3D stochastic modeling of karst aquifers using a pseudo-genetic methodology

thesis presented at the

Faculty of Sciences

Institute of Hydrogeology and Geothermics (CHYN)

University of Neuchâtel, Switzerland

for the degree of

Doctor of natural sciences

presented by

Andrea Borghi

accepted on the recommendation of

Prof. Philippe Renard
Prof. Guillaume Caumon
Prof. Philip Brunner
Prof. Fabien Cornaton
Dr Pierre-Yves Jeannin
Arnaud Levannier

Defended on February 22, 2013

unine
UNIVERSITÉ DE NEUCHÂTEL

Faculté des sciences
Secrétariat-décanat de Faculté
Rue Emile-Argand 11
2000 Neuchâtel - Suisse
Tél: + 41 (0)32 718 2100
E-mail: secretariat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Monsieur Andrea BORGHI

Titre: 3D stochastic modeling of karst aquifers using a pseudo-genetic methodology

sur le rapport des membres du jury:

- Prof. Philippe Renard, Université de Neuchâtel, directeur de thèse
- Prof. Philip Brunner, Université de Neuchâtel
- Prof. Guillaume Gaumon, Université de Lorraine, France
- Prof. Dr Fabien Cornaton, WASY-DHI, Berlin, Allemagne
- Dr Pierre-Yves Jeannin, ISSKA, La Chaux-de-Fonds, Suisse

Neuchâtel, le 9 avril 2013

Le Doyen, Prof. P. Kropf

Tous les compléments de temps, sans préposition, qui indiquent la durée et la répétition, sont à l'accusatif. Merci à feu Prof. Eric Roth, pour cette règle de grammaire Allemande que je n'oublierai probablement jamais, ainsi que pour le reste.

Contents

Acknowledgments

Français

Je tiens à remercier tout d'abord Prof. Philippe Renard, mon directeur de thèse. Je n'aurais jamais fait une thèse sans lui, j'en suis bien convaincu. Durant mon travail de master, il a su, par son intelligence, sa curiosité et sa sympathie, allumer la flamme de la recherche en moi. C'est aussi lui qui avait déjà commencé à discuter avec SWS d'une thèse sur le karst, bien avant que je n'envisage la voie accadémique, et qui a accepté de me prendre sous son aile, à la foi carburant (par ses idées) et gouvernail de ma recherche. Son amabilité est à la fois sa plus grande qualité et sa malédiction; j'espère qu'il saura conserver cette qualité tout en réussissant à trouver des espaces pour poursuivre ses propres recherches, ou coder un truc génial par-ci par là. J'esprère aussi qu'il réussira à profiter comme il se doit du bâteau à voile dont il vient de devenir l'heureux propriétaire. Bon vent mon capitaine ;)

Je tiens ensuite à remercier Schlumberger Water Services (SWS) pour avoir financé cette thèse, et en particulier tous ses membres qui m'ont suivi durant ces années. En tout premier je voudrais remercier Jean-Pierre Delhomme, qui a été intéressé par ce sujet dès le début, et qui à poussé pour que cette thèse voie le jour. Je souhaite ensuite remercier Sandra Jenni, pour toute son énergie, et pour les nombreuses suggestions au début de ma thèse. La fatalité à voulu que le groupe de SWS de Paris soit fermé, et que Sandra suive un autre chemin. Dans cette période de "battement", j'ai été suivi par Gilles Mathieu, que je remercie vivement pour son amabilité, et pour avoir trouvé une solution qui me permette de terminer ma thèse cette quatrième année avec un financement de SWS Pérou. Mon dernier remerciement du côté de SWS va à Arnaud Levannier, mon dernier superviseur, qui a su valoriser grandment mon travail, en particulier le simulateur de conduits karstiques, après la réalisation de la thèse.

Merci à Fabien Cornaton, pour les innombrables heures qu'il a passées à m'aider et pour sa disponibilité à me laisser utiliser son code GROUNDWATER. Je n'oublierai jamais la première fois que, jeune étudiant de master, j'ai ouvert la porte de son bureau pour lui poser une question, alors qu'il n'était pas forcément de sa meilleure humeur... J'ai depuis toujours conservé une forme de timidité subsonsciente quand il s'agissait d'aller l'embêter (de personne ou par e-mail) pour lui demander conseil.

Merci à Jahouer Kerrou pour m'avoir souvent passée des conseils bien utiles et sa disponibilité. Un grand merci aussi au Prof. Pierre Perrochet pour tous les conseils hydrodynamiques, et au Prof. Philip Brunner, pour les discussions sur le problème inverse et pour ses innombrables relectures d'articles et morceaux de ma thèse.

Je remercie aussi tous les autres collaborateurs du groupe stochastique. A Julien Straubhaar, en particulier, un grand merci pour tous les conseils algorithmiques et sa patience. Sans lui mon apprentissages de certains méandres obscurs de la simulation et de la programmation auraient sans doute été beaucoup plus difficles.

Je remercie de tout coeur Damian Glenz, collègue et collocataire, pour les innombrables heures à discuter nos difficultés réciproques, et à trouver ensemble des solutions (parfois peut-être pas totalement scientifiques). Je le remercie également pour toute son amitié, pour les longues nuits insomniaques passées ensemble à l'unimail, et pour avoir toujours su garder le sourire, en supportant mes excès d'humeur.

Je souhaite aussi remercier Pierre-Yves Jeannin, directeur de l'ISSKA, pour m'avoir accueilli dans son institut à plusieurs reprises, ainsi que pour tous les conseils et critiques constructives qui ont grandement influencé ma thèse. Merci aussi à tous les collaborateurs de l'ISSKA pour leur amabilité et leur amitié.

Merci aussi au premier rapporteur de ma thèse, Prof. Guillaume Caumon, pour ses commentaires et pour m'avoir offert un postdoc dans son institut, merci aussi à Pauline Collon-Drouaillet pour être sur la même longueur d'onde que moi. J'espère que notre collaboration future nous sera profitable à tous.

Merci aux 4 étudiants de master qui ont eu la patience de me supporter en tant que superviseur: Guillaume de Scheppers, Cybèle Cholet, Cécile Vuillleumier, et Louis Fournier. Ces 4 étudiants ont été les premiers (heureux) utilisateurs de mon code, et ont tous beaucoup contribué à l'améliorer. Je les remercie beaucoup pour leur sympathie, leur sourire, et leur imprudence en m'acceptant en tant que superviseur. J'espère sincèrement qu'ils aient appris de moi autant que j'ai appris d'eux.

A Guillaume Pirot, qui sera pour moi toujours *Guillaume Cravatte*, un grand merci pour m'avoir supporté ces deux dernières années dans le bureau, et avoir su garder le sourire même en ayant pour voisin un ours brun comme moi.

Une thèse dans une université est une expérience scientifique et humaine avant tout. Je remercie sincèrement tous les collègues et étudiants, pour leur amitié et leur présence, pour les pauses café, pour tous les bon mements passés ensemble. Je garderai toujours un bon souvenir des soupers et apéros au bord du lac, et j'en profite pour remercier de leur amitié bon nombre des étudiants avec qui j'ai eu l'occasion de trinquer (alors que ma thèse ne m'avait pas encore totalement fagocité). Merci pour son amitié et sa bonne humeur à Mehdi Al-Zoubaidy, dont le charisme et la sympathie sont indiscutables, et qui a habité dans notre salon pendant les quelques mois de fin de son master (Il pourra d'ailleur passer récupérer sa brosse à dents un jour). Merci au groupe des "pintades", à savoir Lorienne Thueler, Anaïs Eschbach, Sabrina Gonseth, et Chantal Crisinel, pour leur sympathie, les apéros, les sorties, et les pauses. Vous m'avez (sauf Lorienne bien sûr) bien manqué ces 2 dernières années de thèse.

Merci a François *"Paco"* Negro, dernière résistance de la géologie Neuchâteloise, retranché dans son bureau du 3è étage. Il a toujours su m'illuminer sur mes questions de géologie strucurale. Merci aussi pour toutes les discussions un peu *geek* sur les éditeurs géologiques, en particulier Geomodeller3D®. Qu'il soit aussi remercié pour son aptitude à animer les soirées des assitants, il restera dans ma mémoire, comme celui qui personifie *l'âme de la fête.*

Merci aussi à nos responsables informatiques: Gianni Gasparaotto, Romain Bessire et Eric de Carvalho. Nous avons toujours eu des discussions enrichissantes, et j'espère avoir un peu contribué à présever leur nerfs durant certaines intéractions entre le CHYN et le SITEL. Ils m'ont toujours aidé à trouver des solutions lorsque des probèmes surgissaient, et toujours au plus vite.

Merci à Manu, notre concierge, hélas retraité. Si parfois il était un peu distrait avec le nettoyage, il nettoyait toujours nos humeurs avec grand soin. Sa retraite a porté un coup assez dur à mon humeur, qu'il savait toujours améliorer quand il passait dans le bureau. Merci, ô concierge *de l'humeur.*

Merci infiniment aussi à mes amis proches, à mon *clan* pour ainsi dire: Anka Aubaret, Jean-Marc Buntschu, Daniel Burns, Patrick Oldendorf, Giovanni Dalle Nogare, Thomas D'Eysmond et Matteo Guggiari pour leur amitié durable, pour la confiance que j'ai en eux, et pour le soutient qu'il m'ont toujours apporté.

J'ai bien peur qu'une page ne soit pas suffisante pour remercier aussi tous les autres. Je ne voudrais oublier personne dans ces remerciement, alors encore merci à tous pour ces belles années de collaboration et d'amitié.

English

I would like to particularly thank the group of SWS Perou, and SWS Chile, for having funded my last year of research, and for having allowed me to test my methodology on one of their projects. In particular thanks to John Rupp and Loïc Ragas who followed this project, for their disponibility and kindliness. I also want to thank Wayne Hesh from SWS-software in Waterloo, for his interest in my code and all his suggestions.

Italiano

Devo ringraziare innanzitutto i miei famigliari, tutti quanti, per avermi sopportato (in tutti i sensi della parola) durante questa fase alquanto intensa della mia vita. In particolar modo grazie ai miei genitori, Marco e Gibi, per la loro infinita (ma proprio infinita, davvero) pazienza e mia sorella Angela per lo stesso motivo. Grazie a mia cugina Ninna per tutte le volte che mi ha sfamato durante i primi periodi della tesi, e per il suo affetto.

Un grazie ENORME alla mia compagna, Nora, per tutta la pazienza e amore che è riuscita, e che riesce a darmi, anche nei miei periodi insopportabili e stressati. Spero di riuscire a darle tutta la felicità che lei dà a me.

Infine, grazie ad Alessandro Comunian, mio primo *coinquilino* di ufficio (se si può dire). I due anni in sua presenza sono stati veramente interessanti, e ha saputo trasmettermi un sacco di trucchi, la sua *nerdaggine* di Linux, e l'interesse per Python, per l'opensource, per la *purezza* del bash shell. Lo ringrazio anche per le innumerevoli serate all'uni, che sarebbero sembrate molto piu tetre senza la sua presenza.

Keywords

stochastic, karst modeling, inverse problem, flow and transport simulation, pseudo-genetic, Fast Marching Algorithm, fracture modeling, geological modeling

Parole chiavi

stocastico, modellizzazione di karst, problema inverso, simulazione di flussi e contaminanti, pseudo genetico, Fast Marching Algorithm, modellizzazione di fratturazione, modellizzazione geologica

Mots-clé

stochastique, modelisation de karst, probleme inverse, simulation d'écoulement et transport, pseudo-genetique, Fast Marching Algorithm, modelisation de fractures, modelisation géologique

Stichworte

stochastisch, Modellierung von Karstaquiferen, inversen Ansatz, pseudogenetisch, Fast Marching Algorithm, Kluftmodellierung, geologisches Modell

Abstract

English

The focus of this thesis is the development of a methodology to model karst aquifers. First the geometry of the karst conduits is simulated. Their geometry is controlled by the geology at large scale (geological model) and by the fracturation at smaller scale (stochastic model of fractures).

Secondly, these geometrical models (3D geological formation together with conduits) are used as base for flow and transport simulation.

Finally, the karst conduit generator called SKS (*Stochastic Karst Simulator*) is coupled with the physical simulation (flow and transport) to investigate an inverse approach which allows to use this methodology in an uncertainty analysis.

The karst conduits simulation methodology is called *pseudo genetic* because it mimic the results of the speleogenetic processes, without simulating all the complex kinetic of karst systems genesis, like reactive transport, calcite dissolution and precipitation. In this approach, the karst conduits are simulated by approaching the physical principle of minimization of energy using a Fast Marching Algorithm. This algorithm allow to compute the minimum effort path, which is assumed to be the one used by water, and consequently the preferential dissolution.

Italiano

Lo scopo di questa tesi, è lo svipuppo di una metodologia di modellizzazione realistica degli acquiferi carsici. Nella prima parte viene modellizzata la geometria dei condotti carsici. La geometria di questi ultimi, è controllata a larga scala dalla geologia (modello geologico) e a piu piccola scala dalla fratturazione (modello di fratturazione stocastico).

In seguito questi modelli geometrici vengono usati come base per la simulazione di fluidi e trasporto di contaminanti. E infine, il simulatore di condotti carsici SKS (*Stochastic Karst Simulator*) è usato insieme alla simulazione di fluidi e trasporto per valutare l'applicabilità di un approcio inverso che permetta di utilizzare questa metodologia di simulazione.

La metodologia di simulazione dei condotti carsici è detta "*pseudo genetica*", perchè tenta di approssimare la fisica complessa della speleogenesi (come il trasporto reattive, la dissoluzione/precipitazione della calcite) senza dover risolverla numericamente. In effetti si basa sul principio fisico della minimizzazione dell'energia, usando un algoritmo di Fast Marching per calcolare il cammino di minor resistenza (cioè quello che dovrebbe seguire l'acqua). In pratica questa metodologia simula dei sistemi carsici maturi, direttamente nel loro stato finale.

Français

Le but de cette thèse est le développement d'une méthodologie de modélisation des aquifères karstiques. Premièrement, la géométrie des conduits karstiques est simulée. La géométrie de ces conduits est contrôlée à large échelle par la géologie (modèle géologique), et à plus petite échelle par la fracturation (modèle stochastique de fracturation).

Deuxièmement, ces modèles géométriques (formations 3D et conduits ensemble) sont utilisés comme base pour la simulation d'écoulement et de transport.

En dernière partie, le simulateur de conduits SKS (*Stochastic Karst Simulator*) est couplé avec la simulation d'écoulement et transport pour investiguer une approche inverse qui permette d'utiliser cette méthodologie dans une étude d'incertitude.

La méthodologie de simulation des conduits karstiques développée ici se base sur une approche dite "*pseudo-génétique*", càd qui mime les résultats des processus de spéléogénèses, sans pour autant simuler toute la dynamique complexe de ces processus, comme la dissolution et le transport réactif de calcite, etc. Dans cette approche *pseudo-génétique* les conduits karstiques sont simulés par une physique approchée, qui se base sur le principe de minimisation de l'énergie. L'eau se déplace dans un milieu en cherchant le chemin de moindre résistance. Ce principe est utilisé ici, par l'utilisation d'un algorithme de Fast Marching, qui permet de calculer le chemin de moindre effort.

Deutsch

Das Ziel dieser Dissertation besteht in der Entwicklung einer Methode zur Modellierung von Karstaquiferen. In einem ersten Schritt wird die Geometrie des Karstsystems simuliert. Im grossen Massstab wird die Geometrie des Karstsystems durch die Geologie bestimmt (geologisches Modell), in kleinerem Massstab durch Klüfte (stochastisches Modell zur Kluftgenese). In einem zweiten Schritt wird dieses Modell als Grundlage fr die Modellierung von Grundwasserstrmung und Stoffransport im Karstsystem verwendet.

Schliesslich wird die Simulation des Karstsystems mittels des Karst-Simulator (SKS, *Stochastic Karst Simulator*) mit Simulationen von Grundwasserstrmung und Stoffransport gekoppelt, um einen inversen Ansatz zu untersuchen, der es erlauben wrde diese Methode im Rahmen einer Unsicherheitsanalyse zu verwenden.

Die hier entwickelte Methodik zur Simulierung von Karstaquiferen basiert auf einem sogenannt ”*pseudogenetischen*” Ansatz, weil sie speleogenetische Prozesse nachbildet, ohne die komplexe Dynamik dieser Prozesse, wie Auflösung und reaktiver Transport von Kalzit etc., im Detail zu simulieren. Die Karstgenese wird in dieser Methodik vielmehr durch einen Ansatz angenähert, der auf dem Prinzip der Energieminimierung beruht. Dabei wird ein *Fast Marching* Algorithmus verwendet, um zwischen den Stellen mit Wasserzuflüssen ins Karstsystem und den Karstquellen den Weg des geringsten Widerstandes zu berechnen.

Extended Abstract

Karst aquifers constitute an important source for drinking water, even though they are highly vulnerable to contamination. Protecting karst water resources is in many ways challenging. The dissolution of calcite by infiltrating water produces preferential flow paths (conduit networks) within these aquifers. A key difficulty in understanding and protecting karst aquifers is that the temporal dynamics of karst systems are often dominated by these large conduits. However, the location of and structure of these dominant features are typically unknown, undermining quantitative, predictive approaches are inevitable for a solid risk management, the knowledge gaps on karst systems constitute a major impediment for a reliable protection of these critical resources. Stochastic approaches have proven their efficiency in uncertainty estimation in many other fields of application.

The present thesis proposes a new methodology to model karst aquifers, in a stochastic framework. The resulting models are consistent with the actual knowledge about these aquifers and their particular genesis processes (speleogenesis).

The proposed methodology is divided in several steps: first a geological model is built, to model the large scale geometry of the aquifer. Secondly the initial heterogeneities (inception horizons) of the aquifer formation are modeled stochastically, because they influence the small scale geometry of the conduits network. Then the stochastic conduit networks are simulated using the physical assumption that the water follows a minimum effort path while minimizing the total energy of a system. Finally, these models are used to simulate flow and transport. They can then be used in an inverse framework to estimate the uncertainty.

To build the geological model we use an already existing geological editor: Geomodeller3D®. This geological editor uses an implicit approach. The output of this kind of tools, provides on the one hand the information about the 3D localization of karstifiable formations, and on the other hand information about their orientation. This last information is used to generate stochastic Discrete Fracture Networks (DFN) with the following approach: the theoretical families of fractures that are expected in a folded environment are modeled with consistency with the local geological orientation. The fractures (simple rectangular planar objects) are first modeled in an unfolded environment, and then placed in the geological model and rotated according to the local orientation of the geology. We use the gradient of the geological potential field of the implicit geological model to retrieve the information about the geological orientation, and then to build a new vectorial base, which will be used as rotation matrix. The fracture number and size are distributed according to a power law as defined in the Universal Fracture Model (UFM) for the dense regime.

The stochastic conduit networks are modeled using the *pseudo genetic* methodology developed in this thesis: after the geological model, and the initial heterogeneities are built, a study of the regional hydrology / hydrogeology is conducted to identify the potential inlets and outlets of the system, the base levels and the possibility of having different phases of karstification. The conduits are then generated in an iterative manner using a fast marching algorithm. A probabilistic model can be used to represent the knowledge available and the remaining uncertainty depending on the data at hand. The conduits are assumed to follow minimum effort paths in the heterogeneous medium from sinkholes or dolines toward springs. The search of the shortest path is performed using a fast marching algorithm. This process can be iterative, allowing to account for the presence of already simulated conduits and to produce a hierarchical network. The final result is a stochastic ensemble of 3D karst reservoir models that are all constrained by the regional geology, the local heterogeneities, and the regional hydrogeological/hydrological conditions. Several levels of uncertainty can be considered (large scale geological structures, local heterogeneity, position of possible inlets and outlets, phases of karstification). Compared to other techniques, this method is fast, accounts for the main factors controlling the 3D geometry of the network, and can be conditioned to available field observations.

These networks are then used to simulate flow and transport with the finite element code GROUNDWATER. The conduits are meshed as pipe element that follow the edges of the 3D elements of the matrix. In the conduits, turbulent flow is simulated using the Manning-Strickler equations. The radius of the conduits is computed as a function of their hierarchical order. A lumped parameter model is used to model the recharge function. The flow boundary conditions that are used are: prescribed head $[m]$ boundary conditions for the springs, prescribed nodal flux $[m^3/s]$ boundary conditions for the point inlets of the aquifer (dolines, sinkholes), and prescribed flux at element faces $[m/s]$ boundary conditions or a source term $[s^{-1}]$ for the diffuse recharge. The diffuse recharge represent only a small percentage of the total recharge. The point recharge is distributed over all the inlets depending on their catchment area. Transport transient simulations are used to simulate the tracer tests. A standard workflow is used: first the head results of a flow steady state simulation are used as initial head for the transient long term (years) flow simulation. Secondly, several transport simulations (their number depend on the number of tracer tests that have to be simulated) are run. Their initial head field is given by the outputs of the transient flow simulation. Their duration is short compared to the transient flow simulation (only several weeks, or months, depending on the rear tracer tests results).

The flow and transport simulations and the geometrical realizations can be combined in an inverse framework. Here the proposed inverse workflow is to produce many (hundreds) geometrical realizations (conduits networks), and then to optimize the flow parameters of each one and study the uncertainty on this base. In this thesis, two numerical tests have been done to estimate the feasibility of this framework. The goal of the first test is to investigate the possibility to retrieve physical and geometrical parameters using an inverse framework. A systematic analysis is made using 2D karst simulations. The conduits are represented by equivalent porous medium (EPM) cells *crossed* by a conduit of a given radius. 150 geometrical realization (150 *pseudo-genetic* models) have been generated. For each realization 120 possible combination of two physical parameters are systematically tested (radius of the conduits and matrix hydraulic conductivity). This ensemble of 18'000 simulations is then compared to a reference simulation to investigate the importance of including transport simulation (tracer tests) for the computation of the objective function ϕ. The results of this test show that including the transport results in the computation of ϕ can improve consistently the inverse results. To test the robustness of the approach, every simulation is successively used as reference simulation. This shows that in general good parameters are found. They also show that the parameter that is best characterized (with less uncertainty) is the conduit radius. The second test has been made with the aim of investigating the ability of a standard parameter estimation algorithm based on a Levenberg-Markquardt method to retrieve the good physical parameters of a realization. In this case, only transient flow is simulated. The results shows that this kind of algorithm can be used in this case, and that it allows to save a considerable CPU time in comparison with systematic search.

Chapter 1

Introduction

1.1 Motivations and state of the research

Karst aquifers are a very important freshwater resource for mankind. Biondic et al. (1995) show that they provide 35% of Europe drinking water. According to Ford and Williams (2007) they cover about 25% of the total drinkable water of the world. In Switzerland, this correspond to around 18% of the drinking water supply and 66% of Canton Neuchâtel (Spreafico and Weingartner, 2005). Karst aquifers are highly heterogeneous: they are composed on the one hand by low permeability limestone matrix, and on the other hand by highly conductive features created by the preferential dissolution of calcite (Ford and Williams, 2007) which are known as karst conduits and caves. These features may lead to geotechnical hazards especially in civil engineering, e.g. in tunneling or dam construction in karst regions (Milanovic, 2004). The conduit networks are very efficient in draining water, which leads to very high risks of pollution of spring water (Daly et al., 2002; Andreo et al., 2006). Most of the vulnerability assessment techniques used nowadays for the protection of karst system are based on GIS procedures (Ravbar, 2007). The current tools do not allow to consider the uncertainty (and consequently the risk neither) associated to karst aquifers. In a typical karst system, the matrix is assumed to have an important storage role, because it represents more than 90% of the total volume of these aquifers. The high permeability features represent a small total volume, but because they are highly interconnected they are responsible for more than 90% of the discharge of karst aquifers (Worthington et al., 2000). This highly conductive aspect can be problematic also in petroleum industry for paleo-karstic reservoirs (Corre, 1994).

The high permeability features in karst aquifers are produced by speleogenetic processes. Filipponi et al. (2009) have shown that that preferential flow paths do not occur randomly in the aquifer. They are very often initialized on *inception horizons*, which are preexisting discontinuities within the karstifiable formations, before the speleogenetical processes started. These discontinuities are usually particular bedding planes and fractures (Filipponi et al., 2009). The preferential dissolution of calcite along these horizons often lead to the formation of well developed karst conduits networks or huge caves (Palmer, 2007). Again, the currently available modeling techniques for karst aquifers, do not (or not sufficiently) consider these particular horizons, and are difficult to condition with consistency to a complex geological model.

One way to characterize aquifers in other hydrogeological environments (e.g. in porous media aquifers), and to produce forecasts, is to use analytical or numerical models. Unfortunately, the very complex nature of karst makes them very difficult to model (Bakalowicz, 2005). Several authors (e.g. Bakalowicz, 2005; Ford and Williams, 2007; Palmer, 2007) have stated that modeling karst aquifers including conduit networks and realistic physics that would pragmatically represent existing karst aquifers, represent a challenge that is not realistically achievable or difficultly believable in term of results. Nonetheless many authors have already worked on karst modeling.

Many authors have already developed models to simulate the physics of flow and transport in karst (Király et al., 1995; Király, 2003; Eisenlohr et al., 1997; Cornaton and Perrochet, 2002; Jeannin, 2001; De Rooij, 2007; Shoemaker, 2010; Reimann et al., 2011). According to Ghasemizadeh et al. (2012) who give an exhaustive review of the main existing techniques, two main types of approaches exist: lumped and distributed models. Lumped models do not try to solve the complex physics of the karst systems; they aim to properly reproduce time series (usually spring discharges) by black-box models, transfer functions or neural networks models. These models have proven efficiency when the goal is only to reproduce the time-series within the range of already observed events, but they do not provide any spatial information about the distribution of physical parameters and processes. They are also less reliable when used outside of their calibration range because they are not based on any physical ground (de Marsily, 1994). On the opposite, distributed parameter models aim at reproducing the physics of karst aquifers, such as mixed turbulent flow in the conduits and laminar flow in the matrix. A third type of models aims at understanding the speleogenetic processes. They are called *genetic* models, and include all the complex physics and chemistry of karst conduits formation (Kaufmann, 2009; Dreybrodt et al., 2005; Gabrovsek and Dreybrodt, 2010; Worthington et al., 2000; Annable, 2003). These models are done to study the processes, but they are not suitable for uncertainty quantification, because they are computationally too demanding to produce a multitude of realizations of conduit networks.

For the authors that focus on the simulation of flow and transport using distributed parameter models, one of the biggest challenges is the ability of producing a satisfactory conduit network (De Rooij, 2007). Several authors have already developed numerical tools to generate the conduits. Because of the uncertainty related to the localization of the conduits, several stochastic methods have been developed. Jeannin (1996) reviews several existing techniques that have been used to model the geometry of karst conduits. These techniques included geostatistics, fractal simulation, and random-walk techniques. Jeannin (1992) used a fractal generator to develop realistic karst paths. His simulator is an iterative algorithm, which subdivide the length between several points with respect to the fractal dimension, and generate conduits of variable directions between them. Jaquet and Jeannin (1994) used a geostatistical approach to model the conduits, using the same conduit density like a real karst network. One mature random-walk based model is the one developed by Jaquet (1998), i.e. a lattice gaz automaton simulator based on the Langevin transport equation: walkers are sent through the medium and erode it progressively. The hierarchical structure is modeled by attracting the walkers by the conduit system simulated at the previous iteration. This simulator was coupled later with flow simulation (Jaquet et al., 2004) and showed the benefits of coupling a stochastic karst generator with flow simulation. Ronayne and Gorelick (2006) propose the use of a nonlooping invasion percolation model to simulate branching channels (analogous to karst conduits). Bailly et al. (2008) are principally interested in flow simulation and generate a fissured porous media using Boolean algorithms, with 3D objects (fissures/karst) embedded in a porous matrix.

Fournillon et al. (2010) uses sequential indicator simulations in g◯cad. Their method is based on a variogram for the conduits and a proportion between conduit/no-conduit facies, the conduits are simulated as an indicator function. To enhance the results they provide an alteration map of carbonates to increase the probability of generating karst locally. This approach has the advantage to be easy to apply using already existing tools, but is not optimal to produce realistic networks. Furthermore it lacks the ability of connecting inlets and outlets.

Henrion et al. (2008) and Henrion et al. (2010b) propose to model complex cave geometries by generating first a stochastic discrete fracture network and then combining it with a truncated multi-Gaussian field in order to represent the 3D geometry of the caves around the conduit network. They focus on the geometrical generation of caves. In their approach, they first generate karst skeleton objects (line objects), then they compute the Euclidian distance from the skeleton, and finaly they use a random threshold value to truncate the distance map. In this way they obtain caves with irregular shapes, which depend on the parameters used for the random threshold. Their approach is very interesting from a geometrical point of view, but it appears difficultly applicable for fluid simulation. It would need to use flow simulators that are able to solve the flow in these geometries. This means solving complex turbulent flow in the caves, considering their small scale geometry.

Mariethoz and Renard (2011) consider karst conduits as 1D tortuous lines composed by segments of given length. They infer the statistics of strike and dip of the segments from real caves data and simulate them using high order discrete Markov processes. This approach can be conditioned to known inlets and outlets, and also to geological properties within a certain limit. Collon-Drouaillet et al. (2012) developed a method that allow also to connect inlets and outlets; they use another approximated physic (*pseudo-genetic*) karst simulator, similarly to the one developed in this thesis. Their generator takes the inlets and outlets of the system as input data, and search for the minimum effort paths between them. Instead of using a fast-Marching algorithm as proposed in this thesis, they use the $\mathcal{A}\star$ algorithm (Hart et al., 1968), but the general concept is very similar.

Pardo-Igúzquiza et al. (2011) first generate statistically several karst segments (their number depends on the simulation size), and *bind* them toghether later by using a modified diffusion-limited aggregation method. They consider vertical histograms for the presence of conduits (*inception horizons*), but do not include real geological observations on specific layers in the model. Finally, Erzeybek Balan (2012) develops a vectorial based multiple-point geostatistics (MPS) method that can be used in a non-gridded domain to reproduce the statistics of karst conduit networks, based on point datasets. First, she infers the MPS statistics from real cave datasets, and then reproduces the pattern. She applied the results of this kind of model to a paleokarst oil reservoir with good results. This example shows that karst modeling can be useful also in oil industry. The application of the last approach in a complex geological model appears difficult. Moreover, the inclusion of the knowledge about *inception horizons* (Filipponi et al., 2009) and other known important concepts for speleogenesis are not accounted for.

Unfortunately, the majority of these authors do not present flow or transport simulations based on their models, which is important because the resulting networks should produce realistic response when trying to match with observed data in the field. On that purpose Kovács (2003) inferred the characteristics of a karst aquifer by an inverse approach. Even if the latter used 2D simulations, he shows the benefits of coupling physical simulation with geometrical consideration to better understand a system. Moreover, the techniques seen until these days, have difficulties to reproduce the high connectivity that is typical of conduit networks, and are often very difficult to condition to field observation; karst conduit models need to be conditioned to observed presence of springs and inlets (sinkhole, dolines).

1.2 Goal of the thesis

The goal of this thesis is to develop a methodology to model realistic karst aquifers. The proposed methodology should model a whole karst aquifer; including geological modeling, and flow simulation. Ideally, the karst simulator should be coupled with flow and transport simulation and inserted in an inverse problem framework. This binding could allow to better quantify the uncertainty related with the models and allow to produce several models based on different assumptions. It should rely as far as possible on already existing modeling techniques, especially for flow and transport simulation, because these methods have been extensively studied (Király et al., 1995; Király, 2003; Eisenlohr et al., 1997; Cornaton and Perrochet, 2002; Jeannin, 2001; De Rooij, 2007; Shoemaker, 2010; Reimann et al., 2011).

The previous considerations imply that the first objective of the thesis is to develop a stochastic karst conduit generator that allows to produce efficiently (in term of computer resources needs) a multitude of equiprobable realizations of karst networks. The stochastic component of the approach is important to allow quantifying the uncertainty especially for those karst systems (the majority actually) that are not explorable by human beings. Uncertainty analysis is crucial for the quantification of risk in many civil engineering projects. Furthermore the final conduit models must be consistent with the actual knowledge about karst aquifers, and easy to condition to field observations. The proposed method should therefore mimic the results of speleogenetic models (Dreybrodt et al., 2005) but without solving the all the complex physics and kinetics occurring in karst systems: speleogenetic models are very demanding in terms of computational needs because they solve iteratively transient flow and transport equations, coupled with thermodynamic and chemistry. The new approach that has to be developed here, should therefore be a *pseudo-genetic* (pseudo speleogenetic) simulator. The resulting geometrical models will be constrained by geological data (such as geological model), and include the actual knowledge about the initialization of the conduits based on the *inception horizons*, such as fractures and bedding planes (Filipponi et al., 2009).

1.3 Structure of the thesis

The thesis is structured in six chapters (chapters 2 to 7). All along the thesis, the different methodologies that are developed are illustrated on the site of the Valley of *les Ponts-de-Martel*. The site context is explained in chapter 2, from regional, geological and hydrogeological perspectives. This chapter also describes the already existing geological model (Borghi, 2008) that is used as *base model* in all the following chapters of the thesis.

Chapter 3 presents the modeling technique that is used to generate the initial discontinuities of the karstic geological formation. It presents the fracture generator called *FRiGID* (*FRacture Generator based on an Implicit geomoDel*), which produces stochastic simplified discrete fracture networks (DFN) whose orientation is consistent with the orientation of the geological layers. This is done using the information about the bedding planes orientation extracted from an implicit geological model.

The following two chapters (4 and 5) describe the *pseudo genetic* methodology to generate realistic karst networks, and how it is implemented in the *Stochastic Karst Simulator* (SKS). These two chapters represent the most innovative part of the thesis.

The next chapter (6) describes how the geometrical models produced with the *Pseudo-genetic* methodology can be used to simulate flow and transport. In particular, an already existing flow simulator called GROUNDWATER (Cornaton, 2007) is used with the geometrical realizations. A particular focus is given to the kind of physics that can be used in these models, and on the assignment of boundary conditions to models for flow and transport in karst.

The last chapter of the thesis (chapter 7) presents the results of some numerical experiments related to the resolution of the inverse problem in karst modeling. In this Chapter, a possible approach to investigate the karst by inverse modeling is discussed. The first part of the chapter shows a systematic analysis which aims at analyzing the identifiability of both physical and geometrical parameters in a very simplified setup. The second part of this chapter shows how to use PEST (Doherty et al., 1994) for the estimation of the physical parameters for a given geometry.

The thesis is followed by a series of appendices. Appendix A presents a couple of source code examples that are cited in the thesis. Appendix B is a compilation of all the known tracer tests that have been realized in the Valley of *les Ponts-de-Martel* site.

Appendix C is a paper published in Hydrogeology Journal presenting the application of a modified version of the *pseudo-genetic* methodology to an impressive karst aquifer in the Yucatan peninsula (Mexico). In this paper a huge amount of aerial geophysics are used to constrain the geometrical karst models.

Chapter 2

Conceptual model and study site

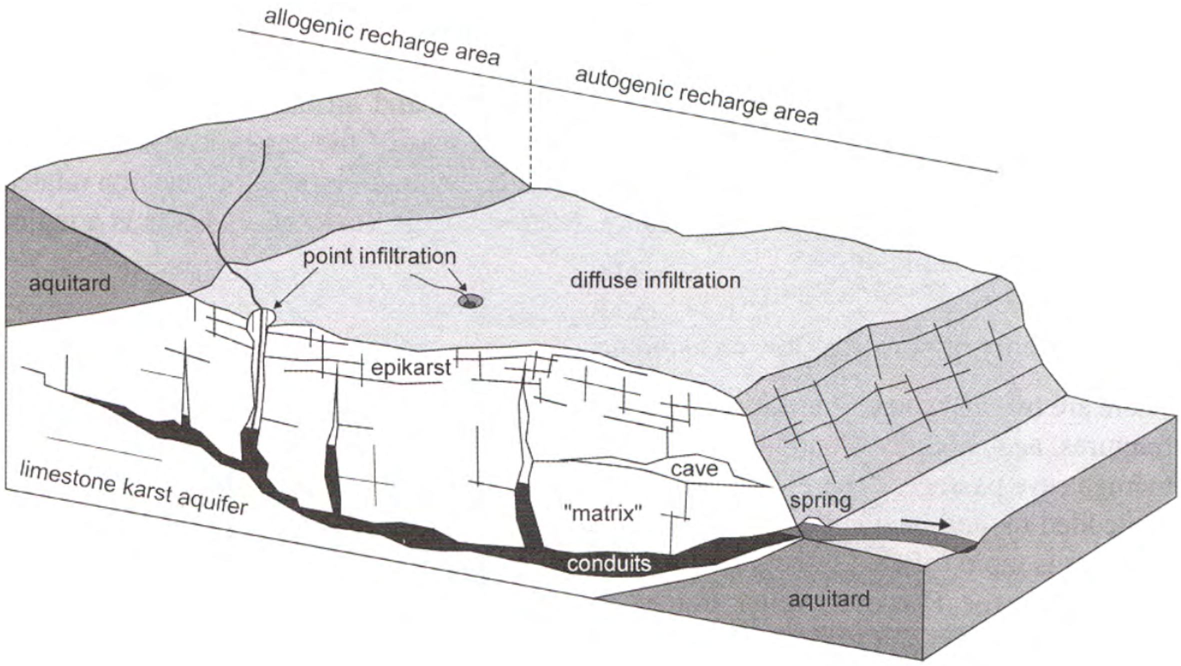


Figure 2.1: Epigenetic karst conceptual model that has to be modeled in this thesis. Modified from Goldscheider and Drew (2007)

2.1 Introduction and conceptual model

The conceptual model of epigenetic karst systems (Figure 2.1) states that Karst conduits are formed by preferential dissolution of given *inception horizons*. They can only form in karstifiable formations: there is a geologic control over the large scale structure of the aquifer. Moreover, the recharge processes of these aquifers are quite complex: on the one hand, a certain percentage of the precipitation infiltrates through the low permeability volumes, but on the other hand a great part of the recharge occurs on the form of concentrated (point) infiltration (Király et al., 1995). The concentrated infiltration has two forms: autogenic recharge and allogenic recharge. The autogenic recharge occurs when the precipitations occurring directly on the karst formation are concentrated by the epikarst towards the conduit system in the dissolutions forms of *dolines*. Allogenic recharge occurs on the contact between an aquitard formation on which rivers can flow and a karst formation, where the given river infiltrates directly into the karst aquifer conduit system through *sinkholes*. Moreover, at least two distinct zones are present in a karst conduit network: the vadose (or unsaturated) zone and the phreatic zone (or saturated). The conduits in the vadose zone tend to follow a subvertical path, while the saturated conduits follow a preferential subhorizontal path. The flow path between the inlets (sinkholes, dolines) and the outlets (springs) is supposed to follow a minimum effort path by the physical principle of minimization of energy.

In the present section, the Valley of *les Ponts-de-Martel* (Jura Mountains, Switzerland) is presented, because it illustrates a typical epigenetical karst system that can be found in the Jura mountains. The Valley of *les Ponts-de-Martel* is underlained by a karstic aquifer whose main outlet is the spring of *Noiraigue*. It is located about 20 km East of Neuchâtel city, in a highly tectonized area of the Jura Mountains in Switzerland. The site has been studied for more than a century (Desor, 1865; Daubrée, 1887; Schardt, 1906, 1908) and very soon the first conceptual models about the spring of *Noiraigue* behavior have appeared (Figure 2.2). The first known tracer test in the region to determine the hydrogeolgical link between the spring of *Noiraigue* and the sinkhole *la Perte Du Voisinage* was performed by Desor (1865).

Starting in the 70's the first modern studies were carried out by the Center of Hydrogeology and Geothermics (IGH-CHYN) of the University of Neuchâtel (Király, 1973; Morel, 1976; Gogniat, 1995a; Atteia et al., 1996; Valley, 2002; Gillmann, 2007; Borghi, 2008; Fournier, 2012). One reason for this interest is probably related to the proximity of the site with the CHYN.

In this thesis, this site has been chosen to test the proposed methodology. The reasons for this choice

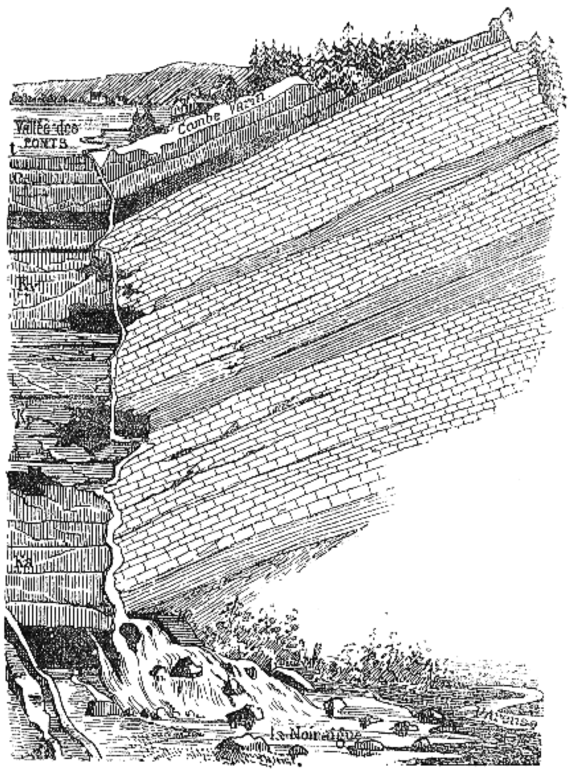


Figure 2.2: One the first known conceptual models about the spring of *Noiraigue* behavior and recharge. From Daubrée (1887).

were:

1. Existence of a large dataset related to the geology, hydrogeology and climate. This allowed to apply the methodology without having to acquire too much additional data.

2. Proximity of the site, allowing to collect new data if needed.

3. High level of complexity, allowing to test the possibility of building a realistic model even in a complex setting.

The methodology proposed in this thesis is fully explained in the following chapters. The present chapter is intended to give the background information related to the Valley of *les Ponts-de-Martel* site to minimize redundancy within the rest of the text. Section 2.2 summarizes the geographical, geological and hydrogeological settings and section 2.3 describes the geological model of the Valley of *les Ponts-de-Martel* that was built earlier (Borghi, 2008) and that will be used further in the present thesis.

2.2 Description of the site

2.2.1 Regional situation

The chain of the Jura mountains extends for approximately 350 km in its internal part and 400 km in its external part. The chain presents an arched shape, and its width varies between a few kilometers to 65 km in its wider part between Besançon and Neuchâtel. The Jura mountains are surrounded by tertiary sedimentary basins: in the North by the Rhein river graben, in the South by the Swiss Molasse basin and in the West by the "Fossé de la Bresse" basin.

The chain is subdivided in 2 main regions, the internal and the external Jura. The Valley of *les Ponts-de-Martel* belongs to the internal Jura. The internal Jura consists in a succession of well formed folds. It is also called *"la Haute Chaîne"* (i.e. the High Chain). The characteristic deformations of this region are folds, overthrusts and strike-slip faults. The folds are closely related to thrusts. In many cases, the folds are *fault-propagation folds* and *fault-bend folds*. The strike-slip faults are mainly sinistral and cut the folds with a high angle. Their orientation is mainly NE-SW in the Eastern part of the chain and rotate progressively to NE-SW in the western part of the chain. The southern limit of the chain plunges under the Molasse basin (Sommaruga, 1996).

The Jura mountains have been studied for more that three centuries. Valley (2002) gives an exhaustive overview of the available literature. One of the biggest controversies is the implication or not of the crystalline basement in the folding process. According to Sommaruga (1996) the basement is not involved, and the entire chain has moved over an evaporitic formation (mainly gypsum) at the base of the sedimentary cover in the Triassic formations.

2.2.2 Geographical situation

The Valley of *les Ponts-de-Martel* has an overall orientation of N45, and is situated in the Jura Mountain in the Canton of Neuchâtel (Figure 2.3). Its length is about 20 km and its mean width is about 3.5 km, its surface is 70 km^2. The central plain of the valley represent 2/3 of its surface, and the surrounding mountains represent 1/3 of the surface. The altitude ranges between 980 masl and 1439 masl with a mean altitude of 1065 masl.

The main village of the Valley of *les Ponts-de-Martel* is "Les Ponts-de-Martel". There are several other smaller villages called "La Sagne", "La Corbatière", "Martel Dernier" and "Brot Dessus" (Figure 2.3). The karst spring of *Noiraigue* is located in the village with the same name, in the *Val de Travers* valley. The outflowing water of the spring of *Noiraigue* flows in an important river of the canton Neuchâtel, i.e. the *Areuse* river.

2.2.3 Geological settings

The main geological structures are detachment folds that evolved into fault-propagation folds after breakthrough of thrusts, with progressive deformation (Sommaruga, 1996). The orientation of their axis are commonly SW-NE.

The Valley of *les Ponts-de-Martel* is determined in its NW side by a long anticline which extends from the SW to the NE very regularly (N60). This anticline forms the chain of *Sommartel* mounts with a maximal altitude of 1338 masl (from nb 1 to nb 8 in figure 2.4).

The southern side is more complex. The main SE anticline is quite regular in its NE part (N40), but has a complex relay structure near the the spring of *Noiraigue* (point 2 in figure 2.4). In this relay structure, the anticline rotate from N40 to N110 in a highly fracturated and faulted zone. After this relay structure, the anticline *closes* the valley against the NW anticline towards the SE. This anticline forms the chain of *Plamboz-Mont-Racine-Tête de Ran* mounts (max altitude: 1439 masl).

These two anticline are both overthrusting: the SW anticline is on a regular overthrust with a SE dip, while the NW anticline is above a back-thrust structure with a NW dip (Borghi, 2008).

Between the two anticlines, the syncline of the Valley of *les Ponts-de-Martel* presents a mostly subhorizontal dip, and a tertiary sediment filling (Molasse sandstone) of about 300m in the thicker part. Figure 2.5 present 3 cross-sections across the site. As it can be noticed, the valley tend to narrow in the NE part because the two anticlines are very close on to each other.

The study site crosses 4 geological maps: i.e. the 1:25'000 map of the Swiss geological atlas *Val de Ruz* (nb. 1144), *Neuchâtel* (nb. 1164), *Travers* (nb. 1163, unpublished) and the 1:50'000 map of the French geological atlas *Morteau*.

The stratigraphy is very well known in the region, after 2 centuries of geological studies (Figure 2.6). In the study area, the outcropping formations go from the top of the Dogger (more precisely the *Callovian* formation) to the tertiary (Molasse). After the last ice age, glacial quaternary deposits have been deposited in the center of the valley (Valley, 2002). The most relevant formations for this study belong to the Malm. They form the succession of the basal aquitard (*Argovian* marl formation) and karst aquifer (*Sequanian, Kimmeridgian and Tithonian (Portlandian)* limestone formations). Figure 2.6 gives a detailed log of the local stratigraphy[1].

[1]For more details about the local geology, refer to Sommaruga (1996); Valley (2002); Borghi (2008); Fournier (2012)

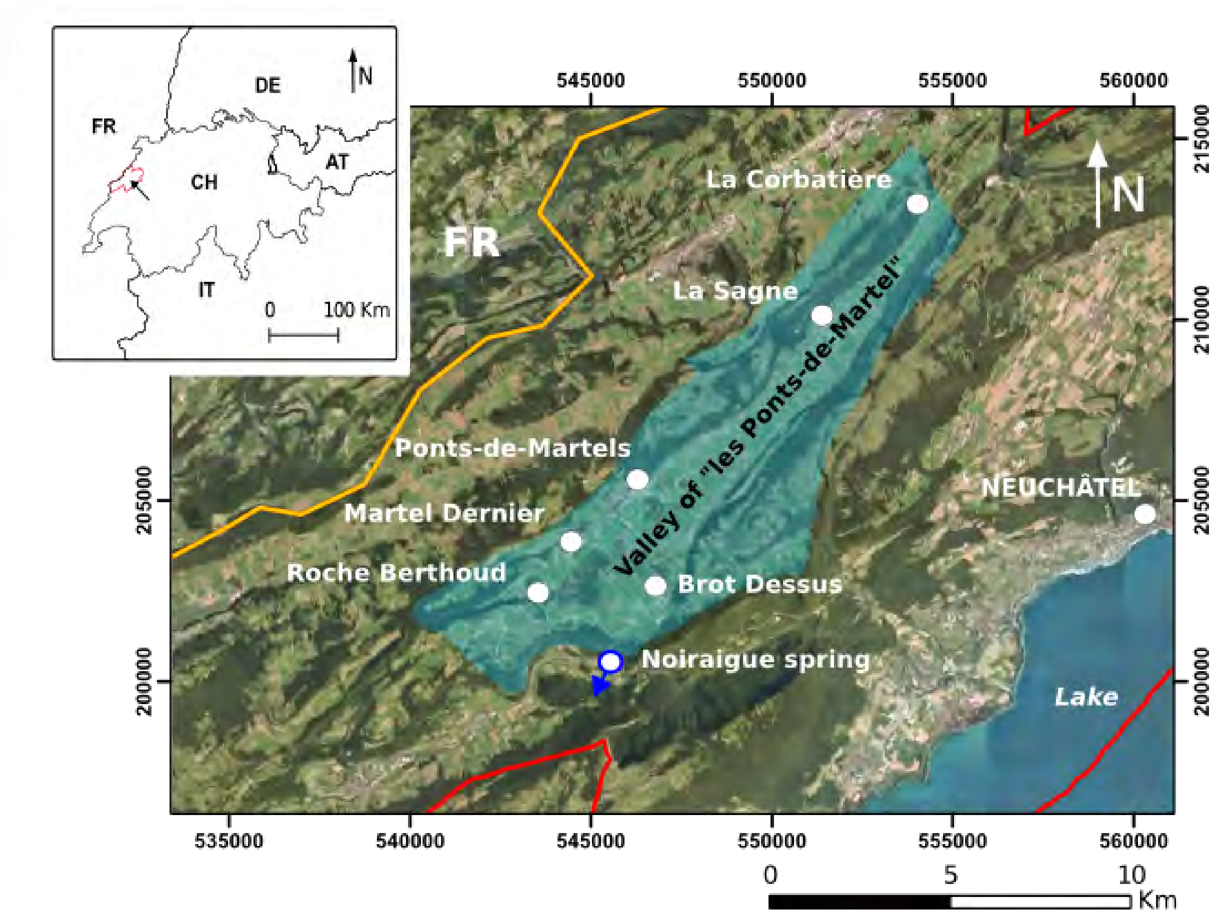


Figure 2.3: Geographical situation of the Valley of *les Ponts-de-Martel*. Yellow line: Swiss border; red line: Canton Neuchâtel border; Satellite image taken from Google Earth. Modified from Fournier (2012)

2.2.4 Hydrogeological settings

Noiraigue spring (CH-1903: 545'688/200'964, altitude 737 masl) represents the main outlet the Valley of *les Ponts-de-Martel* aquifer (Figure 2.3). From a geomorphological point of view, the Valley of *les Ponts-de-Martel* is a small endoheric catchment; it is in fact a *polje* whose totality of the precipitations infiltrates in the karst aquifer. The outcropping rocks of the catchment are composed by 22.5km^2 of quaternary sediments (31%), 2.9km^2 of Molasse and Cretaceous (4.1%), 42.3km^2 of Malm limestones (60.5%), and about 3% of "other". Morel (1976) affirms that the water of the spring of *Noiraigue* comes by 65% from the sub-outcropping limestones, 31% from the peat bogs of the Valley of *les Ponts-de-Martel*, and 3% from the Argovian combes (where Dogger may sometimes outcrop).

The three main aquifers of the region are the Dogger, Malm and Cretaceous aquifer. They are all karstified. The main aquifer of Noiraigue is the Malm aquifer (section 2.2.3). The mean discharge of the spring of *Noiraigue* is 2.02m^3/s, with a minimum of 0.15m^3/s and a maximum of about 15m^3/s (Király, 1973; Morel, 1976; Fournier, 2012).

The infiltration of the precipitation occur in two ways: *diffuse* and *point* infiltration. Diffuse recharge occurs on the outcropping Malm limestone, while point recharge is localized in the numerous dolines and sinkholes around the valley (Figure 2.7). Two small rivers drain the plain (aquitard molasse) of the Valley of *les Ponts-de-Martel*: The "Grand Bied" river, which drains the North-Eastern part of the plain, and the "Petit Bied" river, which drains the SW part of the plain. Both rivers converge into the "Bied du Voisinage" river, which infiltrate directly into the sinkhole *la Perte Du Voisinage* (CH-1903: 545'725/204'900) which is the main sinkhole of the plain (Atteia et al., 1996) located near to *Les Ponts-de-Martel* village (Figure 2.3).

The spring of *Noiraigue* is located in a highly fractured/faulted zone. Many faults are correlated with

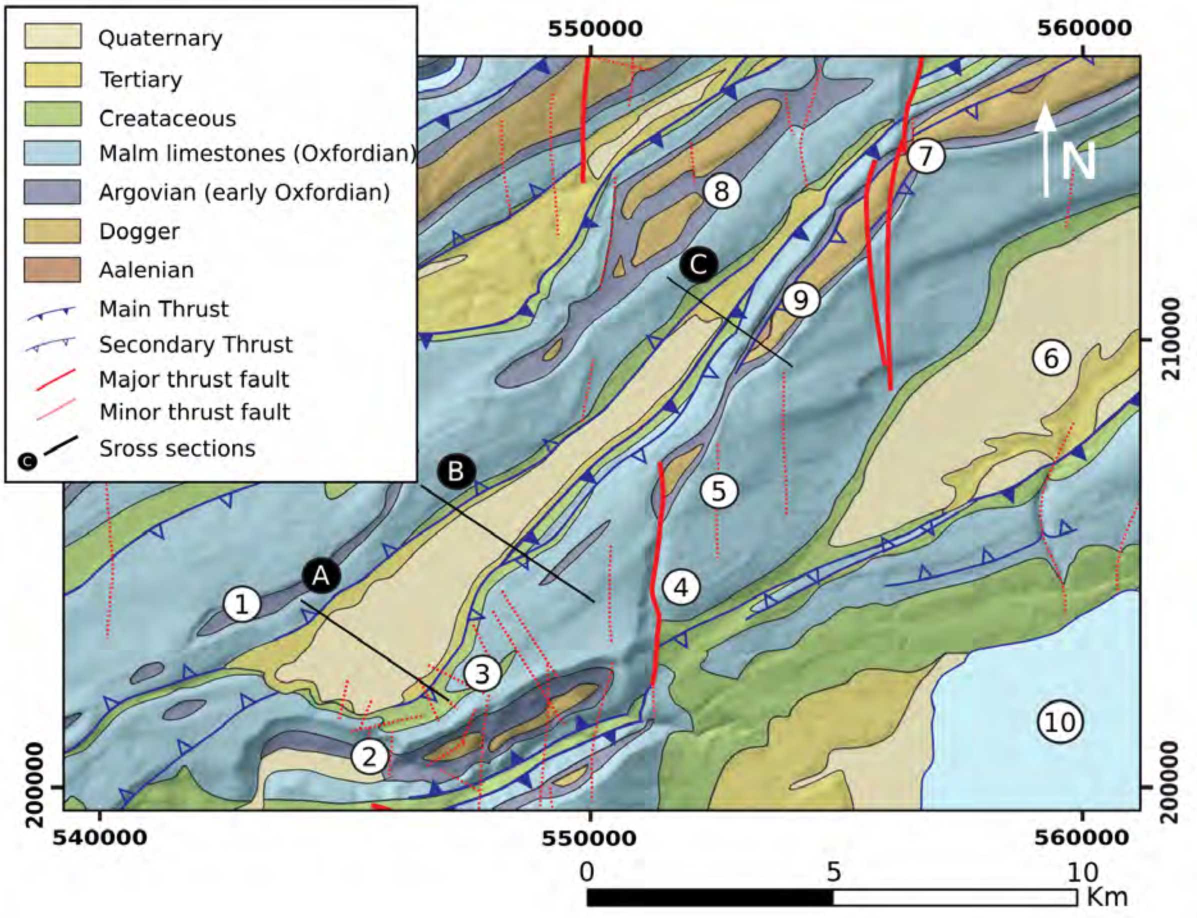


Figure 2.4: Tectonic situation of the Valley of *les Ponts-de-Martel*. Legend: 1. Combes Dernier; 2. Noiraigue; 3. Solmont; 4. La Tourne; 5. Mt Racine; 6. Val de Ruz; 7. Vue-des-Alpes; 8. Somartel; 9. Combe Cugnets; 10: Neuchâtel lake. Modified from PGN (2008).

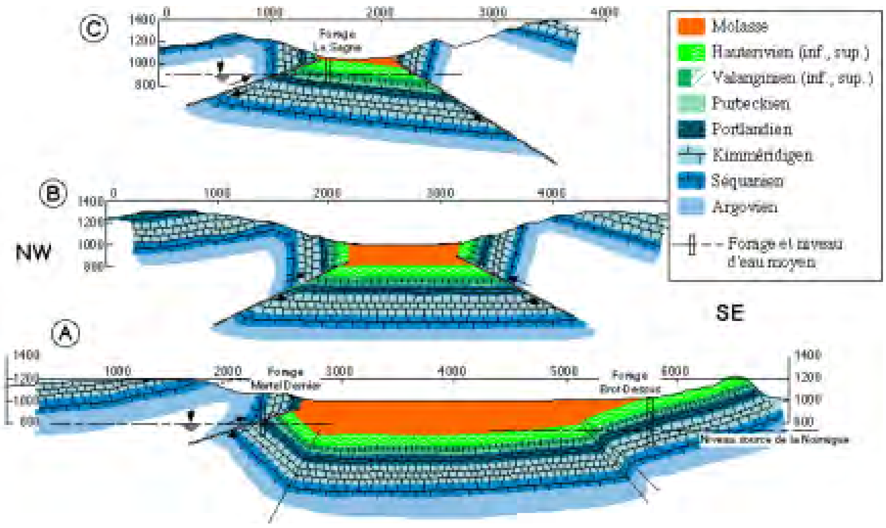


Figure 2.5: Tectonic cross section of the Valley of *les Ponts-de-Martel*. The positions of these cross-sections are located on figure 2.4. source: Atteia et al. (1996)

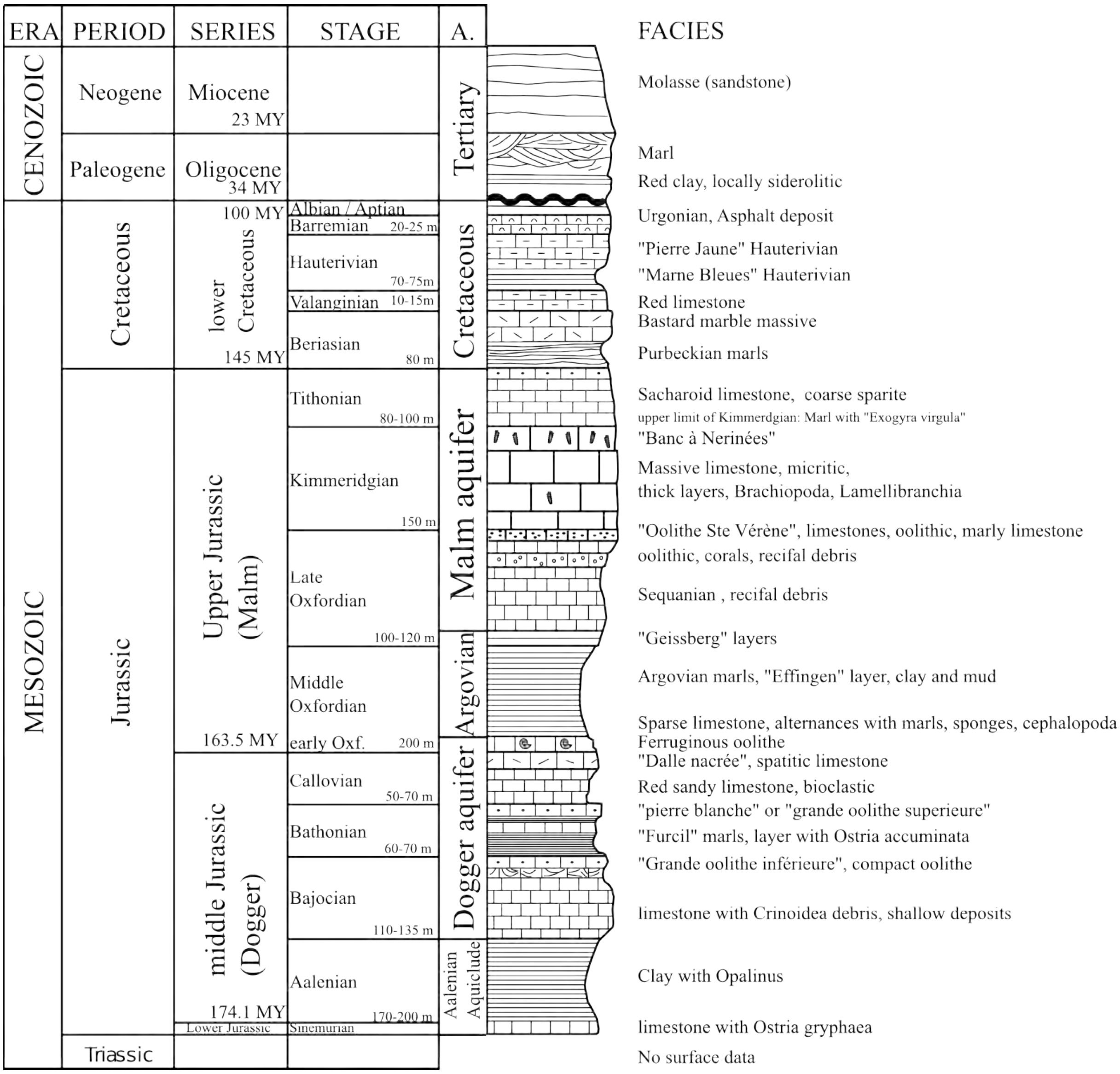


Figure 2.6: Stratigraphy of Canton Neuchâtel. Compiled from earlier works (Sommaruga, 1997; Valley, 2002; Fournier, 2012) and using the updated international stratigraphic chart of 2012. (*A.* means *Aquifer* or *Aquiclude*. The Malm and Dogger limestones are considered aquifer formations, the Cretaceous is considered an aquifer as well, but only at regional scale. Tertiary molasse, Argovian marls and Aalenian marls are considered aquicludes or aquitard formations.

the rotation of the “Solmont” anticline (Figure 2.4) which can explain the presence of big karstic conduits in this zone (Gogniat, 1995a). Two small springs are related to the spring of *Noiraigue*: the “Epinette” and the “Libarde” springs. These small springs are temporary overflow springs which are active only during high water. Their discharge are far less than the discharge of the spring of *Noiraigue* (a couple of orders of magnitude). These three springs are situated in the Malm reservoir, and located at the contact between the aquiclude *Argovian* marls and the bottom of the *Sequanian* limestones.

The limit of the catchment have first been defined by Király (1973) on the base of the outcrop map of the aquiclude *Argovian* marl formation (74km^2). A multitude of tracer tests has been performed on the Valley of *les Ponts-de-Martel* (presented in appendix B). These tests allow to slightly redefine the original catchment defined by Király (1973) as shown in Figure 2.7. The new catchment definition proposed by Gogniat (1995a) takes into account the dye test made in 1986 near *La Corbatière* (CH-1903: 554727/213601) which reappeared 10km to the N-NE and 25km to the NE from the injection point, but not at the spring of *Noiraigue*. In addition, the tracer test performed by (Matthey, 1992) near *la Tourne*

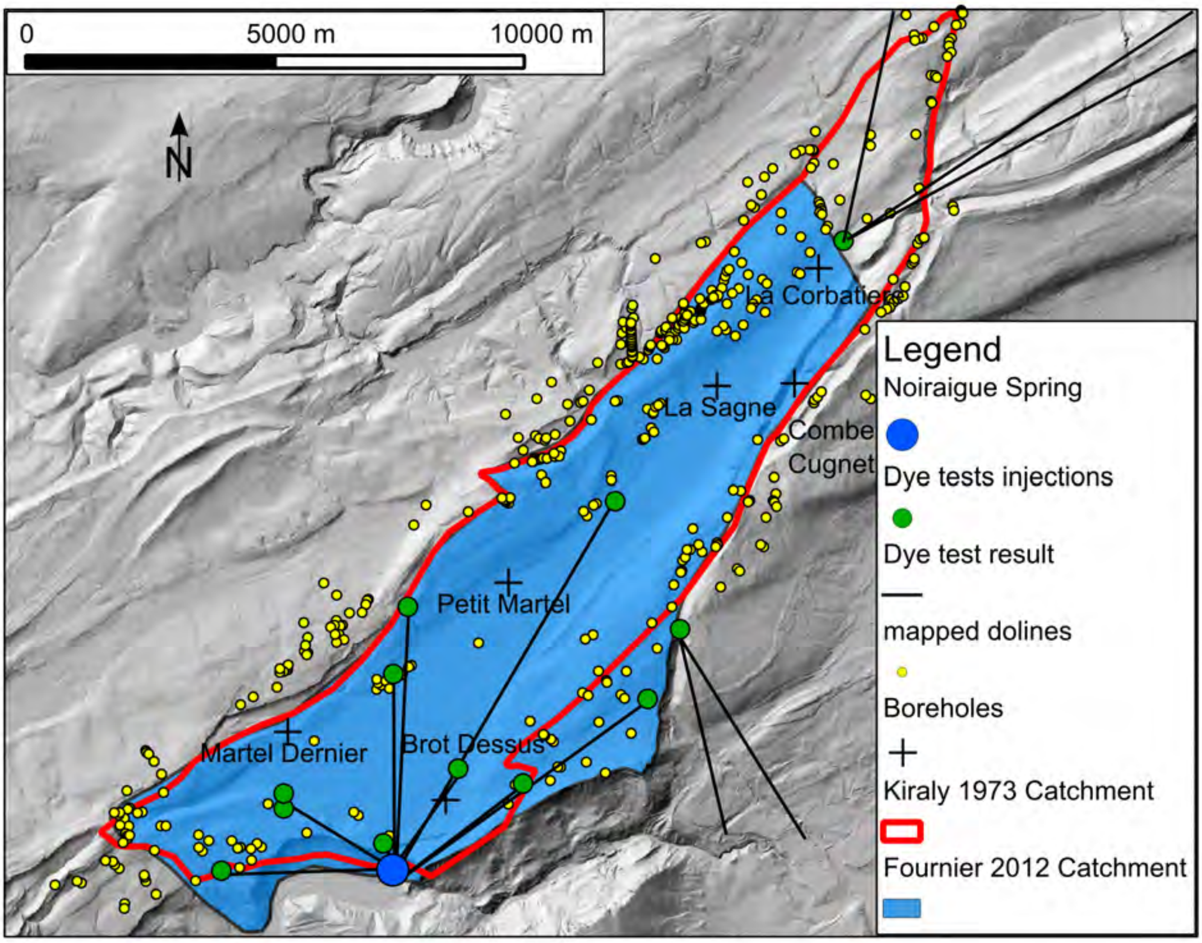


Figure 2.7: Noiraigue Spring catchment limit definition. (base map: 20m DEM)

(CH-1903: 550777/204413) shows a connection between this zone, which is unexpectedly connected to the spring. The hydraulic connection crosses the Argovian aquiclude somehow, probably because this region is highly fracturated. The retained version of the catchment definition is the one proposed by Fournier (2012) (blue surface in Figure 2.7). With respect to the catchment of Király (1973), its surface does not change consistently: Király (1973) catchment extension was 74km^2 and the new one is 73,5km^2. This is important to keep the total water-balance consistent. One dye test is located outside the spring of *Noiraigue* catchment approximately 7km to the North of the spring. It has not been considered in the definition of the catchment because its author advanced the hypothesis of the contamination of the samples (appendix B) and because successive dye tests from the same location didn't confirm this connection (Gogniat, 1995a).

All around the Valley of *les Ponts-de-Martel* a multitude of dolines are present. Fournier (2012) has mapped them using GIS processing of the new high resolution (1m) DEM of the region and manual verification. Their local catchment has also been computed by computer assisted techniques, and will be used later to evaluate the flow boundary conditions in section 6.3.5. These dolines occur often in the contact zone between the Malm aquifer and the aquitard tertiary molasse fulfillment around the Valley of *les Ponts-de-Martel*, and also in contact zone between the Malm aquifer and the argovian aquiclude (Gogniat, 1995a). The dolines are also visible in Figure 2.7.

2.2.5 Available hydrogeological data

Spring hydrograph

When studying karst aquifers, one of the most important set of observation data is the spring hydrograph. The spring hydrograph during the interest period had several missing data, because of instrument failures.

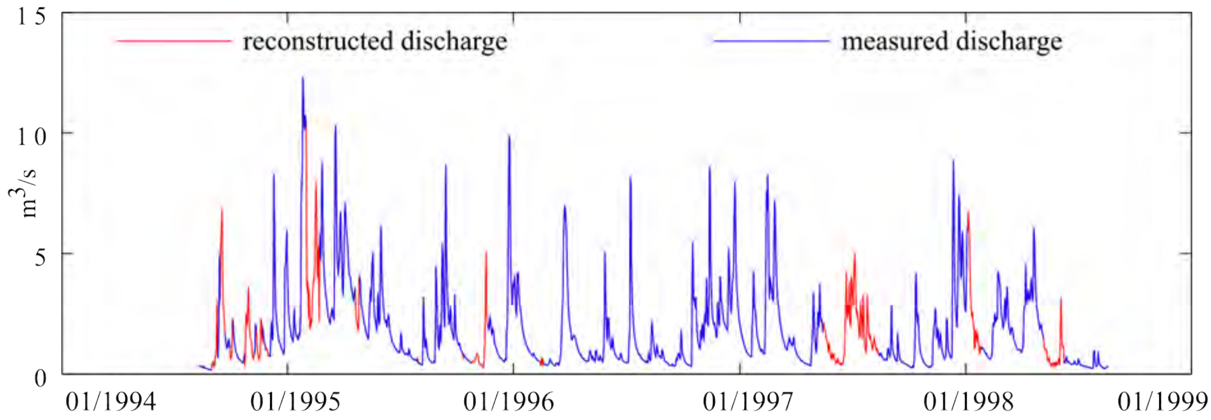


Figure 2.8: Measured and reconstructed discharge at spring of *Noiraigue*

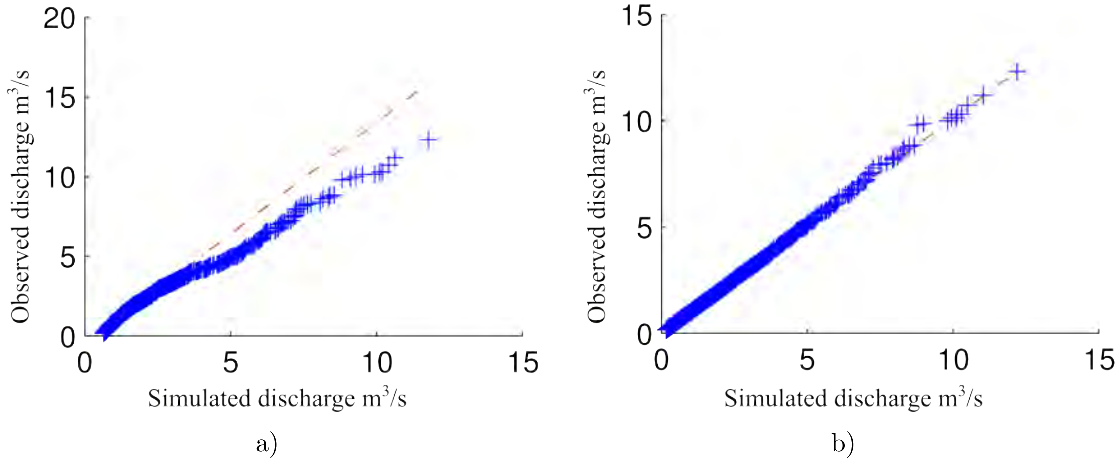


Figure 2.9: data reconstruction technique validation for the spring of *Noiraigue* discharge. a) reconstruction using linear correlation. b) reconstruction using MPDS

Fortunately, the *Areuse* river hydrograph was known for these periods. Gogniat (1995a) shows a good correlation between the *Areuse* discharge measured in *Saint Sulpice* village and the discharge of the spring of *Noiraigue*.

In this chapter, a method that uses the Multiple Point Direct Sampling (MPDS) geostatistical algorithm is used to reconstruct the missing time series part. The MPDS algorithm is described in Mariethoz et al. (2010) and its use for the reconstruction of missing data is described in Mariethoz and Renard (2010).

Unlike other multiple-point geostatistics algorithms (e.g. Strebelle, 2002; Straubhaar et al., 2011; Zhang et al., 2012) which allows to simulate only categorical variables (e.g. facies), the MPDS algorithm allows to simulate continuous variables, like e.g. the porosity of an heterogeneous aquifer. The algorithm needs a Training Image (TI) to "learn" how the variable is spatially distributed. Furthermore, an additional *auxiliary* variable can be provided. This variable provides to the algorithm a trend that can influence the simulation of the primary variable. In common geostatistical applications, the primary variable can be e.g. a flow property (like the porosity) and the auxiliary variable a geophysical measurement. If both TI (i.e. a TI for the primary property and a TI of the corresponding auxiliary variable) are provided, and if the auxiliary property is known for the simulation domain, the MPDS simulation results are more accurate than without the auxiliary variable.

Here, the primary variable is the spring discharge hydrograph. The training image is the spring hydrograph where it is known. The known part of the hydrograph is also used as conditioning data for the reconstruction simulation. The auxiliary variable is a stream or spring in the neighborhood, which is well correlated with the current spring. The auxiliary variable has to be defined over the domain that has to be reconstructed, and on the training image domain. The known part of the spring of *Noiraigue* hydrograph is used as training image and as conditioning data, while the hydrograph of the *Areuse* is used as auxiliary variable. The role of the auxiliary variable is to *influence* the results of the simulation where there is a great lack of conditioning points in the primary variable, which is the variable that is simulated (in this case the spring of *Noiraigue* hydrograph). The resulting hydrograph is shown in Figure 2.8. In this case, only one equiprobable realization of the MPDS algorithm is shown.

Another way to retrieve the missing data has been tested as "reference method": the computation of the missing data by linear correlation. The correlation straight line between the spring of *Noiraigue* discharge Q_{Nrg} (m^3/s) and the *Areuse* river discharge Q_{Ar} is given by:

$$Q_{Nrg} = 0.30 \cdot Q_{Ar} + 0.47 \tag{2.1}$$

with a linear correlation coefficient R^2 of 0.84.

To validate the resulting reconstructed data for both techniques the available data from January 1st 1996 to January 1st 1997 of the spring hydrograph has been simulated. The hydrograph simulated with the MPDS algorithm shows a significantly better results (Figure 2.9) than the linear correlation method, which in this case systematically overestimate the discharge, while the MPDS gives a very good fit. The reason of this better fit is that it is very well influenced by the auxiliary variable.

Borehole data

In the 70's, 6 deep boreholes have been realized in the Valley of *les Ponts-de-Martel*. Their depth is comprised between 200 and 450m. They have been realized on the border of the tertiary cover to investigate the limestone formations (Morel, 1976). Five of these wells have been monitored by Morel (1976) from 1971 to 1974. Gogniat (1995a) reactivated 4 of these wells in the first half of the 90's and monitored one of them again. The data about these boreholes are presented in tabular 2.1. These two studies give interesting hypotheses about the hydraulic connections and behavior of the Noiraigue spring and these boreholes. Fournier (2012) tried also to reactivate the borehole of Martel Dernier, but it was stuck at 150m depth.

According to Morel (1976) and Gogniat (1995a), the boreholes of *Brot-Dessous* (BD), *Petit-Martel* (PM), and *Combe Cugnet* (CC), present a very similar hydraulic behavior (head variations), which are very closely correlated with the spring hydrograph. A direct hydraulic connection between these boreholes and the spring of *Noiraigue* is therefore very probable. The location of these boreholes is visible in Figure 2.7.

