



Nouvelles approches de tomographies hydrauliques en aquifère hétérogène : théories et applications en milieu karstique et fracturé.

Pierre Fischer

► To cite this version:

Pierre Fischer. Nouvelles approches de tomographies hydrauliques en aquifère hétérogène : théories et applications en milieu karstique et fracturé.. Sciences de la Terre. Normandie Université, 2018. Français. NNT : 2018NORMR059 . tel-01955328

HAL Id: tel-01955328

<https://theses.hal.science/tel-01955328>

Submitted on 14 Dec 2018

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

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



Normandie Université

THÈSE

Pour obtenir le diplôme de doctorat

Spécialité Sciences de l'Univers

Préparée au sein de l'Université de Rouen Normandie

Nouvelles approches de tomographies hydrauliques en aquifère hétérogène
Théories et applications en milieu karstique et fracturé

Présentée et soutenue par
Pierre FISCHER

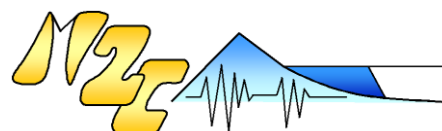
Thèse soutenue publiquement le 21 novembre 2018
devant le jury composé de

Mr Frederick DELAY	Pr / LHyGeS / Université de Strasbourg	Rapporteur
Mr Olivier BOUR	Pr / Géosciences Rennes / Université de Rennes 1	Rapporteur
Mr Philippe RENARD	DR / CHYN / Université de Neuchâtel	Rapporteur
Mr André REVIL	DR / ISTerre / Université Grenoble Alpes	Examineur
Mme Mickaële LE RAVALEC	DR / IFPEN Rueil-Malmaison	Examinatrice
Mr Michel QUINTARD	DR / IMFT / Université Toulouse 3	Examineur
Mr Abderrahim JARDANI	MCF HDR / M2C / Université de Rouen Normandie	Directeur de thèse
Mr Nicolas LECOQ	MCF HDR / M2C / Université de Rouen Normandie	Codirecteur de thèse

Thèse dirigée par Abderrahim JARDANI et Nicolas LECOQ, laboratoire Morphodynamique Continentale et Côtière (M2C)



Normandie de Biologie Intégrative,
Santé, Environnement



**Nouvelles approches de tomographies hydrauliques en aquifère
hétérogène**

Théories et applications en milieu karstique et fracturé

New hydraulic tomography approaches in heterogeneous aquifer

Theories and applications in karst and fractured fields

par

Pierre FISCHER

**Manuscrit de thèse rédigé et présenté en vue
d'obtenir le diplôme de doctorat**

Spécialité Sciences de l'Univers

**Thèse préparée au sein de l'Université de Rouen Normandie
2018**

RÉSUMÉ

Ce manuscrit de thèse présente une nouvelle approche pour caractériser qualitativement et quantitativement la localisation et les propriétés des structures dans un aquifère fracturé et karstique à l'échelle décamétrique. Cette approche est basée sur une tomographie hydraulique menée à partir de réponses à une investigation de pompages et interprétée avec des méthodes d'inversions adaptées à la complexité des systèmes karstiques. L'approche est appliquée sur un site karstique d'étude expérimental en France, une première fois avec des signaux de pompage constants, et une deuxième fois avec des signaux de pompage harmoniques. Dans les deux cas, l'investigation a fourni des réponses de niveaux d'eau de nappe mesurés pendant des pompages alternés à différentes positions. L'interprétation quantitative de ces jeux de réponses consiste à les reproduire par un modèle avec un champ de propriété réaliste adéquat généré par inversion. Les méthodes d'inversions proposées dans ce manuscrit permettent de reconstruire un champ de propriétés hydrauliques réaliste en représentant les structures karstiques soit par un réseau généré par automates cellulaires, soit par un réseau discrétisé. Les résultats d'interprétations obtenus sur le site d'étude expérimental permettent d'imager les structures karstiques sur une carte et de « lire » leur localisation. De plus, les résultats obtenus avec les réponses à des pompages harmoniques tendent à montrer le rôle de la fréquence du signal sur les informations portées par les réponses. En effet, les fréquences plus élevées caractérisent mieux les structures les plus conductrices, alors que les fréquences plus faibles mobilisent des écoulements également dans des structures karstiques moins conductrices.

ABSTRACT

This thesis manuscript presents a novel approach to characterize qualitatively and quantitatively the structures localization and properties in a fractured and karstic aquifer at a decametric scale. This approach relies on a hydraulic tomography led from responses to a pumping investigation and interpreted with inversion methods adapted to the complexity of karstic systems. The approach is applied on a karstic experimental study site in France, a first time with constant pumping signals, and a second time with harmonic pumping signals. In both applications, the investigation resulted in groundwater level responses measured during alternated pumping tests at different locations. The quantitative interpretation of these sets of responses consists in reproducing these responses through a model with an adequate realistic property field generated by inversion. The inversion methods proposed in this manuscript permit to reconstruct a realistic hydraulic property field by representing the karstic structures either through a network generated by cellular automata, or through a discretized network. The interpretation results obtained on the experimental study site permit to image the karstic structures on a map and to ‘read’ their localization. Furthermore, the results obtained with the responses to harmonic pumping tests tend to show the role of the signal frequency on the information carried by the responses. In fact, higher frequencies better characterize the most conductive structures, while lower frequencies mobilize flows also in less conductive karstic structures.

REMERCIEMENTS

Je vais commencer par remercier, dans cette partie, mes directeurs de thèses, Abderrahim Jardani et Nicolas Lecoq. Merci de m'avoir fait confiance, dès le début, en me prenant en doctorant sans m'avoir vraiment réellement rencontré et en montant un financement non forcément prévu à la base. Merci également pour la confiance et la liberté accordée durant ma thèse qui m'ont permis de publier des articles, et notamment sur un concept que je voulais y développer, les automates cellulaires. Ensuite je vais remercier les autres membres de mon jury, André Revil, Frederick Delay, Olivier Bour, Philippe Renard, Mickaële Le Ravalec et Michel Quintard. Merci d'avoir accepté de lire mon manuscrit, d'avoir assisté à ma présentation de soutenance et d'avoir discuté, commenté et critiqué mes travaux afin de me montrer de nouvelles perspectives.

Mes prochains remerciements, plus personnels, iront à mes proches et membres de ma famille. Tout d'abord merci ma chérie Louise, pour m'avoir soutenu tout le long de ce doctorat. Tu m'as d'abord encouragé dans la période de transition entre la fin de mon TFE et le début du doctorat 10 mois après. Tu m'as suivi à Rouen malgré la difficulté à y trouver un travail. Puis tu m'as soutenu, apaisé et parfois secoué dans les hauts et les bas de ces trois années. Enfin tu m'as grandement aidé dans la dernière ligne droite pour organiser toute ma journée de soutenance. Pour tout ça je te remercie, mais aussi plus généralement pour tout le bonheur que tu m'as apporté. Il n'y a pas une seule seconde depuis qu'on est ensemble durant laquelle l'amour que j'ai pour toi et la fierté de t'avoir comme compagne n'a cessé de grandir.

Merci à mes parents, Christine et Gottfried, pour m'avoir soutenu dans l'idée de réaliser un doctorat et d'être régulièrement venu nous rendre visite à Rouen. Merci également pour l'aide apportée à la réalisation de mon pot de soutenance. Merci à mes grands frères, Julien et Stephan, pour m'avoir « ouvert la voie » dans le système éducatif et avoir influencé mes différentes décisions de formations. Merci Stephan pour tes nombreux conseils, ton aide précieuse quand j'en avais besoin, d'être venu des Etats-Unis pour ma soutenance, et pour les games de LoL épiques les week-ends. Merci à mes autres proches qui se sont déplacés pour assister à ma soutenance, Mamie, Papi, Laurence, Agnès et Jacques. Je remercie également Oma et Opa, qui ne pouvaient malheureusement pas se déplacer pour ma soutenance mais qui ont eu l'occasion de m'adresser tous leurs encouragements.

Merci enfin aux autres membres de mes proches (plus particulièrement Marie-Christine, Nicole, Guillaume, Ysabault, Charlotte) pour leur constante bienveillance envers moi et leurs encouragements tout au long de mon doctorat. Pour finir, merci Snow, même si tu ne sais pas (encore) lire, je suis sûr que tu apprécieras au moins mes papouilles.

Ensuite, je vais remercier les autres personnes qui m'ont accompagné, de près ou de loin durant ces trois années de doctorat. Je commencerai par remercier les doctorants que j'ai régulièrement côtoyé au labo M2C « antenne de Rouen ». Merci Arnaud, Aziz, Théo, Manu, Flavie, Marie, Antonin, Thomas, Adeline, David, Mohamad, Nazih, Ruth, Manon, Mickaël, Léa, Valentin, Edward, Antoine, Raphaël. Merci à cette « ligue de doctorants extraordinaires », très soudée et joviale, qui aura grandement participé à rendre mes trois années de doctorats très agréables, que ce soit au labo durant les passionnants et passionnés échanges du repas du midi, ou dans les nombreuses soirées.

Je remercie également les membres permanents du labo. Merci d'abord à Michel, pour ton aide sur la préparation et la réalisation de ma campagne de terrain, mais également évidemment pour ta bonne humeur que tu apportes au labo. Merci également à Abderrahim, Nicolas L., Jean-Paul (avec une pensée particulière), Maria, Nicolas M., Matthieu, Robert, Julien, Maxime et Yoann, pour vos conseils et l'aide que vous avez pu régulièrement m'apporté en colloque ou au labo. Merci à tous les autres membres du labo pour votre accueil qui m'a fait me sentir bien au sein du laboratoire (à tel point que je me suis mis à m'y promener en chaussettes).

Je vais finir par remercier les personnes qui m'ont décidé à faire un doctorat et celles qui m'ont aidé à le réaliser. Merci Fabrice Golfier pour le projet de Master qui m'a appris et donné le goût des modèles et de la recherche. Merci David Labat de m'avoir fait découvrir les karsts et leurs études. Merci Hervé Jourde pour l'aide apportée sur les applications de mes méthodes sur le site du Terrieu. Merci Michael Cardiff pour l'aide apportée sur les modèles fréquentiels et les signaux harmoniques. Merci enfin à tous mes autres co-auteurs, Xiaoguang, Stéphane, Mohamad, Abdellahi, pour leurs aides et leurs corrections sur mes articles.

SOMMAIRE

Liste des figures	10
Liste des tableaux (tables)	20
1 Introduction générale	23
2 Tomographie hydraulique en milieu alluvial avec inversion géostatistique	37
2.1 Contexte	39
2.2 Application of large-scale inversion algorithms to hydraulic tomography in an alluvial aquifer	41
2.2.1 Introduction	42
2.2.2 Principal component geostatistical approach	45
2.2.3 Application to an experimental site	49
2.2.4 Results	52
2.2.5 Conclusion	61
3 Un nouvel outil d'inversion structurale pour les modèles équivalents milieux poreux	63
3.1 Contexte	65
3.2 A cellular automata-based deterministic inversion algorithm for the characterization of linear structural heterogeneities	67
3.2.1 Introduction	68
3.2.2 Parameterization of inverse problem using cellular automaton	70
3.2.3 Optimization process	81
3.2.4 Applications	85
3.2.5 Discussion and conclusion	99
4 Un nouvel outil d'inversion structurale pour les modèles discrets couplés	103
4.1 Contexte	105
4.2 Hydraulic tomography of discrete networks of conduits and fractures in a karstic aquifer by using a deterministic inversion algorithm	107
4.2.1 Introduction	108
4.2.2 Algorithm framework	111
4.2.3 Validation of the DNDI algorithm on hypothetical study cases	121
4.2.4 Discussion	130
4.2.5 Conclusion	131
5 Caractérisation des écoulements dans le site du Terrieu par pompage à débit constant	135
5.1 Contexte	137
5.2 Identifying flow networks in a karstified aquifer by application of the cellular automata-based deterministic inversion method (Lez aquifer, France)	139
5.2.1 Introduction	140
5.2.2 Methodology	142
5.2.3 Application	148
5.2.4 Conclusion	161

6	Analyse hydraulique des réponses à un pompage harmonique en milieu karstique et fracturé	163
6.1	Contexte	165
6.2	Hydraulic Analysis of Harmonic Pumping Tests in Frequency and Time Domains for Identifying the Conduits Networks in a Karstic Aquifer	167
6.2.1	Introduction	168
6.2.2	Theoretical background	171
6.2.3	Synthetic application	175
6.2.4	Example of harmonic pumping investigation	189
6.2.5	Discussion and conclusion	195
7	Application de tomographie hydraulique en mode harmonique sur le site du Terrieu	199
7.1	Contexte	201
7.2	Harmonic Pumping Tomography Applied to Image the Hydraulic Properties and Interpret the Connectivity of a Karstic and Fractured Aquifer (Lez Aquifer, France)	203
7.2.1	Introduction	204
7.2.2	Field investigation	207
7.2.3	Modeling methodology	215
7.2.4	Modeling application	223
7.2.5	Discussion	237
7.2.6	Conclusion	239
8	Conclusion générale et perspectives	241
	Bibliographie (References)	251
	Annexes (Appendix)	263

LISTE DES FIGURES

Figure 1.1: Gauche : Répartition des grands hydrosystèmes souterrains français (BRGM 2015). Droite : Répartition des formations karstifiables en France (Marsaud 1996).	26
Figure 1.2: Schématisation des écoulements dans trois types de milieux pouvant être aquifère : milieux poreux, fracturés et karstiques (modifié d'après Heath 1998).	26
Figure 1.3: Schéma de la caractérisation des propriétés par interprétation des réponses à un signal de sollicitation.	27
Figure 1.4: Représentation schématisée d'une investigation par pompage avec débit à signal constant d'un aquifère fracturé et karstique (modifié d'après Goldscheider et Drew 2007).	28
Figure 1.5: Schéma de la reconstitution des propriétés à partir d'un modèle direct et de l'inversion des réponses mesurées.	30
Figure 1.6: Schéma entre la liberté attribuée dans la reconstitution des réponses dans le processus d'inversion par rapport aux contraintes établies sur les solutions.	31
Figure 1.7: Haut : Schéma des différentes techniques de modélisation distribuées issu de Ghasemizadeh et al. 2012 (EPM = Equivalent Milieu Poreux ; DPM = Milieu Double Porosité ; DFN = Réseau de Fractures Discrètes ; DCN = Réseau de Conduits Discrets ; HM = Modèle Hybride). Bas : Exemples de modélisations d'aquifères karstiques par (a) EPM (Saller et al. 2013), (b) DPM (Kordilla et al. 2012), et (c) HM (Kovacs 2003).	32
Figure 2.1: Location of the studied experimental site 'La Céreirède' (Map and aerial photography from geoportail.fr) occupying an area of 720 m ² . It is situated in the South of France, near the town of Montpellier and the Mediterranean Sea.	49
Figure 2.2: Schematic geological section of the experimental site 'La Céreirède'. Three aquifers formations have been characterized in the sands and silts alluvium, in the gravels and in the clayey sands.	50
Figure 2.3: Well pattern on the experimental site 'La Céreirède' (circles represent the 10 measurement wells and triangles represent the 2 pumping wells). As hydraulic drawdowns in the pumping wells are not measured, the tomography provided 20 observed data.	51
Figure 2.4: Maps of the log-transmissivity (a, on the left) and parameter's a posteriori standard deviation (b, on the right) for a PCGA inversion method with 25,600 parameters, 20 observed data and a covariance matrix decomposition of order $K=128$ applied to the experimental site. The transmissivities vary around a mean of 10^{-5} m ² /s which is consistent with transmissivity values estimated from pumping test analysis. The aquifer is less transmissive in the eastern part and more in the western part especially in a zone around PZ 7 (see Figure 2.3). But we got a better precision in zones with more information: at the center and the western part of the	

map, while in the eastern part where we didn't have piezometers, the results show a larger standard deviation. 53

Figure 2.5: Graph showing the differences between the 20 observed drawdowns and modeled drawdowns after convergence of a PCGA inversion method with 25,600 parameters and a covariance matrix decomposition of order $K=128$ applied to the experimental site. The drawdowns are globally well reproduced. 54

Figure 2.6: Covariance matrix singular values decrement curve for 10,000 parameters. Three decomposition order (a to c) corresponding to truncation error of 1, 0.1 and 0.01 have been chosen for the results comparison of the PCGA inversion method (see 0.02 Figure 2.7). 55

Figure 2.7: Maps of the log-transmissivity for a PCGA inversion method with 10,000 parameters, 20 observed data and three different covariance matrix decomposition applied to the experimental site. The map (a) was obtained for $K=69$, the map (b) for $K=313$ and the map (c) for $K=1,532$ (see Figure 2.6). The results obtained for these three decomposition are relatively the same (same transmissivity values, same zones) so, for this site, there is no significant loss of information when using a truncation order corresponding to an error of 1 (map (a)) for the covariance matrix which allows us to reduce the computation time of the inversion without decreasing the accuracy of the results. 56

Figure 2.8: Maps of the log-transmissivity (a, c, e) and parameter's a posteriori standard deviation (b, d, f) for three different inversion methods with 10,000 parameters and 20 observed data applied to the experimental site. The maps (a) and (b) were obtained with the GA adjoint-state method, the maps (c) and (d) with the GA finite-difference method and the maps (e) and (f) with the PCGA method with a covariance matrix decomposition of order $K=69$. The results between the three methods are relatively the same for this site, except for the map a which presents a slightly higher contrast of the transmissivity distribution leading to a better data matching (see Figure 2.9), though the PCGA inversion method is much more efficient for the calculation time (see Table 2.2). 57

Figure 2.9: Graphs showing the differences between the 20 observed drawdowns and modeled drawdowns after convergence of three different inversion methods with 10,000 parameters applied to the experimental site. The graph (a) was obtained with the GA adjoint-state method, the graph (b) with the GA finite-difference method and the graph (c) with the PCGA method with a covariance matrix decomposition of order $K=69$. Regarding the mathematical norm 2 the GA adjoint-state method has a slightly better convergence on the data than the other methods but the PCGA inversion method is much more efficient for the calculation time (see Table 2.2). 59

Figure 3.1: Scheme explaining how the CA are used in the CADl model. In the figure grey occurs for state 'background' and white for state 'structure'. The model is partitioned in m_{CA} independent CA subspaces (here $m_{CA}=9$). During the generation process the structure will go through different CA subspaces (a) and will be generated in the local direction assigned by the structural parameters piloting these CA (b). Along the generation direction the CA will modify the property values of the model cells it controls (represented by the squares lattice in Figure 3.1c). 70

Figure 3.2: An example of the dual-radius neighborhood considered in our CA definition. The black highlighted cell is the cell under consideration in this example (each cell of the lattice would alternatively be considered during a full CA time step). The green highlighted cells are considered as its 'activators' neighbors in the transition rule and the orange highlighted cells as its 'inhibitors' neighbors. These cells are selected by an inner and an outer circle (in bold) with configurable radius

which permit the configuration of the neighborhood. In this example, the inner circle has a radius = 2 and the outer circle has a radius = 3. Additionally, the neighborhood is split into 2×8 sectors (by the radial lines) which permit a more configurable weighting definition (see Figure 3.4). 73

Figure 3.3: Time evolution of a CA configured with a neighborhood weighting defining a horizontal structure generation (see Figure 3.4). After the sixth time step the CA has converged and its geometry is stable over the following steps. Here grey occurs for state 'background' and white for state 'structure'. 74

Figure 3.4: Presentation of 8 different stable structures started by a unique centered cell, and their associated CA neighborhood configuration. The greyed cell in the neighborhood configuration is a given cell considered during the CA process. It is surrounded by its neighbor cells, which are not shown for reasons of readability. Its neighborhood is split in 8 internal 'activator' sectors and 8 external 'inhibitor' sectors, each one being assigned to a given weight. A '++' occurs for a positive weight for the neighbor cells in the area, a '++' weight is twice higher than a positive weight represented by a single '+'. A '--' occurs for a negative weight for the neighbor cells in the area, a '--' weight is twice higher than a negative weight represented by a single '-'. An empty part of the neighborhood occurs for a null weight, meaning that cells in the area are not considered in the transition rule. Here we present the CA configuration leading to 8 different structure directions which will be considered as sub-orientation of the global structure in the model. In the structural map, grey occurs for state 'background' and white for state 'structure'. 75

Figure 3.5: Operating scheme for the Cellular Automata-based Deterministic Inversion (CADI) algorithm. After an initialization of \mathbf{P}_N and \mathbf{P}_p with chosen directions and property values for each subspace, the algorithm begins an iterative process. It will firstly optimize the geometry of the structure in the model by iteratively updating the structural model using the CA generation process. Once the objective function has converged to a local minimum on the structure, it will lead a second optimization on the values of the properties for the previously inverted structure, until the objective function converges to a local minimum again. Finally, the uncertainties on the structure and the properties of the model are estimated. 80

Figure 3.6: Presentation of the 4 different structures tested in the 6 study cases in this paper. The case 1 is a linear inversion of a simple geometry (a) to show how the optimization works. The case 2 is a linear inversion of a more complex geometry (b). The case 3 is a linear inversion of a complex multi-directional linear structure (c). The cases 4, 5 and 6 are linear, non-linear and joint inversion of a geostatistical generated geometry (d), appearing as a more natural structure. 86

Figure 3.7: Result of the linear inverse modeling of the case study 1. The inversion finished after 4 iterations. This figure shows all different iterations of the inversion, from initial model (a), to inverted model (e). The true structure is shown in (f). The figure (d) corresponds to the structural optimization and the figure (e) to the properties optimization for this structure. The different CA subspaces of the model are highlighted by the black lines. 87

Figure 3.8: Comparison of the optimal structure found by inversion (in white) and the true structure (bold boundaries) for the case study 1. For this simple geometry, the inverse algorithm could easily reproduce the structure. 88

Figure 3.9: Result of the linear inverse modeling of the case study 2. The convergence is performed with 21 iterations. This figure shows some different iterations of the inversion, from initial model (a), to inverted model (e). The true structure is shown in (f). We noted that the optimization on the property values permits to

balance the structural inversion errors. For example, in this case, the structural additions in the center of the model in (e) were optimized by a light augmentation of its seismic velocity (0.5 km/s instead 0.26 km/s). 89

Figure 3.10: Comparison of the optimal structure found by inversion (in white) and the true structure (bold boundaries) for the case study 2. The optimization process reproduced a good structural inversion. The few inversion errors in the center of the model were lightly balanced by the inversion on the properties (see Figure 3.9). 90

Figure 3.11: Results of the linear inverse modeling of the case study 3. This figure shows two inversions with different initial models (a and d), their results (b and e) and the comparison of these results to the true geometry boundaries in red (c and f). The convergence is performed with 26 iterations in the first inversion and 30 iterations in the second. We noted that the information contained in the initial model could slightly modify the result of the inversion but even with a very simple initial case (a) the optimization process permits to find the main shapes and trends of the true structure (c). 92

Figure 3.12: Map of the positioning of the wells for the hydraulic tomography inversion for the study case 4. The circles are the position of the measurement piezometers and the triangles are the position of the pumping wells. 92

Figure 3.13: Result of the non-linear inverse modeling of the study case 4. The inversion finished after 7 iterations. This figure shows the initial model (a), the inverted model (b) and the true structure (c). The inversion process found an optimized equivalent structure to the initial property value. The true transmissivities were found during the properties optimization. 93

Figure 3.14: Result of the linear inverse modeling of the study case 5. The inversion finished after 4 iterations. This figure shows the initial model (a), the inverted model (b) and the true structure (c). The structural optimization was limited by the initial properties and by its constant aperture generation to reproduce a variable aperture true structure. In this case, the optimization on the property values permits to balance the initial information and the structural inversion aperture limitations. The properties optimization balanced this limitation by globally decreasing the seismic velocity of the background to a lower value than the true one. 94

Figure 3.15: Result of the joint inverse modeling of the study case 6. The inversion finished after 7 iterations. This figure shows the hydraulic model (a), the seismic model (c) and the true models (b,d). The geometry of the structure in the models was optimized through a joint inversion of seismic and hydraulic data. 96

Figure 3.16: Pixel-wise comparison of the optimal structures found by inversion (in white) and the true structure (bold boundaries) for the study cases 4 (a), 5 (b) and 6 (c). Both hydraulic and seismic data permitted to find a geometry of the global trends of the true structure but the joint inversion resulted to a better model regarding the structure and also the convergence on the data, which avoided the difficulties encountered by the simple hydraulic inversion and the simple seismic inversion. 96

Figure 3.17: Uncertainties analysis for the joint inversion of the study case 6. The structural constraint in (a) indicates where the structure of the model is well-constrained by a low value, and at the opposite, a high value indicates an uncertainty for its subspace direction. The properties uncertainties for (b) the hydraulic transmissivity and (c) the seismic velocity are quantified by a standard deviation on the inverted values. 97

Figure 4.1: Example of a simulated distribution of hydraulic heads (here drawdowns) by solving the forward problem f (Equation (4.1)) for a steady state pumping in a given coupled discrete-continuum distributed model $\Gamma(\mathbf{P}_{Dir}, \mathbf{P}_{Prop})$. 112

Figure 4.2: Schema of the node-to-node generation process in the DNDI method with six subspaces. An activated node in the top subspaces (a) starts the generation of the structure. The structure generates to the nodes in the bottom of these subspaces, following the local direction defined in the subspaces through the encoding rules. These reached nodes then become activated (b). The subspaces in which the structure has generated become inhibited to another generation (shown as greyed number in this figure). The structure then continues its generation from its newly activated node if the subspaces structural parameters permit it (c) – (d). 113

Figure 4.3: Parameterization of a model in the DNDI method. For each subspace of the model there are six local direction possibilities (see encoding in Figure 4.2) that are used to parameterize a network structure in the model (a). The structure (in red) is then generated, following a node-to-node rule, from the set of structural parameters in (a) and a chosen starting point at a node between subspaces (b). Finally a set of property values (transmissivities), also defined for each subspace, is assigned to the structural model (c). 114

Figure 4.4: A flowchart of the inversion steps used in the DNDI algorithm. After the initialization of the parameters, a sequential iterative optimization is led on the structure geometry and on the property values in order to minimize both objective functions (Equations (4.5) and (4.6)). An eventual re-run of the inversion process (multi-scale option) using the result as new initial model can be performed in order to improve this result. 118

Figure 4.5: Initial and inverted models for an inversion using drawdown data produced from a true model (on the right) with a homogeneous matrix. The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data. The inverted model permits to localize approximatively the karstic network connections but in this case the amount of data is insufficient to have a proper imagery. 123

Figure 4.6: Initial (a) and inverted (b) models for an inversion using drawdown data produced from a true model with a homogeneous matrix, and associated map of the conduit properties posterior standard deviations (c). The inverted model in (b) permits a good localization the true karstic network. It also reduced locally the initial transmissivity ($0.06 \text{ m}^2/\text{s}$ to $0.01 \text{ m}^2/\text{s}$) of the conduits connected to the primary drain in the bottom right part of the model (the conduit thickness is proportional to its transmissivity value). The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data. 124

Figure 4.7: Initial and inverted models for an inversion using drawdown data produced from a true model (on the right) with a homogeneous matrix. The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data, primarily located in the matrix. The inverted model permits to almost reproduce the karstic network even if only two measurement points are located in the true network. 125

Figure 4.8: Initial and inverted models for an inversion using drawdown data produced from a true model (on the right) with a heterogeneous matrix. The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data. A first inverted model (a) permits to localize the true karstic network but also generates conduits to simulate the more transmissive part of the true model. A second inversion (b) starting from the previous inverted model permits to correct the geometry and produces an inverted model matching more accurately the true model. 126

Figure 4.9: Maps of the conduit and matrix transmissivities posterior standard deviations. The matrix higher transmissivity zones in the inverted model (bottom left) have a higher uncertainty value than the lower transmissivity zones (top right). On the contrary, the uncertainty on the transmissivities of the conduits of the primary drain is higher than the secondary conduits.

127

Figure 4.10: Initial and inverted models for an inversion using drawdown data generated from a true model (on the right) with a homogeneous matrix. The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data. A first inverted model (a), starting from a simple initial model, permits to localize approximately the true network geometry. A second inversion (b), starting from a more detailed initial model, permits to produce a more precise network geometry.

128

Figure 4.11: Maps of the posterior uncertainties of the network local directions for the Cases a and b. In the Case a, started from a simple initial model, the highest uncertainties are distributed uniformly over the inverted network. In the Case b, started from a more detailed initial model, the highest uncertainties are located in the periphery of the model.

129

Figure 5.1: Scheme of the 8 different weighting distributions N possibilities to parameterize the CA subspaces. Each distribution defines a different direction for the conduit-state generation shown by the arrows. The dual radius neighborhood is described here for a given cell in grey (the other cells are not shown for a reason of readability). In the configurations $N_i, i \in [1, 4]$ the circles are defined by an inner circle of radius 2 cells and an outer circle of radius 6 cells, and in the configurations $N_i, i \in [5, 8]$ the circles are defined by an inner circle of radius 4 cells and an outer circle of radius 5 cells. The neighbor cells of the greyed cell are split in 8 internal 'activator' weighting sectors and 8 external 'inhibitor' weighting sectors represented by the two radially split circles. A neighbor cell in state matrix can be associated (given its position in the neighborhood) to a positive weight '+ +' which is twice higher than a '+' weight, or to a negative weight '- -' which is twice higher than a '-' weight, or to a null weight in the empty sectors and beyond the neighborhood.

143

Figure 5.2: Presentation of a model in the CADI algorithm. Here the model is partitioned in 9 subspaces controlled by CA. The model is parameterized by a structural parameter \mathbf{P}_N (here $\mathbf{P}_N(5) = N_1$; $\mathbf{P}_N(4) = \mathbf{P}_N(6) = N_4$ and

$\mathbf{P}_N(1) = \mathbf{P}_N(9) = N_3$ (see Figure 1)) and a property values parameter \mathbf{P}_β

(here every subspace is defined by the same β but it could vary in each subspace).

Initially the whole model is considered as matrix, except an initial conduit cell. Within the CA time process the conduit is generated from this initial cell and propagates through the model depending on the subspaces structural parameters until it reaches a global converged geometry.

144

Figure 5.3: (a) Map indicating the location of the experimental site. The black square indicates the location of the Lez aquifer in which the Terrieu site is included.

(b) Distribution of twenty-two boreholes of the Terrieu experimental site. The red dots indicate the boreholes where the pumping tests were performed while the grey dots indicate the measurement boreholes. (c) Pumping rates (red captions). Inferred principal flow path connectivity (blue dotted lines) and local karstic conduits (green lines) based on downhole videos, well logs, and packer tests. The orientation of the green lines indicates the orientation of local karstic features observed on downhole videos. A green dot indicates that no karstic features were seen in this borehole.

149

Figure 5.4: Schematic showing the sequential series of inversions led to obtain the final flow network model. The initial model was partitioned with 4×6 subspaces for its inversion. The inverted flow network model was then used as a new initial model for an inversion with 8×12 subspaces. The same operation was repeated on last time so that our final flow network has a partitioning of 16×24 subspaces. 151

Figure 5.5: (a) Comparison of the observed drawdowns to the drawdowns modeled by the inverted flow model. (b) Resultant model of the inversion modeling showing the heterogeneous distribution of the transmissivities. (c) Comparison of the result model with the known preferential flow path connectivity (interpreted in the model in dotted blue lines). (d) Superimposition of the known local conduits direction (shown as blue lines) presented in Figure 5.3c. 153

Figure 5.6: Maps of hydraulic drawdowns calculated from the result flow network model. The drawdowns are shown for each of the pumping wells (white triangles) used for the hydraulic tomography (the pumping rate is indicated in each figure). The drawdowns can have very different forms depending on the localization of the borehole in a conduit or in the matrix, highlighting the heterogeneity of the model. Pumping in the matrix (P2, P10, P17) results in a very local drawdown, while pumping in a conduit (P0, P5, P11, P16, P21) produces a more global drawdown in the whole model (in these cases the area the most impacted by the pumping is delimited by white dotted lines). 154

Figure 5.7: Schematic representation of the modeled karstic structure at the Terrieu experimental site, considering the geological information, the hydraulic tomography investigation, and the flow network produced by inversion with the CADI method. The red lines indicate the boreholes where the pumping tests were performed, while the grey lines indicate the measurement boreholes. 155

Figure 5.8: The map of the network structural uncertainties (left) shows that the network geometry is well constrained especially in a zone between each borehole in the center of the model, and compared to the map of transmissivities standard deviation (right), the hydraulic data permitted to constrain more the conduits position than the matrix. 157

Figure 5.9: Maps of the pumped water velocities calculated by the result model for a pumping in borehole P0 and in borehole P21 (the two most productive pumping). The pumping boreholes are indicated by white triangles. For a reason a better readability of the low velocities, the scale has been fixed on a maximal velocity of 10^{-3} m/s, thus in the blackest zones, the velocity can be higher than this value (up to 10^{-2} m/s near the pumping point for P0). 158

Figure 5.10: Comparison of the inversion result produced by the CADI method and by the SNOPT method (Wang et al. 2016) at the same scale of the Terrieu field site and with same hydraulic dataset. The initial models are shown on the left and the inverted models are presented on the right. 159

Figure 6.1: The theoretical synthetic case used to study the responses of a harmonic pumping in a karstic field. A karstic network (in blue) composed of a large conduit (LC) and two thin conduits (TC) crosses a homogeneous matrix (in white). All conduits are 1-D features in the model, but shown with conductivity-weighted thicknesses for clarity. Eight different boreholes are positioned in the model and represent pumping or measurement points. 175

Figure 6.2: Drawdown responses h in each borehole to a harmonic pumping in P3 in a time domain simulation. If the greyed portion of the time series is not considered, these drawdown responses can be described as the sum of a linear signal $h_{lin.}$ and a purely oscillatory signal $h_{osc.}$. 178

Figure 6.3: Oscillatory signals responses in each borehole for a harmonic pumping in P3, for a frequency domain simulation and a time domain simulation (avoiding the first signal period). One sees that these signals are almost the same for the two simulations. 179

Figure 6.4: Relative amplitude (% , in blue) and relative phase offset ($^{\circ}$, in orange) values in the oscillatory responses in each borehole for different harmonic pumping locations (P4, P7, P6, P3). A dash represents an absence of oscillatory response (< 1 mm). The pumping location is indicated by 'P' and its drawdown oscillatory signal is considered as a 100% amplitude signal with a 0° phase offset. 181

Figure 6.5: Differences in relative amplitude (in blue, in %) and relative phase offset (in orange, in $^{\circ}$) values in the oscillatory responses by decreasing from a 5 min period signal to a 1 min period signal for two different harmonic pumping locations (P6, P3). A dash represents an absence of oscillatory response (< 1 mm). The pumping location is indicated by 'P'. The main signal differences appear for the boreholes located in the matrix, near to a conduit (P1, P4) (dual connection). 182

Figure 6.6: Comparison of the oscillatory relative responses for a harmonic pumping in P3 for a 1min period signal during 6 min (full line) and a 5min period signal during 30 min (dotted line). The measurement boreholes have been separated regarding their location: in a conduit (P2, P5, P6, P8) or in the matrix (P1, P4, P7). The main signal differences appear for the boreholes located in the matrix, near to a conduit (P1, P4) (dual connection). 183

Figure 6.7: Maps of distribution of the amplitude value in the responses to a harmonic pumping signal with a 5 min period at different locations: in the matrix near a conduit (P4), in the matrix (P7), in a large conduit (P6), in a thin conduit (P3). 185

Figure 6.8: Maps of distribution of the phase offset value in the responses to a harmonic pumping signal with a 5 min period at different locations: in the matrix near a conduit (P4), in the matrix (P7), in a large conduit (P6), in a thin conduit (P3). 186

Figure 6.9: Comparative maps of distribution of the amplitude and absolute phase offset values in the responses to a harmonic pumping at two different locations (in the matrix near a conduit (P4), in a large conduit (P6)) for a 5 min period (left) and 1 min period (right) signal. 188

Figure 6.10: Boreholes locations on the Terrieu site. The colors for P2, P9, P10 and P15 refer to the colors used to designate these boreholes in Figure 6.11. The blue line indicates a conduit connectivity assessed from previous investigations (Dausse 2015; Wang et al. 2016). The boreholes in light grey were not measured during the harmonic pumping test. 189

Figure 6.11: Example of different type of responses registered during the 5 min period harmonic pumping test in P15 on the Terrieu site. The top graph shows the complete responses and the bottom graph shows the purely oscillatory responses after having subtracted the linear signal. 191

Figure 6.12: Registered oscillatory responses for each measurement borehole compared to the T= 5 min period pumping borehole signal (full lines) and the

interpreted signals for an equation form of Equation (6.10) with variables amplitude and phase offset values (dotted lines). 192

Figure 6.13: Example of a possible conduits network (inside the zone delineated by violet dotted boundaries) interpreted from the boreholes connectivity by applying the same analysis than in the synthetic case. The captions represent the relative amplitude (in blue, in %) and relative phase offset (in orange, in °) values in the oscillatory responses in each measured borehole. A dash represents an absence of oscillatory response (< 1 mm). The pumping location is indicated by 'P'. The blue line indicates a conduit connectivity known from previous investigations (Jazayeri Noushabadi 2009; Dausse 2015). 193

Figure 7.1: Maps of localization of the Terrieu site in France (left) and well pattern on the site (right). Boreholes used as pumping and measurement points are indicated using red triangles, and boreholes used only as measurement points are indicated using grey circles. Boreholes indicated by solid black points were not used during the investigation. The blue dotted line delineates a preferential flow path identified by previous studies (Jazayeri Noushabadi 2009 and Dausse 2015), which shows a connectivity between P2, P8, P11, P12, P15 and P20. 208

Figure 7.2: Left: Measured drawdown curves for a selection of boreholes (P2, P10, P11, P15) during a pumping in P15 with a 2 min and a 5 min period. Right: Zoom-in view of three oscillation cycles after removing the linear part from the drawdown curves. 210

Figure 7.3: Zoom-in on the oscillatory responses extracted from the drawdown measured in P2, P10, P11 and P15 during pumping tests in P15 with a 2 min (left) and a 5 min (right) signal periods and FFT results of the interpreted amplitude (Amp.) and phase offset (P.-O.) responses. Solid lines represent the measured signals, dotted lines represent the interpreted signals ($h_{osc.}$ in Equation (7.2)) reconstructed from the amplitudes and phase offsets interpreted by FFT. For interpreted amplitudes smaller than 1 mm (for example here in P10), we considered the oscillatory responses to be negligible. The blue lines represent the interpreted pumping signals (P15) and are presented for each borehole for a better visualization of the interpreted phase offset responses. 212

Figure 7.4: Connectivity maps interpreted from the amplitude (in blue) and phase offset (in orange) responses to a pumping in P15 with a 2 min (left) and a 5 min (right) period of signal. The areas within the dotted lines delineate a possible area where boreholes are connected through a direct conduit connectivity. Dashes indicate negligible oscillatory responses. 213

Figure 7.5: Schema of the parameterization of a model with the CADI method. \mathbf{P}_N contains the encoded (see Encoding) structural directions of generation associated to each subspace which permits to generate, from an initial 'conduit' cell, a network of conduits in the matrix. \mathbf{P}_β contains the conduit (C) and matrix (M) transmissivity and storativity values associated to each subspace. $\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)$ designates the model produced by applying the property values from \mathbf{P}_β to the network generated from \mathbf{P}_N . 218

Figure 7.6: Schematization of the complete multi-scale inversion process. Starting from an initial model, firsts inversions were led for a 6×4 partitioning (shown by the grid). The results were refined to 12×8 subspaces and used for new inversions. Finally, joint inversion were led starting from the results of the previous separate inversion. 224

Figure 7.7: Comparison of some measured and simulated (with the property distributions presented in Figure 7.9) responses signals in observation points P2 (green), P10 (orange), P11 (red) and pumping points P3, P9, P15, P20 (each time in blue), for pumping signals with a 2 min (left) and a 5 min (right) period. In the case of the pumping in P3 we present in blue the signal in P0, located 1 m away from P3 (which was not measured during the investigation). For a better readability the responses are presented separately for a pumping in P15 with their amplitude (A. in cm) and their phase offset (P. in °) values. For the pumping in P3, P9 and P20 the responses are presented on a same graph.	226
Figure 7.8: Maps of simulated spatial amplitude (Amp.) and phase offset (P.-O.) with the models in Figure 7.9 for a pumping in P15 with a signal period of 2 min and 5 min.	227
Figure 7.9: Maps of the distributions of transmissivity (T) and storativity (S) found by separate inversions of the responses to periods of 2 min and 5 min.	229
Figure 7.10: Maps of the distributions of transmissivity found by inversions of the responses to the 2 min and 5 min periods, and joint inversions started with the 2 min result (2 min (+5 min)), and with the 5 min result (5 min (+2 min)).	230
Figure 7.11: Maps of the connectivity responses associated to each borehole from the networks (shown in background in black) inverted with the joints inversions. Boreholes in blue are associated to a conduit connectivity, in orange to a dual connectivity, and in red to a matrix connectivity. The red lines show flow paths in the models which show a same connectivity as the field preferential flow path highlighted in Jazayeri Noushabadi (2009) and Dausse (2015).	233
Figure 7.12: Structural uncertainty values from the results found for separate inversions of the 2 min and 5 min responses, and joint inversions started with the 2 min result (2 min (+5 min)), and with the 5 min result (5 min (+2 min)).	234
Figure 7.13: Transmissivity (T) and storativity (S) standard deviation values of the results found for separate inversions of the 2 min and 5 min responses, and joint inversions started with the 2 min result (2 min (+5 min)), and with the 5 min result (5 min (+2 min)).	236
Figure 8.1: Résultats d'imageries des transmissivités obtenus sur le site du Terrieu à partir d'un jeu de réponses en domaine permanent de pompages à débit constants (2009), et d'un jeu de réponses en domaine fréquentiel de pompages à débits harmoniques (2017) avec deux périodes (Per.) différentes. Ces résultats ont été obtenus à partir de différentes méthodes d'inversions, notamment pour les données de pompages à débits constants.	244
Figure 8.2: Interprétation des résultats de tomographie harmonique sur le site du Terrieu en terme de connectivité des forages au réseau de conduits principaux grâce aux pompages à « hautes fréquences », et en terme de localisation des structures karstiques (conduits, fractures, fissures) grâce aux pompages à « basses fréquences ».	247

LISTE DES TABLEAUX (TABLES)

Tableau 1.1: Méthodes « classiques » d'investigation (Butler 2005 ; Bechtel et al. 2007 ; Hartmann et al. 2014). Le tableau présente le type de sollicitation et les réponses communément associées, ainsi que l'information qui peut en être retirée.	28
Tableau 1.2: Exemples de contraintes pouvant être appliquées sur des inversions afin de produire des distributions de propriétés adaptées aux connaissances des milieux.	31
Table 2.1: Values of variables used to perform the PCGA inversion on a model of the site for 25,600 parameters and 20 observed data. Results of this inversion are shown in Figures 2.4 and 2.5.	52
Table 2.2: Comparison of the efficiency between three algorithm of geostatistical inversion methods (GA adjoint-state, GA finite-difference and PCGA) on a same under-determined modeling. Results of these inversions are shown in Figures 2.8 and 2.9. The convergence on data was slightly better for an adjoint-state method but the calculation time was considerably reduced by using a PCGA method.	60
Table 2.3: Convergence times for different methods using different grid sizes. An Intel Xeon QuadCore 2.8GHz with 12Go RAM has been used to perform the computations. The PCGA method (with a truncation error of approximately 1) is always the fastest because it involves less forward problems than the GA finite-difference method and that the Gauss-Legendre resolution of the integral in the GA adjoint-state method requires a calculation of a number of nodes proportional to the number of cells in the grid in each forward problem.	60
Table 3.1: Inversion results obtained for the 6 different study cases. This table includes the inversion type, the number of cells of the model, the partitioning and the observed data considered in the inverse modeling, the error variance of data, the number of iteration necessary to the convergence of the inversion process, the proximity between inverted data and observed data (R^2) and between the inverted structure and the true one pixel wise (structural similarity), and the inversion time. In Case 3, Nolnit=Initial simple model and Init=Initial more complex model. In Case 6, NL=Non-linear and L=Linear.	85
Table 3.2: This table presents the main advantages provided by the CADl algorithm. The limits of the methods are also listed with a suggested solution for each limit.	100
Table 4.1: Parameters used in the inversion study cases.	122
Table 5.1: List of the inversion parameter values chosen for the final inversion (16×24 subspaces).	151
Table 6.1: Coordinates, position and pumping signal parameters for the eight boreholes. For the positioning M=Matrix, TC=Thin Conduit, LC=Large Conduit and NC=Near Conduit. The pumping signal parameters are the amplitude (Q_A) and the mean flow rate (Q_m) (see Equation (6.1)).	176

Table 6.2: Table of the relative amplitude and phase offset values in the oscillatory responses of P1 (1 m away from the network) and P4 (50 cm away from the network) to harmonic pumping in P3 and P6 and for increasing signal periods. In this table Amp.=Amplitude, P.O.=Phase Offset, TC=Thin Conduit, LC=Large Conduit, NTC=Near Thin Conduit and NLC=Near Large Conduit. Values in parentheses signify phase offsets greater than one cycle ($>360^\circ$).	184
Table 7.1: Harmonic pumping rates registered for each pumping point during the investigation. Q_A and Q_m refer to Equation (7.1).	209
Table 7.2: Parameters used for the inversion process.	224
Table 7.3: RMSEs on the amplitude (Amp.) and phase offset (P.-O.) values for the different inversion results. RMSEs values in brackets represent responses that were simulated through models generated for another period of signal (i.e. 5 min responses simulated with a model generated specifically for the 2 min responses and vice versa).	231
Table 7.4: Positioning or connectivity response of each borehole interpreted from the qualitative estimations (Figure 7.4), the separate inversion results (Figure 7.9), and the joint inversion results (Figure 7.10).	232
Tableau 8.1: Principaux avantages et limites des méthodes CADI et DNDI.	245

1 INTRODUCTION GÉNÉRALE

A l'échelle du globe, il est estimé que uniquement 2,5 % de l'ensemble de l'eau présente sur Terre est de l'eau douce. Les eaux souterraines représentent près du tiers (30,1 %) de cette eau douce, bien plus que les eaux de surface (1,3 %). La plus grande partie de l'eau douce est contenue dans les glaciers (*Shiklomanov 1993*). La pérennisation du bon état quantitatif et qualitatif des masses d'eaux douces exploitables (eau souterraine et eau de surface) est un enjeu crucial d'un point de vue sanitaire, économique et écologique (et parfois même géopolitique). En France, l'état des lieux des eaux souterraines mené en 2013 dans le cadre de la Directive Cadre Européenne sur l'eau de 2000 montre la dégradation de celles-ci. Ainsi, si seules 9,4 % des masses d'eaux souterraines du territoire français sont en mauvais état quantitatif, elles sont pour 32,8 % d'entre elles en mauvais état chimique (*Agences et offices de l'Eau, Onema, Ministère en charge de l'environnement 2013*). Une meilleure gestion locale de ces masses d'eaux, notamment vis-à-vis de leur vulnérabilité à la pollution, requerrait une connaissance plus approfondie des aquifères dans lesquels ces eaux circulent.

La présence de grands hydrosystèmes souterrains concernent une grande partie des sous-sols français (Figure 1.1). Ces aquifères sont globalement catégorisés en trois grands types structurels (*Heath 1998*) : les milieux poreux dans des sols non consolidés ou des roches poreuses, les milieux fracturés dans des roches consolidées ayant subi une fracturation et les milieux karstiques dans des roches carbonatées ayant subi une dissolution par action d'eau acide (Figure 1.2).

Ces trois types d'aquifères sont caractérisés par différentes structures géologiques qui génèrent des écoulements de plutôt diffus et lents au travers des pores en milieu poreux, à plutôt rapides et préférentiellement localisés dans les conduits en milieu karstique. De part ces différences de fonctionnements hydrauliques et hydrologiques, les études des aquifères requièrent de bien intégrer également leurs particularités morphologiques.

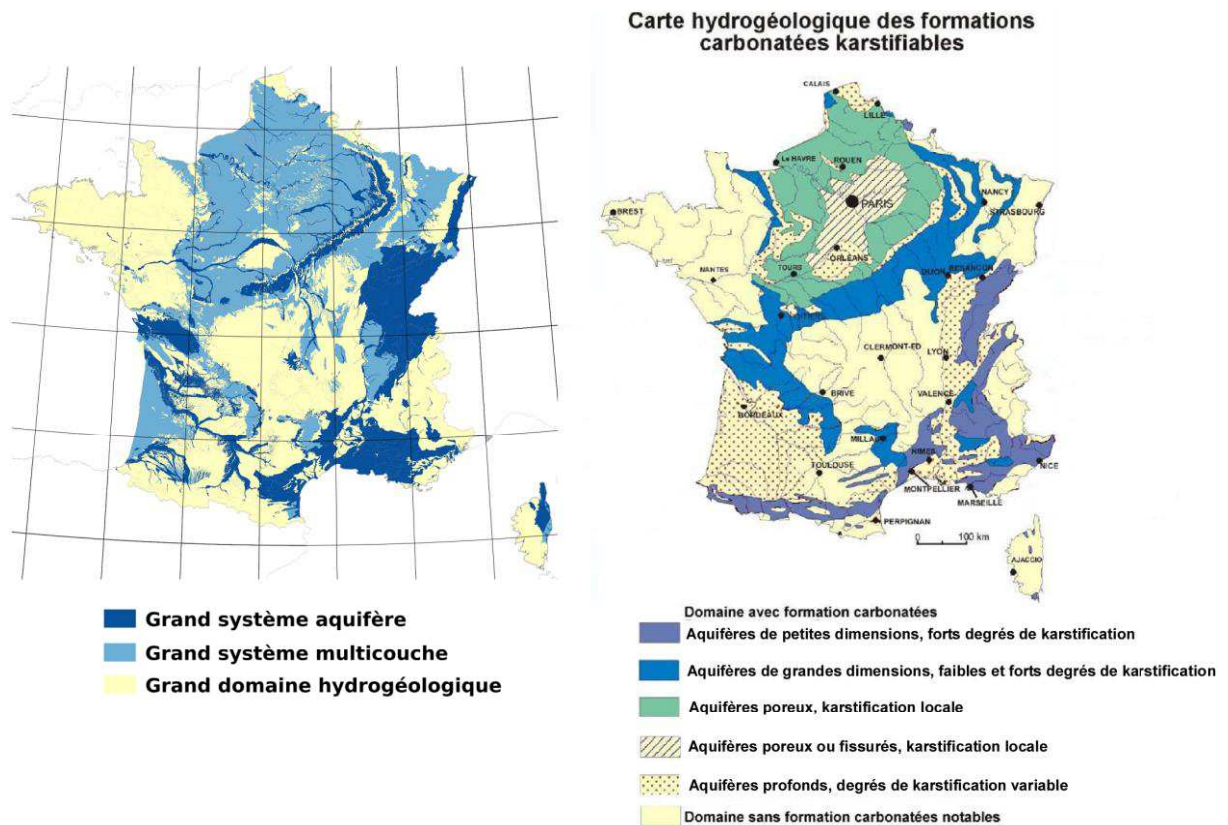


Figure 1.1: Gauche : Répartition des grands hydrosystèmes souterrains français (BRGM 2015). Droite : Répartition des formations karstifiables en France (Marsaud 1996).

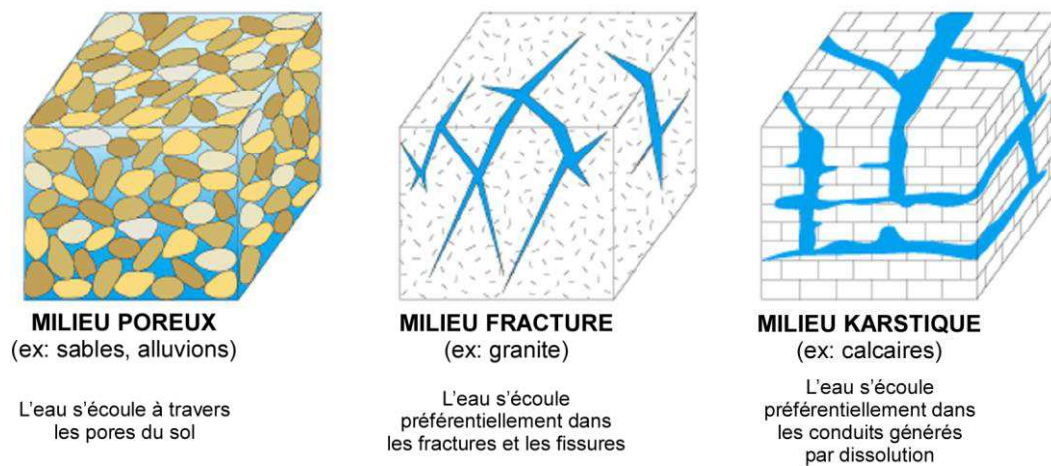


Figure 1.2: Schématisation des écoulements dans trois types de milieux pouvant être aquifère : milieux poreux, fracturés et karstiques (modifié d'après Heath 1998).

L'étude du fonctionnement hydrodynamique d'un aquifère repose sur la caractérisation de ses propriétés hydrauliques telles que la conductivité hydraulique et le coefficient d'emménagement. Ces propriétés contrôlent les modalités de transferts hydriques et des contaminants. Du fait de l'impossibilité de pouvoir accéder directement aux propriétés du milieu aquifère, excepté par des analyses d'échantillons de roches issus de forages qui ne fournissent que des informations locales, il convient d'estimer indirectement ces propriétés. Il est ainsi possible de les estimer, à partir d'investigations in situ, en interprétant les réponses d'un aquifère à divers types de sollicitations (pompage, injection traçage,...), puisque ces réponses dépendent de ces propriétés hydrodynamiques et reflètent leur hétérogénéité dans le milieu (Figure 1.3).

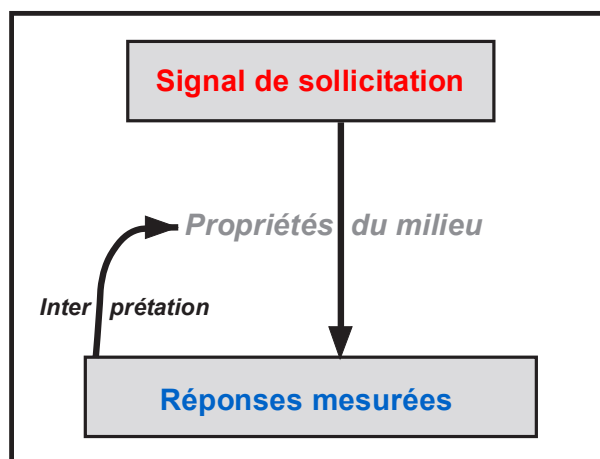


Figure 1.3: Schéma de la caractérisation des propriétés par interprétation des réponses à un signal de sollicitation.

Le choix du signal de sollicitation utilisé, de la quantité et de la répartition de réponses mesurées, et de la méthode d'interprétation de ces dernières sont autant de variables qui permettent d'adapter la caractérisation au type de milieu investigué et au type d'information qui est recherché (Butler 2005). Les types de sollicitation et de réponses mesurées sont liés par leurs interactions via les propriétés du sol. Il est ainsi possible d'investiguer un aquifère par pompage dans la nappe et de mesurer les réponses de rabattements (Figure 1.4) afin d'en estimer les conductivités hydrauliques.

Les approches d'investigation sont nombreuses et en constante évolution. Les méthodes « classiques » d'investigation d'aquifère sont présentées dans le Tableau 1.1.

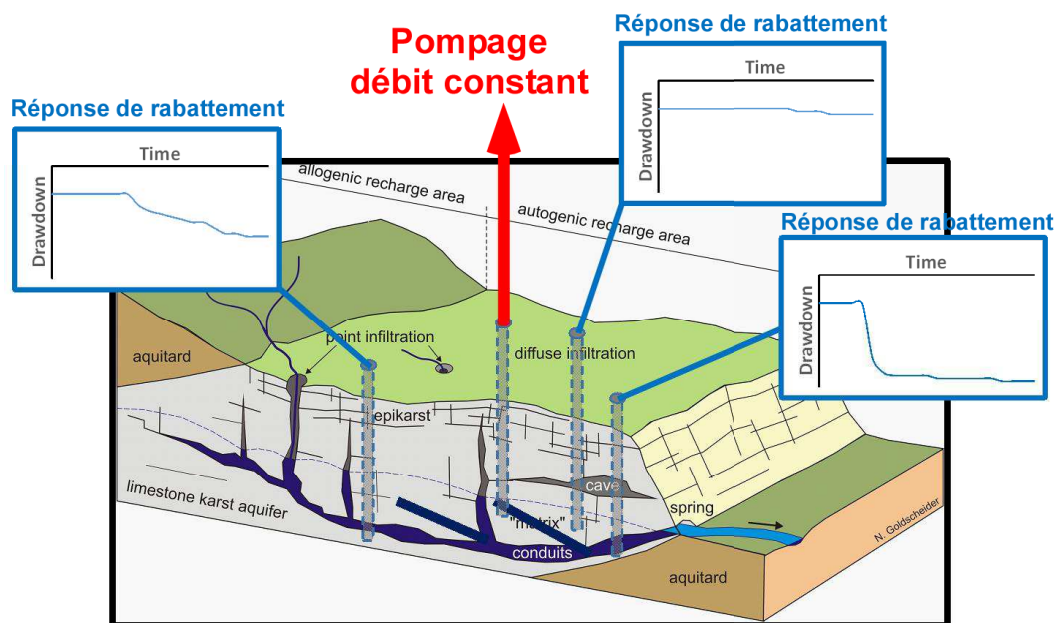


Figure 1.4: Représentation schématisée d'une investigation par pompage avec débit à signal constant d'un aquifère fracturé et karstique (modifié d'après Goldscheider et Drew 2007).

Tableau 1.1: Méthodes « classiques » d'investigation (Butler 2005 ; Bechtel et al. 2007 ; Hartmann et al. 2014). Le tableau présente le type de sollicitation et les réponses communément associées, ainsi que l'information qui peut en être retirée.

Sollicitation	Réponses mesurées	Informations interprétables
Pompage débit constant	Rabattement	Conductivités hydrauliques moyennées du volume impacté
Slug test	Rabattement	Conductivités hydrauliques locales autour du forage
Traçage	Concentrations	Connectivités et vitesses d'écoulements
Pluie	Variations de charges, débit, concentrations	Réactivité physico-chimique du système, types d'écoulements
Courant électrique	Différence de potentiel	Résistivités, saturation des matériaux
Onde sismique	Vitesse d'onde	Structures géologiques sous-terraines
Champ gravimétrique	Micro-variation du champ gravitationnel	Localisation de vides
Electromagnétisme – Radar pénétrant	Variation de champ magnétique	Identification d'anomalies dans le sol
Potentiel spontané	Potentiel électrique	Identification et localisations des mouvements d'eaux

Il existe différentes méthodes pour interpréter quantitativement les propriétés d'un aquifère à partir des réponses mesurées. Globalement ces méthodes peuvent être classées en deux grandes catégories : les solutions analytiques et les solutions numériques. Les solutions analytiques sont basées sur des expressions mathématiques sous forme analytique, c'est-à-dire construites à partir des opérations arithmétiques basiques, qui se prêtent aisément au calcul. Ces solutions analytiques excluent néanmoins le caractère hétérogène des propriétés hydrauliques. Les solutions numériques, au contraire, permettent de pouvoir modéliser les réponses à des équations plus complexes par des méthodes de résolution des opérations différentielles (parmi lesquelles la méthode de différence finie, volume fini ou encore élément fini) en prenant en compte la variabilité spatiale des propriétés hydrodynamiques. Ainsi, un modèle distribué, basé sur une solution numérique, permet de simuler spatialement les réponses d'un aquifère à partir d'un champ de propriétés, en discrétisant et résolvant les équations d'écoulements sur un maillage. L'interprétation des réponses mesurées sur le terrain consiste dans ce cas à retrouver un champ de propriétés qui serait capable de reproduire, par simulation dans le modèle, les données de terrain.

La caractérisation numérique distribuée des aquifères s'est orientée vers des techniques de tomographies hydrauliques (*Carrera et Neuman 1986b ; De Marsily et al. 1995 ; Yeh et Liu 2000 ; Delay et al. 2007*). A l'instar des tomographies utilisées dans le domaine médical, elles permettent d'imager, en plan ou en volume, les propriétés d'un milieu à partir de mesures ponctuelles de surface. La tomographie hydraulique repose d'abord sur la définition d'un problème direct, décrivant les lois associées aux phénomènes physiques qui sont employées pour la résolution du modèle. Puis, le problème inverse, basé sur le problème direct, optimise le champ de propriétés du modèle afin de reproduire les réponses mesurées lors de l'investigation. Schématiquement, le problème direct définit le lien entre les réponses au signal de sollicitation, alors que le problème inverse permet de reconstruire les propriétés du milieu dans le modèle, correspondant à l'interprétation des réponses recherchées (Figure 1.5). La tomographie hydraulique repose également sur la qualité et la quantité des données (réponses de l'aquifère), qui doivent permettre de bien caractériser le champ de propriété généré par le problème inverse (*Yeh et Lee 2007*). Si ces conditions sont réunies, la tomographie hydraulique peut alors être vue comme un outil d'imagerie des aquifères, qui peut ensuite servir l'interprétation et la discussion scientifique.

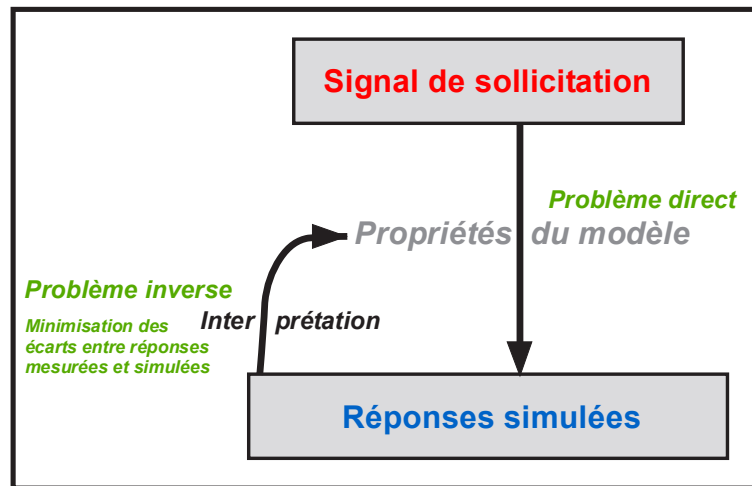


Figure 1.5: Schéma de la reconstitution des propriétés à partir d'un modèle direct et de l'inversion des réponses mesurées.

L'inversion est un processus mathématique, stochastique ou déterministe, qui aboutit à une ou des solutions possibles de champs de propriétés minimisant l'écart entre les données simulées et observées (*Tarantola et Valette 1982*). Le processus d'inversion stochastique consiste à générer itérativement un grand nombre de modèles de propriétés différents et d'en garder au final les meilleurs, qui approchent des solutions globales au problème (en s'appuyant sur la loi des grands nombres). Le processus déterministe se base plutôt sur une optimisation itérative d'un modèle initial, basée sur une étude de sensibilité, et convergeant au final vers un modèle de propriétés approchant une solution locale au problème (dépendante du modèle initial).

Néanmoins, dans les deux cas, ce processus purement mathématique est mal posé, avec une solution qui n'est pas unique, et peut donc aboutir à des résultats reproduisant les données de terrain, mais irréalistes physiquement ou structurellement. Il devient intéressant de pouvoir contraindre l'inversion à partir des connaissances a priori sur les propriétés du milieu, afin de limiter ses résultats à des solutions réalistes vis-à-vis du milieu. Ainsi de nombreux travaux ont visé à développer des manières de contraindre l'inversion, dont quelques-unes sont présentées dans le Tableau 1.2.

Tableau 1.2: Exemples de contraintes pouvant être appliquées sur des inversions afin de produire des distributions de propriétés adaptées aux connaissances des milieux.

Type de contrainte	Type de variation de distribution de propriétés	Exemple de référence
Géostatistique	Lisse	Geostatistical approach (Hoeksema et Kitanidis 1984)
Mathématique (loi de distribution)	Hétérogène	Total variation prior (Lee et Kitanidis 2013)
Structuration du modèle direct	Structurée	Level set (Lu et Robinson 2006)
Guidage par image	Structurée	Training image (Lochbühler et al. 2015) Image guided (Soueid Ahmed et al. 2015)
Processus physique	Structurée	Karstic network formation (Jaquet et al. 2004)
Résolution du champ de propriétés	Hétérogène	Multi-scale (Grimstadt et al. 2003)

Les contraintes dans l'inversion visent à orienter les solutions vers une certaine forme de distribution spatiale des propriétés. Les solutions proposant des distributions trop éloignées des règles établies par les contraintes, sont écartées. Les contraintes peuvent être vues, schématiquement, à un contrepoids informatif à la reproduction purement mathématique des réponses, tel que représenté en Figure 1.6.

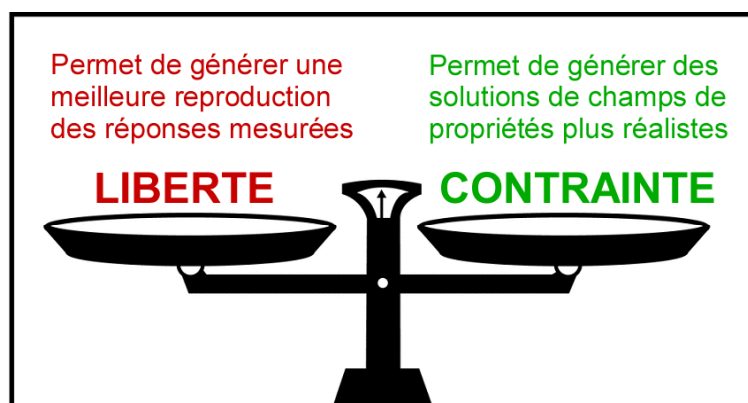


Figure 1.6: Schéma entre la liberté attribuée dans la reconstitution des réponses dans le processus d'inversion par rapport aux contraintes établies sur les solutions.

Les aquifères fracturés et karstiques génèrent des écoulements suivant des chemins préférentiels : via le réseau de fissures et fractures de la roche dans le premier cas, et via le réseau de conduits formés par dissolution de la matrice calcaire dans le second cas. Ces aquifères requièrent donc une attention toute particulière lors de l'investigation et dans l'interprétation numérique afin de caractériser le haut degré de contraste spatial existant dans leurs propriétés hydrodynamiques et qui provoque ces écoulements préférentiels.

De nombreux travaux de caractérisation d'écoulements se sont déjà intéressés à imager ces contrastes dans des modèles équivalents milieux poreux (*Larocque et al. 1999 ; Abusaada et Sauter 2013 ; Saller et al. 2013 ; Wang et al. 2016*), dans des modèles à double continuum (*Zimmerman et al. 2013 ; Kordilla et al. 2012*), ou dans des modèles couplés discret-continu (*Kovacs 2003 ; Jaquet et al. 2004 ; Saller et al. 2013*). Des exemples de ces modèles sont présentés en Figure 1.7.

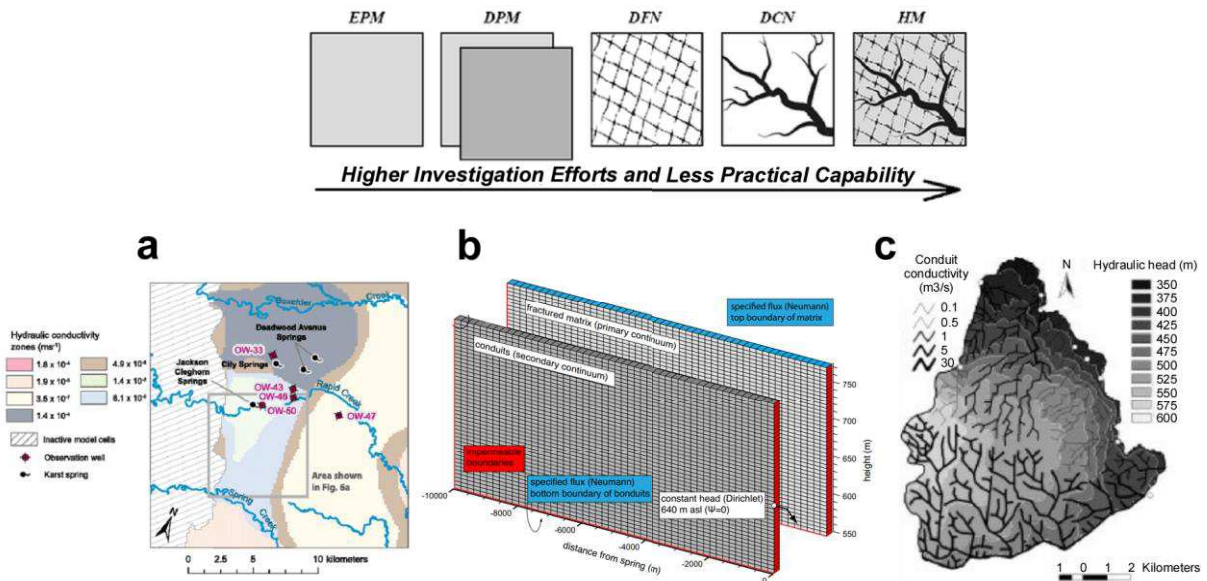


Figure 1.7: Haut : Schéma des différentes techniques de modélisation distribuées issu de Ghasemizadeh et al. 2012 (EPM = Equivalent Milieu Poreux ; DPM = Milieu Double Porosité ; DFN = Réseau de Fractures Discrètes ; DCN = Réseau de Conduits Discrets ; HM = Modèle Hybride). Bas : Exemples de modélisations d'aquifères karstiques par (a) EPM (Saller et al. 2013), (b) DPM (Kordilla et al. 2012), et (c) HM (Kovacs 2003).

Néanmoins, il apparaît que la caractérisation spatiale des propriétés des milieux fracturés et karstiques requiert toujours le développement de nouvelles méthodes d'investigation et de modélisation qui permettront de mieux caractériser leurs écoulements très

contraints et d'apporter, pour ces milieux, des outils d'imagerie hydraulique pour une meilleure compréhension de leurs fonctionnements hydrodynamiques.

Les travaux menés dans le cadre de cette thèse visent à répondre à ce verrou en proposant de nouvelles approches tomographiques de caractérisation adaptées aux aquifères fracturés et karstiques.

Ces approches tomographiques reposent, d'une part, sur l'acquisition de données lors d'une phase d'investigation basée sur une méthode particulière de sollicitation de la nappe, par pompage harmonique, et de traitement et d'analyse des réponses oscillatoires du niveau hydraulique. L'utilisation de pompes harmoniques pour la caractérisation des milieux fracturés a été récemment étudiée par *Renner et Messar 2006* et *Guiltinan et Becker 2015*. Leurs résultats montrent que cette technique d'investigation peut également être prometteuse pour les milieux karstiques, pour lesquelles elle n'a pas encore été particulièrement étudiée.

Ces approches tomographiques sont basées, d'autre part, sur une interprétation quantitative en utilisant des concepts de modélisation numérique des écoulements souterrains (type équivalent milieu poreux ou couplé discret continu) et de nouvelles méthodes d'inversions adaptées afin de représenter plus fidèlement les contrastes existant dans ces milieux en déterminant l'architecture spatiale des conduits karstiques.

Tout au long du manuscrit les résultats produits par applications de ces différentes nouvelles approches servent à alimenter une discussion sur les écoulements et échanges existant dans les aquifères fracturés et karstiques à l'échelle décamétrique.

Ce manuscrit est construit autour de six articles produits au cours de cette thèse :

- Chapitre 2 : Application of large-scale inversion algorithms to hydraulic tomography in an alluvial aquifer, *Groundwater*, 2017.
- Chapitre 3 : A cellular automata-based deterministic inversion algorithm for the characterization of linear structural heterogeneities, *Water Resources Research*, 2017.

- Chapitre 4 : Hydraulic tomography of discrete networks of conduits and fractures in a karstic aquifer by using a deterministic inversion algorithm, *Advances in Water Resources*, 2018.
- Chapitre 5 : Identifying flow networks in a karstified aquifer by application of the cellular automata-based deterministic inversion method (Lez aquifer, France), *Water Resources Research*, 2017.
- Chapitre 6 : Hydraulic analysis of harmonic pumping tests in frequency and time domains for identifying the conduits networks in a karstic aquifer, *Journal of Hydrology*, 2018.
- Chapitre 7 : Harmonic pumping tomography applied to image the hydraulic properties and interpret the connectivity of a karstic and fractured aquifer (Lez aquifer, France), *Advances in Water Resources*, 2018.

Dans un premier temps, dans le chapitre 2, une méthode de tomographie pour la caractérisation d'un milieu poreux, allant de l'investigation à la modélisation et l'imagerie, est présentée en discutant sur la résolution du problème inverse et la contrainte géostatistique qui est choisie dans ce cas.

Les chapitres 3 et 4 présentent deux nouvelles méthodes d'inversions contraintes structurellement, qui sont plus adaptées à l'imagerie de milieux à écoulements localisés. La première méthode, la méthode CADI, se base sur la génération d'un réseau d'écoulement dans un milieu équivalent à un milieu poreux. La deuxième méthode, la méthode DNDI, assimile les réseaux d'écoulements préférentiels à des lignes discrétisées, couplées à un milieu poreux représentant la roche de l'aquifère.

Une application de la méthode CADI pour une tomographie du site d'étude de milieux karstiques du Terrieu est présentée dans le chapitre 5. La tomographie consiste dans ce cas en plusieurs pompages à débit constant alternés dans différents forages. Les résultats d'imagerie produits par cette méthode sont comparés aux connaissances sur les écoulements de ce site pour discuter de la pertinence et de la précision des résultats.

Le chapitre 6 a pour but d'étudier les réponses théoriques d'un milieu à écoulement préférentiel aux signaux oscillatoires à partir d'un cas synthétique, basé sur un modèle de réseau karstique simple. Ce cas synthétique permet de mettre en place un traitement et une interprétation des réponses nécessaires à l'application sur le site d'étude du Terrieu qui est présentée en chapitre 7. Ce dernier chapitre s'appuie sur les méthodes nouvellement proposées, la tomographie harmonique et l'interprétation CADI, pour caractériser le site du Terrieu et les connectivités hydrauliques y existant à l'aide des résultats d'imagerie. Ces résultats permettent également de discuter le rôle de la période d'oscillation du signal de pompage pour la caractérisation des différents chemins d'écoulements existant dans les milieux fracturés et karstiques.

2 TOMOGRAPHIE HYDRAULIQUE EN MILIEU ALLUVIAL AVEC INVERSION GÉOSTATISTIQUE

2.1 Contexte

Ce chapitre présente, et compare entre eux, différents résultats de tomographies pour imager les propriétés d'un milieu peu hétérogène au travers d'un cas d'application dans un aquifère alluvial.

Le site d'étude investigué, le site expérimental de la Céréirède, est localisé dans le sud de la France, près de Montpellier, et proche de la rivière du Lèz. La tomographie réalisée sur ce site se base sur un jeu de réponses de rabattements à deux différents pompages à débit constant.

La majeure partie de ce chapitre se focalise plus spécifiquement sur le problème inverse, en en présentant la théorie et les équations dans un premier temps. Des inversions déterministes sont réalisées avec des contraintes géostatistiques, particulièrement pertinentes et communément utilisées dans ce genre de cas d'étude. Trois méthodes d'inversions géostatistiques sont comparées à travers ce cas : une méthode géostatistique à résolution en états adjoints, une méthode géostatistique à résolution en différence finie et une méthode géostatistique en analyse en composantes principales. Cette dernière méthode permet une résolution plus rapide pour les milieux aux distributions de propriétés très « lissées », en réduisant les tailles des matrices de covariances du problème inverse pour n'en conserver que les composantes principales.

Les résultats d'inversions permettent une bonne reproduction des réponses, ainsi qu'une image réaliste des propriétés hydrodynamiques (conductivité) pour ce type de milieu. Les différentes méthodes d'inversions utilisées aboutissent à des images de propriétés semblables. Elles diffèrent principalement par leur efficacité en temps de calculs, la méthode la plus efficace étant celle en analyse en composantes principales qui permet de diviser le temps de calcul par 10 par rapport à une résolution avec les états adjoints et par 70 par rapport à une résolution avec différence finie.

Néanmoins si la méthode géostatistique en analyses en composantes principales, qui repose sur une approximation lissant la distribution des propriétés dans le modèle, peut être efficace dans les cas de milieux peu hétérogènes dans lesquels elle génère une erreur minime, elle devient inefficace pour les milieux hétérogènes aux propriétés contrastées. Donc, cette méthode d'inversion est particulièrement adaptée aux aquifères alluviaux par exemple, mais ne

pourrait pas du tout être utilisée pour des aquifères fracturés et karstiques. Une caractérisation d'aquifère réalisée par tomographie nécessite non seulement le jeu de données permettant cette caractérisation, mais aussi le choix d'une méthode d'inversion adaptée au milieu, notamment concernant les contraintes. L'utilisation des méthodes à résolution par états adjoints et différences finies est envisageable en aquifère karstique, mais sans assurance d'aboutir à une distribution réaliste des propriétés, très contrastée dans ces milieux. De ce fait, dans le cas des aquifères karstiques et fracturés, des méthodes d'inversions plus adaptées que celles décrites dans ce chapitre seront présentées dans les autres chapitres de ce manuscrit et seront nécessaires afin d'aboutir à des représentations réalistes des champs de propriétés pour ces milieux.

2.2 Application of large-scale inversion algorithms to hydraulic tomography in an alluvial aquifer

Cette partie est composée de l'article « Application of large-scale inversion algorithms to hydraulic tomography in an alluvial aquifer », publié dans le journal *Groundwater* en mars 2017 (Fischer et al. 2017a). Le texte a été remis en forme au format du manuscrit. La version originale de l'article est donnée en Annexe 1.

Groundwater

Vol. 55, No. 2 – Groundwater – March–April 2017 (pages 208–218)
doi: 10.1111/gwat.12457

Application of Large-Scale Inversion Algorithms to Hydraulic Tomography in an Alluvial Aquifer

by P. Fischer¹, A. Jardani², A. Soueid Ahmed², M. Abbas², X. Wang³, H. Jourde⁴, and N. Lecoq²

¹Corresponding author: UNIROUEN, UNICAEN, CNRS, M2C, Normandie University, 76000 Rouen, France

²UNIROUEN, UNICAEN, CNRS, M2C, Normandie University, 76000 Rouen, France.

³Department of Earth Science and Engineering, Imperial College London, London SW7 2AZ, UK.

⁴CNRS, Laboratoire Hydrosiences, Université de Montpellier, 34000 Montpellier, France.

2.2.1 Introduction

In hydrogeology, the assessment of hydraulic properties of subsurface aquifers, such as transmissivity, storage coefficient and solute transport parameters, is a key issue to an adequate management and protection of groundwater resources. Generally, aquifer characterization is based on the interpretation of hydraulic observations data collected during pumping, infiltration, or tracer tests (*Carrera and Neuman 1986b ; Rao et al. 2003 ; Lee and Kitanidis 2014 ; Pool et al. 2015*). Therefore, hydraulic tomography is considered as one of the most effective approaches for characterizing the spatial distribution of hydraulic transmissivity of an aquifer (*Cardiff et al. 2009 ; Berg and Illman 2013 ; Cardiff et al. 2013a ; Soueid Ahmed et al. 2015 ; Zha et al. 2015 ; Wang et al. 2016*). This method relies on a set of hydraulic head responses recorded during cross-hole pumping experiments. Then, the interpretation can be achieved through the use of an inverse algorithm to image the spatially varying hydraulic properties in the subsurface.

The inverse problem for estimating hydraulic parameters involves a formulation of a forward problem, which sets up the link between the hydraulic observations and the unknown hydraulic parameters (*Tarantola and Valette 1982*). For a hydraulic tomography inversion, the forward modeling is based on a numerical method (e.g. the finite element, finite difference and finite volume methods) used to solve the partial differential equation of the groundwater flow. The forward problem operator can be formulated as:

$$\mathbf{d} = f(\mathbf{s}) + \boldsymbol{\eta} , \quad (2.1)$$

where \mathbf{d} represents the hydraulic responses of the model, \mathbf{s} is the logarithm of the m unknown hydraulic transmissivities, to be estimated from a set of n observed data \mathbf{d}_{obs} and a nonlinear forward modeling application $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$. $\boldsymbol{\eta}$ is an additive noise of the numerical modeling.

In a probabilistic framework, the inverse problem maximizes a posterior probability density function $\pi_{\text{post}}(\mathbf{s}|\mathbf{d})$. Generally, the problem is ill-posed and the solution is non-unique, therefore additional *a priori* information is required in order to find one physically meaningful solution (*Carrera and Neuman 1986a*). Furthermore, in our case, the problem is under-determined, it involves a small number of measurements but a large number of unknown parameters ($n \ll m$). The inverse problem corresponds to recovering the ‘best fitting’ model parameters $\hat{\mathbf{s}}$ which allow the model responses to match the observed data, and at the same

time respect the constraints imposed by the *a priori* information on the model. Using the Bayes' formula, the posterior probability density function can be expressed as (Elsheikh *et al.* 2014):

$$\pi_{\text{post}}(\mathbf{s}|\mathbf{d}) \propto \exp \left[-\frac{1}{2}(\mathbf{f}(\mathbf{s}) - \mathbf{d})^T \mathbf{R}^{-1}(\mathbf{f}(\mathbf{s}) - \mathbf{d}) - \frac{1}{2}(\mathbf{s} - \mathbf{s}_{\text{prior}})^T \mathbf{Q}^{-1}(\mathbf{s} - \mathbf{s}_{\text{prior}}) \right], \quad (2.2)$$

where $\mathbf{s}_{\text{prior}}$ denotes the *a priori* model, \mathbf{Q} is a $m \times m$ covariance matrix of the model parameters, which can include geostatistical information about the distribution trend or pattern of the unknowns parameters (in that case \mathbf{Q} is defined in a matrix where elements of the matrix represent the variogram function associated with the distance between cells of the model) and \mathbf{R} denotes a $n \times n$ diagonal covariance matrix which accounts for the errors occurred in the data measurements.

The aim of the inversion problem is to find a set of parameter which maximizes the density of probability, $\pi_{\text{post}}(\mathbf{s}|\mathbf{d})$. This corresponds to a model of high probability with respect to the measurements and the imposed *a priori* model (Tarantola and Valette 1982). To solve the inverse problem, two main groups of iterative methods are often employed: (1) the deterministic methods which assume that the algorithm converges to a local minimum by performing a linearization of an objective function and (2) the stochastic methods which converge to a global minimum by selecting different randomly generated parameters fields as probable solutions to the model (Pool *et al.* 2015). Several deterministic and stochastic inversion algorithms have been widely applied in hydrogeology, but they are time- and memory-expensive, particularly for the cases which involve a fine discretization of the parameter grid and/or a large amount of observational data (Kitanidis and Lee 2014). Thereby, one of the main trends in the development of inversion theory during the last few years was to develop algorithms which are able to solve inversion models with a large number of unknown parameters and data.

Recently, several time- and memory-saving methods have been developed to reduce the memory cost and the computation time of the usually large matrices involved in the inversion algorithms. One way for reducing the computational and memory demands is to use some approximation methods on matrix operations, such as the Fast Multipole Method (FMM) (Greengard and Rokhlin 1987), which is based on Legendre polynomial expansions and spherical harmonics. The FMM was associated with the Hierarchical matrices approach

(*Hackbusch and Börm 2002*) to compute matrix-vector products for a large-scale application in seismic imaging (*Ambikasaran et al. 2013*). Another way is through using the MINRES Krylov subspace method (*Paige and Saunders 1975*) which can be combined with the Fast Fourier Transform (FFT) (*Nowak et al. 2003*) to iteratively solve inversions of large matrix systems. This method has recently been applied to 3-D large-scale transient hydraulic tomography problems (*Liu et al. 2014*). The two approaches mentioned above avoid the calculation of the full Jacobian matrix of the forward model at each iteration. A new method has recently been developed by *Lieberman et al. (2013)*, in an application of an inversion algorithm for a large-scale 3-D transient contaminant transport. The authors used a Proper Orthogonal Decomposition method (POD) to compute a projection basis with the eigenvectors associated to the highest eigenvalues of the Hessian matrix of the forward problem. The inverse problem was then solved in a reduced projected subspace.

In this paper, we have attempted to achieve an efficient site hydraulic characterization (i.e. to obtain high-resolution transmissivity fields at a low computational cost) by applying a recently developed geostatistical inversion method, the Principal Component Geostatistical Approach (PCGA) (*Lee and Kitanidis 2014*). This method can considerably reduce the computation time and the memory cost of inversions by using an approximation of the covariance matrix \mathbf{Q} based on a Singular Value Decomposition method (SVD), and by avoiding the computation of the Jacobian matrix through the use of a matrix-free product based on a finite difference method.

This paper presents an application of the PCGA method, combined with a hydraulic tomography investigation, for a large-scale inverse modeling of the hydraulic transmissivity field of an alluvial aquifer. First, we present the methodology of the Geostatistical Approach (GA) algorithm, and the modifications for large-scale application (PCGA). Then, we describe the hydrogeological background of the experimental field site, from which the hydraulic measurements were taken, and present the numerical model setup. Finally, we show our inversion results. In particular, we compare these results to those obtained by applying the classical GA method with two different methods in Jacobian matrix computation (i.e., an adjoint state and a finite difference method), which do not use a covariance matrix approximation. We have evaluated the computation times, and the sensitivities and accuracies of the inversion results for the three different methodologies. Using a hydraulic tomography field application on a porous aquifer, we show the advantages of the PCGA inversion method for efficient large-scale inverse modeling in hydrogeology.

2.2.2 Principal component geostatistical approach

In the Geostatistical Approach (GA) (*Kitanidis and Vomvoris 1983 ; Hoeksema and Kitanidis 1984 ; Kitanidis 1995*), the prior probability density function of the m model parameters \mathbf{s} is set as a multivariate Gaussian with a mean $E(\mathbf{s})=\mathbf{X}\boldsymbol{\beta}$ where \mathbf{X} is an $m \times p$ known matrix and $\boldsymbol{\beta}$ a $p \times 1$ vector to be determined during the inversion process (generally $p = 1$), and a covariance, $E[(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})^T] = \mathbf{Q}(\boldsymbol{\theta})$.

The posterior probability density function $\Psi = -\ln[\pi(\mathbf{s}|\mathbf{d})]$ (also called in inversion problems the objective function) then becomes:

$$\Psi = \frac{1}{2}(\mathbf{f}(\mathbf{s}) - \mathbf{d})^T \mathbf{R}^{-1}(\mathbf{f}(\mathbf{s}) - \mathbf{d}) + \frac{1}{2}(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})^T \mathbf{Q}^{-1}(\mathbf{s} - \mathbf{X}\boldsymbol{\beta}) . \quad (2.3)$$

The best approximation $\hat{\mathbf{s}}$ for the model parameters, taking into account the *a priori* information and the observed data, can be found as being the model maximizing the density of probability in Equation (2.2), which is also equivalent to minimize the argument of its exponential. Thus, $\hat{\mathbf{s}}$ is found by minimizing the objective function Ψ (Equation (2.3)). This minimization can be achieved by using a Newton linearization iterative approach on \mathbf{s} . The iterative process initializes at a reasonable \mathbf{s}_0 . Then, at iteration step $j+1$, the new value \mathbf{s}_{j+1} is found in the vicinity of the previous model \mathbf{s}_j using a first order Taylor approximation:

$$\mathbf{f}(\mathbf{s}_{j+1}) = \mathbf{f}(\mathbf{s}_j) + \mathbf{F}_j(\mathbf{s}_{j+1} - \mathbf{s}_j) . \quad (2.4)$$

Here \mathbf{F}_j is the $n \times m$ Jacobian matrix of the forward problem \mathbf{f} for \mathbf{s}_j : $\mathbf{F}_j = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{s}} \right|_{\mathbf{s}=\mathbf{s}_j}$.

After some matrix manipulations, the updated solution of the parameters in the iterative process, found by minimizing the objective function, can be written as (*Kitanidis 1995*):

$$\mathbf{s}_{j+1} = \mathbf{X}\boldsymbol{\beta}_j + \mathbf{Q}\mathbf{F}_j^T \boldsymbol{\xi}_j , \quad (2.5)$$

where the $p \times 1$ matrix β_j and the $n \times 1$ matrix ξ_j are solutions of the following matrix system (Kitanidis 1995):

$$\begin{bmatrix} \mathbf{F}_j \mathbf{Q} \mathbf{F}_j^T + \mathbf{R} & \mathbf{F}_j \mathbf{X} \\ (\mathbf{F}_j \mathbf{X})^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \xi_j \\ \beta_j \end{bmatrix} = \begin{bmatrix} \mathbf{d} - f(\mathbf{s}_j) + \mathbf{F}_j \mathbf{s}_j \\ \mathbf{0} \end{bmatrix}. \quad (2.6)$$

Note that here ' $\mathbf{0}$ ' represents a $p \times p$ matrix of zeros on the left-hand side and a $p \times 1$ matrix of zeros on the right-hand side.

At the end of the iterative process, to quantify the model parameter's uncertainty after optimization, we can compute the posterior covariance of \mathbf{s} derived as:

$$\mathbf{Q}_{\text{post}} = \mathbf{Q} - \begin{bmatrix} \mathbf{F}_j \mathbf{Q} \\ \mathbf{X}^T \end{bmatrix}^T \left(\begin{bmatrix} \mathbf{F}_j \mathbf{Q} \mathbf{F}_j^T + \mathbf{R} & \mathbf{F}_j \mathbf{X} \\ (\mathbf{F}_j \mathbf{X})^T & \mathbf{0} \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{F}_j \mathbf{Q} \\ \mathbf{X}^T \end{bmatrix}. \quad (2.7)$$

The GA method as presented above needs the computation of the Jacobian matrix \mathbf{F} for each iteration in order to solve the system Equation (2.6), which can usually be done by solving the forward problem $m + 1$ times using a finite difference method, or $n + 1$ times using an adjoint state method. Even if the adjoint state method may considerably decrease the computation time for under-determined problems (see *Cardiff and Kitanidis 2008* for a comparison of the finite difference and adjoint state method computation times), it is not efficient for large scale problems with a large number of measurements and parameters. Another problem which appears in the GA method is that when the number of data and/or parameters is high, it requires a significant computational power for the calculation and storage of the covariance matrix \mathbf{Q} (which can be alleviated by FFT, H-matrices or FMM). To overcome these difficulties, Kitanidis and Lee have developed the Principal Component Geostatistical Approach (PCGA), on the basis of the GA method.

In this new approach, the computational and memory costs associated with the manipulation of matrix \mathbf{Q} are reduced using a low-rank approximation of a chosen truncation order $K \ll m$ through a Singular Value Decomposition (SVD):

$$\mathbf{Q}_K = \mathbf{U} \mathbf{S} \mathbf{V}^T, \quad (2.8)$$

where \mathbf{S} is a $K \times K$ diagonal matrix containing the singular values of \mathbf{Q} sorted in descending order, \mathbf{U} is a $m \times K$ matrix and \mathbf{V} is an $m \times K$ matrix. As \mathbf{Q} is defined as a symmetric matrix, its SVD simplifies to: $\mathbf{Q}_K = \mathbf{V}\mathbf{S}\mathbf{V}^T$.

This decomposition can also be written as:

$$\mathbf{Q}_K = \sum_{i=1}^K \zeta_i \zeta_i^T \quad \text{with } \zeta_i = \sqrt{\lambda_i} \mathbf{V}_i. \quad (2.9)$$

Here λ_i is the i th singular value (also $\mathbf{S}(i,i)$) and \mathbf{V}_i is the i th column vector of \mathbf{V} associated to λ_i . The error arising from this K -rank decomposition equals to the $K+1$ th singular value (λ_{K+1}) of \mathbf{Q} .

However, this decomposition is a good approximation only for a matrix \mathbf{Q} in which the most of its information is contained in its few highest eigenvalues and eigenvectors, meaning a relatively smooth pattern. One can also use an eigen-decomposition if \mathbf{Q} is a matrix defined by positive eigenvalues or a randomized decomposition approach which is efficient for high-dimensional matrices with $m \sim 1,000,000$ (Halko et al. 2011).

In addition, the PCGA method also avoids the full Jacobian matrix calculation at each iteration. When performing a matrix product, such as $\mathbf{F}_j \mathbf{u}$ (where \mathbf{F}_j is the $n \times m$ Jacobian matrix and \mathbf{u} is a $m \times 1$ vector), instead of computing it directly, the method finds an approximation to its accurate form using a first order Taylor series :

$$f\left(\mathbf{s}_j + \frac{\delta \|\mathbf{s}_j\|}{\|\mathbf{u}\|} \mathbf{u}\right) = f(\mathbf{s}_j) + \frac{\delta \|\mathbf{s}_j\|}{\|\mathbf{u}\|} \mathbf{F}_j \mathbf{u} + \sigma(\delta^2), \quad (2.10)$$

$$\mathbf{F}_j \mathbf{u} \approx \frac{\|\mathbf{u}\|}{\delta \|\mathbf{s}_j\|} \left[f\left(\mathbf{s}_j + \frac{\delta \|\mathbf{s}_j\|}{\|\mathbf{u}\|} \mathbf{u}\right) - f(\mathbf{s}_j) \right], \quad (2.11)$$

where δ is a finite-difference interval and $\|\mathbf{u}\|$ and $\|\mathbf{s}_j\|$ are the norm of the vector \mathbf{u} and \mathbf{s}_j .

Thus, in the matrix system Equation (2.6) of the GA algorithm the calculation of the full Jacobian \mathbf{F} is avoided by approximating the computation of $\mathbf{F}_j \mathbf{s}_j$, $\mathbf{F}_j \mathbf{X}$, $\mathbf{F}_j \mathbf{Q} \mathbf{F}_j^T$ and $\mathbf{Q} \mathbf{F}_j^T$ with (Kitanidis and Lee 2014):

$$\mathbf{F}_j \mathbf{s}_j = \frac{1}{\delta} [f(\mathbf{s}_j + \delta \mathbf{s}_j) - f(\mathbf{s}_j)] , \quad (2.12)$$

$$\mathbf{F}_j \mathbf{X}_i = \frac{\|\mathbf{X}_i\|}{\delta \|\mathbf{s}_j\|} \left[f\left(\mathbf{s}_j + \frac{\delta \|\mathbf{s}_j\|}{\|\mathbf{X}_i\|} \mathbf{X}_i\right) - f(\mathbf{s}_j) \right] \text{ with } \mathbf{X}_i \text{ the } i \text{ th column of } \mathbf{X} , \quad (2.13)$$

$$\mathbf{F}_j \mathbf{Q} \mathbf{F}_j^T \approx \mathbf{F}_j \mathbf{Q}_K \mathbf{F}_j^T = \mathbf{F}_j \sum_{i=1}^K \zeta_i \zeta_i^T \mathbf{F}_j^T = \sum_{i=1}^K (\mathbf{F}_j \zeta_i) (\mathbf{F}_j \zeta_i)^T = \sum_{i=1}^K \boldsymbol{\eta}_i \boldsymbol{\eta}_i^T , \quad (2.14)$$

$$\mathbf{Q} \mathbf{F}_j^T \approx \mathbf{Q}_K \mathbf{F}_j^T = \sum_{i=1}^K \zeta_i \zeta_i^T \mathbf{F}_j^T = \sum_{i=1}^K \zeta_i (\mathbf{F}_j \zeta_i)^T = \sum_{i=1}^K \zeta_i \boldsymbol{\eta}_i^T , \quad (2.15)$$

$$\text{where } \boldsymbol{\eta}_i = \mathbf{F}_j \zeta_i = \frac{\|\zeta_i\|}{\delta \|\mathbf{s}_j\|} \left[f\left(\mathbf{s}_j + \frac{\delta \|\mathbf{s}_j\|}{\|\zeta_i\|} \zeta_i\right) - f(\mathbf{s}_j) \right] .$$

If we now consider the number of forward model evaluations needed per iteration for the calculation of the Jacobian matrix, there are $K + p + 2$ runs. One run is needed for evaluating $f(\mathbf{s}_j)$, one is needed for assessing $f(\mathbf{s} + \delta \mathbf{s}_j)$, K runs are needed for calculating $\boldsymbol{\eta}_i$ and p runs are needed for computing \mathbf{X}_i . It can be observed that, with this method, the number of forward model runs per iteration is no longer dependent on m or n . Hence, the number of parameters and observed data can increase without increasing the run time of the algorithm. However, if the number of parameter increases, the low-rank approximation order K might also need to be increased slightly in order to maintain a small truncation error for \mathbf{Q}_K .

The algorithm iteratively updates the parameters set in Equation (2.5) by solving the matrix system of Equation (2.6) with the PCGA approximations until the optimum $\hat{\mathbf{s}} = \mathbf{s}_{\text{post}}$ is achieved, that is, the objective function has iteratively converged to a local minimum.

2.2.3 Application to an experimental site

We have applied the PCGA large-scale method as presented in the previous part to an experimental site, named 'la Céreirède'. The field site is located in Montpellier in the South of France, on the alluvium of the Lez river, which flows towards the Mediterranean Sea a few kilometers downstream (Figure 2.1).

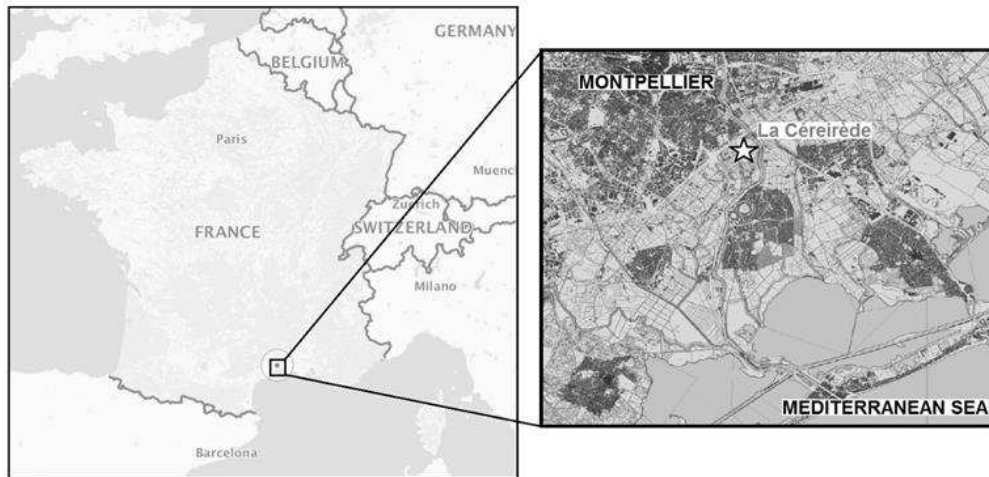


Figure 2.1: Location of the studied experimental site 'La Céreirède' (Map and aerial photography from geoportail.fr) occupying an area of 720 m². It is situated in the South of France, near the town of Montpellier and the Mediterranean Sea.

At the field site, the alluvial deposit is composed of a 12 m thick formation of unconsolidated sands and silts lying on a 2 m thick layer of pebbles and gravels. Beneath these Quaternary formations, there exist clayey sands of the Pliocene, marls of the Miocene, and limestone of the Jurassic (Figure 2.2). Three porous aquifer formations have been characterized in this site:

- a low permeability aquifer in the upper part of the alluvium (unconsolidated sands and silts),
- a semi-confined aquifer in the pebbles and gravel,
- a semi-confined to confined aquifer in the clayey sands of Pliocene.

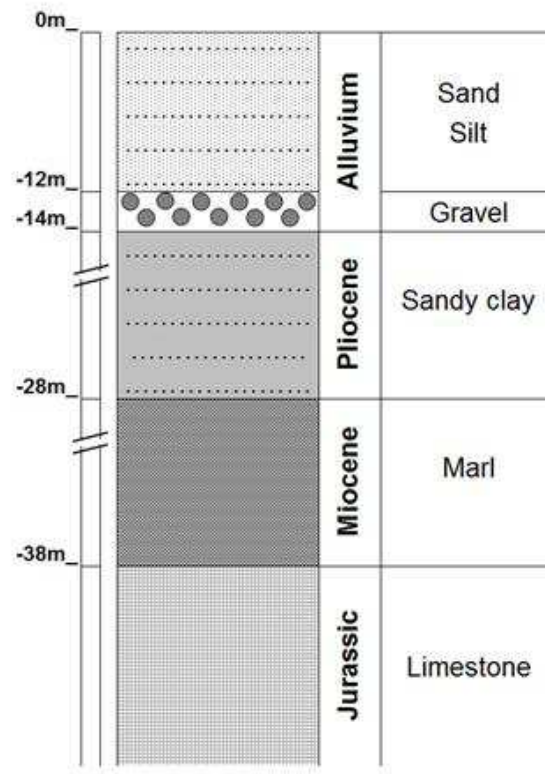


Figure 2.2: Schematic geological section of the experimental site 'La Céreirède'. Three aquifers formations have been characterized in the sands and silts alluvium, in the gravels and in the clayey sands.

The field site comprises 12 wells which fully penetrate the three aforementioned aquifers in an area of $36 \times 20 \text{ m}^2$ (Figure 2.3). The hydraulic data were collected by performing two pumping tests in PZ 2 (5 L/min), and in PZ 11 (5 L/min), respectively, while measuring hydraulic head variations in the other 10 wells.

The pumping were performed at the depth of the pebbles and gravel layer, which is the most productive aquifer, because its transmissivity is considerably higher than the transmissivities in silts and clayed sands. We considered that the contribution of the two others aquifers to the water pumped is negligible compared to the contribution from the pebbles and gravel aquifer. From a classic hydrogeological analysis of the soil of each well, we could also estimate values of the field transmissivity at these points.

