR laissent espérer des données plus précises à l'avenir. Notre étude se basant principalement sur les hauteurs et largeurs mesurées des conduits, via l'assimilation de leur section à une ellipse, pouvoir analyser la géométrie réelle des conduits pourrait offrir de nouvelles informations.

Il convient également d'améliorer les méthodes développées pour notre analyse. La définition actuelle des branches est purement topologique (entre deux nœuds de degré différent de 2) et celles-ci ne peuvent être considérées comme réellement représentatives des conduits eux-mêmes. Cela limite la pertinence des variogrammes intrabranches et interbranches car plusieurs parties d'un même conduit, spéléogénétiquement parlant, peuvent correspondre à des branches différentes. Utiliser une nouvelle définition est

envisageable mais il risque d'être compliqué de développer un algorithme offrant des résultats pertinents et ayant de faibles temps de calcul. Pour autant, il devrait être possible de définir manuellement les différentes branches du réseau.

L'algorithme de hiérarchisation créé lors de cette thèse est également très limité et devrait être amélioré afin de pouvoir être appliqué aux réseaux présentant des siphons.

Simulation des propriétés géométriques des conduits

La méthode de SGS 1D-curvilinéaire permet de reproduire le variogramme global. La méthode de SGS 1D-curvilinéaire *branchwise* devrait théoriquement reproduire à la fois les variogrammes intrabranches, interbranches et globaux mais les tests nous ont démontré qu'il est compliqué de respecter plusieurs comportements simultanément.

Sur les grands réseaux, les calculs peuvent être longs et coûteux en mémoire. Optimiser le code permettrait de réduire l'importance de ces problèmes.

Au risque de moins bien respecter les paramètres d'entrée, il serait intéressant d'analyser les résultats de simulations effectuées en se servant d'un krigeage ordinaire et non d'un krigeage simple comme c'est le cas actuellement. Cela permettrait potentiellement d'obtenir une meilleure cohérence locale des résultats.

Il a été vu que pour des variogrammes et une distribution d'entrée identique, les résultats de la méthode *branchwise* changent fortement en fonction des autres paramètres d'entrée que sont la proportion *prop* et le nombre de nœuds *nbnœudsmax* simulés dans chaque branche avec le variogramme interbranche. Donner plus de poids au variogramme interbranche ou au variogramme intrabranche lors de la simulation fait naturellement pencher la balance vers l'un ou l'autre dans les résultats. Une analyse de sensibilité serait évidemment un atout pour mieux identifier les paramètres qui influencent le plus les résultats.

Comme dit précédemment, la définition des branches, qui impacte directement les variogrammes intrabranches et interbranches, est actuellement purement topologique. Un éventuel changement de définition des branches à la fois pour le calcul des variogrammes et pour la méthode de simulation *branchwise* pourrait potentiellement améliorer les résultats. Les conduits étant moins nombreux et plus longs, on pourrait espérer des variogrammes intrabranches et interbranches plus distincts et plus simples à reproduire simultanément : le comportement intrabranche devrait être plus marqué même pour une proportion de points simulés avec le variogramme interbranche plus importante.

Ces méthodes sont théoriquement applicables à n'importe quelle propriété d'un objet curvilinéaire à condition qu'une certaine continuité des valeurs soit observable à courte distance et quantifiable avec des variogrammes 1D-curvilinéaires. L'utilisation de ces méthodes dans d'autres cadres d'études pourrait s'avérer intéressante.

Les difficultés à reproduire plusieurs variogrammes simultanément sont directement liées à l'implémentation actuelle de notre méthode de SGS 1D-curvilinéaire *branchwise*. Il pourrait être intéressant de changer son concept même, voire d'adapter d'autres méthodes de simulation utilisées en géostatistique telles que les méthodes de simulations spectrales, de moyennes mobiles ou de *turning bands* au cas présent. Ces méthodes sont en effet connues pour permettre une bonne reproduction des variogrammes d'entrée, bien que les méthodes de SGS semblent particulièrement bien adaptées pour la reproduction

de variogrammes ayant une faible portée et un effet pépité (Pinto et Deutsch, 2020). Bien qu'elles ne permettent pas de reproduire les variogrammes, le développement de méthodes basées sur les Simulation Multipoints (MPS) est également envisagé.

Simulation du squelette des réseaux

La méthode de simulation du squelette des réseaux est actuellement loin d'être complète. La priorité concerne la première étape, c'est à dire la génération des points. Le travail à effectuer se décompose suivant deux principaux aspects qui devraient être combinés : i) l'intégration des données de terrain et des informations géologiques et ii) le remplissage automatique de la grille. Il nous faut actuellement définir une méthode pertinente d'intégration des données. Les conduits connus devraient être directement intégrés comme partie du réseau final mais il est nécessaire d'évaluer la probabilité de présence d'autres conduits autour de ces derniers afin que la grille de probabilité soit cohérente. De la même manière, il est nécessaire de déterminer les bonnes valeurs de probabilité à assigner aux cellules dans lesquelles se trouvent des objets géologiques susceptibles de favoriser la formation de conduits. Il est également nécessaire de définir une méthode capable de remplir de manière cohérente les cellules de la grille n'étant pas associées à des données, voire la grille entière, dans le cas où aucune donnée ne soit disponible. Il serait de plus souhaitable de contraindre la grille en fonction du type morphologique de réseau (e.g., arborescent, anastomosé). L'utilisateur devra être en mesure de trouver une résolution adaptée pour la grille de probabilité. Une résolution grossière pourrait conduire à la génération de points éloignés des surfaces d'inception, tandis qu'une résolution trop fine pourrait résulter en de grands temps de calcul. L'implémentation de cette étape s'annonce donc comme un véritable défi et il sera nécessaire de procéder à de nombreux tests avant de trouver la meilleure manière de calculer les probabilités.

Lors de l'étape de calcul de la connectivité, la topologie du réseau est principalement contrainte par le ratio intersections/extrémités utilisé comme paramètre d'entrée. Bien que ce dernier ait pu être étudié sur notre base de données (Annexe B), les résultats obtenus lors de l'analyse des différents motifs morphologiques ne peuvent pas être qualifiés de robustes. L'étude d'une base de données plus grande, comme proposé précédemment, permettrait d'améliorer la compréhension de ce paramètre et d'obtenir des résultats plus proches des observations de terrain. Dans l'état actuel de l'algorithme, le ratio intersections/extrémités et les autres paramètres sont principalement donnés à titre indicatif et il est complexe de les reproduire avec l'algorithme actuel. Trouver d'autres paramètres permettant de bien contraindre le type de réseau voulu s'avèrerait bénéfique. Les défauts majeurs de l'étape de connexion sont le manque de respect de la connectivité calculée lors de l'étape précédente, le manque de connectivité et le manque de cohérence locale des résultats d'un point de vue géologique. Améliorer ces trois points augmenterait grandement le réalisme des résultats.

Cette méthode devrait également être adaptée pour permettre la simulation de conduits dont l'existence est connue, notamment grâce à des essais de traçage, mais n'ayant pas pu être explorés. Il devrait être possible de définir les points d'entrée et de sortie, de délimiter une certaine zone autour d'eux et de simuler le chemin les reliant.

Enfin, nous souhaiterions être en mesure de simuler plusieurs morphologies de réseaux (e.g., arborescent, anastomosé). Une fois cela pleinement maîtrisé, il nous faudra

adapter la méthode de sorte qu'une simulation séquentielle ou régionalisée nous permette de combiner ces motifs dans un même réseau.

Le chemin à parcourir reste donc long et les verrous scientifiques nombreux avant de parvenir à une méthode opérationnelle.

Évaluation et utilisation des réseaux simulés

Un aspect important des simulations est l'évaluation de leurs résultats. Nous souhaiterions simuler des réseaux *réalistes* mais cette notion peut être difficile à évaluer de façon objective. La manière la plus simple pour juger le réalisme des modèles obtenus est de procéder à leur analyse statistique et de comparer les résultats à ceux de réseaux connus. Les résultats des méthodes de simulation des propriétés présentées dans le Chapitre 3 sont directement comparés aux paramètres d'entrée. L'analyse statistique présentée dans le Chapitre 2 peut servir à la fois de référence pour les paramètres d'entrée des simulations et pour l'évaluation des résultats. Il sera à priori plus compliqué d'évaluer les réseaux obtenus par la méthode de simulation des squelettes présentée dans le Chapitre 4 car il ne s'agit pas d'une méthode purement statistique. Plusieurs méthodes comparant certains paramètres d'entrée (e.g., les orientations des conduits) aux résultats (e.g., Pardo-Igúzquiza et al. [2012]; Fournillon et al. [2012]). Dans notre cas, exception faite du nombre de nœuds, du nombre d'extrémités et du ratio intersections/extrémités, qui ne peuvent pas toujours être respectés, les autres paramètres d'entrée seront à priori optionnels, ce qui limitera les comparaisons possibles. Il est néanmoins envisageable d'analyser différentes métriques (e.g., Renard et al., 2013; Collon et al., 2017), principalement topologiques, sur notre base de données et de s'en servir comme références pour évaluer le réalisme des réseaux simulés. Étant donné que nous souhaiterions combiner plusieurs motifs dans un même réseau simulé, il nous faudra également être capables de pouvoir les séparer et de les analyser individuellement. Il serait intéressant de développer une méthode permettant d'évaluer et de comparer plusieurs réalisations, obtenues par des méthodes et des paramètres d'entrée identiques ou différents, selon des critères précis, comme cela avait pu être fait pour les chenaux par Rongier et al. [2016]. Tous les paramètres caractérisant un réseau ne pouvant être quantifiés, l'avis de l'utilisateur sera également important.

La principale application des réseaux simulés sera de servir de support à des simulations d'écoulement (e.g., SWMM, Epanet, Modflow-CFP). Il sera très intéressant d'étudier l'impact des paramètres d'entrée, notamment les dimensions des conduits, sur les flux simulés. Les métriques *statiques*, servant à étudier le réalisme des réseaux des réseaux générés, évoquées précédemment pourront ici être complétées par des métriques *dynamiques*, traduisant leur connectivité hydrodynamique. Il est même envisageable d'intégrer ces différents paramètres dans une boucle d'inversion afin de les calibrer pour obtenir les résultats hydrodynamiques souhaités.

Annexes

Annexe A

Réseaux utilisés lors de l'étude statistique et résultats

Sont présentées ici les informations générales des réseaux étudiés dans le Chapitre 2, ainsi que le détail des résultats statistiques obtenus sur le rayon équivalent et le ratio largeur-hauteur. D'autres paramètres n'ayant pas été évoqués dans le reste de cette thèse sont également calculés et comparés pour pouvoir servir de référence, comme l'extension du réseau dans ses 3 dimensions spatiales, x , y et z , et le volume de la zone englobant le réseau, obtenu par la multiplication de son extension dans les 3 dimensions. Nous calculons également le volume des conduits connus V_c et regardons la proportion qu'il occupe au sein de la zone d'étude. Pour cela, les volumes de chaque arête dont les deux nœuds sont associés à une valeur de rayon sont calculés et sommés de la façon suivante :

$$V_c = \sum_{e=1}^{N_e} l_i \pi \left(\frac{(Re(i_e) + Re(j_e))}{2} \right)^2 \quad (\text{A.1})$$

où N_e est le nombre d'arêtes dont les deux nœuds ont une valeur de rayon, l_i la longueur de l'arête considérée, $Re(i_e)$ la valeur du rayon équivalent du premier nœud de l'arête et $Re(j_e)$ la valeur du rayon équivalent du deuxième nœud de l'arête. L'indice de verticalité I_v qui reflète le développement vertical du réseau (e.g., Piccini, 2011) est également calculé :

$$I_v = \frac{Z_{ext}}{L} \quad (\text{A.2})$$

Z_{ext} étant l'extension verticale du réseau et L la longueur totale des conduits. Enfin, les indices de karstification des réseaux I_k , qui indiquent l'avancée du processus de karstification dans la zone étudiée (e.g., Pardo-Iguzquiza et al., 2011) sont également présentés :

$$I_k = \frac{L}{V_{tot}^{\frac{1}{3}}} \quad (\text{A.3})$$

où L est la longueur totale des conduits et V_{tot} le volume de la zone englobant le réseau.

Annexe A. Réseaux utilisés lors de l'étude statistique et différents résultats

Nom	Localisation	Nombre d'arêtes	Nombre de nœuds	Nombre de nœuds sans valeur	Distance moyenne entre les points (m)
Abisso Chimera	Italie (Alpes apuanes)	1134	1128	129	8.5
Agas	France (Gard)	614	615	35	7.8
Arphidia Robinet	France (Pyrénées-Atlantiques)	1483	1479	54	9.0
Arrestelia	France (Pyrénées-Atlantiques)	6012	6000	1717	9.7
Aspirateur	France (Gard)	268	269	4	7.0
Autran	France (Vaucluse)	1669	1656	142	6.0
Barrage Gardon	France (Gard)	705	703	16	6.1
Baume	France (Hérault)	178	171	49	3.5
Baume Galiniere	France (Vaucluse)	53	50	18	3.8
Baume Salene	France (Gard)	212	213	3	10.0
Bez	France (Gard)	189	190	2	7.5
Bornegre	France (Gard)	257	257	7	7.7
Bret	France (Gard)	240	238	14	5.1
Brun	France (Gard)	650	649	58	9.8
Buse	France (Gard)	54	55	0	6.9
Calles	France (Gard)	122	123	5	11.4
Cazilhac	France (Hérault)	528	525	16	6.2
Ceberi	France (Hautes-Pyrénées)	619	618	36	9.1
Cellagua	Espagne (Cantabrie)	44	45	1	7.2
Chamois	France (Alpes-de-Haute-Provence)	1224	1190	463	10.7
Chapo	Espagne (Cantabrie)	162	163	12	9.2
Charentais Heche	France (Hautes-Pyrénées)	1985	1978	237	6.7
Cocaliere	France (Gard)	275	276	2	8.8
Cougnat	France (Gard)	148	148	8	9.3
Due Dong	Vietnam	75	76	10	9.4
Eau Relie	France (Var)	204	204	172	3.3
Everest	France (Gard)	94	94	2	5.7
Foussoubie	France (Ardèche)	2664	2624	218	8.7
Gardies	France (Hérault)	142	143	1	7.8
Garma Ciega	Espagne (Cantabrie)	585	586	10	12.2
Lali	Italie (Alpes apuanes)	421	419	59	11.5
Malaval	France (Lozère)	2269	2262	185	5.3
Mazo Chico	Espagne (Cantabrie)	1930	1924	77	6.2
Nam Pakan	Vietnam	364	362	81	8.0
Peyrejal	France (Ardèche)	102	102	4	7.5
Pigette 1	France (Alpes-de-Haute-Provence)	88	87	16	3.3
Pigette 2	France (Alpes-de-Haute-Provence)	144	141	31	3.2
Plantayrol	France (Hérault)	213	214	6	9.1
Rubicera	Espagne (Cantabrie)	383	384	7	11.2
Saint Benoit	France (Alpes-de-Haute-Provence)	399	397	56	5.4
Saint Marcel	France (Ardèche)	4459	4456	655	10.6
Saint Sebastien	France (Alpes-de-Haute-Provence)	81	79	26	4.0
Sakany	France (Ariège)	1829	1822	145	4.1
Satan	France (Gard)	88	89	0	5.4
Sauvas	France (Ardèche)	63	64	11	8.5
Sergent General	France (Hérault)	258	259	2	10.4
Sieben Hengste	Suisse	15570	15340	196	5.4
Souchon	France (Gard)	131	132	11	4.3
Wakulla	USA	477	474	6	11.5