Borehole	X	Y	Z	Depth.	Years*	H_{min}	H_{mean}	H_{max}
Petit Martel	548013	206735	1007	365	72	741	749	768
Brot Dessus	546750	203740	1019	450	72 / 97-98	738	741	760
La Sagne	552192	210674	1038	216	72	850	875	970
La Corbatière	554231	213037	1080	300	72	NaN	890	NaN
Martel Dernier	543584	203740	1023	305	72	831	840	855

* 72: period since 01.01.1972 to 31.12.1972; 97-98: period since 13.08.1997 to 11.02.1998.

Table 2.1: Malm aquifer piezometry. X,Y: geographic coordinates (CH-1903) of the well heads. Z: altitude of well heads. Depth: expressed in meters from the well head. Years: only one borehole could be measured by Gogniat (1995a) in 97-98. H_{min}: minimal head value (m); H_{mean}: mean head value (m); H_{max}: maximal head value (m).

With the analysis of the borehole data, Morel (1976) and Gogniat (1995a) have identified three distinct flow systems:

1. The system of *La Sagne - La Corbatière*
2. The system of *Martel Dernier*
3. The *main* system

The system of *La Sagne - La Corbatière* shows head level which are far superior to the ones measured downstream (Brot-Dessus,Petit-Martel), and with major variations. Moreover, Morel (1976) has shown that by low water conditions, both levels tend to a stable value around to 850 masl. This level seems to him too high to be explainable only by head losses (Morel, 1976, p.69). In fact, the hydraulic gradient, if the spring of *Noiraigue* is assumed to be the outlet (distance of 18km), is of about 0.6% ($\partial H/\partial d = (850 - 737)/18000 = 0.006$), which is still compatible with head losses.

The head in *Martel-Dernier* borehole has a variation amplitude which is much higher than the boreholes of *Petit-Martel* and *Brot Dessus*. These variations can exceed 100m. Moreover, the maximal value of this boreholes seem to reach a maximal threshold value of 853 masl. Morel (1976) and Gogniat (1995a) give to possible explanations of this phenomena: Morel (1976) thinks that this borehole shows the head values in the captive part of the reservoir (under the cover of the tertiary aquitard fulfillment), while the other two boreholes show the free aquifer zone. Gogniat (1995a) thinks that an hydraulic connection with the *Doux* basin is possible because of an higher level of conduits, which activates by high flow conditions. Unfortunately, this second hypothesis has not been confirmed by a dye test for the moment.

The *main* system is near to the spring of *Noiraigue*. The boreholes in this zone (*Petit-Martel* and *Brot Dessus*) show fast head variations which are synchronous with the spring discharge variations. Moreover, their mean level is closer to the altitude of the spring (737 masl).

2.3 Geological modeling

The model presented here has been developed by Borghi (2008)[2]. It has been used without modifications in this thesis. It has been developed from the beginning with the aim to be used in a flow model. The model has been built with the geological editor Geomodeller3D® which uses an implicit approach to model the geology (section 3.2.1). This section describes briefly the main modeling choices that have been made; the model limit and local coordinate system (section 2.3.1) and the simplified stratigraphy that has been modeled to take into account the main hydrofacies (section 2.3.2). Finally the resulting model is shown in section 2.3.3.

2.3.1 Model limit and local coordinate system

The axis of the Valley of *les Ponts-de-Martel* is mainly orientated N45. Therefore, to avoid building an excessively large model to include the whole catchment of the spring of *Noiraigue*, the model has been built in a local coordinate system. The limits of the model have been chosen in order to include the catchment defined by Gogniat (1995a) and corrected by Fournier (2012) as shown in figure 2.10. The new coordinate system is rotated (rotation of 45°, rotation center = CH-1903 (0;0)). It is a metric coordinate system which origin (0,0) is the lower left corner of the model. The maximal extension of the model is (19'150,11'050).

2.3.2 Modeled stratigraphy

Because the goal of the geological model is to provide 3D information about the location of the hydrogeological formations, only the the main hydrogeological units are considered in the model. The stratigraphic log displayed in Figure 2.6 shows a column noted as *"A."* which define the main hydrostratigraphy. The Tertiary molasse is considered as one unique aquitard unit. The Cretaceous succession of karstified limestones and marl horizons is considered as an aquifer formation on a regional scale. In fact the Cretaceous is poorly defined in the region because of a great sedimentary discontinuity between the Tertiary base and the top of the Cretaceous formations (66 My). The main aquifer of the study site, which is also the reservoir of the spring of *Noiraigue*, is the *Malm* aquifer, with an overall thickness of around 400m. This aquifer comprises the *Tithonian* (called also *Portlandian* in old terminology), the *Kimmeridgian* and upper *Oxfordian* (also called *Sequanian*). The middle *Oxfordian* marls form the *Argovian* aquiclude (from local old terminology), with a thickness of 200m. These last two formations are the most relevant for this model. Deeper the model includes two additional formations, the *Dogger* aquifer which locally outcrops in the anticline combes (only the *Callovian* stage is present as outcrop in the region), and the *Aalenian* aquiclude, whose petro-physical properties are quite unknown in the region because of the lack of outcrops. The *Triassic* series is the bottom formation of the model (Borghi, 2008).

[2]see Borghi (2008), chap. 2 for a complete description of the model

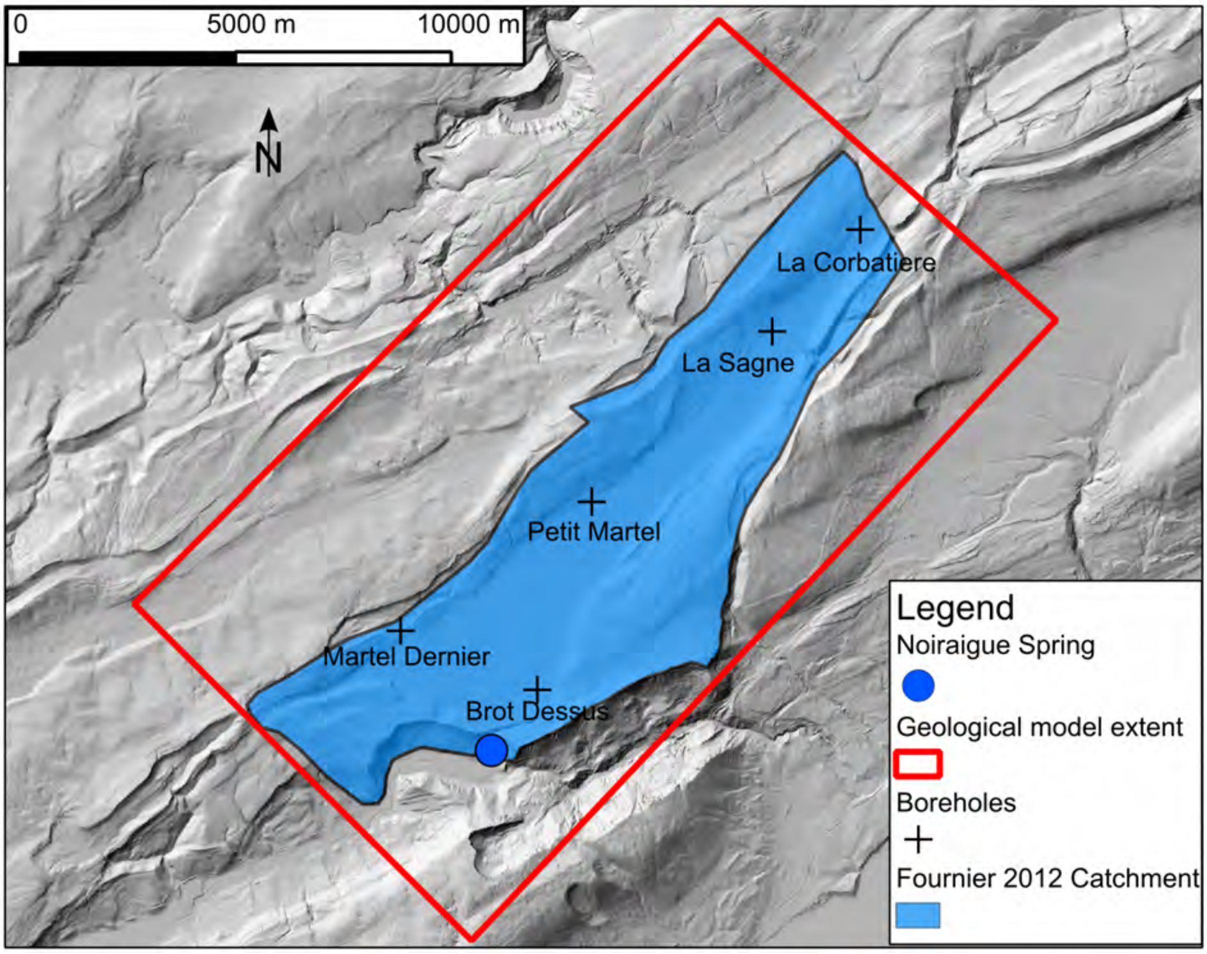


Figure 2.10: Geological model extension based on actual hydrogeological knowledge. (base map: 20m DEM)

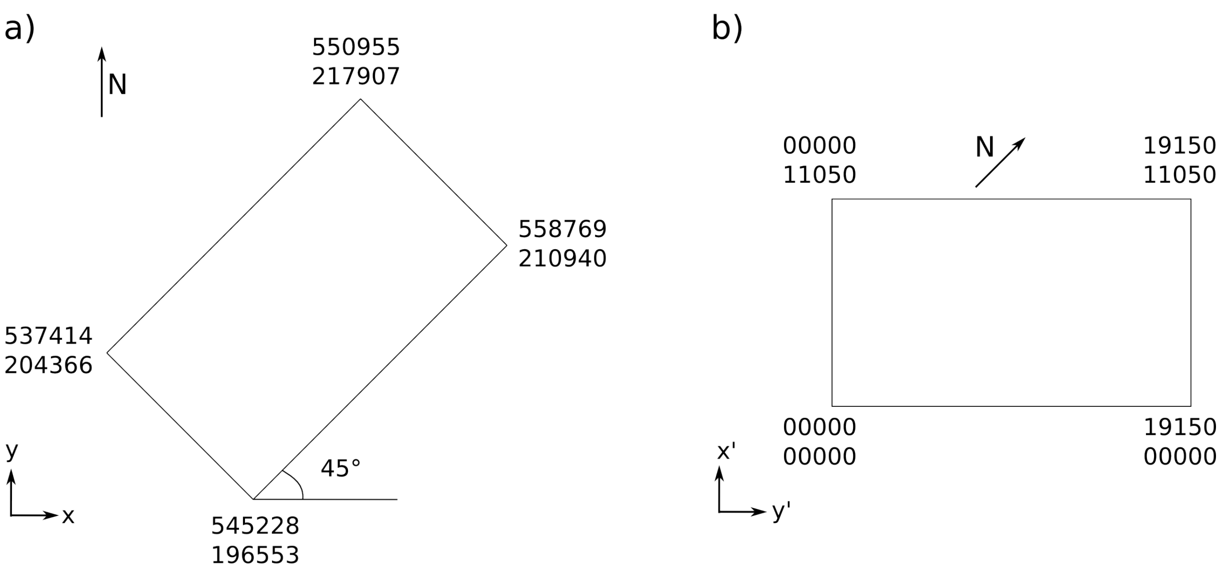


Figure 2.11: Geological model limits. a) CH-1903 coordinates. b) local metric coordinates.

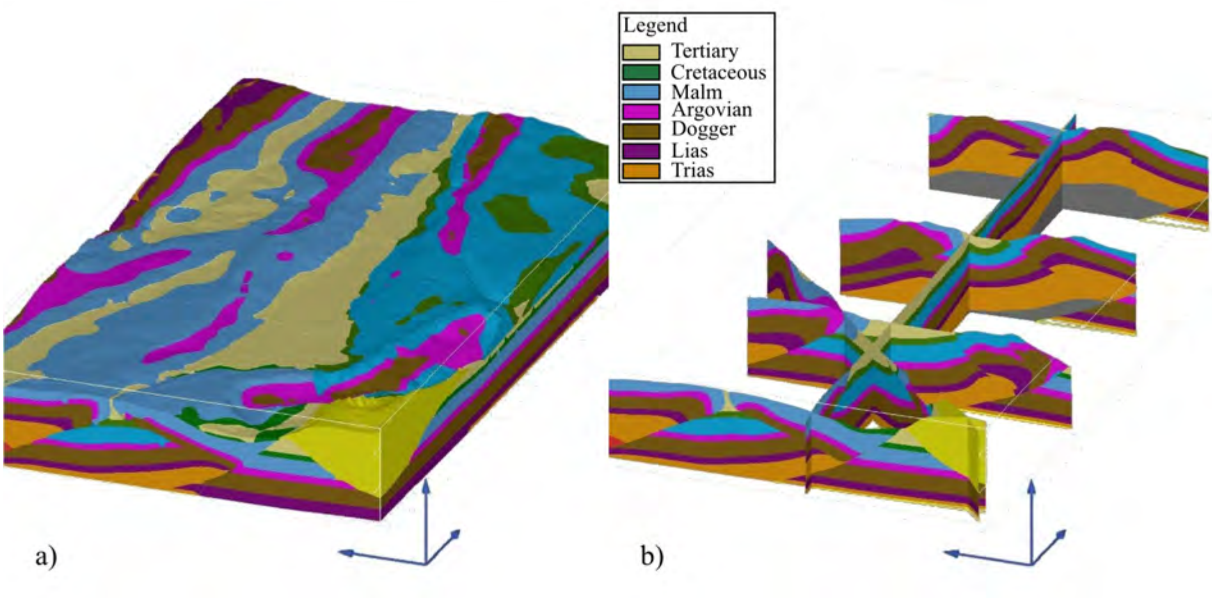


Figure 2.12: Geological model of the Valley of *les Ponts-de-Martel*

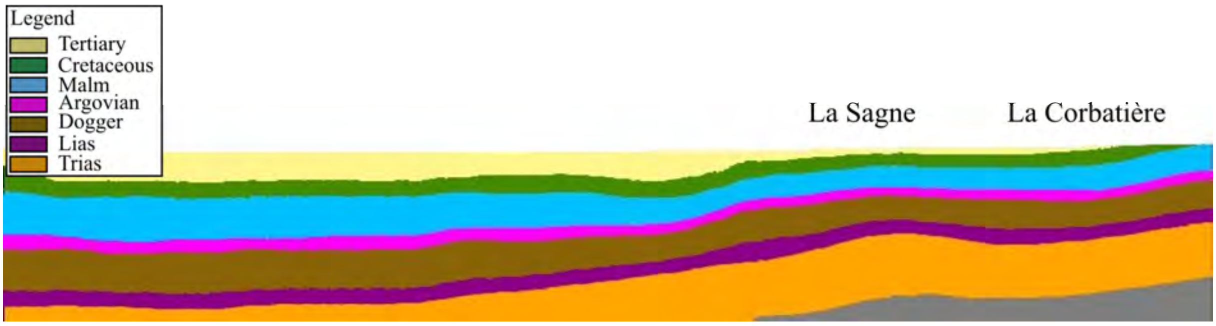


Figure 2.13: Longitudinal cross-section showing a threshold near to *La Sagne*

2.3.3 Model construction

The geological model of the Valley of *les Ponts-de-Martel* has been built using a couple of seismic profiles published in Sommaruga (1996), all the cross-sections that were available at that time (recently new cross-section have been made for a geothermics project in the last two years (PGN, 2008). This geological model includes the information resulting from the geological model built by Valley (2002), who focused on the southern part of the Valley of *les Ponts-de-Martel*. In particular, the model of Valley (2002) has been useful for the highly tectonized zone near the spring of *Noiraigue*. The thickness of the strata has been inferred from outcrop field observations. More than 50 conditioning cross-sections have been necessary to obtain a satisfactory result. The logs of the 6 deep boreholes that have been drilled in the 70's (Morel, 1976) have also been used.

The 3D model (figure 2.12) allows to better understand and visualize the complex tectonic structure of the Valley of *les Ponts-de-Martel*. The central syncline of the valley is bordered by the two main anticlines. The NE anticline is quite regular at the model scale. It is a backthrusting anticline. The Southern anticline is very complex. A great part of this complexity comes from the rotation relay structure that it presents in the zone near to the spring of *Noiraigue*. This zone is highly fractured and faulted.

Probable role of the geology on water flow

The top of the *Argovian* aquiclude is considered by Király (1973) as the base level of the *Malm* karstic aquifer. Its geometry is therefore a factor that strongly determines the possible underground flow and karstification. Király (1973) mapped the first catchment of the spring of *Noiraigue* using the geological

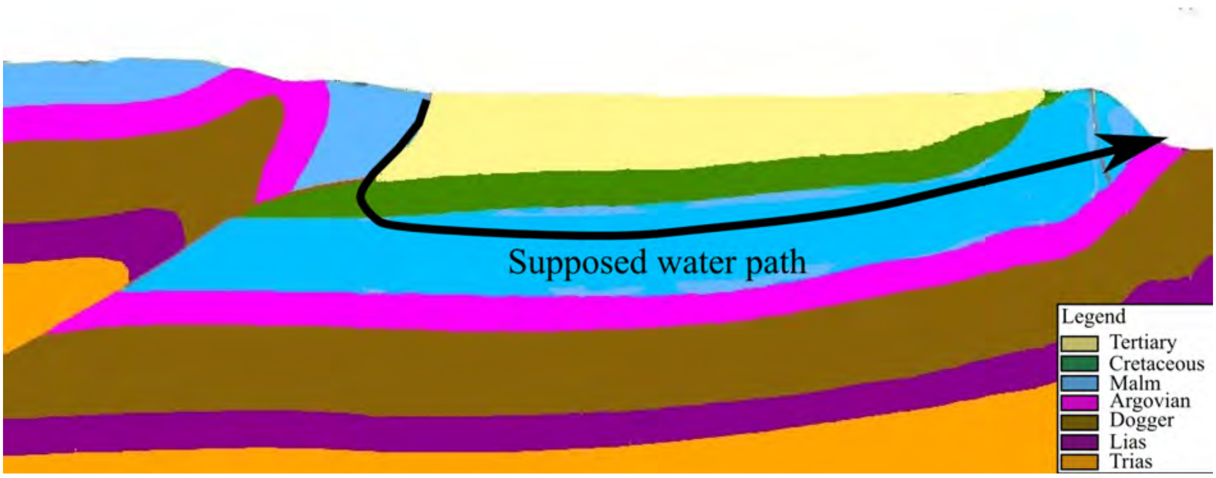


Figure 2.14: Cross section between the sinkhole *la Perte Du Voisinage* and the spring of *Noiraigue*

map. He thought that the axes of the anticlines where the Argovian formation is outcropping were an hydrogeological limit for the flow. While this hypothesis seems to be confirmed in most of the region, in some cases it has been corrected after successive tracer tests (section 2.2.4).

Borghi (2008) shows with a longitudinal cross-section in the Valley that the threshold seen by Gogniat (1995a) and Morel (1976) in the head of the boreholes of *la Sagne* and *la Corbatière* could be explained by subtile a geological threshold (Figure 2.13). Another geometrical observation on a transversal cross-section between the sinkhole *la Perte Du Voisinage* and the spring of *Noiraigue* (Figure 2.14) is that the *Argovian* aquiclude presents a sub-horizontal dip in the center of the syncline and progressively raises a the level of the spring of *Noiraigue*. These two observations on the geometry of the aquiclude shows that its geometry tends to concentrate the flow in the region of Noiraigue in both longitudinal and transversal directions. Moreover Borghi (2008) shows that the spring of *Noiraigue* is located in the lowest outcrop of the *Argovian* aquitard and that this zone has a particular shape that tends to concentrate the flow.

The role of the North backthrust surface is not completely clear. There is an evident hydraulic connection between the sinkhole *la Perte Du Voisinage* and the spring of *Noiraigue*, which means that the water crosses or uses this surface. It is not clear if this surface is a preferential karstifiable feature, or if it simply brings into contact the *Malm* aquifer of the North anticline, with the Cretaceous formation in the core of the syncline (as shown in figure 2.14). In this case, the water (an consequently the karst networks) are connected between the Cretaceous and the Malm under the backthrust surface. The *Malm* aquifer is partially captive under the *Tertiary* aquitard.

2.4 Conclusion

This chapter has provided an overview of the study site that will be used in this thesis as a *test site*. The study site is the catchment of the karstic spring of *Noiraigue*, i.e. the Valley of *les Ponts-de-Martel*. The catchment extension is about 74km^2. The spring discharge is comprised between 150 l/s and 16m^3/s with a mean value of 2m^3/s. The three most important hydrogeological formations are the molasse (sandstone) tertiary aquitard with a thinkness of 300m, the *Malm* limestone karstic aquifer (400m) and the *Argovian* marly aquiclude (200m). Several dolines and sinkholes surround the Valley of *les Ponts-de-Martel*, especially near to the contacts between the tertiary aquitard and the *Malm* formations. The most important sinkhole is the sinkhole *la Perte Du Voisinage*, where the *Bied* river infiltrates into the karstic formations. In the 70's, six deep wells have been bore.

The great and reactive discharge variations of the spring of *Noiraigue* and the results of the tracer tests clearly indicate that this site is a karst aquifer. Unfortunately, the studies that have already been done in the region did not allow to precisely localize the karst conduits. This particular issue, shows the need of development of a methodology to simulate the conduits to allow to determine their localization, their number and their size on the base of a modeling approach.

Chapter 3

Stochastic fracture generation accounting for the stratification orientation in a folded environment based on an implicit geological model[1]

[1]This chapter is planned to be submitted as article to the journal *Computer and Geosciences*.

3.1 Introduction

Natural rock systems are fractured. Fracturing has multiple origins; it occurs when the rock is exposed to sufficient constraints to break it. Fractures have a significant impact on the hydraulic properties of the rocks, as they are often preferential flow paths. For this reason, the study of fractured systems has a great importance in many fields of application: petroleum industry, hydrogeology, waste disposal, geothermics, etc. Many authors have studied fractured systems, and noticed the relationship between fracture orientation and the orientation of folded structures (Twiss and Moores, 1992; Bazagette, 2004; Bellahsen et al., 2006; Zahm and Hennings, 2009). This chapter explains a methodology that can be used to generate a stochastic Discrete Fracture Network (DFN), in which the fracture orientations are consistent with the orientation (deformation) of the geological formations.

A wide literature already exists about DFN modeling, and several reviews have been published (Chilès, 2005; Dershowitz et al., 2004; Jing, 2003; Sassi et al., 2012). The fracture "objects" in the DFN models are often assumed to be planes (often rectangular or elliptic). In some cases a regional/local strain analysis similar to Priest (1993) is used (Kloppenburg et al., 2003), or 3D maps of orientation and density derived from seismic orientation (Freudenreich et al., 2005). Other authors use geomechanical models in which the fractures "grow" during the deformation process, and the complex interactions between them (Olson, 1993) can be modeled (Maerten et al., 2000), but these models are very difficult to condition by field observations, and need high computational resources. More recently, hybrid (stochastic and geomechanical at the same time) models have been proposed (Mace, 2006), as well as *pseudo-genetic* models (Bonneau et al., 2013). These techniques mimic the complex interactions between fractures without solving all the complexity of the fracturing physics.

While the theory has been an extensive topic of research, it is difficult to find a DFN generator free to use and independent from an existing geological editor such as e.g. g◯cad (Mallet, 1992) or PETREL®. There are some DFN generator sponsored by the US government like e.g. Smith et al. (2006), or *FRAC* (Reeves et al., 2009) which produce 2D oversimplified DFN for MODFLOW. Unfortunately those fracture generators are often difficult to condition to a geological model, where the fractures are consistent with the geological knowledge. The fracture generator presented in this chapter is called *FRiGID*, which stands for: *FRacture Generator based on an Implicit geomoDel*.

The proposed code aims to model DFN composed by very simple "fracture objects", i.e. the fractures are supposed to be rectangular. These fractures are modeled in folded environments. The resulting DFN will be generated according to the local orientation of the stratification: the orientation "map" is extracted from an implicit geological model. In implicit geological modeling, the geology is not modeled as a multitude of distinct surfaces, but as iso-values of one or several continuous abstract scalar field. Nowadays there are several implicit geological methods, but two are the most common: 1) implicit modeling based on the interpolation of a so-called *potential field* (Lajaunie et al., 1997; Calcagno et al., 2008), and 2) the *GeoChron model* (Moyen, 2005). In this methodology the potential field method is used (section 3.2.1).

The code described in this chapter considers that the potential field is known. It is assumed that the orientation of the fractures will follow a conceptual model such as the one described by Price and Cosgrove (1990) or Twiss and Moores (1992) (section 3.2.4). The distribution of the size of the fractures is assumed to follow the Universal Fracture Model (UFM) (Davy et al., 2010) as explained in section 3.2.2. The algorithm first generates the fractures in an unfolded environment (section 3.2.4) and then moves and rotates them into the folded model according to the orientation of the stratigraphy (section 3.2.5). The chapter ends with a synthetic 3D example of a conform fold as proof of concept for the use of an implicit geological model as a base for the fracture orientation consistency with the geological folding orientation (section 3.3) and a demonstrative application to the Valley of *les Ponts-de-Martel* study site. This fracture generator will also be used later in the thesis. The spatial discretization of the the fractures is essential for many numerical applications (section 3.2.6) as they will be used later (chapter 5).

3.1.1 Terminology convention used in this chapter

Geological surfaces (fractures or bedding planes) can be described using different terminology conventions To avoid any ambiguity, the conventions that are used in this chapter are the following:

- **strike**: orientation of the line representing the intersection between the geological plane and the horizontal plane, given in degrees (°) with respect to the North, the values of strike can be comprised between 0°and 360°.

- **dip**: orientation of a geological plane with respect to the horizontal plane, expressed in degrees (°), the values of dip are comprised between 0°(horizontal) and 90°(vertical).
- **orientation**: normal vector to the given plane , if projected on the horizontal plane, it is perpendicular to the strike vector, it is also perpendicular to the dip.

3.2 Methodology

In this section we describe the methodology that we use to generate the fractures. Our fracture generator is based on a stochastic approach, the dimensions and locations of the faults are generated randomly. The probability distribution function (PDF) for the generation of fractures orientations and dip must be defined by the user. The code allows to use a combination of purely theoretical families of fractures, and families based on field observations. It is therefore possible to provide a model that is consistent with the fold orientation, without having field data. This give a lot of freedom to the user.

Several authors have described the different families of fractures that are present in a fold (Price, 1966; Price and Cosgrove, 1990; Twiss and Moores, 1992; Bazagette, 2004; Zahm and Hennings, 2009). Our approach is based on the Twiss and Moores (1992) theoretical families of fractures (section 3.2.4). The generation of the fractures is divided into 3 main steps:

- definition of the number and length for the fractures based on a power function, section 3.2.2
- construction of the unrotated fractures , section 3.2.4
- placement of the fractures into the geological model and rotation according to the local orientation of the structure, section 3.2.5

The information of the orientation of the strata is extracted from an implicit geological model; before going into the detail of the fracture generator, we give a short explanation of the implicit geological model that we use (section 3.2.1). This explanation is intended to be just a short overview and is based on the papers of Lajaunie et al. (1997) and Calcagno et al. (2008).

3.2.1 Implicit geological model

This modeling method is described in Lajaunie et al. (1997) and in Calcagno et al. (2008). This section is a brief synthesis of these papers which helps understanding the subsequent work.

The implicit modeling approach is based on a potential field P defined over the model domain $\boldsymbol{\Omega} \in \mathbb{R}^3$. This potential field $P(\boldsymbol{x})\ \forall \boldsymbol{x} \in \boldsymbol{\Omega}$, is a scalar function defined for each point $\boldsymbol{x}$ of the domain. The gradient $\boldsymbol{G}$ of the potential can be defined as follow:

$$\boldsymbol{G_i} = \begin{bmatrix} G_x^i \\ G_y^i \\ G_z^i \end{bmatrix} = \begin{bmatrix} \frac{\delta P}{\delta x}(\boldsymbol{x_i}) \\ \frac{\delta P}{\delta y}(\boldsymbol{x_i}) \\ \frac{\delta P}{\delta z}(\boldsymbol{x_i}) \end{bmatrix} \tag{3.1}$$

In practice, all the points having the same value of the potential $P_j(\boldsymbol{x_j})$ belong to the same geological interface, and the orientation of this interface is given by the gradient of the potential $\boldsymbol{G_j}(\boldsymbol{x_j})$. In other words, the gradient of the potential is the normal vector of the geological interfaces. These two pieces of information are very important, because they are the only ones that are normally known from field data.

The potential field is obtained from field observations by a special type of co-kriging, based on the differences between 2 values of the potential field $\boldsymbol{D_i} = \boldsymbol{P}(\boldsymbol{x_i}) - \boldsymbol{P}(\boldsymbol{x_i'})$. The geological interfaces (e.g. the interfaces between geological formations) are defined as all the points with the same value of $\boldsymbol{P}()$ such as $\boldsymbol{P}(\boldsymbol{x_a}) - \boldsymbol{P}(\boldsymbol{x_b}) = 0$ if a and b are on the same geological interface. The gradient of the potential field is normal to the geological interface and is computed with a co-kriging function of the form:

$$\boldsymbol{P^*}(\boldsymbol{x}) - \boldsymbol{P^*}(\boldsymbol{x_{ref}}) = \sum_j \lambda_j G_j^u + \sum_i \mu_i D_i \tag{3.2}$$

where $*$ refers to the estimate, $\boldsymbol{x_{ref}}$ is an arbitrary reference point with an arbitrary value of $\boldsymbol{P}$. The only important thing is to compute the relative difference of the value of $\boldsymbol{P}$ for this reference point, and any other point of $\boldsymbol{\Omega}$. λ_j and μ_i are the weights assigned to the gradient and to the difference of potential, respectively. Furthermore, to obtain a geology field that shows a succession of several formations, a non stationary mean is added to the co-kriging function:

$$m(\boldsymbol{x}) = E[\boldsymbol{P^*}(\boldsymbol{x})] = \sum_{l=1}^{L} \gamma_l f^l(\boldsymbol{x}) \tag{3.3}$$

where γ_l is the weight assigned to the value of $f^l(\boldsymbol{x})$, i.e. the function that will be used as non stationnary mean, whose can be of degree 0 (constant), 1 (plane) or polynomial.

Finally, merging eq. 3.3 and eq. 3.2 gives:

$$\boldsymbol{P^*}(\boldsymbol{x}) - \boldsymbol{P^*}(\boldsymbol{x_{ref}}) = \sum_j \lambda_j G_j^u + \sum_i \mu_i D_i + \sum_{l=1}^{L} \gamma_l f^l(\boldsymbol{x}) \tag{3.4}$$

which is the system to be solved for the interpolation of the potential field.

Figure 3.1 illustrate the concept: in Figure 3.1(a) some conditioning points are given (blue and red points). These are the points where the position of the geological interface is known. All the points belonging to the same interface have the same value of P. The gradients of the potential (i.e. the orientation of the formation) are also given at certain locations, which do not need to be near the interface points. Looking at figure 3.1(b), we see that the computed potential follows a trend (non stationary mean), it is constrained by the given values of the potential, and its orientation is locally influenced by the conditioning orientation data. Figure 3.2 shows a 3D illustration of the orientation of the geological gradient $\boldsymbol{G}$ in a folded environment.

In this work, we used the software *Geomodeller3D*®; the resulting geological model is exported on a regular grid mesh, which contains the information about lithology type and orientation for every voxel.

3.2.2 Using the UFM model for fracture length distribution

One of the most difficult parameters to measure in the field, and subsequently to model, is the fracture length. Three statistical models for the fracture length (negative exponential, uniform and normal law) are commonly used in the literature (Priest and Hudson, 1981; La Pointe and Hudson, 1985; Priest, 1993). Recent statistical studies show that the negative exponential approach is often appropriate (Odling, 1997; Bour et al., 2002; Davy et al., 2010). For this fracture generator, the fracture length model proposed by Davy et al. (2010) has been chosen: a multi-scale fracture length distribution model called *UFM* (Universal Fracture Model). This model states that most fractured systems can have 2 distinct regimes: a diluted regime and a dense one. In the diluted regime, the fracture length distribution of the fractures is not influenced by the presence of the other fractures, while in the dense regime their length is heavily constrained by the presence of other fractures, i.e. there are interactions between the different fractures. This implies that their length follows some specific density distribitution. In the present chapter, the modeled system is supposed to follow the dense regime. In fact, the UFM model is not properly applied here, as only the power-law about the fracture length distribution derived from the UFM model is used, but the interactions between neighboring fractures is not properly modeled. The power law of the UFM proposed by Davy et al. (2010) provides the number of fractures n of a given length l:

$$n(l, L) = \alpha l^{-a} L^D \tag{3.5}$$

where α (-) is a density parameter, L (m) is the size of the system and D (-) is the fractal dimension of the observed system . The parameter D is limited by the dimensionality of the system (i.e. it cannot exceed the dimension of the system, i.e. 2 for a plane and 3 for a volume) but it can be smaller than the dimensionality of the system and is not necessary an integer. The exponent a is given by the relation $a = D + 1$.

Davy et al. (2010) explain that the increasing fracture density during the fracturing process imposes that the α parameter cannot be considered universal in the *diluted regime* (the regime with small fractures, following the terminology of (Davy et al., 2010)) and depends on the maturity of the system. In the dense regime the interaction between the different fractures greatly constrain the length of the fractures, because the bigger fractures inhibit the extension of the smaller. In this regime, the law can be considered universal and α is fully determined by the fractal dimension D. A new parameter γ (dimensionless) controls the

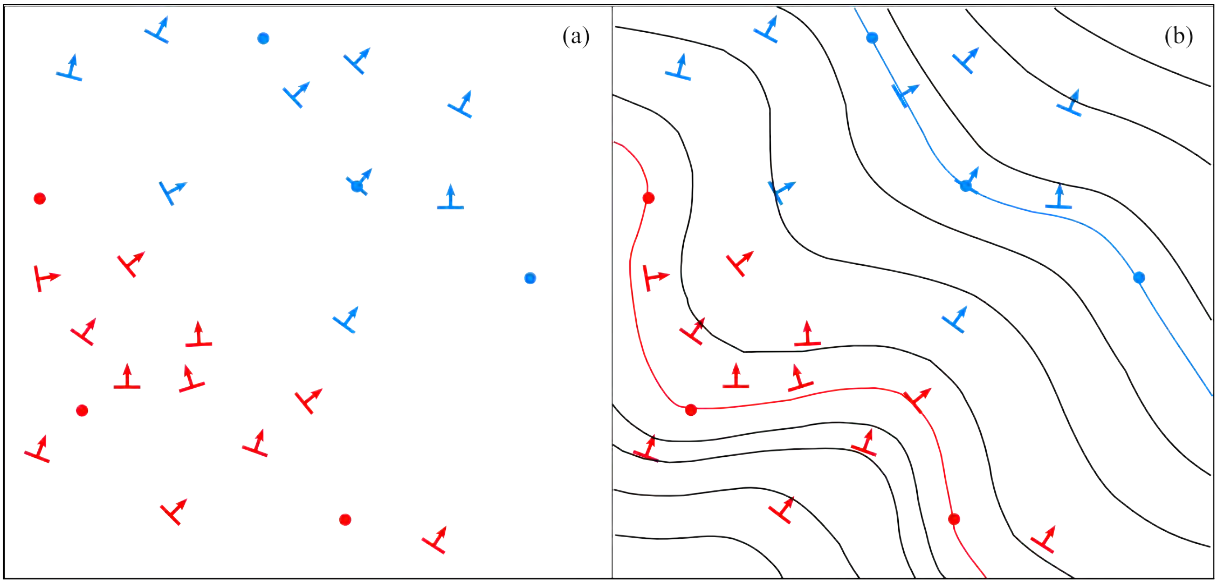


Figure 3.1: 2D example of the potential field interpolation, based on orientation data (arrows) and contact data (points) (modified from Calcagno et al., 2008), (a) conditioning points, (b) iso-values of the resulting potential field are shown. We notice that the iso-values curves are normal to the orientation given by the gradient of the potential.

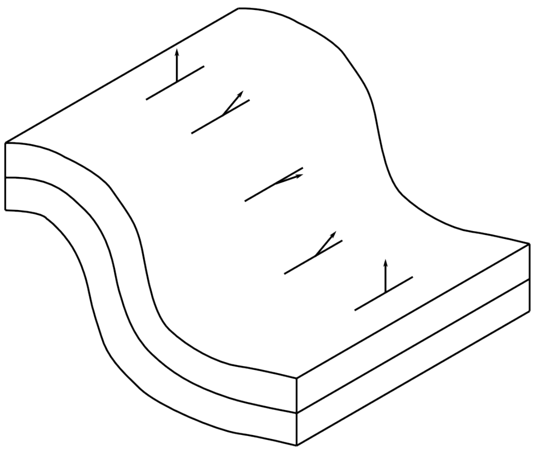


Figure 3.2: In a 3D environment, the geological gradient $\boldsymbol{G}$ (arrows) is a vector that is normal to the stratification.

influence of the relative orientation between the fractures. In this form the law for the dense regime is defined as:

$$n_{dense}(l, L) = D\gamma^D l^{-(D+1)} L^D \tag{3.6}$$

The statistical data collected by Davy et al. (2010) show that the possible range of values for γ is very restrained (between 1 and 2).

3.2.3 Consideration and integration of field data in the proposed approach

The aim of the fracture model is to be able to reproduce the families of fractures observed in the field. Because the proposed model rotates the fractures according to the local stratigraphic orientation, the field observations need to be treated before inserting them in the model. The observed fractures are therefore *rotated* into the reference coordinate system (i_0, j_0, k_0) using the inverse of the rotation matrix R defined by the local stratigraphy orientation (Figure 3.3). To do that, first the rotation matrix must be defined as follows:

$$R = (\boldsymbol{s_i}, \boldsymbol{s_j}, \boldsymbol{s_k}) \tag{3.7}$$

with $\boldsymbol{s_i}$, the vector with a dip of $-\theta$ (from the horizontal plane) and an azimut of $-\alpha$ (from the X axis):

$$\boldsymbol{s_i} = \begin{bmatrix} s_{ix} \\ s_{iy} \\ s_{iz} \end{bmatrix} = \begin{bmatrix} \cos\alpha\cos\theta \\ -\sin\alpha\cos\theta \\ -\sin\theta \end{bmatrix} \tag{3.8}$$

$\boldsymbol{s_j}$, horizontal normal vector to $\boldsymbol{s_i}$:

$$\boldsymbol{s_j} = \begin{bmatrix} s_{jx} \\ s_{jy} \\ s_{jz} \end{bmatrix} = \begin{bmatrix} \cos(\pi/2 - \alpha) \\ \sin(\pi/2 - \alpha) \\ 0 \end{bmatrix} \tag{3.9}$$

and $\boldsymbol{s_k}$ the normal vector to the plane, computed by vector product of $\boldsymbol{s_i}$ and $\boldsymbol{s_j}$:

$$\boldsymbol{s_k} = \boldsymbol{s_i} \times \boldsymbol{s_j}, \tag{3.10}$$

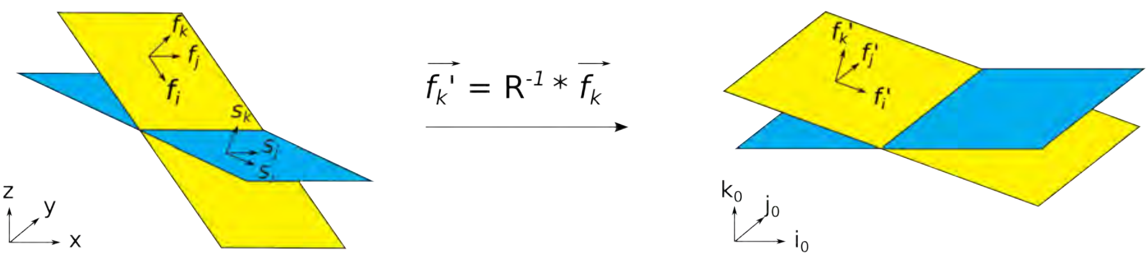


Figure 3.3: The fractures measured on the field are rotated in the (i_0, j_0, k_0) base using the inverse of the rotation matrix R defined by the local stratigraphy orientation.

The fracture plane system $(\boldsymbol{f_i}, \boldsymbol{f_j}, \boldsymbol{f_k})$ is computed in the same way. It is used to retrieve the rotated (*unfolded*) fracture dip θ_f and dip azimut α_f by first computing the values of the *unfolded* normal vector to the fracture plane $\boldsymbol{f'_k}$:

$$\boldsymbol{f'_k} = R^{-1}\boldsymbol{f_k} \tag{3.11}$$

which is used to compute the horizontal orthogonal vector $\boldsymbol{f'_j}$:

$$\boldsymbol{f'_j} = \begin{bmatrix} -f'_{ky} \\ f'_{kx} \\ 0 \end{bmatrix} \tag{3.12}$$

from which α_f can be computed:

$$\alpha_f = \arccos\left(\frac{f'_{jx}}{||\boldsymbol{f'_j}||}\right) - \frac{\pi}{2} \tag{3.13}$$

θ_f can be computed from $\boldsymbol{f'_i}$:

$$\boldsymbol{f'_i} = \boldsymbol{f'_k} \times \boldsymbol{f'_j} \tag{3.14}$$

in this way:

$$\theta_f = -\arcsin(f'_{iz}) \tag{3.15}$$

So, all the fractures measured in the field can be *normalized*, which allows to have fracturation statistics independently from the orientation of the stratigraphy. This is useful to build fracture networks that are consistent with the observed fractures families even if the site is very folded.

3.2.4 Construction of the initial set of fractures

The second step of the proposed approach is to build a number n of rectangular fractures of length l as defined above (section 3.2.2).

Definition of the fracture families in the model

In our model, we use the theoretical families of fractures in a fold as proposed by Twiss and Moores (1992). This conceptual model indicates that different fracture families will occur, depending on the position within the fold (Figure 3.4a). Even if this is only a conceptual representation of natural systems; it provides some guidelines for the definition of theoretical fracture families, which are often actually observed on the field.

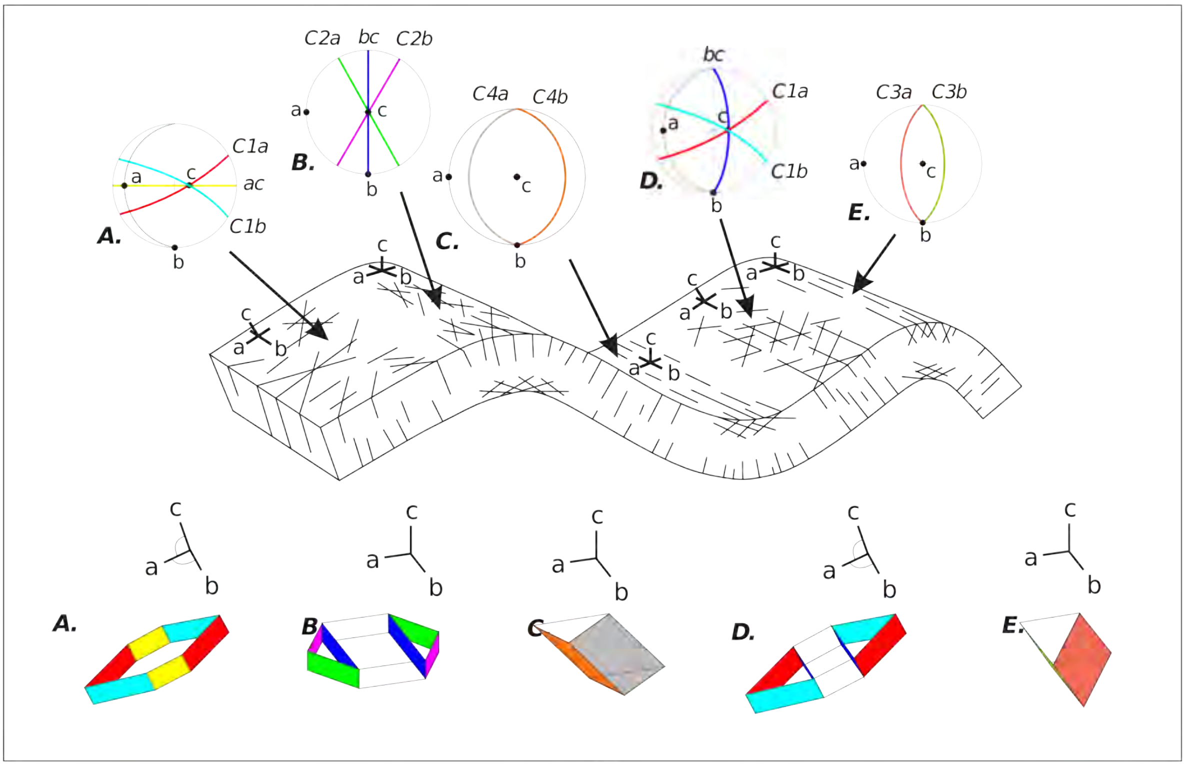
- Conjugate system C1: This system is globally perpendicular to the fold axis: 2 conjugated fracture system slanting to the strike of the bedding planes and perpendicular with respect to its dip: C1a and C1b (situation A and D in figure 3.4a).
- Conjugate system C2: This system is globally parallel to the fold axis: 2 subvertical conjugate fracture systems slanting to the strike of the bedding planes and perpendicular with respect to its dip: C2a and C2b (situation B in figure 3.4a).
- Conjugate system C3: This system is globally parallel to the strike of the stratigraphy: 2 inclined conjugate fracture systems (C3a and C3b) and slanting with respect to its dip (situation E in figure 3.4a).
- Conjugate system C4: 2 inclined conjugate fracture system (C4a and C4b) parallel to the bedding planes strike and slanting with respect to its dip (situation C in figure 3.4a).
- Fractures with ac orientation: Vertical fractures that have their strike perpendicular to the fold axis and which dip is perpendicular to the bedding planes dip.
- Fractures with bc orientation: Inclined fractures whit a strike that is parallel to the axis of the fold and which dip is perpendicular to the bedding planes dip.

The fractures systems C1 and ac are typical for fold limbs. The systems C2,C3 and bc are typical for the fold extrados (extension), while the system C4 is characteristic for the fold intrados (compression).

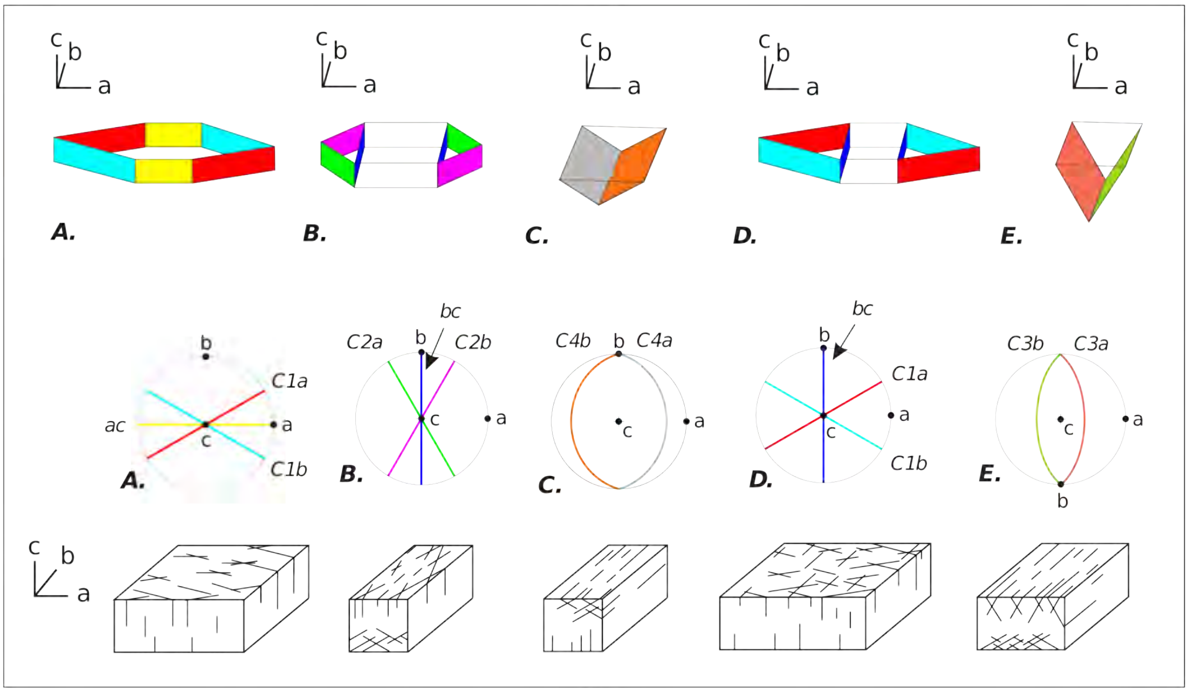
Figure 3.4a shows the fractures families in the folded environment, Figure 3.4b shows the families in an unfolded environment, i.e. in the environment that will be used do generate them as explained in the following sections.

Strike and dip PDF definition

Each of the fracture families described above are given for a theoretical case. In practice, the fractures are never perfectly parallel between them. Therefore in our approach, for every family the user has to define a particular PDF. In the case of an uniform distribution, upper and lower bound values for both strike and dip have to be provided, and in the case of a normal distribution, the mean value and standard deviation have to be provided. When each fracture composing the initial set is built, a random value of



(a)



(b)

Figure 3.4: Illustration of the fracture families in: a) the folded environment (modified from Twiss and Moores, 1992), b) in a unfolded environment (see section 3.2.4). The colors are only meant to identify easily the families in the stereoplots.

dip and strike is retrieved from the PDF of the family which the fracture belongs to. Moreover, if field observation are available, the fracture families that have been observed can be imported in this model following the methodology described is section 3.2.3.

Inital coordinates of the fracture vertices

Each initial fracture is generated after its length and its orientation information are defined. As shown in Figure 3.5, the fractures are defined as being rectangular, and their initial vertexes (the 4 bounds of the fracture) are computed. Their length on the horizontal plane l is given by the *UFM* (section 3.2.2), and their second length (in the dip direction) is given by $d = a * l$, with a a user-defined (or random) ratio. The vertices coordinates $\boldsymbol{x_i}$ with $i \in 0, 1, 2, 3$:

$$\boldsymbol{x_i} = \boldsymbol{x_0} + \overrightarrow{\boldsymbol{x_0 x_i}} = \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} \tag{3.16}$$

and are computed as follow:

$$\boldsymbol{x_0} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{x_1} = \begin{bmatrix} dx_1 \\ dy_1 \\ 0 \end{bmatrix}, \quad \boldsymbol{x_2} = \begin{bmatrix} dx_2 \\ -dy_2 \\ -dz \end{bmatrix}, \quad \boldsymbol{x_3} = \begin{bmatrix} dx_1 + dx_2 \\ dy_1 - dy_2 \\ dz \end{bmatrix} \tag{3.17}$$

with $dx_1 = \cos(\alpha) * l$ and $dy_1 = \sin(\alpha) * l$ (where $-\alpha$ is the strike of the fracture with respect to the $\boldsymbol{i}$ axis) are the components of $\overrightarrow{\boldsymbol{x_0 x_1}}$. $dx_2 = \sin(\alpha) * d'$, $dy_2 = \cos(\alpha) * d'$, and $dz = \sin(\theta)$ are the components of $\overrightarrow{\boldsymbol{x_0 x_2}}$ (with $d' = \cos(\theta) * d$ and θ the dip of the fracture). Finally $\overrightarrow{\boldsymbol{x_0 x_3}} = \overrightarrow{\boldsymbol{x_0 x_1}} + \overrightarrow{\boldsymbol{x_0 x_2}}$.

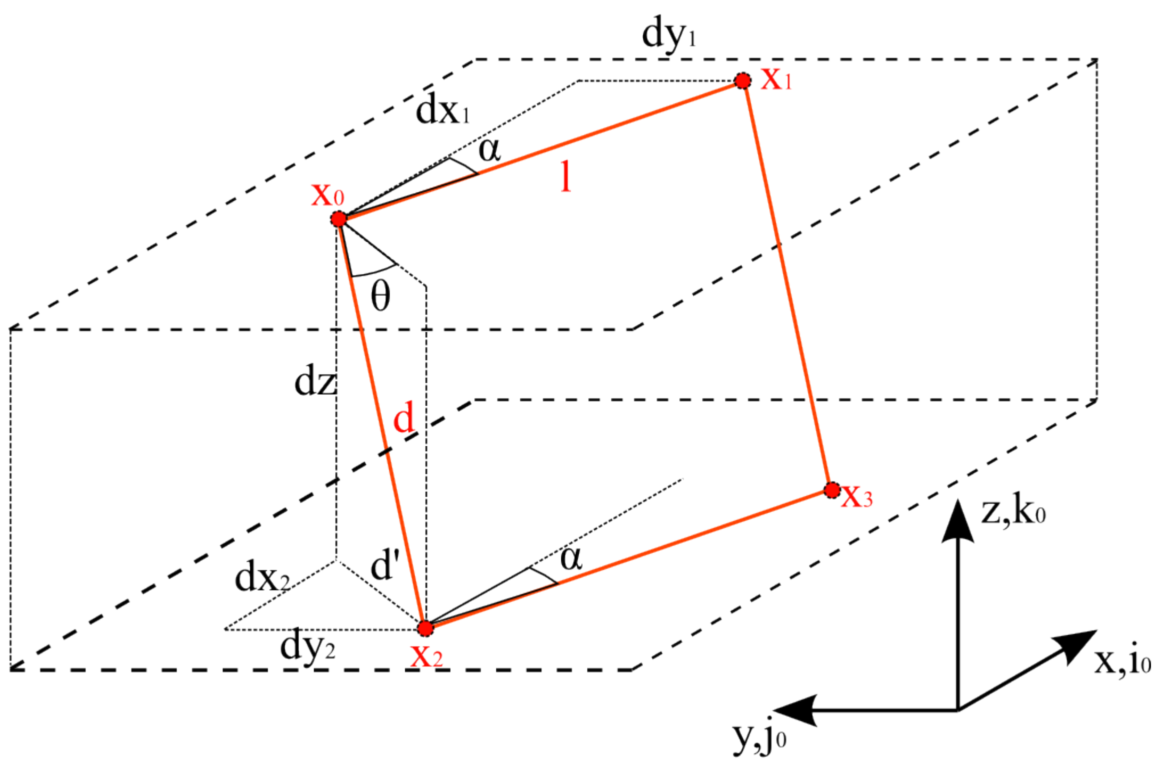


Figure 3.5: The 4 initial points of the basic "fracture object", x_i with $i \in [0, 1, 2, 3]$, see the text in section 3.2.4 for the explanation of the other symbols

3.2.5 Placement and rotation of the fractures

After the initial fractures are built, they are translated one by one into the 3D model, and then rotated according to the local stratigraphy orientation. Clearly, if a geological editor that do not allow to know the structure orientation is not usef, this step is not possible, and the fractures would be kept with the original orientation.

Translation of the fractures inside the model

Let us consider one fracture previously generated. Before moving it, its gravity center $\boldsymbol{c}$ is computed:

$$\boldsymbol{c} = \frac{1}{4}\sum_{i=0}^{3} \boldsymbol{x_i} \tag{3.18}$$

as well as the lag vectors $\boldsymbol{v_i}$, which are the vectors between the gravity center $\boldsymbol{c}$ and each point $\boldsymbol{x_i}$ of the fracture:

$$\boldsymbol{v_i} = \boldsymbol{x_i} - \boldsymbol{c} \tag{3.19}$$

To be able to use the orientation information, we export our model from *Geomodeller3D®* in a voxel grid. For every voxel, the information available is a code indicating the geological formation F_g (where $g \in 1, 2, ..., n_g$ with n_g the number of geological formations in the model), and the gradient $\boldsymbol{G}$ of the potential in the voxel. The combination of these 2 pieces of information allow us to define several rules for the generation of the fractures. The first rule is a given family of fractures will develop only into certain specific formations, e.g. the fractures can be allowed to be generated only in hard rock formations (like limestones or sandstones) but not into soft rocks (shales). This is done by first indexing the voxels that belong to the different formations with an indicator function I_g that is defined for every voxel $\boldsymbol{i} = (i, j, k)$ of the model and for every formation F_g:

$$I_g(\boldsymbol{i}) = \begin{cases} 1 & \text{if } F(\boldsymbol{i}) = F_g \\ 0 & \text{if not} \end{cases} \tag{3.20}$$

In the same way we can define the indicator of brittle rocks $I_{brittle}$:

$$I_{brittle}(\boldsymbol{i}) = \begin{cases} 1 & \text{if } F(\boldsymbol{i}) \text{ is brittle} \\ 0 & \text{if not} \end{cases} \tag{3.21}$$

The initial probability of having a fracture $\rho(\boldsymbol{i})$ in the voxel $\boldsymbol{i}$ has to be defined. Using the indicator function we can define the final probability of having a fracture $\rho'(\boldsymbol{i})$:

$$\rho'(\boldsymbol{i}) = \rho(\boldsymbol{i}) * I_{brittle}(\boldsymbol{i}) \quad \text{or} \quad \rho'(\boldsymbol{i}) = \rho(\boldsymbol{i}) * I_g(\boldsymbol{i}) \tag{3.22}$$

depending if we want to use the indicator function of the formations I_g or the indicator function of the brittle formations $I_{brittle}$ (which is normally the case). After that, we select randomly one of the cells $\boldsymbol{i_r}$ and test if it is possible to place a fracture in this voxel. Knowing the origin of the grid $\boldsymbol{o} = (o_x, o_y, o_z)$, and the dimensions of the cells dx, dy, dz, then a random point $\boldsymbol{x_r}$ is generated inside the cell using a uniform distribution:

$$\boldsymbol{x_r} = \begin{bmatrix} o_x + (j_c - 1)dx + dxrand() \\ o_y + (i_c - 1)dy + dyrand() \\ o_z + (k_c - 1)dz + dzrand() \end{bmatrix} \tag{3.23}$$

where i_c, j_c, k_c are the matrix indices of the randomly selected cell, and $rand()$ is a random function that generates a random floating point number in the interval $[0, 1]$.

Finally the displacement vector $\boldsymbol{d}$ between $\boldsymbol{c}$ (gravity center of the fracture) and the random point $\boldsymbol{x_r}$ is computed. $\boldsymbol{x_r}$ becomes the new gravity center of the fracture:

$$\boldsymbol{d} = \boldsymbol{x_r} - \boldsymbol{c} \tag{3.24}$$

The new points of the fracture $\boldsymbol{x_i'}$ are computed with their respective $\boldsymbol{v_i}$ vectors or by using the displacement vector $\boldsymbol{d}$:

$$\boldsymbol{x_i'} = \boldsymbol{x_r} + \boldsymbol{v_i} = \boldsymbol{x_i} + \boldsymbol{d} \tag{3.25}$$

Rotation according to the local geological orientation

After the fracture has been placed into the medium, a rotation is applied to account for the local geological orientation. The rotation of the fracture is made in 3 steps. First a new local orthonormal basis consistent with the orientation of the stratigraphy is built. Secondly the rotation matrix is computed, and finally every point of the fracture is rotated using this rotation matrix.

The new orthonormal basis ($\boldsymbol{i}$,$\boldsymbol{j}$,$\boldsymbol{k}$) is defined as follows: $\boldsymbol{k}$ is the orthonormal vector with the same orientation as the local gradient $\boldsymbol{G}$ (section 3.2.1), $\boldsymbol{j}$ is the orthonormal vector that is horizontal and $\boldsymbol{i}$ is the orthonormal vector of $\boldsymbol{k}$ and $\boldsymbol{j}$, i.e. $\boldsymbol{i}$ is parallel to the dip, as shown in figure 3.6:

$$\boldsymbol{k} = \begin{bmatrix} g_x \\ g_y \\ g_z \end{bmatrix} /\|\boldsymbol{G}\| \ , \quad \boldsymbol{j} = \begin{bmatrix} -k_y \\ k_x \\ 0 \end{bmatrix} /\|\overrightarrow{k_y k_x}\| \ , \quad \boldsymbol{i} = \boldsymbol{j} \times \boldsymbol{k} \tag{3.26}$$

Where $g_{x,y,z}$ are the components of $\boldsymbol{G}$. The rotation matrix $\boldsymbol{R}$ is defined as follows:

$$\boldsymbol{R} = [\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}] \tag{3.27}$$

We can now use this rotation matrix to generate the four rotated vertexes of the fracture $\boldsymbol{x'_{ir}}$:

$$\boldsymbol{x'_{ir}} = \boldsymbol{x_r} + \boldsymbol{R}\boldsymbol{v_i} \tag{3.28}$$

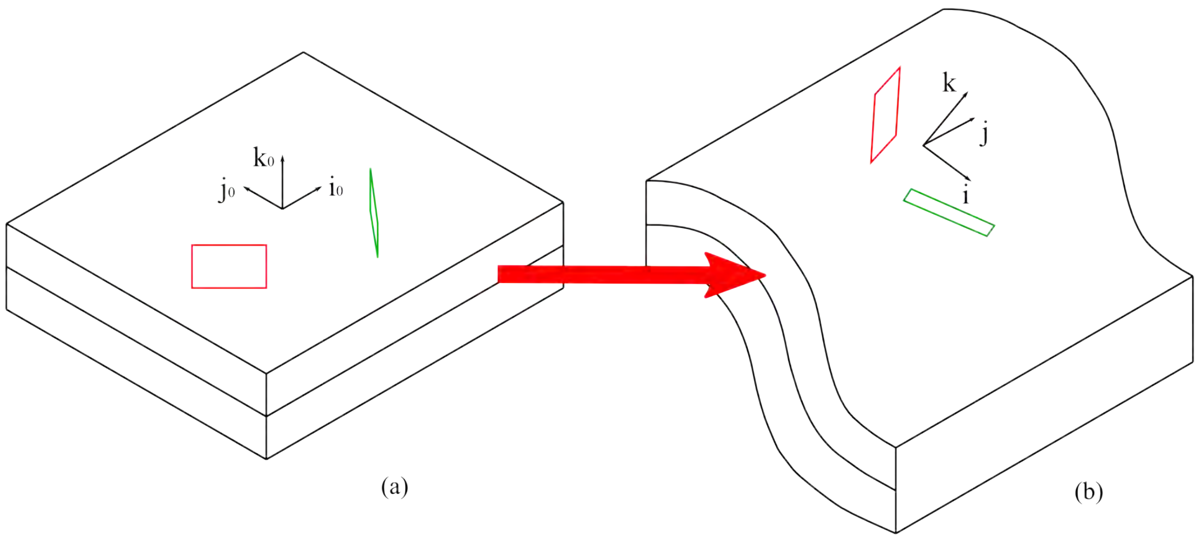


Figure 3.6: The original fractures (green and red objects) are first generated in an unfolded environment (a) and secondly rotated according to the local geological orientation (b), this is equivalent to a basis transformation from the initial (i_0, j_0, k_0) basis to the local (i, j, k) basis.

3.2.6 Discretization of the fractures in a regular grid

This section describes how the fractures that have been previously generated can be discretized in a regular grid in a simple and efficient manner. For each fracture f, an indicator function I_f is used. This indicator function can be used as input for any process requiring such format (Section 5.3.4). The computation of this discrete indicator function takes several distinct steps . First of all the centroids of the grid cells are computed. Three matrices $\boldsymbol{X}$,$\boldsymbol{Y}$, and $\boldsymbol{Z}$ contain the values of x, y and z respectively for all the elements of the model domain $\boldsymbol{\Omega}$:

$$\begin{array}{ll} \boldsymbol{X}(i,j,k) = o_x + d_x/2 + j d_x & \forall j \in [1, ..., n_j] \\ \boldsymbol{Y}(i,j,k) = o_y + d_y/2 + i d_y & \forall i \in [1, ..., n_i] \\ \boldsymbol{Z}(i,j,k) = o_z + d_z/2 + k d_z & \forall k \in [1, ..., n_k] \end{array} \tag{3.29}$$

For each fracture, the matrix indexes $\boldsymbol{i_b}$,$\boldsymbol{j_b}$, and $\boldsymbol{k_b}$ of the bounding box $\boldsymbol{\Omega_s}$ of the fracture are computed:

$$\left.\begin{aligned} i_{b_c} &= int\left(\frac{y_c - o_y}{d_y}\right) \\ j_{b_c} &= int\left(\frac{x_c - o_x}{d_x}\right) \\ k_{b_c} &= int\left(\frac{z_c - o_z}{d_z}\right) \end{aligned}\right\} \forall c \in [1, ..., 4] \tag{3.30}$$

These indexes are used to select $\boldsymbol{X_s}$,$\boldsymbol{Y_s}$, and $\boldsymbol{Z_s}$ which are the sub ensemble of grid cells that belongs to the bounding box $\boldsymbol{\Omega_s}$ limits:

$$\begin{aligned} \boldsymbol{X_s} &= X(\boldsymbol{i_b}, \boldsymbol{j_b}, \boldsymbol{k_b}) \\ \boldsymbol{Y_s} &= Y(\boldsymbol{i_b}, \boldsymbol{j_b}, \boldsymbol{k_b}) \\ \boldsymbol{Z_s} &= Z(\boldsymbol{i_b}, \boldsymbol{j_b}, \boldsymbol{k_b}) \end{aligned} \tag{3.31}$$

Then the equation of the plane is obtained by first computing the normal vector $\boldsymbol{n}$ which is given by the vectorial product of $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$:

$$\begin{aligned} \boldsymbol{v_1} &= \overrightarrow{\boldsymbol{x'_0 x'_1}} \\ \boldsymbol{v_2} &= \overrightarrow{\boldsymbol{x'_0 x'_2}} \\ \boldsymbol{n} &= \boldsymbol{v_1} \times \boldsymbol{v_2} \end{aligned} \tag{3.32}$$

and d is computed as:

$$d = -\boldsymbol{n} \cdot \boldsymbol{x_0}^{\intercal} \tag{3.33}$$

After that, the centroids $\boldsymbol{p} = \{\boldsymbol{X_s}, \boldsymbol{Y_s}, \boldsymbol{Z_s}\}$ of the subspace $\boldsymbol{\Omega_s}$ are projected on the plane of the fracture, with $\boldsymbol{x'_0}$ as new origin, i.e. the projected coordinates are expressed in plane coordinates local system (Figure 3.7):

$$\boldsymbol{p'} = \boldsymbol{p} - d_n \cdot \boldsymbol{n_0} - \boldsymbol{x'_0} \quad \text{with:} \quad \boldsymbol{n_0} = \frac{\boldsymbol{n}}{||\boldsymbol{n}||} \tag{3.34}$$

with d_n being the distance between the point and the plane. Then the projected points $\boldsymbol{p'}$ are expressed in the basis $(\boldsymbol{v_1}, \boldsymbol{v_2})$ in this form:

$$\boldsymbol{p'} = \theta_x \boldsymbol{v_1} + \theta_y \boldsymbol{v_2} \tag{3.35}$$

with θ_x and θ_y:

$$\begin{pmatrix} \theta_x \\ \theta_y \end{pmatrix} = \begin{pmatrix} \frac{\det(\boldsymbol{p'}, \boldsymbol{v_1})}{\det(\boldsymbol{v_1}, \boldsymbol{v_2})} \\ \frac{\det(\boldsymbol{v_2 p'})}{\det(\boldsymbol{v_1}, \boldsymbol{v_2})} \end{pmatrix} \tag{3.36}$$

Finally, the indicator function I_f for the fracture is defined as follow:

$$I_f(\boldsymbol{x}, f) = \begin{cases} 1, & \text{if } (\theta_x, \theta_y) \in [0,1] \times [0,1] \;\&\; d_n < max_d \\ 0, & \text{otherwise} \end{cases} \tag{3.37}$$

where max_d is the maximal distance allowed between the plane and the centroids (it must be given by the user).

While this method is more computationally demanding compared e.g. to the Bresenham algorithm (Bresenham, 1965), it has the advantage of allowing to better connect the discrete cells around the fracture. This is suitable in many applications, especially when using finite differences fluid solver, or the FMA algorithm (Sethian, 1996) as done in the next chapters (4 and 5). Moreover, when using a very fine grid, it is also possible to take into account for the aperture of the fractures, by defining a projection distance depending on the aperture.

3.3 Examples of application

This section shows the results of the proposed methodology on two 3D examples. The first one is a synthetic 3D model representing an anticline, with all the theoretical families of fractures. The second is an application of the model, based on field observations, to the Valley of *les Ponts-de-Martel* study site (Chapter 2).

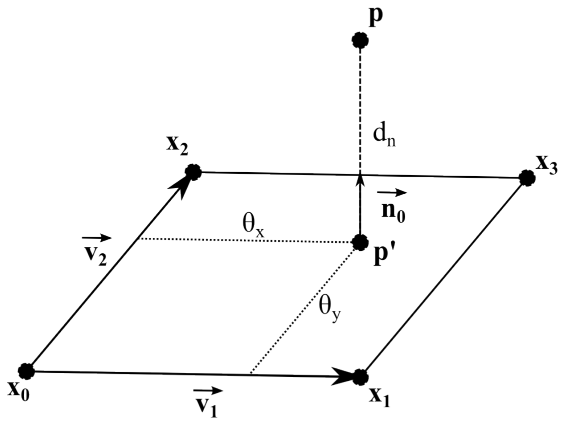


Figure 3.7: The centroids $\boldsymbol{p}$ of the surrounding elements are projected on the plane of the fracture, and expressed in the new basis $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$ by their θ_x, θ_y coordinates.

3.3.1 Synthetic case

For this synthetic example, we consider a simple anticline and 2 geological formations. The upper one is considered to be a brittle hard-rock. The second one is a ductile rock that has a more plastic behavior, and it is not fractured.

The dimensions of the synthetic model are 1000m x 1000m x 1000m. In the brittle formation (not shown in the figure) we modeled all the families of fractures proposed by Twiss and Moores (1992) and listed in section 3.2.4. For a purpose of visibility we modeled only the fractures whose length are between 25m and 100m. One realization of this fracture model is presented in figure 3.8.

As it can be seen, all the families of fractures keep their relative orientation to the folding direction. This is satisfactory for every direction of folding, except when the geological gradient is exactly vertical, in this case the fractures may be rotated in any direction, as the information provided by the direction of the gradient is insufficient to avoid confusion.

In the additional material of the online paper, the reader will find a 3D pdf file, which can be opened with adobe reader, and contain the 3D model. Every family of fractures is a layer of the 3D pdf file, so it is possible so explore it more significantly than the printed figure itself. In the 3D pdf, the name of every family is characterized by *S***D***, where *S**** means *Strike degrees* and *D** Dip degrees.*

3.3.2 Application to the Valley of *les Ponts-de-Martel*

To use fracturation data extracted from field data, we measured the fractures on the field, and also the orientation of the stratigraphy (dip and strike). Successively we computed the relative orientation of the fractures with respect to the fold orientation, retrieved the families that had to be modeled, and used these new statistics to inform the fracture generator and build only the fractures families observed in the field as explained in section 3.2.3.

Observed families of fractures

The fracturation in three sites has been measured in the Valley of *les Ponts-de-Martel* (Fournier, 2012). Two scanlines have been interpreted near to *Les Ponts-de-Martel* (Figure 3.9 F2 and F3) and a new tunnel for the water supply of the city of *La Chaux-de-Fond* (Figure 3.9 F1). These data have been *unfolded* in the way described in Section 3.2.3. The observed fracture families in the reference system (i_0, j_0, K_0) are shown in Table 3.1. These are used to build the fracturation model presented in Figure 3.10. In this example, the theoretical families of fractures have not been included in the model.

3.4 Discussion and perspectives

The use of an implicit model, instead of a geomechanical one, gives promising results while avoiding the computational load of geomechanical models. But this results in a probably over-simplified DFN.

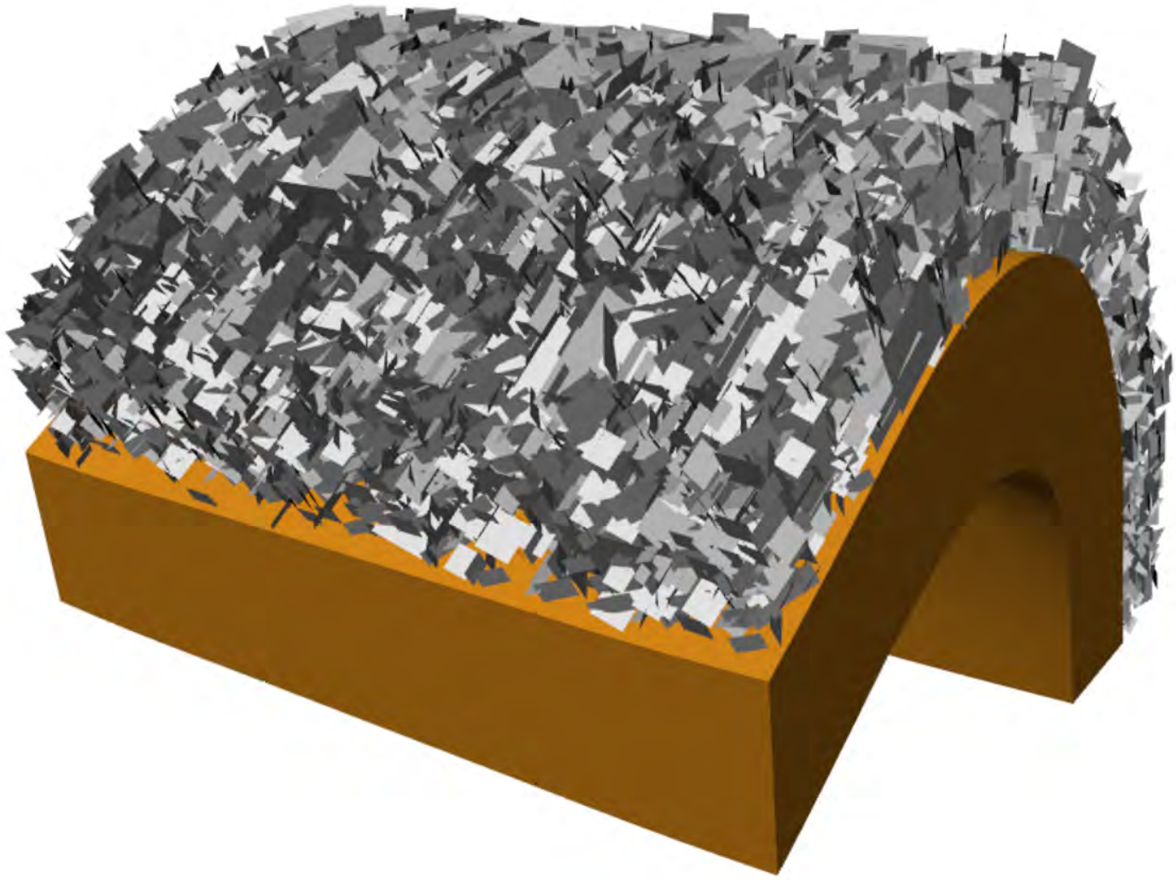


Figure 3.8: 3D view of the synthetic example. all the families of fractures are present in the fracturable formation (which is not shown). In the additional material a 3D pdf is available for this figure, the reader will find a specific layer for every family of fractures.

As it has been shown in the literature (e.g. Mace, 2006; Bonneau et al., 2013), geomechanical control over the fracture growth, relative position, and interaction are very important. In our approach, there are no such rules that control the generation of fractures. Moreover the UFM model is applied in a consistent manner in this approach. Only the statistics about the fracture length distributions are used, but the interactions between the fractures are not modeled as opposed to the "true" UFM approach. One important rule, in order to develop more realistic fracture networks would be to generate first the largest fractures, and successively the smallest, paying attention that the smallest ones should stop on the largest. This could be achieved e.g. by making the fractures grow iteratively in a similar way as Bonneau et al. (2013), who developed pseudo-genetic DFN algorithm, where the fractures grow iteratively, and the complex interactions of the fractures terminations are considered. Further studies may be required to investigate if the fracture orientation within a folded structure may vary according to the stratification orientation within a single fracture plane, or if its orientation depends on the orientation at the fracture center.