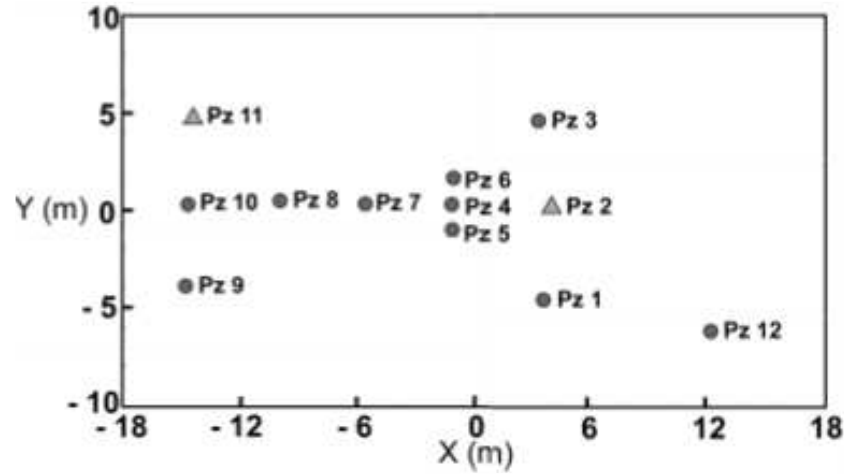


Figure 2.3: Well pattern on the experimental site 'La Cérèrède' (circles represent the 10 measurement wells and triangles represent the 2 pumping wells). As hydraulic drawdowns in the pumping wells are not measured, the tomography provided 20 observed data.

Using these punctual values of transmissivities and their positions in the field as input in a MATLAB variogram routine ('variogramfit' by W. Schwanghart), we were able to obtain the transmissivity field variogram function, which will be used in the inversion algorithm to create the covariance matrix \mathbf{Q} (Table 2.1). The variogram function is of type exponential with a sill of 0.11 and a range of 8 m. The PCGA inversion algorithm was implemented in MATLAB and connected to the flow modeling software COMSOL Multiphysics, which solves the forward problem. The inversion algorithm performs, at the end of each iteration, a linear minimum research ('fminsearch' MATLAB function) to accelerate the convergence of Equation (2.3) to a local minima.

The 2-D flow model is discretized as a grid of $\sqrt{m} \times \sqrt{m}$ transmissivity cells in a rectangular zone. This local model is enclosed by a larger buffer zone of $100 \times 100 \text{ m}^2$ with a constant transmissivity of $10^{-5} \text{ m}^2/\text{s}$ (average value of locally estimated transmissivities at the field site) and a 0 m constant head condition at the boundaries (no drawdown induced by the pumping wells). This buffer zone was set up in order to minimize the impact of the boundary conditions. The flow simulations were performed under steady-state conditions. The inversion of the model is set up using the 2×10 drawdown observed during the pumping tests (10 measurement wells for each of the 2 pumping tests) representing the observed data in the inversion algorithm. The inversion aims to reconstruct the spatially varying T distribution in the local region producing the observational data set.

2.2.4 Results

We have applied the PCGA to assess the equivalent transmissivity field of the multi-layered aquifer at the Céreirède field site. The most transmissive part of the aquifer is the pebbles and gravels part, but the alluvium and the clayey sands might also be the cause of some variations in the estimated equivalent transmissivity field. The inputs to the inversion models are given in Table 2.1.

Table 2.1: Values of variables used to perform the PCGA inversion on a model of the site for 25,600 parameters and 20 observed data. Results of this inversion are shown in Figures 2.4 and 2.5.

Geometry	$X \text{ (m)} = [-18, 18] ; Y \text{ (m)} = [-10, 10]$
Grid (number of parameters)	160×160 cells
Uncertainty on data	$\sigma = 0.001 \text{ m} ; \mathbf{R} = \sigma \cdot Id(n)$
Experimental variogram function	$Var(\text{distance}) = 0.11 \times \left(1 - \exp\left(-\frac{\text{distance}}{8}\right) \right)$
Finite-difference step	$\delta = 10^{-5}$

The inversions were performed on a uniform fine-scale discretization grid (160×160). A constant initial transmissivity field was considered in the inverse models. The K th order of truncation for the covariance matrix \mathbf{Q} was selected such that the first truncated singular value of the matrix (the low rank approximation error) fall below 1. This corresponds to an order of $K = 128$. We will show that this choice of truncation order is acceptable and allows the significant information about the prior model structure to be preserved. The low-rank covariance matrix was calculated and then imported to the inversion algorithm. The parallelized computation of the low-rank decomposition takes only a few minutes (it even takes less than 1 minute with a 32 cores parallelization). The inversion then converged in 2h 45min on an Intel Xeon QuadCore 2.8GHz with 12Go RAM.

The results from the inversion are presented in Figures 2.4 and 2.5. Firstly, from the distribution of model parameters (given as negative log transmissivity in Figure 2.4), it can be seen that the value of the inverted transmissivity takes the mean value $10^{-5} \text{ m}^2/\text{s}$, which is the mean of the transmissivity measurements on the site. A clear contrast in T is observed between the two regions on the east and west sides of boreholes PZ1, PZ2 and PZ3. Overall, the eastern

part which is closer to the Lez river, is slightly less transmissive than the western part (T are approximately $2 \times 10^{-6} \text{ m}^2/\text{s}$ and $1 \times 10^{-5} \text{ m}^2/\text{s}$, respectively). An area with a highest T ($3 \times 10^{-5} \text{ m}^2/\text{s}$) is also highlighted within the western part, around PZ 7. But it has to be noticed that the boreholes and pumping wells are not homogeneously distributed over the site, thus some parts of the site (especially on the eastern side) might give more uncertain results. Therefore, it is interesting to estimate the values of the transmissivity field uncertainties.

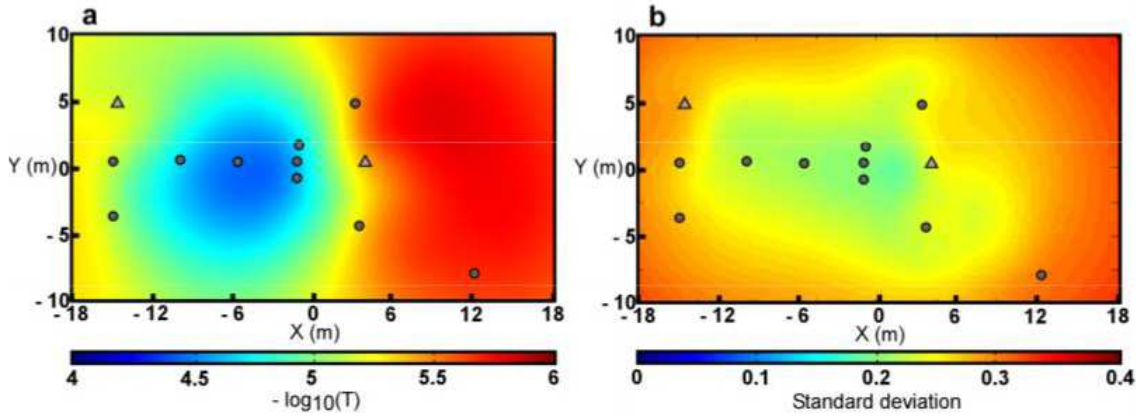


Figure 2.4: Maps of the log-transmissivity (a, on the left) and parameter's a posteriori standard deviation (b, on the right) for a PCGA inversion method with 25,600 parameters, 20 observed data and a covariance matrix decomposition of order $K = 128$ applied to the experimental site. The transmissivities vary around a mean of $10^{-5} \text{ m}^2/\text{s}$ which is consistent with transmissivity values estimated from pumping test analysis. The aquifer is less transmissive in the eastern part and more in the western part especially in a zone around PZ 7 (see Figure 2.3). But we got a better precision in zones with more information: at the center and the western part of the map, while in the eastern part where we didn't have piezometers, the results show a larger standard deviation.

Figure 2.4b presents a map of the uncertainty for each parameter value (given by the diagonal entries of the posterior covariance matrix \mathbf{Q}_{post}). The standard deviation for the log-transmissivity varies between 0.2 for the parameter cells near the investigation wells, and 0.33 in both the area with very few information and close to the model's boundaries. In particular, the uncertainty in the eastern part, where the number of wells is small, is much higher.

Good correlations between the calculated and measured hydraulic heads were obtained (Figure 2.5). The root mean square error calculated at the end of the inversion was computed as 0.194 m. The hydraulic heads with the most significant difference between the inverted and the measured values are observed on PZ1 and PZ9.

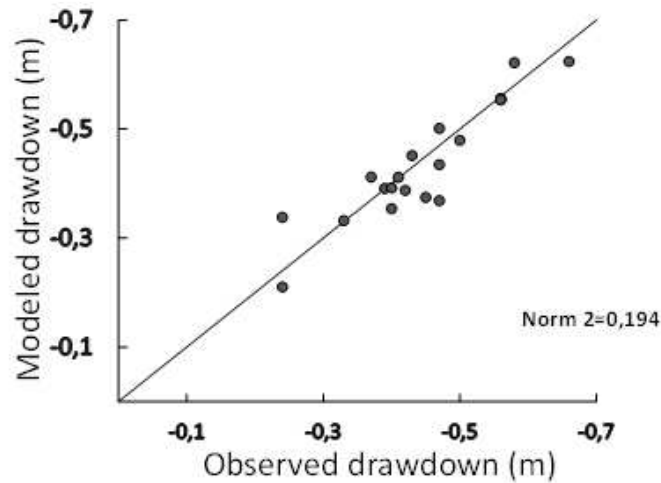


Figure 2.5: Graph showing the differences between the 20 observed drawdowns and modeled drawdowns after convergence of a PCGA inversion method with 25,600 parameters and a covariance matrix decomposition of order $K = 128$ applied to the experimental site. The drawdowns are globally well reproduced.

It is clear that the PCGA method is reliable for the modeling of the porous aquifer considered in this work because it produces a set of good inversion results with a high resolution (each cell represents a rectangle of $22.5 \text{ cm} \times 12.5 \text{ cm}$ on the site) from a few measurements (20 observed data for 720 m^2) in less than 3 hours. However, the main problem of this method is that the SVD of the covariance matrix needs a considerably large amount of time and memory. The computational demands increase squarely with the number of model parameters, m . In this work, the decomposition was only performed once before the inversion, and the resulting low-rank covariance matrix was used at all iteration steps throughout the inversion (i.e., the variogram function remaining the same). Otherwise, if a variable variogram model is desired at different iteration steps, the computation of the covariance matrix decomposition can also be accelerated by specific linear algebra methods (FFT, FMM and H-matrices) and a parallelization on several cores to achieve a reasonable computational time (Lee and Kitaniidis 2014).

2.2.4.1 Comparison of results between PCGA simulations with different decomposition order, and between PCGA and GA simulations

In addition to the PCGA experimental application presented above, several other inversions were also conducted for the same field site but with a smaller number of parameters and using different methods for the computation of sensitivity matrix (input for each inversion is shown in Table 2.1). These numerical experiments allows us to compare (1) the results from

PCGA with varied low-rank truncation K ; and (2) the results from PCGA to those obtained from the GA method with a 1st order finite-difference Jacobian matrix computation, and an adjoint-state Jacobian matrix computation (the integral was solved using the Gauss-Legendre quadratic method as described in *Soueid Ahmed et al. 2014*) The inversion results are compared with respect to the total computation time and relative accuracy of the results.

2.2.4.2 Comparison of inversion results for using different decomposition order

First, we assess the effect of the chosen truncation order for the covariance matrix. Three PCGA inversions, with 10,000 parameters and using different K -order truncations of the covariance matrix corresponding to singular values (truncation errors) of $\lambda_{K+1} \approx 1$ ($K = 69$), $\lambda_{K+1} \approx 0.1$ ($K = 313$) and $\lambda_{K+1} \approx 0.01$ ($K = 1,532$), were performed. Figure 2.6 shows a relationship between the singular value and the truncation order, on which the position of the three K orders that were adopted in our inversions are indicated. Figure 2.7 shows the results obtained from these inversions.

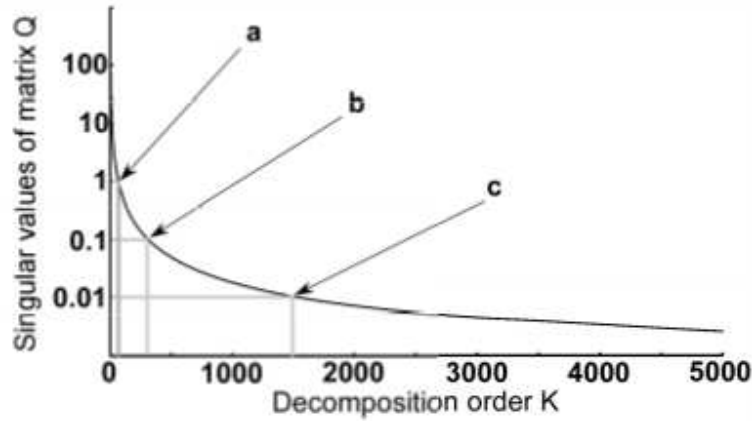


Figure 2.6: Covariance matrix singular values decrement curve for 10,000 parameters. Three decomposition order (a to c) corresponding to truncation error of 1, 0.1 and 0.01 have been chosen for the results comparison of the PCGA inversion method (see Figure 2.7).

Note that very similar T fields (e.g. similar trend and location of high T zones) were obtained from the inversions using different truncation order. Therefore, the influence of reducing the K -order, as long as the truncation error is below 1, on the inversion results is mild.

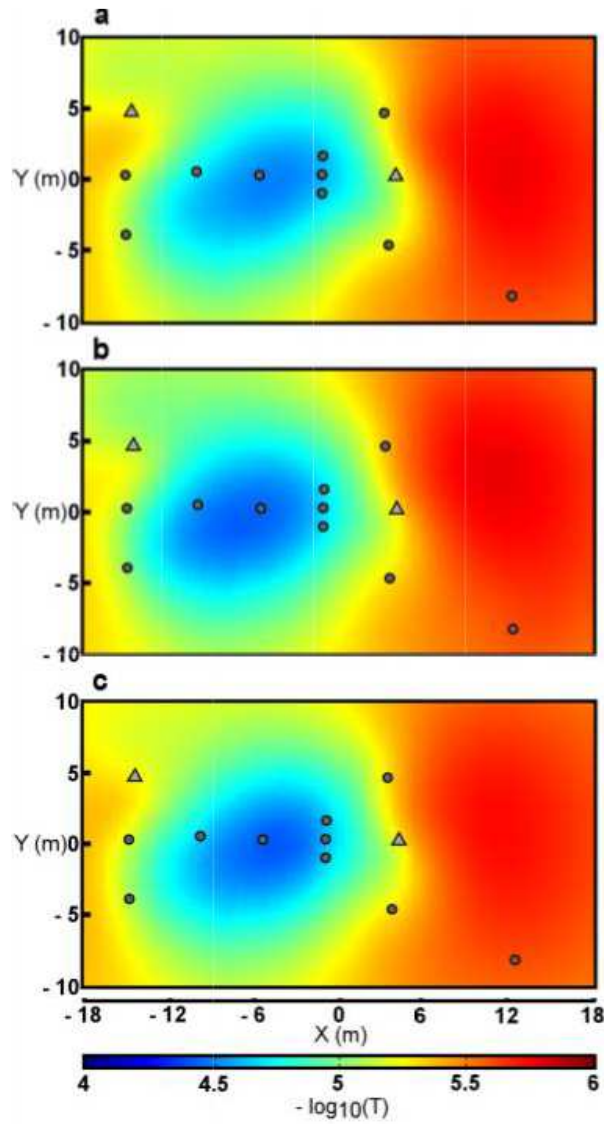


Figure 2.7: Maps of the log-transmissivity for a PCGA inversion method with 10,000 parameters, 20 observed data and three different covariance matrix decomposition applied to the experimental site. The map (a) was obtained for $K = 69$, the map (b) for $K = 313$ and the map (c) for $K = 1,532$ (see Figure 2.6). The results obtained for these three decomposition are relatively the same (same transmissivity values, same zones) so, for this site, there is no significant loss of information when using a truncation order corresponding to an error of 1 (map (a)) for the covariance matrix which allows us to reduce the computation time of the inversion without decreasing the accuracy of the results.

The most information regarding the spatial structure of the prior model is preserved in its first few singular values, so it is acceptable to consider a truncation order for a truncation error $\lambda_{K+1} \approx 1$ for the covariance matrix.

2.2.4.3 PCGA and GA results comparison

In this section, we compare the inversion results obtained for the PCGA method using an approximated \mathbf{Q} matrix, with those of the GA method using two different methods for

Jacobian matrix computation (i.e., the finite-difference and adjoint state methods), where the entirety of the \mathbf{Q} matrix were used. All the inverse simulations were performed on a 100×100 grid. Figure 2.8 shows the inverted transmissivity distributions and the corresponding distributions of standard deviation of each model parameter.

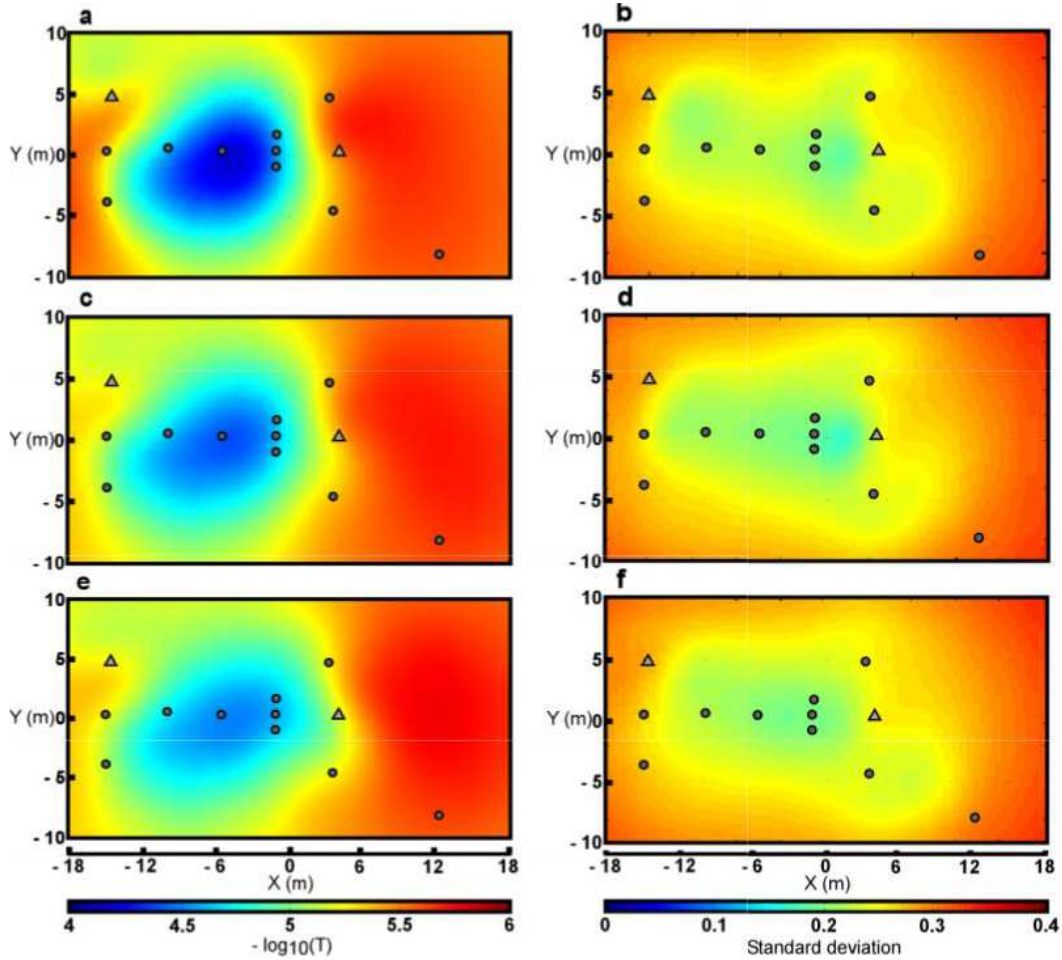


Figure 2.8: Maps of the log-transmissivity (a, c, e) and parameter's a posteriori standard deviation (b, d, f) for three different inversion methods with 10,000 parameters and 20 observed data applied to the experimental site. The maps (a) and (b) were obtained with the GA adjoint-state method, the maps (c) and (d) with the GA finite-difference method and the maps (e) and (f) with the PCGA method with a covariance matrix decomposition of order $K = 69$. The results between the three methods are relatively the same for this site, except for the map a which presents a slightly higher contrast of the transmissivity distribution leading to a better data matching (see Figure 2.9), though the PCGA inversion method is much more efficient for the calculation time (see Table 2.2).

It can be seen that in the three transmissivity fields, the calculated T value for each cell varies around the mean $10^{-5} \text{ m}^2/\text{s}$. In general, the three approaches produced similar spatial distribution of the transmissivity, however the range of the inverted transmissivities from the GA-adjoint state method (i.e., 2×10^{-6} to $1 \times 10^{-4} \text{ m}^2/\text{s}$) is larger than that of the PCGA and that

of the GA-finite difference method (i.e., 2×10^{-6} to 3×10^{-5} m²/s). Since the PCGA method is based on a finite difference matrix vector product approximation, it is expected to obtain a similar parameter range from this method and the GA-finite difference method. The difference between the results from these two methods and the GA-adjoint state method may come from the definition of the finite-difference step δ . In addition, the resolved transmissivity field from PCGA is smoother compared to that from the method where an adjoint state method is used to compute the sensitivity matrix. This is caused by the low-rank truncation in PCGA and also the finite difference approach, which tends to reduce the heterogeneity of the inverted T field.

The spatial distributions of the standard deviation of the inverted parameters is presented in Figure 2.8. For each method, the uncertainty of the reconstruction is mainly dependent on the number and position of the wells. The correlation between inverted and measured hydraulic head data for the three inversion models are presented in the cross-plots of Figure 2.9.

It can be seen that, with the same inversion inputs, the GA-adjoint-state method generated a slightly better correlation compared to the other two methods, which is reflected by the smaller RMSE values (0.165 m compared to 0.182 m and 0.188 m). The performance of the three methods in terms of simulation time is compared in Table 2.2.

A significant reduction in computational time is observed for the GA-adjoint-state method compared to the GA-finite difference method (Table 2.3). This reduction is mainly related to the calculation of the Jacobian sensitivity matrix (*Cardiff and Kitanidis 2008*). As the grid discretization increases, the significance of reduction in computational time of the GA-adjoint-state method compared to the GA-finite difference based method become more apparent. However, an even more significant time reduction was observed in using the PCGA method. Note that in Table 2.3 the computation time for PCGA includes the time from both the covariance matrix decomposition and the inversion calculation. The computational time of PCGA is observed to be 10 time less than that of the GA-adjoint state method and 70 time less than that of the GA-finite difference method. Altogether, the advantage of PCGA in obtaining a fast solution without compromising the inversion quality makes it a promising candidate in solving large-scale inversion problems.

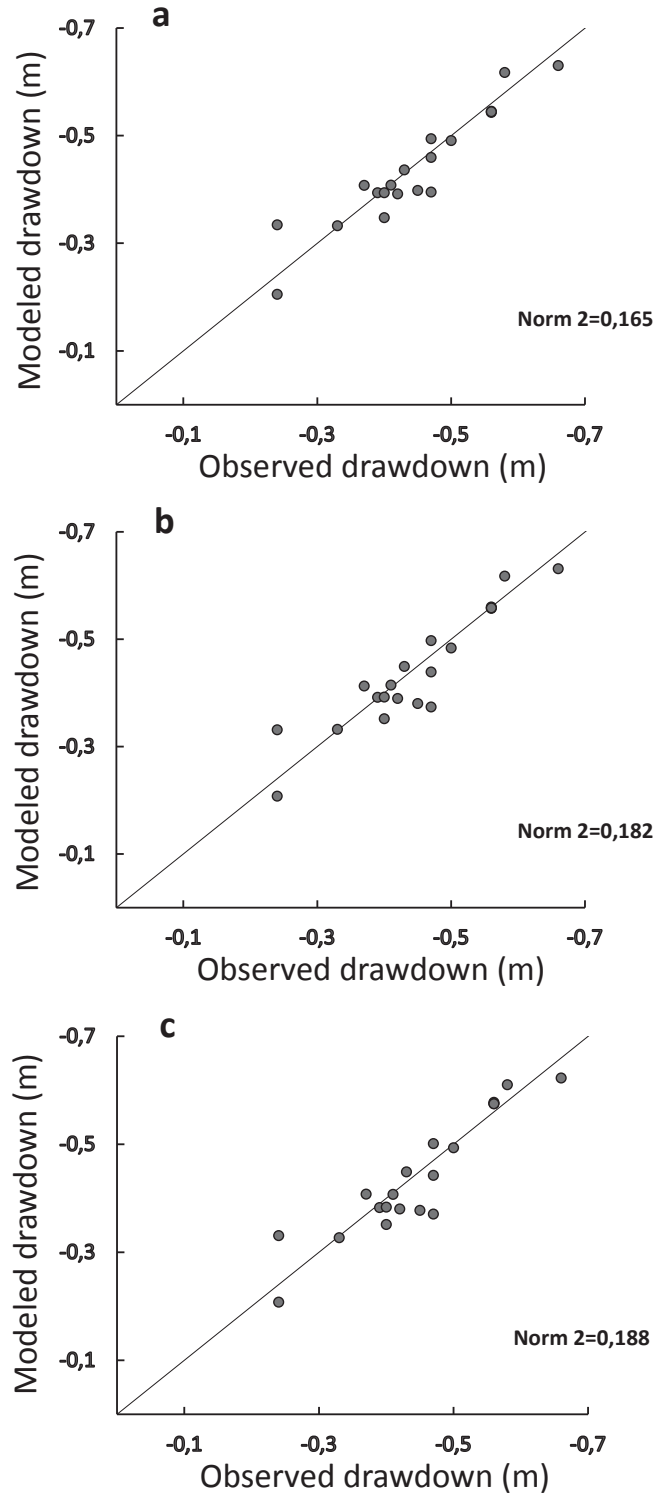


Figure 2.9: Graphs showing the differences between the 20 observed drawdowns and modeled drawdowns after convergence of three different inversion methods with 10,000 parameters applied to the experimental site. The graph (a) was obtained with the GA adjoint-state method, the graph (b) with the GA finite-difference method and the graph (c) with the PCGA method with a covariance matrix decomposition of order $K = 69$. Regarding the mathematical norm 2 the GA adjoint-state method has a slightly better convergence on the data than the other methods but the PCGA inversion method is much more efficient for the calculation time (see Table 2.2).

Table 2.2: Comparison of the efficiency between three algorithm of geostatistical inversion methods (GA adjoint-state, GA finite-difference and PCGA) on a same under-determined modeling. Results of these inversions are shown in Figures 2.8 and 2.9. The convergence on data was slightly better for an adjoint-state method but the calculation time was considerably reduced by using a PCGA method.

	GA adjoint-state method	GA finite-difference method	PCGA method (truncation order $K = 69$)
Number of parameters	100×100	100×100	100×100
Computation time	10h 43min	72h 10min	1h 33min
Value of objective function to be minimized after convergence	23.0602	26.6526	27.4641
ℓ_2 -norm between observed and modeled data	0.165 m	0.182 m	0.188 m

Table 2.3: Convergence times for different methods using different grid sizes. An Intel Xeon QuadCore 2.8GHz with 12Go RAM has been used to perform the computations. The PCGA method (with a truncation error of approximately 1) is always the fastest because it involves less forward problems than the GA finite-difference method and that the Gauss-Legendre resolution of the integral in the GA adjoint-state method requires a calculation of a number of nodes proportional to the number of cells in the grid in each forward problem.

Grid resolution	GA adjoint-state method	GA finite-difference method	PCGA method (truncation error $\lambda_{K+1} \simeq 1$)
10×10	5min	16min	1min
30×30	9min	1h 40min	3min
50×50	45min	8h 21min	6min
80×80	3h 19min	41h 44min	29min
100×100	10h 43min	72h 10min	1h 33min

2.2.5 Conclusion

The GA finite-difference method is useful and straightforward for inversions with a few parameters and a large number of observational data. In contrast, the GA adjoint-state method is advantageous in dealing with inversion models with a few observational data but a relatively large parameters set. On the contrary, the PCGA is an efficient method for both cases. It is also helpful for extremely under-determined problems where a large number of unknown parameters is present. In fact, the time and memory required by this method to perform the iterative process of the inversion is less sensitive to the number of parameters or measurements, but more dependent on the approximation order of the covariance matrix chosen by the modeler. A higher order approximation will lead to higher computational costs but the error introduced in the inversions will be much smaller.

The comparison of different methodologies has shown that PCGA approach appears to be the most efficient strategy for carrying out large-scale inversions in porous aquifers. We noticed that, in the studied case, the errors introduced by the approximation in the PCGA methods were not significant. Thus, we obtained the same accuracy in results from the PCGA inverse modeling compared to the GA inverse modeling. Thereby, with this strategy, only the principal components of the covariance matrix are kept in the inversion process, and the computational and memory costs necessary for the inversion algorithm are optimized. Additionally, the PCGA method significantly reduces the computational time. With the PCGA method we divided the computation time by seven compared to the GA adjoint state method, and by 50 compared to the GA finite-difference method.

In summary, by applying the PCGA for a hydraulic tomography in a porous aquifer, we found an especially adapted strategy, which produces accurate inversion results with a good resolution in a reduced time, and which manages optimally the computer memory involved in the inversion algorithm. Nevertheless, the PCGA method is efficient specifically for models with a smooth distribution of the targeted parameters (which could typically be used for a good average representation of porous aquifers) so that the covariance matrix can be approximated by much smaller matrices.

3 UN NOUVEL OUTIL D'INVERSION STRUCTURALE POUR LES MODÈLES ÉQUIVALENTS MILIEUX POREUX

3.1 Contexte

A partir du constat du chapitre précédent que les méthodes d'inversion géostatistiques n'étaient pas pertinentes pour les milieux fracturés et karstiques, et que l'état de l'art actuel est globalement limité en proposition de méthodes adaptées à ce genre de distribution très contrasté des propriétés, de nouvelles méthodes d'inversions ont été développées et sont présentées dans ce manuscrit de thèse.

Une première méthode d'inversion déterministe basée sur le concept mathématique des automates cellulaire est présentée dans ce chapitre. Les automates cellulaires permettent de générer dans un modèle équivalent milieu poreux, un réseau de structures linéaires se dessinant sur un fond. Le modèle binaire ainsi généré permet de présenter un réseau de fractures ou conduits dans la roche mère matricielle, en attribuant à chaque domaine des valeurs de conductivités très contrastées. Cette méthode permet une optimisation à la fois des valeurs de propriétés, mais également de la forme du réseau généré. La liberté apportée dans l'optimisation est permise par un découpage du modèle en sous-espace, chaque sous-espace étant piloté par un automate cellulaire indépendant permettant d'y modifier facilement la direction locale de génération de la structure linéaire et également des valeurs locales de propriétés.

Ce chapitre présente d'abord le paramétrage du modèle et des automates cellulaires, suivi par les équations du problème inverse permettant l'optimisation de la forme de réseau de structures et des valeurs de propriétés, puis par la quantification d'incertitudes *a posteriori*. La méthode d'inversion est ensuite testée et validée au travers de six applications sur des cas d'études synthétiques et théoriques à partir de réponses sismiques et hydrauliques. Les limites de la méthode, liées à la binarité structurale du modèle et aux ouvertures de la structure qui sont non optimisables sont contrebalancées par l'optimisation des propriétés, qui permettent de représenter en équivalence une variation d'ouverture.

Cette méthode permet d'imager des structures fines en réseaux complexes à partir de réponses d'un milieu à une sollicitation, ce qui en fait une méthode intéressante pour les milieux fracturés et karstiques. Les cas présentés dans ce chapitre permettent certes de valider la méthode, mais demeurent des cas synthétiques, non représentatifs de la complexité rencontrée dans un cas réel. L'application de cette méthode à partir de données de terrain, présentée plus tard dans ce manuscrit, permettra de discuter plus spécifiquement de l'applicabilité de cette méthode dans des cas concrets.

3.2 A cellular automata-based deterministic inversion algorithm for the characterization of linear structural heterogeneities

Cette partie est composée de l'article « A cellular automata-based deterministic inversion algorithm for the characterization of linear structural heterogeneities », publié dans le journal *Water Resources Research* en mars 2017 (*Fischer et al. 2017b*). Le texte a été remis en forme au format du manuscrit. La version originale de l'article est donnée en Annexe 2.



Water Resources Research

RESEARCH ARTICLE

Citation:

Fischer, P., A. Jardani, and N. Lecoq (2017), A cellular automata-based deterministic inversion algorithm for the characterization of linear structural heterogeneities, *Water Resour. Res.*, 53, 2016–2034, doi:10.1002/2016WR019572.

A cellular automata-based deterministic inversion algorithm for the characterization of linear structural heterogeneities

P. Fischer¹, A. Jardani¹, and N. Lecoq¹

¹Normandie Université, UNIROUEN, UNICAEN, CNRS, M2C, Rouen, France

3.2.1 Introduction

In geophysics, the inverse method is an efficient way for mapping the geological structures by assessing the physical properties of the subsurface (such as hydraulic conductivity, electrical resistivity, magnetic susceptibility, volumetric density, porosity, etc.) from a set of observed data. These observed data represent the responses of the investigated area to solicitations applied during the geophysical surveys (pumping tests, electrical resistivity tomography, electromagnetic, gravimetry, seismic, etc.). Commonly, the inverse problems are undetermined, with non-uniqueness of the solution, leading to provide doubtful interpretations of the geophysical surveys. Thus, the addition of a priori information on the properties to estimate is a necessity for avoiding the physically unrealistic models. Most often, geostatistical constraints are used to reconstruct the physical properties of a soil that can be modeled by smooth spatial variabilities (*Hoeksema and Kitanidis 1984*). However, when the parameters have a high spatial variability the use of statistical characteristics as a priori information becomes ineffective and inadequate to locate the discontinuities of the physical properties. Therefore several algorithms have been proposed to deal with the "structural" inversion, considering both, the estimation of physical properties and reconstruction of boundaries between different heterogeneities.

Among these approaches, we cite those which incorporate structural information in the model parameterization of the inverse problem, such as the multiscale method that rests on an increasing resolution of the parameterization during the optimization sequences (*Grimstad et al. 2003*). The adaptive multiscale method permits to reduce the number of unknown parameters by a local refinement of the parameterization where the heterogeneity is the most important, to avoid an over-parameterization. *Tsai et al. (2003)* used the Voronoi zonation with a pilot-point parameterization method to identify parameters structures in a model.

For the approaches using no specific parameterization of the model, *Lelièvre and Oldenburg (2009)* have proposed to incorporate constraints to the inversion objective function in terms of some structural information such as orientation to obtain more realistic solutions. The spatial distribution of the unknown parameters can also be approximated by a sparse set of coefficients to be identified in a compressed sensing sparsity-promoting inversion (*Jafarpour et al. 2009*) which promotes sparse solutions. Finally the inversion with total variation prior (*Lee and Kitanidis 2013*) uses a Laplace prior instead of a Gaussian in a Bayesian inversion in order to delimitate the shapes of discrete structures piloted by some hyperparameters determined during the inversion using an expectation-maximization approach.

In another register, *Lochbühler et al. (2015)* used the training image method in the inverse formulation to represent the structural characteristics of a field as prior information to eliminate inversion artefact and improve the estimate of the parameters. *Hale (2009)* and *Soueid Ahmed et al. (2015)* have proposed the guided image method in which the structural features of the domain is presented graphically and used as a priori information to guide the inversion by refining the model sensitivity at boundaries between different zones. It permits a better estimate of the intra-structure parameter variabilities and location of different features in a model. The level set is an alternative approach to detect the interfaces between different facies thanks to the use of extensible boundaries that move during the inversion process to fit the observed data (*Lu and Robinson 2006 ; Cardiff and Kitanidis 2009*). *Haber and Oldenburg (1997)* have identified the profits that could bring joint inversion to structural identification and have presented a protocol to run a joint inversion in geophysics by constraining the results with a unique structural consideration. Since then, several other structural joint inversions tools have been developed which were summarized in a review proposed by *Gallardo and Meju (2011)*. This review presents the recent techniques of structural joint inversions and the upcoming challenges of such inversions in the next years.

However, regarding the imaging of linear structures, which are characterized by an aperture significantly lower than their length (such as karst conduits and fractures), the deterministic inversion remains, according to our knowledge, an unexplored subject. The inverse modeling of such structures requires a large-scale parameterization, which makes the computation very heavy particularly in the case of stochastic or global optimization algorithms. (*Pardo-Igúzquiza et al. 2012 ; Reeves et al. 2013 ; Bruna et al. 2015 ; Javadi et al. 2016*). In this paper we propose a new method called Cellular Automata-based Deterministic Inversion (CADI), adapted for the inversion of linear structures. This approach is based on a Bayesian formulation with the use of Cellular Automaton (CA) concept to parameterize the model. The dynamic structural optimization in the algorithm is controlled by the CA, which allows the drive of an entire discretized system with only some local configurable interaction rules. After a global presentation of the model parameterization, we will detail the interaction rules chosen for the CA in this CADI algorithm to reproduce linear structures (section 3.2.2.1). Then we will present the structure of the inverse problem (section 3.2.2.2) and the protocol of optimization (section 3.2.3). Several examples conducted with the CADI algorithm on linear problems (seismic) and non-linear problems (hydraulic tomography) are described in the last part of this paper (section 3.2.4).

3.2.2 Parameterization of inverse problem using cellular automaton

3.2.2.1 Parameterization of the model

In the CADI method, the distribution of the properties in the model is structurally generated by several CA. In previous works, CA have already been coupled to global optimization algorithms such as genetic algorithms (*Dewri and Chakraborti 2005 ; Ghosh et al. 2009*). However, in the CADI method we wanted to couple the possibilities offered by CA to a deterministic inverse process. Therefore, the model (discretized in m cells) is partitioned in m_{CA} CA subspaces (with $m_{CA} \ll m$), each one being monitored by an independent CA configured by its neighborhood definition (Figure 3.1). Thus the CA subspaces and their parametrization are pilot zones for the model, which permits to avoid an over-parameterization of the inverse problem.

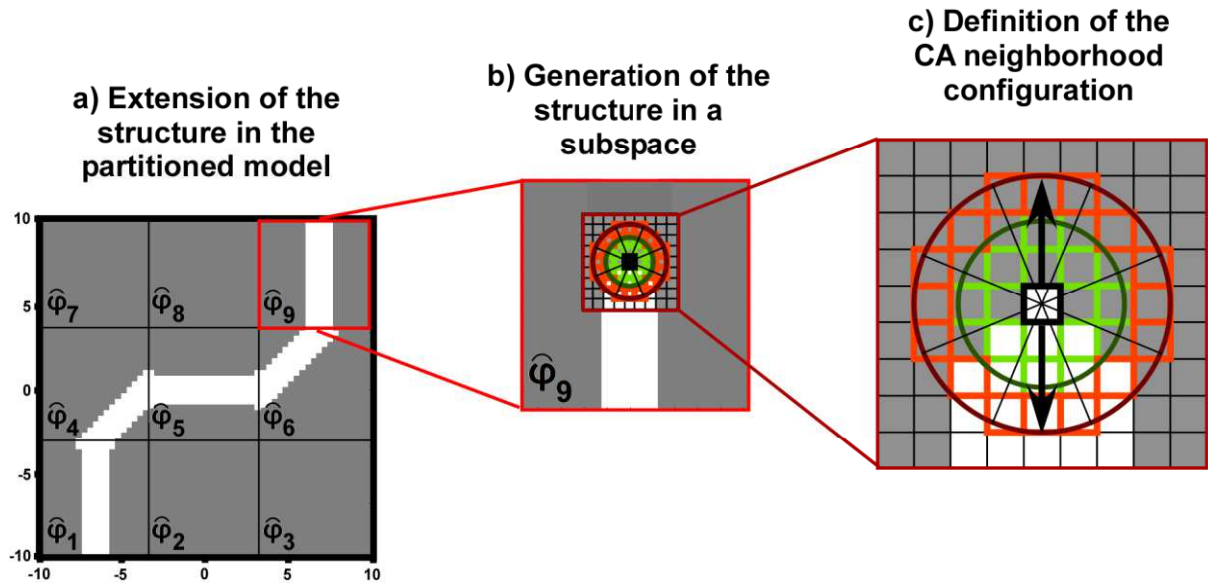


Figure 3.1: Scheme explaining how the CA are used in the CADI model. In the figure grey occurs for state 'background' and white for state 'structure'. The model is partitioned in m_{CA} independent CA subspaces (here $m_{CA} = 9$). During the generation process the structure will go through different CA subspaces (a) and will be generated in the local direction assigned by the structural parameters piloting these CA (b). Along the generation direction the CA will modify the property values of the model cells it controls (represented by the squares lattice in Figure 3.1c).

The CA is a widely used mathematical system to generate discrete dynamic models. It has been applied to diverse fields of modeling such as Random Number Generators (*Tan and Guan 2007*), chemical reactions (*Van der Weeën et al. 2011*), solid-solid phase transformation

during heating (*Halder et al. 2014, 2015*) or cooling (*Dewri and Chakraborti 2005 ; Ghosh et al. 2009 ; Jin and Cui 2012*), fluid flow through fractures (*Pan et al. 2011*), or transport in fluid flow (*Chopard and Masselot 1999*). The CA is a popular method due to its capacity to model complex systems by using simple rules. In fact, the evolution of the entire system is driven by some configurable local interaction rules.

The CA is a discrete time-evolving system in which a cell of the system is defined by its state and its neighborhood. At a CA time step $t_{CA} + 1$, the state of each cell will simultaneously evolve to a new one depending on a constant transition rule involving the state of a cell and the states of the cells in its neighborhood at the time step t_{CA} . Thus, a CA subspace of the model proposed in the CADI algorithm can be described as a quadruple $\langle \phi, Z, N, T \rangle$ (*Sun et al. 2011 ; Van der Weeën et al. 2011*):

- ϕ is a subspace of the global model, discretized as a lattice of c cells,
- Z is a function returning the states values for each selected cells of the subspace at a specified time step (with two possible states: $Z_{t_{CA}}(c_i) = \beta_{background}$ or $Z_{t_{CA}}(c_i) = \beta_{structure}$, where $\beta_{background}$ and $\beta_{structure}$ are the parameterized values assigned to the properties of the structure and the background in the subspace),
- N is a neighborhood function that selects among all cells of the subspace the subset of cells that are considered in the neighborhood of a given cell c_i ,
- T is a function of cell-state transition rule. Thus a transition in the CA for a given cell c_i is expressed as $Z_{t_{CA}+1}(c_i) = T(Z_{t_{CA}}(N(c_i)))$, and a full transition in the CA process (considering all cells of the subspace lattice) is $\phi_{t_{CA}+1} = Z_{t_{CA}+1}(c_i), \forall i$.

The choice of N and T for the CA in this work will be detailed in the following paragraphs. The CA will be used to produce a spatial linear structure in the model. The global model is partitioned in several subspaces, each one being discretized as a lattice of c squared cells (Figure 3.1c). Each cell of a subspace ϕ can be in only two possible different states: state ‘background’ which take a value $\beta_{background}$, or state ‘structure’ which take a value $\beta_{structure}$. So,

a subspace has a binary distribution. $\beta_{\text{structure}}$ is homogeneous within a subspace but can vary among the different subspaces.

Commonly, CA use neighborhood sequencing such as the Moore or the Von Neumann neighborhood rules (see Appendix 2) (*Moore 1962 ; Von Neumann and Burks 1966*). But here we chose N as a dual-radius neighborhood definition as presented in Figure 3.2. Two circles of cells, defined by their cell-radius R_{inner} and R_{outer} , are centered on a given cell c_i (for a full CA time step transition c_i would be alternatively each cell of the CA subspace). The inner circle defines the ‘activator’ cells for c_i (green on Figure 3.2) and the outer circle defines the ‘inhibitor’ cells for c_i (orange on Figure 3.2). The terms ‘activators’ and ‘inhibitors’ are relative only to the cells in ‘background’ state: the cells in state ‘background’ in the ‘activator’ neighborhood will tend to transform the cell c_i in a state ‘background’ while the cells in state ‘background’ in the ‘inhibitor’ neighborhood will tend to transform the cell c_i in a state ‘structure’. The balance of the ratio of ‘background’ cells in each ‘activator’ and ‘inhibitor’ neighborhoods can be disturbed by the existence of cells in state ‘structure’ (for example the presence of a cell in state ‘structure’ in the ‘activator’ neighborhood of a cell c_i will reduce the amount of cells in state ‘background’ in this ‘activator’ neighborhood compared to in the ‘inhibitor’ neighborhood, and thus the cell c_i would then become ‘structure’). The cells outside of the circles are not considered for the state transition of the cell c_i . This definition can be seen as an extension of the Von Neumann rules, by adding inhibitors neighbors to activators cells that follow a Von Neumann neighborhood. Additionally, each circle of the neighborhood definition was radially split into 8 sectors to allow for spatially variable weighting of activators and inhibitors cells, in order to privilege particular directions during the generation (see Figure 3.4).

During the transition to $t_{\text{CA}} + 1$, the transition rule T defines alternatively for each cell of the lattice the new state of a cell c_i by considering the equilibrium of activators and inhibitors cells in state background ($\beta_{\text{background}}$) in its neighborhood $N(c_i)$ at the instant t_{CA} . Therefore the neighborhood configuration associates cells values in the activator zone to a positive weighting (+) and cells values in the inhibitor zone to a negative weighting (-).

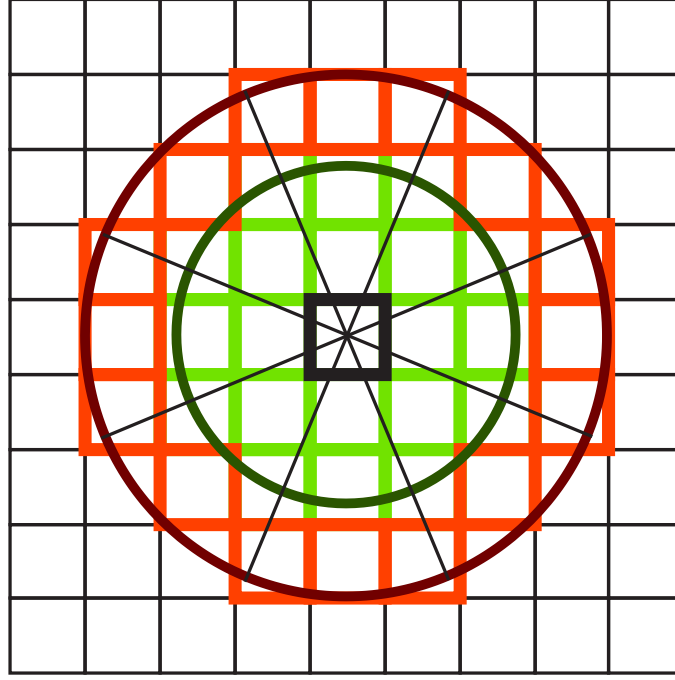


Figure 3.2: An example of the dual-radius neighborhood considered in our CA definition. The black highlighted cell is the cell under consideration in this example (each cell of the lattice would alternatively be considered during a full CA time step). The greened highlighted cells are considered as its 'activators' neighbors in the transition rule and the orange highlighted cells as its 'inhibitors' neighbors. These cells are selected by an inner and an outer circle (in bold) with configurable radius which permit the configuration of the neighborhood. In this example, the inner circle has a radius = 2 and the outer circle has a radius = 3. Additionally, the neighborhood is split into 2×8 sectors (by the radial lines) which permit a more configurable weighting definition (see Figure 3.4).

The weighted values in each of 8 activator and inhibitor sectors are then also corrected by an additional balancing weight (ratio between the number of cells in a sector and the total number of cells), in order to have the same consideration between each sector of the neighborhood. In fact, each sector does not contain the same amount of cells, due to the consideration of deformations of circles in a lattice of squares.

Finally the transition rule T sums the weighted values from all cells in state $\beta_{\text{background}}$ in the neighborhood $N(c_i)$.

If the total weight of activators in state $\beta_{\text{background}}$ is higher (the sum is positive), the cell c_i will take the value $\beta_{\text{background}}$ ('background'), if the total weight of inhibitor in state $\beta_{\text{background}}$ is higher (the sum is negative), the cell c_i will take the value $\beta_{\text{structure}}$ ('structure'):

$$T(c_i) = \sum_{k=1}^{ncell} Z_{t_{CA}}^{\text{weight}}(c_k) \text{ for } c_k \in [N(c_i)] \cap [Z_{t_{CA}}(c_k) = \beta_{\text{background}}] , \quad (3.1)$$

$$Z_{t_{CA}+1}(c_i) = \beta_{\text{background}} \text{ if } T(c_i) \geq 0 , Z_{t_{CA}+1}(c_i) = \beta_{\text{structure}} \text{ if } T(c_i) < 0$$

where $ncell$ denotes the total number of cells in $N(c_i)$ and $Z_{t_{CA}}^{\text{weight}}$ is the function returning the state value taking into account the weighting parametrization from the neighborhood.

After sufficient time steps of the CA with the same transition rule, the subspace ϕ will converge to a stable geometry $\hat{\phi}$ (the geometry will not change over increasing CA time steps anymore) depending on the weighting parametrization given to the neighborhood definition N (Figure 3.3).

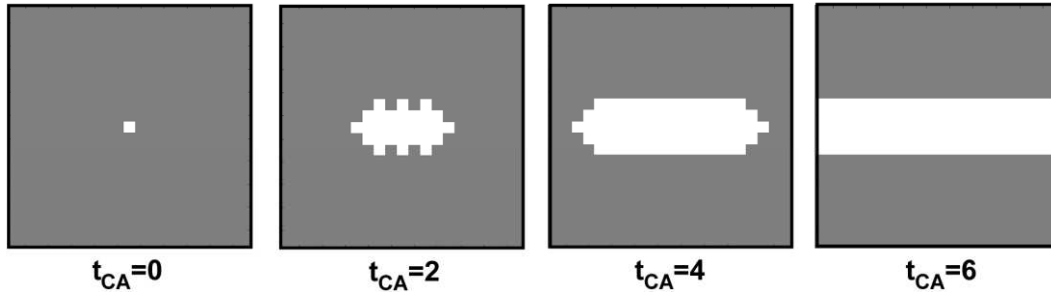


Figure 3.3: Time evolution of a CA configured with a neighborhood weighting defining a horizontal structure generation (see Figure 3.4). After the sixth time step the CA has converged and its geometry is stable over the following steps. Here grey occurs for state 'background' and white for state 'structure'.

Thus, playing on the weighting distribution in the divided activator and inhibitor sectors, and on the radius R_{inner} and R_{outer} of the neighborhood definition N , the CA can produce linear structures in 8 directions from a unique starting cell, as shown in Figure 3.4. The weighting distribution defining each direction has been empirically specified.

On the presented configurations, the starting cell is considered in the center of the lattice. The neighborhood weighting permits to modify the direction of the structure and the radius values modify its aperture.

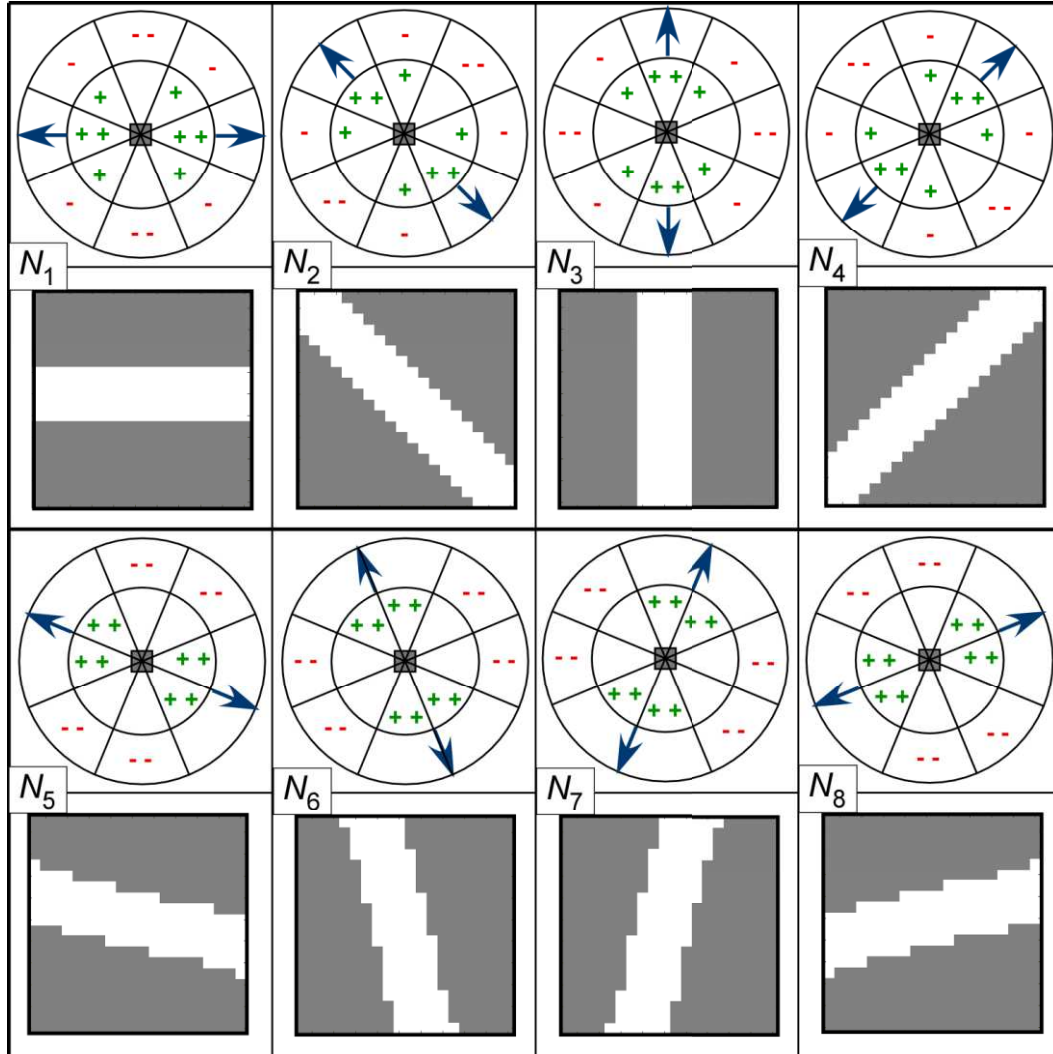


Figure 3.4: Presentation of 8 different stable structures started by a unique centered cell, and their associated CA neighborhood configuration. The greyed cell in the neighborhood configuration is a given cell considered during the CA process. It is surrounded by its neighbor cells, which are not shown for reasons of readability. Its neighborhood is split in 8 internal 'activator' sectors and 8 external 'inhibitor' sectors, each one being assigned to a given weight. A '++' occurs for a positive weight for the neighbor cells in the area, a '++' weight is twice higher than a positive weight represented by a single '+'. A '- -' occurs for a negative weight for the neighbor cells in the area, a '- -' weight is twice higher than a negative weight represented by a single '-'. An empty part of the neighborhood occurs for a null weight, meaning that cells in the area are not considered in the transition rule. Here we present the CA configuration leading to 8 different structure directions which will be considered as sub-orientation of the global structure in the model. In the structural map, grey occurs for state 'background' and white for state 'structure'.

These 8 weighted neighborhood configuration functions $N_i, i \in [1, 8]$ will be considered as the different configuration possibilities in the subspaces parameterization in the dynamic structural optimization process of the inversion algorithm (presented in the section 3.2.3.1). Thus, in the CADI algorithm a CA subspace of the model is parametrized by two parameters: its structural direction (neighborhood configuration N) among the 8 possible and its values of property $\beta = [\beta_{\text{background}}, \beta_{\text{structure}}]$. Therefore, a converged configuration of a parameterized subspace will be expressed as $\hat{\phi}(N, \beta)$ (using one of the 8 different direction configuration N_i , as presented in Figure 3.4).

The geometry of a structure over the entire model Γ (composed of all converged subspaces $\hat{\phi}_k(N, \beta), k \in [1, m_{\text{CA}}]$) can be defined in the CA generation process with only two ‘pilot’ vectors containing each subspaces parameters. This piloted model can be expressed as $\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)$, where \mathbf{P}_N is a m_{CA} vector containing the direction of generation (configured by the weighted neighborhood configurations $N_i, i \in [1, 8]$) assigned to each CA subspace of the model, and \mathbf{P}_β a $m_{\text{CA}} + 1$ vector containing the $\beta_{\text{structure}}$ values assigned to the ‘structure’ cells in each of the m_{CA} CA subspaces and also the $\beta_{\text{background}}$ value (the background being considered, in this paper, as uniform, but it could also be possible to consider a $\beta_{\text{background}}$ value for each subspace). Thus, by piloting the CA generation process with only \mathbf{P}_N and \mathbf{P}_β as parameters we can generate the whole model as shown in Figure 3.1. The aperture all along the structure is considered as constant, and can be configured with the CA neighborhood radius values and the partitioning of the model.

The CA generation process of the structure starts from an entire ‘background’ ($\beta_{\text{background}}$) state model with only one or several selected cell(s) of the model in state structure ($\beta_{\text{structure}}$) which are considered as the starting point(s) of the structure. At the firsts CA time steps, the structure will be generated only in the subspaces where initial structure cells are defined. Each boundary cells state at the edge of a CA subspace is symmetrically transferred to the boundary cells of the adjacent CA. Therefore when the structure arrives to the limit of its first subspace, it can enter a new CA subspace by local symmetry at the boundary limit between them. The new CA subspace the structure has entered has potentially another neighborhood definition, thus the structure will follow a new direction from there. Once the structure has been

generated in a subspace, this subspace becomes 'inhibited' to another generation (the structure can enter only one time each subspace). And so the structure will propagate within the model, through the increasing CA time steps, until it reaches a stable geometry $\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)$ (see Figure 3.1).

3.2.2.2 Statement of inverse problem

The inverse problem involves a formulation of the forward problem which links the spatial properties of the model to the data:

$$\mathbf{d} = f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)) + \boldsymbol{\varepsilon}, \quad (3.2)$$

where $\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)$ is the spatial distribution of the m properties cells in the model. The cells of the model take their values from a finite set \mathbf{P}_β and are structured by the CA directions \mathbf{P}_N , \mathbf{d} is a vector of n modeled data, f is a forward problem application $\mathbb{R}^m \rightarrow \mathbb{R}^n$, and $\boldsymbol{\varepsilon}$ represents the observed data error.

In a probabilistic framework, the aim of the inverse problem is to find the most probable models considering \mathbf{P}_N and \mathbf{P}_β as parameters constrained by the observed data and the prior information on both parameters. This inverse issue can be treated as a sequential inversion. First, for a given \mathbf{P}_β , we determine the geometry of the structure via the estimation of \mathbf{P}_N , which is then used in the second time to infer the values of \mathbf{P}_β . Using a Bayesian approach on Gaussian probability density functions, the problem can be formulated by two posterior probability densities $\rho_{\text{structure}}(\mathbf{P}_N | \mathbf{d}, \mathbf{P}_\beta)$ and $\rho_{\text{properties}}(\mathbf{P}_\beta | \mathbf{d}, \mathbf{P}_N)$ in order to image the geometry of structure controlled by \mathbf{P}_N and their physical property values controlled by \mathbf{P}_β :

$$\begin{aligned} \rho_{\text{structure}}(\mathbf{P}_N | \mathbf{d}_{\text{obs}}, \mathbf{P}_\beta) &\propto \rho(\mathbf{d}_{\text{obs}} | \mathbf{P}_N, \mathbf{P}_\beta) \cdot \rho(\mathbf{P}_N) \\ &\propto \exp\left(-\frac{1}{2}(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)))^T \mathbf{C}_d^{-1} (\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)))\right), \quad (3.3) \\ &\quad \times \exp\left(-\frac{1}{2}(\mathbf{P}_{N,\text{prior}} - \mathbf{P}_N)^T \mathbf{C}_{\mathbf{P}_N}^{-1} (\mathbf{P}_{N,\text{prior}} - \mathbf{P}_N)\right) \end{aligned}$$

$$\begin{aligned}
\rho_{\text{properties}}(\mathbf{P}_\beta | \mathbf{d}_{\text{obs}}, \mathbf{P}_N) &\propto \rho(\mathbf{d}_{\text{obs}} | \mathbf{P}_\beta, \mathbf{P}_N) \cdot \rho(\mathbf{P}_\beta) \\
&\propto \exp\left(-\frac{1}{2}(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)))^T \mathbf{C}_d^{-1} (\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)))\right) \cdot \\
&\quad \times \exp\left(-\frac{1}{2}(\mathbf{P}_{\beta, \text{prior}} - \mathbf{P}_\beta)^T \mathbf{C}_{\mathbf{P}_\beta}^{-1} (\mathbf{P}_{\beta, \text{prior}} - \mathbf{P}_\beta)\right)
\end{aligned} \quad (3.4)$$

with ρ denotes the Gaussian probability density function. $\rho(\mathbf{d}_{\text{obs}} | \mathbf{P}_N, \mathbf{P}_\beta)$ is the likelihood function. $\rho(\mathbf{P}_N)$ and $\rho(\mathbf{P}_\beta)$ represent a priori information on the parameters \mathbf{P}_N and \mathbf{P}_β . \mathbf{d}_{obs} is the $(1 \times n)$ vector of observed data. $\mathbf{P}_{\beta, \text{prior}}$ and $\mathbf{P}_{N, \text{prior}}$ are the prior models (parameter assumptions) on the unknown parameters \mathbf{P}_β and \mathbf{P}_N , chosen by the modeler to constrain the inversion. \mathbf{C}_d ($n \times n$), $\mathbf{C}_{\mathbf{P}_N}$ ($m_{\text{CA}} \times m_{\text{CA}}$) and $\mathbf{C}_{\mathbf{P}_\beta}$ ($(m_{\text{CA}} + 1) \times (m_{\text{CA}} + 1)$) are covariance matrices of the expected uncertainties on data and the prior models. $\mathbf{C}_{\mathbf{P}_N}$ and $\mathbf{C}_{\mathbf{P}_\beta}$ can also be seen as weighting values in the objective function to constrain the inversion result to have subspaces property values and directions remaining close to the parameters chosen in the prior models. The maximization of the posterior probability densities (Equations (3.3) and (3.4)) can be achieved by a minimization of the following objective functions in the inversion process (Tarantola and Valette 1982):

$$\begin{aligned}
\psi_{\text{structure}}(\mathbf{P}_N) &= \frac{1}{2}(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)))^T \mathbf{C}_d^{-1} (\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta))) \\
&\quad + \frac{1}{2}(\mathbf{P}_{N, \text{prior}} - \mathbf{P}_N)^T \mathbf{C}_{\mathbf{P}_N}^{-1} (\mathbf{P}_{N, \text{prior}} - \mathbf{P}_N)
\end{aligned} \quad (3.5)$$

$$\begin{aligned}
\psi_{\text{properties}}(\mathbf{P}_\beta) &= \frac{1}{2}(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)))^T \mathbf{C}_d^{-1} (\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta))) \\
&\quad + \frac{1}{2}(\mathbf{P}_{\beta, \text{prior}} - \mathbf{P}_\beta)^T \mathbf{C}_{\mathbf{P}_\beta}^{-1} (\mathbf{P}_{\beta, \text{prior}} - \mathbf{P}_\beta)
\end{aligned} \quad (3.6)$$

This minimization can be achieved iteratively with sequential optimizations on the geometry of the structure and on the values taken by the properties. The convergence of these two objective functions to their minimal values depend on the parameterization of the model and its initialization. A global minimization is not guaranteed, as the result of the inversion depends of the initial model. However the optimum can be explored by leading several inversion starting from different initial models. The different steps of the CADI algorithm are presented in Figure 3.5. The structural and property values parameters in \mathbf{P}_N and \mathbf{P}_β are first

initialized to generate the initial model. This initialization consists in assigning an initial reasonable direction of generation and initial property values $\beta_{\text{structure}}$ and $\beta_{\text{background}}$ to each CA subspaces in the model. After this initialization part, a sequential inversion process will firstly conduct an iterative structural optimization in which the CA structural generation process will regenerate the model with the updated parameters at each iteration. Once this optimization is completed, the inversion will continue with an optimization of the property values for the previously inverted structure. The process is then ended by an estimation of uncertainties on the structure geometry and on the property values. These different parts of the inversion process are detailed in the following sections.

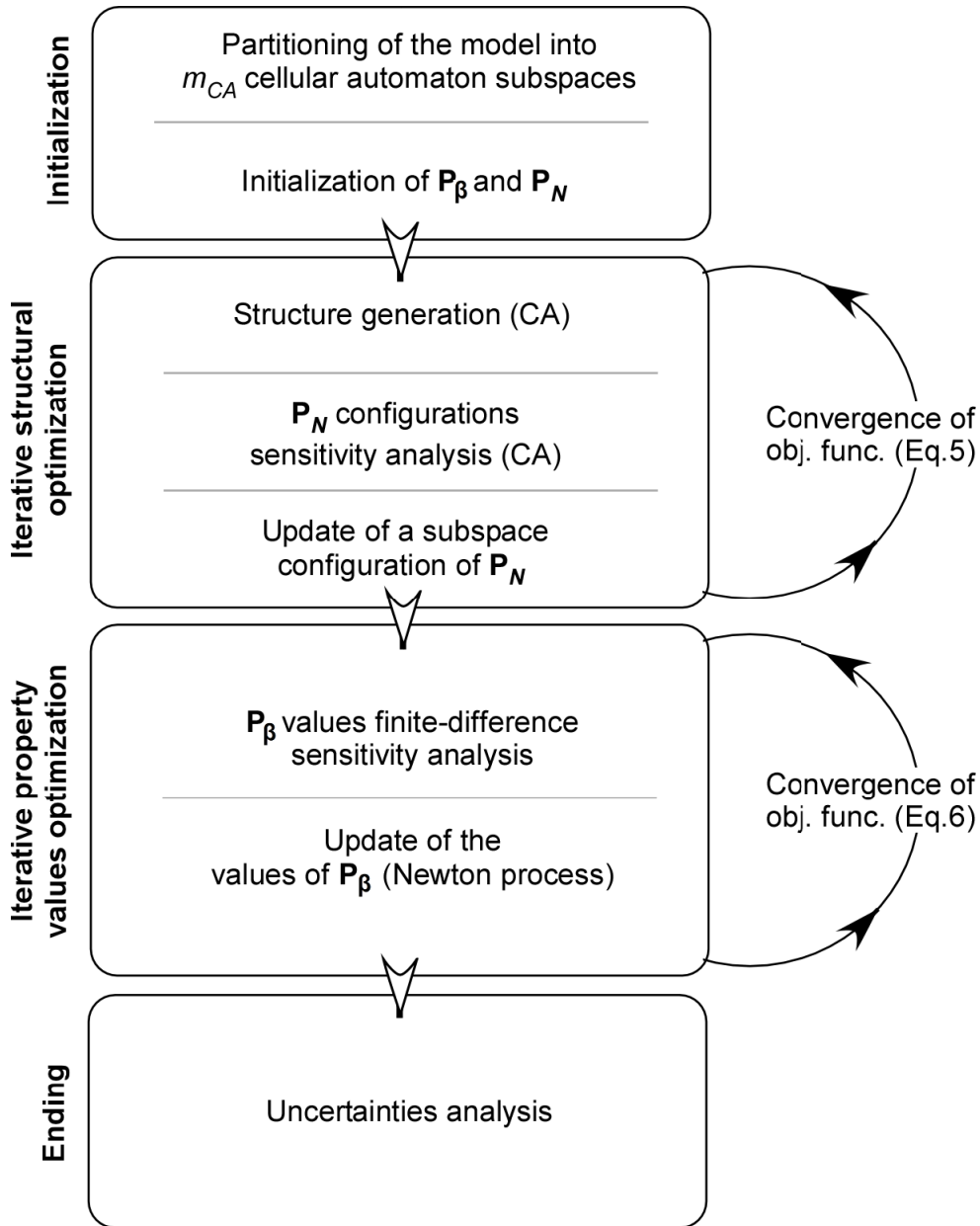


Figure 3.5: Operating scheme for the Cellular Automata-based Deterministic Inversion (CADI) algorithm. After an initialization of \mathbf{P}_N and \mathbf{P}_β with chosen directions and property values for each subspace, the algorithm begins an iterative process. It will firstly optimize the geometry of the structure in the model by iteratively updating the structural model using the CA generation process. Once the objective function has converged to a local minimum on the structure, it will lead a second optimization on the values of the properties for the previously inverted structure, until the objective function converges to a local minimum again. Finally, the uncertainties on the structure and the properties of the model are estimated.

3.2.3 Optimization process

3.2.3.1 Structural optimization

Initially, a chosen set of probable property values $\mathbf{P}_{\beta,ini}$ and chosen direction configurations $\mathbf{P}_{N,ini}$ are assigned to the piloted model to build an initial model. The aim of the structural optimization will be to modify iteratively the structure piloted by \mathbf{P}_N for a given distribution of \mathbf{P}_β until the convergence of the objective function (Equation (3.5)). The modification of the configuration \mathbf{P}_N is defined through a sensitivity analysis.

At an iteration step k , the sensitivity on the structure is estimated by introducing ‘perturbations’ in the generation and by analyzing the responses by solving the forward problem. Here a perturbation consists in a modification of the configuration N in a CA subspace of the model (i.e. a local direction modification), the other subspaces configurations remaining unchanged. The structural inversion sensitivity analysis tests the 8 configuration possibilities $N_i, i \in [1, 8]$ as parameter in each CA subspace in order to optimize the geometry of the structure regarding the objective function. This sensitivity analysis is led on the entire model to create a $8 \times m_{CA}$ sensitivity matrix \mathbf{S} . Thus, at the k th iteration and for a perturbation using a configuration N_i in a subspace j , the element (i, j) of the sensitivity matrix is defined as:

$$\begin{aligned} \mathbf{S}^k(i, j) = & \frac{1}{2} \left(\mathbf{d}_{obs} - f \left(\Gamma \left(\mathbf{P}_N^k \Big|_{\mathbf{P}_N^k(j)=N_i}, \mathbf{P}_\beta \right) \right) \right)^T \mathbf{C}_d^{-1} \left(\mathbf{d}_{obs} - f \left(\Gamma \left(\mathbf{P}_N^k \Big|_{\mathbf{P}_N^k(j)=N_i}, \mathbf{P}_\beta \right) \right) \right) \\ & + \frac{1}{2} \left(\mathbf{P}_{N,prior}(j) - N_i \right)^T \mathbf{C}_{\mathbf{P}_N}^{-1} \left(\mathbf{P}_{N,prior}(j) - N_i \right) \end{aligned} \quad (3.7)$$

where $f \left(\Gamma \left(\mathbf{P}_N^k \Big|_{\mathbf{P}_N^k(j)=N_i}, \mathbf{P}_\beta \right) \right)$ represents the modeled data through this perturbation for a given model of \mathbf{P}_β , and $\mathbf{P}_{N,prior}(j) - N_i$ represents the angular gap between the prior sub-direction and the perturbation direction. Here, the sensitivity analysis does not involve variations in \mathbf{P}_β , it determines all possible variations of the objective function for a single modification in the structure geometry.

The best improvement is found with the index $(i, j)_{\min}$ in matrix \mathbf{S} , representing the minimal value in the matrix which will give the best improvement for the minimization of the objective function (i gives the updated configuration N_i for the CA in the subspace j of the model). Thus, from a structural parameter set \mathbf{P}_N^k , the optimized set \mathbf{P}_N^{k+1} is built as $\mathbf{P}_N^{k+1} = \mathbf{P}_N^k$ except for its index j : $\mathbf{P}_N^{k+1}(j) = N_i$. By updating the subspace which gives the best structural improvement, a new structure will be generated for the iteration $k+1$. The algorithm reproduces the same sensitivity analysis for each iteration until the convergence of the objective function. The total number of forward problems evaluations for a structural iteration is $8 \times m_{CA} + 1$ (with $8 \times m_{CA}$ evaluations for the sensitivity analysis and 1 for the updated objective function calculation).

At the end of the inversion process, the uncertainties on the inverted structure are estimated through an uncertainties analysis on each subspace of the structure. Due to difficulty to infer the posterior covariance matrix of the structural inversion, this analysis is done for each subspace by inverting the difference between the posterior objective function and the sum of sensitivity values for all CA configuration possibilities, and the prior uncertainties for the subspace j :

$$\mathbf{S}^{post}(j) = \left(\frac{1}{8} \sum_{i=1}^8 \mathbf{S}(i, j) - \Psi_{structure}^{post} + \mathbf{C}_{\mathbf{P}_N}^{-1}(j, j) \right)^{-1}, \quad (3.8)$$

with $\mathbf{S}^{post}(j)$ is the posterior structural uncertainty for a subspace j in the model, i denotes the different rows of the matrix \mathbf{S} of the last iteration and $\Psi_{structure}^{post}$ is the value of the minimized objective function after convergence. If a subspace is well-constrained, its value \mathbf{S}^{post} should be low (another structure direction would have a negative impact in the minimization of the objective function), and if not, this value should be high (another structure direction would be quite neutral in the minimization of the objective function).

3.2.3.2 Property values optimization

Once the structure is optimized, the property parameters \mathbf{P}_p taken by the CA subspaces of the model are then also iteratively optimized, for the inverted structure; using a finite difference approach for the sensitivity analysis (for $m_{CA} + 1$ unknown property values to

optimize, including m_{CA} $\mathbf{p}_{\text{structure}}$ values plus one common value for $\mathbf{p}_{\text{background}}$). The Jacobian sensitivity matrix \mathbf{J} ($n \times (m_{CA} + 1)$), for an index (i, j) is defined as:

$$\mathbf{J}(i, j) = \left. \frac{\partial f_i}{\partial \mathbf{p}_j} \right|_{\mathbf{p}_j(j) = \mathbf{p}_j(j) + \Delta \mathbf{p}_j}, \quad (3.9)$$

with f_i the forward problem on a data i for a variation $\Delta \mathbf{p}_j$ of $\mathbf{p}_j(j)$. Here, $\Delta \mathbf{p}_j$ is the finite difference step.

The new values \mathbf{p}_j^{k+1} from a previous set \mathbf{p}_j^k are calculated from a linearization of Equation (3.6):

$$f(\mathbf{p}_j^{k+1}) \approx f(\mathbf{p}_j^k) + \mathbf{J}^k \cdot (\mathbf{p}_j^{k+1} - \mathbf{p}_j^k). \quad (3.10)$$

The optimization of \mathbf{p}_j is achieved via a Newton iterative process, initialized at a reasonable $\mathbf{p}_{j,\text{ini}}$ (Tarantola and Valette 1982). For the $k+1$ th step in the iterative process:

$$\begin{aligned} \mathbf{p}_j^{k+1} = & \mathbf{p}_j^k + \left((\mathbf{J}^k)^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{J}^k + \mathbf{C}_{p_j}^{-1} \right)^{-1} \cdot (\mathbf{J}^k)^T \cdot \mathbf{C}_d^{-1} \cdot (\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{p}_N, \mathbf{p}_j^k))) \\ & + \mathbf{C}_{p_j}^{-1} \cdot (\mathbf{p}_{j,\text{prior}} - \mathbf{p}_j^k) \end{aligned} \quad (3.11)$$

The total number of forward problems in an iteration for the property value sensitivity evaluation in a model with a uniform background will be $m_{CA} + 1$, while for a model with a varying background among the CA subspaces it would require $2 \times m_{CA}$ evaluations. Then the compute of the updated objective function requires one more forward problem evaluation.

The uncertainties on the values of properties, calculated at the end of the inversion process, are given by the diagonal entries of the posterior covariance matrix:

$$\mathbf{C}_{\mathbf{p}}^{post} = \left(\left(\mathbf{J}^{post} \right)^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{J}^{post} + \mathbf{C}_{\mathbf{p}}^{-1} \right)^{-1} . \quad (3.12)$$

These values represent variances of the properties. Then, the square root of the diagonal entries represent their standard deviation.

3.2.4 Applications

The CADI algorithm has been tested on six theoretical study cases for a linear inversion of a simple structure (Study case 1), a more complex structure (Study case 2), a complex multi-directional structure (Study case 3) and for a linear, non-linear and joint inversions (Study cases 4, 5 and 6) of a geostatistical generated structure. For these different examples we did not use any prior information on the structure in $\mathbf{P}_{N,prior}$ but we incorporated constant measurement errors in a diagonal matrix $\mathbf{C}_d = \sigma_{data}^2 \cdot Id(n)$, and prior background and structure property values in $\mathbf{P}_{\beta,prior}$ with their covariances in a diagonal matrix $\mathbf{C}_p = \sigma_{\beta}^2 \cdot Id(m_{CA} + 1)$. These six study cases and their results are presented in the Table 3.1, and the theoretical true structures to be reproduced are presented in Figure 3.6.

Table 3.1: Inversion results obtained for the 6 different study cases. This table includes the inversion type, the number of cells of the model, the partitioning and the observed data considered in the inverse modeling, the error variance of data, the number of iteration necessary to the convergence of the inversion process, the proximity between inverted data and observed data (R^2) and between the inverted structure and the true one pixel wise (structural similarity), and the inversion time. In Case 3, NoInit=Initial simple model and Init=Initial more complex model. In Case 6, NL=Non-linear and L=Linear.

	Inversion type	Number of cells (m)	Number of data	Number of iteration	Data R^2	Structural similarity	Inversion time
		CA grid	Error variance				
Case 1 (Figure 3.6a)	Linear	3,600 (3 × 3)	358 $\sigma_{data}^2 = 1$ ms	4	0.99	99.7 %	4 min
Case 2 (Figure 3.6b)	Linear	10,000 (5 × 5)	598 $\sigma_{data}^2 = 1$ ms	21	0.96	97.9 %	1.3 h
Case 3 (Figure 3.6c)	Linear	48,400 (11 × 11)	1,318 $\sigma_{data}^2 = 1$ ms	NoInit: 26 Init: 30	0.91 0.96	97.1 % 98.3 %	13.2 h 18.8 h
Case 4 (Figure 3.6d)	Non-linear	3,600 (3 × 3)	128 $\sigma_{data}^2 = 0.1$ m	7	0.98	85.8 %	13 min
Case 5 (Figure 3.6d)	Linear	3,600 (3 × 3)	358 $\sigma_{data}^2 = 1$ ms	4	0.99	82.6 %	4 min
Case 6 (Figure 3.6d)	Joint: Linear + Non-linear	3,600 (3 × 3)	486 $\sigma_{data}^2 = 1$ ms, $\sigma_{data}^2 = 0.1$ m	7	NL=0.98 L=0.99	88.2 %	20 min

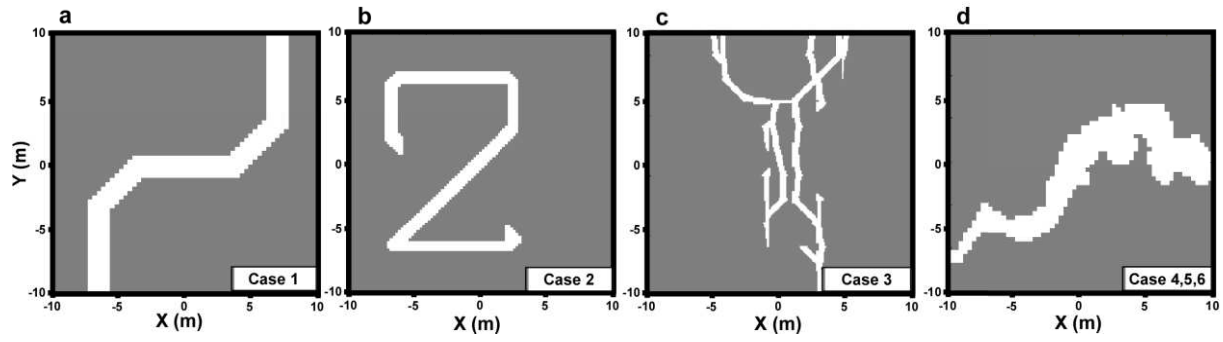


Figure 3.6: Presentation of the 4 different structures tested in the 6 study cases in this paper. The case 1 is a linear inversion of a simple geometry (a) to show how the optimization works. The case 2 is a linear inversion of a more complex geometry (b). The case 3 is a linear inversion of a complex multi-directional linear structure (c). The cases 4, 5 and 6 are linear, non-linear and joint inversion of a geostatistical generated geometry (d), appearing as a more natural structure.

3.2.4.1 Study case 1

The first study case is a linear inversion of a simple structure. The purpose of this study case is essentially to illustrate how the optimization within the CA subspaces in the structural inversion works. For the linear inversion, we considered seismic data. The properties taken into account in the model are the seismic velocities. Our simple synthetic model (Figure 3.6a) is set up as a field of $20 \times 20 \text{ m}^2$ with a perfectly uniform matrix (background) with a seismic velocity of 3.33 km/s, and an empty conduit (structure) of 0.26 km/s. The model properties are discretized in a 60×60 regular grid. Seismic transmitters and receptors are set up around the theoretical model, at the beginning and end of each row, column and diagonal of the model grid, which would correspond to a device every 33 cm. The observed data consists on travel time from seismic waves traveling through the model. The seismic wave travel time is calculated by summing the products between inverse of seismic velocity and distance traveled in each cell swept by the wave following the shortest path in the grid (in this case by summing cells in rows, columns and diagonals). If each cell of the grid swept divides the path followed by the wave, the total travel time of the wave is:

$$t_{\text{wave}} = \sum_{i=1}^{\eta} \frac{1}{s_i} \cdot \Delta x_i, \quad (3.13)$$

where t_{wave} designs the travel time of the wave (in ms), $i \in [1, \eta]$ identifies the different η cells swept by the wave during its travel, s_i is their seismic velocities (m/ms or km/s) and Δx_i the distance traveled by the wave through these cells (m).

Thereby, 358 observed data were generated from the theoretical model and will be used for the inversion process. For this simple geometry structure, the inversion algorithm was conducted using a relatively coarse 3×3 CA partitioning, and by considering a simple straight structure initially (Figure 3.7a) with seismic velocities of 2 km/s for the background and 0.5 km/s in the structure. The covariance matrix C_p was generated with a seismic velocity variance of $\sigma_p^2 = 1$ km/s and the seismic velocities of the initial model were also taken as prior values in $P_{p,prior}$. The inversion converged in 4 iterations.

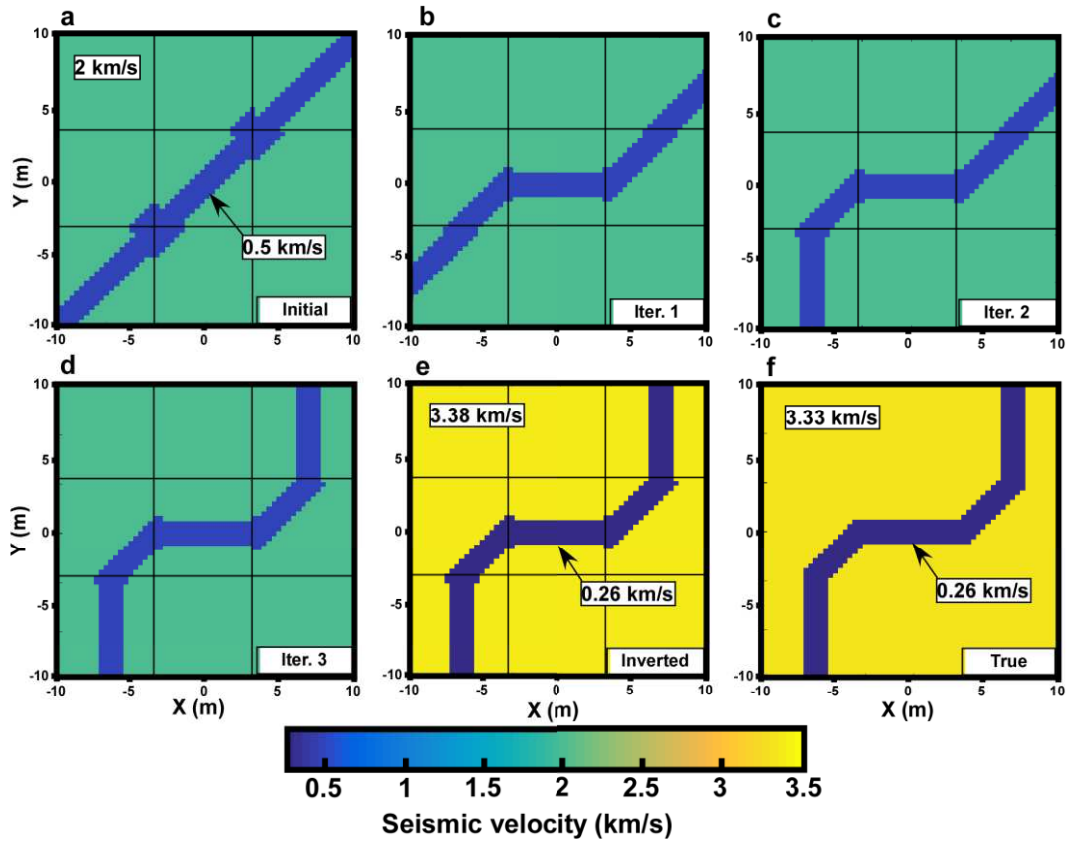


Figure 3.7: Result of the linear inverse modeling of the case study 1. The inversion finished after 4 iterations. This figure shows all different iterations of the inversion, from initial model (a), to inverted model (e). The true structure is shown in (f). The figure (d) corresponds to the structural optimization and the figure (e) to the properties optimization for this structure. The different CA subspaces of the model are highlighted by the black lines.

This case permits to understand how the CADI algorithm works. Each different step of the optimization of the model are presented in Figure 3.7. Starting from the initial structure in Figure 3.7a, for each next steps the optimization process tries to find new sub-directions improving the initial structure over the partitioning of the model (shown as a black grid in

Figure 3.7). At the first step (Figure 3.7b) the initial model was improved in its central part, and at the two next steps (Figure 3.7c and 3.7d) the angles of the lower left and upper right parts of the true model were found. The last step of the inversion in Figure 3.7e corresponds to the properties optimization in order to improve the objective function and find the true properties. With this parameterization of the inverse problem, the result for this study case reproduces the true structure (Figure 3.8) and the observed data (Table 3.1) well.

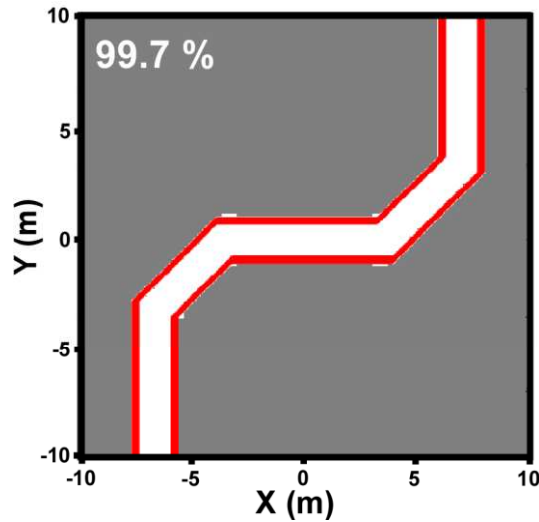


Figure 3.8: Comparison of the optimal structure found by inversion (in white) and the true structure (bold boundaries) for the case study 1. For this simple geometry, the inverse algorithm could easily reproduce the structure.

This simple case is useful to show how the CADI algorithm modifies at each step the geometry of the initial structure and thus to understand why the information and the partitioning chosen in the initial model will considerably influence the deterministic process (in term of time but also in term of result as we will see in study case 3).

3.2.4.2 Study case 2

The CADI algorithm was then applied on a more complex study case to test its capacities to reproduce complex geometries. The theoretical structure under consideration for this case is presented in Figure 3.6b. As in the first study case, the linear inversion is led by using seismic data, but with other seismic velocities for the matrix (2.5 km/s) and the conduit (0.26 km/s). The model was discretized as a regular grid of 100×100 cells. The observed data consist in 598 seismic time travel calculated in the same way than explained for the case study 1. This time,

in order to give more possibilities to the structural inverse process, the inverse model was partitioned in a 5×5 CA subspaces with the true property values initially known. The covariance matrix C_{β} was generated with a seismic velocity variance of $\sigma_{\beta}^2 = 1$ km/s and the seismic velocities of the initial model were also taken as prior values in $P_{\beta, \text{prior}}$.

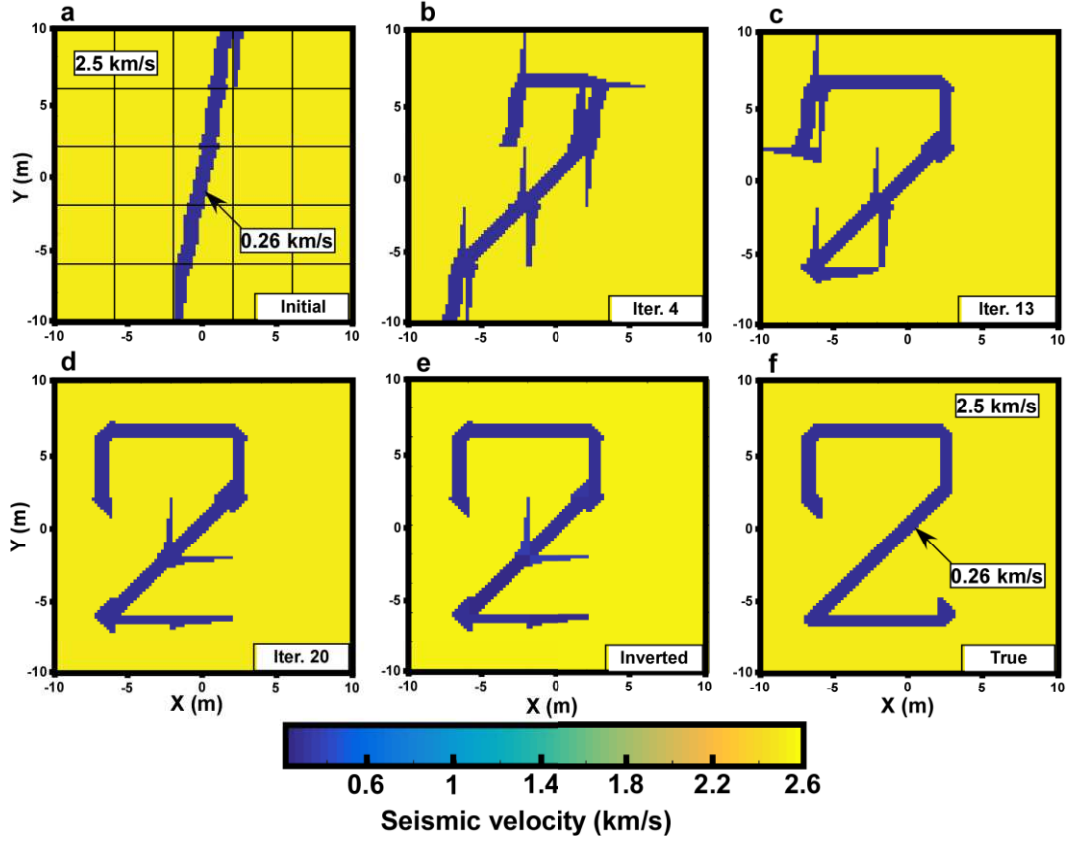


Figure 3.9: Result of the linear inverse modeling of the case study 2. The convergence is performed with 21 iterations. This figure shows some different iterations of the inversion, from initial model (a), to inverted model (e). The true structure is shown in (f). We noted that the optimization on the property values permits to balance the structural inversion errors. For example, in this case, the structural additions in the center of the model in (e) were optimized by a light augmentation of its seismic velocity (0.5 km/s instead 0.26 km/s).

Initially, we set up a straight linear structure (Figure 3.9a). The algorithm then converged in 21 iterations (it took approximately 1 hour with a computer with 2 processors Intel Xeon 2.4GHz of 16 cores). The Figure 3.9 shows several steps of the inversion process (Figure 3.9b, 3.9c and 3.9d), the optimized inverse model in Figure 3.9e and the true model in Figure 3.9f.

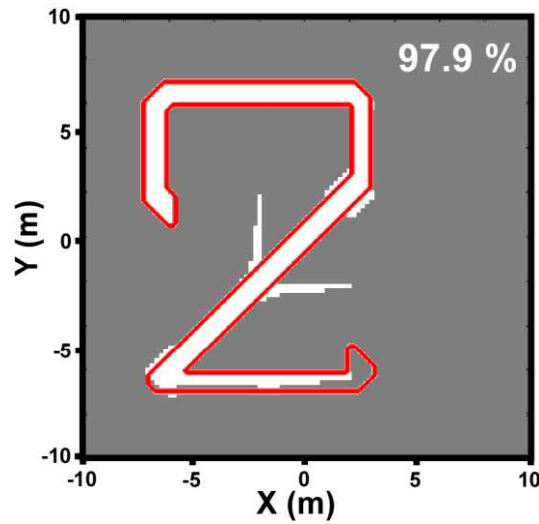


Figure 3.10: Comparison of the optimal structure found by inversion (in white) and the true structure (bold boundaries) for the case study 2. The optimization process reproduced a good structural inversion. The few inversion errors in the center of the model were lightly balanced by the inversion on the properties (see Figure 3.9).

For this case, the structural inversion is close to the real one but not perfect as in first study case (Figure 3.10) and we can show in Figure 3.9e that these imperfections are balanced by the properties optimization. Thus, the zone of the structure with an inverted structural part which doesn't exist in the true structure (central part) is corrected by a light increase of its seismic velocity in order to minimize the differences between calculated and observed data (0.5 km/s instead of 0.26 km/s). This correction tends to locally slightly approach the structural seismic velocity to the matrix seismic velocity and thus slightly reducing the existence of this local part of the structure in the model. This property values optimization permits a better convergence on the objective function and is, in some cases, useful to counterbalance the approximations of the structural optimization when the property values are initially well known.

3.2.4.3 Study case 3

In this third study case we applied the CADI algorithm on a complex multi-linear structures network. The study was done to show all the capacities of the CADI method to model fractured fields, which are equivalent to linear structures dispersing in multiple direction among the space. The theoretical structure under consideration for this case is presented in Figure 3.6c. The linear inversion is led by using seismic data generated with given seismic velocities for the matrix (2.5 km/s) and the conduit (0.33 km/s). The model was discretized as a regular properties grid of 220×220 cells. The observed data consist in 1,318 seismic time travel calculated in the

same way than explained for the case study 1. The inverse model was partitioned in an 11×11 CA subspaces with the true property values initially known to investigate the structural optimization capacities.

In this case we compared the sensitivity of the result to the initial model. We set two inversions with two different initial models, the first one being very simple and incorporating only a global direction of generation (Figure 3.11a) and the second one incorporating a bit more information on the structural shape (Figure 3.11d). The Figure 3.11 shows these initial models, the inverted models they have produced and the comparison between the inverted models and the true one (Figure 3.6c). For the first inversion the result is already close to the true model, especially considering the simplicity of the initial model. A simple straight structure can become a more complex multi-directional structure through the optimization process and find the main shapes and trends of a complex structure geometry which shows the possibilities given by the parameterization in the CADI algorithm. Starting from a different initial structure in the second inversion we arrived to a slightly better result on the geometry which becomes really close to the true one. It highlights the importance of incorporating some information in the initial model for the inversion process, but however if no information are known, the first inversion shows that even a very simple initial assumption can produce a good result.

3.2.4.4 Study case 4

For the fourth study case, a non-linear inversion has been led on a structure generated by a geostatistical technique using a directionally oriented variogram function with the package *gstat* in R (Figure 3.6d), which appears to be more natural than the previous structures. The steady state observed data have been produced by a hydraulic tomography with 4 alternate pumping wells and 32 measurement wells (for a total of 128 observed data) regularly distributed over the model. The positioning of the wells is presented in the model in Figure 3.12, this model is enclosed in a larger buffer zone ($1,000 \times 1,000 \text{ m}^2$) defined with a constant head (no drawdown) on its lateral boundary condition and a uniform 'background' transmissivity value.

The hydraulic transmissivities are considered as the unknown properties to be inverted in a model with a 60×60 cells grid. The theoretical model is set up as a matrix with a transmissivity of $1.6 \times 10^{-7} \text{ m}^2/\text{s}$ and a structure with an equivalent transmissivity of $5 \times 10^{-4} \text{ m}^2/\text{s}$.

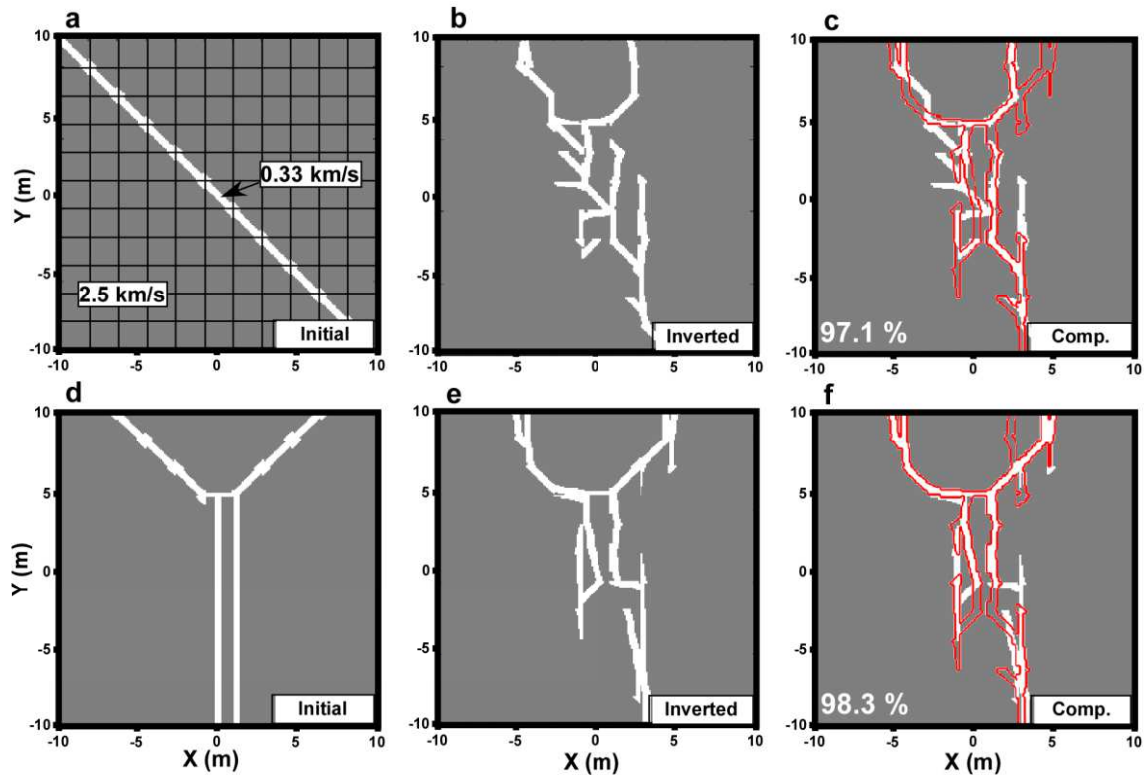


Figure 3.11: Results of the linear inverse modeling of the case study 3. This figure shows two inversions with different initial models (a and d), their results (b and e) and the comparison of these results to the true geometry boundaries in red (c and f). The convergence is performed with 26 iterations in the first inversion and 30 iterations in the second. We noted that the information contained in the initial model could slightly modify the result of the inversion but even with a very simple initial case (a) the optimization process permits to find the main shapes and trends of the true structure (c).

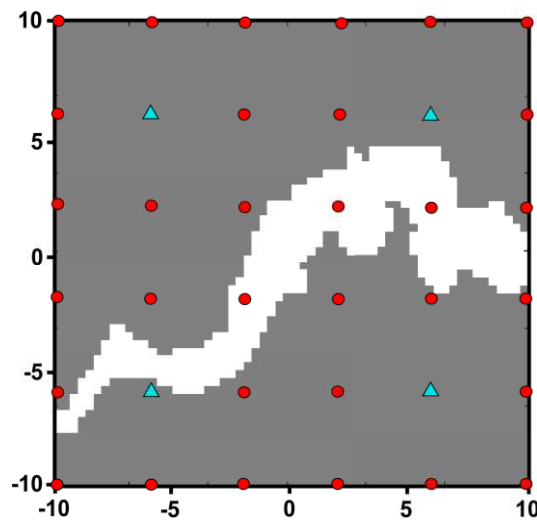


Figure 3.12: Map of the positioning of the wells for the hydraulic tomography inversion for the study case 4. The circles are the position of the measurement piezometers and the triangles are the position of the pumping wells.

A 3×3 CA partitioning was chosen for the inverse modeling with a good a priori information on the background properties and a structure transmissivity value of $10^{-3} \text{ m}^2/\text{s}$. The covariance matrix \mathbf{C}_β was generated with a multiplicative variance on the transmissivity of the form $10^{\pm\sigma_\beta^2}$ where $\sigma_\beta^2 = 1$ (± 1 variance on the transmissivity exponent) and the transmissivities of the initial model were also taken as prior values in $\mathbf{P}_{\beta,\text{prior}}$.

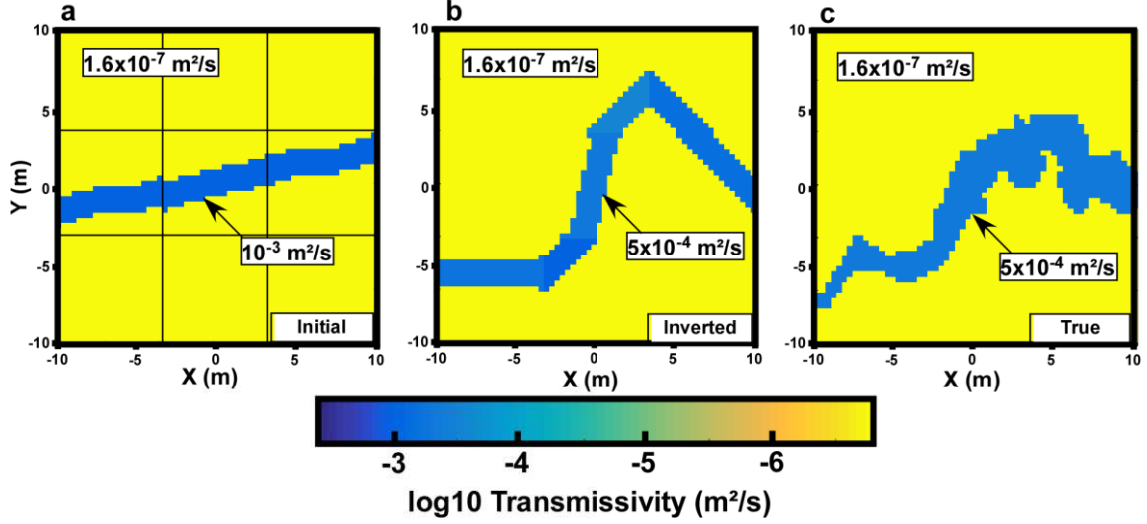


Figure 3.13: Result of the non-linear inverse modeling of the study case 4. The inversion finished after 7 iterations. This figure shows the initial model (a), the inverted model (b) and the true structure (c). The inversion process found an optimized equivalent structure to the initial property value. The true transmissivities were found during the properties optimization.