Nom	Numéro du réseau	Rayon moyen	Rayon médian	Écart type rayon	Coeff. Pearson rayon	Longueur totale (m)
Abisso Chimera	1	2.26	1.37	2.92	0.81	9616
Agas	2	1.26	1.02	0.82	0.54	4811
Arphidia Robinet	3	1.72	1.50	0.83	0.89	13250
Arrestelia	4	1.68	1.20	1.59	0.71	58429
Aspirateur	6	1.81	1.38	1.43	0.49	1886
Autran	5	1.49	1.12	1.49	0.68	9993
Barrage Gardon	7	1.20	1.06	0.75	0.52	4294
Baume	8	0.77	0.57	0.59	0.69	603
Baume Galiniere	9	0.63	0.56	0.31	0.81	189
Baume Salene	10	2.19	1.40	2.34	0.66	2135
Bez	11	1.65	1.62	0.91	0.51	1432
Bornegre	12	0.94	0.95	0.41	0.71	1973
Bret	13	1.26	0.93	1.06	0.49	1206
Brun	14	1.66	1.17	1.61	0.57	6351
Buse	15	1.79	1.20	1.74	0.50	379
Calles	16	4.81	3.69	4.37	0.82	1404
Cazilhac	17	1.29	0.99	1.14	0.57	3231
Ceberi	18	1.69	1.22	1.46	0.70	5606
Cellagua	19	3.32	2.83	1.71	0.68	324
Chamois	20	1.74	1.22	1.41	0.81	12691
Chapo	21	2.60	2.12	2.21	0.85	1494
Charentais Heche	22	1.54	1.30	0.99	0.68	13329
Cocaliere	23	2.49	2.32	1.18	0.61	2427
Cougnet	24	1.47	1.11	0.93	0.69	1383
Due Dong	25	5.02	4.80	2.47	0.71	718
Eau Relie	26	1.02	1.00	0.48	0.88	673
Everest	27	1.64	1.33	1.22	0.39	531
Foussoubie	28	1.10	1.00	0.66	0.91	22950
Gardies	29	1.26	1.15	0.58	0.40	1113
Garma Ciega	30	3.78	3.24	2.65	0.88	7125
Lali	31	1.97	1.58	1.40	0.67	4827
Malaval	32	1.62	1.40	1.01	0.68	11961
Mazo Chico	33	2.14	1.84	1.22	0.83	11994
Nam Pakan	34	1.86	1.51	2.56	0.85	2897
Peyrejal	35	1.19	1.02	0.71	0.64	762
Pigette 1	36	0.82	0.71	0.54	0.45	290
Pigette 2	37	0.84	0.71	0.50	0.71	456
Plantayrol	38	1.35	1.08	0.93	0.51	1947
Rubicera	39	3.07	2.29	2.97	0.76	4282
Saint Benoit	40	0.72	0.57	0.50	0.79	2125
Saint Marcel	41	1.56	1.22	1.28	0.73	47249
Saint Sebastien	42	0.42	0.39	0.21	0.78	314
Sakany	43	1.19	0.94	1.08	0.73	7528
Satan	44	1.04	0.81	0.81	0.28	478
Sauvas	45	2.66	2.24	1.23	0.80	546
Sergent General	46	1.84	1.69	0.99	0.54	2683
Sieben Hengste	47	1.22	0.92	1.09	0.70	82238
Souchon	48	1.60	1.42	1.20	0.51	569
Wakulla	49	2.09	1.87	1.59	0.88	5440

Nom	WH moyen	WH médian	Écart type WH	Coeff. Pearson WH	Indice de verticalité	Indice de karstification
Abisso Chimera	0.92	0.50	1.73	0.51	1.08E-01	1.43E+01
Agas	1.17	1.00	1.01	0.34	3.37E-02	1.11E+01
Arphidia Robinet	1.21	1.00	0.48	0.86	4.44E-02	1.51E+01
Arrestelia	0.74	0.50	0.82	0.45	7.12E-03	3.72E+01
Aspirateur	1.18	0.72	1.39	0.35	5.65E-02	9.74E+00
Autran	0.71	0.25	1.40	0.42	6.17E-02	1.08E+01
Barrage Gardon	1.21	0.86	1.30	0.35	3.03E-02	1.23E+01
Baume	1.90	1.48	1.47	0.42	5.74E-02	1.19E+01
Baume Galiniere	1.08	0.91	0.94	0.53	5.73E-02	9.32E+00
Baume Salene	2.47	1.50	3.39	0.67	4.12E-02	7.05E+00
Bez	1.02	0.67	1.05	0.60	1.28E-01	3.66E+00
Bornegre	1.82	1.61	1.10	0.45	2.84E-02	9.32E+00
Bret	1.52	1.19	1.63	0.37	7.52E-02	1.08E+01
Brun	1.52	1.00	1.70	0.49	4.19E-02	9.70E+00
Buse	1.00	0.83	0.74	0.21	2.55E-01	3.70E+00
Calles	1.33	1.00	1.32	0.43	7.91E-02	4.34E+00
Cazilhac	1.23	1.02	1.00	0.28	3.57E-02	7.96E+00
Ceberi	1.08	0.67	1.41	0.46	1.99E-02	1.32E+01
Cellagua	0.25	0.13	0.46	0.68	7.44E-01	5.07E+00
Chamois	1.87	1.26	2.30	0.44	2.55E-02	1.38E+01
Chapo	0.40	0.20	0.60	0.28	2.30E-01	8.23E+00
Charentais Heche	1.18	0.60	1.70	0.67	3.44E-02	1.32E+01
Cocaliere	2.02	1.38	1.96	0.57	1.08E-02	9.29E+00
Cougnat	2.42	2.16	1.71	0.46	3.30E-02	8.79E+00
Due Dong	1.40	1.25	0.96	0.45	3.55E-02	7.09E+00
Eau Relie	0.91	1.00	0.70	0.34	1.94E-01	4.25E+00
Everest	1.02	0.71	1.28	0.04	1.43E-01	8.12E+00
Foussoubie	1.32	1.20	1.08	0.58	5.94E-03	2.34E+01
Gardies	1.42	1.04	1.21	0.26	9.61E-03	9.68E+00
Garma Ciega	0.53	0.36	0.66	0.71	1.16E-01	5.93E+00
Lali	1.51	1.20	1.18	0.46	5.86E-02	9.13E+00
Malaval	0.81	0.35	1.59	0.39	1.91E-02	1.40E+01
Mazo Chico	0.54	0.17	0.92	0.76	6.00E-02	1.13E+01
Nam Pakan	1.84	1.43	1.79	0.53	1.26E-02	1.29E+01
Peyrejal	2.11	1.55	1.70	0.27	2.25E-02	9.03E+00
Pigette 1	1.75	1.50	1.17	0.45	1.21E-01	8.67E+00
Pigette 2	1.61	1.19	1.55	0.25	1.11E-01	8.76E+00
Plantayrol	1.05	0.82	0.91	0.39	9.93E-02	5.38E+00
Rubicera	1.43	0.67	2.48	0.72	1.10E-01	5.21E+00
Saint Benoit	2.77	2.00	2.24	0.56	3.60E-02	1.25E+01
Saint Marcel	2.28	1.67	2.16	0.54	5.01E-03	3.33E+01
Saint Sebastien	1.63	1.50	0.76	0.80	2.67E-02	1.11E+01
Sakany	0.85	0.60	1.05	0.62	1.91E-02	3.63E+01
Satan	1.34	1.00	1.30	0.14	1.08E-01	4.02E+00
Sauvas	0.92	0.40	1.55	0.41	8.87E-02	3.77E+00
Sergent General	1.11	0.74	1.16	0.54	4.87E-02	6.11E+00
Sieben Hengste	0.82	0.35	1.30	0.51	1.20E-02	4.27E+01
Souchon	1.12	0.77	1.29	0.00	1.37E-01	7.16E+00
Wakulla	1.57	1.48	1.05	0.43	5.23E-03	1.70E+01

Nom	Ext. x (m)	Ext. y (m)	Ext. z (m)	Volume massif (m^3)	Volume conduits (m^3)	Prop. volume conduits (%)
Abisso Chimera	7.18E+02	4.08E+02	1.04E+03	3.05E+08	4.43E+05	0.15
Agas	3.66E+02	1.39E+03	1.62E+02	8.24E+07	3.51E+04	0.04
Arphidia Robinet	1.15E+03	1.00E+03	5.88E+02	6.77E+08	1.67E+05	0.02
Arrestelia	2.44E+03	3.81E+03	4.16E+02	3.86E+09	1.08E+06	0.03
Aspirateur	2.39E+02	2.85E+02	1.06E+02	7.27E+06	4.70E+04	0.65
Autran	1.06E+03	1.20E+03	6.17E+02	7.82E+08	9.69E+04	0.01
Barrage Gardon	5.17E+02	6.31E+02	1.30E+02	4.25E+07	3.34E+04	0.08
Baume	6.53E+01	5.74E+01	3.46E+01	1.30E+05	5.24E+02	0.40
Baume Galiniere	2.84E+01	2.71E+01	1.08E+01	8.33E+03	6.54E+01	0.78
Baume Salene	7.57E+02	4.18E+02	8.79E+01	2.78E+07	9.96E+04	0.36
Bez	5.61E+02	5.83E+02	1.83E+02	6.00E+07	1.67E+04	0.03
Bornegre	3.47E+02	4.88E+02	5.60E+01	9.49E+06	6.92E+03	0.07
Bret	1.08E+02	1.41E+02	9.07E+01	1.38E+06	1.23E+04	0.89
Brun	9.90E+02	1.07E+03	2.66E+02	2.81E+08	9.26E+04	0.03
Buse	6.80E+01	1.64E+02	9.68E+01	1.08E+06	1.20E+04	1.11
Calles	6.03E+02	5.07E+02	1.11E+02	3.39E+07	2.46E+05	0.73
Cazilhac	8.73E+02	6.64E+02	1.15E+02	6.69E+07	4.02E+04	0.06
Ceberi	1.66E+03	4.09E+02	1.12E+02	7.59E+07	1.08E+05	0.14
Cellagua	4.96E+01	2.18E+01	2.41E+02	2.61E+05	2.07E+04	7.94
Chamois	1.55E+03	1.53E+03	3.24E+02	7.72E+08	1.28E+05	0.02
Chapo	2.29E+02	7.61E+01	3.43E+02	5.98E+06	5.03E+04	0.84
Charentais Heche	1.81E+03	1.23E+03	4.59E+02	1.02E+09	1.33E+05	0.01
Cocaliere	7.16E+02	9.53E+02	2.62E+01	1.78E+07	5.52E+04	0.31
Cougnat	5.21E+02	1.64E+02	4.56E+01	3.89E+06	1.70E+04	0.44
Due Dong	1.78E+02	2.30E+02	2.55E+01	1.04E+06	6.18E+04	5.95
Eau Relie	1.90E+02	1.61E+02	1.30E+02	3.97E+06	2.12E+02	0.01
Everest	7.25E+01	5.10E+01	7.58E+01	2.80E+05	1.01E+04	3.59
Foussoubie	2.39E+03	2.90E+03	1.36E+02	9.44E+08	1.16E+05	0.01
Gardies	7.71E+02	1.84E+02	1.07E+01	1.52E+06	7.10E+03	0.47
Garma Ciega	2.13E+03	9.84E+02	8.27E+02	1.74E+09	6.48E+05	0.04
Lali	4.77E+02	1.10E+03	2.83E+02	1.48E+08	6.85E+04	0.05
Malaval	8.74E+02	3.14E+03	2.28E+02	6.27E+08	1.34E+05	0.02
Mazo Chico	1.07E+03	1.54E+03	7.20E+02	1.19E+09	2.48E+05	0.02
Nam Pakan	8.73E+02	3.58E+02	3.64E+01	1.14E+07	3.93E+04	0.35
Peyrejal	2.45E+02	1.43E+02	1.72E+01	6.01E+05	5.33E+03	0.89
Pigette 1	2.98E+01	3.57E+01	3.51E+01	3.74E+04	4.61E+02	1.23
Pigette 2	5.54E+01	5.03E+01	5.06E+01	1.41E+05	5.72E+02	0.40
Plantayrol	2.91E+02	8.43E+02	1.93E+02	4.74E+07	1.35E+04	0.03
Rubicera	1.35E+03	8.73E+02	4.72E+02	5.56E+08	2.86E+05	0.05
Saint Benoit	1.58E+02	4.03E+02	7.65E+01	4.88E+06	4.78E+03	0.10
Saint Marcel	3.73E+03	3.22E+03	2.37E+02	2.85E+09	7.08E+05	0.02
Saint Sebastien	7.53E+01	3.63E+01	8.39E+00	2.29E+04	5.53E+01	0.24
Sakany	2.87E+02	2.17E+02	1.43E+02	8.94E+06	6.30E+04	0.70
Satan	1.94E+02	1.67E+02	5.19E+01	1.69E+06	2.47E+03	0.15
Sauvas	3.00E+02	2.09E+02	4.84E+01	3.03E+06	1.14E+04	0.38
Sergent General	4.72E+02	1.38E+03	1.31E+02	8.49E+07	4.17E+04	0.05
Sieben Hengste	3.35E+03	2.16E+03	9.88E+02	7.14E+09	9.42E+05	0.01
Souchon	9.24E+01	6.97E+01	7.79E+01	5.02E+05	7.63E+03	1.52
Wakulla	8.93E+02	1.30E+03	2.84E+01	3.29E+07	1.50E+05	0.46

Annexe B

Analyse statistique des motifs morphologiques

Annexe B. Résumé de l'analyse statistique de 40 parties de réseaux dont le motif est connu

Motif	Nombre de réseaux	Nombre de nœuds médian	Nombre de nœuds sans valeur médian	Moyenne des moyennes des rayons	Médiane des médianes des rayons	Écart type médian des rayon
Vadose	16	183.5	8	2.01	1.22	1.51
Water-Table cave	10	90	27	2.07	1.35	1.30
Looping cave	9	350	101	1.36	1.00	0.98
Hypogene	5	87	26	0.70	0.57	0.50

Attention, ces résultats ne sont pas statistiquement significatifs et ne devraient pas être surinterprétés.

Motif	Moyenne des moyennes des WH	Médiane des médianes des WH	Écart type médian des WH	Longueur totale médiane (m)	Médiane de la distance moyenne entre les nœuds (m)
Vadose	0.68	0.45	0.71	912	6.98
Water-Table cave	2.11	1.56	2.31	1199	9.02
Looping cave	1.64	1.20	1.48	3601	8.75
Hypogene	1.59	1.48	1.17	314	3.53

Attention, ces résultats ne sont pas statistiquement significatifs et ne devraient pas être surinterprétés.