The information provided by the implicit model (potential, gradient, but also curvature, etc) could also be further exploited. For example, the different zones of the fold (intrados, extrados, limbs) could be extracted, and then only the fractures relative to the position within the fold could be generated. This information could be extracted by analysis of the relative change of orientation of the fold: locally the gradient may vary a lot (intrados) or not at all (limbs), and in the middle we find the extrados. This could be achieved by differentiating the different zones with two variables: the value of the potential, which indicate in which part of the formation the fracture is localized, and the second derivative of the isohypses, whose sign should indicate if it is an extension or a compression region in the model. This could be an interesting starting point for further developments. The initial PDF of fracture within the

	Families	N	strike [°]	Dip [°]
Tunnel	ac I	22	089	84
	ac II	23	100	89
	bc I	23	015	76
	bc II	31	351	79
	C1b	19	311	81
	C2b	17	038	75
	strati	160	000	00
	random	72	-	-
Scanlines	ac I	21	083	86
	ac II	16	100	90
	C2a I	7	337	80
	C2a II	10	137	87
	strati	74	000	00
	random	39	-	-

Table 3.1: Orientation of the fractures observed at the Valley of *les Ponts-de-Martel*. These data are expressed in an *unfolded* referential (Section 3.2.3). N: number of observations

fold could also be extracted by the relative differences of the gradient orientation (curvature). In this way we could assert that the most folded zones are the most fractured.

3.5 Conclusions

In this chapter we propose a methodology and a code that allows to use the information provided by an implicit geological model about the deformation within a folded structure to generate realistically oriented discrete fracture networks (DFN). The methodology is simple: the gradient of the potential is extracted from the implicit model. It is used to rotate the fractures within the fold. The fracture families can either be taken from a theoretical model such as Twiss and Moores (1992) or provided by the user in the undeformed space. Some tools are provided to transform the field data in such undeformed space. For the length distribution, the code uses the UFM model (Universal Fracture Model) proposed by Davy et al. (2010). The resulting DFN models can be rasterized if needed for further use in flow, transport or karst modeling applications.

The proposed code could be further improved by adding “geomechanical rules” to the DFN that are generated with this approach, of better accounting for the position of the fractures within the structure. It would be useful to be able of simulating the different theoretical fracture families in consistency with the zone of the fold in which they occurs (compression in the intrados and extension in the extrados).

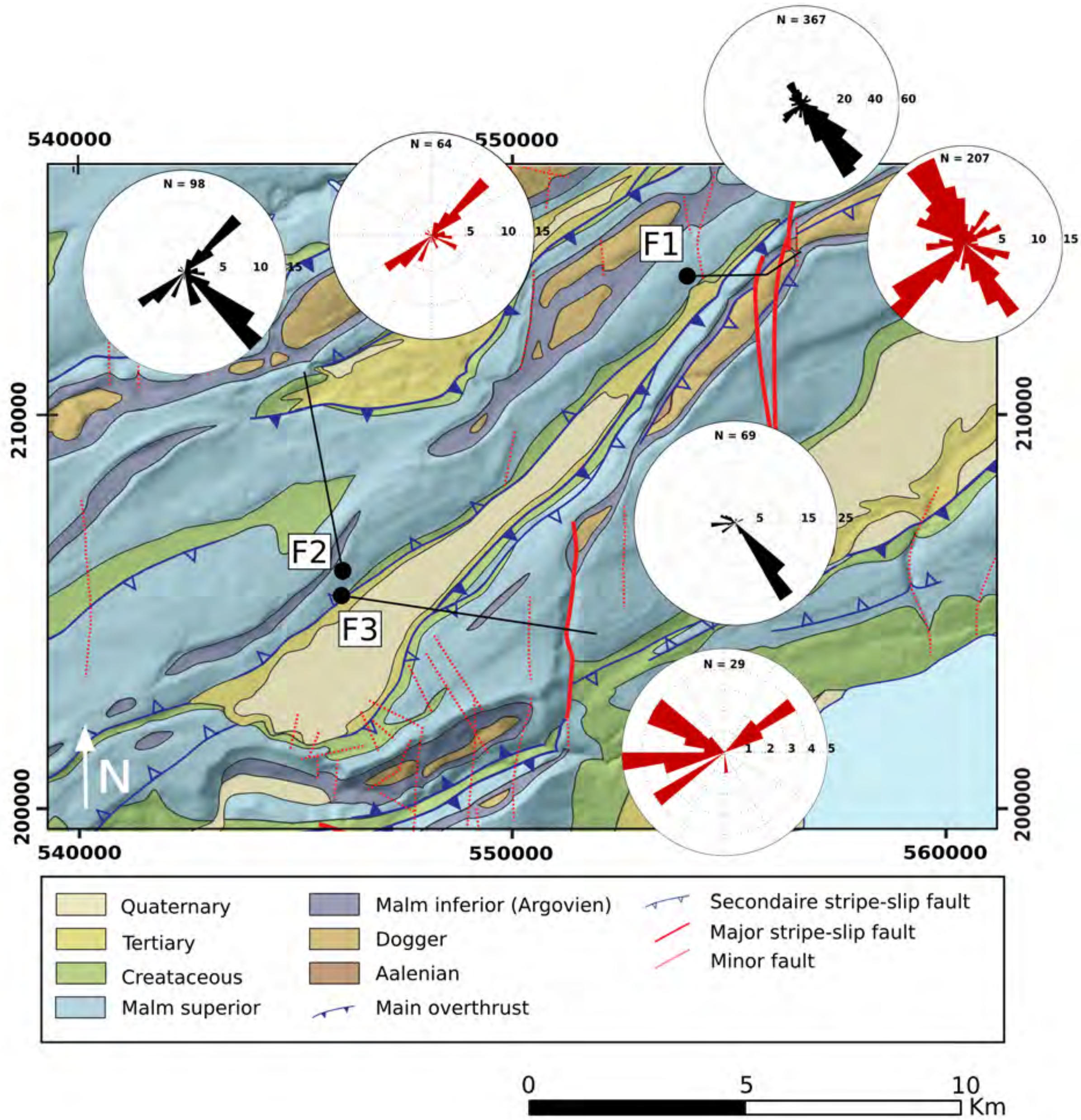


Figure 3.9: Observed fracture families in the Valley of *les Ponts-de-Martel*. F1: tunnel of *La Corbatière*. F2 and F3: scanlines. In black: bedding planes and fractures; red: only fractures. (modified from Fournier, 2012)

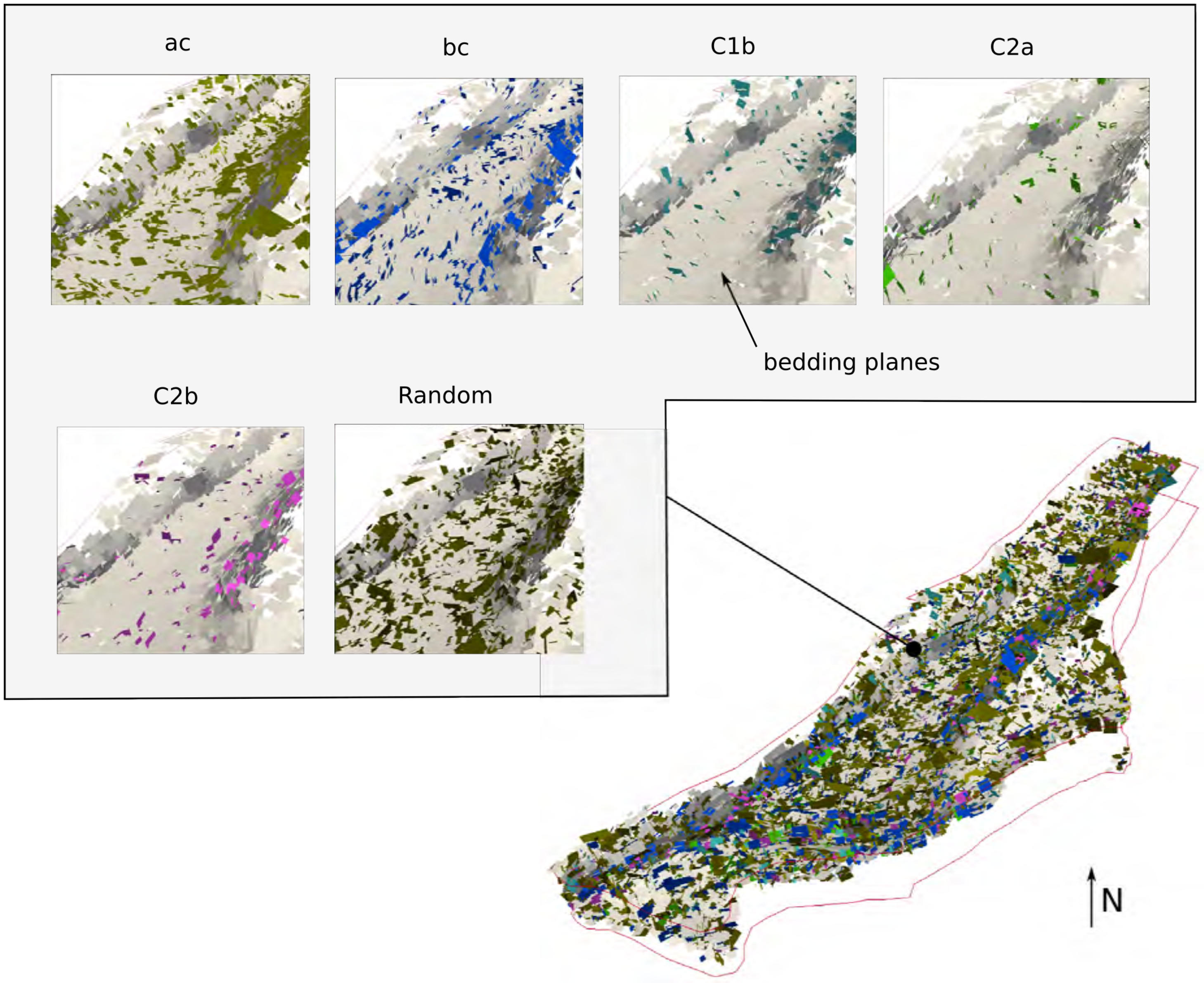


Figure 3.10: Fracturation model of the Valley of *les Ponts-de-Martel*. (modified from Fournier, 2012)

Chapter 4

A pseudo-genetic stochastic model to generate karstic networks [1]

[1]This paper is published as: *Andrea Borghi, Philippe Renard, and Sandra Jenni, 2012, A pseudo-genetic stochastic model to generate karstic networks. Journal of Hydrology, 414-415(0):516-529, ISSN 0022-1694, doi: 10.1016/j.jhydrol.2011.11.032.*

Abstract

In this paper, we present a methodology for the stochastic simulation of 3D karstic conduits accounting for conceptual knowledge about the speleogenesis processes and accounting for a wide variety of field measurements. The methodology consists of 4 main steps. First, a 3D geological model of the region is built. The second step consists in the stochastic modeling of the internal heterogeneity of the karst formations (e.g. initial fracturation, bedding planes, inception horizons, etc). Then a study of the regional hydrology / hydrogeology is conducted to identify the potential inlets and outlets of the system, the base levels and the possibility of having different phases of karstification. The last step consists in generating the conduits in an iterative manner using a fast marching algorithm. In most of these steps, a probabilistic model can be used to represent the degree of knowledge available and the remaining uncertainty depending on the data at hand. The conduits are assumed to follow minimum effort paths in a heterogeneous medium from sinkholes or dolines toward springs. The search of the shortest path is performed using a fast marching algorithm. This process can be iterative, allowing to account for the presence of already simulated conduits and to produce a hierarchical network. The final result is a stochastic ensemble of 3D karst reservoir models that are all constrained by the regional geology, the local heterogeneities and the regional flow conditions. These networks can then be used to simulate flow and transport. Several levels of uncertainty can be considered (large scale geological structures, local heterogeneity, position of possible inlets and outlets, phases of karstification). Compared to other techniques, this method is fast, to account for the main factors controlling the 3D geometry of the network, and to allow conditioning from available field observations.

Keywords: karst, conduit network, stochastic simulation, geological modeling, reservoir modeling, fractures modeling, process-based modeling, pseudo-genetic approach, fast-marching algorithm (FMA)

4.1 Introduction

Karst formations are estimated to cover about 20% of the land surface of the world (Ford and Williams, 2007) and 35% of Europe (Biondic et al., 1995). According to these last authors, they constitute the main water resource in Europe and have therefore a crucial importance. However, they are difficult to characterize because of their extreme heterogeneity induced by the presence of sparse conduits embedded in a fractured carbonate matrix. Worthington et al. (2000) reviews of 4 typical karst aquifers in which more than 90% of the flow occurs in the conduits while more than 90% of the water is stored in the fractured matrix. It is therefore crucial to account both for the matrix and the conduits when modeling these aquifers for groundwater protection, management, or to investigate contamination issues. Including the conduits in the formulation of the numerical models has been pioneered by Kiraly (1979; 1988) and is still a topic of active research (Cornaton and Perrochet, 2002; Eisenlohr et al., 1997; Hill et al., 2010; Jourde et al., 2002; Sauter et al., 2006). These authors provide various numerical techniques to account for the conduits and discuss the different types of flow and transport equations that can be used in the conduits and in the matrix.

In this paper, we take a different perspective. Indeed, a major difficulty when modeling karst aquifers is that the location of the conduits is often unknown since they are seldom hit by boreholes and are difficult to detect by geophysical techniques when their diameter is small. We therefore propose a method to model the possible occurrence of the conduits. Once the conduits are modeled, they can be used as input for flow and transport models. Modeling the possible occurrence of the conduits can also be used directly to assess the risk of conduit occurrence in mining or geotechnical projects involving for example tunnels or dams (e.g. Filipponi et al., 2010). All these potential applications will not be discussed in detail or tested by the present paper.

So far, only a few techniques have been developed to model the structure of karst conduit networks. The most advanced ones are based on modeling the physics and chemistry of the speleogenesis processes (Dreybrodt et al., 2005; Kaufmann and Braun, 2000). This approach allows understanding the kinetics of karst evolution, and give precious information about the speleogenesis processes. However it is difficult to apply to model in order to characterize an actual system, mainly because it is difficult to reconstruct the paleoclimatic and geological conditions that were prevailing during its formation. Consequently, these models are extremely difficult to condition to field observations. In addition, they are very CPU demanding, because they are based on the numerical solution of highly non linear and coupled systems of equations. For this reason, these models are usually not applied in a Monte-Carlo framework allowing site characterization and risk analysis. Another approach is to build a purely statistical model. For example, Henrion et al. (2008) and Henrion et al. (2010a) propose to model complex cave geometries by generating first a stochastic discrete fracture network and then combining it with a truncated multi-Gaussian field in order to represent the 3D geometry of the caves around the conduit network. An intermediate type of model is proposed by Jaquet et al. (2004). They use a modified lattice-gas automaton for the discrete simulation of karstic networks. By this approach, walkers are sent through the medium and erode it progressively. The hierarchical structure is modeled by attracting the walkers by the conduit system simulated at the previous iteration. This technique is also very difficult to condition to field data because of the random walk technique. Ronayne and Gorelick (2006) propose to simulate branching channels (analogous to karst conduits) using a nonlooping invasion percolation model. This method is used to investigate the effective property of media presenting such type of structures but it is not designed to model a specific site and to be conditioned by field observations (known presence of conduits in boreholes, dolines, sinkholes, etc).

In this paper, we propose a new methodology that aims at simulating the geometry of karstic conduits at a regional scale constrained by the regional geology, by the regional hydrological and hydrogeological knowledge, and by a conceptual knowledge of the genetic factors (like e.g. fracturation, bedding planes) that control the conduit genesis, without solving the physical and chemical equations of the speleogenesis. The motivation is to include those processes in a simplified manner to ensure that the modeled geometry is realistic while being sufficiently numerically efficient to allow a stochastic analysis. The proposed methodology is based on the assumption that water and karst conduits will follow a minimum effort path which is computed using a fast-marching algorithm (Sethian, 1996). The generation of the conduits is performed in a stochastic framework allowing to generate multiple equiprobable realizations that are all conditioned by field observations (dolines, sinkholes, presence of caves in boreholes, known interconnected springs and sinkholes, etc.).

After describing the principle of the methodology in section 4.2 and its implementation in section 3, the paper presents two illustrations of the method. One in a coastal aquifer along the Mediterranean coast and one for the Noiraigue spring catchment in the Swiss Jura Mountains, close to the city of Neuchâtel. This last example has been chosen because an important data set is available from previous studies (Atteia et al., 1998; Gogniat, 1995b; Morel, 1976) and because its structural geology is complex (Valley et al., 2004).

4.2 Overview of the approach

Before describing the details of our approach, we provide an overview of the underlying conceptual model and of the rationale behind the proposed methodology. Different conceptual models have been developed in the last century to describe the factors that control the formation of karst structures; they are reviewed by Ford and Williams (2007). According to Klimchouk (2007), one can distinguish two main types of karsts. Hypogenetic karst aquifers are formed under deep conditions mainly by hydrothermal waters while epigenetic karst aquifers are formed under surface conditions by the circulation of rainfall water.

The methodology discussed in the present paper focuses on epigenetic karsts. They can be divided in 3 distinct zones. The first is the unsaturated (vadose) zone. It can be very deep for mature karst aquifers (according to Ford and Williams, 2007, several hundreds of meters). In this zone, karst conduits are mainly subvertical because gravity is the most influent factor. Then these subvertical conduits reach the phreatic (by definition saturated) zone. They become more and more concentrated leading to a hierarchical structured network (Jeannin, 1996) which leads toward one or more springs. The epiphreatic (intermittently saturated) zone is located between the vadose and phreatic zones. It is the zone which is saturated under high flow conditions. It can be formed by old phreatic conduits or by subvertical vadose conduits (Ford and Williams, 2007). Because of their organized structure, the connectivity of the karst conduit network is very strong, and therefore crucial to understand the flow and transport phenomena. Furthermore karst aquifers have a superficial zone, called epikarst (a few meters below topographic surface). This superficial zone collects rainfall water and redirect it toward the vadose conduits. Typically, dolines are points where epikarstic water is concentrated in a preferential inlet.

The epikarstic layer will not been taken into account in this methodology, because it focuses on the geometrical modeling of the main karst conduits. In epigenetic karst aquifers, the density and geometry of the karst conduits depend on two main factors. The first is the geology that controls their shape and distribution: conduits are formed by dissolution. This process is predominant in limestones and

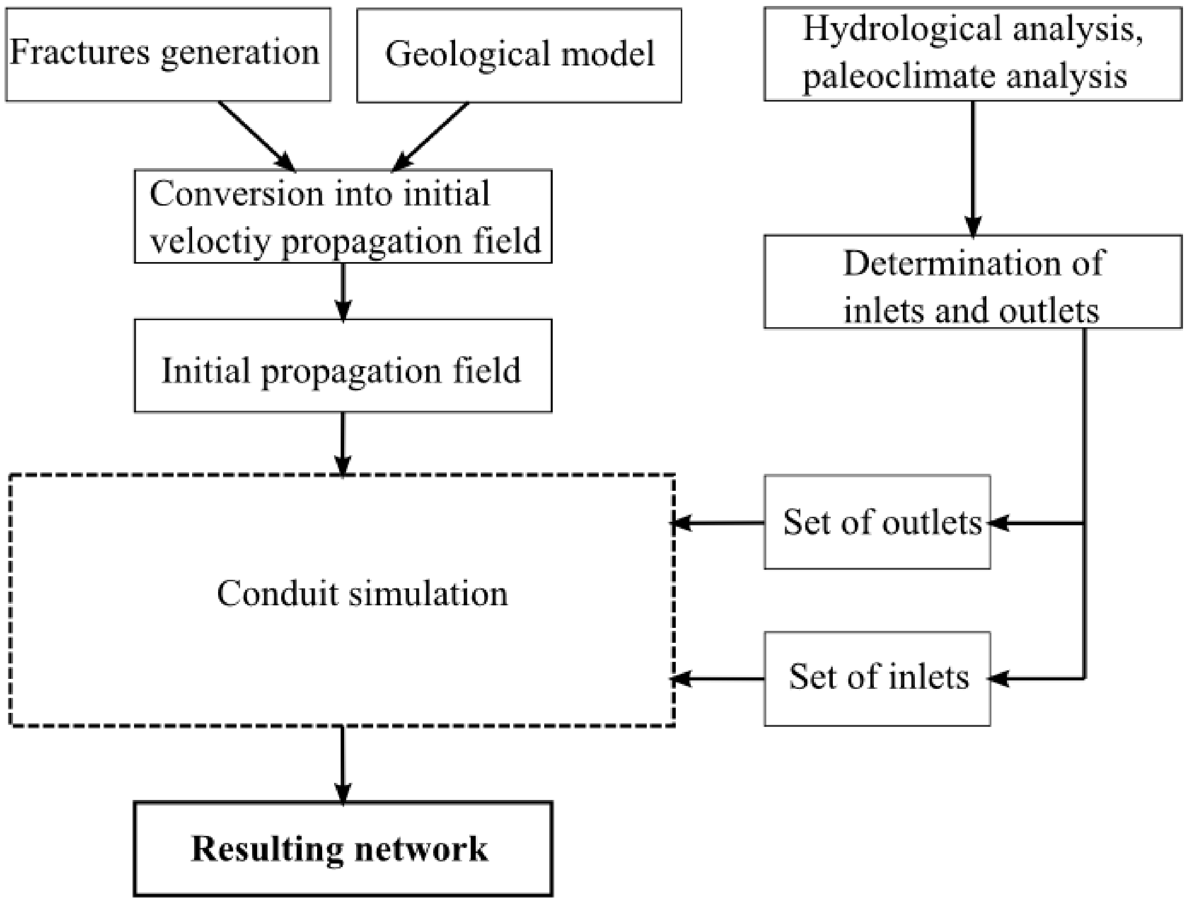


Figure 4.1: workflow of the complete algorithm, the input are the fracture model, the geological model and the hydrogeological analysis. After a preprocessing step they pass through the conduit simulation box (see Figure 4.1) and finally the resulting network is obtained.

other highly soluble rocks like evaporites and are preferentially located along bedding planes and/or initial fracturation (Filipponi et al., 2009; Palmer, 1991; Worthington et al., 1999). The second factor is the hydrology that controls recharge and the position of recent or paleo base levels (Bakalowicz, 2005). These two factors interact during the formation of the karst network leading a final distribution of conduit orientations in 3D that is different from the orientations of the fractures alone (Király et al., 1971). The simulation of the speleogenesis processes has shown in addition (Dreybrodt et al., 2005) that the increased conductivity of the conduits, that are formed progressively during the evolution of the system, re-orient the flow field and the preferential areas of dissolution leading then to a hierarchical structure. Overall, to account for all the constraints described above in a simplified manner, the proposed methodology consists of 4 main steps (Figure 4.1): modeling the regional geology, modeling the fracturation and bedding planes within the regional geology, analyzing the hydrological control, and finally modeling the conduits. The concepts underlying these 4 steps and the choices that have been made are presented in the following subsections while their implementation is described in section 4.3.

4.2.1 Regional Geology modeling

The first major control on the formation of karstic network is the regional geology (i.e. the location of the aquitards and carbonate formations). The first step consists therefore in modeling the regional scale structures, i.e. modeling the sedimentary and tectonic structure. The extension of the model has to be decided carefully. It must take into account both the topographical and hydrogeological catchment. Furthermore, it is better to extend the model to take into account for possible complex 3D flow paths.

As a side remark, the geological model can be used to estimate more accurately the catchment extension in surface (Borghi, 2008). For example, if an aquitard formation underlies a karstic formation, the axis of the anticline on the top of the aquitard can be used to draw a reasonable estimation of the boundary of the catchment.

During that step, an accurate study of the stratigraphic or sedimentary units is fundamental because the conduits will develop only in carbonates and other highly soluble formations. Once this information is acquired, it allows defining how the stratigraphic units will be grouped. The aim is to construct a geological model of the hydrogeological formations based on their hydrogeological properties rather than their age.

Once the hydro-stratigraphy is defined, the geological structure has to be modeled because the karstic network will be strongly constrained by the structure (folds, faults and thrusts). A good structural interpretation is therefore a key prerequisite to build a realistic 3D geological model.

From a technical point of view, different 3D geological modeling techniques and software are available (Calcagno et al., 2008; Caumon et al., 2009; Lajaunie et al., 1997; Mallet, 2002). The important point is that one should select a technique which ensures that the modeled volumes are coherent. Models that are only based on the 3D modeling of the surfaces bounding the formation may not ensure this coherence if they do not account for topological constraints. They may for example not ensure the exact definition of closed volumes and will not be applicable in the following steps of the methodology. This is particularly important in highly deformed areas such as the the Vallée des Ponts-de-Martel" (section 4.4.2) where thrusts, folds and faults can create hydraulic contact or barrier between 2 different aquifers that would normally be separated by an aquitard. In summary, one should select a 3D modeling technique allowing to account for the degree of complexity of the geological structure present in the studied site while ensuring the 3D volume coherence.

4.2.2 Fracturation and bedding planes

As described in Filipponi et al. (2009), the karst conduits shape and density are strongly influenced by the properties of the geological formations in which they develop. These characteristics, fracturation and bedding planes, can be inferred from borehole data (for deep reservoirs) or outcrop analysis when they are expected to be representative of the fractures at depth (Sharp et al., 2006).

Fracturation

The fracturation of the rock influences the orientation of the conduits and their path (Palmer, 1991). From a geometrical point of view, a fracture network is usually decomposed in a number of fracture families. Each family is then described statistically by a density or spacing, ranges of orientation, length or size distribution, relationship between the fracture density and stratigraphy. The types of geometrical relations between the fracture families can also be important (Kaufmann, 2009). All those properties control the connectivity of the system and have a strong influence on the initialization of the karst conduits (Ghosh and Mitra, 2009). In addition, according to (Gillespie et al., 1993) two additional characteristics influence the flow properties of the medium: fracture roughness and fracture aperture.

Once the field data have been acquired, a classical approach is to use stochastic discrete fracture network models. Different methods have been developed, from the most simple Boolean model of uniformly distributed fractures of identical size (monodisperse case) to complex models allowing the inclusion of several fracture families, variable density in space and various types of relations between the families (Bour et al., 2002; Castaing et al., 2002; Gringarten, 1998). Among the most advanced techniques, some authors propose to estimate the local density and orientation of the fractures by using numerical codes allowing to compute rock deformation and geomechanical stresses (Camac and Hunt, 2009).

Inception horizons

Filipponi et al. (2009) shows that particular horizons within a sedimentary formation have a very strong importance during the activation process of the cave genesis. These horizons, called inception horizons, present either a particular capability to be dissolved or a specific contrast with the surrounding rock (chemical, mineralogical or rheological). The inception horizons are therefore specific stratification joints or bedding plane. As a result, like for the fracturation, the location of the inception horizons or stratification joints within a karstifiable formation is a key heterogeneity factor that controls the geometry and position of the conduits. They must therefore be accounted for in the model either in a statistical or deterministic manner. In our approach, the location of the inception horizons is inferred from the geological model. They are considered to be parallel to the main sedimentary structure. In the field, the location of these horizons can be identified and measured along outcrops.

4.2.3 Hydrodynamic control

While the two factors described above are related to the static structure of the medium, the third main factor controlling the geometry of a Karstic network is related to the hydrological and (paleo) climatic situation at the site (see classification by Palmer, 1991). The amount of recharge controls the degree of karstification (Bakalowicz, 2005) and its spatial distribution controls the type of karstic network that can be expected. The presence of a base level controls the position of the outlets and the limit between the vadose and phreatic zones. In presence of seawater, the position of the interface with freshwater will be a zone of preferential dissolution due to chemical inequilibrium resulting from the mixing of fresh and salt waters (Gabrovsek and Dreybrodt, 2010, p.457). All these constraints can change in time and space leading to different stages of karstification. Consequently a good knowledge of the paleo and actual conditions is required to understand and simulate the different stages of karstification.

The information concerning the hydrodynamic control can be obtained from classical hydrogeological analysis, tracers tests, recharge and geomorphological studies. At the end of this analysis one should know: the extension of the catchment, the location of the actual and possible paleo outlet(s), the type of recharge (diffuse - punctual), the map of the zones of recharge (known point, possible points, and diffuse areas) and the overall water balance of the catchment basin.

Hydrogeological analysis

In that part of the work, the aim is to collect all the information available related to the position of the known inlets and outlets in order to define where they are located and how they should be modeled (punctual or diffuse).

The paleo (several hundreds of thousands of years, i.e. glacial periods) recharge and base level conditions should be estimated as well, at least approximatively, in addition to the actual ones. This allows to identify the possible existence of different stages of karstification and to model them. For example, the "Sieben Hengste" karstic network was formed at least under three different hydrological conditions Jeannin (1996). The change of the regional hydrologic base level has then triggered the generation and reactivation of different families of conduits resulting in a complex structure that cannot be understood without knowing these different stages (Filipponi et al., 2009).

Recharge analysis

In the proposed procedure, the water balance of the catchment must be analyzed for two reasons. First, the water balance of the catchment allows to check whether the size of the catchment is in agreement with the total discharge at the outlets. On the other hand, the water balance can also be used to check whether all the outlets of a given catchment have been identified. If there is more inflow than outflow, one can be sure that there are other outlets which may be hidden or diffuse and more hydrogeological research should be done to clarify the situation before starting to make a model. There is a second reason to analyze recharge: it is that the type of networks that are expected at a given location will be different depending on the type of recharge. Palmer (1991) considers mainly 3 different types of recharge: via karst depressions, diffuse (through porous rocks) and hypogenic (acids). In the first case one may expect a branchwork type of network, in the second case it should be an anastomotic or network maze, while on the third case one would expect more a spongework maze or ramiform pattern. Identifying the different types of recharge can be based on field observations and supported by water balance calculation. Numerical modeling can help in this process. For example, Borghi (2008) used a finite element method to solve the steady state flow based on mean annual recharge values using an equivalent hydraulic conductivity for the different hydrogeological units (based on a 3D geological model). This allowed identifying the different recharge/discharge zones of the Noiraigues catchment. He showed in that case that a significant proportion of the recharge (30%) was diffuse trough the sandstone aquitard located over the karstic aquifer.

Geomorphological analysis

While the most important inlet and outlet are often known from field observations and their relation confirmed by tracer tests, many smaller inlets are not known precisely. Still some must be modeled and their location should be estimated. This is why we propose to complement the classical hydrogeological analysis with a geomorphological study. The underlying concept is that punctual inlets (dolines and sinkholes) create geomorphological depressions; an accurate geomorphological analysis of the landscape allows identifying them. This can be partly automated using high resolution digital elevation model and numerical algorithms (Lamelas et al., 2007), but it needs to be controlled in the field if possible.

Tracer tests

Tracer tests are often used in karst studies Käss et al. (1998); Leibundgut et al. (2009). Here we use their results in a two manners. First, they are used to delineate the extension of the catchment by showing which inlets are connected to the outlets of interest. Second, this connectivity information is used during the simulation process to ensure that points known to be connected from field observations will also be connected in the model. The model should however be flexible enough to account for the fact that the same sinkhole can be connected to many springs and vice versa (i.e. spring connected to many sinkholes).

4.2.4 Karst conduit simulation

The final step of the proposed methodology is the simulation of the karstic network. The method is inspired by speleogenetic process based models Dreybrodt et al. (2005); Kaufmann and Braun (2000). In order to accelerate the simulation, we use a proxy allowing to mimic the final result (mature system) without modeling the physical and chemical processes.

The proxy is based on the assumption that water, and thus karst conduits, will follow the minimum effort path, i.e. the path that is the most convenient for flow between inlets and outlets. In a homogeneous medium, it would be a straight line between sinkholes and the spring. In reality, however, it depends on the pre-existing geological heterogeneity and follows more complex paths.

The computation of these paths is based on a Fast-Marching Algorithm (FMA) (Sethian, 1996, 1999, 2001). We assume the karst spring to be the starting location of the FMA. The result is a map of the minimum time required to reach the spring. We then we perform a particle tracking starting from the sinkholes (or dolines) toward the spring. The particle follows the shortest path and "corrode" the model in a manner similar to Jaquet et al. (2004). The resulting path becomes one of the karst conduits. Section 4.3.4 explains the implementation in detail.

4.3 The proposed algorithm and its implementation

This section describes in detail the four main steps of the methodology: geological modeling, heterogeneity modeling, hydrogeological analysis and simulation of karst networks. We use two 2D synthetic cases (one vertical cross-section and one horizontal map) for illustration. The simulation framework is composed of a deterministic and a stochastic component. To differentiate them, we use the superscript "S" to denote the variables that are realizations of a stochastic process.

4.3.1 3D geological model

The first step of the methodology is to build a 3D geological model of the catchment area. Mathematically, and in a very general manner, we consider that the 3D model is represented by a set of indicator functions allowing to know which is the geology present in any position $\boldsymbol{x} = (x, y, z)$

$$I_G(\boldsymbol{x}, g) = \begin{cases} 1, & \text{if geology is present at location } \boldsymbol{x} \\ 0, & \text{otherwise} \end{cases} \tag{4.1}$$

Where $g = 1, ..., N_g$ is the geological facies, with N_g the number of lithological facies. This function can be obtained from a wide variety of interpolation methods and modeling tools. In the example and the case study that are discussed in section 4.4, we used the software Geomodeller3D (Calcagno et al., 2008). It is based on an implicit method (Lajaunie et al., 1997) particularly well suited for the complex geology of the Jura because it allows accounting for the structural data (strikes and dips) not necessarily measured along a geological interface.

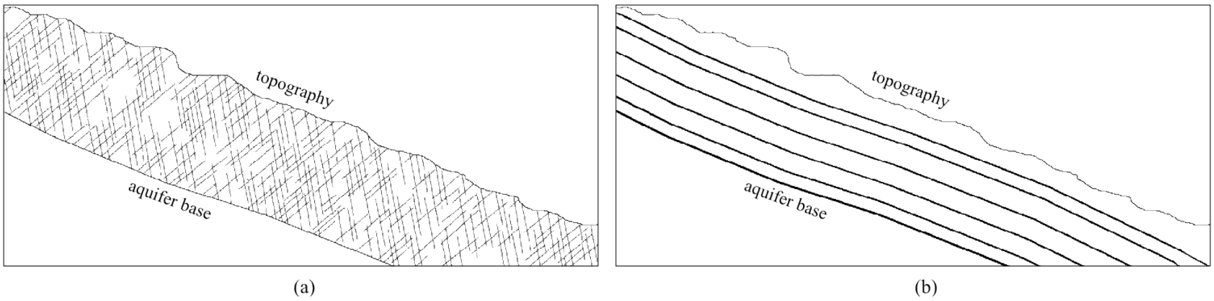


Figure 4.2: 2D cross-section: a) simple fracture model with 2 families of fractures, the fracture model is applied only in the karstifiable formation, b) bedding planes in the karstifiable formation

4.3.2 Internal structure of the carbonates

Here, the aim is to obtain an indicator function of the internal heterogeneities of the carbonates due to fracturation and bedding planes. As for the geology, we define a set of indicator functions.

$$I_F(\boldsymbol{x}, f) = \begin{cases} 1, & \text{if fracture } f \text{ is present at location } \boldsymbol{x} \\ 0, & \text{otherwise} \end{cases} \tag{4.2}$$

Where $f = 1, ..., N_f$ is the fracture or bedding plane family, N_f the number of codes defining which class of heterogeneity is present and the superscript S means that it is an equiprobable realization of a stochastic model (Figure 4.2). There are many ways to generate the fracture network. In the following examples, we use a simple Boolean object model for the fractures and rasterize it on a regular rectangular grid. The fracture model could be much more complex, but this would not change the overall methodology. More precisely, we assume that there is a finite number of fracture families. For each fracture family, the parameters of our model are a probability of occurrence (density), and a range of dip, strike, and fracture length. In practice, the coordinates of the centroid of every fracture is randomly sampled in a uniform distribution within the 3D model. The fracture length is uniformly sampled between a minimum and a maximum values as well as the dip and the strike. The number of families and their statistical parameters are provided by the user.

Furthermore, we consider bedding planes according to Filipponi et al. (2009) as inception horizons. These horizons are considered in the same way as fractures, i.e. as cells with increased permeability. They can also be modeled stochastically. In this case, they are assumed to be the surfaces of iso-thickness dividing the karstic formation. Their value of iso-thickness is sampled randomly as a relative thickness of the karstic formation.

4.3.3 Hydrogeological analysis

Determining the inlets and outlets of a karst aquifer is a key step in the procedure because the karst conduits will be constructed (section 4.3.4) as the paths connecting the inlets and outlets.

Inlets of the system

A hydrogeological analysis of the system must be carried out to identify the inlets of the system (dolines, sinkholes). These points are located on the topography. Among them some are known from the hydrogeological studies: main sinkhole for example or points identified by tracer tests. Their location is deterministic. However, we consider that there is also a number of potential inlets on the topography whose location has not been confirmed by field observations. For those points, we use a stochastic point process model. The set of N_D deterministic inlets is denoted by $\Im_D$.

$$\Im_D = (x_i, y_i, z_i) \quad \text{with} \quad i \in (1, ..., N_D) \tag{4.3}$$

For each realization, the set of N_D potential inlet locations $\Im_P^S$ is generated randomly based on a map of probability of occurrence $Pi_{\Im}(x, y)$ derived from field observations and from analysis of the recharge process (section 4.2.3).

$$\Im_P^S = (x_i, y_i, z_i) \quad \text{with} \quad i \in (1, ..., N_P) \tag{4.4}$$

Finally the complete set of inlets $\Im^S$ for each realization S is the union of $\Im_D$ and $\Im_P^S$

$$\Im^S = \Im_D \cup \Im_P^S \tag{4.5}$$

Furthermore, we mark each inlet point with an importance factor $\omega_i^{\Im}$. It is proportional to the upstream drained area C_i computed using the Digital Elevation Model of the area and the matlab TopoToolbox (Schwanghart and Kuhn, 2010). We then define the importance factor for every single inlet $\omega_i^{\Im}$ as the percentage of the total catchment area C_{tot} that it drains.

$$\omega_i^{\Im} = \frac{C_i}{C_{tot}} \quad \forall \boldsymbol{i} \in (1, ..., N_{\Im}) \mid N_{\Im} = N_P + N_D \tag{4.6}$$

Outlets of the system

The second aim of the hydrogeological analysis is to identify the major outlets of the system (springs and spring zones). As well as the inlets, the deterministic outlets will be identified as a set of O_D points defined as follows:

$$O_D = (x_i, y_i, z_i) \text{ with } i \in (1, ..., N_{OD}) \tag{4.7}$$

And the set of N_{OP} potential outlets O_P^S is generated randomly based on a probability map of occurrence $Pi_O(x, y)$ derived from field observations and from topographical analysis of the catchment.

$$O_P^S = (x_i, y_i, z_i) \text{ with } i \in (1, ..., N_{OP}) \tag{4.8}$$

Finally the complete set of outlets O^S for each realization S is the union of O_D and O_P^S

$$O^S = O_D \cup O_P^S \tag{4.9}$$

4.3.4 Karstic network modeling

In every simulation, the karstic network is generated using walkers that follow a shortest path to the spring. The path accounts for the local heterogeneity and is computed using a fast marching algorithm (FMA). In the following, we describe how the simulation medium is obtained and how the locations of the conduits are computed.

Characterizing the simulation medium

To use the FMA, it is necessary to convert the geological information I_g (eq. 4.1) and the fracturation information I_F (eq. 4.2) in a velocity field $K(\boldsymbol{x})$ [L/T] for each simulation of the fracture network and inception horizons.

We therefore convert the set of indicator functions obtained previously into a single velocity field. For that purpose, a velocity is defined for each of the different hydrogeological features: fractures and bedding planes are more conductive (high velocity) than matrix (low velocity). Aquitard formations have very low velocity. These velocities have the same units as hydraulic conductivities (m/s), but their values should be interpreted only as model parameters allowing to create contrasts between the different components of the propagation medium (matrix, fractures, inceptions horizons, etc). In other words, these velocities control the shape of the network and must be adapted by the user to obtain a realistic geometry. A lower contrast leads to a network that will be mainly controlled by the position of the inlets and outlet, on the opposite a higher contrast leads to a network that will be mainly controlled by the shape of the initial heterogeneity.

Once these parameters are defined by the user, the velocity field for each simulation is obtained by computing the sum of the various components multiplied by their respective indicator functions:

$$K_0^S(\boldsymbol{x}) = \sum_{f=1}^{N_f}[I_F(\boldsymbol{x}, f)\dot{K_F}(f)] + (1 - \sum_{f=1}^{N_f} I_F(\boldsymbol{x}, f)) * [\sum_{g=1}^{N_g}(I_G(\boldsymbol{x}, g)\dot{K_G}(g))] \tag{4.10}$$

where $K_G = \{\kappa_1, ..., \kappa_{N_g}\}$ and $K_F = \{\kappa_{F1}, ..., \kappa_{FNf}\}$ are 2 arrays of length N_G and N_F (number of geological formations and number of fractures families) defining respectively the hydraulic conductivity (velocity) of every geological formation and every fracturation family. Equation (10) assigns the velocity of the fractures at their simulated location within the medium and assigns otherwise the value associated to the corresponding geological formation.

Conduit simulation, basic algorithm

The conduits are generated using a particle tracking approach within a 3D map of shortest time to the outlet. This map is computed using the Fast-Marching Algorithm (FMA) developed by Sethian (1996; 2001) and implemented by G. Peyré within the matlab Toolbox Fast Marching (www.mathworks.com). The FMA is a front propagation algorithm. Here, we use the spring(s), i.e. the known outlet(s) of the aquifer, as the starting point for the propagation algorithm. The velocity of the front propagation at any point of the domain is given initially by the field $K_0^S(\boldsymbol{x})$. The result of the FMA is a 3D map of the time $T_0^S(\boldsymbol{x})$ it has taken for the front to arrive in any point (pixels in 2D, voxels in 3D) of the domain from the outlets:

$$T_0^S(\boldsymbol{x}) = t(K_0^S, O^S) \tag{4.11}$$

where $t()$ represents the Fast Marching Algorithm. The next step consists in using the set of starting points locations $\Im^S$ (eq. 4.5) generated previously. For each point in this set, we compute the path from the inlet toward the outlet by following the gradient of the time map $T_0^S(\boldsymbol{x})$. In this way, the algorithm performs a particle tracking from points situated at the land surface (”starting points”) and follow the shortest path from its starting point to the spring.

The tracks of the particles are stored via an indicator function $I_c(\boldsymbol{x}, \kappa)$ and describe the 3D karstic network through the domain:

$$I_c(\boldsymbol{x}, \kappa) = \begin{cases} 1, & \text{if karst } \kappa \text{ is present at location } \boldsymbol{x} \\ 0, & \text{otherwise} \end{cases} \tag{4.12}$$

By applying this algorithm iteratively, one can update the initial velocity field $K_0^S(\boldsymbol{x})$ with the indicator function of karst conduits $I_c(\boldsymbol{x}, \kappa)$ and by defining K_{karst}, the velocity associated to the karst conduits. This allows generating a hierarchical karstic network that can include several phases of karstification. The resulting new velocity field is called K_i^S where i is the iteration number.

$$K_i^S(\boldsymbol{x}) = I_c(\boldsymbol{x}, \kappa) * K_{karst} + (1 - I_c(\boldsymbol{x}, \kappa)) * (K_{i-1}) \tag{4.13}$$

Consequently the 3D time map is updated for every iteration in this way:

$$T_i^S(\boldsymbol{x}) = t(K_i^S(\boldsymbol{x}), O^S) \tag{4.14}$$

As illustrated in Figure 4.3b the first generation of conduits is influenced only by the initial heterogeneity of the media (Figure 4.2a). Figure 4.3c shows the network that is generated after 3 iterations. The hierarchical structure of karst conduits is respected because the pre-simulated network is a preferential path for the next walkers.

Point or Diffuse outlet

In real applications, it may be difficult to know on the field every single outlet of a karstic system. Sometimes outlets are located in diffuse areas and are therefore not represented by a single spring. To simulate this feature, it is possible to initialize the fast marching algorithm to start from a region instead of a single karst spring. i.e. O^S (eq. 4.9) is defined as a set of points defining a region and not as a single or a few points. This option can be interesting if one want to identify possible outlets of a region. Figure 4.4 illustrate the results obtained using a single deterministic spring (located at the center of the inferior border of the image, Figure 4.4a) and a diffuse spring zone (corresponding to the whole inferior border of the image Figure 4.4b). Both simulations use the same fracture network and the same starting points for comparison. We notice that even when a diffuse outlet zone is used, the resulting networks still show a hierarchical structure. Because the locations of the springs are not imposed, they depend on fracture positions and interconnections only. These locations vary from one stochastic simulation to the next.

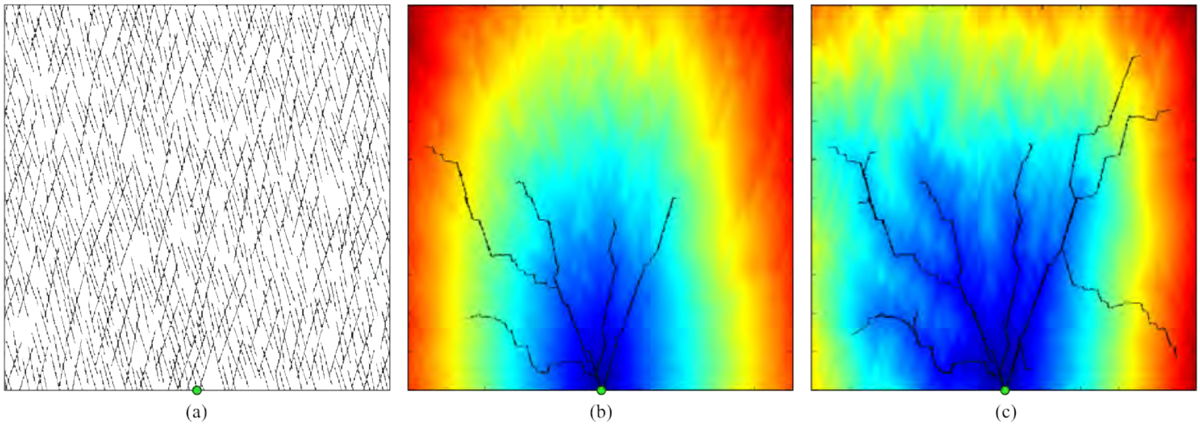


Figure 4.3: a) initial fracture network, b) first iteration of the algorithm, initial time map resulting of FMA from the spring location (green point) to every point of the domain. The black lines are the first conduits generated following the time map; color: blue (low values) to red (high values). c) Third iteration of the algorithm

All the previous steps are summarized in Figure 4.5 and constitutes the basic algorithm that we propose. It allows creating realistic conduit networks accounting for the regional geology and the internal heterogeneity. But this is not yet sufficient since it also needed to account for the unsaturated zone and for more complex situations such as multiple phases of karstifications. These questions are treated in the next sections.

Accounting for the unsaturated zone

In 3D, the basic algorithm (Figure 4.5) needs to be extended to account for the difference of structures in the unsaturated and saturated parts of the karstic systems. In the unsaturated zone (vadose zone), the main controlling factor is gravity, and the conduits are mainly collectors that drain the water concentrating from the epikarst (Huntoon et al., 1999). Therefore conduits are essentially vertical in the vadose zone, whereas in the saturated zone the conduits are essentially driven by the regional base level and the spring locations.

To model those feature, the simulation process is divided in 2 steps: first the vadose conduits are generated and second the saturated ones (Figure 4.6). Each step uses the basic algorithm described above (Figure 4.5) but with different inlets and outlets.

For the first step, the initial front propagation points (diffuse outlet) correspond to a 2D surface: the base level that has been extended all over the 3D model. For a 2D cross section, it is a line or a curve (see dotted line in Figure 4.6a):

$$O^S_{step1} = \{\boldsymbol{x_{i1}} = (x_{i1}, y_{i1}, z_{i1}) \text{ with } i1 \in (1, ..., N_{XY}\} \tag{4.15}$$

where N_{XY} is the total number of points of the grid for one XY plane (i.e. $N_{XY} = nx * ny$, with nx and ny the grid dimensions in X and Y directions). Then, the FMA is used to compute the map of arrival times starting from the base plane O^S_{step1}:

$$T^S_{i1}(\boldsymbol{x}) = t(K^S_{i1}(\boldsymbol{x}), O^S_{step1}) \tag{4.16}$$

Particles are sent from the starting points $\Im^S$ located on the topography. The result of the particle tracking is an indicator function of the presence of vadose conduits (Figure 4.6a).

$$I_{vadose}(\boldsymbol{x}, vc) = \begin{cases} 1, & \text{if vadose conduits } vc \text{ is present at location } \boldsymbol{x} \\ 0, & \text{otherwise} \end{cases} \tag{4.17}$$

At the end of step 1, the set of coordinates of arrival locations of the particles on the base level O^S_{step1} is used as the set of starting points $\Im^S_{step2}$ for the phreatic conduits in step 2.

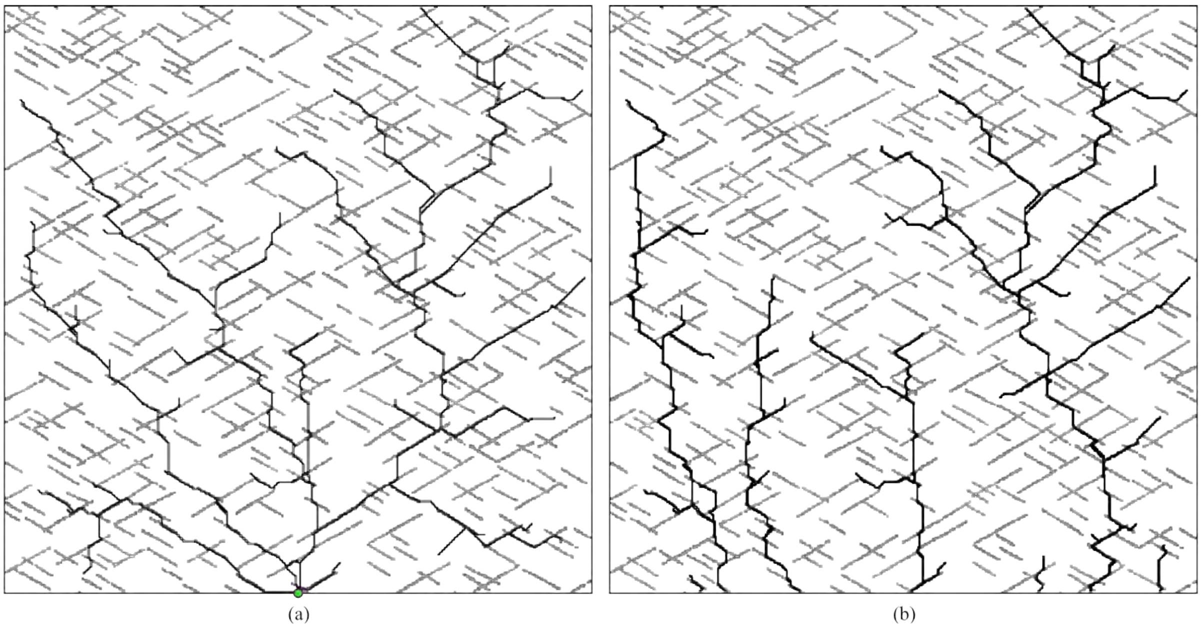


Figure 4.4: results of simulations on a 2D map. a) point deterministic spring (green point), b) the whole bottom of the image is a zone of possible outlets, the real position of the outlet vary from one realization to the other. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Step 2 uses the FMA again, but now the starting locations for the front propagation are the the actual outlets locations (O^S) observed in the field. The resulting time map $T^S_{i2}(\boldsymbol{x}) = t(Ki2^S(\boldsymbol{x}), O^S)$ is used to compute the paths from the starting points $\Im^S_{step2}$ obtained in step 1. The result is the indicator function of the phreatic conduits:

$$I_{phreatic}(\boldsymbol{x}, pc) = \begin{cases} 1, & \text{if phreatic conduits } pc \text{ is present at location } \boldsymbol{x} \\ 0, & \text{otherwise} \end{cases} \tag{4.18}$$

We associate then the indicator function of vadose conduits $I_{vadose}(\boldsymbol{x}, vc)$ (eq. 4.17) with the one of phreatic conduits $I_{phreatic}(\boldsymbol{x}, pc)$ (eq. 4.18) to obtain the complete indicator function of the karst conduits $I_c(\boldsymbol{x}, \kappa)$ (Figure 4.6b):

$$I_c(\boldsymbol{x}, \kappa) = I_{vadose}(\boldsymbol{x}, vc) \cup I_{phreatic}(\boldsymbol{x}, pc) \tag{4.19}$$

Complex connections between inlets and oulets

Karst aquifers are often characterized by springs alimented by several sinkholes. But it can also occur that some of these sinkholes are connected to other springs as well. To simulate this type of relationship, the easiest way is to first simulate the network for one specific spring ($O^S[1]$), and then for the second ($O^S[2]$) and so on, keeping the result of every iteration as initial velocity field ($K^S_i(\boldsymbol{x})$) for the next. In this way, it is possible to connect several springs $O^S[1, , n]$ to several sinkholes in a very simple way (Figure 4.7).

Multiple phases of karstification

If different phases of karstification have to be accounted for due to the existence of different paleo hydrological conditions, the generation of karst networks can be repeated several times using the pre-simulated network as an initial velocity field ($K^S_i(\boldsymbol{x})$) and changing the FMA initial front propagation points O^S for every change in the paleo conditions. Figure 4.8 illustrate this method on a simple 2D system assuming the presence of two phases of karstification with 2 successive springs corresponding to two distinct base levels.

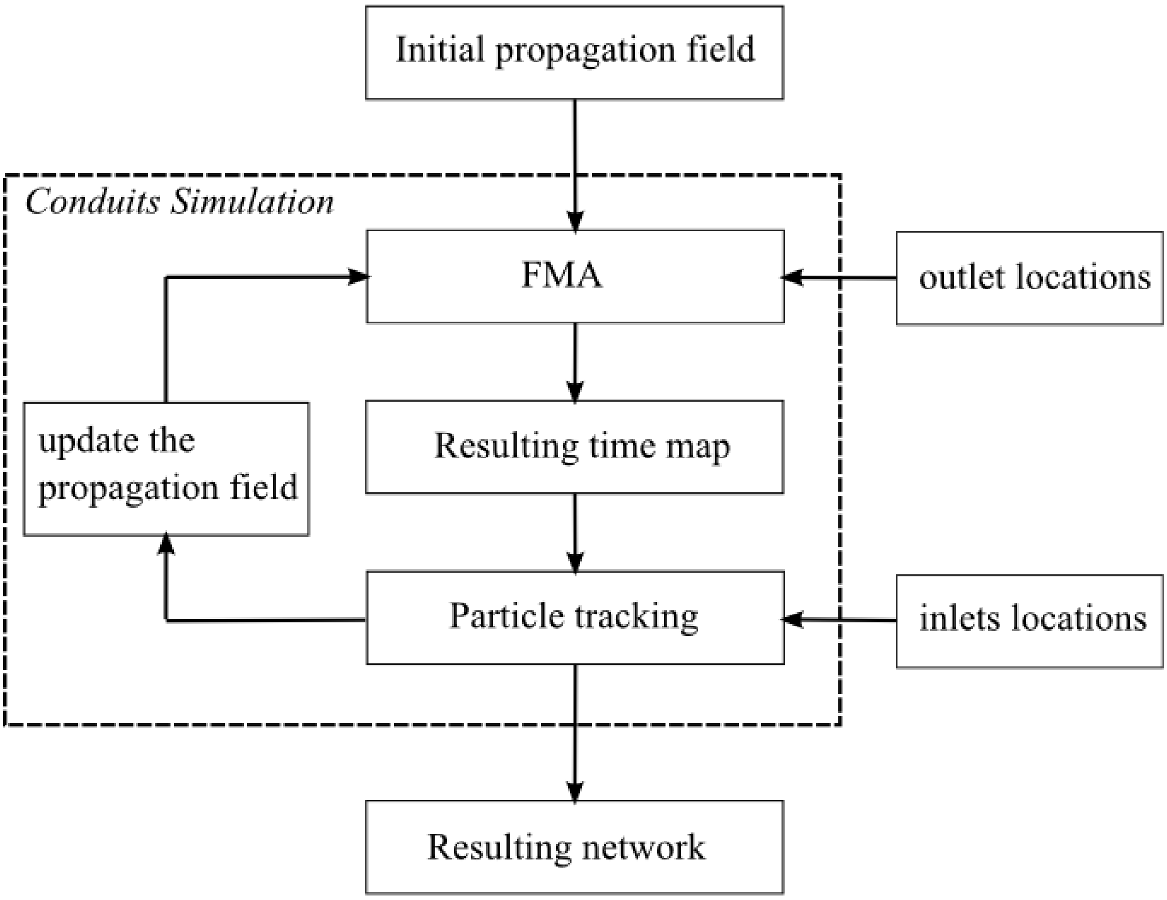


Figure 4.5: workflow of the basic algorithm for the generation of the karstic conduits using the Fast-Marching Algorithm (FMA)

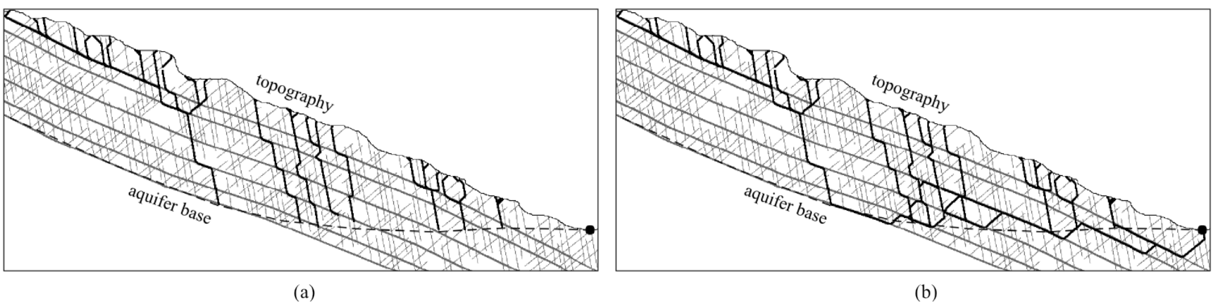


Figure 4.6: result of one 2D cross-section simulation: (a) the conduits (black lines) within the vadose zone are generated toward a base level first, (b) and then the saturated conduits are simulated toward the spring (b). The dotted line represent the phreatic surface. The filled circle represent the spring, the gray lines represent the fractures and the bedding planes

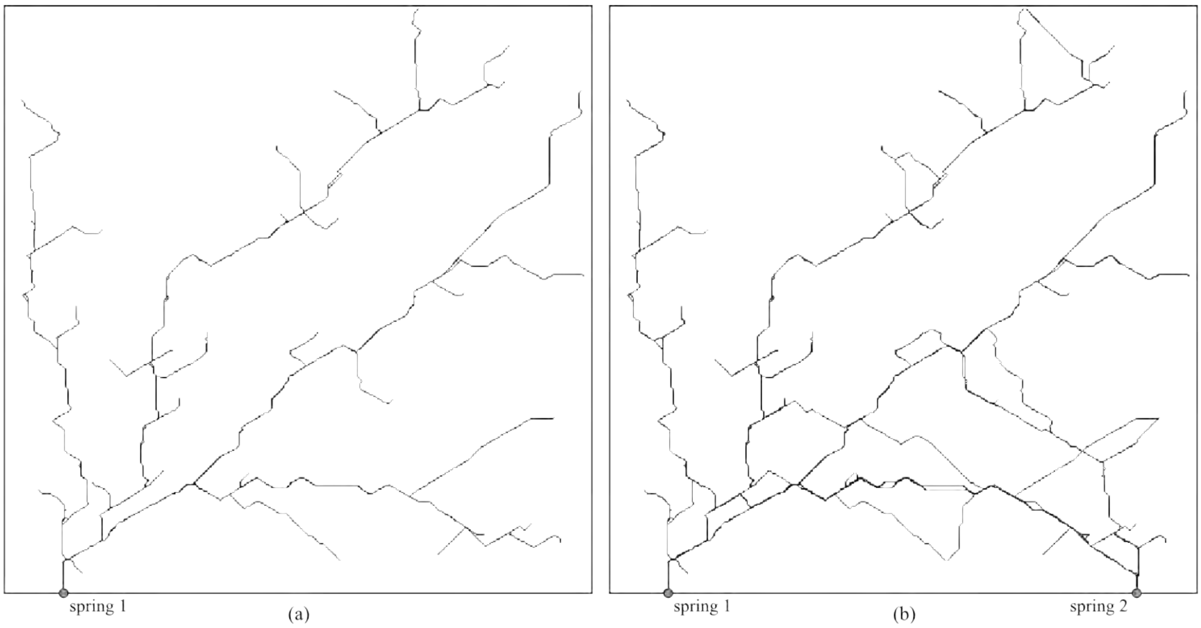


Figure 4.7: 2D map. To simulate multiple connections we first develop a whole network for the first spring (a) and then the one for the second spring (b), keeping the first one to create a complex multiconnected network.

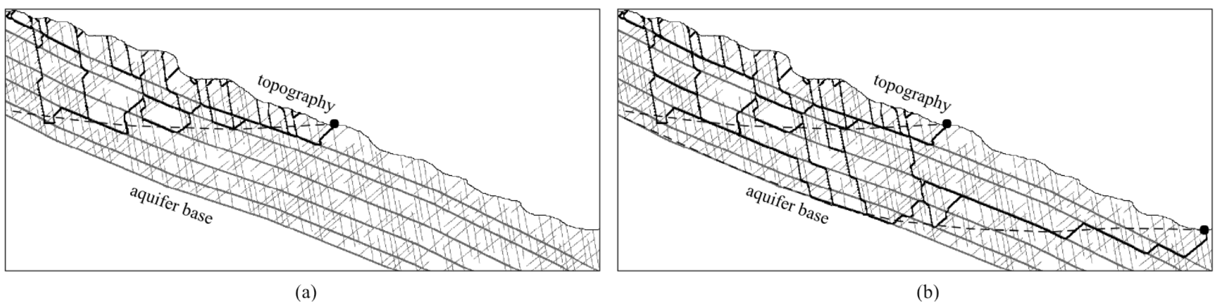


Figure 4.8: 2D cross-section of multi-phase karstification. 2 successive speleogenetic phases are illustrated. The dotted line represent both phreatic surfaces, the old (a) and the actual (b). The filled circles represent the springs, the gray lines represent the fractures and the bedding planes, black lines represent the conduits

Hierarchical ordering of the conduit

The algorithm described above allows to compute a conduit order, which depends on the total catchment area drained by the conduit. The whole network is divided in a set of topological segments. Each segment is defined as a portion of the network bounded by two branching or diverging connections without any internal connection. Every segment j inherit its order ω_j (expressed in %) from the upstream segments. It can be computed as the sum of the indicator functions of presence for each conduit I_{ci} (i.e. its full path from the inlet toward the spring) of the upstream conduits multiplied by the importance factor $\omega_j^{\Im}$ (see eq. 4.6) of its corresponding inlet.

$$\omega_j = \sum_{i=1}^{N_{\Im}} I_{ci} * \omega_j^{\Im} \quad \forall \ j \in (1, ..., N_S) \tag{4.20}$$

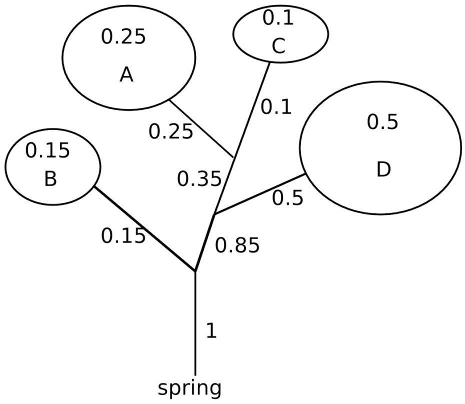


Figure 4.9: schematic illustration of the conduit ordering. Every conduit inherit the rank (in % of total catchment drained surface) of the upstream conduit or inlet (A,B,C,D).

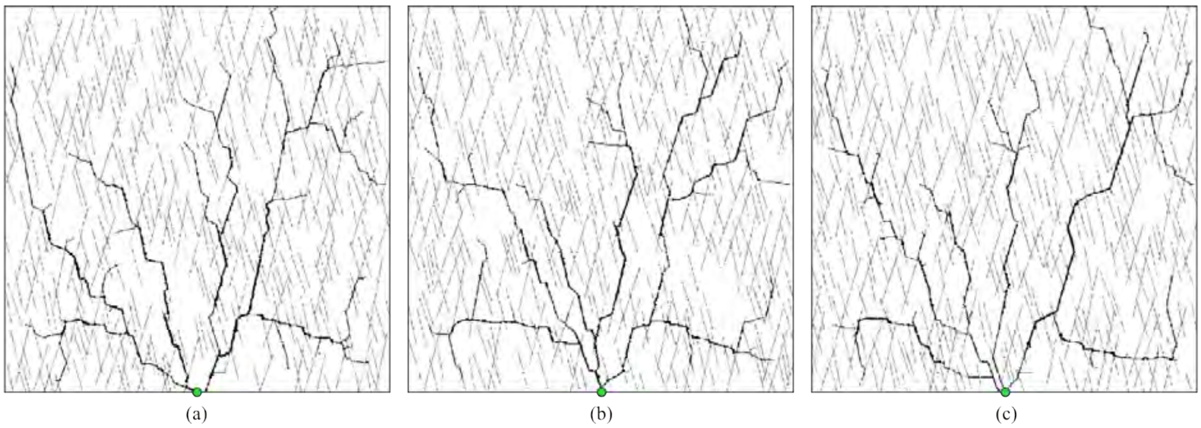


Figure 4.10: 3 different equiprobable realizations (2D map), the outlet (green point) and the inlets (starting points of the conduits, black lines) are the same for all the realizations. The different equiprobable fracture network (gray lines) influence the connectivity of the resulting network.

4.3.5 Stochasticity of the generated networks

Despite of the influence of the geological model which controls the regional shape of the networks, the generation of conduits networks has three main levels of stochasticity.

The first level depends on the stochastic model of inlet locations $\Im_P^S$. Since not all the inlets of the model are precisely known, they are modeled by a point process which can be parametrized by a map of a probability of occurrence of an inlet $Pi(x, y)$ (section 4.3.3).

The second level depends on the stochastic model of fractures. Figure 4.10 illustrate equiprobable realizations obtained using the same inlets and outlet, but using 3 equiprobable realizations for fractures (gray lines). The paths of the resulting conduits (black lines) and their interconnections vary significantly. Furthermore, one can test several models of fractures to investigate the uncertainty related to the fracturation model itself.

Finally, the third level depends on the stochastic model of the outlets of the system O_P^S. Despite they are often known, one can generate models also in the case where there are no information about the outlets of the system, like in the Damour example (section 4.4.1).

4.4 3D demonstrative examples

This section provides two illustrations of the applications of the proposed methodology. The aim of the paper is to develop a conceptual framework and a numerical method but not to provide a detailed and accurate model of the regions that are used for the test. This is why we show the following examples only to demonstrate the applicability of the method and we use them to identify the work that remains to be done to represent as accurately as possible a real system.

4.4.1 Case 1, coastal aquifer

This first example illustrates the application of the methodology to a synthetic case, representing a typical kind of karstic coastal aquifer on the Mediterranean basin. The basic geology is inspired by the Damour region in Lebanon. Even if that specific region is known to be only moderately karstified, its geological structure is a good analog of many other aquifers along the Mediterranean that are well karstified and often present submarine karstic springs (Fleury et al., 2007). We show how to use the proposed methodology, with multiple phases of karstification due to the variation of the base level caused by glaciations. The lowest sea level known below the present occurred during the Wrm glacial maximum (Ford and Williams, 2007). It was between -120 to -140 m. Thus, we modeled 2 distinct phases of karstification, one with the sea (base) level at -140m and one at 0m. The network resulting from the first simulation is kept for the second stage, so that conduits will be reactivated during the second stage.

The geology of the site is relatively regular. In the western part of the model (i.e. seaside), the structure is mainly homoclinal, dipping Westward at an angle between 30 and 50 following the topography. Towards the mountains, the structure rises to the surface at an angle of 90 (Figure 4.11a).

We had no field information about fracturation, so (because it is a synthetic case) we assumed 2 families of fractures due to the compression that resulted in the folding. Considering a syncline axis (Eastern part of the model, in the mountains) oriented N-S, we can assume 2 families of conjugated faults oriented at 60 degrees one from each other: one NW-SE (strike 345N) and one SW-NE (strike 15N). The fractures are assumed to be vertical, their mean length is 250m and their mean vertical extension is 150m (Figure 4.11a).

For every karstification stage, the spring location has been chosen randomly on the coastline at a depth of 15 m below the respective base level, like one of the springs of the analog aquifer of Gekka, Lebanon (El-Hajj et al., 2006). Because it is a synthetic case, we simply assume that the probability distribution for the starting points of the conduits $\Im_P^S$ (eq. 4.4) is uniform at the land surface where karstic formations are outcropping. The number of starting points for the conduits is , i.e. 5 for the first iteration, 10 for the second and 20 for the last one. We assume for this example that the paleo inlets are the same as the actual ones. Figure 4.11b shows one resulting network.

4.4.2 Case 2, aquifer in folded tectonics

The second example corresponds to a karstic system developed in a more complex tectonic setting. The studied site is the catchment of the Noiraigue spring in the Swiss Jura mountains, close to city of Neuchâtel. This karstic system cannot be explored by speleologist. The site has however been studied for decades at the university of Neuchâtel by hydrogeologists (Gillmann, 2007; Atteia et al., 1996; Morel, 1976) and geologists (Sommaruga, 1996; Valley et al., 2004). It is characterized by a highly folded geology with complex overthrusting. The catchment of the Noiraigue spring is a dry syncline valley, called "Vallée des Ponts-de-Martel"; all the precipitation falling on the valley infiltrates into the karst system and the only known outlet of the system is the Noiraigue spring, so we defined it deterministically as the only output point of the domain.. The catchment area is approximately 70 square km and is delimited by 2 overthrusting anticlines.

To build the 3D model of the regional geology of the catchment (Figure 4.12a), we used more than 50 interpreted cross-sections (Sommaruga, 1996; Valley et al., 2004), part of 3 seismic profiles, 6 boreholes and 4 different geological maps (interpretation) of the region (Borghi, 2008). These data were imported in the software Geomodeller3D (Calcagno et al., 2008) and used to interpolate the 3D geometry of the main structures including the lithological formations, folds, thrusts and faults. There are several families of conjugated faults that can be observed in the field due to folding, but we retained only the major ones, i.e. WNW-ESE (strike = 300°N) and N-S (strike = 0°N) according to Sommaruga (1996) for the

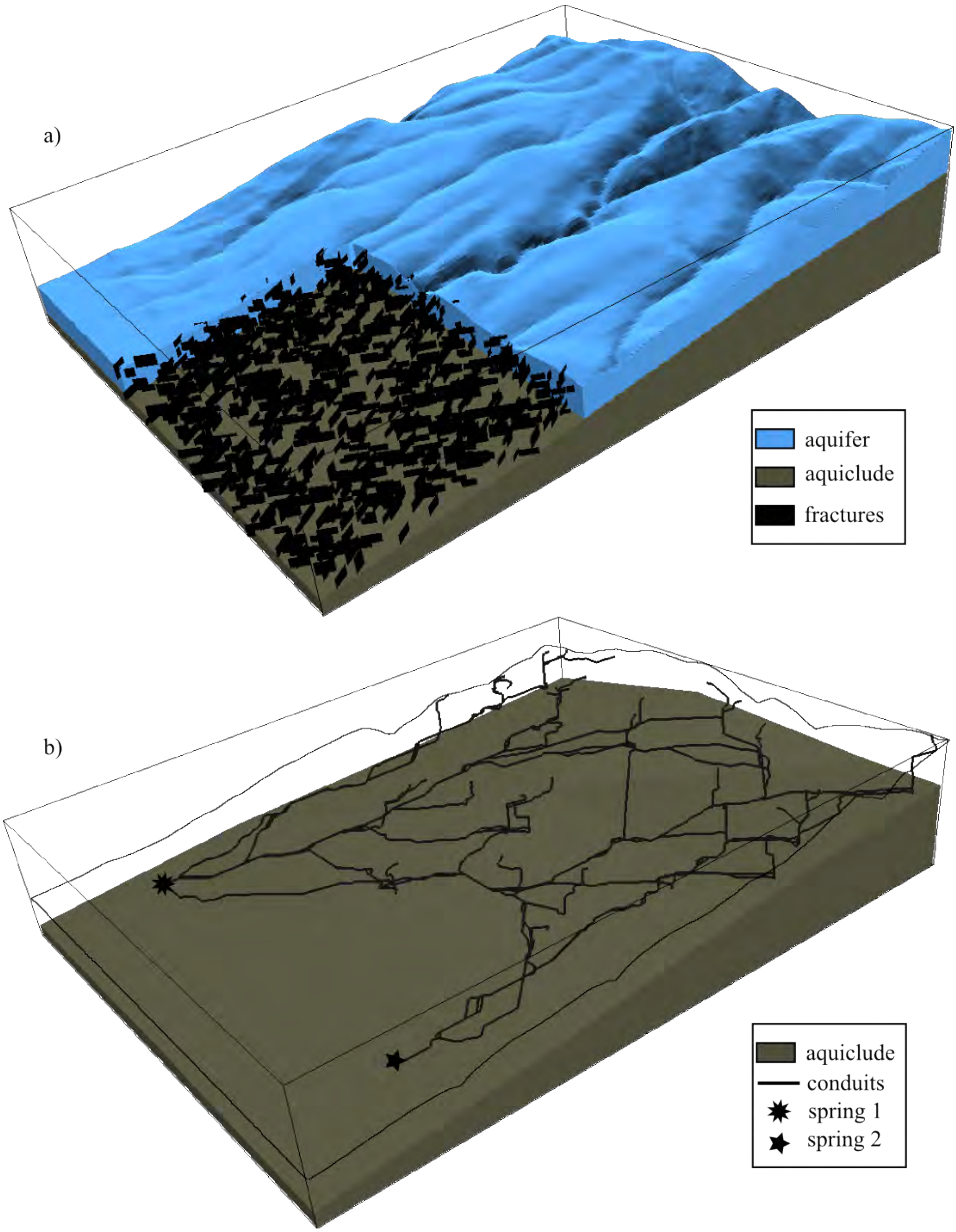


Figure 4.11: coastal aquifer example. (a) geological model with fractures, showing also both basel levels (-140m and 0m) (b), karst system for paleo base level (-150m), karst system under actual conditions (base level = 0m)

model. Similarly to example 1 (section 4.4.1) we used a Boolean technique to generate the fractures (Figure 4.12b).

The probability of occurrence of the starting points of the conduits is uniformly distributed at the land surface within the catchment but only where karstic formations are outcropping. In addition, 8 major known inlets (sinkholes and dolines) are defined deterministically.

We then simulated the network by first simulating the vadose conduits and second the phreatic ones (see section 4.3.4). For the vadose conduits, we assumed the base level to be the a plane at the same altitude of the spring and crossing the whole catchment.

As expected, the resulting networks follow the geometry of the main geological structures (Figure 4.12c). They are well constrained by the known locations of inlets and the outlet. The main conduits cross the aquifer underneath the tertiary aquitard as supposed by the existing hydrogeological analyses of the region (Atteia et al., 1996; Morel, 1976). We therefore obtain a reasonnable 3D model of the conduits in a very complex geological setting.

To increase the reliability of this model, we would need to better constrain the fracturation model and use a more advanced tool for the generation of the fracturation that would account for the known folds and thrusts. We would also need to improve the stochastic model of potential inlet locations since geomorphological information is available in surface. This work will certainly constitute a direction for future research.

4.5 Discussion

In this paper, we present a new stochastic methodology to model karst reservoirs. This method is significantly different from the ones found in literature. Its main strength is that it allows integrating a broad range of data and conceptual knowledge into a single stochastic model. The regional geology can be accounted for even when it is very complex. The method allows to condition the karst conduit model with actual observations. If inlets or outlets are known on the field it is sufficient to set them as starting or ending points in a deterministic way. Borehole data confirming the presence of conduits can be used as well. The method can handle the existence of different phases of karstification and the distinction between the vadose and phreatic zones. When some information is only available in a probabilistic manner, the method allows to account for it and is able to generate an ensemble of realizations representing the uncertainty on the true position or structure of the network.