From a straight linear initial structure (Figure 3.13a), the inversion converged in 7 iterations and produced the model presented in Figure 3.13b. The global trends of the true structure were found although the CADI, as presented previously, produces structures which have a constant aperture. Thus, the inversion process found the best constant-aperture equivalent structure which reproduced the true model for the initial parameters. The properties optimization has permitted to find the true structure property value. However, the initial hydraulic properties did not permit to find the best fitting structure to the true model (Figure 3.13c). We will show in study case 6 that with the same initial model in a joint inversion, we can have both the true property values and a better fitting structure geometry.

3.2.4.5 Study case 5

A linear inversion has been led on the same geostatistical-generated structure than in study case 4 (Figure 3.6d). This time, observed data have been produced by seismic, as presented in study case 1 (producing 358 observed data). The seismic velocities are considered as the properties to be inverted in a model with a 60×60 cells grid. The theoretical model is set up as a matrix with a seismic velocity of 2 km/s and a structure with a seismic velocity of 0.26 km/s. A 3×3 CA partitioning was chosen for the inverse modeling with close initial property values (0.2 km/s for the structure and 2.5 km/s for the background). The covariance matrix C_β was generated with a seismic velocity variance of $\sigma_\beta^2 = 1$ km/s and the seismic velocities of the initial model were also taken as prior values in $P_{\beta, \text{prior}}$.

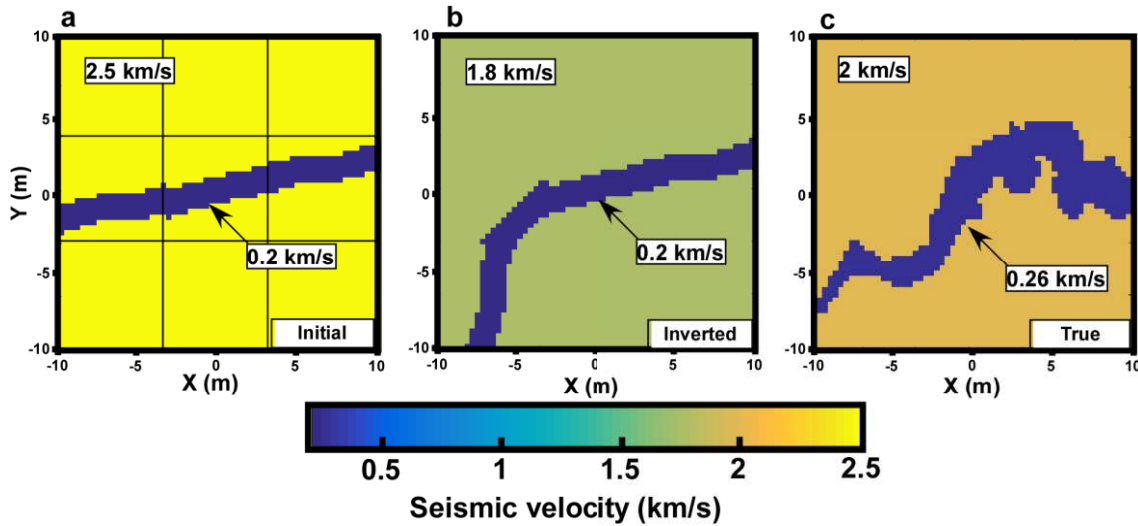


Figure 3.14: Result of the linear inverse modeling of the study case 5. The inversion finished after 4 iterations. This figure shows the initial model (a), the inverted model (b) and the true structure (c). The structural optimization was limited by the initial properties and by its constant aperture generation to reproduce a variable aperture true structure. In this case, the optimization on the property values permits to balance the initial information and the structural inversion aperture limitations. The properties optimization balanced this limitation by globally decreasing the seismic velocity of the background to a lower value than the true one.

The inversion converged in 4 iterations, the results are presented in the Figure 3.14. With the same initial structure (Figure 3.14a) than the hydraulic inversion, the seismic inversion produced a slightly different equivalent structure which approximatively reproduces global trends of the true structure but is not the best fitting possibility. The properties optimization (Figure 3.14b) has permitted to balance the structural approximations caused by the limits of

the initial information and the generation of a constant-aperture structure. Thus, in the properties optimization, the seismic velocity of the background was decreased under the value of the true one to counterbalance the lower aperture of the generated structure. Thereby, the properties optimization part can bring more flexibility to the algorithm, which is constrained in its structural part by the prior information in the initial condition and its constant aperture. However the inversion process could not truly reproduce the structure and the properties of the true model for the initial parameters. As for the previous non-linear inversion, we will show in the next study case that a joint inversion permits to reproduce both the property values and a better structure geometry for the same initial parameters.

3.2.4.6 Study case 6

For the last study case, a joint inversion has been led on the same structure generated by a geostatistical approach than in study cases 4 and 5 (Figure 3.6d). The joint inversion is a simultaneous inversion of different data sets with a same unique inverted structure which has to be able to reproduce the information contained in all different data sets. The information brought by different investigation techniques will reduce the non-uniqueness of the inverse solution, each techniques bringing different information on the parameters (*Haber and Oldenburg 1997*). We have jointly inverted the hydraulic data from study case 4 and the seismic data from study case 5. The joint objective function in this case is a weighted linear combination of the seismic properties objective function and the hydraulic properties objective function. We chose a weighting in order to have initially approximately the same value for each of the two parts of the joint objective function. The observed data were produced by hydraulic tomography and seismic (for a total of 486 observed data). The hydraulic transmissivities and seismic velocities are considered as the properties to be inverted in a model with a 60×60 cells grid. The theoretical model is set up with the same property values as presented in study case 4 for the hydraulic properties (1.6×10^{-7} m²/s for the matrix and 5×10^{-4} m²/s for the structure) and in study case 5 for the seismic properties (2 km/s for the matrix and 0.26 km/s for the structure). A 3×3 CA partitioning was chosen for the inverse modeling with the same initial parameterized model than in study case 4 and 5. The covariance matrices \mathbf{C}_β were generated with a seismic velocity variance of $\sigma_\beta^2 = 1$ km/s and a multiplicative variance on the transmissivity of the form $10^{\pm\sigma_\beta^2}$ where $\sigma_\beta^2 = 1$ (± 1 variance on the transmissivity exponent), and the seismic velocities and hydraulic transmissivities of the initial models were also taken as prior values in $\mathbf{P}_{\beta,\text{prior}}$.

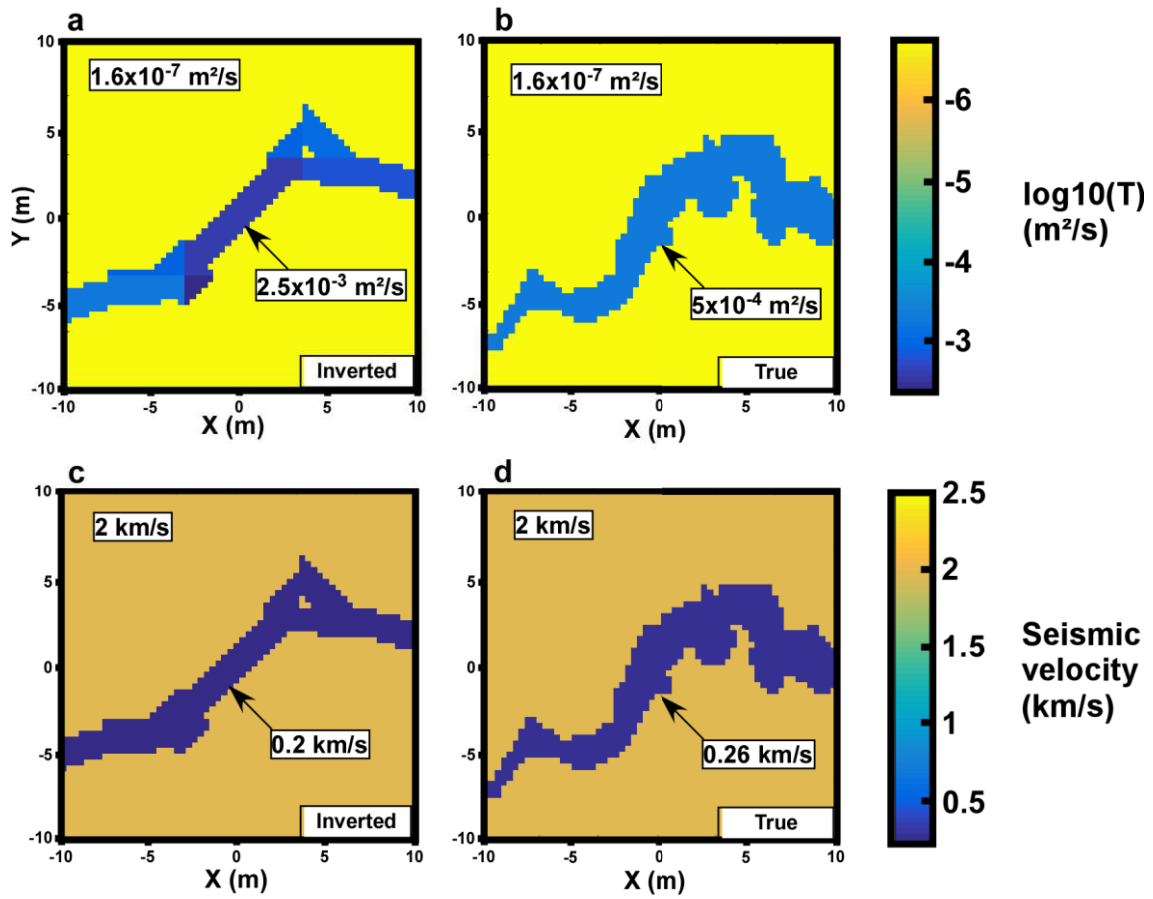


Figure 3.15: Result of the joint inverse modeling of the study case 6. The inversion finished after 7 iterations. This figure shows the hydraulic model (a), the seismic model (c) and the true models (b,d). The geometry of the structure in the models was optimized through a joint inversion of seismic and hydraulic data.

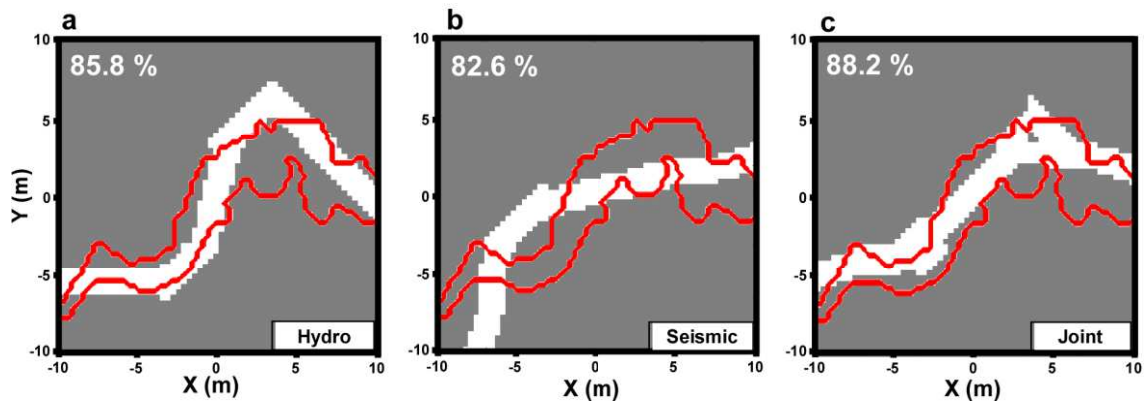


Figure 3.16: Pixel-wise comparison of the optimal structures found by inversion (in white) and the true structure (bold boundaries) for the study cases 4 (a), 5 (b) and 6 (c). Both hydraulic and seismic data permitted to find a geometry of the global trends of the true structure but the joint inversion resulted to a better model regarding the structure and also the convergence on the data, which avoided the difficulties encountered by the simple hydraulic inversion and the simple seismic inversion.

The inversion converged in 7 iterations. The results of the inversion for each method and the true models are presented in Figure 3.15. Compared to the hydraulic and seismic separate inversion, the joint inversion produced better results on the data (Table 1) and on the inverted structure (Figure 3.16) which is closer to the true structure. The structural joint inversion permits to combine the hydraulic and seismic data to find the best structure. The optimized property values are also better in the joint inversion than in the separate ones. The properties optimization permits to counterbalance the limitations of a constant aperture structure by keeping a modeled higher value of transmissivity for the structure regarding the true value to simulate a thicker structure. The lower left part of the structure was optimized with a transmissivity close to the real one because the true structure is thinner in this part. The seismic velocity of the structure was also kept at a lower value than the true one to counterbalance to constant aperture geometry.

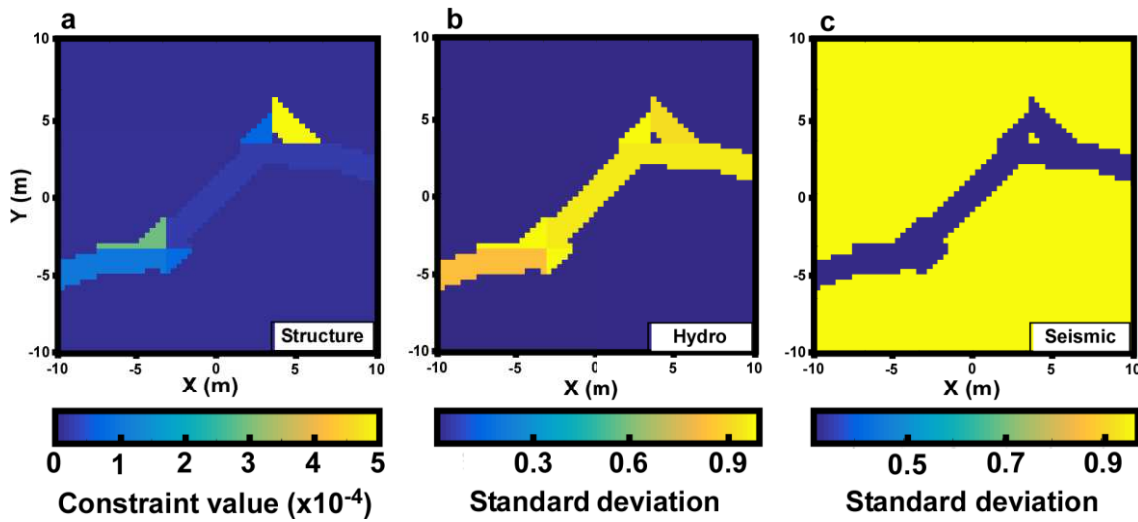


Figure 3.17: Uncertainties analysis for the joint inversion of the study case 6. The structural constraint in (a) indicates where the structure of the model is well-constrained by a low value, and at the opposite, a high value indicates an uncertainty for its subspace direction. The properties uncertainties for (b) the hydraulic transmissivity and (c) the seismic velocity are quantified by a standard deviation on the inverted values.

The uncertainties analysis on the structure and the property values are reported in Figure 3.17. The structure is well constrained by the data, except in the lower left part where the true structure is thinner and the upper right part for the angle of the structure. This means that another close sub-direction of the inverted structure would not affect significantly the results. In the parameterization of the joint inversion more accuracy was given to the hydraulic data, therefore the uncertainties on the seismic properties are more important than those on the

hydraulic ones. The uncertainties on the hydraulic property values vary locally within the structure. The lower left part has fewer uncertainties on the properties because its aperture is closer to the true one and therefore this part is globally closer to the true structure. The background value is well constrained because the true value was considered as a priori known. The seismic properties uncertainties are more important for its background. This is caused by the high properties difference between structure and background and because the background property was initially not known.

3.2.5 Discussion and conclusion

The Cellular Automata-based Deterministic Inversion (CADI) algorithm is an especially adapted method for linear structure geometries. The inversion process is based on a Bayesian approach and a sequentially optimization of the structure geometry and property values. The structural optimization is monitored by cellular automaton to generate the structure, and by a configurable partitioning of the model into subspaces which permits a monitoring of the complexity of the inverted structure. One can choose a coarse subspace partitioning for simple structures and for a fast inversion process, or a fine subspace partitioning for inversion of more complex structures. The property values optimization brings more flexibility to the inversion by slightly modifying the values of the properties in the structure. This optimization permits to counterbalance some approximations in the structural optimization and some constraints from the initial information.

The CADI algorithm parametrization is mainly focused on the structural optimization, therefore it considers only two units: a constant-aperture structure and a background, which is considered as a unique uniform unit or with an intern variability which is negligible regarding the variability with the structure. Therefore, and as for any other inversion methods, it is especially effective for specific structural cases. Furthermore, the limits of the CADI algorithm have to be clearly identified in order to make a good use out of it and to have a critical view on the results it can produce. For this purpose, the main advantages and limits of the CADI algorithm have been summarized in Table 3.2. For each limit of the algorithm, an appropriate solution has been suggested.

In this paper we promote the potential of the CADI algorithm to image the complex linear structures, exploiting its capacity to reproduce large-scale structures in a relatively short time. As far as we know, the CADI algorithm is the first algorithm which permits the deterministic inversion of linear structures (global direction-oriented structure characterized by an aperture significantly lower than its length) with a dynamic structural optimization. This first attempt is mainly focused on the general presentation of the method and the theory of the algorithm ; but, we believe that this method can be improved and inspire other ones in various domain. For example, with the same algorithm structure, and by changing the cellular automaton configuration rules, it is conceivable to generate other types of forms than linear structures.

Table 3.2: This table presents the main advantages provided by the CADI algorithm. The limits of the methods are also listed with a suggested solution for each limit.

Advantages	Limits	Solutions
<ul style="list-style-type: none"> The complexity of the structural optimization can be monitored with a configurable partitioning of the model. 	<ul style="list-style-type: none"> The cellular automata parameterization permits only the formation of structure with a constant aperture all along in the structural optimization. 	<ul style="list-style-type: none"> ➤ The property values optimization permits to digitally balance some local variation of aperture. Thus, some results on property values can be structurally interpreted.
<ul style="list-style-type: none"> The model properties are monitored by pilot cellular automaton, which permits to easily handle with large-scale modeling and makes a sensitivity analysis possible to accelerate the optimization. 	<ul style="list-style-type: none"> Only binary pattern are considered (structure and background). The background is considered as invariable regarding the variation structure / background and intra-structure. 	<ul style="list-style-type: none"> ➤ If the algorithm is applied on a case where the background has significant intern variability, a particular attention should be paid on setting an appropriate equivalent background.
<ul style="list-style-type: none"> The convergence of the inversion is constraint to a local solution regarding the prior information which can be easily incorporated in the objective function. 		

We also plan further works with the presented algorithm, especially for improvements on its capacities (by adding a third ‘micro-structures’ state within the background) and for field application cases, with a higher consideration on additional prior information (as the tortuosity factor of the structure) and on sensitivity analysis of the method.

4 UN NOUVEL OUTIL D'INVERSION STRUCTURALE POUR LES MODÈLES DISCRETS COUPLÉS

4.1 Contexte

Une seconde méthode d'inversion, alternative à celle développée dans le chapitre précédent, est présentée dans ce chapitre 4. Contrairement à la méthode précédente, reposant sur un modèle équivalent à un milieu poreux, cette nouvelle méthode se base sur un modèle couplé dans lequel le réseau de fractures et conduits est représenté sous forme discrète par des lignes, et le fond, représentant la matrice, est représenté par un continuum équivalent milieu poreux. Ce modèle permet la représentation des contrastes inhérents aux milieux fracturés et karstiques en attribuant des distributions de valeurs de propriétés spécifiques au réseau discret et au continuum de fond.

L'optimisation des valeurs de propriétés et de la forme du réseau discret dans cette nouvelle méthode est permise par un découpage du modèle en sous-espace. La propagation du réseau discret dans le modèle suit une règle simple de « nœud à nœud » et chaque sous-espace du modèle pilote indépendamment la direction de propagation locale du réseau, ainsi que les valeurs locales de propriétés du réseau et du continuum.

Dans un premier temps, ce chapitre aborde le problème direct et le paramétrage du modèle couplé, ainsi que la règle de propagation du réseau discret. Ensuite, les équations du problème inverse sont présentées avec le processus d'optimisation et la quantification des incertitudes *a posteriori*. Le processus d'optimisation permet des modifications à la fois des valeurs distribuées de propriétés, mais également de la forme du réseau discret. La méthode est finalement testée sur trois applications de cas d'études synthétiques à partir de réponses hydrauliques. A partir des différents résultats des cas tests, une discussion est menée sur les conditions nécessaires pour produire une tomographie satisfaisante à partir de cette méthode. Elle aborde notamment la question de la quantité et la répartition des forages de mesures et les stratégies de modélisation inverse, en particulier l'intérêt d'intégrer des informations au modèle initial et la possibilité de coupler cette méthode à une inversion multi-échelle afin d'améliorer la qualité du résultat.

Cette méthode, comme la méthode précédente, représente un outil intéressant pour l'interprétation de réponses de milieux fracturés et karstiques. La discussion menée dans ce chapitre concernant le choix du modèle initial et les limites d'une inversion séquentielle permettra également de définir une stratégie adaptée pour l'application de l'inversion dans le cas réel à partir des données de terrain, présenté dans le chapitre suivant de ce manuscrit.

4.2 Hydraulic tomography of discrete networks of conduits and fractures in a karstic aquifer by using a deterministic inversion algorithm

Cette partie est composée de l'article « Hydraulic tomography of discrete networks of conduits and fractures in a karstic aquifer by using a deterministic inversion algorithm », publié dans le journal *Advances in Water Resources* en février 2018 (*Fischer et al. 2018a*). Le texte a été remis en forme au format du manuscrit. La version originale de l'article est donnée en Annexe 3.



4.2.1 Introduction

In hydrogeological studies, the choice of the management and protection strategies of the groundwater resources is mainly based on the characterization of the hydraulic properties of the aquifer, such as hydraulic conductivity and specific storage. This characterization is most often carried out from pumping, slug and tracer tests by intrusively recording the aquifer responses, such as hydraulic pressure and tracer concentration at a set of boreholes (*Butler 2005*). The reliability of these techniques for capturing the spatial heterogeneity of the hydrodynamic properties is particularly conditioned by the amount and spatial disposition of wells used during the investigation, and the procedure applied to analyze the hydraulic data (*Yeh and Lee 2007*). In karstic and/or fractured aquifers the hydrodynamic properties (such as the hydraulic conductivity) can vary significantly from 10^{-10} m/s to 10^{-1} m/s, even at small scales (*Wang et al. 2016*). This heterogeneity mainly depends on the apertures, connectivity and density of the conduits and fractures network in the medium, making the groundwater flow path complex (*Eisenlohr et al. 1997 ; Kovacs et al. 2005 ; Borghi et al. 2016 ; Ronayne 2013*). In this complex context, the hydraulic flow pattern is spatially disconnected and principally focused in the transmissive fissures and fractured zones, wherein the geometrical features and hydraulic flow regime (turbulent or laminar) are usually difficult to identify, especially with a limited number of wells, or with the use of oversimplified assumptions for interpreting the piezometric data to infer the hydrodynamic parameters.

In the hydrosience literature, several different modeling approaches based on the physical theories have already been tested in order to simulate the dynamics of karstic flows for the prediction of hydraulic properties (*Hartmann et al. 2014a*). Among them, the equivalent porous media model, also called the single continuum model, in which the discrete features of fractures and karstic conduits are conceptualized as a porous media with continuous hydraulic properties (*Larocque et al. 1999 ; Illman 2014 ; Wang et al. 2016*). This simplifies the description of heterogeneity of karstic aquifers because it does not require an accurate knowledge on the architecture of fractures and conduits networks for simulating the groundwater flows. In such concept, it is sufficient to assign high hydraulic conductivity values to fractured zones and very low conductivity for intact rock. Otherwise, the coupled discrete-continuum distributed approach is of great interest thanks to its ability to imitate the dual hydrodynamic behaviors in the fractured aquifers by using Discrete Channel or Fracture Networks (DCN/DFN) for the conduits and fractures, and equivalent porous media for representing the matrix blocks (*Teutsch 1993 ; Liedl et al. 2003 ; De Rooij et al. 2013*). In contrast to the equivalent porous media model, the discrete-continuum approach requires a good

knowledge on the geometry of the karstic and fracture networks. The influence of the discrete network geometry on the hydraulic simulations and the benefits of a coupled discrete-continuum approach compared to the equivalent porous media have been widely discussed in the literature (*Kovacs 2003 ; Painter and Cvetkovic 2005 ; Ghasemizadeh et al. 2012 ; Hartmann et al. 2014a*). One of these advantages is its efficiency to reproduce numerically the hydraulic fluctuations of karst spring discharge, while an equivalent porous media systematically generated lower values than the ones measured (*Kovacs 2003*).

The hydraulic tomography is a useful tool to predict rigorously the spatial distribution of the hydraulic properties, or the structural architecture of the fractures and conduits and their properties. It involves the use of inverse algorithms to analyze jointly a set of hydraulic data collected during multiple pumping tests (*Carrera et al. 2005 ; Cliffe et al. 2011 ; Zhou et al. 2014*). In this framework, various inversion algorithms were successfully applied for characterizing the hydraulic properties of fractured and heterogeneous aquifers using both concepts of parametrization discussed in the previous paragraph: the equivalent porous media and coupled discrete-continuum approach.

(1) Inversions in an equivalent porous media were led by using geostatistical approaches in which the statistical characteristics of hydraulic properties are used as a priori information to constraint the inversion. Among these tools we cite: sequential successive linear estimator (*Yeh and Liu 2000 ; Ni and Yeh 2008 ; Hao et al. 2008 ; Illman et al. 2009 ; Sharmeen et al. 2012*), pilot-point (*Lavenue and de Marsily 2001*), transitional-probability (*Wang et al. 2017*), anisotropy directions (*Meier et al. 2001*), multi-scale resolution (*Ackerer and Delay 2010*), or structural approaches: probability perturbation method (*Caers and Hoffman 2006*), image-guided (*Soueid Ahmed et al. 2015*), and cellular automata-based (*Fischer et al. 2017b*).

(2) On the other hand, the parameterization of hydraulic tomography using a distributed discrete-continuum approach is less flexible than the concept of the equivalent porous medium because the discrete-continuum model relies on the establishment of the architecture of the conduits and fractures, and their hydraulic properties. Several works have already brought some solutions to these difficulties. One solution would be to generate stochastically patterns of networks with various constraints: statistical constraints (*Li et al. 2014 ; Le Coz et al. 2017*), mechanical constraints (*Bonneau et al. 2013 ; Jaquet et al. 2004*), geological and speleogeological metrics information (*Collon et al. 2017 ; Pardo-Iguzquiza et al. 2012*), or flows hierarchical identification (*Le Goc et al. 2010*). More recently, *Borghi et al. (2016)* have

combined the use of a generator of karstic networks, based on sets of fractures stochastically generated, with a gradient-based parameters optimization in order to reconstruct a discrete network able to reproduce a set of tracer test hydraulic data.

In this present article we propose a novel strategy for dealing with hydraulic tomography of fractured and karstic aquifers, which we will shorten as the Discrete Network Deterministic Inversion (DNDI). The DNDI algorithm permits to map the architecture of fractures and conduits networks, their hydraulic properties, and the distribution of the transmissivity in the hard rock (matrix).

The DNDI approach relies on the use of a coupled discrete-continuum concept to simulate water flows through a karstic and fractured aquifer and a deterministic optimization algorithm to invert a set of observed piezometric data recorded during multiple pumping tests. The model is partitioned in several subspaces, each one being piloted locally by a set of parameters including: the orientations of the conduit/fracture, their equivalent transmissivity values, and the transmissivity of the rock matrix. This partitioning makes it possible to locally modify the directions of the fracture network and to iteratively update the geometry of the global network in order to minimize the objective function in the inverse process. The method is tested on several hypothetical and simplified karstic aquifers with simple to more complex conduit networks and with homogeneous or heterogeneous transmissivity in the matrices.

4.2.2 Algorithm framework

4.2.2.1 Forward problem and model parameterization

We represent a confined karstic and fractured aquifer in a two-dimensional model Γ with an equivalent porous media Γ_M (for representing the water flows where the rock is intact) and a discrete network Γ_N (for simulating the water flows in the fracture/conduit networks). The numerical simulation of groundwater flows are governed by a steady state continuity equation associated to Darcy's law, considering a laminar flow in both the matrix domain and the discrete networks:

$$\begin{cases} \nabla \cdot (-e_M K_M \cdot \nabla h) = \frac{e_M Q_M}{V_{el.}} & \text{in the matrix } \Gamma_M \\ \nabla_T \cdot (-e_N K_N \cdot \nabla_T h) = \frac{e_N Q_N}{V_{el.}} & \text{in the network } \Gamma_N \end{cases}, \quad (4.1)$$

where Q_M and Q_N are punctual water extraction or injection rates per unit of thickness ($m^3/s/m$) applied on the matrix and network respectively, K_M denotes the matrix hydraulic conductivities (m/s), K_N denotes the fractures or conduits equivalent conductivities (m/s), h is the piezometric level (m) common to both domains Γ_M and Γ_N , e_M (m) is the thickness of the matrix block, e_N (m) is the aperture of the network, and $V_{el.}$ is an elementary volume at the pumping location (m^3). We mention that Darcy's law formulation in the matrix domain is described in 2D, and in 1D for fractured networks at the internal network boundaries. That's why we use the tangential gradient operator $\nabla_T = \vec{\nabla} \cdot \vec{\ell}$ (where $\vec{\ell}$ is a local directional unit vector of the network) to solve the hydraulic equation at the network. In the study cases presented later in this article, we have chosen to simulate laminar flows as presented in Equation (4.1) in a network of conduits. However, the property values K_N in the network can be more specifically adapted to the behavior of turbulent conduit flows or fracture flows through other empirical laws (eventually related to an aperture variable).

The forward problem consists in solving numerically Equation (4.1) by using a finite element technique with a triangular meshing. It links the hydraulic head data simulated continuously over the coupled model to the spatial distribution in the model of the conduits or fractures with their properties in 1D, and the hydraulic transmissivities of the matrix in 2D (Figure 4.1). The forward problem can be formulated as:

$$\mathbf{d} = f\left(\Gamma\left(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}}\right)\right) + \boldsymbol{\varepsilon}, \quad (4.2)$$

where \mathbf{d} is a vector of simulated hydraulic data ($n \times 1$), f is a forward operator that calculates the hydraulic data field from a model $\Gamma\left(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}}\right)$ defined by given parameters of network geometry \mathbf{P}_{Dir} and hydraulic properties \mathbf{P}_{Prop} , $\boldsymbol{\varepsilon}$ is a null mean Gaussian noise to add the uncertainties associated to the numerical discretization technique and the hydraulic experimental data. The model is enclosed in a large buffer zone associated to an equivalent porous media mean transmissivity. This zone permits to limit the influence of the boundary conditions. The DNDI inversion algorithm was coded using Matlab and is linked to the COMSOL Multiphysics software which runs the forward problems.

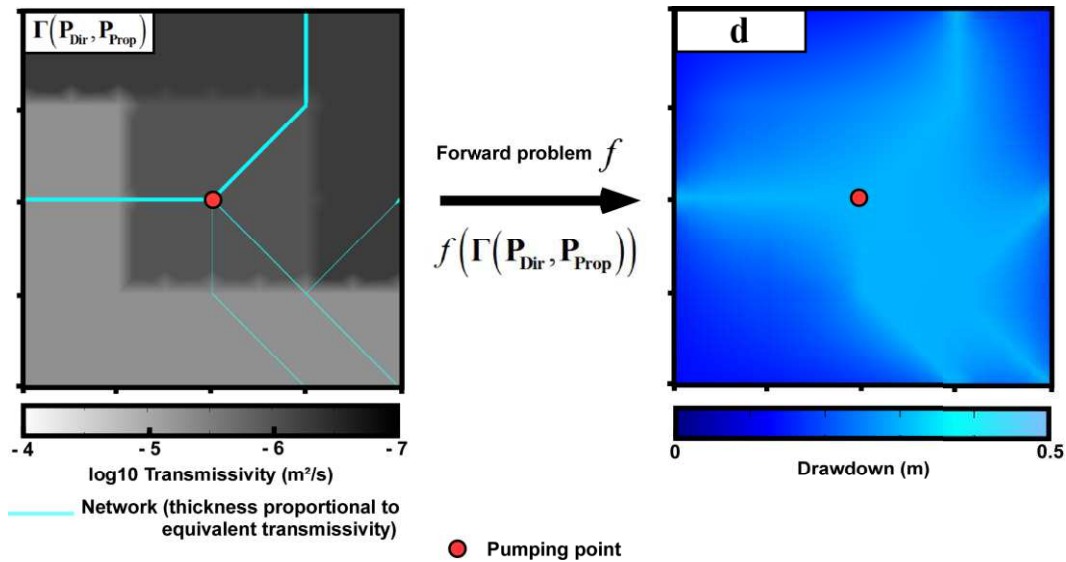


Figure 4.1: Example of a simulated distribution of hydraulic heads (here drawdowns) by solving the forward problem f (Equation (4.1)) for a steady state pumping in a given coupled discrete-continuum distributed model $\Gamma\left(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}}\right)$.

For the DNDI algorithm, the model domain Γ is partitioned in p_x squared subspaces along the X-axis and p_y along the Y-axis. The total of subspaces of the whole domain is then $p = p_x \times p_y$. Three parameters are assigned to each subspace (Figure 4.3): (1) the local direction of the conduit/fracture network, (2) the local conduit/fracture equivalent transmissivity value, (3) and the local matrix transmissivity value.

The geometry of the network follows the local direction in each subspace by a node-to-node principle. The network structure enters a subspace by activating one of its four nodes (corners of the square) and the subspace direction parameter will define to each other node of the subspace the structure will generate. This other node will then be activated itself and permits to the structure to include new subspaces. A subspace in which the structure has already been generated becomes inhibited to another generation from the same network. The generation process is schematized in Figure 4.2.

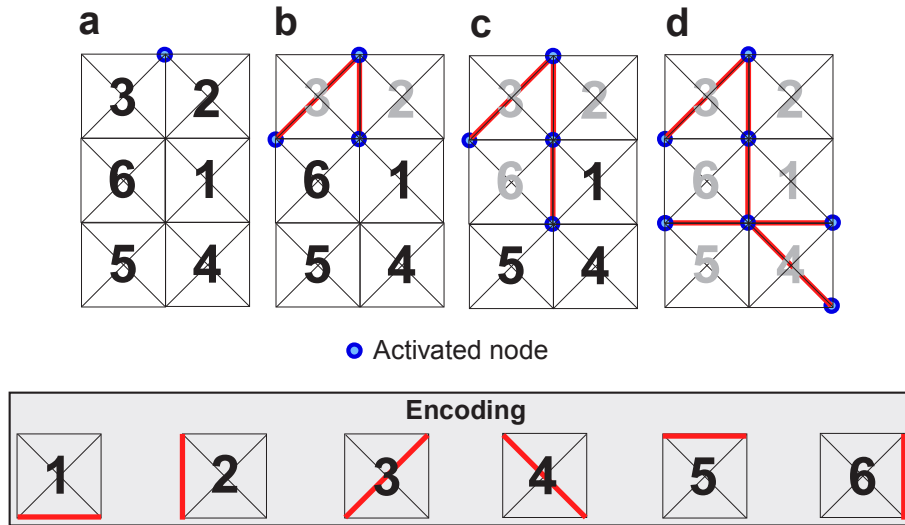


Figure 4.2: Schema of the node-to-node generation process in the DNDI method with six subspaces. An activated node in the top subspaces (a) starts the generation of the structure. The structure generates to the nodes in the bottom of these subspaces, following the local direction defined in the subspaces through the encoding rules. These reached nodes then become activated (b). The subspaces in which the structure has generated become inhibited to another generation (shown as greyed number in this figure). The structure then continues its generation from its newly activated node if the subspaces structural parameters permit it (c) – (d).

In order to perform this node-to-node generation, an initially activated node has to be specified in the model (starting node in Figure 4.3). The model geometry in COMSOL is built

as a discrete grid including all network possibilities (a grid of squares and diagonals as presented in Figure 4.3). This whole geometry is initially disabled in the COMSOL physical part. When different network geometries are tested in the inversion process, only the associated parts in the model grid are enabled for the solver computation. This avoids the creation of a new model geometry for each modification of the network and permits to reduce the computing time in the inversion.

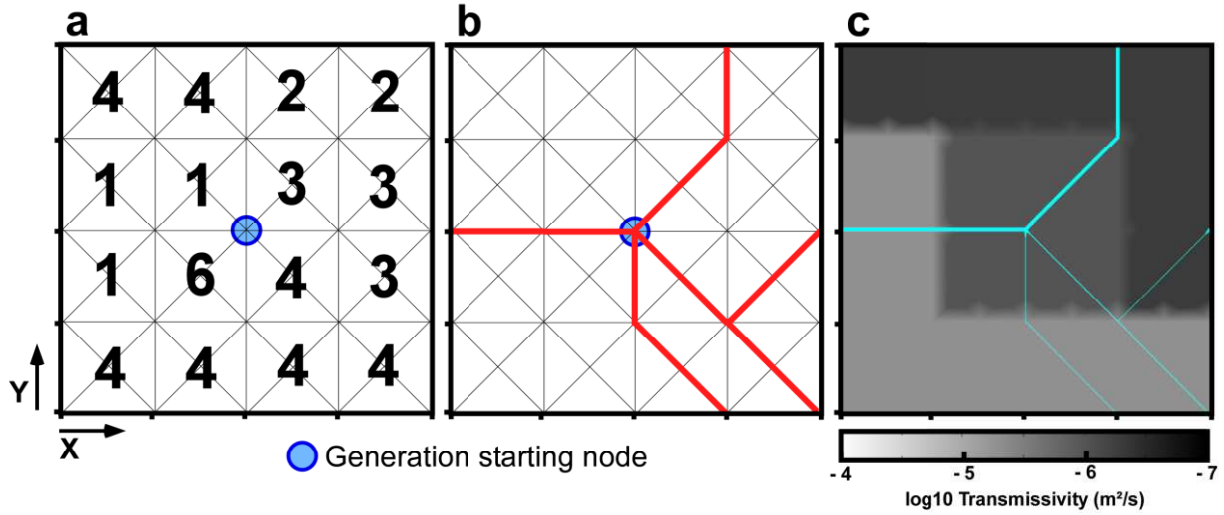


Figure 4.3: Parameterization of a model in the DNDI method. For each subspace of the model there are six local direction possibilities (see encoding in Figure 4.2) that are used to parameterize a network structure in the model (a). The structure (in red) is then generated, following a node-to-node rule, from the set of structural parameters in (a) and a chosen starting point at a node between subspaces (b). Finally a set of property values (transmissivities), also defined for each subspace, is assigned to the structural model (c).

The parameterization of the whole model is contained in two vectors piloting the subspaces:

(1) The local direction in a subspace is selected among six possibilities (see Figure 4.3) as a structural parameter $\text{Dir} \in \{1, 2, 3, 4, 5, 6\}$. The set of structural parameters for all subspaces in the model is contained in a $(p \times 1)$ vector \mathbf{P}_{Dir} . It is also possible to generate more than one network, but this would add more unknown structural parameters. For example if one would want to generate 3 independent networks in the model, each subspace would need to define 3 local directions instead of one. Thus, the structural vector of direction parameters \mathbf{P}_{Dir} would become a $(3p \times 1)$ vector.

(2) The local equivalent transmissivity of the structure in a subspace is defined as a property parameter T_N and the matrix transmissivity as a property parameter T_M . The set of property parameters for all subspaces in the model is contained in a $(2p \times 1)$ vector \mathbf{P}_{Prop} with the p parameters T_N followed by the p parameters T_M . The entire model Γ is thus piloted only by two parameter vectors: \mathbf{P}_{Dir} and \mathbf{P}_{Prop} , and can be noted $\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})$.

4.2.2.2 Inverse problem

The inversion process in the DNDI algorithm consists in retrieving a model of network of conduit/fracture and of spatial distribution of the transmissivities of the network and matrix which permits to maximize two probability density functions ρ_{network} and $\rho_{\text{properties}}$. Following the theory described by *Tarantola and Valette (1982)* for a least square criterion resolution of the inverse problem, we calculate ρ_{network} and $\rho_{\text{properties}}$ with the Bayes theorem, by considering Gaussian laws for the probability density functions ρ , and $\rho(\mathbf{d}_{\text{obs}})$ as certain:

$$\begin{aligned} \rho_{\text{network}}(\mathbf{P}_{\text{Dir}} | \mathbf{d}_{\text{obs}}, \mathbf{P}_{\text{Prop}}) &= \rho(\mathbf{d}_{\text{obs}} | \mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}}) \cdot \rho(\mathbf{P}_{\text{Dir}}) / \rho(\mathbf{d}_{\text{obs}}) \\ \rho_{\text{network}}(\mathbf{P}_{\text{Dir}} | \mathbf{d}_{\text{obs}}, \mathbf{P}_{\text{Prop}}) &\propto \exp\left(-\frac{1}{2}(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})))^T \mathbf{C}_d^{-1}(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})))\right), \\ &\quad \times \exp\left(-\frac{1}{2}(\mathbf{P}_{\text{Dir}, \text{prior}} - \mathbf{P}_{\text{Dir}})^T \mathbf{C}_{\mathbf{P}_{\text{Dir}}}^{-1}(\mathbf{P}_{\text{Dir}, \text{prior}} - \mathbf{P}_{\text{Dir}})\right) \end{aligned} \quad (4.3)$$

$$\begin{aligned} \rho_{\text{properties}}(\mathbf{P}_{\text{Prop}} | \mathbf{d}_{\text{obs}}, \mathbf{P}_{\text{Dir}}) &= \rho(\mathbf{d}_{\text{obs}} | \mathbf{P}_{\text{Prop}}, \mathbf{P}_{\text{Dir}}) \cdot \rho(\mathbf{P}_{\text{Prop}}) / \rho(\mathbf{d}_{\text{obs}}) \\ \rho_{\text{properties}}(\mathbf{P}_{\text{Prop}} | \mathbf{d}_{\text{obs}}, \mathbf{P}_{\text{Dir}}) &\propto \exp\left(-\frac{1}{2}(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})))^T \mathbf{C}_d^{-1}(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})))\right), \\ &\quad \times \exp\left(-\frac{1}{2}(\mathbf{P}_{\text{Prop}, \text{prior}} - \mathbf{P}_{\text{Prop}})^T \mathbf{C}_{\mathbf{P}_{\text{Prop}}}^{-1}(\mathbf{P}_{\text{Prop}, \text{prior}} - \mathbf{P}_{\text{Prop}})\right) \end{aligned} \quad (4.4)$$

where \propto represents a proportionality relation, $\rho_{\text{network}}(\mathbf{P}_{\text{Dir}} | \mathbf{d}_{\text{obs}}, \mathbf{P}_{\text{Prop}})$ represents the a posteriori probability density function of the discrete fracture network model for a given hydraulic observed data \mathbf{d}_{obs} and the transmissivity model of the network and matrix \mathbf{P}_{Prop} .

$\rho_{\text{properties}}(\mathbf{P}_{\text{Prop}} | \mathbf{d}_{\text{obs}}, \mathbf{P}_{\text{Dir}})$ is the a posteriori probability density function of the spatial distribution of the transmissivity parameters for a given hydraulic observed data \mathbf{d}_{obs} and network model \mathbf{P}_{Dir} . $\rho(\mathbf{d}_{\text{obs}} | \mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})$ and $\rho(\mathbf{d}_{\text{obs}} | \mathbf{P}_{\text{Prop}}, \mathbf{P}_{\text{Dir}})$ represent the probability density functions of the network structure and property models, which permit to evaluate the ability of the network structure and property models to reproduce the observed data via the use of the forward operator. $\rho(\mathbf{P}_{\text{Dir}})$ and $\rho(\mathbf{P}_{\text{Prop}})$ represent prior distributions for the unknown parameters. It is well known that, on one hand, the piezometric data are insufficient to cope with the non-uniqueness of the solution of an inverse process, and on another hand, that a deterministic inversion process leads to a single local solution dependent to the initial model. For these reasons, and in order to additionally constrain the inversion to a more realistic solution in relation to the field knowledges, it can be interesting to incorporate prior distributions for the unknown parameters in $\rho(\mathbf{P}_{\text{Dir}})$ and $\rho(\mathbf{P}_{\text{Prop}})$.

The maximization of the a posteriori probability density functions ρ_{network} and $\rho_{\text{properties}}$ is equivalent to a minimization of the arguments of the exponentials in Equations (4.3) and (4.4). This is what we aim to minimize during the inversion process in the following objective functions Ψ :

$$\begin{aligned} \Psi_{\text{network}}(\mathbf{P}_{\text{Dir}}) = & \frac{1}{2} \left(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})) \right)^T \mathbf{C}_d^{-1} \left(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})) \right) \\ & + \frac{1}{2} (\mathbf{P}_{\text{Dir,prior}} - \mathbf{P}_{\text{Dir}})^T \mathbf{C}_{\mathbf{P}_{\text{Dir}}}^{-1} (\mathbf{P}_{\text{Dir,prior}} - \mathbf{P}_{\text{Dir}}) \end{aligned} \quad (4.5)$$

$$\begin{aligned} \Psi_{\text{properties}}(\mathbf{P}_{\text{Prop}}) = & \frac{1}{2} \left(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})) \right)^T \mathbf{C}_d^{-1} \left(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})) \right) \\ & + \frac{1}{2} (\mathbf{P}_{\text{Prop,prior}} - \mathbf{P}_{\text{Prop}})^T \mathbf{C}_{\mathbf{P}_{\text{Prop}}}^{-1} (\mathbf{P}_{\text{Prop,prior}} - \mathbf{P}_{\text{Prop}}) \end{aligned} \quad (4.6)$$

where \mathbf{d}_{obs} is a vector of observed data ($n \times 1$), \mathbf{P}_{Dir} and \mathbf{P}_{Prop} are the unknown parameters to estimate for imaging the geometry of the network (in \mathbf{P}_{Dir}) and the hydraulic properties (defined here by the equivalent transmissivity of the conduits/fractures and the transmissivity of the matrix in \mathbf{P}_{Prop}). $\mathbf{P}_{\text{Dir,prior}}$ ($p \times 1$) and $\mathbf{P}_{\text{Prop,prior}}$ ($2p \times 1$) are the prior information on the geometry and on the property parameters employed to constrain the inverse problem for overcoming the unrealistic solutions, \mathbf{C}_d ($n \times n$) is a covariance matrix on the observed data

that permits to include the uncertainties of the hydraulic data in the inversion process. $\mathbf{C}_{\mathbf{P}_{\text{Dir}}}$ ($p \times p$) and $\mathbf{C}_{\mathbf{P}_{\text{Prop}}}$ ($2p \times 2p$) are the covariance matrices on the structural and property parameters respectively.

This separated formulation of the probability density functions between network and properties permits to sequentially estimate the two dependent unknown models \mathbf{P}_{Dir} and \mathbf{P}_{Prop} . In a first step, we focus on the characterization of the network with the piezometric data by fixing the model of the transmissivity distributions in the conduits/fractures and matrices. The model of network resulting from the first step will then be used in a second step as known parameter to infer the transmissivity pattern.

4.2.2.3 Optimization and uncertainty analysis

The minimization of Equations (4.5) and (4.6) can be done by optimizing the network geometry and the property values during two sequential iterative processes. These optimizations consist in successively modifying the structural and property parameters \mathbf{P}_{Dir} and \mathbf{P}_{Prop} .

The inversion process is led as a sequential optimization (Figure 4.4) of (1) the structural geometry (considering as fixed the initially chosen property values), (2) and the property values T_N, T_M (considering as fixed the previously inverted structural geometry).

The structural optimization is performed iteratively by modifying the structural parameter \mathbf{P}_{Dir} through a structural sensitivity analysis and by considering the hydraulic properties \mathbf{P}_{Prop} as fixed. At a given iteration step k , the sensitivity analysis of the network geometry toward the observed data is recorded into a $(6 \times p)$ sensitivity matrix \mathbf{J}_n^k .

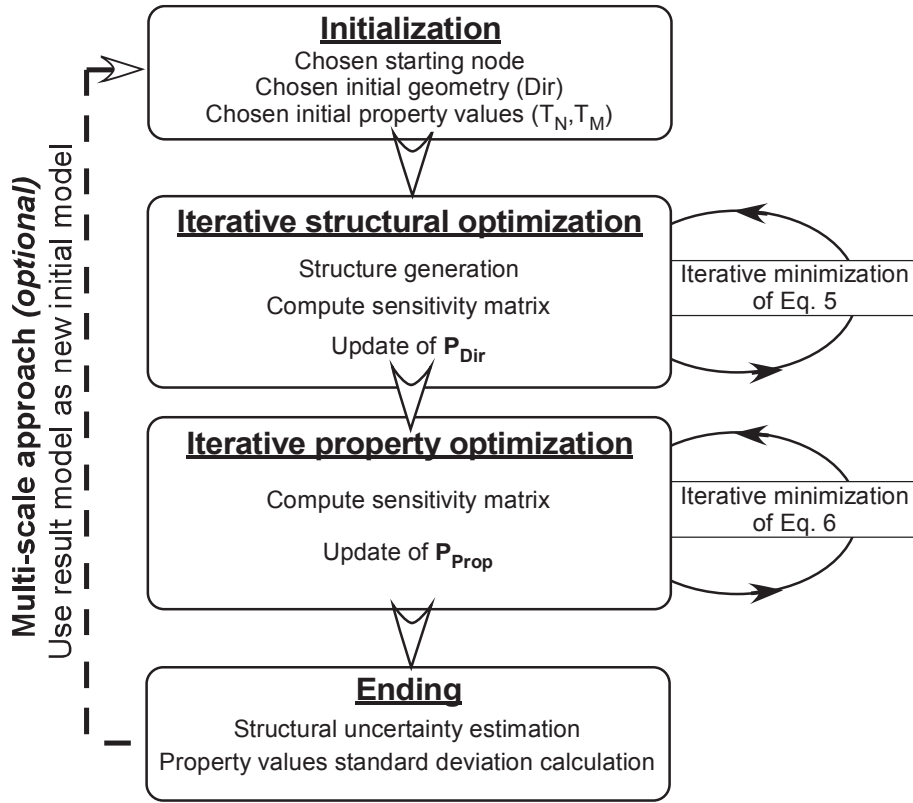


Figure 4.4: A flowchart of the inversion steps used in the DNDI algorithm. After the initialization of the parameters, a sequential iterative optimization is led on the structure geometry and on the property values in order to minimize both objective functions (Equations (4.5) and (4.6)). An eventual re-run of the inversion process (multi-scale option) using the result as new initial model can be performed in order to improve this result.

For a local direction $i \in [1, 6]$ and for a subspace $j \in [1, p]$ the element of the matrix \mathbf{J}_n^k is calculated as:

$$\mathbf{J}_n^k(i, j) = \frac{1}{2} \left(\mathbf{d}_{\text{obs}} - f \left(\Gamma \left(\mathbf{P}_{\text{Dir}}^k \Big|_{\mathbf{P}_{\text{Dir}}(j)=i}, \mathbf{P}_{\text{Prop}} \right) \right) \right)^T \mathbf{C}_d^{-1} \left(\mathbf{d}_{\text{obs}} - f \left(\Gamma \left(\mathbf{P}_{\text{Dir}}^k \Big|_{\mathbf{P}_{\text{Dir}}(j)=i}, \mathbf{P}_{\text{Prop}} \right) \right) \right) + \frac{1}{2} \left(\mathbf{P}_{\text{Dir}, \text{prior}}(j) - i \right)^T \mathbf{C}_{\mathbf{P}_{\text{Dir}}}^{-1} \left(\mathbf{P}_{\text{Dir}, \text{prior}}(j) - i \right) \quad (4.7)$$

with $\mathbf{P}_{\text{Dir}}^k \Big|_{\mathbf{P}_{\text{Dir}}(j)=i}$ the structural geometry parameter at the iteration k considering a modified local direction in the subspace j , and $\left(\mathbf{P}_{\text{Dir}, \text{prior}}(j) - i \right)$ the gap between the prior local direction of subspace j and the modified local direction.

Thus, the sensitivity matrix guides the evolution of the objective function in Equation (4.5) by testing successively the modification of the network with all possible local direction in each subspace. The minimal value in the matrix $\mathbf{J}_n^k(i_{\min}, j_{\min})$ designates the local direction i_{\min} in the subspace j_{\min} which would produce the best decrease in the objective function.

Then the parameters set $\mathbf{P}_{\text{Dir}}^k$ is updated from the previous set $\mathbf{P}_{\text{Dir}}^k = \mathbf{P}_{\text{Dir}}^{k-1}$ by taking into account the sensitivity analysis minimum $\mathbf{P}_{\text{Dir}}^k(j_{\min}) = i_{\min}$ in order to minimize the objective function at each step of the optimization.

Once the sensitivity analysis cannot find any more \mathbf{P}_{Dir} configuration decreasing the objective function, the iterative structural optimization is stopped. The last structural iteration represents the local solution, dependent to the initial model. The uncertainty analysis of the inverted network geometry can be inferred from the computation of the posterior covariance matrix as:

$$\mathbf{C}_{\text{P}_{\text{Dir}}}^{\text{post}}(j) = \left(\frac{1}{6} \sum_{i=1}^6 \mathbf{J}_n^{\text{post}}(i, j) - \Psi_{\text{network}}^{\text{post}} + \mathbf{C}_{\text{P}_{\text{Dir}}}^{-1}(j, j) \right)^{-1}, \quad (4.8)$$

where $\mathbf{C}_{\text{P}_{\text{Dir}}}^{\text{post}}(j)$ is the posterior structural uncertainty value for the local direction in the subspace j , $\mathbf{J}_n^{\text{post}}$ is the last iteration structural sensitivity matrix and $\Psi_{\text{network}}^{\text{post}}$ is the value of the minimized structural objective function. If the structural uncertainty value is low, then another direction in the subspace would lead to a deterioration of the reproduction of the data. On the other hand if the structural uncertainty value is high, then the structure in the subspace could have another local direction without significantly degrading the reproduction of the observed data.

Following the network geometry optimization, the property parameters optimization will iteratively modify the transmissivities with the previously inverted geometry in order to minimize the objective function in Equation (4.6). The network equivalent transmissivities and the matrix transmissivities are optimized simultaneously.

At a given iteration step k , the parameters set \mathbf{P}_{Prop} , which contains the transmissivities for both the network and the matrix, is updated by linearizing Equation (4.6), which can be formulated as:

$$\mathbf{P}_{\text{Prop}}^{k+1} = \mathbf{P}_{\text{Prop}}^k + \left(\left(\mathbf{J}_{\text{p}}^k \right)^T \cdot \mathbf{C}_{\text{d}}^{-1} \cdot \mathbf{J}_{\text{p}}^k + \mathbf{C}_{\text{P}_{\text{Prop}}}^{-1} \right)^{-1} \cdot \left(\mathbf{J}_{\text{p}}^k \right)^T \cdot \mathbf{C}_{\text{d}}^{-1} \cdot \left(\mathbf{d}_{\text{obs}} - f \left(\Gamma \left(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}}^k \right) \right) \right) + \mathbf{C}_{\text{P}_{\text{Prop}}}^{-1} \cdot \left(\mathbf{P}_{\text{Prop,prior}} - \mathbf{P}_{\text{Prop}}^k \right), \quad (4.9)$$

where \mathbf{J}_{p}^k is the Jacobian matrix ($n \times 2p$) that holds the sensitivity for each modeled data f_i (at the positions of the observed data) toward the property values in the matrix and the network. This Jacobian matrix can be calculated by using a finite difference approach, with a finite difference step $\Delta \mathbf{P}_{\text{Prop}}$:

$$\mathbf{J}_{\text{p}}^k(i, j) = \left. \frac{\partial f_i}{\partial \mathbf{P}_{\text{Prop}}^k} \right|_{\mathbf{P}_{\text{Prop}}^k(j) = \mathbf{P}_{\text{Prop}}^k(j) + \Delta \mathbf{P}_{\text{Prop}}}. \quad (4.10)$$

Finally, once the objective function has iteratively converged to a minimum, the property optimization is stopped. The posterior covariance matrix on the inversion of the property values can be calculated as:

$$\mathbf{C}_{\text{P}_{\text{Prop}}}^{\text{post}} = \left(\left(\mathbf{J}_{\text{p}}^{\text{post}} \right)^T \cdot \mathbf{C}_{\text{d}}^{-1} \cdot \mathbf{J}_{\text{p}}^{\text{post}} + \mathbf{C}_{\text{P}_{\text{Prop}}}^{-1} \right)^{-1}, \quad (4.11)$$

where $\mathbf{C}_{\text{P}_{\text{Prop}}}^{\text{post}}$ is the posterior covariance matrix and $\mathbf{J}_{\text{p}}^{\text{post}}$ is the Jacobian matrix of the last iteration step. The diagonal entries of the posterior matrix represent the variances on the property values of each subspace.

4.2.3 Validation of the DNDI algorithm on hypothetical study cases

The DNDI inversion algorithm has been tested on three hypothetical and simplified confined karstic fields with network of conduits:

- in a first case, we treat a simple network case with heterogeneity in the equivalent transmissivity of the conduits and a homogenous transmissivity assigned to the matrix,
- a second case is similar to the first one but adding a transmissivity variability also in the matrix,
- in a third case, we seek to image a complex network geometry with the use of two different initial models to start the inverse problem.

We considered in the forward problem (Equation (4.1)) a unit thickness for the matrix (2D modeling) and a unit aperture for the network (but with a variable equivalent transmissivity). The buffer zone boundaries were associated to a $h_{\text{bound.}} = 0$ m Dirichlet condition and the hydraulic heads were set to $h_0 = 0$ m initially over the model. These theoretical study cases were used to produce 2,401 hydraulic drawdown data from 49 pumping/measurement boreholes (a pumping test is performed alternatively in each borehole) distributed homogeneously over the 100×100 m² models. The pumping rates were set to 0.6 L/min for a borehole in the matrix and 5 L/s for one in the conduit network.

In these different cases, for the inversion of the geometry of the network, no a priori information has been added. On the other hand, we have constrained the inversion of the hydraulic properties with a priori values. The a priori models on the properties are used also as initial model to launch the inversion process. For the property optimization, for the matrix we took the $-\log_{10}(T)$ as transmissivity parameter T_M (for example a transmissivity parameter equal to 6 represents in the model a 10^{-6} m²/s transmissivity value) and for the network we took directly the T value as transmissivity parameter T_N . The covariance matrices C_d and $C_{P_{\text{Prop}}}$ are built as diagonal matrices with a constant variance value σ^2 (in the case of the matrix transmissivity the variance value $\sigma_{T_M}^2$ applies to the exponent of the transmissivity, in the case of the network transmissivity the variance value $\sigma_{T_N}^2$ applies to the transmissivity). The

partitioning of the models and the chosen inversion parameters values for each study case are given in Table 4.1. The different study cases inversions were led on a 64Go RAM PC on 2 processors of 16 cores.

Table 4.1: Parameters used in the inversion study cases

	Study case 1	Study case 2	Study case 3
Partitioning	4×4	4×4	8×8
A priori T_N	$0.06 \text{ m}^2/\text{s}$	$0.04 \text{ m}^2/\text{s}$	$0.1 \text{ m}^2/\text{s}$
A priori T_M	$10^{-6} \text{ m}^2/\text{s}$	$10^{-6} \text{ m}^2/\text{s}$	$10^{-6} \text{ m}^2/\text{s}$
Data cov. matrix $C_d = \sigma_{\text{data}}^2 \cdot Id(n)$	$\sigma_{\text{data}} = 10^{-2} \text{ m}$	$\sigma_{\text{data}} = 10^{-2} \text{ m}$	$\sigma_{\text{data}} = 10^{-2} \text{ m}$
Property cov. matrix $C_{P_{\text{Prop}}} = \sigma_{T_N/T_M}^2 \cdot Id(2p)$	$\sigma_{T_N} = 10^{-6} \text{ m}^2/\text{s}$ $\sigma_{T_M} = 10^{-6}$	$\sigma_{T_N} = 10^{-6} \text{ m}^2/\text{s}$ $\sigma_{T_M} = 10^{-1}$	$\sigma_{T_N} = 10^{-6} \text{ m}^2/\text{s}$ $\sigma_{T_M} = 10^{-6}$

4.2.3.1 Study case 1

In a first study case, we have tested the ability of the inversion method to reproduce a network geometry with variable conduit equivalent transmissivities in a homogeneous matrix. We generated drawdown data from a theoretical model with a $10^{-6} \text{ m}^2/\text{s}$ matrix transmissivity and a principal conduit associated to a $0.1 \text{ m}^2/\text{s}$ transmissivity and secondary conduits associated to a $0.01 \text{ m}^2/\text{s}$ transmissivity. Firstly we tested an inversion with a small set of data (100 drawdown data from 10 boreholes, see the ‘True model’ in Figure 4.5).

We started the inversion from a simple initial model with a single horizontal $0.06 \text{ m}^2/\text{s}$ conduit and a homogeneous $10^{-6} \text{ m}^2/\text{s}$ matrix transmissivity. The structural optimization converged in 10 iterations and the properties optimization in 1 iteration.

The inverted model reproduces the data set ($R^2 = 0.97$) and approximately the connectivity between the points in the network, however this reconstruction remains distant from the true geometry. This is due to a lack of data to correctly identify the shape of the conduit network.

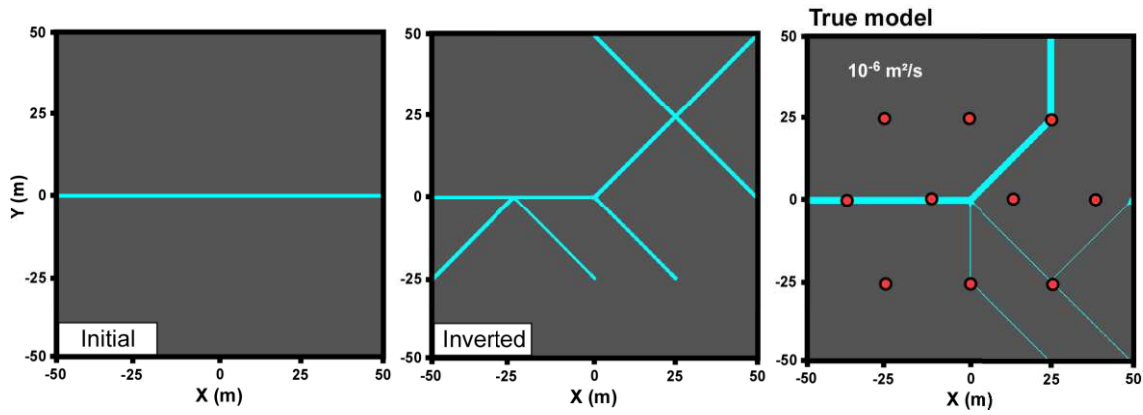


Figure 4.5: Initial and inverted models for an inversion using drawdown data produced from a true model (on the right) with a homogeneous matrix. The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data. The inverted model permits to localize approximately the karstic network connections but in this case the amount of data is insufficient to have a proper imagery.

Therefore, the efficiency of the inversion for mapping the heterogeneity of the hydraulic parameters and retrieving the principal karstic conduits is highly dependent to the number of wells and their locations. In the next test, we used a denser distribution of wells (49 wells) for providing a better spatial resolution in order to image the heterogeneity of the aquifer presented for the same ‘True model’ in Figure 4.6.

The structural optimization converged in 11 iterations and the properties optimization in 1 iteration. The inverted model reproduces now the data set ($R^2 = 0.95$) and also a very good representation of the true geometry (Figure 4.6). The property optimization permitted to correct the initial equivalent transmissivity of $0.06 \text{ m}^2/\text{s}$ to $0.01 \text{ m}^2/\text{s}$ for the conduits connected in the bottom right area of the network. It permits to reduce the flow rates coming to this zone and enhance the reproduction of the true cones of depression. The flows in this zone are mainly conditioned by the properties of the conduits connected directly to the primary drain. This affirmation can be supported by the conduit transmissivity standard deviation map produced from Equation (4.11) (Figure 4.6), that shows that the properties of the conduits directly connected to the primary drain have lower uncertainties than the primary drain itself in the center of the inverted model. The conduit in the bottom right periphery of the inverted model does not image correctly the true model. But as the data reproduction is perfect, this periphery zone might not be sufficiently described by the data to permit a very good reproduction. The uncertainty map confirms that this part of the network has a more uncertain transmissivity value than the rest of the network.

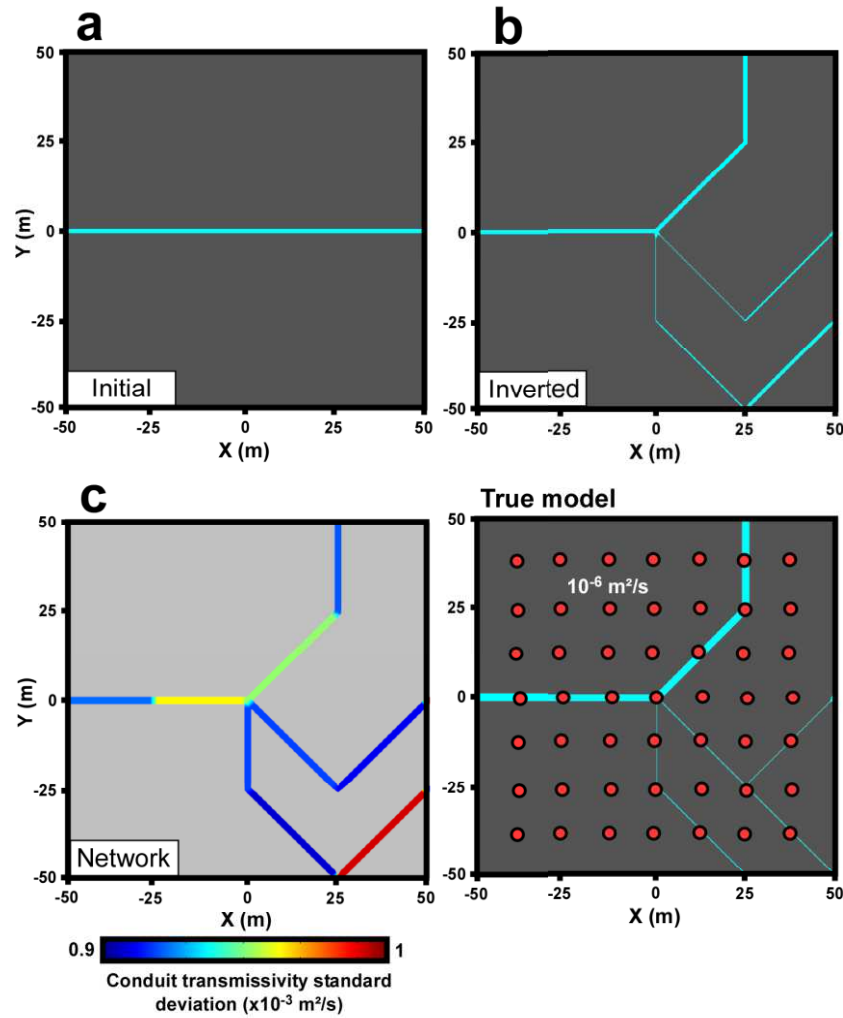


Figure 4.6: Initial (a) and inverted (b) models for an inversion using drawdown data produced from a true model with a homogeneous matrix, and associated map of the conduit properties posterior standard deviations (c). The inverted model in (b) permits a good localization the true karstic network. It also reduced locally the initial transmissivity ($0.06 \text{ m}^2/\text{s}$ to $0.01 \text{ m}^2/\text{s}$) of the conduits connected to the primary drain in the bottom right part of the model (the conduit thickness is proportional to its transmissivity value). The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data.

We have also tested another configuration with more available boreholes than in the case in Figure 4.5, but in which only two boreholes intersect the true karstic network. The true model and the inversion result are presented in Figure 4.7.

The inverted model can almost reproduce the true network geometry, which shows that boreholes, even in the matrix, can provide information about the localization of nearby conduits. This is especially true for the thin conduits which appear in the inverted model although no boreholes are intersecting them.

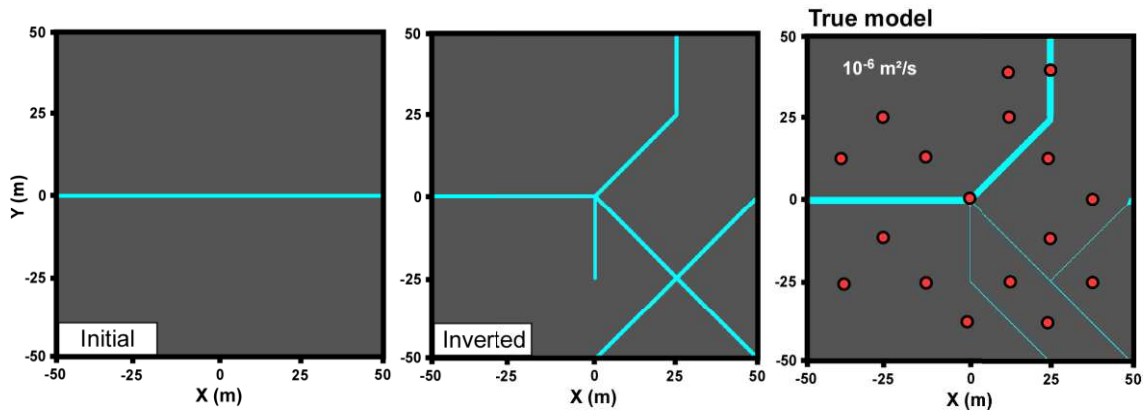


Figure 4.7: Initial and inverted models for an inversion using drawdown data produced from a true model (on the right) with a homogeneous matrix. The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data, primarily located in the matrix. The inverted model permits to almost reproduce the karstic network even if only two measurement points are located in the true network.

Therefore the DNDI method can be used with dataset with only a few boreholes intersecting the conduits as long as there are a sufficient number of other boreholes, in the matrix, in suitable locations for characterizing the nearby conduit network.

4.2.3.2 Study case 2

A second study case was led to test the ability of this inversion method to reproduce the data in a case of a karstic network with various conduit properties developed in a heterogeneous matrix. We simulated the piezometric data from a theoretical model with the same karstic network than in study case 1, but in a matrix with a transmissivity varying from 5×10^{-6} m²/s to 5×10^{-7} m²/s (Figure 4.8).

We started the inversion from a simple initial model with a single horizontal 0.04 m²/s conduit and a homogeneous 10^{-6} m²/s matrix transmissivity. The structural optimization converged in 10 iterations and the properties optimization in 3 iterations (Figure 4.8a).

The structural optimization permitted to retrieve the true geometry of the conduits network, but it also added conduits in the bottom left part of the model to reproduce the drawdown data of the more transmissive area of the matrix. Then the property optimization could reproduce the true transmissivity values distribution in the matrix.

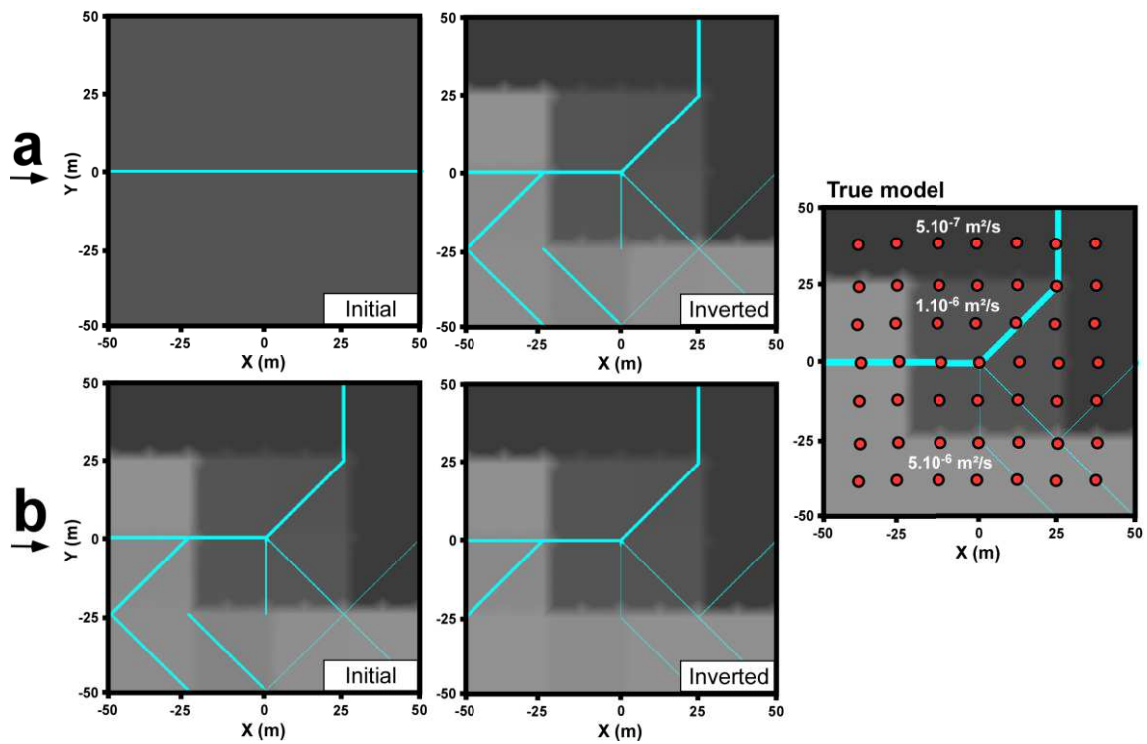


Figure 4.8: Initial and inverted models for an inversion using drawdown data produced from a true model (on the right) with a heterogeneous matrix. The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data. A first inverted model (a) permits to localize the true karstic network but also generates conduits to simulate the more transmissive part of the true model. A second inversion (b) starting from the previous inverted model permits to correct the geometry and produces an inverted model matching more accurately the true model.

In the end the inverted model can reproduce the true drawdowns data, but its network geometry incorporates parts, inexistent in the true model, that has been generated in order to simulate a more transmissive area of the matrix before the matrix transmissivity values could be optimized.

We started a second inversion using the previously inverted model (indicated in Figure 4.4 as the ‘multi-scale option’). The structural optimization converged in 2 iterations and the parameter property optimization in 1 iteration (Figure 4.8b). The only changes were made during the structural optimization step, with an important improvement in the identification of the shape of the conduits. In this case the inverted model reproduces the drawdowns data ($R^2 = 0.99$) but is also a good representation of the true network geometry.

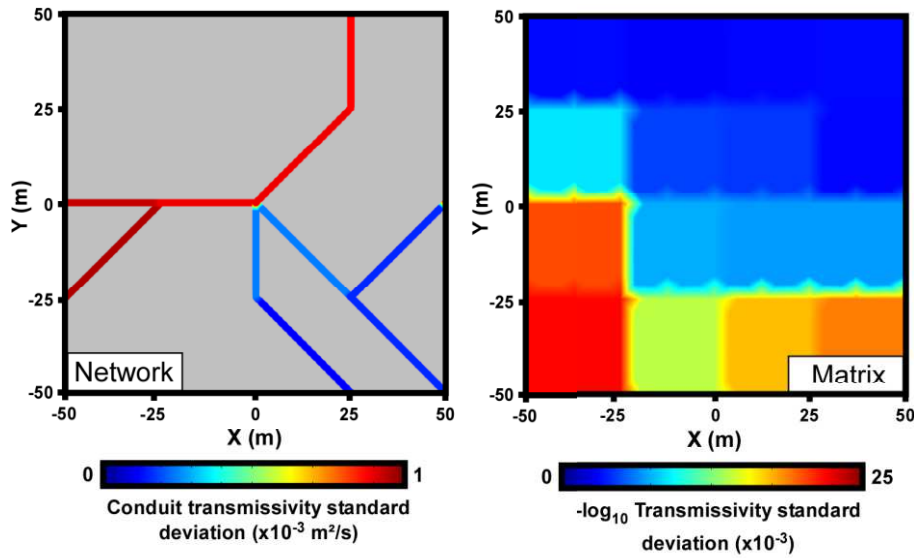


Figure 4.9: Maps of the conduit and matrix transmissivities posterior standard deviations. The matrix higher transmissivity zones in the inverted model (bottom left) have a higher uncertainty value than the lower transmissivity zones (top right). On the contrary, the uncertainty on the transmissivities of the conduits of the primary drain is higher than the secondary conduits.

The posterior standard deviation maps produced from Equation (4.11) (Figure 4.9) show, for the conduit property values, a smaller uncertainty for the secondary conduits and a higher uncertainty for the primary drain, especially for the part of the network on the left of the model. Concerning the matrix transmissivity property values, the highest uncertainty are located mostly in the most transmissive areas.

4.2.3.3 Study case 3

Finally, a third study case was led to test the ability of this inversion method to reproduce the data in a case of a complex karstic network geometry. We generated drawdown data from a theoretical model with a karstic network with a constant equivalent transmissivity of $0.1 \text{ m}^2/\text{s}$ in a homogeneous matrix with a transmissivity of $10^{-6} \text{ m}^2/\text{s}$ (Figure 4.10).

We started an inversion from a simple initial model with a single vertical $0.1 \text{ m}^2/\text{s}$ conduit and a homogeneous $10^{-6} \text{ m}^2/\text{s}$ matrix transmissivity. The structural optimization converged in 33 iterations and the parameter optimization in 1 iteration (Figure 4.10a). The inverted model permits to fit the data set approximately ($R^2 = 0.78$) and represents the global geometry of the conduits network of the true model. Regarding the simplicity of the initial model, the result model remains satisfying.

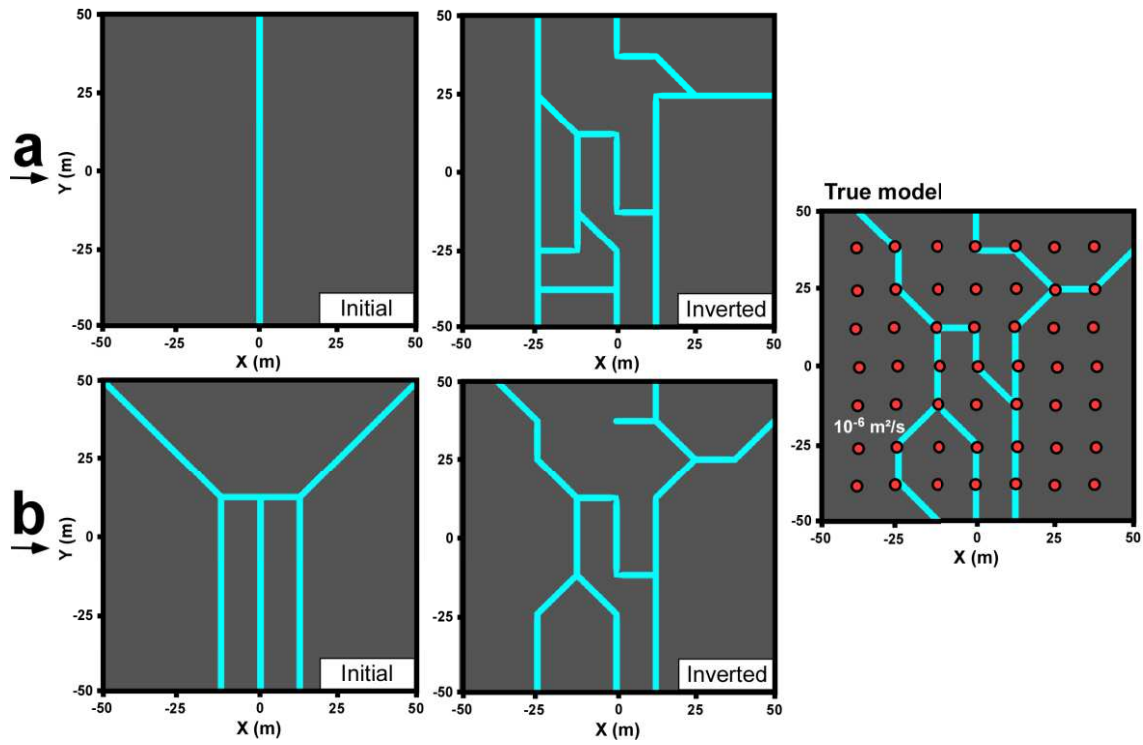


Figure 4.10: Initial and inverted models for an inversion using drawdown data generated from a true model (on the right) with a homogeneous matrix. The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data. A first inverted model (a), starting from a simple initial model, permits to localize approximately the true network geometry. A second inversion (b), starting from a more detailed initial model, permits to produce a more precise network geometry.

We also started an inversion from a more complex initial model with two vertical $0.1 \text{ m}^2/\text{s}$ conduits diverging in the upper part of the model in a homogeneous $10^{-6} \text{ m}^2/\text{s}$ matrix transmissivity. This initial model geometry (representing a simple approximation of the true geometry) can be associated to a priori field knowledge information. The structural optimization converged in 17 iterations and the parameter optimization in 1 iteration (Figure 4.10b). In this case, the inverted geometry of the discrete network permits a good reproduction of the data ($R^2 = 0.97$) and is closer to the real network than the case in Figure 4.10a.

The structural posterior uncertainty maps produced from Equation (4.8) are presented in Figure 4.11. These maps show that, in the Case a, the highest uncertainties are distributed relatively uniformly among the inverted model, while in Case b, they are mostly located in the periphery of the model. Here, the structural posterior uncertainties are giving important information about the local validity of the different inverted networks.

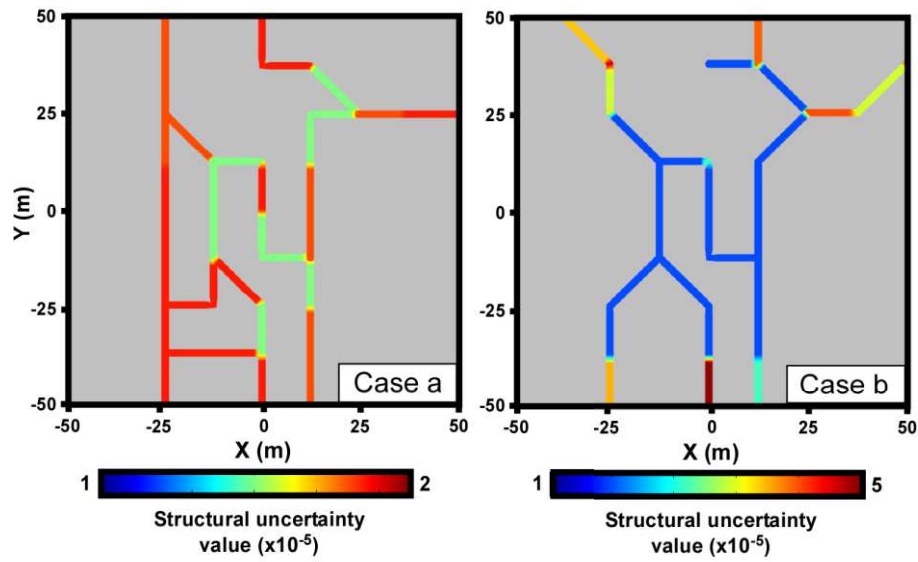


Figure 4.11: Maps of the posterior uncertainties of the network local directions for the Cases a and b. In the Case a, started from a simple initial model, the highest uncertainties are distributed uniformly over the inverted network. In the Case b, started from a more detailed initial model, the highest uncertainties are located in the periphery of the model.

4.2.4 Discussion

We have successfully tested the DNDI method on three theoretical and simplified study cases with steady state drawdowns. However as we have seen, an inversion process is limited by the non-uniqueness of its solution. Therefore using the DNDI method requires several prerequisites and the modeler needs to be critical toward the result.

As we have seen in the first study case, the efficiency of the inversion is dependent to the hydraulic data set, and in particular the number and the localization of observation wells on the field. We note that even wells in the matrix can provide information on nearby conduits for the inversion. Globally it appears that the most important point about a steady-state dataset is to have a homogeneous and sufficiently dense distribution of wells on the site, in order to characterize successfully the network.

Concerning the inversion process itself, we note, in the third study case, the ability of the DNDI method to image complex networks. However, as the inversion is deterministic, the precision of the result model is dependent to the initial model. The inversion process will converge to a local solution dependent to the initial model. In fact, in Figure 4.10 we show that a simple initial model permitted to reproduce a satisfying global representation of the true model, but with local approximations, while a more complex initial model permitted a more accurate reproduction of the true model and a faster convergence. Therefore an inverted model using the DNDI method should be analyzed critically, like any deterministic inverse methods, depending from the initial model. The study of the computed structural and property uncertainty values (with Equations (4.8) and (4.11)) can supply this critical analysis on the result model.

The second study case also illustrates some limits of the sequential optimization of the method, especially when starting from a too simple initial model. Therefore the amount of a priori information introduced in the initial model is important for the accuracy of the result model. Otherwise, as we demonstrate in Figure 4.8, a simple possible operation would be to re-run the inversion with a first inversion result to slightly improve the result. We would also recommend the coupling of this inversion method to a multiscale method (*Grimstadt et al. 2003*) which consists in a re-run of the inversion starting from a previous result with a refinement of the partitioning. It permits to lead several inversions with an initial model each time more precise while saving time as we initially start with a coarsely partitioned model.

4.2.5 Conclusion

We present in this paper a novel deterministic inversion method that permits to characterize, in a partitioned model, the karst conduits and fractures network geometry and their hydraulic properties, including the transmissivity distribution of the matrices. The DNDI method let the modeler choose the partitioning of the model for the inversion. This ‘cursor’ permits to define either an inversion with a coarse partitioning for a quick approximation model, or with a fine partitioning and a longer computation time for a better fitting model. The use of a discrete network model permits to associate a specific behavior to the flows in the network and thus, produces more realistic models than an equivalent porous media model. This method can be easily adapted for channels or fractures network models by modifying the properties associated to the discrete network (these properties can also be directly linked to an aperture value, by choosing an adapted law). Therefore we believe that the DNDI method is an interesting new imagery tool for the distributed modeling associated to a set of data from an investigation in a karstic and/or fractured aquifer.

We have realized different tests in three theoretical and simplified study cases with an increasing complexity, and the DNDI could always produce satisfying results, both on the reproduction of the generated data and on finding the network geometry and property values from the true model. As we have seen in the first study case, the result of the structural inversion is dependent on the positioning and the amount of observed data. This is true for any inversion, but is especially important in the case of highly heterogeneous aquifers for delineating the position of the heterogeneities. Therefore, the result of the inversion has to be interpreted critically regarding the set of data used for it. A first critical analysis can be performed from the maps of posterior uncertainties on the structure or on the property values that can be produced by using the formulas we propose in this paper. The a priori information on the geometry of the network and on the property values is also a way to constrain the inversion in addition to the data. This information can be inferred from general field knowledges (geological and geophysical information, conduits observation in wells through video camera, other studies, etc.).

Because this method is deterministic, the choice of the initial model should be based on a relatively coherent possibility and should not be too far from the real solution in order to produce a good result. Therefore, we propose to couple the DNDI method to a multi-scale method. This consists in a first inversion started from an initial model which is followed by a new one that starts from the first inversion solution with a refined partitioning. This strategy

permits to start from a simple initial model and to progressively make the model more complex and improve the solution.

An application of this method for mapping the conduits and fractures network with real data from a karstic field is planned for future works. These works will be more specifically focused on the sensitivity of the method to the spatial distribution of the measurement boreholes and on delineating the preferential flow paths in the network.

5 CARACTÉRISATION DES ÉCOULEMENTS DANS LE SITE DU TERRIEU PAR POMPAGE À DÉBIT CONSTANT

5.1 Contexte

Une application réelle de la méthode CADI afin de caractériser un milieu fracturé et karstique par tomographie hydraulique est détaillée dans ce chapitre. Le milieu investigué est le site expérimental du Terrieu, un site d'étude karstique situé à environ 15 km de Montpellier. Ce site d'étude fait partie du réseau du Service National d'Observation des karsts et de l'observatoire MEDYCYSS. L'investigation menée sur ce site a généré un jeu de réponses de rabattements de la nappe qui ont été mesurées dans vingt-et-un forages lors de huit différents pompages à débit constant. Ce jeu de réponses est utilisé, avec la méthode CADI, afin de générer un modèle de propriétés structuré en réseau de conduits capable de le reproduire.

La méthode CADI est brièvement rappelée en début de chapitre, avant une présentation des connaissances géologiques et hydrogéologiques du site du Terrieu, puis d'une brève partie sur le paramétrage et l'initialisation de l'inversion des données. Les résultats d'imagerie produits par la méthode CADI sont finalement détaillés dans la dernière partie. Ils permettent de visualiser une possible distribution du champ de propriétés (conductivités) du milieu, mais également des cartes d'incertitudes associées à ce champ, et des cartes de rabattements et de vitesses d'écoulements simulés pour les différents pompages. La reproduction des réponses produite par le modèle de champ de propriétés inversé est bonne, excepté éventuellement sur les faibles rabattements plus sensibles aux micro-fractures, qui sont non représentées dans le modèle. Le champ de propriété permet néanmoins une bonne lisibilité du positionnement du réseau karstique et peut être ainsi validé par des connaissances de connectivités entre forages et d'observations de conduits in-situ.

La méthode CADI produit une interprétation sous forme d'imagerie qui, par rapport à d'autres méthodes d'inversion non structurée, permet la lecture de connectivité du réseau de conduits et la localisation des chemins préférentiels d'écoulement. Son application pour une tomographie dans ce cas est d'autant plus efficace que sur le site du Terrieu le réseau de forages de mesures est dense et bien réparti. Néanmoins, la tomographie présentée dans ce chapitre ne permet pas de hiérarchiser les conduits du réseau, c'est-à-dire de différencier ceux représentant les conduits les plus importants de ceux représentant des fissures plus petites. En effet, bien que les pompages à débit constant atteignant des régimes permanents génèrent des rabattements distribués en fonction des chemins d'écoulements préférentiels, ils mobilisent tout de même de l'eau provenant de la matrice et des petites fissures qui ont tendance à cacher la localisation précise des conduits plus importants.

5.2 Identifying flow networks in a karstified aquifer by application of the cellular automata-based deterministic inversion method (Lez aquifer, France)

Cette partie est composée de l'article « Identifying flow networks in a karstified aquifer by application of the cellular automata-based deterministic inversion method (Lez aquifer, France) », publié dans le journal *Water Resources Research* en décembre 2017 (*Fischer et al. 2017c*). Le texte a été remis en forme au format du manuscrit. La version originale de l'article est donnée en Annexe 4.



Water Resources Research

RESEARCH ARTICLE

Citation:

Fischer, P., Jardani, A., Wang, X., Jourde, H., & Lecoq, N. (2017). Identifying flow networks in a karstified aquifer by application of the Cellular Automata-based Deterministic Inversion Method (Lez aquifer, France). *Water Resources Research*, 53, 10.506-10.522. <https://doi.org/10.1002/2017WR020921>

Identifying Flow Networks in a Karstified Aquifer by Application of the Cellular Automata-Based Deterministic Inversion Method (Lez Aquifer, France)

P. Fischer¹ , A. Jardani¹, X. Wang², H. Jourde², and N. Lecoq¹

¹Normandie Univ, UNIROUEN, UNICAEN, CNRS, Rouen, France, ²Laboratoire Hydrosiences, Université de Montpellier, CNRS, Montpellier, France

5.2.1 Introduction

The numerical modeling of groundwater flows within heterogeneous aquifers and the assessment of their hydrodynamic properties (such as the hydraulic conductivity and specific storage coefficient) remain actually an important and complex research challenge (*White 2002 ; Hartmann et al. 2014a*). The main difficulties faced in the modeling of these types of aquifers are due to the high contrast in the hydraulic properties at small spatial scale, at the limits between conduits, fractures and matrix. These heterogeneities lead to complex and discontinuous patterns of groundwater flows that are mainly controlled by the geometric characteristics of the fracture or conduit networks (spatial locations, apertures, sizes, densities). Most often, fractured and karstified aquifers are modeled by using a lumped simulation method (*Dreiss 1982 ; Labat et al 1999 ; Long and Derickson 1999 ; Arfib and Charlier 2016 ; Kong-A-Siou et al. 2015 ; Ladouche et al. 2014 ; Hartmann et al. 2014b*), in which the whole hydrosystem is considered as a grey-box (or black-box) in order to study the responses of the system in an output signal by conceptualizing some physical processes. This method can be useful to describe the global responses of a system to a rainfall signal but it does not give precise information on the flow behavior within the aquifer. A distributed hydrodynamic simulation method is more adequate to describe the mechanistic processes of water flows within a heterogeneous aquifer. The distributed hydrodynamic simulations can be categorized in three main approaches (*Kovacs and Sauter 2007 ; Ghasemizadeh et al. 2012 ; Hartmann et al. 2014a*): (1) the equivalent porous media (*Wang et al. 2016 ; Abusaada and Sauter 2013*) in which the hydraulic features of the fractured areas are approached with equivalent continuous hydraulic properties, (2) the double continuum (*Zimmerman et al. 1993 ; Kordilla et al. 2012*), in which the model is conceptualized with two porous continuum media (matrix and conduit) that have distinct hydraulic properties, and (3) the combined discrete-continuum (*Saller et al. 2013 ; Jaquet et al. 2004*) in which the discrete fractures are defined by their geometries and their local apertures, their interactions with the porous matrix media are included by using exchanging flow terms.

The characterization of the spatial distribution of the hydraulic properties is commonly provided from an inversion process coupled to a hydraulic tomography approach (*Yeh and Liu 2000 ; Bolhing et al. 2002 ; Zhu and Yeh 2005 ; Cardiff and Kitanidis 2009 ; Wang et al. 2017*). This approach consists of a joint analysis of a set of piezometric data collected as the responses of a water extraction during multiple pumping tests. Both steady-state and transient hydraulic tomography have been considered in previous works. In a transient hydraulic experiment, both hydraulic conductivity and specific storage influence hydraulic head distribution (*Zhu and Yeh*

2005 ; Castagna *et al.* 2011 ; Sharmeen *et al.* 2012), thus, more unknown parameters need to be estimated compared to a steady-state experiment where drawdown data depend exclusively on hydraulic conductivity (Yeh and Liu 2000 ; Cardiff *et al.* 2009 ; Wang *et al.* 2016).

The efficiency of the characterization depends on the number and spatial disposition of the boreholes used in the investigation. However, in the practice cases, the number of wells is usually insufficient to reduce the uncertainty and uniqueness of the solution. To overcome these difficulties, *a priori* knowledge is used to constrain the inverse problem. The geostatistical approach is widely applied to constrain the hydraulic tomography particularly for porous aquifers with a moderate variability in hydraulic conductivity (Fischer *et al.* 2017a ; Lee and Kitanidis 2014 ; Hoeksema and Kitanidis 1984). To deal with discrete spatial patterns, the method was advanced by replacing the Gaussian prior model by a Laplace prior (or total variation prior) in the Bayesian framework (Lee and Kitanidis 2013).

However, in the case of highly heterogeneous and complex aquifers, the use of geostatistical constraints can lead to unrealistic models dominated by a relatively smooth variability of the hydraulic properties.

In this paper, we apply a novel structural inversion method, the Cellular Automata-based Deterministic Inversion (CADI), to invert the steady state hydraulic head data recorded during a hydraulic tomography to image the spatial distribution of the hydraulic transmissivities in the fractured and karstified Lez aquifer (Southern France). The theoretical aspects of the CADI method have been developed in a previous article (Fischer *et al.* 2017b). This method is based on the Cellular Automata (CA) concept to parameterize the model. It permits a deterministic inversion of linear structures which is interesting for the modeling of fractures and karstic conduits. The paper is developed as follows: in the first section, we present the CADI algorithm and the concept used to parameterize the model and the inverse problem. Then, in the second section we present investigations on the experimental site. Finally, we discuss the results of the application of the proposed inverse method to map the hydrodynamic properties of a karstified and fractured aquifer.

5.2.2 Methodology

5.2.2.1 Model parameterization

We dedicate this section to describe briefly the main concepts of the CADI method. For more details about the theory of this method, we invite the readers to refer to *Fischer et al. (2017b)*.

A Cellular Automata (CA) is a mathematical concept that permits the generation of discrete time-evolving cells grids. At a given CA time-step (t_{CA}), the states of the cells is simultaneously modified following a global transition rule which involves the states of the cells in the neighborhood of each cell of the grid (*Von Neumann and Burks 1966*).

The model in the CADI method is built as a lattice space Γ discretized in m squared property cells (in our case transmissivities) which are grouped in m_{CA} different CA pilot subspaces noted $\phi_i, i \in [1, m_{CA}]$ (with $m_{CA} \ll m$). The cells in a CA subspace can have two possible states: (1) state ‘matrix’ with a transmissivity value β_{matrix} or (2) state ‘conduit’ with a transmissivity value $\beta_{conduit}$. The global structural distribution of the transmissivities in the model Γ is, thus, monitored by the different CA subspaces. Each subspace is piloted by a weighting distribution assigned to the neighborhood of each cell of the subspace (noted N). This distribution is set up among an inner circle of ‘activator’ ‘matrix’ neighbor cells and an outer circle of ‘inhibitor’ ‘matrix’ neighbor cells (Figure 5.1). An ‘activator’ ‘matrix’ neighbor will tend to transform a given cell of the subspace (in grey in Figure 5.1) in a state ‘matrix’, while an ‘inhibitor’ ‘matrix’ neighbor will tend to transform this cell in a state ‘conduit’. These circles are also radially split in 8 weighting sectors for a higher weight distribution possibility. At a CA time-step t_{CA} , a global transition rule compares, alternatively for each cell of the subspace, the cells in state ‘matrix’ in the ‘activator’ and ‘inhibitor’ sectors of its neighborhood. For example, for a given cell, if the sum of ‘matrix’-state weights in its activator sector is higher than the ‘matrix’-state weights in its inhibitor sector, then this cell will become ‘matrix’ in the next time step $t_{CA} + 1$, and in the opposite case this cell will become ‘conduit’. In the CADI algorithm we configured 8 different weighting distribution possibilities $N_i, i \in [1, 8]$ which define 8 different directions of propagation for the conduit in a CA subspace (see the 8 configurations in Figure 5.1). After several successive CA time steps transitions, a subspace ϕ will converge to a stable geometry noted $\hat{\phi}$ which depends on the weighting distribution chosen for N .

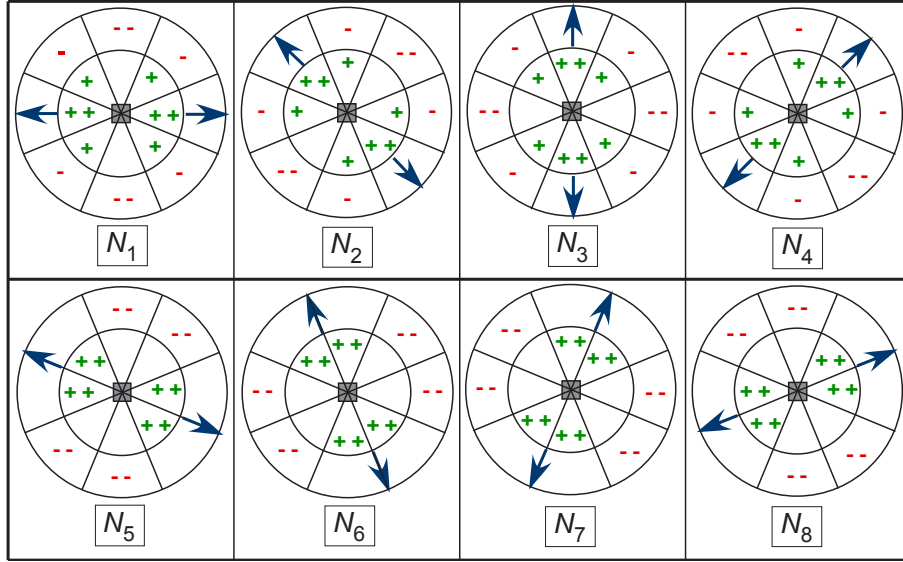


Figure 5.1: Scheme of the 8 different weighting distributions N possibilities to parameterize the CA subspaces. Each distribution defines a different direction for the conduit-state generation shown by the arrows. The dual radius neighborhood is described here for a given cell in grey (the other cells are not shown for a reason of readability). In the configurations $N_i, i \in [1, 4]$ the circles are defined by an inner circle of radius 2 cells and an outer circle of radius 6 cells, and in the configurations $N_i, i \in [5, 8]$ the circles are defined by an inner circle of radius 4 cells and an outer circle of radius 5 cells. The neighbor cells of the greyed cell are split in 8 internal 'activator' weighting sectors and 8 external 'inhibitor' weighting sectors represented by the two radially split circles. A neighbor cell in state matrix can be associated (given its position in the neighborhood) to a positive weight '+' which is twice higher than a '-' weight, or to a negative weight '-' which is twice higher than a '+' weight, or to a null weight in the empty sectors and beyond the neighborhood.

The 8 configurations aforementioned are considered as the different possibilities for the CA subspace parameterization during the structural optimization of the inversion process. Thus, in the CADI method a converged CA subspace $\hat{\Phi}(N, \beta)$ is parameterized by one of the eight weighted neighborhood configuration $N_i, i \in [1, 8]$ as structural parameter and by the transmissivity values $\beta = [\beta_{\text{matrix}}, \beta_{\text{conduit}}]$ as property parameter. The global partitioned model composed of all converged CA subspaces $\hat{\Phi}_i, i \in [1, m_{\text{CA}}]$ will be referred to as $\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)$ with \mathbf{P}_N a m_{CA} vector of the different structural parameters piloting each CA subspace and \mathbf{P}_β a $2m_{\text{CA}}$ vector of the β_{conduit} and β_{matrix} transmissivity values in each CA subspace of the model.