Annexe C

Analyse du degré des nœuds

Annexe C. Résumé de l'étude du degré des nœuds dans les différents réseaux

Nom	Nombre de nœuds	Degré 1	Degré 2	Degré 3+	Ratio int/ext
Abisso Chimera	1128	70	985	73	1.04
Agas	615	44	536	35	0.80
Arphidia Robinet	1479	92	1294	93	1.01
Arrestelia	6000	460	5142	398	0.87
Aspirateur	269	42	188	39	0.93
Autran	1656	85	1479	92	1.08
Barrage Gardon	703	66	573	64	0.97
Baume	171	41	86	44	1.07
Baume Galiniere	50	17	20	13	0.76
Baume Salene	213	22	171	20	0.91
Bez	190	7	178	5	0.71
Bornegre	257	7	243	7	1.00
Bret	238	38	167	33	0.87
Brun	649	30	590	29	0.97
Buse	55	3	51	1	0.33
Calles	123	9	108	6	0.67
Cazilhac	525	31	458	36	1.16
Ceberi	618	42	533	43	1.02
Cellagua	45	2	43	0	0.00
Chamois	1190	147	858	185	1.26
Chapo	163	14	138	11	0.79
Charentais Heche	1978	117	1739	122	1.04
Cocaliere	276	4	270	2	0.50
Cognet	148	20	109	19	0.95
Due Dong	76	7	64	5	0.71
Eau Relie	204	9	187	8	0.89
Everest	94	18	60	16	0.89
Foussoubie	2624	135	2282	207	1.53
Gardies	143	3	139	1	0.33
Garma Ciega	586	11	568	7	0.64
Lali	419	38	342	39	1.03
Malaval	2262	179	1910	173	0.97
Mazo Chico	1924	85	1751	88	1.04
Nam Pakan	362	47	268	47	1.00
Peyrejal	102	7	88	7	1.00
Pigette 1	87	15	58	14	0.93
Pigette 2	141	23	89	29	1.26
Plantayrol	214	10	196	8	0.80
Rubicera	384	9	368	7	0.78
Saint Benoit	397	50	295	52	1.04
Saint Marcel	4456	310	3869	277	0.89
Saint Sebastien	79	24	33	22	0.92
Sakany	1822	216	1402	204	0.94
Satan	89	8	75	6	0.75
Sauvas	64	5	56	3	0.60
Sergent General	259	8	246	5	0.63
Sieben Hengste	15340	843	13337	1160	1.38
Souchon	132	19	96	17	0.89
Wakulla	474	23	423	28	1.22

Annexe D

Résultats de simulations de propriétés géométriques

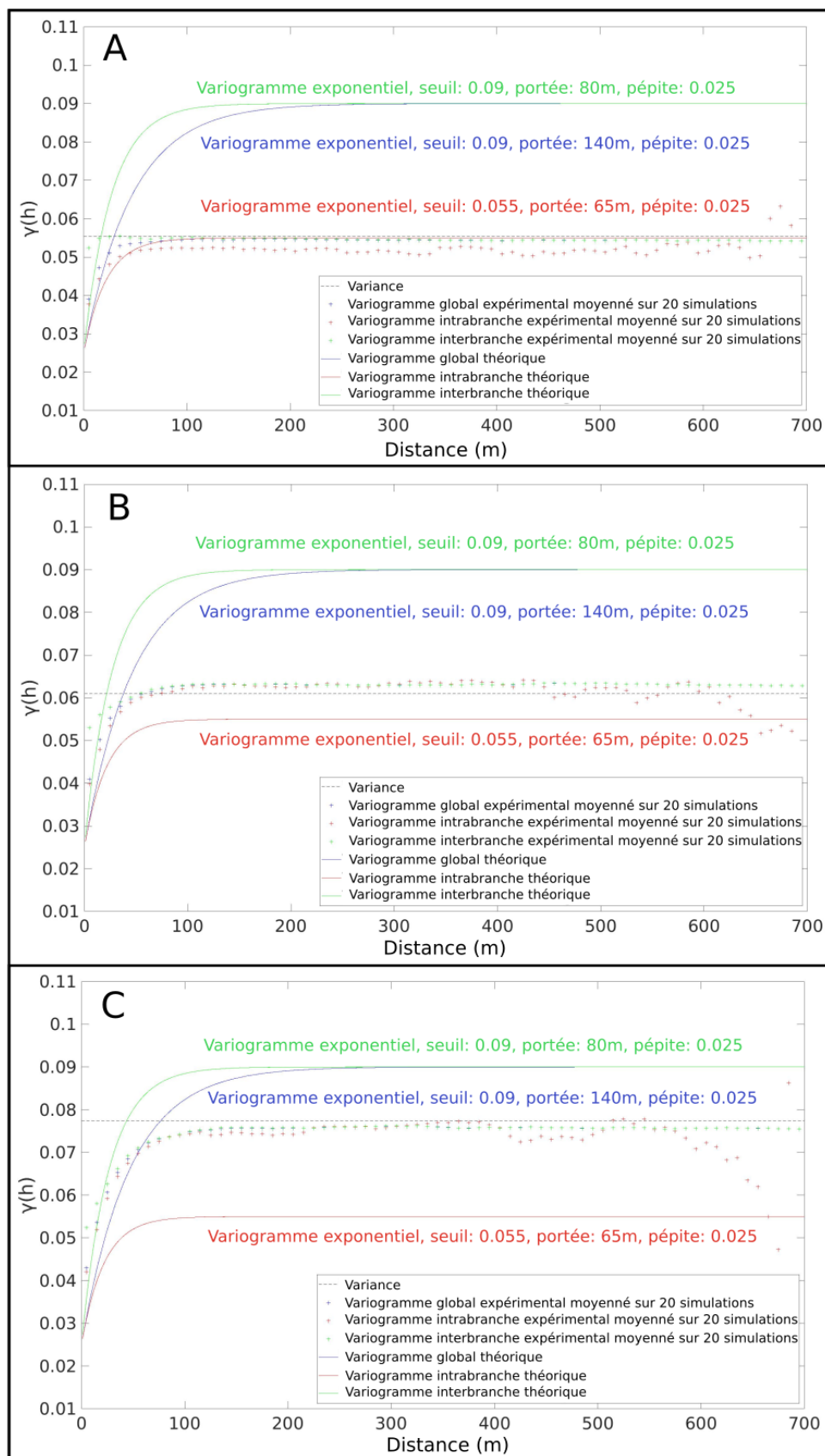


Figure D.1 – Variogrammes expérimentaux moyennés des résultats de 20 simulations non-conditionnelles effectuées avec la méthode de SGS 1D-curvilinéaire *branchwise* sur le réseau d’Arrestelia, pour un nombre maximal de nœuds simulés via le variogramme interbranch non défini et une proportion variable. A : 1%. B : 20%. C : 50%.

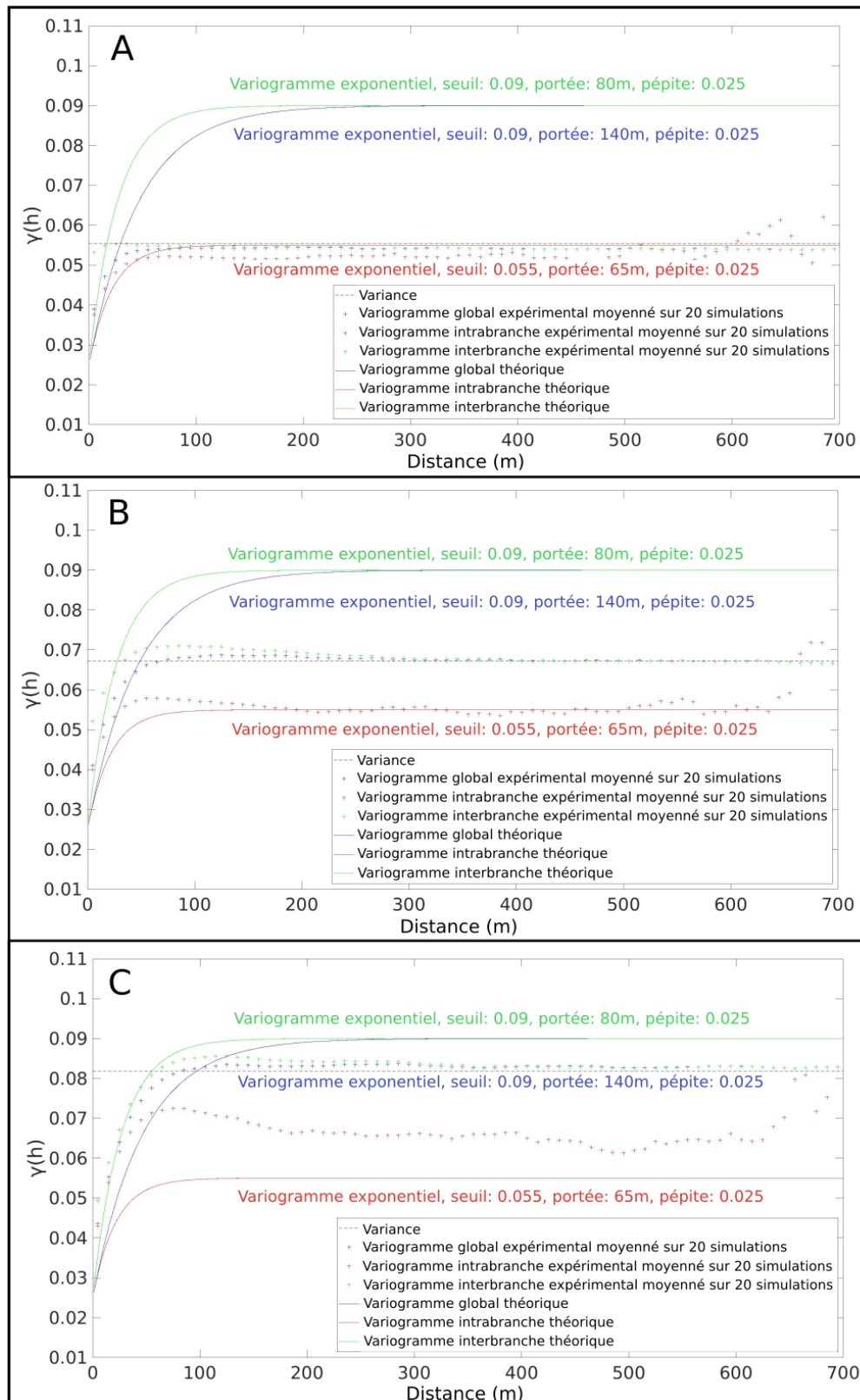


Figure D.2 – Variogrammes expérimentaux moyennés des résultats de 20 simulations non-conditionnelles effectuées avec la méthode de SGS 1D-curvilinéaire *branchwise* sur le réseau d'Arrestelia, pour une proportion de 100% et un nombre maximal de nœuds simulés via le variogramme interbranche variable. A : 1 nœud. B : 7 nœuds. C : 30 nœuds.

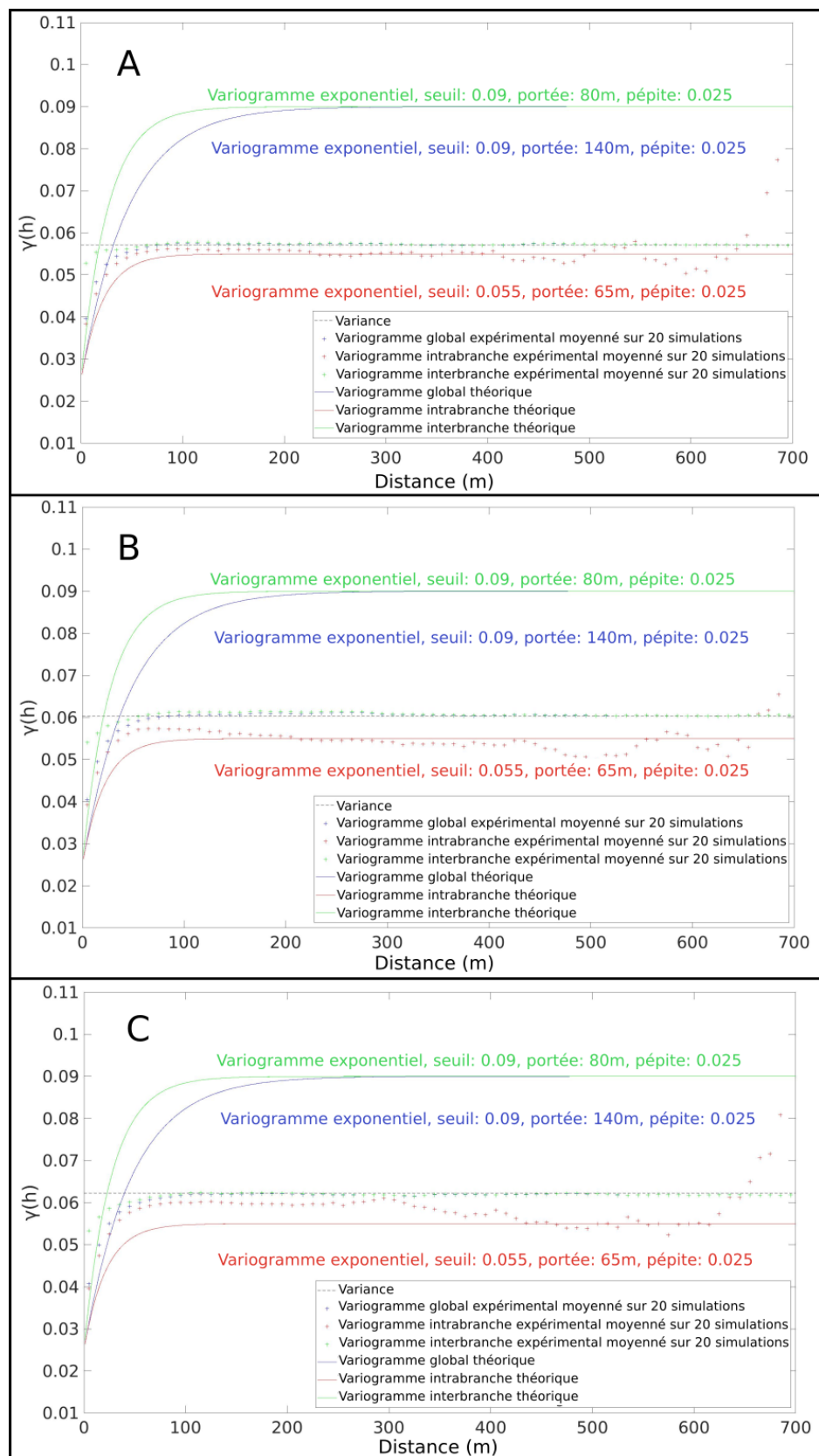


Figure D.3 – Variogrammes expérimentaux moyennés des résultats de 20 simulations non-conditionnelles effectuées avec la méthode de SGS 1D-curvilinéaire *branchwise* sur le réseau d'Arrestelia, pour une proportion un nombre maximal de nœuds simulés via le variogramme interbranche variables. A : 10% et 7 nœuds . B : 20% et 7 nœuds. C : 20% et 15 nœuds.