For the moment, the geological model is defined in a deterministic way. The resulting model reflects the vision that the modeler has of a specific site. It is therefore a good practice to build several geological models if different assumptions can be made related to the geological structure. By doing different models one can use the proposed approach to test the plausibility of different assumptions, and evaluate their influence on the subsequent processes. In addition, it is very important to have a good model for the fracturation and other heterogeneities like the bedding planes. This requires accurate field measurements and observations as well as a numerical tool able to include those observations into the model. Some additional work is required in that direction. But it does not change the principle of the proposed model.

The heart of the methodology relies on the use of the Fast Marching Algorithm. Alternatively, one could use groundwater flow simulations and compute streamlines to obtain similar results. Both techniques have pros and cons. As a first guess, using a flow simulation may seem better. Flow is the physical process controlling the speleogenesis and could therefore seem easier to justify. But in fact, it is more difficult to use because one has to define the flow parameters of the different geological materials (matrix, fractures and conduits) and the boundary conditions. The FMA is much simpler to apply. To evaluate the impact of the two options, we made several tests in 2D. One of the comparison between FMA and flow simulation is show in Figure 4.13. The flow simulation was computed with a constant and uniform source term on the whole domain, a constant head boundary condition at the spring location and a no-flow boundary on the sides. A difference of 5 orders of magnitude is prescribed between the hydraulic conductivity of the matrix and of the fractures and one additional order of magnitude between the fractures and the conduits. The flow paths are influenced by this strong heterogeneity and so do the conduits. In this case, the comparison of Figure 4.13(a) and Figure 4.13(b) shows that the shape of the conduits obtained from the flow simulation is less influenced by the fracture position than those obtained from the FMA (Figure 4.13a). Furthermore the flow simulation is also strongly influenced by the no-flow boundaries on the sides of the model. In this situation, using the flow simulation generates more geometrical artifacts than the FMA. The flow simulations are also more CPU demanding, one simulation

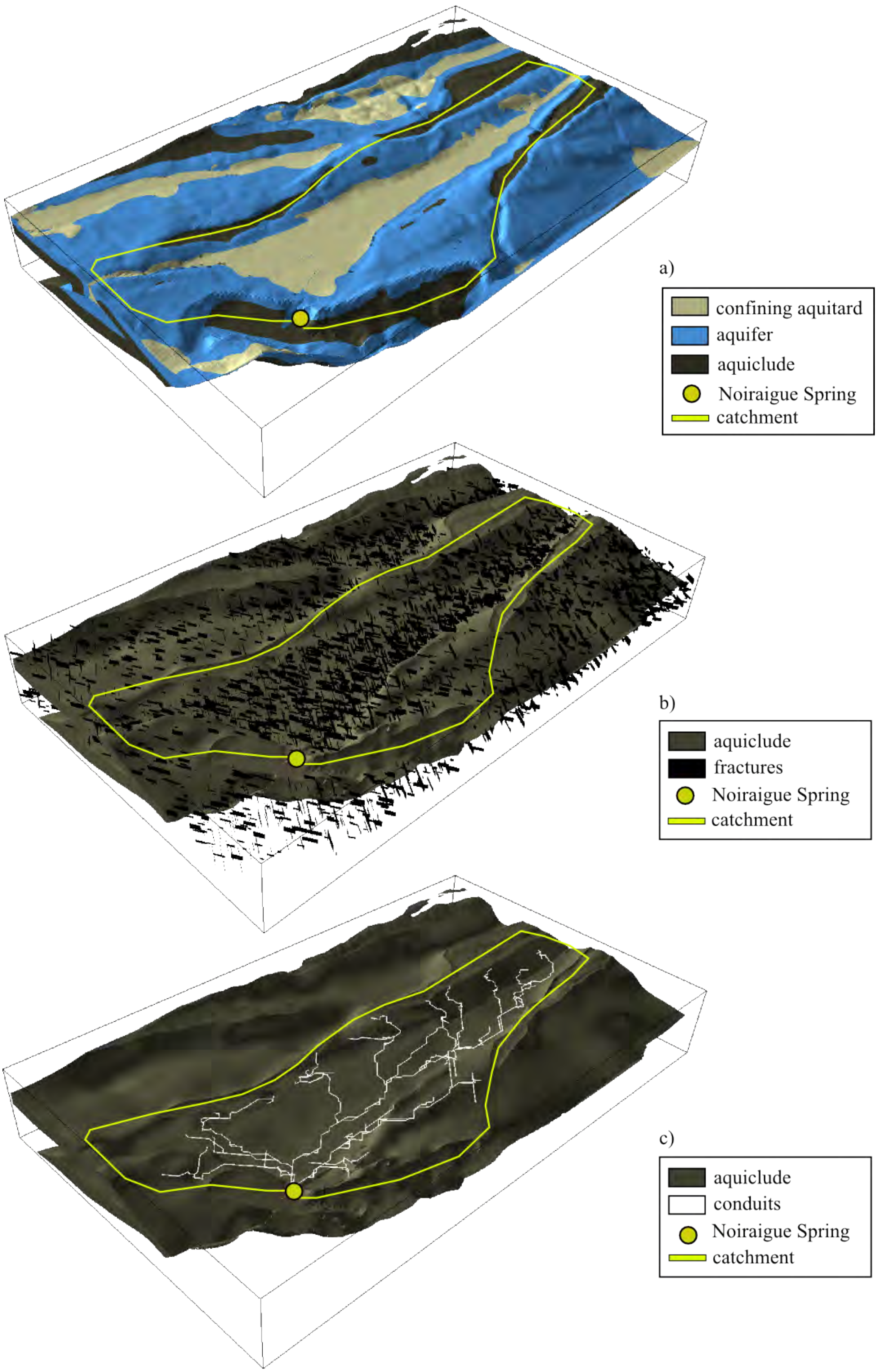


Figure 4.12: folded aquifer. (a) geological model, (b) fracture model, (c) resulting network, (d) resulting network with aquitard. The the conduits follow the regional geology and the local fracturation

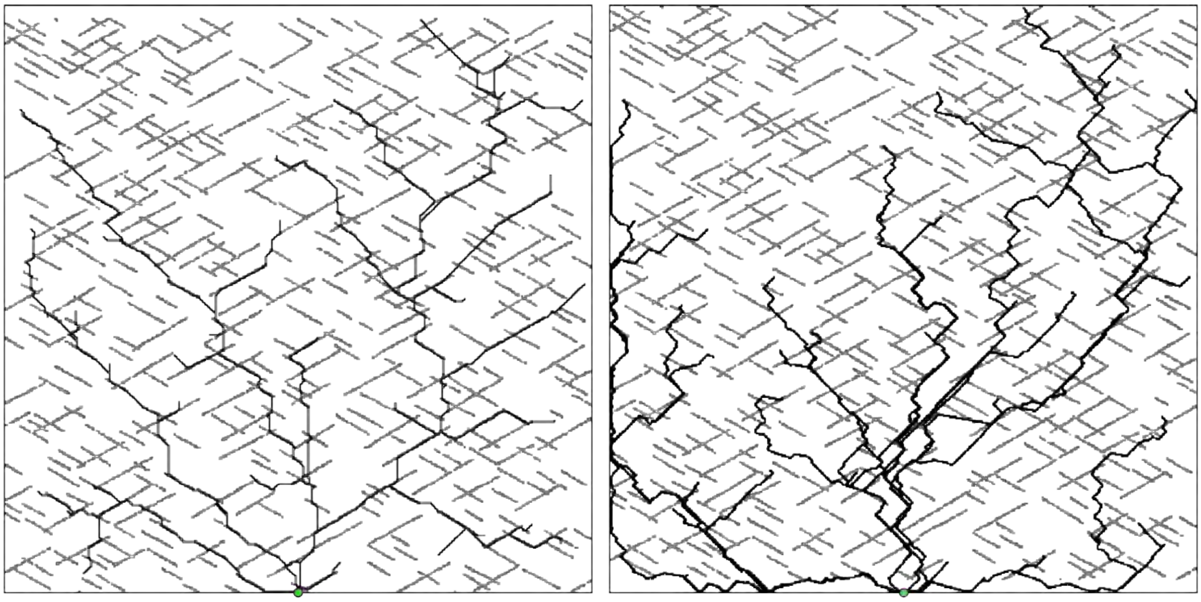


Figure 4.13: conduits generated using: (a) the Fast Marching Algorithm and (b) a steady state flow simulation. For case b an uniform source term has been applied on all the model surface, and an constant head boundary condition has been applied at spring location. We notice that that using a flow simulation we obtain a network which is slightly more "rounded" and more influenced by the border of the model (no flow borders).

for this 500x500 2D simulation takes less than 10 seconds using the FMA (PC: dual-core 2.4 gHz, 2Gb Ram) and about 3 minutes using a steady state flow simulation (without writing the output files). In 3D, accounting for the vadose zone would be even more complicated since it would require solving the unsaturated flow equations on very large grids. Due to the non-linearity of those equations, this can become extremely cpu intensive and not applicable in a stochastic context as shown by Brunner et al. (2009). For all those reasons, we prefer to use the Fast Marching Algorithm which simply computes the shortest path. This is certainly a reasonable physical approximation for our purpose.

As the FMA is only an approximation, it has to be applied carefully. For example, artifacts can occur if the karst formation is folded. The worst artifacts that we obtained were conduits going upward. It may occur if a starting point has been defined in a hole, and so following the time map computed by the FMA, the conduit may go upward. To avoid this kind of artifacts, a great attention has to be accorded to the definition of the starting points probability map.

Another problem happens in 3D when conduits follow the "gravity" in the vadose zone (step 1 explained in section 4.3.4). If the base level for the FMA Sl_1 (eq. 4.11) is badly defined, or if it is located outside of the aquifer limit, the conduits can be "attracted" toward aberrant zones. The modeler has to accurately investigate the resulting networks, and a good interpretation is essential. Actually, this feature can be used to test the extension of a catchment: points located outside the catchment will generate conduits that will leave the catchment.

Moreover, in real karst systems, the development of the unsaturated and saturated zone are synchronous and follow several successive phases of maturation (Filipponi et al., 2009). In our approach, we approximate this behavior by simulating only mature karst systems, that have already reached the equilibrium. With this assumption, simulating first all the vadose conduits and in a second step the phreatic ones leads to a geometrical reconstruction of the mature karst that is coherent with the actual conceptual model (Ford and Williams, 2007).

As we already mentioned in the introduction,the proposed approach aims at generating karst conduits networks and not cave geometry like Henrion et al. (2008). In real karst aquifers, huge caves can be present, but we do not consider them yet in our approach because we work on regular grids. At a regional scale, the resolution of the grid is often coarser than the dimensions of the caves (cell resolution $\geqslant$ 50m). Therefore, the caves can be represented only by "conduit" elements. Instead of trying to model the detailed geometry of the caves, we think that it is more reasonable to pursue our work in the future by defining a model of the radius (or rugosity) of the conduits that will be computed all along the conduit networks. A reasonable model for the radius will be to relate statistically the conduit order defined above (eq. 4.20) with its radius using a power law. For similar reasons, we do not simulate the epikarst and its internal detailed heterogeneity using discrete features. Instead, in a regional flow model we think that it is more reasonable to represent the epikarst by a highly conductive layer (that can be heterogeneous) at the topographical surface of the model.

Another important point that needs to be discussed is that the number of conduits and the different phases of karstification are for the moment controlled by the user. He has to define those parameters. As speleogenetic process based simulations show (Dreybrodt et al., 2005) the dissolution in natural system is very active when the head gradient is high. Because of the dissolution, the hydraulic resistance of the medium tends to decrease and the gradient decreases (the rate of precipitation or inflow is supposed to remain constant) until the system reaches a breakthrough point. This determines the moment after which the karst network is in equilibrium with the inflow, i.e. all the infiltrated water is drained easily by the system. After the breakthrough, the dissolution is more active in the epikarst and the geometry of the main network does not change that much anymore. Using the pseudo-genetic approach, we do not account for this factor because we do not model the dissolution processes. Instead, the user has to control that the parameters that were used to build a reasonable network. In this process, the importance factor (section 4.3.3 and 4.3.4) can be used as a criteria controlling the density of the inlets: using a DEM analysis tool, as the drainage surface of every inlet is already estimated; the algorithm could proceed by adding inlets on the surface until the sum of the drained surface is equal to a user defined threshold. This idea needs to be tested in detail and is therefore not yet implemented in the code.

4.6 Conclusions and outlooks

We propose an integrated method to model epigenetic karst reservoirs. It allows honoring conceptual knowledge existing about karst and speleogenesis processes as well as most of the available field data. As far as we know this is the fisrt stochastic method allowing to account for several phases of karstification and to distinguish the saturated and unsaturated zones. The method is very flexible and numerically efficient thanks to the use of approximated physics concepts and it belongs therefore to the pseudo-genetic class of stochastic models. The use of the Fast Marching Algorithm to perform the search of minimum effort paths allows to simulate realistic conduits in a very CPU efficient manner.

The proposed model is a tool allowing to test different assumptions concerning either the regional geology, the fracturation, the inception horizon or the hydrogeological control of the system. It generates conduits that can be used to generate finite element or finite differences mesh which can be used to model flow and transport in the karst system. The model is stochastic and allows to investigate the uncertainty related to a current state of knowledge and data availability. As any model, its underlying parameters have to be inferred from field observations. Some parameters can be directly derived from the measurements (location of the springs for example), some other parameters are not necessarily directly measurable (number of major stages of karstification for example). Depending on the parameters used in the model one can obtain a very wide range of structures from curvilinear branching network to anastomotic mazes.

For the moment, a trial and error procedure is used to select the parameters. By testing the impact of those parameters and checking visually if the resulting networks are reasonable, one can obtain the type of network that is expected. In the future, the methodology will have to include a calibration procedure based on a number of statistical measures describing the topology and geometry of the network. Those measures will be used to compare the simulated networks with observed ones in order to test the validity of the selected model parameters.

Acknowledgements

This research was funded by Schlumberger Water Services. We thank J.-P. Delhomme who initiated the project and always gave passionate feed-back, G. Mathieu who continuously supported the project. A. Comunian, P. Brunner and three anonymous reviewers helped us making the manuscript more understandable.

Chapter 5

Extensions and improvements of the pseudo genetic karst model

5.1 Introduction

This chapter describes three modifications of the *pseudo genetic* algorithm made after the publication of the previous chapter. They have been made 1) to allow creating labyrinthic conduits, 2) to account for geophysical data, and 3) to avoid unrealistic conduits going upward. In this chapter, we describe these additions from a theoretical point of view. Furthermore this chapter provides some detailed information about the fast marching algorithm (FMA) and illustrates how the algorithm is sensitive to its input parameters. This is important because it enables the user to better understand the *Stochastic Karst Simulator* (SKS).

Many SKS options allows to generate karst networks that are consistently different one from another. As explained in Section 4.3.4, SKS uses a Fast Marching Algorithm (FMA) (Sethian, 1996, 1999, 2001) to compute the minimum effort paths between karst inlets (sinkholes, dolines) and outlets (springs). The aim of this approach is to model mature karst networks, with the possibility to condition them to field observation. In this case, a karst system is considered *mature* when it reaches the sufficient state of enlargement of karstic voids (conduits) that allows them to drain the majority of the infiltrating water.

SKS generates only static networks, i.e. it is not possible to simulate the evolution of these networks as it would be done in speleogenetical models (like e.g. Kaufmann and Braun, 2000; Kaufmann, 2009; Dreybrodt et al., 2005; Gabrovsek and Dreybrodt, 2010) because the subsequent philosophy of the approach is to generate stochastic networks (section 4.2). The geometry of the resulting networks is strongly related to the FMA, and to the way it is used. Therefore the first section of the present chapter gives a short explanation of the FMA (Section 5.2). The second section describes how the FMA is used by SKS and how the parameters for the FMA can be used to generate different networks (Section 5.3). The last part shows the new additional parameters that help to get more realistic karst networks (sections 5.4 and 5.5) and how geophysical data can be used to control the formation of karst conduits (section 5.7).

5.2 Description of the Fast Marching Algorithm

This section gives an explanation of the Fast Marching (FM) method, which is at the core of SKS. This explanation is intended to ease the understanding of SKS and its implementation. The FM is particular a case of boundary value front propagation methods. The reference book for these kind of techniques is Sethian (1999) and its re-edition of 2008 (Sethian, 2008). The explanation of the FM that is given here is taken from this reference. In this section, we explain first the general boundary value front propagation methods (Section 5.2.1), then the particular case of the Fast Marching Methods (Section 5.2.2).

5.2.1 Boundary value front propagation methods

Front propagation methods are used to compute the temporal evolution of a front interface in a medium. These fronts are propagated in the medium according to a speed function (i.e. at which speed the front is evolving). This speed function K can be noted[1]:

$$K = K(L, G, I) \tag{5.1}$$

It depends on several factors:

- L local properties
- G global front properties
- I front independent properties

The local properties are mainly related to the geometry of the model (mesh) and to the geometry of its associated properties. The global front properties are related to the partial differential equations that are associated to the front evolution and how they are integrated. The independent parameters are called this way, because they are independent from the shape of the front. In the case of a transport problem, they can be related for example to the velocity of the fluid that transports the front of concentration. There are many ways to formulate these speed functions, which will not be discussed here[2].

[1] see also Section 4.3.4
[2] refer to Sethian (2008) for further information

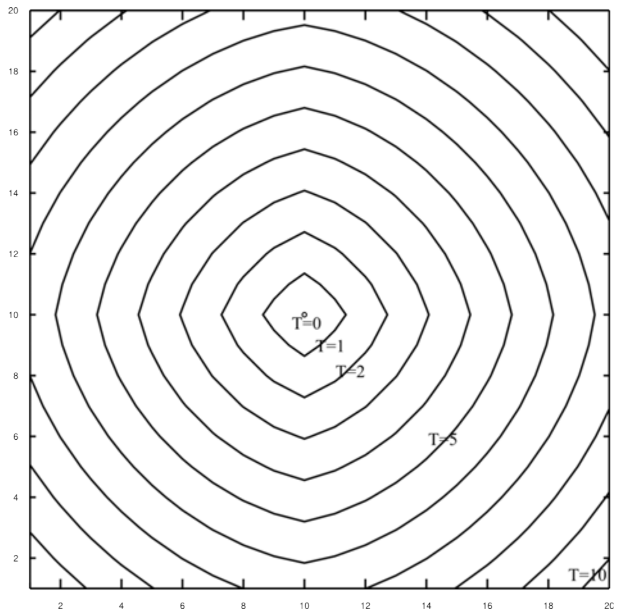


Figure 5.1: Evolution of the time T in an homogeneous FM speed field where $K = 1$ using the FMA implementation made by Peyre (2008). The coarse grid discretization ($20 * 20$ elements) gives a small bias to the shape of the concentric equidistant T circles in the horizontal and vertical direction. This is a known artifact of this implementation of the FMA.

In the boundary value front propagation methods, the speed function is considered as always positive (i.e. $K > 0$), which means that the front propagates always "forward". The position of the interface of the front can be known by computing the time of arrival $T(\boldsymbol{x})$ of the front when it reaches the position $\boldsymbol{x} = (x, y, z)$ (see also Section 4.3.4). The speed function can be derived considering that $d = KT$ (i.e. $distance = speed * time$):

$$K\frac{\delta T}{\delta x} = 1 \tag{5.2}$$

which gives in many dimensions:

$$|\nabla T|K = 1 \tag{5.3}$$

with $T = 0$ at $\boldsymbol{x_0}$, the initial locations of the front, and where $|\nabla T|$ is the euclidian norm of ∇T:

$$|\nabla T| = \sqrt{\left(\frac{\partial T}{\partial x}\right)^2 + \left(\frac{\partial T}{\partial y}\right)^2 + \left(\frac{\partial T}{\partial z}\right)^2} \tag{5.4}$$

∇T is orthogonal to the level set of T and its magnitude is inversely proportional to the speed K. According to Sethian (2008) this is the standard definition of the boundary value front propagation methods, which usually needs iterative algorithms to be solved. The Fast Marching Algorithm is different because it allows to solve this type of problems without iterations as explained in the next section.

Figure 5.1 shows an example of the evolution of T with $K = 1$. The front propagates in concentric form, because the propagation medium (speed field) is homogeneous and constant. In this figure one small artifact is visible: the concentric circles are slightly longer in the vertical and horizontal direction. The reason of this is the preferential time T using the implementation of the FMA that has been used (Peyre, 2008), which privileges the diagonal direction.

5.2.2 Fast marching methods

The Fast Marching Algorithm (FMA) is an algorithm that allows to solve boundary value front propagation problems without iterations, by an optimal ordering of the nodes of the grid that supports the computation. The system $|\nabla T|K = 1$ is then solved by one unique computation. Rouy and Tourin (1992) (cited in Sethian 2008) propose the following convenient one sided upstream scheme:

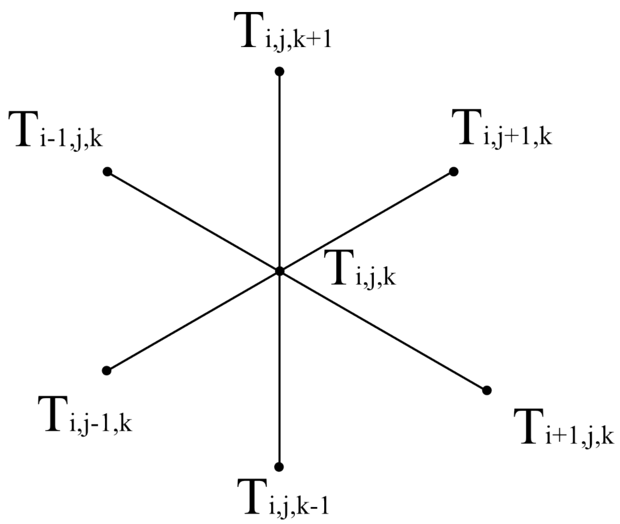


Figure 5.2: Updating T scheme, modified from Sethian (2008)

$$[max(D^{-x}_{ijk}T, -D^{+x}_{ijk}T, 0) + max(D^{-y}_{ijk}T, -D^{+y}_{ijk}T, 0) + max(D^{-z}_{ijk}T, -D^{+z}_{ijk}T, 0)]^{1/2} = \frac{1}{K_{ijk}} \tag{5.5}$$

where D^{d}_{ijk} are the partial derivatives of T in the given d direction at position (i, j, k) (Figure 5.2). The FMA allows to solve this scheme systematically using only the upstream values. The upstream differences structures of equation 5.5 allow to propagate the information only "one way", i.e. from the smallest values of T toward the higher ones. The FM methods are based on the construction of the solution of T in the downstream direction by *freezing* the upstream values, which are already computed (i.e. the *accepted* values). The key point of the method relies on the way of selecting the grid nodes that have to be computed[3].

Figure 5.3 shows the selection of the grid points to be computed in a downstream scheme. The first step is the initialization (Figure 5.3a). In this step the initial value $T = 0$ is given to the starting position $\boldsymbol{x_0}$. The initial frozen value is represented by a black point. In the second step (Figure 5.3b), the T values for the neighbors of $\boldsymbol{x_0}$ are computed. After that, the minimal value is searched (Figure 5.3c). In this case we assume that the point A has the smallest value. Then (Figure 5.3d) the values of the neighbors of A are computed and then the algorithm looks for the smallest value (Figure 5.3e) in the active points (gray points), in this case point C. The value at point A is then frozen. The values of the neighbors of point C are updated following the rules defined in eq. 5.5. This procedure is repeated until the value of T at all points has been simulated. The frozen points (black points) are the accepted values, the gray points are the *active points*, i.e. the points which value may change during the update process given by the front evolution.

5.3 SKS usage of the Fast Marching Algorithm

This section explains in detail how the FMA is implemented and used in SKS. The first subsection explains the use of a regular grid to implement the FMA (Section 5.3.2), because this has some implications on the final shape of the conduit networks, but also on the efficiency of the execution. In Section 4.2.4 it is explained that the FMA is used in SKS to build a time of flight map, that will be used by *walkers* to find the minimum effort paths between the inlets and outlets of the karst system. The present section describes also the implementation of these walkers (Section 5.3.3) and how they are used to build the karst indicator function $I_c(\boldsymbol{x}, \kappa)$ described in Section 4.3.4 (eq. 4.12). It also describes how the contrast between the different speed medium (Section 4.3) can influence the geometry of the final network (Section 5.3.4), and how the rules on the choice of the inlets can have an impact on the final network hierarchy (Section 5.3.5).

[3]Several implementations exist to enhance the efficiency of the algorithms, the reader is suggested to refer to Sethian (2008) for further details.

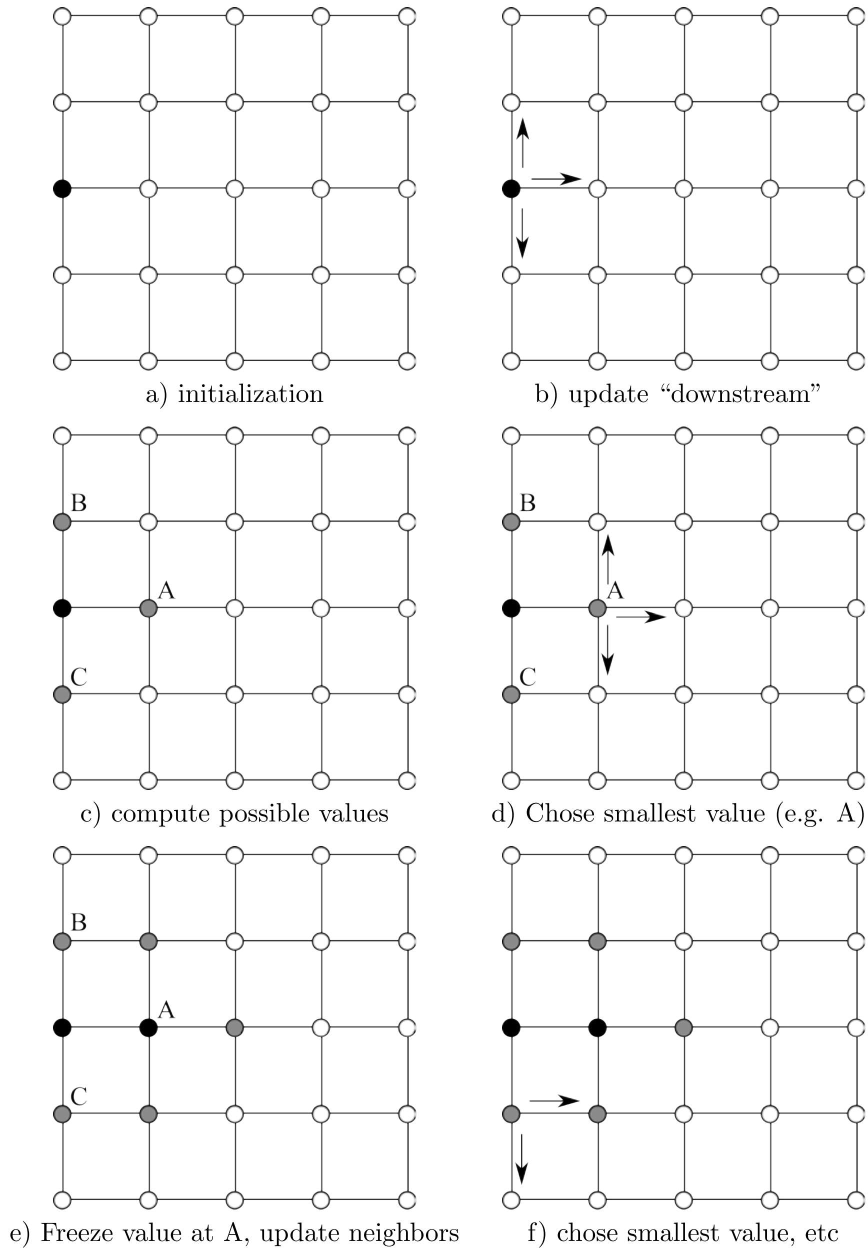


Figure 5.3: Fast Marching Algorithm downstream computing scheme in a 2D grid. Dark points: frozen values, gray points: active points; white points: downstream not computed points

5.3.1 Matlab programming language choice

SKS is written in MATLAB®. This programming language has been retained for the development of this prototype, because it is a high level interpreted language, with a multitude of already built-in functions (especially for visualizing early results). For the development of a prototype algorithm it is very useful to be able to quickly test new ideas. Furthermore, MATLAB® allows an easy use of matrix operations, which is a very practical tool when using indicator functions as it is done here. Moreover, an existing implementation of the Fast marching Algorithm for MATLAB® (*Toolbox Fast Marching*, Peyre, 2008) was available when the project started in 2008.

5.3.2 Regular grid to support the computation

As Sethian (2008) explains, there are many different ways to implement the Fast Marching Methods. The most advanced ones work with unstructured meshes. In the case of unstructured meshes, the time T

```
while 1
  % find the minimal value in the surrounding elements
  minT = min(T(i-1:i+1,j-1:j+1,k-1:k+1));
  % retrieve the index of the element
  ind = find(T(i-1:i+1,j-1:j+1,k-1:k+1) == minT);
  % stop if reaches the minimal value of T
  if T(ind) == min(T);
    break
  else
      % update the index of the walker
    [i,j,k] = update_current_index(i,j,k,ind);
  end
end
```

Listing 5.1: Pseudo-matlab code of the basic walker

front evolution depends not only on the speed function K, but also on the distance between the different nodes of the grid. The algorithms working with 3D unstructured meshes may have to integrate the speed function for several geometries of elements (like tetrahedrons, pyramids) which can slow down the execution. Depending on the implementation, this additional complexity can be felt in the efficiency of the algorithm.

Even if the *Toolbox Fast Marching* (Peyre, 2008) provides the functionalities to work with unstructured meshes, SKS has been coded on a regular grid for 3 reasons:

1. Efficiency
2. Simplicity of use
3. conversion from/to other softwares

Using a regular grid, the FMA gains in efficiency, because the distance between nodes is implicitly known with the mesh discretization, and there is no need to consider it in the time computation. Furthermore, a regular mesh is easier to handle (especially in MATLAB®) as it can be treated like a 2D/3D matrix, where the properties at given matrix indexes directly correspond with the elements. In addition, the use of a regular mesh allows to ease the model transition from the geological editor to SKS, and from SKS to the flow simulators (see Chapter 6).

5.3.3 Walkers

The concept of using walkers on a time field T to build the karst indicator function $I_c(\boldsymbol{x}, \kappa)$ (eq. 4.12) has been proposed in section 4.3.4. The present section gives a brief clarification on the different ways the walkers have been implemented in SKS.

The first attempts used walkers that used the gradient of the time field T. This kind of walkers had several problem to reach the the minimum of the time map, because the gradient of this kind of maps is very discontinuous, with very sharp value differences between the cels, if the contrast of speed K are high. Finally the walkers implemented in SKS are very simple (and fast): in an infinite *while loop* the walker scans the values of the neighbors elements and moves to the element with the minimal value. If the minimal value of the complete field is found ($T = 0$), the loop ends. Listing 5.1 shows a pseudo-code (pseudo-matlab) of this simple implementation.

The stop conditions of the main loop depend on the kind of conduits being simulated. E.g. if the conduits that are being simulated are the ones corresponding to the vadose indicator function $I_{vadose}(\boldsymbol{x}, vc)$ (Section 4.3.4, eq. 4.17) then the walkers have to stop also if a non-karstifiable formation is reached. This is important during this step to avoid conduits to cross non-karstifiable formations, while they are being simulated mainly vertically. Another stop condition is when the walker cannot find a lower value in the surrounding cells. This happened quite often with gradient-based walkers but with the simple scan it is very rare. The FMA implementation made by Peyre (2008) has a predilection for the diagonal directions. If we use a walker that looks for the minimal value in its complete surrounding cube (i.e. the 26 surrounding cells in 3D, and the 8 surrounding pixels in 2D), in a homogeneous K field (figure 5.4a),

a) b) c)

Figure 5.4: Differences in the way walkers find the path in the FMA time field. a) time field in an homogeneous speed field. b) using corner and faces connectivity. c) using only face connectivity

the walker will go straight forward to the T_0 value as shown in figure 5.4b. It can be noticed in this figure that the pixels are connected from the corner. If the simulated conduits have to be used in flow simulation (Chapter 6) it is convenient to avoid this kind of connection, and prefer a face to face connectivity (i.e. in 3D the elements will be connected by the faces, and by the edges in 2D) as shown in figure 5.4c. This kind of connectivity is particularly necessary for finite-differences solvers like MODFLOW. The path shown in figure 5.4c is not straight forward, this is because the walkers have to find the minimum locally, and because in an homogeneous field like this one the preferential values are on the diagonal, there are confusing points with the same value. Therefore, if more surrounding values of T are equal (and are the local minimum) the walker select one of them randomly. This adds a little additional variability to the simulated networks.

5.3.4 Speed function definition

As explained above (Section 5.2.1) all the front propagation methods need the definition of a speed function which may be very complex (eq. 5.1). The speed function in SKS is very simple and depends only on local parameters. In fact, the use of the FMA that is made in SKS is quite particular, because in this case, the return value of T is only used as a fast approximation allowing to integrate geological information and to provide the minimum effort paths between inlets and outlets of the karst network (see Section 4.2). What is relevant is that the time field T within the model is influenced by geological features. Therefore, the speed function K in SKS is only a speed associated to the different facies (aquifer, aquitard, fractures). Using the FMA in this way means that only the contrasts of speed in the different facies influence the propagation of the time field.

Figure 5.5 illustrates the effect of different contrasts in speed for a very simple synthetic case. In this 2D example, a more conductive formation (in black, figure 5.5a) is added into an uniform $K = 1$ speed field (white). Two speed contrasts are illustrated. Figure 5.5b and figure 5.5c show the time map T and the path of the walker for a contrast of 5 ($K_{black} = 5 * K_{white}$) respectively. Figure 5.5d and figure 5.5e shows the same results for a contrast of 2 ($K_{black} = 2 * K_{white}$). In figure 5.5c, the path of the walker passes preferentially through the black formation, while in Figure 5.5e its influence is smaller. As shown by this small example, the relative contrast in speed between different objects can have an important impact on the shape of the final minimum effort paths.

As explained in Section 4.3.1, the geological information, such as the presence of karstic formation and aquicludes is implemented in SKS using indicator functions ($I_G(\boldsymbol{x}, g)$ for the geology eq 4.1, where g is the given geological formation and $\boldsymbol{x}$ is a given point of the model). To each geological formation g correspond a value for the FMA speed $K_G = \{\kappa_1, ..., \kappa_{N_g}\}$ (N_g is the number of geological formations). The contrast between the values corresponding to the different formations is the key factor that influences the large scale distribution of the conduit networks. As shown in figure 5.5, the conduits will develop mainly in the formations with a high value of K. For practical reasons, the K value of the aquifer formation is noted K_a here after.

In practice, the karstifiable formations should have a relative high speed K value and the non karstifiable formation a low value. To produce consistent models, the outlets and inlets of the system are

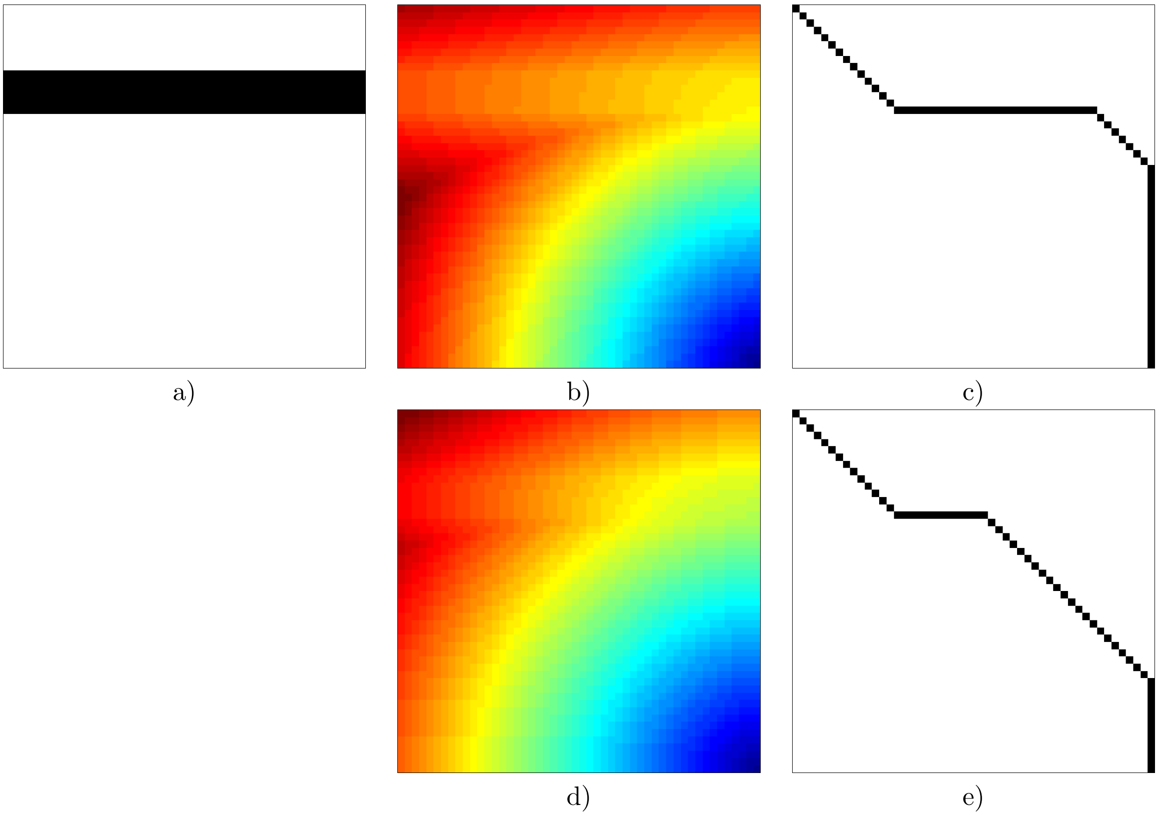


Figure 5.5: Effect of a higher velocity layer. a) black: the more conductive formation. b) T field with a contrast of 5 between the facies. c) path of the walker with a contrast of 5. d) T with a contrast of 2. e) path of the walker with a contrast of 2.

both put on aquifer formations. If the constrast is high, the conduits will be highly constrained by the karstifiable formations (tests have demonstrated that a contrast of $K_{aquiclude} = K_a/100$ is a good order of magnitude).

One solution to simulate the case when there are several aquifer formations and when one is known to be more karstified, is to put a small contrast between them. A contrast of 2 or 3 is in most cases sufficient to constrain all the conduits into the karstifiable formation. For example, let's assume two karstifiable formations: a_1 and a_2. a_2 being the most karstified formation, then $K_{a_2} \geqslant 2K_{a_1}$. In this way, the conduits will still cross the less karstifiable formation a_1, but will be slightly preferentially attracted by the second aquifer a_2. Big discontinuities, such as faults and thrust structures, may be preferential karstifiable zones, because the rocks are highly tectonized in these features. These are treated as a particular case in SKS which is described in the next paragraph.

The smaller scale (local) shape of the conduits is controlled by the discontinuities of the karstic aquifer. These are the particular horizons identified by Filipponi et al. (2009) as *inception horizons*. They include fractures and bedding planes. In the last implementation of SKS, these features are now considered in a slightly different manner compared to Chapter 4, as described in the following paragraphs.

Faults

Large scale tectonic accidents, like major faults, but also thrust structures, can have a role in speleogenetic processes. Because these accidents are highly fractured, the karstification can occur preferentially in this kind of features (Palmer, 2007). They are included in SKS by discretizing the fault and defining an indicator function I_{faults} for all the voxels of the model that are touched by the fault:

$$I_{faults}(\boldsymbol{x}, F) = \begin{cases} 1, & \text{if fault } F \text{ is present at location } \boldsymbol{x} \\ 0, & \text{otherwise} \end{cases} \tag{5.6}$$

This discretization is simply done by exporting the fault as a vectorial point set with the form

$\boldsymbol{x_{Faults}} = (x, y, z)$ from the geological editor and by computing the corresponding indexes (i, j, k) of the elements that are touched by these points in the regular grid. To use this simple way of discretizing a complex surface, it has to be output sufficiently finely from the geological editor to connect all the elements that are "crossed" by it.

If these faults have a regional scale (e.g. the SW thrust of the Valley of *les Ponts-de-Martel*, Chapter 2) the contrast between their associated speed K_{faults} and the aquifer speed K_a should not be too high, because otherwise they could lead to strong artifacts, like conduits running around the whole model before reaching the spring. For the Valley of *les Ponts-de-Martel* model, a contrast of 5 has been used between the aquifer speed and the faults speed (i.e. $K_{faults} = 5K_a$). If faults are smaller, i.e. their size is similar to the biggest fractures in the model (Section 3.2.2), it is possible to treat them as fractures.

Fractures, bedding planes and inception horizons

Fracture and bedding planes play an important role in the initialization of karst network as they can be considered *inception horizons* (Filipponi et al., 2009). As explained in Section 4.3.4, they are considered using the same indicator function for local discontinuities I_F (eq: 4.2). In the present section the term "fracture" include both fractures and bedding planes.

From the initial implementation of SKS described in Chapter 4, considerable work has been done in enhancing the fracture model, and the way there are used in SKS. This effort lead to the development of the fracture generator presented in Chapter 3. After the fractures are built with this fracture generator (the bedding planes are modeled as fractures parallel to the stratification), they are discretized in the model as described in Section 3.2.6.

Using this new version of the fracture generator, the fracture indicator function I_F described in section 4.2 is now used differently. In the previous model (equation 4.10), the speed field $K_0^S(\boldsymbol{x})$ was given as a sum of the indicator function for the geology I_G multiplied by possible values for the geological formations $K_G(g)$ and the indicator function of fractures $I_F(f)$ (f being the fracture family) as shown again here:

$$K_0^S(\boldsymbol{x}) = \sum_{f=1}^{N_f} I_F(\boldsymbol{x}, f) K_F(f) + \left(1 - \sum_{f=1}^{N_f} I_F(\boldsymbol{x}, f)\right) \sum_{g=1}^{N_g} \left(I_G(\boldsymbol{x}, g) K_G(g)\right) \tag{5.7}$$

where $K_G = \{\kappa_1, ..., \kappa_{N_g}\}$ and $K_F = \{\kappa_{F1}, ..., \kappa_{FNf}\}$ are 2 vectors of length N_G and N_F (number of geological formations and number of fractures families). This implementation excluded the possibility of having multiple families of fractures. When $I_F(f)$ was defined for many families f, the value of K would become excessively high because of the sum of their values. Therefore, the implementation of the indicator function $I_F(f)$ has been redefined. Now a property matrix $\mathcal{F}$ is defined for the entire model domain as follows:

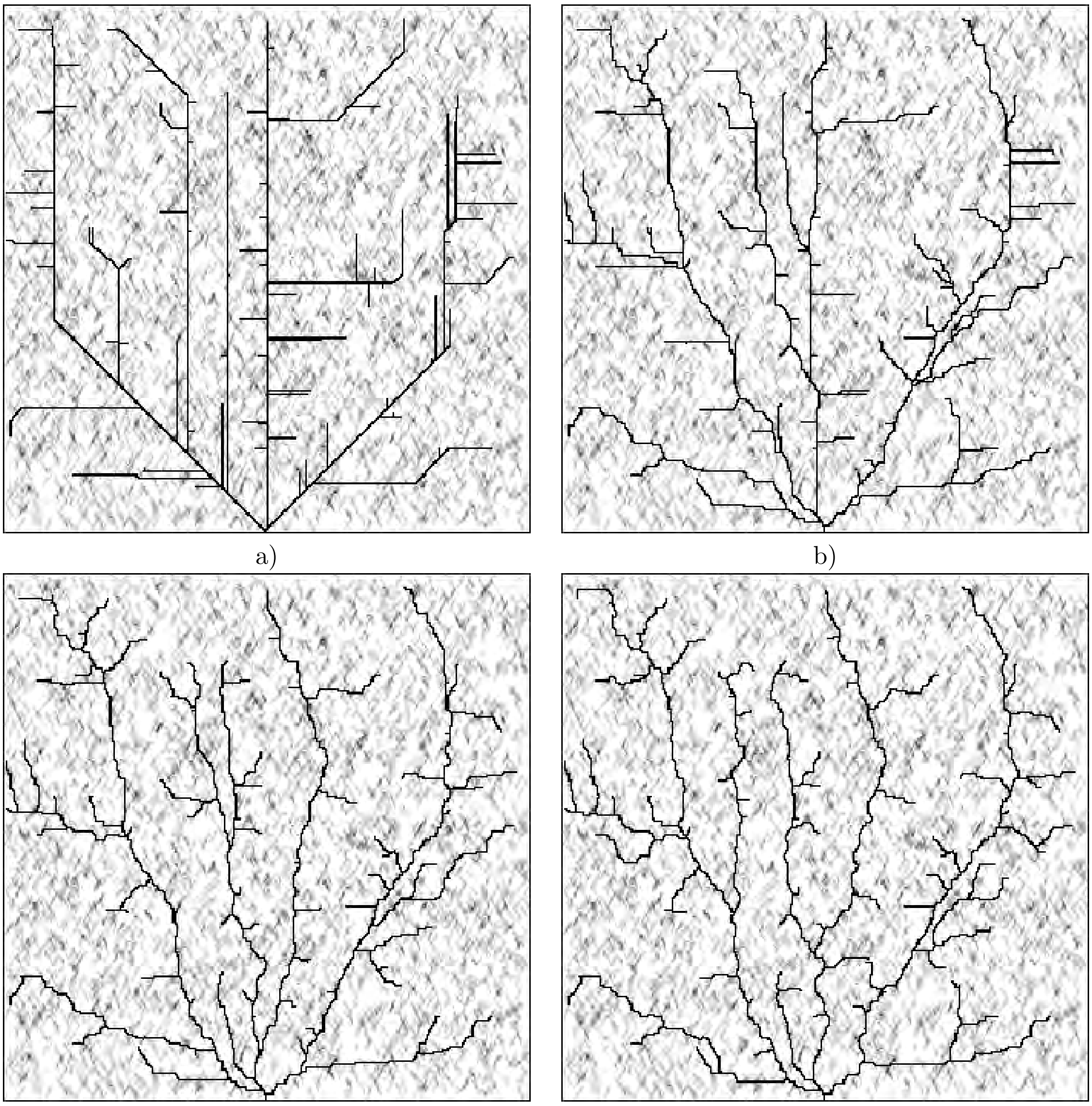
$$\mathcal{F}(\boldsymbol{x}) = \sum_{f=1}^{N_f} I_F(\boldsymbol{x}, f) \tag{5.8}$$

where f represents now *each fracture*, and N_F the total number of fractures. The value of K_F at location $\boldsymbol{x}$ ($\forall \boldsymbol{x} \in \{I_F = 1\}$) is now defined as:

$$K_F(\boldsymbol{x}) = \frac{\mathcal{F}(\boldsymbol{x})}{\max(\mathcal{F})}(K_{Fmax} - K_a) + K_a \tag{5.9}$$

where K_{Fmax} is the user defined maximal possible value for the FM speed field of fractures K_F and K_a the aquifer speed. This formulation has been chosen, because the role of fractures is principally to enhance the ability of the karstifiable formation to be karstified. In this way, the elements with a higher number of fractures will be more preferential than the elements with less fractures. For the model of the Valley of *les Ponts-de-Martel* a contrast of about 5 to 10 is usually employed (i.e. $K_{Fmax} = cK_a \mid c \in [5, ..., 10]$). Figure 5.6 shows the effect of an increasing speed contrast between the aquifer and the fracture facies. As it can be seen, with an increasing contrast, the fractures control more the local shape of the conduits. Moreover, this example also shows that the topology of the modeled karst network (the way the conduits are connected together) changes in consequence to this increasing contrast.

The fracture density has also an influence on the topology and shape of the conduits. Figure 5.7 shows the effect of increasing fracture density over the modeled network. It can be noticed, with an increasing fracture density, the local shape of the models is increasingly driven by the fractures and



a) b)

c) d)

Figure 5.6: Effect of the FMA speed contrast between the FMA value for the fractures K_f (gray facies) and the FMA value for the aquifer (white facies) $_a$. a) $K_{Fmax}/K_a = 1$ (homogeneous), b) $K_{Fmax}/K_a = 3$, c) $K_{Fmax}/K_a = 6$, d) $K_{Fmax}/K_a = 9$. We notice that with increasing contrast, the facies *fractures* has a higher influence over the conduit shape (black facies).

follows the orientation of the fractures. In fact, if the fracture density becomes very large (Figure 5.7i), the simulated network tend to homogeneous case again (Figure 5.7a), because there is quite no more contrast between the aquifer and the fractures. The maximal tortuosity is obtained in the intermediate cases.

Also the orientation of the fractures has a strong influence on the conduits orientation. The effect of the fracture orientation can be seen in figure 5.8: two sets of discrete fracture networks have been modeled. In figure 5.8a, the two main fractures families have an orientation of N0° and N90° respectively, and in figure 5.8b the two main families have an orientation of N60° and N120° respectively.

In the case where some specific layers of the karst formation can be considered as preferential for karstification (altered layers), they can be included in the model directly on the base of the geological model, as shown in figure 5.10: the K value of theses specific layers can be set with a higher value with respect to the one of the aquifer and the karst network will preferentially occur into them.

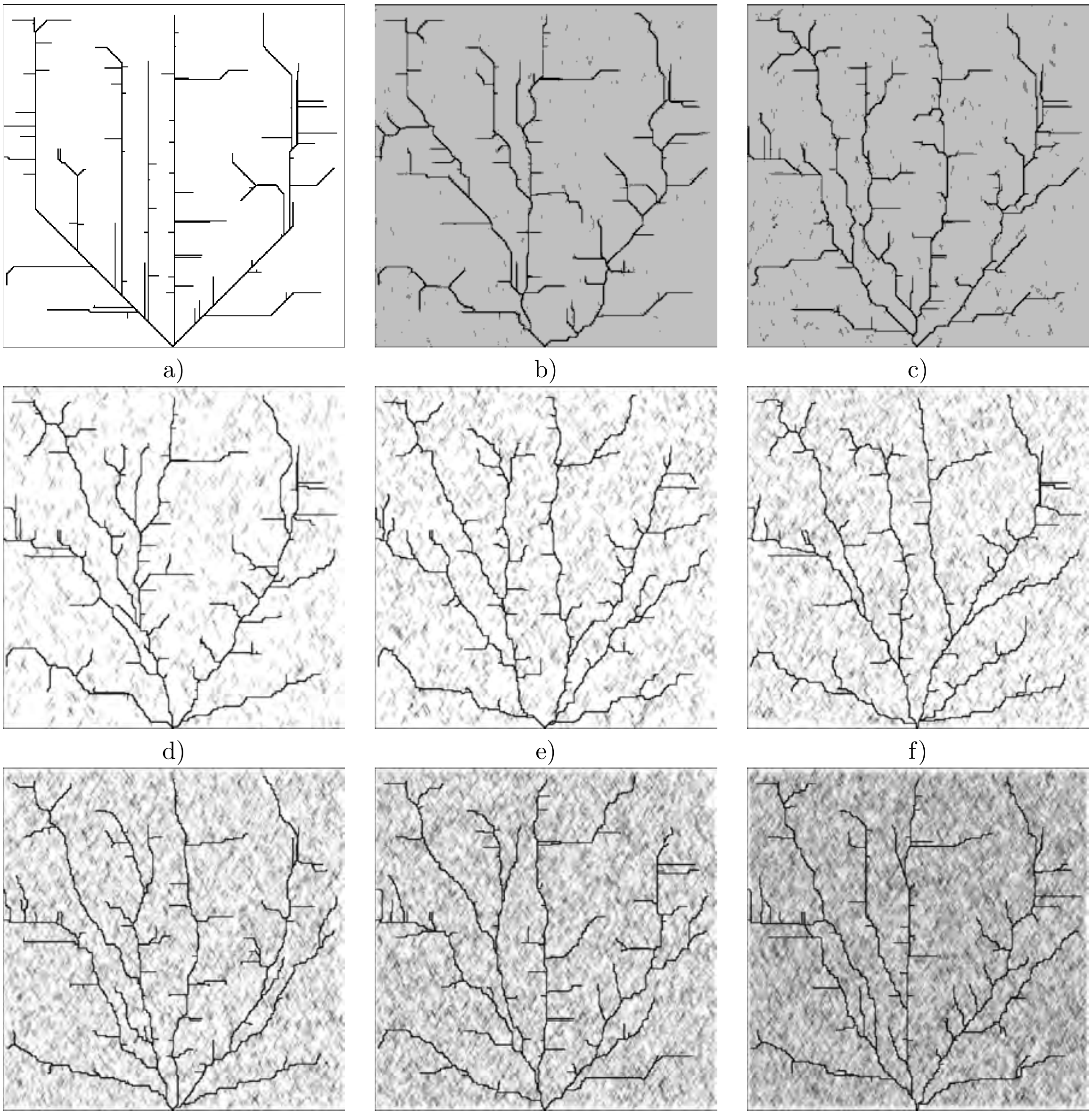


Figure 5.7: Effect of the fracture density over the final shape of the network. The fracture density f_d in this case is computed as $f_d = N_f/N_e$, with N_f the number of elements with fractures and N_e the number of elements. a) $f_d = 0$ (homogeneous), b) $f_d = 0.03$, c) $f_d = 0.06$, d) $f_d = 0.13$, e) $f_d = 0.24$, f) $f_d = 0.43$, g) $f_d = 0.68$, h) $f_d = 0.88$, i) $f_d = 0.99$

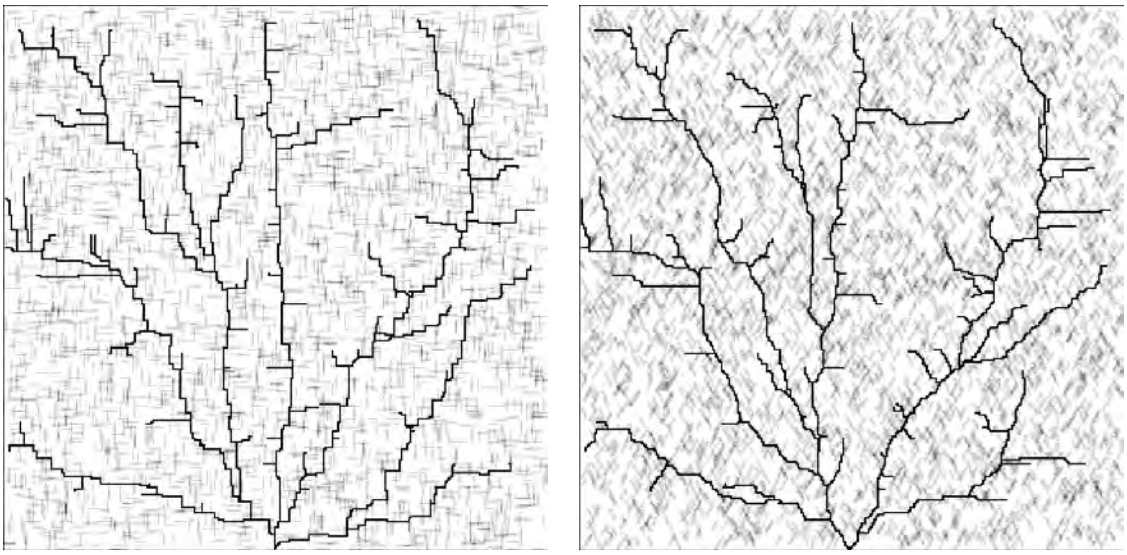
5.3.5 Impact of the inlet selection over the final shape of the network

To build a hierarchic network of conduits, SKS uses several iterations. This concept has already been explained in Section 4.3.4 and is used also by Collon-Drouaillet et al. (2012). The present subsection clarifies the rules that govern these iterations.

Equation 4.13 shows that the K_i^S field of each iteration θ is updated using the indicator function I_c of the presence of conduits (eq. 4.12). The key point is that these *iterations* are made partitioning the set of starting point locations $\Im^S$ (eq. 4.5) in several subsets $\Im_\theta^S$ as follows:

$$\Im^S = \left\{\Im_1^S, ..., \Im_\theta^S, ..., \Im_{N_\theta}^S\right\} \quad , \quad \forall\, \theta \in 1, \ldots, N_\theta \tag{5.10}$$

with N_θ the total number of iterations. The number of elements in the subset $\Im_\theta^S$ is the integer part



a) b)

Figure 5.8: Effect of the fracture orientation over the final shape of the network. This is shown with two different fracture models composed by two fractures families: a) N0° and N90°b) N60° and N120°. We notice that the orientation of the fracture families influence the the conduit orientation. In both cases the preferential orientation of the conduits is the same as the fracture families orientation.

of the fixed proportion of the total number of inlets $N^{\Im}$:

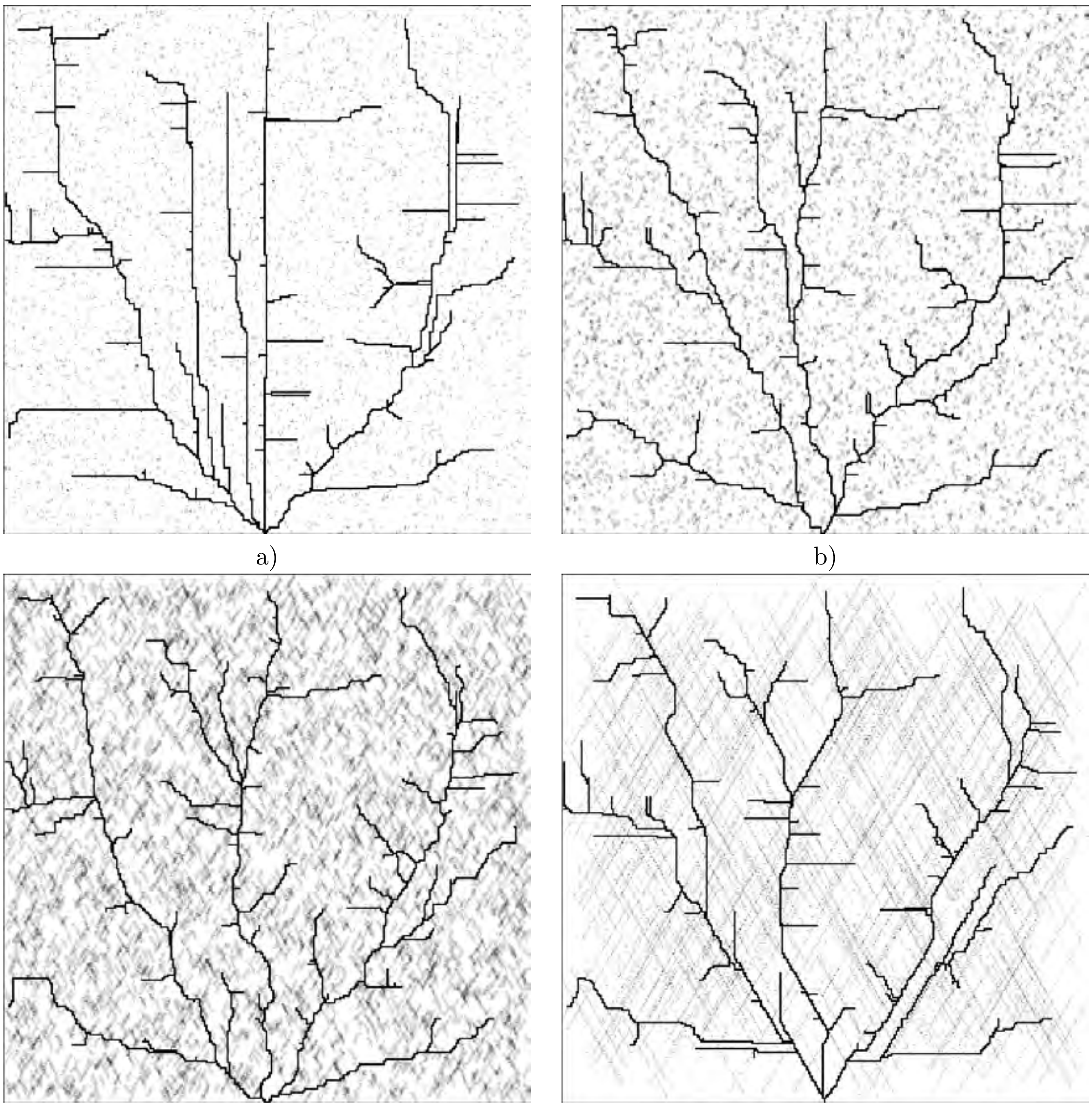
$$N_{\theta}^{\Im} = \left\lfloor \frac{N_{\Im}}{\sum_{i=1}^{N_\theta} \mathcal{F}_i} \mathcal{F}_\theta \right\rfloor \tag{5.11}$$

where $\boldsymbol{\mathcal{F}} = \{\mathcal{F}_1, ..., \mathcal{F}_{N_\theta}\}$ is a vector called *importance factor* of length N_θ. This vector is provided by the user and contains the percentage of inlets that have to be simulated at each iteration θ. The number of inlets (i.e. the number of conduits) $N_{\theta}^{\Im}$ at each iteration is a portion of the total number of inlets corresponding to the θ^{th} element of $\boldsymbol{\mathcal{F}}$.

The networks presented in figure 5.11 shows the effect induced by the FM speed K contrast between the aquifer (K_a) and the fractures (K_f), combined with the hierarchic ordering of the inlets controlled by the parameter *importance factor* $\boldsymbol{\mathcal{F}}$ (eq. 5.11) on the network connectivity. In the first column, all the conduits are simulated in only one iteration of the *pseudo genetic* algorithm. This lead to poorly hierarchized network. In the second column, one third of the inlets is used in the first iteration, and two thirds are used in the second iteration. The firstly simulated conduits *attract* the next ones. The third column show the effect of a more hierachized network, where only 1/7 of the inlets are used in the first iteration, 2/7 in the second, and the majority (4/7) are simulated in the last iteration. The conduits generated in the first iterations now have a strong attracting role on the next ones.

Figure 5.12 shows 2D simulation results obtained with 35 inlets. To add more variability in the simulations, the order of the inlets can be *mixed*. The way it can be done is by randomly permuting the order of the inlet coordinates vectors as shown in Listing 5.2. Figure 5.12 show the effect of randomly permuting the starting point coordinates (inlets) on the same inlet points set. It can be noticed that the main conduits are different, they depend on the first that are simulated. This simple example illustrates the strong *attracting* effect of the early simulated conduits on the final result. If some inlets are considered as more important than others, they can be fixed explicitly in the earlier iterations by the user.

This example shows that SKS offers a high degree of flexibility to generate karst networks, but the user has to be aware that the various input parameters will interact in several manners. Their interactions will determine the degree of tortuosity of the networks, their topology, their orientations, etc. This flexibility allows the user to test many different hypothesis very quickly, and allows to generate very different kinds of networks easily.



a) b) c) d)

Figure 5.9: Effect of the fracture length on the shape of the simulated network. All the fracturation models have 2 fracture families (N60° and N120°). Four length distributions are shown: a) from 1 to 20m, b) from 10 to 500m, c) from 50 to 1000, and d) from 500 to 2000. With increasing length of the fractures, the conduits are more continuous according to the orientation of the fractures.

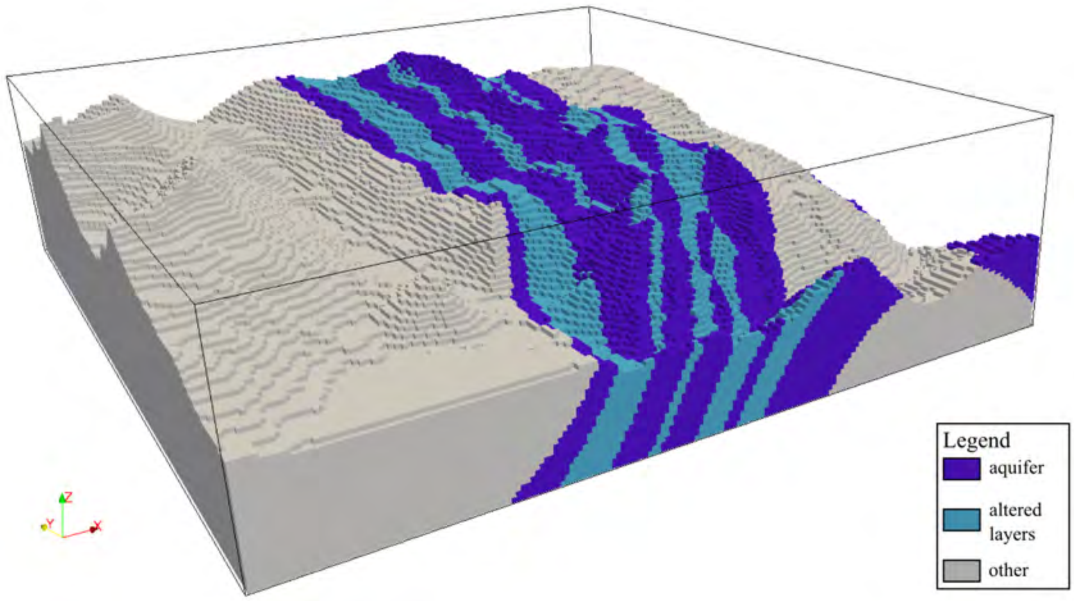


Figure 5.10: Some layers of the karst system may be more altered, and consequently used as preferential for the karst formation, by assigning them a higher K

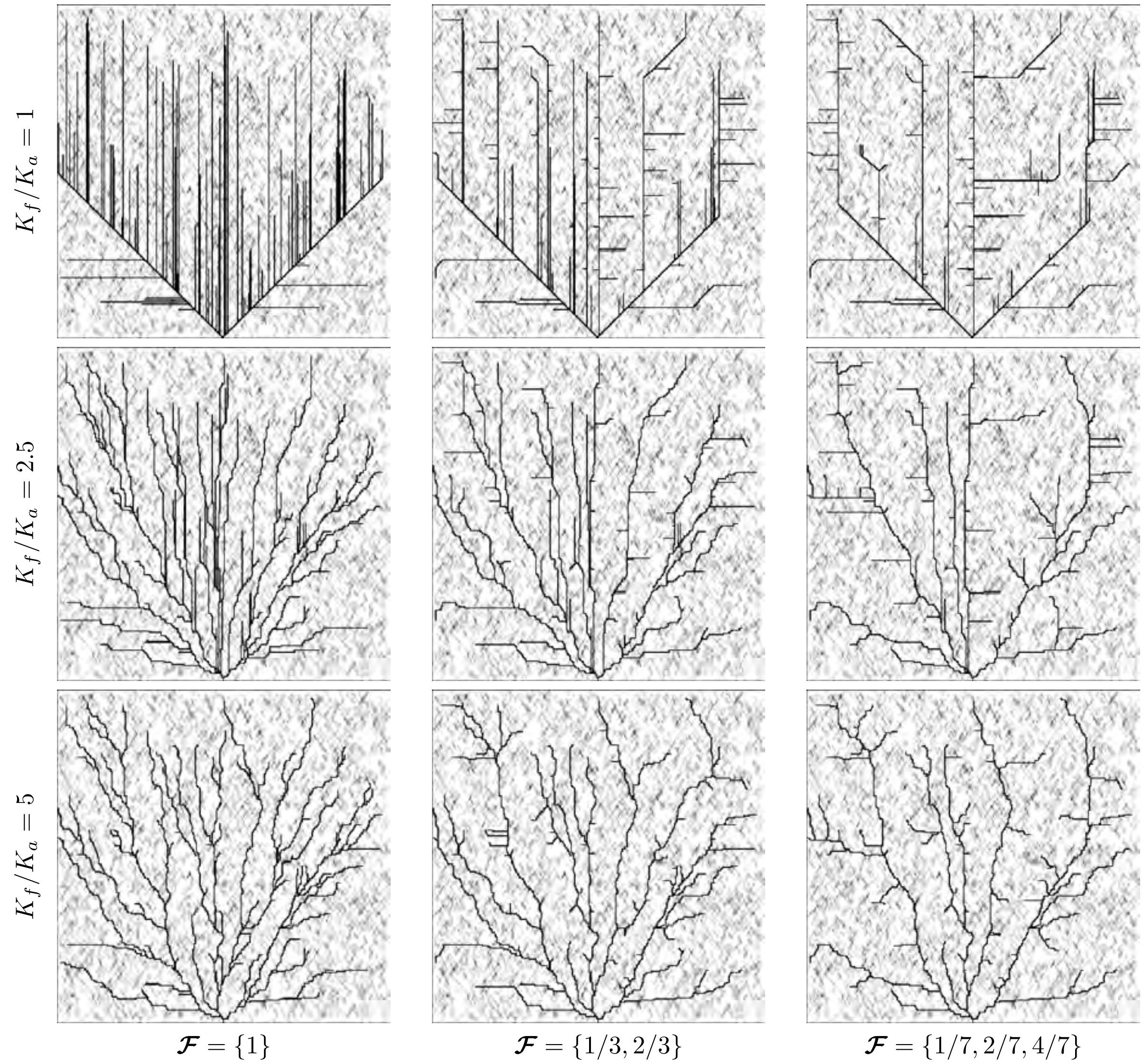


Figure 5.11: Effect of hierarchic inlet selection combined with the FMA speed K increasing contrast between the FM value for aquifer K_a and the FM maximal value for fractures K_f. The same inlet points are used for all realizations. The hierarchy is determined by the parameter *importance factor* $\mathcal{F}$ (eq. 5.11) which length correspond to the number of iterations of the algorithm and the values correspond to the fraction of the total inlet points that are simulated during each iteration.

```
n=size(input,1); % size of the input coordinate triple array (x,y,
    z)
c=n-1; % counter for the position of the array that have been
    treated

% permute the current element with one randomly beginning from the
    end
for i=n:-1:2
    tmp = input(i,:);
    id = ceil(rand*c);
    input(i,:)=input(id,:);
    input(id,:)=tmp;
    c=c-1;
end
output = input;
```

Listing 5.2: permutation of the coordinates of the inlets to add more variability in the simulations

a) b) c)

Figure 5.12: Effect of permuting the inlet selection. The same inlet points are used for all realizations, the fracture network and speed K contrast are fixed. Permuting the computation order of each conduit leads to differently connected networks.

```
if k > kMax % if the walker tries to go upwards
  % use the subset of the original propagation medium that
  % is below the max k value
  subMedium = medium;
  % above the max k value, set K to 0
  subMedium(:,:,kMaxPossible+1:end) = 0;
  % re-run fast marching algorithm
  T = perform_fast_marching(subMedium,spring');
  continue
end
```

Listing 5.3: If the conduit tries to go upwards, the FMA is run again on a subset of the propagation medium

5.4 Forcing the conduits to go downstream

One artifact that has been observed using the pseudo genetic approach was the presence of conduits going upward. In Section 4.3.4 it was shown that to simulate the vadose conduits, i.e. to take into account the gravity effects that control the unsaturated flow in the vadose zone, the simulation is done in two distinct steps: the first from the land surface toward a baselevel, and the second from the intersection of the baselevel toward the spring. While this approach already solved partially the problem, additional work has been made to avoid some small artifacts that where still present. It could happen in the unsaturated zone. In the worst cases, the maximal altitude could be superior to the altitude of the corresponding inlet. This can happen if the vadose zone is not previously simulated.

Section 4.3.4 explains that the use of two steps, the first to simulate the vadose zone, and the second for the phreatic conduit solves this issue. This is only partially true. Further tests show that when simulating the vadose conduits in this way (i.e. first toward a baselevel) produce another kind of artifacts: the conduits first descend toward the baselevel and stop, but later they can “escape from the bottom” and “climb” again above the base level, sometimes using the vadose conduits for a while. This behavior is shown in figure 5.13a.

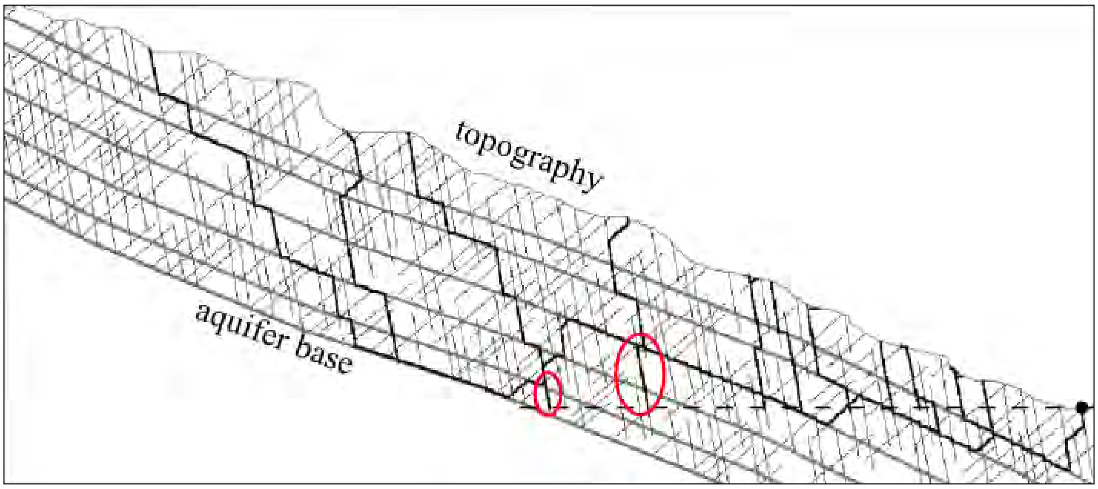
To avoid this behavior, a test is added now to the walker: if the walker tries to go above its starting altitude, then the part of the model that is above the starting altitude is removed from the model and the FMA is applied again but only into the submodel (Listing 5.3). This means that if the k index of the walker is higher than the k index of the inlet (vadose step), or if it is higher than the k index of the baselevel (phreatic step), the model is truncated above the starting altitude of the walker and the FMA is applied again. In this way, the preferential paths will follow only the geometry of the geological formations present below the starting point (figure 5.13b).

5.5 Diverging conduits

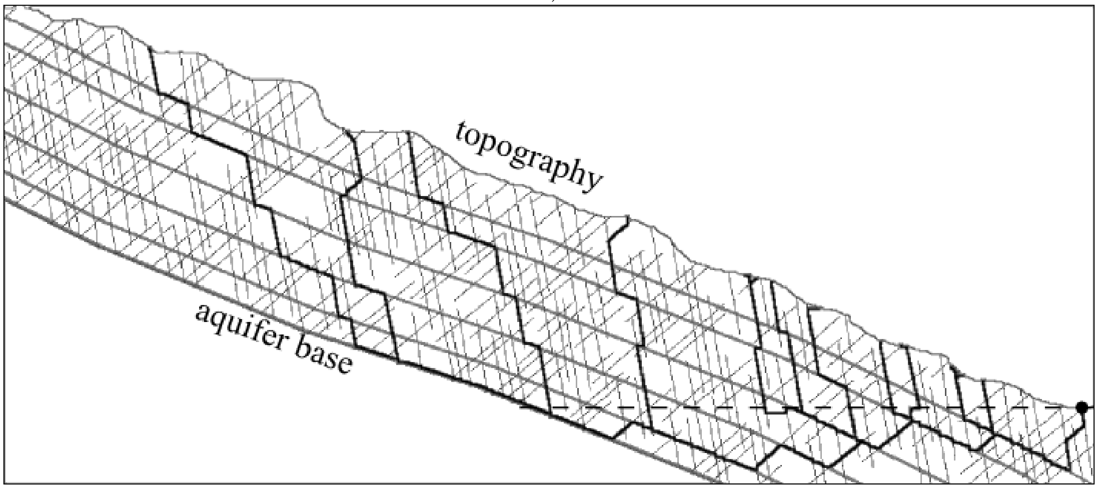
The connectivity of karst system is very complex. It is commonly observed that conduits have several interconnections (Palmer, 2007), like e.g. diverging conduits, multiple springs, multiple level of karstification, etc. The solution to the problem of several levels of karstification and multiple springs has already been solved in Section 4.3.4 and Section 4.3.4 respectively. One missing feature in the resulting networks produced with the first implementation described in Chapter 4 are the diverging conduits, which are typical of karst networks (Ford and Williams, 2007; Palmer, 2007). Like anastomotic and braided rivers, karst conduits can have diverging flow, i.e. in some part of the network the conduits do not converge, but diverge. With the approach described in Chapter 4 all the conduits can only converge, because after the first conduits are simulated, the next ones will be *attracted* by them (Section 4.3.4).