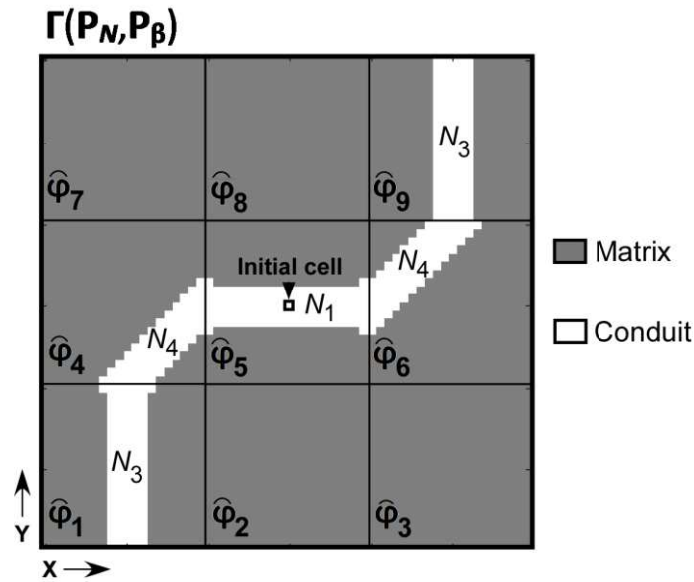


Figure 5.2: Presentation of a model in the CAD algorithm. Here the model is partitioned in 9 subspaces controlled by CA. The model is parameterized by a structural parameter \mathbf{P}_N (here $\mathbf{P}_N(5) = N_1$; $\mathbf{P}_N(4) = \mathbf{P}_N(6) = N_4$ and $\mathbf{P}_N(1) = \mathbf{P}_N(9) = N_3$ (see Figure 5.1)) and a property values parameter \mathbf{P}_β (here every subspace is defined by the same β but it could vary in each subspace). Initially the whole model is considered as matrix, except an initial conduit cell. Within the CA time process the conduit is generated from this initial cell and propagates through the model depending on the subspaces structural parameters until it reaches a global converged geometry.

Initially in the CA temporal process, the whole model starts in an entire matrix state with only a single cell in state conduit. In the first CA time step the conduit appears in the subspace of this initial cell, and as it arrives to the limit of this subspace it will potentially enter new subspaces (by local symmetry at the boundary between these subspaces) with another generation direction. Thus, from a unique conduit cell, the different subspaces permit the generation of a complex conduit network model at the end of the convergence of all CA (Figure 5.2).

5.2.2.2 Inverse problem

In the Bayesian framework, the inverse problem retrieves a model matching the observed data and respecting the priori information (Tarantola and Valette 1982).

The optimization of the unknown parameters (here structural and property parameters) can be achieved by using an iterative deterministic algorithm that minimizes sequentially the following objective functions:

$$\begin{aligned} \Psi_{\text{structure}}(\mathbf{P}_N) = & \frac{1}{2} \left(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)) \right)^T \mathbf{C}_d^{-1} \left(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)) \right) \\ & + \frac{1}{2} (\mathbf{P}_{N,\text{prior}} - \mathbf{P}_N)^T \mathbf{C}_{\mathbf{P}_N}^{-1} (\mathbf{P}_{N,\text{prior}} - \mathbf{P}_N) \end{aligned} \quad (5.1)$$

$$\begin{aligned} \Psi_{\text{properties}}(\mathbf{P}_\beta) = & \frac{1}{2} \left(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)) \right)^T \mathbf{C}_d^{-1} \left(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)) \right) \\ & + \frac{1}{2} (\mathbf{P}_{\beta,\text{prior}} - \mathbf{P}_\beta)^T \mathbf{C}_{\mathbf{P}_\beta}^{-1} (\mathbf{P}_{\beta,\text{prior}} - \mathbf{P}_\beta) \end{aligned} \quad (5.2)$$

with \mathbf{d}_{obs} the vector of $n \times 1$ observed data (such as hydraulic responses from pumping tests), $\mathbf{P}_{N,\text{prior}}$ and $\mathbf{P}_{\beta,\text{prior}}$ the *a priori* information to constrain the inversion of the structural and property parameters, \mathbf{C}_d a $(n \times n)$ covariance matrix of uncertainties on data, and $\mathbf{C}_{\mathbf{P}_N}$ a $(m_{\text{CA}} \times m_{\text{CA}})$ and $\mathbf{C}_{\mathbf{P}_\beta}$ a $(2m_{\text{CA}} \times 2m_{\text{CA}})$ covariance matrices of uncertainties on prior parameters. $f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta))$ denotes the nonlinear forward problem operator that links the hydraulic head data and the transmissivity field. $\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)$ is the spatial partition of the transmissivity model Γ that is parameterized by the CA via the structural parameters \mathbf{P}_N and its property parameters \mathbf{P}_β .

The inversion process is conducted sequentially. First, the parameters \mathbf{P}_N and \mathbf{P}_β are initialized with reasonably chosen structural directions and transmissivity values for each subspace of the model. Then, the structural parameter \mathbf{P}_N is iteratively estimated with the fixed initial transmissivity model \mathbf{P}_β . Afterward, once the structural geometry is optimized, the spatial distribution of the transmissivity parameters \mathbf{P}_β is reconstructed considering this optimized structural geometry.

5.2.2.3 Optimization and uncertainties estimation

The optimization process begins with a sensitivity analysis: a local ‘One-factor-At-the-Time’ (OAT) perturbation method, according to the classification in *Pianosi et al. (2016)*. The

structural sensitivity analysis considers the eight different neighborhood configuration (Figure 5.1) as structural parameters possibilities in each CA subspace in order to modify the conduit network and minimize the difference between the modeled data to the observed data. This sensitivity analysis establishes a $(8 \times m_{CA})$ sensitivity matrix \mathbf{S} .

At a k th iteration, for a modification in a CA subspace j by testing a configuration N_i , the element (i, j) of the matrix \mathbf{S} is calculated as:

$$\mathbf{S}^k(i, j) = \left(\mathbf{d}_{\text{obs}} - f\left(\Gamma\left(\mathbf{P}_N^k \Big|_{\mathbf{P}_N^k(j)=N_i}, \mathbf{P}_\beta\right)\right) \right)^T \mathbf{C}_d^{-1} \left(\mathbf{d}_{\text{obs}} - f\left(\Gamma\left(\mathbf{P}_N^k \Big|_{\mathbf{P}_N^k(j)=N_i}, \mathbf{P}_\beta\right)\right) \right) + \frac{1}{2} \left(\mathbf{P}_{N,\text{prior}}(j) - N_i \right)^T \mathbf{C}_{\mathbf{P}_N}^{-1} \left(\mathbf{P}_{N,\text{prior}}(j) - N_i \right), \quad (5.3)$$

where $f\left(\Gamma\left(\mathbf{P}_N^k \Big|_{\mathbf{P}_N^k(j)=N_i}, \mathbf{P}_\beta\right)\right)$ represents the data modeled through the local subspace parameter perturbation and $\mathbf{P}_{N,\text{prior}}(j) - N_i$ represents the gap between the local prior direction and the perturbation direction.

The new structural parameter \mathbf{P}_N^{k+1} is updated with the minimal value found in the matrix \mathbf{S} at the index $(i, j)_{\min}$ (where i represents the new configuration N_i for the subspace j). The updated parameter is built as an unique local improvement: $\mathbf{P}_N^{k+1} = \mathbf{P}_N^k$ except for $\mathbf{P}_N^{k+1}(j_{\min}) = N_{i_{\min}}$. This update will generate a new structure for the iteration $k+1$. The same sensitivity analysis is repeated at each iteration until there is no more possible improvements in the structure.

Then, the uncertainties of the inverted structure are estimated from the last iteration sensitivity matrix, the prior uncertainties and the value of the objective function:

$$\mathbf{C}_{\mathbf{P}_N}^{\text{post}}(j) = \left(\frac{1}{8} \sum_{i=1}^8 \hat{\mathbf{S}}(i, j) - \Psi_{\text{structure}}^{\text{post}} + \mathbf{C}_{\mathbf{P}_N}^{-1}(j, j) \right)^{-1}, \quad (5.4)$$

with $\mathbf{C}_{\mathbf{P}_N}^{post}(\mathbf{j})$ the structural uncertainty for the subspace \mathbf{j} , $\hat{\mathbf{S}}$ the last iteration sensitivity matrix and $\Psi_{structure}^{post}$ the value of the objective function after optimization. If a subspace conduit is well-constrained its value $\mathbf{C}_{\mathbf{P}_N}^{post}$ should be low.

After the structural optimization, a second optimization is led on the subspaces conduit and matrix transmissivity values \mathbf{P}_β , for the previously inverted structure. This optimization is an iterative process using a finite difference sensitivity analysis. The $(n \times 2m_{CA})$ Jacobian sensitivity matrix is defined for its index (i, j) as:

$$\mathbf{J}(i, j) = \left. \frac{\partial f_i}{\partial \mathbf{P}_\beta} \right|_{\mathbf{P}_\beta(j) = \mathbf{P}_\beta(j) + \Delta \mathbf{P}_\beta}, \quad (5.5)$$

with $\Delta \mathbf{P}_\beta$ the finite difference step, f_i the forward problem variation on a data i for a variation on $\mathbf{P}_\beta(j)$.

At an iteration k , the updated values \mathbf{P}_β^{k+1} are calculated from a Newton linearization:

$$\begin{aligned} \mathbf{P}_\beta^{k+1} = & \mathbf{P}_\beta^k + \left(\left(\mathbf{J}^k \right)^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{J}^k + \mathbf{C}_{\mathbf{P}_\beta}^{-1} \right)^{-1} \cdot \left(\mathbf{J}^k \right)^T \cdot \mathbf{C}_d^{-1} \cdot \left(\mathbf{d}_{obs} - f \left(\Gamma \left(\mathbf{P}_N, \mathbf{P}_\beta^k \right) \right) \right) \\ & + \mathbf{C}_{\mathbf{P}_\beta}^{-1} \cdot \left(\mathbf{P}_{\beta, prior} - \mathbf{P}_\beta^k \right) \end{aligned} \quad (5.6)$$

This iterative process continues until an acceptable convergence of the objective function is achieved.

Then, the uncertainties of the property values are estimated from the last iteration Jacobian matrix \mathbf{J}^{post} with the posterior covariance matrix $\mathbf{C}_{\mathbf{P}_\beta}^{post}$:

$$\mathbf{C}_{\mathbf{P}_\beta}^{post} = \left(\left(\mathbf{J}^{post} \right)^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{J}^{post} + \mathbf{C}_{\mathbf{P}_\beta}^{-1} \right)^{-1}. \quad (5.7)$$

The diagonal entries of this matrix give the variances of the property values. The square root values of these entries represent the standard deviation of the uncertainties on the estimated transmissivity field.

5.2.3 Application

5.2.3.1 Site presentation

The Terrieu experimental site is located in the North of Montpellier, Southern France (Figure 5.3a). The site has been performed for the hydrodynamic studies of the Lez aquifer that is mainly composed of Early Cretaceous and Late Jurassic limestones. This site is one of the sites of the French research network SO Karst which was developed to monitor the karstic aquifers in France (*Jourde et al. 2011* ; www.sokarst.org).

The Terrieu site sits on a local monocline structure that trends NE-SW and dips at about 15 to 20 degrees towards NW. The surface area of the experimental site is about 2,400 m² (40 m by 60 m). Detailed fracture mappings, conducted at the ground surface of the site and nearby outcrops, have indicated that two major fracture sets (trending ENE-WSW and NW-SE, respectively) are present in the study area (*Jazayeri Noushabadi 2009* ; *Wang et al. 2016*).

Twenty-two boreholes have been drilled at the site to study the local-scale hydraulic behavior of the aquifer (Figure 5.3b). These boreholes are vertical and have a mean total drilled depth of 55 m. Downhole videos recorded in some boreholes have shown that the upper 30-40 m of the drilled formations are largely comprised of thin-layered, marly, early Cretaceous limestones while the lower part mainly consists of purer, massive and non-aquifer late Jurassic limestones. The early Cretaceous limestone has a low permeability therefore it forms a confined upper boundary for the aquifer existing at the interface between these two units. A number of well-developed karstic conduits, with apertures up to 50 cm, have also been identified on downhole video logs. These karstic conduits were found to be present at a depth between 35 m to 40 m (*Jazayeri Noushabadi et al. 2011* ; *Wang et al. 2016*) at the interface of the aforementioned two units. The observed local orientation of the karstic conduits is indicated as green lines in Figure 5.3c. The extent of the lower rock unit is unknown due to limited drilled depth (maximum of 60 m). Well logs (temperature and electrical conductivity) and straddle packer tests have shown that a preferential flow path (blue line in Figure 5.3c) exists along the major bedding plane corresponding to the interface between the two major rock units (*Jourde et al 2002* ; *Dausse 2015*). All the observations from downhole videos, well logs, and packer tests have led to the conceptual model that a network of interconnected karstic conduits developed along an important bedding plane comprises the main flow paths of the experimental site (*Wang et al. 2016*).

Eight cross-hole pumping tests, in the form of hydraulic tomography, have been performed at the experimental site. The applied flow rate of each pumping test ranges from 0.2 to 53 m³/h depending on the well productivity and whether the well is connected to a high-permeability feature (Figure 5.3c). The drawdown of water table level generated by the pumping tests did not reach to the depth of the important bedding plane where main flow occurs; this means that the karstic network was saturated during the pumping tests. All tests reached steady state. The field-scale hydraulic tomography yielded a total number of 168 drawdown steady-state measurements, which is the main dataset of the present work.

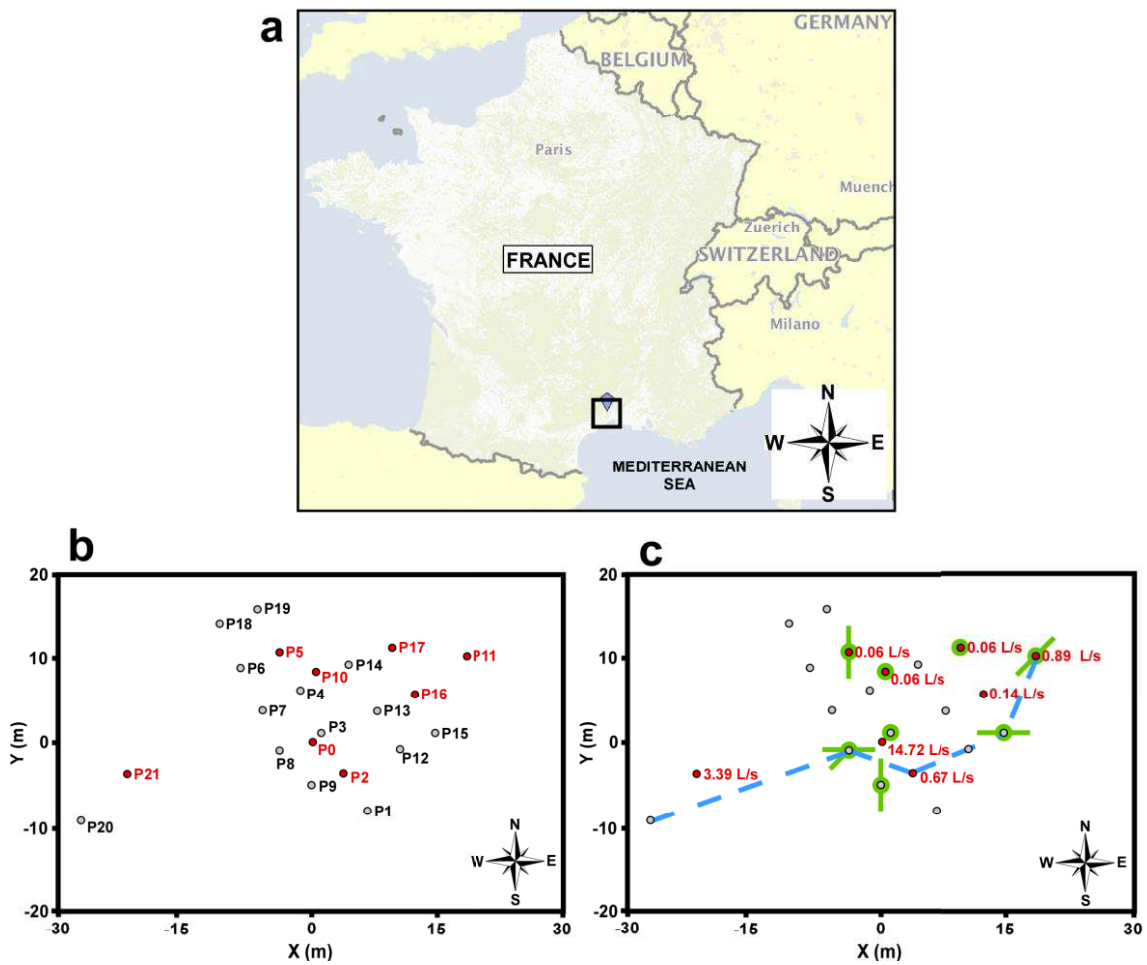


Figure 5.3: (a) Map indicating the location of the experimental site. The black square indicates the location of the Lez aquifer in which the Terrieu site is included. (b) Distribution of twenty-two boreholes of the Terrieu experimental site. The red dots indicate the boreholes where the pumping tests were performed while the grey dots indicate the measurement boreholes. (c) Pumping rates (red captions). Inferred principal flow path connectivity (blue dotted lines) and local karstic conduits (green lines) based on downhole videos, well logs, and packer tests. The orientation of the green lines indicates the orientation of local karstic features observed on downhole videos. A green dot indicates that no karstic features were seen in this borehole.

This set of drawdown responses represents the observed data used in the inversion process, while the inferred preferential flow path and local conduits direction information presented in Figure 5.3c are not taken into account in the inversion but will be used for evaluating the effectiveness of the inversion results.

5.2.3.2 Modeling method

We have applied the CADI method to image the hydraulic transmissivity distribution at the Terrieu site from the joint inversion of the set of 168 steady state drawdown data. A two dimensional equivalent porous medium parameterization was adopted to model the domain. The porous medium was distributed in two states: bedding plane fissured matrix (CADI matrix state) or karstic conduit (CADI conduit state). We adopted an evolving meshing, that is, for each iteration, refined preferentially at the boundaries of the conduits. This avoids an over-meshing of the model, especially in the matrix zones and, thus, reduces the computation time of the forward problem solver. The forward problem (i.e. steady-state diffusion equation in saturated porous media) was solved using a commercial software (COMSOL Multiphysics). The inversion process was implemented as a MATLAB code that is connected to the COMSOL solvers via a local server. The local flow network model was enclosed in a large regional buffer zone ($1,000 \times 1,000 \text{ m}^2$) to reduce the influence of the boundaries conditions.

Based on a multiscale inversion method (*Grimstadt et al. 2003*), the flow network model was sequentially partitioned during the inversion process (see Figure 5.4).

The different inversion parameters chosen for the final inversion (16×24 CA subspaces) are presented in the Table 5.1. The property value parameter β was chosen to define the exponent of the transmissivity ($\log_{10}(T)$). Thus, the prior standard deviation also applies to the exponent ($T \times 10^{\pm\sigma_T}$). The initial transmissivities values were chosen according to the values found during the field characterizations done by *Jazayeri Noushabadi (2009)* and *Dausse (2015)*. We did not make any prior assumption on the structure directions in \mathbf{P}_N and $\mathbf{C}_{\mathbf{P}_N}$.

We began the inversion with a model partitioned in a large discretization 4×6 CA subspaces. In the flow network model, the karstic conduit was generated from an initial fracture cell placed at the P8 borehole (Figure 5.4).

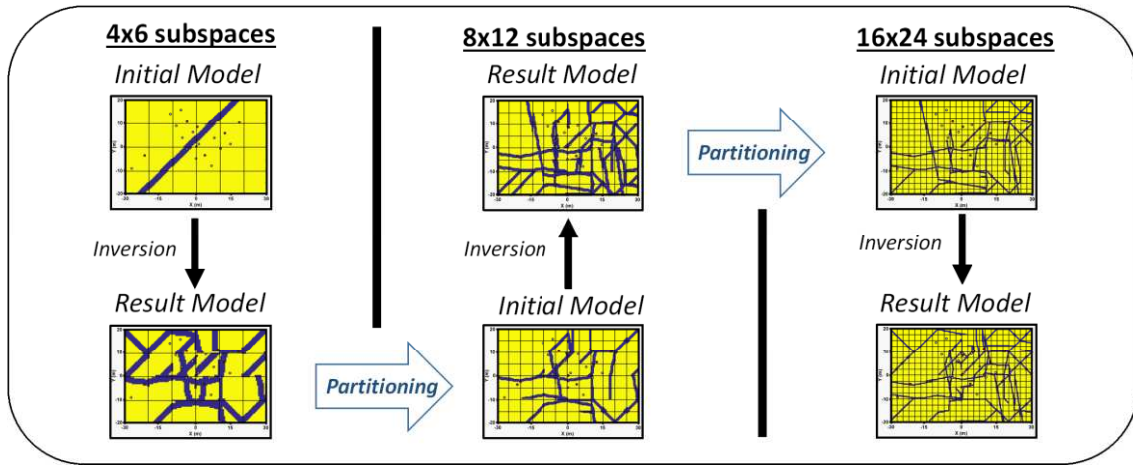


Figure 5.4: Schematic showing the sequential series of inversions led to obtain the final flow network model. The initial model was partitioned with 4×6 subspaces for its inversion. The inverted flow network model was then used as a new initial model for an inversion with 8×12 subspaces. The same operation was repeated on last time so that our final flow network has a partitioning of 16×24 subspaces.

Table 5.1: List of the inversion parameter values chosen for the final inversion (16×24 subspaces)

Parameter	Value
Model dimension	$X = [-30, 30] \text{ m} ; Y = [-20, 20] \text{ m}$
Partitioning / Total cells amount ($Y \times X$)	$16 \times 24 / 320 \times 480$
Conduit aperture	50 cm
Data uncertainties	$C_d = \sigma_{\text{data}}^2 \times Id(n) ; \sigma_{\text{data}}^2 = 0.01 \text{ m}$
\log_{10} transmissivity prior standard deviation	$C_{p_p} = \sigma_T^2 \times Id(2m_{CA})$ $\sigma_T^2 = 0.001$ for the conduits ; $\sigma_T^2 = 1$ for the matrix
Initial \log_{10} transmissivities	$\beta_{\text{conduit}} = -1 (10^{-1} \text{ m}^2/\text{s}) ; \beta_{\text{matrix}} = -8 (10^{-8} \text{ m}^2/\text{s})$ $\beta_{\text{regional}} = -2 (10^{-2} \text{ m}^2/\text{s})$
Finite difference step	$\Delta P_\beta = 10^{-4}$
Observation data (d_{obs})	168 steady state drawdown responses

This choice was made based on field hydrogeological knowledge that P8 has the highest probability of intersecting the main flow path. The initial network was a simple conduit with a direction choice based on the principal flow path direction and with an aperture of 2 m. The remaining parts of the model were assigned with a transmissivity of matrix. As the structural

geometry for this initial conduit was optimized, we added a second conduit in the model, orthogonally to the principal flow path and intersecting P8, and we led the structural inversion with these two principal conduits directions. When the inversion process found a 4×6 subspaces result model, we refined the obtained solution to a model partitioned in 8×12 subspaces (a CA subspace was divided into four new ones with the same parameters) and used it as initial model for a more precise inversion solution with conduits with an aperture of 1 m. Finally, we partitioned once more this new 8×12 solution into a 16×24 subspaces model which was once again used as the initial model for a final inversion solution with conduits with an aperture of 50 cm. This sequential modeling reduces the inversion time and allows interpretation of the importance of the flow paths found according to their emergence at different scales during the inversion process and thus their influence on the convergence of the objective function.

5.2.3.3 Results and discussion

The simulated hydraulic heads from the final emergent flow network model show a high correlation to the field measurements (Figure 5.5a). It can be seen that a higher degree of difference between the simulated and measured hydraulic head exists for small drawdowns. This may be due to the fact that the small hydraulic drawdowns are more sensible to micro-fractures that were embedded in the equivalent porous matrix in the current model. The resulting karstic flow-network model from the inversion process is presented in Figure 5.5b. The hydraulic transmissivity of the bedding plane fissured matrix was only slightly perturbed (i.e. a small deviation from the initial value of $10^{-8} \text{ m}^2/\text{s}$) during the entire inversion process (Figure 5.5b), the drawdown data were essentially reproduced only by the karstic network geometry. The emergent flow network shows a high consistency with field observations shown in Figure 5.3c in terms of connectivity between boreholes (Figure 5.5c) and local direction of karstic conduits (Figure 5.5d). The simulated drawdown maps from the resultant flow network model are presented in Figure 5.6.

The simulated karstic network structure in the Terrieu experimental site investigated through the hydraulic tomography can be schematized using the emergent flow network model as shown in Figure 5.7. This schema conceptualizes the site in three dimensions taking into account the slope of the bedding plane.

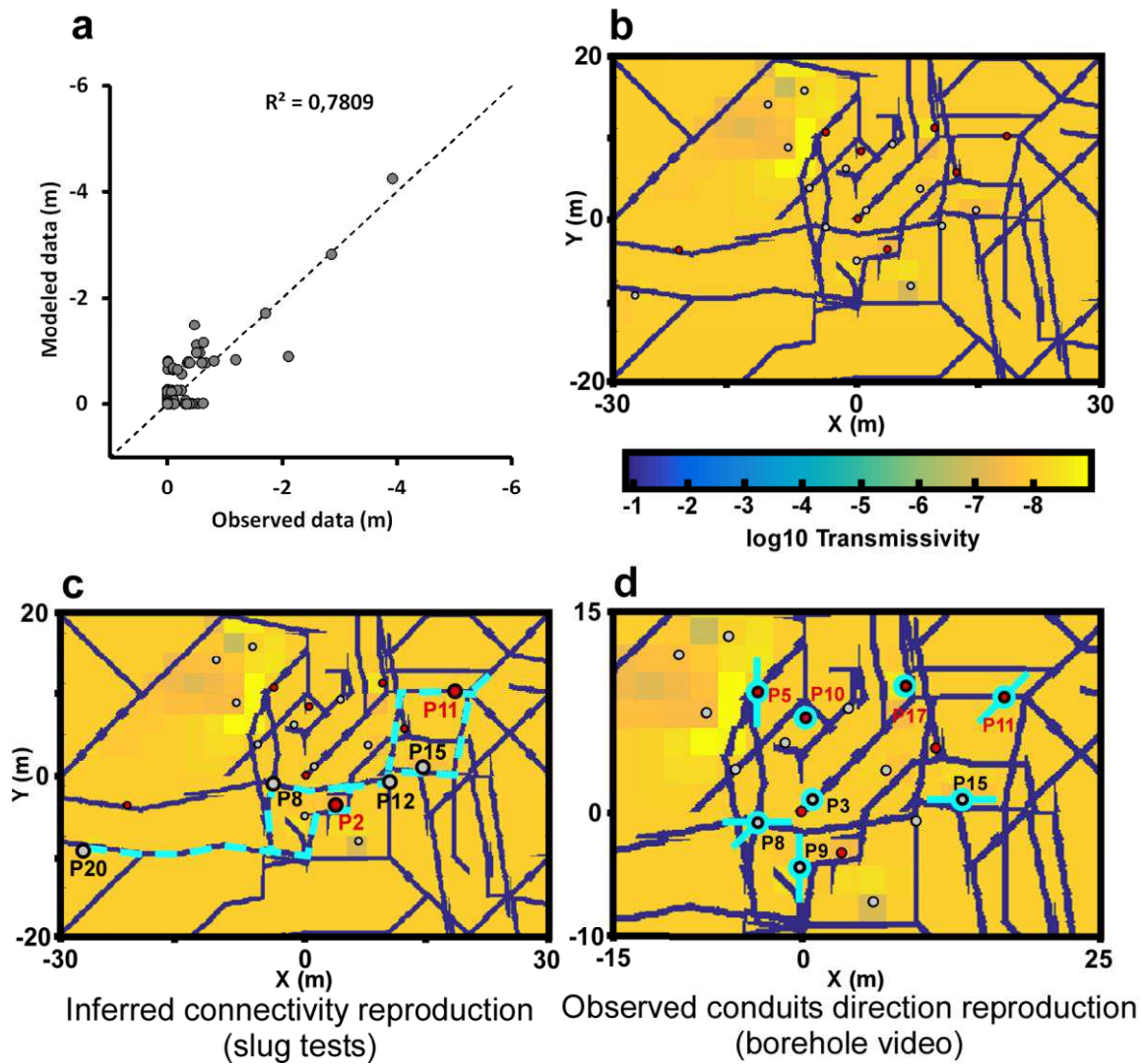


Figure 5.5: (a) Comparison of the observed drawdowns to the drawdowns modeled by the inverted flow model. (b) Resultant model of the inversion modeling showing the heterogeneous distribution of the transmissivities. (c) Comparison of the result model with the known preferential flow path connectivity (interpreted in the model in dotted blue lines). (d) Superimposition of the known local conduits direction (shown as blue lines) presented in Figure 5.3c.

The maps of the simulated drawdowns (Figure 5.6) using the final flow network model highlight the high degree of heterogeneity of the experimental site. The steady-state drawdown cone of each pumping test is highly irregular, and the shape of the drawdown cone can have very different geometry in different pumping tests.

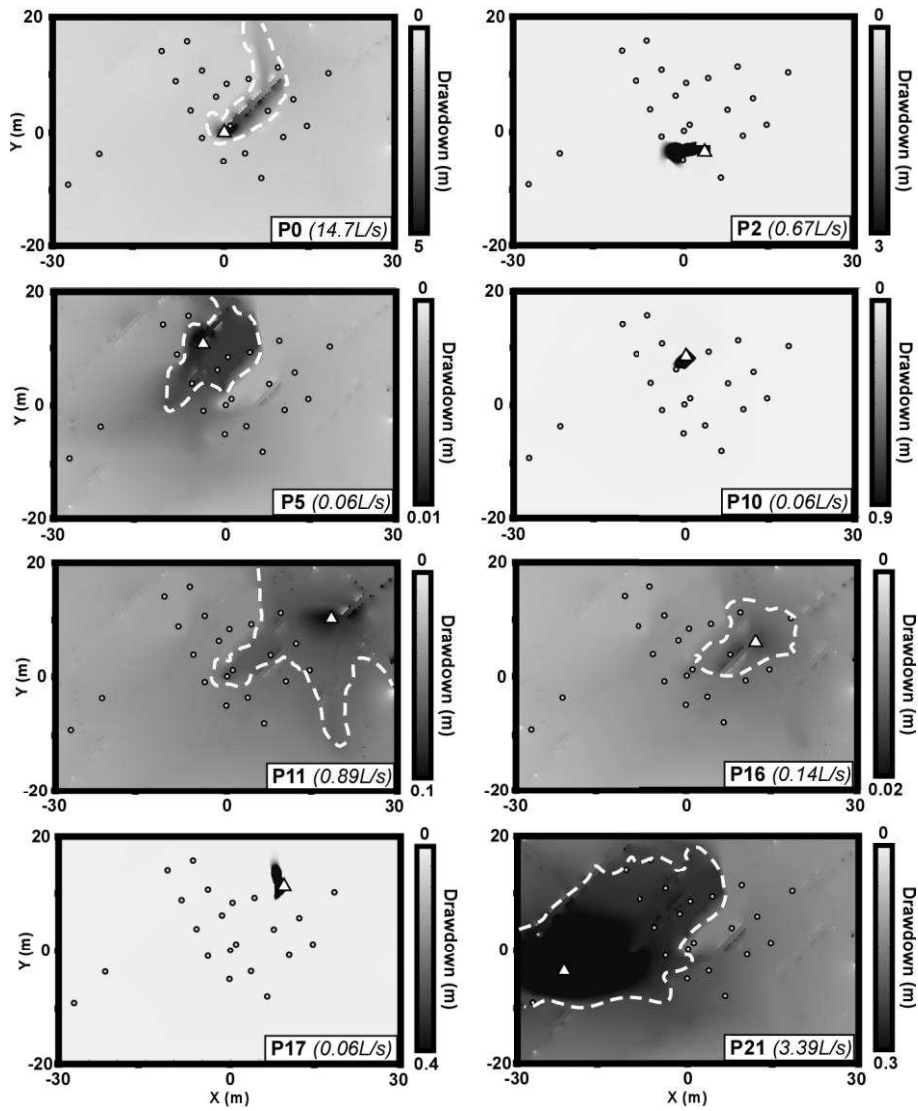


Figure 5.6: Maps of hydraulic drawdowns calculated from the result flow network model. The drawdowns are shown for each of the pumping wells (white triangles) used for the hydraulic tomography (the pumping rate is indicated in each figure). The drawdowns can have very different forms depending on the localization of the borehole in a conduit or in the matrix, highlighting the heterogeneity of the model. Pumping in the matrix (P2, P10, P17) results in a very local drawdown, while pumping in a conduit (P0, P5, P11, P16, P21) produces a more global drawdown in the whole model (in these cases the area the most impacted by the pumping is delimited by white dotted lines)

Three types of behaviors can be distinguished: (1) pumping in the non-karstified bedding plane matrix (P2, P10, P17) associated with a low pumping rate and a small influence zone of drawdown localized around the pumping borehole ; (2) pumping in low-productivity conduits (P5, P11, P16) associated with a low pumping rate and a large influence zone of small drawdowns covering the entire field ; and (3) pumping in high-productivity conduits (P0, P21) associated with a high pumping rate and a large influence zone of drawdowns impacting the whole field.

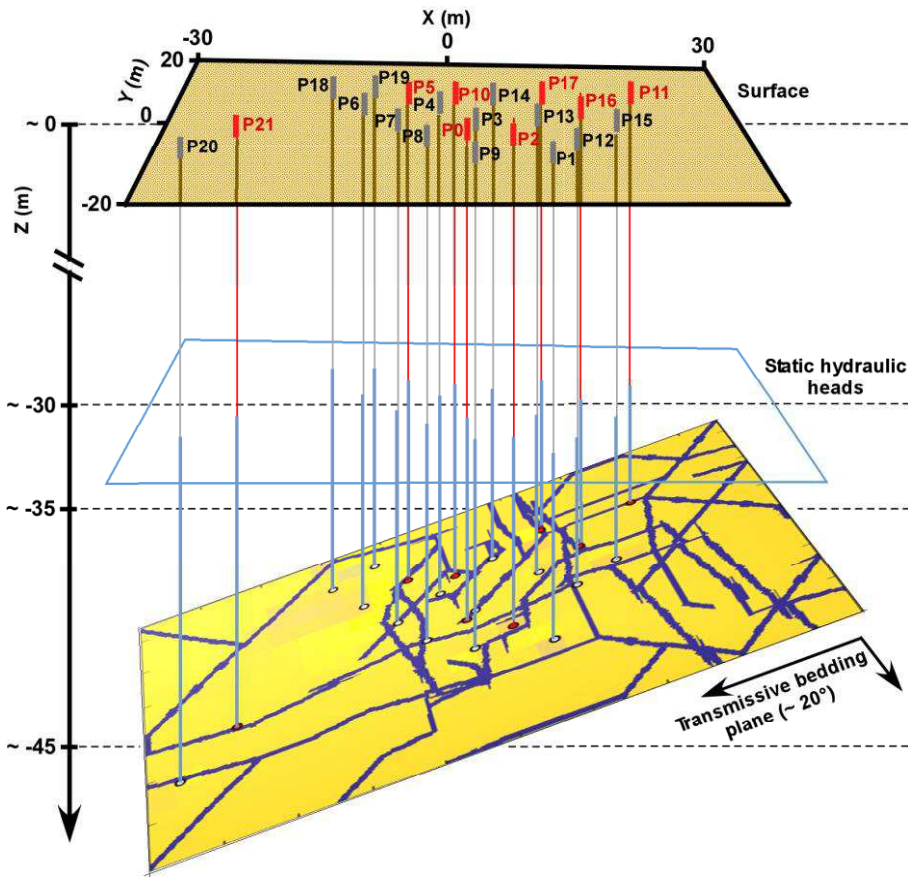


Figure 5.7: Schematic representation of the modeled karstic structure at the Terrieu experimental site, considering the geological information, the hydraulic tomography investigation, and the flow network produced by inversion with the CADI method. The red lines indicate the boreholes where the pumping tests were performed, while the grey lines indicate the measurement boreholes.

In our model, pumping in the matrix only induces a response in the vicinity of the pumped well, while pumping in the conduit network induces responses impacting significantly a much larger impact area (delineated by dotted white lines in Figure 5.6). This shows that the global flows in the model are controlled by the karstic conduits, which can be linked to the real behavior of the site in which flows follow the important discontinuities (*Jourde et al 2002*). The modeled steady-state drawdowns were compared to the field observations made by *Jazayeri Noushabadi (2009)* and *Dausse (2015)*, and they appear to be similar to the estimated drawdowns from the field pumping tests.

To verify the results, we have compared the inverted flow network to the mapped connectivity and local direction knowledges gained from slug tests and downhole videos (Figure 5.3c). This information was not taken into account during the inversion process, but they permit to assess the effectiveness of the CADI approach with the steady state drawdown responses.

The observed connectivity described in Figure 5.3c between P20, P8, P2, P12, P15 and P11 can be reproduced by the inverted flow network model (Figure 5.5c). However, the drawdown map for P2 (Figure 5.6), shows that it is not directly connected to the main flow network in the model. We noticed that the establishment of this connectivity appeared quickly during the multi-scale structural inversion (Figure 5.4). It indicates the importance of the connectivity between these boreholes toward the reproduction of the global drawdown distribution on the site.

Regarding the knowledge about local conduit direction (Figure 5.5d), in the inverted model, P8 and P15 are in fact intersecting conduits following the same local direction as observed from downhole videos. In the cases of P5 and P11, where conduits have been observed in the field, the inverted flow network map shows that they intersect preferential flow paths, but not following exactly the same direction as field observations. From the drawdown maps (Figure 5.6), we can see that simulated pumping in P10 and P17 in our model behave like pumping in the matrix; thus, P10 and P17 are not connected to the karstic network. This is consistent with the information highlighted by camera observations. Only two inverted local flow structures do not reproduce downhole videos observations: P3 seems to be connected to the karstic network; however, no karstic conduits were identified on its downhole video. On the contrary, P9 seems to be located in the matrix in the inverted model, while a conduit has been observed in its borehole video.

In general, the inverted flow network model reconstructs most of the knowledges that we have about the site, even if this information was not used in the inversion process. Thus, it seems that the drawdown data set alone provides sufficient information to reproduce the true network connectivity between boreholes.

Figure 5.8 presents uncertainty maps, which reflects the reliability of the inversion results, and the areas of the model that are covered by the information given from the different local responses in boreholes. The posterior structural uncertainty map indicates that the structure uncertainty is lower in the center of the model domain than in the periphery. The highest uncertainty in the inverted flow structures (in orange/yellow in Figure 5.8) occurs consistently in the regions beyond the one constrained by the borehole pattern, where no hydraulic information is available. The most reliable parts of the structure (in blue/green) are located in the middle of the domain, where the borehole pattern is dense, and therefore provides more hydraulic information. The map of the posterior standard deviations of the \log_{10}

transmissivity shows that the matrix transmissivity value is well-constrained only near some boreholes (e.g. P0, P2, P4, P9, P10, P13 and P17). Thus, the pumping tests do not permit a characterization of the bedding plane transmissivity very far from the boreholes.

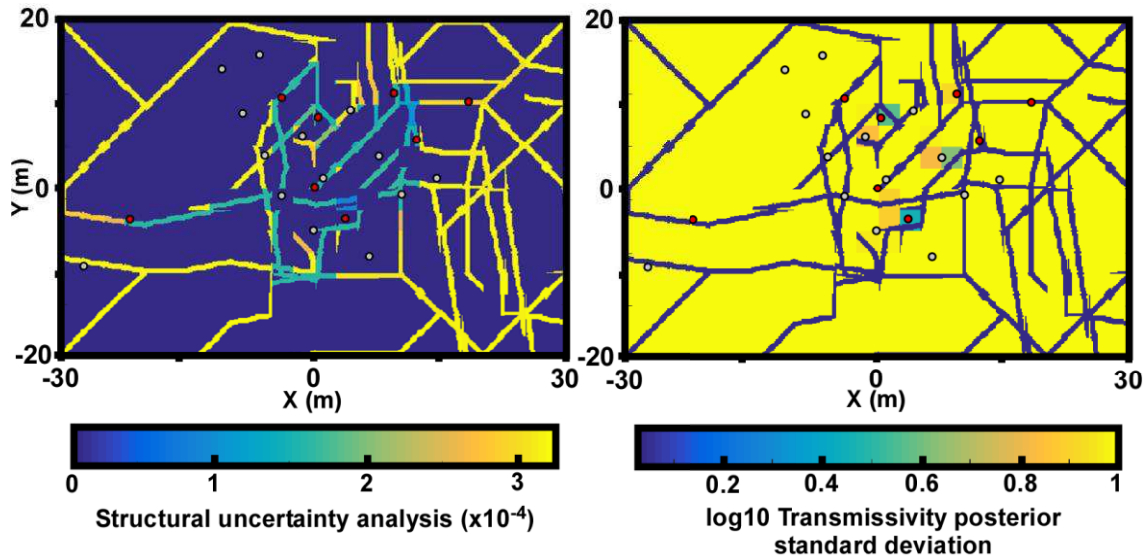


Figure 5.8: The map of the network structural uncertainties (left) shows that the network geometry is well constrained especially in a zone between each borehole in the center of the model, and compared to the map of transmissivities standard deviation (right), the hydraulic data permitted to constrain more the conduits position than the matrix.

Figure 5.9 shows maps of flow velocities for pumping in P0 and P21, the two most productive boreholes, intersected by karstic conduits. The distribution of the modeled flow velocities during each pumping test are controlled by the high-transmissivity karstic features in the model. It appears clearly that the water pumped in these boreholes is mobilized in the karstic network, while the velocity of the water in the bedding plane matrix is low. Even if the CADI method allows a certain imagery of the flow paths that honors the observed hydraulic connectivity to be determined, it may not produce a geomorphologically realistic model. This is caused by the non-unique nature of inverse problem. To provide more realistic karst networks, the hydraulic data should be inverted jointly with other sources of data coming from geophysical and tracer tests or/and constrained by geomorphological and geological information. By comparing the velocity maps for P0 and for P21 we observed that a larger pumping rate (P0) induces an increase of the flow velocities around the borehole, but it also generates a mobilization of water from conduits located further from the pumping borehole, and thus, characterization of a larger area. On the contrary, a lower pumping rate (P21) will

more specifically characterize the karstic conduits near the borehole. This implies that the use of transient modeling of a variable pumping rate to characterize both small scale and large scale karstic network would be useful. The velocities maps also indicate that the North-South oriented conduit in the center of the model ($X = -5\text{ m}$) is less solicited by the pumping in P0 and P21. This same conduit is, however, characterized more specifically by a pumping in P5, which would indicate a dual flow direction in the karstic network of the model (N-S and E-W).

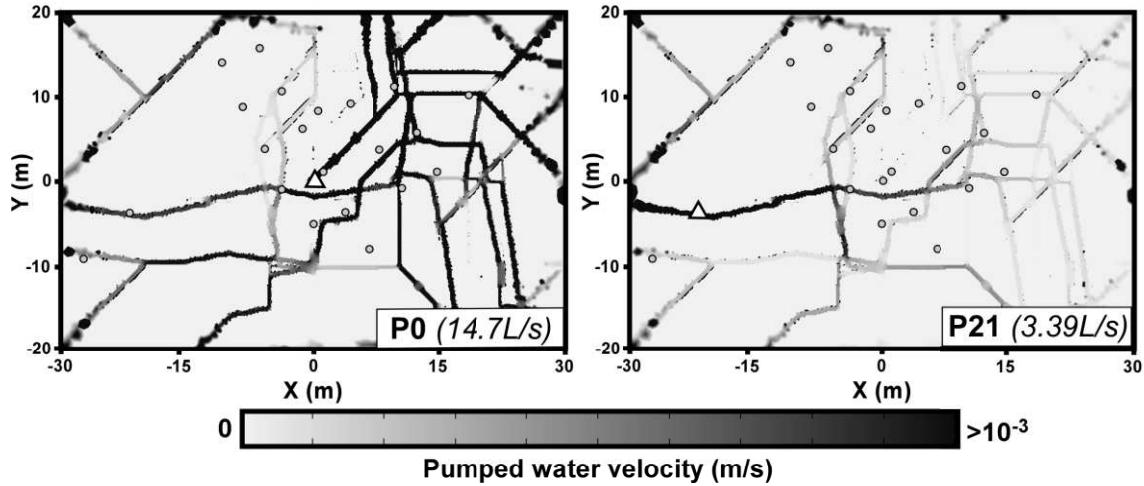


Figure 5.9: Maps of the pumped water velocities calculated by the result model for a pumping in borehole P0 and in borehole P21 (the two most productive pumping). The pumping boreholes are indicated by white triangles. For a reason a better readability of the low velocities, the scale has been fixed on a maximal velocity of 10^{-3} m/s , thus in the blackest zones, the velocity can be higher than this value (up to 10^{-2} m/s near the pumping point for P0).

In Figure 5.10, the inversion result produced by the CADI method is compared to the one produced by the Sparse Non-linear OPTimizer (SNOPT) method at the same field site and with the same data set (Wang *et al.* 2016). The SNOPT method is a classical efficient algorithm for non-linear large-scale inverse problems. The model of Wang *et al.* (2016) is composed of a grid of squared cells of $0.5 \times 0.5\text{ m}^2$; each cell is assigned with a transmissivity value, which is optimized in the inversion process. The optimization of the transmissivity values was constrained by upper and lower limits (10^{-1} and $10^{-8}\text{ m}^2/\text{s}$, respectively).

While the CADI method, associated to the multi-scale process, can be initialized with a simple model to achieve an inverted flow network consistent with field observations, the SNOPT method required some connectivity information included in the initial model to converge to a coherent model. Overall, the SNOPT and CADI methods generate similar results

in terms of medium morphology, where a highly conductive conduit network is embedded in a background matrix. The SNOPT method let more freedom for the optimization of both conduit and matrix transmissivities instead of primarily constraining the flows in a network as done by the CADI method. Thus, it allows establishing a smoother transient of transmissivity from flow paths to the background medium. This is important, in some cases, to reproduce the diffusive behavior of local micro-fissures, which cannot be modelled by the channelized flow in conduits. For this reason, the SNOPT result permits a better reproduction of the measured drawdowns ($R^2 \sim 0.9$ of SNOPT compared to the $R^2 \sim 0.78$ of CADI). However, the result by the CADI method is more consistent with the field knowledge of the local conduit orientation and borehole connectivity presented in Figure 5.3c. The constraints imposed in the CADI method also present the advantage to represent a network in an equivalent porous media result to mimic the hydraulic behaviors in the karstic and fractured environment, thus allowing an easier interpretation of the preferential flow paths positioning and of the connectivity between boreholes.

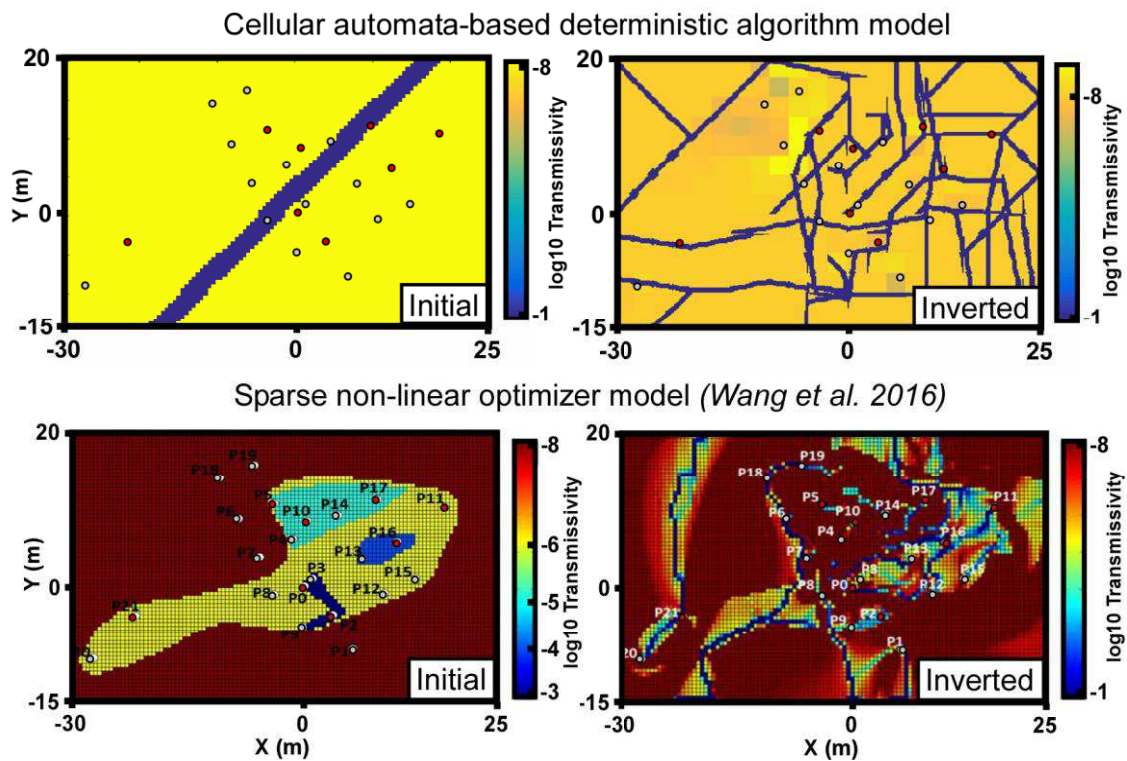


Figure 5.10: Comparison of the inversion result produced by the CADI method and by the SNOPT method (Wang et al. 2016) at the same scale of the Terrieu field site and with same hydraulic dataset. The initial models are shown on the left and the inverted models are presented on the right.

Although the combination of the CADI method and steady-state hydraulic tomography technique appears effective in the identification of flow networks in karstic fields due to its ability of constraining the flows in a network, some limits can be highlighted.

First, we point out that even though the inverted flow paths near the boreholes show a high consistency to field observations, it is possible that some of the inverted conduits may represent the assemble hydraulic behavior of highly fractured areas, while in reality conduits may not be present in those areas. In the same idea, the network in the model is generated with a constant aperture, and thus does not permit to identify the flow paths associated to developed conduits to those associated to fractured zones. To achieve a more realistic network configuration, this method needs to be further developed to include a variable aperture model.

Second, the use of steady state head data may also limit the identification of the preferential paths. In fact, a pumping reaching the steady state in a karstic field mobilizes water in both the conduit network and the fractured rock matrix. The latter tends to blur the hydraulic responses of the conduit network. Using transient data could be useful to improve the inversion result and to reduce the non-uniqueness of the inversion. However, transient inversions may increase significantly the computational demand. A more efficient solution may be to use harmonic pumping (e.g. *Soueid Ahmed et al. 2016 ; Rabinovich et al. 2015*). The harmonic data permit to perform inversions in the frequency domain and may highlight the influence of flow from specific components of the system, and therefore would achieve a better characterization of the conduit network.

5.2.4 Conclusion

The use of the CADI method for a site scale karstic field characterization shows its capabilities in identifying the preferential flow network. The inverted distributed karstic flow-network model reproduces the observed data while maintaining the realism of a highly heterogeneous flows distribution. The constraints imposed for the inversion process result in a model that localizes the conduits of the karstic network, which in turn control the main drawdowns direction.

The use of hydraulic tomography data in the inversion permits the characterization of the highly heterogeneous discrete conduit network. However, this requires a large number of boreholes, and, especially boreholes intersecting the karstic network. Results of the flow velocities in P0 and P21 show that pumping in boreholes intersecting the karstic network with different pumping rate could characterize the network at different scales (locally or more globally). This promotes the use, in the future, of the recently developed harmonic pumping technique at the Terrieu experimental site to identify more precisely the geometry of the karstic network with less pumping tests.

Some necessary conditions were met in the case of the Terrieu experimental site to successfully apply the CADI method. The karstic network at this site was mainly constrained within a bedding plane, which permits a two-dimensional modeling. Also, the subsurface flow in the Terrieu site are constrained mainly by the karstic conduits, while the matrix has a very low transmissivity. This permitted the model to start from an assumption of a two domain site (matrix and conduit), distinguished by a high contrast in transmissivity. In the future, the current method will be further developed to deal with transient hydraulic data and/or to use harmonic pumping responses, which are more sensitive to the karst connectivity, for improving the characterization of the geometry of the conduit networks.

6 ANALYSE HYDRAULIQUE DES RÉPONSES À UN POMPAGE HARMONIQUE EN MILIEU KARSTIQUE ET FRACTURÉ

6.1 Contexte

Le chapitre précédent souligne les limites de l'utilisation des réponses à des pompages à débit constant en régime permanent avec la méthode CADI afin de réaliser une tomographie des milieux karstiques. En effet, ces réponses ne permettent pas de hiérarchiser les chemins d'écoulements dans l'interprétation des résultats. Afin de proposer une alternative au pompage à débit constant et d'améliorer la caractérisation des milieux karstiques, ce chapitre propose une analyse synthétique et théorique des réponses hydrauliques d'un milieu fracturé et karstique à des pompages harmoniques, c'est-à-dire à débits oscillatoires.

Cette analyse est menée à la fois dans le domaine temporel et le domaine fréquentiel, dont les équations d'écoulements sont présentées en début de chapitre. Elle est d'abord effectuée à partir d'un modèle théorique simplifié couplant une matrice équivalente à un milieu poreux et un réseau discret représentant des structures karstiques. Des points de mesures sont positionnés à différents endroits représentatifs et complémentaires afin d'en étudier les réponses. L'analyse des réponses dans le domaine temporel montre que celles-ci sont composées d'une partie oscillatoire et d'une partie approximativement linéaire. La même partie oscillatoire des réponses peut être simulée beaucoup plus rapidement dans le domaine fréquentiel, ce qui le rend plus intéressant pour une inversion que le domaine temporel.

L'amplitude et le déphasage des parties oscillatoires des réponses dépend du degré de connectivité existant entre le point de mesure et le point de pompage, ce qui permet de caractériser le type d'écoulement existant entre deux points parmi trois possibilités : connexion par conduits, connexion par matrice ou connexion double. Une modification de la fréquence de l'oscillation du débit de pompage permet de préciser l'interprétation des réponses catégorisées en « connexion double », une période plus courte caractérisant plutôt les points connectés directement par conduits et une période plus longue caractérisant également les points proches de conduits mais localisés dans la matrice, où se produit un échange plus lent.

Une application sur le site du Terrieu est finalement effectuée. Les réponses mesurées lors d'un pompage harmonique sur site ont en effet des comportements proches de celles observées sur le cas théorique, ce qui rend possible une interprétation qualitative des réponses sur la base de l'analyse du modèle théorique simplifié. La connectivité de conduits interprétée par cette méthode sur le site du Terrieu apparaît comme cohérente avec la connectivité déjà établie par des études préalables.

6.2 Hydraulic Analysis of Harmonic Pumping Tests in Frequency and Time Domains for Identifying the Conduits Networks in a Karstic Aquifer

Cette partie est composée de l'article « Hydraulic Analysis of Harmonic Pumping Tests in Frequency and Time Domains for Identifying the Conduits Networks in a Karstic Aquifer », publié dans le journal *Journal of Hydrology* en avril 2018 (*Fischer et al. 2018b*). Le texte a été remis en forme au format du manuscrit. La version originale de l'article est donnée en Annexe 5.



Research papers

Hydraulic analysis of harmonic pumping tests in frequency and time domains for identifying the conduits networks in a karstic aquifer

P. Fischer^{a,*}, A. Jardani^a, M. Cardiff^b, N. Lecoq^a, H. Jourde^c

^a Normandie Univ, UNIROUEN, UNICAEN, CNRS, M2C, 76000 Rouen, France

^b Department of Geoscience, University of Wisconsin-Madison, Madison, WI, USA

^c Hydrosiences Montpellier (HSM), Univ. Montpellier, CNRS, IRD, Montpellier, France

* Corresponding author.

6.2.1 Introduction

Characterization of the hydrodynamic properties of porous or fractured aquifers is a common challenge in many areas including: exploitation and protection of water resources, oil production, geothermal energy extraction, civil engineering, remediation engineering and storage of radioactive waste. Among many approaches (tracer tests, slug tests, geophysical investigations, etc.), this characterization can be accomplished using pumping tests, in which the values of the transmissivity and storativity parameters are derived from the analysis of the hydraulic responses to the aquifer stimulation (*Butler 2005*). However, in field conditions, many noise sources are unmanageable hydraulic contributions that can contaminate the signal induced by pumping, such as: aquifer recharge, river-aquifer interactions, evapotranspiration by plants, tidal fluctuations, or unplanned / unknown pumping at nearby water supply boreholes. For example, in unconfined aquifers evapotranspiration occurring during pumping tests may result in a spurious “signal” (see, e.g., *Cardiff et al. 2009*).

To address this difficulty, harmonic pumping tests have been suggested as an efficient approach to characterize the hydraulic properties, by making the hydraulic signal exploitable even for low signal amplitudes and noises corruptions (*Cardiff and Barrash 2015*). The oscillations in the hydraulic responses caused by a harmonic excitation with a known frequency can, in fact, be more easily extracted from the ambient noises, by applying filtering techniques (*Bakhos et al. 2014*), than the hydraulic responses generated by a constant rate pumping. Harmonic pumping tests also offer the possibility of avoiding the non-linear regimes of groundwater flow by controlling the characteristics of the periodic excitation. This excitation can be conducted through a pumping-reinjection system (*Rasmussen et al. 2003*), or even without any pumping or injection, by only using a moving mass (*Guiltinan and Becker 2015*). Harmonic pulse tests by using cyclic injection rates have also been conducted by *Sun et al. (2015)* and *Sun et al. (2016)* in a deep subsurface aquifer for leakage detection. Harmonic pumping tests have been especially investigated for their abilities for characterizing aquifers properties. The analysis and interpretation of the harmonic data for a characterization are most often realized with analytical solutions (*Renner and Messar 2006 ; Dagan and Rabinovich 2014 ; Rabinovich et al. 2015*). Among these attempts, *Black and Kipp (1981)* proposed solutions for analytical models, under the homogeneity assumption, to derive the average transmissivity and storativity parameters from harmonic test analysis. *Rasmussen et al. (2003)* applied an analytical model expressed in frequency form to interpret the piezometric data generated from harmonic pumping tests in a porous contaminated aquifer, in order to estimate the average values of the hydraulic properties. The numerical models have also been used

particularly in the frequency domain instead of the time domain in order to reduce the computing time and to take into account the spatial heterogeneity of the hydraulic properties (*Black and Kipp 1981 ; Cardiff et al. 2013b*). Modeling and imaging potential of harmonic pumping tests data were also already addressed in the literature for the reconstruction of the spatial variability of hydraulic properties (*Lavenue and de Marsily 2001 ; Cardiff et al. 2013b; Zhou et al. 2016 ; Soueid Ahmed et al. 2016*).

Some works have focused more specifically on the capacity of harmonic pumping for characterizing fractured aquifers. *Renner and Messar (2006)* applied harmonic pumping tests on a fractured sandstone aquifer to deduce the average values of transmissivity and storativity fields using a cyclical hydraulic excitation by alternating pumping, no flow and injection periods. At the same study site *Maineult et al. (2008)* used the self-potential method to monitor, remotely at the ground surface, the piezometric fluctuations caused by these harmonic excitations. More recently, *Guiltinan and Becker (2015)* similarly conducted periodic slug tests on isolated fractures without any extraction or injection, only by oscillating a slug at different depths of the water column in the well, to characterize the hydraulic connectivity of the fracture using analysis of phase shift and attenuation of the signal with an analytical model. *Schuite et al. (2017)* used tilt data recorded at the ground surface to follow the oscillatory deformations induced by harmonic pulses performed in a fractured aquifer.

Among the works previously cited and dealing with this subject, none have sought to model the impact of spatial high contrasts of hydraulic property variations on oscillatory testing. In karstic aquifers, where low-transmissivity host rocks are directly adjacent to sparsely-distributed, high-transmissivity fractures and conduits, such a scenario is clearly present and leads to flow paths particularly constrained. Wells in karstic aquifers therefore may be expected to have drastically different responses depending on whether they are located on or near a fracture or a conduit. In fact, the main flow paths in subsurface karstic field follow the conduit and fracture network and one would need to have an idea of its positioning for understanding the flow behavior (*White 2002 ; Saller et al. 2013*). Commonly, karstic fields are described by their connectivity, whether the flows between wells follows a highly conductive path or not. *Jazayeri et al. (2011)* defined three types of hydrodynamic response to pumping and pulse tests, as a function of the degree of conductivity of the flow path network in the investigated karstic aquifer. The authors assumed that a short time lag and a high amplitude hydrodynamic response (conduit type hydrodynamic response) was the consequence of a high permeability and high connectivity between the observation well and the main flow path network (karst conduits that

generate the large scale permeability of the aquifer), while ‘fracture type’ and ‘matrix type’ hydrodynamic responses were related to both a lower permeability and lower connectivity between the observation wells and the main flow path network. To be more consistent with previous works, this connectivity must be defined in term of property distribution as a ‘static’ or ‘topological’ connectivity, and in term of physical flow/transport processes as a ‘dynamical’ connectivity (*Renard and Allard 2013 ; Tyukhova and Willmann 2016*).

In this work, we propose an oscillatory signal analysis method, based on a synthetic and simplified model, for qualitatively interpreting hydraulic responses of a karstic field to a harmonic pumping test performed at different frequencies. This method consists in interpreting the ‘dynamical’ connectivity information from the responses in terms of ‘topological’ connectivity within the karstic field. In the first section of the manuscript, we present the mathematical framework used to simulate numerically, in both temporal and frequency domains, the groundwater flow responses to a sinusoidal excitation in a fractured aquifer. These numerical approaches are applied in the second section to a hypothetical karstic aquifer characterized by the presence of a simple karstic network. In the third section, we perform signal analysis by comparing point measurements of spatial amplitude decay and phase shift values with respect to the source of the oscillating signal at the pumping well to determine a conduit (conduit network), dual (fissure, conduit proximity) or matrix flow connectivity. Finally, we apply the same analysis method on real field data acquired on a karstic field located in Southern France. This signal analysis method permits an easy, fast and coherent interpretation of the preferential flow paths’ location on this site.

6.2.2 Theoretical background

In this article we use several terms such as harmonic, oscillatory, periodic, or sinusoidal signals, which all refer to the same idea of a signal defined by amplitude, a mean value and a period, repeating over time.

The harmonic pumping signal Q used in this manuscript is mathematically defined as:

$$Q(t) = -Q_A \cos(\omega t) + Q_m, \quad (6.1)$$

where t is the time (s), Q is the time-dependent flow rate signal (m^3/s), Q_A is the amplitude of the oscillatory portion of the flowrate signal (m^3/s), Q_m is the flow rate signal mean value (m^3/s), and $\omega = \frac{2\pi}{T}$ is the angular frequency where T denotes the period of the pumping signal (s). We note that this flow formulation reproduces only the extraction of water with a flow rate sinusoidally fluctuating around the mean value Q_m . This flow rate signal can also be rewritten in a complex form:

$$Q(t) = -Q_{\text{osc.}}(t) + Q_m, \quad (6.2)$$

with $Q_{\text{osc.}}(t) = \text{Re}(Q_A e^{i\omega t})$, and i representing the imaginary unit.

In order to describe the hydraulic drawdown responses of a confined karstic aquifer in 2D to a harmonic pumping signal, we represent the aquifer in a model domain Ω with a coupled discrete-continuum concept (for more details about the coupled discrete-continuum modeling, see *Teutsch 1993 ; Liedl 2003*).

We combine Darcy's law and the law of conservation of mass in a 2D matrix domain (intact rock) containing a 1D discrete conduits domain, which gives us the following partial differential equations, both defined in the Cartesian coordinates dimension:

$$\begin{cases} S_{S,mat} \frac{\partial h}{\partial t} - \nabla \cdot (K_{mat} \nabla h) = \frac{-Q(t)}{V_{el.}} \delta(x - x_s) & \text{in the 2D matrix} \\ S_{S,cond} \frac{\partial h}{\partial t} - \nabla_T \cdot (K_{cond} \nabla_T h) = \frac{-Q(t)}{V_{el.}} \delta(x - x_s) & \text{in the 1D conduits} \end{cases}, \quad (6.3)$$

where h represents the hydraulic drawdown which links the 1D and the 2D domains (m), Q is the harmonic pumping rate (m^3/s) in an elementary volume V_{el} (m^3) of the matrix or the fractures, $\delta(x - x_s)$ represents the Dirac distribution where x_s denotes the coordinates of the pumping well that can be positioned in the matrix domain or in a conduit. $S_{S,mat}$ and $S_{S,cond}$ are the specific storages in the matrix and the conduits (m^{-1}), K_{mat} and K_{cond} are the conductivities in the matrix and the conduits (m/s), and ∇_T is the tangential gradient operator for the hydraulic equation in the conduits described as discrete elements at the internal boundaries of the domain.

Both governing equations can be solved numerically by following no-drawdowns initial and boundary conditions:

$$\begin{cases} h(x, y, t) = 0 & \forall (x, y) \in \Omega \quad \text{when } t = 0 \\ h(x, y, t) = 0 & \forall t \geq 0 \quad \text{when } (x, y) \in \Omega_{boundary} \end{cases}. \quad (6.4)$$

A sufficiently large equivalent porous media buffer zone encloses the model, in order to reduce the effects of the boundaries $\Omega_{boundary}$ on the flows within the model area of interest.

The nature of the excitation signal applied at the pumping borehole creates hydraulic responses $h(t)$ composed of a sinusoidal signature $h_{osc.}(t)$ and a non-sinusoidal signature $h_{lin.}(t)$, which must be eliminated to deal only with the harmonic component.

$$h(x, y, t) = h_{osc.}(x, y, t) + h_{lin.}(x, y, t). \quad (6.5)$$

The hydraulic periodic signature $h_{osc.}$ can be expressed in a complex formulation:

$$h_{osc.}(x, y, t) = Re(H_{\omega}(x, y)e^{i\omega t}) , \quad (6.6)$$

where H_{ω} is a complex number representing the wave phasor at the given frequency.

Therefore we can compute numerically the harmonic responses due to the sinusoidal signal of the pumping rate by a reformulation of the time domain groundwater equation in a frequency domain (*Black and Kipp 1981; Cardiff et al. 2013b*):

$$\begin{cases} i\omega S_{s,mat} H_{\omega} - \nabla \cdot (K_{mat} \nabla H_{\omega}) = \frac{Q_A}{V_{el.}} \delta(x - x_s) & \text{in the 2D matrix} \\ i\omega S_{s,cond} H_{\omega} - \nabla_T \cdot (K_{cond} \nabla_T H_{\omega}) = \frac{Q_A}{V_{el.}} \delta(x - x_s) & \text{in the 1D conduits} \end{cases} , \quad (6.7)$$

with H_{ω} the complex field variable which links the 1D and the 2D domains and describes the harmonic signal responses spatially.

In the frequency domain the initial and boundary conditions become:

$$\begin{cases} H_{\omega}(x, y) = 0 & \forall (x, y) \in \Omega \text{ initially} \\ H_{\omega}(x, y) = 0 & \text{when } (x, y) \in \Omega_{boundary} \end{cases} . \quad (6.8)$$

The amplitude and phase offset values of the response to the harmonic pumping signal at a given position (x, y) are calculated from the complex variable H_{ω} value:

$$\begin{aligned} \text{Amplitude: } A(x, y) &= \sqrt{(Re H_{\omega}(x, y))^2 + (Im H_{\omega}(x, y))^2} \text{ in m} \\ \text{Phase offset: } \Phi(x, y) &= \frac{180}{\pi} atan2(-Im H_{\omega}(x, y), Re H_{\omega}(x, y)) \text{ in } ^{\circ}, \end{aligned} \quad (6.9)$$

with Re the real part of H_{ω} , Im the imaginary part of H_{ω} , and $atan2$ the two-arguments inverse tangent function in radian mode (rad).

Then the time-dependent periodic response can be reconstructed as:

$$h_{\text{osc.}}(x, y, t) = A(x, y) \cos\left(\omega t - \Phi(x, y) \frac{\pi}{180}\right). \quad (6.10)$$

Thus the frequency domain solver permits to reproduce periodic time-dependent responses while avoiding the use of a time domain solver.

6.2.3 Synthetic application

6.2.3.1 Model structure

In order to interpret drawdown responses to harmonic pumping tests in a real karstic aquifer (Terrieu, Southern France), we have studied the spatial drawdown responses from a synthetic case model of dimensions 50 m \times 50 m. This synthetic case model was built in consideration of our current understanding of the Terrieu karstic field (presented more specifically in section 6.2.4.1) and our pumping signals during investigations at the field site. A simple karstic network is represented in the model as a 1D discrete geometry in a 2D continuum homogeneous matrix (Figure 6.1).

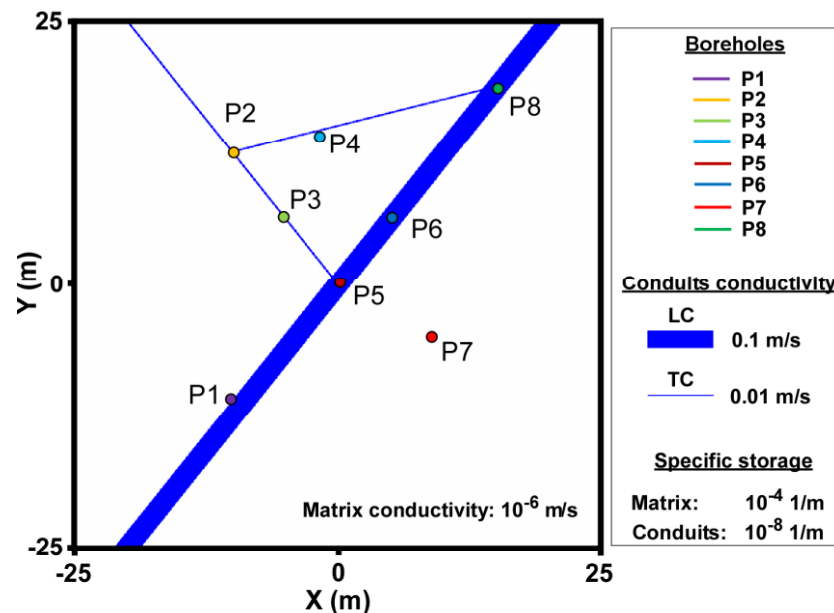


Figure 6.1: The theoretical synthetic case used to study the responses of a harmonic pumping in a karstic field. A karstic network (in blue) composed of a large conduit (LC) and two thin conduits (TC) crosses a homogeneous matrix (in white). All conduits are 1-D features in the model, but shown with conductivity-weighted thicknesses for clarity. Eight different boreholes are positioned in the model and represent pumping or measurement points.

The homogeneous matrix is associated with a 10^{-6} m/s conductivity value and a 10^{-4} m $^{-1}$ specific storage value. The karstic network is composed of a large conduit with a 0.1 m/s equivalent hydraulic conductivity and two thin conduits with a 0.01 m/s equivalent hydraulic conductivity. The whole conduit network is associated with a 10^{-8} m $^{-1}$ specific storage.

Eight different boreholes have been placed in this model at different strategic positions (Table 6.1). Three boreholes are located in the large conduit (LC: P5, P6, P8), two in the thin conduit (TC: P2, P3), two in the matrix but near to the conduit network (M-NC: P1, P4) and one in the matrix, distant from the conduit network (M: P7). We applied a 5 min period harmonic pumping signal (as defined in Equation (6.1)) in each of these boreholes alternatively, while measuring the drawdown responses in the seven other boreholes. Different values of pumping amplitudes and mean flow rates were chosen according to the positioning of the pumping borehole (in a conduit or the matrix) in order to simulate a difference of productivity at each location.

Table 6.1: Coordinates, position and pumping signal parameters for the eight boreholes. For the positioning M=Matrix, TC=Thin Conduit, LC=Large Conduit and NC=Near Conduit. The pumping signal parameters are the amplitude (Q_A) and the mean flow rate (Q_m) (see Equation (6.1)).

	P1	P2	P3	P4	P5	P6	P7	P8
X ; Y (m)	-10 ; -11	-10 ; 12.5	-5 ; 6.25	-2 ; 14	0 ; 0	5 ; 6.25	9 ; -5	15 ; 18.75
Position	M-NC (1 m from LC)	TC	TC	M-NC (50 cm from TC)	LC	LC	M	LC
Q_A / Q_m (L/s)	0.5 / 1	2 / 4	2 / 4	0.5 / 1	2 / 4	2 / 4	0.5 / 1	2 / 4

The synthetic case model is enclosed in a 1,000 m \times 1,000 m buffer zone with a global 10^{-3} m/s conductivity value and a 10^{-4} m $^{-1}$ specific storage value. The boundaries of the buffer zone are associated with a no-drawdown condition. Thus the buffer zone reduces boundaries effects on drawdowns simulated within the central area of the model.

The governing equation (Equation (6.3) and Equation (6.7)) for the simulation in the model were solved with the software COMSOL Multiphysics using a finite element method on a triangular adaptive grid (with a mesh refinement around the model discrete structures: the linear conduits and the boreholes points presented in Figure 6.1) considering the initial and boundary conditions described in Equation (6.4) and Equation (6.8).

6.2.3.2 Modeling in time domain

We firstly solved the synthetic case model in the time domain (Equation (6.3)). The top graphic in Figure 6.2 shows the time domain drawdown response in all boreholes for a harmonic pumping in P3 (in a thin conduit). Except for P7, we notice a periodic signal in the responses of each borehole. Moreover, past the first signal period, the drawdown responses can be represented as the sum of a linear drawdown (extracted by linear regression) and a purely oscillatory signal (Figure 6.2). This oscillatory signal can be parameterized, for each borehole, by a value of amplitude and a value of phase offset. In the particular case of P7, its oscillatory signal is almost null (< 1 mm), thus its drawdown response is composed almost exclusively of a linear signal.

From the observations made in Figure 6.2, the drawdown response at a given position, after the first signal period, can be mathematically approximated as the sum of a linear function and an oscillatory function:

$$h(t) \simeq \underbrace{-at - h_0}_{h_{\text{lin.}}(t)} + \underbrace{A \cos\left(\frac{2\pi}{T}t - \Phi \frac{\pi}{180}\right)}_{h_{\text{osc.}}(t)} + H_0, \quad (6.11)$$

with h (in m) the time domain drawdown response, $h_{\text{lin.}}$ (in m) the linear part of the response with its slope a (in m/min) and its intercept values h_0 (in m), $h_{\text{osc.}}$ (in m) the oscillatory part of the response with A (in m), Φ (in $^\circ$) and T (in min) its amplitude, phase offset and period values, and H_0 (in m) the initial water level (in our synthetic case, this value is considered as 0).

For the interpretation of the responses we will be more specifically interested in the oscillatory response $h_{\text{osc.}}$.

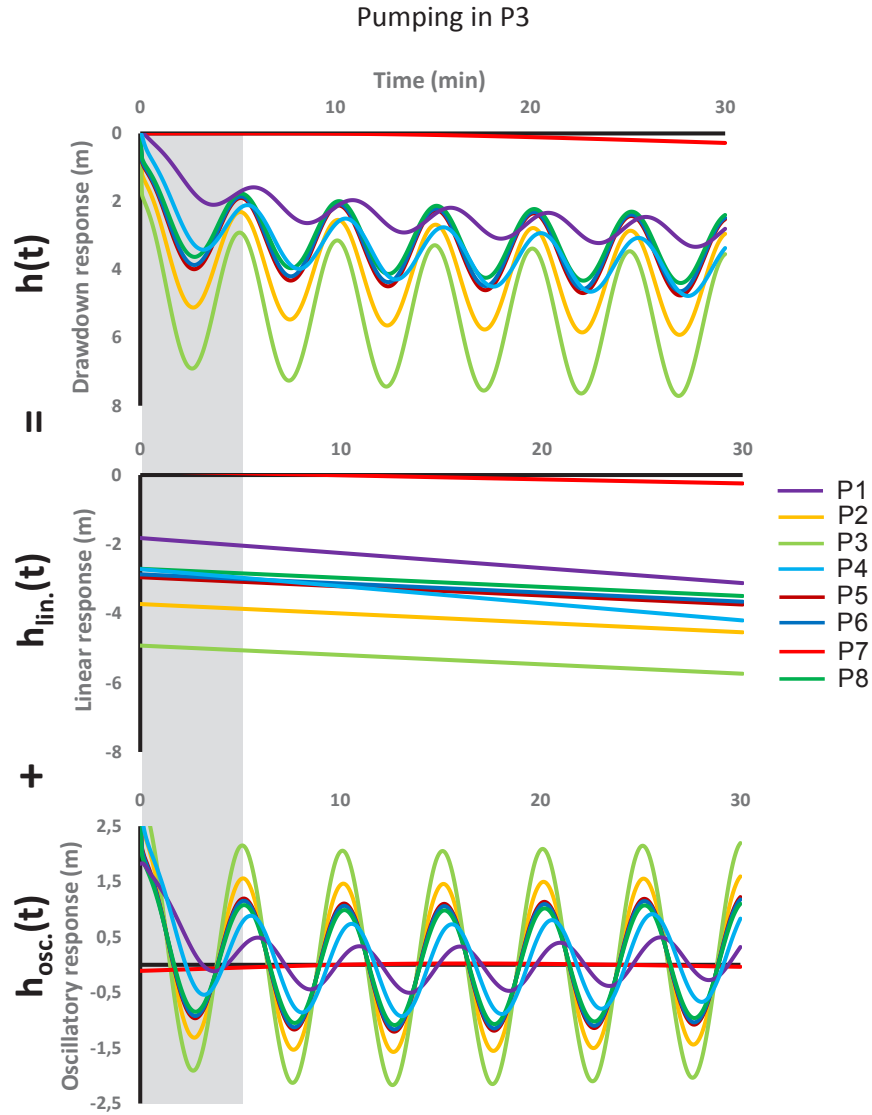


Figure 6.2: Drawdown responses h in each borehole to a harmonic pumping in P3 in a time domain simulation. If the greyed portion of the time series is not considered, these drawdown responses can be described as the sum of a linear signal $h_{lin.}$ and a purely oscillatory signal $h_{osc.}$.

6.2.3.3 Modeling in frequency domain

6.2.3.3.1 Frequency / Time domains comparison

We also solved the synthetic case in the frequency domain (Equation (6.7)). In this way, instead of spatial drawdown values, we calculate the spatial variation of amplitude and phase offset values in the response signal. We can then reconstruct for a given position the oscillatory signal $h_{osc.}$ in the drawdown response from these amplitude and phase offset values as described in Equation (6.10). By comparing the reconstructed frequency domain signals to the time domain oscillatory ones, we show that they are almost identical (see the example for

pumping in P3 in Figure 6.3 and Appendix 5), except for the first signal period (see Figure 6.2), in which the time domain signals have not reached a stationary behavior (because of the pumping signal oscillating around a non-null mean rate).

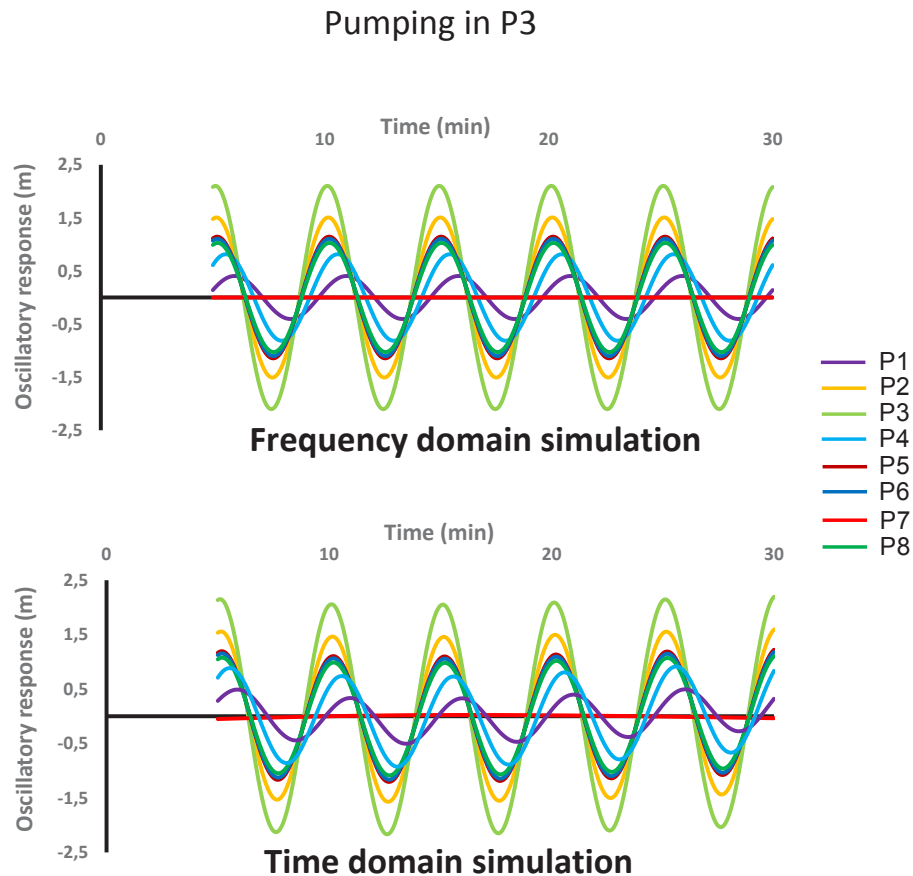


Figure 6.3: Oscillatory signals responses in each borehole for a harmonic pumping in P3, for a frequency domain simulation and a time domain simulation (avoiding the first signal period). One sees that these signals are almost the same for the two simulations.

We can simulate the same oscillatory responses signals to a harmonic pumping test either with a time domain solver, or with a frequency domain solver (if we put aside the first period where the time domain responses have not reach their stationary behavior, and thus are not reproducible in the frequency domain). However, in our case, the frequency-domain solver is more useful, as it provides directly the spatial distribution of the oscillatory responses amplitude and phase offset values that interest us for the interpretation part. There is no need of signal decomposition as in the time domain simulations. The frequency domain solver is also faster than the time domain one as we don't have to simulate different time steps (for the

simulations on the presented synthetic case the frequency domain solver was approximately 120 times faster than the time domain one).

Therefore, for the interpretation of the responses oscillatory signals $h_{osc.}$, we have used the frequency domain simulations, and more specifically the responses' amplitude and phase offset values at the position of the different boreholes.

6.2.3.3.2 Analysis of the harmonic hydraulic responses

In order to interpret the spatial responses to harmonic pumping at a point in the karstic synthetic case, we have specifically studied the amplitude and phase offset values in these oscillatory signal responses. We have observed that, depending on the pumping borehole location, the responses could be highly variable. From Figure 6.4 we first state that the oscillatory signal amplitude response is not proportional to the distance between the measurement points and the pumping point, as it would be expected in a purely homogeneous aquifer. In our synthetic case, the spatial responses were controlled by the degree of connectivity between the pumping point, the flow path induced by the pumping, and the measurement point. In particular cases, if the pumping borehole is positioned in the matrix, no oscillatory responses can be seen in the other boreholes (see case P7 in Figure 6.4).

Four representative examples of response are presented in Figure 6.4, with the complete table of responses amplitudes and phases offsets for these cases presented in Appendix 5. From the analysis of the 8 different pumping cases in the synthetic model, if we pump in or near to a conduit, three degree of connectivity can be interpreted by comparing the spatial oscillatory responses relatively to the pumping signal:

- *Amplitude responses associated with a low phase shift* (see P2, P3, P5, P8 responses for a pumping in P3 or P6 in Figure 6.4): the pumping point is in a conduit and the measurement point has a *conduit connection* to it through a flow path in the karstic network,

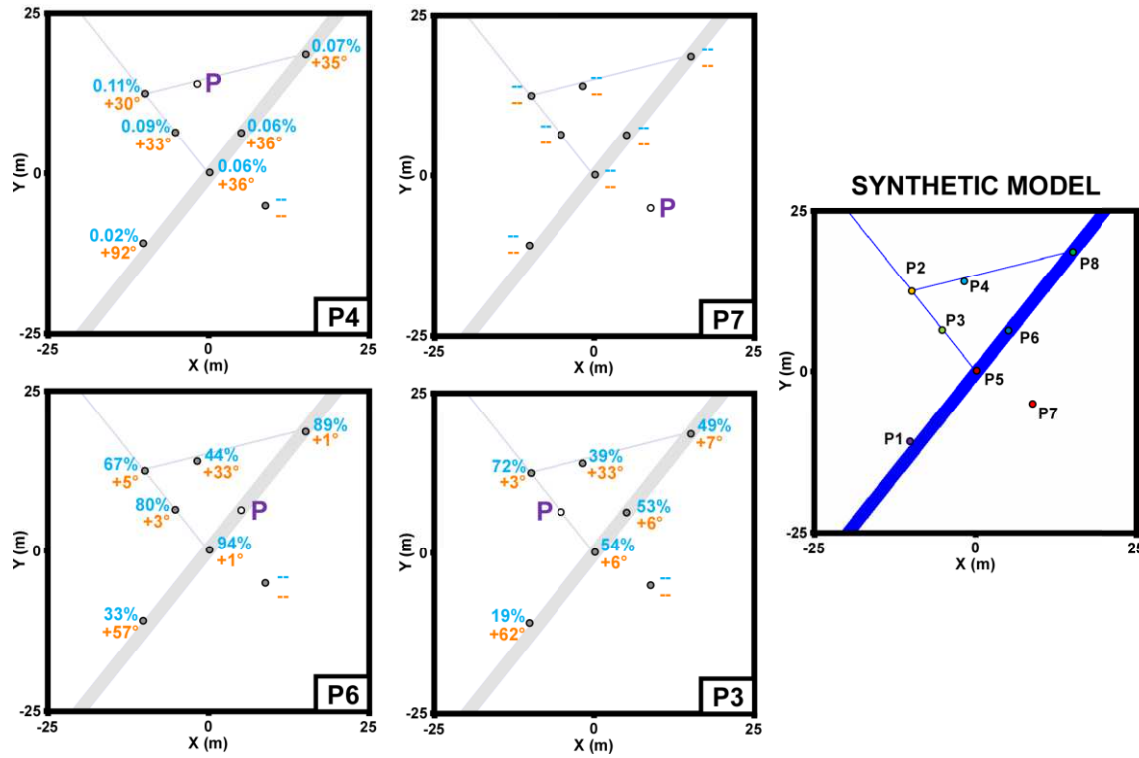


Figure 6.4: Relative amplitude (% , in blue) and relative phase offset (° , in orange) values in the oscillatory responses in each borehole for different harmonic pumping locations (P4, P7, P6, P3). A dash represents an absence of oscillatory response (< 1 mm). The pumping location is indicated by 'P' and its drawdown oscillatory signal is considered as a 100% amplitude signal with a 0° phase offset.

- *Amplitude responses associated with a high phase shift*: the measurement point and the pumping point have a *dual connection*. A prevailing part of the flow path follows the karstic network and another part is in the matrix. The response phase offset value is proportional to the matrix flow path importance. Then, either the pumping point is in the matrix near to a karstic network (see P2, P3, P5, P8 responses for a pumping in P4 in Figure 6.4), or the measurement point is in the matrix near to the network (see P1 and P4 responses for a pumping in P3 or P6), or both with a higher phase offset (see P1 response for a pumping in P4),
- *Negligible amplitude responses (almost no oscillatory signal)* (see P7 responses in Figure 6.4): the prevailing part of the flow path between the measurement point and the pumping point is located in the matrix, it generates a *matrix connection* response.

Further information can be interpreted from the relative amplitude value of the measured signals for the responses with the same phase offset. If the flow path follows a thin conduit in the network, the amplitude of the signal will decrease along the flow path away from the

pumping point (see the responses in the thin conduit for the pumping case P6). The rate of decay in the responses amplitude becomes less important when the signal reaches more conductive conduits of the network (see the amplitude decay of the responses in the large conduit for the pumping case P3). The decay rate is inversely proportional to the conductivity of the conduits.

The choice of the harmonic pumping period is important for a good interpretation of the oscillatory responses. We show, in Figure 6.5, the spatial responses' signal differences for harmonic pumping in P6 and P3 with a same signal amplitude ($Q_A = 2 \text{ L/s}$) but two different periods (1min – 5min).

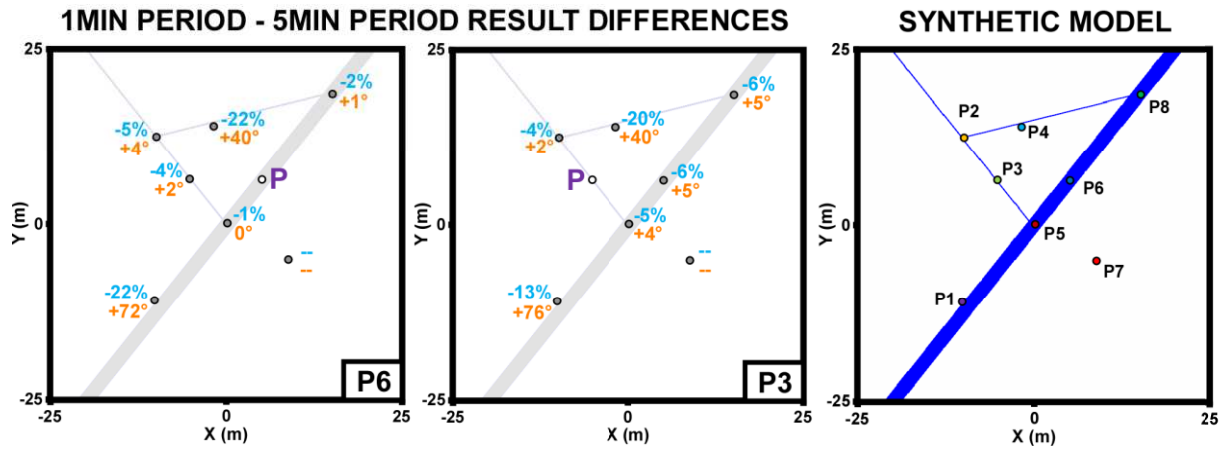


Figure 6.5: Differences in relative amplitude (in blue, in %) and relative phase offset (in orange, in °) values in the oscillatory responses by decreasing from a 5 min period signal to a 1 min period signal for two different harmonic pumping locations (P6, P3). A dash represents an absence of oscillatory response (< 1 mm). The pumping location is indicated by 'P'. The main signal differences appear for the boreholes located in the matrix, near a conduit (P1, P4) (dual connection).

The main effect of changing the period duration affects specifically the measurement boreholes that have a dual connection to the pumping borehole, while the other boreholes will have little consequent response differences. Decreasing the period from 5 min to 1 min will produce a relative decrease of the amplitude and a relative phase offset increase in the response signal of the boreholes with a dual conduit/matrix connection. Globally, this would tend to bring a 'dual connectivity' response closer to a 'matrix connectivity' response. This can be clearly seen in the oscillatory responses reconstructed from Equation (6.10) (Figure 6.6). For a 1 min pumping period, there is no significant signal changes for the boreholes in the conduits, but the oscillatory signals for the boreholes in the matrix tend to disappear compared to their responses

to a 5 min pumping period signal (especially for P1 that becomes closer to the P7 ‘matrix connectivity’ signal).

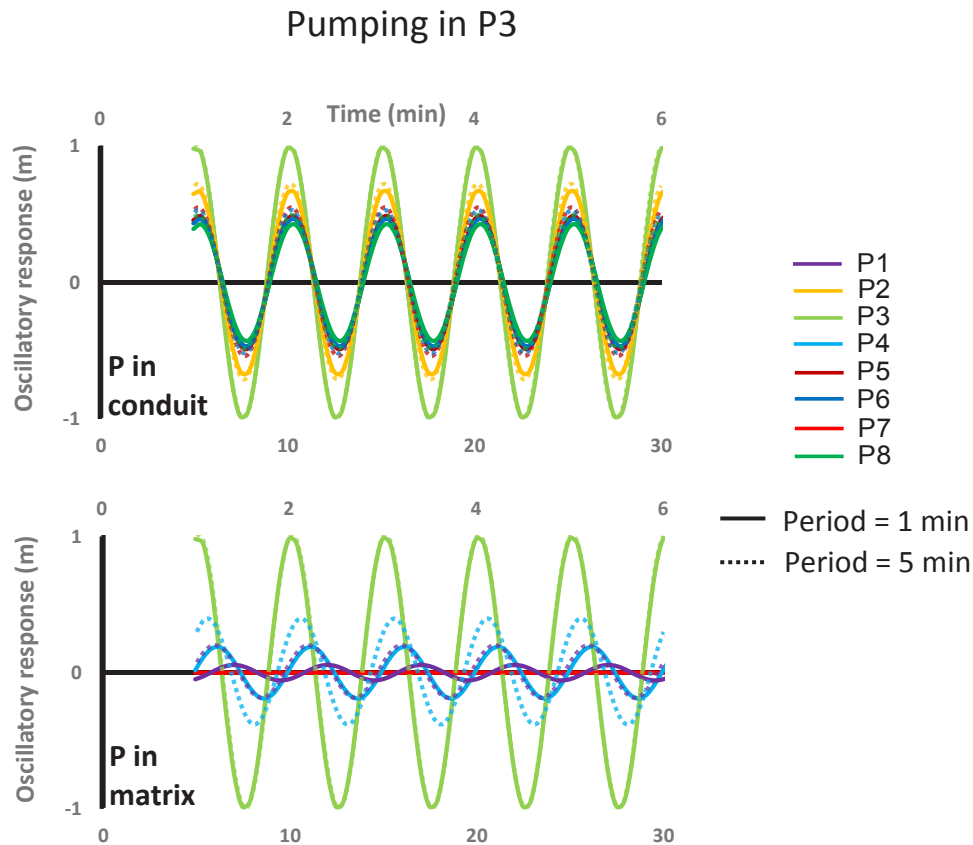


Figure 6.6: Comparison of the oscillatory relative responses for a harmonic pumping in P3 for a 1 min period signal during 6 min (full line) and a 5 min period signal during 30 min (dotted line). The measurement boreholes have been separated regarding their location: in a conduit (P2, P5, P6, P8) or in the matrix (P1, P4, P7). The main signal differences appear for the boreholes located in the matrix, near to a conduit (P1, P4) (dual connection).

From the previous observation, we have studied the evolution of the relative amplitude and phase offset responses more specifically in the points with a dual connection to the network (P1, P4) for increasing periods (Table 6.2). It appears that the most important factor impacting the responses in these points is their distances to the network. Independently from the importance of the conduit in which the pumping is performed, the more a measurement point is distant from the network the more its phase offset response will evolve with a period change (see in Table 6.1 P1, distant from 1 m, compared to P4, distant from 50 cm). On the contrary, the relative amplitude of the response seems to be related to the productivity of the pumping location, but is therefore less interesting for delineating the position of the measurement point.

Globally, for a good characterization of the conduits positioning, it is important to choose at least two different periods in order to compare the evolution of the phase offset in the responses. These periods should be sufficiently high to avoid the risks related to a too low amplitude response (unreadable response) or a phase offset value exceeding one cycle (see Table 6.2) that may lead to incorrect interpretations.

Table 6.2: Table of the relative amplitude and phase offset values in the oscillatory responses of P1 (1 m away from the network) and P4 (50 cm away from the network) to harmonic pumping in P3 and P6 and for increasing signal periods. In this table Amp.=Amplitude, P.O.=Phase Offset, TC=Thin Conduit, LC=Large Conduit, NTC=Near Thin Conduit and NLC=Near Large Conduit. Values in parentheses signify phase offsets greater than one cycle ($>360^\circ$).

Pump. point	Meas. point		Period						
			10 s	30 s	1 min	2 min	5 min	10 min	30 min
P3 (TC)	P1 (NLC)	Amp.	2%	3%	6%	11%	19%	26%	37%
		P.O.	(-14°)	(+128°)	+138°	+96°	+62°	+45°	+27°
	P4 (NTC)	Amp.	4%	11%	19%	28%	39%	46%	55%
		P.O.	(+79°)	+107°	+73°	+52°	+33°	+24°	+16°
P6 (LC)	P1 (NLC)	Amp.	4%	6%	11%	19%	33%	44%	57%
		P.O.	(-20°)	(+124°)	+129°	+90°	+57°	+40°	+23°
	P4 (NTC)	Amp.	5%	13%	22%	32%	44%	53%	63%
		P.O.	(+51°)	+106°	+73°	+51°	+33°	+24°	+15°

To summarize, when pumping in the network, studying the amplitude of the responses provides some information about the conductivity of the conduits along the flow path, while studying the phase offset of the responses permits to characterize the degree of connectivity between the measurement points. A low period pumping (high frequency) tends to highlight more specifically the boreholes directly connected to the pumping point by conduits (network flow propagation). A high period pumping (low frequency) tends to generate responses in boreholes with a conduit or a dual connectivity to the pumping point (part of network and part of diffusive flows propagation).

In order to locate only the boreholes directly connected through the conduit network, a high frequency pumping is necessary. But adding in combination responses from a lower

frequency permits the identification of boreholes close to this network, which is useful information for imaging the karstic network arrangement.

6.2.3.3.3 Spatial analysis on simulation maps

The previous interpretation of the oscillatory signal responses to pumping signal with a 5 min period can be generalized in maps of spatial distribution of the amplitude (Figure 6.7) and phase offset (Figure 6.8) responses.

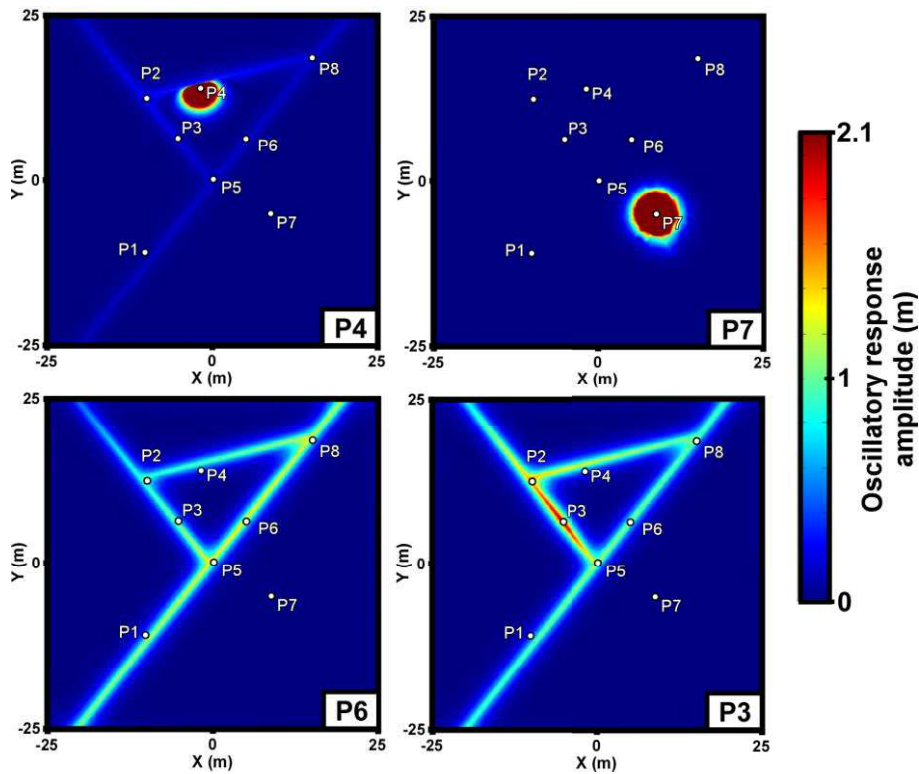


Figure 6.7: Maps of distribution of the amplitude value in the responses to a harmonic pumping signal with a 5 min period at different locations: in the matrix near a conduit (P4), in the matrix (P7), in a large conduit (P6), in a thin conduit (P3).

The case of a pumping in P7 (in the matrix) in Figure 6.7 shows a typical response map that would be expected from a homogeneous aquifer with an amplitude response decreasing with the distance to the pumping point forming a circle. In the case of a pumping in P4 (in the matrix), this response circle reaches the karstic network and the signal can propagate in the conduits with a subdued amplitude. If the pumping point is directly located in a conduit of the network (P6 and P3), the oscillatory signal propagates uniquely through the flows of the

conduits, highlighting the karstic network. A linear decrease of the signal amplitude is visible in the thin conduits along the flow path, but in the most conductive conduits, the signal easily propagates at a site scale without attenuation.

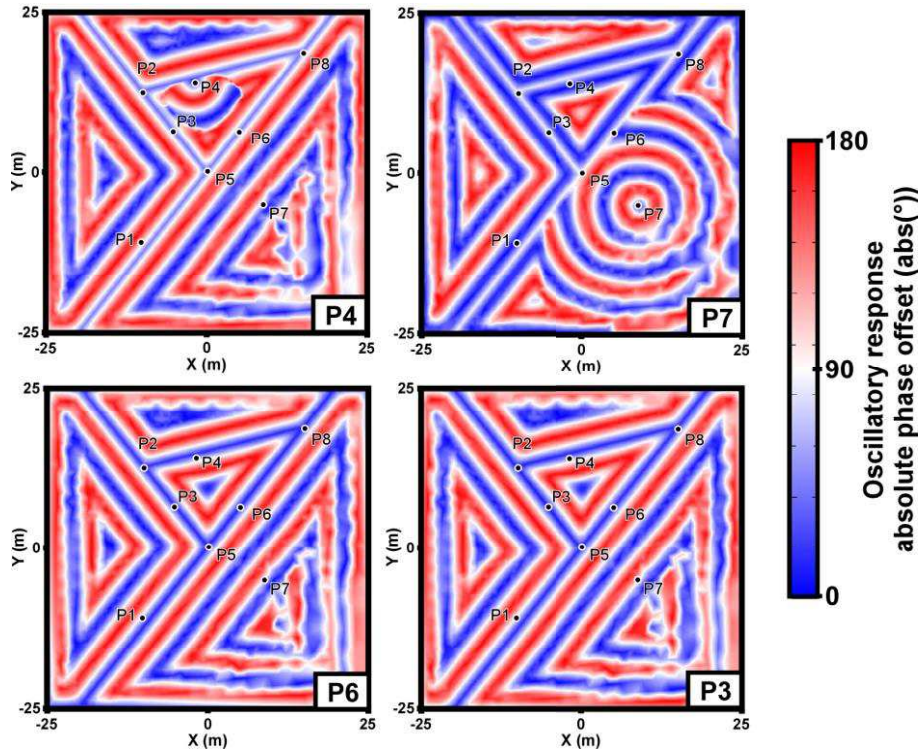


Figure 6.8: Maps of distribution of the phase offset value in the responses to a harmonic pumping signal with a 5 min period at different locations: in the matrix near a conduit (P4), in the matrix (P7), in a large conduit (P6), in a thin conduit (P3).