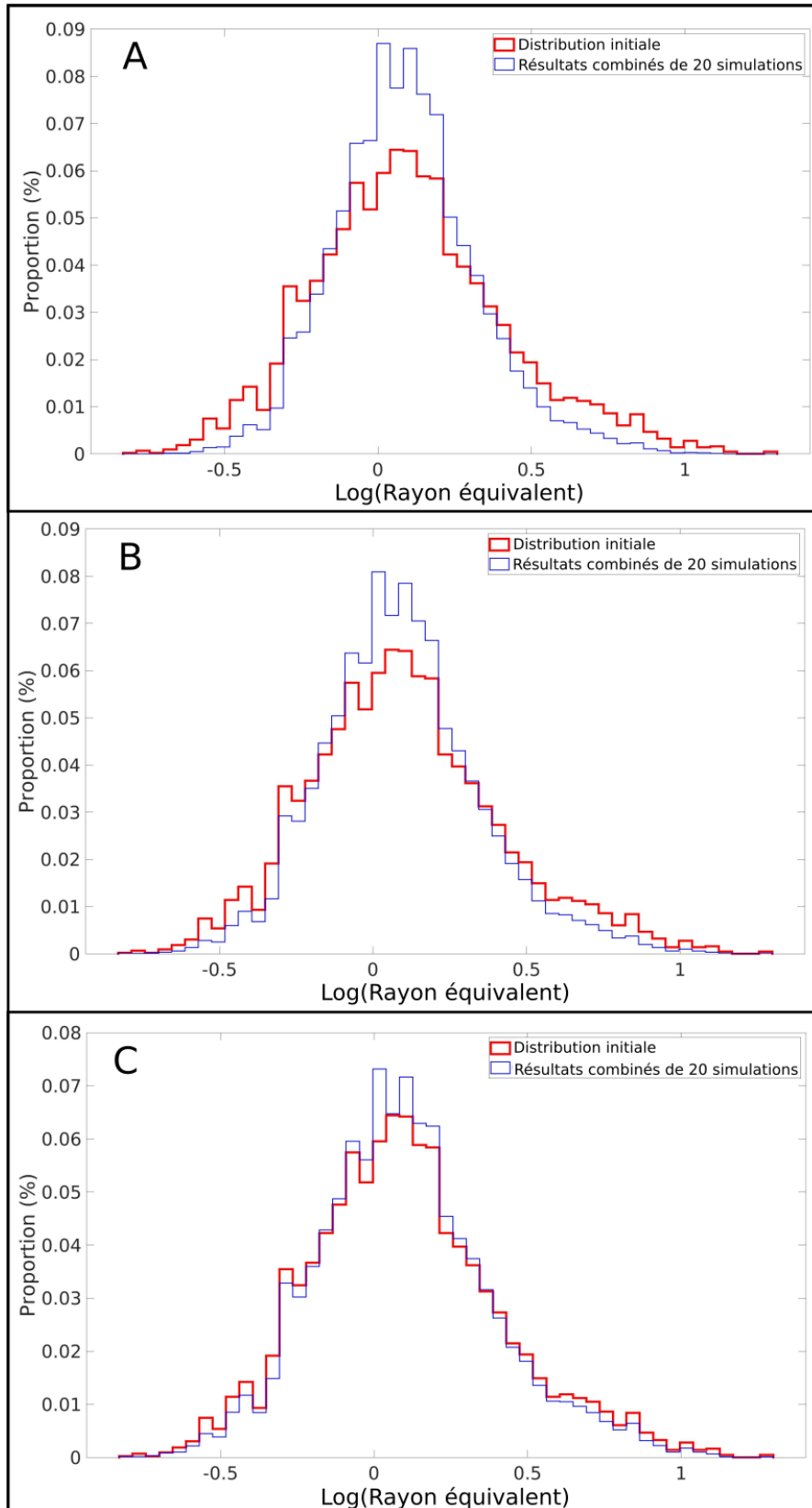


Figure D.4 – Histogramme de la somme des résultats de 20 simulations non-conditionnelles effectuées avec la méthode de SGS 1D-curvilinéaire *branchwise* sur le réseau d'Arrestelia, pour une proportion de 100% et un nombre maximal de nœuds simulés via le variogramme interbranche variable. A : 1 nœud. B : 7 nœuds. C : 30 nœuds.

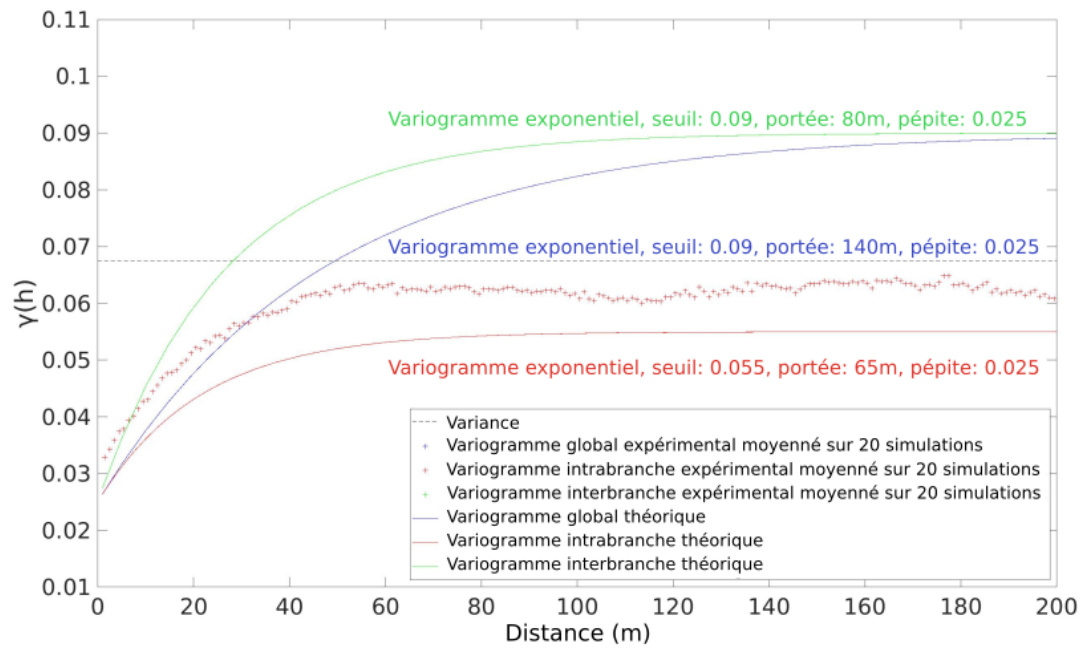


Figure D.5 – Variogrammes expérimentaux moyennés des résultats de 20 simulations non-conditionnelles effectuées sur le cas synthétique A (401 nœuds) avec la méthode de SGS 1D-curvilinéaire *branchwise* et variogramme théorique du logarithme décimal du rayon équivalent. Le variogramme intrabrancher et le variogramme global sont confondus, tandis que le variogramme interbrancher n'existe pas. Bien que ce cas ne soit constitué que d'une unique branche, la prise en compte du variogramme interbrancher théorique en plus du variogramme intrabrancher résulte en des variogrammes intrabranches expérimentaux localisés au-dessus du variogramme théorique.

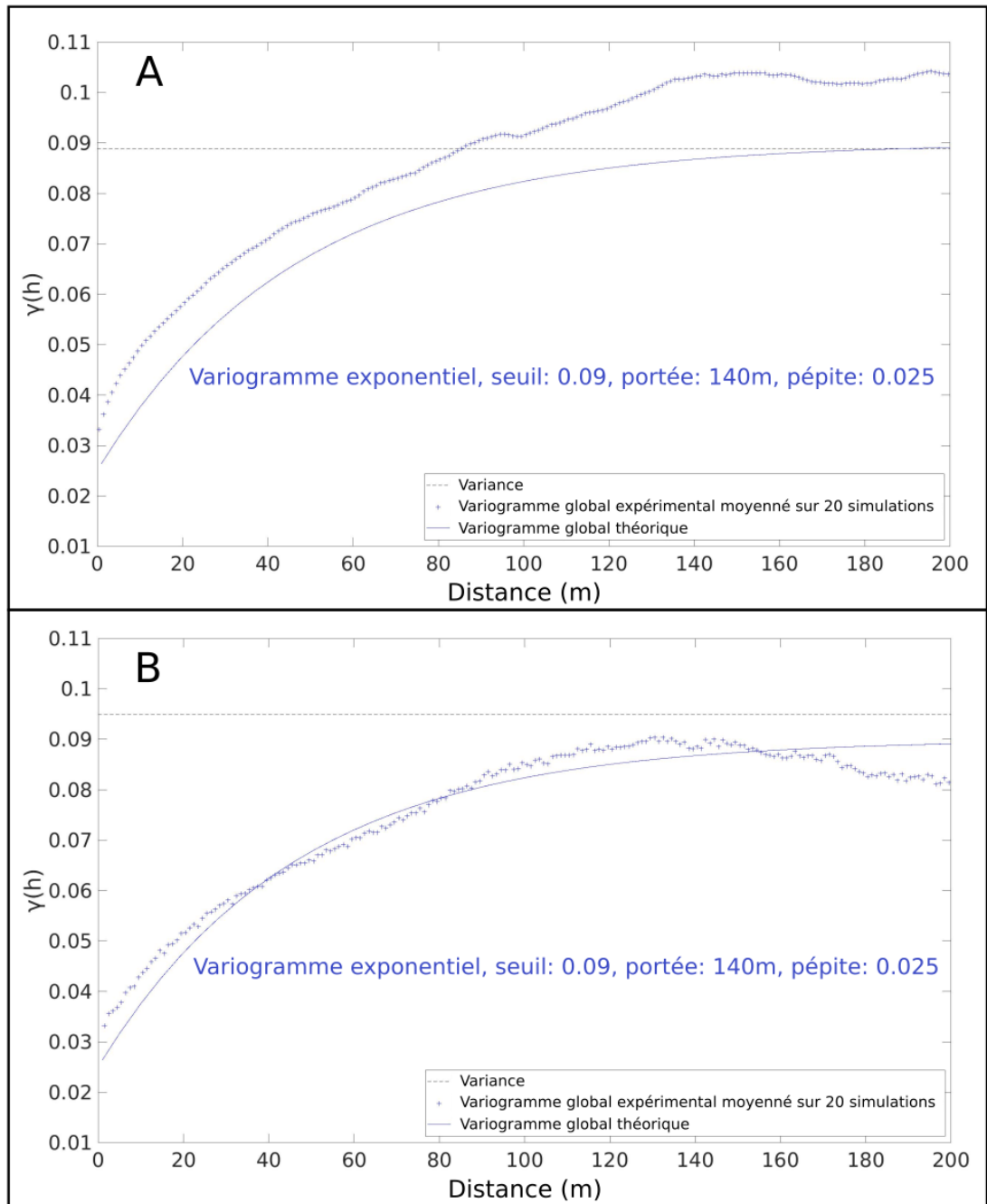


Figure D.6 – Variogramme expérimental moyenné des résultats de 20 simulations non-conditionnelles effectuées sur le cas synthétique A avec la méthode de SGS 1D-courvilinearé et variogramme théorique du logarithme décimal du rayon équivalent. Deux densités de nœuds différentes sont comparées. A : Densité d'un nœud tous les 0.1m, 4001 nœuds au total. B : Densité d'un nœud tous les 1m, 401 nœuds au total. Une densité de nœuds plus importante ne produit pas de meilleurs résultats sur ce cas.

Bibliographie

- T. W. Anderson et D. A. Darling. Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes. *Annals of Mathematical Statistics*, 23(2): 193–212, 1952. doi: 10.1214/aoms/1177729437. (Cited p. 60)
- T. W. Anderson et D. A. Darling. A Test of Goodness of Fit. *Journal of the American Statistical Association*, 49(268): 765–769, 1954. doi: 10.2307/2281537. (Cited p. 60)
- P. Anquez, G. Rongier, et P. Collon. Object-based method for stochastic simulations of karst networks with anastomotic pattern based on Lindenmayer systems. In *35th Gocad-Meeting*, p. 1–14, 2015. (Cited p. 41, 43)
- P. Audra. Hypogene Caves in France. In *Hypogene Karst Regions and Caves of the World*, p. 61–83. 2017. doi: 10.1007/978-3-319-53348-3_3. (Cited p. 26, 27)
- P. Audra et A. N. Palmer. The Vertical Dimension of Karst: Controls of Vertical Cave Pattern. In *Treatise on Geomorphology*, vol. 6, p. 186–206. Elsevier, 2013. ISBN 9780080885223. doi: 10.1016/B978-0-12-374739-6.00098-1. (Cited p. 15, 18, 22, 23, 24, 28, 29, 30)
- P. Audra et A. N. Palmer. Research frontiers in speleogenesis. Dominant processes, hydrogeological conditions and resulting cave patterns. *Acta Carsologica*, 44(3): 315–348, 2015. doi: 10.3986/ac.v44i3.1960. (Cited p. 24)
- P. Audra, F. Hoblea, J.-Y. Bigot, et J.-C. Nobécourt. The Role of Condensation Corrosion in Thermal Speleogenesis : Study of a Hypogenic Sulfidic Cave in Aix-les-Bains, France. *Acta Carsologica*, 36(2): 185–194, 2007. ISSN 0583-6050. doi: 10.3986/ac.v36i2.186. (Cited p. 27)
- P. Audra, L. Mocochain, J.-Y. Bigot, et J.-C. Nobécourt. The association between bubble trails and folia: A morphological and sedimentary indicator of hypogenic speleogenesis by degassing, example from Adaouste Cave (Provence, France). *International Journal of Speleology*, 38(2): 93–102, 2009a. ISSN 1827806X. doi: 10.5038/1827-806X.38.2.1. (Cited p. 24, 26, 27)
- P. Audra, L. Mocochain, J.-Y. Bigot, et J.-C. Nobécourt. Morphological indicators of speleogenesis: Hypogenic speleogens. *Hypogene speleogenesis and karst hydrogeology of artesian basins*, 1(January): 23–32, 2009b. (Cited p. 26)
- P. Audra, L. Mocochain, J.-Y. Bigot, et J.-C. Nobécourt. The pattern of hypogenic caves. In *Proceedings of 15th International Congress of Speleology*, p. 795–800, 2009c. ISBN 9781879961340. (Cited p. 26, 27)