One first way of creating diverging conduits is to simulate successively the conduits toward different springs, as it is explained in section 4.3.4 and used in appendix C.

The present section shows how it is possible to create diverging conduits using *plugs* in the model between the different iterations of the algorithm. The idea comes from the field observations (in active and paleo conduits), that in some cases sediments can obstruct the conduit, and the water will find another path. While this is not always the case in real karst system, i.e. there are active diverging



a)



b)

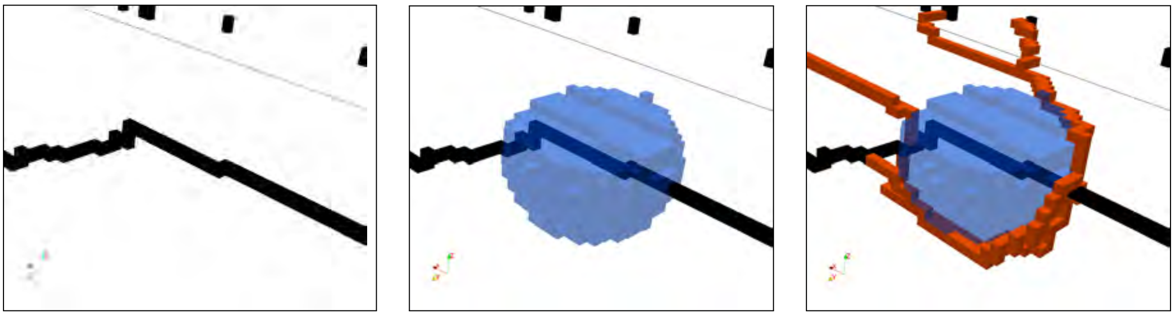
Figure 5.13: phreatic conduits (2D cross-section simulation results). Black point: spring. Dotted line: base level. Gray: inception horizons. Black line: karst conduits.. a) old version: artifact (evidenced by red circles): the phreatic conduits raise up again and are higher than the base level. b) new version: the phreatic conduits are constrained by the base level.

conduits that generate spontaneously, the way it is implemented in SKS is inspired from the plugging processes.

After the first subset of conduits are generated, plugs with $K = K_{aquiclude}$ speed are added temporary to avoid the next simulated conduits to pass through them. Numerical experiments have shown that it is not sufficient to simply plug the conduits, because the next ones will pass very close to the first ones. The aim here is to build conduits that will find alternative paths, which may diverge consequently from the initial ones. Therefore the plugs are not only added to the conduits, but to the propagation medium in a certain volume around the conduit. This low-K feature forces the new conduits to by-pass it.

Two different kinds of plugs have been implemented: spheric plugs (section 5.5.1) and gaussian plugs (section 5.5.2). Both ways will produce an indicator function for the presence of plugs I_p, defined as follows:

$$I_p(\boldsymbol{x}, p) = \begin{cases} 1, & \text{if plug } p \text{ is present at location } \boldsymbol{x} \\ 0, & \text{otherwise} \end{cases} \tag{5.12}$$



a) b) c)

Figure 5.14: a) previous conduit, b) spherical plug added on the conduit, c) next conduit by-passing the plug

5.5.1 Spheric plugs

Spheric plugs have been implemented as first trial. A given N number of plugs is modeled, for each plug, the center is chosen on a valid IJK_list index (Figure 5.14a), i.e. an index of the model where the indicator function for karst $I_c(\boldsymbol{x}, \kappa)$ is valid (Section 4.3.4, eq. 4.12), i.e. $I_c(\boldsymbol{x_{valid}}, \kappa) = 1$, with $\boldsymbol{x_{valid}}$ the valid IJK indices. The radius of the sphere is randomly sampled between two user defined values (r_{min} and r_{max}).

Then the distance matrix D between the selected center $c_{i,j,k}$ and all the other elements of the model is computed; X,Y, and Z being 3 matrices that contain the gravity center coordinates of all the elements in x,y, and z dimensions respectively. Finally all the elements that satisfy the condition $D < r$ (i.e. the elements included in the sphere radius) are flagged as *true* in a boolean output matrix M that will be used later as indicator function I_p for the plugs (eq. 5.12). Figure 5.14b shows one of these plugs in a 3D model. The conduit gently follows the sphere limit, and produce therefore a diverging conduit that reconnects to the original conduit later (Figure 5.14c). The spherical shape of the plugs is strictly followed by the conduit. In addition also Gaussian plugs have been implemented to add variability in the shape of the plugs.

5.5.2 Gaussian plugs

The goal of this second method is the same as the spherical plugs methods, i.e. to add *obstacles* into the FM propagation medium to force the conduits to by-pass them, creating more realitic networks. The only difference is that in this case, the shape of the plugs is not spherical, but based on the simulation of a gaussian object by fast Fourier transform (Dietrich and Newsam, 1993) with a normal law and a spherical variogram. The creation of these plugs works as follows:

1. 3D Gaussian simulation
2. quantile selection, extract binary matrix
3. identification of the geobodies
4. randomly sample one of them
5. compute gravity center of the selected geobody
6. "patch" the selected geobody "on" the conduit

The code relative to this section is shown in appendix A.1. The first step is to build a 3D non conditional Gaussian simulation of a dimensionless variable v (Figure 5.15a), having a mean $\langle v \rangle = 0$ and variance $\sigma^2 = 1$. It is then used to produce discrete geobodies by selecting only the region where v is greater than a user-defined quantile q (step 2, Figure 5.15b). In practice, the user can define the range of the variogram in the 3 dimensions (lx, ly, and lz), as well as a threshold, corresponding to a given quantile q of the variable v. This produce an indicator function of the form:

$$I_q(\boldsymbol{x}, v, q) = \begin{cases} 1, & \text{if the value of } v \text{ is superior to its quantile } q \text{ at location } \boldsymbol{x} \\ 0, & \text{otherwise} \end{cases} \tag{5.13}$$

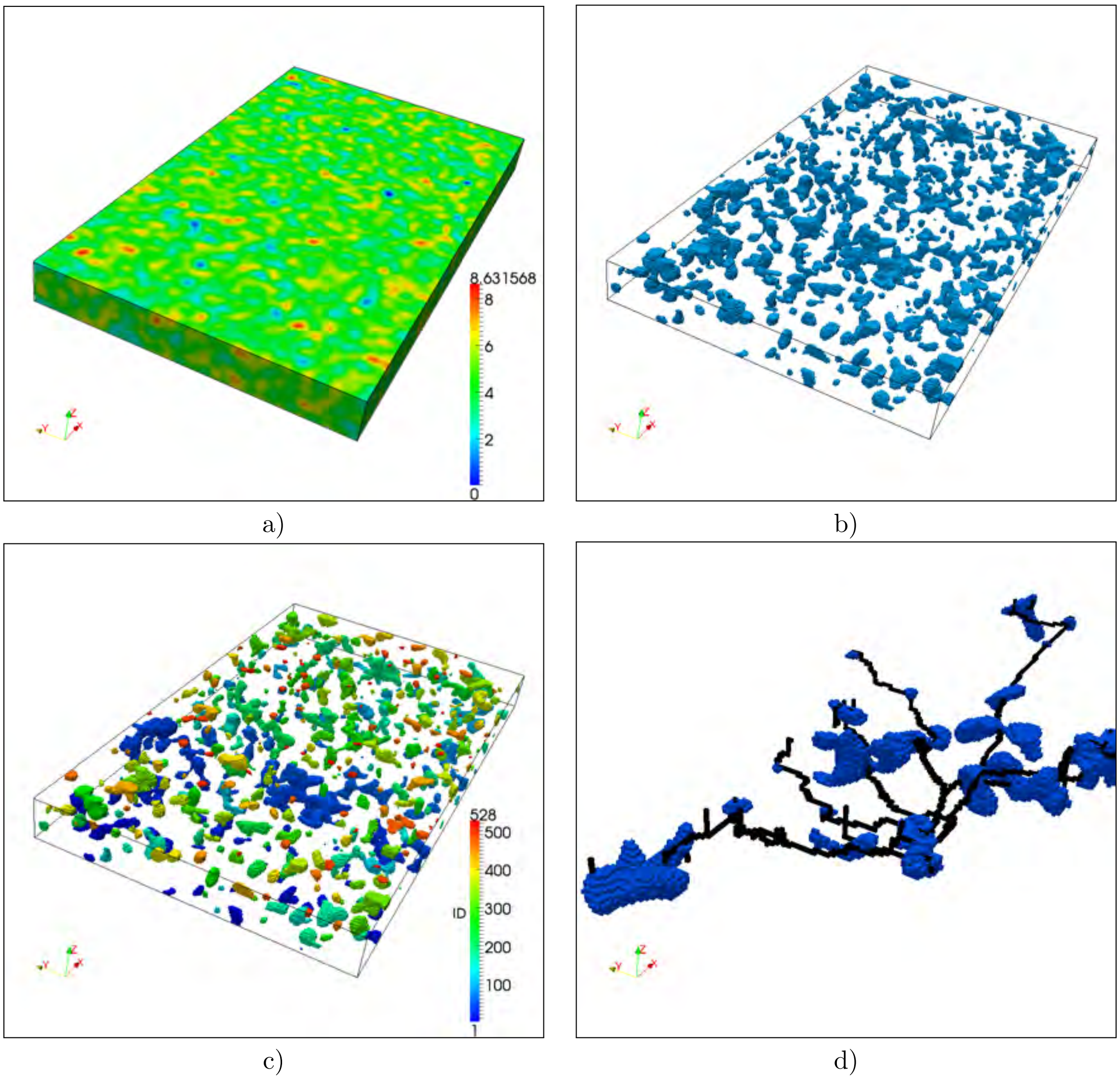


Figure 5.15: a) gaussian field, b) $v > q$ selected objects, c) identified geobodies (the color correspond to the ID of the geobodies), d) geobodies patched on the already existing conduits to plug the medium

The result of this step is a 3D logical matrix (Figure 5.15b). The third step is the identification of the geobodies defined by I_q; the MATLAB® *Image Processing Toolbox* is used for this step. After this step, each geobody is identified by an unique identifier which is an increasing integer (Figure 5.15c). One of them is selected randomly. Then the gravity center of the selected geobody is computed (step 5) as the mean of the X, Y and Z matrices where I_q is defined. If the shape of the geobody is not convex, which is highly probable with this kind of objects, its gravity center may not be included into the geobody. In this case the point which is included into the geobody and is the nearest to the gravity center is used as gravity center. This is important, because in the last step (step 6), the selected geobody is used as plug into the conduit model, and it is used to define the indicator function I_p (eq. 5.12). To do this, the plugs are *patched* on a randomly selected point of the previously simulated conduits (Figure 5.15d). The point of the network where the plug has to be patched is selected like in the case of spherical plugs, here above. To ensure that the plug *will effectively plug* the conduit, it is patched on its gravity center (or on nearest point to the gravity center included into the plug).

5.5.3 Discussion

The idea of adding plugs into the medium to generate labyrinthic networks is arguable because it cannot be fully related to natural processes. It is inspired by the fact that conduits can be filled by sediments, but in natural karst systems this is probably not the dominant process and there other kinds of diverging conduits. For example, multiple connections may occur depending on the local variations of hydraulic boundary conditions, but simulating this process is not achievable in a simple manner in SKS, unless many springs are used (section 4.3.4). The algorithm does not consider temporal variations in hydraulic regime, which may occur in real systems (e.g. local storms that provide water not uniformly, etc).

Adding plugs, as implemented in SKS, is one simple way of forcing a conduit to diverge from another pre-existing one using this *minimum effort path search* with the FMA. Therefore this feature has to be considered as an algorithmic proposition to increase the complexity of the generated networks, but without a valid conceptual justification. Further research and developments are needed to overcome this limitation.

5.6 Application to the Valley of *les Ponts-de-Martel*

This section shows the application of SKS to the karst aquifer of the Valley of *les Ponts-de-Martel* (Chapter 2). The only known spring of for this regional aquifer is the spring of *Noiraigue*. Many inlets (sinkholes and dolines) have given positive results when tracer tests have been performed (appendix B). Fournier (2012) has mapped all the inlets of the system (figure 2.7). The site is very tectonized and many regional faults and thrusts have been identified. These features have been identified as possible preferential karstification zones. Based on the geomorphological study, three karst modeling scenarios are proposed and used to demonstrate the use of the code.

The cells are cubic and have a size of 50m in the three dimensions. The conduits that are modeled are only the macro collectors (i.e. the largest conduits), which are supposed to have a regional influence.

Three models have been generated with similar parameters, but different sets of inlets depending on the assumption. All the models have been generated with 3 levels of hierarchy: the number of second order conduits is twice the first order and the number of conduits of third order is 4 times the number of conduits of first order, i.e. the *importance factor* of the simulations is $\boldsymbol{\mathcal{F}} = \{1, 2, 4\}$ (eq 5.11). Faults of regional importance are used in this model, together with the fracture model described in Section 3.3.2. The FMA speed values that have been used are listed in table 5.1.

name	$K_{aquiclude}$	$K_{aquifer}$	K_{faults}	K_{Fmax}	$K_{conduit}$
value	1	10	50	100	200

Table 5.1: FMA speed values used for the Valley of *les Ponts-de-Martel* aquifer model.

The three tested hypotheses are the following. In the first case (figure 5.16a), all the mapped inlets are used as true inlets of the system. In the second case (figure 5.16b), the smallest inlets are grouped to form *inlet zones* and in the third case, only the most important inlets (which have also been traced by dye tests) are used (figure 5.16c). The resulting networks are shown in figure 5.17.

The decision of these three models relies on the way flow will be simulated (Chapter 6). In the first case, all the conduits are considered relevant for flow, even the smallest. In the other two, the conduits that are generated represent principally the macro-collectors, while the karstic matrix is see as *fractured matrix*, which hydraulic parameters will be principally modeled as an equivalent porous medium.

As the Valley of *les Ponts-de-Martel* is a very complex and folded environment, several authors (Morel, 1976; Gogniat, 1995a; Borghi, 2008) discuss the relative importance of the different tectonic accidents (thrusts, see figure 2.4) present in the region. Using SKS it is possible to create different models that consider these thrusts. Figure 5.18 shows the results of these different possibilities: In the first model, only the southern thrust is considered; in the second model both Souther and Northern thrusts are considered. The resulting models show how the conduits are influenced by these large scale discontinuities.

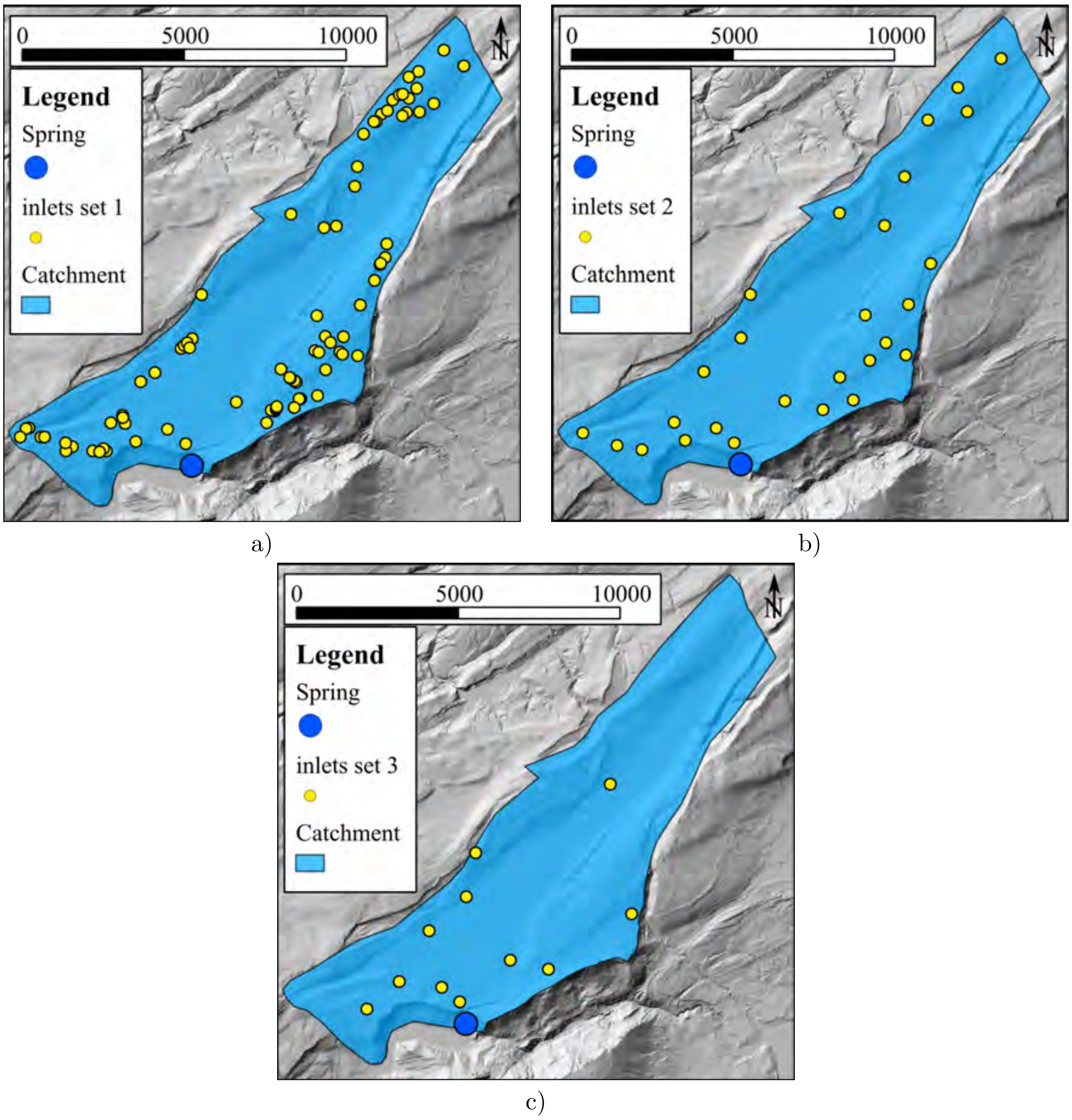


Figure 5.16: Three tested hypothesis regarding the inlets, a: set 1, all the mapped dolines; b: set 2, major infiltration zones; c: set 3, only the biggest (traced) inlets (from Fournier, 2012)

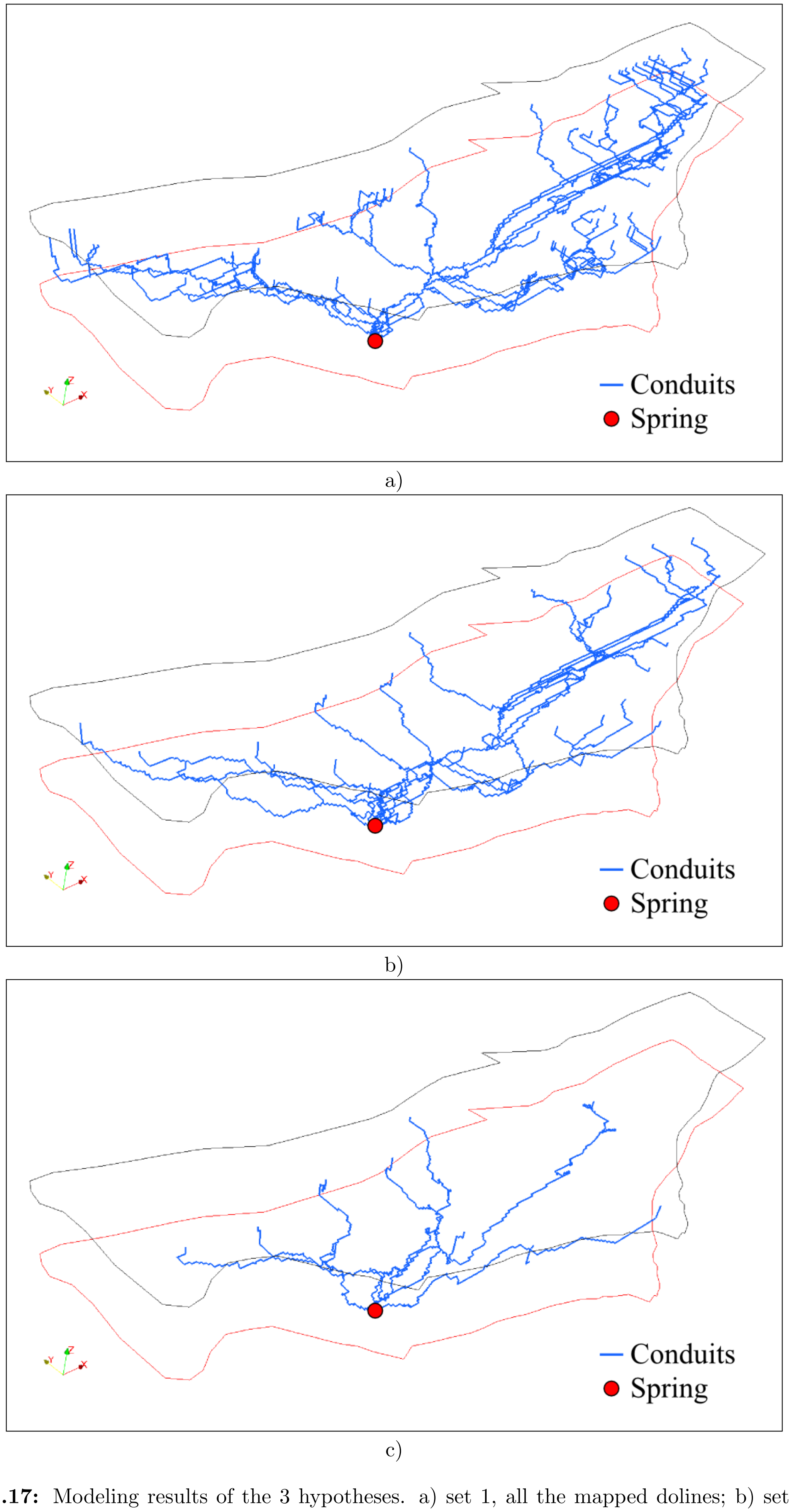


Figure 5.17: Modeling results of the 3 hypotheses. a) set 1, all the mapped dolines; b) set 2, major infiltration zones; c) set 3, only the biggest (traced) inlets (from Fournier, 2012)

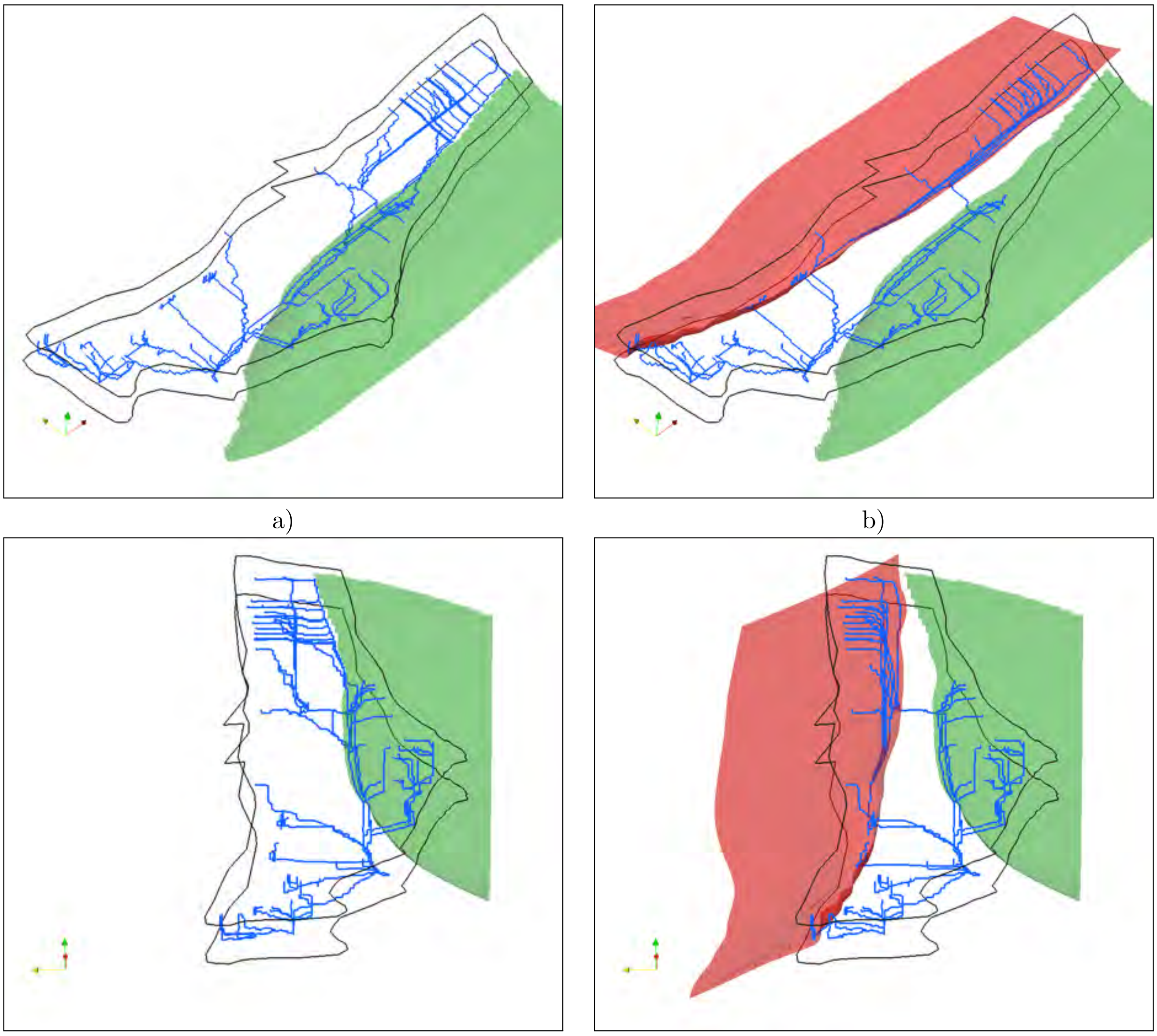


Figure 5.18: Two possible realizations, depending on the consideration of the large scale thrusts of the Valley of *les Ponts-de-Martel*. a) and c) show the model where only the southern thrust (green) is considered as possibly karstified; b) and d) show a model where also the Northern backthrust (red) is considered as possibly karstified. The consideration of the thrusts in the model result also in topological changes of the network: the way the conduits (blue lines) are interconnected is different, because they follow the different thrusts.

5.7 Consideration of geophysical data

SKS as been modified for the study of the karst aquifer of the Yucatan peninsula (Mexico) presented in appendix C. The procedure that has been used to simulate the conduits in the study case of the Yucatan peninsula is deeply explained in appendix C. This section shows only the SKS relevant considerations.

A great amount of aerial geophysical measurements (electro magnetic) was available for this case study, and the geophysical anomalies are present where the known caves have been mapped (figure 5.19). The caves are huge (several tens of meters of diameter in many cases) and filled with salty water. Therefore the geophysical signal is very strong.

The idea is to use this geophysical signal as speed function K. Because the geophysical anomalies are both positive and negative, the first step is to convert them only to positive values by summing the minimum of the electromagnetic map $E_m(\boldsymbol{x})$ and use it as speed field K_E:

$$K_E(\boldsymbol{x}) = E_m(\boldsymbol{x}) + \min E_m(\boldsymbol{x}) \tag{5.14}$$

but the results obtained with this method showed that the contrast needed to be enhanced because the conduits were not sufficiently controlled by the electromanetic anomalies (figure 5.20a). Therefore a power law was used to strengthen the contrast and gives the final K function:

$$K(\boldsymbol{x}) = b^{K_E(\boldsymbol{x})} \tag{5.15}$$

which gave better results (figure 5.20b).

In addition, in the case study of the Yucatan peninsula many conduits realizations were created, and this allowed to produce probability of occurrence of the karst conduits. Figure 5.21 shows two probability of occurrence maps that were generated using SKS with the geophysical data as K function[4].

[4] The reader is invited to read the paper in appendix C for further informations

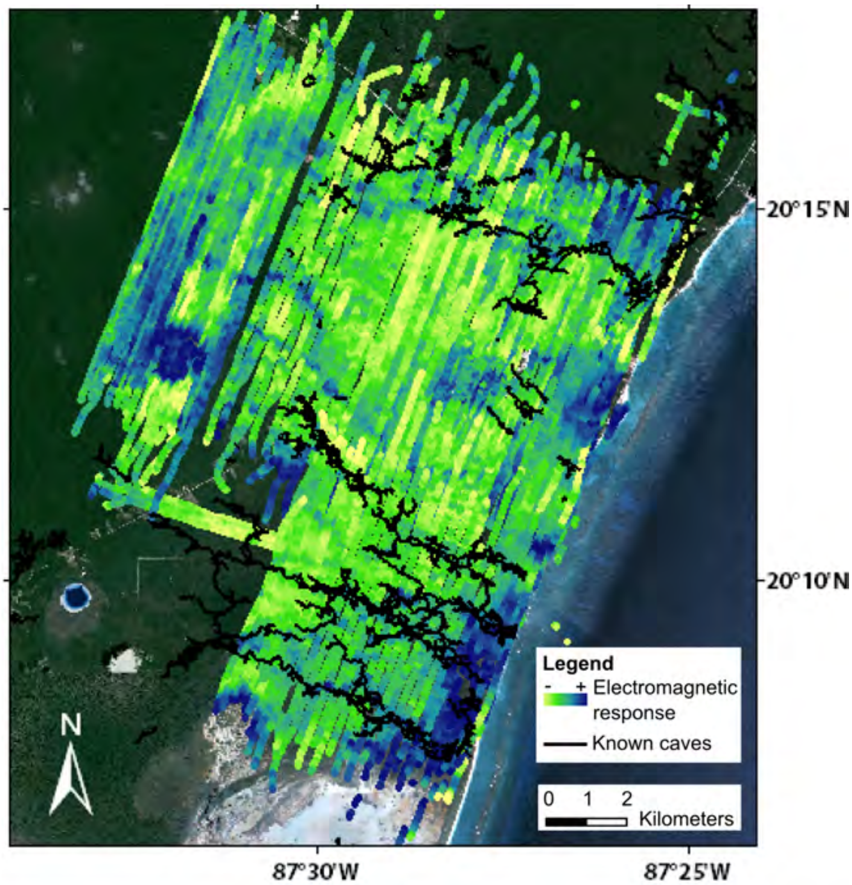


Figure 5.19: A lot of aerial electromagnetic measurements are available. Where the conduits have been explored, the electro magnetic signal often present negative anomalies. (from Vuilleumier et al., 2012, appendix C)

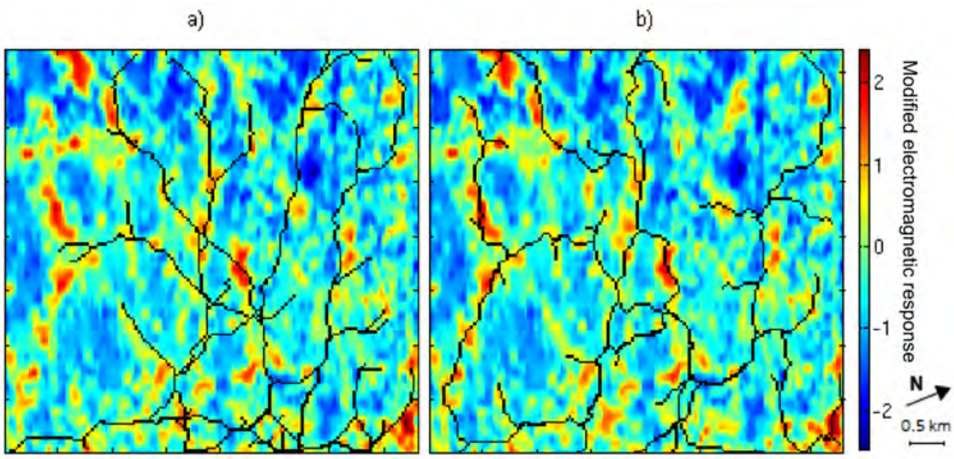


Figure 5.20: Using the geophysical anomaly as speed function K allows to simulate conduits that are consistent with the geophysical data. In a, the electromagnetic component of the velocity medium is the electromagnetic map with positively shifted values. In b, the electromagnetic control on the karst simulation is enhanced by multiplying the value of $E_m(\boldsymbol{x})$ by a factor of 3 of and setting b (eq. 5.15) to 3. (from Vuilleumier et al., 2012, appendix C)

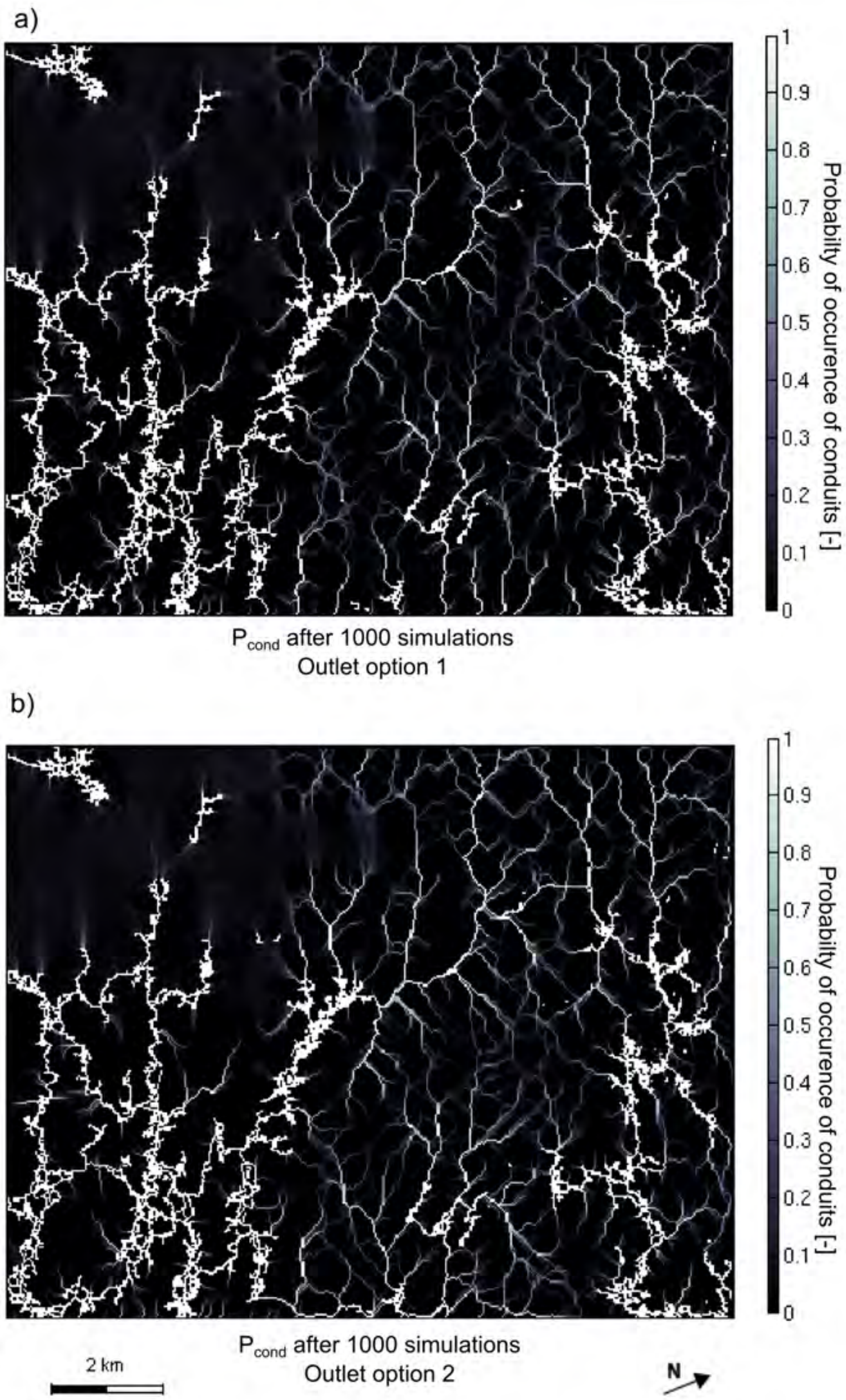


Figure 5.21: Probability of occurrence of karst in the model. (from Vuilleumier et al., 2012, appendix C)

5.8 Conclusions

In this chapter, we have proposed some modifications and extensions of the pseudo genetic approach. First the core of the methodology is clarified, and several synthetic examples are shown to illustrate the effect of the different parameters over the final results. Then the new modifications are explained, like e.g. the modification to produce more realistic conduits which can display labyrinthic structures and the one that allow to avoid conduits going upward. Moreover, a technique to integrate geophysical measurements has been proposed.

Chapter 6

Practical flow and transport modeling for karst aquifers using pseudo-genetic karst networks to build finite element models

6.1 Introduction

Karst aquifers are extremely heterogeneous and difficult to characterize; this makes them very different from other aquifers (Bakalowicz, 2005). Their heterogeneity is induced by the presence of highly permeable preferential flow paths created by dissolution of the surrounding rock. The preferential flow paths are often fractures that are enlarged by dissolution, and under certain conditions they often result into karstic conduits, which are organized in hierarchical networks.

In the introduction of his PhD, Annable (2003) gives a very complete overview of the evolution of the conceptual models about speleogenetic processes over the last two centuries. The conceptual model that is considered in the present study is based on concepts described by Worthington et al. (2000): the karst aquifer model is composed by 2 main hydrofacies: the matrix, which represents more than 90% of the volume of the aquifer, and has an important storage role, and the conduits which represent a very small volume, but which have a very high importance for flow, more precisely they are considered to be responsible for more or less 90% of the total flow.

Although karst aquifers have already been studied for more than a century they are still very difficult to understand (Annable, 2003). Over the last 3 decades, several numerical modeling techniques have been developed in order to provide better understanding of these aquifers, but also prediction tools for water quantity and quality management. A review of these techniques is provided by Ghasemizadeh et al. (2012). As they explain in their review, there are several ways that have been developed to simulate karst aquifers.

The aim of this chapter is to discuss how one can model groundwater flow and transport in karst aquifers in practice. First we provide a brief overview of the main available modeling techniques in Section 6.2. We show that they can give complementary results, and we propose a practical methodology to include the benefits of several different techniques in one unique model. This review shows that only a few attempts have been developed to simulate karstic systems in practice with a distributed and physically based model. When using a physically based model for a karstic system, one has to model first the conduit geometry (which is often unknown), develop a concept for the distribution of the radii of the conduits, have a meaningful model for the recharge accounting for the epikarst. Often numerical problems occur because of the high contrasts of the hydraulic parameters between the matrix (e.g. low hydraulic conductivity) and the conduits (high hydraulic conductivity). In this chapter, we try to provide practical answers to these questions using existing modeling techniques and concepts. The underlying assumptions and considerations that lead to our choice are explained in detail in Section 6.2, while Section 6.3 explains more in detail the implementation.

In a first step, we use the pseudo-genetic *Stochastic Karst Simulator* (SKS) developed previously (chapters 4 and 5) to build the conduit network (Section 6.3.1). The conduits are meshed as pipe elements (2 nodes) which follow the edges of the cubic elements representing the matrix (Section 6.3.2). For comparison, the flow is also modeled using an Equivalent Porous Media approach similar to Worthington (2003), i.e. specific "karst" elements have an increased porosity and permeability induced by the presence of conduits (Section 6.3.7). The networks have a hierarchical structure and distribution of their radius (Section 6.3.3) as discussed by Király (2003). Both flow and transport (tracer tests) are modeled. The flow in the conduits can be modeled as laminar, or turbulent (Jeannin, 2001). In the matrix the flow is Darcian. The finite element code GROUNDWATER(Cornaton, 2007) is used to solve the flow and transport problem. The epikarst effect is simulated using an approximated model based on a transfer function (black-box) model of recharge using the watershed simulator RS2012 (edric.ch, 2012) (Section 6.3.4). The calibration of these models needs a model-run workflow with successive steps as explained in Section 6.3.6.

In this chapter we show how to bind all these elements together and discuss the benefits of combining the different approaches.

6.2 Overview of the existing models

6.2.1 Black-box models

Mathematical black-box model assume the whole aquifer as an abstract schematic mathematical description. It can be a reservoir model, i.e. several "reservoirs" that are connected with mathematical rules, like e.g. global parameter models (e.g. Mangin, 1975); or they can be complex neural networks (e.g. Hu et al., 2008). Unfortunately, as explained by de Marsily (1994), while these models may be a sufficient tool to predict spring hydrographs, they do not provide any spatial information about the karstic system and their parameters do not have a real physical significance for the description of physical processes inside the karstic aquifer. Black-box models have been applied successfully in many cases when the aim of the model is to provide a prediction tool for a function. A good review of these kind of models can be found in Ghasemizadeh et al. (2012). Recently, Hartmann et al. (2013) propose a way of determining the better conceptual model structure for lumped parameters models using chemical signatures and multiple objective functions for model calibration.

6.2.2 Distributed parameters models

Differently than black-box models, distributed parameters models try to solve the physics of the system, with varying degrees of complexity. These models discretize the model domain into sub-units. Each sub-unit has homogeneous parameters in the space that it delimits. They are organized in a grid, which can also be called a mesh. The process that allows to divide the model domain into a grid is called *discretization* (Section 6.3.2). This step is fundamental, because it allows to calculate the fluid flow and transport equation for each cell easily. As Ghasemizadeh et al. (2012) say, *the challenge of distributed modeling approaches to represent karst groundwater systems is to cope with the high spatial heterogeneity of karst aquifers*. Many authors have already modeled karst aquifer using parameter distributed models. The easiest way is to assume the whole karst aquifer as an equivalent porous medium (EPM), where matrix, fractures and conduit are brought toghether in an equivalent hydrofacies (e.g. Lapcevic et al., 1999; Borghi, 2008). Unfortunately, EPM models have a very low applicability in very kartified fields and may lead to catastrophic situations like the case of Walkerton (Ontario, Canada) where, in May 2000, 7 people died from a bacterial contamination of the municipal water supply because the spring protection zone where based on a MODFLOW EPM model which gave much higher transit time than what was observed (later) by field tracer tests (Worthington et al., 2002; Goldscheider and Drew, 2007, cited in Kresic and Stevanovic (2009)).

To avoid this kind of issues, other authors have developed models which use the available information about the conduit geometry to add heterogeneity in their model. Worthington (2003) models the Mammoth Cave aquifer using MODFLOW, and he defines the mesh elements where karst conduits were explored as cells with an higher hydraulic conductivity. He also had to increase the hydraulic conductivity according to the hierarchy of the conduits to be able of simulating realistic head distributions. Király et al. (1975); Király (1988) modeled the conduits as discrete 1D or 2D features embedded in a 3D matrix, the flow in conduits is laminar using a discrete-continuum approach. The discrete-continuum approach is often used to develop speleogenesis models (Annable, 2003; Dreybrodt et al., 2005; Gabrovsek and Dreybrodt, 2010). These speleogenesis models are useful to understand the complex kinetics of karst aquifer. In addition, as Jeannin (2001) pointed out, turbulence is often observed in conduit flow. Nowadays there are computer codes that allow to use turbulent flow equations in discrete conduits, as MODFLOW-CFP (Conduit Flow Process, Kuniansky and Shoemaker, 2008) or GROUNDWATER (Cornaton, 2007). De Rooij (2007) developed a model that is able to simulate also unsaturated flow in pipes, but he points out that it is difficult to obtain the karst network geometry.

6.2.3 Network modeling

As pointed out by De Rooij (2007), a major difficulty when simulating karst aquifers, is to have precise information about the conduit geometries. To use realistic conduits, one could use the networks resulting from speleogenetic models (Annable, 2003; Dreybrodt et al., 2005; Gabrovsek and Dreybrodt, 2010), but the computation of these models is very heavy and they are difficult to condition to field data. Geostatistical based models can be used too to obtain a discrete conduit models like Fournillon et al. (2010). They are easier to condition to field data, but it is difficult to produce well connected paths. Jaquet et al. (2004) propose to use a modified lattice-gas automaton for the discrete simulation of karstic networks, which can produce hierarchical structures, but which is also difficult to condition to field data. Ronayne and Gorelick (2006) propose to simulate branching channels (analogous to karst conduits) using a nonlooping invasion percolation model, which produces nice features, but which are based on a pixel discretization to represent the enlargement of the conduits, and are therefore very difficult to use as base for a flow simulation. Finally, a new approach is to use a *pseudo-genetic* algorithm (Henrion et al., 2008; Collon-Drouaillet et al., 2012; Borghi et al., 2012), which mimics the resulting structures of karst networks produced by speleogentical models, but without solving the complete kintetics of the system. This kind of model use approximated physics to generate the karst networks. Both approaches are based on the assumption that water (and consequently conduit development) will follow a minimum effort path as suggested by Groves and Howard (1994). This is the kind of model that will be used to generate the conduit networks in the present chapter.

6.2.4 Recharge/epikarst model

Karst aquifers present several complex recharge phenomena, which are caused by the presence of a highly altered superficial "skin" (of a few meters) on the top of the aquifers, called *epikarst* (Ford and Williams, 2007). Epikarst has a complex role on aquifer recharge, because it concentrates the rainfall water into several point inlets (dolines), which leads directly to the karstified network. Moreover, it is quite common that surface water streams infiltrate directly into the conduit networks through other point inlets called *sinkholes*.

The conceptual model of Mangin (1975) considers that a consistent part of the rainfall is quickly drained by the epikarstic toward the conduit network, and the remaining part of water will be infiltrated diffusely on the low-permeability fractured volumes. The epikarst is a sort of "skin" of karst aquifers, which rapidly drains and concentrates the rainfall water into several infiltration points known as dolines (concentrate recharge). Király et al. (1995) use a nested 2D model to simulate the epikarst, coupled with a 3D model where the karstic network was modeled. The epikarst model outputs are used as input for the concentrated recharge (sinkholes and dolines). Their recharge function for the flow model underlying the epikarst layer is computed as a proportion of diffuse and concentrated recharge. The concentrated recharge was computed with the 2D epikarst model. Their final conclusions admit that in *most open karst aquifers more than 40% of the infiltration should be drained rapidly into the karst channels.* This proportion may be greater if we consider the common assumption that in mature karst aquifers about 90% of the flow occurs in conduits (Worthington et al., 1999). This work showed the benefits of correctly modeling the effect of epikarst and the recharge, but the use of a 2D model for this topic is difficult to apply if the topography is not an horizontal plane like in their work. Authors like Ofterdinger et al. (2013) use a complex hydrological model that computes the recharge contribution and the runoff of precipitation on the base of several parameter, like the slope, altitude and soil cover. They obtain satisfactory results, but this kind of model can be very difficult to correctly parametrize, especially if there are only a few measurements stations to calibrate the model.

The use of black-box models to estimate the recharge of a karst aquifer has been successfully demonstrated by Weber et al. (2012). Therefore in the present paper, a reservoir black-box model is used to simulate the recharge function for the finite element model, as explained in Section 6.3.4.

6.3 Methodology

This section describes, step by step, how we propose to model flow (and transport) in a karst aquifer with a distributed physical model.

6.3.1 Modeling the karstic network by pseudo-genetic approach

To model a karstic system using a distributed physical model, the first required steps are to construct a model of the geometry of the system and of the conduits. For the conduits we use the pseudo-genetic methodology described in chapter 4 and 5. This approach is fully described in the previous chapters. In summary, first a 3D geological model is built, then a stochastic fracture model is used to add heterogeneity into the 3D geological model, and finally the conduits of the karstic network are generated using the pseudo-genetic approach, which uses a *Fast Marching Algorithm* (FMA) to compute minimum effort paths between karstic springs and inlets (like dolines and sinkholes). This minimum effort path computation is based on the assumption that the water (and consequently the conduits generated by dissolution) will preferentially flow inside the more conductive discontinuities like fractures and bedding planes: the *inception horizons* as defined by Filipponi et al. (2009).

The conduits are generated iteratively. Every conduit that is already simulated will influence the next ones: it is considered as a new preferential path. This lead to a hierarchical network, where the conduits that start from several inlet points (like dolines and sinkholes) converge toward the outlets of the system. The *Stochastic Karst Simulator* (SKS) allows also cycles and complex interconnections to be built into the model.

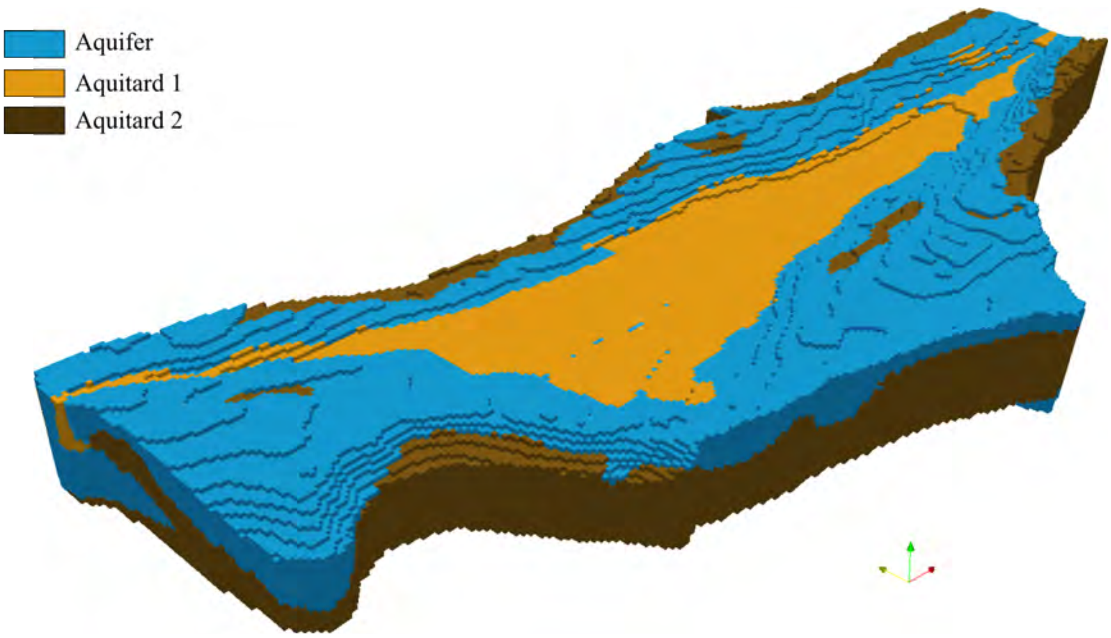


Figure 6.1: Example of a meshed 3D model, an example of the Vallee des Ponts-De-Martel, Jura Mountains, Switzerland. We notice that the model is meshed as a 3D voxel unstructured grid, which is cut according to a delimiting polygon and according to the topography. The colors represent the three hydrofacies.

6.3.2 Meshing

Like many other flow property generators, the *Stochastic Karst Simulator* (SKS) works with a regular grid (voxels). To keep the procedure simple, the grid of the flow model is also constrained to regular elements (8-nodes exahedrons). The geological and karst model are built on a larger parallelepiped region, which is cut later to keep only the relevant cells. The elements above the topography and the ones outside the catchment limits (defined by a polygon) are eliminated. The elements below the lowest aquitard level are also eliminated, because they do not have a significant contribution to the flow in the model. The result of these cuttings is an unstructured grid; its borders and the topography are "lego-brick" shaped. Figure 6.1 shows an example of this kind of mesh on the model of the Valley of *les Ponts-de-Martel* (Jura Mountains, Switzerland). If this kind of model has to be used with a finite differences solver that does not support unstructured grids like MODFLOW or MODFLOW-CFP, the cells outside the model limit and above the topography can simply be deactivated.

If an Equivalent Porous Medium approach is used to model the flow (Section 6.3.7), the cells corresponding to the conduits are only "flagged" as karst elements. Otherwise, if a flow simulation that uses laminar or turbulent 1D flow equation for discrete conduits (1D element) has to be perfomed (Section 6.3.8), the conduits are meshed as *pipe elements*, which are composed by 2 nodes and follow the edges of the 3D voxels. As Figure 6.2 shows, the pipe nodes are consistent with the 3D finite element grid, which allows to use a discrete continuum approach to simulate the interactions between matrix and pipes.

6.3.3 Hierarchy of the networks

Karst networks often show a degree of hierarchy. As Annable (2003) demonstrates, karst conduits often show an organization of the general diameter of the conduits that increases according with the hierarchy of the network, i.e. that the biggest conduits are located downstream as they collect the water of their affluent conduits in a similar way as surface rivers. Differently from streams, karst conduits can also diverge, and in this case the flow is divided between the different diffluent conduits (Annable, 2003). To consider these particular characteristics of karst networks, we have proposed a simple model (Vuilleumier et al., 2012, appendix C) in which the conduit radius r is computed using a power law analog to Horton's

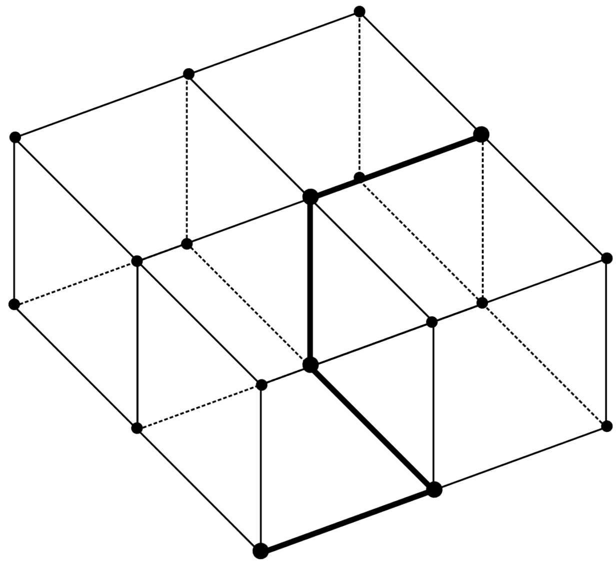


Figure 6.2: The conduits are meshed as 1D pipes that follow the edges of the 3D voxels (matrix)

law:

$$r(u) = \alpha e^{u\beta} \tag{6.1}$$

where α and β are two parameters that have to be calibrated, and u is the order of the conduit. But unlike Collon-Drouaillet et al. (2012) and Vuilleumier et al. (2012) who used a Horton's order based on the rivers Strahler classification, the order u of each conduit is computed here as the ratio of the catchment surface S that is drained by the given conduit over the total catchment surface that is drained by karst conduits as suggested in Section 4.3.4:

$$u_i = \frac{\sum\limits_{j \in A_i} S_j}{\sum\limits_{k=1}^{N} S_k} \tag{6.2}$$

where N is the total number of inlets of the system, S_k represents the total surface of the catchment that is drained by karst conduits, and A_i represents the set of indices that drain into the conduit i (figure 6.3 and figure 6.4).

SKS uses walkers to compute the paths of the conduits. To consider the diverging conduits, SKS keeps the information of the origin of the walkers for each paths its computes. In this way it is possible to keep the divergence of conduits consistent with their radius.

It is clear that this approach correspond to a strong simplification of natural conduits. In reality, the shape of the conduits can vary (a lot) even if their topology remains constant. Because of anisotropic preferential dissolution, some parts of the same conduit can present local variation of the radius. Moreover, the simple fact of assuming the conduits to be cylindrical pipes is already a simplification. However, these simplifications are necessary, because on the one hand it is not yet clear which rules should be followed to characterize these local variations, and on the other hand because it would be practically impossible to validate them. In the proposed approach, the idea is to create the *minimal hydraulic necessary* dimensions to allow the system to drain the water of the system.

6.3.4 Recharge and epikarst model

Two kinds of recharge are considered: diffuse and concentrated. Diffuse recharge is directly infiltrated on topographic surface nodes of the model, and concentrated recharge is infiltrated into the nodes corresponding to dolines or sinkholes. The proportion λ_c of the total recharge that will be considered as concentrated has to be defined by the user and is used to compute both the total concentrated recharge R_c and the total diffuse recharge R_d:

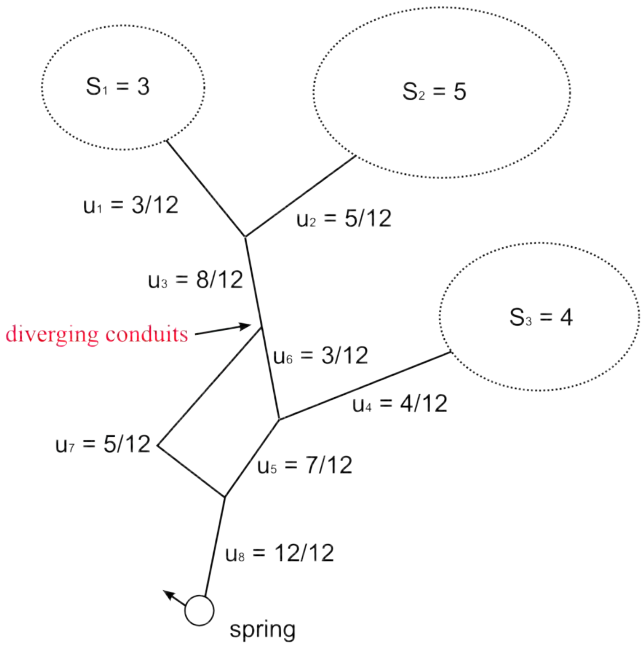


Figure 6.3: The order of each conduit u is the ratio of total catchment surface S that it drains with the total karstic catchment (equation 6.2).

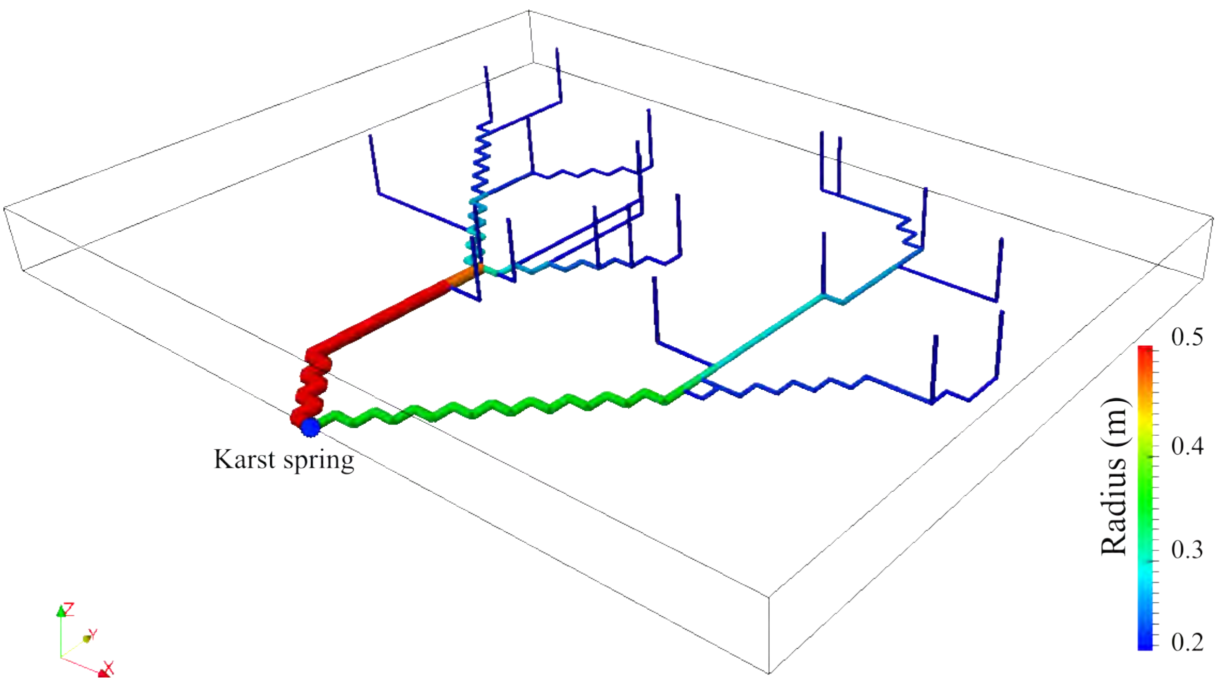


Figure 6.4: Distribution of the radius of the conduits in a synthetic case. The radius increases with the order of the conduits. For a visual purpose, the apparent conduit radius has been exaggerated.

$$R_c = \lambda_c R \quad , \quad R_d = R - R_c \tag{6.3}$$

where R is the total recharge. Then the concentrated recharge R_c^i for each inlet i is calculated with respect to their estimated catchment surface S_i over the total N dolines catchment surface:

$$R_c^i = R_c \frac{S_i}{\sum_k^N S_k} \tag{6.4}$$

R is a function that varies with time and depends on the meteorological conditions. Here it is estimated using the software *RS2012*, which is a hydrology routing system intended for watershed management. Its functioning is based on several objects that are connected with routing functions. Among these objects (or reservoirs) some are *virtual meteorological stations*. These are the points where the local meteorological conditions are computed according to real meteorological stations in the neighborhood: an influence radius has to be provided for the interpolation with a list of surrounding stations and their relative database. The software then computes several meteorological factors; the most important ones, which mostly influence the available quantity of water are:

- rainfall
- evapotranspiration
- snow budget (melting vs deposition)

RS2012 computes them as a function of input parameters[1]. The most important ones are those that control the infiltration rate and the temperature evolution with respect to the altitude. The evolution of the temperature can be used to model the differences of exposure in some part of the model, i.e. a zone that is oriented to the North, will be considered “colder” than the opposite, so that the snow will melt later on this side of the model. The infiltration coefficient will influence the quantity of available water for the karst aquifer. The results of the recharge model are shown if figure 6.5.

6.3.5 Boundary and initial conditions

Flow boundary conditions

For the spring zones, and in general the discharge area, a fixed head boundary condition is used; the outlets of the system are supposed to be at fixed altitudes, and the outflows are dependent only on the head variations inside the model. The springs are usually in the lower regions of the model, but in several cases, some of them are perched temporary springs and they have a higher altitude than the other. In this case a conditional boundary condition must be applied to prevent these springs to become an infiltration node. The condition must deactivate the fixed head boundary condition if the flow is greater than 0, i.e. if it is inflowing into the model. In this way, it becomes possible to simulate also perched temporary springs.

Three kind of boundary conditions can be used for recharge:

- Prescribed flux (Neumann) at the boundary of an element
- source term within an element
- Prescribed flux at a node (infiltrating well)

The Neumann type boundary conditions are flux (m/s) boundary conditions. They are defined on the nodes of the elements, and the finite element solver (GROUNDWATER in this case) integrates the flux on the area of the face of the element that is surrounded by these boundary condition nodes. The flux is perpendicular to the surface. GROUNDWATER can use a slightly different Neumann boundary condition for recharge, which uses the inflowing flux that is vertical, i.e. the vertical component of the flux.

A source term can also be used to model recharge. It correspond to the direct infection of water into the corresponding element, i.e. the elements at the topographic surface. To compute the value of the source term $q_s[s^{-1}]$ the inflowing recharge $Q_{in}[m^3/s]$ has to be divided by the volume $V_{el}[m^3]$ of the element:

$$q_s = Q_{in}/V_{el} \tag{6.5}$$

[1]the reader can find the documentation of the software at http://e-dric.ch/index.php/en/software-en/rs2012

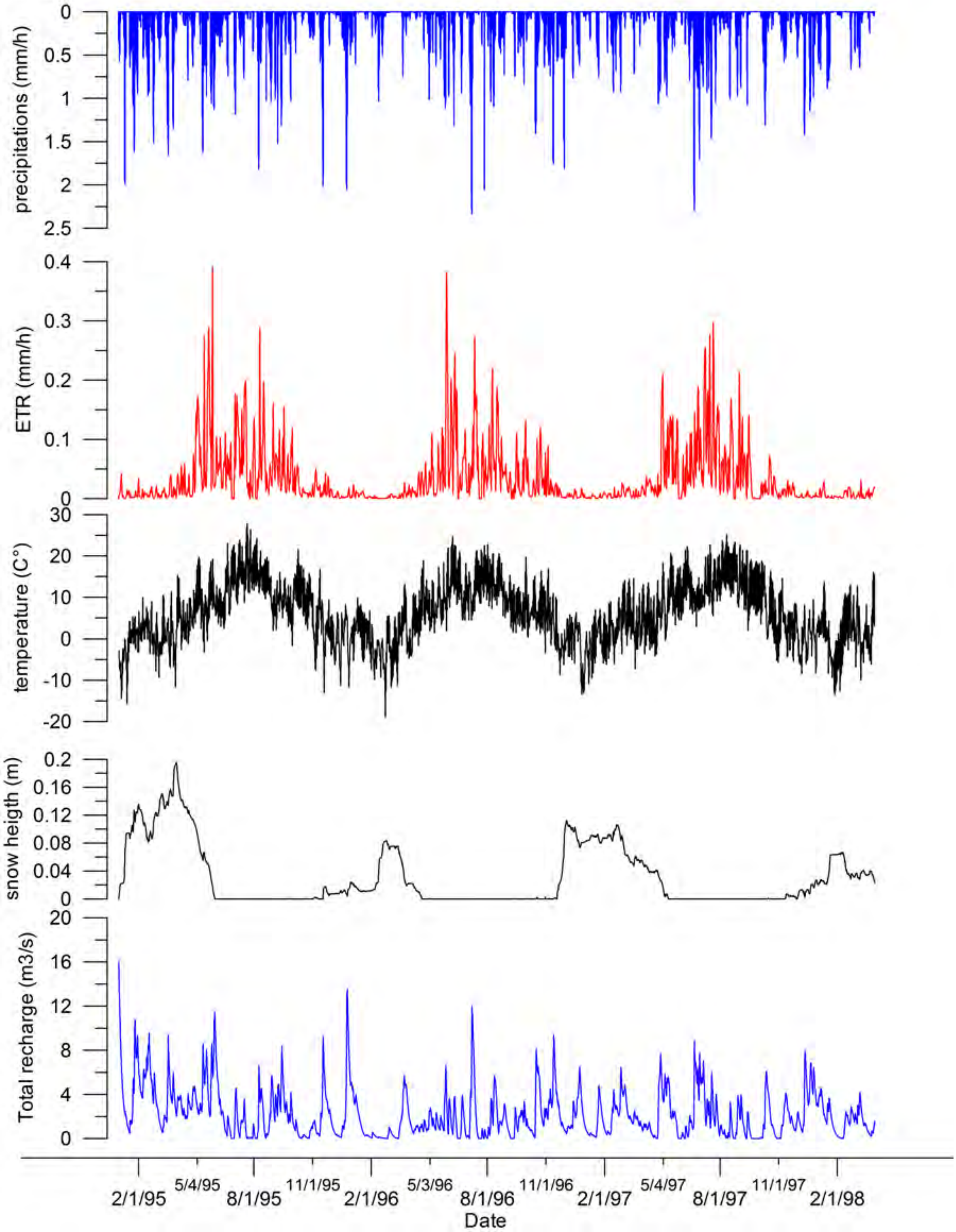


Figure 6.5: Recharge model computed with RS2012.

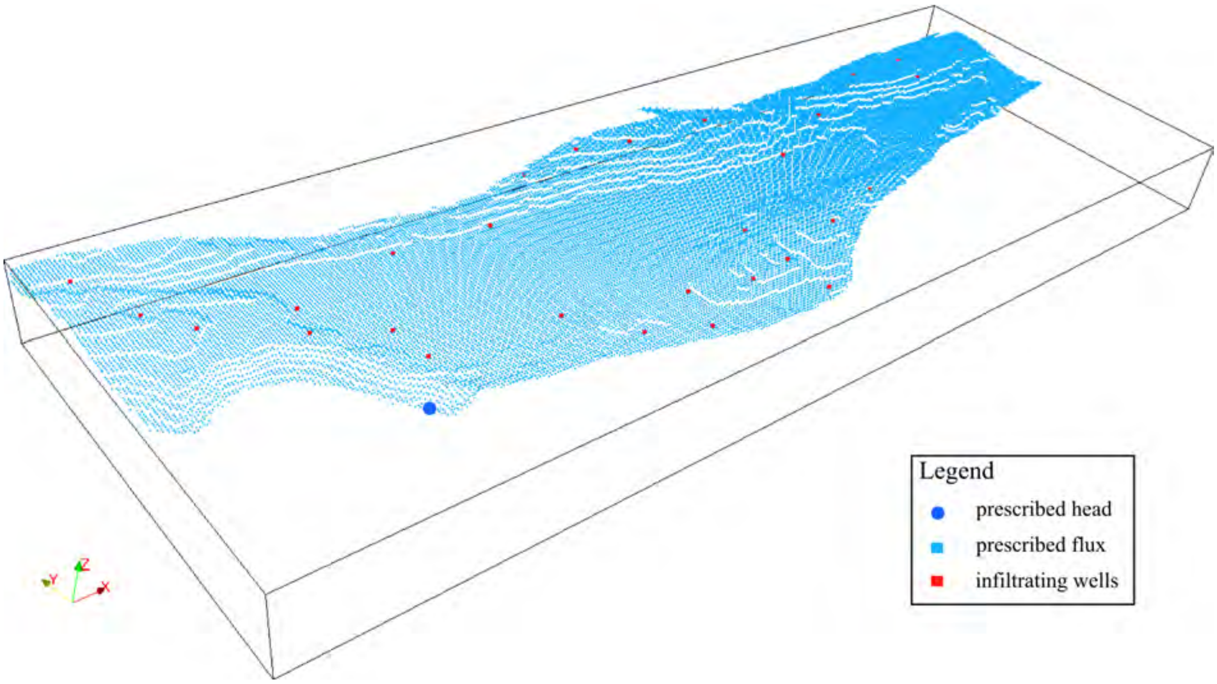


Figure 6.6: Flow boundary condition localization for one model of the Valley of *les Ponts-de-Martel*. The spring of *Noiraigue* is modeled with a prescribed head boundary condition, the inlets with infiltrating wells and the diffuse recharge by prescribed flux at element boundaries.

In a 2D problem the inflowing recharge Q_{in} has to be divided by the area of the element (m^2) and the source-sink term is therefore expressed in m/s.

An infiltrating well boundary condition, can be used to impose a recharge condition (m^3/s) at a single node. It must be used carefully, as it should need a local mesh refinement around the well (the "well" node) if the representation of the head in the surrounding of the node needs to be calculated accurately. These boundary conditions are assigned to the nodes that correspond to sinkholes and dolines. In this way, it is possible to infiltrate the corresponding value of recharge into the nodes (see Section 6.3.4.

The flow boundary condition localization for one model of the Valley of *les Ponts-de-Martel* are shown in figure 6.6. The values assigned to the different boundary conditions are consistent with the recharge model described in section 6.3.4. The prescribed discharge $[m^3/s]$ at the nodes corresponding to the inlets is proportional to their catchment. All the nodes corresponding to diffuse recharge nodes have a prescribed flux $[m/s]$ boundary condition, which is the same for all of them. The spring of *Noiraigue* has been modeled with a prescribed head boundary condition corresponding to its altitude, i.e. 737 masl.

Transport boundary conditions

Boundary conditions are also needed for the simulations of tracer tests. The most common injection tests in karst hydrology are *instantaneous tracer injection tests*.

In our workflow, the tracer tests are simulated separately: every one starts at injection time. In this way it is possible to initialize the problem with an initial concentration of tracer at the injection node, instead of assigning a transport boundary condition. GROUNDWATER, as other flow simulators, computes the initial mass present into the system from the initial concentration. Therefore, if a known mass has to be injected, its corresponding initial concentration C_{in} has to be computed as follow:

$$C_{in} = M/V_p \tag{6.6}$$

where M is the injected mass (kg), and V_p is the *porous volume* influenced by the given node (i.e. the node of the injection). The porous volume of a given node in this case:

$$V_p = \sum_{i=1}^{N_i} \frac{\phi_i V_e}{8} \tag{6.7}$$

where i runs on all the cubic element that are connected to the given node, V_e is the cubic element volume, ϕ_i is the porosity of each cubic element. If the simulation uses also 1D *pipe* elements, and that the given node is also connected with a 1D element, eq. 6.7 becomes:

$$V_p = \sum_{i=1}^{N_i} \frac{\phi_i V_e}{8} + \sum_{j=1}^{N_j} \frac{\pi r_j^2}{2} \tag{6.8}$$

where j are the 1D elements connected to the node, and r_j is the radius of these elements.

These equations show that the initial concentration is strongly dependent on the discretization and on the elements that are connected to the given node, and also on their properties (i.e. their porosity). Moreover, if recharge is applied to the injection node, it may additionally dilute the solute. To take into account for this dilution, the inflowing recharge Q_t at the time t with duration of dt has to be summed to the porous volume V_p as follows:

$$V_p^t = V_p + \frac{Q_t(t)}{dt} \tag{6.9}$$

where V_p^t represent the "porous" volume at time t, i.e. the water volume that dilute the tracer at time t.

6.3.6 Model run workflow

As explained before, this kind of model is planned to be used in an inverse problem framework (chapter 7), therefore the model run workflow is defined in order to minimize the total necessary CPU time. The workflow must take into account transient flow conditions, and transient transport. When studying karst aquifers, several tracer tests may be performed. Therefore, the model should match all of them. Unfortunately, the transport simulation, especially with strong heterogeneity in the flow medium, may lead to very high computational times. The proposed model run workflow is intended to reduce the necessary CPU time as much as possible, when simulating flow and transport at transient state. This workflow can be used for all kinds of simulations described hereafter, i.e. simulating the conduits with an equivalent porous medium approach (Section 6.3.7), or using discrete pipes embedded between porous matrix elements (Section 6.3.8). The workflow is separated in three main steps:

1. steady state flow simulation
2. long transient state (years) flow simulation
3. short transient state (months) flow *and* transport simulations for each tracer

The first step of the model run is a steady state flow simulation using the average flow conditions (i.e. the average discharge of the springs) as input for the model. The steady state hydraulic head result is then used as initial conditions for the transient flow simulation. In this way the initial conditions for the transient simulation are more realistic. The steady state simulation has to be run in the same hydrological conditions as the beginning of the transient simulation, e.g. if the transient simulation starts in a low-flow period, the steady state simulation must be run under low flow condition too, and vice-versa.

The transient simulation is done for the whole period of interest, which must include the periods of the tracer tests, and which may last a few years. During this simulation, the hydraulic head results at every tracer test injection time is stored to initialize the flow field of the short flow and transport simulations, in the same way as it is done between the steady state simulation and the transient one. The transport simulations are performed only over a short period of time (a few days or weeks, depending on the duration of the experiments), because they are heavier to compute. The advantage of this workflow is that the tracer simulations are consistent with the whole transient flow simulation, without needing to solve the whole transient simulation (years) including transport. Figure 6.7 shows the flowchart of the model run.