Concerning the spatial phase offset distribution in Figure 6.8, the case of a pumping in P7 in the matrix highlights what would be expected from a homogeneous aquifer. The phase offset value varies in function of the distance to the pumping point, forming a circle in the figure. As this circle reaches the network, its homogeneous behavior gets stopped, but at this position the signal has already lost its whole amplitude (see P7 in Figure 6.7). The same ‘homogeneous behavior’ can also be seen locally in the case of a pumping in P4 (in the matrix), but the phase offset variation gets quickly ‘controlled’ by the conduit’s disposition. The phase offset remains then roughly constant along the network geometry, with a value dependent on the pumping point’s distance to the network. For a direct pumping in the network (P6, P3), independently of the conduits conductivities (as long as these conductivities are significantly higher than the matrix), there is no significant phase shifting along the conduits of the network.

For all cases, once the signal has reached the karstic network, its phase shifting value will stay constant along the conduits, but it will increase rapidly in flow paths orthogonal to the conduit (conduit to matrix flows).

The effect of the signal period, already described in the previous part, can be generalized by using these same maps (Figure 6.9).

Figure 6.9 shows that a shorter period (high frequency) highlights more specifically the conduit network: the signal amplitude quickly decreases and the phase offset quickly increases when the signal enters the matrix. Thus, the propagation of the signal through diffusive flows in the matrix is almost absent. On the contrary a longer period signal can propagate at longer distances along matrix diffusive flows and with less phase shifting, which permits the boreholes in the matrix to respond. By comparing the cases of P6 in the figure, for a 5 min period the diffusive behavior of this signal seems to ‘blur’ the karstic network. In fact, the responses associated to boreholes located near the conduits become undistinguishable from the responses of boreholes located directly in the conduits, and thus a precise localization of the conduit will become more difficult. Concretely this behavior is seen for P1 or P4: their responses are close to the ones of in the conduits network for a 5 min period, but their responses for a 1 min period clearly show that they are located in the matrix.

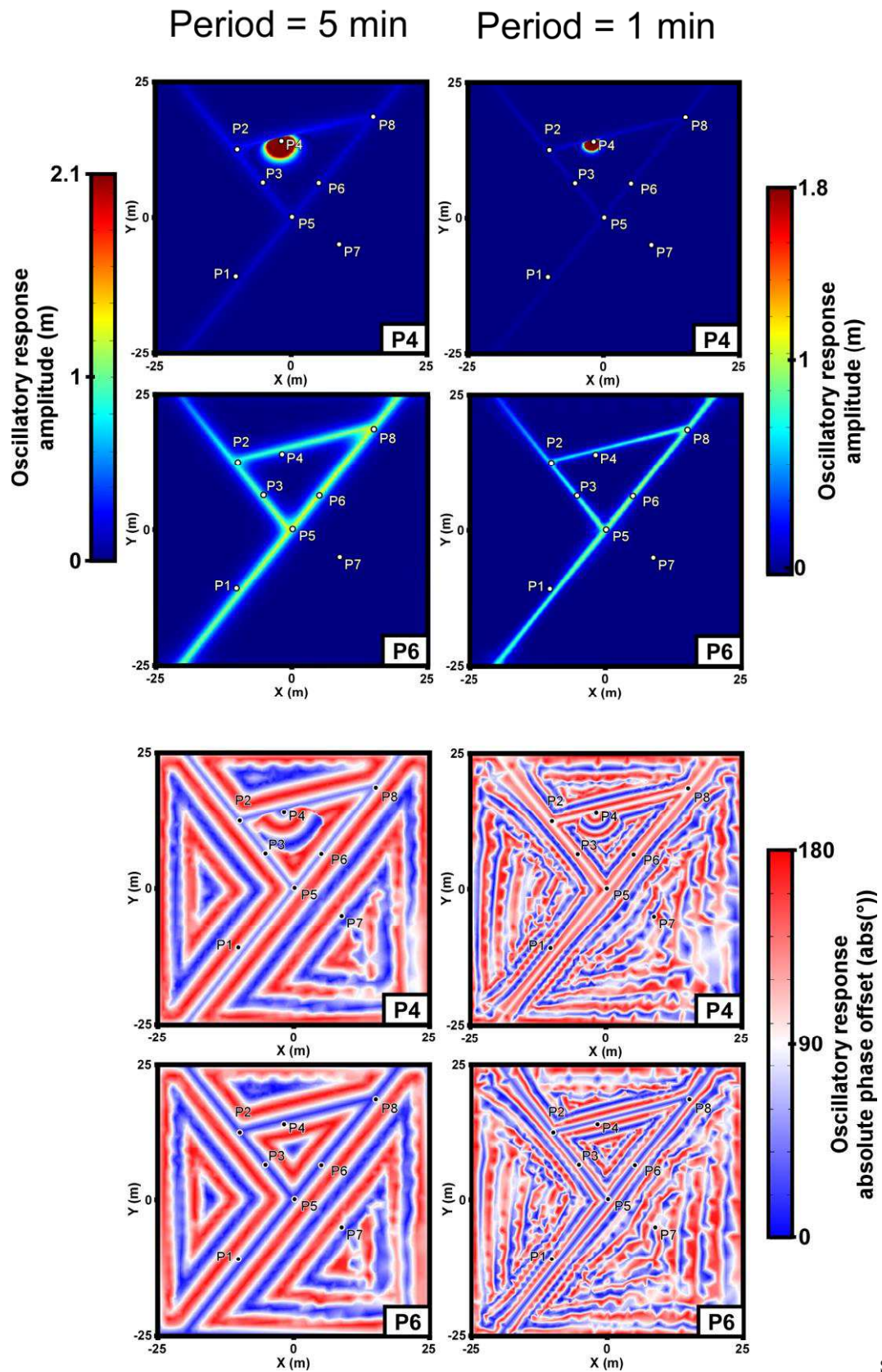


Figure 6.9: Comparative maps of distribution of the amplitude and absolute phase offset values in the responses to a harmonic pumping at two different locations (in the matrix near a conduit (P4), in a large conduit (P6)) for a 5 min period (left) and 1 min period (right) signal.

6.2.4 Example of harmonic pumping investigation

6.2.4.1 Site presentation

We now apply the lessons learned from our analysis of synthetic data trends to a true field case. We have performed an oscillatory pumping test on the well-studied Terrieu karstic field (approximately 2,500 m²) near to Montpellier, in Southern France. The oscillations in the pumping rate were controlled by a programmable electrical device based on a dimmer, linked to the pump. This site has been studied in two recent thesis (*Jazayeri Noushabadi 2009* ; *Dausse 2015*) and several articles (*Jourde et al. 2002*; *Jazayeri Noushabadi et al. 2011*; *Wang et al. 2016* ; *Wang et al. 2017*). The property values used in the synthetic model are inspired from conduits, fractures, matrix and regional property estimations issued from investigations presented in *Dausse (2015)*. The Terrieu karstic field is part of the MEDYCYSS observatory (*Jourde et al. 2011*), and an experimental site of the French National Karst Observatory (SNO Karst - www.sokarst.org).

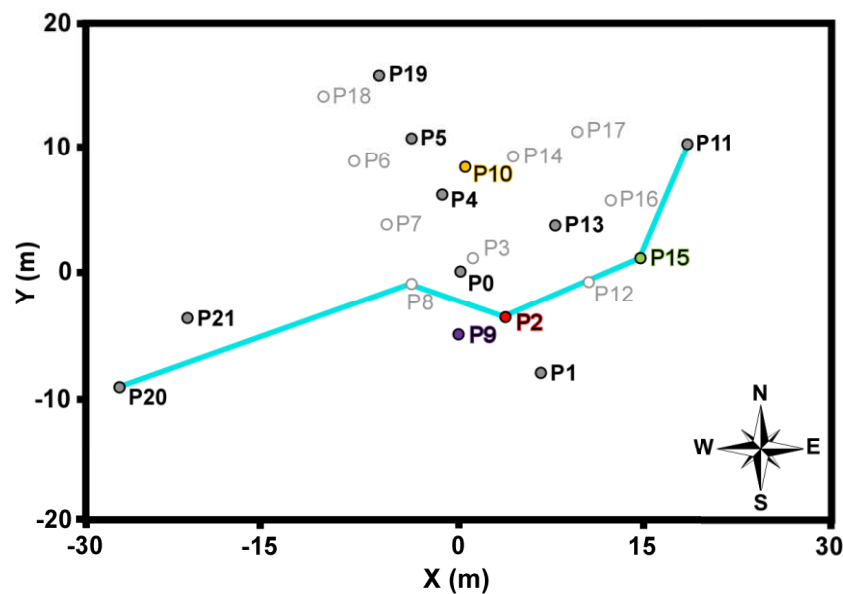


Figure 6.10: Boreholes locations on the Terrieu site. The colors for P2, P9, P10 and P15 refer to the colors used to designate these boreholes in Figure 6.11. The blue line indicates a conduit connectivity assessed from previous investigations (*Dausse 2015*; *Wang et al. 2016*). The boreholes in light grey were not measured during the harmonic pumping test.

This site is investigated through 22 boreholes (Figure 6.10) and lies on a confined aquifer. Observed conduits through downhole videos (*Jazayeri Noushabadi et al. 2011*), located between 35 m and 45 m under the surface, have been generated at a sloped and fractured

interface between marly Cretaceous and massive Jurassic limestones. Both units have very low permeability, which permits to consider the karstic aquifer to be confined. At the time of the field investigation karst features located at the sloped interface were fully saturated. These conduits have an aperture that can reach 20 to 50 cm. Previous field investigations (packer tests, temperature and electrical conductivity logging) have permitted to highlight a preferential flow path (see blue line in Figure 6.10) within this karstified interface crossing several boreholes (*Jazayeri Noushabadi 2009; Dausse 2015*). This preferential flow path could be the consequence of a network of conduits directly connecting these points.

A pumping investigation was conducted with an electronic automata device connected to the pump and generating oscillations in the pumping rate. A harmonic pumping test was led for 30 min on the borehole P15 with a period of 5 min with flow rates that varied between 3.2 m³/h and 7.4 m³/h. In other words, the signal from this pumping can be represented as a constant-rate pumping test of magnitude 5.3 m³/h (Q_m) convolved with an oscillatory (net zero) pumping signal with period 5 minutes and amplitude 2.1 m³/h (Q_A). The signal drawdown responses were measured in 12 other boreholes, additionally to a measurement in the pumping borehole itself.

6.2.4.2 Example of typical responses

The upper graph in Figure 6.11 shows the drawdown measurement in the pumping borehole (P15), and example of measured responses in three other boreholes (P2, P9 and P10). These three boreholes are approximately at the same distance from the pumping point.

This graph shows that the field measured responses have the same behavior as the theoretical ones, with additional noise. The drawdown responses, if we pass over the first pumping period, can be approximately decomposed as an addition of a purely linear signal and a purely oscillatory signal of the form of Equation (6.10) (shown in the lower graph of Figure 6.11). The amplitude and phase offset analysis of the oscillatory signal of the three chosen measured responses examples also show that we have the same type of responses in this field case as seen in the theoretical case: responses with measurable amplitude and a low phase offset (P2), responses with similar amplitude but a high phase offset (P9), and responses that contain no measurable oscillatory component (P10). Therefore, it seems acceptable to test the same interpretation that we made on the theoretical case for the field data.

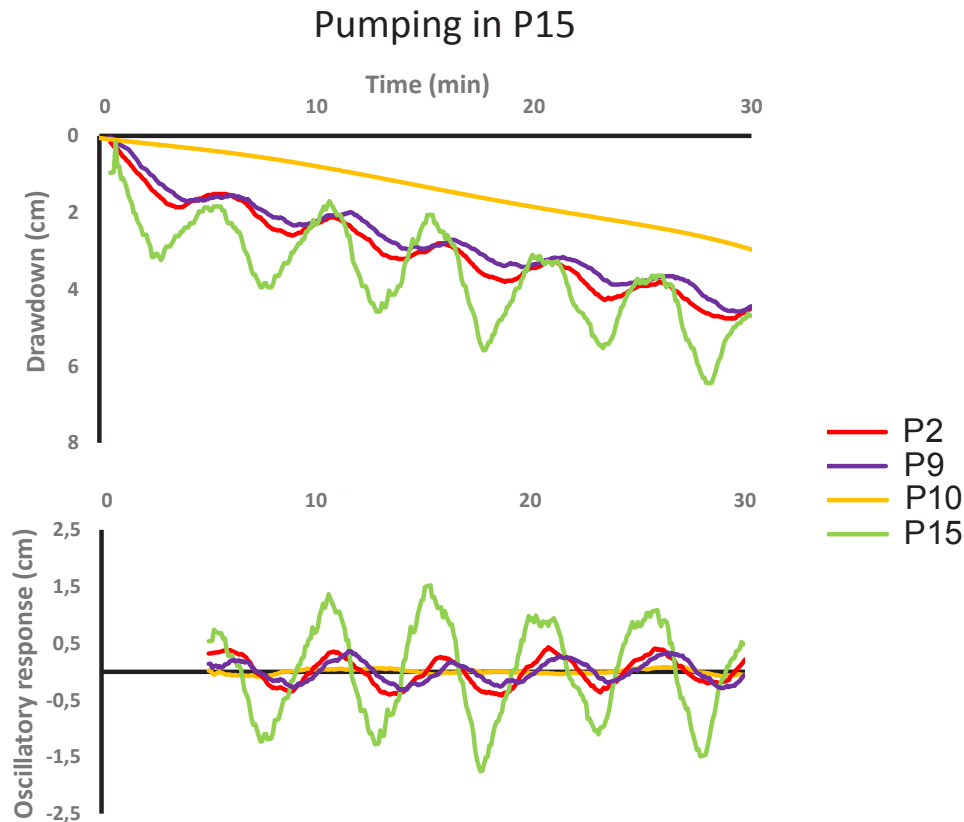


Figure 6.11: Example of different type of responses registered during the 5 min period harmonic pumping test in P15 on the Terrieu site. The top graph shows the complete responses and the bottom graph shows the purely oscillatory responses after having subtracted the linear signal.

6.2.4.3 Interpretation of the responses

We decomposed the entire set of measured drawdowns (following Equation (6.11)) in order to keep only their oscillatory signal (the oscillatory responses in each borehole are presented in Figure 6.12). We have then fit these oscillatory signals to function of the form of Equation (6.10) with amplitude and phase offset as variable parameters.

As for the interpretation of the amplitude and phase offset values from the frequency domain modeling of the theoretical case in Part 6.2.3.3.2, we have produced in Figure 6.13 a map of the measured amplitude and phase offset values from the oscillatory responses to pumping at a point within the karstic network.

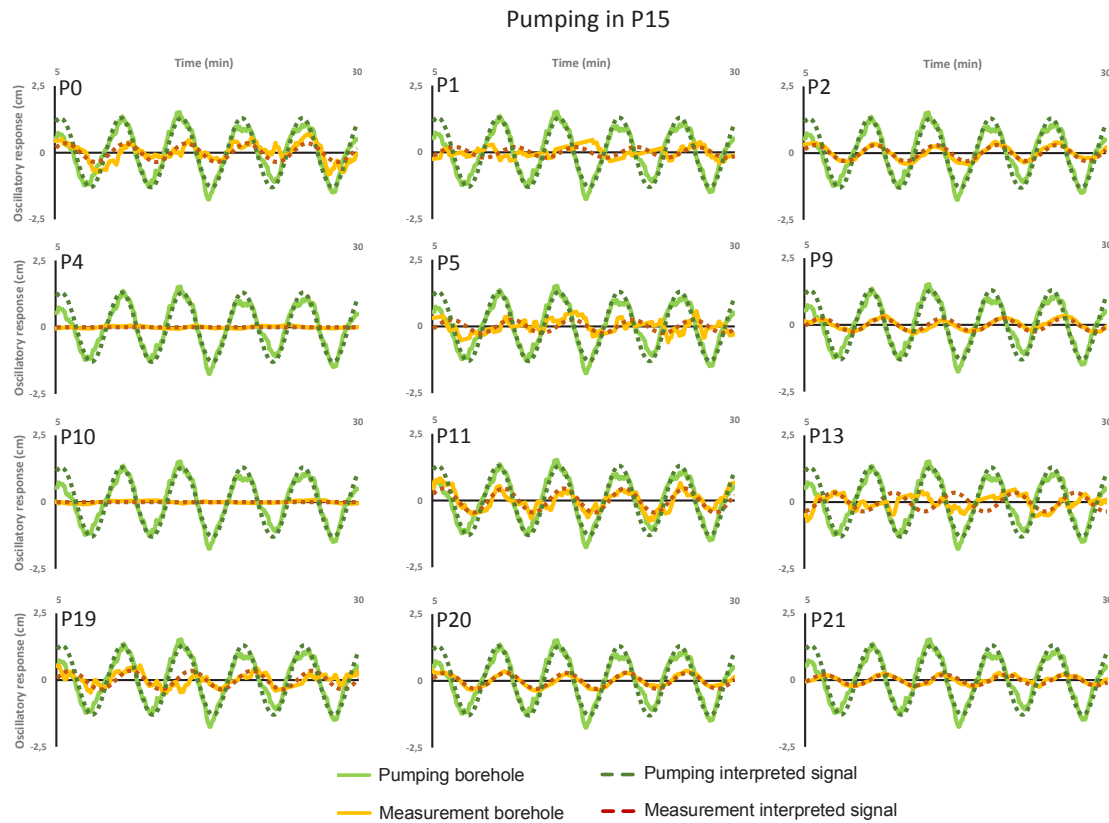


Figure 6.12: Registered oscillatory responses for each measurement borehole compared to the $T = 5$ min period pumping borehole signal (full lines) and the interpreted signals for an equation form of Equation (6.10) with variables amplitude and phase offset values (dotted lines).

The relative amplitude values of the responses vary between 0% (no oscillatory response) and 35%, and the relative phase offset values vary between $+30^\circ$ and $+140^\circ$ (relatively to the pumping signal). From downhole observations and productivity estimations on the pumping well P15 (Jazayeri Noushabadi et al. 2011), we know that this well is located in a conduit of the network.

By analyzing the phase offset value, we note that the lower phase offset values are around $+30^\circ/+50^\circ$, and that several points have this same phase shift. This would indicate the presence of a network of conduits directly connecting these boreholes ('conduit' connectivity: P0, P2, P11, P19, P20) to the pumping point P15 (a possible network of conduits is proposed, within the zone delineated by the violet dotted boundaries, in Figure 6.13). Then, the other boreholes ('dual' connectivity: P1, P5, P9, P21; 'matrix' connectivity': P4, P10) would have coherent responses toward this conduit network, similar to those seen in the study of the synthetic case ('dual' or 'matrix' connectivity).

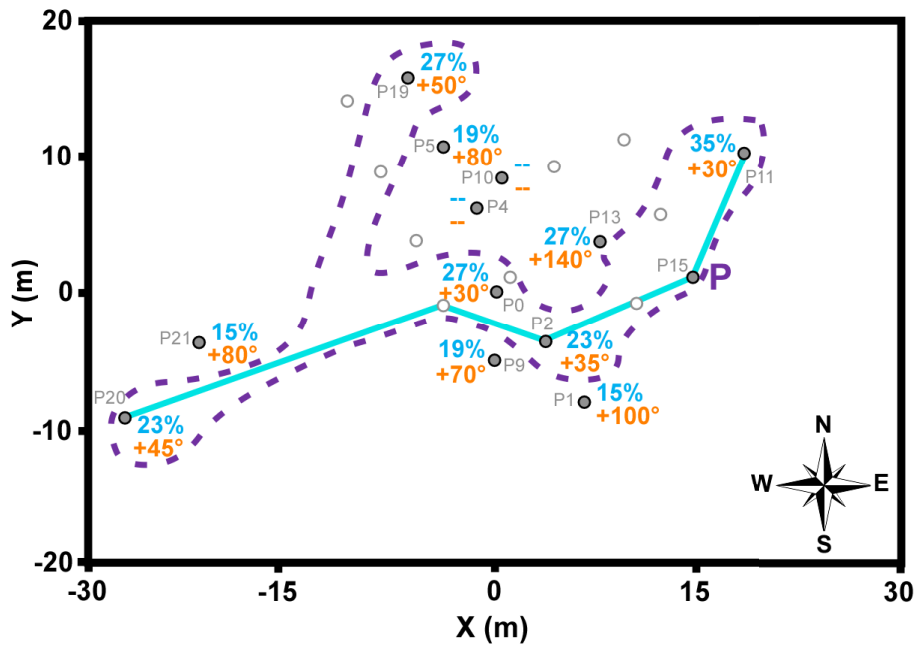


Figure 6.13: Example of a possible conduits network (inside the zone delineated by violet dotted boundaries) interpreted from the boreholes connectivity by applying the same analysis than in the synthetic case. The captions represent the relative amplitude (in blue, in %) and relative phase offset (in orange, in °) values in the oscillatory responses in each measured borehole. A dash represents an absence of oscillatory response (< 1 mm). The pumping location is indicated by 'P'. The blue line indicates a conduit connectivity known from previous investigations (Jazayeri Noushabadi 2009 ; Dausse 2015).

The more a borehole is distant to this possible conduit network, the more its response signal relatively to the pumping signal has a low amplitude and a high phase shift (for the more distant boreholes P4 and P10, the oscillatory signal disappear which indicate a 'matrix' connectivity). Only P13 would present an incoherent response (with both a high amplitude response and a high phase shift), but its true signal was too noisy to permit a good amplitude and phase offset parameters fitting (see P13 in Figure 6.12). This incoherent result can be attributed to a bad signal measurement. Except this point, the other 11 measured responses on the field are coherent with the behaviors interpreted in the theoretical study case.

The conduit connectivity highlighted from the method presented in this article is coherent with the direct connectivity already highlighted from previous investigations (blue line in Figure 6.13), which tend to strengthen the validity of this interpretation. Furthermore, P19, which was not found as connected in the previous investigation, appears to be directly connected in this work. However, this connectivity information would require a supplementary validation in the next field campaigns.

We have finally made quantitative estimation on the property values in the field, relatively to each response, in order to compare these estimations to the previously found degree of connectivity. These estimations were made by finding, for each response, an equivalent homogeneous model (K_{eq} and $S_{S,eq}$) reproducing the response amplitude and phase offset. The optimization of K_{eq} and $S_{S,eq}$ in each case was performed through a classical least square minimization criterion inversion. It appears that P0, P2, P11, P20 (interpret as connected through conduits), and P13 require high K_{eq} values (10^{-2} to 10 m/s) and low $S_{S,eq}$ (10^{-8} to 10^{-7} m⁻¹). P5, P9, P21 (interpret as dual connectivity responses), and P19 (interpreted as directly connected) require high K_{eq} values (10^{-3} to 10^{-2} m/s) and high $S_{S,eq}$ (10^{-4} to 10^{-3} m⁻¹). Finally, P1 (interpreted as dual connectivity response), and P4 and P10 (interpreted as a prevailing matrix connection) require low K_{eq} values (10^{-6} m/s) and high $S_{S,eq}$ (10^{-4} m⁻¹). Globally three categories of estimated equivalent properties can also be highlighted and would correspond to the previously described degree of connectivity: conduit connectivity (high K_{eq} , low $S_{S,eq}$), dual connectivity (high K_{eq} , high $S_{S,eq}$), and matrix connectivity (low K_{eq} , high $S_{S,eq}$). In this case only P19 appears to be rather connected in a ‘dual’ manner and P1 rather as a matrix point, which nuance the interpretation of some responses at the limit of the defined categories.

From these results, we believe that the qualitative method of interpretation of harmonic signal responses developed in this work can facilitate interpreting the degree of connectivity of karstic field responses to an oscillatory signal, especially in order to characterize and localize the conduits network through the boreholes connectivity. This interpretation can be led manually as shown in this field example. However the next step would be to use this interpretation in an inverse modeling approach (associated to a frequency domain modeling and by considering the amplitude and phase offset values of the observed and simulated responses). This approach becomes, in fact, very useful when the amount and complexity of responses to interpret is important. The Cellular Automata-based Deterministic Inversion developed in *Fischer et al. (2017b)* would provide in this case interesting results, as it would permit to generate both adequate conduits network and property distributions (conductivity and specific storage) for the reproduction of the responses, and thus go further in the quantitative analysis of the harmonic data in karstic fields.

6.2.5 Discussion and conclusion

We have studied the responses to harmonic pumping in a theoretical synthetic study, by applying a time domain solver and a frequency domain solver. We have firstly demonstrated that the harmonic result from the frequency domain simulation was very similar to the time domain oscillatory part, and thus, as the frequency domain solver is much faster, it is more useful for the simulation of periodic responses.

From an analysis of the amplitude and the phase offset of the response relative to the pumping signal and its positioning in the model, we have proposed a global method for qualitatively interpreting a degree of hydraulic connectivity between each borehole. The amplitude and phase offset values permit to distinguish either a conduit connectivity between boreholes (flow path in the conduit network), or a dual connectivity (flow path partly in the matrix for boreholes near to a conduit but not directly in it), or a matrix connectivity (majority of flow associated with flow in the matrix). By modifying the period of the pumping signal, we can dissociate more precisely the conduit connectivity and the connectivity and obtain some information about the distance of a measurement point relatively to the karstic network. A high frequency signal will more specifically highlight the conduit flows, while a low frequency signal will give more importance to the matrix diffusive flows.

In previous works *Renner and Messar (2006)* and *Guiltinan and Becker (2015)* used analytical solutions to show that, in a fractured aquifer, increasing the pumping period decreased the estimates of effective hydraulic diffusivity, due to the increase of the values of estimated storativity (while the transmissivity estimations remained almost unchanged), which was associated to a mobilization of the surrounding fracture void spaces. In our karstic model this observation would be related to the behavior of a well in the matrix near to a conduit, where the signal can better propagate within the high storage that represents the matrix when the pumping period is increased. As it has been noticed by *Rabinovich et al. (2015)* with the use of a heterogeneity model for a porous aquifer, flows (by extension the harmonic signal) will preferentially propagate in the most conductive media, especially at lower pumping period, which is also what we observe in our karstic model between conduits and matrix flows, when pumping at different periods. The results we present in this article between amplitude ratio and phase shifting and the degree of connectivity of boreholes show equivalent relations (in a more accentuated way) than those presented by *Guiltinan and Becker (2015)* in a fractured aquifer. The harmonic signal will arrive with an important attenuation and phase shift between areas

badly connected and, on the contrary, almost unchanged when fractures (or in our case a conduit) connect two boreholes.

In order to test the validity of our signal analysis method, we have confronted the same interpretation for a set of responses from a true karstic field to a harmonic pumping. The interpretation permitted to delineate the main flow paths easily and quickly by satisfying all measurement. This method could be generalized for a manual interpretation of a set of field responses. The benefits brought here by the harmonic tests compared to the other connectivity investigations done on the same site (packers, temperature and electrical conductivity) and to constant-rate pumping can be found in the easier extraction of the signal in the responses (with filtering techniques), even when responses are noised, and from the possibility to simulate the responses with a modeling in a frequency domain, much quicker than the time domain. Furthermore, the possibility of simulating the amplitude and phase offset values in the responses using a frequency domain modeling permits a more affordable application of this interpretation method in an inversion approach.

We have presented in the field part only the results from one pumping test associated to one period, selected from a campaign of pumping tests in which several different boreholes and different period values were tested. In order to interpret simultaneously all responses from all harmonic pumping tests, we have to use an inverse modeling. This technique is already widely used to characterize and quantify the heterogeneity in fractured and karstic fields, by interpreting the responses to constant-rate pumping hydraulic tomography (*Hao et al. 2008 ; Illman et al. 2009 ; Castagna et al. 2011 ; Illman 2014 ; Zha et al. 2016 ; Wang et al. 2017*). In a future work, we plan to focus more specifically our discussion on the quantitative analysis of the harmonic pumping responses in a karstic field, by associating them to a conduit network hydraulic imagery, based on an inverse modeling approach and its ability to reproduce the complete set of responses with a given distribution of properties.

7 APPLICATION DE TOMOGRAPHIE HYDRAULIQUE EN MODE HARMONIQUE SUR LE SITE DU TERRIEU

7.1 Contexte

Ce chapitre final aborde la caractérisation par tomographie hydraulique du site karstique expérimental du Terrieu à partir des nouvelles méthodes d'investigation et de modélisation présentées dans ce manuscrit, à savoir l'investigation par pompes harmoniques et l'imagerie par méthode CADI.

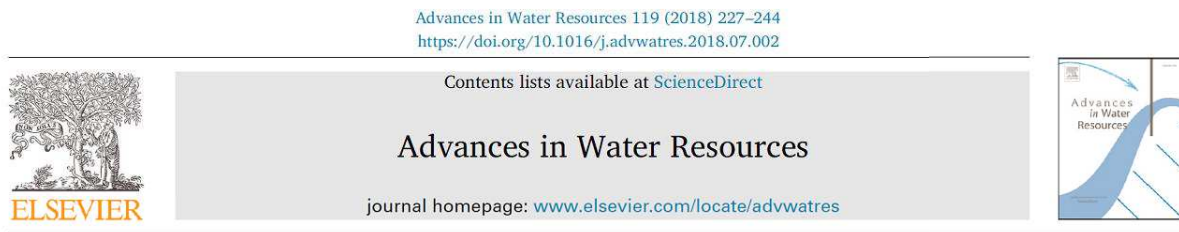
L'investigation du site consiste en une série de pompes harmoniques dans quatre forages différents, à chaque fois menés pour deux différentes fréquences de signal. Les réponses de rabattements sont mesurées à chaque fois dans treize forages de mesures. Le traitement des signaux de réponses est le même que celui présenté dans le chapitre précédent : seules les parties oscillatoires des réponses sont préservées afin d'en interpréter les amplitudes et déphasages.

La modélisation inverse, basée sur la méthode CADI en domaine fréquentiel, cherche à reproduire les amplitudes et déphasages des réponses mesurées sur site. Des inversions séparées sont d'abord réalisées avec les réponses liées à chacune des deux fréquences de signal de pompage, puis des inversions jointes tentant de reproduire l'ensemble des réponses en même temps. Les résultats d'inversion montrent que les réponses aux différentes fréquences de signal peuvent être bien reproduites séparément par la méthode CADI, mais la minimisation des écarts entre réponses simulées et mesurées devient moins bonne lorsque toutes les réponses doivent être reproduites simultanément. Il apparaît dans les images des champs de propriétés produites que les réponses à un signal à plus basse fréquence sont reproduites par un réseau de conduits plus dense que les réponses à un signal à plus haute fréquence. Une analyse des connectivités interprétées à partir de ces imageries suggère de plus que les réponses aux signaux à plus haute fréquence permettent de mieux caractériser les connectivités par conduits entre les forages.

Finalement, l'étude des cartes d'incertitudes sur les champs de propriétés générés, associée aux observations faites précédemment, permet de discuter sur les informations qu'apportent chaque fréquence de signal dans la caractérisation du milieu fracturé et karstique. Les fréquences plus élevées de pompage seraient associées à des écoulements plus spécifiquement de conduits, qui permettent de mieux caractériser les connectivités des forages situés dans ces réseaux de conduits. Les fréquences plus faibles mobilisent des écoulements plus variés autour du point de pompage, mobilisant également des structures moins conductrices, et permettent de mieux caractériser les chemins d'écoulements préférentiels autour du forage de pompage.

7.2 Harmonic Pumping Tomography Applied to Image the Hydraulic Properties and Interpret the Connectivity of a Karstic and Fractured Aquifer (Lez Aquifer, France)

Cette partie est composée de l'article « Harmonic Pumping Tomography Applied to Image the Hydraulic Properties and Interpret the Connectivity of a Karstic and Fractured Aquifer (Lez Aquifer, France) », publié dans le journal *Advances in Water Resources* en septembre 2018 (Fischer *et al.* 2018c). Le texte a été remis en forme au format du manuscrit. La version originale de l'article est donnée en Annexe 6.



Harmonic pumping tomography applied to image the hydraulic properties and interpret the connectivity of a karstic and fractured aquifer (Lez aquifer, France)

P. Fischer^{a,*}, A. Jardani^a, H. Jourde^b, M. Cardiff^c, X. Wang^b, S. Chedeville^b, N. Lecoq^a

^a Normandie Univ, UNIROUEN, UNICAEN, CNRS, M2C, 76000 Rouen, France

^b Laboratoire Hydrosiences, Université de Montpellier, CNRS, 34000 Montpellier, France

^c Department of Geoscience, University of Wisconsin-Madison, Madison, WI, USA

* Corresponding author.

7.2.1 Introduction

The protection and the management of the water resources involve the identification of the preferential flow paths in the ground. Therefore, one needs to characterize the spatial distribution of the hydraulic properties in the field subsurface. A common way to assess the hydraulic properties of a field, such as conductivity and specific storage, is the analysis of the drawdown responses to a pumping test from which local or average properties can be inferred from analytical equations that relate the hydraulic response to the hydraulic properties (*Butler 2005*).

However, in the case of karstic aquifers, the assessment of the hydraulic properties is challenging (*White 2002 ; Hartmann et al. 2014a*) as the hydraulic properties in this type of aquifer can vary by several orders of magnitude within a short distance (*Wang et al. 2016*). This makes the characterization of the karstic fields very complex. To face this difficulty, it is then necessary to interpret the responses of the field by taking into account the positioning of the conduits network, which determines the preferential flow paths (*Kovacs 2003 ; Ghasemizadeh et al. 2012 ; Saller et al. 2013*).

The hydraulic tomography concept has been widely applied to map the spatial variability of hydraulic properties, in both type of aquifers (porous and fractured), by performing a joint interpretation of hydraulic data recorded simultaneously at several wells, as responses to extraction/injection of water (*Yeh and Liu 2000 ; Bolhing et al. 2002 ; Zhu and Yeh 2005 ; Yeh and Lee 2007 ; Cardiff et al. 2009 ; Castagna et al. 2011 ; Berg and Illman 2013 ; Cardiff et al. 2013a ; Zha et al. 2015 ; Zha et al. 2016 ; Zha et al. 2017*). This approach relies on a numerical technique (such as finite difference, finite element and finite volume) to solve the groundwater flow equation, and the inverse process to reconstruct the heterogeneity of the hydraulic conductivities and the storage properties by fitting the piezometric responses. The inversion process usually provides a non-unique solution which can produce an ambiguous interpretation of the hydraulic data. To overcome this issue, a prior information on the distribution of the properties can be used to constrain and guide the inversion to a more realistic solution (*Carrera and Neuman 1986a*). In the case of aquifers with a low heterogeneity, the geostatistical constraints remain the most simple and efficient way to find accurate solutions (*Hoeksema and Kitanidis 1984 ; Kitanidis 1995 ; Fischer et al. 2017a*). In the context of fractured and karstic aquifers, the definition of the a priori model, or even the strategy for the numerical groundwater flow simulation (which can be performed by using various techniques such as equivalent porous media or discrete fractures networks), remain the subject of open

debates among hydrogeologists. In fact, for a successful interpretation of hydraulic responses of karstic aquifers, the ‘classical’ geostatistical inversion method would require a dense network of measurement and a significant resolution of model parameterization because of the high contrasts existing in the distribution of the spatial properties. Recently, several inversion methods have been developed for characterizing karst networks. One way is to directly incorporate a discrete geometry within a background model using a discrete-continuum forward model (*Teutsch 1993 ; Liedl et al. 2003 ; de Rooij et al. 2013*). In this case, the parameterization of the inverse problems usually relies on a stochastic generation of discrete fracture networks that are conditioned to statistical (*Li et al. 2014 ; Le Coz et al. 2017*), mechanical (*Jaquet et al. 2004 ; Bonneau et al. 2013*), or structural data (*Pardo-Iguzquiza et al. 2012 ; Collon et al. 2017*). Another way is based on a deterministic optimization of the geometry of discrete networks (*Borghi et al. 2016 ; Fischer et al. 2018a*).

Previous studies have shown that equivalent porous media models are able to reproduce the hydraulic flows in karstic aquifers at a kilometric scale (*Larocque et al. 1999 ; Abusaada and Sauter 2013*) or a decametric scale (*Wang et al. 2016*). However, if the scale of investigation is too small, this type of model can become unreliable for the characterization of the properties of fractured rocks, extremely contrasted and structured at a small scale (*Illman 2014*). Although the classical geostatistical inverse approaches were originally proposed for inversion of hydraulic fields, they can be made adaptive to discrete geometries with special treatments to the prior model (e.g. the total variation prior model, *Lee and Kitanidis 2013*), or using an iterative procedure (e.g. the sequential successive linear estimator, *Ni and Yeh 2008 ; Hao et al. 2008 ; Illman et al. 2009 ; Sharmeen et al. 2012*). Other methods for inversion of complex discrete structures involve introducing constraints of a priori knowledge to the inverse model using a guided image (*Hale 2009 ; Soueid Ahmed et al. 2015*), a training image (*Lochbühler et al. 2015*), a probability perturbation (*Caers and Hoffman 2006*), a transition probability distribution (*Wang et al. 2017*), a multi-scale resolution (*Ackerer and Delay 2010*), a level-set method (*Lu and Robinson 2006 ; Cardiff and Kitanidis 2009*), or based on cellular automata (*Fischer et al. 2017b*).

Apart from these challenges in modeling techniques, a further difficulty in karst aquifer characterization raises from characteristics of hydraulic tests. Due to the integration nature of pressure diffusion, the steady state responses of long-term constant-rate pumping tests in a karst aquifer represent the combined effect of the multiple media (conduits, fissures, matrix) rather than specific components. Although the interpretation of transient responses may provide

additional information about the relative importance of each aquifer components, the inclusion of such data in a modeling in the time domain is also computationally demanding. Recently, harmonic pumping tests have been introduced as an alternative for the identification of the networks of preferential groundwater flows. Harmonic characterization designates an investigation in which an oscillatory/sinusoidal signal is used to disturb the water level of an aquifer. Different ways to produce such signals have already been proposed: a pumping-reinjecting system (*Rasmussen et al. 2003 ; Renner and Messar 2006*), a moving mass at the water table interface (*Guiltinan and Becker 2015*), or a controlled pumping system (*Lavenue and de Marsily 2001*). Then, the response signals among the aquifer contain an oscillatory part (characterized by an amplitude and a phase offset values) that can be easily isolated from the ambient noise (*Bakhos et al. 2014 ; Cardiff and Barrash 2015*). Harmonic characterization has already been successfully applied to a large range of complex cases such as contaminated aquifer (*Rasmussen et al. 2003*), leakage detection (*Sun et al. 2015*), or fractured aquifers (*Renner and Messar 2006 ; Maineult et al. 2008 ; Guiltinan and Becker 2015*). The theoretical aspects of the application of harmonic pumping to karstic aquifers have also been developed in *Fischer et al. (2018b)*. The imagery potential of harmonic investigations has been studied for mapping the distribution of hydraulic properties in heterogeneous aquifers with models solved in the time domain (*Lavenue and de Marsily 2001 ; Soueid Ahmed et al. 2016*) or in the frequency domain (*Cardiff et al. 2013b ; Zhou et al. 2016*).

In this article we will present a field characterization of karst network based on a harmonic pumping tomography. Hydraulic data were obtained from the Terrieu experimental site located in Montpellier, in Southern France. At the same site, results of hydraulic tomography, in which hydraulic responses were generated with constant-rate pumping tests, have already been presented and discussed in our previous works (*Wang et al. 2016 ; Wang et al. 2017 ; Fischer et al. 2017c*). In this new work, we rely our analysis on a set of responses to harmonic pumping tests with different oscillation periods, to characterize the karst network. We describe in section 7.2.2 the experimental study site, the harmonic pumping investigation led on it, and the processing made on the measured field responses for the later interpretation. Then, in section 7.2.3 we briefly introduce the numerical model setup and the Cellular Automata-based Deterministic Inversion (CADI) algorithm. Further details of our inverse algorithm can be found in *Fischer et al. (2017b)*. In section 7.2.4 we present the inversion results obtained with the CADI method at the Terrieu field site and the efficiency of the method in reproducing the observed hydraulic responses. Finally, section 7.2.5 presents a discussion of the effect of the harmonic signal period on the inversion results.

7.2.2 Field investigation

7.2.2.1 Experimental site presentation

The Terrieu experimental site is located ~15 km in north of the town of Montpellier in southern France. The site consists of 22 vertical boreholes drilled within a surface area of approximately 2,500 m² (40 m × 60 m) and permits the study of karstic flows at a local scale (Figure 7.1). As a part of the network of the French Karst Observatory (SNO Karst, www.sokarst.org) and the Medycyss network (*Jourde et al. 2011*), the site has been used as a field laboratory for testing new field hydraulic methods and tools developed for the characterization of karstic aquifers (*Jourde et al. 2002 ; Jazayeri Noushabadi 2009 ; Jazayeri Noushabadi et al. 2011 ; Dausse 2015 ; Wang et al. 2016 ; Wang et al. 2017 ; Fischer et al. 2017c*).

The geological logs collected from the different boreholes shows that the subsurface of the field is composed of 35 to 45 m of thin-layered marly Cretaceous limestones, deposited on pure and massive Jurassic limestones. The interface between these two units is a sloped monocline bedding plane dipping at 20° toward Nord-West (*Wang et al. 2016*).

The Terrieu field is located in the Lez regional aquifer. Temperature and electrical conductivity measurements, and packer tests in boreholes presented in previous works (*Jazayeri Noushabadi 2009 ; Dausse 2015*) have shown the existence of preferential flow paths (shown in Figure 7.1) along the bedding plane between the Cretaceous and Jurassic limestones. Downhole videos in the boreholes show, that, at this interface, karstic conduits with aperture up to 50 cm are present (*Jazayeri Noushabadi et al. 2011*).

The massive Jurassic limestones are non-aquifer and the Cretaceous limestones have a low permeability. This causes the aquifer to be confined at the interface between these two layers, in the bedding plane. A network of karstic conduits has developed preferentially on this bedding plane, and controls the fluid circulation within the aquifer.

7.2.2.2 Harmonic pumping investigation

The main dataset used in this study was collected from an investigation using harmonic pumping tests performed at the Terrieu site. We have conducted pumping sequentially in four different boreholes while recording the water-level responses in 13 selected observational

boreholes (see Figure 7.1). The water-level responses were also measured in the pumping wells P9, P15 and P20 but not in P3.

The static water level before the hydraulic investigation was at a depth of 20 m. The maximal drawdown generated by the pumping tests was 4 m. Therefore the karstic conduits (located at a depth of 35 to 45 m) were saturated during the entire duration of hydraulic tests.

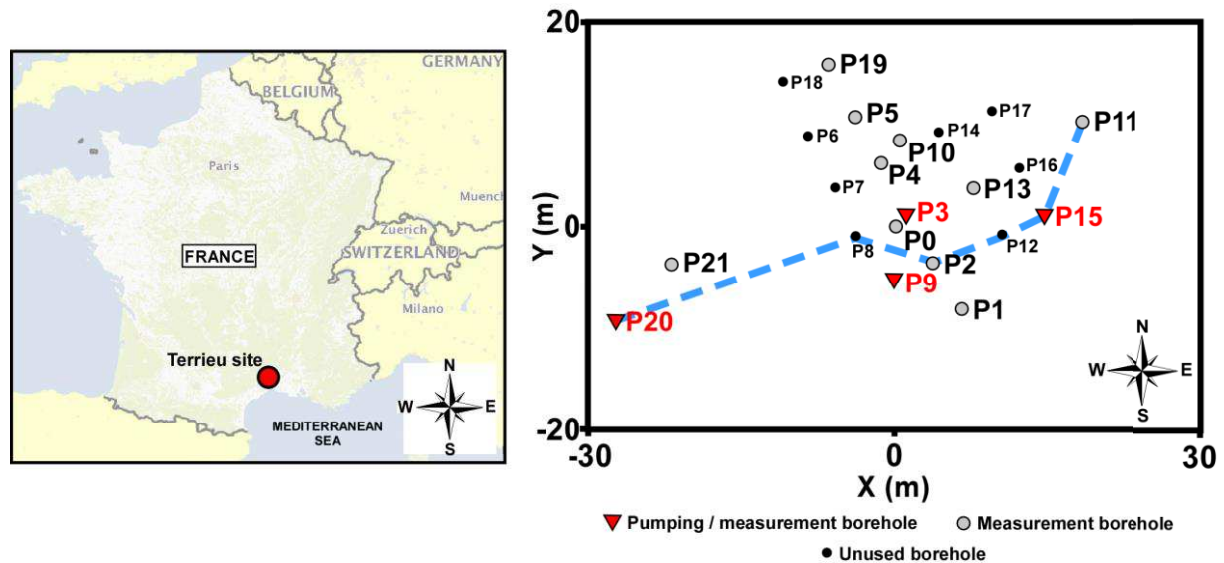


Figure 7.1: Maps of localization of the Terrieu site in France (left) and well pattern on the site (right). Boreholes used as pumping and measurement points are indicated using red triangles, and boreholes used only as measurement points are indicated using grey circles. Boreholes indicated by solid black points were not used during the investigation. The blue dotted line delineates a preferential flow path identified by previous studies (Jazayeri Noushabadi 2009 and Dausse 2015), which shows a connectivity between P2, P8, P11, P12, P15 and P20.

The harmonic pumping tests were performed with a configurable electronic device, specially designed for this study by electronics engineers. This device controls a flow rate variator linked to the pump, which can generate a pumping signal with a sinusoidal shape around a mean value. The period and amplitude of the sinusoid signal can be configured with the device.

The generated pumping rate can be described by:

$$Q(t) = Q_m - Q_A \cos(\omega t) , \quad (7.1)$$

where Q is the output pumping signal (m^3/s), Q_m the mean pumping rate (m^3/s), Q_A the oscillatory signal amplitude (m^3/s), and $\omega = \frac{2\pi}{T}$ the pulsation (rad.s^{-1}) with T the period (s).

Different signal amplitudes and mean values were independently applied in each different pumping borehole according to its productivity (see Table 7.1).

Table 7.1: Harmonic pumping rates registered for each pumping point during the investigation. Q_A and Q_m refer to Equation (7.1).

Pumping well:		P3	P9	P15	P20
Pumping rate (m^3/h)	Amp. Q_A	1	0.22	2.1	2.5
	Mean Q_m	4.1	0.35	5.3	3.8

For each pumping location, two pumping tests with different periods ($T = 2$ min and $T = 5$ min) were conducted during 30 min (15 cycles for a 2 min period, 6 cycles for a 5 min period). Water-level variations were continuously measured with digital pressure sensors installed in the measurement wells.

Overall, this investigation permitted to record 104 drawdown curves (13 measurements for each 2 different periods of signal applied in each 4 pumping wells).

7.2.2.3 Data processing

In order to interpret the harmonic signal in the drawdown curves, we have performed the same signal decomposition as proposed in *Fischer et al. (2018b)*. This decomposition consists in removing the linear part, induced by the mean pumping signal Q_m , from the drawdown curve (through a linear regression) to keep only the oscillatory response. This

operation is feasible only after an early transient period (accordingly we truncate the first cycle of the recorded responses). As we show in Figure 7.2 for the pumping in P15, the operation of removing the linear part is acceptable, as the resulting signals appear to be purely oscillatory. Some natural noises and vibrations induced by the pumping appear as high frequencies fluctuations in the oscillatory responses.

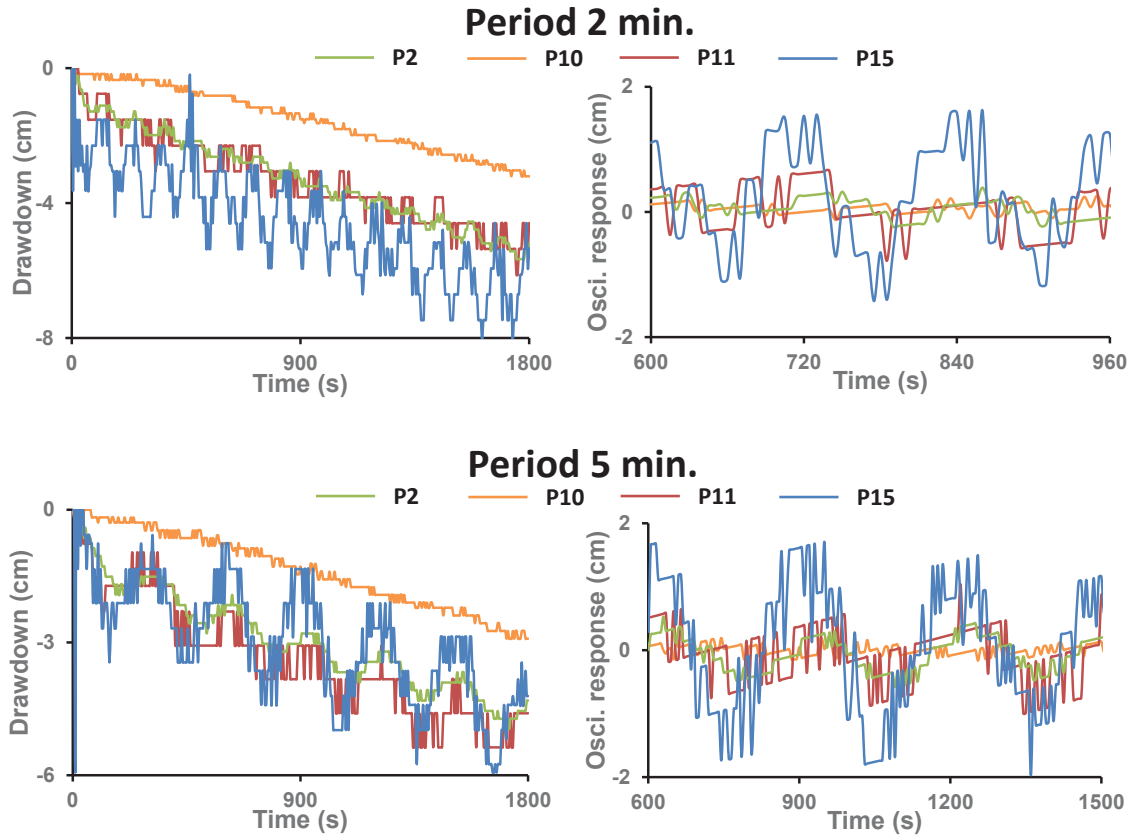


Figure 7.2: Left: Measured drawdown curves for a selection of boreholes (P2, P10, P11, P15) during a pumping in P15 with a 2 min and a 5 min period. Right: Zoom-in view of three oscillation cycles after removing the linear part from the drawdown curves.

Mathematically the drawdown curves generated by the harmonic pumping tests can be approximated as a sum of a linear signal and an oscillatory signal applied on the initial water table level:

$$h(t) \simeq h_{lin.}(t) + h_{osc.}(t) + H_0$$

$$\text{with } h_{lin.}(t) \simeq -at - h_0 \quad \text{and} \quad h_{osc.}(t) \simeq A \cos\left(\frac{2\pi}{T}t - \Phi \frac{\pi}{180}\right), \quad (7.2)$$

where h represents the drawdown (m) over time, $h_{lin.}$ is a linear signal described by its slope a (m/s) and its intercept h_0 (m) (whose values can be retrieved by linear regression), $h_{osc.}$ is an oscillatory signal described by its amplitude A (m), its period T (s) and its phase shift Φ ($^\circ$), and H_0 represents the initial water table level (m) (in our case we considered $H_0 = 0$ m).

The linear signal $h_{lin.}$ can be easily estimated in a first approximation through a linear regression performed on each drawdown curve. After removing this linear trend, the amplitude and phase offset of the remaining signal of each borehole can be determined by a Fast Fourier Transform (FFT) on their oscillatory signals. The FFT permits to extract the main oscillatory components of a signal, to denoise it, and to interpret its parameters. Figure 7.3 presents the FFT results for the oscillatory signals of three representative boreholes (P10, P11, P2) during a pumping in P15 with the two periods (2 min and 5 min). The interpretation results of amplitude and phase offset for the entire dataset are presented in Appendix 6.

7.2.2.4 Preliminary analysis

The different responses of amplitude and phase offset interpreted in P10, P11 and P2 highlight three distinct flow behaviors (*Fischer et al. 2018b*). The responses in P10, having a negligible amplitude (< 1 mm) relatively to the pumping signal, which we interpreted as a negligible oscillatory response, is associated to a ‘matrix connectivity’ between the pumping and the observational well. In contrast, the response in P11 has a significant amplitude and an almost invariable phase relatively to the pumping signal for the two different periods. This behavior is associated to a ‘conduit connectivity’ response, meaning that P15 and P11 would be connected through a karstic conduit network. The response in P2 has a lower amplitude response than P11, and its phase offset relatively to the pumping signal increases as the pumping period decreases ($+71^\circ$ for a 2 min signal, $+38^\circ$ for a 5 min signal). This third behavior is associated to a ‘dual connectivity’ response, which corresponds to an inter-well connection either through fissures or when the observation borehole is located in the matrix but close to a conduit.

Following the method described in *Fischer et al. (2018b)* and through the integration of the amplitude and phase offset results interpreted for each pumping-observation well pair, it is possible to obtain a map of inter-well connectivity which contains qualitative information

regarding the spatial distribution of the conduit network and the relative position of boreholes to the network.

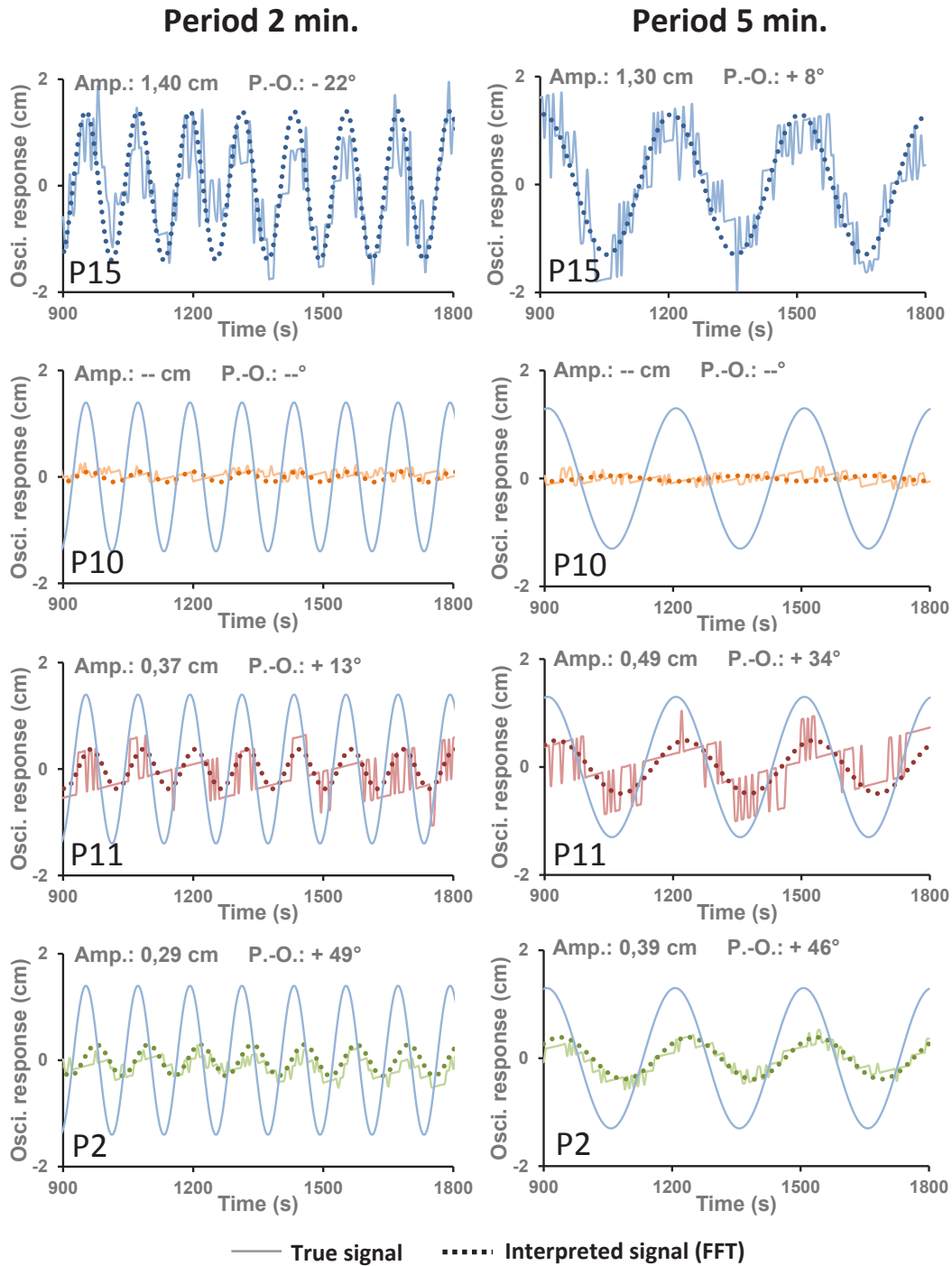


Figure 7.3: Zoom-in on the oscillatory responses extracted from the drawdown measured in P2, P10, P11 and P15 during pumping tests in P15 with a 2 min (left) and a 5 min (right) signal periods and FFT results of the interpreted amplitude (Amp.) and phase offset (P.-O.) responses. Solid lines represent the measured signals, dotted lines represent the interpreted signals ($h_{osc.}$ in Equation (7.2)) reconstructed from the amplitudes and phase offsets interpreted by FFT. For interpreted amplitudes smaller than 1 mm (for example here in P10), we considered the oscillatory responses to be negligible. The blue lines represent the interpreted pumping signals (P15) and are presented for each borehole for a better visualization of the interpreted phase offset responses.

We can first link, on the map, the boreholes with a low phase shift relatively to the pumping signal, to represent a conduit connection. From this conduit connectivity we can then establish a possible conduit network, and then verify that the other boreholes responses would be adequate toward their position to the interpret network (dual connectivity with higher phase shift for boreholes close to the network and matrix connectivity with negligible responses for the others).

Possible connectivity maps interpreted with the responses to the harmonic pumping test in P15 for a period of 2 min. and 5 min. are proposed in Figure 7.4.

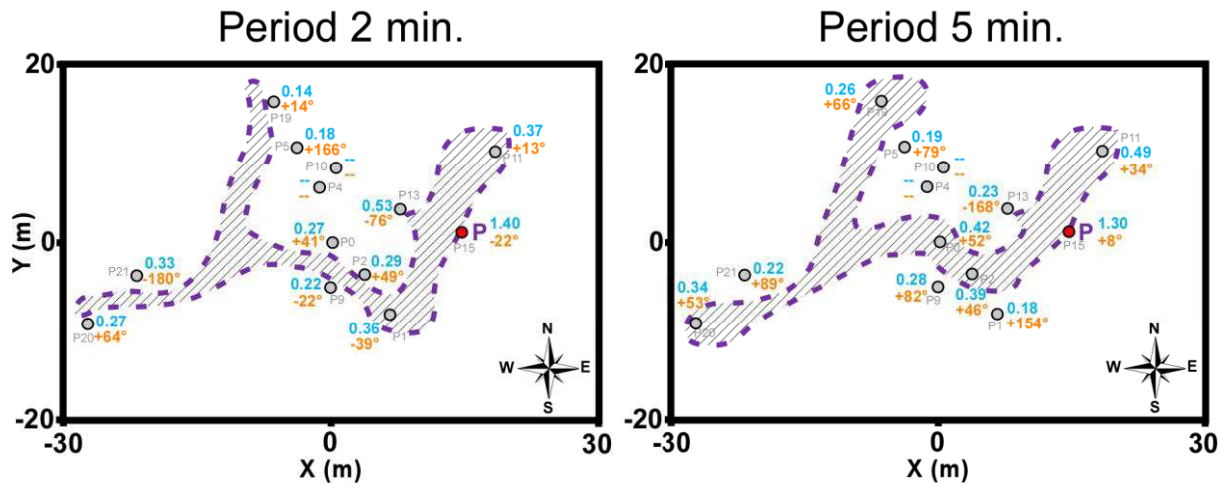


Figure 7.4: Connectivity maps interpreted from the amplitude (in blue) and phase offset (in orange) responses to a pumping in P15 with a 2 min (left) and a 5 min (right) period of signal. The areas within the dotted lines delineate a possible area where boreholes are connected through a direct conduit connectivity. Dashes indicate negligible oscillatory responses.

The comparison between the two connectivity maps in Figure 7.4 shows that the period of the harmonic pumping signal may have a slight impact on the connectivity interpretation. For example in the pumping test with a 2 min period P1 can be interpreted as connected to the pumping well through conduits, but not in the pumping test with a 5 min period. This implies that a change in the period of the pumping signal modifies the flow field induced by the pumping.

Furthermore, the manual interpretation is possible only when the amount of hydraulic data to deal with is limited (13 responses for each period in Figure 7.4). Therefore, although such a qualitative analysis through manual interpretation of inter-well connectivity could

provide some important guidance to hydrogeological investigation, such as indicating the general trend of the main conduits and relative inter-well connectivity, to obtain a quantitative hydrodynamics characterization and to integrate a larger amount of hydraulic measurements (104 responses from 4 different pumping locations with each time two different periods) an inverse modelling is required.

7.2.3 Modeling methodology

7.2.3.1 Forward problem and model parameterization

Inverse modeling involves the use of the forward problem in order to simulate, for a given model of hydraulic properties, the hydraulic responses. In this section we present a 2D model that represents the property field along the bedding plane interface, in which the karstic network has developed on the Terrieu site.

As seen in the previous section, the variation of the piezometric level among the site, in response to a harmonic pumping, can be approximated by the sum of a linear drawdown and an oscillatory drawdown (Equation (7.2)). We simulate in the model only the oscillatory part of the drawdown responses $h_{osc.}$. The inversion aims to reproduce the values of amplitude and phase offset of the oscillatory part in the measured responses. This oscillatory part can be described as a signal in the frequency domain in the model:

$$h_{osc.}(x, y, t) = Re(\mathcal{H}_\omega(x, y)e^{i\omega t}) . \quad (7.3)$$

with \mathcal{H}_ω a complex parameter holding the amplitude and phase offset responses over space (x, y) , Re the function returning the real part of a complex value, $\omega = \frac{2\pi}{T}$ the pulsation ($rad.s^{-1}$) and i the imaginary unit.

This oscillatory feature of the hydraulic signal permits to rewrite the time domain form of the groundwater flow equation into a frequency form, in order to reduce the computation time of the forward problem.

In a 2D, porous, isotropic and saturated domain Γ the groundwater flow equation based on the Darcy's law in a frequency domain can be expressed as:

$$i\omega S_s \mathcal{H}_\omega - \nabla \cdot (K \nabla \mathcal{H}_\omega) = \frac{Q_A}{V_{el.}} \delta(x - x_s, y - y_s) , \quad (7.4)$$

with S_s the specific storage distribution (m^{-1}), K the conductivity distribution (m/s), Q_A the pumping amplitude (m^3/s), $V_{el.}$ an elementary volume of the finite element grid in the model,

and $\delta(x - x_s, y - y_s)$ the Dirac distribution where x_s, y_s represents a pumping location. As the system is 2D, with a unit thickness, conductivity K and transmissivity T are of same value, as well as specific storage S_s and storativity S . In this study we considered the Darcy's law to be acceptable for representing the flows generated in the karstic structures. In fact, it appears from the previous studies on the Terrieu site that the flows in the conduits have a low velocity, inducing a low Reynolds value, even for higher pumping rates than the ones used during this new investigation.

The initial and boundary conditions used for solving Equation (7.4) are:

$$\begin{aligned} \mathcal{H}_\omega(x, y) &= 0 \quad \forall (x, y) \in \Gamma \text{ as initial condition} \\ \mathcal{H}_\omega(x, y) &= 0 \quad \text{when } (x, y) \in \Gamma_{\text{bound.}} \text{ as boundary condition} \end{aligned} \quad (7.5)$$

The spatial distribution of the complex parameter \mathcal{H}_ω permits the reconstruction of the oscillatory responses simulated among the model, through the calculation of their amplitude and phase offset values:

$$\left| \begin{array}{l} - \text{Amplitude: } A(x, y) = \sqrt{(Re \mathcal{H}_\omega(x, y))^2 + (Im \mathcal{H}_\omega(x, y))^2} \quad \text{in m} \\ - \text{Phase offset: } \Phi(x, y) = \frac{180}{\pi} \text{atan2}(-Im \mathcal{H}_\omega(x, y), Re \mathcal{H}_\omega(x, y)) \quad \text{in } ^\circ \end{array} \right. , \quad (7.6)$$

where Re and Im are the functions returning the real and imaginary parts of a complex value, and $atan2$ is the function returning the inverse tangent value in radian mode from two arguments.

The simulated response signals are then reconstructed temporally and spatially:

$$h_{\text{osc.}}(x, y, t) = A(x, y) \cos\left(\omega t - \Phi(x, y) \frac{\pi}{180}\right). \quad (7.7)$$

In a karstic medium, the spatial response signals are very dependent to the highly heterogeneous distribution of the field properties T and S along the karstic conduits. Therefore this heterogeneity has to be taken into account in the distribution of these properties in the model in order to simulate a realistic responses behavior. For this purpose we chose to

apply as parameterization for our model and inverse problem the Cellular Automata-based Deterministic Inversion (CADI) method, developed and detailed in *Fischer et al. (2017b)*. For a detailed description of the CADI method we refer the reader to *Fischer et al. (2017b)*, as we will only briefly summarize the concept in this article.

The CADI method uses a particular parameterization of the property field in the model to generate linear structures (conduits) over a background (matrix). The field is composed of a grid of cells, each cell being assigned to a value of transmissivity and storativity. This grid of cells is divided in m_{CA} subspaces, each one being controlled by a cellular automaton piloting the part of the cells inside of its subspace (Figure 7.5). The cellular automata concept is a mathematical tool which permits to generate structures within a grid with simple neighborhood and transition rules (*Von Neumann and Burks 1966*). The cellular automata control the local direction of generation of the conduit in the different subspaces. They are piloted through eight different neighborhood definition $N_i, i \in [1,8]$ that permit to define eight different directions. One of these eight direction possibilities is assigned to each cellular automaton in the subspaces. The conduit network is generated by first assigning a state ‘matrix’ or ‘conduit’ to each cell. The whole cells are initially in state ‘matrix’, except an initial cell of the grid in state ‘conduit’ which designates the starting point and starting subspace for the generation of the conduits. The network of conduits then generates following the different local direction affected to each subspace it crosses. The generation ends once each part of the network has reached an end (either the limit of the model or a subspace in which the network has already generated).

Then property values are assigned to the cells depending on their state (‘matrix’ or ‘conduit’) and their localization (subspace). Each subspace defines locally a value for T_{mat} and S_{mat} for its cells in state ‘matrix’ and a value for T_{cond} and S_{cond} for its cells in state ‘conduit’.

In order to be able to easily pilot and modify the configuration of the model through this parameterization, the structural directions and properties assigned to each subspace are defined in two parameter vectors: \mathbf{P}_N and \mathbf{P}_β . \mathbf{P}_N is a m_{CA} -vector containing the directions of generation $N_i, i \in [1,8]$ assigned to each subspace. Several independent networks can be generated in the same model with different directions parameters for each network. In this case \mathbf{P}_N becomes a $(m_{CA} \times frac)$ matrix where *frac* represents the amount of independent networks in the model. Each column contains the subspaces directions for each network.

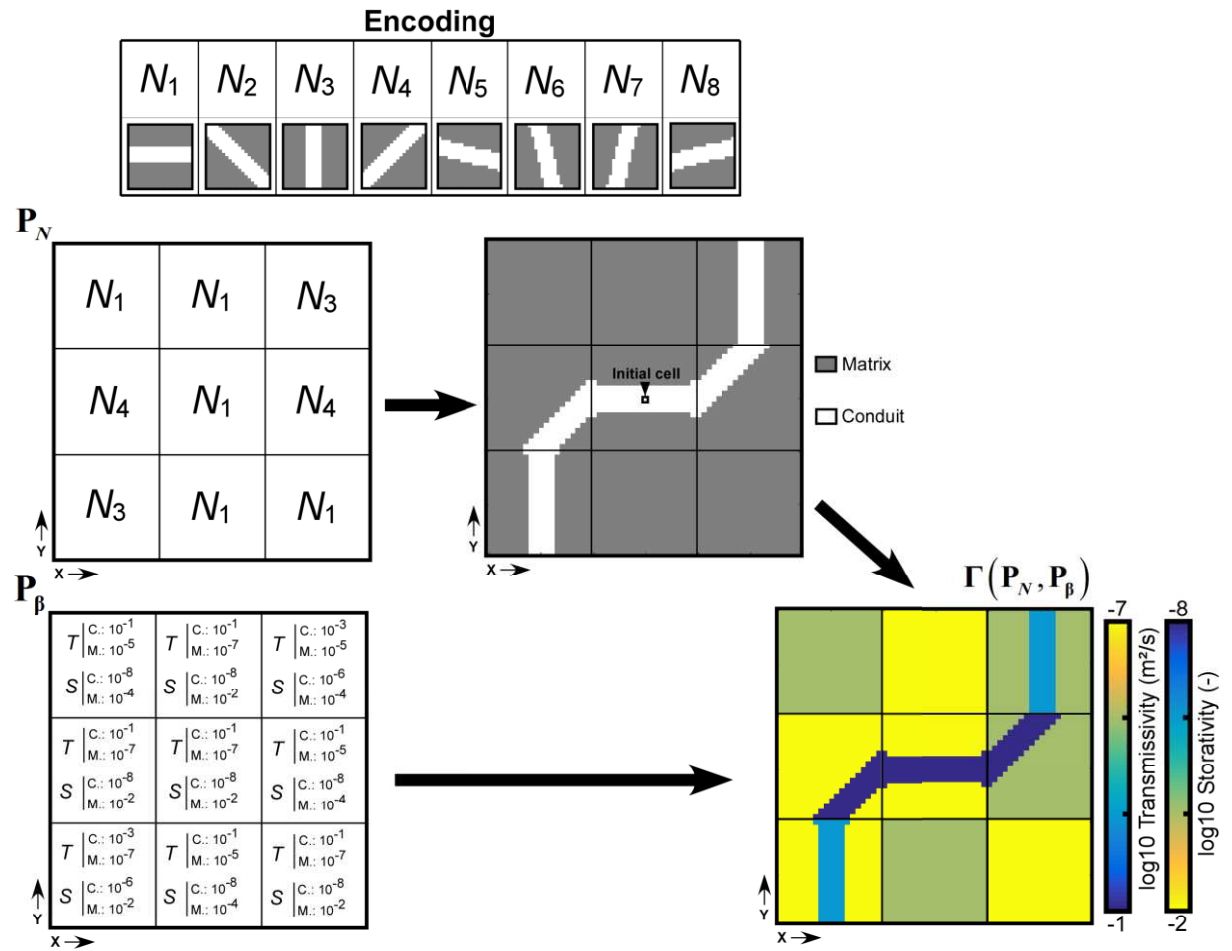


Figure 7.5: Schema of the parameterization of a model with the CADI method. P_N contains the encoded (see Encoding) structural directions of generation associated to each subspace which permits to generate, from an initial 'conduit' cell, a network of conduits in the matrix. P_β contains the conduit (C) and matrix (M) transmissivity and storativity values associated to each subspace. $\Gamma(P_N, P_\beta)$ designates the model produced by applying the property values from P_β to the network generated from P_N .

P_β is a $4m_{CA}$ -vector containing the T_{mat} and S_{mat} and the T_{cond} and S_{cond} values assigned to each subspace. In this way the parameters controlling the configuration of the model, and thus the model itself, can be easily and locally modified. P_N and P_β represent the parameters to be optimized in the inverse problem in order to reproduce the observed data (amplitude and phase offset responses) through a suitable model.

7.2.3.2 Inverse problem

The inverse problem consists in retrieving the best values for the parameters contained in \mathbf{P}_N and \mathbf{P}_β regarding the minimization of the gap between the simulated data and the observed data (amplitude and phase offset responses at the measurement points for the different pumping tests). This inverse algorithm contains two steps in which we seek to minimize two objective functions sequentially, $\Psi_{\text{structure}}$ for the optimization of the structural parameter \mathbf{P}_N and $\Psi_{\text{properties}}$ for the optimization of the property parameter \mathbf{P}_β (Tarantola and Valette 1982):

$$\begin{aligned} \Psi_{\text{structure}}(\mathbf{P}_N) = & \frac{1}{2} \left(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)) \right)^T \mathbf{C}_d^{-1} \left(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)) \right) \\ & + \frac{1}{2} \left(\mathbf{P}_{N,\text{prior}} - \mathbf{P}_N \right)^T \mathbf{C}_{\mathbf{P}_N}^{-1} \left(\mathbf{P}_{N,\text{prior}} - \mathbf{P}_N \right) \end{aligned} \quad (7.8)$$

$$\begin{aligned} \Psi_{\text{properties}}(\mathbf{P}_\beta) = & \frac{1}{2} \left(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)) \right)^T \mathbf{C}_d^{-1} \left(\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)) \right) \\ & + \frac{1}{2} \left(\mathbf{P}_{\beta,\text{prior}} - \mathbf{P}_\beta \right)^T \mathbf{C}_{\mathbf{P}_\beta}^{-1} \left(\mathbf{P}_{\beta,\text{prior}} - \mathbf{P}_\beta \right) \end{aligned} \quad (7.9)$$

where \mathbf{d}_{obs} is a n -vector containing the n measured responses, $f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta))$ is a n -vector containing the responses simulated with the model at the same positions than in \mathbf{d}_{obs} , \mathbf{C}_d is a $(n \times n)$ matrix of covariance on the data, $\mathbf{P}_{N,\text{prior}}$ is a m_{CA} -vector holding *a priori* structural parameters for \mathbf{P}_N , $\mathbf{P}_{\beta,\text{prior}}$ is a $4m_{\text{CA}}$ -vector holding *a priori* property values for \mathbf{P}_β , and $\mathbf{C}_{\mathbf{P}_N}$ and $\mathbf{C}_{\mathbf{P}_\beta}$ are $(m_{\text{CA}} \times m_{\text{CA}})$ and $(4m_{\text{CA}} \times 4m_{\text{CA}})$ matrices of prior covariance on the parameters \mathbf{P}_N and \mathbf{P}_β .

At the beginning of the inversion process, the variable parameters \mathbf{P}_N and \mathbf{P}_β are initialized with *a priori* conduit directions and property values in order to create the initial model. Then a sequential and deterministic process optimizes firstly the structural parameters in \mathbf{P}_N (considering the initial properties in \mathbf{P}_β as invariable), and then, in second step, the property parameters in \mathbf{P}_β (considering the previously optimized \mathbf{P}_N as invariable). Finally, after the optimization process, the posterior uncertainties on the structural and property parameters are estimated.

7.2.3.2.1 Optimization of the structural parameters

The optimization of the structural parameters in \mathbf{P}_N is an iterative process in which a sensitivity matrix is computed, at each iteration step, to minimize the objective function in Equation (7.8). This analysis requires the computation of a $(8 \times m_{CA})$ sensitivity matrix \mathbf{S} . At a given iteration step k , each element (i, j) of the matrix is calculated as follow:

$$\mathbf{S}^k(i, j) = \left(\mathbf{d}_{\text{obs}} - f\left(\Gamma\left(\mathbf{P}_N^k \Big|_{\mathbf{P}_N^k(j)=N_i}, \mathbf{P}_\beta\right)\right) \right)^T \mathbf{C}_d^{-1} \left(\mathbf{d}_{\text{obs}} - f\left(\Gamma\left(\mathbf{P}_N^k \Big|_{\mathbf{P}_N^k(j)=N_i}, \mathbf{P}_\beta\right)\right) \right) + \frac{1}{2} \left(\mathbf{P}_{N,\text{prior}}(j) - N_i \right)^T \mathbf{C}_{\mathbf{P}_N}^{-1} \left(\mathbf{P}_{N,\text{prior}}(j) - N_i \right) \quad (7.10)$$

where $f\left(\Gamma\left(\mathbf{P}_N^k \Big|_{\mathbf{P}_N^k(j)=N_i}, \mathbf{P}_\beta\right)\right)$ represents the responses simulated with the modified direction N_i in the subspace $\mathbf{P}_N^k(j)$, and $\mathbf{P}_{N,\text{prior}}(j) - N_i$ denotes the angular gap between the modified direction N_i and the *a priori* direction $\mathbf{P}_{N,\text{prior}}(j)$.

The coordinates (i_{\min}, j_{\min}) of the minimal value in the sensitivity matrix provide the subspace to be optimized (j_{\min}) and the direction to apply (i_{\min}) in order to minimize the objective function during this iteration. At the end of an iteration, the value of the structural objective function is recalculated. This iterative optimization ends when the objective function has reached a minimum (no more structural modification can decrease the objective function). The optimized parameters in \mathbf{P}_N will then be considered as invariable during the optimization of \mathbf{P}_β .

After the convergence of the objective function in the structural optimization, the uncertainties on the local directions of the geometry of the network can be estimated from the posterior structural covariance:

$$\mathbf{C}_{\mathbf{P}_N}^{\text{post}}(j) = \left(\frac{1}{8} \sum_{i=1}^8 \mathbf{S}^{\text{post}}(i, j) - \Psi_{\text{structure}}^{\text{post}} + \mathbf{C}_{\mathbf{P}_N}^{-1}(j, j) \right)^{-1} \quad (7.11)$$

where $\mathbf{C}_{\mathbf{P}_N}^{post}(\mathbf{j})$ denotes the uncertainty associated to the direction of the subspace \mathbf{j} , \mathbf{S}^{post} is the sensitivity matrix of the last iteration, and $\Psi_{structure}^{post}$ is the value of the minimized objective function associated to the last iteration. The higher the uncertainty value associated to a subspace is, the more the direction of the subspace is uncertain. In the contrary a low value denotes a well constrained direction.

7.2.3.2.2 Optimization of the property parameters

In the second step, we estimate the hydraulic properties in \mathbf{P}_β through an iterative optimization process and by considering the previously optimized \mathbf{P}_N as invariable. The process relies on a linearization of the objective function in Equation (7.9), through the computation of the Jacobian matrix. At an iteration step k , the values of the properties in \mathbf{P}_β are updated from the previous ones as follow:

$$\mathbf{P}_\beta^{k+1} = \mathbf{P}_\beta^k + \left(\left(\mathbf{J}^k \right)^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{J}^k + \mathbf{C}_{\mathbf{P}_\beta}^{-1} \right)^{-1} \cdot \left(\mathbf{J}^k \right)^T \cdot \mathbf{C}_d^{-1} \cdot \left(\mathbf{d}_{obs} - f\left(\Gamma\left(\mathbf{P}_N, \mathbf{P}_\beta^k\right)\right) \right) + \mathbf{C}_{\mathbf{P}_\beta}^{-1} \cdot \left(\mathbf{P}_{\beta,prior} - \mathbf{P}_\beta^k \right), \quad (7.12)$$

where \mathbf{J} is the $(n \times 4m_{CA})$ Jacobian matrix calculated with a finite difference method:

$$\mathbf{J}(i, j) = \left. \frac{\partial f_i}{\partial \mathbf{P}_\beta} \right|_{\mathbf{P}_\beta(j) = \mathbf{P}_\beta(j) + \Delta \mathbf{P}_\beta} \quad \text{with } \Delta \mathbf{P}_\beta \text{ a finite difference step.}$$

The value of the properties objective function is recalculated at the end of each iteration. This iterative optimization ends when the objective function converges to a minimum value.

The uncertainties on the T and S values can then be estimated through the computation of the posterior properties covariance matrix.

The square root values of the diagonal entries of this matrix represent the standard deviation associated to each property value:

$$\mathbf{C}_{\mathbf{P}_\beta}^{post} = \left(\left(\mathbf{J}^{post} \right)^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{J}^{post} + \mathbf{C}_{\mathbf{P}_\beta}^{-1} \right)^{-1}, \quad (13)$$

where $\mathbf{C}_{\mathbf{P}_\beta}^{post}(j, j)$ denotes the variance of the j^{th} property parameter in \mathbf{P}_β , and \mathbf{J}^{post} is the Jacobian matrix of the last iteration.

7.2.4 Modeling application

7.2.4.1 Modeling strategy

We have applied the CADI method on the 2D parameterized model presented in the previous section, in order to find a structured property distribution able to reproduce the measured oscillatory responses (amplitudes and phase offsets) presented in the section 7.2.2.3. We considered the oscillatory responses interpreted as negligible (amplitude < 1 mm) to be null for the inversion. We have coded the CADI algorithm with Matlab and used Comsol Multiphysics to solve the model in the frequency domain (see Equation (7.5)). This resolution was led using a finite element method on a triangular mesh, refined around the boundaries between the matrix and the conduit network represented by the equivalent porous media properties distributed over the model. This mesh refinement is performed with the Comsol *mphimage2geom* function on the conduit network image generated by the cellular automata.

The distributed model is constructed as a 40×60 m² rectangle included in a $1,000 \times 1,000$ m² buffer zone. The external boundaries of the buffer zone are built with imposed Dirichlet boundary conditions as presented in Equation (7.5). Thus, this zone permits to limit the effect of the boundary conditions on the parameterized model.

The values of the parameters chosen for the model parametrization and the inversion process are presented in Table 7.2. The initial values for the inversion were chosen accordingly to estimates from previous studies on the Terrieu site (*Jazayeri Noushabadi 2009 ; Dausse 2015 ; Wang et al. 2016*). In the inverse problem the properties values β in \mathbf{P}_β were associated to the exponent of the transmissivity and the storativity. The initial standard deviation values on the data (σ_{data}) and on the property parameters (σ_T, σ_S) are used to construct the covariance matrices as diagonal matrices: $\mathbf{C}_d = \sigma_{\text{data}}^2 \times Id(n)$ and $\mathbf{C}_{\mathbf{p}_\beta} = \sigma_{\text{prop.}}^2 \times Id(4m_{\text{CA}})$. No *a priori* information were considered for the structure local directions in $\mathbf{P}_{N,\text{prior}}$ and $\mathbf{C}_{\mathbf{p}_N}$.

The inversion was led following a multi-scale method (*Grimstadt et al. 2003*), as described in Figure 7.6. The multi-scale inversion consists in performing an inversion first for a coarse resolution of the model, and then use the inversion result as a new initial model with a higher resolution for a new inversion process.

Table 7.2: Parameters used for the inversion process.

Parameter	Value
Final partitioning	12×8
Final grid size	$\Delta x : 0.25 \text{ m} ; \Delta y : 0.25 \text{ m}$
Final network thickness	1 m
T_{init}	$10^{-1} \text{ m}^2/\text{s}$ for the conduits ; $10^{-6} \text{ m}^2/\text{s}$ for the matrix
S_{init}	10^{-8} for the conduits ; 10^{-4} for the matrix
$T_{\text{buffer}} ; S_{\text{buffer}}$	$10^{-2} \text{ m}^2/\text{s} ; 10^{-3}$
σ_{data}^2	0.01 on amplitude values 10 on phase offset values
$\sigma_{\text{prop.}}^2$	0.1 (applied on the exponent: $10^{\beta \pm \sigma_{\text{prop.}}}$)

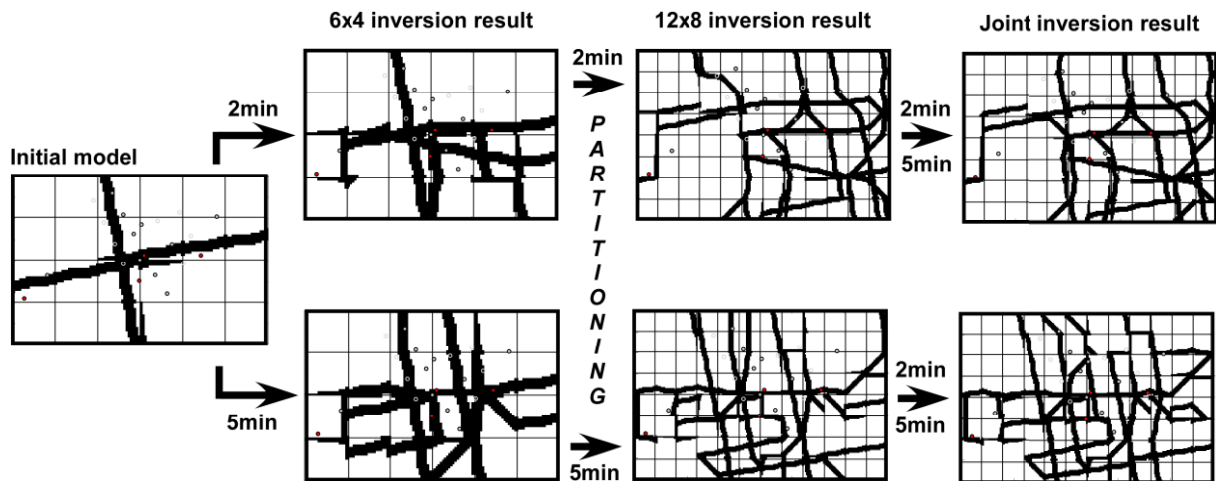


Figure 7.6: Schematization of the complete multi-scale inversion process. Starting from an initial model, firsts inversions were led for a 6×4 partitioning (shown by the grid). The results were refined to 12×8 subspaces and used for new inversions. Finally, joint inversion were led starting from the results of the previous separate inversion.

This permits to progressively reduce the size of the discretization cells for the property field during the inversion, which can be interesting for studies on heterogeneous fields with no prior information on the property distribution. This was already done with the CADI method in *Fischer et al. (2017c)*.

The initial model has been constructed with two unidirectional conduits with uniform property values (see Table 7.2) as a coarse approximation of the manual estimation made in Figure 7.4.

A first set of inversions were led by separating the 2 min and 5 min period responses, and with a 6×4 subspaces partitioning of the model (with a conduit thickness of 2 m). The results of these separate inversions have then been repartitioned in 12×8 subspaces models (with a conduit thickness of 1 m), which were used as initial models for a new inversions with the same datasets. This partitioning of the models permits to give more liberty to the inversion process, while starting from ‘not too far’ solutions, which is especially interesting for a deterministic process. The inversion results for the 12×8 models with 2 min and 5 min data separated will be presented in section 7.2.4.2.

A final inversion process has consisted in starting from these 12×8 separate results, with the same partitioning, by adding the 5 min data to the 2 min inversion result and the 2 min data to the 5 min results, for joint inversions. The results of these joint inversions will be discussed in section 7.2.4.3.

7.2.4.2 Modeling results

Figure 7.9 shows the spatial distribution of hydraulic transmissivity and storativity inverted using the responses to the 2 min and 5 min periods, respectively. The comparison between the measured and simulated hydraulic responses is presented in Figure 7.7 and Table 7.3. It can be seen that the simulated responses match the measured ones quite well.

The simulated responses to the P15 pumping appear to be slightly overestimated implying the existence of a very productive conduit in P15 that could not be simulated in the presented model. Otherwise, the responses proportions and behaviors for each pumping are well respected with the simulated signals. Overall, the Root Mean Square Error (RMSE) of the amplitude is 1.1 cm for the 2 min response signals, and 0.5 cm for the 5 min response signals. The RMSE of the phase offset is 56° for the 2 min response signals, and 66° for the 5 min response signals (see Table 3). The difficulties in reproducing the phase offset data with the CADI method may be contributed to the high degree of variation of the phase shift within the low transmissivity matrix (Figure 8).

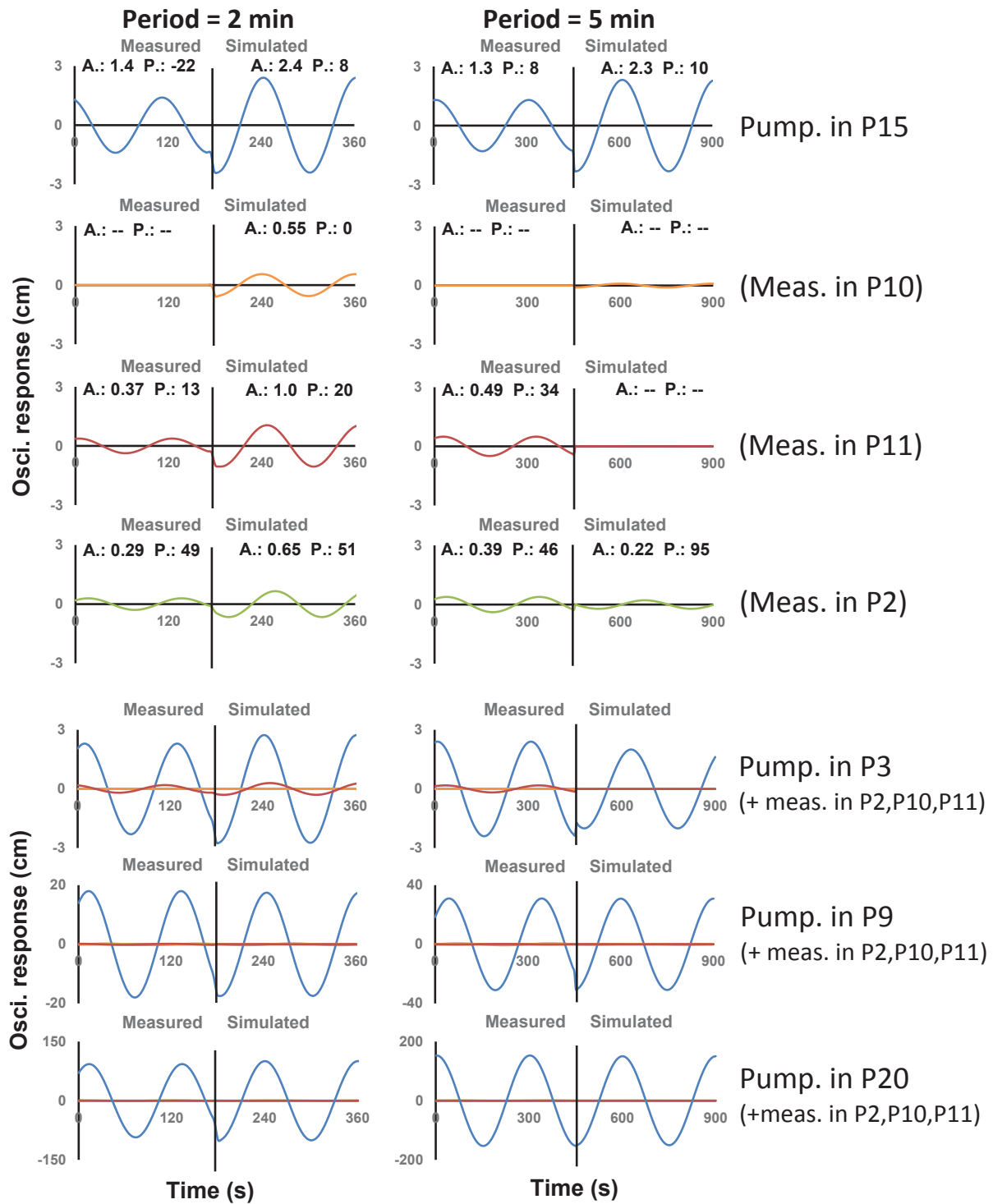


Figure 7.7: Comparison of some measured and simulated (with the property distributions presented in Figure 7.9) responses signals in observation points P2 (green), P10 (orange), P11 (red) and pumping points P3, P9, P15, P20 (each time in blue), for pumping signals with a 2 min (left) and a 5 min (right) period. In the case of the pumping in P3 we present in blue the signal in P0, located 1 m away from P3 (which was not measured during the investigation). For a better readability the responses are presented separately for a pumping in P15 with their amplitude (A. in cm) and their phase offset (P. in °) values. For the pumping in P3, P9 and P20 the responses are presented on a same graph.

A small displacement of a certain wellbore location by 1 m in the matrix can modify the phase offset by a value of 90° . Thus, our phase offset RMSEs remain under the variations produced by a 1 m displacement on the field, which is acceptable at our scale.

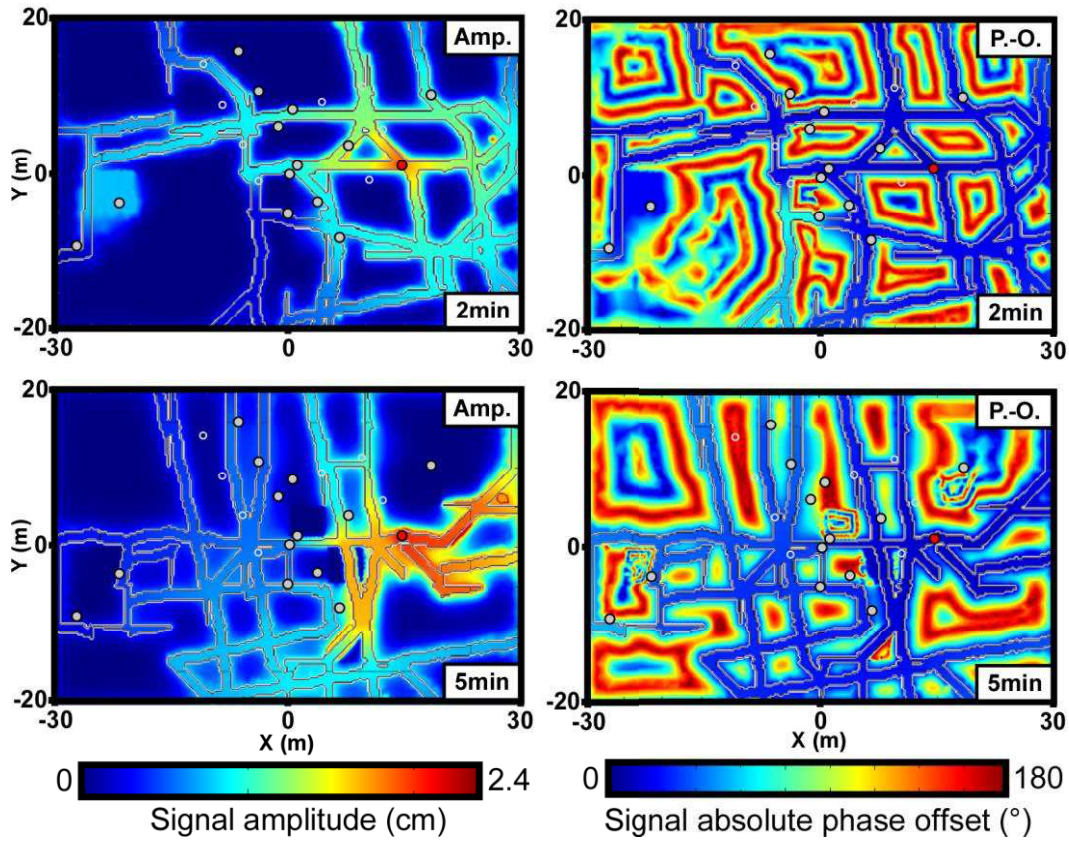


Figure 7.8: Maps of simulated spatial amplitude (Amp.) and phase offset (P.-O.) with the models in Figure 7.9 for a pumping in P15 with a signal period of 2 min and 5 min.

The amplitude of the signal is also decaying very fast in the matrix (Figure 7.8), thus the amplitude of a response is already a good information to characterize the proximity of a borehole with a conduit of the karst network. According to *Fischer et al. (2018b)* and the maps in Figure 7.8 produced with the CADI method, the amplitude in the responses signals of a karstic aquifer permit to distinguish the boreholes in (or near) conduits from the ones in the matrix. On the other side, the phase offset response permit to characterize more precisely the distance of a responding borehole to conduits, as it varies orthogonally to the direction of the conduit and stays very low within the network.

One advantage of using the CADI method in this work is that the optimized conduits networks can be clearly distinguished from the matrix in the inverted fields in Figure 7.9. If these optimized conduits networks represent only one possible geometry among other likely models, they permit to interpret the relative positioning of each borehole (in a conduit, close to a conduit, or in the matrix) and thus a degree of connectivity between them. The models of networks produced by inversions of pumping tests of different periods are very different, indicating that the both sets of responses provide different hydraulic information of the aquifer. The reconstructed network for the dataset of a period of 5 min is denser than that for a period of 2 min. The 5 min period dataset seems to carry information about karstic structures of different scales (conduits, fractures, fissures) around the measurement points, while the 2 min period dataset tends to characterize more specifically the most conductive karstic structures over the field scale. This hypothesis is supported by the maps of the amplitude in the models. Amplitude responses to a 2 min pumping signal in P15 quickly decrease around the borehole but remain visible in the coarse network over almost the entire field, while the ones to a 5 min pumping signal stay high in the dense karstic network around P15, and decrease beyond (Figure 7.8).

Concerning the property values, it appears that the reproduction of the responses required more modifications on the transmissivity values than on the storativity, especially for the conduits. Some similarities appear between the results to both periods, for example the existence of conduits near P9 with lower transmissivities, which can then be assimilated more likely to large fissures rather than conduits. Also in both results the conduits around P15 have been associated to very high transmissivities ($> 1 \text{ m}^2/\text{s}$), which tends to indicate the existence of a very productive conduit at this position. This information could already been deduced by the fact that the responses induced by a pumping in P15 in the model were slightly overestimated (Figure 7.7).

The fact that the two periods of signal lead to different solutions to the inverse problem, even by starting from a same initial model, tends to indicate that different periods of pumping signal induce different flow fields in the tested karst aquifer. In order to better understand the benefits of each period in a harmonic pumping characterization, one would need to study the results of the joint inversions, led with the responses to both periods simultaneously.

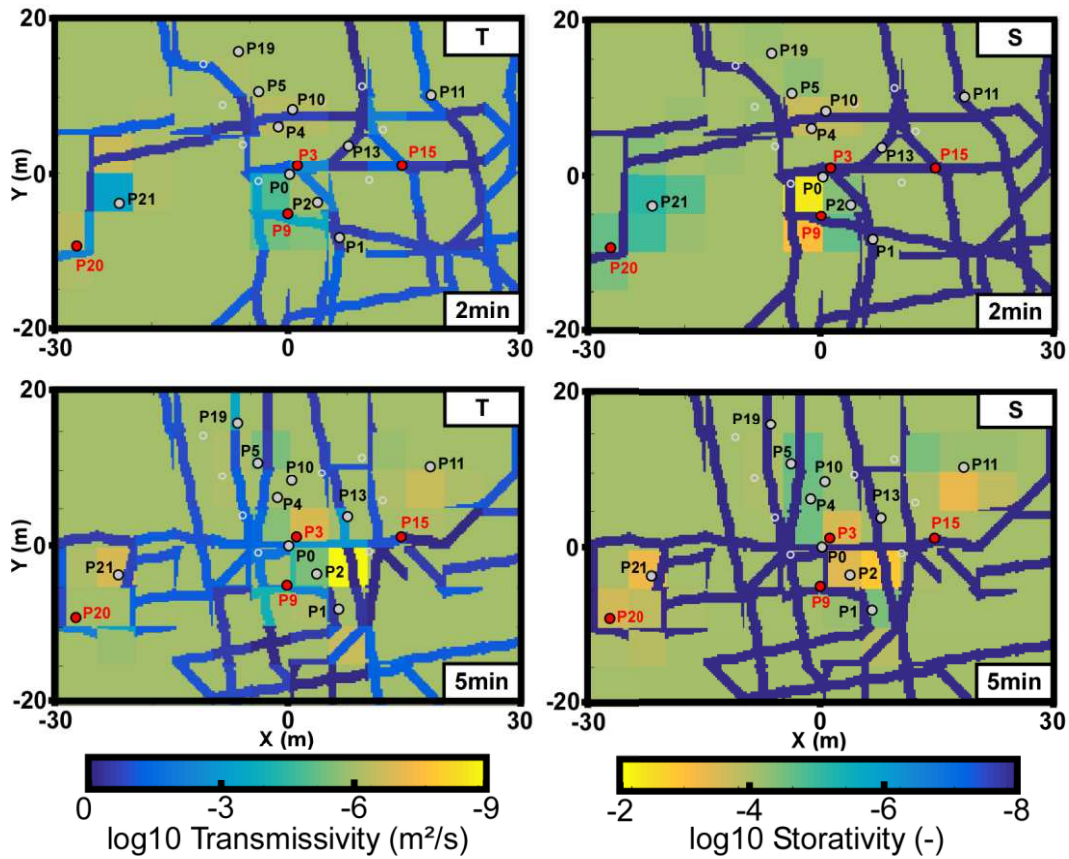


Figure 7.9: Maps of the distributions of transmissivity (T) and storativity (S) found by separate inversions of the responses to periods of 2 min and 5 min.

7.2.4.3 Effect of the period of pumping signal on the inversion results

It appears, in fact (see Table 7.3), that the separate models can badly simulate the responses associated to the signal period not used in each inversion (i.e. 5 min responses simulated through the model generated with 2 min responses, and conversely), suggesting that each set of response contains different and complementary information for the characterization of the field. Therefore, new inversions were started from the results presented in Figure 7.9 as initial models, by joining the missing responses to the ‘observed responses’ dataset in the inverse process (see the joint inversion results in Figure 7.6). For a better understanding, we will mention as ‘2 min’ and ‘5min’ separate results the model results in Figure 7.9 produced from the inversion of the responses to the 2 min and 5 min periods separately. The models produced by inversions of the joint datasets started with the ‘2 min’ model and with the ‘5 min’ model will be respectively mentioned as the ‘2 min (+5 min)’ and ‘5 min (+2 min)’ joint results. The results of the joint inversions are presented in Figure 7.10.