- M. Bakalowicz. Les Processus de karstification et les différents types de karst associés. *Mémoires de la Société géologique de France*, 169: 363–371, 1996. (Cited p. 18)
- M. Bakalowicz. Guide Technique N° 3 Connaissance et gestion des ressources en eau souterraines dans les régions karstiques., 1999. (Cited p. 10)
- M. Bakalowicz. Karst groundwater: A challenge for new resources. *Hydrogeology Journal*, 13(1): 148–160, 2005. ISSN 14312174. doi: 10.1007/s10040-004-0402-9. (Cited p. 12)
- M. Bakalowicz. Epikarst Processes. In *Treatise on geomorphology*, chapter 6.15, p. 164–171. Academic Press Inc, 2013. (Cited p. 24)
- M. E. Barrett et R. J. Charbeneau. A Parsimonious Model for Simulation of Flow and Transport in a Karst Aquifer. Technical report, Center for Research in Water Resources, 1996. (Cited p. 36)
- C. Baudement, B. Arfib, N. Mazzilli, J. Jouvès, T. Lamarque, et Y. Guglielmi. Groundwater management of a highly dynamic karst by assessing base flow and quick flow with a rainfall-discharge model (Dardennes springs , SE France). *Bulletin De La Societe Geologique De France*, 188(6): 20, 2017. doi: <https://doi.org/10.1051/bsgf/2017203>. (Cited p. 36)
- F. Beaud, G. E. Flowers, et S. Pimentel. Seasonal-scale abrasion and quarrying patterns from a two-dimensional ice- flow model coupled to distributed and channelized subglacial drainage. *Geomorphology*, 219: 176–191, 2014. ISSN 0169-555X. doi: 10.1016/j.geomorph.2014.04.036. (Cited p. 13)
- J. Berkson. Some Difficulties of Interpretation Encountered in the Application of the Chi-Square Test. *Journal of the American Statistical Association*, 33(203): 526–536, 1938. doi: 10.2307/2279690. (Cited p. 82)
- A. Bini, P. Tognini, et L. Zuccoli. Rapport entre karst et glaciers durant les glaciations dans les vallées préalpines du Sud des Alpes. *Karstologia*, 32: 7–26, 1998. doi: <https://doi.org/10.3406/karst.1998.2422>. (Cited p. 13)
- S. Birk, R. Liedl, M. Sauter, et G. Teutsch. Hydraulic boundary conditions as a controlling factor in karst genesis : A numerical modeling study on artesian conduit development in gypsum. *Water Resources Research*, 39(1): 1–14, 2003. doi: 10.1029/2002WR001308. (Cited p. 2, 40)
- D. Bittner, T. S. Narany, B. Kohl, M. Disse, et G. Chiogna. Modeling the hydrological impact of land use change in a dolomite- dominated karst system. *Journal of Hydrology*, 567: 267–279, 2018. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2018.10.017. (Cited p. 37)
- A. Bonte. Les remplissages karstiques. *Sedimentology*, 2(4): 333–340, 1962. doi: <https://doi.org/10.1111/j.1365-3091.1963.tb01223.x>. (Cited p. 16)

- A. Borghi, P. Renard, et S. Jenni. A pseudo-genetic stochastic model to generate karstic networks. *Journal of Hydrology*, 414-415: 516–529, 2012. ISSN 00221694. doi: 10.1016/j.jhydrol.2011.11.032. (Cited p. 2, 3, 40, 43, 44, 45, 110)
- A. Borghi, P. Renard, et F. Cornaton. Can one identify karst conduit networks geometry and properties from hydraulic and tracer test data? *Advances in Water Resources*, 90: 99–115, 2016. ISSN 03091708. doi: 10.1016/j.advwatres.2016.02.009. (Cited p. 31, 46, 55, 61, 83, 86)
- P. Boston, M. N. Spilde, D. E. Northup, M. D. Curry, L. A. Melim, et L. Rosales-Lagarde. *Microorganisms as speleogenetic agents: geochemical diversity but geomicrobial unity*. 2009. (Cited p. 18)
- L. Bresinsky, J. Kordilla, M. Sauter, et I. Engelharft. Investigation of the impact of changed precipitation patterns on groundwater recharge processes in a semi-arid carbonate aquifer by an integrated surface-subsurface, multi-continuum model. In *Proceedings of IAH2019, the 46th Annual Congress of the International Association of Hydrogeologists*, p. 588, 2019. (Cited p. 38)
- G. A. Brook, M. E. Folkoff, et E. O. Box. A world model of soil carbon dioxide. *Earth Surface process and landforms*, 8: 79–88, 1983. doi: <https://doi.org/10.1002/esp.3290080108>. (Cited p. 15)
- L. Bruxelles, Y. Quinif, et M. Wiénin. How can ghost rocks help in karst development? In *15th International Congress of Speleology*, p. 814–818, 2009. (Cited p. 16)
- D. Buhmann et W. Dreybrodt. The kinetics of calcite dissolution and precipitation in geologically relevant situations of karst areas: 1. Open System. *Chemical Geology*, 48(1-4): 189–211, 1985. doi: [https://doi.org/10.1016/0009-2541\(85\)90046-4](https://doi.org/10.1016/0009-2541(85)90046-4). (Cited p. 14)
- C. W. Campbell et S. M. Sullivan. Simulating time-varying cave flow and water levels using the Storm Water Management Model. *Engineering Geology*, 65(2-3): 133–139, 2002. doi: 10.1016/S0013-7952(01)00120-X. (Cited p. 38)
- R. Cassidy, J.-c. Comte, J. Nitsche, C. Wilson, R. Flynn, et U. Ofterdinger. Combining multi-scale geophysical techniques for robust hydro-structural characterisation in catchments underlain by hard rock in post-glacial regions. *Journal of Hydrology*, 517: 715–731, 2014. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2014.06.004. (Cited p. 33)
- D. Chapellier. *Well Logging in Hydrogeology*. A A Balkema Publishers, 1992. (Cited p. 35)
- Z. Chen et N. Goldscheider. Modeling spatially and temporally varied hydraulic behavior of a folded karst system with dominant conduit drainage at catchment scale , Hochifen – Gottesacker, Alps. *Journal of Hydrology*, 514: 41–52, 2014. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2014.04.005. (Cited p. 38, 46, 54, 86)
- Z. Chen, A. S. Auler, M. Bakalowicz, D. Drew, F. Griger, J. Hartmann, G. Jiang, N. Moosdorf, A. Richts, Z. Stevanovic, G. Veni, et N. Goldscheider. The World Karst

- Aquifer Mapping project: concept, mapping procedure and map of Europe. *Hydrogeology Journal*, 25(3): 771–785, 2017. ISSN 14350157. doi: 10.1007/s10040-016-1519-3. (Cited p. 8, 9)
- J.-P. Chilès et P. Delfiner. *Geostatistics: Modeling Spatial Uncertainty*. John Wiley & Sons, 2009. (Cited p. 87)
- P. Collon, D. Bernasconi, C. Vuilleumier, et P. Renard. Statistical metrics for the characterization of karst network geometry and topology. *Geomorphology*, 283: 122–142, 2017. ISSN 0169555X. doi: <https://doi.org/10.1016/j.geomorph.2017.01.034>. (Cited p. 50, 51, 52, 53, 54, 55, 56, 58, 136)
- P. Collon-Drouaillet, V. Henrion, et J. Pellerin. An algorithm for 3D simulation of branchwork karst networks using Horton parameters and A* Application to a synthetic case. *Geological Society, London, Special Publications*, 370(1): 295–306, 2012. ISSN 0305-8719. doi: 10.1144/SP370.3. (Cited p. 2, 3, 40, 42, 44, 45, 46, 110)
- W. J. Conover. *Practical Nonparametric Statistics*. John Wiley & Sons, third edition, 1999. (Cited p. 60)
- R. L. Curl. Caves as a measure of karst. *The Journal of Geology*, 74(5): 798–830, 1966. doi: <https://doi.org/10.1086/627212>. (Cited p. 50)
- C. de Fouquet. *La modélisation géostatistique des milieux anthropisés*. PhD thesis, Université Pierre et Marie Curie, 2006. (Cited p. 59)
- M. Delle Rose, A. Federico, et M. Parise. Sinkhole genesis and evolution in Apulia, and their interrelations with the anthropogenic environment. *Natural Hazards and Earth System Science*, 4: 747–755, 2004. ISSN 1684-9981. doi: 10.5194/nhess-4-747-2004. (Cited p. 11)
- C. V. Deutsch et A. G. Journel. *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, 1998. (Cited p. 41, 87)
- L. Dewaide. *Hydrodynamics of solute transport in karst conduits with contrasted geometries: Karstogenesis description and tracer test modelling in the Waulsortian mudmounds (Furfooz, Belgium)*. PhD thesis, 2018. (Cited p. 31)
- E. W. Dijkstra. A Note on Two Problems in Connexion with Graphs. *Numerische Mathematik*, 1(1): 269–271, 1959. (Cited p. 59)
- S. J. Dreiss. Regional scale transport in a Karst Aquifer 2. Linear systems and time moment analysis. *Water Resources Research*, 25(1): 126–134, 1989. (Cited p. 36)
- W. Dreybrodt. The role of dissolution kinetics in the development of karst aquifers in limestone: A model simulation of karst evolution. *The Journal of Geology*, 98(5): 639–655, 1990. doi: 10.1086/629431. (Cited p. 18)
- W. Dreybrodt, F. Gabrovsek, et D. Romanov. *Processes of Speleogenesis: A Modeling Approach*. Založba ZRC, 2005. (Cited p. 2, 40)

- C. Dubois, Y. Quinif, J.-M. Baele, L. Barriquand, A. Bini, L. Bruxelles, G. Dandurand, C. Havron, O. Kaufmann, B. Lans, R. Maire, J. Martin, J. Rodet, M. D. Rowberry, P. Tognini, et A. Vergari. The process of ghost-rock karstification and its role in the formation of cave systems. *Earth Science Reviews*, 131: 116–148, 2014. ISSN 0012-8252. doi: 10.1016/j.earscirev.2014.01.006. (Cited p. 16)
- H. R. Duchene, A. N. Palmer, M. V. Palmer, J. M. Queen, V. J. Polyak, D. D. Decker, C. A. Hill, M. Spilde, P. A. Burger, D. W. Kirkland, et P. Boston. *Hypogene Speleogenesis in the Guadalupe Mountains, New Mexico and Texas, USA*. Cave and Karst Systems of the World. Springer International Publishing, Cham, 2017. ISBN 978-3-319-53347-6. doi: 10.1007/978-3-319-53348-3. (Cited p. 81)
- A. Dufoyer, N. Massei, N. Lecoq, J.-C. Marechal, D. Thiery, D. Pennequin, et P.-Y. David. Links between karst hydrogeological properties and statistical characteristics of spring discharge time series : a theoretical study. *Environmental Earth Sciences*, 78(14): 1–13, 2019. ISSN 1866-6299. doi: 10.1007/s12665-019-8411-0. (Cited p. 38)
- X. Emery. Testing the correctness of the sequential algorithm for simulating Gaussian random fields. *Stochastic Environmental Research and Risk Assessment*, 18: 401–413, 2004. doi: <https://doi.org/10.1007/s00477-004-0211-7>. (Cited p. 105)
- A. S. Engel, L. A. Stern, et P. C. Bennett. Microbial contributions to cave formation: New insights into sulfuric acid speleogenesis. *Geology*, 32(5): 369–372, 2004. ISSN 00917613. doi: 10.1130/G20288.1. (Cited p. 18, 26)
- S. Erzeybek, S. Srinivasan, et X. Janson. Multiple-Point Statistics in a Non-gridded Domain : Application to Karst/Fracture Network Modeling. In *Geostatistics Oslo 2012*, p. 221–237. 2012. ISBN 9789400741539. doi: 10.1007/978-94-007-4153-9. (Cited p. 41)
- A. R. Farrant et P. L. Smart. Role of sediment in speleogenesis; sedimentation and paragenesis. *Geomorphology*, 134(1-2): 79–93, 2011. ISSN 0169-555X. doi: 10.1016/j.geomorph.2011.06.006. (Cited p. 24, 31)
- M. Filipponi, P.-Y. Jeannin, et L. Tacher. Evidence of inception horizons in karst conduit networks. *Geomorphology*, 106(1-2): 86–99, 2009. ISSN 0169-555X. doi: 10.1016/j.geomorph.2008.09.010. (Cited p. 18)
- M. Filipponi, P.-Y. Jeannin, et L. Tacher. Understanding cave genesis along favourable bedding planes. The role of the primary rock permeability. *Zeitschrift für Geomorphologie*, 54(2): 91–114, 2010. doi: 10.1127/0372-8854/2010/0054S2-0006. (Cited p. 18)
- Filipponi, Marco et P.-Y. Jeannin. What makes a bedding plane favourable to karstification ? – The role of the primary rock permeability. In *Proceeding of 4th European Speleological Congress*, p. 32–37, 2008. (Cited p. 18, 19)
- P. Fleury. *Sources sous-marines et aquifères côtiers méditerranéens : fonctionnement et caractérisation*. PhD thesis, 2005. (Cited p. 19)

- P. Fleury, V. Plagnes, et M. Bakalowicz. Modelling of the functioning of karst aquifers with a reservoir model : Application to Fontaine de Vaucluse (South of France). *Journal of Hydrology*, 345(1-2): 38–49, 2007. doi: 10.1016/j.jhydrol.2007.07.014. (Cited p. 36)
- D. Ford et P. Williams. *Karst Hydrogeology and Geomorphology*. Springer Netherlands, john wiley edition, 2007. ISBN 9780470849965. (Cited p. 1, 8, 10, 11, 14, 15, 19, 20, 21, 33, 34, 35)
- A. Fournillon. *Modélisation géologique 3D et hydrodynamique appliquées aux réservoirs carbonatés karstiques : caractérisation des ressources en eau souterraine de l'Unité du Beausset (SE France)*. PhD thesis, 2012. (Cited p. 2, 35)
- A. Fournillon, S. Abelard, S. Viseur, B. Arfib, et J. Borgomano. Characterization of karstic networks by automatic extraction of geometrical and topological parameters: comparison between observations and stochastic simulations. *Geological Society, London, Special Publications*, 370: 247–264, 2012. ISSN 0305-8719. doi: 10.1144/SP370.8. (Cited p. 2, 40, 41, 43, 44, 45, 50, 53, 110, 136)
- A. Francesconi, F. Bigoni, F. Galatà, G. Leoni, R. Berto, F. Contento, et M. Miarelli. Carbonate Reservoir Fields: karst features recognition and modeling. In *Eurokarst 2018*, p. 174, 2018. (Cited p. 33, 35)
- Y. Frantz, P. Collon, P. Renard, et S. Viseur. Statistical analysis and simulation of conduit dimensions within karstic systems. In *Annual Conference IAMG*, 2019. (Cited p. 5)
- Y. Frantz, P. Collon, P. Renard, et S. Viseur. Analysis and stochastic simulation of geometrical properties of conduits in karstic networks. *Geomorphology*, 377, 2021. ISSN 0169-555X. doi: 10.1016/j.geomorph.2020.107480. (Cited p. 5, 50, 86, 131)
- A. R. Freeze et J. A. Cherry. *Groundwater*. Prentice-Hall, 1979. ISBN 0133653129. (Cited p. 35)
- S. Ā. Frisia et A. Borsato. Karst. In *Developments in Sedimentology*, vol. 61, p. 269–318. 2010. doi: 10.1016/S0070-4571(09)06106-8. (Cited p. 15)
- A. Frumkin et I. Fischhendler. Morphometry and distribution of isolated caves as a guide for phreatic and confined paleohydrological conditions. *Geomorphology*, 67(3-4): 457–471, 2005. ISSN 0169555X. doi: 10.1016/j.geomorph.2004.11.009. (Cited p. 50, 53, 54)
- K. R. Gabriel et R. R. Sokal. A new statistical approach to geographic variation analysis. *Systematic Biology*, 18(3): 259–270, 1969. doi: 10.2307/2412323. (Cited p. 128, 129)
- F. Gabrovšek, P. Häuselmann, et P. Audra. 'Looping caves' versus 'water table caves' : The role of base-level changes and recharge variations in cave development. *Geomorphology*, 204: 683–691, 2014. doi: 10.1016/j.geomorph.2013.09.016. (Cited p. 40)