6.3.7 Simulating flow and transport by equivalent medium "karst" elements

The first tests with the hierarchical conduit networks have been done using an equivalent porous medium (EPM) approach, even if this do not allow to simulate the full complex karstic flow (e.g. turbulent flow in the conduits). Still, some authors have obtained satisfactory results in some cases such as Scanlon et al. (2003) who use equivalent properties for entire regions of their model. In this section, the EPM approach

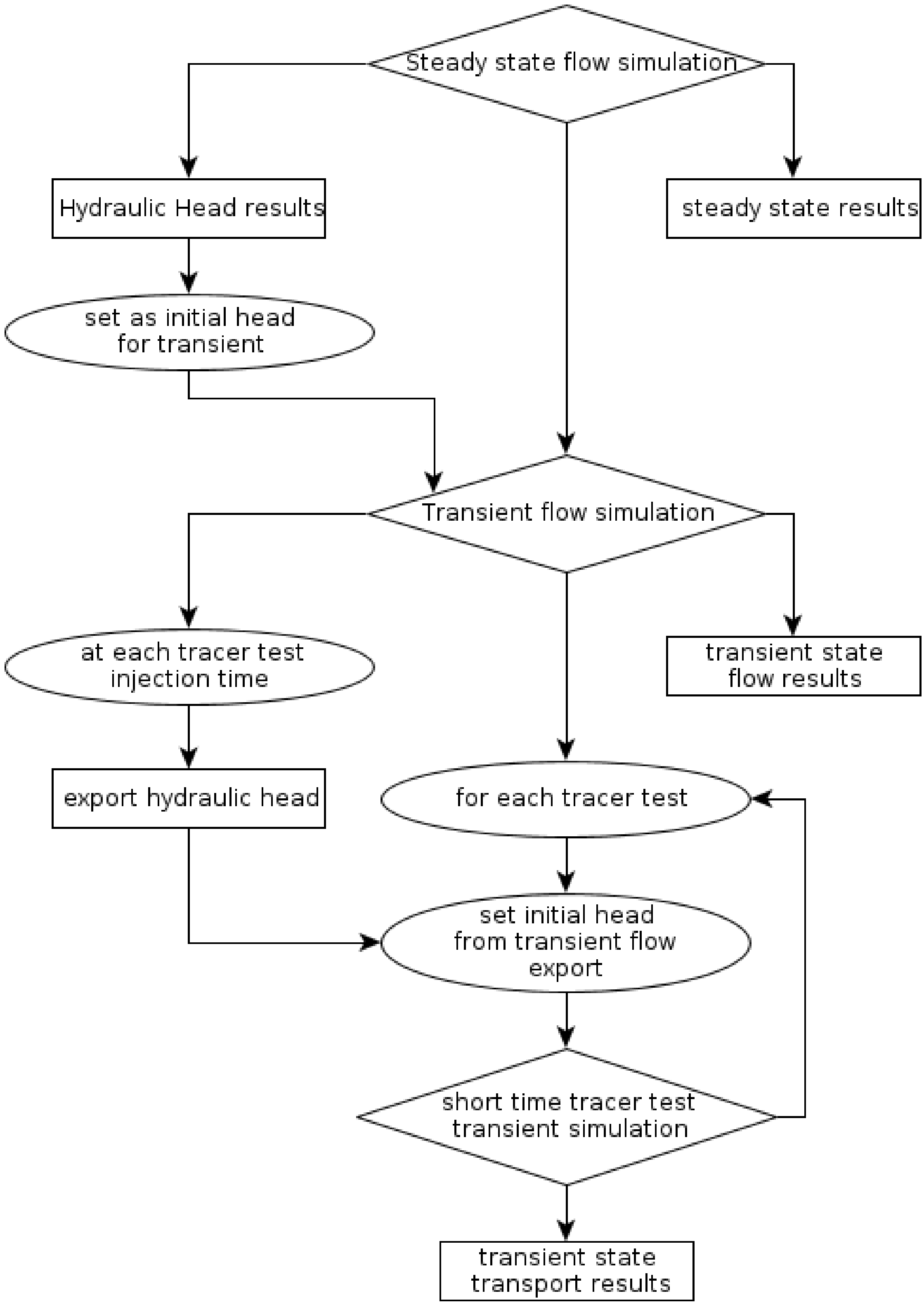


Figure 6.7: Flowchart of the model run

is used by assigning specific flow and transport properties to the karst conduit elements, similarly to Worthington (2003). This method has the advantage of being simple and fast to solve as compared to the use of discrete elements with non linear flow equations as described in section 6.3.8.

The properties of the conduit elements are derived from the radius of the conduits and the mesh size in the following manner: the equivalent porosity ϕ_{eq} (-) is computed by estimating the volume of void in the cell divided by the value of the cell:

$$\phi_{eq} = \frac{V_{void}}{V_{cell}} = \frac{V_{conduit} + (V_{cell} - V_{conduit}) \cdot \phi_{matrix}}{V_{cell}} \tag{6.10}$$

where $V_{conduit}$ is the volume of the conduit (m^3), i.e. $\pi r^2 d$, with d the length in the cell (m), ϕ_{matrix} is the porosity of the matrix (-).

The equivalent permeability is computed knowing the total discharge $Q_{tot}(m^3/s)$ flowing through one cell in a given direction by:

$$Q_{tot} = Q_m + Q_c \quad | \quad Q_c = -\frac{\pi r^4}{8}\frac{\rho g}{\mu}\nabla H \quad , \quad Q_m = -\kappa_m \frac{\rho g}{\mu} \cdot \frac{(A - \pi r^2)}{A}\nabla H \tag{6.11}$$

where $A[m^2]$ is the area of the cell side crossed by the flux, πr^2 the area of the conduit, $Q_c[m^3/s]$ is the discharge through the conduit that crosses the element (using the Poiseuille law), $Q_m[m^3/s]$ is the discharge flowing through the matrix surrounding the conduit, κ_m is the permeability of the matrix (m^2), g is the acceleration of the gravity ($m \cdot s^{-2}$), ρ is the water density ($kg \cdot m^{-3}$) and μ is the water dynamic viscosity ($kg \cdot m^{-1} \cdot s^{-1}$) and ∇H is the hydraulic gradient. Therefore the total flux q_{tot} (m/s) is equal to:

$$q_{tot} = \frac{Q_{tot}}{A} = -\left[\frac{\pi r^4}{8A} + \kappa_m \frac{(A - \pi r^2)}{A}\right]\frac{\rho g}{\mu}\nabla H = -\kappa_{eq}\frac{\rho g}{\mu}\nabla H \tag{6.12}$$

The equivalent permeability of the cell κ_{eq} (m^2) is given by:

$$\kappa_{eq} = \kappa_m + \frac{\pi r^2}{A}(\frac{r^2}{8} - \kappa_m) \tag{6.13}$$

The use of an equivalent porous medium approach is not particularly suited for karst modeling, especially for the transport, as we need to put unrealistically low porosity to the karst elements so that the flow pore velocity becomes satisfactory in term of transport (Kresic and Stevanovic, 2009). Therefore, the use of special conduit elements with other physical equations has been chosen.

6.3.8 Simulating flow and transport in 1D pipes embedded in a 3D porous matrix

Now, let's consider that the conduits are meshed with 1D pipe elements.

The whole model is meshed using a regular grid as explained in Section 6.3.2. The pipes elements are located on the edges of the 3D elements (hexaedrons) of the karstic matrix. There is a direct hydraulic connection between the pipes and the matrix, as the nodes of both types of elements are the sames.

To model the flow in the conduits two different approaches can be used: laminar and turbulent flow. The linear flow equation used in pipes is the Darcy-Poiseuille formula, the hydraulic conductivity for the pipes elements in this case is:

$$\kappa = \frac{\rho g}{\mu}\frac{r^2}{8} \tag{6.14}$$

The flow equation used in pipes for turbulent flow is the Manning-Strickler (Turbulent flow conditions). The hydraulic conductivity for the pipes is therefore:

$$\kappa = \frac{\phi \frac{r}{2}}{\sqrt{||\nabla H||}} \tag{6.15}$$

where ϕ is the friction coefficient ($m^{1/3}s^{-1}$). Both cases are used with the same law for flow ($\vec{v} = -K\nabla H$) where only the definition of K varies from one case to the other ($K = \kappa\frac{\rho g}{\mu}$).

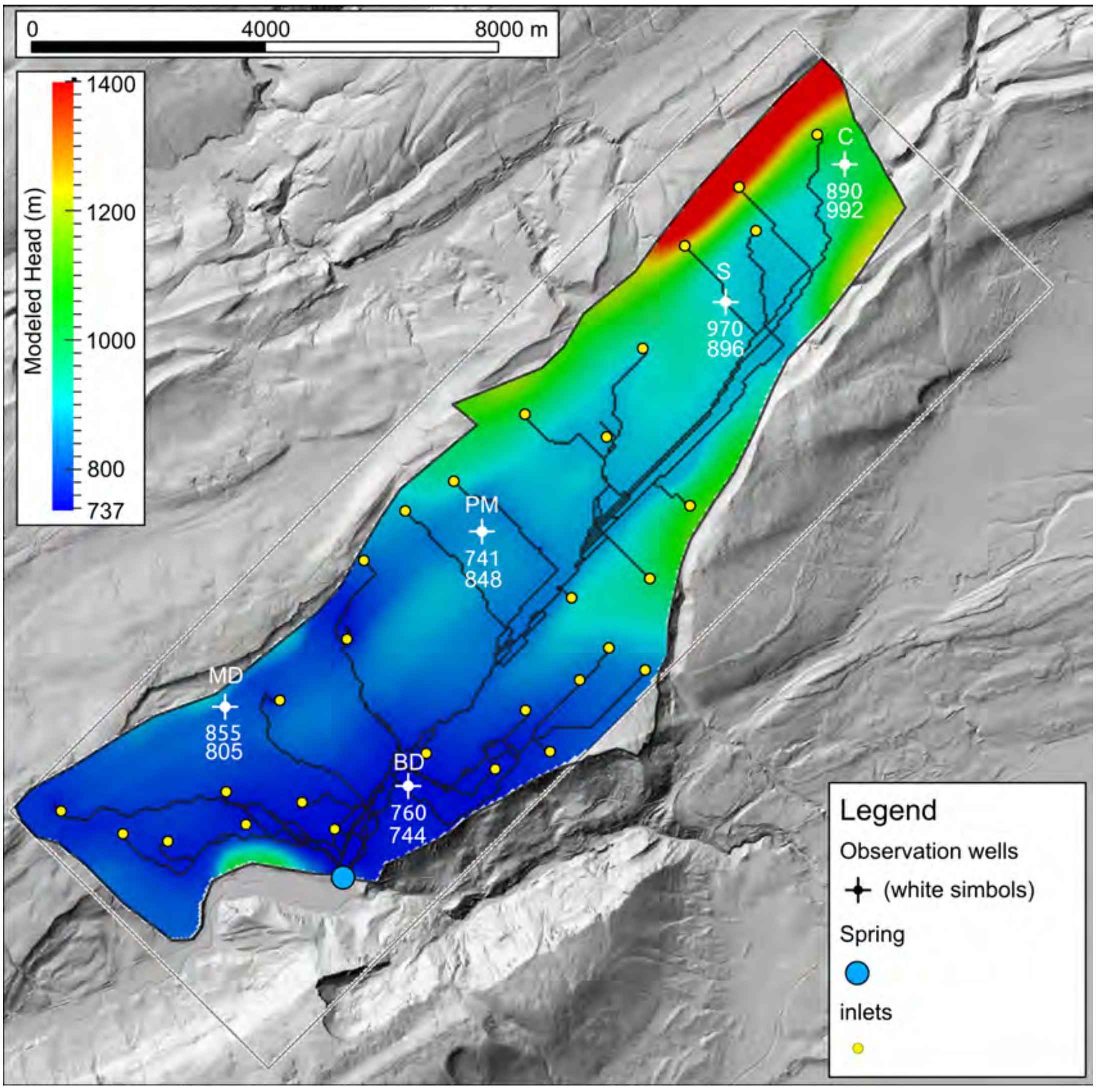


Figure 6.8: Simulated high flow heads (steady state flow). This figure shows an horizontal cross-section at the altitude of the spring (737 masl). The observation wells label are: line 1, the ID of the well; line 2, measured maximal head; line 3, simulated head. The head field is clearly influenced by the high conductive conduits (black lines).

This approach has been tested for the Valley of *les Ponts-de-Martel* case study site (figure 6.8). The simulated head field is clearly influenced by the conduits. The modeled conduits influence the head field widely upstream, which was not possible only with an EPM model (Borghi, 2008). In this example, the head in the Norther part of the model is to high because too high recharge is applied over an aquiclude formation in the border of the model.

6.3.9 Transport simulation

The transient transport simulations are used to model the tracer tests that have been performed on the field. GROUNDWATER allows only one species of tracer to be simulated for each transport simulation. It is therefore necessary to perform a separate simulation for each tracer test.

The injection of the tracer is modeled as an initial concentration at the beginning of the transport simulation. The initial flow field of the simulation is extracted from the transient long period flow simulation. In this way, all the transport simulations are consistent with the long period flow simulation. Ideally, the transient flow field of the transient flow simulation should be used as flow field for the transport simulation, to reduce the computational time, but this is not possible, because the time-steps evolution depends also from the transport simulation.

According to many authors (Bear, 1972; Bear and Jacobs, 1965; Bear et al., 1993), it is necessary to pay attention to two parameters: the *Peclet* number Pe and the *current number* Cr. The Peclet number is defined as:

$$Pe = \frac{vdx}{D} \tag{6.16}$$

where v is the flow velocity, i.e. the pore velocity v_p multiplied by the porosity ϕ ($v = v_p\phi$). dx is the mesh dimension in the flow direction (in this case $dx = dy = dz$) and D is the dispersion coefficient (m^2/s). The current number is defined as:

$$Cr = \frac{v\delta t_{max}}{dx} \tag{6.17}$$

where δt_{max} is maximal time step size (s) for the transient simulation. The numerical solution can be considered stable if these two conditions are satisfied:

$$Pe < 2 \quad , \quad Cr < 1 \tag{6.18}$$

These two inequalities can be expressed in order to isolate the dispersion D and the maximum time step δt_{max}:

$$\frac{vdx}{2} < D \quad , \quad \delta t_{max} < \frac{dx}{v} \tag{6.19}$$

The dispersion is a parameter that depends on the scale of the problem, as it can be used to assume the dispersion of a pollutant induced by pore-scale heterogeneities in an equivalent porous medium. As explained in Rausch et al. (2005) the effect of other kinds of heterogeneities (like karst features in this case) may strongly influence the behavior of the solute in the simulation. In this case, the dispersion has to be interpreted as the dispersion occurring at cell scale, and depends therefore strongly on the mesh dimensions.

Moreover, following the conditions of equation 6.19, it appears that particular attention has to be payed to mesh size and on the values to the dispersion coefficient. The time discretization is also very important. However when simulating karst aquifers, with a very strong heterogeneity, it can be difficult to compute the value of the maximal flow velocity in the medium *a priori*, which can be very fast in some cases. To guess these parameters before running the full transient transport simulation, a steady state flow simulation can be performed in high flow conditions.

6.4 Conclusion

This chapter shows how to use a combination of already existing tools to simulate flow and transport in karst aquifers with distributed parameter models. The karstic conduits are meshed as 1D pipes that follow the edges of the 3D elements used for the matrix. Turbulent flow equations are used to simulate flow in the conduits; this adds physical consistency to the models, but complicates the solution of the numerical problem. The radius of the conduits depends on their hierarchical order. Major conduits have largest radiuses, while secondary conduits are smaller. The conduits are hierarchically classified by the ratio of their catchment surface over the total catchment surface. The recharge function applied on the inlets also depend on their catchment surface. The greatest part of the recharge is applied directly in the inlets and only a small percentage is applied on the rest of the model. This allows to approximate the effect of epikarst in convoying water to the inlets.

The estimation of the physical parameters of these simulations can be difficult to achieve by *manual* trial and error approach. Therefore the calibration of these models should profit from the available automatic parameter estimation techniques. Moreover an inverse framework could be applied to evaluate the uncertainty related to these models. These two aspects are the topic of the next and last chapter.

Chapter 7

Framing the inverse problem for karst modeling

7.1 Introduction

Many techniques exist, in traditional inverse problem, that use gradient based methods that gently modify the hydraulic conductivities field of a flow simulation until the value of a so-called objective function reaches a minimum or stabilizes at an asymptotic value (Tarantola, 2005; Carrera et al., 2005; Caers and Hoffman, 2006; Carrera et al., 1988). The problem with karst aquifers is that, because of the presence of conduits which act as preferential flow paths, it is not possible to *gradually* modify the parameter field: when modifying the geometry of the conduits many connections/reconnections are possible, and this leads to a very discontinuous objective function. It is difficult to find studies in the literature that have applied an inverse method for karst aquifer distributed parameter model. Larocque et al. (1999; 2000) and Panagopoulos (2012) have used an inverse method to calibrate karst hydrogeological model. Both do not include karst conduits in their model and assume the karst aquifer as a 2D equivalent porous medium aquifer. Moreover, it is even more difficult to find someone who tested an inverse problem technique in karst aquifer using also transport simulations.

Our proposition is to first generate a great amount of karst conduit realizations using the *pseudo-genetic* algorithm to build them (chapter 4 and 5), and then calibrate the physical properties of each one of them using a "classical" optimization algorithm. The geometries that do not allow to obtain a reasonable fit would be rejected. In this way a set of *best fitting* models could be used for uncertainty analysis and predictive modeling.

Unfortunately, it is not yet clear which parameters (like the hydraulic conductivity and the porosity, or geometrical like the shape and number of conduits) have the strongest impact on simulation results, and therefore the value of the objective function. These parameters can be inferred based on common sense, and it can be assumed that a parameter like the radius of the conduits or the conductivity of the matrix will have a strong impact on flow, but as far as we know, there has not yet been any systematic research with this aim in the literature.

Moreover, the ability of a "classical" optimization technique to effectively (and possibly efficiently) calibrate the physical properties of karst aquifers (matrix hydraulic conductivity, conduit radius) has also to be tested when simulating complex systems, i.e. simulating Darcy laminar flow in the matrix and turbulent Manning-Strickler flow in discrete pipes as done in chapter 6.

The present chapter is intended as a first step to investigate the proposed inverse technique (i.e. many karst realizations, optimize each of them). Two distinct numerical experiments related to the resolution of the inverse problem in karst modeling are shown here. The first test present a sensitivity analysis about the influence of the variation of the karst conduits radius and the matrix hydraulic conductivity on the simulation results. It is performed by comparing the results of 18'000 2D flow and transport simple (equivalent porous medium) finite elements simulations (sections 7.2). The second test investigates the ability of an inverse algorithm such as PEST (Doherty et al., 1994) to retrieve optimal physical parameters when using a more complex model, i.e. a 3D model with the karst conduits meshed as 1D pipe elements.

7.2 Investigation about the identifiable karst geometrical and physical properties from tracer test data

It is well known that inverse problems are usually ill posed (Hadamard, 1932) and have either no solution or an infinity of solutions. A way to overcome these limitations is to pose the inverse problem in a perspective where sources of information are combined (Tarantola and Valette, 1982). In that part of the work, we analyze which geometric or physical properties of a karstic system can be identified using an inverse approach. To answer this question, we perform numerical experiment in which we generate an ensemble of different geometries and parameter sets. Then we search systematically which other geometry or parameter set produce similar results.

7.2.1 Model description

The case study considered in this numerical experiment is a synthetic 2D finite element model. The domain has a size of 200x200 meters with a grid resolution (dx=dy) of 2m. The flow and transport equations are solved using the finite element code GROUNDWATER (Cornaton, 2007).

The model has 2 hydrofacies: "karst" and "matrix". Flow in karst and matrix elements follows Darcy's law (laminar flow). The model assumes steady-state and transient transport. The synthetic case mimics a small karst aquifer, for which only 2 inlets (sinkholes) and the spring are known. For both inlets, we know the mean head and a tracer test has been performed. The tracer test is a punctual injection of a unitary mass at time 0 of simulation. The flow boundary conditions are: no flow boundary on every side of the model, a fixed head at the spring node and a constant recharge over the model. 90% of the total recharge is directly infiltrated in the inlets of the model and only 10% on the "matrix" elements. Leaving 10% of recharge on the other elements allows to reproduce a gradient from every element of the model toward the spring. Figure 7.1 resumes the model setup.

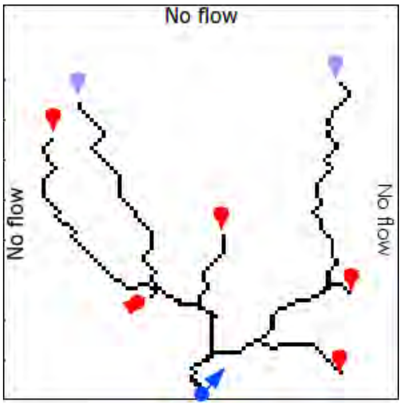


Figure 7.1: model description: in white the matrix elements, in black the conduits; flow boundary conditions: the blue arrow is the spring (constant head), the inlets have 90% of total recharge, the remaining 10% is distributed on the matrix elements; transport: injection of 1g of tracer in 2 inlets (violet point)

7.2.2 Geometrical karst realizations

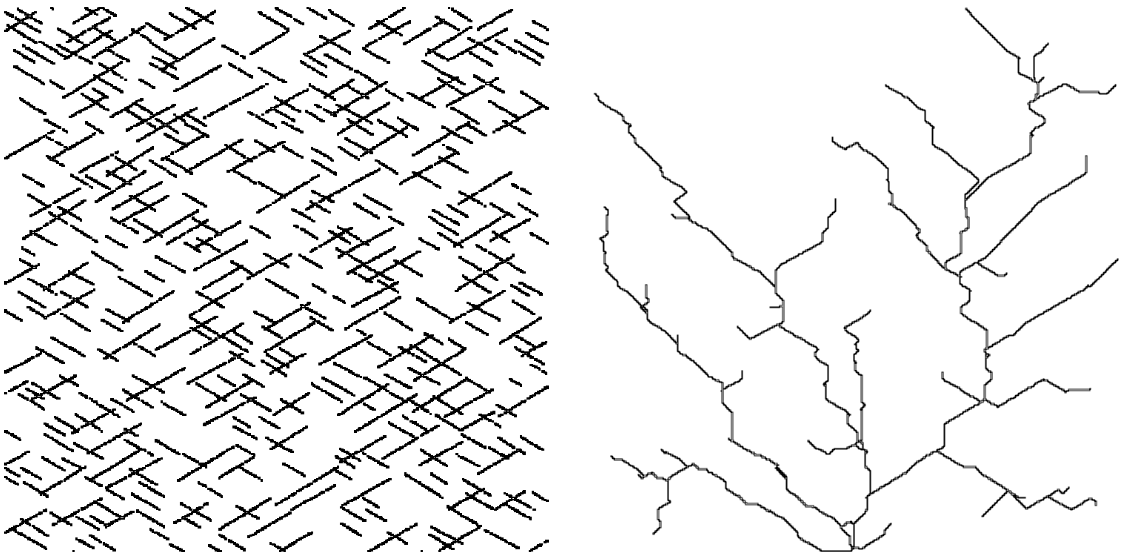
150 geometrical realizations are constructed using the *pseudo genetic* methodology described in chapters 4 and 5 (Figure 7.3). All the realizations can be grouped in 3 families based on the number of inlets (50 realizations for every family) and are used for the simulation. The first family has only 2 inlets (i.e. the 2 known inlets), the second 7 (i.e. 2 of known location, and the others with random locations) and the third 14 inlets (of which 2 are deterministic). The *pseudo genetic* methodology requires the existence of an initial fracture set to control the variability of the realizations of karst networks. Here, we use a different equiprobable realization of a fracture model for every karst realization. We use the same statistics about fractures for every realization so that the final karst realizations still remain comparable. Figure 7.2 shows one realization of the fracturation model (a) and the resulting stochastic karst conduits model (b).

7.2.3 Flow parameters for each karst realization

For every karst realization, we tested 120 different combinations of 2 properties. The first is the hydraulic conductivity of the matrix K, varying from $5 \cdot 10^{-7}$ m/s to 10^{-3} m/s (8 values with regular steps in a log_{10} space) and the second is the "radius" r of the conduits in "karst" elements, varying from 0.1 to 0.8 m (15 values with an increment of 0.05m). It is important to note that, as we use an equivalent flow parameter field, changing the radius of the conduits leads to differences in both porosity (equation 6.10) and equivalent permeability (equation 6.13). Figure 7.4 shows the variation of both porosity n and equivalent hydraulic conductivity K_{eq} with an increasing radius r from 0.1 m to 0.8 in 2D cells of 2m length (i.e. the tested values in this chapter).

7.2.4 Reference simulation selection

After having solved the direct problem for both transient tracer tests on every combination of parameters of every karst realization (which means 150x120x2=36'000 simulations) we choose one reference simulation. This will be selected from the 18'000 flow parameter fields (150x120). This reference simulation will be considered as our "reality". The results of this simulation are 2 observations of head H^i at both inlets, and 2 tracer test results C^i at the spring.



a) b)

Figure 7.2: A stochastic fracture network (a) is the basis for the karst conduits model (b)

These observations will be compared to the results of all the other simulations. The comparison is based on the following objective function:

$$e = \lambda e_c + (1 - \lambda) e_H \tag{7.1}$$

with:

$$e_c = \frac{\frac{1}{N_c}\sum_{i=1}^{N_c} |C^i_{obs} - C^i_{calc}|}{\max(\frac{1}{N_c}\sum_{i=1}^{N_c} |C^i_{obs} - C^i_{calc}|)} \tag{7.2}$$

and

$$e_H = \frac{\frac{1}{N_H}\sum_{i=1}^{N_H} |H^i_{obs} - H^i_{calc}|}{\max(\frac{1}{N_H}\sum_{i=1}^{N_H} |H^i_{obs} - H^i_{calc}|)} \tag{7.3}$$

where λ is a weighting percentage, N_c are the number of time measurements of concentration, and N_H the number of head observations.

7.2.5 Inverse problem solving

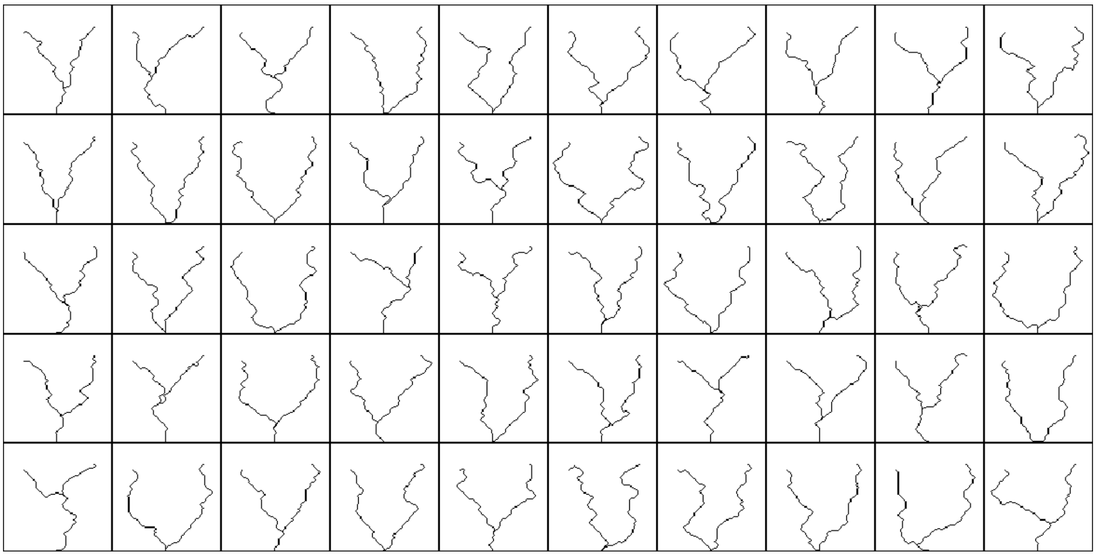
Having a given reference data set, solving the inverse problem consists here in finding the configurations that match the observations. In practice, we compute the error e using equation 7.1 for all the other models and select the models for which e is lower than a given threshold. The result is an ensemble of realizations that match the data and describe statistically the remaining uncertainty.

Figure 7.5 illustrates this procedure. In figure 7.5a, the ensemble of all the 18'000 recovery curves are displayed with the one that is considered the reference. We observe huge differences between the reference result and most breakthrough curves. Figure 7.5b shows the selected "best fit" simulations, in this case all the simulations that have an error e (section 7.2.4) lower than 1%.

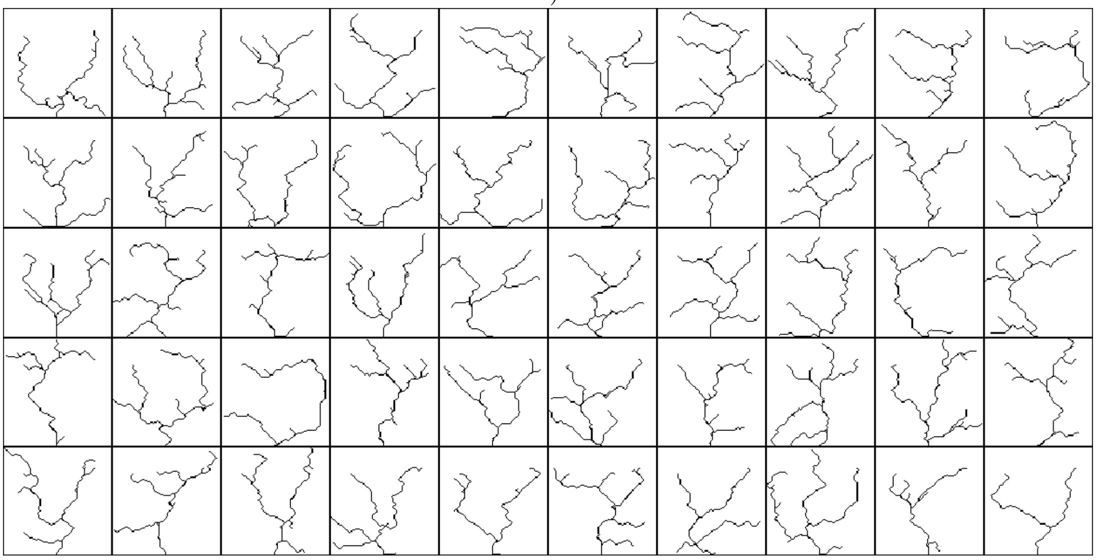
Figure 7.6a shows the histogram of e_c and Figure 7.6b the histogram of e_H in this case. We notice that the response for transport (Figure 7.6a) is much more discriminant than the one for flow (Figure 7.6b). The flow response shows that it is quite easy to obtain a good fit, as many simulations have a very small value of e_H. On the contrary, transport results show a well distributed error e_c, which is induced by a higher sensitivity of transport to the effective flow paths. Figure 7.6c shows the histogram of the objective function with a value of λ equal to 1/2 (equation 7.1), i.e. giving the same weight to both e_c and e_H.

7.2.6 Statistical analysis of the error, switching the reference

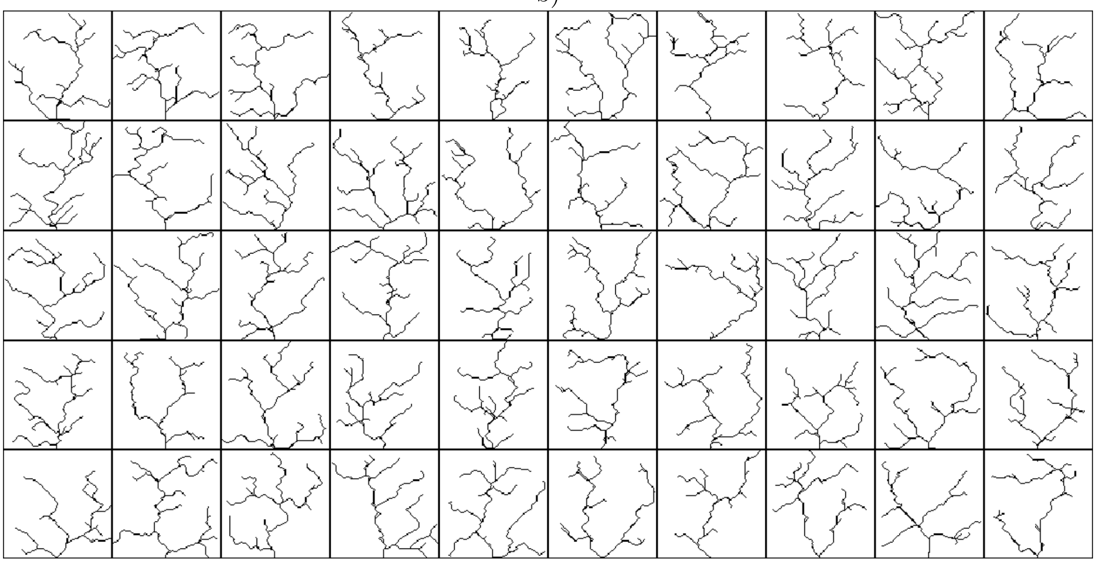
The previous section has shown that it is possible to find models matching a certain reference. Here, we use systematically all models as reference, solve the inverse problem for each one (as explained in section



a)



b)



c)

Figure 7.3: stochastic karst realizations used for this study, 50 realizations with a) 2 conduits, b) 7 conduits, and c) 14 conduits

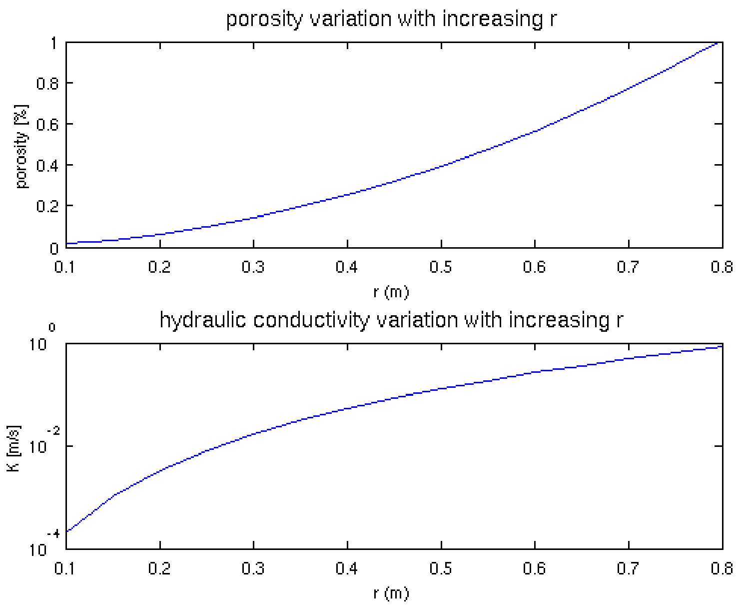


Figure 7.4: variation of equivalent porosity [-] and hydraulic conductivity [m/s] for the chosen values of r, the radius of the conduits

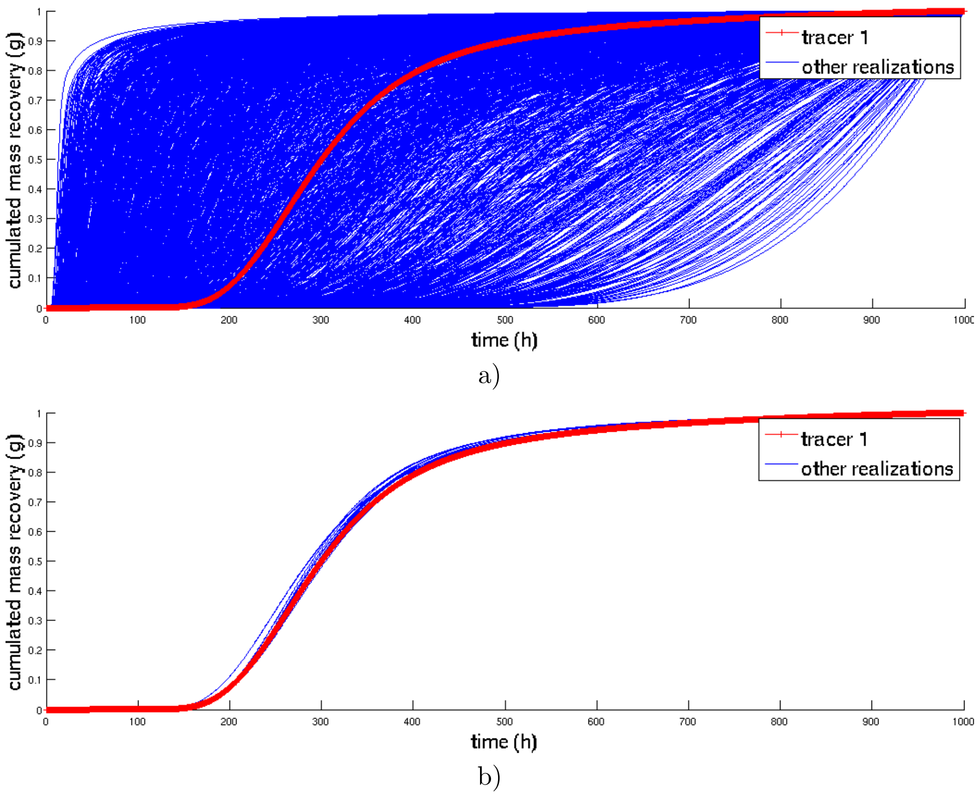


Figure 7.5: Tracer test results (cumulated mass at the spring). a) In red we see the tracer test result for tracer one on the reference simulation, in blue the results of all the other simulations. b) selection of the "best fit" simulations, i.e. the ones under 1% of error

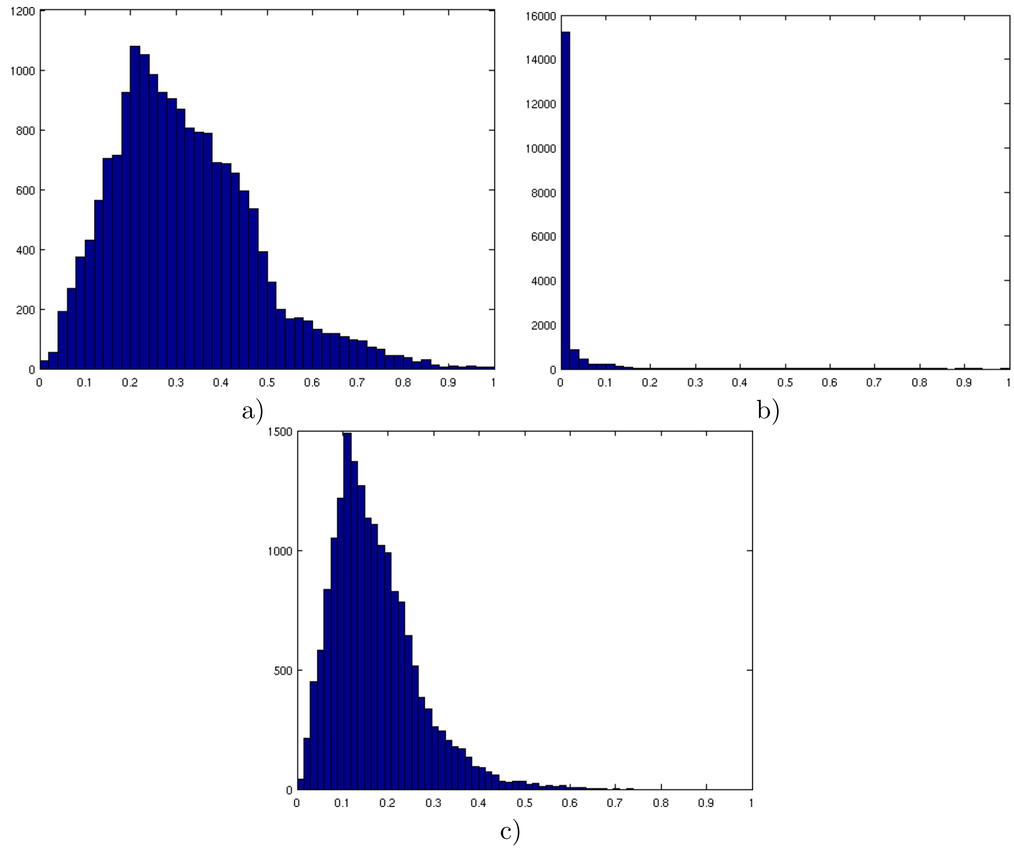


Figure 7.6: pdf of normalized error, a) on tracer test results(e_c). b) on flow results (e_H). c) using both flow and transport (e), with $\lambda = 0.5$.

7.2.5), and analyze in a statistical manner if the parameters are well identified.

To represent the results, we estimate the probability distribution of the estimated parameter values as compared to the true parameter values in the form of log ratios: $\log_{10}(r_{est}/r_{ref})$ and $\log_{10}(K_{est}/K_{ref})$. These probabilities are represented conditional to a certain value of the reference parameters:

$$P(\log_{10}(r_{est}/r_{ref}), \log_{10}(K_{est}/K_{ref}) \mid N_{ref} = x) \tag{7.4}$$

where x is the possible number of conduits. When $\log_{10}(r_{est}/r_{ref})$ or $\log_{10}(K_{est}/K_{ref})$ are equal to 0, it means that the estimated radius r_{est} and the estimated matrix conductivity K_{est} are equal to the reference. We also compute the probability of having N_{est} number of estimated conduits, knowing the N_{ref} number of reference conduits:

$$P(N_{est} \mid N_{ref}) \tag{7.5}$$

which allows to understand the probability to identify properly the number N of conduits as the references. We also compute the conditional joint pdf of the ratios: $\log_{10}(r_{est}/r_{ref})$ and $\log_{10}(K_{est}/K_{ref})$ knowing N_{ref} and N_{est} the number of conduits of the possible solutions:

$$P(\log_{10}(r_{est}/r_{ref}), \log_{10}(K_{est}/K_{ref}) \mid N_{est} = y, N_{ref} = x) \tag{7.6}$$

where x and y are the possible number of conduits.

Figure 7.7 and 7.8 suggest that the radius of the conduit seems much better identified by the inverse procedure than the hydraulic conductivity of the matrix. The probability of having an estimated radius larger or smaller than twice the real radius if very small. On the opposite, the probability to have an estimated hydraulic conductivity for the matrix around 3 orders of magnitude larger or smaller than the truth is large. This is consistent with our conceptual model, for which 90% of the flow and possibly all the tracer are supposed to pass through the conduits.

One interesting thing that can be seen on figure 7.7 is that when the reference has only 2 conduits ($N_{ref} = 2$), the probability $P(N_{est} = 2|N_{ref} = 2)$ is extremely high as compared to $P(N_{est} =$

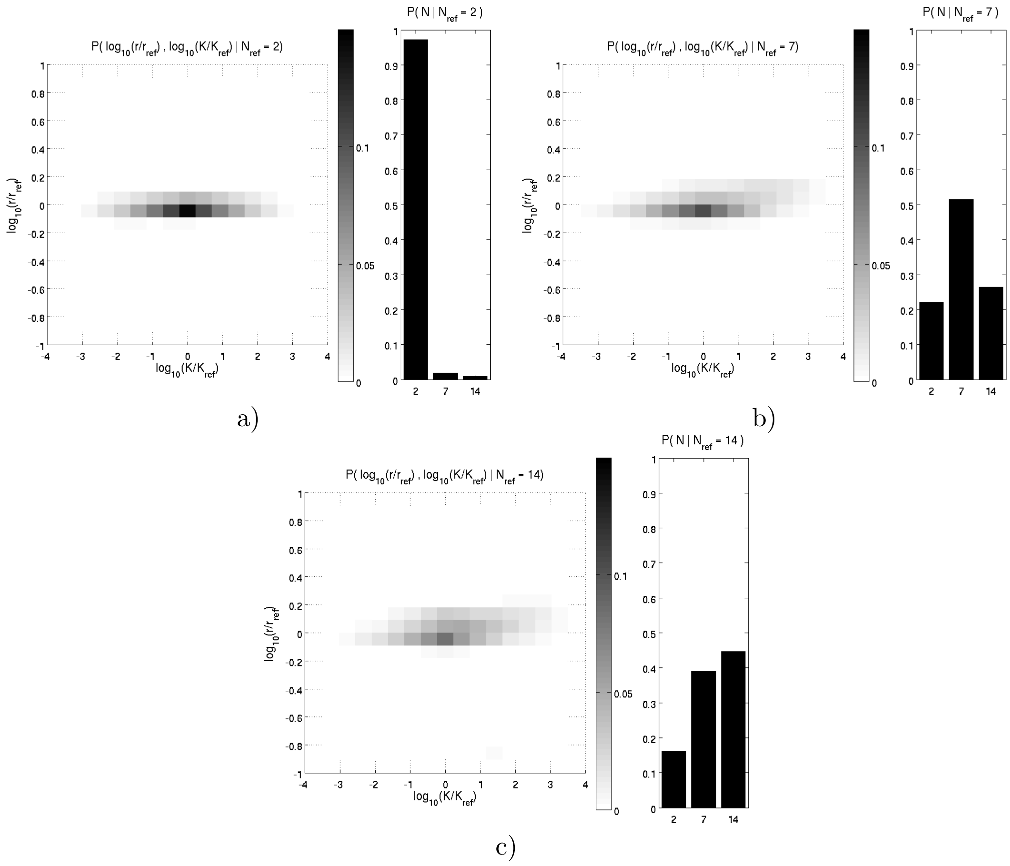


Figure 7.7: Joint probability distribution function using the 3 different families of geometries (a=2, b=7, and c=14 conduits respectively).

$7, 14|N_{ref} = 2)$. In this case the inverse problem is quite robust and this shows that when the reference has 2 inlets and that 2 tracer tests are available, we are able to identify confidently the fact that the model only has 2 inlets. On the opposite, when the reference has 7 inlets ($N_{ref} = 7$), the probability to estimate that the model should have 7 inlets is slightly higher than the probability of having 2 or 14 conduits. This is even more difficult when the reference has 14 conduits ($N_{ref} = 14$). Probably, this is related to the fact that only 2 tracer tests are available in this synthetic example.

Looking more in detail at figure 7.8 we notice that the estimated radius r_{est} and matrix hydraulic conductivity K_{est} are very well identified for the simulations for which $N_{est} = N_{ref}$ (figure 7.8a,e, and i), i.e. the value of the $\log_{10}$ of both ratios are near to (0;0). When $N_{est} < N_{ref}$, the estimated values of r_{est} and K_{est} are overestimated. It makes sense because the model needs to "drain more water" relatively to the number of conduits. On the opposite, when $N_{est} > N_{ref}$ the estimated values of r_{est} and K_{est} are underestimated for the same reason.

Overall, this exercise has shown that it is possible to find simulated breakthrough curves that match the reference. It also shows that it in many cases the parameters are correctly represented, especially the radius of the conduits. We expect this identifiability to be related with the complexity of the system and the quantity of available data.

7.3 Possibility of acceleration of the search using PEST (Doherty et al., 1994)

In the previous section, we searched systematically the best fit by rerunning the model for all the possible parameter values. This was possible because the model was small and simple. Using a desktop PC (2Gb Ram, intel duo -core 2.2 GHz), it takes 20 minutes to run each of them, while the generation of the karst network is about 5 seconds.

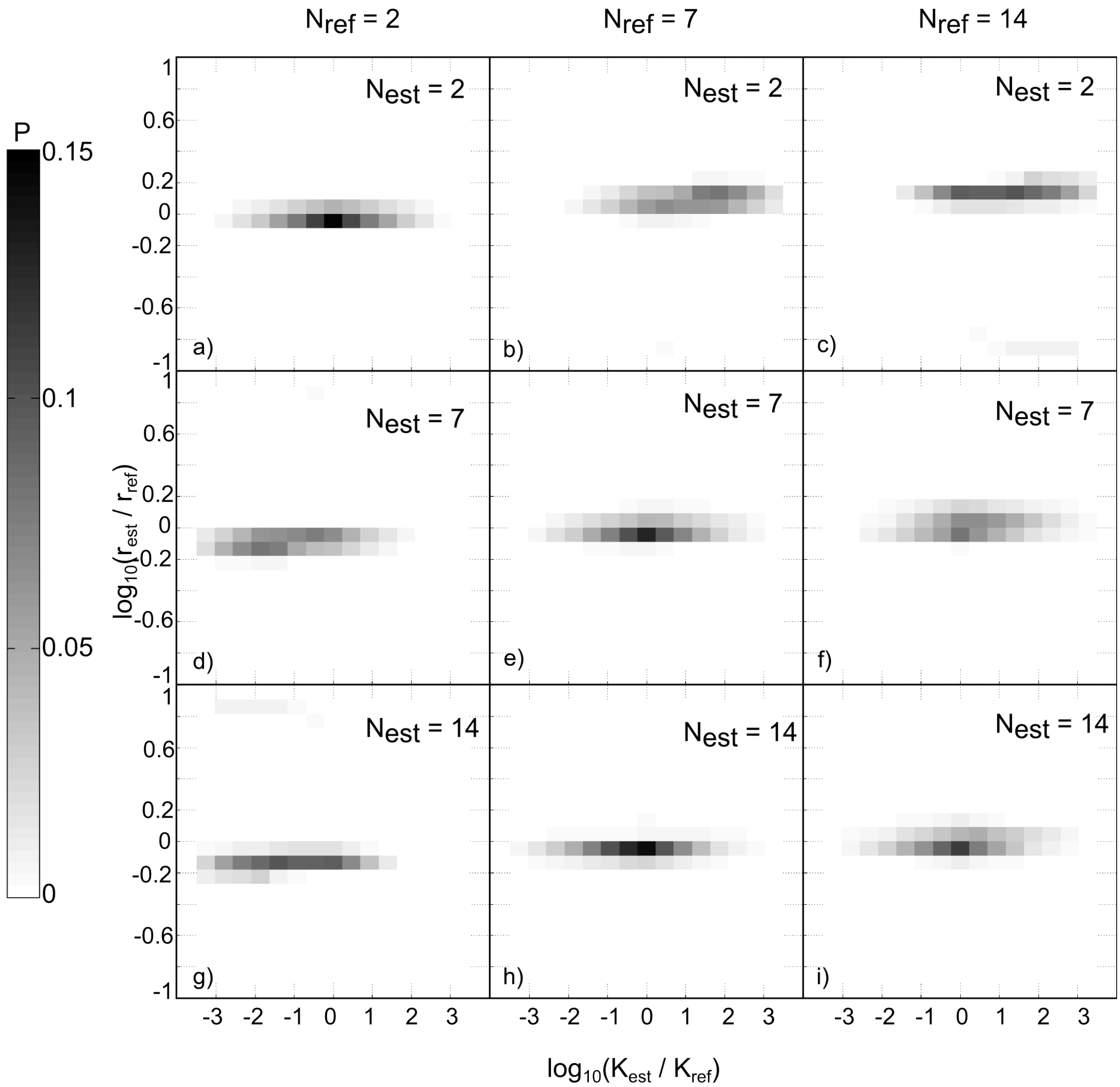


Figure 7.8: conditional joint probability distribution functions using the 3 different families of geometries; references: a) 2 conduits, b) 7 conduits, c) 14 conduits); first row: $N = 2$, second row: $N = 7$, third row: $N = 14$.

For a real large scale model like the model of the Valley of *les Ponts-de-Martel* shown in chapter 6, this is not possible and a faster method is required. This is why we test in this section the feasibility to use a gradient based approach to accelerate the search of the best physical parameters when the geometry is fixed. For that purpose, we run a systematic parameter search to known exhaustively the objective function and then test the efficiency of the gradient search using the software PEST (Doherty et al., 1994) (Doherty et al., 1994) on a synthetic case.

7.3.1 Model description

In this section we use a 3D realization of a synthetic karstic aquifer generated using the *pseudo genetic* methodology. The 3D mesh is quite small to reduce the CPU consumption; it is composed by 40 elements in x direction, 30 in y direction and 3 in z direction. The cells are cubic, with a width of 50m. Therefore the surface of the model is $2 \times 1.5 = 3(km^2)$. The flow model that has been used here is similar to the one described in section 6.3.8, i.e. a 3D finite element model, with the karst conduits meshed as 1D pipes along the matrix 3D cubes (figure 6.2). The flow is simulated in transient state using GROUNDWATER

ID	X	Y
H1	1000	750
H2	650	500
H3	650	1000
H4	1350	500
H5	1350	1000

Table 7.1: observation wells coordinates

parameter	reference	min	max	nb
$\alpha\,[-]$	0.2	0.05	1.5	16
$\beta\,[-]$	1.5	0.1	5	14
$K_M\,[m/s]$	$10^{-5.5}$	10^{-7}	10^{-4}	16

Table 7.2: Ranges of parameter values for the systematic search. *min/max*: bounds for the parameter; *nb*: number of intervals. α and K_M increments have been computed in a log10 scale.

(Cornaton, 2007).

The synthetic model could represent a small karst aquifer in the Jura mountains. The recharge function that has been used is the one described for the Valley of *les Ponts-de-Martel* aquifer (section 6.3.4). The reference simulation geometry is shown in figure 6.4. The reference parameters ($\alpha = 0.2$, $\beta = 1.5$, equation 6.1; and $K_M = 10^{-5.5}[m/s]$, the hydraulic conductivity of the matrix) have been chosen on the base of the flow simulations results. These parameters allow to represent satisfactorily a realistic karst behavior for the synthetic model (Fig. 7.9). To monitor the head in the model, five observation wells are placed at the base of the model at coordinates given in table 7.1. The spring is located at (1000;500) and its altitude is one meter (head is equal to the altitude).

7.3.2 Systematic search

The systematic search of best parameters is performed particularly to investigate the shape of the objective function ϕ. The three parameters α, β and K_M have been systematically tested. K_M is sampled in a logarithmic scale, because it can vary over several orders of magnitude. The parameter ranges that have been tested are given in table 7.2: 3584 simulations have been made.

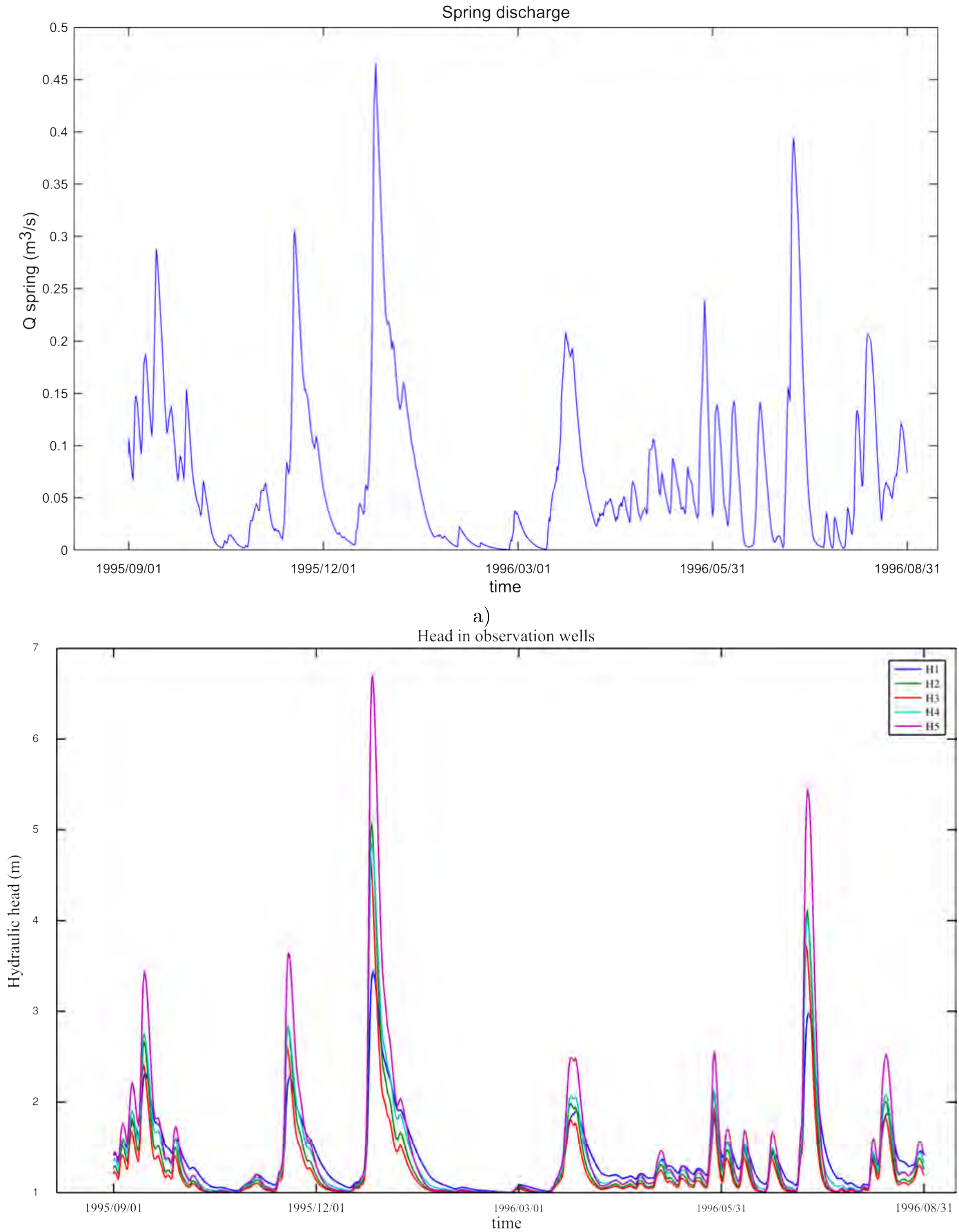
The objective function ϕ for this test is a sum of squared weighted residual. It has been chosen because it is the same objective function as that implemented in PEST (Doherty et al., 1994):

$$\phi = \sum_{i=1}^{n_o} (w_i \cdot r_i)^2 \tag{7.7}$$

where n_o are the number of observations, r_i are the residuals and w_i are the weight assigned to each observation i. The observations are taken from 6 different time series. The first one is the spring discharge, and the other 5 are the heads in the observation wells. To give the same weight to the discharge observation and to the head observation, a weight of $1/\sigma$ is given to the discharge observation (where σ is the variance of each time series) and a weight of $1/\sigma * 1/5$ is given to the head observations.

The results of this systematic search are shown in figure 7.10. The objective function ϕ is very sensitive to variations in the hydraulic conductivity of the matrix. Looking in detail to the shape of ϕ in the K_M dimension, it can be noticed that the value of K_M corresponding to the minimum of ϕ is well defined, but it is far less defined for its high values. This can be understood, because when the hydraulic conductivity of the matrix is too low, the heads in the model will be much too high and unrealistic, therefore ϕ is extremely high. But when the hydraulic conductivity of the matrix allows to drain the diffuse recharge that is infiltrated directly on it, this parameter becomes far less sensitive to the variations.

Another feature of ϕ is that several combinations of α and β can give similar results. The shape of the minimum values of ϕ is quite well defined with respect to K_M, but much less clearly for α and β. This is not surprising because both of these parameters influence the radius r of the conduits. α influences the *base radius* principally, while β influences the *size distribution*, i.e. the differences between the biggest



b)

Figure 7.9: Transient flow results for reference simulation. a) Spring discharge and b) calculated head in observation wells

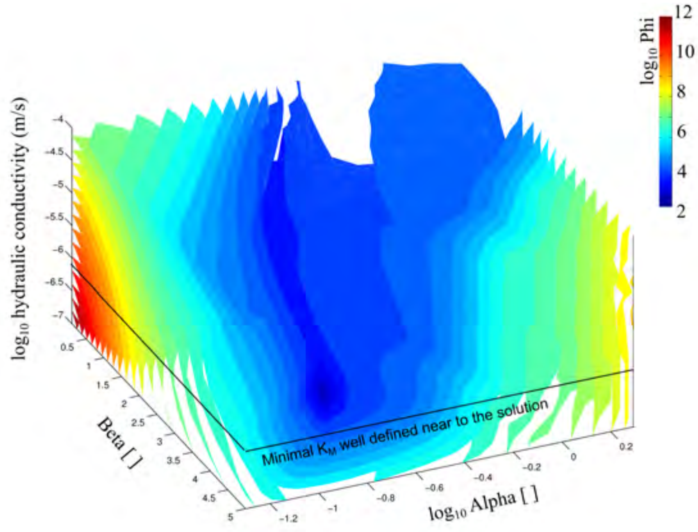


Figure 7.10: Systematic search result. This figure represent some isovalues surfaces of the objective function ϕ in the 3 parameter space. The axis of the parameter α and the hydraulic conductivity of the matrix are displayed in log scale for more clarity.

and smallest conduits. Therefore the objective function ϕ has the shape of a *valley*. Several combinations of α and β values provide a good fit. When α is small the good fits are obtained with large values of β: the smallest conduits are small, and the biggest conduits must be sufficiently large to drain all the system. When α is large, the contrast of size between the smallest and biggest conduits is less strong, i.e. all the conduits have more similar sizes, but they are all bigger than in the first case, and the system can be drained equally efficiently.

7.3.3 PEST optimization

PEST (Doherty et al., 1994)is an iterative parameter estimation software. It is one of the most used parameter estimation software in the industry, it is free and open-source. Here we just use it as a gradient based optimization technique. The idea is to test the ability of this kind of methods to retrieve the optimal physical parameters for the karst realizations. The way PEST works is conceptually simple:

- compute the model
- compute the value of ϕ
- compute the gradient of ϕ for each parameter
- update the parameters using the Levenberg-Marquardt method (Marquardt, 1963; More, 1978)
- compute the model again
- and so on, until a convergence criterion is reached

To test the ability of PEST (Doherty et al., 1994) to find the best physical parameters for one karst realization, eight optimization runs have been done, starting approximatively from all the corners of the cube (in the parameter space) defined by the parameters of table 7.2. To show the results of these optimization runs, the paths defined by the parameter variations (in the parameter space) have been displayed in figure 7.11. The value of the objective function ϕ for each parameter combination is displayed as colored balls along the paths. Moreover, the same paths are plotted together with the results of section 7.3.2 to show the paths within the plot of ϕ in figure 7.12.

The results of this test indicate that from the eight optimization runs, one run did not converge toward the right parameters at all, and 5 of them approached very closely the exact reference parameters. The others 2 where *stuck* in a local minimum of ϕ, as it can be noted in figure 7.12. This is a consequence of

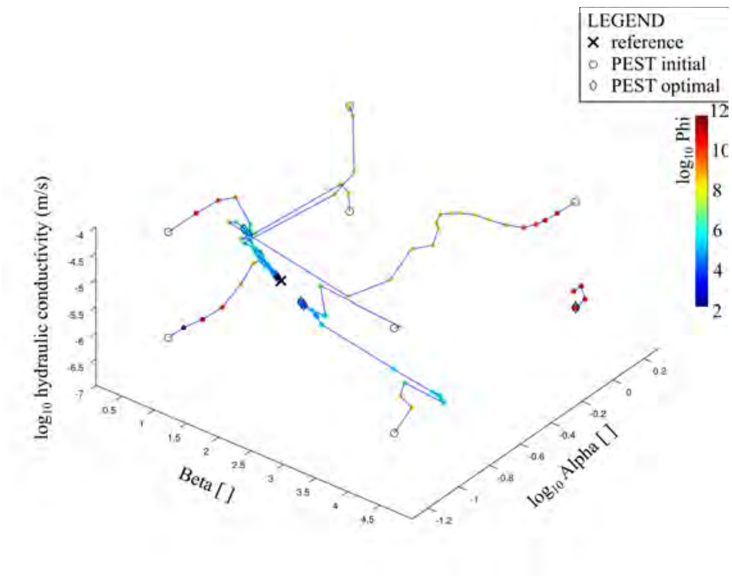


Figure 7.11: Eight parameter estimation runs using PEST (Doherty et al., 1994), displayed in the parameter space. 7 runs have given promising results and found a parameter set combination similar to the reference.

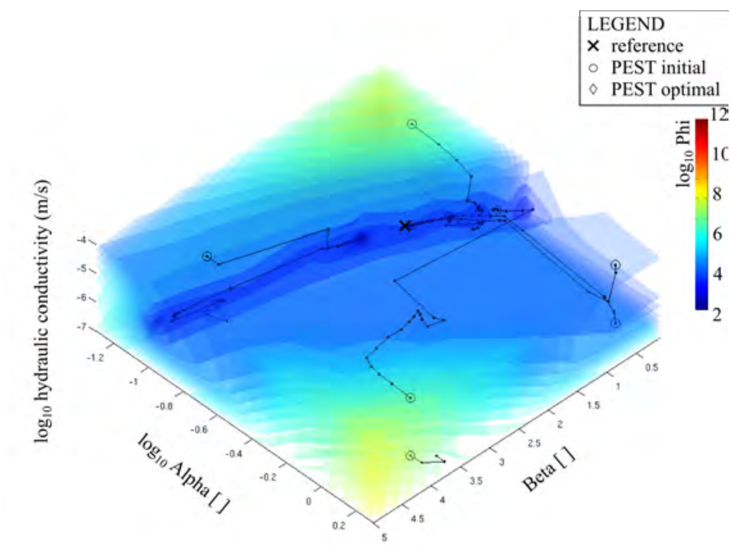


Figure 7.12: combination of figure 7.10 (with some transparency) and figure 7.11. We notice that some PEST (Doherty et al., 1994) optimial parameter sets are in local minima of the objective function.

the shape of ϕ described in the previous section (7.3.2). The model run time for the forward problem is approximately 5 minutes, depending on the parameters. The PEST (Doherty et al., 1994) runs needed between 34 and 274 model runs to reach convergence in this case (except one optimization, which failed). Compared to the 3584 simulations of the systematic search it is a gain in efficiency of 10-100 times.

These test results show that it is possible to use this kind of approach to find the best parameters for the karst realizations. As the results of the optimization depend on the starting positions (in the parameter space), one possibility to use this method could be to run several parameter estimations and look for the ones with the best values of ϕ.

7.4 Conclusions and further outlooks

In this chapter, several aspects of the inverse problem resolution in karst modeling have been studied. The numerical experiments did not aim at solving the inverse problem, but they allow to better understand the underlying problems. The inverse framework that we propose at the beginning of the chapter, even if it has not been tested completely, seems to be applicable. Our proposition is simple: a set of many different (the most different as possible) karst geometrical realizations are first generated using the *pseudo-genetic* methodology (chapters 4 and 5), and secondly for each one of them the best physical parameters are found using a parameter estimation technique. This set of optimized simulations is then used for statistical analysis, and for uncertainty analysis. The present study has investigated the possibility to consider feasible a similar approach in order to be able, in the future, to use inverse problem frameworks to calibrate models of karst aquifers.

The first test has investigated the concepts of the optimization of multiple realizations systematically. It shows that both physical and geometrical parameter can be identified in an inverse framework. This has been achieved by statistically comparing the results of 36'000 flow and transport simulations (18'000 distinct flow fields with 2 tracers). It shows that the transport simulation results have a very strong impact on determining the geometrical properties of a karst aquifer. This is not possible when looking only at flow simulation results. Indeed, transport is strongly related to the exact flow paths, and can provide a very distinctive response that flow alone does not.

This suggests therefore that any tracer test (or maybe also natural tracers) information should be used to calibrate a karst model and in any inverse problem framework. The objective function should be a composite function with a strong weight on transport. We observed that with our model setup there was a larger uncertainty on the hydraulic conductivity of the limestone matrix than on conduit radius. Furthermore we noticed that karst system with a very small number of drains (2 conduits) where more identifiable than other less extreme systems, like those with 7 or 14 conduits.

The second test is mainly intended to investigate the possibility of improving the effectiveness of the algorithm that we propose by accelerating the search for the optimal parameters when the geometry is fixed using e.g. gradient methods. In our test, a standard optimization technique like the Levenberg-Marquardt method (Marquardt, 1963) (such as implemented in PEST) is used to optimize the flow parameter of on single karst realization. The model that has been used for this test considers more complicated physics than the one of the first test. In this case, the flow in the conduits is simulated as turbulent flow using the Manning/Strickler equation. A systematic analysis of the objective function ϕ shows that several combinations of the parameters that influence the radius of the conduits (α and β) can give more or less equally satisfactorily results. The hydraulic conductivity of the matrix is very sensitive to the low conductivity values, but as soon as it is sufficient to infiltrate the recharge and lead this water to the conduits, ϕ is far less sensitive to its variations. This test also shows that PEST (Doherty et al., 1994) can be considered as optimization tool for the current problem, because it allows to reduce the number of needed model runs by at least one order of magnitude with respect to the systematic search.

Chapter 8

Conclusions

8.1 Main results

This thesis proposes a methodology to model the geometry of karstic conduits in a stochastic framework, and to include them directly in a complete modeling workflow. The workflow includes geological modeling, fracture modeling, conduit modeling, meshing for finite element models, estimating the boundary conditions, and optimizing the parameters in an inverse procedure. The proposed methodology can now be used as a starting point to model karst aquifers. Furthermore, the thesis provides several practical considerations that should help to build new karst models in the practice.

The major innovation in this methodology is the *pseudo genetic* Stochastic Karst Simulator (SKS). The main progress that has been realized here is the coupling of speleogenetic considerations with a fast stochastic algorithm. The related tool allows generating a large number of different models to account for the great uncertainty of karst conduits geometry. The importance of the stochastic approaches in karst modeling was already pointed out by Jaquet (1998) to take into account for the great variability of karst: *"cette variabilité finit toujours par éclater, au hasard, et en pleine figure, au cours de nos recherches"*. SKS allows also to quickly test different hypotheses on the main controlling factors for the speleogenesis. It is very versatile and allows to quickly include new available knowledge about a system.

In addition, the approach of considering the conduits like the minimum effort paths between inlets and outlets allows to condition the models very easily, which was not the case with the previously existing methods. The use of a fracture model and bedding planes to simulate the initial heterogeneities of the aquifer add consistency to the approach with respect to the most up to date conceptual model about speleogenetic processes and karst initialization.

Another important aspect of SKS, is that it allows to exploit directly all the data that are usually available when studying karst aquifers. It allows to use geological data, fracturation observations and geomorphological observations (inlets, springs). It allows also to use the results of tracer tests and other time series, like spring hydrographs to control the connectivity of the final models.

Concerning flow simulation, the new concept that has been developed is the integration of the hierarchy of the conduits to increase their order, which is used to compute their radius distribution within the model. Moreover, the use of a complex recharge function gives promising results, even if it has not been tested extensively: the distribution of the recharge directly into the inlets according to the surface that they drain allows to approximate the role of the epikarst in a simple manner.

Finally, the numerical tests that have been done to investigate the possibility to use inverse problem methods in the context of karst simulation gives promising results. They show on the one hand that it should be possible to retrieve both geometrical parameters and physical parameters using an inverse approach. On the other hand, the use of standard optimizations techniques like the Levenberg-Marquardt algorithm for the optimization of the flow properties of each geometrical realization seems feasible. Therefore, the proposed procedure: first generate a lot of realizations and then optimize each of them, should work even if it could be very computationally demanding.

8.2 Further outlooks

The *pseudo genetic* approach described in this thesis has the advantage to have a high computational efficiency. This allows to build quickly a multitude of realizations of karst networks that are equivalently conditioned to field data and observations. This methodology is an ongoing research and further enhancements can be expected for future work. In particular, two aspects have to be further investigated: the validation of the methodology, and the improvements of the *Stochastic Karst Simulator* (SKS).

It is difficult to validate the proposed karst models by comparison to the reality or to speleogenetical models. On the one hand, the degree of realism of the karst network produced by SKS cannot yet be defined, because it is not yet clear what is the meaning of "reality" for karst features. Further research is clearly needed to understand which metrics have to be used to "measure" the karst systems. For sure some basic statistics about the orientations of the conduits can be used, but this will be strongly dependent on the fracturation model results. Topological indices have to be found, to allow a better characterization of the different karst network types. This topic is an ongoing research here at the university of Neuchâtel, but unfortunately it goes beyond the focus of this thesis, which is principally the development of the modeling tools. This further step in the research is fundamental, on the one hand to validate the methodology developed in this thesis, and on the other hand to be able to directly generate different karst types (anastomous, branchwork, etc, according to the terminology Palmer (1991)).

Another way of validating the SKS results, could be the comparison with speleogenetical models (Worthington et al., 2000; Annable, 2003; Dreybrodt et al., 2005; Kaufmann, 2009; Gabrovsek and Dreybrodt, 2010). Unfortunately these models are not intended to be applied on real systems, but to understand the physical processes underlying the karst genesis. To be able to compare them with SKS, one should first find a way to compare them to reality quantitatively, and this lead to the same shortcoming as written above, i.e. one need to define the metrics to be used to measure the karst systems.

Several aspects of the *Stochastic Karst Simulator* (SKS) could be further enhanced. The particle tracking method that is used in SKS is very simple, and has shown satisfactory results. A further improvement could be to try a different particle tracking technique, than the actual one, which is based only on the scan of the minimal local values of T. A more sophisticated particle tracking technique could be for example based on the gradient ∇T of the time map. This could be an interesting further research topic. A more sophisticated particle tracking technique could probably reduce the artifacts due to the implementation of FMA, which prefers diagonal direction by computing the particle trajectories in a vectorial space, taking into account a wider region sub-region around the walker. Particle tracking techniques based on something similar to Lagrangian approaches (see e.g. Ouellette et al., 2006) could also be tested and would probably lead to better walkers paths, but could lead to higher computation time.

The speed function as defined in SKS does not consider the gravity effects. This issue has been solved in Chapter 5 by cutting the propagation medium (Section 5.4). Further research on the speed function K could probably enhance the ability of the code to simulate gravity effects. First of all an anisotropic speed function could be used, with a preferential direction in the vertical direction. Even if this approach has not been tested, it does not appear to be a very efficient way, because this would add a bias in the simulations, and it is not clear what could happen to the saturated (sub-horizontal) conduits. Alternatively, the speed function could be influenced by a cost function similarly to Collon-Drouaillet et al. (2012). In their approach this gives promising results, but the shape of the conduits simulated by Collon-Drouaillet et al. appear to be very "straight". In the vadose zone they appear to be almost perfectly vertical, while in the phreatic zone they are completely horizontal. Furthermore, the computation of the cost function $\mathcal{A}\star$ is less efficient, compared to a FMA. But the concept of penalizing unrealistic features by minimizing an error function could be added in the SKS approach. Further research is needed to implement this feature. It is not clear also, if it should be inserted in the speed function definition directly or in the walker algorithm during the simulation of karst paths. Another way to enhance the realism of the speed function, could be to use a speed function influenced by a flow field (or flow velocity field) resulting from a flow simulation. While this approach would be computationally inefficient, because flow simulation (which are much longer compared to FMA) should be run before the FMA, it could (probably) enhance the results because the flow paths would already provide an indication of the main flow direction, which include gravity effects and preferential flow paths through aquifer formations. One more investigation that could be achieved is to add a stochastic component to the speed K field in the matrix, to add more variability withing the aquifer formation, which is actually considered as homogeneous.

The coupling of conduit-matrix flow based on imposing a head continuity used in this study may be critized because it has been shown that with this approach, simulated exchange fluxes across conduit-matrix interface are sensitive to the spatial discretization of the matrix (De Rooij, 2007). In a recent study, De Rooij et al. (2012) introduce a more accurate approach for coupling conduit-matrix flow based on a so-called Peaceman well-index (Peaceman, 1990). They demonstrate that the use of a well-index lead to less sensitivity of the model to spatial discretization. Except for coupling conduit-matrix flow, their model also includes coupling with channel and overland flow. Moreover the model can handle variably saturated conditions. De Rooij et al. (2013) explain in detail the methodology: they use a cell centered finite-differences discretization scheme; multiple one-dimensional cells (conduits) can be placed inside of the matrix cells, and they do not need to be aligned with the gridlines of the matrix grid. Moreover their vertices do not need to coincide with the cell centers of the matrix cells. In practice, the conduits can be placed anywhere in the model. It would be very interesting to couple such a flow simulator with a geometrical conduit generator like SKS.

One issue for the physical simulations is their large computation time. It can make the inverse problem insolvable. One could therefore try to reduce the size of the simulation by upscaling the flow properties and diminishing the number of elements of the simulations. Bi et al. (2009) shows the benefits of upscaling in a carbonate (paleokarst) petroleum reservoir. In their model, multiphasic flow simulation and complex transport are simulated, but using upscaling they can compute the models in a reasonable time. Vitel (2007) developed a very powerful upscaling technique for fractured reservoirs: she reduced the regular grid of her model (several millions of elements) to a very light irregular mesh (few thousands of elements) by keeping only the connected features for the fractures and a few elements for the matrix. A similar technique could be applicable also to karst simulations, because the most relevant elements for flow are the conduits. One could think also to consider only the pipes without the matrix, as proposed by Jeannin (2001). Further research is needed to understand in which cases this oversimplification of the model could be applied, but the resolution of the flow model would become significantly lighter in computational demands, because of the consistent reduction of the number of nodes (more or less 1% of the complete 3D mesh).

The above considerations show that an additional effort should also be put in the improvement of the meshing techniques. It is still difficult to find a sufficiently robust tool to produce complex finite element mesh in very complicated geological settings. Even sofisticated tools like LaGriT or CGAL are difficultly able to produce meshes that include tetrahedrons (or other flexible geometry elements) for the 3D matrix, triangles for the surfaces (like faults and other geological discontinuities) and 1D pipes for the conduits if the geometries are very complex. Sethian (2008) propose to use level-set methods (similar to FMA) to build progressive-refined meshes. This could be an interesting direction to be able to produce a final mesh with a local refinement near the conduits and a progressive coarsening in the matrix, where too much detail in the mesh is not needed.