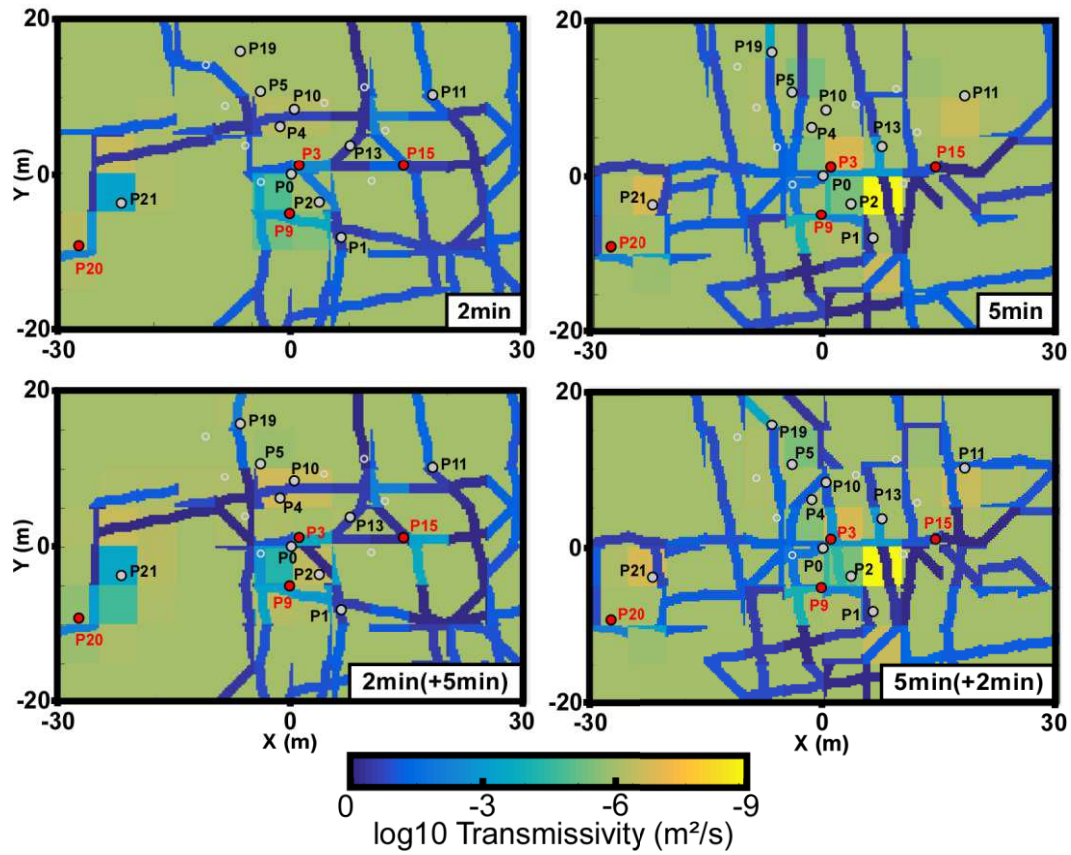


Figure 7.10: Maps of the distributions of transmissivity found by inversions of the responses to the 2 min and 5 min periods, and joint inversions started with the 2 min result (2 min (+5 min)), and with the 5 min result (5 min (+2 min)).

The ‘2 min (+5 min)’ result, solution to the inversion started from the ‘2 min’ result, is very close to its initial model. The ‘5 min (+2 min)’ result, solution to the inversion started from the ‘5 min’ result, shows some modifications on the periphery of its network (P4, P5, P10, P11). Both joint solutions, however, do not permit a reproduction of the measured signals as good as the one generated by the ‘2 min’ and ‘5 min’ separate results (see Table 7.3).

Although the phase offsets RMSEs are almost the same for the joint results and the separate results, the amplitudes RMSEs are multiplied by 5 to 10 with the joint results. This shows that, even if the amplitude responses can be well reproduced by separate models for each period, they cannot be reproduced very well with a unique model. This tends to validate the hypothesis that different periods of pumping signal induce different flow fields in the aquifer, which need to be characterized separately.

Table 7.3: RMSEs on the amplitude (Amp.) and phase offset (P.-O.) values for the different inversion results. RMSEs values in brackets represent responses that were simulated through models generated for another period of signal (i.e. 5 min responses simulated with a model generated specifically for the 2 min responses and vice versa).

Results RMSEs		'2 min'	'5 min'	'2 min (+5 min)'	'5 min (+2 min)'
Amp.	2 min	1.1 cm	(11 cm)	5.4 cm	5.1 cm
	5 min	(9 cm)	0.5 cm	5 cm	5.3 cm
P.-O.	2 min	56 °	(112 °)	60 °	60 °
	5 min	(85 °)	66 °	67 °	66 °

In fact, the CADI method is limited in its ability to represent a variation of aperture in the generated network, which can partly explain why the joint inversions are less good than the separate ones if the flows mobilized with each period occur in structures of different apertures.

However, even if the preferential flows change among the field for different pumping periods, the relative distance of each borehole toward the karstic network does not depend on the hydrodynamic but on the morphology of the karst structures, and thus their connectivity behavior should logically remain the same as the karst structure does not vary with the period of pumping. This is also what the reproduction of the phase offset values in the joint results would tend to indicate. In fact, according to Figure 7.8, if the area of propagation of the amplitude response is dependent to the density of the conduit network at that scale, it is not true for its phase offset which remains null in the conduits regardless of the network geometry.

In Table 7.4 we present the interpretation of the position of each borehole relatively to the conduits network, with the data processing estimations, the separate modeling results and the joint modeling results. First, it is interesting to point out that the estimations made manually in Figure 7.4 for the P15 pumping match for 62 % in term of position (in the conduit, close to a conduit, in the matrix) the separate modeling results '2 min' and '5 min'. Taking account that the estimation were made with only 13 responses over the 52 available, it shows that the manual interpretation method described in section 7.2.2.4 and in Fischer et al. (2018b) can already provide rather interesting and fast estimations.

Table 7.4: Positioning or connectivity response of each borehole interpreted from the qualitative estimations (Figure 7.4), the separate inversion results (Figure 7.9), and the joint inversion results (Figure 7.10).

Notation:

× : in a conduit = conduit connectivity response

⊗ : close to a conduit (< 0.5 m) = conduit connectivity response

○ : close to a conduit (< 2 m) = dual connectivity response

— : in the matrix = matrix connectivity response

	PERIOD = 2 MIN			PERIOD = 5 MIN			
	Est.	Sep.	Joint	Joint	Sep.	Est.	
P0	○	⊗	⊗	×	×	×	P0
P1	×	×	×	⊗	⊗	○	P1
P2	○	○	○	○	○	×	P2
P3	×	×	×	×	⊗	×	P3
P4	—	○	○	○	—	—	P4
P5	○	—	—	—	○	○	P5
P9	○	⊗	⊗	×	×	○	P9
P10	—	⊗	⊗	×	—	—	P10
P11	×	×	×	×	—	×	P11
P13	×	×	×	×	×	○	P13
P15	×	×	×	×	×	×	P15
P19	○	—	×	×	×	×	P19
P20	○	○	○	○	○	×	P20
P21	○	—	—	⊗	⊗	○	P21

According to Figure 7.8, the amplitude and phase offset responses of an observable point very close to a conduit (< 0.5 m) are almost undistinguishable to the ones directly in a conduit. Therefore we consider these points to also have a conduit connectivity response in Table 7.4. Figure 7.11 maps represent the interpreted connectivity of each borehole obtained from the joint inversion results. This figure also shows that these results reproduce the schema of connectivity of the preferential flow path established in *Jazayeri Noushabadi (2009)* and *Dausse (2015)* between P2, P11, P15, and P20 (see Figure 7.1).

Taking into account the connectivity response, 57 % of the boreholes show a similar behavior between the ‘2 min’ and the ‘5 min’ separate results. This degree of similarity increases to 93 % when comparing the boreholes connectivity responses from the two joint results. The only behavior difference between the two results comes from P21, which appears as connected through the matrix in the ‘2 min (+5 min)’ result and as connected through conduits in the ‘5 min (+2 min)’ result. However the property distribution of the ‘2 min (+5 min)’ result in Figure 7.10 shows that its matrix transmissivity near the P21 point has a value close to a conduit transmissivity.

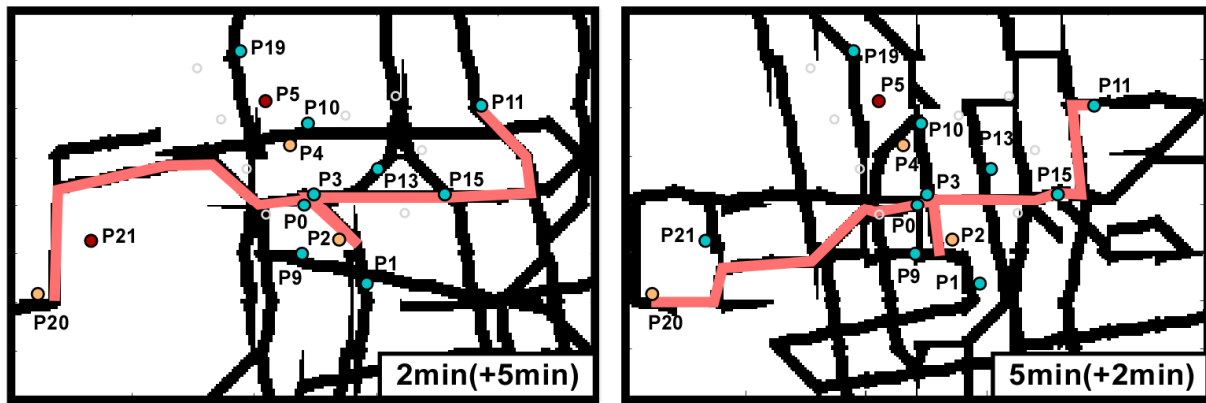


Figure 7.11: Maps of the connectivity responses associated to each borehole from the networks (shown in background in black) inverted with the joints inversions. Boreholes in blue are associated to a conduit connectivity, in orange to a dual connectivity, and in red to a matrix connectivity. The red lines show flow paths in the models which show a same connectivity as the field preferential flow path highlighted in Jazayeri Noushabadi (2009) and Dausse (2015).

This permits to induce a conduit connectivity behavior for P21, even if the conduit is distant from the borehole. Therefore, we can consider that P21 should have a ‘conduit connection behavior’. The fact that the degree of similarity of the boreholes behavior has increased to almost 100 % for two joint result clearly shows that the reproduction of the whole responses dataset requires to delineate a unique connectivity relation between each borehole, even if it is not sufficient to reproduce the amplitudes of the measured responses.

The remaining question is: which period of signal contains this information, as the separate results show only a 57 % similarity ? The conduits network almost didn’t change during the joint inversion started from the ‘2 min’ result, and in fact, the separate ‘2 min’ result and the joint ‘2 min (+5 min)’ result show a 93 % similarity in their borehole behaviors, while the ‘5 min’ result has a degree of similarity of 71 % with ‘5 min (+2 min)’. Then, most of the borehole behaviors found in the joint results were already present in the result of the 2 min signal period, which indicates that the 2 min responses contain the most information about the connectivity of each borehole.

If the two sets of amplitude responses are not well reproduced by the joint inversion while they both delineate a similar type of response for each borehole, it shows that the characterized property distribution (*in extenso* the induced flow paths network) is different for each period. In fact, the responses to a 5 min period require a dense flow network to be reproduced, while the responses to a 2 min period require a much more dispersed network. The structural posterior uncertainty maps in Figure 7.12 indicate that the network of the ‘2 min’

result is very uncertain compared to the network of the '5 min' result. This indicates that the responses to a 5 min period contain more information on the localization of the flow paths around the measurement points than the responses to a 2 min period. These last ones seem to provide less precise information on the localization of the flows between the boreholes among the field. When inverting jointly the responses from the two period, the networks in the joint results show overall lower uncertainties. Globally, the study of the structural uncertainties tends to indicate that, while, as seen before, responses to a the lower period contain more information on the type of connectivity of each borehole, the responses to the higher period contain more information on the position of the preferential flow paths around the boreholes. However, the fact that these both sets of responses cannot be well reproduced simultaneously also indicates that the flows highlighted in the '5 min' result do not exist with a period of 2 min. Therefore, while the lower period essentially mobilizes water from the most conductive karstic structures among the field, the dense flow field highlighted by the higher period can be assimilated to a mobilization of water also in less conductive karstic structures.

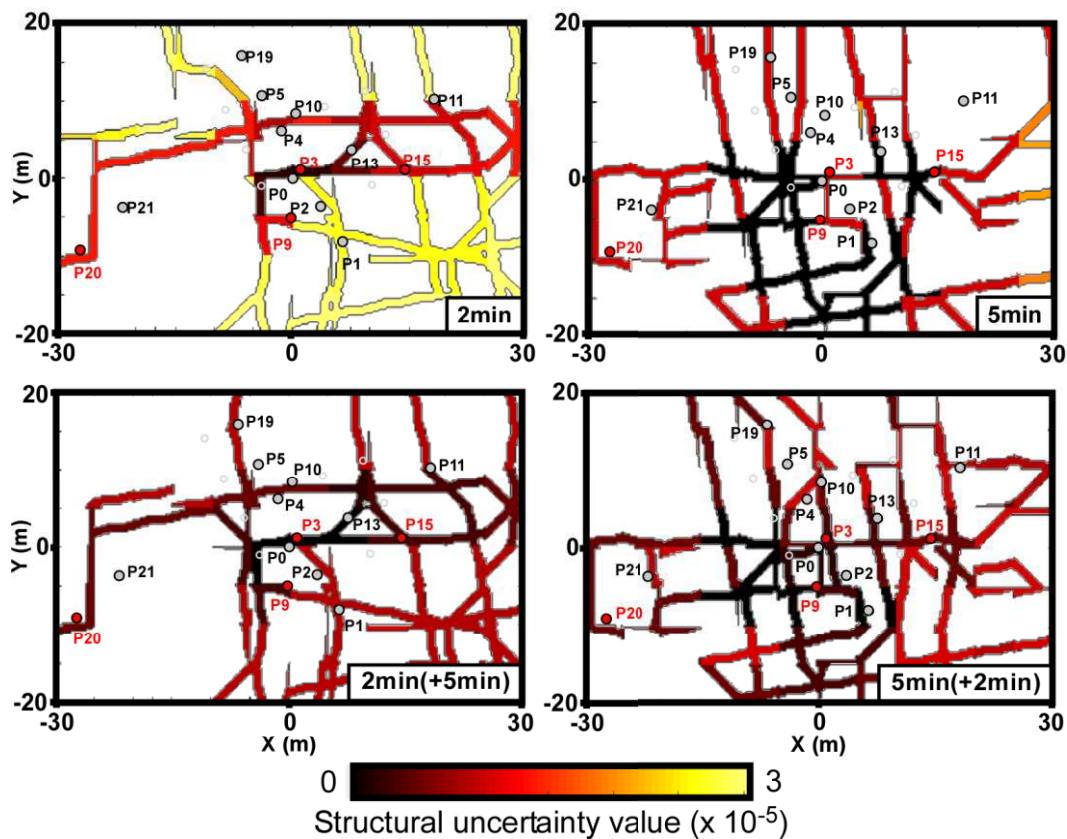


Figure 7.12: Structural uncertainty values from the results found for separate inversions of the 2 min and 5 min responses, and joint inversions started with the 2 min result (2 min (+5 min)), and with the 5 min result (5 min (+2 min)).

The maps of the posterior uncertainties on the transmissivity values in Figure 7.13 tend to show that the responses to a 2 min period give more information on the transmissivity of conduits locally around the pumping points (average ± 0.1 on the transmissivity exponent) than the responses to a 5 min period (± 0.2 on the transmissivity exponent). The uncertainties on the storativity values remain high in the whole conduit network and for both periods (± 0.2 on the storativity exponent). Further from the pumping boreholes, both property values in the conduits network (transmissivity and storativity) globally remain uncertain (± 0.3 on the exponent), even in the joint results. The transmissivity and storativity values in the matrix are more constrained around the measurement points in the matrix. These information indicate that, for both periods, the characterization of the property values with oscillatory responses remains local, around the boreholes. Oscillatory responses provide more information on the global connectivity and the localization of preferential flows rather than on the property values of the matrix and the conduits and fissures in the aquifer.

Overall, regarding our results, it appears that the flow paths generated by periods of 2 min and a 5 min in the pumping signals in this karstic field are very different. It appears in fact that, at our field scale, higher frequency signals (here a period of 2 min) activate principally the most conductive flow paths over the field, mostly located in the conduits network. These frequencies permit to better characterize the distance between each borehole and the most conductive karstic structures. Therefore, they allow a better interpretation of a degree of connectivity between boreholes, through the network of karstic conduits. Lower frequency signals (here a period of 5 min) activate, at our field scale, both conductive and also less conductive structures. Therefore, these frequencies permit to better characterize the existence and the localization of networks of fractures and fissures around the boreholes.

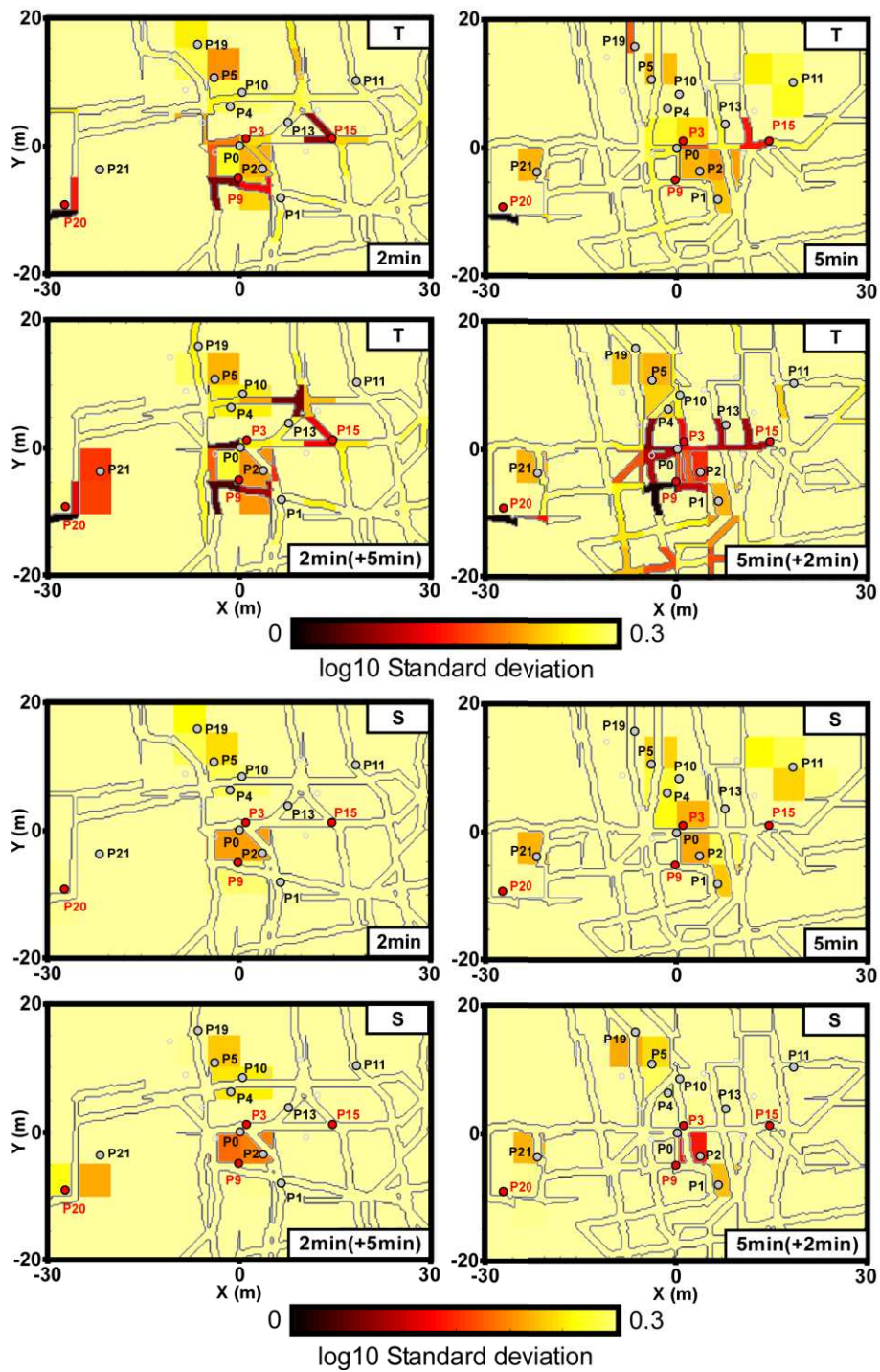


Figure 7.13: Transmissivity (T) and storativity (S) standard deviation values of the results found for separate inversions of the 2 min and 5 min responses, and joint inversions started with the 2 min result (2 min (+5 min)), and with the 5 min result (5 min (+2 min)).

7.2.5 Discussion

In *Fischer et al. (2018b)*, the authors have described a qualitative method for interpreting inter-well connectivity from the responses to harmonic pumping tests in karstic aquifers, by categorizing the extracted oscillatory responses in three types (conduit, dual, matrix connectivity). In this work the method is further developed through the integration of a quantitative interpretation with an inverse algorithm, the CADI method, that can handle a large number of measured data simultaneously and generate complex distributions of properties. The integrated approach permits to produce spatial distributions of amplitude and phase offset responses consistent with those studied by *Fischer et al. (2018b)*.

The comparison of inverted conduit networks from periods of 2 min and 5 min indicates different pumping frequency generate different flow fields. A higher frequency will permit to better characterize the flows in highly conductive structures, and the conduit connectivity at a field scale. This finding is consistent with the work of *Rabinovich et al. (2015)* where the authors show that the flow paths tend to follow the most conductive media especially at lower period. On the contrary, a lower frequency will permit to better characterize the flows in less conductive structures, and thus the localization of networks of smaller conduits, fractures and fissures. Each frequency of signal permits to generate responses holding different and complementary information on karst structures. There is therefore, according to our results, no ‘best’ choice of period for the characterization of a karstic field. This choice should be made accordingly to what structures one would most likely characterize. The important impact of the period of the pumping signal on the ratio of conduit/matrix flows has already been highlighted with a simplified study case in *Fischer et al. (2018b)*, but our new work also shows a more complex role of the different structures (conduit, fracture, fissure, matrix) on the generation of different flow fields with the different frequencies of pumping signal.

To summarize, if at a regional scale one would imagine that the difference between a lower and a higher frequency pumping would principally concern the zone of influence in the aquifer (with a larger zone for a lower frequency), this difference implies more specific behaviors at a smaller site scale. In fact, at this scale, higher frequencies mobilize water essentially in the most conductive structures, while lower frequencies permit to reach also a mobilization of water in the less conductive structures and media. On another hand oscillatory responses do not provide precise information on the conductivity and specific storage values in the conduits and the matrix very far from the pumping point. However the same observation was made on the same field for the inversion of steady-state responses to eight pumping tests

at constant rates in *Fischer et al. (2017c)*. This previous article also showed that steady state data to constant rate pumping (comparable to an infinitely high oscillatory period) were more sensitive to fracture flows and required a dense inverted network to be reproduced, which is in agreement with our observation for a higher period of signal in this work. Furthermore, the uncertainty analysis from this previous article indicated that the constant rate data, as the data for a period of 5 min in the pumping signal in this work, permitted to better characterize the flow structures in areas where we had measurement points. The inverted networks in the results from our new work and the ones in *Fischer et al. (2017c)* both reproduce the established connectivity of the Terrieu field (*Jazayeri Noushabadi 2009 ; Dausse 2015*), but in the case of the constant rate data, the inversions were led with responses to an investigation of eight pumping points instead of four with the harmonic investigation.

The results of our new work, associated to the previous ones in *Fischer et al. (2017c)*, tend to indicate that the steady state responses to a constant pumping rate would blur the connectivity associated to the most conductive conduits among the field, while better characterizing the karstic structures of different scales in the areas around the pumping wells. Therefore it would require several well distributed pumping points in order to characterize the whole karstic network. Responses to an oscillatory pumping rate, on contrary, allow an already good characterization of the karstic connectivity from a unique pumping in a borehole in the karstic network (as P15 in this work).

7.2.6 Conclusion

In this work we have extended the qualitative method presented in *Fischer et al. (2018b)*, for interpreting a karstic network connectivity from the hydraulic responses to harmonic pumping tests, to a quantitative analysis by combining these responses with an inversion algorithm. The integrated approach is able to deal with a large set of data simultaneously and to construct structurally contrasted distributions of hydraulic properties conditioned to the measured tomographic harmonic pumping responses.

Our results show that tomographic harmonic pumping tests performed with different signal frequencies led to a characterization of different structures of the karstic network. Higher frequency signals tend to assist in interpreting a degree of connectivity between each borehole of the field and the most conductive structures, while lower frequency signals are more useful in the localization of less conductive features, such as small fractures and fissures.

The CADI method, as imaging tool, shows limitations in its ability to represent complex structures of different aperture simultaneously, as already noticed in *Fischer et al. (2017c)*, which can partly explain the less good results of the joint inversions compared to the separate inversions in this work. However we believe that the combination of the CADI method with tomographic harmonic pumping tests appears as a promising methodology for a quantitative characterization of the hydraulic properties and the hydraulic connectivity in karstic aquifers.

8 CONCLUSION GÉNÉRALE ET PERSPECTIVES

Des nouvelles méthodes et outils de tomographies hydrauliques sont présentés et discutés dans ce manuscrit, afin de mieux caractériser qualitativement et quantitativement les écoulements dans des aquifères fracturés et karstiques.

Dans la phase d'acquisition de données de terrain, l'utilisation de signaux harmoniques, notamment via des pompages, permet de mieux comprendre les chemins d'écoulements formés par les différentes structures karstiques (conduits, fissures, fractures et matrice). Afin d'interpréter quantitativement les réponses à un (ou des) pompage(s) harmonique(s) en milieu fracturé karstique, il est nécessaire de recourir à une méthode d'inversion adaptée à ces milieux. Les méthodes d'inversions développées dans ce manuscrit, CADI et DNDI, répondent à cette problématique en proposant une interprétation des données de terrain sous forme d'une imagerie des champs de propriétés. La représentation des structures karstiques dans ces méthodes d'imagerie se fait par une génération dynamique et optimisable de réseaux, produite par des automates cellulaires (CADI) ou avec des éléments discrets (DNDI).

Les applications de la méthode CADI sur des données d'investigations sur le site fracturé et karstique du Terrieu (près de Montpellier, en France), permet d'apporter des résultats de caractérisation intéressants, permettant de mieux localiser les chemins d'écoulements et les connectivités qu'ils génèrent. La Figure 8.1 présente un comparatif de différents résultats d'imagerie sur le même site du Terrieu, à partir de différentes méthodes d'inversions et deux jeux de données (investigation de 2009 avec pompages à débits constants et inversion en domaine permanent, et investigation de 2017 avec pompages harmoniques et inversion en domaine fréquentiel).

Par rapport à la méthode CADI, les méthodes utilisées dans *Wang et al. (2016)* et *Wang et al. (2017)* permettent plus de liberté de variations dans la représentation des champs de propriétés, ce qui permet d'avoir une meilleure reproduction des réponses, mais néanmoins au détriment de la lisibilité et de la possibilité d'interprétation des résultats telles qu'offertes par la méthode CADI. Les résultats obtenus sur un même jeu de données (pompage à débit constant) montrent en effet que la méthode CADI met plus en avant les contrastes existants entre les structures karstiques et la matrice, et donc localise plus clairement ces structures.

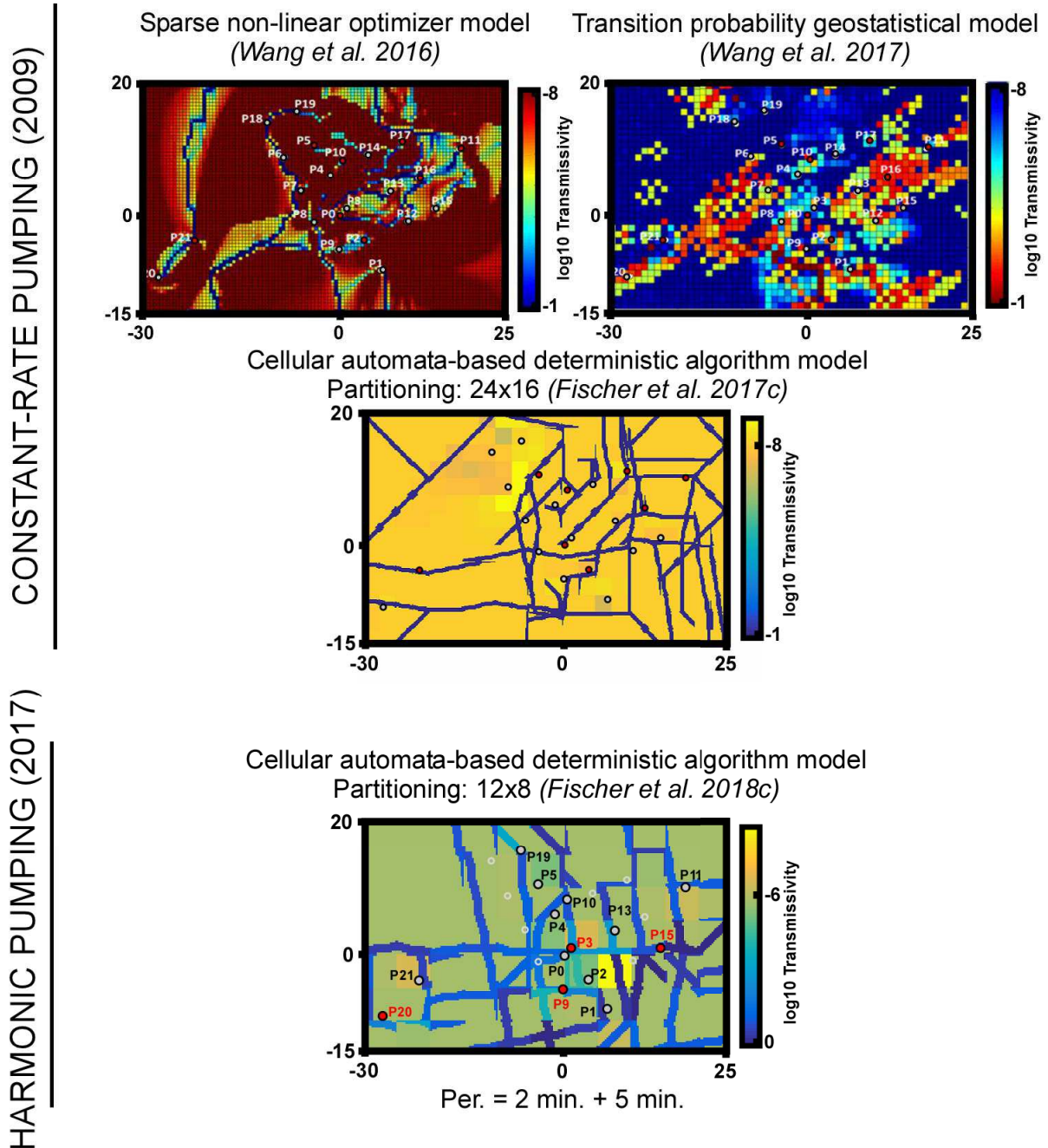


Figure 8.1: Résultats d'imageries des transmissivités obtenus sur le site du Terrieu à partir d'un jeu de réponses en domaine permanent de pompages à débits constants (2009), et d'un jeu de réponses en domaine fréquentiel de pompages à débits harmoniques (2017) avec deux périodes (Per.) différentes. Ces résultats ont été obtenus à partir de différentes méthodes d'inversions, notamment pour les données de pompages à débits constants.

Les résultats d'applications de la méthode CADI dans les chapitres 5 et 7 peuvent même être validés par des informations de connectivité ou des observations de terrain, par simple lecture des images des champs de propriétés produits.

Néanmoins, comme souligné dans les chapitres 3, 4 et 5, les méthodes CADI et DNDI, présentées dans ce manuscrit, possèdent des limites qu'il convient de connaître afin de pouvoir bien les utiliser. Les principaux avantages et limites de ces deux méthodes sont présentés dans le Tableau 8.1. Certaines de ces limites sont directement inhérentes au principe de la tomographie telles que la nécessité de collecter des données sur des forages spatialement bien répartis sur le site, ou la dépendance au choix du modèle initial dû au processus déterministe (comme discuté dans le chapitre 4). D'autres limites sont liées aux méthodes elles-mêmes, telle que le biais créé par le processus d'optimisation séquentielle (qui peut néanmoins être contrebalancé par un couplage à une méthode d'inversion multi-échelle, comme décrit dans les chapitres 4 et 5).

Tableau 8.1: Principaux avantages et limites des méthodes CADI et DNDI.

	Méthode CADI	Méthode DNDI
AVANTAGES	Représentation plus contrastée des structures karstiques	
	Réseau dynamique de structures optimisables	
	Flexibilité des automates cellulaires	Optimisation de l'ouverture des conduits et fractures possible
	Résolution rapide des modèles équivalents milieux poreux	Possibilité d'introduire une physique spécifique au réseau de structures karstiques
LIMITES		Sensibilité au modèle initial
		Optimisation séquentielle
	Pas d'optimisation de l'ouverture des structures	Moins de flexibilité de génération que les automates cellulaires
	Pas de physique particulière pour le réseau karstique (équivalent milieu poreux)	Résolution plus exigeante des réseaux discrets

Les méthodes CADI et DNDI apparaissent comme complémentaires vis-à-vis de leurs avantages et limites, c'est-à-dire que les limites de l'une des méthodes représentent les avantages apportés par l'autre. L'utilisation simultanée de ces deux méthodes pour une même étude peut donc apparaître comme intéressante afin d'en comparer les résultats. Néanmoins ceci n'a pas été effectué dans ce manuscrit et demande encore à être testé dans le futur.

La comparaison des résultats d'inversions produites par la méthode CADI entre les deux jeux de données (pompage à débit constant et pompage harmonique) permet de discuter des informations portées par chacun des signaux de sollicitation de nappe. Avant tout il est important de noter que les résultats d'inversions présentés en Figure 8.1 proviennent de modélisation en domaines non temporels, c'est-à-dire en régime quasi-statique pour les données de pompages à débits constants et en régime fréquentiel pour les pompages harmoniques. Il est donc possible que les informations contenues dans les réponses à un pompage à débit constant soient semblables à celles contenues dans les réponses à un pompage harmonique si on en effectue une analyse temporelle. Néanmoins une analyse quantitative temporelle exigerait un temps de résolution plus long dans la modélisation, et donc un temps de calcul global beaucoup plus long pour l'inversion ; ce qui rend l'analyse fréquentielle des réponses aux pompages harmoniques plus intéressante. Ensuite, il est intéressant d'observer également qu'un signal constant est comparable à un signal harmonique à période infinie (fréquence extrêmement faible), ce qui fait du pompage à débit constant un cas particulier des pompages harmoniques. Les résultats obtenus dans ce manuscrit tendent à montrer que les informations contenues dans les réponses à des pompages, à signal harmonique de différentes fréquences et à signal constant en régime quasi-statique, sont différentes et spécifiques.

Le rôle de la fréquence du signal oscillatoire de pompage sur les informations contenues dans les réponses mesurées en aquifère fracturé et karstique est déjà souligné sur l'étude théorique menée en chapitre 5, et étudié plus en détails sur une application sur le site du Terrieu en chapitre 6. L'utilisation d'un signal oscillatoire à haute fréquence permet en effet de mobiliser les écoulements plus spécifiquement dans les conduits karstiques les plus conducteurs, alors qu'un signal oscillatoire à basse fréquence va également entraîner des écoulements dans des structures moins conductrices, telles que des conduits moins importants et des fissures. Ce type d'investigation permet de mettre en évidence et de hiérarchiser les chemins d'écoulements des plus importants aux moins importants. Les signaux à fréquence plus élevée permettraient ainsi de mieux caractériser les degrés de connexion de chaque point de mesure au réseau de conduits principaux, alors que les fréquences plus basses fourniraient plus d'informations sur la localisation des chemins d'écoulements formés par l'ensemble des structures karstiques (conduits, fractures, fissures). La Figure 8.2 montre ainsi les informations interprétées de chacune des fréquences de signal utilisées sur le site du Terrieu.

HIGHER FREQUENCY INFORMATION :

Borehole connectivity to the main conduits network

- In or close to (< 50 cm) the conduits network
- 1 or 2 m away from the conduits network
- More than 2 m away from the conduits network

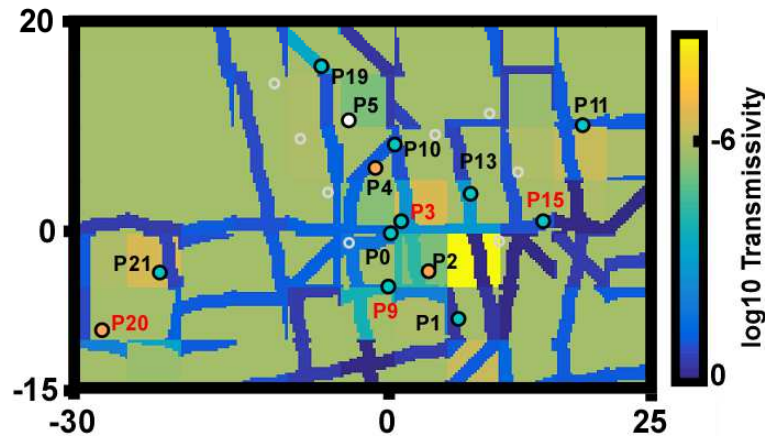
LOWER FREQUENCY INFORMATION :Localization of the karstic structures
(conduit, fissures, fractures,...)

Figure 8.2: Interprétation des résultats de tomographie harmonique sur le site du Terrieu en terme de connectivité des forages au réseau de conduits principaux grâce aux pompages à « hautes fréquences », et en terme de localisation des structures karstiques (conduits, fractures, fissures) grâce aux pompages à « basses fréquences ».

Il est d'ailleurs intéressant d'observer que, sur le site du Terrieu, les chemins d'écoulements interprétés à partir des pompages à débits constants sont assez proches de ceux interprétés à partir des pompages harmoniques à la période la plus élevée (5 min). Ceci montre que le pompage à débit constant apporte une information plus proche d'un signal oscillatoire à période élevée (donc fréquence plus faible), comme souligné plus haut. Néanmoins les signaux à fréquence plus faible, qui mobilisent des écoulements dans des structures karstiques plus ou moins conductrices, ont tendance à « cacher » les chemins d'écoulements les plus importants, liés aux plus gros conduits. Si l'on s'intéresse plus spécifiquement à ces derniers, il convient donc plutôt d'utiliser une fréquence plus élevée.

Les méthodes employées dans l'approche de tomographie harmonique proposée dans ce manuscrit ouvrent des possibilités d'exploration et d'approfondissement dans d'autres domaines. L'approche elle-même demande encore à être améliorée.

En effet, les applications de tomographies présentées dans ce manuscrit traitent exclusivement de données hydrauliques, notamment de niveaux piézométriques. Le site du Terrieu se présente comme un site idéal pour l'étude des tomographies hydrauliques, puisqu'il comporte de nombreux forages, répartis idéalement sur une surface relativement petite (2500 m² environ). Néanmoins l'approche de tomographie harmonique pour milieux fracturés et karstiques telle que présentée dans ce manuscrit s'exporte difficilement sur des sites moins bien pourvus en forages, si l'on souhaite recueillir des données hydrauliques essentiellement. Une solution plus adaptée pour des sites moins équipés serait de coupler cette approche à des systèmes de mesures en surface, notamment de géophysique. Certaines techniques de mesures géophysiques de surface ont en effet montré leur efficacité pour capter les réponses à des signaux oscillatoires dans des aquifères, notamment le potentiel spontané (*Revil et al. 2008 ; Maineult et al. 2008 ; Soueid Ahmed et al. 2016*), ou l'inclinométrie (*Schuite et al. 2017*).

Concernant les techniques de pompage à signal harmonique telles qu'utilisées dans l'approche proposée dans ce manuscrit, elles apportent des avantages intéressants (modélisation en régime fréquentiel, hiérarchisation des écoulements en faisant varier la fréquence de signal). Elles apportent, notamment dans le milieu karstique étudié, la possibilité de caractériser différents processus d'écoulements. Des signaux harmoniques sont également générés naturellement, dans la marée par exemple, et permettent de caractériser un aquifère sans avoir à générer un signal artificiellement. Cet apport pourrait être étudié et développé plus spécifiquement dans d'autres domaines de l'hydrogéologie. Il pourrait être intéressant d'étudier l'apport des pompes harmoniques pour, par exemple, la détection d'une possible continuité hydraulique entre deux réservoirs distincts (*Sun et al. 2015*), la caractérisation des zones non saturées (*Revil et al. 2008*), la caractérisation des aquifères côtiers (*Alcolea et al. 2007 ; Jardani et al. 2012*), ou encore l'interprétation du positionnement et des propriétés de structures anthropiques enterrées. Il peut être également intéressant d'étudier plus en détails les informations contenues plus spécifiquement dans l'amplitude et dans le déphasage des réponses aux signaux oscillatoires pour chaque type d'aquifère existant (alluvial, hétérogène, karstique, etc.). Toutes ces possibilités représentent autant de développements scientifiques possibles dans le futur, à partir des informations apportées par ce manuscrit et des autres travaux déjà effectués dans ces divers domaines.

BIBLIOGRAPHIE (REFERENCES)

A

Abusaada, M. and M. Sauter. 2013. Studying the flow dynamics of a karst aquifer system with an equivalent porous medium model. *Groundwater* 51: 641-650

Ackerer, P. and F. Delay. 2010. Inversion of a set of well-test interferences in a fractured limestone aquifer by using an automatic downscaling parameterization technique. *Journal of Hydrology* 389: 42-56.

Agences et offices de l'Eau, Onema, Ministère en charge de l'environnement [En ligne]. 2013 [dernière mise à jour : 02/11/2015]. DCE - État des eaux.
<http://www.statistiques.developpement-durable.gouv.fr/lessentiel/ar/240/0/dce-etat-eaux.html>

Alcolea, A., E. Castro, M. Barbieri, J. Carrera, S. Bea. 2007. Inverse Modeling of Coastal Aquifers Using Tidal Response and Hydraulic Tests. *Groundwater* 45: 711-722.

Ambikasaran, S., J.Y. Li, P.K. Kitanidis, E. Darve. 2013. Large-scale stochastic linear inversion using hierarchical matrices. *Computers & Geosciences* 17: 913-927.

Arfib, B. and J.B. Charlier. 2016. Insights into saline intrusion and freshwater resources in coastal karstic aquifers using a lumped Rainfall-Discharge-Salinity model (the Port-Miou brackish spring, SE France). *Journal of Hydrology* 540: 148-161.

B

Bakhos, T., M. Cardiff, W. Barrash, P.K. Kitanidis. 2014. Data processing for oscillatory pumping tests. *Journal of Hydrology* 511: 310-319.

Bechtel, T., F. Bosch, M. Gurk. 2007. Geophysical methods in karst hydrogeology. In: Goldscheider, N. and D. Drew. *Methods in Karst Hydrogeology*. Taylor & Francis. London, U.K.

Berg, S.J. and W.A. Illman. 2013. Field study of subsurface heterogeneity with steady-state hydraulic tomography. *Groundwater* 51: 29-40.

Black, J.H., and K.L. Kipp Jr. 1981. Determination of hydrogeological parameters using sinusoidal pressure tests: A theoretical appraisal. *Water Resources Research* 17: 686-692.

Bohling, G.C., X. Zhan, J.J. Butler Jr., L. Zheng. 2002. Steady shape analysis of tomographic pumping tests for characterization of aquifer heterogeneities. *Water Resources Research* 38. doi: 10.1029/2001WR001176.

Bonneau, F., V. Henrion, G. Caumon, P. Renard, J. Sausse. 2013. A methodology for pseudo-genetic stochastic modeling of discrete fracture networks. *Computer & Geosciences* 56: 12-22.

Borghi, A., P. Renard, F. Cornaton. 2016. Can one identify karst conduit networks geometry and properties from hydraulic and tracer test data? *Advances in Water Resources* 90: 99-115.

BRGM [En ligne]. 2015 [dernière mise à jour : 08/10/2015]. Le référentiel hydrogéologique français BDLISA.
<http://www.brgm.fr/projet/referentiel-hydrogeologique-francais-bdlisa>

Bruna, P.O., Y. Guglielmi, S. Viseur, J. Lamarche, O. Bildstein. 2015. Coupling fracture facies with in-situ permeability measurements to generate stochastic simulations of tight carbonate aquifer properties: Example from the lower cretaceous aquifer, Northern Provence, SE France. *Journal of Hydrology* 529: 737-753.

Butler, J.J. 2005. Hydrogeological methods for estimation of spatial variations in hydraulic conductivity. In: Rubin, Y. and S.S. Hubbard. Hydrogeophysics. *Water Science and Technology Library* (vol. 50). Springer, Dordrecht.

C

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

Cardiff, M. and P.K. Kitanidis. 2008. Efficient solution of nonlinear, underdetermined inverse problems with a generalized PDE model. *Computers & Geosciences* 34: 1480-1491.

Cardiff, M. and P.K. Kitanidis. 2009. Bayesian inversion for facies detection: An extensible level set framework. *Water Resources Research* 45. doi: 10.1029/2008WR007675.

Cardiff, M., W. Barrash, P.K. Kitanidis, B. Malama, A. Revil, S. Straface, E. Rizzo. 2009. A potential-based inversion of unconfined steady-state hydraulic tomography, *Ground Water* 47: 259-270.

Cardiff, M., W. Barrash, P.K. Kitanidis. 2013a. Hydraulic conductivity imaging from 3-D transient hydraulic tomography at several pumping/observation densities. *Water Resources Research* 49: 7311-7326.

Cardiff, M., T. Bakhos, P.K. Kitanidis, W. Barrash. 2013b. Aquifer heterogeneity characterization with oscillatory pumping: Sensitivity analysis and imaging potential. *Water Resources Research* 49: 5395-5410.

Cardiff, M. and W. Barrash. 2015. Analytical and semi-analytical tools for the design of oscillatory pumping tests. *Groundwater* 53: 896-907.

Carrera, J. and S.P. Neuman. 1986a. Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information. *Water Resources Research* 22: 199-210.

Carrera, J. and S.P. Neuman. 1986b. Estimation of aquifer parameters under transient and steady state conditions: 3. Application to synthetic and field data. *Water Resources Research* 22: 228-242.

Carrera, J., A. Alcolea, A. Medina, J. Hidalgo, L.J. Slooten. 2005. Inverse problem in hydrogeology. *Hydrogeology Journal* 13: 206-222.

Castagna, M., M.W. Becker, A. Bellin. 2011. Joint estimation of transmissivity and storativity in a bedrock fracture. *Water Resources Research* 47. doi: 10.1029/2010WR009262.

Chopard, B. and A. Masselot. 1999. Cellular automata and lattice Boltzmann methods: a new approach to computational fluid dynamics and particle transport. *Future Generation Computer Systems* 16: 249-257.

Cliffe, K., D. Holton, P. Houston, C. Jackson, S. Joyce, A. Milne. 2011. Conditioning discrete fracture network models of groundwater flow. *International Journal of Numerical Analysis and Modeling* 8: 543-565.

Collon, P., D. Bernasconi, C. Vuilleumier, P. Renard. 2017. Statistical metrics for the characterization of karst network geometry and topology. *Geomorphology* 283: 122-142.

D

Dagan, G., and A. Rabinovich. 2014. Oscillatory pumping wells in phreatic, compressible, and homogeneous aquifers. *Water Resources Research* 50: 7058-7066.

Dausse, A. 2015. Facteurs d'échelle dans la hiérarchisation des écoulements au sein d'un aquifère karstique : Analyse multi-échelles des propriétés hydrodynamiques et de transport de l'aquifère de Lez. PhD Thesis, Université de Montpellier. French.

De Marsily, G., A.M. LaVenue, B.S. RamaRao, M.G. Marietta. 1995. Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields: 2.Application. *Water Resources Research* 31 : 495-516.

De Rooij, R., P. Perrochet, W. Graham. 2013. From rainfall to spring discharge: coupling conduit flow, subsurface matrix flow and surface flow in karst systems using a discrete-continuum model. *Advances in Water Resources* 61: 29-41.

Delay, F., A. Kaczmaryk, P. Ackerer. 2007. Inversion of interference hydraulic pumping tests in both homogeneous and fractal dual media. *Advances in Water Resources* 30 : 314-334.

Dewri, R. and N. Chakraborti. 2005. Simulating recrystallization through cellular automata and genetic algorithms. *Modelling and Simulation in Materials Science and Engineering* 13: 173-183.

Dreiss, S.J. 1982. Linear Kernels for Karst Aquifers. *Water Resources Research* 18: 865-876.

E

Eisenlohr, L., L. Kiraly, M. Bouzelboudjen, Y. Rossier. 1997. Numerical simulation as a tool for checking the interpretation of karst spring hydrographs. *Journal of Hydrology* 193: 306-315.

Elsheikh, A.H., M.F. Wheeler, I. Hoteit. 2014. Hybrid nested sampling algorithm for Bayesian model selection applied to inverse subsurface flow problems. *Journal of Computational Physics* 258: 319-337.

F

Fischer, P., A. Jardani, A. Soueid Ahmed, M. Abbas, X. Wang, H. Jourde, N. Lecoq. 2017a. Application of large-scale inversion algorithms to hydraulic tomography in an alluvial aquifer. *Groundwater* 55: 208-218.

Fischer, P., A. Jardani, N. Lecoq. 2017b. A cellular automata-based deterministic inversion algorithm for the characterization of linear structural heterogeneities. *Water Resources Research* 53: 2016-2034.

Fischer, P., A. Jardani, X. Wang, H. Jourde, N. Lecoq. 2017c. Identifying flow networks in a karstified aquifer by application of the cellular automata-based deterministic inversion method (Lez aquifer, France). *Water Resources Research* 53: 10508-10522.

Fischer, P., A. Jardani, N. Lecoq. 2018a. Hydraulic tomography of discrete networks of conduits and fractures in a karstic aquifer by using a deterministic inversion algorithm. *Advances in Water Resources* 112: 83-94.

Fischer, P., A. Jardani, M. Cardiff, N. Lecoq, H. Jourde. 2018b. Hydraulic analysis of harmonic pumping tests in frequency and time domains for identifying the conduits networks in a karstic aquifer. *Journal of Hydrology* 559: 1039-1053.

Fischer, P., A. Jardani, H. Jourde, M. Cardiff, X. Wang, S. Chedeville, N. Lecoq. 2018c. Harmonic pumping tomography applied to image the hydraulic properties and interpret the connectivity of a karstic and fractured aquifer (Lez aquifer, France). *Advances in Water Resources* 119: 227-244.

G

Gallardo, L.A. and M.A. Meju. 2011. Structure-coupled multiphysics imaging in geophysical sciences. *Reviews of Geophysics* 49. doi: 10.1029/2010RG000330.

Ghasemizadeh, R., F. Hellweger, C. Butscher, I. Padilla, D. Vesper, M. Field, 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* 20: 1441-1464.

Ghosh, S., P. Gabane, A. Bose, N. Chakraborti. 2009. Modeling of recrystallization in cold rolled copper using inverse cellular automata and genetic algorithms. *Computational Materials Science* 45: 96-103.

Goldscheider, N. and D. Drew. 2007. Methods in karst hydrogeology. *Taylor & Francis*, London, U.K.

Greengard, L. and V. Rokhlin. 1987. A fast algorithm for particle simulations. *Journal of Computational Physics* 73: 325-348.

Grimstadt, A.A., T. Mannseth, G. Naevdal, H. Urkedal. 2003. Adaptive multiscale permeability estimation. *Computational Geosciences* 7: 1-25.

Guiltinan, E., and M.W. Becker. 2015. Measuring well hydraulic connectivity in fractured bedrock using periodic slug tests. *Journal of Hydrology* 521: 100-107.

H

Haber, E. and D. Oldenburg. 1997. Joint inversion: a structural approach. *Inverse Problems* 13: 63-77.

Hackbusch, W. and S. Börm. 2002. Data-sparse approximation by adaptive \mathcal{H}^2 -matrices. *Computing* 69: 1-35.

Halder, C., L. Madej, M. Pietrzyk. 2014. Discrete micro-scale cellular automata model for modelling phase transformation during heating of dual phase steels. *Archives of Civil and Mechanical Engineering* 14: 96-103.

Halder, C., D. Bachniak, L. Madej, N. Chakraborti, M. Pietrzyk. 2015. Sensitivity analysis of the finite difference 2-D cellular automata model for phase transformation during heating. *Journal of the Iron and Steel Institute of Japan* 55: 285-292.

Hale, D. 2009. Structure-oriented smoothing and semblance. *Center for Wave Phenomena Report* 635.
<http://inside.mines.edu/~dhale/papers/Hale09StructureOrientedSmoothingAndSemblance.pdf>

Halko, N., P.G. Martinsson, J.A. Tropp. 2011. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review* 53: 217-288.

Hao, Y., T.C.J. Yeh, J. Xiang, W.A. Illman, K. Ando, K.C. Hsu, C.H. Lee. 2008. Hydraulic tomography for detecting fracture zone connectivity. *Ground Water* 46: 183–192.

Hartmann, A., N. Goldscheider, T. Wagener, J. Lange, M. Weiler. 2014a. Karst water resources in a changing world: Review of hydrological modeling approaches. *Reviews of Geophysics* 52: 218-242.

Hartmann, A., M. Mudarra, B. Andreo, A. Marin, T. Wagener, J. Lange. 2014b. Modeling spatiotemporal impacts of hydroclimatic extremes on groundwater recharge at a Mediterranean karst aquifer. *Water Resources Research* 50: 6507-6521.

Heath, R.C. 1998. Basic ground-water hydrology: Water-Supply Paper 2220 (ninth printing). *U.S. Geological Survey*.

Hoeksema, R.J. and P.K. Kitanidis. 1984. An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resources Research* 20: 1003-1020.

I

Illman, W.A., X. Liu, S. Takeuchi, T.C.J. Yeh, K. Ando, H. Saegusa. 2009. Hydraulic tomography in fractured granite: Mizunami underground research site, Japan. *Water Resources Research* 45. doi: 10.1029/2007WR006715.

Illman, W.A. 2014. Hydraulic tomography offers improved imaging of heterogeneity in fractured rocks. *Groundwater* 52: 659-684.

J

Jafarpour, B., V.K. Goyal, D.B. McLaughlin, W.T. Freeman. 2010. Compressed History Matching: Exploiting Transform-Domain Sparsity for Regularization of Nonlinear Dynamic Data Integration Problems. *Mathematical Geosciences* 42: 1-27.

Jaquet, O., P. Siegel, G. Klubertanz, H. Benabderhamane. 2004. Stochastic discrete model of karstic networks. *Advances in Water Resources* 27: 751-760.

Jardani, A., J.P. Dupont, A. Revil, N. Massei, M. Fournier, B. Laignel. 2012. Geostatistical inverse modeling of the transmissivity field of a heterogeneous alluvial aquifer under tidal influence. *Journal of Hydrology* 472-473: 287-300.

Javadi, M., M. Sharfizadeh, K. Shahriar. 2016. Uncertainty analysis of groundwater inflow into underground excavations by stochastic discontinuum method: Case study of Siah Bisheh pumped storage project, Iran. *Tunnelling and Underground Space Technology* 51: 424-438.

Jazayeri Noushabadi, M.R. 2009. Characterization of relationship between fracture network and flow-path network in fractured and karstic reservoirs: Numerical modeling and field investigation (Lez aquifer, Southern France). PhD Thesis, Université de Montpellier. English.

Jazayeri Noushabadi, M.R., H. Jourde, G. Massonnat. 2011. Influence of the observation scale on permeability estimation at local and regional scales through well tests in a fractured and karstic aquifer (Lez aquifer, Southern France). *Journal of Hydrology* 403: 321-336.

Jin, Z. and Z. Cui. 2012. Investigation on dynamic recrystallization using a modified cellular automaton. *Computational Materials Science* 63: 249-255.

Jourde, H., F. Cornaton, S. Pistre, P. Bidaux. 2002. Flow behavior in a dual fracture network. *Journal of hydrology* 266: 99-119.

Jourde, H., C. Batiot-Guilhe, V. Bailly-Comte, C. Bicalho, M. Blanc, V. Borrell, C. Bouvier, J.F. Boyer, P. Brunet, M. Cousteau, C. Dieulin, E. Gayard, V. Guinot, F. Hernandez, L. Kong-A-Siou, A. Johannet, V. Leonardi, N. Mazzilli, P. Marchand, N. Patris, S. Pistre, J.L. Seidel, J.D. Taupin, S. Van-Exter. 2011. The MEDYCYSS observatory, a multi scale observatory of flood dynamics and hydrodynamics in karst (Mediterranean border Southern France). In: Lambrakis, N., G. Stournaras, K. Katsanou. *Advances in the research of aquatic environment. Environmental Earth Sciences*. Springer, Berlin, Heidelberg.

K

Kitanidis, P.K. and E.G. Vomvoris. 1983. A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations. *Water Resources Research* 19: 677-690.

Kitanidis, P.K. 1995. Quasi-linear geostatistical theory for inversing. *Water Resources Research* 31: 2411-2419.

Kitanidis, P.K. and J. Lee. 2014. Principal component geostatistical approach for large-dimensional inverse problem. *Water Resources Research* 50: 5428-5443.

Kong-A-Siou, L., A. Johannet, V. Borrell Estupina, S. Pistre. 2015. Neural networks for karst groundwater management: case of the Lez spring (Southern France). *Environmental Earth Sciences* 74: 7617-7632.

Kordilla, J., M. Sauter, T. Reimann, T. Geyer. 2012. Simulation of saturated and unsaturated flow in karst systems at catchment scale using a double continuum approach. *Hydrology of Earth System Sciences* 16: 3909-3923.

Kovacs, A. 2003. Estimation of conduits network geometry of a karst aquifer by the means of groundwater flow modeling (Bure, Switzerland). *Boletin Geologico y Minero* 114: 183-192.

Kovacs, A., P. Perrochet, L. Kiraly, P.Y. Jeannin. 2005. A quantitative method for the characterization of karst aquifers based on spring hydrograph analysis. *Journal of Hydrology* 303: 152-164.

Kovacs, A. and M. Sauter. 2007. Modelling karst hydrodynamics. In: Goldscheider, N. and D. Drew. *Methods in karst hydrogeology*. Taylor and Francis, London, UK.

L

Labat, D., R. Ababou, A. Mangin. 1999. Linear and nonlinear input/output models for karstic springflow and flood prediction at different time scales. *Stochastic Environmental Research and Risk Assessment* 13: 337-364.

Ladouche, B., J.C. Marechal, N. Dorfliger. 2014. Semi-distributed lumped model of a karst system under active management. *Journal of Hydrology* 509: 215-230.

Larocque, M., O. Banton, P. Ackerer, M. Razack. 1999. Determining karst transmissivities with inverse modeling and an equivalent porous media. *Ground Water* 37: 897-903.

Lavenue, M. and G. de Marsily. 2001. Three-dimensional interference test interpretation in a fractured aquifer using the pilot point inverse method. *Water Resources Research* 37: 2659-2675.

Le Coz, M., J. Bodin, P. Renard. 2017. On the use of multiple-point statistics to improve groundwater flow modeling in karst aquifers: a case study from the hydrogeological experimental site of Poitiers, France. *Journal of Hydrology* 545: 109-119.

Le Goc, R., J.R. de Dreuzy, P. Davy. 2010. An inverse problem methodology to identify flow channels in fractured media using synthetic steady-state head and geometrical data. *Advances in Water Resources* 33: 782-800.

Lee, J. and P.K. Kitanidis. 2013. Bayesian inversion with total variation prior for discrete geologic structure identification. *Water Resources Research* 49: 7658-7669.

Lee, J. and P.K. Kitanidis. 2014. Large-scale hydraulic tomography and joint inversion of head and tracer data using the Principal Component Geostatistical Approach (PCGA). *Water Resources Research* 50: 5410-5427.

Lelièvre, P.G. and D.W. Oldenburg. 2009. A comprehensive study of including structural orientation information in geophysical inversions. *Geophysical Journal International* 178: 623-637.

Li, Z.Y., J.H. Zhao, X.H. Qiao, Y.X. Zhang. 2014. An automated approach for conditioning discrete fracture network modelling to in situ measurements. *Australian Journal of Earth Sciences* 61: 755-763.

Lieberman, C., K. Fidkowski, K. Willcox, B. van Bloemen Waanders. 2013. Hessian-based model reduction: large-scale inversion and prediction. *International Journal for Numerical Methods in Fluids* 71: 135-150.

Liedl, R., M. Sauter, D. Huckinghaus, T. Clemens, G. Teutsch. 2003. Simulation of the development of karst aquifers using a coupled continuum pipe flow model. *Water Resources Research* 39. doi: 10.1029/2001WR001206.

Liu, X., Q. Zhou, P.K. Kitanidis, J.T. Birkholzer. 2014. Fast iterative implementation of large-scale nonlinear geostatistical inverse modeling. *Water Resources Research* 50: 198-207.

Lochbühler, T., J.A. Vrugt, M. Sadegh, N. Linde. 2015. Summary statistics from training images as prior information in probabilistic inversion. *Geophysical Journal International* 201: 157-171.

Long, A.J. and R.G. Derickson. 1999. Linear systems analysis in a karst aquifer. *Journal of Hydrology* 219: 206-217.

Lu, Z. and B.A. Robinson. 2006. Parameter identification using the level set method. *Geophysical Research Letters* 33. doi: 10.1029/2005GL025541.

M

Maineult, A., E. Strobach, J. Renner. 2008. Self-potential signals induced by periodic pumping tests. *Journal of Geophysical Research* 113. doi: 10.1029/2007JB005193.

Marsaud, B. 1996. Structure et fonctionnement de la zone noyée des karsts à partir de résultats expérimentaux. PhD Thesis, Université Orsay Paris Sud. French.

Meier, P., A. Medina, J. Carrera. 2001. Geostatistical inversion of cross-hole pumping tests for identifying preferential flow channels within a shear zone. *Groundwater* 39: 10-17.

Moore, E.F. 1962. Machine models of self-reproduction. *The American Mathematical Society* 14: 17-33.

N

Ni, C.F. and T.C.J. Yeh. 2008. Stochastic inversion of pneumatic cross-hole tests and barometric pressure fluctuations in heterogeneous unsaturated formations. *Advances in Water Resources* 31: 1708-1718.

Nowak, W., S. Tenkleve and O.A. Cirpka. 2003. Efficient computation of linearized cross-covariance and auto-covariance matrices of interdependent quantities. *Mathematical Geology* 35: 53-66.

P

Paige, C.C. and M.A. Saunders. 1975. Solution of sparse indefinite systems of linear equations. *SIAM Journal on Numerical Analysis* 12: 617-629.

Painter, S. and V. Cvetkovic. 2005. Upscaling discrete fracture network simulations: An alternative to continuum transport models. *Water Resources Research* 41. doi: 10.1029/2004WR003682.

Pan, P.Z., X.T. Feng, D.P. Xu, L.F. Shen, J.B. Yang. 2011. Modelling fluid flow through a single fracture with different contacts using cellular automata. *Computers and Geotechnics* 38: 959-969.

Pardo-Igúzquiza, E., P.A. Dowd, C. Xu, J.J. Durán-Valsero. 2012. Stochastic simulation of karst conduit networks. *Advances in Water Resources* 35: 141-150.

Pianosi, F., K. Beven, J. Freer, J.W. Hall, J. Rougier, D.B. Stephenson, T. Wagener. 2016. Sensitivity analysis of environmental models: a systematic review with practical workflow. *Environmental Modelling & Software* 79: 214-232.

Pool, M., J. Carrera, A. Alcolea, E.M. Bocanegra. 2015. A comparison of deterministic and stochastic approaches for regional scale inverse modeling on the Mar del Plata aquifer. *Journal of Hydrology* 531: 214-229.

R

Rabinovich, A., W. Barrash, M. Cardiff, D.L. Hochstetler, T. Bakhos, G. Dagan, P.K. Kitanidis. 2015. Frequency dependent hydraulic properties estimated from oscillatory pumping tests in an unconfined aquifer. *Journal of Hydrology* 531: 2-16.

Rao, S.V.N., B.S. Thandaveswara, S. Murthy Bhallamudi, V. Srinivasulu. 2003. Optimal groundwater management in deltaic regions using simulated annealing and neural networks. *Water Resources Management* 17: 409-428.

Rasmussen, T.C., K.G. Haborak, M.H. Young. 2003. Estimating aquifer hydraulic properties using sinusoidal pumping at the Savannah River site, South California, USA. *Hydrogeology Journal* 11: 466-482.

Reeves, D.M., R. Parashar, G. Pohll, R. Carroll, T. Badger, K. Willoughby. 2013. The use of discrete fracture network simulations in the design of horizontal hillslope drainage networks in fractured rock. *Engineering Geology* 163: 132-143.

Renard, P., and D. Allard. 2013. Connectivity metrics for subsurface flow and transport. *Advances in Water Resources* 51: 168-196.

Renner, J., and M. Messar. 2006. Periodic pumping tests. *Geophysical Journal International* 167: 479-493.

Revil, A., C. Gevaudan, N. Lu, A. Maineult. 2008. Hysteresis of the self-potential response associated with harmonic pumping tests. *Geophysical Research Letters* 35. doi:10.1029/2008GL035025.

Ronayne, M.J. 2013. Influence of conduit network geometry on solute transport in karst with a permeable matrix. *Advances in Water Resources* 56: 27-34.

S

Saller, S.P., M.J. Ronayne, A.J. Long. 2013. Comparison of a karst groundwater model with and without discrete conduit flow. *Hydrogeology Journal* 21: 1555-1566.

Schuite, J., L. Longuevergne, O. Bour, N. Guihéneuf, M.W. Becker, M. Cole, T.J. Burbey, N. Lavenant, F. Boudin. 2017. Combining periodic hydraulic tests and surface tilt measurements to explore in situ fracture hydromechanics. *Journal of Geophysical Research: Solid Earth* 122: 6046-6066.

Sharmeen, R., W.A. Illman, S.J. Berg, T.C.J. Yeh, Y.J. Park, E.A. Sudicky, K. Ando. 2012. Transient hydraulic tomography in a fractured dolostone: laboratory rock block experiments. *Water Resources Research* 48. doi: 10.1029/2012WR012216.

Shiklomanov, I. 1993. World fresh water resources. In: Gleick, P.H. Water in crisis: A guide to the world's fresh water resources. *Oxford University Press*. New York.

Soueid Ahmed, A., A. Jardani, A. Revil, J.P. Dupont. 2014. Hydraulic conductivity field characterization from the joint inversion of hydraulic heads and self-potential data. *Water Resources Research* 50: 1-21.

Soueid Ahmed, A., J. Zhou, A. Jardani, A. Revil, J.P. Dupont. 2015. Image-guided inversion in steady-state hydraulic tomography. *Advances in Water Resources* 82: 83-97.

Soueid Ahmed, A., A. Jardani, A. Revil, J.P. Dupont. 2016. Joint inversion of hydraulic head and self-potential data associated with harmonic pumping tests. *Water Resources Research* 52: 6769-6791.

Sun, A.Y., J. Lu, S. Hovorka. 2015. A harmonic pulse testing method for leakage detection in deep subsurface storage formations. *Water Resources Research* 51: 4263-4281.

Sun, A.Y., J. Lu, B.M. Freifeld, S.D. Hovorka, A. Islam. 2016. Using pulse testing for leakage detection in carbon storage reservoirs: A field demonstration. *International Journal of Greenhouse Gas Control* 46: 215-227.

Sun, X., P.L. Rosin, R.R. Martin. 2011. Fast rule identification and neighbourhood selection for cellular automata. *IEEE Transactions on Systems, Man and Cybernetics. Part B: Cybernetics* 41: 749-760.

T

Tan, S.K. and S.U. Guan. 2007. Evolving cellular automata to generate nonlinear sequences with desirable properties. *Applied Soft Computing* 7: 1131-1134.

Tarantola, A. and B. Valette. 1982. Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics and Space Physics* 20: 219-232.

Teutsch, G. 1993. An extended double-porosity concept as a practical modeling approach for a karstified terrain. *Hydrogeological Processes in Karst Terranes* 207: 281-292.

Tsai, F.T.C., N.Z. Sun, W.W.G. Yeh. 2003. Global-local optimization for parameter structure identification in three-dimensional groundwater modeling. *Water Resources Research* 39. doi: 10.1029/2001WR001135.

Tyukhova, A.R., and M. Willmann. 2016. Connectivity metrics based on the path of smallest resistance. *Advances in Water Resources* 88: 14-20.

V

Van der Weeën, P., J.M. Baetens, B. de Baets. 2011. Design and parameterization of a stochastic cellular automaton describing a chemical reaction. *Journal of Computational Chemistry* 32: 1952-1961

Von Neumann, J. and A.W. Burks. 1966. Theory of self-reproducing automata. University of Illinois Press, USA. ISBN 0-598-37798-0.

cba.mit.edu/events/03.11.ASE/docs/VonNeumann.pdf

W

Wang, X., A. Jardani, H. Jourde, L. Lonergan, J. Cosgrove, O. Gosselin, G., Massonnat. 2016. Characterisation of the transmissivity field of a fractured and karstic aquifer, Southern France. *Advances in Water Resources* 87: 106-121.

Wang, X., A. Jardani, H. Jourde. 2017. A hybrid inverse method for hydraulic tomography in fractured and karstic media. *Journal of Hydrology* 551: 29-46.

White, W.B. 2002. Karst hydrology: recent developments and open questions. *Engineering Geology* 65: 85-105.

Y

Yeh, T.C.J. and S. Liu. 2000. Hydraulic tomography: Development of a new aquifer test method. *Water Resources Research* 36: 2095–2105.

Yeh, T.C.J. and C.H. Lee. 2007. Time to change the way we collect and analyze data for aquifer characterization. *Groundwater* 45: 116-118.

Z

Zha, Y., T.C.J. Yeh, W.A. Illman, T. Tanaka, P. Bruines, H. Onoe, H. Saegusa. 2015. What does hydraulic tomography tell us about fractured geological media? A field study and synthetic experiments. *Journal of Hydrology* 531:17-30.

Zha, Y., T.C.J. Yeh, W.A. Illman, T. Tanaka, P. Bruines, H. Onoe, H. Saegusa, D. Mao, S. Takeuchi, J.C. Wen. 2016. An application of hydraulic tomography to a large-scale fractured granite site, Mizunami, Japan. *Groundwater* 54: 793-804.

Zha, Y., T.C.J. Yeh, W.A. Illman, H. Onoe, C.M.W. Mok, J.C. Wen, S.Y. Huang, W. Wang. 2017. Incorporating geologic information into hydraulic tomography: A general framework based on geostatistical approach. *Water Resources Research* 53. doi: 10.1002/2016WR019185.

Zhou, H., J. Gomez-Hernandez, L. Li. 2014. Inverse methods in hydrogeology: evolution and recent trends. *Advances in Water Resources* 63: 22-37.

Zhou, Y., D. Lim, F. Cupola, M. Cardiff. 2016. Aquifer imaging with pressure waves - Evaluation of low-impact characterization through sandbox experiments. *Water Resources Research* 52: 2141-2156.

Zhu, J. and T.C.J. Yeh. 2005. Characterization of aquifer heterogeneity using transient hydraulic tomography. *Water Resources Research* 41. doi: 10.1029/2004WR003790.

Zimmerman, R.W., G. Chen, T. Hadgu, G.S. Bodvarsson. 1993. A numerical dual-porosity model with semianalytical treatment of fracture/matrix flow. *Water Resources Research* 29: 2127–2137.

ANNEXES (APPENDIX)

Annexe / Appendix 1:

Article "Application of Large-Scale Inversion Algorithms to Hydraulic Tomography in an Alluvial Aquifer"

Groundwater 55 (2017): 208-218

P. Fischer, A. Jardani, A. Soueid Ahmed, M. Abbas, X. Wang, H. Jourde, N. Lecoq

Annexe / Appendix 2:

Article "A Cellular Automata-based Deterministic Inversion Algorithm for the Characterization of Linear Structural Heterogeneities"

Water Resources Research 53 (2017): 2016-2034

P. Fischer, A. Jardani, N. Lecoq

Annexe / Appendix 3:

Article "Hydraulic Tomography of Discrete Networks of Conduits and Fractures in a Karstic Aquifer by Using a Deterministic Inversion Algorithm"

Advances in Water Resources 112 (2018): 83-94

P. Fischer, A. Jardani, N. Lecoq

Annexe / Appendix 4:

Article "Identifying Flow Networks in a Karstified Aquifer by Application of the Cellular Automata-based Deterministic Inversion Method (Lez Aquifer, France)"

Water Resources Research 53 (2017): 10508-10522

P. Fischer, A. Jardani, X. Wang, H. Jourde, N. Lecoq

Annexe / Appendix 5:

Article "Hydraulic Analysis of Harmonic Pumping Tests in Frequency and Time Domains for Identifying the Conduits Networks in a Karstic Aquifer"

Journal of Hydrology 559 (2018): 1039-1053

P. Fischer, A. Jardani, M. Cardiff, N. Lecoq, H. Jourde

Annexe / Appendix 6:

Article "Harmonic Pumping Tomography Applied to Image the Hydraulic Properties and Interpret the Connectivity of a Karstic and Fractured Aquifer (Lez Aquifer, France)"

Advances in Water Resources 119 (2018): 227-244

P. Fischer, A. Jardani, H. Jourde, M. Cardiff, X. Wang, S. Chedeville, N. Lecoq

Application of Large-Scale Inversion Algorithms to Hydraulic Tomography in an Alluvial Aquifer

by P. Fischer¹, A. Jardani², A. Soueid Ahmed², M. Abbas², X. Wang³, H. Jourde⁴, and N. Lecoq²

Abstract

Large-scale inversion methods have been recently developed and permitted now to considerably reduce the computation time and memory needed for inversions of models with a large amount of parameters and data. In this work, we have applied a deterministic geostatistical inversion algorithm to a hydraulic tomography investigation conducted in an experimental field site situated within an alluvial aquifer in Southern France. This application aims to achieve a 2-D large-scale modeling of the spatial transmissivity distribution of the site. The inversion algorithm uses a quasi-Newton iterative process based on a Bayesian approach. We compared the results obtained by using three different methodologies for sensitivity analysis: an adjoint-state method, a finite-difference method, and a principal component geostatistical approach (PCGA). The PCGA is a large-scale adapted method which was developed for inversions with a large number of parameters by using an approximation of the covariance matrix, and by avoiding the calculation of the full Jacobian sensitivity matrix. We reconstructed high-resolution transmissivity fields (composed of up to 25,600 cells) which generated good correlations between the measured and computed hydraulic heads. In particular, we show that, by combining the PCGA inversion method and the hydraulic tomography method, we are able to substantially reduce the computation time of the inversions, while still producing high-quality inversion results as those obtained from the other sensitivity analysis methodologies.

Introduction

In hydrogeology, the assessment of hydraulic properties of subsurface aquifers, such as transmissivity, storage coefficient and solute transport parameters, is a key issue to an adequate management and protection of groundwater resources. Generally, aquifer characterization is based on the interpretation of hydraulic observations data collected during pumping, infiltration, or tracer tests (Carrera and Neuman 1986b; Rao et al. 2003; Lee and Kitanidis 2014; Pool et al. 2015). Therefore, hydraulic tomography is considered as one of the most effective approaches for characterizing the spatial distribution of hydraulic transmissivity of an aquifer (Cardiff et al. 2009; Berg and

Illman 2013; Cardiff et al. 2013; Soueid Ahmed et al. 2015; Zha et al. 2015; Wang et al. 2016). This method relies on a set of hydraulic head responses recorded during cross-hole pumping experiments. Then, the interpretation can be achieved through the use of an inverse algorithm to image the spatially varying hydraulic properties in the subsurface.

The inverse problem for estimating hydraulic parameters involves a formulation of a forward problem, which sets up the link between the hydraulic observations and the unknown hydraulic parameters (Tarantola and Valette 1982). For a hydraulic tomography inversion, the forward modeling is based on a numerical method (e.g., the finite element, finite-difference and finite volume methods) used to solve the partial differential equation of the groundwater flow. The forward problem operator can be formulated as:

$$\mathbf{d} = f(\mathbf{s}) + \boldsymbol{\eta} \quad (1)$$

where \mathbf{d} represents the hydraulic responses of the model, \mathbf{s} is the logarithm of the m unknown hydraulic transmissivities, to be estimated from a set of n observed data \mathbf{d}_{obs} and a nonlinear forward modeling application $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$. $\boldsymbol{\eta}$ is an additive noise of the numerical modeling.

In a probabilistic framework, the inverse problem maximizes a posterior probability density function $\pi_{\text{post}}(\mathbf{s}|\mathbf{d})$. Generally, the problem is ill-posed and the

¹Corresponding author: UNIROUEN, UNICAEN, CNRS, M2C, Normandie University, 76000 Rouen, France; +33(0)235147181; fax: +33(0)235147022; pierre.fischer1@univ-rouen.fr

²UNIROUEN, UNICAEN, CNRS, M2C, Normandie University, 76000 Rouen, France.

³Department of Earth Science and Engineering, Imperial College London, London SW7 2AZ, UK.

⁴CNRS, Laboratoire Hydrosociétés, Université de Montpellier, 34000 Montpellier, France.

Article impact statement: Efficiency comparison of different largescale inversion methodologies associated with hydraulic tomography to model an alluvial aquifer.

Received March 2016, accepted August 2016.

© 2016, National Ground Water Association.

doi: 10.1111/gwat.12457

solution is non-unique, therefore additional a priori information is required in order to find one physically meaningful solution (Carrera and Neuman 1986a). Furthermore, in our case, the problem is under-determined, it involves a small number of measurements but a large number of unknown parameters ($n \ll m$). The inverse problem corresponds to recovering the “best fitting” model parameters $\hat{\mathbf{s}}$ which allow the model responses to match the observed data, and at the same time respect the constraints imposed by the a priori information on the model. Using the Bayes’ formula, the posterior probability density function can be expressed as (Elsheikh et al. 2014):

$$\pi_{\text{post}}(\mathbf{s}|\mathbf{d}) \propto \exp \left[-\frac{1}{2} (\mathbf{f}(\mathbf{s}) - \mathbf{d})^T \mathbf{R}^{-1} (\mathbf{f}(\mathbf{s}) - \mathbf{d}) - \frac{1}{2} (\mathbf{s} - \mathbf{s}_{\text{prior}})^T \mathbf{Q}^{-1} (\mathbf{s} - \mathbf{s}_{\text{prior}}) \right] \quad (2)$$

where $\mathbf{s}_{\text{prior}}$ denotes the a priori model; \mathbf{Q} is a $m \times m$ covariance matrix of the model parameters, which can include geostatistical information about the distribution trend or pattern of the unknowns parameters (in that case \mathbf{Q} is defined in a matrix where elements of the matrix represent the variogram function associated with the distance between cells of the model) and \mathbf{R} denotes a $n \times n$ diagonal covariance matrix which accounts for the errors occurred in the data measurements.

The aim of the inversion problem is to find a set of parameter which maximizes the density of probability, $\pi_{\text{post}}(\mathbf{s}|\mathbf{d})$. This corresponds to a model of high probability with respect to the measurements and the imposed a priori model (Tarantola and Valette 1982). To solve the inverse problem, two main groups of iterative methods are often employed: (1) the deterministic methods which assume that the algorithm converges to a local minimum by performing a linearization of an objective function and (2) the stochastic methods which converge to a global minimum by selecting different randomly generated parameters fields as probable solutions to the model (Pool et al. 2015). Several deterministic and stochastic inversion algorithms have been widely applied in hydrogeology, but they are time- and memory-expensive, particularly for the cases which involve a fine discretization of the parameter grid and/or a large amount of observational data (Kitanidis and Lee 2014). Thereby, one of the main trends in the development of inversion theory during the last few years was to develop algorithms which are able to solve inversion models with a large number of unknown parameters and data.

Recently, several time- and memory-saving methods have been developed to reduce the memory cost and the computation time of the usually large matrices involved in the inversion algorithms. One way for reducing the computational and memory demands is to use some approximation methods on matrix operations, such as the fast multipole method (FMM) (Greengard and Rokhlin 1987), which is based on Legendre polynomial expansions and spherical harmonics. The FMM was associated with the hierarchical matrices approach (Hackbusch and Börm

2002) to compute matrix-vector products for a large-scale application in seismic imaging (Ambikasaran et al. 2013). Another way is through using the MINRES Krylov subspace method (Paige and Saunders 1975) which can be combined with the fast fourier transform (FFT) (Nowak et al. 2003) to iteratively solve inversions of large matrix systems. This method has recently been applied to 3-D large-scale transient hydraulic tomography problems (Liu et al. 2014). The two approaches mentioned above avoid the calculation of the full Jacobian matrix of the forward model at each iteration. A new method has recently been developed by Lieberman et al. (2013), in an application of an inversion algorithm for a large-scale 3-D transient contaminant transport. The authors used a proper orthogonal decomposition method (POD) to compute a projection basis with the eigenvectors associated to the highest eigenvalues of the Hessian matrix of the forward problem. The inverse problem was then solved in a reduced projected subspace.

In this paper, we have attempted to achieve an efficient site hydraulic characterization (i.e., to obtain high-resolution transmissivity fields at a low computational cost) by applying a recently developed geostatistical inversion method, the principal component geostatistical approach (PCGA) (Lee and Kitanidis 2014). This method can considerably reduce the computation time and the memory cost of inversions by using an approximation of the covariance matrix \mathbf{Q} based on a singular value decomposition method (SVD), and by avoiding the computation of the Jacobian matrix through the use of a matrix-free product based on a finite-difference method.

This paper presents an application of the PCGA method, combined with a hydraulic tomography investigation, for a large-scale inverse modeling of the hydraulic transmissivity field of an alluvial aquifer. First, we present the methodology of the geostatistical approach (GA) algorithm, and the modifications for large-scale application (PCGA). Then, we describe the hydrogeological background of the experimental field site, from which the hydraulic measurements were taken, and present the numerical model setup. Finally, we show our inversion results. In particular, we compare these results to those obtained by applying the classical GA method with two different methods in Jacobian matrix computation (i.e., an adjoint-state and a finite-difference methods), which do not use a covariance matrix approximation. We have evaluated the computation times, and the sensitivities and accuracies of the inversion results for the three different methodologies. Using a hydraulic tomography field application on a porous aquifer, we show the advantages of the PCGA inversion method for efficient large-scale inverse modeling in hydrogeology.

Principal Component Geostatistical Approach

In the GA (Kitanidis and Vomvoris 1983; Hoeksema and Kitanidis 1984; Kitanidis 1995), the prior probability density function of the m model parameters \mathbf{s} is set as a multivariate Gaussian with a mean $E(\mathbf{s}) = \mathbf{X}\boldsymbol{\beta}$ where \mathbf{X}

is an $m \times p$ known matrix and β a $p \times 1$ vector to be determined during the inversion process (generally $p = 1$), and a covariance, $E[(s - X\beta)(s - X\beta)^T] = Q(\theta)$.

The posterior probability density function $\Psi = -\ln[\pi(s|\mathbf{d})]$ (also called in inversion problems the objective function) then becomes:

$$\Psi = \frac{1}{2} (f(s) - \mathbf{d})^T \mathbf{R}^{-1} (f(s) - \mathbf{d}) + \frac{1}{2} (s - X\beta)^T \mathbf{Q}^{-1} (s - X\beta) \quad (3)$$

The best approximation \hat{s} for the model parameters, taking into account the a priori information and the observed data, can be found as being the model maximizing the density of probability in (Equation 2), which is also equivalent to minimize the argument of its exponential. Thus, \hat{s} is found by minimizing the objective function Ψ (Equation 3). This minimization can be achieved by using a Newton linearization iterative approach on s . The iterative process initializes at a reasonable s_0 . Then, at iteration step $j+1$, the new value s_{j+1} is found in the vicinity of the previous model s_j using a first-order Taylor approximation:

$$f(s_{j+1}) = f(s_j) + \mathbf{F}_j (s_{j+1} - s_j) \quad (4)$$

Here \mathbf{F}_j is the $n \times m$ Jacobian matrix of the forward problem f for s_j : $\mathbf{F}_j = \left. \frac{\partial f}{\partial s} \right|_{s=s_j}$.

After some matrix manipulations, the updated solution of the parameters in the iterative process, found by minimizing the objective function, can be written as (Kitanidis 1995):

$$s_{j+1} = X\beta_j + \mathbf{Q}\mathbf{F}_j^T \xi_j \quad (5)$$

where the $p \times 1$ vector β_j and the $n \times 1$ vector ξ_j are solutions of the following matrix system (Kitanidis 1995):

$$\begin{bmatrix} \mathbf{F}_j \mathbf{Q} \mathbf{F}_j^T + \mathbf{R} & \mathbf{F}_j \mathbf{X} \\ (\mathbf{F}_j \mathbf{X})^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \xi_j \\ \beta_j \end{bmatrix} = \begin{bmatrix} \mathbf{d} - f(s_j) + \mathbf{F}_j s_j \\ \mathbf{0} \end{bmatrix} \quad (6)$$

Note that here $\mathbf{0}$ represents a $p \times p$ matrix of zeros on the left-hand side and a $p \times 1$ vector of zeros on the right-hand side.

At the end of the iterative process, to quantify the model parameter's uncertainty after optimization, we can compute the posterior covariance of s derived as:

$$\mathbf{Q}_{\text{post}} = \mathbf{Q} - \begin{bmatrix} \mathbf{F}_j \mathbf{Q} \\ \mathbf{X}^T \end{bmatrix}^T \left(\begin{bmatrix} \mathbf{F}_j \mathbf{Q} \mathbf{F}_j^T + \mathbf{R} & \mathbf{F}_j \mathbf{X} \\ (\mathbf{F}_j \mathbf{X})^T & \mathbf{0} \end{bmatrix} \right)^{-1} \begin{bmatrix} \mathbf{F}_j \mathbf{Q} \\ \mathbf{X}^T \end{bmatrix} \quad (7)$$

The GA method as presented above needs the computation of the Jacobian matrix \mathbf{F} for each iteration in order to solve the system (Equation 6), which can usually be done by solving the forward problem $m+1$ times using a finite-difference method, or $n+1$ times using an

adjoint-state method. Even if the adjoint-state method may considerably decrease the computation time for under-determined problems (see Cardiff and Kitanidis 2008 for a comparison of the finite-difference and adjoint-state method computation times), it is not efficient for large-scale problems with a large number of measurements and parameters. Another problem which appears in the GA method is that when the number of data and/or parameters is high, it requires a significant computational power for the calculation and storage of the covariance matrix \mathbf{Q} (which can be alleviated by FFT, H-matrices, or FMM). To overcome these difficulties, Kitanidis and Lee have developed the PCGA, on the basis of the GA method.

In this new approach, the computational and memory costs associated with the manipulation of matrix \mathbf{Q} are reduced using a low-rank approximation of a chosen truncation order $K \ll m$ through a SVD:

$$\mathbf{Q}_K = \mathbf{U} \mathbf{S} \mathbf{V}^T \quad (8)$$

where \mathbf{S} is a $K \times K$ diagonal matrix containing the singular values of \mathbf{Q} sorted in descending order, \mathbf{U} is a $m \times K$ matrix and \mathbf{V} is an $m \times K$ matrix. As \mathbf{Q} is defined as a symmetric matrix, its SVD simplifies to: $\mathbf{Q}_K = \mathbf{V} \mathbf{S} \mathbf{V}^T$.

This decomposition can also be written as:

$$\mathbf{Q}_K = \sum_{i=1}^K \xi_i \xi_i^T \text{ with } \xi_i = \sqrt{\lambda_i} \mathbf{V}_i \quad (9)$$

where λ_i is the i th singular value (also $\mathbf{S}(i,i)$) and \mathbf{V}_i is the i th column vector of \mathbf{V} associated to λ_i . The error arising from this K -rank decomposition equals to the $K+1$ th singular value (λ_{K+1}) of \mathbf{Q} .

However, this decomposition is a good approximation only for a matrix \mathbf{Q} in which the most of its information is contained in its few highest eigenvalues and eigenvectors, meaning a relatively smooth pattern. One can also use an eigen-decomposition if \mathbf{Q} is a matrix defined by positive eigenvalues or a randomized decomposition approach which is efficient for high-dimensional matrices with $m \sim 1,000,000$ (Halko et al. 2011).

In addition, the PCGA method also avoids the full Jacobian matrix calculation at each iteration. When performing a matrix product, such as $\mathbf{F}_j \mathbf{u}$ (where \mathbf{F}_j is the $n \times m$ Jacobian matrix and \mathbf{u} is a $m \times 1$ vector), instead of computing it directly, the method finds an approximation to its accurate form using a first-order Taylor series:

$$f\left(s_j + \frac{\delta \|\mathbf{s}_j\|}{\|\mathbf{u}\|} \mathbf{u}\right) = f(s_j) + \frac{\delta \|\mathbf{s}_j\|}{\|\mathbf{u}\|} \mathbf{F}_j \mathbf{u} + \sigma(\delta^2) \quad (10)$$

$$\mathbf{F}_j \mathbf{u} \approx \frac{\|\mathbf{u}\|}{\delta \|\mathbf{s}_j\|} \left[f\left(s_j + \frac{\delta \|\mathbf{s}_j\|}{\|\mathbf{u}\|} \mathbf{u}\right) - f(s_j) \right] \quad (11)$$

where δ is a finite-difference interval and $\|\mathbf{u}\|$ and $\|\mathbf{s}_j\|$ are the norm of the vectors \mathbf{u} and \mathbf{s}_j .

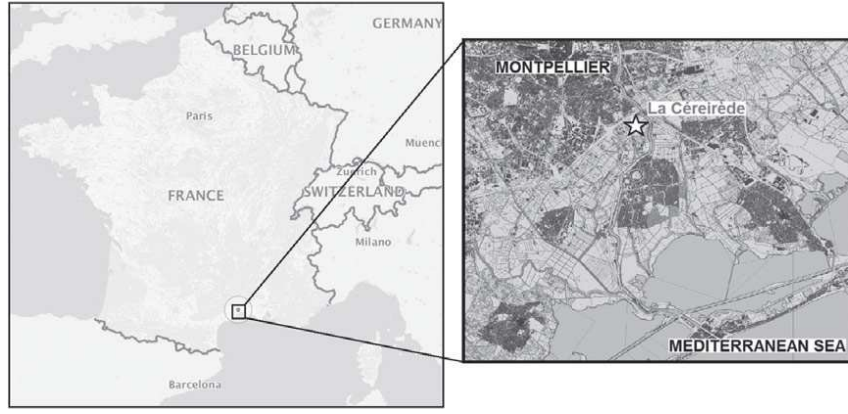


Figure 1. Location of the studied experimental site “La Céreirède” (map and aerial photography from geoportail.fr) occupying an area of 720 m². It is situated in the South of France, near the town of Montpellier and the Mediterranean Sea.

Thus, in the matrix system (Equation 6) of the GA algorithm the calculation of the full Jacobian \mathbf{F} is avoided by approximating the computation of $\mathbf{F}_j \mathbf{s}_j$, $\mathbf{F}_j \mathbf{X}$, $\mathbf{F}_j \mathbf{Q} \mathbf{F}_j^T$ and $\mathbf{Q} \mathbf{F}_j^T$ with (Kitanidis and Lee 2014):

$$\mathbf{F}_j \mathbf{s}_j = \frac{1}{\delta} [f(\mathbf{s}_j + \delta \mathbf{s}_j) - f(\mathbf{s}_j)] \quad (12)$$

$$\mathbf{F}_j \mathbf{X}_i = \frac{\|\mathbf{X}_i\|}{\delta \|\mathbf{s}_j\|} \left[f\left(\mathbf{s}_j + \frac{\delta \|\mathbf{s}_j\|}{\|\mathbf{X}_i\|} \mathbf{X}_i\right) - f(\mathbf{s}_j) \right] \quad (13)$$

with \mathbf{X}_i the i th column of \mathbf{X}

$$\begin{aligned} \mathbf{F}_j \mathbf{Q} \mathbf{F}_j^T &\approx \mathbf{F}_j \mathbf{Q}_K \mathbf{F}_j^T = \mathbf{F}_j \sum_{i=1}^K \zeta_i \zeta_i^T \mathbf{F}_j^T \\ &= \sum_{i=1}^K (\mathbf{F}_j \zeta_i) (\mathbf{F}_j \zeta_i)^T = \sum_{i=1}^K \eta_i \eta_i^T \end{aligned} \quad (14)$$

$$\begin{aligned} \mathbf{Q} \mathbf{F}_j^T &\approx \mathbf{Q}_K \mathbf{F}_j^T = \sum_{i=1}^K \zeta_i \zeta_i^T \mathbf{F}_j^T \\ &= \sum_{i=1}^K \zeta_i (\mathbf{F}_j \zeta_i)^T = \sum_{i=1}^K \zeta_i \eta_i^T \end{aligned} \quad (15)$$

where $\eta_i = \mathbf{F}_j \zeta_i = \frac{\|\zeta_i\|}{\delta \|\mathbf{s}_j\|} \left[f\left(\mathbf{s}_j + \frac{\delta \|\mathbf{s}_j\|}{\|\zeta_i\|} \zeta_i\right) - f(\mathbf{s}_j) \right]$.

If we now consider the number of forward model evaluations needed per iteration for the calculation of the Jacobian matrix, there are $K + p + 2$ runs. One run is needed for evaluating $f(\mathbf{s}_j)$, one is needed for assessing $f(\mathbf{s}_j + \delta \mathbf{s}_j)$, K runs are needed for calculating η_i and p runs are needed for computing \mathbf{X}_i . It can be observed that, with this method, the number of forward model runs per iteration is no longer dependent on m or n . Hence, the number of parameters and observed data can increase without increasing the run time of the

algorithm. However, if the number of parameter increases, the low-rank approximation order K might also need to be increased slightly in order to maintain a small truncation error for \mathbf{Q}_K .

The algorithm iteratively updates the parameters set in Equation 5 by solving the matrix system of Equation 6 with the PCGA approximations until the optimum $\hat{\mathbf{s}} = \mathbf{s}_{\text{post}}$ is achieved, that is, the objective function has iteratively converged to a local minimum.

Application to an Experimental Site

We have applied the PCGA large-scale method as presented in the previous part to an experimental site, named “la Céreirède.” The field site is located in Montpellier in the South of France, on the alluvium of the Lez river, which flows toward the Mediterranean Sea a few kilometers downstream (Figure 1).

At the field site, the alluvial deposit is composed of a 12-m thick formation of unconsolidated sands and silts lying on a 2-m thick layer of pebbles and gravels. Beneath these Quaternary formations, there exist clayey sands of the Pliocene, marls of the Miocene, and limestone of the Jurassic (Figure 2). Three porous aquifer formations have been characterized in this site:

- a low permeability aquifer in the upper part of the alluvium (unconsolidated sands and silts),
- a semi-confined aquifer in the pebbles and gravel,
- a semi-confined to confined aquifer in the clayey sands of Pliocene.

The field site comprises 12 wells which fully penetrate the three aforementioned aquifers in an area of 36×20 m² (Figure 3). The hydraulic data were collected by performing two pumping tests in PZ 2 (5 L/min), and in PZ 11 (5 L/min), respectively, while measuring hydraulic head variations in the other 10 wells. The pumping was performed at the depth of the pebbles and gravel layer, which is the most productive aquifer, because its transmissivity is considerably higher than the

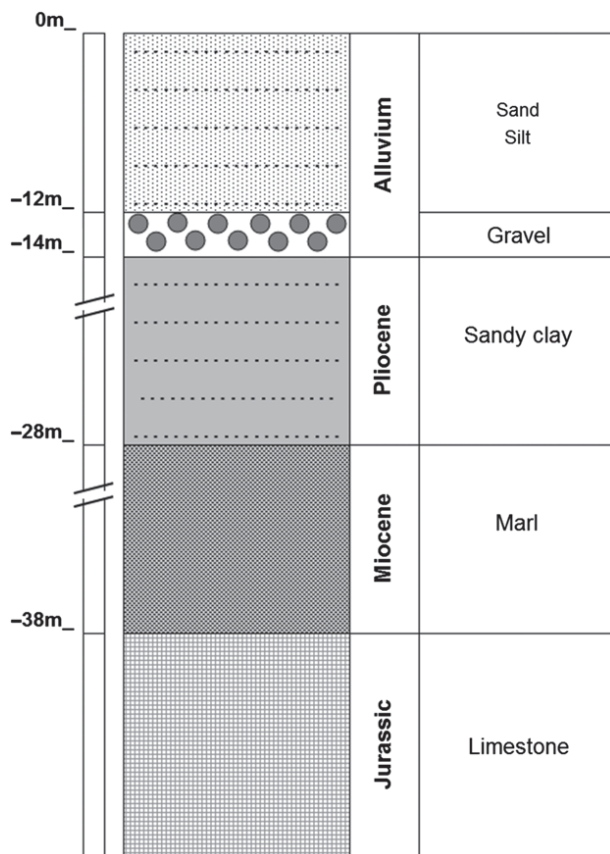


Figure 2. Schematic geological section of the experimental site “La Céreirède.” Three aquifers formations have been characterized in the sands and silts alluvium, in the gravels and in the clayey sands.

transmissivities in silts and clayed sands. We considered that the contribution of the two others aquifers to the water pumped is negligible compared to the contribution from the pebbles and gravel aquifer. From a classic hydrogeological analysis of the soil of each well, we could also estimate values of the field transmissivity at these points. Using these punctual values of transmissivities and their positions in the field as input in a MATLAB variogram routine (“variogramfit” by W. Schwanghart), we were able to obtain the transmissivity field variogram function, which will be used in the inversion algorithm to create the covariance matrix **Q** (Table 1). The variogram function is of type exponential with a sill of 0.11 and a range of 8 m.

The PCGA inversion algorithm was implemented in MATLAB and connected to the flow modeling software COMSOL Multiphysics, which solves the forward problem. The inversion algorithm performs, at the end of each iteration, a linear minimum research (“fminsearch” MATLAB function) to accelerate the convergence of Equation 3 to a local minima.

The 2-D flow model is discretized as a grid of $\sqrt{m} \times \sqrt{m}$ transmissivity cells in a rectangular zone. This local model is enclosed by a larger buffer zone of 100×100 m² with a constant transmissivity of 10^{-5} m²/s (average value of locally estimated transmissivities at the field

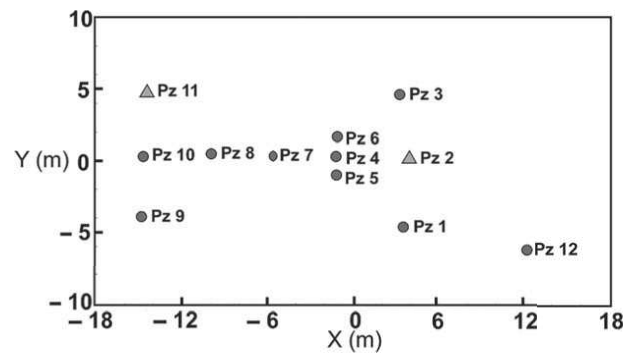


Figure 3. Well pattern on the experimental site “La Céreirède” (circles represent the 10 measurement wells and triangles represent the 2 pumping wells). As Hydraulic draw-downs in the pumping wells are not measured, the tomography provided 20 observed data.

Table 1
Values of Variables Used to Perform the PCGA Inversion on a Model of the Site for 25,600 Parameters and 20 Observed Data. Results of This Inversion Are Shown in Figures 4 and 5

Geometry	$X \text{ (m)} = [-18, 18];$ $Y \text{ (m)} = [-10, 10]$
Grid (number of parameters)	160×160 cells
Uncertainty on data	$\sigma = 0.001 \text{ m};$ $\mathbf{R} = \sigma \cdot \mathbf{Id} \text{ (n)}$
Experimental variogram function	$\text{Var}(\text{distance}) = 0.11 \times \left[1 - \exp\left(-\frac{\text{distance}}{8}\right) \right]$
Finite-difference step	$\delta = 10^{-5}$

site) and a 0 m constant head condition at the boundaries (no drawdown induced by the pumping wells). This buffer zone was set up in order to minimize the impact of the boundary conditions. The flow simulations were performed under steady-state conditions. The inversion of the model is set up using the 2×10 drawdown observed during the pumping tests (10 measurement wells for each of the 2 pumping tests) representing the observed data in the inversion algorithm. The inversion aims to reconstruct the spatially varying T distribution in the local region producing the observational data set.

Results

We have applied the PCGA to assess the equivalent transmissivity field of the multi-layered aquifer at the Céreirède field site. The most transmissive part of the aquifer is the pebbles and gravels part, but the alluvium and the clayey sands might also be the cause of some variations in the estimated equivalent transmissivity field. The inputs to the inversion models are given in Table 1.

The inversions were performed on a uniform fine-scale discretization grid (160×160). A constant initial transmissivity field was considered in the inverse models. The K th order of truncation for the covariance matrix **Q**

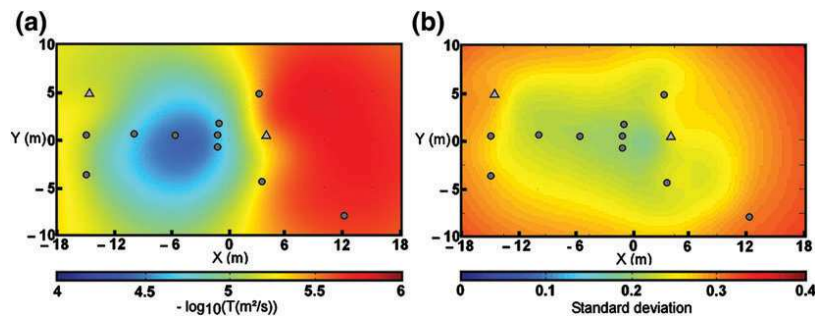


Figure 4. Maps of the log-transmissivity (a, on the left) and parameter's a posteriori standard deviation (b, on the right) for a PCGA inversion method with 25,600 parameters, 20 observed data and a covariance matrix decomposition of order $K = 128$ applied to the experimental site. The transmissivities vary around a mean of 10^{-5} m²/s which is consistent with transmissivity values estimated from pumping test analysis. The aquifer is less transmissive in the eastern part and more in the western part especially in a zone around PZ 7 (see Figure 3). But we got a better precision in zones with more information: at the center and the western part of the map, while in the eastern part where we did not have piezometers, the results show a larger standard deviation.

was selected such that the first truncated singular value of the matrix (the low-rank approximation error) fall below 1. This corresponds to an order of $K = 128$. We will show that this choice of truncation order is acceptable and allows the significant information about the prior model structure to be preserved. The low-rank covariance matrix was calculated and then imported to the inversion algorithm. The parallelized computation of the low-rank decomposition takes only a few minutes (it even takes less than 1 min with a 32 cores parallelization). The inversion then converged in 2 h 45 min on an Intel Xeon QuadCore 2.8GHz with 12Go RAM.