- S. Galdenzi, M. Cocchioni, L. Morichetti, V. Amici, et S. Scuri. Sulfidic ground-water chemistry in the Frasassi caves, Italy. *Journal of Cave and Karst Studies*, 70(2): 94–107, 2008. (Cited p. 15)
- S. J. Gale. The hydraulics of conduit flow in carbonate aquifers. *Journal of Hydrology*, 70: 309–327, 1984. doi: [https://doi.org/10.1016/0022-1694\(84\)90129-X](https://doi.org/10.1016/0022-1694(84)90129-X). (Cited p. 36)
- E. Galin, A. Peytavie, E. Guérin, et B. Benes. Authoring Hierarchical Road Networks. *Computer Graphics Forum*, 30(7), 2011. doi: 10.1111/j.1467-8659.2011.02055.x. (Cited p. 128, 129)
- I. Gams. Origin of the term "karst," and the transformation of the Classical Karst (kras). *Environmental Geology*, 21: 110–114, 1993. doi: 10.1007/BF00775293. (Cited p. 8)
- R. Ghasemizadeh, F. Hellweger, C. Butscher, I. Padilla, D. Vesper, M. Field, et A. Alshwabkeh. Review: Groundwater flow and transport modeling of karst aquifers, with particular reference to the North Coast Limestone aquifer system of Puerto Rico. *Hydrogeology Journal*, 20(8): 1441–1461, 2012. ISSN 1431-2174. doi: 10.1007/s10040-012-0897-4. (Cited p. 39)
- R. Ghobadi. *Geostatistical modelling and simulation of karst systems*. PhD thesis, University of Adelaide, 2016. (Cited p. 41)
- J. D. Gibbons et S. Chakraborti. *Nonparametric Statistical Inference*. Marcel Dekker, fourth edition, 2003. ISBN 0824740521. (Cited p. 58, 60)
- L. W. Gill, P. Schuler, L. Duran, P. Morrissey, et P. M. Johnston. An evaluation of semidistributed-pipe-network and distributed-finite-difference models to simulate karst systems. *Hydrogeology Journal*, 29: 259–279, 2020. doi: 10.1007/s10040-020-02241-8. (Cited p. 38)
- P. Goovaerts. *Geostatistics for Natural Resource Evaluation*. Oxford University Press, 1997. ISBN 0195115384. (Cited p. 41, 87)
- W. R. Halliday. Pseudokarst in the 21st century. *Journal of Cave and Karst Studies*, 69(1): 103–113, 2007. (Cited p. 8)
- A. Hartmann, N. Goldscheider, T. Wagener, J. Lange, et M. Weiler. Karst water resources in a changing world: Review of hydrological modeling approaches. *Review of Geophysics*, 52: 218–242, 2014. ISSN 0096-3941. doi: 10.1002/2013RG000443. (Cited p. 1, 8, 11, 20, 37, 39)
- P. Häuselmann, P.-Y. Jeannin, et M. Monbaron. Role of epiphreatic flow and soutirages in conduit morphogenesis : the Bärenschacht example (BE, Switzerland). *Zeitschrift für Geomorphologie*, 47(2): 171–190, 2003. doi: 10.1127/zfg/47/2003/171. (Cited p. 23)
- R. Heinsoo et A. Collins. *Dungeons & Dragons 4.0 : Outreterre*. Play Factory, © Wizards of the Coast, 2010. (Cited p. 1)

- M. Hendrick et P. Renard. Subnetworks of Percolation Backbones to Model Karst Systems Around Tulum, Mexico. *Frontiers in Physics*, 4: 1–9, 2016a. ISSN 2296-424X. doi: 10.3389/fphy.2016.00043. (Cited p. 2, 3, 40, 43, 44, 110)
- M. Hendrick et P. Renard. Fractal Dimension, Walk Dimension and Conductivity Exponent of Karst Networks around Tulum. *Frontiers in Physics*, 4: 1–8, 2016b. ISSN 2296-424X. doi: 10.3389/fphy.2016.00027. (Cited p. 54, 78)
- V. Henrion, J. Pellerin, et G. Caumon. A Stochastic Methodology for 3D Cave Systems Modeling. In *VIII International Geostatistics Congress*, 2008. (Cited p. 42)
- V. Henrion, G. Caumon, et N. Cherpeau. ODSIM: An Object-Distance Simulation Method for Conditioning Complex Natural Structures. *Mathematical Geosciences*, 42 (8): 911–924, 2010. ISSN 1874-8961. doi: 10.1007/s11004-010-9299-0. (Cited p. 46)
- F. Herman et J. Braun. Evolution of the glacial landscape of the Southern Alps of New Zealand : Insights from a glacial erosion model. *Journal of Geophysical Research*, 113, 2008. doi: 10.1029/2007JF000807. (Cited p. 13)
- M. Hollander et D. A. Wolfe. *Nonparametric Statistical Methods*. John Wiley & Sons, second edition, 1999. (Cited p. 58)
- C. Holler. Pseudokarst. In *Encyclopedia of Caves*, p. 836–849. Elsevier Inc., 3 edition, 2019. ISBN 9780128141243. doi: 10.1016/B978-0-12-814124-3.00101-1. (Cited p. 8)
- R. E. Horton. Erosional development of streams and their drainage basins, hydrophysical approach to quantitative morphology. *Bulletin of the Geological Society of America*, 56: 275–370, 1945. (Cited p. 42, 50)
- N. Houillon. *La dynamique du carbone inorganique dans le continuum sol-épikarst-cavité du site de la Grotte de Lascaux (Dordogne , France) : apports des monitorings hydrogéochimique et microclimatique continus pour l'étude de l'aérogologie et le développement d'une m.* PhD thesis, 2016. (Cited p. 15, 23)
- A. D. Howard. Quantitative measures of cave patterns. *Caves and karst*, 13(1): 1–7, 1971. (Cited p. 50)
- C. Hu, Y. Hao, T.-c. J. Yeh, B. Pang, et Z. Wu. Simulation of spring flows from a karst aquifer with an artificial neural network. *Hydrological Processes*, 22: 596–604, 2008. doi: 10.1002/hyp.6625. (Cited p. 37)
- P. W. Huntoon. Is It Appropriate to Apply Porous Media Groundwater Circulation Models to Karstic Aquifers? In *Groundwater Models for Resources Analysis and Management*, p. 339–358. Lewis Publishers, 1995. (Cited p. 17, 23)
- E. Husson. *Interaction géodynamique/karstification et modélisation géologique 3D des massifs carbonatés : Implication sur la distribution prévisionnelle de la karstification . Exemple des paléokarts crétacés à néogènes du Languedoc montpelliérain.* PhD thesis, 2014. (Cited p. 18, 20)

- O. Jaquet, P. Siegel, G. Klubertanz, et H. Benabderrhamane. Stochastic discrete model of karstic networks. *Advances in Water Resources*, 27(7): 751–760, 2004. ISSN 03091708. doi: 10.1016/j.advwatres.2004.03.007. (Cited p. 3, 40, 110)
- J. W. Jaromczyk et G. T. Toussaint. Relative neighborhood graphs and their relatives. In *Proceedings of IEEE 80*, p. 1502–1517, 1992. doi: 10.1109/5.163414. (Cited p. 128, 129)
- C. M. Jarque et A. K. Bera. A Test for Normality of Observations and Regression Residuals. *International Statistical Review*, 55(2): 163–172, 1987. doi: <https://doi.org/10.2307/1403192>. (Cited p. 60)
- P.-Y. Jeannin. *Structure et comportement hydraulique des aquifères karstiques*. PhD thesis, Université de Neuchâtel, 1996. (Cited p. 2, 35, 50, 53)
- P.-Y. Jeannin. Modeling flow in phreatic and epiphreatic karst conduits in the Hölloch cave (Muotatal, Switzerland). *Water Resources Research*, 37(2): 191–200, 2001. doi: 10.1029/2000WR900257. (Cited p. 38)
- P.-Y. Jeannin, A. Malard, D. Rickerl, et E. Weber. Assessing karst-hydraulic hazards in tunneling — the Brunnum spring system — Bernese Jura, Switzerland. *Environmental Earth Sciences*, 74(12): 7655–7670, 2015. ISSN 1866-6299. doi: <https://doi.org/10.1007/s12665-015-4655-5>. (Cited p. 38, 46, 55, 86)
- P.-Y. Jeannin, G. Artigue, C. Butscher, Y. Chang, J.-B. Charlier, L. Duran, L. Gill, A. Hartmann, A. Johannet, H. Jourde, A. Kavousi, T. Liesch, Y. Liu, M. Lüthi, A. Malard, N. Mazzilli, E. Pardo-igúzquiza, D. Thiéry, T. Reimann, P. Schuler, W. Thomas, et A. Wunsch. Karst modelling challenge 1: Results of hydrological modelling. *Journal of Hydrology*, 600, 2021. doi: 10.1016/j.jhydrol.2021.126508. (Cited p. 36, 37, 38)
- M. Jébrak et É. Marcoux. *Géologie des ressources minérales*. Québec : Ministère des ressources naturelles et de la faune, 2008. (Cited p. 2)
- D. B. Johnson. Efficient Algorithms for Shortest Paths in Sparse Networks. *Journal of the Association for Computing Machinery*, 24(1): 1–13, 1977. doi: <https://doi.org/10.1145/321992.321993>. (Cited p. 87)
- J. Jouvès. *Origine, caractérisation et distribution prédictive des structures karstiques*. PhD thesis, 2018. (Cited p. 2, 8, 15, 16, 24, 25, 29, 35, 36, 39, 40, 42)
- J. Jouvès, S. Viseur, B. Arfib, C. Baudement, H. Camus, P. Collon, et Y. Guglielmi. Speleogenesis, geometry, and topology of caves: A quantitative study of 3D karst conduits. *Geomorphology*, 298: 86–106, 2017. ISSN 0169555X. doi: 10.1016/j.geomorph.2017.09.019. (Cited p. 28, 29, 45, 50, 51, 52, 54, 55, 56, 57, 66, 75, 78, 110)
- P. Kambesis. The importance of cave exploration to scientific research. *Journal of Cave and Karst Studies*, 69(1): 46–58, 2007. (Cited p. 30)

- G. Kaufmann et J. Braun. Karst aquifer evolution in fractured, porous rocks. *Water Resources Research*, 36(6): 1381–1391, 2000. doi: <https://doi.org/10.1016/j.jhydrol.2016.10.049>. (Cited p. 2, 40, 45)
- R. Khaledialidusti et J. Kleppe. Significance of the Kinetics of Minerals in Reactive-Transport Problems. In *79th EAGE Conference and Exhibition*, 2017. doi: 10.3997/2214-4609.201701629. (Cited p. 14)
- L. Kiraly. Régularisation de l’Areuse (Jura Suisse) simulée par modèle mathématique. *Hydrogeology of karstic terrains*, 1(3): 94–99, 1984. (Cited p. 38)
- L. Kiraly. Modelling karst aquifers by the combined discrete channel and continuum approach. *Bulletin du Centre d’Hydrogéologie de l’Université de Neuchâtel*, 15: 77–98, 1998. (Cited p. 38)
- A. Klimchouk. Towards defining, delimiting and classifying epikarst : Its origin, processes and variants of geomorphic evolution. *Speleogenesis and Evolution of Karst Aquifers*, 2(1): 1–13, 2004. (Cited p. 22)
- A. Klimchouk. *Hypogene Speleogenesis: Hydrogeological and Morphogenetic Perspective*. National Cave and Karst Research Institute, Carlsbad, 2007. ISBN 9780979542206. (Cited p. 24)
- M. Knez. The influence of bedding-planes on the development of karst caves (a study of Velika Dolina at Skocjanske Jame Caves, Slovenia). *Carbonates and Evaporites*, 13(2): 121–131, 1998. ISSN 08912556. doi: 10.1007/BF03176585. (Cited p. 18, 30)
- C. E. Koltermann et S. M. Gorelick. Heterogeneity in Sedimentary Deposits: A Review of Structure-Imitating, Process-Imitating, and Descriptive Approaches. *Water Resources Research*, 32(9): 2617–2658, 1996. doi: 10.1029/96WR00025. (Cited p. 40)
- L. Kong A Siou, A. Johannet, V. Borrell, et S. Pistre. Complexity selection of a neural network model for karst flood forecasting : The case of the Lez Basin (southern France). *Journal of Hydrology*, 403(3-4): 367–380, 2011. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2011.04.015. (Cited p. 37)
- L. Kong-A-Siou, P. Fleury, A. Johannet, V. B. Estupina, S. Pistre, et N. Dörfli. Performance and complementarity of two systemic models (reservoir and neural networks) used to simulate spring discharge and piezometry for a karst aquifer. *Journal of Hydrology*, 519: 3178–3192, 2014. doi: 10.1016/j.jhydrol.2014.10.041. (Cited p. 37)
- J. Kordilla, M. Sauter, T. Reimann, et T. Geyer. Simulation of saturated and unsaturated flow in karst systems at catchment scale using a double continuum approach. *Hydrology and Earth System Sciences*, 16(10): 3909–3923, 2012. ISSN 10275606. doi: 10.5194/hess-16-3909-2012. (Cited p. 38)
- A. Kranjc. The Origin and evolution of the term “Karst”. *Procedia Social and Behavioral Sciences*, 19: 567–570, 2011. doi: 10.1016/j.sbspro.2011.05.170. (Cited p. 8)