The inverse problem also needs furthest additional investigations. The proposed way of solving the inverse problem in this thesis is computationally demanding. Scheidt et al. (2011) propose to use a distance-based procedure, which relies on the rapid construction of multiple models. They construct a large set of high-resolution models constrained to dynamic data. Their methodology allows to select a subset of reservoir models reflecting the same uncertainty in flow response as the full set. The gain in computational demand is very high with this kind of approach: they demonstrated the efficiency of such an approach on multi-Gaussian fields. Another idea to improve the effectiveness of the algorithm could be to use approximate physics instead of the true physical solver (a so-called *proxy*) instead of solving the direct problem for each realization. Then estimate an error function to model the difference between the proxy result and the direct problem. The error function can be estimated by solving the complete problem for a few realizations and then using interpolations techniques like kriging to map the error function as done in Ginsbourger et al. (2013).

To conclude, the *pseudo genetic* methodology implemented in the *Stochastic Karst Simulator* needs now to be applied and tested for a rather large range of contexts possibly coupled with simulation of flow and transport, this is fundamental to validate its applicability on real problems and to determine the additional needs in terms of research. The aspect that mostly needs further tests and research is the inverse problem because the proposed inverse methodology could not yet be sufficiently tested due to time limitations; we expect that a significant amount of work has still to be done in gaining efficiency for large scale problems.

Bibliography

Bibliography

Andreo, B., N. Goldscheider, I. Vadillo, J. M. Vias, C. Neukum, M. Sinreich, P. Jimenez, J. Brechenmacher, F. Carrasco, H. Hotzl, M. J. Perles, and F. Zwahlen (2006). Karst groundwater protection: First application of a pan-european approach to vulnerability, hazard and risk mapping in the sierra de libar (southern spain). *Science of the Total Environment 357*(1), 54–73.

Annable, W. K. (2003). *Numerical analysis of conduit evolution in karstic aquifers.* Ph. D. thesis, University of Waterloo, Canada.

Atteia, O., S. Gogniat, and R. Kozel (1996). Apport de l'hydrologie, de la géophysique et des essais de traçage à la connaissance de l'aquifère karstique de la Noiraigue (Jura Neuchâtelois, Suisse). *Bulletin d'Hydrogéologie / Université de Neuchâtel* (15), 33–61.

Atteia, O., D. Perret, T. Adatte, R. Kozel, and P. Rossi (1998). Characterization of natural collodis from a river and spring in a karstic basin. *Environmental Geology 34*, 257–269.

Bailly, D., R. Ababou, and M. Quintard (2008). Characterization and hydraulic behaviour of highly fissured 3d geologic media (karsts). In *Flow and Transport in Heterogeneous Subsurface Formations: Theory, Modelling & Applications*, Istambul, Turkey.

Bakalowicz, M. (2005). Karst groundwater: A challenge for new resources. *Hydrgeology Journal 13*(1), 148–160.

Bazagette, L. (2004). *Relations plissement/fracturation multi échelle dans les multicouches sédimentaires du domaine élastique/fragile: accommodation discontinue de la courbure par la fracturation de petite échelle et par les articulations. Possibles implications dynamiques dans les écoulements des réservoirs.* Memoires geosciences-Montpellier. Institut des Sciences de la Terre de l'Eau et de l'Espace de Montpellier.

Bear, J. (1972). *Dynamics of Fluids in Porous Media.* Dover publications.

Bear, J. and M. Jacobs (1965). On the movement of water bodies injected into aquifers. *Journal of Hydrology 3*, 37–57.

Bear, J., C.-F. Tsang, and G. de Marsily (1993). *Flow and contaminant transport in fractured rock.* Academic Press, San Diego, CA (United States).

Bellahsen, N., P. Fiore, and D. Pollard (2006). The role of fractures in the structural interpretation of sheep mountain anticline, wyoming. *Journal of Structural Geology 28*(5), 850–867.

Bi, L., G. Qin, P. Popov, Y. Efendiev, and M. Espedal (2009). An efficient upscaling process based on a unified fine-scale multi-physics model for flow simulation in naturally fracture carbonate karst reservoirs. In *SPE/EAGE Reservoir Characterization and Simulation Conference.*

Biondic, B., M. Bakalowicz, and E. Commission (1995). Cost action 65 - hydrogeological aspects of groundwater protection in karstic areas. final report. Technical report.

Bonneau, F., V. Henrion, G. Caumon, P. Renard, and J. Sausse (2013). A methodology for pseudo-genetic stochastic modeling of discrete fracture networks. *Computers & Geosciences in press.*

Borghi, A. (2008). Modélisation 3d d'un aquifère karstique, la Vallée des Ponts-de-Martel et la source de la Noiraigue. Master's thesis. Université de Neuchâtel, Switzerland.

Borghi, A., P. Renard, and S. Jenni (2012). A pseudo-genetic stochastic model to generate karstic networks. *Journal of Hydrology 414 – 415*(0), 516 – 529.

Bour, O., P. Davy, C. Darcel, and N. E. Odling (2002). A statistical scaling model for fracture network geometry, with validation on a multiscale mapping of a joint network (hornelen basin, norway). *Journal of Geophysical Research 107.*

Bresenham, J. E. (1965). Algorithm for computer control of a digital plotter. *IBM Systems journal 4*(1), 25–30.

Brunner, P., C. Simmons, and P. Cook (2009). Spatial and temporal aspects of the transition from connection to disconnection between rivers, lakes and groundwater. *Journal of Hydrology 376*(1), 159–169.

Caers, J. and T. Hoffman (2006). The probability perturbation method: A new look at bayesian inverse modeling. *Mathematical Geology 38*, 81–100.

Calcagno, P., J. Chilès, G. Courrioux, and A. Guillen (2008). Geological modelling from field data and geological knowledge: Part i. modelling method coupling 3d potential-field interpolation and geological rules. *Physics of the Earth and Planetary Interiors 171*(14), 147 – 157. Recent Advances in Computational Geodynamics: Theory, Numerics and Applications.

Camac, B. A. and S. R. Hunt (2009). Predicting the regional distribution of fracture networks using the distinct element numerical method. *AAPG Bulletin 93*.

Carrera, J., A. Alcolea, A. Medina, J. Hidalgo, and L. Slooten (2005). Inverse problem in hydrogeology. *Hydrogeology Journal 13*, 206–222.

Carrera, J., E. Custodio, A. Gurgui, J. Lobo Ferreira, A. NATO, C. Series, et al. (1988). State of the art of the inverse problem applied to the flow and solute transport equations. *Groundwater flow and quality modelling.*, 549–583.

Castaing, C., A. Genter, B. Bourgine, J. Chilès, J. Wendling, and P. Siegel (2002). Taking into account the complexity of natural fracture systems in reservoir single-phase flow modelling. *Journal of Hydrology 266*, 15.

Caumon, G., P. Collon-Drouaillet, C. LeCarlierdeVeslud, S. Viseur, and J. Sausse (2009). Surface-based 3d modeling of geological structures. *Mathematical Geosciences 41*, 927–945.

CGAL. Computational geometry algorithms library. Geometry Factory, http://www.cgal.org/.

Chilès, J.-P. (2005). Stochastic modeling of natural fractured media: A review. In O. Leuangthong and C. V. Deutsch (Eds.), *Geostatistics Banff 2004*, Volume 14 of *Quantitative Geology and Geostatistics*, pp. 285–294. Springer Netherlands. 10.1007/978-1-4020-3610-1_29.

Collon-Drouaillet, P., V. Henrion, and J. Pellerin (2012). An algorithm for 3d simulation of branchwork karst networks using horton parameters and ⊣⋆ application to a synthetic case. *Geological Society, London, Special Publications 370*, pp. 295–306.

Cornaton, F. (2007). Ground water : A 3-d ground water and surface water flow, mass transport and heat transfer finite element simulator.

Cornaton, F. and P. Perrochet (2002). Analytical 1d dual-porosity equivalent solutions to 3d discrete single-continuum models. application to karstic spring hydrograph modelling. *Journal of Hydrology 262*, 165–176.

Corre, B. (1994). Diverses approches de simulations des écoulements dans un karst: application à rospo mare (italie) ;various approaches for flow simulations in a karst: application to rospo mare field (italy). *Bulletin des centres de recherches exploration-production Elf-Aquitaine 18*, 299–313.

Daly, D., A. Dassargues, D. Drew, S. Dunne, N. Goldscheider, S. Neale, I. C. Popescu, and F. Zwahlen (2002). to karst-groundwater-vulnerability assessment and mapping. *Hydrogeology Journal 10*, 340–345.

Daubrée, A. (1887). *Les eaux souterraines à l'époque actuelle, leur régime, leur température, leur composition.* Dunod éditeurs.

Davy, P., R. Le Goc, C. DARCEL, O. Bour, J.-R. De Dreuzy, and R. Munier (2010). A likely universal model of fracture scaling and its consequence for crustal hydromechanics. *Journal of Geophysical Research B: Solid Earth 115*, B10411. 13 p. PL02518 PL02518.

de Marsily, G. (1994). Quelques réflexions sur l'utilisation des modèles en hydrologie. *Revue des sciences de l'eau / Journal of Water Science 7*, 219–234.

De Rooij, R. (2007). *Towards improved numerical modeling of karst aquifers: coupling turbulent conduit flow and laminar matrix flow under variably saturated conditions.* Ph. D. thesis, University of Neuchâtel, Switzerland.

De Rooij, R., W. D Graham, and P. Perrochet (2012). H11f-1252: A newly designed numerical tool for advanced flow and solute transport modeling in large and complex karst systems. In *AGU fall meeting*.

De Rooij, R., P. Perrochet, and W. Graham (2013). From rainfall to spring discharge: Coupling conduit flow, subsurface matrix flow and surface flow in karst systems with a discrete-continuum model. paper in preparation.

Dershowitz, W., P. L. Pointe, and T. W. Doe (2004). Advances in discrete fracture network modeling. In *US EPA/NGWA fractured rock conference, Portland, Maine, USA*.

Desor, E. (1865). Expériences faites pour connaître le temps de parcours des eaux depuis les ponts jusqu'aux sources de la noiraigue. *Bulletin de la société neuchâteloise des sciences naturelles 7*, 37–40.

Dietrich, C. and G. Newsam (1993). A fast and exact method for multidimensional gaussian stochastic simulations. *Water Resources Research 29*(8), 2861–2869.

Doherty, J., L. Brebber, and P. Whyte (1994). Pest: Model-independent parameter estimation. *Watermark Computing, Corinda, Australia 122*.

Dreybrodt, W., F. Gabrovsek, and R. Douchko (2005). *Processes of speleogenesis: a modelling approach*, Volume 4. Zalozhba.

edric.ch (2012). Rs_user_guide. `http://edric.ch/images/techpdf/RS_USERS_GUIDE.pdf`.

Eisenlohr, L., M. Bouzelboudjen, L. Király, and Y. Rossier (1997). Numerical versus statistical modelling of natural response of a karst hydrogeological system. *Journal of Hydrology 202*, 244–262.

Eisenlohr, L., L. Király, M. Bouzelboudjen, and Y. Rossier (1997). Numerical simulation as a tool for checking the interpretation of karst spring hydrographs. *Journal of Hydrology 193*, 306–315.

El-Hajj, A., M. Bakalowicz, and W. Najem (2006). Hydrogéologie des calcaires crétaces du nord liban. In *The 3rd International Conference on the Water Resources in the Mediterranean Basin*, pp. 210.

Erzeybek Balan, S. (2012). *Characterization and Modeling of Paleokarst Reservoirs Using Multiple-Point Statistics on a Non-Gridded Basis.* Ph. D. thesis, The University of Texas at Austin.

Filipponi, M., P.-Y. Jeannin, B. Andreo, F. Carrasco, J. Duran, J. LaMoreaux, and J. LaMoreaux (2010). *Karst-ALEA: a Scientific Based Karst Risk Assessment for Underground Engineering*, pp. 435–440. Springer.

Filipponi, M., P. Y. Jeannin, and L. Tacher (2009). Evidence of inception horizons in karst conduit networks. *Geomorphology 106*, 86–99. Sp. Iss. SI.

Filipponi, M., D. A. Parriaux, and D. P.-Y. Jeannin (2009). *Spatial analysis of karst conduit networks and determination of parameters controlling the speleogenesis along preferential lithostratigraphic horizons*. Ph. D. thesis, EPFL–Ecole Polytechnique Fédérale de Lausanne, Switzerland.

Fleury, P., V. Plagnes, and M. Bakalowicz (2007). Modelling of the functioning of karst aquifers with a reservoir model: Application to fontaine de vaucluse (south of france). *Journal of Hydrology 345*, 38–49. Times Cited: 0.

Ford, D. and P. Williams (2007). *Karst hydrogeology and geomorphology.* Chichester, England: John Wiley and Sons, Ltd.

Fournier, L. (2012). Modélisation hydrogéologique 3d d'un aquifère kasrtique intégrant la fracturation de structures plissées. Master's thesis, University of Neuchâtel.

Fournillon, F., S. Viseur, B. Arfib, and J. Borgomano (2010). Insights of 3d geological modelling in distributed hydrogeological models of karstic carbonate aquifers. *Advances in Research in Karst Media*, 257–262.

Freudenreich, Y., A. A. D. Monte, E. Angerer, C. Reiser, and C. Glass (2005). Fractured reservoir characterisation from seismic and well analysis: a case study. In *67th EAGE conference and exhibition, madrid Spain.*

Gabrovsek, F. and W. Dreybrodt (2010). Karstification in unconfined limestone aquifers by mixing of phreatic water with surface water from a local input: A model. *Journal of Hydrology 386*, 130–141.

g◯cad. Geological object computer aided design. Paradigm™, http://www.gocad.org.

Geomodeller3D®. Editeur geologique 3d. INTREPID Geophysics, http://www.geomodeller.com.

Ghasemizadeh, R., F. Hellweger, C. Butscher, I. Padilla, D. Vesper, M. Field, and A. Alshawabkeh (2012). Review: Groundwater flow and transport modeling of karst aquifers, with particular reference to the north coast limestone aquifer system of puerto rico. *Hydrogeology Journal*, 1–21. 10.1007/s10040-012-0897-4.

Ghosh, K. and S. Mitra (2009). Two-dimensional simulation of controls of fracture parameters on fracture connectivity. *AAPG Bulletin 93*, 1517–1533.

Gillespie, P., C. Howard, J. Walsh, and J. Watterson (1993). Measurement and characterisation of spatial distributions of fractures. *Tectonophysics 226*, 28.

Gillmann, A. (2007). Etude du comportement des substances particulaires et solubles dans un système karstique de la source de la Noiraigue. Master's thesis, Université de Neuchâtel, Switzerland.

Ginsbourger, D., B. Rosspopoff, G. Pirot, N. Durrande, and P. Renard (2013). Distance-based kriging relying on proxy simulations for inverse conditioning. *Advances in Water Resources 52*(0), 275 – 291.

Gogniat, S. (1995a). aquifère de la Noiraigue (Neuchâtel, Suisse), approche par jeaugeages et multitraçages. Master's thesis.

Gogniat, S. (1995b). Etude des infiltrations rapides de l'aquifère de la Noiraigue (neuchâtel, suisse). Travail de Diplôme, Université de Neuchatel.

Goldscheider, N. and D. Drew (2007). *Methods in Karst Hydrogeology: IAH: International Contributions to Hydrogeology, 26*, Volume 26. Taylor & Francis.

Gringarten, E. (1998). Fracnet: Stochastic simulation of fractures in layered systems. *Computer and Geosciences 24*, 8.

Groves, C. and A. Howard (1994). Early development of karst systems: 1. preferential flow path enlargement under laminar flow. *Water Resources Research 30*(10), 2837–2846.

Hadamard, J. (1932). *Le problème de Cauchy et les équations aux dérivées partielles linéaires hyperboliques.*

Hart, P., N. Nilsson, and B. Raphael (1968). A formal basis for the heuristic determination of minimum cost paths. *Systems Science and Cybernetics, IEEE Transactions on 4*(2), 100–107.

Hartmann, A., T. Wagener, A. Rimmer, J. Lange, H. Brielmann, and M. Weiler (2013). Testing the realism of model structures to identify karst system processes using water quality and quantity signatures. *Water Resources Research*.

Henrion, G. Caumon, and N. Cherpeau (2010a). *Conditioning Complex Natural Structures*, Volume 42.

Henrion, V., G. Caumon, and N. Cherpeau (2010b). Odsim: An object-distance simulation method for conditioning complex natural structures. *Mathematical Geosciences 42*(8), 911–924.

Henrion, V., J. Pellerin, and G. Caumon (2008). *A stochastic methodology for 3D cave system modeling*, Volume Vol 2, pp. 525–534.

Hill, M. E., M. T. Stewart, and A. Martin (2010). - evaluation of the MODFLOW-2005 conduit flow process. - *48*, – 559. - 549.

Hu, C., Y. Hao, T.-C. J. Yeh, B. Pang, and Z. Wu (2008). Simulation of spring flows from a karst aquifer with an artificial neural network. *Hydrological Processes 22*(5), 596–604.

Huntoon, P., A. N. Palmer, M. Palmer, and I. Sasowsky (1999). Karstic permeability: Organized flow pathways created by circulation. Volume 5.

Jaquet, O. (1998). *Modèle stochastique de la géométrie des réseaux karstiques.* Ph. D. thesis, Université deLausanne, Switzerland.

Jaquet, O. and P. Jeannin (1994). *Geostatistical simulations*, Chapter Modelling the karstic medium: a geostatistical approach. Kluwer Academic Publisher, London.

Jaquet, O., P. Siegel, G. Klubertanz, and H. Benabderrhamane (2004). Stochastic discrete model of karstic networks. *Advances in Water Resources 27.*

Jeannin, D. P.-Y. (2001). Modeling flow in phreatic and epiphreatic karst conduits in the hoelloch cave(muotatal, switzerland). *Water Resources Research 37, no 2*, pp. 191–200.

Jeannin, P. (1992). Géometrie des réseaux de drainage karstique: approche structurale, statistique et fractale. *Ann. Sci. de lUniversité de Besançon*, 1–8.

Jeannin, P. Y. (1996). *Structure et comportement hydraulique des aquifères karstiques.* Ph. D. thesis.

Jing, L. (2003). A review of techniques, advances and outstanding issues in numerical modelling for rock mechanics and rock engineering. *International Journal of Rock Mechanics and Mining Sciences 40*(3), 283 – 353.

Jourde, H., F. Cornaton, S. Pistre, and P. Bidaux (2002). Flow behavior in a dual fracture network. *Journal of Hydrology 266*, 99–119.

Käss, W., H. Behrens, T. Himmelsbach, and H. Hötzl (1998). *Tracing technique in geohydrology.* Balkema Rotterdam.

Kaufmann, G. (2009). Modelling karst geomorphology on different time scales. *Geomorphology 106*, 62–77.

Kaufmann, G. and J. Braun (2000). Karst aquifer evolution in fractured, porous rocks. *Water Resources Research 36*, 1381–1391. Times Cited: 25.

Király, L. (1973). Notice explicative de la carte hydrogéologique du canton de neuchâtel. Supplément du bulletin de la société neuchâteloise des sciences naturelles, tome 96. Institut de géologie, Université de Neuchâtel, Suisse.

Király, L. (1979). écoulement. *Hydrogéologie 3*, 155–167.

Király, L. (1988). *Large scale 3-D groundwater flow modelling in highly heterogeneous geologic medium*, pp. 761–775.

Király, L. (2003). Karstification and groudwater. *Evolution of karst nf*, pp. 155–190.

Király, L., A. Burger, and L. Dubertret (1975). *État actuel des connaissances dans le domaine des caractères physiques des roches karstiques [Report on the present knowledge of the physical characters of karstic rocks]*, Volume Series 8 - Number 3, pp. 53–67.

Király, L., B. Mathey, and J.-P. Tripet (1971). Fissuration et orientation des cavités souterraines: région de la grotte de milandre (jura tabulaire). *Bull. Soc. Neuchâteloise Sc. Nat. 94*, 15.

Király, L., P. Perrochet, and Y. Rossier (1995). Effect of the epikarst on the hydrograph of karst sprigs: a numerical approach. *Bulletin d'hydrogéologie 14*.

Klimchouk, A. (2007). *Hypogene Speleogenesis: Hydrogeological and Morphologenetic Perspectives.* National cave and karst research institute.

Kloppenburg, A., J. C. Alzate, and G. R. Charry (2003). Building a discrete fracture network based on the deformation history: A case study from the guaduas field, colombia. In *VIII Simposio Bolivariano - Exploracion Petrolera en las Cuencas Subandinas.*

Kovács, A. (2003). *Geometry and hydraulic parameters of karst aquifers: a hydrodynamic modeling approach.* Ph. D. thesis, University of Neuchâtel, Switzerland.

Kresic, N. and Z. Stevanovic (2009). *Groundwater Hydrology of Springs ,Engineering, Theory, Management and Sustainability.* Springer.

Kuniansky, E. and W. Shoemaker (2008). Usgs releases conduit flow process (cfp) for modflow-2005. *2008 Joint Meeting of The Geological Society of America, Soil Science Society of America, American Society of Agronomy, Crop Science Society of America, Gulf Coast Association of Geological Societies with the Gulf Coast Section of SEPM*.

La Pointe, P. and J. A. Hudson (1985). Characterization and interpretation of rock mass joint patterns. *Special paper / Geological Society of America; 199 199*, 1–38.

LaGriT. Meshing for geological applications. Los Alamos National Laboratoy, https://meshing.lanl.gov/.

Lajaunie, C., G. Courrioux, and L. Manuel (1997). Foliations field and 3d cartography in geology: Principles of a method based on potential interpolation. *Mathematical Geology 29*, 571–584.

Lamelas, M. T., O. Marinoni, A. Hoppe, and J. de la Riva (2007). Doline probability map using logistic regression and gis technology in the central ebro basin (spain). *Environ Geol 54*, 15.

Lapcevic, P., K. Novakowski, and E. Sudicky (1999). Groundwater flow and solute transport in fractured media. *The Handbook of Groundwater Engineering 17*, 17–39.

Larocque, M., O. Banton, P. Ackerer, and M. Razack (1999). Determining karst transmissivities with inverse modeling and an equivalent porous media. *Ground Water 37*(6), 897–903.

Larocque, M., O. Banton, and M. Razack (2000). Transient-state history matching of a karst aquifer ground water flow model. *Ground Water 38*(6), 939–946.

Leibundgut, C., P. Maloszewski, and C. Külls (2009). *Tracers in Hydrology* (Blackwell and Wiley ed.).

Mace, L. (2006). *Caractérisation et modèlisation numérique tridimensionnelle des réseaux de fractures naturelles.* Ph. D. thesis, INPL, Nancy, France.

Maerten, L., D. D. Pollard, and R. Karpuz (2000). How to constrain 3-d fault continuity and linkage using reflection seismic data: A geomechanical approach. *AAPG Bulletin 84(9)*, 1311–1324.

Mallet, J. (1992). Gocad: A computer aided design program for geological applications. *Three-dimensional modeling with geoscientific information systems 354*, 123–142.

Mallet, J.-L. (2002). *Geomodeling.* Oxford University Press, USA.

Mangin, A. (1975). *Contribution a l'étude hydrodynamique des aquifères karstiques.* Ph. D. thesis, Université de Dijon.

Mariethoz, G. and P. Renard (2010). Reconstruction of incomplete data sets or images using direct sampling. *Mathematical Geosciences 42*(3), 245–268.

Mariethoz, G. and P. Renard (2011). Simulation of karstic networks using high order discrete markov processes. In *Proceeding of the 9th Conference on Limestone Hydrogeology, Besançon*, pp. 1–4.

Mariethoz, G., P. Renard, and J. Straubhaar (2010, November). The direct sampling method to perform multiple-point geostatistical simulations. *Water Resources Research 46*, W11536.

Marquardt, D. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial & Applied Mathematics 11*(2), 431–441.

MATLAB®. The language of technical computing.

Matthey, B. (1992). Déterminiation des zones de protection des captages des gorges de l'areuse. résultats des essais de traçage à thomasset, la tourne et rochefort, le 18 mai 1992. Rapport du bureau Bernard Matthey Ingénieurs-Conseil SA.

Matthey, B. (1994). Zones de protection des captages de l'areuse. multitraçage de juillet 1993. Rapport du bureau Bernard Matthey Ingénieurs-Conseil SA.

Milanovic, P. T. (2004). *Water resources engineering in karst.* CRC.

More, J. (1978). The levenberg-marquardt algorithm: implementation and theory. *Numerical analysis*, 105–116.

Morel, G. (1976). Étude hydrogéologique du bassin de la source de la noiraigue (jura). Université de Neuchâtel, Suisse.

Moyen, R. (2005). *Paramétrisation géo-chronologique.* Ph. D. thesis, INPL, Nancy, France.

Odling, N. (1997). Scaling and connectivity of joint systems in sandstones from western norway. *Journal of Structural Geology 19*(10), 1257–1271.

Ofterdinger, U., P. Renard, and S. Loew (2013). Hydraulic subsurface measurements and hydrodynamic modelling as indicators for groundwater flow systems in the Rotondo granite, central alps (Switzerland). *Hydrological Processes*.

Olson, J. (1993). Joint pattern development: Effects of subcritical crack growth and mechanical crack interaction. *Journal of Geophysical Research 98(B7)*, 12251–12265.

Ouellette, N., H. Xu, and E. Bodenschatz (2006). A quantitative study of three-dimensional lagrangian particle tracking algorithms. *Experiments in Fluids 40*(2), 301–313.

Palmer, A. (2007). *Cave geology*, Volume 454. Cave books Dayton, Ohio.

Palmer, A. N. (1991). Origin and morphology of limestone caves. *Geological Society of America Bulletin 103.*

Panagopoulos, G. (2012). Application of modflow for simulating groundwater flow in the trifilia karst aquifer, greece. *Environmental Earth Sciences 67*, 1877–1889.

Pardo-Igúzquiza, E., P. Dowd, C. Xu, and J. Durán-Valsero (2011). Stochastic simulation of karst conduit networks. *Advances in Water Resources*.

Peaceman, D. (1990). Interpretation of wellblock pressures in numerical reservoir simulation: Part 3–off-center and multiple wells within a wellblock. *SPE Reservoir Engineering 5*(2), 227–232.

PETREL®. seismic-to-simulation software (mark of Schlumberger).

Peyre, G. (2008). Toolbox fast marching. *MATLAB File exchange*.

PGN (2008). Évaluation du potentiel géothermique du canton de neuchâtel (pgn). CREGE 11-08/02. Rapport du groupe de travail PGN. Centre de recherche en géothermie. Université de Neuchátel, Suisse.

Price, N. (1966). *Fault and joint development in brittle and semi-brittle rock*, Volume 1. Pergamon Press London.

Price, N. and J. Cosgrove (1990). *Analysis of geological structures.* Cambridge University Press.

Priest, S. (1993). *Discontinuity analysis for rock engineering.* Chapman & Hall.

Priest, S. D. and J. A. Hudson (1981). Estimation of discontinuity spacing and trace length using scanline surveys. *International Journal of Rock Mechanics and Mining Science & Geomechanics Abstracts 18*(3), 183–197.

Rausch, R., W. Schäfer, R. Therrien, C. Wagner, et al. (2005). *Solute transport modelling: an introduction to models and solution strategies.* Gebrüder Borntraeger Verlagsbuchhandlung.

Ravbar, N. (2007). *The protection of karst waters.* Institut za raziskovanje krasa ZRC SAZU= Karst Research Institute at ZRC SAZU.

Reeves, D., Y. Zhang, G. Pohll, and D. Benson (2009). Frack: A freeware flow and transport suite for fractured media. *Proceedings of MODFLOW and MORE 2008: Ground Water and Public Policy*, 67–71.

Reimann, T., C. Rehrl, W. Shoemaker, T. Geyer, and S. Birk (2011). The significance of turbulent flow representation in single-continuum models. *Water Resources Research 47*(9), W09503.

Ronayne, M. J. and S. M. Gorelick (2006). Effective permeability of porous media containing branching channel networks. *Phys Rev E Stat Nonlin Soft Matter Phys 73*, 026305. Times Cited: 0.

Rouy, E. and A. Tourin (1992). A viscosity solutions approach to shape-from-shading. *SIAM Journal on Numerical Analysis 29*(3), 867–884.

Sassi, W., M. Guiton, Y. Leroy, J. Daniel, and J. Callot (2012). Constraints on bed scale fracture chronology with a fem mechanical model of folding: The case of split mountain (utah, usa). *Tectonophysics*.

Sauter, M., A. Kovacs, T. Geyer, and G. Teutsch (2006). Modelling karst groundwater hydraulics - an overview. *Grundwasser 11*, 143–156. Sauter, Martin Kovacs, Attila Geyer, Tobias Teutsch, Georg.

Scanlon, B. R., R. E. Mace, M. E. Barrett, and B. Smith (2003). 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*(14), 137 – 158.

Schardt, H. (1906). Deux coupes générales à travers la chaîne du jura. *Arch. Sci .Phy. Nat. Genève XXIII.*

Schardt, H. (1908). Les causes du plissement et les chevauchements dans le jura. *Eclogae Geologicae Helvetiae X*, 484–488.

Scheidt, C., J. Caers, Y. Chen, and L. Durlofsky (2011). A multi-resolution workflow to generate high-resolution models constrained to dynamic data. *Computational Geosciences*, 1–19.

Schwanghart, W. and N. Kuhn (2010). Topotoolbox: A set of matlab functions for topographic analysis. *Environmental Modelling & Software 25*(6), 770–781.

Sethian, J. A. (1996). A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences of the United States of America 93*, 1591–1595. Times Cited: 214.

Sethian, J. A. (1999). *Level Set Methods and Fast Marching Methods Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science.*

Sethian, J. A. (2001). Evolution, implementation, and application of level set and fast marching methods for advancing fronts. *Journal of Computational Physics 169*, 503–555. Times Cited: 59.

Sethian, J. A. (2008). *Level Set Methods and Fast Marching Methods Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science (re-edition).*

Sharp, I., P. Gillespie, A. Lonoy, S. Horn, and D. Morsalnezhad (2006). Outcrop characterization of fractured cretaceous carbonate reservoirs, zagros moountains, iran.

Shoemaker, W. (2010). Simulating solute transport within karst using the conduit flow process for modflow-2005. *2010 GSA Denver Annual Meeting*.

Smith, D., M. McKoy, and E. Boyle (2006). Fracgen stochastically generates fracture networks consistent with data. Technical report, National Energy Technology Laboratory (NETL), Pittsburgh, PA, and Morgantown, WV.

Sommaruga, A. (1996). *Geology of the central Jura and the molasse basin: new insight into an evaporite -based foreland fold and thrust belt.* Ph. D. thesis.

Sommaruga, A. (1997). *Geology of the Central Jura and the Molasse Basin: New Insight Into an Evaporite-based Foreland Fold and Thrust Belt.* Number vol. 1 in Mémoires de la Société Neuchâteloise des Sciences Naturelles. Société neuchâteloise des Sciences naturelles.

Spreafico, M. and R. Weingartner (2005). Hydrologie der schweiz - ausgewählte aspekte und resultate. Berichte des BWG, Serie Wasser, Nr. 7, Bern.

Straubhaar, J., P. Renard, G. Mariethoz, R. Froidevaux, and O. Besson (2011). An improved parallel multiple–point algorithm using a list approach. *Math Geosci 43*(3), 305–328. DOI: 10.1007/s11004-011-9328-7.

Strebelle, S. (2002). *Sequential simulation drawing structures from training images.* Ph. D. thesis.

Tarantola, A. (2005). *Inverse Problem Theory and methods for parameter estimation.*

Tarantola, A. and B. Valette (1982). Inverse problems = quest for information. *Journal of Geophysics 50*, 159–170.

Thierrin, J. (1987). Zones de protection du puits de captage de la brévine. commune de la brévine. Rapport du centre d'hydrogéologie de Neuchâtel. Université de Neuchâtel, Suisse. Non publié.

Twiss, R. and E. Moores (1992). *Structural geology.* W.H. Freeman.

Valley, B. (2002). La vallée des ponts-de-martel, rétro-déformation 3-d d'une structure complexe dans le jura neuchâtelois. Travail de diplôme. Université de Neuchâtel, Suisse.

Valley, B., M. Burkhard, and P.-A. Schnegg (2004). Dépliage 3-d des anticlinaux bordant le synclinal fermé de la vallée des ponts, jura central, suisse. *Eclogae Geologicae Helvetiae 97*, 279–291.

Vitel, S. (2007). *Méthodes de discretization et de changement d'échelle pour les rśervoirs fracturés 3D.* Ph. D. thesis, Institut national Polytechnique de Lorraine, France.

Vuilleumier, C., A. Borghi, P. Renard, D. Ottowitz, A. Schiller, R. Supper, and F. Cornaton (2012). 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*.

Weber, E., P.-Y. Jeannin, A. Malard, J. Vouillamoz, and F. Jordan (2012). A pragmatic simulation of karst spring discharge with semi-distributed models. advantages and limits for assessing the effect of climate change. In *Actes du 13e Congrès national de SpActes du 13e Congrès national de Spéléologie*.

Worthington, S. (2003). Characterization of the mammoth cave aqifer. In *Significance of caves in watershed management and protection in florida*.

Worthington, S., D. Ford, P. Beddows, A. Klimchouk, D. Ford, A. Palmer, and W. Dreybrodt (2000). *Porosity and permeability enhancement in unconfined carbonate aquifers as a result of solution*, pp. 527.

Worthington, S., A. Palmer, M. Palmer, and I. Sasowsky (1999). *A comprehensive strategy for understanding flow in carbonate aquifers*, Volume Special Publication 5.

Worthington, S., C. Smart, and W. Ruland (2002). Assessment of groundwater velocities to the municipal wells at walkerton. In *Proceedings of the 2002 joint annual conference of the canadian geotechnical society and the canadian chapter of the IAH*.

Zahm, C. and P. H. Hennings (2009). Complex fracture development related to stratigraphic architecture; challenges for structural deformation prediction, tensleep sandstone at the alcova anticline, wyoming. *AAPG Bulletin 93*, pp. 1427–1446.

Zhang, T., S. I. Pedersen, C. Knudby, and D. McCormick (2012). Memory–efficient categorical multi–point statistics algorithms based on compact search trees. *Math Geosci 44* (8), 863–879. DOI 10.1007/s11004-012-9412-7.

Appendix A

Source code details

```
% perform a gaussian 3D simulation
G = gaussian3D(meshDims,lx,ly,lz, mu, sigma2);

% select the voxels where the value of G is superior to a given "
    portion" quantile
selected = G > quantile(reshape(G,[],1),portion);

% compute the geobodies with a 6 node connection (i.e. connected
    by the faces)
geob = bwconncomp(selected,6);

% output matrix
M = false(ni,nj,nk);

% create N plugs
for n = 1:N
    % conduit element on which the plug will be patched
    centerId=randi([1,size(IJK_list,1)],1);
    cI=IJK_list(centerId,1); cJ=IJK_list(centerId,2); cK=IJK_list(
        centerId,3);

    % randomly chose one geobody index
    randGeob=randi([1,geob.NumObjects],1);

    % gravity center of the geobody
    geobIds = geob.PixelIdxList{randGeob};
    gravityCenter = mean([reshape(X(geobIds),[],1),reshape(Y(
        geobIds),[],1),reshape(Z(geobIds),[],1)],1);
    [gI,gJ,gK] = assignElement2Point(gravityCenter(1),
        gravityCenter(2),gravityCenter(3),meshDims);
    [geobI,geobJ,geobK] = ind2sub(size(M),geobIds);

    % search the index of the geobody that is the nearest to its
        gravity center
    D = sqrt(((((X(geobIds)-X(gI,gJ,gK)).*(X(geobIds)-X(gI,gJ,gK))
        )+(Y(geobIds)-Y(gI,gJ,gK)).*(Y(geobIds)-Y(gI,gJ,gK)))+(Z(
        geobIds)-Z(gI,gJ,gK)).*(Z(geobIds)-Z(gI,gJ,gK))));
    minDistIndex = find(D == min(min(min(D))));
    minDistIndex = minDistIndex(randi([1,length(minDistIndex)]));

    % replace gravity center with its index
    gx=X(geobIds); gy=Y(geobIds); gz=Z(geobIds);
    [gI,gJ,gK]=assignElement2Point(gx(minDistIndex),gy(
        minDistIndex),gz(minDistIndex),meshDims);

    % patch the plug on the conduit by translation
    transI=cI-gI; transJ=cJ-gJ; transK=cK-gK;

    % new indexes
    newGeobI=geobI+transI; newGeobJ=geobJ+transJ; newGeobK=geobK+
        transK;

    % place all the elements of the geobody in the model, if the
        elements are inside the grid
    for i_id=1:size(newGeobI,1)
        % if the indexes are not coerent don't keep them
        if (~(newGeobI(i_id)<1 || newGeobI(i_id)>ni || newGeobJ(
            i_id)<1 || newGeobJ(i_id)>nj || newGeobK(i_id)<1 ||
            newGeobK(i_id)>nk));
            M(newGeobI(i_id),newGeobJ(i_id),newGeobK(i_id))=true;
        end
    end
end
```

Listing A.1: Matlab code to produce gaussian plugs on the conduits, refer to the text for explanation.

Appendix B

Summary the tracer test done in the Valley of *les Ponts-de-Martel*

injection point	Date	Tracer	Dist. (m)	inj. M (kg ou Pfu)	mod. speed (m/h)	C_{max} ($\mu g/L$ ou Pfu)	Rest.(%) (%)
Voisinage	30.09.1864 [2]	SulfoG	3925	7	500 [a]	?	8
	04.05.1901 [2]	SulfoG		3	20	?	50
	26.07.1901 [2]	Uranine		5	26	?	49
	16.05.1995 [4]	Fluo.		3	59/35	3.59	96
	16.05.1995 [4]	H40		$5.47\cdot10^{14}$	59/35	1333.3	93
	16.05.1995 [4]	H4		$3.32\cdot10^{14}$	59/35	807.7	66
	28.07.1995 [1]	Fluo.		3	8.3/5.6	2.3	41.8
	20.09.1995 [1]	Fluo.		3	29/17.4	8.2	95.9
	15.09.2006 [2]	Uranine		3	38	2.09	47.65
	15.09.2006 [2]	H40		$4.95\cdot10^{14}$	38	80	8.48
	15.09.2006 [2]	Microsph.		$2.275\cdot10^{11}$	40	?	?
M. Dernier	28.07.1995 [1]	H40	3100	$5.1\cdot10^{14}$	6.2	120	3.4
	20.09.1995 [1]	H40		$5.1\cdot10^{14}$	14.1	500	30.9
R. Berthoud	09.06.1976 [2]	Uranine	2440	3	8	?	?
	28.07.1995 [1]	SulfoG		3.18	11.1	2	9.8
	20.09.1995 [1]	SulfoG		3.18	58.1	14.2	61.3
C. Varin	28.07.1995 [1]	T7	1300	$5.2\cdot10^{14}$	2.6	22	2.5
	29.09.1995 [1]	T7		$4.3\cdot10^{14}$	38.2	80	1.7
Naturalistes	28.07.1995 [1]	H6	700	$1.7\cdot10^{13}$	-	-	0.6
	20.09.1995[1]	H6		$3.3\cdot10^{13}$	14	28	22
B. Dessus	28.07.1995 [1]	H4	2375	$3.7\cdot10^{13}$	6.0	100	49.6
	28.07.1995 [1]	Duasine		2.8	5.7/4.7	7	80.3
	20.09.1995 [1]	H4		$1.7\cdot10^{14}$	37.1	410	33.6
	20.09.1995 [1]	Duasine		3	36.0	25	91.2
Rotel	15.07.1993 [3,b]	Fluo.	~3500	10	29	4	5
Coeudres	01.04.1986 [2]	Uranine	~8600	10	132	?	1
C. Sagnettes	23.06.1990 [2]	?	~4600	?	25	?	?
Thomasset	18.05.1992 [3]	Fluo.	~3000	10	11-13	?	?
La Tourne	18.05.1992 [5,c]	Napht.	~6000	50	?	?	?
	18.05.1992 [5,d]	T7		?	30	?	?
Ch.-du-Milieu	02.04.1987 [2,e]	Uranine	6930	1	45	?	?

[a] Uncertain value due to a probable contamination (Desor, 1865).
[b] Slow infiltration speed (two days to infiltrate the tracer) (Matthey, 1994).
[c] Difficult interpretation because of contamination (Matthey, 1992).
[d] no mentioning by Matthey (1992) who realized a tracer test the same day at the same place
[e] Uncertain connection (Thierrin, 1987).
[1,2,3,4] Sources: 1 : Atteia et al. (1996); 2: Gillmann (2007); 3: Matthey (1994); 4 : *Noiraigue project* database (Atteia et al., 1996; Gogniat, 1995a); 5: Matthey (1992).

Table B.1: Synthesis of the tracer test realized until nowadays (2012) with positive result at the spring of *Noiraigue*. Dist. : distance between the injection point and the spring of *Noiraigue*; inj. M. : Injected mass; mod. speed. : transit time of the peak; C_{max} : maximal concentration; Pfu : phages per ml.; Rest. : restitution percentage; SulfoG : sulforhodamin G; Fluo. : Fluorescin; Napht. : naphtionate; Microsph. : fluorescent microspheres ; H40, T7, H6 and H4 : bacteriophages. Modified from Fournier (2012)

Appendix C

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)[1]

[1]This paper is published as Vuilleumier et al. (2012)

Published in Hydrogeology Journal, 1-16, 2013
which should be used for any reference to this work

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)

C. Vuilleumier · A. Borghi · P. Renard · D. Ottowitz · A. Schiller · R. Supper · F. Cornaton

Abstract The eastern coast of the Yucatan Peninsula, Mexico, contains one of the most developed karst systems in the world. This natural wonder is undergoing increasing pollution threat due to rapid economic development in the region of Tulum, together with a lack of wastewater treatment facilities. A preliminary numerical model has been developed to assess the vulnerability of the resource. Maps of explored caves have been completed using data from two airborne geophysical campaigns. These electromagnetic measurements allow for the mapping of unexplored karstic conduits. The completion of the network map is achieved through a stochastic pseudo-genetic karst simulator, previously developed but adapted as part of this study to account for the geophysical data. Together with the cave mapping by speleologists, the simulated networks are integrated into the finite-element flow-model mesh as pipe networks where turbulent flow is modeled. The calibration of the karstic network parameters (density, radius of the conduits) is conducted through a comparison with measured piezometric levels. Although the proposed model shows great uncertainty, it reproduces realistically the heterogeneous flow of the aquifer. Simulated velocities in conduits are greater than $1 \mathrm{cm\,s^{-1}}$, suggesting that the reinjection of Tulum wastewater constitutes a pollution risk for the nearby ecosystems.

Keywords Karst · Numerical modeling · Coastal aquifers · Geophysical methods · Mexico

C. Vuilleumier (✉) · A. Borghi · P. Renard
Centre of Hydrogeology and Geothermics,
University of Neuchâtel,
11 Rue Emile Argand, 2000 Neuchâtel, Switzerland
e-mail: cecile.vuilleumier@unine.ch

D. Ottowitz · A. Schiller · R. Supper
Department of Geophysics,
Geological Survey of Austria,
Neulinggasse 38, 1030 Vienna, Austria

F. Cornaton
Instituto Tecnológico de Monterrey,
Ave. Eugenio Garza Sada 2501 Sur Col. Tecnológico, 64849
Monterrey, Nuevo León, Mexico

Present Address:
F. Cornaton
DHI-WASY GmbH,
WaltersdorferStraße 105, 12526 Berlin, Bohnsdorf, Germany

Introduction

The Caribbean coast of the Yucatan Peninsula, Mexico (Fig. 1), is known to contain an important karst system, which includes several of the longest underwater caves in the world (Gulden and Coke 2011). It has been explored and mapped by cave divers for decades (Fig. 2a). The aquifer consists of a freshwater lens of thicknesses between <10 and 100 m on the top of a saline-water intrusion, which reaches several tens of kilometers inland (Bauer-Gottwein et al. 2011). The relief is low and surface runoff is absent from the peninsula (Beddows 2004). The karst system appears to be very well developed: the deepest explored cave lies at 119 m depth (Smart et al. 2006) and the largest conduit diameters reach some 70 m.

Because of the limited resource of the freshwater lens and the growing demand, water supply in the Yucatan Peninsula is becoming problematic (Bauer-Gottwein et al. 2011), especially in the area of Tulum (Quintana Roo State) where environmental concerns are growing because of the planned urban development that includes the building of a hotel complex (Supper et al. 2009), an airport and a highway (SCT 2010). Moreover, it is common practice in the Yucatan Peninsula to reinject wastewater into the aquifer without previous treatment (Marin et al. 2000). This pollution risk is a major threat for the nearby Sian Ka'an Biosphere Reserve and the large coral reefs located near the shore (Fig. 1b), which are host to ecosystems that are highly dependent on the karst aquifer (Gondwe 2010). To assess the vulnerability of the aquifer, a finite-element flow model has been built (Fig. 2a). As the network geometry is a major constraint on flow paths and velocities in karstic aquifers (Worthington and Smart 2003), it has to be incorporated in the model, which is done by using a finite-element

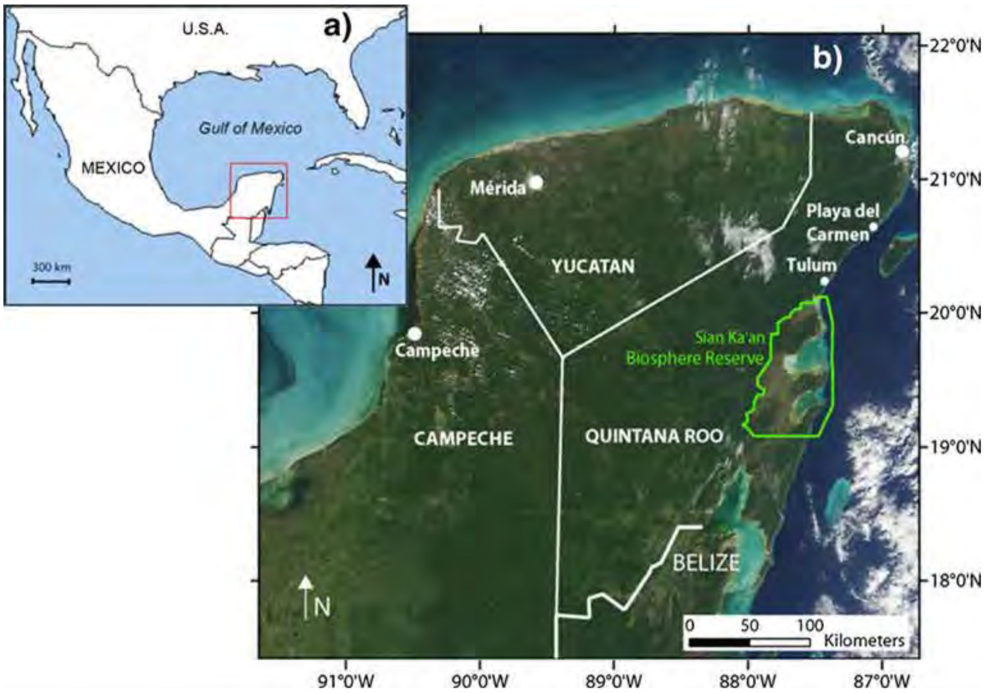


Fig. 1 **a** Location of the study area (modified after Google Earth 2011). **b** Geographical situation of the Yucatan Peninsula. Modified after NASA (2010)

method accounting both for the matrix and the conduits. The conduits are included as one-dimensional (1D)-pipes where flow is modeled by the Manning-Strickler formula allowing for representative turbulent flow. Attempts to calibrate this model are made using high-resolution global positioning system (GPS) water-level measurements.

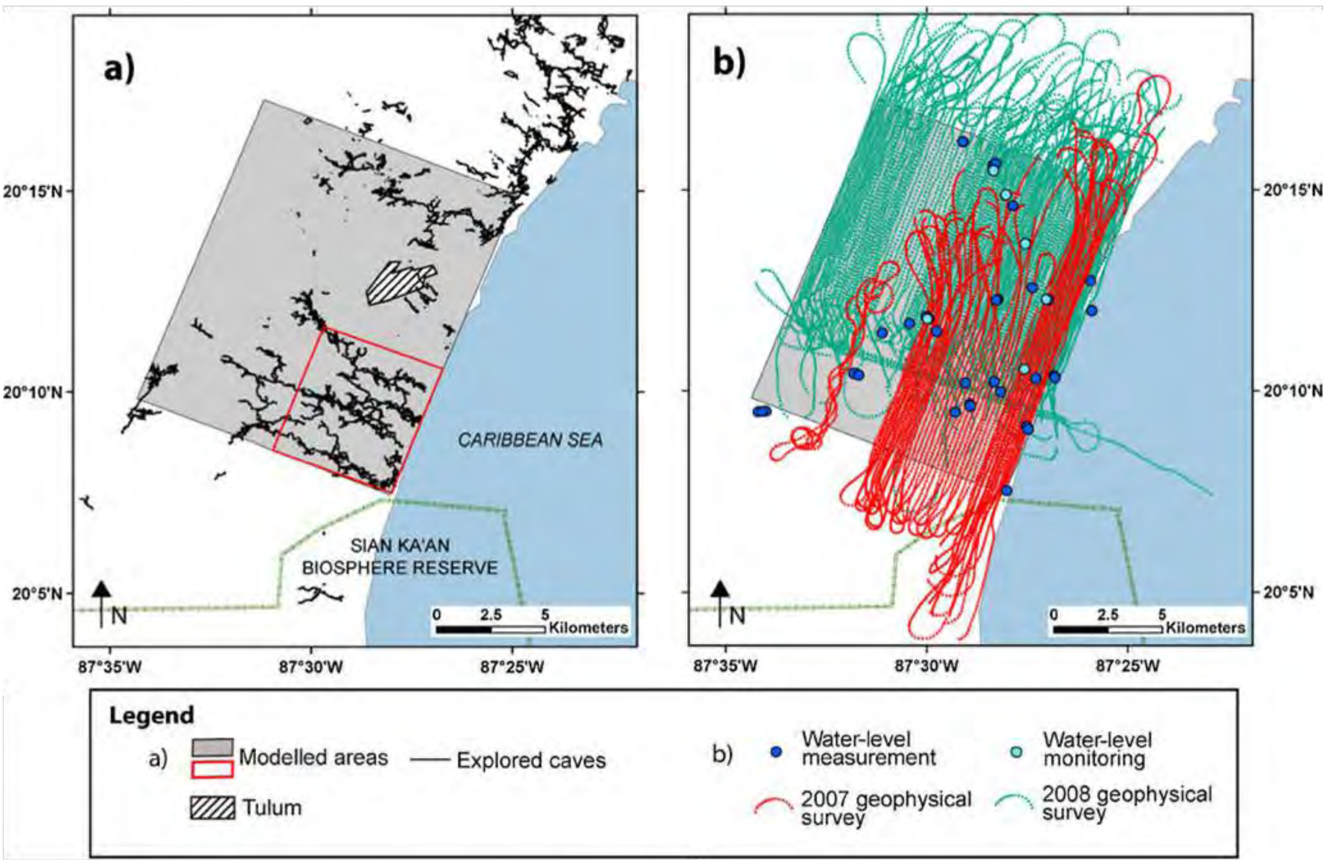


Fig. 2 Location of the **a** modeled area and **b** available data

However, the main focus of this paper is not on the modeling of this specific site, but rather on the extension of a new pseudo-genetic method to model the geometry of the karstic system. Indeed, even if the karst conduits have been extensively mapped in this region of the world, they are not known exhaustively. The same situation occurs in most karstic systems in the world, which are only partially explored. For modeling groundwater flow and transport, it is therefore necessary to construct a realistic model of the unexplored parts, which is why, in recent years, several techniques have been developed to build stochastic karstic network models (Collon-Drouaillet et al. 2012; Fournillon et al. 2010; Henrion et al. 2008; Jaquet et al. 2004; Pardo-Igúzquiza et al. 2012). Here, the method of Borghi et al. (2012) is extended to account for geophysical data when available. The application of this new method is illustrated using airborne electromagnetic measurements collected by the Geological Survey of Austria (Supper et al. 2009), revealing the presence of underwater karstic conduits thanks to the strong electrical conductivity contrast between water-filled caves and the limestone matrix. Several equiprobable karstic network geometries are constructed with this method. It is then proposed to consider that the radius of the conduits follows a power law relating the radius with the order of the conduit. This assumption is based on an analogy with Horton's law for rivers and provides a simple model relating the flow properties with the geometry. The parameters involved in this formulation are the targets for the calibration of the flow model.

Overall, the aim of this paper is to present this new methodology and to test if it is applicable on a real data set. The resulting model is considered as a preliminary step allowing a better understanding of how such systems could be modeled in the future.

Description of the study site

The Yucatan Peninsula aquifer

The Yucatan Peninsula is a 300-km-wide carbonated platform located between the Gulf of Mexico and the Caribbean Sea (Fig. 1). The overall topography is relatively flat with maximum elevations reaching 300 m above sea level (asl). The mean annual temperature is 26 °C and annual precipitation is between 1,000 and 1,400 mm y^{-1}, the major part falling during the wet season from May to October (Héraud-Piña 1996). Surface runoff is absent (Beddows 2004) and, according to Gondwe et al. (2010), in the southeastern peninsula, 17 % of the precipitation recharges the aquifer, while the remaining undergoes evapotranspiration.

The aquifer is densely stratified: its upper part consists of freshwater flowing toward the sea, while the lower part consists of warmer saline water (Beddows et al. 2007). General groundwater circulation is organized in a concentric flow from the center of the peninsula towards the coast. The regional hydraulic gradient is relatively low with estimates lying between 1 and 10 $cm\,km^{-1}$ for coastal plains (Bauer-Gottwein et al. 2011).

The groundwater hydrodynamics is characteristic of strongly karstified aquifers. According to the review of Bauer-Gottwein et al. (2011), estimations of the hydraulic conductivity in the Yucatan Peninsula aquifer vary widely depending on the scale of interest. Testing of core samples gives values of 10^{-6}–$5{\cdot}10^{-2}\,ms^{-1}$, while calibration of flow models (at a scale of hundreds of kilometers) yields effective hydraulic conductivity values in the range of 10^{-1}–$10^{2}\,ms^{-1}$. Those very high equivalent conductivities are related to the presence of large karst conduits that were not described explicitly in the calibrated model. As for groundwater velocity, Moore et al. (1992) measured values in fractures increasing coastward from 1 to 12 $cm\,s^{-1}$, while in the limestone matrix, they obtain values of approximately $10^{-2}\,cm\,s^{-1}$. Beddows (2004) provides measurements for two conduits that were monitored for several months. These data reveal an important and rapid effect of sea level variations on the conduit flow velocity. Most of the recorded values are in the range of a few centimeters per second, with a maximum of ~20 $cm\,s^{-1}$ at a coastal site. Cave divers report that the flow is sometimes too strong to swim against, which suggests that velocities can reach several tens of centimeters per second.

The freshwater lens constituting the Yucatan Peninsula aquifer is relatively thin: its maximum thickness is approximately 100 m (Bauer-Gottwein et al. 2011). The depth of the interface between the freshwater lens and the saline water increases inland following a linear trend rather than the Dupuit-Ghyben-Herzberg model, possibly because of the influence of the karstic network (Beddows 2004). In southern Quintana Roo, the freshwater lens discharges to the sea at a rate of 0.27–0.73 $m^3 s^{-1}$ per km of coastline, depending on estimations (Bauer-Gottwein et al. 2011).

Setting of the karst network

The Yucatan Peninsula consists of Mesozoic and Cenozoic sediments on top of a Paleozoic basement. According to seven drillings realized in the northern peninsula at depths between 1.5 and 3.5 km, there is no low-solubility geological formation that could constrain the formation of karst at large scale (Ward et al. 1995). The upper hundreds of meters are sub-horizontally oriented limestones and dolomites (SGM 2007) that show a high solubility (Marin et al. 2000). These carbonates are indeed highly karstified, containing one of the most developed cave system in the world. Karstification may have commenced as early as the late Eocene, when the peninsula emerged (Iturralde-Vinent and MacPhee 1999).

Smart et al. (2006) presented a study of conduit geometry and distribution on the Caribbean coast of the peninsula, as well as assumptions about the processes ruling cave development in this area. According to them, the Yucatan cave system represents an intermediate type between telogenetic and eogenetic karst, following the classification proposed by Vacher and Mylroie (2002). The first type includes typical continental karst, where speleogenesis proceeds under the influence of water flowing from infiltration points toward base level. Cave development is mainly influenced by recharge through secondary porosity, which creates preferential flow paths. In contrast, eogenetic karst development, also known as the flank-margin model, is typical of small carbonated islands. In this case, speleogenesis occurs in diagenetically immature sediments presenting a high primary porosity. Mixing corrosion at the interface between fresh and saline water is the major process ruling carbonate dissolution. The resulting karst system consists of isolated chambers developing along the coast.

Following the observations of Smart et al. (2006), mixing corrosion has a major influence on the development of the Quintana Roo caves. Their study of conduit depths shows a correlation with the halocline position. Many of the conduit cross sections are enlarged at the freshwater/saline-water interface, which is in some locations very sharp (Beddows et al. 2007). Caves located above or below the interface often show speleothems or recrystallization features that suggest that dissolution is not active there anymore. Moreover, the authors were able to link these paleo-karst horizons to previous sea level low or high stands. However, the flank-margin model cannot explain on its own the wide extent of the Yucatan karst system. Caves are organized into large anastomotic systems discharging to the sea, extending up to 8–12 km inland. On the one hand, this morphology suggests the action of an important discharge toward the sea in the process of cave development; and, on the other hand, the lack of preferential orientation in conduit direction suggests their development in a high-porosity matrix rather than in a heterogeneous fractured matrix.

Karst network modeling

Method

The modeling of the karst network geometry is based on the stochastic pseudo-genetic method proposed by Borghi et al. (2012). The method proceeds in four main steps: (1) modeling the three-dimensional (3D) geology of the site (folds, thrusts, faults, etc.) to define the location of the karstifiable formations; (2) modeling the internal heterogeneity of the karstifiable formations (fractures, bedding planes, etc.); (3) map or model (with a stochastic point process) the locations of inlets such as sinkholes or dolines and outlets such as springs; (4) use a fast marching algorithm (Sethian 1996) to compute the fastest path between the inlets and outlets. This procedure is repeated iteratively to generate a hierarchical network. One has just to update the heterogeneity described in step 2 with the result of step 4 and repeat steps 3 and 4. Other variations and improvements are proposed in Borghi et al. (2012) to, for example, account for the unsaturated zone.

At the heart of the method, the fast marching algorithm allows a very efficient computation of the paths between inlets and outlets. This algorithm requires as input a 3D map of maximum velocities. The velocities represent a characteristic of the medium and are used as a parameter to control the contrast between the different geological heterogeneities and their ability to get karstified. It is thus a representation of potential locations for preferential speleogenesis. For typical continental karst systems, this field is built by means of a geological model with higher velocities assigned to soluble formations, faults and/or fractures. Other input parameters are the inlet and outlet points of the network. They can be positioned either deterministically or stochastically. If the precise location of springs is unknown, it is possible to select a diffuse spring zone. In this case, the algorithm selects the most probable outlet points in the given area. The last main parameter to set is the number of iterations and the number of generated conduits per iteration.

For the Yucatan coastal system that was described in the previous section, the caves develop mainly toward the sea in a sub-horizontal plane. Thus, there is no need to build a complex 3D geological model; furthermore, there is no evidence for a clear control by small-scale fracturation or bedding. Steps 1 and 2 of Borghi et al. (2012) must, therefore, be adapted to account for the specificity of the site. In addition, two sources of data are available at that site and must be considered. First, the area has been extensively explored by two airborne geophysical campaigns. Supper et al. (2009) and Ottowitz (2009) already mentioned the ability of these electromagnetic measurements to reveal karstic conduits. This source of information needs to be used to control the simulation of the karstic conduits. Here, the proposal is to directly rescale the geophysical anomaly and use it as one component of the input heterogeneity field resulting from step 2 of the algorithm. Second, extensive maps of conduits are available and they should constrain the karst network model; again, that information will be added during step 2 of the algorithm. The following sections present how the pseudo-genetic karst simulator algorithm has been modified and how these new data have been processed to obtain an ensemble of realistic network models.

Processing the geophysical data

Method and data

The airborne geophysical measurements were carried out by the Geological Survey of Austria during two campaigns in 2007 and 2008. They cover an area of approximately 140 km^2 around the town of Tulum and are distributed along flight paths oriented N22 with a spacing of 20–100 m (Fig. 2b). They were obtained by means of an active frequency-domain electromagnetic method. In principle, the generated primary field induces eddy currents in the subsurface, which themselves create a secondary magnetic field. The amplitude and phase of the secondary field depends on the electrical resistivity of the subsurface. Thanks to a strong contrast in resistivity between water-filled conduits and the surrounding limestone matrix, anomalous responses are expected where conduits are located. This method is thought to be particularly efficient in the Tulum area because of a very flat topography, a thin soil cover and the relatively shallow position of the main karst features (Supper et al. 2009).

The main part of the measurement system consists of a modified GEOTECH-"Bird" of 5.6 m length and 140 kg weight (Motschka 2001). It is towed on a cable 30 m below the helicopter, and transmitter coils inside the probe generate primary electromagnetic fields of four frequencies (340, 3,200, 7,190 and 28,850 Hz). The resultant secondary field is recorded by the corresponding receiver coils. For every frequency there are two measurement values, the in-phase (no phase shift between primary and secondary field) and the out-phase component (90° phase shift). Results are given in parts per million (ppm), which is the ratio of the secondary field amplitude over the primary field amplitude. The data from the two field campaigns that give the best insight on the karstic conduits are the in-phase (north) and out-phase (south) component of the measured signal for a primary alternating field of the third frequency (7,190 Hz), which were chosen on the basis of a visual analysis. Indeed, a comparison with the map of explored conduits (Fig. 3) shows a potentially good agreement between conduits and anomalous electromagnetic responses: positive anomalies are measured around known conduits. The electromagnetic map reveals as well potential unexplored conduits.

The variation of the signal amplitude is induced by several factors. It is influenced by groundwater salinity, which varies across the study area. Thus, a global rise in conductivity toward the coast is observed which is linked to the shallower depth of the halocline. Other identified sources of noises are a circular lagoon west of Tulum, the town of Tulum itself, and some shift in the values between the flight paths (Fig. 3). In addition, the sharpness of conduit-induced anomaly highly depends on the conduit size, depth and on the resistivity contrast between the water filling the conduit and the surrounding matrix. For these reasons, available inversion does not work properly to map the small anomalies of caves since they are mostly 1D inversion algorithms (assuming a horizontally layered Earth) and cave structures show at least a 2D structure (if not 3D). It was therefore decided to use the geophysical signal after some processing to guide the simulation of the karstic network in a statistical manner. The details of these steps are described in the following paragraphs.

Processing the data

The initial data consists of the value of the electromagnetic signal at each measurement point of the 2007 and 2008 field campaigns. The first steps were to apply (1) an altitude correction (method from Huang (2008)) and (2) a median filter. The correction reduces the influence of altitude variation of the bird, while the median filter, on the other hand, enhances conduit-induced anomalies. Figure 4 shows the histograms of the base 10 logarithm of the resulting values of the filtered signal. One can see on Fig. 4 that the distributions of the signal are significantly different for the two surveys and display a significant proportion of extreme values making these two distributions clearly different and not Gaussian. The extreme values are not correlated to the presence of conduits and tend to mask the information contained in the rest of the data (Fig. 5a). These points were therefore

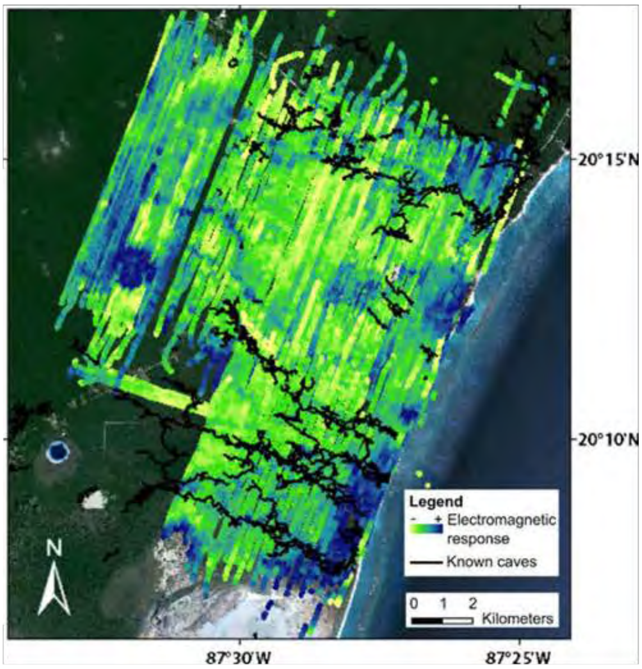


Fig. 3 Measured amplitude (after altitude correction) of the field induced by a frequency of 7,190 Hz. In the northern area (2008 survey, see Fig. 2b), the measured out-phase component is shown and the southern area (2007 survey) it is the in-phase component. *Color scale* is the equalized histogram, ranging from 2 to 20 ppm for the northern area and from 14 to 57 ppm for the southern

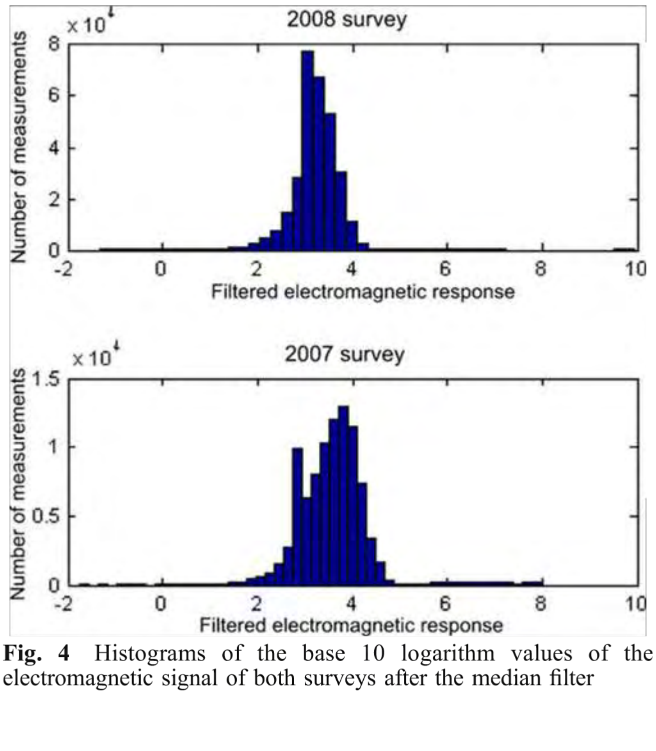


Fig. 4 Histograms of the base 10 logarithm values of the electromagnetic signal of both surveys after the median filter

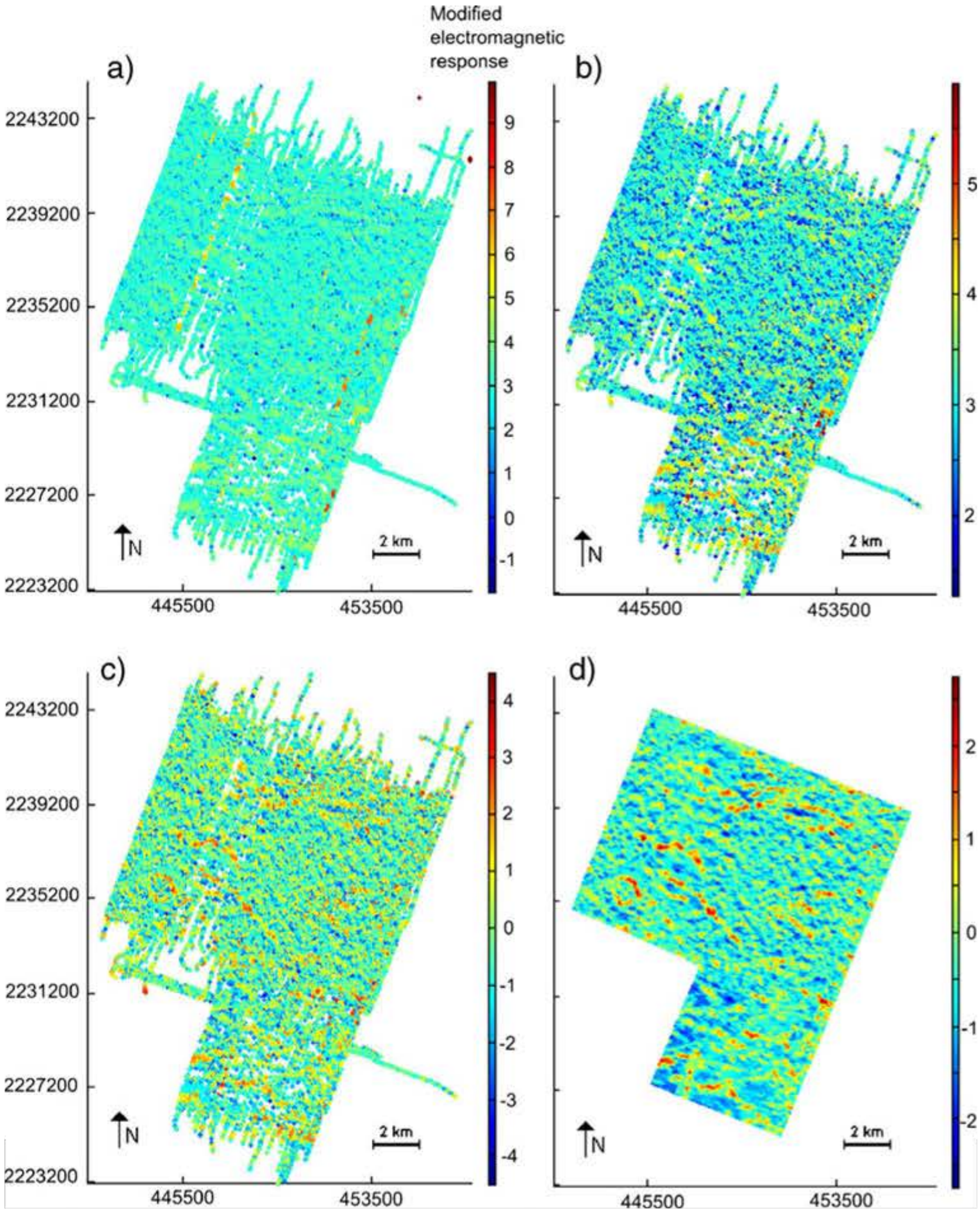


Fig. 5 **a** Logarithm of the magnitude of the out-phase (north) and in-phase (south) component of 7,190 Hz after altitude correction and median filter; **b** after removal of extreme values; **c** after normal score transform on both data fields; **d** after kriging and moving average. Coordinate system is UTM Zone 16 N, WGS84

considered as outliers; by selecting the values ranging beyond the interval of three standard deviations from the mean value they are removed from the data sets (Fig. 5b). To combine the 2007 and 2008 data sets into a single one, a normal score transform—see e.g. Chilès and Delfiner (1999)—has been applied separately on the two distributions after eliminating the extreme values. The resulting data set is depicted in Fig. 5c. The next step was carried out using the Isatis software (Geovariances 2010). The resulting variable has been interpolated using kriging on a regular grid of 20 × 20-m cells. The origin of the grid (southern vertex) is positioned at (451, 100 m; 2,225,280 m) in the UTM Zone 16 N coordinate system (datum WGS84). Its dimensions are 11.8 × 14.8 km and its orientation is N22 (Fig. 2). First, the experimental variogram of the processed variable has been computed for twenty distance lags of 20 m each. A variogram model has been fitted to the experimental one; it is an exponential model with a range of 70 m and a sill of 0.95. Some residual flight path noise still appears on the interpolated field. To reduce it, a moving average filter with a window of 100 × 20 m has been applied—the longer axis of the ellipsoid being oriented perpendicular to the flight paths (i.e. N68). The explored cave data were been taken into account for the interpolation. The final map of normalized electromagnetic anomalies is shown in Fig. 5d.

Completing the electromagnetic map

A final step in processing the geophysical data is to complete the southwestern part of the model domain.

Indeed, this zone has not yet been surveyed by geophysical measurements and maps of the geophysical anomalies should be provided over the complete domain. For that purpose, an ensemble of equiprobable maps of geophysical anomalies have been simulated in the region using a geostatistical technique. To guide the geostatistical simulation in this area, available cave maps have been used (Fig. 2). Furthermore, the anomalous signal caused by a lagoon northwest of Tulum has also been replaced by simulated values.

The simulation of the geostatistical anomaly in the parts where the signal is missing has been achieved using the direct sampling multiple point algorithm (Mariethoz et al. 2010). The reconstruction method is presented in detail in Mariethoz and Renard (2010). The simulation grid is the same as defined in the preceding. The available cave map allows the realization of a bivariate simulation, variable 1 being the geophysical anomaly and variable 2 the presence of conduits (categorical variable, 1=conduit is present; 0=conduit is absent). By doing so, the simulation of the geophysical signal accounts for the presence of known caves. The area of the Ox Bel Ha system was chosen as a training image, for it has been extensively surveyed by cave divers and by the 2007 geophysical campaign. It was assumed that, in this zone, all the caves were known, i.e. the conduits are considered as absent where no information is provided. This is a reasonable assumption at the resolution of the model since the density of the mapped caves is very high in this small area. The conditioning data consists of, for variable 1, the whole geophysical data (Fig. 6a). Variable 2 was built with the cave map: it is equal to 1 where a known cave is located. The rest of the map was set as being not informed (Fig. 6b). The search radius has been set to 50 m and the distance threshold to 0.01. Twice as much weight has been given to variable 1 as to variable 2 for the distance calculation. One realization is illustrated in Fig. 6c, and the mean simulated signal for 100 equiprobable realizations is shown in Fig. 6d. A positive anomaly is calculated around known conduits, while the remainder of the map shows strong variability, reflecting the uncertainty on the data in this area.

Simulating the karst

Building the heterogeneity field

Two pieces of information have been used to create the velocity field describing the geological heterogeneity for the karst simulator: the processed electromagnetic anomaly map $E_{\mathrm{m}}(\mathbf{x})$ and the map of explored caves $I_{\mathrm{c}}(\mathbf{x})$.