The results from the inversion are shown in Figures 4 and 5. First, from the distribution of model parameters (given as negative log transmissivity in Figure 4), it can be seen that the value of the inverted transmissivity takes the mean value 10^{-5} m²/s, which is the mean of the transmissivity measurements on the site. A clear contrast in T is observed between the two regions on the east and west sides of boreholes PZ1, PZ2, and PZ3. Overall, the eastern part which is closer to the Lez river, is slightly less transmissive than the western part (T are approximately 2×10^{-6} m²/s and 1×10^{-5} m²/s, respectively). An area with a highest T (3×10^{-5} m²/s) is also highlighted within the western part, around PZ 7. But it has to be noticed that the boreholes and pumping wells are not homogeneously distributed over the site, thus some parts of the site (especially on the eastern side) might give more uncertain results. Therefore, it is interesting to estimate the values of the transmissivity field uncertainties.

Figure 4b presents a map of the uncertainty for each parameter value (given by the diagonal entries of the posterior covariance matrix \mathbf{Q}_{post}). The standard deviation for the log-transmissivity varies between 0.2 for the parameter cells near the investigation wells, and 0.33 in both the area with very few information and close to the model's boundaries. In particular, the uncertainty in the eastern part, where the number of wells is small, is much higher.

Good correlations between the calculated and measured hydraulic heads were obtained (Figure 5). The root

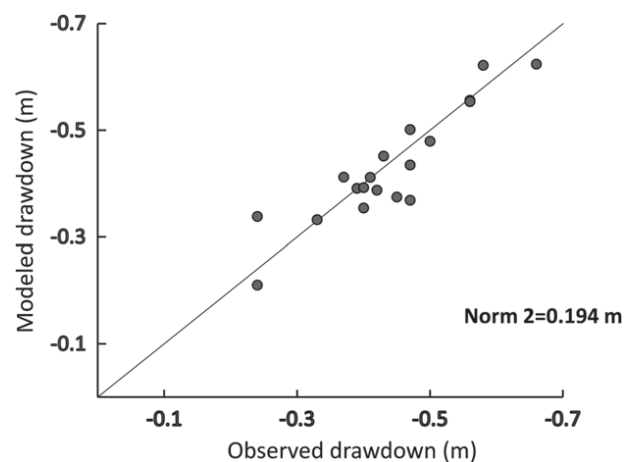


Figure 5. Graph showing the differences between the 20 observed drawdowns and modeled drawdowns after convergence of a PCGA inversion method with 25,600 parameters and a covariance matrix decomposition of order $K = 128$ applied to the experimental site. The drawdowns are globally well reproduced.

mean square error calculated at the end of the inversion was computed as 0.194 m. The hydraulic heads with the most significant difference between the inverted and the measured values are observed on PZ1 and PZ9.

It is clear that the PCGA method is reliable for the modeling of the porous aquifer considered in this work because it produces a set of good inversion results with a high resolution (each cell represents a rectangle of 22.5cm \times 12.5 cm on the site) from a few measurements (20 observed data for 720 m²) in less than 3 h. However, the main problem of this method is that the SVD of the covariance matrix needs a considerably large amount of time and memory. The computational demands increase squarely with the number of model parameters, m . In this work, the decomposition was only performed once before the inversion, and the resulting low-rank covariance matrix was used at all iteration steps throughout the inversion (i.e., the variogram function remaining the same). Otherwise, if a variable variogram model is desired

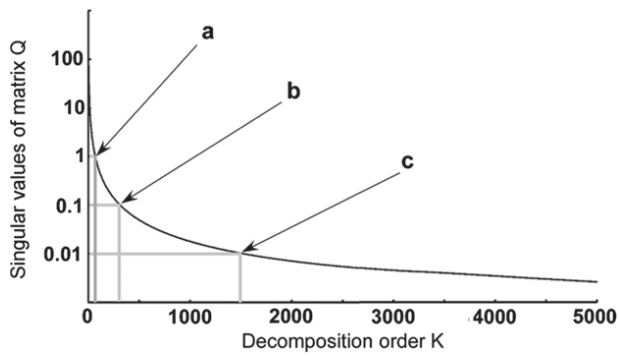


Figure 6. Covariance matrix singular values decrement curve for 10,000 parameters. Three decomposition order (a to c) corresponding to truncation error of 1, 0.1, and 0.01 have been chosen for the results comparison of the PCGA inversion method (see Figure 7).

at different iteration steps, the computation of the covariance matrix decomposition can also be accelerated by specific linear algebra methods (FFT, FMM, and H-matrices) and a parallelization on several cores to achieve a reasonable computational time (Lee and Kitanidis 2014).

Comparison of Results between PCGA Simulations with Different Decomposition Order, and between PCGA and GA Simulations

In addition to the PCGA experimental application presented above, several other inversions were also conducted for the same field site but with a smaller number of parameters and using different methods for the computation of sensitivity matrix (input for each inversion is shown in Table 1). These numerical experiments allows us to compare (1) the results from PCGA with varied low-rank truncation K and (2) the results from PCGA to those obtained from the GA method with a first-order finite-difference Jacobian matrix computation, and an adjoint-state Jacobian matrix computation (the integral was solved using the Gauss-Legendre quadratic method as described in Soueid Ahmed et al. 2014). The inversion results are compared with respect to the total computation time and relative accuracy of the results.

Comparison of Inversion Results for Using Different Decomposition Order

First, we assess the effect of the chosen truncation order for the covariance matrix. Three PCGA inversions, with 10,000 parameters and using different K -order truncations of the covariance matrix corresponding to singular values (truncation errors) of $\lambda_{K+1} \simeq 1$ ($K = 69$), $\lambda_{K+1} \simeq 0.1$ ($K = 313$) and $\lambda_{K+1} \simeq 0.01$ ($K = 1532$), were performed. Figure 6 shows a relationship between the singular value and the truncation order, on which the position of the three K orders that were adopted in our inversions are indicated. Figure 7 shows the results obtained from these inversions. Note that very similar T fields (e.g., similar trend and location of high T zones) were obtained from the inversions using different truncation order. Therefore, the influence of reducing the

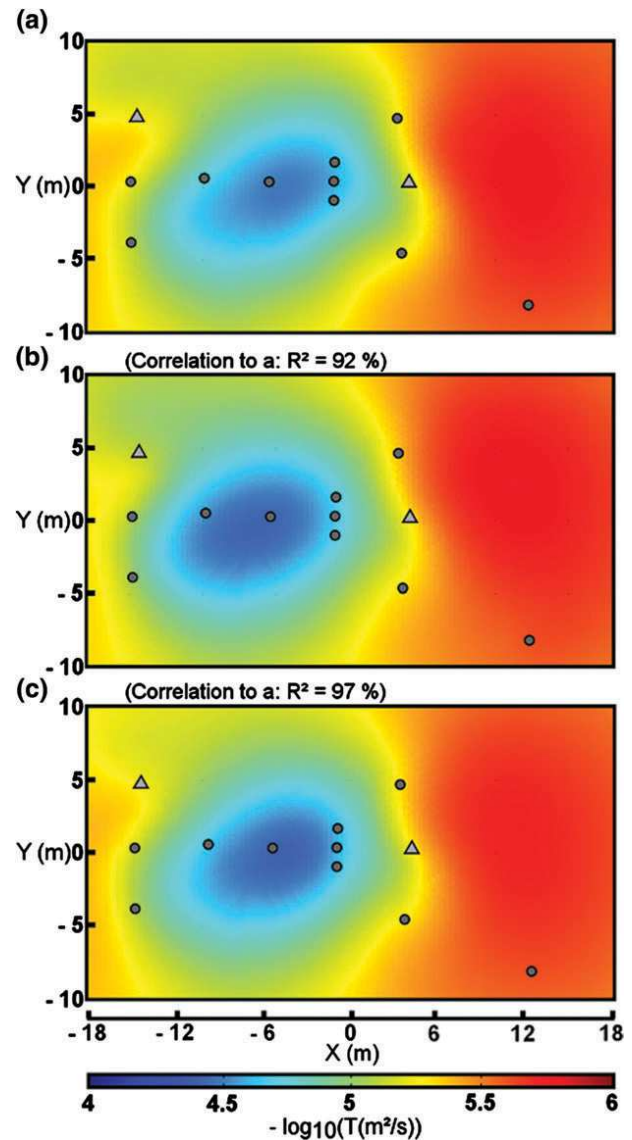


Figure 7. Maps of the log-transmissivity for a PCGA inversion method with 10,000 parameters, 20 observed data and three different covariance matrix decomposition applied to the experimental site. The map (a) was obtained for $K = 69$, the map (b) for $K = 313$ and the map (c) for $K = 1532$ (see Figure 6). The results obtained for these three decomposition are relatively the same (same transmissivity values, same zones) so, for this site, there is no significant loss of information when using a truncation order corresponding to an error of 1 (map (a)) for the covariance matrix which allows us to reduce the computation time of the inversion without decreasing the accuracy of the results.

K -order, as long as the truncation error is below 1, on the inversion results is mild. The most information regarding the spatial structure of the prior model is preserved in its first few singular values, so it is acceptable to consider a truncation order for a truncation error $\lambda_{K+1} \simeq 1$ for the covariance matrix.

PCGA and GA Results Comparison

In this section, we compare the inversion results obtained for the PCGA method using an approximated Q matrix, with those of the GA method using two different

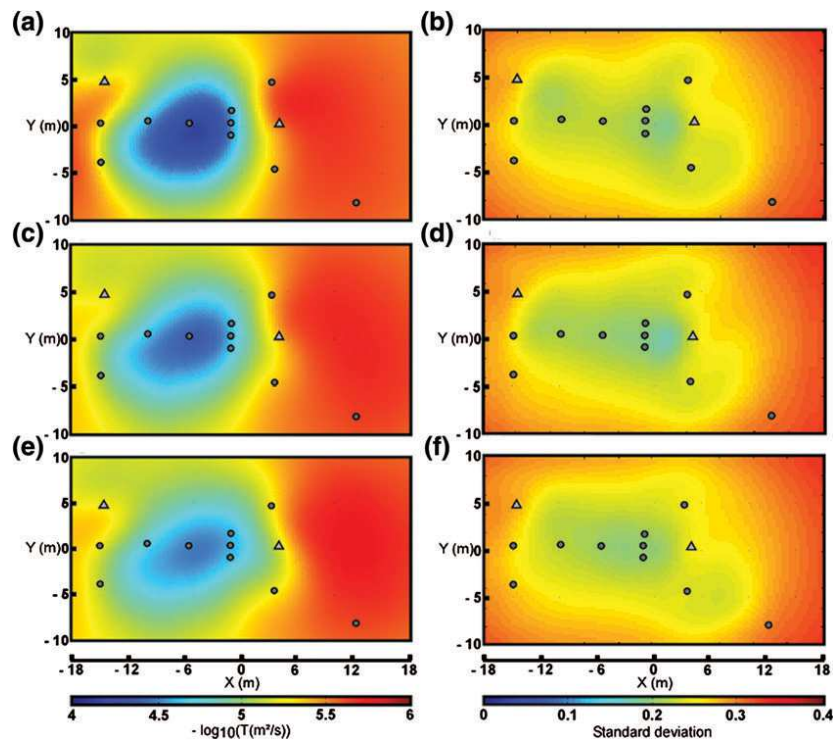


Figure 8. Maps of the log-transmissivity (a, c, e) and parameter's a posteriori standard deviation (b, d, f) for three different inversion methods with 10,000 parameters and 20 observed data applied to the experimental site. The maps (a) and (b) were obtained with the GA-adjoint-state method, the maps (c) and (d) with the GA-finite-difference method and the maps (e) and (f) with the PCGA method with a covariance matrix decomposition of order $K = 69$. The three methods provide relatively the same results for this site. The GA-adjoint-state method leads to a better data matching (see Figure 9) due to its slightly higher contrast of transmissivity distribution, but regarding the calculation time the PCGA inversion is much more efficient (see Table 2).

methods for Jacobian matrix computation (i.e., the finite-difference and adjoint-state methods), where the entirety of the \mathbf{Q} matrix were used. All the inverse simulations were performed on a 100×100 grid. Figure 8 shows the inverted transmissivity distributions and the corresponding distributions of standard deviation of each model parameter. It can be seen that in the three transmissivity fields, the calculated T value for each cell varies around the mean $10^{-5} \text{ m}^2/\text{s}$. In general, the three approaches produced similar spatial distribution of the transmissivity; however, the range of the inverted transmissivities from the GA-adjoint-state method (i.e., 2×10^{-6} to $1 \times 10^{-4} \text{ m}^2/\text{s}$) is larger than that of the PCGA and that of the GA-finite-difference method (i.e., 2×10^{-6} to $3 \times 10^{-5} \text{ m}^2/\text{s}$). Since the PCGA method is based on a finite-difference matrix vector product approximation, it is expected to obtain a similar parameter range from this method and the GA-finite-difference method. The difference between the results from these two methods and the GA-adjoint-state method may come from the definition of the finite-difference step δ . In addition, the resolved transmissivity field from PCGA is smoother compared to that from the method where an adjoint-state method is used to compute the sensitivity matrix. This is caused by the low-rank truncation in PCGA and also the finite-difference approach, which tends to reduce the heterogeneity of the inverted T field.

The spatial distributions of the standard deviation of the inverted parameters are shown in Figure 8. For each method, the uncertainty of the reconstruction is mainly dependent on the number and position of the wells. The correlation between inverted and measured hydraulic head data for the three inversion models are shown in the cross-plots of Figure 9. It can be seen that, with the same inversion inputs, the GA-adjoint-state method generated a slightly better correlation compared to the other two methods, which is reflected by the smaller RMSE values (0.165 m compared to 0.182 and 0.188 m). The performance of the three methods in terms of simulation time is compared in Table 2.

A significant reduction in computational time is observed for the GA-adjoint-state method compared to the GA-finite-difference method (Table 3). This reduction is mainly related to the calculation of the Jacobian sensitivity matrix (Cardiff and Kitanidis 2008). As the grid discretization increases, the significance of reduction in computational time of the GA-adjoint-state method compared to the GA-finite-difference-based method become more apparent. However, an even more significant time reduction was observed in using the PCGA method. Note that in Table 3 the computation time for PCGA includes the time from both the covariance matrix decomposition and the inversion calculation. The computational time of PCGA is observed to be 10

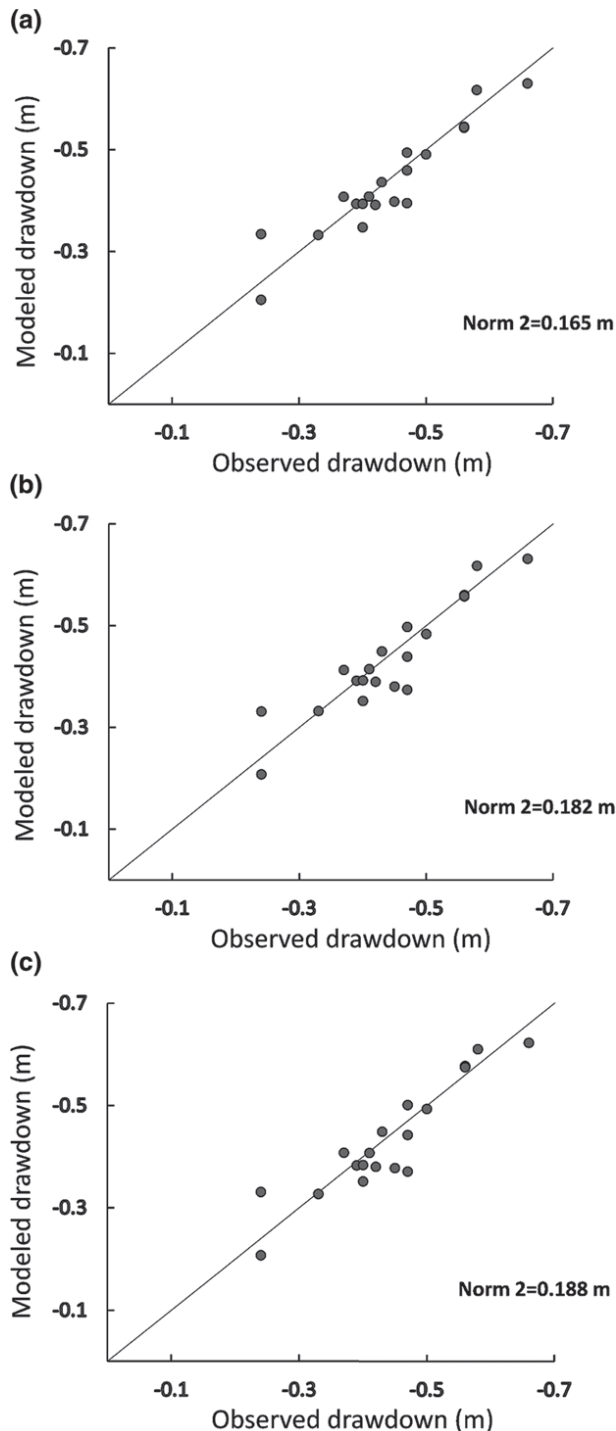


Figure 9. Graphs showing the differences between the 20 observed drawdowns and modeled drawdowns after convergence of three different inversion methods with 10,000 parameters applied to the experimental site. The graph (a) was obtained with the GA-adjoint-state method, the graph (b) with the GA-finite-difference method, and the graph (c) with the PCGA method with a covariance matrix decomposition of order $K = 69$. Regarding the mathematical norm 2 the GA-adjoint-state method has a slightly better convergence on the data than the other methods but the PCGA inversion method is much more efficient for the calculation time (see Table 2).

Table 2
Comparison of the Efficiency between Three Algorithm of Geostatistical Inversion Methods (GA-Adjoint-State, GA-Finite-Difference, and PCGA) on a Same Under-Determined Modeling. Results of These Inversions Are Shown in Figures 8 and 9. The Convergence on Data Was Slightly Better for an Adjoint-State Method but the Calculation Time Was Considerably Reduced by Using a PCGA Method

	GA-Adjoint-State Method	GA-Finite-Difference Method	PCGA Method (Truncation Order $K = 69$)
Number of parameters	100×100	100×100	100×100
Computation time	10 h 43 min	72 h 10 min	1 h 33 min
Value of objective function to be minimized after convergence	23.0602	26.6526	27.4641
ℓ_2 -norm between observed and modeled data	0.165 m	0.182 m	0.188 m

Table 3
Convergence Times for Different Methods Using Different Grid Sizes. An Intel Xeon QuadCore 2.8GHz with 12Go RAM Has Been Used to Perform the Computations. The PCGA Method (with a Truncation Error of Approximately 1) Is Always the Fastest because It Involves Less Forward Problems than the GA-Finite-Difference Method and That the Gauss-Legendre Resolution of the Integral in the GA-Adjoint-State Method Requires a Calculation of a Number of Nodes Proportional to the Number of Cells in the Grid in Each Forward Problem

Grid Resolution	GA-Adjoint-State Method	GA-Finite-Difference Method	PCGA Method (Truncation Error $\lambda_{K+1} \simeq 1$)
10×10	5 min	16 min	1 min
30×30	9 min	1 h 40 min	3 min
50×50	45 min	8 h 21 min	6 min
80×80	3 h 19 min	41 h 44 min	29 min
100×100	10 h 43 min	72 h 10 min	1 h 33 min

times less than that of the GA-adjoint-state method and 70 times less than that of the GA-finite-difference method. Altogether, the advantage of PCGA in obtaining a fast solution without compromising the inversion quality makes it a promising candidate in solving large-scale inversion problems.

Conclusion

The GA-finite-difference method is useful and straightforward for inversions with a few parameters and a large number of observational data. In contrast, the GA-adjoint-state method is advantageous in dealing with inversion models with a few observational data but a relatively large parameters set. On the contrary, the PCGA is an efficient method for both cases. It is also helpful for extremely under-determined problems where a large number of unknown parameters is present. In fact, the time and memory required by this method to perform the iterative process of the inversion is less sensitive to the number of parameters or measurements, but more dependent on the approximation order of the covariance matrix chosen by the modeler. A higher order approximation will lead to higher computational costs but the error introduced in the inversions will be much smaller.

The comparison of different methodologies has shown that PCGA approach appears to be the most efficient strategy for carrying out large-scale inversions in porous aquifers. We noticed that, in the studied case, the errors introduced by the approximation in the PCGA methods were not significant. Thus, we obtained the same accuracy in results from the PCGA inverse modeling compared to the GA inverse modeling. Thereby, with this strategy, only the principal components of the covariance matrix are kept in the inversion process, and the computational and memory costs necessary for the inversion algorithm are optimized. Additionally, the PCGA method significantly reduces the computational time. With the PCGA method we divided the computation time by seven compared to the GA-adjoint-state method, and by 50 compared to the GA-finite-difference method.

In summary, by applying the PCGA for a hydraulic tomography in a porous aquifer, we found an especially adapted strategy, which produces accurate inversion results with a good resolution in a reduced time, and which manages optimally the computer memory involved in the inversion algorithm. Nevertheless, the PCGA method is efficient specifically for models with a smooth distribution of the targeted parameters (which could typically be used for a good average representation of porous aquifers) so that the covariance matrix can be approximated by much smaller matrices.

Acknowledgments

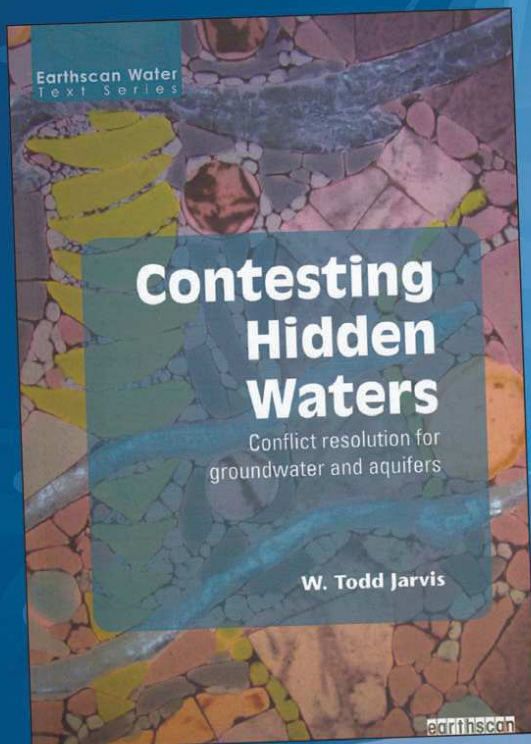
We would like to thank Pr. Michael Cardiff and two others anonymous reviewers for their relevant comments

on this article. We also thank the region Normandy for financially supporting the PhD of Pierre Fischer.

References

- Ambikasaran, S., J.Y. Li, P.K. Kitanidis, and E. Darve. 2013. Large-scale stochastic linear inversion using hierarchical matrices. *Computers & Geosciences* 17: 913–927.
- Berg, S.J., and W.A. Illman. 2013. Field study of subsurface heterogeneity with steady-state hydraulic tomography. *Groundwater* 51, no. 1: 29–40.
- Cardiff, M., and P.K. Kitanidis. 2008. Efficient solution of nonlinear, underdetermined inverse problems with a generalized PDE model. *Computers & Geosciences* 34: 1480–1491.
- Cardiff, M., W. Barrash, P.K. Kitanidis, B. Malama, A. Revil, S. Straface, and E. Rizzo. 2009. A potential-based inversion of unconfined steady-state hydraulic tomography. *Groundwater* 47, no. 2: 259–270.
- Cardiff, M., W. Barrash, and P.K. Kitanidis. 2013. Hydraulic conductivity imaging from 3-D transient hydraulic tomography at several pumping/observation densities. *Water Resources Research* 49: 7311–7326.
- Carrera, J., and S.P. Neuman. 1986a. Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information. *Water Resources Research* 22, no. 2: 199–210.
- Carrera, J., and S.P. Neuman. 1986b. Estimation of aquifer parameters under transient and steady state conditions: 3. Application to synthetic and field data. *Water Resources Research* 22, no. 2: 228–242.
- Elsheikh, A.H., M.F. Wheeler, and I. Hoteit. 2014. Hybrid nested sampling algorithm for Bayesian model selection applied to inverse subsurface flow problems. *Journal of Computational Physics* 258: 319–337.
- Greengard, L., and V. Rokhlin. 1987. A fast algorithm for particle simulations. *Journal of Computational Physics* 73, no. 2: 325–348.
- Hackbusch, W., and S. Börm. 2002. Data-sparse approximation by adaptive \mathcal{H}^2 -matrices. *Computing* 69, no. 1: 1–35.
- Halko, N., P.G. Martinsson, and J.A. Tropp. 2011. Finding structure with randomness: probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review* 53, no. 2: 217–288.
- Hoeksema, R.J., and P.K. Kitanidis. 1984. An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resources Research* 20, no. 7: 1003–1020.
- Kitanidis, P.K. 1995. Quasi-linear geostatistical theory for inverting. *Water Resources Research* 31, no. 10: 2411–2419.
- Kitanidis, P.K., and J. Lee. 2014. Principal component geostatistical approach for large-dimensional inverse problem. *Water Resources Research* 50, no. 7: 5428–5443.
- Kitanidis, P.K., and E.G. Vomvoris. 1983. A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations. *Water Resources Research* 19, no. 3: 677–690.
- Lee, J., and P.K. Kitanidis. 2014. Large-scale hydraulic tomography and joint inversion of head and tracer data using the principal component geostatistical approach (PCGA). *Water Resources Research* 50, no. 7: 5410–5427.
- Lieberman, C., K. Fidkowski, K. Willcox, and B. van Bloemen Waanders. 2013. Hessian-based model reduction: large-scale inversion and prediction. *International Journal for Numerical Methods in Fluids* 71: 135–150.
- Liu, X., Q. Zhou, P.K. Kitanidis, and J.T. Birkholzer. 2014. Fast iterative implementation of large-scale nonlinear geostatistical inverse modeling. *Water Resources Research* 50: 198–207.

- Nowak, W., S. Tenkleve, and O.A. Cirpka. 2003. Efficient computation of linearized cross-covariance and auto-covariance matrices of interdependent quantities. *Mathematical Geology* 35, no. 1: 53–66.
- Paige, C.C., and M.A. Saunders. 1975. Solution of sparse indefinite systems of linear equations. *SIAM Journal on Numerical Analysis* 12, no. 4: 617–629.
- Pool, M., J. Carrera, A. Alcolea, and E.M. Bocanegra. 2015. A comparison of deterministic and stochastic approaches for regional scale inverse modeling on the Mar del Plata aquifer. *Journal of Hydrology* 531: 214–229.
- Rao, S.V.N., B.S. Thandaveswara, S. Murthy Bhallamudi, and V. Srinivasulu. 2003. Optimal groundwater management in deltaic regions using simulated annealing and neural networks. *Water Resources Management* 17: 409–428.
- Soueid Ahmed, A., A. Jardani, A. Revil, and J.P. Dupont. 2014. Hydraulic conductivity field characterization from the joint inversion of hydraulic heads and self-potential data. *Water Resources Research* 50: 1–21.
- Soueid Ahmed, A., J. Zhou, A. Jardani, A. Revil, and J.P. Dupont. 2015. Image-guided inversion in steady-state hydraulic tomography. *Advances in Water Resources* 82: 83–97.
- Tarantola, A., and B. Valette. 1982. Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics and Space Physics* 20, no. 2: 219–232.
- Wang, X., A. Jardani, H. Jourde, L. Lonergan, J. Cosgrove, O. Gosselin, and G. Massonnat. 2016. Characterisation of the transmissivity field of a fractured and karstic aquifer, Southern France. *Advances in Water Resources* 87: 106–121.
- Zha, Y., T.J. Yeh, W.A. Illman, T. Tanaka, P. Bruines, H. Onoe, and H. Saegusa. 2015. What does hydraulic tomography tell us about fractured geological media? A field study and synthetic experiments. *Journal of Hydrology* 531, no. 1: 17–30.



Discover options for addressing the unique challenges of groundwater resource management

Unlike other books about conflict resolution and negotiations over water resources, this volume is unique in focusing exclusively on conflicts over groundwater and aquifers. Exploring the specific challenges presented by these “hidden” resources, the book describes issues surrounding the governance and management of groundwater, the various parties involved in conflicts and negotiations over them, and offers options for addressing the challenges and issues through a transdisciplinary approach.

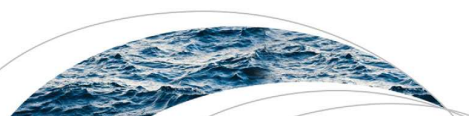
Catalog #T1107

NGWA member price \$39.00

Nonmember price \$49.00

Order your copy today.

NGWA.org/Bookstore • (800) 551-7379 • (614) 898-7791



RESEARCH ARTICLE

10.1002/2016WR019572

Key Points:

- A novel inverse approach is developed to reconstruct the structural heterogeneities
- The cellular automaton method is used to parameterize the inverse problem
- The inverse algorithm is validated on the hydrogeological and geophysical data

Correspondence to:

A. Jardani,
abderrahim.jardani@univ-rouen.fr

Citation:

Fischer, P., A. Jardani, and N. Lecoq (2017), A cellular automata-based deterministic inversion algorithm for the characterization of linear structural heterogeneities, *Water Resour. Res.*, 53, 2016–2034, doi:10.1002/2016WR019572.

Received 27 JUL 2016

Accepted 11 FEB 2017

Accepted article online 15 FEB 2017

Published online 11 MAR 2017

A cellular automata-based deterministic inversion algorithm for the characterization of linear structural heterogeneities

P. Fischer¹, A. Jardani¹, and N. Lecoq¹
¹Normandie Université, UNIROUEN, UNICAEN, CNRS, M2C, Rouen, France

Abstract Inverse problem permits to map the subsurface properties from a few observed data. The inverse problem can be physically constrained by a priori information on the property distribution in order to limit the nonuniqueness of the solution. The geostatistical information is often chosen as a priori information; however, when the field properties present a spatial locally distributed high variability, the geostatistical approach becomes inefficient. Therefore, we propose a new method adapted for fields presenting linear structures (such as a fractured field). The Cellular Automata-based Deterministic Inversion (CADI) method is, as far as we know when this paper is produced, the first inversion method which permits a deterministic inversion based on a Bayesian approach and using a dynamic optimization to generate different linear structures iteratively. The model is partitioned in cellular automaton subspaces, each one controlling a different zone of the model. A cellular automata subspace structures the properties of the model in two units (“structure” and “background”) and control their dispensing direction and their values. The partitioning of the model in subspaces permits to monitor a large-scale structural model with only a few pilot-parameters and to generate linear structures with local direction changes. Thereby, the algorithm can easily handle with large-scale structures, and a sensitivity analysis is possible on these structural pilot-parameters, which permits to considerably accelerate the optimization process in order to find the best structural geometry. The algorithm has been successfully tested on simple, to more complex, theoretical models with different inversion techniques by using seismic and hydraulic data.

1. Introduction

In geophysics, the inverse method is an efficient way for mapping the geological structures by assessing the physical properties of the subsurface (such as hydraulic conductivity, electrical resistivity, magnetic susceptibility, volumetric density, porosity, etc.) from a set of observed data. These observed data represent the responses of the investigated area to solicitations applied during the geophysical surveys (pumping tests, electrical resistivity tomography, electromagnetic, gravimetry, seismic, etc.). Commonly, the inverse problems are undetermined, with nonuniqueness of the solution, leading to provide doubtful interpretations of the geophysical surveys. Thus, the addition of a priori information on the properties to estimate is a necessity for avoiding the physically unrealistic models. Most often, geostatistical constraints are used to reconstruct the physical properties of a soil that can be modeled by smooth spatial variabilities [Hoeksema and Kitanidis, 1984]. However, when the parameters have a high spatial variability, the use of statistical characteristics as a priori information becomes ineffective and inadequate to locate the discontinuities of the physical properties. Therefore, several algorithms have been proposed to deal with the “structural” inversion, considering both the estimation of physical properties and reconstruction of boundaries between different heterogeneities.

Among these approaches, we cite those which incorporate structural information in the model parameterization of the inverse problem, such as the multiscale method that rests on an increasing resolution of the parameterization during the optimization sequences [Grimstad et al., 2003]. The adaptive multiscale method permits to reduce the number of unknown parameters by a local refinement of the parameterization where the heterogeneity is the most important, to avoid an overparameterization. Tsai et al. [2003] used the Voronoi zonation with a pilot-point parameterization method to identify parameters structures in a model.

For the approaches using no specific parameterization of the model, Lelièvre and Oldenburg [2009] have proposed to incorporate constraints to the inversion objective function in terms of some structural information such as orientation to obtain more realistic solutions. The spatial distribution of the unknown parameters

can also be approximated by a sparse set of coefficients to be identified in a compressed sensing sparsity-promoting inversion [Jafarpour *et al.*, 2010] which promotes sparse solutions. Finally, the inversion with total variation prior [Lee and Kitanidis, 2013] uses a Laplace prior instead of a Gaussian in a Bayesian inversion in order to delimitate the shapes of discrete structures piloted by some hyperparameters determined during the inversion using an expectation-maximization approach.

In another register, Lochbühler *et al.* [2015] used the training image method in the inverse formulation to represent the structural characteristics of a field as prior information to eliminate inversion artifact and improve the estimate of the parameters. Hale [2009] and Soueid Ahmed *et al.* [2015] have proposed the guided image method in which the structural features of the domain is presented graphically and used as a priori information to guide the inversion by refining the model sensitivity at boundaries between different zones. It permits a better estimate of the intrastructure parameter variabilities and location of different features in a model. The level set is an alternative approach to detect the interfaces between different facies thanks to the use of extensible boundaries that move during the inversion process to fit the observed data [Lu and Robinson, 2006; Cardiff and Kitanidis, 2009]. Haber and Oldenburg [1997] have identified the profits that could bring joint inversion to structural identification and have presented a protocol to run a joint inversion in geophysics by constraining the results with a unique structural consideration. Since then, several other structural joint inversions tools have been developed which were summarized in a review proposed by Gallardo and Meju [2011]. This review presents the recent techniques of structural joint inversions and the upcoming challenges of such inversions in the next years.

However, regarding the imaging of linear structures, which are characterized by an aperture significantly lower than their length (such as karst conduits and fractures), the deterministic inversion remains, according to our knowledge, an unexplored subject. The inverse modeling of such structures requires a large-scale parameterization, which makes the computation very heavy particularly in the case of stochastic or global optimization algorithms [Pardo-Igúzquiza *et al.*, 2012; Reeves *et al.*, 2013; Bruna *et al.*, 2015; Javadi *et al.*, 2016].

In this paper we propose a new method called Cellular Automata-based Deterministic Inversion (CADI), adapted for the inversion of linear structures. This approach is based on a Bayesian formulation with the use of Cellular Automaton (CA) concept to parameterize the model. The dynamic structural optimization in the algorithm is controlled by the CA, which allows the drive of an entire discretized system with only some local configurable interaction rules. After a global presentation of the model parameterization, we will detail the interaction rules chosen for the CA in this CADI algorithm to reproduce linear structures (section 2.1). Then we will present the structure of the inverse problem (section 2.2) and the protocol of optimization (section 3). Several examples conducted with the CADI algorithm on linear problems (seismic) and nonlinear problems (hydraulic tomography) are described in the last part of this paper (section 4).

2. Parameterization of Inverse Problem Using Cellular Automaton

2.1. Parameterization of the Model

In the CADI method, the distribution of the properties in the model is structurally generated by several CA. In previous works, CA have already been coupled to global optimization algorithms such as genetic algorithms [Dewri and Chakraborti, 2005; Ghosh *et al.*, 2009]. However, in the CADI method we wanted to couple the possibilities offered by CA to a deterministic inverse process. Therefore, the model (discretized in m cells) is partitioned in m_{CA} CA subspaces (with $m_{CA} \ll m$), each one being monitored by an independent CA configured by its neighborhood definition (Figure 1). Thus, the CA subspaces and their parameterization are pilot zones for the model, which permits to avoid an overparameterization of the inverse problem.

The CA is a widely used mathematical system to generate discrete dynamic models. It has been applied to diverse fields of modeling such as Random Number Generators [Tan and Guan, 2007], chemical reactions [Van der Weeën *et al.*, 2011], solid-solid phase transformation during heating [Halder *et al.*, 2014, 2015] or cooling [Dewri and Chakraborti, 2005; Ghosh *et al.*, 2009; Jin and Cui, 2012], fluid flow through fractures [Pan *et al.*, 2011], or transport in fluid flow [Chopard and Masselot, 1999]. The CA is a popular method due to its capacity to model complex systems by using simple rules. In fact, the evolution of the entire system is driven by some configurable local interaction rules.

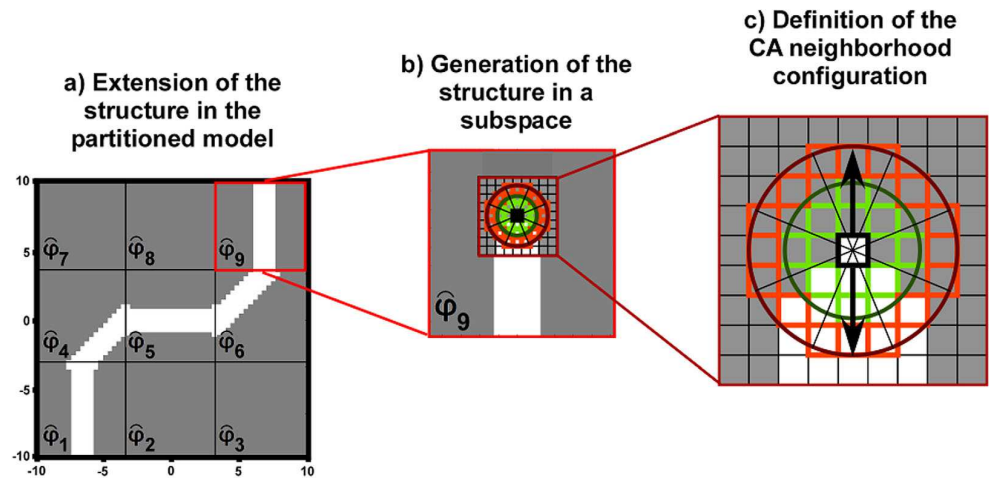


Figure 1. Scheme explaining how the CA are used in the CADI model. In the figure gray occurs for state “background” and white for state “structure.” The model is partitioned in m_{CA} independent CA subspaces (here $m_{CA} = 9$). During the generation process the structure will go through different CA subspaces (a) and will be generated in the local direction assigned by the structural parameters piloting these CA (b). Along the generation direction the CA will modify the property values of the model cells it controls (represented by the squares lattice in Figure 1c).

The CA is a discrete time-evolving system in which a cell of the system is defined by its state and its neighborhood. At a CA time step $t_{CA} + 1$, the state of each cell will simultaneously evolve to a new one depending on a constant transition rule involving the state of a cell and the states of the cells in its neighborhood at the time step t_{CA} . Thus, a CA subspace of the model proposed in the CADI algorithm can be described as a quadruple $\langle \varphi, Z, N, T \rangle$ [Sun et al., 2011; Van der Weeën et al., 2011]:

1. φ is a subspace of the global model, discretized as a lattice of c cells.
2. Z is a function returning the states values for each selected cells of the subspace at a specified time step (with two possible states: $Z_{t_{CA}}(c_i) = \beta_{background}$ or $Z_{t_{CA}}(c_i) = \beta_{structure}$, where $\beta_{background}$ and $\beta_{structure}$ are the parameterized values assigned to the properties of the structure and the background in the subspace).
3. N is a neighborhood function that selects among all cells of the subspace the subset of cells that are considered in the neighborhood of a given cell c_i .
4. T is a function of cell-state transition rule. Thus, a transition in the CA for a given cell c_i is expressed as $Z_{t_{CA}+1}(c_i) = T(Z_{t_{CA}}(N(c_i)))$, and a full transition in the CA process (considering all cells of the subspace lattice) is $\varphi_{t_{CA}+1} = Z_{t_{CA}+1}(c_i), \forall i$.

The choice of N and T for the CA in this work will be detailed in the following paragraphs. The CA will be used to produce a spatial linear structure in the model. The global model is partitioned in several subspaces, each one being discretized as a lattice of c squared cells (Figure 1c). Each cell of a subspace φ can be in only two possible different states: state “background” which take a value $\beta_{background}$, or state “structure” which take a value $\beta_{structure}$. So, a subspace has a binary distribution. $\beta_{structure}$ is homogeneous within a subspace but can vary among the different subspaces.

Commonly, CA use neighborhood sequencing such as the Moore or the Von Neumann neighborhood rules (see Appendix 1) [Moore, 1962; Von Neumann and Burks, 1966]. But here we chose N as a dual-radius neighborhood definition as presented in Figure 2. Two circles of cells, defined by their cell-radius R_{inner} and R_{outer} , are centered on a given cell c_i (for a full CA time step transition c_i would be alternatively each cell of the CA subspace). The inner circle defines the “activator” cells for c_i (green in Figure 2) and the outer circle defines the “inhibitor” cells for c_i (orange in Figure 2). The terms “activators” and “inhibitors” are relative only to the cells in “background” state: the cells in state “background” in the “activator” neighborhood will tend to transform the cell c_i in a state “background” while the cells in state “background” in the “inhibitor” neighborhood will tend to transform the cell c_i in a state “structure.” The balance of the ratio of “background” cells in each “activator” and “inhibitor” neighborhoods can be disturbed by the existence of cells in state “structure” (for example the presence of a cell in state “structure” in the “activator” neighborhood of a cell c_i).

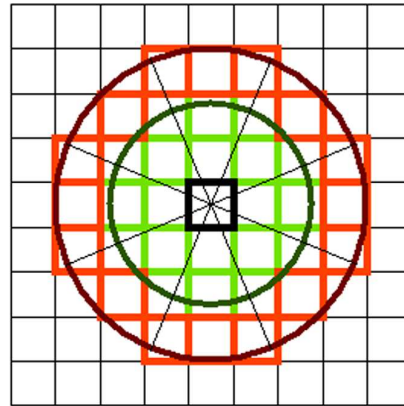


Figure 2. An example of the dual-radius neighborhood considered in our CA definition. The black highlighted cell is the cell under consideration in this example (each cell of the lattice would alternatively be considered during a full CA time step). The green highlighted cells are considered as its “activators” neighbors in the transition rule and the orange highlighted cells as its “inhibitors” neighbors. These cells are selected by an inner and an outer circle (in bold) with configurable radius which permit the configuration of the neighborhood. In this example, the inner circle has a radius = 2 and the outer circle has a radius = 3. Additionally, the neighborhood is split into 2×8 sectors (by the radial lines) which permit a more configurable weighting definition (see Figure 4).

will reduce the amount of cells in state “background” in this “activator” neighborhood compared to in the “inhibitor” neighborhood, and thus the cell c_i would then become “structure”). The cells outside of the circles are not considered for the state transition of the cell c_i . This definition can be seen as an extension of the Von Neumann rules, by adding inhibitors neighbors to activators cells that follow a Von Neumann neighborhood. Additionally, each circle of the neighborhood definition was radially split into eight sectors to allow for spatially variable weighting of activators and inhibitors cells, in order to privilege particular directions during the generation (see Figure 4).

During the transition to $t_{CA} + 1$, the transition rule T defines alternatively for each cell of the lattice the new state of a cell c_i by considering the equilibrium of activators and inhibitors cells in state background ($\beta_{background}$) in its neighborhood $N(c_i)$ at the instant t_{CA} . Therefore, the neighborhood configuration associates cells values in the activator zone to a positive weighting (+) and cells values in the inhibitor zone to a negative weighting (−). The weighted values in each of eight activator and inhibitor sectors are then also corrected by an additional balancing weight (ratio between the number of cells in a sector and the total number of cells), in order to have the same consideration between each sector of the neighborhood. In fact, each sector does not contain the same amount of cells, due to the consideration of deformations of circles in a lattice of squares. Finally, the transition rule T sums the weighted val-

ues from all cells in state $\beta_{background}$ in the neighborhood $N(c_i)$. If the total weight of activators in state $\beta_{background}$ is higher (the sum is positive), the cell c_i will take the value $\beta_{background}$ (“background”), if the total weight of inhibitor in state $\beta_{background}$ is higher (the sum is negative), the cell c_i will take the value $\beta_{structure}$ (“structure”)

$$T(c_i) = \sum_{k=1}^{ncell} Z_{t_{CA}}^{weight}(c_k) \text{ for } c_k \in [N(c_i)] \cap [Z_{t_{CA}}(c_k) = \beta_{background}] \quad (1)$$

$$Z_{t_{CA}+1}(c_i) = \beta_{background} \text{ if } T(c_i) \geq 0, \quad Z_{t_{CA}+1}(c_i) = \beta_{structure} \text{ if } T(c_i) < 0,$$

where $ncell$ denotes the total number of cells in $N(c_i)$ and $Z_{t_{CA}}^{weight}$ is the function returning the state value taking into account the weighting parameterization from the neighborhood.

After sufficient time steps of the CA with the same transition rule, the subspace φ will converge to a stable geometry $\hat{\varphi}$ (the geometry will not change over increasing CA time steps anymore) depending on the weighting parameterization given to the neighborhood definition N (Figure 3).

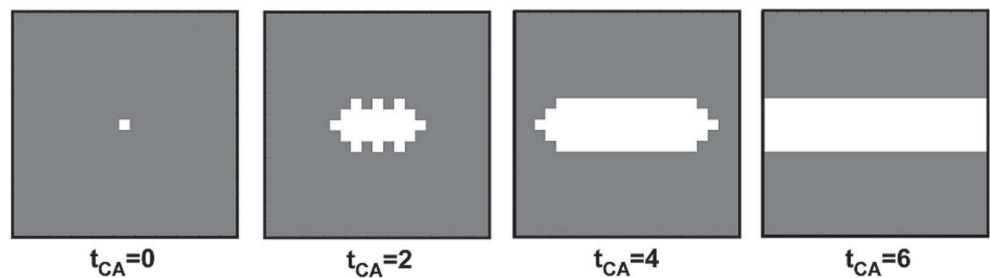


Figure 3. Time evolution of a CA configured with a neighborhood weighting defining a horizontal structure generation (see Figure 4). After the sixth time step the CA has converged and its geometry is stable over the following steps. Here gray occurs for state “background” and white for state “structure.”

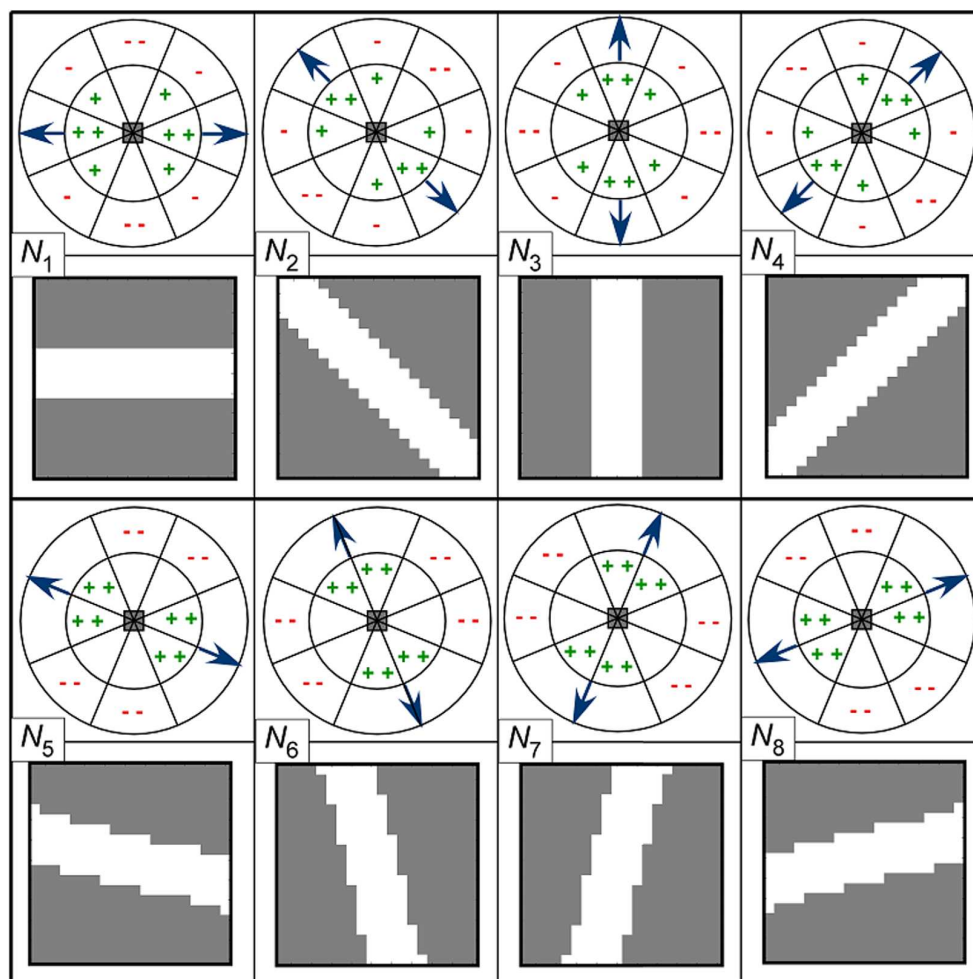


Figure 4. Presentation of eight different stable structures started by a unique centered cell, and their associated CA neighborhood configuration. The grayed cell in the neighborhood configuration is a given cell considered during the CA process. It is surrounded by its neighbor cells, which are not shown for reasons of readability. Its neighborhood is split in eight internal “activator” sectors and eight external “inhibitor” sectors, each one being assigned to a given weight. A “+ +” occurs for a positive weight for the neighbor cells in the area, a “+ +” weight is twice higher than a positive weight represented by a single “+.” A “- -” occurs for a negative weight for the neighbor cells in the area, a “- -” weight is twice higher than a negative weight represented by a single “-.” An empty part of the neighborhood occurs for a null weight, meaning that cells in the area are not considered in the transition rule. Here we present the CA configuration leading to eight different structure directions which will be considered as suborientation of the global structure in the model. In the structural map, gray occurs for state “background” and white for state “structure.”

Thus, playing on the weighting distribution in the divided activator and inhibitor sectors, and on the radius R_{inner} and R_{outer} of the neighborhood definition N , the CA can produce linear structures in eight directions from a unique starting cell, as shown in Figure 4. The weighting distribution defining each direction has been empirically specified.

On the presented configurations, the starting cell is considered in the center of the lattice. The neighborhood weighting permits to modify the direction of the structure and the radius values modify its aperture. These eight weighted neighborhood configuration functions $N_i, i \in [1, 8]$ will be considered as the different configuration possibilities in the subspaces parameterization in the dynamic structural optimization process of the inversion algorithm (presented in the section 3.1). Thus, in the CADI algorithm a CA subspace of the model is parameterized by two parameters: its structural direction (neighborhood configuration N) among the eight possible and its values of property $\beta = [\beta_{background}, \beta_{structure}]$. Therefore, a converged configuration of a parameterized subspace will be expressed as $\hat{\varphi}(N, \beta)$ (using one of the eight different direction configuration N_i , as presented in Figure 4).

The geometry of a structure over the entire model Γ (composed of all converged subspaces $\hat{\varphi}_k(N, \beta), k \in [1, m_{CA}]$) can be defined in the CA generation process with only two “pilot” vectors containing each subspaces parameters. This piloted model can be expressed as $\Gamma(P_N, P_\beta)$, where P_N is a m_{CA} vector containing the direction of generation (configured by the weighted neighborhood configurations $N_i, i \in [1, 8]$) assigned to each CA subspace of the model, and P_β a $m_{CA} + 1$ vector containing the $\beta_{structure}$ values assigned to the “structure” cells in each of the m_{CA} CA subspaces and also the $\beta_{background}$ value (the background being considered, in this paper, as uniform, but it could also be possible to consider a $\beta_{background}$ value for each subspace). Thus, by piloting the CA generation process with only P_N and P_β as parameters we can generate the whole model as shown in Figure 1. The aperture all along the structure is considered as constant and can be configured with the CA neighborhood radius values and the partitioning of the model.

The CA generation process of the structure starts from an entire “background” ($\beta_{background}$) state model with only one or several selected cell(s) of the model in state structure ($\beta_{structure}$) which are considered as the starting point(s) of the structure. At the firsts CA time steps, the structure will be generated only in the subspaces where initial structure cells are defined. Each boundary cells state at the edge of a CA subspace is symmetrically transferred to the boundary cells of the adjacent CA. Therefore when the structure arrives to the limit of its first subspace, it can enter a new CA subspace by local symmetry at the boundary limit between them. The new CA subspace the structure has entered has potentially another neighborhood definition; thus, the structure will follow a new direction from there. Once the structure has been generated in a subspace, this subspace becomes “inhibited” to another generation (the structure can enter only one time each subspace). And so the structure will propagate within the model, through the increasing CA time steps, until it reaches a stable geometry $\Gamma(P_N, P_\beta)$ (see Figure 1).

2.2. Statement of Inverse Problem

The inverse problem involves a formulation of the forward problem which links the spatial properties of the model to the data

$$d = f(\Gamma(P_N, P_\beta)) + \varepsilon, \quad (2)$$

where $\Gamma(P_N, P_\beta)$ is the spatial distribution of the m properties cells in the model. The cells of the model take their values from a finite set P_β and are structured by the CA directions P_N , d is a vector of n modeled data, f is a forward problem application $\mathbb{R}^m \rightarrow \mathbb{R}^n$, and ε represents the observed data error.

In a probabilistic framework, the aim of the inverse problem is to find the most probable models considering P_N and P_β as parameters constrained by the observed data and the prior information on both parameters. This inverse issue can be treated as a sequential inversion. First, for a given P_β , we determine the geometry of the structure via the estimation of P_N , which is then used in the second time to infer the values of P_β . Using a Bayesian approach on Gaussian probability density functions, the problem can be formulated by two posterior probability densities $\rho_{structure}(P_N|d, P_\beta)$ and $\rho_{properties}(P_\beta|d, P_N)$ in order to image the geometry of structure controlled by P_N and their physical property values controlled by P_β

$$\begin{aligned} \rho_{structure}(P_N|d_{obs}, P_\beta) &\propto \rho(d_{obs}|P_N, P_\beta) \cdot \rho(P_N) \\ &\propto \exp\left(-\frac{1}{2}(d_{obs} - f(\Gamma(P_N, P_\beta)))^T C_d^{-1}(d_{obs} - f(\Gamma(P_N, P_\beta)))\right), \\ &\times \exp\left(-\frac{1}{2}(P_{N,prior} - P_N)^T C_{P_N}^{-1}(P_{N,prior} - P_N)\right) \end{aligned} \quad (3)$$

$$\begin{aligned} \rho_{properties}(P_\beta|d_{obs}, P_N) &\propto \rho(d_{obs}|P_\beta, P_N) \cdot \rho(P_\beta) \\ &\propto \exp\left(-\frac{1}{2}(d_{obs} - f(\Gamma(P_N, P_\beta)))^T C_d^{-1}(d_{obs} - f(\Gamma(P_N, P_\beta)))\right), \\ &\times \exp\left(-\frac{1}{2}(P_{\beta,prior} - P_\beta)^T C_{P_\beta}^{-1}(P_{\beta,prior} - P_\beta)\right) \end{aligned} \quad (4)$$

with ρ denotes the Gaussian probability density function. $\rho(d_{obs}|P_N, P_\beta)$ is the likelihood function; $\rho(P_N)$ and $\rho(P_\beta)$ represent a priori information on the parameters P_N and P_β . d_{obs} is the $(1 \times n)$ vector of observed data. $P_{\beta,prior}$ and $P_{N,prior}$ are the prior models (parameter assumptions) on the unknown parameters P_β and

P_N , chosen by the modeler to constrain the inversion. C_d ($n \times n$), C_{P_N} ($m_{CA} \times m_{CA}$), and C_{P_β} ($(m_{CA}+1) \times (m_{CA}+1)$) are covariance matrices of the expected uncertainties on data and the prior models. C_{P_N} and C_{P_β} can also be seen as weighting values in the objective function to constrain the inversion result to have subspaces property values and directions remaining close to the parameters chosen in the prior models. The maximization of the posterior probability densities (equations (3) and (4)) can be achieved by a minimization of the following objective functions in the inversion process [Tarantola and Valette, 1982]:

$$\Psi_{structure}(P_N) = \frac{1}{2} (d_{obs} - f(\Gamma(P_N, P_\beta)))^T C_d^{-1} (d_{obs} - f(\Gamma(P_N, P_\beta))) + \frac{1}{2} (P_{N,prior} - P_N)^T C_{P_N}^{-1} (P_{N,prior} - P_N), \quad (5)$$

$$\Psi_{properties}(P_\beta) = \frac{1}{2} (d_{obs} - f(\Gamma(P_N, P_\beta)))^T C_d^{-1} (d_{obs} - f(\Gamma(P_N, P_\beta))) + \frac{1}{2} (P_{\beta,prior} - P_\beta)^T C_{P_\beta}^{-1} (P_{\beta,prior} - P_\beta). \quad (6)$$

This minimization can be achieved iteratively with sequential optimizations on the geometry of the structure and on the values taken by the properties. The convergence of these two objective functions to their minimal values depends on the parameterization of the model and its initialization. A global minimization is not guaranteed, as the result of the inversion depends of the initial model. However, the optimum can be explored by leading several inversions starting from different initial models. The different steps of the CADI algorithm are presented in Figure 5. The structural and property values parameters in P_N and P_β are first initialized to generate the initial model. This initialization consists in assigning an initial reasonable direction of generation and initial property values $\beta_{structure}$ and $\beta_{background}$ to each CA subspaces in the model. After this initialization part, a sequential inversion process will first conduct an iterative structural optimization in which the CA structural generation process will regenerate the model with the updated parameters at each iteration. Once this optimization is completed, the inversion will continue with an optimization of the property values for the previously inverted structure. The process is then ended by an estimation of uncertainties on the structure geometry and on the property values. These different parts of the inversion process are detailed in the following sections.

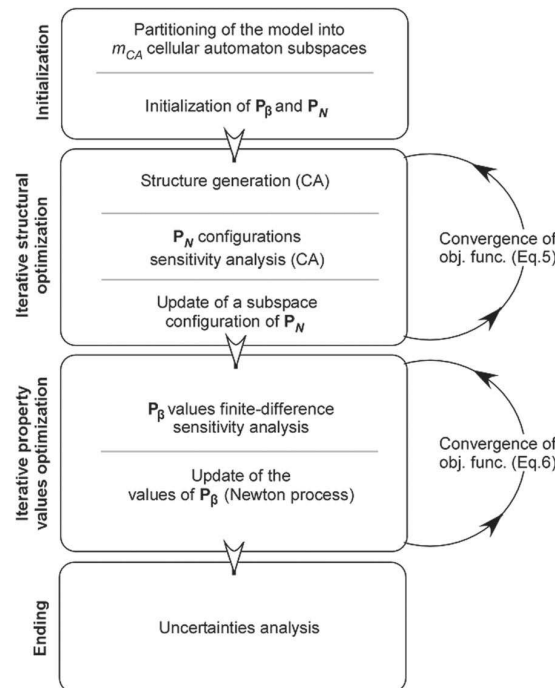


Figure 5. Operating scheme for the Cellular Automata-based Deterministic Inversion (CADI) algorithm. After an initialization of P_N and P_β with chosen directions and property values for each subspace, the algorithm begins an iterative process. It will first optimize the geometry of the structure in the model by iteratively updating the structural model using the CA generation process. Once the objective function has converged to a local minimum on the structure, it will lead a second optimization on the values of the properties for the previously inverted structure, until the objective function converges to a local minimum again. Finally, the uncertainties on the structure and the properties of the model are estimated.

initialized to generate the initial model. This initialization consists in assigning an initial reasonable direction of generation and initial property values $\beta_{structure}$ and $\beta_{background}$ to each CA subspaces in the model. After this initialization part, a sequential inversion process will first conduct an iterative structural optimization in which the CA structural generation process will regenerate the model with the updated parameters at each iteration. Once this optimization is completed, the inversion will continue with an optimization of the property values for the previously inverted structure. The process is then ended by an estimation of uncertainties on the structure geometry and on the property values. These different parts of the inversion process are detailed in the following sections.

3. Optimization Process

3.1. Structural Optimization

Initially, a chosen set of probable property values $P_{\beta,ini}$ and chosen direction configurations $P_{N,ini}$ are assigned to the piloted model to build an initial model. The aim of the structural optimization will be to modify iteratively the structure piloted by P_N for a given distribution of P_β until the convergence of the objective function (equation (5)). The modification of the configuration P_N is defined through a sensitivity analysis.

At an iteration step k , the sensitivity on the structure is estimated by introducing "perturbations" in the generation and by analyzing the responses by solving the forward problem. Here a perturbation consists in a modification of the configuration N in a CA subspace of the model (i.e., a local direction modification), the other subspaces configurations remaining unchanged. The structural inversion sensitivity analysis tests the eight configuration possibilities $N_i, i \in [1, 8]$ as parameter in each CA subspace in order to optimize the geometry of the structure regarding the objective function. This sensitivity analysis is led on the entire model to create a $8 \times m_{CA}$ sensitivity matrix S . Thus, at the k th iteration and for a perturbation using a configuration N_i in a subspace j , the element (i, j) of the sensitivity matrix is defined as

$$S^k(i, j) = \frac{1}{2} \left(d_{obs} - f \left(\Gamma \left(P_N^k |_{P_N^k(j)=N_i}, P_\beta \right) \right) \right)^T C_d^{-1} \left(d_{obs} - f \left(\Gamma \left(P_N^k |_{P_N^k(j)=N_i}, P_\beta \right) \right) \right) + \frac{1}{2} (P_{N,prior}(j) - N_i)^T C_{P_N}^{-1} (P_{N,prior}(j) - N_i), \quad (7)$$

where $f \left(\Gamma \left(P_N^k |_{P_N^k(j)=N_i}, P_\beta \right) \right)$ represents the modeled data through this perturbation for a given model of P_β , and $P_{N,prior}(j) - N_i$ represents the angular gap between the prior subdirection and the perturbation direction. Here the sensitivity analysis does not involve variations in P_β , it determines all possible variations of the objective function for a single modification in the structure geometry.

The best improvement is found with the index $(i, j)_{min}$ in matrix S , representing the minimal value in the matrix which will give the best improvement for the minimization of the objective function (i gives the updated configuration N_i for the CA in the subspace j of the model). Thus, from a structural parameter set P_N^k , the optimized set P_N^{k+1} is built as $P_N^{k+1} = P_N^k$ except for its index j : $P_N^{k+1}(j) = N_i$. By updating the subspace which gives the best structural improvement, a new structure will be generated for the iteration $k + 1$. The algorithm reproduces the same sensitivity analysis for each iteration until the convergence of the objective function. The total number of forward problems evaluations for a structural iteration is $8 \times m_{CA} + 1$ (with $8 \times m_{CA}$ evaluations for the sensitivity analysis and 1 for the updated objective function calculation).

At the end of the inversion process, the uncertainties on the inverted structure are estimated through an uncertainties analysis on each subspace of the structure. Due to difficulty to infer the posterior covariance matrix of the structural inversion, this analysis is done for each subspace by inverting the difference between the posterior objective function and the sum of sensitivity values for all CA configuration possibilities, and the prior uncertainties for the subspace j

$$S^{post}(j) = \left(\frac{1}{8} \sum_{i=1}^8 S(i, j) - \Psi_{structure}^{post} + C_{P_N}^{-1}(j, j) \right)^{-1}, \quad (8)$$

with $S^{post}(j)$ is the posterior structural uncertainty for a subspace j in the model, i denotes the different rows of the matrix S of the last iteration, and $\Psi_{structure}^{post}$ is the value of the minimized objective function after convergence.

If a subspace is well-constrained, its value S^{post} should be low (another structure direction would have a negative impact in the minimization of the objective function), and if not, this value should be high (another structure direction would be quite neutral in the minimization of the objective function).

3.2. Property Values Optimization

Once the structure is optimized, the property parameters P_β taken by the CA subspaces of the model are then also iteratively optimized, for the inverted structure; using a finite difference approach for the sensitivity analysis (for $m_{CA} + 1$ unknown property values to optimize, including m_{CA} $\beta_{structure}$ values plus one common value for $\beta_{background}$). The Jacobian sensitivity matrix $J(n \times (m_{CA} + 1))$, for an index (i, j) is defined as

$$J(i, j) = \left. \frac{\partial f_i}{\partial P_\beta} \right|_{P_\beta(j) = P_\beta(j) + \Delta P_\beta}, \quad (9)$$

with f_i the forward problem on a data i for a variation ΔP_β of $P_\beta(j)$. Here ΔP_β is the finite difference step.

The new values P_β^{k+1} from a previous set P_β^k are calculated from a linearization of (equation (6))

$$f(\mathbf{P}_{\beta}^{k+1}) \approx f(\mathbf{P}_{\beta}^k) + \mathbf{J}^k \cdot (\mathbf{P}_{\beta}^{k+1} - \mathbf{P}_{\beta}^k). \quad (10)$$

The optimization of \mathbf{P}_{β} is achieved via a Newton iterative process, initialized at a reasonable $\mathbf{P}_{\beta, \text{ini}}$ [Tarantola and Valette, 1982]. For the $k + 1$ th step in the iterative process

$$\begin{aligned} \mathbf{P}_{\beta}^{k+1} = & \mathbf{P}_{\beta}^k + \left((\mathbf{J}^k)^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{J}^k + \mathbf{C}_{\mathbf{P}_{\beta}}^{-1} \right)^{-1} \cdot (\mathbf{J}^k)^T \cdot \mathbf{C}_d^{-1} \cdot (\mathbf{d}_{\text{obs}} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_{\beta}^k))) \\ & + \mathbf{C}_{\mathbf{P}_{\beta}}^{-1} \cdot (\mathbf{P}_{\beta, \text{prior}} - \mathbf{P}_{\beta}^k). \end{aligned} \quad (11)$$

The total number of forward problems in an iteration for the property value sensitivity evaluation in a model with a uniform background will be $m_{CA} + 1$, while for a model with a varying background among the CA subspaces it would require $2 \times m_{CA}$ evaluations. Then the compute of the updated objective function requires one more forward problem evaluation.

The uncertainties on the values of properties, calculated at the end of the inversion process, are given by the diagonal entries of the posterior covariance matrix

$$\mathbf{C}_{\mathbf{P}_{\beta}}^{\text{post}} = \left((\mathbf{J}^{\text{post}})^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{J}^{\text{post}} + \mathbf{C}_{\mathbf{P}_{\beta}}^{-1} \right)^{-1}. \quad (12)$$

These values represent variances of the properties. Then, the square root of the diagonal entries represent their standard deviation.

4. Applications

The CADl algorithm has been tested on six theoretical study cases for a linear inversion of a simple structure (Study case 1), a more complex structure (Study case 2), a complex multidirectional structure (Study case 3) and for a linear, nonlinear and joint inversions (Study cases 4–6) of a geostatistical generated structure. For these different examples we did not use any prior information on the structure in $\mathbf{P}_{N, \text{prior}}$ but we incorporated constant measurement errors in a diagonal matrix $\mathbf{C}_d = \sigma_{\text{data}}^2 \cdot \text{Id}(n)$, and prior background and structure property values in $\mathbf{P}_{\beta, \text{prior}}$ with their covariances in a diagonal matrix $\mathbf{C}_{\mathbf{P}_{\beta}} = \sigma_{\beta}^2 \cdot \text{Id}(m_{CA} + 1)$. These six study cases and their results are presented in Table 1, and the theoretical true structures to be reproduced are presented in Figure 6.

4.1. Study Case 1

The first study case is a linear inversion of a simple structure. The purpose of this study case is essentially to illustrate how the optimization within the CA subspaces in the structural inversion works. For the linear inversion, we considered seismic data. The properties taken into account in the model are the seismic

Table 1. Inversion Results Obtained for the Six Different Study Cases^a

	Inversion Type	Number of Cells (m)	Number of Data	Number of Iteration	Data R ²	Structural Similarity	Inversion Time
		CA Grid	Error Variance				
Case 1 (Figure 5a)	Linear	3,600 (3 × 3)	358 $\sigma_{\text{data}}^2 = 1 \text{ ms}$	4	0.99	99.7%	4 min
Case 2 (Figure 5b)	Linear	10,000 (5 × 5)	598 $\sigma_{\text{data}}^2 = 1 \text{ ms}$	21	0.96	97.9%	1.3 h
Case 3 (Figure 5c)	Linear	48,400 (11 × 11)	1,318 $\sigma_{\text{data}}^2 = 1 \text{ ms}$	Nolnit: 26 Init: 30	0.91 0.96	97.1% 98.3%	13.2 h 18.8 h
Case 4 (Figure 5d)	Nonlinear	3,600 (3 × 3)	128 $\sigma_{\text{data}}^2 = 0.1 \text{ m}$	7	0.98	85.8%	13 min
Case 5 (Figure 5d)	Linear	3,600 (3 × 3)	358 $\sigma_{\text{data}}^2 = 1 \text{ ms}$	4	0.99	82.6%	4 min
Case 6 (Figure 5d)	Joint: Linear + nonlinear	3,600 (3 × 3)	486 $\sigma_{\text{data}}^2 = 1 \text{ ms}$, $\sigma_{\text{data}}^2 = 0.1 \text{ m}$	7	NL = 0.98 L = 0.99	88.2%	20 min

^aThis table includes the inversion type, the number of cells of the model, the partitioning and the observed data considered in the inverse modeling, the error variance of data, the number of iteration necessary to the convergence of the inversion process, the proximity between inverted data and observed data (R²) and between the inverted structure and the true one pixel wise (structural similarity), and the inversion time. In case 3, Nolnit = initial simple model and Init = initial more complex model. In case 6, NL = nonlinear and L = linear.

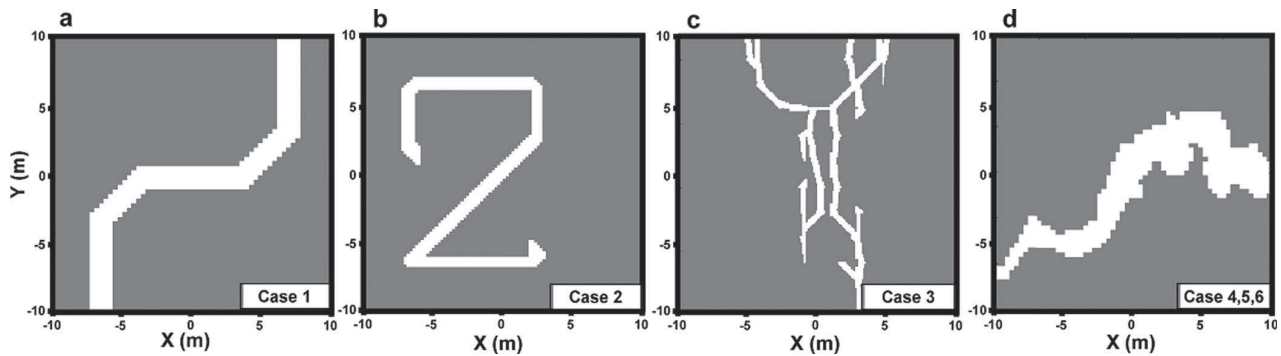


Figure 6. Presentation of the four different structures tested in the six study cases in this paper. (a) The case 1 is a linear inversion of a simple geometry to show how the optimization works. (b) The case 2 is a linear inversion of a more complex geometry. (c) The case 3 is a linear inversion of a complex multidirectional linear structure. (d) The cases 4–6 are linear, nonlinear, and joint inversion of a geostatistical generated geometry, appearing as a more natural structure.

velocities. Our simple synthetic model (Figure 6a) is set up as a field of $20 \times 20 \text{ m}^2$ with a perfectly uniform matrix (background) with a seismic velocity of 3.33 km/s, and an empty conduit (structure) of 0.26 km/s. The model properties are discretized in a 60×60 regular grid. Seismic transmitters and receptors are set up around the theoretical model, at the beginning and end of each row, column and diagonal of the model grid, which would correspond to a device every 33 cm. The observed data consists on travel time from seismic waves traveling through the model. The seismic wave travel time is calculated by summing the products between inverse of seismic velocity and distance traveled in each cell swept by the wave following the shortest path in the grid (in this case by summing cells in rows, columns and diagonals). If each cell of the grid swept divides the path followed by the wave, the total travel time of the wave is

$$t_{\text{wave}} = \sum_{i=1}^{\eta} \frac{1}{s_i} \cdot \Delta x_i, \quad (13)$$

where t_{wave} design the travel time of the wave (in ms), $i \in [1, \eta]$ identifies the different η cells swept by the wave during its travel, s_i is their seismic velocities (m/ms or km/s), and Δx_i the distance traveled by the wave through these cells (m).

Thereby, 358 observed data were generated from the theoretical model and will be used for the inversion process. For this simple geometry structure, the inversion algorithm was conducted using a relatively coarse 3×3 CA partitioning, and by considering a simple straight structure initially (Figure 7a) with seismic velocities of 2 km/s for the background and 0.5 km/s in the structure. The covariance matrix C_{β} was generated with a seismic velocity variance of $\sigma_{\beta}^2 = 1 \text{ km/s}$ and the seismic velocities of the initial model were also taken as prior values in $P_{\beta, \text{prior}}$. The inversion converged in four iterations.

This case permits to understand how the CADI algorithm works. Each different step of the optimization of the model is presented in Figure 7. Starting from the initial structure in Figure 7a, for each next steps the optimization process tries to find new subdirections improving the initial structure over the partitioning of the model (shown as a black grid in Figure 7). At the first step (Figure 7b) the initial model was improved in its central part, and at the two next steps (Figures 7c and 7d) the angles of the lower left and upper right parts of the true model were found. The last step of the inversion in Figure 7e corresponds to the properties optimization in order to improve the objective function and find the true properties. With this parameterization of the inverse problem, the result for this study case reproduces the true structure (Figure 8) and the observed data (Table 1) well.

This simple case is useful to show how the CADI algorithm modifies at each step the geometry of the initial structure and thus to understand why the information and the partitioning chosen in the initial model will considerably influence the deterministic process (in term of time but also in term of result as we will see in study case 3).

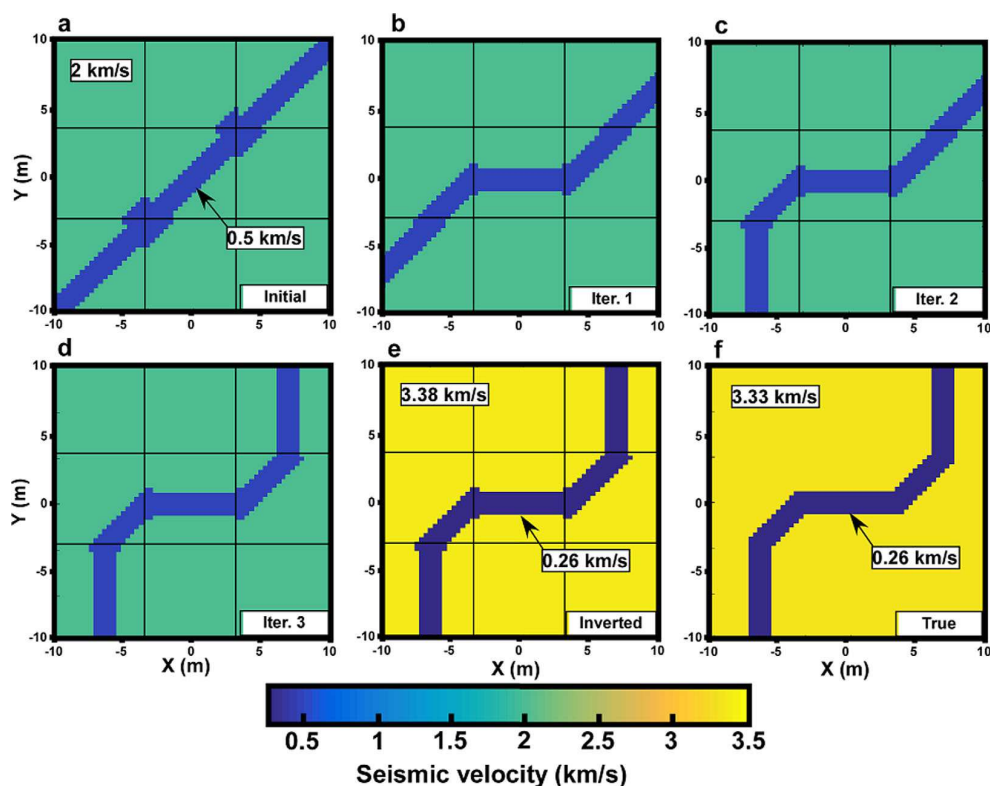


Figure 7. Result of the linear inverse modeling of the case study 1. The inversion finished after four iterations. This figure shows all different iterations of the inversion from (a) initial model to (e) inverted model. The true structure is shown in Figure 7f. Figure 7d corresponds to the structural optimization and Figure 7e to the properties optimization for this structure. The different CA subspaces of the model are highlighted by the black lines.

4.2. Study Case 2

The CADI algorithm was then applied on a more complex study case to test its capacities to reproduce complex geometries. The theoretical structure under consideration for this case is presented in Figure 6b. As in the first study case, the linear inversion is led by using seismic data, but with other seismic velocities for the

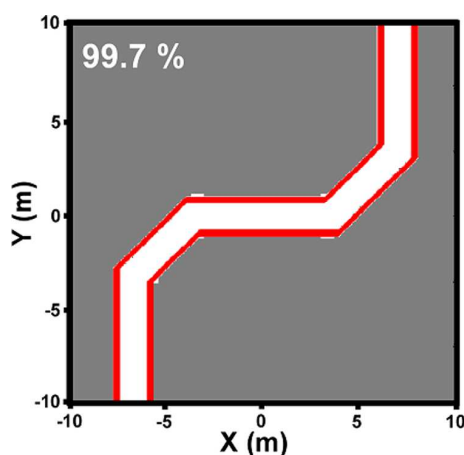


Figure 8. Comparison of the optimal structure found by inversion (in white) and the true structure (bold boundaries) for the case study 1. For this simple geometry, the inverse algorithm could easily reproduce the structure.

matrix (2.5 km/s) and the conduit (0.26 km/s). The model was discretized as a regular grid of 100×100 cells. The observed data consist in 598 seismic time travel calculated in the same way than explained for the case study 1. This time, in order to give more possibilities to the structural inverse process, the inverse model was partitioned in a 5×5 CA subspaces with the true property values initially known. The covariance matrix C_β was generated with a seismic velocity variance of $\sigma_\beta^2 = 1$ km/s and the seismic velocities of the initial model were also taken as prior values in $P_{\beta, \text{prior}}$.

Initially, we set up a straight linear structure (Figure 9a). The algorithm then converged in 21 iterations (it took approximately 1 h with a computer with 2 processors Intel Xeon 2.4GHz of 16 cores). The Figure 9 shows several steps of the inversion process (Figures 9b–9d), the optimized inverse model in Figure 9e and the true model in Figure 9f.

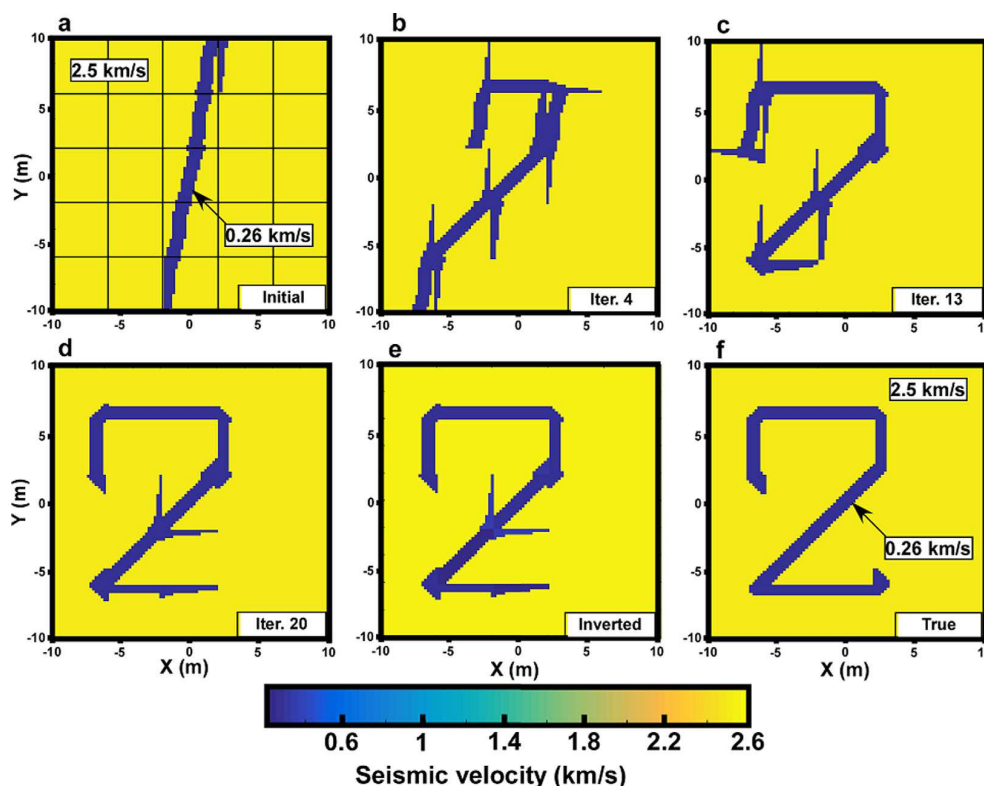


Figure 9. Result of the linear inverse modeling of the case study 2. The convergence is performed with 21 iterations. This figure shows some different iterations of the inversion from (a) initial model to (e) inverted model. The true structure is shown in Figure 9f. We noted that the optimization on the property values permits to balance the structural inversion errors. For example, in this case, the structural additions in the center of the model in Figure 9e were optimized by a light augmentation of its seismic velocity (0.5 km/s instead 0.26 km/s).

For this case, the structural inversion is close to the real one but not perfect as in first study case (Figure 10) and we can show in Figure 9e that these imperfections are balanced by the properties optimization. Thus, the zone of the structure with an inverted structural part which does not exist in the true structure (central part) is corrected by a light increase of its seismic velocity in order to minimize the differences between calculated and observed data (0.5 km/s instead of 0.26 km/s). This correction tends to locally slightly approach the structural seismic velocity to the matrix seismic velocity and thus slightly reducing the existence of this local part of the structure in the model. This property values optimization permits a better convergence on the objective function and is, in some cases, useful to counterbalance the approximations of the structural optimization when the property values are initially well known.

4.3. Study Case 3

In this third study case we applied the CADI algorithm on a complex multilinear structures network. The study was done to show all the capacities of the CADI method to model fractured fields, which are equivalent to linear structures dispersing in multiple direction among the space. The theoretical structure under consideration for this case is presented in Figure 6c. The linear inversion is led by using seismic data generated with given seismic velocities for the matrix (2.5 km/s) and the conduit (0.33 km/s). The model was discretized as a regular properties grid of 220×220 cells. The observed data consist in 1318 seismic time travel calculated in the same way than explained for the case study 1. The inverse model was partitioned in an 11×11 CA subspaces with the true property values initially known to investigate the structural optimization capacities.

In this case we compared the sensitivity of the result to the initial model. We set two inversions with two different initial models, the first one being very simple and incorporating only a global direction of

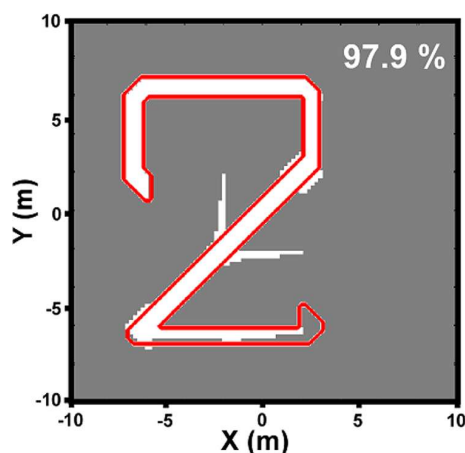


Figure 10. Comparison of the optimal structure found by inversion (in white) and the true structure (bold boundaries) for the case study 2. The optimization process reproduced a good structural inversion. The few inversion errors in the center of the model were lightly balanced by the inversion on the properties (see Figure 9).

generation (Figure 11a) and the second one incorporating a bit more information on the structural shape (Figure 11d). Figure 11 shows these initial models, the inverted models they have produced and the comparison between the inverted models and the true one (Figure 6c). For the first inversion the result is already close to the true model, especially considering the simplicity of the initial model. A simple straight structure can become a more complex multidirectional structure through the optimization process and find the main shapes and trends of a complex structure geometry which shows the possibilities given by the parameterization in the CADI algorithm. Starting from a different initial structure in the second inversion we arrived to a slightly better result on the geometry which becomes really close to the true one. It highlights the importance of incorporating some information in the initial model for the inversion process, but however if no information are known, the first inversion shows that even a very simple initial assumption can produce a good result.

4.4. Study Case 4

For the fourth study case, a nonlinear inversion has been led on a structure generated by a geostatistical technique using a directionally oriented variogram function with the package *gstat* in R (Figure 6d), which appears to be more natural than the previous structures. The steady state observed data have been

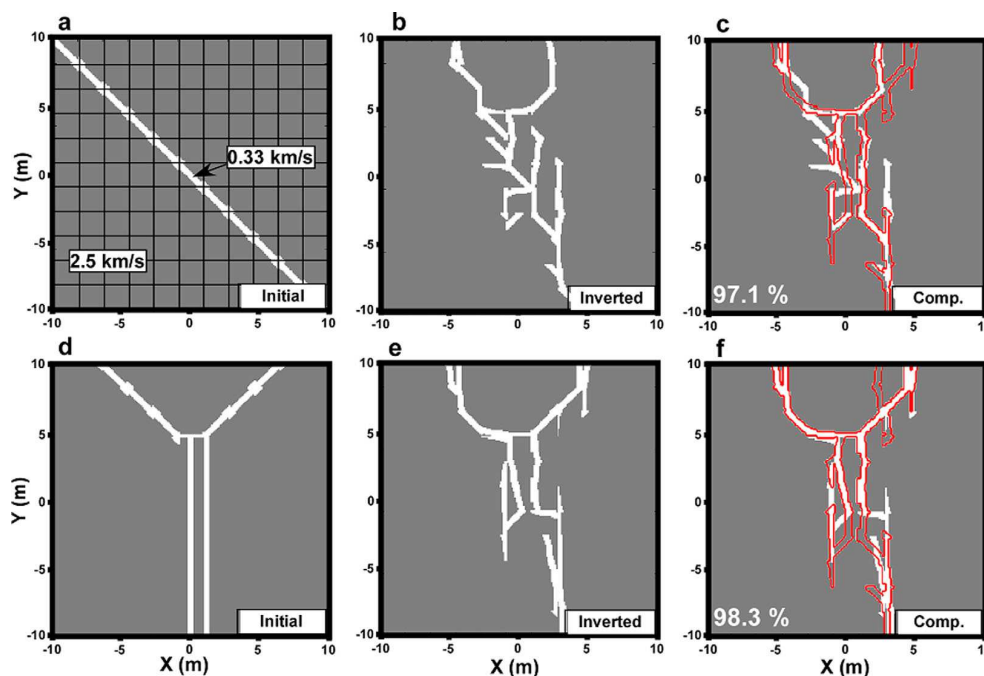


Figure 11. Results of the linear inverse modeling of the case study 3. This figure shows two inversions with (a, d) different initial models, (b, e) their results and (c, f) the comparison of these results to the true geometry boundaries in red. The convergence is performed with 26 iterations in the first inversion and 30 iterations in the second. We noted that the information contained in the initial model could slightly modify the result of the inversion but even with a very simple initial case (a) the optimization process permits to find the main shapes and trends of the true structure (c).

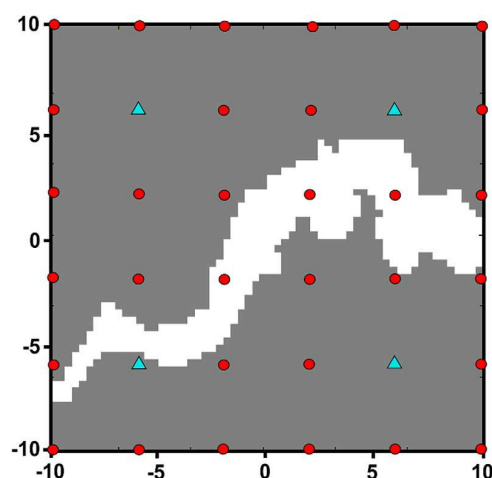


Figure 12. Map of the positioning of the wells for the hydraulic tomography inversion for the study case 4. The circles are the position of the measurement piezometers and the triangles are the position of the pumping wells.

produced by a hydraulic tomography with four alternate pumping wells and 32 measurement wells (for a total of 128 observed data) regularly distributed over the model. The positioning of the wells is presented in the model in Figure 12, this model is enclosed in a larger buffer zone ($1000 \times 1000 \text{ m}^2$) defined with a constant head (no drawdown) on its lateral boundary condition and a uniform “background” transmissivity value.

The hydraulic transmissivities are considered as the unknown properties to be inverted in a model with a 60×60 cells grid. The theoretical model is set up as a matrix with a transmissivity of $1.6 \times 10^{-7} \text{ m}^2/\text{s}$ and a structure with an equivalent transmissivity of $5 \times 10^{-4} \text{ m}^2/\text{s}$. A 3×3 CA partitioning was chosen for the inverse modeling with a good a priori information on the background properties and a structure transmissivity value of $10^{-3} \text{ m}^2/\text{s}$. The covariance matrix C_β was generated with a multiplicative variance on the transmissivity of the form

$10^{\pm \sigma_\beta^2}$ where $\sigma_\beta^2 = 1$ (± 1 variance on the transmissivity exponent) and the transmissivities of the initial model were also taken as prior values in $P_{\beta, \text{prior}}$.

From a straight linear initial structure (Figure 13a), the inversion converged in seven iterations and produced the model presented in Figure 13b. The global trends of the true structure were found although the CADI, as presented previously, produces structures which have a constant aperture. Thus, the inversion process found the best constant-aperture equivalent structure which reproduced the true model for the initial parameters. The properties optimization has permitted to find the true structure property value. However, the initial hydraulic properties did not permit to find the best fitting structure to the true model (Figure 13c). We will show in study case 6 that with the same initial model in a joint inversion, we can have both the true property values and a better fitting structure geometry.

4.5. Study Case 5

A linear inversion has been led on the same geostatistical-generated structure than in study case 4 (Figure 6d). This time, observed data have been produced by seismic, as presented in study case 1 (producing 358

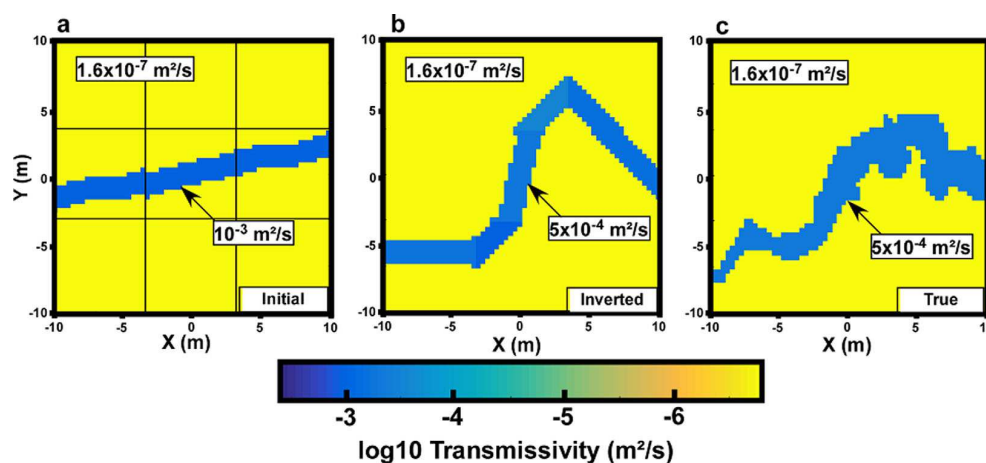


Figure 13. Result of the nonlinear inverse modeling of the study case 4. The inversion finished after seven iterations. This figure shows (a) the initial model, (b) the inverted model, and (c) the true structure. The inversion process found an optimized equivalent structure to the initial property value. The true transmissivities were found during the properties optimization.

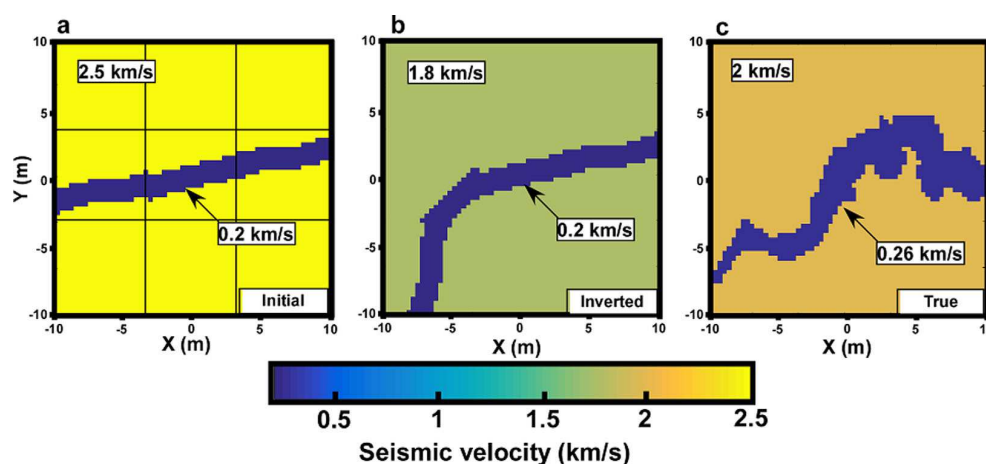


Figure 14. Result of the linear inverse modeling of the study case 5. The inversion finished after four iterations. This figure shows (a) the initial model, (b) the inverted model, and (c) the true structure. The structural optimization was limited by the initial properties and by its constant aperture generation to reproduce a variable aperture true structure. In this case, the optimization on the property values permits to balance the initial information and the structural inversion aperture limitations. The properties optimization balanced this limitation by globally decreasing the seismic velocity of the background to a lower value than the true one.

observed data). The seismic velocities are considered as the properties to be inverted in a model with a 60×60 cells grid. The theoretical model is set up as a matrix with a seismic velocity of 2 km/s and a structure with a seismic velocity of 0.26 km/s. A 3×3 CA partitioning was chosen for the inverse modeling with close initial property values (0.2 km/s for the structure and 2.5 km/s for the background). The covariance matrix C_β was generated with a seismic velocity variance of $\sigma_\beta^2 = 1$ km/s and the seismic velocities of the initial model were also taken as prior values in $P_{\beta, \text{prior}}$.

The inversion converged in four iterations, the results are presented in the Figure 14. With the same initial structure (Figure 14a) than the hydraulic inversion, the seismic inversion produced a slightly different equivalent structure which approximately reproduces global trends of the true structure but is not the best fitting possibility. The properties optimization (Figure 14b) has permitted to balance the structural approximations caused by the limits of the initial information and the generation of a constant-aperture structure. Thus, in the properties optimization, the seismic velocity of the background was decreased under the value of the true one to counterbalance the lower aperture of the generated structure. Thereby, the properties optimization part can bring more flexibility to the algorithm, which is constrained in its structural part by the prior information in the initial condition and its constant aperture. However the inversion process could not truly reproduce the structure and the properties of the true model for the initial parameters. As for the previous nonlinear inversion, we will show in the next study case that a joint inversion permits to reproduce both the property values and a better structure geometry for the same initial parameters.

4.6. Study Case 6

For the last study case, a joint inversion has been led on the same structure generated by a geostatistical approach than in study cases 4 and 5 (Figure 6d). The joint inversion is a simultaneous inversion of different data sets with a same unique inverted structure which has to be able to reproduce the information contained in all different data sets. The information brought by different investigation techniques will reduce the nonuniqueness of the inverse solution, each techniques bringing different information on the parameters [Haber and Oldenburg, 1997]. We have jointly inverted the hydraulic data from study case 4 and the seismic data from study case 5. The joint objective function in this case is a weighted linear combination of the seismic properties objective function and the hydraulic properties objective function. We chose a weighting in order to have initially approximately the same value for each of the two parts of the joint objective function. The observed data were produced by hydraulic tomography and seismic (for a total of 486 observed data). The hydraulic transmissivities and seismic velocities are considered as the properties to be inverted in a model with a 60×60 cells grid. The theoretical model is set up with the same property values as presented in study case 4 for the hydraulic properties (1.6×10^{-7} m²/s for the matrix and 5×10^{-4}

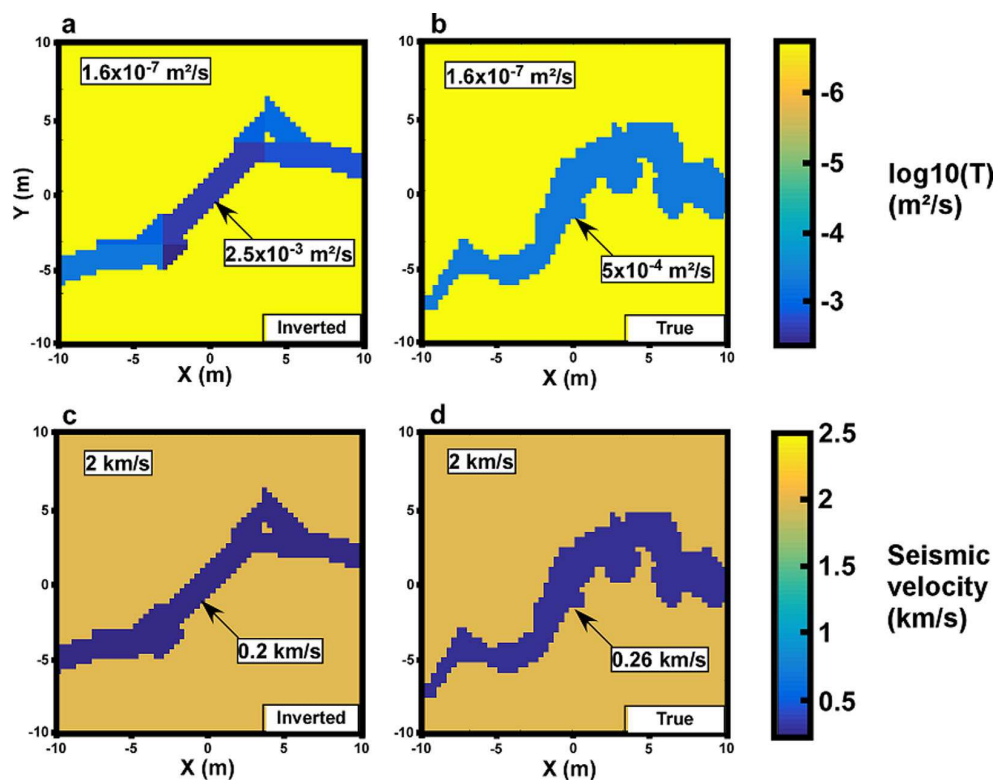


Figure 15. Result of the joint inverse modeling of the study case 6. The inversion finished after seven iterations. This figure shows (a) the hydraulic model, (c) the seismic model, and (b, d) the true models. The geometry of the structure in the models was optimized through a joint inversion of seismic and hydraulic data.

m^2/s for the structure) and in study case 5 for the seismic properties (2 km/s for the matrix and 0.26 km/s for the structure). A 3×3 CA partitioning was chosen for the inverse modeling with the same initial parameterized model than in study cases 4 and 5. The covariance matrices C_β were generated with a seismic velocity variance of $\sigma_\beta^2 = 1 \text{ km/s}$ and a multiplicative variance on the transmissivity of the form $10^{\pm\sigma_\beta^2}$ where $\sigma_\beta^2 = 1$ (± 1 variance on the transmissivity exponent), and the seismic velocities and hydraulic transmissivities of the initial models were also taken as prior values in $P_{\beta, \text{prior}}$.

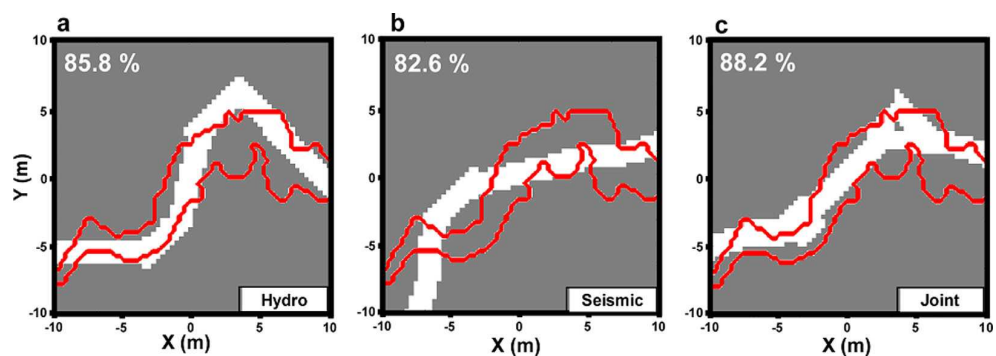


Figure 16. Pixel-wise comparison of the optimal structures found by inversion (in white) and the true structure (bold boundaries) for the study cases (a) 4, (b) 5, and (c) 6. Both hydraulic and seismic data permitted to find a geometry of the global trends of the true structure but the joint inversion resulted to a better model regarding the structure and also the convergence on the data, which avoided the difficulties encountered by the simple hydraulic inversion and the simple seismic inversion.