- N. Kresic. Remote Sensing of Tectonic Fabric Controlling Groundwater Flow in Dinaric Karst. *Remote Sensing of Environment*, 53(2): 85–90, 1995. doi: [https://doi.org/10.1016/0034-4257\(95\)00042-Y](https://doi.org/10.1016/0034-4257(95)00042-Y). (Cited p. 30)
- W. H. Kruskal et W. A. Wallis. Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, 47(260): 583–621, 1952. doi: 10.1080/01621459.1952.10483441. (Cited p. 58)
- E. L. Kuniansky, K. J. Halford, et W. B. Shoemaker. Permeameter Data Verify New Turbulence Process for MODFLOW. *Ground Water*, 46(5): 768–771, 2008. doi: 10.1111/j.1745-6584.2008.00458.x. (Cited p. 38)
- R. Labourdette, I. Lascu, J. Mylroie, et M. Roth. Process-Like Modeling of Flank-Margin Caves: From Genesis to Burial Evolution. *Journal of Sedimentary Research*, 77(11): 965–979, 2007. doi: 10.2110/jsr.2007.086. (Cited p. 3, 40)
- D. Laurent, C. Durllet, G. Barré, P. Sorriaux, P. Audra, P. Cartigny, C. Carpentier, G. Paris, P. Collon, T. Rigaudier, J. Pironon, et E. C. Gaucher. Epigenic vs. hypogenic speleogenesis governed by H₂S/CO₂ hydrothermal input and Quaternary icefield dynamics (NE French Pyrenees). *Geomorphology*, 387, 2021. doi: 10.1016/j.geomorph.2021.107769. (Cited p. 19, 25)
- R. Liedl, M. Sauter, D. Hückinghaus, T. Clemens, et G. Teutsch. Simulation of the development of karst aquifers using a coupled continuum pipe flow model. *Water Resources Research*, 39(3): 1–11, 2003. doi: 10.1029/2001WR001206. (Cited p. 38)
- H. W. Lilliefors. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. *Journal of the American Statistical Association*, 62(318): 399–402, 1967. (Cited p. 60)
- H. W. Lilliefors. On the Kolmogorov-Smirnov Test for the Exponential Distribution with Mean Unknown. *Journal of the American Statistical Association*, 64(325): 387–389, 1969. doi: <http://dx.doi.org/10.1080/01621459.1969.10500983>. (Cited p. 60)
- M. Lin, H. C. Lucas, et G. Shmueli. Too Big to Fail: Large Samples and the p-Value Problem. *Information Systems Research*, 24(4): 1–12, 2013. doi: <http://dx.doi.org/10.1287/isre.2013.0480>. (Cited p. 82)
- A. Lindenmayer. Mathematical Models for Cellular Interactions in Development. I. Filaments with one-sided inputs. *Journal of Theoretical Biology*, 18(3): 280–299, 1968. doi: [https://doi.org/10.1016/0022-5193\(68\)90079-9](https://doi.org/10.1016/0022-5193(68)90079-9). (Cited p. 43)
- A. J. Long et R. G. Derickson. Linear systems analysis in a karst aquifer. *Journal of Hydrology*, 219: 206–217, 1999. doi: [https://doi.org/10.1016/S0022-1694\(99\)00058-X](https://doi.org/10.1016/S0022-1694(99)00058-X). (Cited p. 36)
- G. Lorette. *Fonctionnement et vulnérabilité d'un système karstique multicouche à partir d'une approche multi-traceurs et d'un suivi haute-résolution: application aux Sources du Toulon à Périgueux (Dordogne, France)*. PhD thesis, 2019. (Cited p. 9, 20, 32)

- R. G. Loucks. Paleocave Carbonate Reservoirs: Origins, Burial-Depth Modifications, Spatial Complexity, and Reservoir Implications. *AAPG Bulletin*, 83(11): 1795–1834, 1999. doi: 10.1306/E4FD426F-1732-11D7-8645000102C1865D. (Cited p. 53)
- D. Lowe. *The origin of limestone caverns : an inception horizon hypothesis*. PhD thesis, 1992. (Cited p. 18)
- D. Lowe et J. Gunn. Carbonate speleogenesis: An inception horizon hypothesis. *Acta Carsologica*, 26(2): 457–488, 1997. (Cited p. 18)
- J. Lundberg. Microsculpturing of Solutional Rocky Landforms. In *Treatise on Geomorphology*, vol. 6, p. 121–138. Elsevier, 2013. ISBN 9780123747396. doi: 10.1016/B978-0-12-374739-6.00129-9. (Cited p. 10)
- K. T. B. Macquarrie et E. A. Sudicky. On the Incorporation of Drains into Three-Dimensional Variably Saturated Groundwater Flow Models. *Water Resources Research*, 32(2): 477–482, 1996. doi: 10.1029/95WR03404. (Cited p. 38)
- R. Maire. Les karsts sous-glaciaires et leurs relations avec le karst profond. *Revue de géographie alpine*, 66(2): 139–148, 1978. doi: <https://doi.org/10.3406/rga.1978.2118>. (Cited p. 13)
- J.-L. Mallet. *Geomodeling*. Oxford University Press, 2002. ISBN 0195144600. (Cited p. 59)
- A. Mangin. *Contribution à l'étude hydrodynamique des aquifères karstiques*. PhD thesis, Université de Dijon, 1975. (Cited p. 22, 24, 32, 36)
- J. C. Maréchal, B. Ladouche, N. Dörfliger, et P. Lachassagne. Interpretation of pumping tests in a mixed flow karst system. *Water Resources Research*, 44(5): 1–18, 2008. ISSN 00431397. doi: 10.1029/2007WR006288. (Cited p. 38)
- G. Mariethoz, P. Renard, et J. Straubhaar. The Direct Sampling method to perform multiple-point geostatistical simulations. *Water Resources Research*, 46(11): 1–14, 2010. ISSN 00431397. doi: 10.1029/2008WR007621. (Cited p. 41)
- G. Marsaglia, W. W. Tsang, et J. Wang. Evaluating Kolmogorov's Distribution. *Journal of Statistical Software*, 8(18): 1–4, 2003. doi: 10.18637/jss.v008.i18. (Cited p. 58, 60)
- É.-A. Martel. Contamination, protection et amélioration des sources thermominérales. In *Congrès international des mines, de la métallurgie et de la géologie appliquée, Paris, 7e session*, p. 791–798, 1935. (Cited p. 24)
- F. J. Martínez-Moreno, J. Galindo-Zaldívar, A. Pedrera, T. Teixido, P. Ruano, J. A. Peña, L. González-Castillo, A. Ruiz-Constán, M. López-Chicano, et W. Martín-Rosales. Integrated geophysical methods for studying the karst system of Gruta de las Maravillas (Aracena, Southwest Spain). *Journal of Applied Geophysics*, 107: 149–162, 2014. ISSN 0926-9851. doi: 10.1016/j.jappgeo.2014.05.021. (Cited p. 32, 33)

- F. J. Massey. The Kolmogorov-Smirnov Test for Goodness of Fit. *Journal of the American Statistical Association*, 46(253): 68–78, 1951. doi: <https://doi.org/10.2307/2280095>. (Cited p. 58, 60)
- N. Mazzilli, V. Guinot, H. Jourde, N. Lecoq, D. Labat, B. Arfib, C. Baudement, C. Danquigny, L. D. Soglio, et D. Bertin. KarstMod : A modelling platform for rainfall - discharge analysis and modelling dedicated to karst systems. *Environmental Modelling & Software*, 122: 1–7, 2017. doi: [10.1016/j.envsoft.2017.03.015](https://doi.org/10.1016/j.envsoft.2017.03.015). (Cited p. 2, 36)
- R. McGrath, P. Styles, E. Thomas, et S. Neale. Integrated high-resolution geophysical investigations as potential tools for water resource investigations in karst terrain. *Environmental Geology*, 42(5): 552–557, 2002. doi: [10.1007/s00254-001-0519-2](https://doi.org/10.1007/s00254-001-0519-2). (Cited p. 33)
- MedWater. MedWater. (Cited p. 38)
- L. H. Miller. Table of Percentage Points of Kolmogorov Statistics. *Journal of the American Statistical Association*, 51(273): 111–121, 1956. doi: <https://doi.org/10.2307/2280807>. (Cited p. 58, 60)
- T. Mochales, A. M. Casas, E. L. Pueyo, O. Pueyo, M. T. Roman, A. Pocovi, M. A. Soriano, et D. Anson. Detection of underground cavities by combining gravity, magnetic and ground penetrating radar surveys: a case study from the Zaragoza area, NE Spain. *Environmental Geology*, 53(5): 1067–1077, 2008. doi: [10.1007/s00254-007-0733-7](https://doi.org/10.1007/s00254-007-0733-7). (Cited p. 33)
- P. Mourlanette, P. Anquez, G. Rongier, et P. Collon. Integration of Geological Knowledge in Karstic Network Simulation with L-Systems. *2016 RING Meeting*, 2016. (Cited p. 2, 41, 43)
- M. E. J. Newman. Assortative mixing in networks. *Physical Review Letters*, 89: 1–5, 2002. doi: [10.1103/PhysRevLett.89.208701](https://doi.org/10.1103/PhysRevLett.89.208701). (Cited p. 58, 59)
- R. Noldus et P. V. Mieghem. Assortativity in Complex Networks. *Journal of Complex Networks*, 3(4): 507–542, 2015. doi: <https://doi.org/10.1093/comnet/cnv005>. (Cited p. 58)
- D. E. Northup et K. H. L. Lavoie. Geomicrobiology of Caves: A Review. *Geomicrobiology Journal*, 18(3): 199–222, 2001. ISSN 0149-0451. doi: [10.1080/01490450152467750](https://doi.org/10.1080/01490450152467750). (Cited p. 18)
- R. A. L. Osborne. Palaeokarst deposits in caves: Examples from eastern Australia and Central Europe. *Acta Carsologica*, 46(1): 19–32, 2017. ISSN 15802612. doi: [10.3986/ac.v46i1.4637](https://doi.org/10.3986/ac.v46i1.4637). (Cited p. 12)
- A. N. Palmer. Origin and morphology of limestone caves. *Geological Society of America Bulletin*, 103: 1–21, 1991. ISSN 00167606. doi: [10.1130/0016-7606\(1991\)103<0001:OAMOLC>2.3.CO;2](https://doi.org/10.1130/0016-7606(1991)103<0001:OAMOLC>2.3.CO;2). (Cited p. 13, 14, 17, 23, 24, 25, 26, 28, 29, 30)

- A. N. Palmer. Speleogenesis in carbonate rocks. *Speleogenesis and Evolution of Karst Aquifers*, 1(1): 1–11, 2003. (Cited p. 17)
- A. N. Palmer. *Cave Geology*. Cave Books, 2007. (Cited p. 15)
- A. N. Palmer. The Mammoth Cave system, Kentucky, USA. *Boletín Geológico y Minero*, 127(1): 131–145, 2016. (Cited p. 12)
- E. Pardo-Iguzquiza, J. J. Durán-Valseo, et V. Rodríguez-Galiano. Morphometric analysis of three-dimensional networks of karst conduits. *Geomorphology*, 132(1-2): 17–28, 2011. ISSN 0169555X. doi: 10.1016/j.geomorph.2011.04.030. (Cited p. 31, 50, 54, 55, 59, 78, 82, 138)
- E. Pardo-Igúzquiza, P. A. Dowd, C. Xu, et J. J. Durán-Valseo. Stochastic simulation of karst conduit networks. *Advances in Water Resources*, 35: 141–150, 2012. ISSN 03091708. doi: 10.1016/j.advwatres.2011.09.014. (Cited p. 2, 3, 40, 43, 44, 45, 46, 110, 136)
- E. Pardo-Igúzquiza, P. A. Dowd, J. J. Durán, et P. Robledo-Ardila. A review of fractals in karst. *International Journal of Speleology*, 48(1): 11–20, 2019. doi: <https://doi.org/10.5038/1827-806X.48.1.2218>. (Cited p. 78)
- M. Parquer. *Modélisation rétro-chronologique de systèmes chenalisés méandriiformes à partir d'observations géologiques*. PhD thesis, 2018. (Cited p. 40)
- K. Pearson. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series 5*, p. 157–175, 1900. doi: <https://doi.org/10.1080/14786440009463897>. (Cited p. 60)
- E. W. Peterson et C. M. Wicks. Assessing the importance of conduit geometry and physical parameters in karst systems using the storm water management model (SWMM). *Journal of Hydrology*, 329(1-2): 294–305, 2006. doi: 10.1016/j.jhydrol.2006.02.017. (Cited p. 36, 37, 38, 45, 86)
- N. Peyraube, R. Lastennet, et A. Denis. Geochemical evolution of groundwater in the unsaturated zone of a karstic massif, using the PC 2 — SIc relationship. *Journal of Hydrology*, 430-431: 13–24, 2012. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2012.01.033. (Cited p. 23)
- N. Peyraube, R. Lastennet, A. Denis, et P. Malaurent. Estimation of epikarst air PCO₂ using measurements of water δ^{13} CTDIC, cave air PCO₂ and δ^{13} CCO₂. *Geochimica et Cosmochimica Acta*, 118: 1–17, 2013. ISSN 0016-7037. doi: 10.1016/j.gca.2013.03.046. (Cited p. 23)
- L. Piccini. Recent developments on morphometric analysis of karst caves. *Acta Carsologica*, 40(1): 43–52, 2011. ISSN 0583-6050. doi: 10.3986/ac.v40i1.27. (Cited p. 54, 138)
- F. C. Pinto et C. Deutsch. Comparison of Alternative Gaussian Simulation Algorithms. In *CCG Annual Report 22*, 2020. (Cited p. 106, 135)

- P. Prusinkiewicz et A. Lindenmayer. *The algorithmic beauty of plants*. Springer-Verlag, 1990. (Cited p. 43)
- Y. Quinif. Fantômisation, cryptoaltération et altération sur roche nue, le triptyque de la karstification. In *Actes du colloque européen Karst 99*, p. 159–164, 1999. (Cited p. 15)
- Y. Quinif, J.-M. Baele, C. Dubois, C. Havron, O. Kaufmann, et A. Vergari. Fantômisation : un nouveau paradigme entre la théorie des deux phases de Davis et la théorie de la biorhexistase d’Erhart. *Geologica Belgica*, 17(1): 66–74, 2014. (Cited p. 16)
- T. Reimann et M. E. Hill. MODFLOW-CFP : A New Conduit Flow Process for MODFLOW-2005. *Ground Water*, 47(3): 321–325, 2009. doi: 10.1111/j.1745-6584.2009.00561.x. (Cited p. 38)
- T. Reimann, C. Rehrl, W. B. Shoemaker, T. Geyer, et S. Birk. The significance of turbulent flow representation in single-continuum models. *Water Resources Research*, 47: 1–15, 2011. doi: 10.1029/2010WR010133. (Cited p. 37)
- P. Renard, D. Bernasconi, A. Borghi, P. Collon-Drouaillet, et C. Vuilleumier. Karst network geometry and topology: automatic classification of network geometry. In *15th IAMG Conference*, 2013. (Cited p. 136)
- T. Robineau. *Caractérisation du fonctionnement hydrodynamique d’un aquifère karstique. Traitement du signal et modélisation double milieu des écoulements et du transport*. PhD thesis, 2019. (Cited p. 22)
- D. Romanov, W. Dreybrodt, et F. Gabrovsek. Interaction of Fracture and Conduit Flow in the Evolution of Karst Aquifers. *Speleogenesis and Evolution of Karst Aquifers*, 1(3): 2–7, 2002. (Cited p. 18)
- M. J. Ronayne. Influence of conduit network geometry on solute transport in karst aquifers with a permeable matrix. *Advances in Water Resources*, 56: 27–34, 2013. ISSN 0309-1708. doi: 10.1016/j.advwatres.2013.03.002. (Cited p. 46, 54, 86)
- G. Rongier, P. Collon-Drouaillet, et M. Filipponi. Simulation of 3D karst conduits with an object-distance based method integrating geological knowledge. *Geomorphology*, 217: 152–164, 2014. ISSN 0169555X. doi: 10.1016/j.geomorph.2014.04.024. (Cited p. 46)
- G. Rongier, P. Collon, P. Renard, J. Straubhaar, et J. Sausse. Comparing connected structures in ensemble of random fields. *Advances in Water Resources*, 96: 145–169, 2016. ISSN 03091708. doi: 10.1016/j.advwatres.2016.07.008. (Cited p. 42, 136)
- G. Rongier, P. Collon, et P. Renard. Stochastic simulation of channelized sedimentary bodies using a constrained L-system. *Computers & Geosciences*, 105: 158–168, 2017. doi: 10.1016/j.cageo.2017.05.006. (Cited p. 40)
- L. A. Rossman. *Storm Water Management Model User’s Manual Version 5.1*. Number September. 2015. (Cited p. 38)