For the electromagnetic map, the first step was to add the minimum value taken by $E_{\mathrm{m}}(\mathbf{x})$ to the map:

$$N_{\mathrm{m}}(\mathbf{x}) = E_{\mathrm{m}}(\mathbf{x}) + \min\{E_{\mathrm{m}}(\mathbf{x})\}. \tag{1}$$

A first trial of karst simulation on a small portion of the simulation area using $N_{\mathrm{m}}(\mathbf{x})$ as a velocity field is shown in

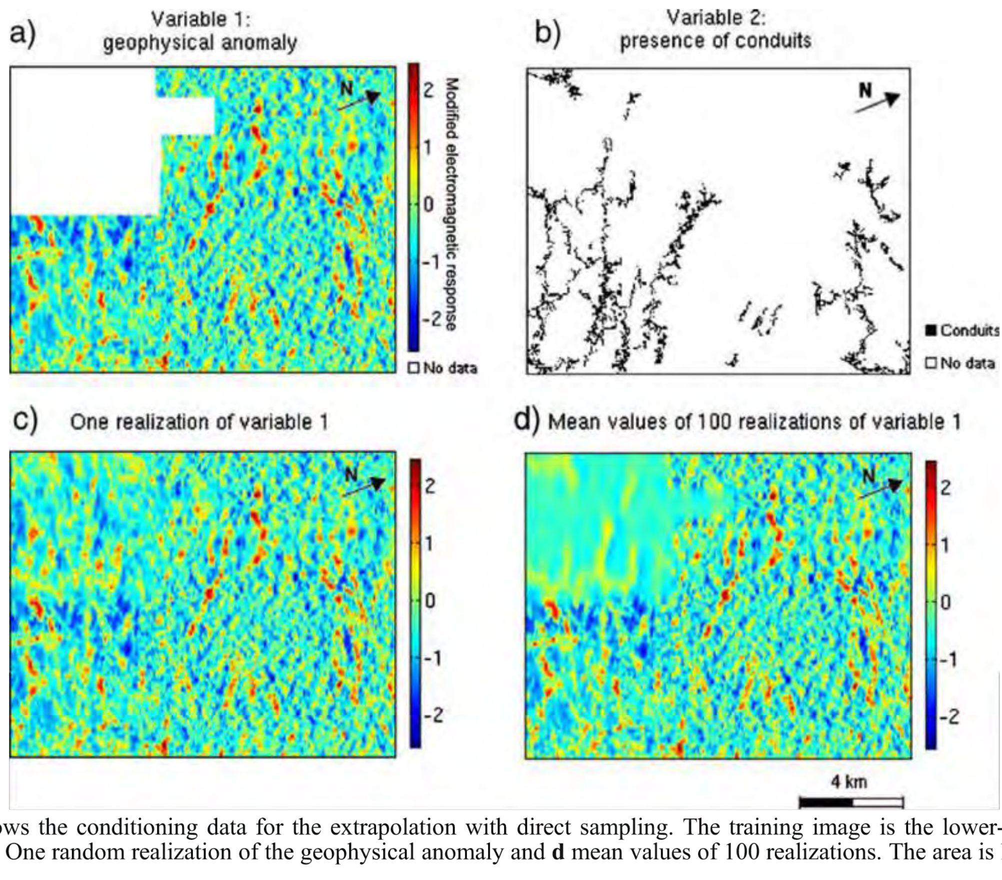


Fig. 6 **a- b** Shows the conditioning data for the extrapolation with direct sampling. The training image is the lower-left quarter of the simulation grid. **c** One random realization of the geophysical anomaly and **d** mean values of 100 realizations. The area is located in the grey box in Fig. 2a

Fig. 7a. A diffuse spring zone has been set at the lower border of the model. Conduit starting points were picked randomly on the map. It appears that the velocity contrast has to be strengthened for a proper control of the electromagnetic map on the karst simulation. This has been realized using a power law relationship between the velocity and the electromagnetic anomaly.

$$V(\mathbf{x}) = b^{N_{\mathrm{m}}(\mathbf{x})} \tag{2}$$

After some trial and errors, a value $b=3$ was chosen as it provides a good constraint to conduit generation (Fig. 7b).

For the explored caves, the map $I_{\mathrm{c}}(\mathbf{x})$ is a Boolean indicator variable constructed by the projection of the cave maps on the simulation grid. It is equal to 1 where caves are present and 0 where no information is provided:

$$I_{\mathrm{c}}(\mathbf{x}) = \begin{cases} 1 \text{ if a cave is present at location } \mathbf{x} \\ 0 \text{ otherwise} \end{cases}. \tag{3}$$

A constant high velocity V_{k} has been assigned to points where $I_{\mathrm{c}}(\mathbf{x})=1$. Finally, the velocity field representing the heterogeneity of the medium is described as:

$$V(\mathbf{x}) = I_{\mathrm{c}}(\mathbf{x}) \cdot V_{\mathrm{k}} + [1 - I_{\mathrm{c}}(\mathbf{x})] \cdot 3^{N_{\mathrm{m}}(\mathbf{x})} \tag{4}$$

Additional settings of the karst simulator

The next step is the determination of in- and outlet points of the simulated karst system. Little information about the conduit inlets was available, as the numerous cenotes present on the surface are not consistently linked with the underground network (Beddows 2004). They were, thus, picked randomly for each conduit simulation in high velocity zones—i.e. where the presence of a conduit is suggested by cave exploration or a geophysical anomaly. The chosen threshold value for the inlet points is the third quartile of the velocity values, so that a quarter of the grid points can be potentially picked as a conduit inlet. An additional constrain was applied to determine the location of the starting points at the first iteration: they must be located on the upper border of the model. This is done in order to ease flow simulation, as a water inflow is located at this specific border (see section Hydrogeological modeling).

As for the outlets of the system, it has been established in section Description of the study site that the aquifer discharges to the sea. Without more precise information on the location of submarine springs, the option of a diffuse spring zone located on the coast was chosen. Two attempts for the configuration of the spring zone were made: (1) all the coast is selected as a spring zone; (2) alternation between both halves of the coast at each iteration. This latest option is expected to increase reconnections between the conduits, in order to reproduce the anastomotic patterns that are observed.

Figure 8 illustrates the map of conduit occurrence probability after 1,000 realizations for both spring zone options. The simulation grid has the same extent as previously defined, but it is discretized in 40 × 40-m squares—note that this resolution is used to represent the overall field and will be used later to solve the flow in the matrix, but the conduits are represented by discrete 1D elements having a radius that can be much smaller than the cell size. Three iterations were computed, with respectively 20, 100 and 1,000 generated conduits per iteration. It appears that the proposed velocity field is a good constrain to the simulation, as some conduits have an occurrence probability of over 0.8. As expected, network branches are more connected with an alternating spring zone (Fig. 8b). Also, conduit orientations show much variability. With a uniform outlet area, conduits are mainly oriented straight coastward and are organized in dendritic patterns (Fig. 8a). Conduit distribution simulated with the second option is, thus, in better agreement with field observations. Furthermore, it can be observed that conduit paths have a wider variability with option 2 (Fig. 8b). With option 1, the most

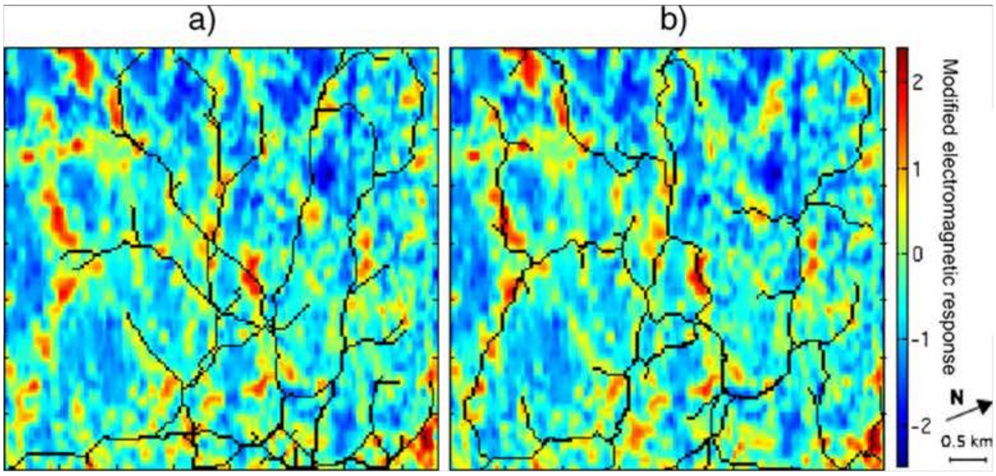


Fig. 7 Two karst simulation attempts on a small portion of the simulation area. In **a**, the electromagnetic component of the velocity medium is the electromagnetic map with positively shifted values (spatial variable e). In **b**, the electromagnetic control on the karst simulation is enhanced by setting its value to 3e. This area is located in the red box in Fig. 2a

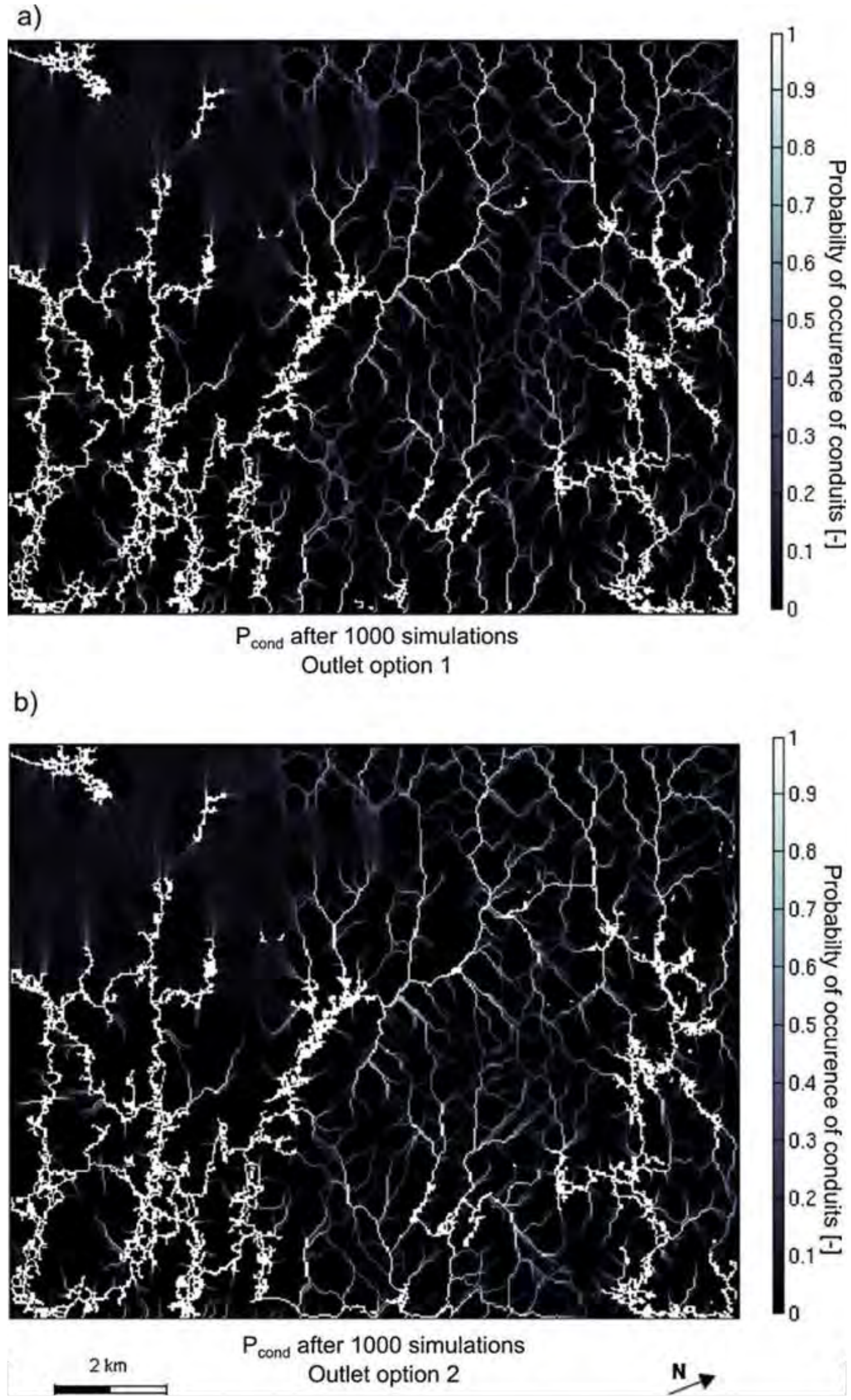


Fig. 8 Map of the probability of occurrence of conduits (P_{cond}) computed on the basis of 1,000 realizations. In **a**, a diffuse spring zone is defined along the coast (*lower border* of the simulation grid). In **b**, one half of the coast line is defined as a diffusive spring zone for the first iteration. The other half is considered for the next iteration, and so on. The area is located in the grey box in Fig. 2a

probable conduits (probability of occurrence of conduits $P_{\mathrm{cond}}>0.8$) are mainly located on the same paths (Fig. 8a). It is another argument to prefer option 2, as it reflects the high uncertainty arising from this cave mapping method. The following simulations presented in this work were thus realized with an alternating spring zone.

The remaining input parameters to set are the number of iterations and the number of conduits generated at each iteration. The hierarchy of the conduits directly arises from the number of computed iterations: when a conduit is simulated on the same path as a conduit resulting from a previous iteration, its order increases by one unit. This is illustrated in Fig. 9. The maximum order is, thus, the total

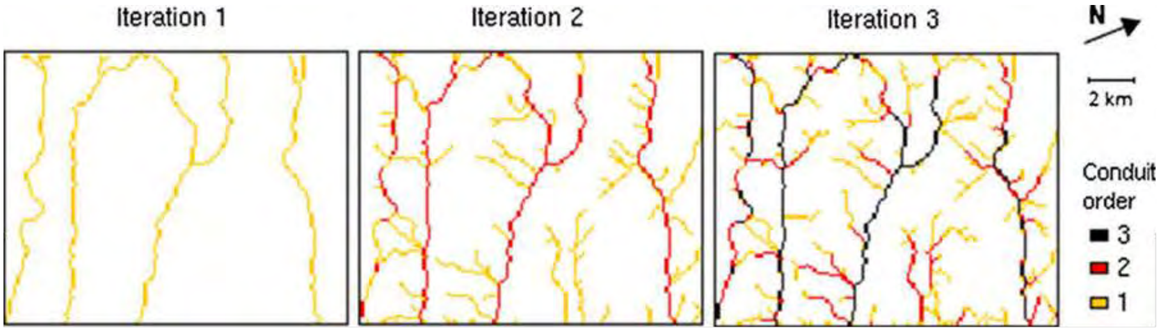


Fig. 9 Simulation of a three-iteration network. Each time a conduit is simulated on the same path as a conduit generated in a previous iteration, its order increases by one unit. The area is located in the grey box in Fig. 2a

number of iterations. As a matter of simplification, the network has been organized in three orders; therefore, three iterations were computed. The known network was then included in the model. If no conduit was simulated on a known conduit, its order is 1; otherwise, the ordering obtained by the iterative process of simulation is kept. The number of generated conduits determines the overall density of the network. As this parameter is unknown, three network densities were generated and tested in the flow model (see section Hydrogeological modeling). Their parameters are described in Table 1 and they are illustrated in Fig. 10.

Hydrogeological modeling

Building the flow model

The hydrogeological modeling was achieved using the finite-element code Ground Water (Cornaton 2007). For the purpose of this preliminary study, the model was restricted to the freshwater lens, which is represented as a two-dimensional (2D) confined aquifer of constant thickness and in steady state. This assumption is justified by the fact that all the simulated conduits are fully saturated. Because only the steady-state long-term behavior of the system is modeled, the variation in saturated thickness in the carbonate matrix is expected to have only a minor effect on the values of the fluxes within the conduits, which is why it is neglected at that stage of the work. The model is discretized in 40 × 40 m squares (4-node elements) and the simulated karstic network is integrated in the mesh as a set of 1D elements (2-node elements)(Fig. 11).

The groundwater flow in matrix elements is represented using Darcy's law, while groundwater flow in the 1D conduit elements is represented using the non-linear Manning-Strickler formula for turbulent flow in pipes:

$$\mathbf{v} = -\Phi_s \cdot \sqrt{[3]}R^2 \frac{\nabla h}{\sqrt{\|\nabla h\|}} \qquad (5)$$

where $\mathbf{v}$ is velocity, Φ_s is the friction coefficient $[L^{1/3}T^{-1}]$ characteristic of the pipe, ∇h is the hydraulic gradient and R [L] the hydraulic radius of the pipe. A homogeneous value of Φ_s=50 $m^{1/3}s^{-1}$ was chosen for the whole network. This value is proposed by Lauterjung and Schmidt (1989) for irregular concrete surfaces. In the model, the pipes are assumed to be saturated, so the hydraulic radius is equal to the physical radius of the conduit. The identification of this parameter is described in the following paragraphs.

A constant-head boundary condition is applied at the downstream border of the model in order to represent the sea boundary. Its value has thus been set to 0 m. The upstream border is defined as an inflow boundary. Direct recharge by rainfall over the site is considered negligible as compared to the upstream flow; it was therefore assumed that the observed coastal outflow is directly flowing from the model upstream limit. In their review, Bauer-Gottwein et al. (2011) state that coastal outflow estimates for southern Quintana Roo range from 0.27 to 0.73 m^3s^{-1} per kilometer of coastline. After comparing these values by means of numerical flow models, they suggest an outflow of 0.3–0.4 m^3s^{-1} km^{-1}. The upstream influx was thus set to 0.35 m^3s^{-1} km^{-1}. A row of very high conductivity elements (10^5ms^{-1}) was added at the upstream border, so that the inflow is distributed between the matrix and the conduits (Fig. 11).

The parameters that are expected to have a significant influence on the flow simulation are: (1) the matrix hydraulic conductivity, (2) the karstic network density and (3) karstic conduit diameters. In order to calibrate the model, output flow fields were compared with 43 GPS groundwater level measurement, which were realized in cenotes and boreholes over the modeled area during February 2011 (Fig. 2b). Six of those points were monitored during the 10 previous months with pressure sensors at a rate of one measurement every 30 min. Monitoring data reveal a tidal influence on the piezometric

Table 1 Input parameters of the three simulated networks

Number of generated conduits	Iteration 1	Iteration 2	Iteration 3
Network 1	20	100	1000
Network 2	20	100	1500
Network 3	20	100	2000

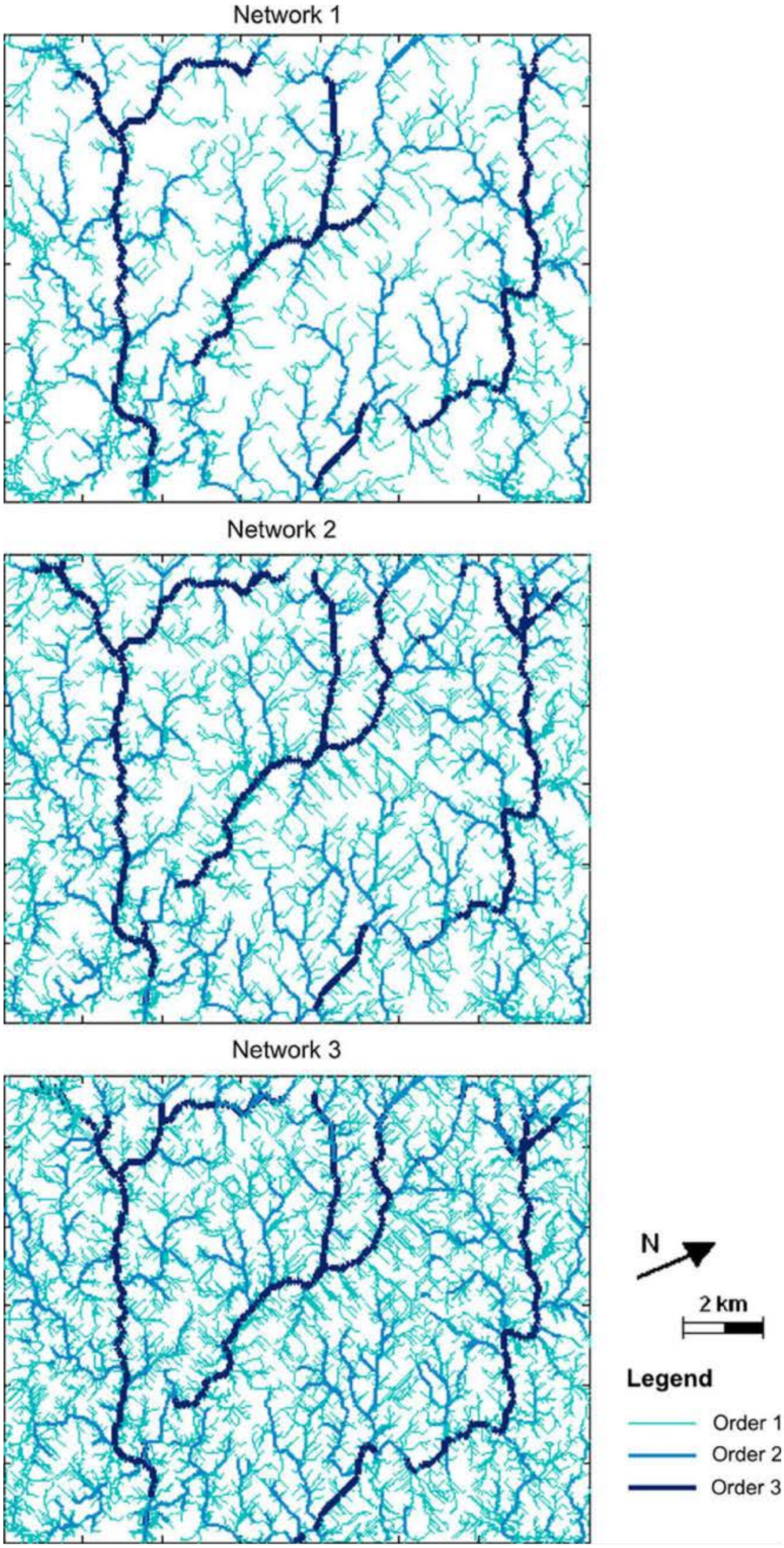


Fig. 10 Three simulated networks that were used in the flow model. Conduit orders are assigned according to the number of times a conduit is generated on the same path. Input parameters of these network realizations are described in Table 1. The area is located in the grey box in Fig. 2a

level in the range of 10 cm, whereas the maximum observed head is 37 cm with respect to the averaged measured sea level (six measurements). On the other hand, the estimated measurement uncertainty is approximately

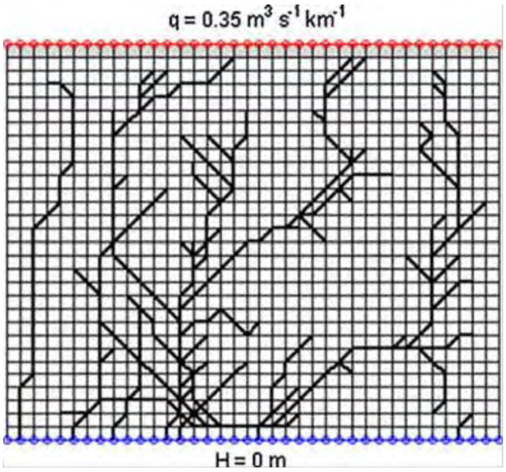


Fig. 11 Simplified scheme representing the finite-element flow model with boundary conditions: *red nodes* represent an inflow boundary condition and *blue* ones represent a constant-head boundary condition. The *bold black lines* are the 1D elements representing karstic conduits; the *white cells* are the limestone matrix elements, and in *gray* are the matrix elements with a high hydraulic conductivity

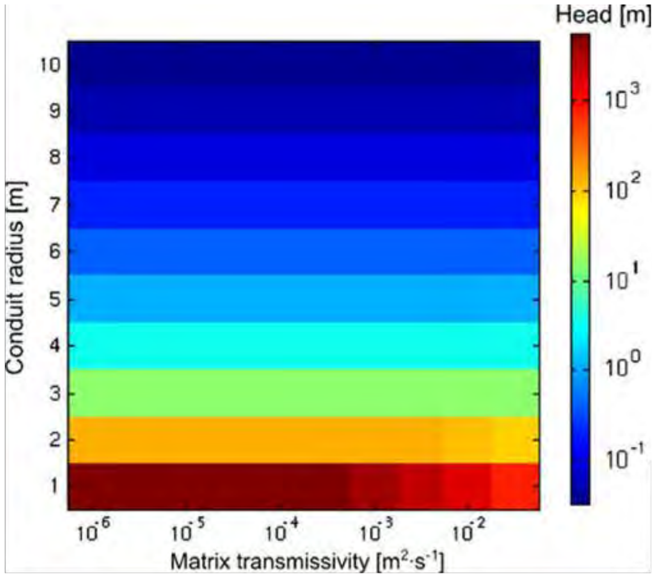


Fig. 13 Maximum simulated heads resulting from a series of simulations testing the influence of the matrix transmissivity and conduit radius

5 cm. These data are thus highly uncertain. Considering this fact and the simplicity of the steady-state model, the simulations were calibrated with the aim of reproducing the overall mean gradient rather than the punctual values. Regardless of tidal fluctuations, minimum and maximum gradients were computed by a linear regression of the punctual observed heads versus the distance to the coast. The 95 % confidence interval of the gradient is 2.74± 0.5 cm$\,$km^{-1}, which is illustrated in Fig. 12.

Sensitivity test

To assess the influence of matrix transmissivity T_{mat} and conduit radius r on the simulated hydraulic heads, a series of simulations were carried out. The same karstic network is used for each simulation. As the hydraulic conductivity contrast between matrix elements and conduit element increases between each simulation, the output flow field of one simulation was used as initial conditions for the next one. The solution is thus more easily computed by the program.

Figure 13 shows the maximum simulated head for T_{mat} ranging from 10^{-6} to $10^{-1.5}\,\mathrm{m^2\,s^{-1}}$ and r from 1 to 10 m. It appears that T_{mat} has a smaller influence on the simulated flow field than conduit radius, which suggests that most of the flow is drained by the conduit network. In the range of transmissivity values that yield realistic hydraulic heads (i.e. less than 1 m) T_{mat} variations have almost no influence on the simulated flow field. Since this parameter has little impact on the evaluation of the model uncertainty, a fixed transmissivity of $10^{-4}\,\mathrm{m^2\,s^{-1}}$ was selected for the following simulations.

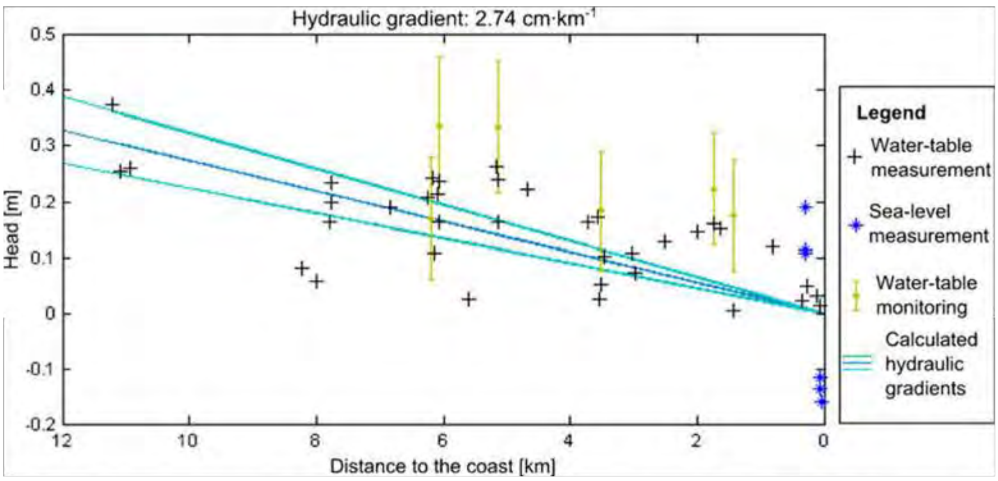


Fig. 12 Measured hydraulic heads versus distance to the coast. *Error bars* on the water-table monitoring points represent their standard deviation

Calibration of karst network parameters

Since only very few measurements were available, a simple model of conduit radius is proposed. The karst network is regarded as an ordered and hierarchical system (Fig. 10). An analogy was then made with river systems and, which assumes that the radius of a conduit can be related to its order via a power law:

$$r = \alpha \cdot e^{\beta u} \tag{6}$$

where r is the conduit hydraulic radius and u the conduit order. α and β are parameters that characterize the system. This type of law is observed in various complex hierarchical systems: for example, they relate many features of river channels such as drainage area, segment length or width, with their order (Horton 1945). Here, it is considered that such a relation can provide a reasonable model in the case of a karst system. It would have to be tested further against field data but for the preliminary model, it provides a simple way to control the variability of the hydraulic radius of the network with only two parameters.

A series of simulations with varying α and β for three increasing network densities (Table 1; Fig. 10) were run. Both α and β range from 0.1 to 1.6. An overview of the maximum head of each simulated flow field is shown in Fig. 14a. It appears that α and β have an important control on the flow simulation; however, the three conduit densities yield similar results. Numerous model configurations with large conduit radius could not be solved numerically (in gray in Fig. 14a), which is probably due to the high conductivity contrast between matrix elements and conduits which prevents the solution from converging.

Based on the calculated gradient of 2.74±0.5 cm km^{-1}, hydraulic heads at the upstream limit of the model (11.8 km from the coast) should lie between 26 and 38 cm (Fig. 12). Simulations that yield maximum heads in this range are shown in Fig. 14b. They are considered to be good approximations of the system. A parameter summary of each simulation is presented in Table 2.

The three network densities can provide equally good fits. The corresponding conduit radiuses range from 2.4 to

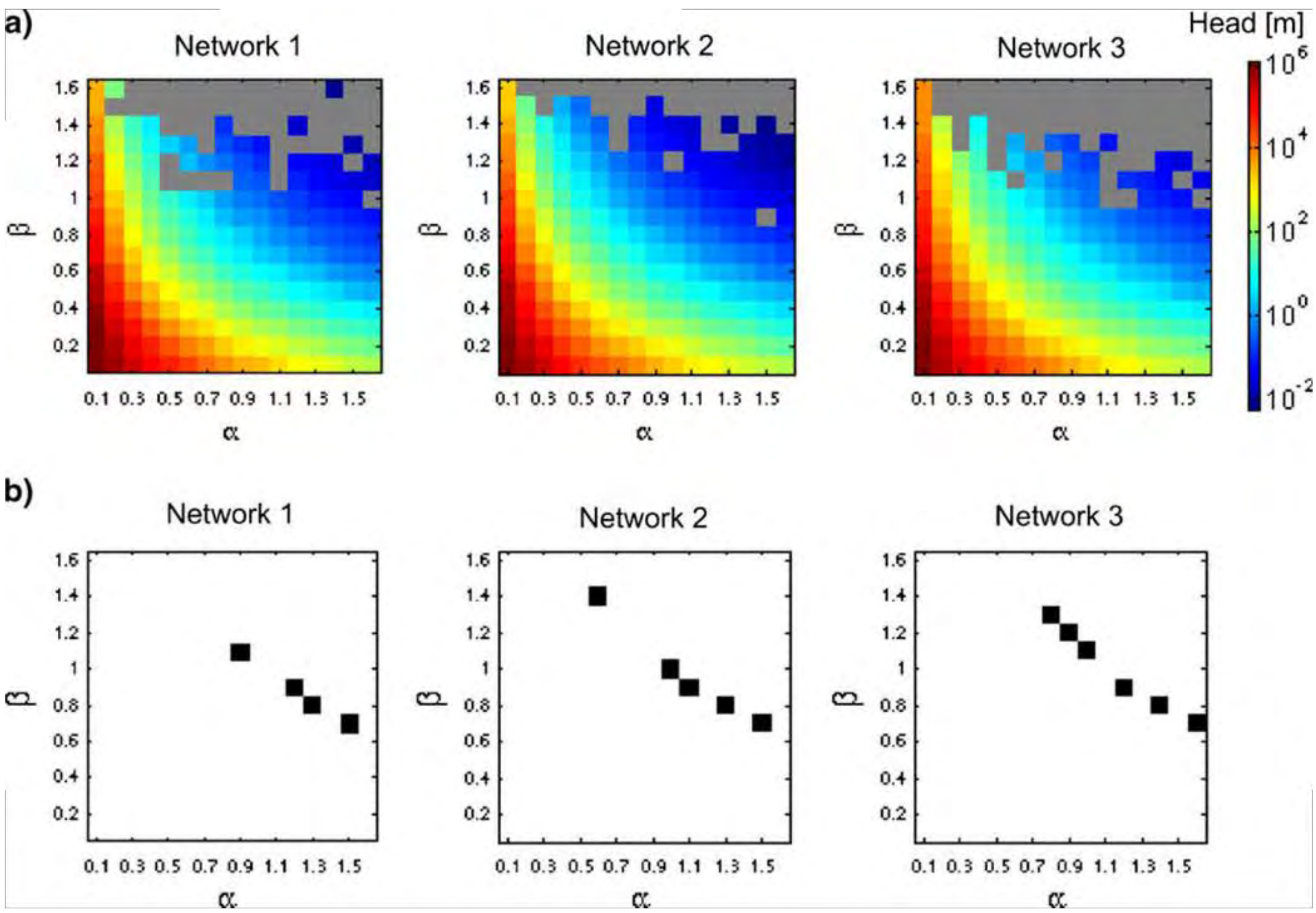


Fig. 14 **a** Maximum simulated hydraulic head (in meters, on a logarithmic scale) for a set of simulations. In *gray* are the models for which the simulator could not compute a solution. The varying parameters are α and β, empirical parameters defining the radius of each conduit order, and the network density (increasing from network 1 to 3, see Fig. 14. **b** Models that yield a maximum head matching the observed gradient

Table 2 Summary of the parameters of the simulations yielding realistic hydraulic gradient (measured gradient: 2.74±0.5 $cm km^{-1}$). Referred network densities are shown in Fig. 14

Order 1 radius [m]	Order 2 radius [m]	Order 3 radius [m]	Network density	Overall simulated gradient [$cm km^{-1}$]	Maximum simulated velocity [$m s^{-1}$]
2.7	8.1	24.4	1	3.21	0.93
3.0	7.3	17.9	1	2.37	0.78
2.9	6.4	14.3	1	3.02	0.82
3.0	6.1	12.2	1	3.02	0.75
2.4	9.9	40.0	2	2.74	0.77
2.7	7.4	20.1	2	2.29	0.71
2.7	6.7	16.4	2	2.73	0.75
2.9	6.4	14.3	2	2.35	0.68
3.0	6.1	12.2	2	2.43	0.64
2.9	10.8	39.5	3	2.69	0.85
3.0	9.9	32.9	3	2.50	0.82
3.0	9.0	27.1	3	2.50	0.81
3.0	7.3	17.9	3	3.03	0.83
3.1	6.9	15.4	3	2.57	0.74
3.2	6.5	13.1	3	2.58	0.68

3.2 m for conduit order 1, from 6.1 to 10.8 m for order 2, and from 12.2 to 40.0 m for order 3. Figure 14b shows that α and β are correlated, which is illustrated in Fig. 15 for all the network types together. A second order least-squares fit of α versus β yield the relation:

$$\beta = 0.37\alpha^2 - 1.60\alpha + 2.27 \cdot \qquad (7)$$

The available data set is therefore not sufficient to constrain the radius model and to provide a uniquely defined set of parameters. This implies a large source of uncertainty and indicates that additional data must be acquired in the future to better constrain that part of the model.

The maximum conduit velocities for the selected simulations are presented in Table 2 and are in the range of 0.6–0.9 $m s^{-1}$, which is in agreement with the in-situ observation by divers of strong currents, but has yet to be confirmed by proper measurements. The output flow and velocity fields of one of these simulations are shown in Fig. 16. The heterogeneity of the piezometric and velocity maps illustrates the strong influence of the conduits on the flow simulation. The velocity map shows that water flows

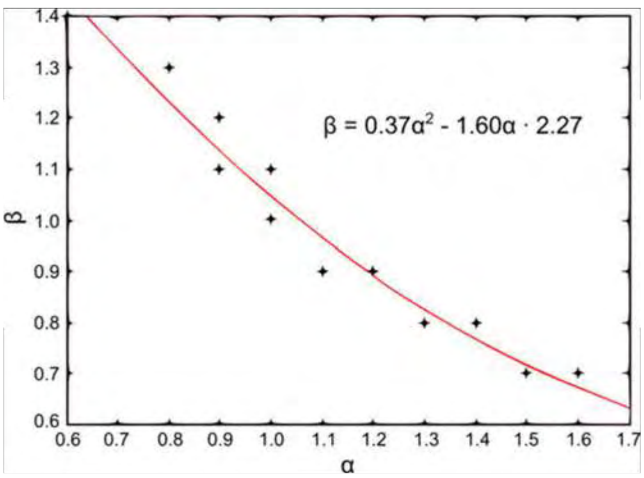


Fig. 15 Second order least-squares fit of the combinations of α and β that yield realistic flow fields

at approximately 1 m y^{-1} in the matrix, while in major conduits, velocities are higher than 1 $cm s^{-1}$, indicating that a large proportion of the flow occurs in the conduits.

Discussion and conclusion

This study presents the development of a finite-element flow model of the karstic aquifer located on the eastern coast of the Yucatan Peninsula. In a first step, the known karstic network, mapped by divers, is extrapolated using a modification of the stochastic pseudo-genetic karst simulator (Borghi et al. 2012), which allows to account for electromagnetic data resulting from extensive airborne geophysical surveys. The resulting conduit networks are modeled as 1D pipes, accounting for turbulent flow in a porous matrix.

However, the resulting hydrogeological models are subject to much uncertainty. A set of 15 tested parameter configurations yield hydraulic fields that match hydraulic head observations, out of more than 700 tested models. In addition, a wide set of stochastic karst networks can be generated with the chosen input parameters. It can, however, be observed that the satisfying combinations of parameters α and β tend to follow a specific trend for each of the three proposed network densities.

The uncertainty of the model arises on the one hand from the lack of conditioning data and on the other hand from the karstic nature of the aquifer. The one-off water-level measurements are difficult to handle because the measurement accuracy is 5 cm, whereas the hydraulic gradient is only a few centimeters per kilometer. In addition, tidal fluctuations observed in boreholes (~10 cm) have not been taken into account for the calculation of the hydraulic gradient. An in-depth study of the tidal wave propagation in the aquifer (delay, amplitude decay) is necessary in order to interpret more precisely the piezometric levels at that scale. Beside, water tables in karstic aquifers are strongly influenced by the position of conduits. Thus, a regional interpretation of localized measurements requires consideration of the karst

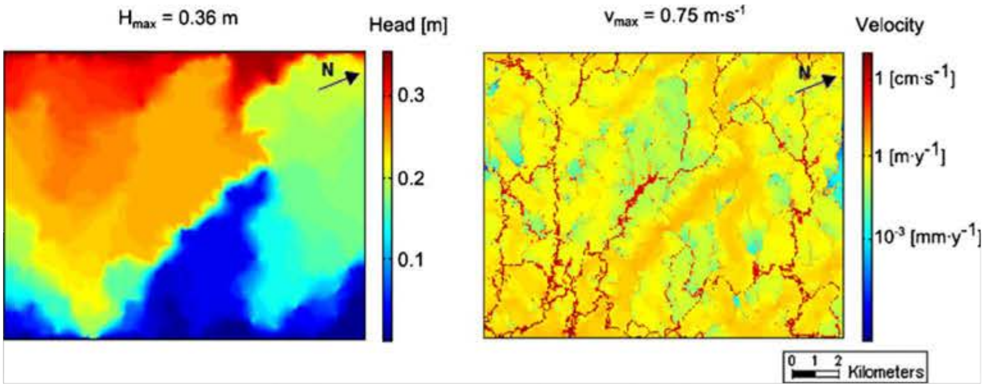


Fig. 16 Piezometric and velocity maps simulated with network 1, α=1.5 and β=0.7. The area is located in the grey box in Fig. 2a

network configuration. Finally, the models do not allow unconnected karst features, which might affect the flow model for they have an important storage capacity but low transmissivity.

Furthermore, a high level of uncertainty is linked with the conduit radius estimates. As no detailed measurements were available, the network was oversimplified in three radius classes. The application of an analogy with Horton laws for radius estimates is questionable, as the ordering does not respect the standard Horton or Strahler ordering systems for river systems. Again, further research is needed here to check the model against data and to elaborate a specific relation for karstic systems.

The necessity to include highly conductive flow channels in karst aquifer models has been pointed out by many authors, for example Kiraly (1998) and Worthington (2009). This cannot go without a high uncertainty since a thorough exploration of the caves is unrealistic, even in this study case where extensive cave maps were available. The method that is proposed here allows for building reasonable models of the cave conduits and exploring the uncertainty. In the specific situation of the Yucatan system, the available electromagnetic measurements constitute a major clue for the completion of the network. Their potential in revealing karstic conduits has been established in previous studies (Ottowitz 2009; Supper et al. 2009).

Although the exact position of the conduits is uncertain, the flow models including the stochastic conduits yield more realistic simulated water tables and flow paths than any homogenous or distributed equivalent porous medium model. Indeed, groundwater flow is widely drained by the cave network. This is in agreement with the estimation provided by Worthington et al. (2000) that 99.7 % of the flow in this aquifer occurs in the conduits. Moreover, the simulation of turbulent flow in conduits allows realistic velocity and travel time estimates even if the whole complexity of the system was not accounted for and even if data were missing to calibrate precisely the distribution of the radius of the conduits. The proposed model gives a global insight on the aquifer behavior. Should it be used to address questions of flow paths and travel times at a more local scale, the output would be very different from one realization to another and thus would allow one to estimate uncertainty in those forecasts. This could be used in an iterative procedure to define where and what type of information is required to better constrain the model and optimally reduce the uncertainty.

Regarding pollution risk, the results of these preliminary flow simulations suggest that the vulnerability of the aquifer is extremely high. Indeed, computed velocities in conduits are on the order of tens of cm s^{-1}, which is unusually rapid for groundwater flow, which infers that wastewater travel-times from the injection point to the outlets are short. Pollutant decay and/or absorption processes that could reduce water toxicity are thus inhibited.

Tulum wastewater is likely to travel to the sea at a very rapid rate. Indeed, cave maps indicate the presence of three explored conduits just beneath the town (Fig. 2a). Based on the proposed conduit probability map (Fig. 8b), it is highly probable that a straight connection exists between Tulum and the sea. According to the velocities computed by flow simulations, water could travel this 3-km distance very rapidly. This is alarming considering the major coral reef that lies near the coast and the lack of wastewater treatment in this area, but needs to be confirmed by in-situ velocity measurements. Regarding hydrologic linkages between Tulum and the Sian Ka'an Biosphere Reserve, a direct karstic connection is possible but has not been proven. More likely, pollution could travel from the sea to Sian Ka'an lagoons, which induces a major pollution risk for the protected ecosystem hosted in this area. To quantify this risk, a better understanding of the karst network and a study of the potential contaminant sources are required.

Acknowledgements The geophysical survey was financially supported by the Austria Science Fund project L524-N10 "XPLORE"; by ETH Zurich and the Austrian Academy of Science within the framework of the MAB program. We would like to sincerely acknowledge the support of Robbie Schmittner, Xibalba Dive Centre, Mario Rebolledo Vieyra, Centro de Investigación Científica de Yucatan (CICY), A.S. Gutierrez of NAFTA for importing and exporting the scientific equipment, and to the Secretearía de Marina—Armada de México. We are very grateful to "The Mexico Cave Exploration Project" and CINDAQ for permission to use the Ox Bel Ha cave line maps. Line maps were kindly provided by the Quintana Roo Speleological Survey,

2006. Principal explorers of the Ox Bel Ha system: A. Alvarez, B. Birnbach, S. Bogaerts, F. Devos, C. Le Maillot. S. Meacham, B. Phillips, S. Richards, D. Riordan, S. Schnittger, R. Schmittner, G. and K. Walten. Survey divers: Franco Attolini, Alejandro Alvarez, Roberto Chavez Arce, J. Jablonski, L. Magheli, A. Marassich, D. Riordan, G. Rocca, P. Thomsen, M. Valotta, J. Sanchez, J. Sherman. The salary of A. Borghi and the development of the technique to model the karstic network were funded by Schlumberger Water Services.

References

Bauer-Gottwein P et al (2011) Review: the Yucatán Peninsula karst aquifer, Mexico. Hydrogeol J 19(3):507–524

Beddows PA (2004) Groundwater hydrology of a coastal conduit carbonate aquifer: Carribean coast of the Yucatán Peninsula, México. PhD Thesis, University of Bristol, UK, 303pp

Beddows PA, Smart PL, Whitaker FF, Smith SL (2007) Decoupled fresh-saline groundwater circulation of a coastal carbonate aquifer: spatial patterns of temperature and specific electrical conductivity. J Hydrol 346(1–2):18–32

Borghi A, Renard P, Jenni S (2012) A pseudo-genetic stochastic model to generate karstic networks. J Hydrol 414–415(1):516–529

Chilès J-P, Delfiner P (1999) Geostatistics: modeling spatial uncertainty. Wiley Series in Probability and Statistics. Wiley, Chichester, UK, 695pp

Collon-Drouaillet P, Henrion V, Pellerin J (2012) An algorithm for 3D simulation of branchwork karst networks using Horton parameters and A*. Geol Soc Lond Spec Pub doi: 10.1144/SP370.3

Cornaton F (2007) Ground water: a 3-D ground water and surface water flow, mass transport and heat transfer finite element simulator, reference manual. University of Neuchâtel, Neuchâtel, Switzerland, 418 pp

Fournillon F, Viseur S, Arfib B, Borgomano J (2010) Insights of 3D geological modelling in distributed hydrogeological models of karstic carbonate aquifers. In: Andreo B, Carrasco F, Durán JJ, LaMoreaux JW (eds) Advances in research in karst media. Springer, Berlin, pp 257–262

Geovariances (2010) Isatis 10.0: technical references. Geovariances, Avon, France, 248 pp

Gondwe BRN (2010) Exploration, modelling and management of groundwater-dependent ecosystems in karst: the Sian Ka'an case study, Yucatan, Mexico. PhD Thesis, Technical University of Denmark, Denmark, 86pp

Gondwe BRN et al (2010) Hydrogeology of the south-eastern Yucatan Peninsula: new insights from water level measurements, geochemistry, geophysics and remote sensing. J Hydrol 389(1–2):1–17

Gulden B, Coke J (2011) World longest underwater caves. In: NSS GEO2 Committee on long and deep caves. http://www.caverbob.com/uwcaves.htm. Cited 12 Sept 2011

Henrion V, Pellerin J, Caumon G (2008) A stochastic methodology for 3D cave systems modeling. In: Ortiz J, Emery X (eds) Eighth Geostatistical Geostatistics Congress, vol 1. Gecamin, Santiago, Chile, December 2008, pp 525–533

Héraud-Piña M-A (1996) Le Karst du Yucatan, Pays des Mayas [The Yucatan Karst, land of the Mayas]. Presses Universitaires de Bordeaux, Bordeaux, France, 284 pp

Horton RE (1945) Erosional development of streams and their drainage basins: hydrophysical approach to quantitative morphology. Geol Soc Am Bull 56(3):275–370

Huang H (2008) Airborne geophysical data leveling based on line-to-line correlations. Geophysics 73(3):F83–F89

Iturralde-Vinent MA, MacPhee RDE (1999) Paleogeography of the Caribbean region: Implications for Cenozoic biogeography. Bull Am Mus Nat Hist (238):1–95

Jaquet O, Siegel P, Klubertanz G, Benabderrhamane H (2004) Stochastic discrete model of karstic networks. Adv Water Resour 27:751–760

Kiraly L (1998) Modelling karst aquifers by the combined discrete channel and continuum approach. Bull Centre Hydrogéol Neuchâtel 16:77–98

Lauterjung H, Schmidt G (1989) Planning of water intake structures for irrigation or hydropower, planning for intake structures. Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ), Bonn, Germany

Mariethoz G, Renard P (2010) Reconstruction of partial images and super-resolution using direct sampling. Math Geosci 42(3):245–268

Mariethoz G, Renard P, Straubhaar J (2010) The direct sampling method to perform multiple-point geostatistical simulations. Water Resour Res 46(11):W11536

Marin L, Steinich B, Pacheco J, Escolero O (2000) Hydrogeology of a contaminated soil-source karst aquifer, Mérida, Yucatán, Mexico. Geofís Int 39(004):359–365

Moore YH, Stoessell RK, Easley DH (1992) Fresh-water sea-water relationship within a groundwater-flow system, northeastern coast of the Yucatan Peninsula. Ground Water 30(3):343–350

Motschka K (2001) Aerogeophysics in Austria. Bull Geol Surv Jpn 52(2/3):83–88

NASA (2010) Yucatan Peninsula. In: Earth science data and information system. http://rapidfire.sci.gsfc.nasa.gov/. Cited 26 Aug 2010

Ottowitz D (2009) 3D-Modellrechnung der Karststrukturen des Ox Bel Ha Höhlensystems zur Methodenevaluierung: Aeroelektromagnetik [3D-modeling of the airborne electromagnetic signal over karst structures of the Ox Bel Ha cave system for method evaluation]. MSc Thesis, University of Vienna, Austria, 123pp

Pardo-Igúzquiza E, Dowd PA, Xu C, Durán-Valsero JJ (2012) Stochastic simulation of karst conduit networks. Adv Water Resour 35:141–150

SCT (2010) Manifestación de Impacto Ambiental Modalidad Regional Construcción del Libramiento de Tulum, en el Estado de Quintana Roo [Environmental Impact Assessment for Tulum Peripheral Road, Quintana Roo State, Road and Transport Ministry, Mexico City]. Secretaría de Comunicaciones y Transportes, Mexico City

Sethian JA (1996) A fast marching level set method for monotonically advancing fronts. Proc Natl Acad Sci U S A 93:1591–1595

SGM (2007) Carta geológico-minera: Estado de Campeche, Quintana Roo y Yucatán [Geology-mining chart: Campeche, Quintana Roo and Yucatan]. Servicio Geológico Mexicano, Oaxaca, Mexico

Smart PL et al (2006) Cave development on the Caribbean coast of the Yucatan Peninsula, Quintana Roo, Mexico. In: Harmon RS, Wicks CM (eds) Perspectives on karst geomorphology, hydrology, and geochemistry: a tribute volume to Derek C. Ford and William B. White. Geol Soc Am Spec Pap, Boulder Co, 404:105–128

Supper R et al (2009) Spatial mapping of submerged cave systems by means of airborne electromagnetics: an emerging technology to support protection of endangered karst aquifers. Near Surf Geophys 7(5–6):613–627

Vacher HL, Mylroie JE (2002) Eogenetic karst from the perspective of an equivalent porous medium. Carbon Evaporit 17(2):182–196

Ward WC, Keller G, Stinnesbeck W, Adatte T (1995) Yucatan subsurface stratigraphy: implications and constraints for the Chicxulub impact. Geology 23(10):873–876

Worthington SRH (2009) Diagnostic hydrogeologic characteristics of a karst aquifer (Kentucky, USA). Hydrogeol J 17(7):1665–1678

Worthington SRH, Smart CC (2003) Groundwater in karst: mathematical models. In: Gunn J (ed) Encyclopaedia of caves and karst science. Routledge, Abingdon, UK, pp 401–403

Worthington SRH, Ford DC, Beddows PA (2000) Porosity and permeability enhancement in unconfined carbonate aquifers as a result of solution. In: Klimchouk A, Ford DC, Palmer AN, Dreybrodt W (eds) Speleogenesis: evolution of karst aquifers. National Speleological Society, Huntsville, AL, pp 463–472

Appendix D

SKS user manual

SKS, The Stochastic Karst Simulator

- User guide -

v.0.3

Andrea Borghi

University of Neuchâtel, Switzerland

March 2013

Table of Contents

1 Introduction

SKS (the "Stochastic Karst Simulator") is a MATLAB® prototype that allows to simulate karst conduits using the *pseudo-genetic* methodology, which is fully described in the PhD thesis of Andrea Borghi (2013). This user guide is not intended to describe the methodology and subsequent implementation, but provides a short description about how to use the SKS software.

The users that are already used to this software may need only to refresh their memory by reading the "quick start check-list" in section 2, while new users may find useful to read also the description of the parameters in section 3. In section 4 a detailed description of the different input and output files is given.

SKS does not have a graphical user interface. To use it, the user needs to prepare the input files, modify the command file called `options_karst.m` and run it in a MATLAB® prompt typing the command "sks". The `options_karst.m` file must be in the current working directory. To install SKS, follow the "readme" instructions provided in the sources folder.

2 “Quick start” check list

Before SKS is run, the user must prepare several input files, depending on the options that he provides in the `options_karst.m` input file. The input file format are described in section 4.1, while the options are described in section 3. The present section gives only a short summary of the most important steps that have to be followed before running the simulator:

- Assign FMA values to `codeOut`, `codeAquifere`, `codeAquiclude` (section 3.3).
- Prepare a geological model `geoleFileName` GSLIB file. Fill also the `geologyIdArray` and `geologyFMArray` variables in the `options_karst.m` input file (section 3.6). These two arrays inform SKS about the correspondence between the geological IDs present in the `geoleFileName` GSLIB file and their FMA value (section 3.3).
- If the inlets of the system have to be loaded from a file (`startPtsType>=1`), prepare the input file with the inlet coordinates `startingPtsInFile` (section 3.4).
- Define the `importanceFactor` variable. This variable is very important because it determines the degree of hierachization of the final network (section 3.4).
- Prepare the outlet coordinates file `springInputXYZFile` (section 3.5).
- If the model has to be cut according to a polygon (`cutModelByPolygon=1`), prepare this coordinate file too (`cutPolygonCsvFile,`section 3.6).
- If a fracture model has to be loaded (`fracturesYoN=1,` section 3.7), prepare the `fracturesFileName` GSLIB file, if the built-in fracture generator has to be used (`fracturesYoN=2`), fill the corresponding section in the `options_karst.m` input file. Assign also a FMA value to `FMvalFrac` (section 3.3).
- If a fault model has to be loaded (`faultsYoN=1,` section 3.8), prepare the `faultsFileName` GSLIB file and assign also a FMA value to `FMvalFaults` (section 3.3).
- Fill the “general output” section in the `options_karst.m` input file (section 3.1).
- when everything is ready, type “sks” in the MATLAB® prompt and the conduit simulation will start.

3 Description of the SKS command file

SKS does not have a user interface. All the options have to be provided in one unique MATLAB® command file called `options_karst.m` . This file has to be in the MATLAB® path, or in the current folder. This file is divided into several sub-sections which are described in detail here below. Moreover, the `options_karst.m` file is highly commented. In practice every option that has to be provided, has several commentary lines that explains it directly into the command file. The present section of the manual is first of all intended to provide guidelines on how to correctly parameter the options of the command file. A lot of SKS options are binary, i.e. the user can assign a value of 0 (deactivate) or 1 (activate) for these options. Other can have multiple values, which will modify the behaviour of SKS. Finally, some other (like the FMA parameters, or the paths to the input files) can have any value, and the user can give them the value that he wants. NB: there are no input check before the SKS execution, therefore if the input parameters are not consistent, the main script will probably encounter issues and exit with errors.

3.1 General output options

This section of `options_karst.m` file informs SKS about the output files that have to be written. See section 4.2 for more information about the output file formats:

- `verbosity:` this option controls the level of "verbosity" of SKS during execution. In practice it defines the quantity of messages that are written to the prompt. 0 means "no output", 1 means "normal information output" (and is the suggested value), 2 and 3 are "debug" and "full output" values, which are useful only for debugging.
- `nomSimul:` this variable is a string. It defines the basename for all the output files of the model, i.e. all the output files will start with this string. If `outputDirYoN` is active, then also a directory with this name will be created to store the resulst.
- `outputDirYoN:` if this option is active (1), then the output files of SKS will be moved in a directory called `<nomSimul>` and this directory will be moved to `<outputDir>.`
- `outputDir:` string variable which gives the path to the directory where the results have to be stored
- `saveWorkspaceYoN:` if this option is active, the MATLAB® workspace will be saved after the execution.
- `VTK_YoN:` if this option is active (1) , the visualization VTK files will be outputted.
- `outputGslib:` if this option is active (1), the result GSLIB file will be outputted.
- `saveAsciiMeshFiles:` this option control the output of the finite element mesh. It can have 4 values: 0, deactivated; 1: only 1D mesh; 2: only 3D mesh; and 3: both meshes.
- `logOptions_YoN:` if this option is active (1), the current options_karst.m file will be kept in the result directory as backup (useful when a lot of tests are made).

- `DTM:` if this option is active (1), the DTM of the model will be outputted.
- `saveStrPts_YoN:` if this option is active (1), the starting points that have been used int the model will be outputted.
- `saveFracturesYoN:` if this option is active (1), the fractures generated by SKS will be outputted for further use in a GSLIB format.

3.2 Mesh dimensions

In this section, the user must provide the mesh dimensions. SKS only works with cartesian regular grids. The grid dimensions must coincide with the size of the geological model that will be loaded. The paramters are:

- `ni,nj,nk:` the number of elements in y,x, and z dimensions respectively. (**NB:** ni → y; nj → x; nk → z)
- `x0,y0,z0:` the coordinates of the origin of the grid. The origin of the grid correspond to the the lower left corner of the grid.
- `dx,dy,dz:` the grid cells size in x,y and z dimensions.

NB: I recommend to work with **regular cubic grids** (i.e. dx=dy=dz) to avoid artifacts.

3.3 Fast marching speed function

The algorithm is based on the assumption that water follows the "minimum effort path" in a karst aquifer, to connect the inlets of the system (dolines, sinkholes) to the karst springs. The karstification will therefore occur along these preferential minimum effort paths. This assumption is simulated using a Fast Marching Algorithm (FMA, Sethian, 2008), which computes the time of the first arrival of a propagation front in each cell of the model. To compute the time map, the FMA needs the information about the speed that the front can follow in each cell (see the PhD, chapter 5, for more information).

There are several parameters in this section. The first parameters are the FMA values for the different "facies" of the model:

- `codeOut:` the FMA value that flags the zone of that are outside the limit of the model (e.g. outside the topgraphy, outside the limit polygon). A value near to 0 is a good value for this parameter (default 0.0011).
- `codeAquifere:` the FMA value for the aquifer, i.e. for the karstifiable formations.
- `codeAquiclude:` the FMA value for the aquicludes (non-karstifiable formations). To be consistent, the value of `codeAquiclude` must be smaller than `codeAquifere.`
- `FMvalFrac:` the FMA value for the fractures. This value is applied directly if `useCumulativeFractures_YoN=0`. If `useCumulativeFractures_YoN=1`, then the FMA value of the fractures is computed, with a minimal value of `codeAquifere` and a maximal value of `FMvalFrac`. The exact value depend on the number of fractures that are present in the cells of the model (min->`codeAquifere`; max->`FMvalFrac`).

- `FMvalFaults:` the FMA value that will be assigned to faults
- `FMvalHorizons:` the FMA value that will be assigned to inception horizons (if specified)
- `multiplicatorConduits:` the FMA value for the conduits is computed as a multiplication of the `FMvalFrac` value with this multiplicator. A good value for this parameter is 2 or 3.

The really important thing that has to be kept in mind with the above parameters, is that the strongest the contrast, the strongest the higher value facies will "attract" the conduits. For example, to produce a model where there is a weak influence of the fractures over the conduit path, the value of `FmvalFrac` can be set to 1.2 or 1.5 the value of `codeAquifere`. On the contrary, to give them a high control over the conduit path, `FMvalFrac` can be set as more or less `codeAquifere*10`. In practice, the user needs to play a little with these parameters until he "feels" the model. The experiments done until now shows that a contrast of x10 between `codeAquifere` and `codeAquiclude` is highly sufficent to constrain all the conduits in the aquifer formation. The same contrast can be set between `FMvalFrac` and `codeAquifere` to produce also good results. The multiplication factor for the conduits will determine how much the already simulated conduits attract the next ones. Here again, a contrast between 1.2 (weak attraction) to 5 (very strong attraction) can be given.

NB: the values associated with these facies **must be different from each other** because they are used later by SKS to flag the facies into the model. If the user wants to give the same value to the different facies (homogenous model) he can do it by adding a very small epsilon to the FMA values (adding 0.0001 works quite well).

Two additional parameters are present in this section of the command file:

- `reRunFMA_YoN:` if this flag is activated (1), the conduits will be constrained below the altitude at wich they started. This option is useful to avoid conduits that "go upward", i.e. the conduits that go against the gravity rules. This option should be always set to 1, but as it may be very long to simulate (needs to re-compute the FMA for each conduit), it can be deactivated for debug reasons, or to gain time.
- `step1:` if this flag is active, the vadose conduits will be simulated. This means that first all the conduits are generated from the topography to the base level (the altitude of the higher spring is taken as base level), and secondly the phreatic conduits will be simulated. If during the simulation of vadose conduits they reach the top of an aquitard formation, they will stop at this level.

3.4 Inlets of the system

This section contains the several options and parameters that are relative to the inlet points of the system. The inlets of the system are very important because they correspond to the starting points of the conduits. The order in which they are simulated will control the final hierarchy of the model, and therefore the final connectivity of the network. Different options are available for the inlet points:

- `startPtsType:` this parameter can have a value comprised between 0 and 4. If a value of 0 is given, the inlet points are generated randomly. SKS will generate a `nb_RandomInlets` number of random inlets. The random inlets coordinates will be placed on the topography surface of the model, where the aquifer is outcropping, in the bounding box provided by the user with the parameters `xMinStart, yMinStart, xMaxStart` and `yMaxStart` (see below). If `startPtsType>=1` the inlets will be loaded from a separate file (see section 4.1). If `startPtsType=2` the inlets are a combination of the loaded inlets from file and the randomly generated. If `startPtsType=3` the inlets are loaded from a file, and their order is recomputed by their respective catchment (3rd column of the input file). If `startPtsType=4` the inlets are a combination of those loaded from the file, and the random ones, but reordered by their catchment extension (3rd column of the file). In this latter case, the random inlets catchment is assumed to be equal to the smallest one found in the input file.

- `startingPtsInFile:` the input file that contains the inlet coordinates and catchment (see section 4.1 for file format)
- `permuteCoord:` if this option is active (1), the order of the loaded points is randomly reordered. This add diveristy to the final connectivity of the model. This option is deactivated by SKS if `startPtsType>=3` (i.e. ordered by catchment).
- `xMinStart, yMinStart, xMaxStart, yMaxStart:` The bounding box coordinates for the randomly generated inlets.
- `nb_RandomInlets:` number of random inlets to be generated.
- `importanceFactor:` this is a very important variable. This variable will highly determine the hierarchy of the network. First of all, the variable is a MATLAB® vector, and the length of this vector determines the number of iterations of the conduit simulation. (Example: if the length of `importanceFactor` is equal to 3, then there will be 3 iterations.) Secondly, the values that are contained in this vector correspond to the proportion of inlets that will be used for each iteration. Example: if `importanceFactor=[1,2,7]` then there will be 3 iterations of the algorithm, in the first one 1/10 of the inlets will be used, in the second 2/10 of the inlets and the remaining 7/10 of the inlets in the last iteration. It is important to remember that the first conduits that are simulated will attract the next ones, therefore very different connectivities can be obtained only by playing with this parameter and using always the same inlets.

3.5 Outlets of the system

This section of the `options_karst.m` file contains the options that correspond to the outlets of the system (i.e. the karst springs):

- `springInputXYZFile:` This variable must contain the path to the file with the coordinates of the karst springs (see section 4.1 for file format).
- `permuteSprings:` If this option is active (1), the order of the springs that are loaded from the file will be randomly changed. This can add variability to the resulting networks.
- `alternateSpringsYoN:` if this option is active (1), and that there are more than one spring, the conduits will be simulated toward each spring separately. This lead to multi-connected inlets. In practice, if this option is deactivated (0), then there will be only one possible connection between the different inlets and the different springs. If the user does not know what

to do, just leave it active (1).

3.6 Geological model

This section of `options_karst.m` file contains the parameters relative to the geological model. These parameters can be divided into 3 subgroups. The main options are:

- `typeOfGeology`: this string variable defines wich kind of geological model will be used. If it is "uniform", then no geological model will be loaded, and a homogeneous geological model with an FMA value of `codeAquifere` (see section 3.3) will be generated. If it is "geomodeller" or "petrel", then SKS loads the GSLIB file containing the geological editor output file. NB: Geomodeller and Petrel output the geological model with a different ordering of the elements than SKS. Therefore in this case, SKS will read it in a special way. The last value that this option can have is "normal", in this case the ordering of the elements in the GSLIB file is already good and SKS can only load it directly.
- `geoleFileName`: this is the GSLIB file name in which the geological model is loaded (see section 4.1 for file format).
- `geologyIdArray`: This array contains the geological ID that can be found in the geological model, and for which a particular FMA value has to be given. In practice, only the ID that correspond to the aquifers and aquiclude need to be listed here. The corresponding FMA value can be given in `geologyFMArray` variable (here below).
- `geologyFMArray`: this array contains the corresponding FMA values that have to be given to the formations contained in "`geologyIdArray`". Both arrays must have the same size, if some formations are not listed in these arrays but are in the geological model, a value of `codeOut` will be assigned to them.

Then there are 2 options that are relevant if the model must be cut according to a polygon that defines the limit. This is useful if the geological model is bigger than what is relevant for flow. This polygon will also determine the deactivated cells for a MODFLOW model and the cells that will be removed by the finite element mesh if GROUNDWATER has to be used as flow simulator:

- `cutModelByPolygon`: If this option is active (1), then the model will be cut by the polygon defined in the `cutPolygonCsvFile` variable.
- `cutPolygonCsvFile`: name of the file containing the coordinates of the limit polygon (see section 4.1 for file format).

And finally, the last 3 parameters of this section are relevant only if some geological indexes have be selected as preferential horizons (so-called "inception" horizons).

- `inception_YoN`: If this option is active (1), the preferential horizons defined by `inceptionIdArray` will be used, and a FMA value of `FmvalHorizons` will be assigned to them (see section 3.3).
- `nb_inception`: number of preferential horizons to be used for each realization (if this number is inferior to the number of possible inception horizons, then `nb_inception` of them will be randomly selected in each realization).

- `inceptionIdArray`: similar to `geologyIdArray`, this variable contains the ID of the facies found in the geological model input GSLIB file that correspond to preferential horizons.

3.7 Fractures model

- `fracturesYoN`: This option can have 3 different values. If it is set to 0 (inactive), there will be no fractures in the model. If it is set to 1, the fracture model is loaded from a GSLIB file (see section 4.1 for file format). If it is set to 2, the built-in fracture generator will be used (see below).
- `fracturesFileName`: this string variable correspond to the input GSLIB file for the fractures (see section 4.1 for file format).

If `fracturesYoN` is set to 2, i.e. that the built-in fracture generator has to be used, the following options must also be filled.

- `chooseNFrac`: the total number of fractures that have to be generated. If `verbosity>=2`, when simulating the fractures, SKS will output the theorical number of fractures that should be used according to the Universal Fracture Model (UFM, Davy, 2010).
- `strikeMinFamilies`: this variable and the next 2 (i.e. `dipMinFamilies` and `importanceFactorFamFrac`) are the 3 vectors of the same length, that define the families of fractures that have to be generated, where each value contained in the vector correspond to the value relative to the given fracture family. This first vector defines the strike of the families. The values are given in degrees from 0 (North) to 364 (North again) according to the strike of the fractures.
- `dipMinFamilies`: This vector defines the dip of the fracture families. The values are given in degrees and can variate between 0 (horizontal) to 90 (vertical).
- `importanceFactorFamFrac`: similarly to `importanceFactor` (section 3.4) this vector gives the relative proportion of fractures in each family.
- `fractureLength`: vector with the lengths for the fractures. This vector will influence the distribution of fracture of given length following the UFM model (Davy, 2010).

3.8 Faults model

The options that are contained in this section are relative to the use of faults as preferential formations for karstifiaction into the model:

- `faultsYoN`: if this option is active (1), then the faults are loaded from a GSLIB file. The FMA value that will be associated to the faults must be provided as `FMvalFaults` (see section 3.3).
- `faultsFileName`: input GSLIB file name with the discretized faults (see section 4.1 for file format).

3.9 Complex conduit paths

The options found in this section will allow the user to generate complex paths for the conduits. The idea is to avoid the conduits to only converging (tree structure) in order to obtain a more labyrinthic structure of the network. Therefore some “plugs” will be added in the propagation medium. These plugs can have a spheric shape, or result from a gaussian field (more variability). These are the possible options:

- `connectDeconnect_YoN:` this option can have 3 values. 0 means “deactivated”, 1 means “spherical plugs” and 2 means “gaussian plugs”. If it is set to 1 (spheres) then `rBalls` must be given, while if it is set to 2, then `paramsGauss` must be given (see below).
- `nPlugsMax:` the maximal number of plugs to be added in the propagation medium between one iteration and the other. The real number of plugs is randomly sampled between `nPlugsMax/2` and `nPlugsMax.`
- `rBalls:` this variable is a vector with 2 elements. The first element correspond to the minimum radius for the plugs, and the second element to the maximal radius for the plugs. The radius value for each plug is then randomly sampled in a uniform distribution between the minimal and the maximal possible values.
- `paramsGauss:` this is a vector providing the information for the gaussian shaped plugs. These information are : `paramsGauss=[lx,ly,lz, mu et sigma2, quantile]` where lx,ly,lz, are the correlation length in x,y and z directions, mu is the mean and sigma2 is the variance of the distribution, quantile is the quantile threshold at which the gaussian property is cut to form the plugs (geobodies).

3.10 Flow simulation files and meshing

There are 4 options relative to the flow simulation files:

- `modflowYoN:` if this option is active (1), the input files for the finite differences solver MODFLOW will be generated. **NB: if modflow has to be used, the next 3 options MUST BE DEACTIVATED (set to 0).**

The next 3 options are relative to finite elements grids and meshing:

- `onlyCutTheMesh:` If this option is active (1), the finite element mesh will be cut where the geological model is set to `codeOut`, i.e. above the topography and around the limit polygon.
- `use1Dconduits:` If this option is active (1), the 1D elements (2 node pipes) will be included into the 3D mesh. The topology will be consistent between the different types of elements (3D hexahedrons for geological formations and 1D pipes for conduits). If this option is active, it will automatically set `onlyCutTheMesh=0.`
- `GW_YoN:` if this option is active (1), the input files for the finite element solver GROUNDWATER (Cornaton, 2007) will be generated.

4 File format

This section describes the input and output file format for SKS.

4.1 Input file format

- **Geological model:** the geological model is given in the standard GSLIB file format (http://www.gslib.com/gslib_help/format.html). The first field of the file has to be an integer field, with the ID of the different geological formations. The 2nd, 3rd and 4th fields are optional. They provide the information about the orientation of the lithology in a vectorial form for x,y, and z. This correspond to the normal vector to the stratigraphy. The SKS variable is: `geoleFileName.` If there are 4 fields (geology and orientation), and if the built-in fracture generator has to be used, then the generated fractures will be rotated according to the orientation found in this file (see section 3.7).

- **Faults model:** the fault model is provided also in standard GSLIB format. The only field of the file is a binary information that inform about the presence or absence of the faults (0: absent ; 1: present). The SKS variable is: `faultsFileName.`

- **Fractures model:** Again a GSLIB file. The only field of the model is a counter of the fractures in each cell of the model. The SKS variable is: `fracturesFileName.`

- **Inlets coordinates:** the inlets are provided in an ascii file with at least 2 columns. The first 2 columns are the X and Y coordinates of the inlets. The altitude must not be provided, as it will be computed by SKS later. Optionally a 3rd column can be provided, with the catchment surface for each inlet. The SKS variable is: `startingPtsInFile.`

- **Outlet coordinates:** The outlets are provided in an ascii file with 3 columns: X,Y and Z coordinates of the outlets of the system. The SKS variable is: `springInputXYZFile.`

- **Polygon for the limit of the model:** if the limit of the model are provided with a polygon, this is given in a 2-column ascii file with the X and Y coordinates of the polygon points. The points must be ordered in the polygon order (i.e. they are read in order of appearance). The SKS variable is: `cutPolygonCsvFile.`

4.2 Output file format

Several output are available. All the output files will begin with a prefix string that is provided by the user in the `nomSimul` variable. If the option `outputDirYoN` is active, an output folder will be created for the output files. It will be called `<nomSimul>.` The available output files are:

- `<nomSimul>.mat`: This MATLAB® file contains the full workspace at the end of simulations. This is the main output file of SKS. Loading later this file in MATLAB® can be usefull to avoid the re-computation of a model. **saveWorkspaceYoN** must be activated to output this file.
- `<nomSimul>_mesh3D.vtk`: This file is an unstructured-grid ascii legacy VTK file (http://www.vtk.org/VTK/img/file-formats.pdf). It is a visualization file that contains the finite element mesh. It can be opened with any VTK-based opensource software such as paraview (www.paraview.org). **VTK_YoN** must be activated to output this file.
- `<nomSimul>_mesh1D.vtk`: Also an unstructured-grid ascii legacy VTK file. This visualization file contains only the 1D (pipes) karst elements. **VTK_YoN** must be activated to output this file.
- `<nomSimul>_DTM.vtk`: Also an unstructured-grid ascii legacy VTK file. This visualization file contains modeled topography with the same resolution as the model. Useful for debug reasons mainly. **DTM** must be activated to output this file.
- `<nomSimul>.gslib`: This is an ascii GSLIB file (http://www.gslib.com/gslib_help/format.html) that contains the information about the geology, the FMA propagation field, the hydraulic material classes and the karst elements. **outputGslib** must be activated to output this file.
- `<nomSimul>_fractures.gslib`: This is an ascii GSLIB file (http://www.gslib.com/gslib_help/format.html) that contains the discrete fracture model. It is useful if the same fracture model has to be kept for multiple realizations. **saveFracturesYoN** must be activated to output this file.
- `<nomSimul>_startPts.dat`: This file contains the inlets coordinates that have been used for the current realization, in the order in which they have been simulated. This is an ascii file with 3 columns: X,Y and catchment of the inlets. This file is useful if the same inlets have to be used for several realizations, because it is already in the right format that can be used as inlet coordinate file (**startingPtsInFile**). To output this file **saveStrPts_YoN** must be activated.
- `<nomSimul>_nodes`: This ascii file contains the finite element mesh nodes. It is a 3 column file with the X,Y and Z coordinates of the nodes. The line number correspond to the ID of the node (i.e. the coordinates set at line 12 correspond to the node 12). **saveAsciiMeshFiles** must be activated to output this file.
- `<nomSimul>_elements`: This ascii file contains the topology of the finite element mesh. Each line correspond to an element, the ID present on each line correspond to the nodes that are connected together. **saveAsciiMeshFiles** must be activated to output this file.
- `<nomSimul>_springNodes`: This ascii file contains the list of the nodes ID that correspond to the outlets of the system (karst spring) in the finite element mesh. **saveAsciiMeshFiles** must be activated to output this file.
- `<nomSimul>_inletNodes`: This ascii file contains the list of the nodes ID that correspond

to inlets of the system (sinkholes, dolines) in the finite element mesh. `saveAsciiMeshFiles` must be activated to output this file.

- **MODFLOW** files: if `modflowYoN` is activated, several input files for MOFLOW-NWT will also be written. They are called `<nomSimul>.<MF-package>`, where `<MF-package>` is the modflow package (BAS, DIS, UPW, BCN, LMT, OC, and NAM files). It also writes a CSV file of ni x nj elements, with the zone information of the surface geology in the model (1 for aquifer, 2 for aquiclude, 3 for karst inlets). The UPW file (flow properties) also contains 3 facies: 1 for aquifer, 2 for aquiclude, 3 for karst inlets, as the recharge CSV. Moreover, the RCH file contains also this same information. For the karst springs, SKS outputs the BCN file, assigning a general head boundary condition (GHB) to the springs cells. The value of these GHB is the altitude of the springs. NB: the UPW file for MF-NWT is equivalent to the LPF file for MF2005.
- GROUNDWATER input files. If `GW_YoN` is activated, the files for this finite element solver will be output. This is a very experimental section and is therefore not commented here.

5 Additional considerations

5.1 Disclaimer

This software is an experimental research prototype. The software is provided by the copyright holders and contributors "as is" without warranty of any kind, express or implied, including, but not limited to, the implied warranties of merchantability and fitness for a particular purpose and non infringement. In no event shall the authors or copyright owner or contributors be liable for any claim, any direct, indirect, incidental, special, exemplary, or consequential damages (including, but not limited to, procurement of substitute goods or services; loss of use, data, or profits; or business interruption) however caused and on any theory of liability, whether in contract, strict liability, or tort (including negligence or otherwise) arising in any way out of the use of or in connection with this software, or other dealing in the software even if advised of the possibility of such damage.

5.2 Acknowledgment

Thanks to Schlumberger Water Services for having funded this research and to the people who have followed this project during 4 years: Sandra Jenni, Gilles Mathieu, and Arnaud Levannier. A particular thanks also to Jean-Pierre Delhomme who believed in this project from the beginning and supported the project with conviction and interest.

Thanks to yWorks company for providing YED for free and to Frans Slothouber for providing ROBODOC under a GPL license. Both software have been incredibly useful to create the SKS code documentation and flowchart.

Alzo lo sguardo l'orizzonte è là
La vita intera non basterà
Nell'ombra vedo te dolce follia
Musa... Indicami la via

Folkstone - Vortici Scuri