- B. Sadier. *3D et géomorphologie karstique : La grotte Chauvet et les cavités des Gorges de l'Ardèche*. PhD thesis, 2013. (Cited p. 30, 51)
- J.-N. Salomon. *Précis de Karstologie*. Presses Universitaires de Bordeaux, 2000. doi: <https://doi.org/10.4000/rge.3952>. (Cited p. 13)
- J.-N. Salomon et M. Pulina. *Les karsts des régions climatiques extrêmes*. Presses Universitaires de Bordeaux, Pessac, 2005. (Cited p. 12, 13, 15)
- S. M. Sarbu et C. Lascu. Condensation Corrosion in Movile Cave, Romania. *Journal of Cave and Karst Studies*, 59(3): 99–102, 1997. (Cited p. 27)
- B. R. Scanlon, R. E. Mace, M. E. Barrett, et B. Smith. Can we simulate regional groundwater flow in a karst system using equivalent porous media models? Case study , Barton Springs Edwards aquifer, USA. *Journal of Hydrology*, 276(1-4): 137–158, 2003. doi: 10.1016/S0022-1694(03)00064-7. (Cited p. 2, 36, 38)
- A. Schiller et P. Renard. An optical laser device for mapping 3D geometry of underwater karst structures : first tests in the Ox Bel ' Ha system, Yucatan, Mexico. *Boletín Geológico y Minero*, 127(1): 99–110, 2016. (Cited p. 30, 46, 52)
- Schlumberger. Schlumberger market analysis. 2007. (Cited p. 2)
- C. Schmelzbach, C. Jordi, D. Sollberger, J. Doetsch, M. Kaufmann, W. Y. Meijer, E. Manukyan, J. Robertsson, H. Maurer, S. Greenhalgh, et H. Horstmeyer. Understanding the Impact of Karst on Seismic Wave Propagation - A Multi-method Geophysical Study. In *EAGE 2015*, 2015. (Cited p. 33)
- V. Sivelle. *Couplage d'approches conceptuelles, systémiques et distribuées pour l'interprétation de traçages artificiels en domaine karstique. Implications pour la détermination de la structure interne des aquifères karstiques*. PhD thesis, 2019. (Cited p. 23)
- R. Ø. Skoglund, S.-E. Lauritzen, et F. Gabrovšek. The impact of glacier ice-contact and subglacial hydrochemistry on evolution of maze caves : A modelling approach. *Journal of Hydrology*, 388(1): 157–172, 2010. ISSN 0022-1694. doi: 10.1016/j.jhydrol.2010.04.037. (Cited p. 13, 30)
- J.-M. Socquet. *Analyse des eaux thermales d'Aix (en Savoie), département du Mont-Blanc*. Cleaz, Chambéry, 1803. (Cited p. 24)
- J. Straubhaar, P. Renard, G. Mariethoz, R. Froidevaux, et O. Besson. An Improved Parallel Multiple-point Algorithm Using a List Approach. *Mathematical Geosciences*, 43(3): 305–328, 2011. ISSN 1874-8961. doi: 10.1007/s11004-011-9328-7. (Cited p. 41)
- F. Sumanovac et M. Weisser. Evaluation of resistivity and seismic methods for hydrogeological mapping in karst terrains. *Journal of Applied Geophysics*, 47(1): 13–28, 2001. doi: [https://doi.org/10.1016/S0926-9851\(01\)00044-1](https://doi.org/10.1016/S0926-9851(01)00044-1). (Cited p. 33)

- V. T. Tam, F. De Smedt, O. Batelaan, et A. Dassargues. Study on the relationship between lineaments and borehole specific capacity in a fractured and karstified limestone area in Vietnam. *Hydrogeology Journal*, 12: 662–673, 2004. doi: 10.1007/s10040-004-0329-1. (Cited p. 30)
- R. Tarjan. Depth-first search and linear graph algorithms. *SIAM Journal on Computing*, 1(2): 146–160, 1972. doi: <https://doi.org/10.1137/0201010>. (Cited p. 119)
- M. Temovski, P. Audra, A. Mihevc, J. Spangenberg, V. Polyak, W. McIntosh, et J.-Y. Bigot. Hypogenic origin of Provalata Cave, Republic of Macedonia: a distinct case of successive thermal carbonic and sulfuric acid speleogenesis. *International Journal of Speleology*, 42(3): 235–246, 2013. ISSN 0392-6672. doi: 10.5038/1827-806X.42.3.7. (Cited p. 25)
- G. Teutsch. An extended double-porosity concept as a practical modelling approach for a karstified terrain. *Hydrological Processes in Karst Terranes*, 207: 281–292, 1993. (Cited p. 2, 38)
- J. Thraillkill. Pipe Flow Models of a Kentucky Limestone Aquifer. *Ground Water*, 12(4): 202–205, 1974. doi: <https://doi.org/10.1111/j.1745-6584.1974.tb03023.x>. (Cited p. 38)
- F. Tian, X. Lu, S. Zheng, H. Zhang, Y. Rong, et D. Yang. Structure and Filling Characteristics of Paleokarst Reservoirs in the Northern Tarim Basin , Revealed by Outcrop , Core and Borehole Images. *Open Geosciences*, 9(1): 266–280, 2017. doi: 10.1515/geo-2017-0022. (Cited p. 33, 34)
- A.-j. Tinet, P. Collon, C. Philippe, L. Dewaide, et V. Hallet. OM-MADE : An open-source program to simulate one-dimensional solute transport in multiple exchanging conduits and storage zones. *Computers and Geosciences*, 127: 23–35, 2019. ISSN 0098-3004. doi: 10.1016/j.cageo.2019.03.001. (Cited p. 31)
- R. Valois, L. Bermejo, R. Guérin, S. Hinguant, R. Pigeaud, et J. Rodet. Karstic Morphologies Identified with Geophysics around Saulges Caves (Mayenne, France). *Archaeological Prospection*, 17(3): 151–160, 2010. doi: 10.1002/arp.385. (Cited p. 33)
- R. Valois, C. Camerlynck, A. Dhemaied, R. Guerin, G. Hovhannissian, V. Plagnes, F. Rejiba, et H. Robain. Assessment of doline geometry using geophysics on the Quercy plateau karst (South France). *Earth Surface process and landforms*, 36: 1183–1192, 2011. doi: 10.1002/esp.2144. (Cited p. 33)
- J. M. Ver Hoef, E. Peterson, et D. Theobald. Spatial statistical models that use flow and stream distance. *Environmental and Ecological Statistics*, 13: 449–464, 2006. doi: 10.1007/s10651-006-0022-8. (Cited p. 59)
- A. Vergari et Y. Quinif. Les paléokarsts du Hainaut (Belgique). *Geodinamica Acta*, 10(4): 175–187, 1997. doi: 10.1080/09853111.1997.11105301. (Cited p. 16)
- J. Verne. *Voyage au centre de la Terre*. Pierre-Jules Hetzel, 1864. (Cited p. 1)

- S. Viseur. Caractérisation de réservoirs turbiditiques : simulations stochastiques basées-objet de chenaux méandriformes. *Bulletin De La Societe Geologique De France*, 175 (1): 11–20, 2004. (Cited p. 40)
- S. Viseur, J. Jouves, A. Fournillon, B. Arfib, et Y. Guglielmi. 3D stochastic simulation of caves : application to Saint-Sébastien case study (SE , France). *Karstologia*, 64: 17–24, 2014. (Cited p. 2, 40, 41, 43, 45, 110)
- C. Vuilleumier, A. Borghi, P. Renard, D. Ottowitz, A. Schiller, R. Supper, et F. Cornaton. A method for the stochastic modeling of karstic systems accounting for geophysical data : an example of application in the region of Tulum , Yucatan Peninsula (Mexico). *Hydrogeology Journal*, 21(3): 529–544, 2013. doi: 10.1007/s10040-012-0944-1. (Cited p. 46, 55, 61, 86)
- S. Vyzhva, V. Onyshchuk, I. Onyshchuk, M. Reva, et O. Shabatura. Application of geophysical methods in the study of karst. In *EAGE*, 2020. (Cited p. 32, 33, 35)
- A. C. Waltham et P. G. Fookes. Engineering classification of karst ground conditions. *Speleogenesis and Evolution of Karst Aquifers*, 3(1): 1–20, 2005. (Cited p. 11)
- J. E. Warren et P. J. Root. The Behavior of Naturally Fractured Reservoirs. *Society of Petroleum Engineers Journal*, 3(3): 245–255, 1963. doi: <https://doi.org/10.2118/426-PA>. (Cited p. 38)
- F. Wilcoxon. Individual Comparisons by Ranking Methods. *Biometrics Bulletin*, 1(6): 80–83, 1945. doi: <https://doi.org/10.2307/3001968>. (Cited p. 58)
- S. R. H. Worthington. A comprehensive strategy for understanding flow in carbonate aquifers. In *Karst Modeling*, p. 30–37. 1999. (Cited p. 34)
- S. R. H. Worthington et D. C. Ford. Self-Organized Permeability in Carbonate Aquifers. *Ground Water*, 47(3): 326–336, 2009. doi: 10.1111/j.1745-6584.2009.00551.x. (Cited p. 17, 23)
- S. R. H. Worthington, D. C. Ford, et P. A. Beddows. Porosity and permeability enhancement in unconfined carbonate aquifers as a result of solution. In *Speleogenesis: evolution of karst aquifers*, p. 463–472. National Speleological Society, 2000. (Cited p. 12, 22)
- D. F. Young, B. R. Munson, T. H. Okiishi, et W. W. Huebsch. *Brief Introduction To Fluid Mechanics*. John Wiley & Sons, 5th edition, 2010. (Cited p. 37)
- Y.-K. Zhang et E.-w. Bai. Simulation of spring discharge from a limestone aquifer in northeastern Iowa with an identification scheme. *Hydrogeology Journal*, 4(4): 41–54, 1996. (Cited p. 36)

Simulation stochastique des réseaux karstiques

Résumé : Difficiles d'accès, les réseaux karstiques ne sont généralement que partiellement décrits. Les techniques d'exploration classiques produisent par ailleurs des données éparées et sujettes à des incertitudes concernant la position des conduits et leurs dimensions, paramètres essentiels aux simulations d'écoulement. Les simulations stochastiques permettent de mieux gérer et évaluer ces incertitudes en proposant plusieurs représentations équiprobables de réseaux karstiques. Le simulateur idéal devrait permettre de construire des réseaux de drains karstiques tridimensionnels, honorant les observations de terrain (indices de karstification), les connaissances apportées par les essais de traçage, les informations recueillies sur des parties accessibles du réseau et celles apportées par l'étude d'autres réseaux. C'est dans ce contexte que s'inscrit cette thèse qui propose 3 principales contributions.

La première contribution est l'analyse statistique d'une base de données de 49 réseaux karstiques. Elle se focalise sur l'étude de la géométrie des conduits, à travers l'analyse de deux métriques : le rayon équivalent et le ratio largeur-hauteur. Aucune loi statistique générique qui décrirait la géométrie des conduits n'a pu être mise en évidence. Néanmoins, la variabilité spatiale des propriétés géométriques à différentes échelles a été caractérisée, notamment grâce au développement de variogrammes 1D-curvilinéaires. L'hypothèse répandue d'une organisation hiérarchique des géométries de conduits a également été

analysée et rejetée.

La deuxième contribution est le développement de deux méthodes de simulation stochastique de propriétés le long de réseaux karstiques s'appuyant sur les résultats de l'analyse statistique. La première méthode se focalise sur la reproduction de la variabilité des propriétés à l'échelle du réseau, tandis que la deuxième se concentre sur la reproduction de la variabilité au sein et entre les branches du réseau. Elles sont toutes deux basées sur les méthodes de Simulations Gaussiennes Séquentielles (SGS), ici adaptées à des objets 1D-curvilinéaires.

La troisième contribution est le prototype d'une méthode ayant pour but de simuler stochastiquement des réseaux karstiques discrets sous forme de graphes (qualifiés de *squelettes des réseaux*). Il est espéré qu'une fois entièrement développée, elle permette de simuler différents types de réseaux et de prendre directement en compte les données de terrain ainsi que les informations géologiques. Elle consiste en trois grandes étapes : i) la génération d'un nuage de points, ii) le calcul de leur connectivité et iii) leur connexion pour former le squelette du réseau karstique.

Ces contributions ouvrent de nouvelles perspectives quant à la simulation de réseaux de drains karstiques utilisables pour la simulation d'écoulement (e.g., SWMM, Epanet, Modflow-CFP), ce qui devrait permettre une meilleure caractérisation des flux en leur sein.

Mots-clés : karst, réseaux karstiques, analyse statistique, géométrie de conduits, simulation stochastique, graphe

Stochastic simulation of karstic networks

Abstract : Despite intensive explorations by speleologists, karstic networks remain only partially described as many conduits are not accessible to humans. The classical exploration techniques produce sparse data subject to uncertainties concerning the conduit position and their dimensions, which are essential parameters for flow simulations. Stochastic simulations make it possible to better handle and assess these uncertainties by offering several equally probable karstic network representations. The ideal simulator should allow for the construction of tridimensional karstic drain networks, respecting the field observations (karstification markers), the knowledge brought by tracer tests, the information collected in the accessible parts of the network and those obtained by the study of other networks. In this context, this PhD thesis offers 3 main contributions.

The first contribution is the statistical analysis of a database of 49 karstic networks. It focuses on the study of conduit geometry, through the analysis of two metrics : the equivalent radius and the width-height ratio. No generic statistical law describing the network geometry was found. Nonetheless, the spatial variability of the geometrical properties at different scales was characterized, mostly through the development of 1D-curvilinear variograms. The widespread hypothesis of a hierarchical organization of the conduit ge-

ometries has also been analysed and rejected.

The second contribution is the development of two methods allowing stochastic simulations of properties along karstic networks and based on the results of the statistical analysis. The first method focuses on the reproduction of the property variability at the network scale, while the second one focuses on the reproduction of the variability within and between the network branches. Both are based on the Sequential Gaussian Simulation methods and are adapted to 1D-curvilinear objects.

The third contribution is the prototype of a method aiming to stochastically simulate discrete karstic networks as graphs (known as *network skeletons*). We hope that once completed, it would allow the simulation of different network types, while taking directly into account field data and geological information. It is divided in three main steps : i) the generation of a point cloud, ii) the computation of the point connectivity and iii) their connection to create the skeleton of the karstic network.

These contributions open new prospects regarding the simulation of karstic drain networks usable for flow simulation (e.g., SWMM, Epanet, Modflow-CFP), which should allow a better characterization of the associated flows.

Keywords : karst, karstic networks, statistical analysis, conduit geometry, stochastic simulation, graph