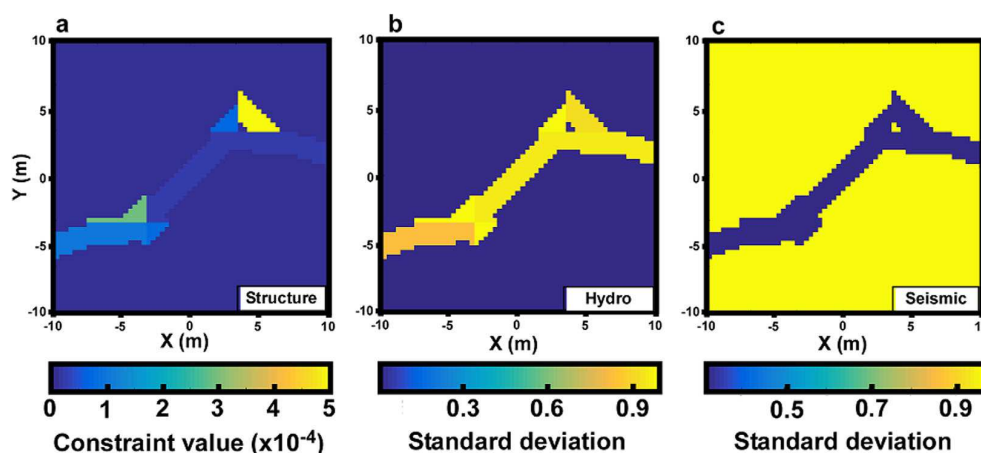


Figure 17. Uncertainties analysis for the joint inversion of the study case 6. The structural constraint in (a) indicates where the structure of the model is well-constrained by a low value, and at the opposite, a high value indicates an uncertainty for its subspace direction. The properties uncertainties for (b) the hydraulic transmissivity and (c) the seismic velocity are quantified by a standard deviation on the inverted values.

The inversion converged in seven iterations. The results of the inversion for each method and the true models are presented in Figure 15. Compared to the hydraulic and seismic separate inversion, the joint inversion produced better results on the data (Table 1) and on the inverted structure (Figure 16) which is closer to the true structure. The structural joint inversion permits to combine the hydraulic and seismic data to find the best structure. The optimized property values are also better in the joint inversion than in the separate ones. The properties optimization permits to counterbalance the limitations of a constant aperture structure by keeping a modeled higher value of transmissivity for the structure regarding the true value to simulate a thicker structure. The lower left part of the structure was optimized with a transmissivity close to the real one because the true structure is thinner in this part. The seismic velocity of the structure was also kept at a lower value than the true one to counterbalance to constant aperture geometry.

The uncertainties analysis on the structure and the property values are reported in Figure 17. The structure is well constrained by the data, except in the lower left part where the true structure is thinner and the upper right part for the angle of the structure. This means that another close subdirection of the inverted structure would not affect significantly the results. In the parameterization of the joint inversion more accuracy was given to the hydraulic data, therefore the uncertainties on the seismic properties are more important than those on the hydraulic ones. The uncertainties on the hydraulic property values vary locally within the structure. The lower left part has fewer uncertainties on the properties because its aperture is closer to the true one and therefore this part is globally closer to the true structure. The background value is well constrained because the true value was considered as a priori known. The seismic properties uncertainties

Table 2. Main Advantages Provided by the CADI Algorithm^a

Advantages	Limits	Solutions
<ol style="list-style-type: none"> 1. The complexity of the structural optimization can be monitored with a configurable partitioning of the model 2. The model properties are monitored by pilot cellular automaton, which permits to easily handle with large-scale modeling and makes a sensitivity analysis possible to accelerate the optimization 3. The convergence of the inversion is constraint to a local solution regarding the prior information which can be easily incorporated in the objective function 	<ol style="list-style-type: none"> 1. The cellular automata parameterization permits only the formation of structure with a constant aperture all along in the structural optimization 2. Only binary pattern are considered (structure and background). The background is considered as invariable regarding the variation structure/background and intrastructure 	<ol style="list-style-type: none"> 1. The property values optimization permits to digitally balance some local variation of aperture. Thus, some results on property values can be structurally interpreted 2. If the algorithm is applied on a case where the background has significant intern variability, a particular attention should be paid on setting an appropriate equivalent background

^aThe limits of the methods are also listed with a suggested solution for each limit.

are more important for its background. This is caused by the high properties difference between structure and background and because the background property was initially not known.

5. Discussion and Conclusion

The Cellular Automata-based Deterministic Inversion (CADI) algorithm is an especially adapted method for linear structure geometries. The inversion process is based on a Bayesian approach and a sequentially optimization of the structure geometry and property values. The structural optimization is monitored by cellular automaton to generate the structure, and by a configurable partitioning of the model into subspaces which permits a monitoring of the complexity of the inverted structure. One can choose a coarse subspace partitioning for simple structures and for a fast inversion process, or a fine subspace partitioning for inversion of more complex structures. The property values optimization brings more flexibility to the inversion by slightly modifying the values of the properties in the structure. This optimization permits to counterbalance some approximations in the structural optimization and some constraints from the initial information.

The CADI algorithm parameterization is mainly focused on the structural optimization, therefore it considers only two units: a constant-aperture structure and a background, which is considered as a unique uniform unit or with an intern variability which is negligible regarding the variability with the structure. Therefore, and as for any other inversion methods, it is especially effective for specific structural cases. Furthermore, the limits of the CADI algorithm have to be clearly identified in order to make a good use out of it and to have a critical view on the results it can produce. For this purpose, the main advantages and limits of the CADI algorithm have been summarized in Table 2. For each limit of the algorithm, an appropriate solution has been suggested.

In this paper we promote the potential of the CADI algorithm to image the complex linear structures, exploiting its capacity to reproduce large-scale structures in a relatively short time. As far as we know, the CADI algorithm is the first algorithm which permits the deterministic inversion of linear structures (global direction-oriented structure characterized by an aperture significantly lower than its length) with a dynamic structural optimization. This first attempt is mainly focused on the general presentation of the method and the theory of the algorithm, but we believe that this method can be improved and inspire other ones in various domain. For example, with the same algorithm structure, and by changing the cellular automaton configuration rules, it is conceivable to generate other types of forms than linear structures. We also plan further works with the presented algorithm, especially for improvements on its capacities (by adding a third "microstructures" state within the background) and for field application cases, with a higher consideration on additional prior information (as the tortuosity factor of the structure) and on sensitivity analysis of the method.

Appendix A

This appendix contains Figure A1.

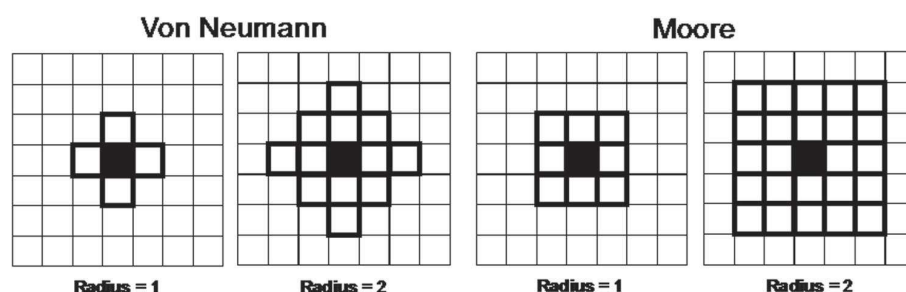


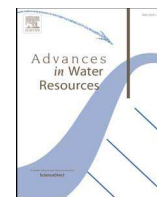
Figure A1. Two mainly used CA neighborhood definition. The two left configuration represent a Von Neumann neighborhood and the two right configuration represent a Moore neighborhood. These neighborhoods are presented for their first two radii. The black-filled cell is the cell under consideration during the CA process, and the black highlighted cells are the cells considered as its neighbors in the transition rule.

Acknowledgments

We would like to thank three anonymous reviewers for their relevant comments which permitted to significantly improve the quality of this article. We also thank the region Normandy for financially supporting the PhD of Pierre Fischer. The data and codes used to produce the results of this paper can be obtained from the first or second author.

References

- Bruna, P.-O., Y. Guglielmi, S. Viseur, J. Lamarche, and O. Bildstein (2015), Coupling fracture facies with in-situ permeability measurements to generate stochastic simulations of tight carbonate aquifer properties: Example from the lower cretaceous aquifer, Northern Provence, SE France, *J. Hydrol.*, **529**, 737–753.
- Cardiff, M., and P. K. Kitanidis (2009), Bayesian inversion for facies detection: An extensible level set framework, *Water Resour. Res.*, **45**, W10416, doi:10.1029/2008WR007675.
- Chopard, B., and A. Masselot (1999), Cellular automata and lattice Boltzmann methods: A new approach to computational fluid dynamics and particle transport, *Future Gener. Comput. Syst.*, **16**, 249–257.
- Dewri, R., and N. Chakraborti (2005), Simulating recrystallization through cellular automata and genetic algorithms, *Modell. Simul. Mater. Sci. Eng.*, **13**(2), 173–183.
- Gallardo, L. A., and M. A. Meju (2011), Structure-coupled multiphysics imaging in geophysical sciences, *Rev. Geophys.*, **49**, RG1003, doi:10.1029/2010RG000330.
- Ghosh, S., P. Gabane, A. Bose, and N. Chakraborti (2009), Modeling of recrystallization in cold rolled copper using inverse cellular automata and genetic algorithms, *Comput. Mater. Sci.*, **45**, 96–103.
- Grimstadt, A.-A., T. Mannseth, G. Naevdal, and H. Urkedal (2003), Adaptive multiscale permeability estimation, *Comput. Geosci.*, **7**(1), 1–25.
- Haber, E., and D. Oldenburg (1997), Joint inversion: A structural approach, *Inverse Problems*, **13**, 63–77.
- Halder, C., L. Madej, and M. Pietrzyk (2014), Discrete micro-scale cellular automata model for modelling phase transformation during heating of dual phase steels, *Arch. Civ. Mech. Eng.*, **14**, 96–103.
- Halder, C., D. Bachniak, L. Madej, N. Chakraborti, and M. Pietrzyk (2015), Sensitivity analysis of the finite difference 2-D cellular automata model for phase transformation during heating, *J. Iron Steel Inst. Jpn.*, **55**(1), 285–292.
- Hale, D. (2009), Structure-oriented smoothing and semblance, *CWP Rep. 635*, Cent. for Wave Phenomena. [Available at <http://inside.mines.edu/~dhale/papers/Hale09StructureOrientedSmoothingAndSemblance.pdf>]
- Hoeksema, R. J., and P. K. Kitanidis (1984), An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling, *Water Resour. Res.*, **20**(7), 1003–1020.
- Jafarpour, B., V. K. Goyal, D. B. McLaughlin, and W. T. Freeman (2010), Compressed history matching: Exploiting transform-domain sparsity for regularization of nonlinear dynamic data integration problems, *Math. Geosci.*, **42**, 1–27.
- Javadi, M., M. Sharfzadeh, and K. Shahriar (2016), Uncertainty analysis of groundwater inflow into underground excavations by stochastic discontinuum method: Case study of Siah Bisheh pumped storage project, Iran, *Tunnelling Underground Space Technol.*, **51**, 424–438.
- Jin, Z., and Z. Cui (2012), Investigation on dynamic recrystallization using a modified cellular automaton, *Comput. Mater. Sci.*, **63**, 249–255.
- Lee, J., and P. K. Kitanidis (2013), Bayesian inversion with total variation prior for discrete geologic structure identification, *Water Resour. Res.*, **49**, 7658–7669.
- Lelièvre, P. G., and D. W. Oldenburg (2009), A comprehensive study of including structural orientation information in geophysical inversions, *Geophys. J. Int.*, **178**, 623–637.
- Lochbühler, T., J. A. Vrugt, M. Sadegh, and N. Linde (2015), Summary statistics from training images as prior information in probabilistic inversion, *Geophys. J. Int.*, **201**, 157–171.
- Lu, Z., and B. A. Robinson (2006), Parameter identification using the level set method, *Geophys. Res. Lett.*, **33**, L06404, doi:10.1029/2005GL025541.
- Moore, E. F. (1962), Machine models of self-reproduction, *Am. Math. Soc.*, **14**, 17–33.
- Pan, P.-Z., X.-T. Feng, D.-P. Xu, L.-F. Shen, and J.-B. Yang (2011), Modelling fluid flow through a single fracture with different contacts using cellular automata, *Comput. Geotech.*, **38**, 959–969.
- Pardo-Igúzquiza, E., P. A. Dowd, C. Xu, and J. J. Durán-Valsero (2012), Stochastic simulation of karst conduit networks, *Adv. Water Resour.*, **35**, 141–150.
- Reeves, D. M., R. Parashar, G. Pohl, R. Carroll, T. Badger, and K. Willoughby (2013), The use of discrete fracture network simulations in the design of horizontal hillslope drainage networks in fractured rock, *Eng. Geol.*, **163**, 132–143.
- Soueid Ahmed, A., J. Zhou, A. Jardani, A. Revil, and J. P. Dupont (2015), Image-guided inversion in steady-state hydraulic tomography, *Adv. Water Resour.*, **82**, 83–97.
- Sun, X., P. L. Rosin, and R. R. Martin (2011), Fast rule identification and neighbourhood selection for cellular automata, *IEEE Trans. Syst. Man Cybern. Part B, Cybern.*, **41**(3), 749–760.
- Tan, S. K., and S.-U. Guan (2007), Evolving cellular automata to generate nonlinear sequences with desirable properties, *Appl. Soft Comput.*, **7**, 1131–1134.
- Tarantola, A., and B. Valette (1982), Generalized nonlinear inverse problems solved using the least squares criterion, *Rev. Geophys. Space Phys.*, **20**(2), 219–232.
- Tsai, F. T.-C., N.-Z. Sun, and W. W.-G. Yeh (2003), Global-local optimization for parameter structure identification in three-dimensional groundwater modeling, *Water Resour. Res.*, **39**(2), 1043, doi:10.1029/2001WR001135.
- Van der Weeën, P., J. M. Baetens, and B. de Baets (2011), Design and parameterization of a stochastic cellular automaton describing a chemical reaction, *J. Comput. Chem.*, **32**, 1952–1961.
- Von Neumann, J., and A. W. Burks (1966), *Theory of Self-Reproducing Automata*, Univ. of Ill. Press. [Available at <http://cba.mit.edu/events/03.11.ASE/docs/VonNeumann.pdf>]



Hydraulic tomography of discrete networks of conduits and fractures in a karstic aquifer by using a deterministic inversion algorithm

P. Fischer*, A. Jardani, N. Lecoq

Normandie Univ, UNIROUEN, UNICAEN, CNRS, M2C, 76000 Rouen, France

ARTICLE INFO

Keywords:

Distributed modeling
Coupled discrete–continuum model
Deterministic inversion
Heterogeneity
Aquifer characterization

ABSTRACT

In this paper, we present a novel inverse modeling method called Discrete Network Deterministic Inversion (DNDI) for mapping the geometry and property of the discrete network of conduits and fractures in the karstified aquifers. The DNDI algorithm is based on a coupled discrete–continuum concept to simulate numerically water flows in a model and a deterministic optimization algorithm to invert a set of observed piezometric data recorded during multiple pumping tests. In this method, the model is partitioned in subspaces piloted by a set of parameters (matrix transmissivity, and geometry and equivalent transmissivity of the conduits) that are considered as unknown. In this way, the deterministic optimization process can iteratively correct the geometry of the network and the values of the properties, until it converges to a global network geometry in a solution model able to reproduce the set of data. An uncertainty analysis of this result can be performed from the maps of posterior uncertainties on the network geometry or on the property values. This method has been successfully tested for three different theoretical and simplified study cases with hydraulic responses data generated from hypothetical karstic models with an increasing complexity of the network geometry, and of the matrix heterogeneity.

1. Introduction

In hydrogeological studies, the choice of the management and protection strategies of the groundwater resources is mainly based on the characterization of the hydraulic properties of the aquifer, such as hydraulic conductivity and specific storage. This characterization is most often carried out from pumping, slug and tracer tests by intrusively recording the aquifer responses, such as hydraulic pressure and tracer concentration at a set of boreholes (Butler 2005). The reliability of these techniques for capturing the spatial heterogeneity of the hydrodynamic properties is particularly conditioned by the amount and spatial disposition of wells used during the investigation, and the procedure applied to analyze the hydraulic data (Yeh and Lee 2007). In karstic and/or fractured aquifers the hydrodynamic properties (such as the hydraulic conductivity) can vary significantly from 10^{-10} m/s to 10^{-1} m/s, even at small scales (Wang et al., 2016). This heterogeneity mainly depends on the apertures, connectivity and density of the conduits and fractures network in the medium, making the groundwater flow path complex (Eisenlohr et al., 1997; Kovacs et al., 2005; Borghi et al., 2016; Ronayne 2013). In this complex context, the hydraulic flow pattern is spatially disconnected and principally focused in the transmissive fissures and fractured zones, wherein the geometrical features

and hydraulic flow regime (turbulent or laminar) are usually difficult to identify, especially with a limited number of wells, or with the use of oversimplified assumptions for interpreting the piezometric data to infer the hydrodynamic parameters.

In the hydrosience literature, several different modeling approaches based on the physical theories have already been tested in order to simulate the dynamics of karstic flows for the prediction of hydraulic properties (Hartmann et al., 2014). Among them, the equivalent porous media model, also called the single continuum model, in which the discrete features of fractures and karstic conduits are conceptualized as a porous media with continuous hydraulic properties (Larocque et al., 1999; Illman 2014; Wang et al., 2016). This simplifies the description of heterogeneity of karstic aquifers because it does not require an accurate knowledge on the architecture of fractures and conduits networks for simulating the groundwater flows. In such concept, it is sufficient to assign high hydraulic conductivity values to fractured zones and very low conductivity for intact rock. Otherwise, the coupled discrete–continuum distributed approach is of great interest thanks to its ability to imitate the dual hydrodynamic behaviors in the fractured aquifers by using Discrete Channel or Fracture Networks (DCN/DFN) for the conduits and fractures, and equivalent porous media for representing the matrix blocks (Teutsch 1993; Liedl et al.,

* Corresponding author.

E-mail address: pierre.fischer1@univ-rouen.fr (P. Fischer).

2003; De Rooij et al., 2013). In contrast to the equivalent porous media model, the discrete–continuum approach requires a good knowledge on the geometry of the karstic and fracture networks. The influence of the discrete network geometry on the hydraulic simulations and the benefits of a coupled discrete–continuum approach compared to the equivalent porous media have been widely discussed in the literature (Kovacs 2003; Painter and Cvetkovic 2005; Ghasemizadeh et al., 2012; Hartmann et al., 2014). One of these advantages is its efficiency to reproduce numerically the hydraulic fluctuations of karst spring discharge, while an equivalent porous media systematically generated lower values than the ones measured (Kovacs 2003).

The hydraulic tomography is a useful tool to predict rigorously the spatial distribution of the hydraulic properties, or the structural architecture of the fractures and conduits and their properties. It involves the use of inverse algorithms to analyze jointly a set of hydraulic data collected during multiple pumping tests (Carrera et al., 2005; Cliffe et al., 2011; Zhou et al., 2014). In this framework, various inversion algorithms were successfully applied for characterizing the hydraulic properties of fractured and heterogeneous aquifers using both concepts of parametrization discussed in the previous paragraph: the equivalent porous media and coupled discrete–continuum approach.

- (i) Inversions in an equivalent porous media were led by using geostatistical approaches in which the statistical characteristics of hydraulic properties are used as a priori information to constraint the inversion. Among these tools we cite: sequential successive linear estimator (Yeh and Liu 2000; Ni and Yeh 2008; Hao et al., 2008; Illman et al., 2009; Sharmeen et al., 2012), pilot-point (Lavenue and de Marsily 2001), transitional-probability (Wang et al., 2017), anisotropy directions (Meier et al., 2001), multi-scale resolution (Ackerer and Delay 2010), or structural approaches: probability perturbation method (Caers and Hoffman 2006), image-guided (Soueid Ahmed et al., 2015), and cellular automata-based (Fischer et al., 2017).
- (ii) On the other hand, the parameterization of hydraulic tomography using a distributed discrete–continuum approach is less flexible than the concept of the equivalent porous medium because the discrete–continuum model relies on the establishment of the architecture of the conduits and fractures, and their hydraulic properties. Several works have already brought some solutions to these difficulties. One solution would be to generate stochastically patterns of networks with various constraints: statistical constraints (Li et al., 2014; Le Coz et al., 2017), mechanical constraints (Bonneau et al., 2013; Jaquet et al., 2004), geological and speleogeological metrics information (Collon et al., 2017; Pardo-Iguzquiza et al., 2012), or flows hierarchical identification (Le Goc et al., 2010). More recently, Borghi et al. (2016) have combined the use of a generator of karstic networks, based on sets of fractures stochastically generated, with a gradient-based parameters optimization in order to reconstruct a discrete network able to reproduce a set of tracer test hydraulic data.

In this present article, we propose a novel strategy for dealing with hydraulic tomography of fractured and karstic aquifers, which we will shorten as the Discrete Network Deterministic Inversion (DNDI). The DNDI algorithm permits to map the architecture of fractures and conduits networks, their hydraulic properties, and the distribution of the transmissivity in the hard rock (matrix).

The DNDI approach relies on the use of a coupled discrete–continuum concept to simulate water flows through a karstic and fractured aquifer and a deterministic optimization algorithm to invert a set of observed piezometric data recorded during multiple pumping tests. The model is partitioned in several subspaces, each one being piloted locally by a set of parameters including: the orientations of the conduit/fracture, their equivalent transmissivity values, and the transmissivity of the rock matrix. This partitioning makes it possible to

locally modify the directions of the fracture network and to iteratively update the geometry of the global network in order to minimize the objective function in the inverse process. The method is tested on several hypothetical and simplified karstic aquifers with simple to more complex conduit networks and with homogeneous or heterogeneous transmissivity in the matrices.

2. Algorithm framework

2.1. Forward problem and model parameterization

We represent a confined karstic and fractured aquifer in a two-dimensional model Γ with an equivalent porous media Γ_M (for representing the water flows where the rock is intact) and a discrete network Γ_N (for simulating the water flows in the fracture/conduit networks). The numerical simulation of groundwater flows are governed by a steady state continuity equation associated to Darcy's law, considering a laminar flow in both the matrix domain and the discrete networks:

$$\begin{cases} \nabla \cdot (-e_M K_M \nabla h) = \frac{e_M Q_M}{V_{el}} & \text{in the matrix } \Gamma_M \\ \nabla_T \cdot (-e_N K_N \nabla_T h) = \frac{e_N Q_N}{V_{el}} & \text{in the network } \Gamma_N \end{cases} \quad (1)$$

where Q_M and Q_N are punctual water extraction or injection rates per unit of thickness ($\text{m}^3/\text{s}/\text{m}$) applied on the matrix and network respectively, K_M denotes the matrix hydraulic conductivities (m/s), K_N denotes the fractures or conduits equivalent conductivities (m/s), h is the piezometric level (m) common to both domains Γ_M and Γ_N , $e_M(\text{m})$ is the thickness of the matrix block, $e_N(\text{m})$ is the aperture of the network, and V_{el} is an elementary volume at the pumping location (m^3). We mention that Darcy's law formulation in the matrix domain is described in 2D, and in 1D for fractured networks at the internal network boundaries. That's why we use the tangential gradient operator $\nabla_T = \vec{\nabla} \cdot \vec{\ell}$ (where $\vec{\ell}$ is a local directional unit vector of the network) to solve the hydraulic equation at the network. In the study cases presented later in this article, we have chosen to simulate laminar flows as presented in Eq. (1) in a network of conduits. However, the property values K_N in the network can be more specifically adapted to the behavior of turbulent conduit flows or fracture flows through other empirical laws (eventually related to an aperture variable).

The forward problem consists in solving numerically Eq. (1) by using a finite element technique with a triangular meshing. It links the hydraulic head data simulated continuously over the coupled model to the spatial distribution in the model of the conduits or fractures with their properties in 1D, and the hydraulic transmissivities of the matrix in 2D (Fig. 1). The forward problem can be formulated as:

$$\mathbf{d} = f(\Gamma(\mathbf{P}_{Dir}, \mathbf{P}_{Prop})) + \varepsilon, \quad (2)$$

where \mathbf{d} is a vector of simulated hydraulic data ($n \times 1$), f is a forward operator that calculates the hydraulic data field from a model $\Gamma(\mathbf{P}_{Dir}, \mathbf{P}_{Prop})$ defined by given parameters of network geometry \mathbf{P}_{Dir} and hydraulic properties \mathbf{P}_{Prop} , ε is a null mean Gaussian noise to add the uncertainties associated to the numerical discretization technique and the hydraulic experimental data. The model is enclosed in a large buffer zone associated to an equivalent porous media mean transmissivity. This zone permits to limit the influence of the boundary conditions. The DNDI inversion algorithm was coded using Matlab and is linked to the COMSOL Multiphysics software which runs the forward problems.

For the DNDI algorithm, the model domain Γ is partitioned in p_x squared subspaces along the X-axis and p_y along the Y-axis. The total of subspaces of the whole domain is then $p = p_x \times p_y$. Three parameters are assigned to each subspace (Fig. 3):

- (i) the local direction of the conduit/fracture network,
- (ii) the local conduit/fracture equivalent transmissivity value,

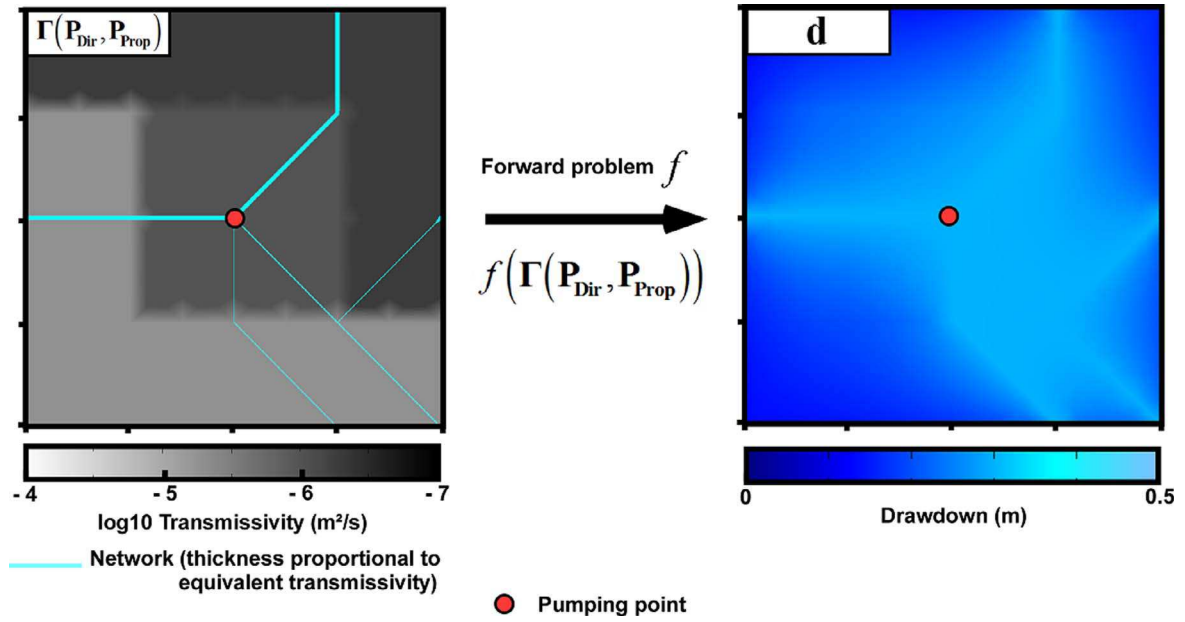


Fig. 1. Example of a simulated distribution of hydraulic heads (here drawdowns) by solving the forward problem f (Eq. (1)) for a steady state pumping in a given coupled discrete–continuum distributed model $\Gamma(\mathbf{P}_{Dir}, \mathbf{P}_{Prop})$.

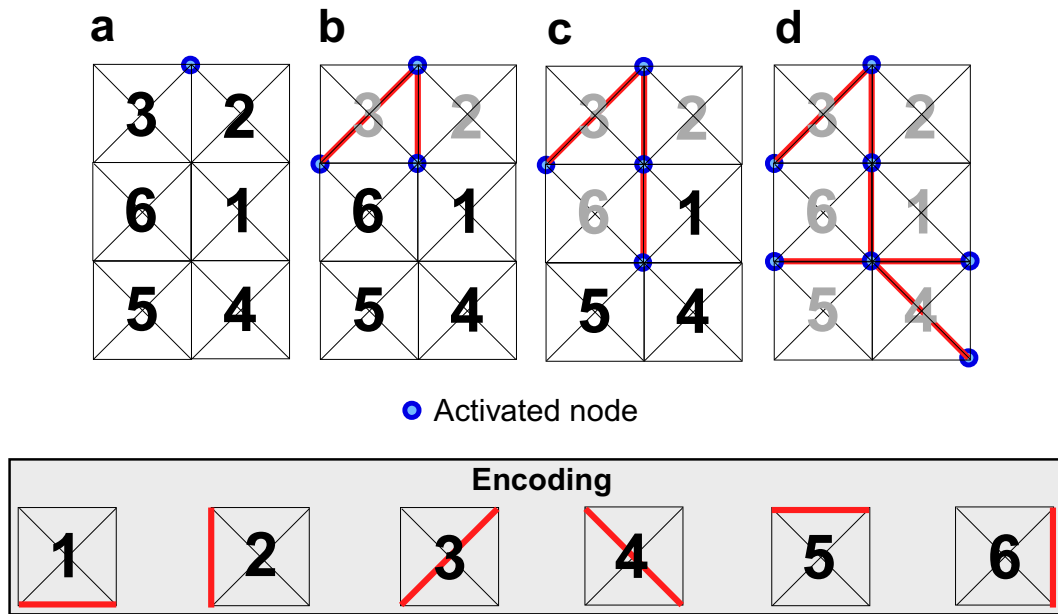


Fig. 2. Schema of the node-to-node generation process in the DNDI method with six subspaces. An activated node in the top subspaces (a) starts the generation of the structure. The structure generates to the nodes in the bottom of these subspaces, following the local direction defined in the subspaces through the encoding rules. These reached nodes then become activated (b). The subspaces in which the structure has generated become inhibited to another generation (shown as greyed number in this figure). The structure then continues its generation from its newly activated node if the subspaces structural parameters permit it (c)–(d).

(iii) and the local matrix transmissivity value.

The geometry of the network follows the local direction in each subspace by a node-to-node principle. The network structure enters a subspace by activating one of its four nodes (corners of the square) and the subspace direction parameter will define to each other node of the subspace the structure will generate. This other node will then be activated itself and permits to the structure to include new subspaces. A subspace in which the structure has already been generated becomes inhibited to another generation from the same network. The generation process is schematized in Fig. 2.

In order to perform this node-to-node generation, an initially activated node has to be specified in the model (starting node in Fig. 3).

The model geometry in COMSOL is built as a discrete grid including all network possibilities (a grid of squares and diagonals as presented in Fig. 3). This whole geometry is initially disabled in the COMSOL physical part. When different network geometries are tested in the inversion process, only the associated parts in the model grid are enabled for the solver computation. This avoids the creation of a new model geometry for each modification of the network and permits to reduce the computing time in the inversion.

The parameterization of the whole model is contained in two vectors piloting the subspaces.

(i) The local direction in a subspace is selected among six possibilities (see Fig. 3) as a structural parameter $Dir \in \{1, 2, 3, 4, 5, 6\}$. The set

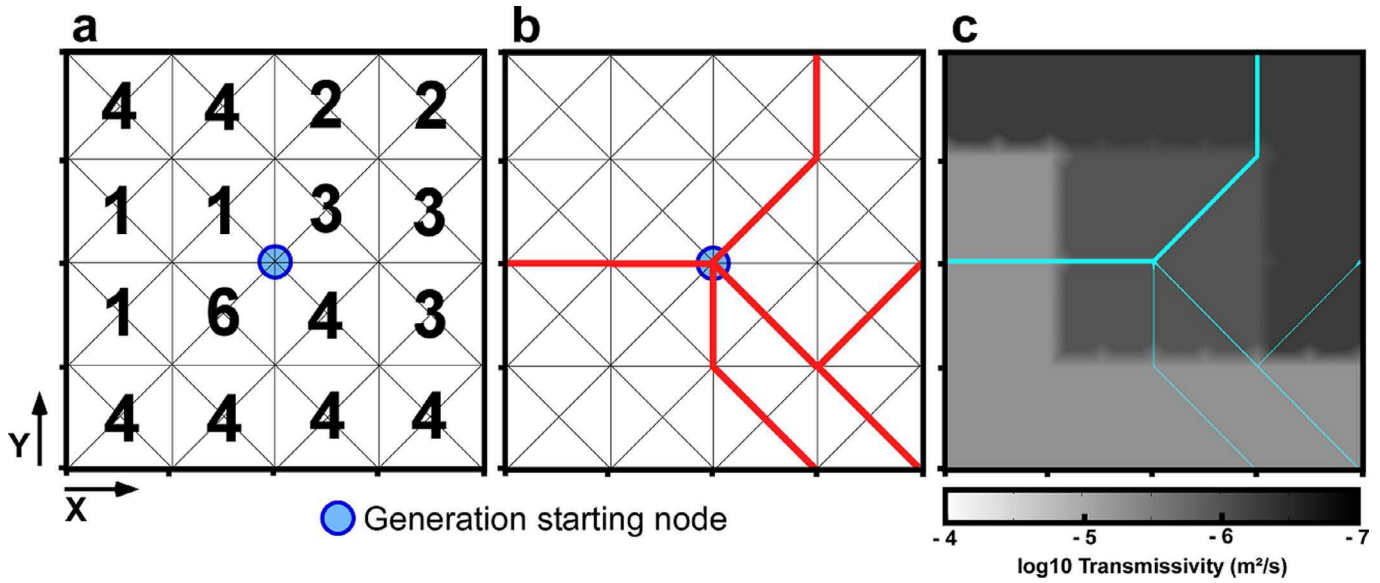


Fig. 3. Parameterization of a model in the DNDI method. For each subspace of the model there are six local direction possibilities (see encoding in Fig. 2) that are used to parameterize a network structure in the model (a). The structure (in red) is then generated, following a node-to-node rule, from the set of structural parameters in (a) and a chosen starting point at a node between subspaces (b). Finally a set of property values (transmissivities), also defined for each subspace, is assigned to the structural model (c). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of structural parameters for all subspaces in the model is contained in a $(p \times 1)$ vector \mathbf{P}_{Dir} . It is also possible to generate more than one network, but this would add more unknown structural parameters. For example if one would want to generate 3 independent networks in the model, each subspace would need to define 3 local directions instead of one. Thus, the structural vector of direction parameters \mathbf{P}_{Dir} would become a $(3p \times 1)$ vector.

- (ii) The local equivalent transmissivity of the structure in a subspace is defined as a property parameter T_N and the matrix transmissivity as a property parameter T_M . The set of property parameters for all subspaces in the model is contained in a $(2p \times 1)$ vector \mathbf{P}_{Prop} with the p parameters T_N followed by the p parameters T_M . The entire model Γ is thus piloted only by two parameter vectors: \mathbf{P}_{Dir} and \mathbf{P}_{Prop} , and can be noted $\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})$.

2.2. Inverse problem

The inversion process in the DNDI algorithm consists in retrieving a model of network of conduit/fracture and of spatial distribution of the transmissivities of the network and matrix which permits to maximize two probability density functions ρ_{network} and $\rho_{\text{properties}}$. Following the theory described by Tarantola and Valette (1982) for a least square criterion resolution of the inverse problem, we calculate ρ_{network} and $\rho_{\text{properties}}$ with the Bayes theorem, by considering Gaussian laws for the probability density functions ρ , and $\rho(\mathbf{d}_{\text{obs}})$ as certain:

$$\begin{aligned} \rho_{\text{network}}(\mathbf{P}_{\text{Dir}}|\mathbf{d}_{\text{obs}}, \mathbf{P}_{\text{Prop}}) &= \rho(\mathbf{d}_{\text{obs}}|\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}}) \cdot \rho(\mathbf{P}_{\text{Dir}})/\rho(\mathbf{d}_{\text{obs}}) \\ \rho_{\text{network}}(\mathbf{P}_{\text{Dir}}|\mathbf{d}_{\text{obs}}, \mathbf{P}_{\text{Prop}}) &\propto \exp\left(-\frac{1}{2}(\mathbf{d}_{\text{obs}} - \mathbf{f}(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})))^T \mathbf{C}_d^{-1}(\mathbf{d}_{\text{obs}} - \mathbf{f}(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})))\right) \\ &\quad \times \exp\left(-\frac{1}{2}(\mathbf{P}_{\text{Dir},\text{prior}} - \mathbf{P}_{\text{Dir}})^T \mathbf{C}_{\mathbf{P}_{\text{Dir}}}^{-1}(\mathbf{P}_{\text{Dir},\text{prior}} - \mathbf{P}_{\text{Dir}})\right) \end{aligned} \quad (3)$$

$$\begin{aligned} \rho_{\text{properties}}(\mathbf{P}_{\text{Prop}}|\mathbf{d}_{\text{obs}}, \mathbf{P}_{\text{Dir}}) &= \rho(\mathbf{d}_{\text{obs}}|\mathbf{P}_{\text{Prop}}, \mathbf{P}_{\text{Dir}}) \cdot \rho(\mathbf{P}_{\text{Prop}})/\rho(\mathbf{d}_{\text{obs}}) \\ \rho_{\text{properties}}(\mathbf{P}_{\text{Prop}}|\mathbf{d}_{\text{obs}}, \mathbf{P}_{\text{Dir}}) &\propto \exp\left(-\frac{1}{2}(\mathbf{d}_{\text{obs}} - \mathbf{f}(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})))^T \mathbf{C}_d^{-1}(\mathbf{d}_{\text{obs}} - \mathbf{f}(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})))\right) \\ &\quad \times \exp\left(-\frac{1}{2}(\mathbf{P}_{\text{Prop},\text{prior}} - \mathbf{P}_{\text{Prop}})^T \mathbf{C}_{\mathbf{P}_{\text{Prop}}}^{-1}(\mathbf{P}_{\text{Prop},\text{prior}} - \mathbf{P}_{\text{Prop}})\right) \end{aligned} \quad (4)$$

where \propto represents a proportionality relation, $\rho_{\text{network}}(\mathbf{P}_{\text{Dir}}|\mathbf{d}_{\text{obs}}, \mathbf{P}_{\text{Prop}})$ represents the a posteriori probability density function of the discrete fracture network model for a given hydraulic observed data \mathbf{d}_{obs} and the transmissivity model of the network and matrix \mathbf{P}_{Prop} . $\rho_{\text{properties}}(\mathbf{P}_{\text{Prop}}|\mathbf{d}_{\text{obs}}, \mathbf{P}_{\text{Dir}})$ is the a posteriori probability density function of the spatial distribution of the transmissivity parameters for a given hydraulic observed data \mathbf{d}_{obs} and network model \mathbf{P}_{Dir} . $\rho(\mathbf{d}_{\text{obs}}|\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})$ and $\rho(\mathbf{d}_{\text{obs}}|\mathbf{P}_{\text{Prop}}, \mathbf{P}_{\text{Dir}})$ represent the probability density functions of the network structure and property models, which permit to evaluate the ability of the network structure and property models to reproduce the observed data via the use of the forward operator. $\rho(\mathbf{P}_{\text{Dir}})$ and $\rho(\mathbf{P}_{\text{Prop}})$ represent prior distributions for the unknown parameters. It is well known that, on one hand, the piezometric data are insufficient to cope with the non-uniqueness of the solution of an inverse process, and on another hand, that a deterministic inversion process leads to a single local solution dependent to the initial model. For these reasons, and in order to additionally constrain the inversion to a more realistic solution in relation to the field knowledges, it can be interesting to incorporate prior distributions for the unknown parameters in $\rho(\mathbf{P}_{\text{Dir}})$ and $\rho(\mathbf{P}_{\text{Prop}})$.

The maximization of the a posteriori probability density functions ρ_{network} and $\rho_{\text{properties}}$ is equivalent to a minimization of the arguments of the exponentials in Eqs. (3) and (4). This is what we aim to minimize during the inversion process in the following objective functions Ψ :

$$\begin{aligned} \Psi_{\text{network}}(\mathbf{P}_{\text{Dir}}) &= \frac{1}{2}(\mathbf{d}_{\text{obs}} - \mathbf{f}(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})))^T \mathbf{C}_d^{-1}(\mathbf{d}_{\text{obs}} - \mathbf{f}(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}}))) \\ &\quad + \frac{1}{2}(\mathbf{P}_{\text{Dir},\text{prior}} - \mathbf{P}_{\text{Dir}})^T \mathbf{C}_{\mathbf{P}_{\text{Dir}}}^{-1}(\mathbf{P}_{\text{Dir},\text{prior}} - \mathbf{P}_{\text{Dir}}) \end{aligned} \quad (5)$$

$$\begin{aligned} \Psi_{\text{properties}}(\mathbf{P}_{\text{Prop}}) &= \frac{1}{2}(\mathbf{d}_{\text{obs}} - \mathbf{f}(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}})))^T \mathbf{C}_d^{-1}(\mathbf{d}_{\text{obs}} - \mathbf{f}(\Gamma(\mathbf{P}_{\text{Dir}}, \mathbf{P}_{\text{Prop}}))) \\ &\quad + \frac{1}{2}(\mathbf{P}_{\text{Prop},\text{prior}} - \mathbf{P}_{\text{Prop}})^T \mathbf{C}_{\mathbf{P}_{\text{Prop}}}^{-1}(\mathbf{P}_{\text{Prop},\text{prior}} - \mathbf{P}_{\text{Prop}}) \end{aligned} \quad (6)$$

where \mathbf{d}_{obs} is a vector of observed data $(n \times 1)$, \mathbf{P}_{Dir} and \mathbf{P}_{Prop} are the unknown parameters to estimate for imaging the geometry of the network (in \mathbf{P}_{Dir}) and the hydraulic properties (defined here by the equivalent transmissivity of the conduits/fractures and the transmissivity of the matrix in \mathbf{P}_{Prop}). $\mathbf{P}_{\text{Dir},\text{prior}}$ ($p \times 1$) and $\mathbf{P}_{\text{Prop},\text{prior}}$ ($2p \times 1$) are the prior information on the geometry and on the property parameters

employed to constrain the inverse problem for overcoming the unrealistic solutions, C_d ($n \times n$) is a covariance matrix on the observed data that permits to include the uncertainties of the hydraulic data in the inversion process. $C_{P_{Dir}}$ ($p \times p$) and $C_{P_{Prop}}$ ($2p \times 2p$) are the covariance matrices on the structural and property parameters, respectively.

This separated formulation of the probability density functions between network and properties permits to sequentially estimate the two dependent unknown models P_{Dir} and P_{Prop} . In a first step, we focus on the characterization of the network with the piezometric data by fixing the model of the transmissivity distributions in the conduits/fractures and matrices. The model of network resulting from the first step will then be used in a second step as known parameter to infer the transmissivity pattern.

2.3. Optimization and uncertainty analysis

The minimization of Eqs. (5) and (6) can be done by optimizing the network geometry and the property values during two sequential iterative processes. These optimizations consist in successively modifying the structural and property parameters P_{Dir} and P_{Prop} .

The inversion process is led as a sequential optimization (Fig. 4) of

- (i) the structural geometry (considering as fixed the initially chosen property values),
- (ii) and the property values T_N, T_M (considering as fixed the previously inverted structural geometry).

The structural optimization is performed iteratively by modifying the structural parameter P_{Dir} through a structural sensitivity analysis and by considering the hydraulic properties P_{Prop} as fixed. At a given iteration step k , the sensitivity analysis of the network geometry toward the observed data is recorded into a $(6 \times p)$ sensitivity matrix J_n^k . For a local direction $i \in [1, 6]$ and for a subspace $j \in [1, p]$ the element of the matrix J_n^k is calculated as:

$$J_n^k(i, j) = \frac{1}{2} (\mathbf{d}_{obs} - \mathbf{f}(\Gamma(P_{Dir}^k |_{P_{Dir}(j)=i}, P_{Prop})))^T C_d^{-1} (\mathbf{d}_{obs} - \mathbf{f}(\Gamma(P_{Dir}^k |_{P_{Dir}(j)=i}, P_{Prop}))) + \frac{1}{2} (P_{Dir, prior}(j) - i)^T C_{P_{Dir}}^{-1} (P_{Dir, prior}(j) - i) \quad (7)$$

with $P_{Dir}^k |_{P_{Dir}(j)=i}$ the structural geometry parameter at the iteration k considering a modified local direction in the subspace j , and $(P_{Dir, prior}(j) - i)$ the gap between the prior local direction of subspace j and the modified local direction.

Thus, the sensitivity matrix guides the evolution of the objective function in Eq. (5) by testing successively the modification of the network with all possible local direction in each subspace. The minimal value in the matrix $J_n^k(i_{min}, j_{min})$ designates the local direction i_{min} in the subspace j_{min} which would produce the best decrease in the objective function.

Then the parameters set P_{Dir}^k is updated from the previous set P_{Dir}^{k-1} by taking into account the sensitivity analysis minimum $P_{Dir}^k(j_{min}) = i_{min}$ in order to minimize the objective function at each step of the optimization.

Once the sensitivity analysis cannot find any more P_{Dir} configuration decreasing the objective function, the iterative structural optimization is stopped. The last structural iteration represents the local solution, dependent to the initial model. The uncertainty analysis of the inverted network geometry can be inferred from the computation of the posterior covariance matrix as:

$$C_{P_{Dir}}^{post}(j) = \left(\frac{1}{6} \sum_{i=1}^6 J_n^{post}(i, j) - \Psi_{network}^{post} + C_{P_{Dir}}^{-1}(j, j) \right)^{-1}, \quad (8)$$

where $C_{P_{Dir}}^{post}(j)$ is the posterior structural uncertainty value for the local direction in the subspace j , J_n^{post} is the last iteration structural

sensitivity matrix and $\Psi_{network}^{post}$ is the value of the minimized structural objective function. If the structural uncertainty value is low, then another direction in the subspace would lead to a deterioration of the reproduction of the data. On the other hand if the structural uncertainty value is high, then the structure in the subspace could have another local direction without significantly degrading the reproduction of the observed data.

Following the network geometry optimization, the property parameters optimization will iteratively modify the transmissivities with the previously inverted geometry in order to minimize the objective function in Eq. (6). The network equivalent transmissivities and the matrix transmissivities are optimized simultaneously. At a given iteration step k , the parameters set P_{Prop} , which contains the transmissivities for both the network and the matrix, is updated by linearizing Eq. (6), which can be formulated as

$$P_{Prop}^{k+1} = P_{Prop}^k + ((J_p^k)^T \cdot C_d^{-1} \cdot J_p^k + C_{P_{Prop}}^{-1})^{-1} \cdot (J_p^k)^T \cdot C_d^{-1} \cdot (\mathbf{d}_{obs} - \mathbf{f}(\Gamma(P_{Dir}, P_{Prop}^k))) + C_{P_{Prop}}^{-1} \cdot (P_{Prop, prior} - P_{Prop}^k) \quad (9)$$

where J_p^k is the Jacobian matrix ($n \times 2p$) that holds the sensitivity for each modeled data f_i (at the positions of the observed data) toward the property values in the matrix and the network. This Jacobian matrix can be calculated by using a finite difference approach, with a finite difference step ΔP_{Prop} :

$$J_p^k(i, j) = \frac{\partial f_i}{\partial P_{Prop}^k} \bigg|_{P_{Prop}^k(j)=P_{Prop}^k(j)+\Delta P_{Prop}} \quad (10)$$

Finally, once the objective function has iteratively converged to a minimum, the property optimization is stopped. The posterior covariance matrix on the inversion of the property values can be calculated as:

$$C_{P_{Prop}}^{post} = ((J_p^{post})^T \cdot C_d^{-1} \cdot J_p^{post} + C_{P_{Prop}}^{-1})^{-1}, \quad (11)$$

where $C_{P_{Prop}}^{post}$ is the posterior covariance matrix and J_p^{post} is the Jacobian matrix of the last iteration step. The diagonal entries of the posterior matrix represent the variances on the property values of each subspace.

3. Validation of the DNDI algorithm on hypothetical study cases

The DNDI inversion algorithm has been tested on three hypothetical and simplified confined karstic fields with network of conduits:

- (i) in a first case, we treat a simple network case with heterogeneity in the equivalent transmissivity of the conduits and a homogenous transmissivity assigned to the matrix,
- (ii) a second case is similar to the first one but adding a transmissivity variability also in the matrix,
- (iii) in a third case, we seek to image a complex network geometry with the use of two different initial models to start the inverse problem.

We considered in the forward problem (Eq. (1)) a unit thickness for the matrix (2D modeling) and a unit aperture for the network (but with a variable equivalent transmissivity). The buffer zone boundaries were associated to a $h_{bound} = 0$ m Dirichlet condition and the hydraulic heads were set to $h_0 = 0$ m initially over the model. These theoretical study cases were used to produce 2401 hydraulic drawdown data from 49 pumping/measurement boreholes (a pumping test is performed alternatively in each borehole) distributed homogeneously over the 100×100 m² models. The pumping rates were set to 0.6 L/min for a borehole in the matrix and 5 L/s for one in the conduit network.

In these different cases, for the inversion of the geometry of the network, no a priori information has been added. On the other hand, we have constrained the inversion of the hydraulic properties with a priori values. The a priori models on the properties are used also as initial model to launch the inversion process. For the property optimization,

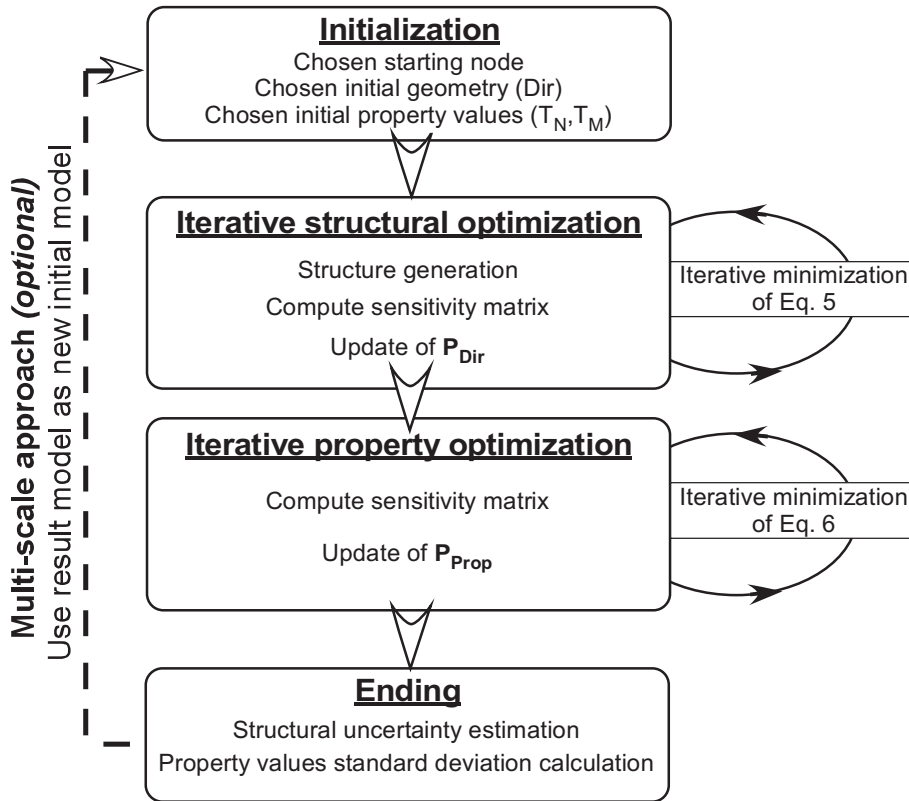


Fig. 4. A flowchart of the inversion steps used in the DNDI algorithm. After the initialization of the parameters, a sequential iterative optimization is led on the structure geometry and on the property values in order to minimize both objective functions (Eqs. (5) and (6)). An eventual re-run of the inversion process (multi-scale option) using the result as new initial model can be performed in order to improve this result.

for the matrix we took the $-\log_{10}(T)$ as transmissivity parameter T_M (for example a transmissivity parameter equal to 6 represents in the model a 10^{-6} m²/s transmissivity value) and for the network we took directly the T value as transmissivity parameter T_N . The covariance matrices C_d and $C_{P_{Prop}}$ are built as diagonal matrices with a constant variance value σ^2 (in the case of the matrix transmissivity the variance value σ_{TM}^2 applies to the exponent of the transmissivity, in the case of the network transmissivity the variance value σ_{TN}^2 applies to the transmissivity). The partitioning of the models and the chosen inversion parameters values for each study case are given in Table 1. The different study cases inversions were led on a 64Go RAM PC on 2 processors of 16 cores.

3.1. Study case 1

In a first study case, we have tested the ability of the inversion method to reproduce a network geometry with variable conduit equivalent transmissivities in a homogeneous matrix. We generated drawdown data from a theoretical model with a 10^{-6} m²/s matrix transmissivity and a principal conduit associated to a 0.1 m²/s transmissivity and secondary conduits associated to a 0.01 m²/s transmissivity. Firstly, we tested an inversion with a small set of data (100 drawdown data from 10 boreholes, see the ‘True model’ in Fig. 5).

We started the inversion from a simple initial model with a single horizontal 0.06 m²/s conduit and a homogeneous 10^{-6} m²/s matrix transmissivity. The structural optimization converged in 10 iterations and the properties optimization in 1 iteration.

The inverted model reproduces the data set ($R^2 = 0.97$) and approximately the connectivity between the points in the network, however this reconstruction remains distant from the true geometry. This is due to a lack of data to correctly identify the shape of the conduit network. Therefore, the efficiency of the inversion for mapping the heterogeneity of the hydraulic parameters and retrieving the principal

karstic conduits is highly dependent to the number of wells and their locations. In the next test, we used a denser distribution of wells (49 wells) for providing a better spatial resolution in order to image the heterogeneity of the aquifer presented for the same ‘True model’ in Fig. 6.

The structural optimization converged in 11 iterations and the properties optimization in 1 iteration. The inverted model reproduces now the data set ($R^2 = 0.95$) and also a very good representation of the true geometry (Fig. 6). The property optimization permitted to correct the initial equivalent transmissivity of 0.06 m²/s to 0.01 m²/s for the conduits connected in the bottom right area of the network. It permits to reduce the flow rates coming to this zone and enhance the reproduction of the true cones of depression. The flows in this zone are mainly conditioned by the properties of the conduits connected directly to the primary drain. This affirmation can be supported by the conduit transmissivity standard deviation map produced from Eq. (11) (Fig. 6), that shows that the properties of the conduits directly connected to the primary drain have lower uncertainties than the primary drain itself in the center of the inverted model. The conduit in the bottom right periphery of the inverted model does not image correctly the true model. But as the data reproduction is perfect, this periphery zone might not be sufficiently described by the data to permit a very good reproduction. The uncertainty map confirms that this part of the network has a more uncertain transmissivity value than the rest of the network.

We have also tested another configuration with more available boreholes than in the case in Fig. 5, but in which only two boreholes intersect the true karstic network. The true model and the inversion result are presented in Fig. 7.

The inverted model can almost reproduce the true network geometry, which shows that boreholes, even in the matrix, can provide information about the localization of nearby conduits. This is especially true for the thin conduits which appear in the inverted model although no boreholes are intersecting them. Therefore, the DNDI method can be

Table 1
Parameters used in the inversion study cases.

	Study case 1	Study case 2	Study case 3
Partitioning	4×4	4×4	8×8
A priori T_N	$0.06 \text{ m}^2/\text{s}$	$0.04 \text{ m}^2/\text{s}$	$0.1 \text{ m}^2/\text{s}$
A priori T_M	$10^{-6} \text{ m}^2/\text{s}$	$10^{-6} \text{ m}^2/\text{s}$	$10^{-6} \text{ m}^2/\text{s}$
Data cov. matrix $C_d = \sigma_{data}^2 \cdot Id(n)$	$\sigma_{data} = 10^{-2} \text{ m}$	$\sigma_{data} = 10^{-2} \text{ m}$	$\sigma_{data} = 10^{-2} \text{ m}$
Property cov. matrix $C_{Pprop} = \sigma_{T_N/T_M}^2 \cdot Id(2p)$	$\sigma_{T_N} = 10^{-6} \text{ m}^2/\text{s}$ $\sigma_{T_M} = 10^{-6}$	$\sigma_{T_N} = 10^{-6} \text{ m}^2/\text{s}$ $\sigma_{T_M} = 10^{-1}$	$\sigma_{T_N} = 10^{-6} \text{ m}^2/\text{s}$ $\sigma_{T_M} = 10^{-6}$

used with dataset with only a few boreholes intersecting the conduits as long as there are a sufficient number of other boreholes, in the matrix, in suitable locations for characterizing the nearby conduit network.

3.2. Study case 2

A second study case was led to test the ability of this inversion method to reproduce the data in a case of a karstic network with various conduit properties developed in a heterogeneous matrix. We simulated the piezometric data from a theoretical model with the same karstic network than in study case 1, but in a matrix with a transmissivity varying from $5.10^{-6} \text{ m}^2/\text{s}$ to $5.10^{-7} \text{ m}^2/\text{s}$ (Fig. 8).

We started the inversion from a simple initial model with a single horizontal $0.04 \text{ m}^2/\text{s}$ conduit and a homogeneous $10^{-6} \text{ m}^2/\text{s}$ matrix transmissivity. The structural optimization converged in 10 iterations and the properties optimization in 3 iterations (Fig. 8(a)).

The structural optimization permitted to retrieve the true geometry of the conduits network, but it also added conduits in the bottom left part of the model to reproduce the drawdown data of the more transmissive area of the matrix. Then the property optimization could reproduce the true transmissivity values distribution in the matrix. In the end, the inverted model can reproduce the true drawdowns data, but its network geometry incorporates parts, inexistent in the true model, that has been generated in order to simulate a more transmissive area of the matrix before the matrix transmissivity values could be optimized.

We started a second inversion using the previously inverted model (indicated in Fig. 4 as the ‘multi-scale option’). The structural optimization converged in 2 iterations and the parameter property optimization in 1 iteration (Fig. 8(b)). The only changes were made during the structural optimization step, with an important improvement in the identification of the shape of the conduits. In this case, the inverted model reproduces the drawdowns data ($R^2 = 0.99$) but is also a good representation of the true network geometry.

The posterior standard deviation maps produced from Eq. (11) (Fig. 9) show, for the conduit property values, a smaller uncertainty for the secondary conduits and a higher uncertainty for the primary drain, especially for the part of the network on the left of the model. Concerning the matrix transmissivity property values, the highest uncertainty are located mostly in the most transmissive areas.

3.3. Study case 3

Finally, a third study case was led to test the ability of this inversion method to reproduce the data in a case of a complex karstic network geometry. We generated drawdown data from a theoretical model with a karstic network with a constant equivalent transmissivity of $0.1 \text{ m}^2/\text{s}$ in a homogeneous matrix with a transmissivity of $10^{-6} \text{ m}^2/\text{s}$ (Fig. 10).

We started an inversion from a simple initial model with a single vertical $0.1 \text{ m}^2/\text{s}$ conduit and a homogeneous $10^{-6} \text{ m}^2/\text{s}$ matrix transmissivity. The structural optimization converged in 33 iterations and the parameter optimization in 1 iteration (Fig. 10(a)). The inverted model permits to fit the data set approximately ($R^2 = 0.78$) and represents the global geometry of the conduits network of the true model. Regarding the simplicity of the initial model, the result model remains satisfying.

We also started an inversion from a more complex initial model with two vertical $0.1 \text{ m}^2/\text{s}$ conduits diverging in the upper part of the model in a homogeneous $10^{-6} \text{ m}^2/\text{s}$ matrix transmissivity. This initial model geometry (representing a simple approximation of the true geometry) can be associated to a priori field knowledge information. The structural optimization converged in 17 iterations and the parameter optimization in 1 iteration (Fig. 10(b)). In this case, the inverted geometry of the discrete network permits a good reproduction of the data ($R^2 = 0.97$) and is closer to the real network than the case in Fig. 10(a).

The structural posterior uncertainty maps produced from Eq. (8) are presented in Fig. 11. These maps show that, in the Case a, the highest

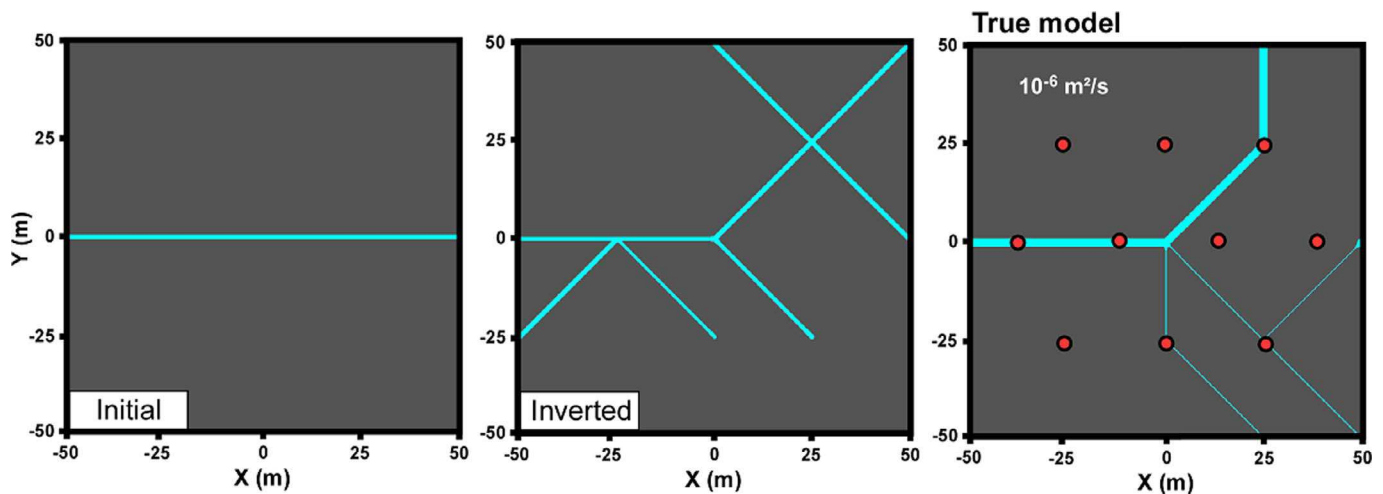


Fig. 5. Initial and inverted models for an inversion using drawdown data produced from a true model (on the right) with a homogeneous matrix. The network on the true model symbolize the pumping/measurement boreholes for the hydraulic data. The inverted model permits to localize approximately the karstic network connections but in this case the amount of data is insufficient to have a proper imagery. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

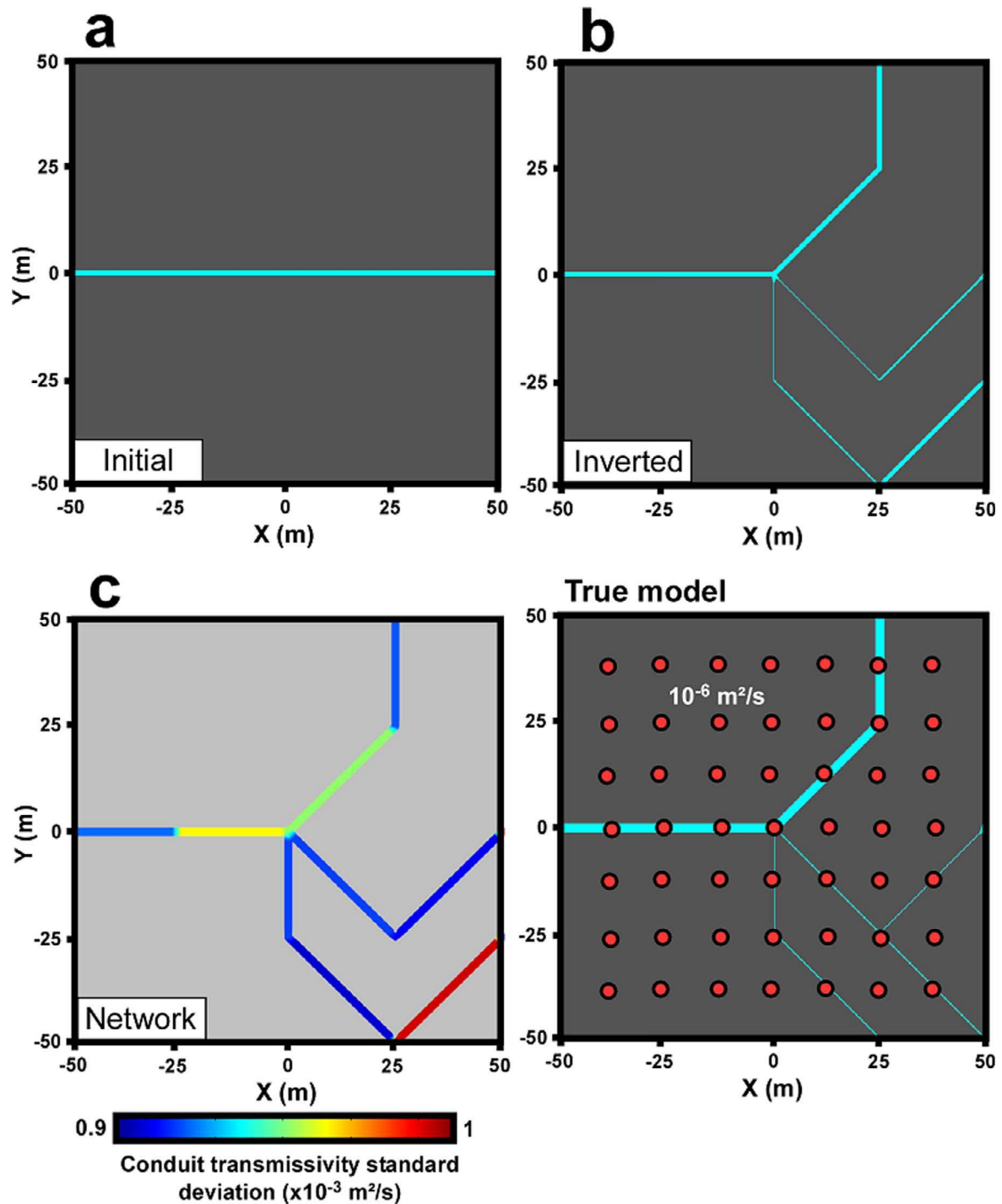


Fig. 6. Initial (a) and inverted (b) models for an inversion using drawdown data produced from a true model with a homogeneous matrix, and associated map of the conduit properties posterior standard deviations (c). The inverted model in (b) permits a good localization the true karstic network. It also reduced locally the initial transmissivity (0.06 m²/s to 0.01 m²/s) of the conduits connected to the primary drain in the bottom right part of the model (the conduit thickness is proportional to its transmissivity value). The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

uncertainties are distributed relatively uniformly among the inverted model, while in Case b, they are mostly located in the periphery of the model. Here, the structural posterior uncertainties are giving important information about the local validity of the different inverted networks.

4. Discussion

We have successfully tested the DNDI method on three theoretical

and simplified study cases with steady state drawdowns. However as we have seen, an inversion process is limited by the non-uniqueness of its solution. Therefore using the DNDI method requires several prerequisites and the modeler needs to be critical toward the result.

As we have seen in the first study case, the efficiency of the inversion is dependent to the hydraulic data set, and in particular the number and the localization of observation wells on the field. We note that even wells in the matrix can provide information on nearby

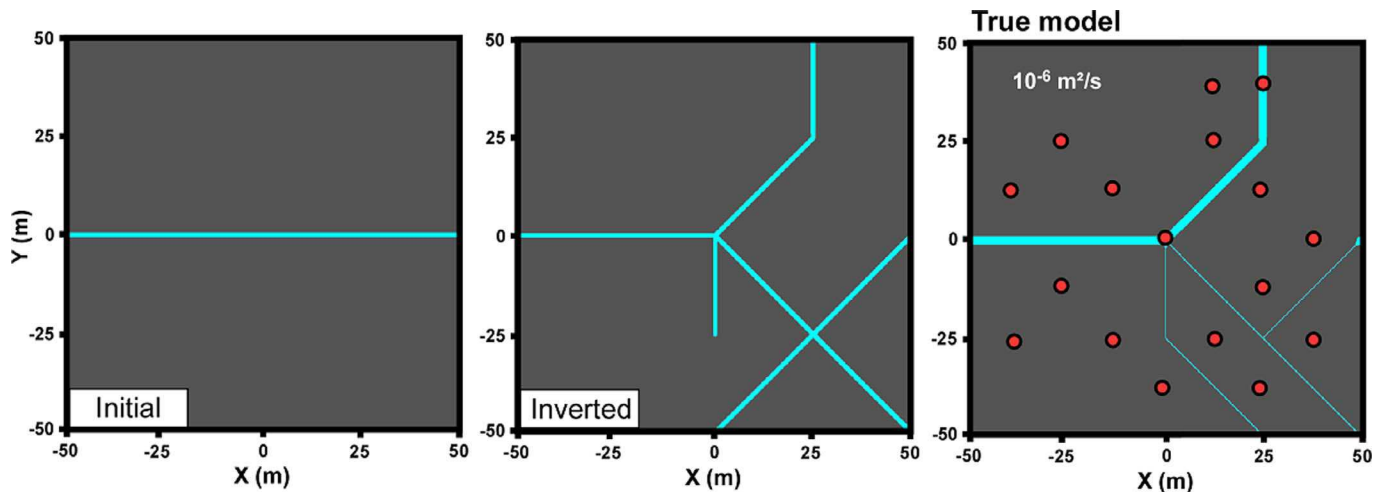


Fig. 7. Initial and inverted models for an inversion using drawdown data produced from a true model (on the right) with a homogeneous matrix. The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data, primarily located in the matrix. The inverted model permits to almost reproduce the karstic network even if only two measurement points are located in the true network. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

conduits for the inversion. Globally it appears that the most important point about a steady-state dataset is to have a homogeneous and sufficiently dense distribution of wells on the site, in order to characterize successfully the network.

Concerning the inversion process itself, we note, in the third study case, the ability of the DNDI method to image complex networks. However, as the inversion is deterministic, the precision of the result model is dependent to the initial model. The inversion process will converge to a local solution dependent to the initial model. In fact, in Fig. 10 we show that a simple initial model permitted to reproduce a satisfying global representation of the true model, but with local

approximations, while a more complex initial model permitted a more accurate reproduction of the true model and a faster convergence. Therefore an inverted model using the DNDI method should be analyzed critically, like any deterministic inverse methods, depending from the initial model. The study of the computed structural and property uncertainty values (from Eqs. (8) to (11)) can supply this critical analysis on the result model.

The second study case also illustrates some limits of the sequential optimization of the method, especially when starting from a too simple initial model. Therefore the amount of a priori information introduced in the initial model is important for the accuracy of the result model.

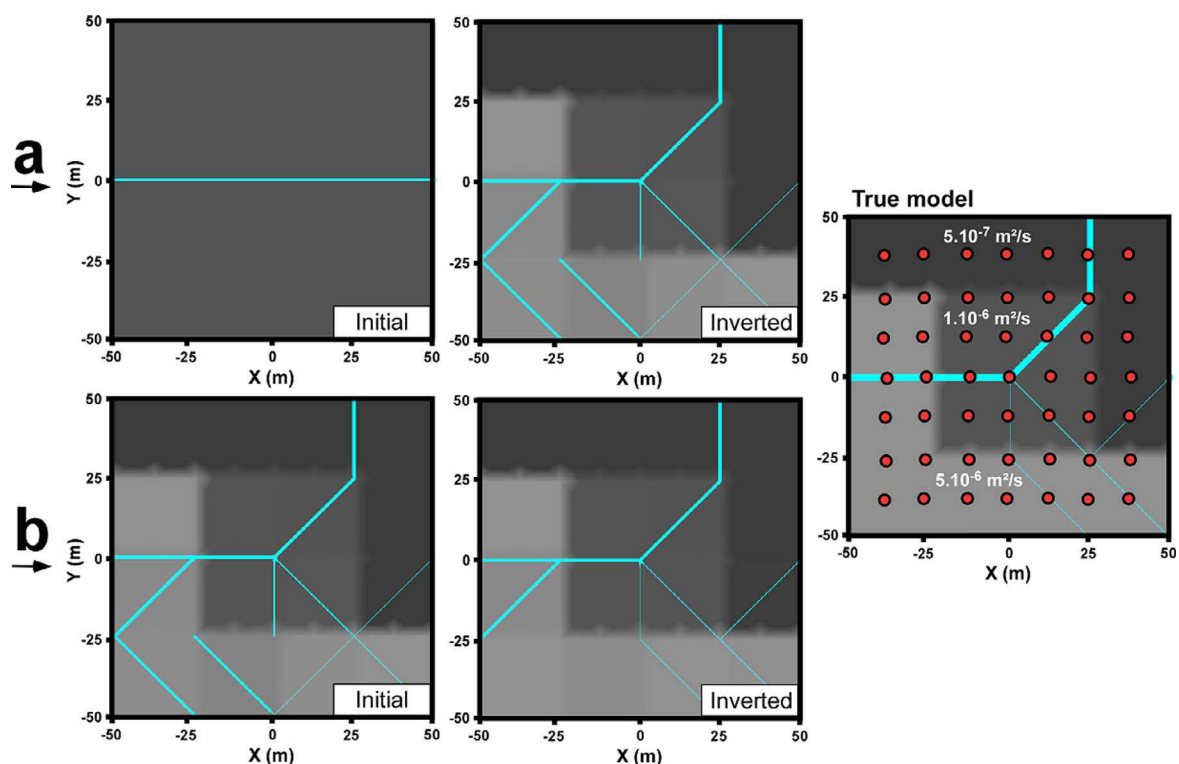


Fig. 8. Initial and inverted models for an inversion using drawdown data produced from a true model (on the right) with a heterogeneous matrix. The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data. A first inverted model (a) permits to localize the true karstic network but also generates conduits to simulate the more transmissive part of the true model. A second inversion (b) starting from the previous inverted model permits to correct the geometry and produces an inverted model matching more accurately the true model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

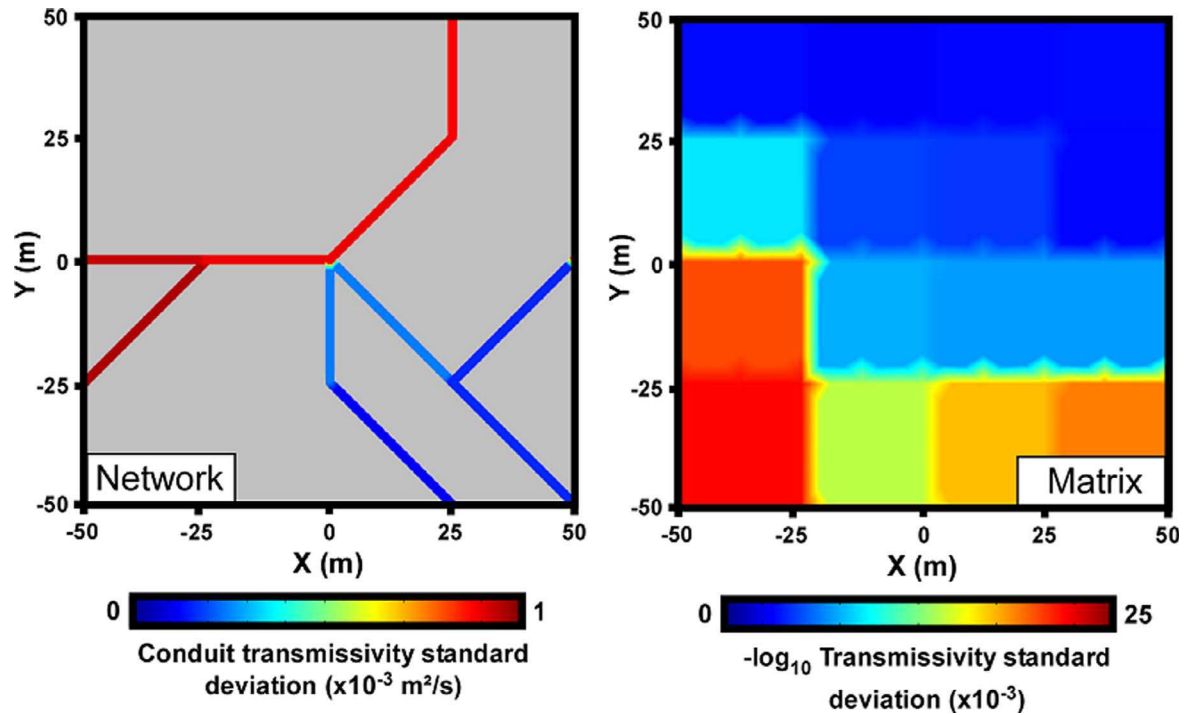


Fig. 9. Maps of the conduit and matrix transmissivities posterior standard deviations. The matrix higher transmissivity zones in the inverted model (bottom left) have a higher uncertainty value than the lower transmissivity zones (top right). On the contrary, the uncertainty on the transmissivities of the conduits of the primary drain is higher than the secondary conduits. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Otherwise, as we demonstrate in Fig. 8, a simple possible operation would be to re-run the inversion with a first inversion result to slightly improve the result. We would also recommend the coupling of this inversion method to a multiscale method (Grimstadt et al., 2003) which

consists in a re-run of the inversion starting from a previous result with a refinement of the partitioning. It permits to lead several inversions with an initial model each time more precise while saving time as we initially start with a coarsely partitioned model.

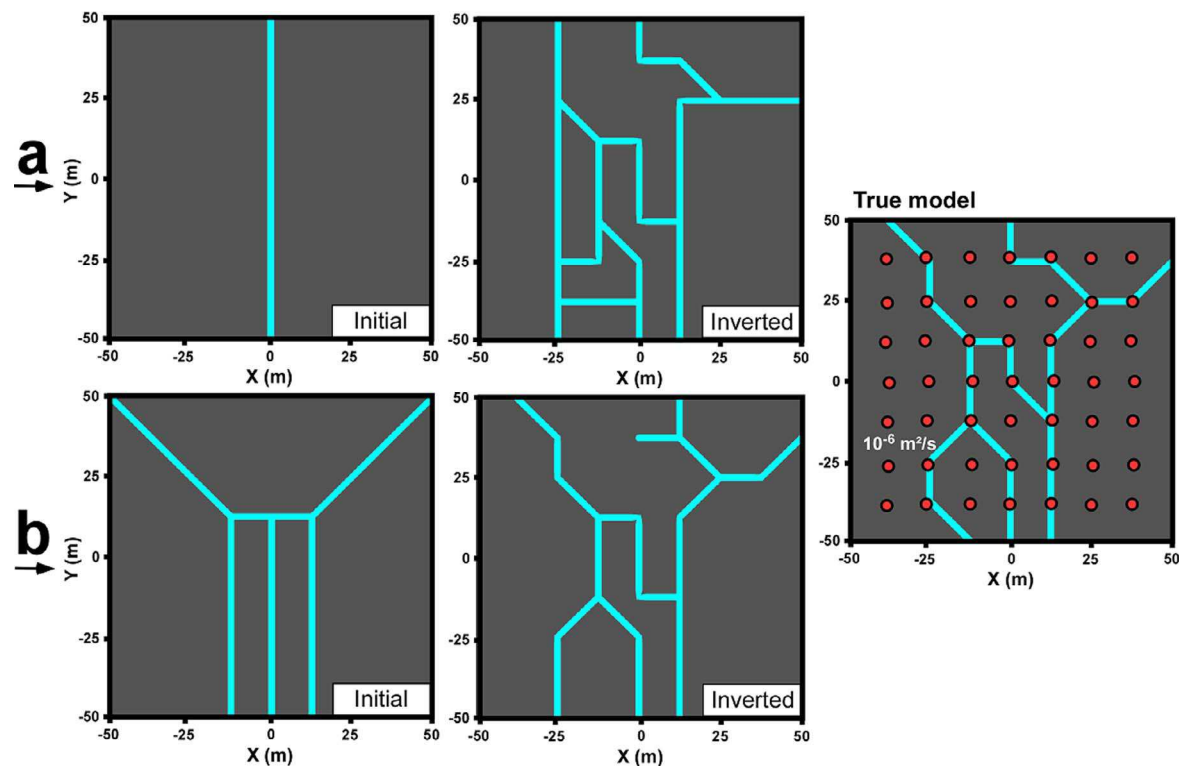


Fig. 10. Initial and inverted models for an inversion using drawdown data generated from a true model (on the right) with a homogeneous matrix. The red dots on the true model symbolize the pumping/measurement boreholes for the hydraulic data. A first inverted model (a), starting from a simple initial model, permits to localize approximately the true network geometry. A second inversion (b), starting from a more detailed initial model, permits to produce a more precise network geometry. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

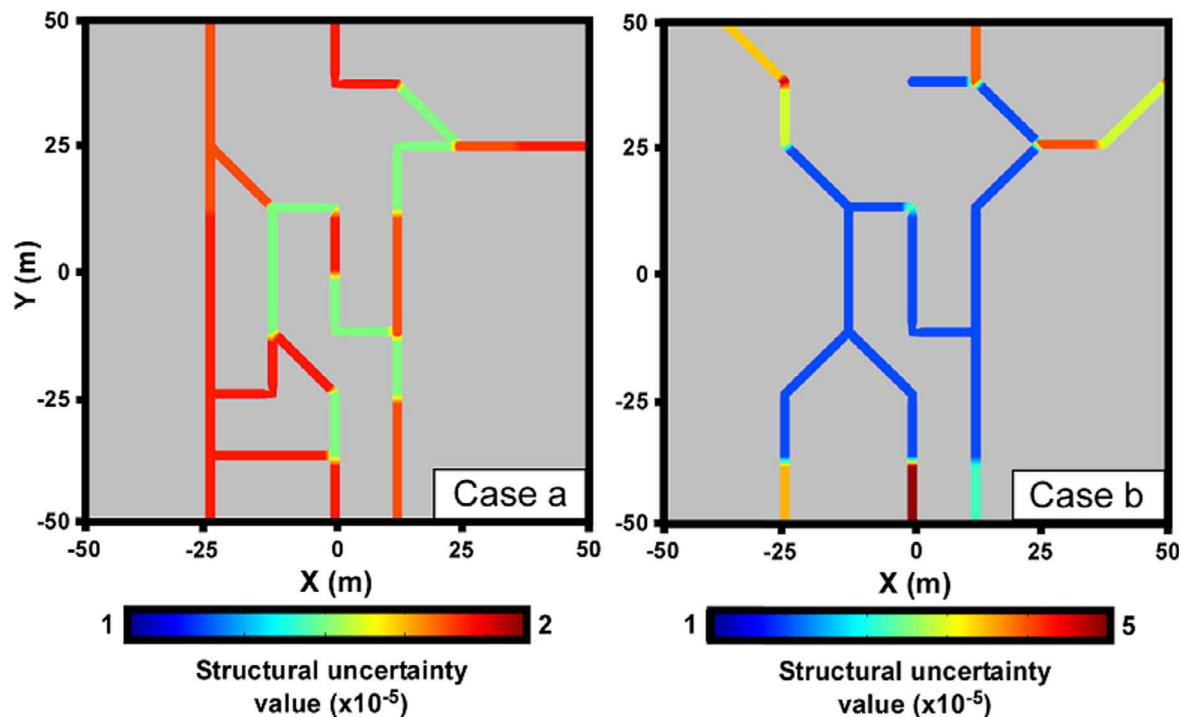


Fig. 11. Maps of the posterior uncertainties of the network local directions for the Cases a and b. In the Case a, started from a simple initial model, the highest uncertainties are distributed uniformly over the inverted network. In the Case b, started from a more detailed initial model, the highest uncertainties are located in the periphery of the model. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Conclusion

We present in this paper a novel deterministic inversion method that permits to characterize, in a partitioned model, the karst conduits and fractures network geometry and their hydraulic properties, including the transmissivity distribution of the matrices. The DNDI method let the modeler choose the partitioning of the model for the inversion. This ‘cursor’ permits to define either an inversion with a coarse partitioning for a quick approximation model, or with a fine partitioning and a longer computation time for a better fitting model. The use of a discrete network model permits to associate a specific behavior to the flows in the network and thus, produces more realistic models than an equivalent porous media model. This method can be easily adapted for channels or fractures network models by modifying the properties associated to the discrete network (these properties can also be directly linked to an aperture value, by choosing an adapted law). Therefore, we believe that the DNDI method is an interesting new imagery tool for the distributed modeling associated to a set of data from an investigation in a karstic and/or fractured aquifer.

We have realized different tests in three theoretical and simplified study cases with an increasing complexity, and the DNDI could always produce satisfying results, both on the reproduction of the generated data and on finding the network geometry and property values from the true model. As we have seen in the first study case, the result of the structural inversion is dependent on the positioning and the amount of observed data. This is true for any inversion, but is especially important in the case of highly heterogeneous aquifers for delineating the position of the heterogeneities. Therefore the result of the inversion has to be interpreted critically regarding the set of data used for it. A first critical analysis can be performed from the maps of posterior uncertainties on the structure or on the property values that can be produced by using the formulas we propose in this paper. The a priori information on the geometry of the network and on the property values is also a way to constrain the inversion in addition to the data. This information can be inferred from general field knowledges (geological and geophysical information, conduits observation in wells through video camera, other studies, etc.)

Because this method is deterministic, the choice of the initial model should be based on a relatively coherent possibility and should not be too far from the real solution in order to produce a good result. Therefore, we propose to couple the DNDI method to a multi-scale method. This consists in a first inversion started from an initial model which is followed by a new one that starts from the first inversion solution with a refined partitioning. This strategy permits to start from a simple initial model and to progressively make the model more complex and improve the solution.

An application of this method for mapping the conduits and fractures network with real data from a karstic field is planned for future works. These works will be more specifically focused on the sensitivity of the method to the spatial distribution of the measurement boreholes and on delineating the preferential flow paths in the network.

Conflict of interest

None.

Acknowledgments

We thank the Normandy region for providing financial support for the PhD of Pierre Fischer. We also would like to thank two anonymous reviewers for their contribution to the improvement of this article.

References

- Ackerer, P., Delay, F., 2010. Inversion of a set of well-test interferences in a fractured limestone aquifer by using an automatic downscaling parameterization technique. *J. Hydrol.* 389, 42–56.
- Ahmed, Soueid, Zhou, A.J., Jardani, A., Revil, A., Dupont, J., 2015. Image-guided inversion in steady-state hydraulic tomography. *Adv. Water Resour.* 82, 83–97.
- Bonneau, F., Henrion, V., Caumon, G., Renard, P., Sausse, J., 2013. A methodology for pseudo-genetic stochastic modeling of discrete fracture networks. *Comput. Geosci.* 56, 12–22.
- Borghi, A., Renard, P., Cornaton, F., 2016. Can one identify karst conduit networks geometry and properties from hydraulic and tracer test data. *Adv. Water Resour.* 90, 99–115.

- Butler, J.J., 2005. Hydrogeological methods for estimation of spatial variations in hydraulic conductivity. In: Rubin, Y., Hubbard, S.S. (Eds.), *Hydrogeophysics: Water Science and Technology Library 50* Springer, Dordrecht.
- Caers, J., Hoffman, T., 2006. The probability perturbation method: a new look at Bayesian inverse modeling. *Math. Geol.* 38 (No. 1), 81–100.
- Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., Slooten, L.J., 2005. Inverse problem in hydrogeology. *Hydrogeol. J.* 13 (No. 1), 206–222.
- Cliffe, K., Holton, D., Houston, P., Jackson, C., Joyce, S., Milne, A., 2011. Conditioning discrete fracture network models of groundwater flow. *Int. J. Numer. Anal. Model.* 8 (No. 4), 543–565.
- Collon, P., Bernasconi, D., Vuilleumier, C., Renard, P., 2017. Statistical metrics for the characterization of karst network geometry and topology. *Geomorphology* 283, 122–142.
- De Rooij, R., Perrochet, P., Graham, W., 2013. From rainfall to spring discharge: coupling conduit flow, subsurface matrix flow and surface flow in karst systems using a discrete-continuum model. *Adv. Water Resour.* 61, 29–41.
- Eisenlohr, L., Kiraly, L., Bouzelboudjen, M., Rossier, Y., 1997. Numerical simulation as a tool for checking the interpretation of karst spring hydrographs. *J. Hydrol.* 193, 306–315.
- Fischer, P., Jardani, A., Lecoq, N., 2017. A cellular automata-based deterministic inversion algorithm for the characterization of linear structural heterogeneities. *Water Resour. Res.* 53, 2016–2034. <http://dx.doi.org/10.1002/2016WR019572>.
- Ghasemizadeh, R., Hellweger, F., Butscher, C., Padilla, I., Vesper, D., Field, M., Alshawabkeh, A., 2012. Review: Groundwater flow and transport modeling of karst aquifers, with particular reference to the North Coast Limestone aquifer system of Puerto Rico. *Hydrogeol. J.* 20, 1441–1461.
- Grimstadt, A.-A., Mannseth, T., Naevdal, G., Urkedal, H., 2003. Adaptive multiscale permeability estimation. *Comput. Geosci.* 7 (No. 1), 1–25.
- Hao, Y., Yeh, T.-C.J., Xiang, J., Illman, W.A., Ando, K., Hsu, K.-C., Lee, C.-H., 2008. Hydraulic tomography for detecting fracture zone connectivity. *Ground Water* 46, 183–192.
- Hartmann, A., Goldscheider, N., Wägenet, T., Lange, J., Weiler, M., 2014. Karst water resources in a changing world: review of hydrological modeling approaches. *Rev. Geophys.* 52 (No. 3), 218–242.
- Illman, W.A., Liu, X., Takeuchi, S., Yeh, T.-C.J., Ando, K., Saegusa, H., 2009. Hydraulic tomography in fractured granite: Mizunami underground research site, Japan. *Water Resour. Res.* 45. <http://dx.doi.org/10.1029/2007WR006715>.
- Illman, W.A., 2014. Hydraulic tomography offers improved imaging of heterogeneity in fractured rocks. *Groundwater* 52 (5), 659–684.
- Jaquet, O., Siegel, P., Klubertanz, G., Benabderrahmane, H., 2004. Stochastic discrete model of karstic networks. *Adv. Water Resour.* 27, 751–760.
- Kovacs, A., 2003. Estimation of conduits network geometry of a karst aquifer by the means of groundwater flow modeling (Bure, Switzerland). *Boletín Geológico y Minero* 114 (No. 2), 183–192.
- Kovacs, A., Perrochet, P., Kiraly, L., Jeannin, P.-Y., 2005. A quantitative method for the characterization of karst aquifers based on spring hydrograph analysis. *J. Hydrol.* 303, 152–164.
- Larocque, M., Banton, O., Ackerer, P., Razack, M., 1999. Determining karst transmissivities with inverse modeling and an equivalent porous media. *Ground Water* 37 (No. 6), 897–903.
- Lavenue, M., de Marsily, G., 2001. Three-dimensional interference test interpretation in a fractured aquifer using the pilot point inverse method. *Water Resour. Res.* 37 (No. 11), 2659–2675.
- Le Coz, M., Bodin, J., Renard, P., 2017. On the use of multiple-point statistics to improve groundwater flow modeling in karst aquifers: a case study from the hydrogeological experimental site of Poitiers, France. *J. Hydrol.* 545, 109–119.
- Le Goc, R., de Dreuz, J.-R., Davy, P., 2010. An inverse problem methodology to identify flow channels in fractured media using synthetic steady-state head and geometrical data. *Adv. Water Resour.* 33, 782–800.
- Li, Z.-Y., Zhao, J.H., Qiao, X.H., Zhang, Y.X., 2014. An automated approach for conditioning discrete fracture network modelling to in situ measurements. *Australian J. Earth Sci.* 61, 755–763.
- Liedl, R., Sauter, M., Hockinghaus, D., Clemens, T., Teutsch, G., 2003. Simulation of the development of karst aquifers using a coupled continuum pipe flow model. *Water Res. Res.* 39, 1057.
- Meier, P., Medina, A., Carrera, J., 2001. Geostatistical inversion of cross-hole pumping tests for identifying preferential flow channels within a shear zone. *Groundwater* 39 (No. 1), 10–17.
- Ni, C.-F., Yeh, T.-C.J., 2008. Stochastic inversion of pneumatic cross-hole tests and barometric pressure fluctuations in heterogeneous unsaturated formations. *Adv. Water Resour.* 31, 1708–1718.
- Painter, S., Cvetkovic, V., 2005. Upscaling discrete fracture network simulations: an alternative to continuum transport models. *Water Resour. Res.* 41. <http://dx.doi.org/10.1029/2004WR003682>.
- Pardo-Iguzquiza, E., Dowd, P.A., Xu, C., Duran-Valsero, J.J., 2012. Stochastic simulation of karst conduit networks. *Adv. Water Resour.* 35, 141–150.
- Ronayne, M.J., 2013. Influence of conduit network geometry on solute transport in karst with a permeable matrix. *Adv. Water Resour.* 56, 27–34.
- Sharmeen, R., Illman, W.A., Berg, S.J., Yeh, T.-C.J., Park, Y.-J., Sudicky, E.A., Ando, K., 2012. Transient hydraulic tomography in a fractured dolostone: laboratory rock block experiments. *Water Resour. Res.* 48. <http://dx.doi.org/10.1029/2012WR012216>.
- Tarantola, A., Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion. *Rev. Geophys. Space Phys.* 20 (2), 219–232.
- Teutsch, G., 1993. An extended double-porosity concept as a practical modeling approach for a karstified terrain. *Hydrogeol. Process. Karst Terranes* 207, 281–292.
- Wang, X., Jardani, A., Jourde, H., Loneran, L., Cosgrove, J., Gosselin, O., Massonat, G., 2016. Characterisation of the transmissivity field of a fractured and karstified aquifer, Southern France. *Adv. Water Resour.* 87, 106–121.
- Wang, X., Jardani, A., Jourde, H., 2017. A hybrid inverse method for hydraulic tomography in fractured and karstic media. *J. Hydrol.* 551, 29–46.
- Yeh, T.-C.J., Liu, S., 2000. Hydraulic tomography: development of a new aquifer test method. *Water Resour. Res.* 36 (No. 8), 2095–2105.
- Yeh, T.-C.J., Lee, C.-H., 2007. Time to change the way we collect and analyze data for aquifer characterization. *Groundwater* 45 (No. 2), 116–118.
- Zhou, H., Gomez-Hernandez, J., Li, L., 2014. Inverse methods in hydrogeology: evolution and recent trends. *Adv. Water Resour.* 63, 22–37.

RESEARCH ARTICLE

10.1002/2017WR020921

Key Points:

- We apply a newly developed inversion method to characterize a karstified aquifer
- The approach permits to map the spatial distribution of the hydraulic conductivities in the fractured and matrix areas
- Our result model can simulate realistic flow behaviors in highly fractured aquifers

Correspondence to:

P. Fischer,
pierre.fischer1@univ-rouen.fr

Citation:

Fischer, P., Jardani, A., Wang, X., Jourde, H., & Lecoq, N. (2017). Identifying flow networks in a karstified aquifer by application of the Cellular Automata-based Deterministic Inversion Method (Lez aquifer, France). *Water Resources Research*, 53, 10,508–10,522. <https://doi.org/10.1002/2017WR020921>

Received 10 APR 2017

Accepted 27 NOV 2017

Accepted article online 30 NOV 2017

Published online 15 DEC 2017

Identifying Flow Networks in a Karstified Aquifer by Application of the Cellular Automata-Based Deterministic Inversion Method (Lez Aquifer, France)

P. Fischer¹ , A. Jardani¹, X. Wang², H. Jourde², and N. Lecoq¹
¹Normandie Univ, UNIROUEN, UNICAEN, CNRS, Rouen, France, ²Laboratoire Hydrosclences, Université de Montpellier, CNRS, Montpellier, France

Abstract The distributed modeling of flow paths within karstic and fractured fields remains a complex task because of the high dependence of the hydraulic responses to the relative locations between observational boreholes and interconnected fractures and karstic conduits that control the main flow of the hydrosystem. The inverse problem in a distributed model is one alternative approach to interpret the hydraulic test data by mapping the karstic networks and fractured areas. In this work, we developed a Bayesian inversion approach, the Cellular Automata-based Deterministic Inversion (CADI) algorithm to infer the spatial distribution of hydraulic properties in a structurally constrained model. This method distributes hydraulic properties along linear structures (i.e., flow conduits) and iteratively modifies the structural geometry of this conduit network to progressively match the observed hydraulic data to the modeled ones. As a result, this method produces a conductivity model that is composed of a discrete conduit network embedded in the background matrix, capable of producing the same flow behavior as the investigated hydrologic system. The method is applied to invert a set of multiborehole hydraulic tests collected from a hydraulic tomography experiment conducted at the Terrieu field site in the Lez aquifer, Southern France. The emergent model shows a high consistency to field observation of hydraulic connections between boreholes. Furthermore, it provides a geologically realistic pattern of flow conduits. This method is therefore of considerable value toward an enhanced distributed modeling of the fractured and karstified aquifers.

1. Introduction

The numerical modeling of groundwater flows within heterogeneous aquifers and the assessment of their hydrodynamic properties (such as the hydraulic conductivity and specific storage coefficient) remain actually an important and complex research challenge (Hartmann et al., 2014a; White, 2002). The main difficulties faced in the modeling of these types of aquifers are due to the high contrast in the hydraulic properties at small spatial scale, at the limits between conduits, fractures, and matrix. These heterogeneities lead to complex and discontinuous patterns of groundwater flows that are mainly controlled by the geometric characteristics of the fracture or conduit networks (spatial locations, apertures, sizes, densities). Most often, fractured and karstified aquifers are modeled by using a lumped simulation method (Arfib & Charlier, 2016; Dreiss, 1982; Hartmann et al., 2014b; Kong-A-Siou et al., 2015; Labat et al., 1999; Ladouche et al., 2014; Long & Derickson, 1999), in which the whole hydrosystem is considered as a grey-box (or black-box) in order to study the responses of the system in an output signal by conceptualizing some physical processes. This method can be useful to describe the global responses of a system to a rainfall signal but it does not give precise information on the flow behavior within the aquifer. A distributed hydrodynamic simulation method is more adequate to describe the mechanistic processes of water flows within a heterogeneous aquifer. The distributed hydrodynamic simulations can be categorized in three main approaches (Ghasemizadeh et al., 2012; Hartmann et al., 2014a; Kovacs & Sauter, 2007): (i) the equivalent porous media (Abusaada & Sauter, 2013; Wang et al., 2016) in which the hydraulic features of the fractured areas are approached with equivalent continuous hydraulic properties, (ii) the double continuum (Kordilla et al., 2012; Zimmerman et al., 1993), in which the model is conceptualized with two porous continuum media (matrix and conduit) that have distinct hydraulic properties, and (iii) the combined discrete-continuum (Jaquet et al., 2004; Saller et al., 2013) in which the discrete fractures are defined by their geometries and their local apertures, their interactions with the porous matrix media are included by using exchanging flow terms.

The characterization of the spatial distribution of the hydraulic properties is commonly provided from an inversion process coupled to a hydraulic tomography approach (Bohling et al., 2002; Cardiff & Kitanidis, 2009; Wang et al., 2017; Yeh & Liu, 2000; Zhu & Yeh, 2005). This approach consists of a joint analysis of a set of piezometric data collected as the responses of a water extraction during multiple pumping tests. Both steady state and transient hydraulic tomography have been considered in previous works. In a transient hydraulic experiment, both hydraulic conductivity and specific storage influence hydraulic head distribution (Castagna et al., 2011; Sharmeen et al., 2012; Zhu & Yeh, 2005), thus, more unknown parameters need to be estimated compared to a steady state experiment where drawdown data depend exclusively on hydraulic conductivity (Cardiff et al., 2009; Wang et al., 2016; Yeh & Liu, 2000).

The efficiency of the characterization depends on the number and spatial disposition of the boreholes used in the investigation. However, in the practice cases, the number of wells is usually insufficient to reduce the uncertainty and uniqueness of the solution. To overcome these difficulties, a priori knowledge is used to constrain the inverse problem. The geostatistical approach is widely applied to constrain the hydraulic tomography particularly for porous aquifers with a moderate variability in hydraulic conductivity (Fischer et al., 2017a; Hoeksema & Kitanidis, 1984; Lee & Kitanidis, 2014). To deal with discrete spatial patterns, the method was advanced by replacing the Gaussian prior model by a Laplace prior (or total variation prior) in the Bayesian framework (Lee & Kitanidis, 2013).

However, in the case of highly heterogeneous and complex aquifers, the use of geostatistical constraints can lead to unrealistic models dominated by a relatively smooth variability of the hydraulic properties.

In this paper, we apply a novel structural inversion method, the Cellular Automata-based Deterministic Inversion (CADI), to invert the steady state hydraulic head data recorded during a hydraulic tomography to image the spatial distribution of the hydraulic transmissivities in the fractured and karstified Lez aquifer (Southern France). The theoretical aspects of the CADI method have been developed in a previous article (Fischer et al., 2017b). This method is based on the Cellular Automata (CA) concept to parameterize the model. It permits a deterministic inversion of linear structures which is interesting for the modeling of fractures and karstic conduits. The paper is developed as follows: in the first section, we present the CADI algorithm and the concept used to parameterize the model and the inverse problem. Then, in the second section, we present investigations on the experimental site. Finally, we discuss the results of the application of the proposed inverse method to map the hydrodynamic properties of a karstified and fractured aquifer.

2. Methodology

2.1. Model Parameterization

We dedicate this section to describe briefly the main concepts of the CADI method. For more details about the theory of this method, we invite the readers to refer to Fischer et al. (2017b).

A Cellular Automata (CA) is a mathematical concept that permits the generation of discrete time-evolving cells grids. At a given CA time step (t_{CA}), the states of the cells is simultaneously modified following a global transition rule which involves the states of the cells in the neighborhood of each cell of the grid (Von Neumann & Burks, 1966).

The model in the CADI method is built as a lattice space Γ discretized in m squared property cells (in our case transmissivities) which are grouped in m_{CA} different CA pilot subspaces noted $\phi_i, i \in [1, m_{CA}]$ (with $m_{CA} \ll m$). The cells in a CA subspace can have two possible states: (i) state "matrix" with a transmissivity value β_{matrix} or (ii) state "conduit" with a transmissivity value $\beta_{conduit}$. The global structural distribution of the transmissivities in the model Γ is, thus, monitored by the different CA subspaces. Each subspace is piloted by a weighting distribution assigned to the neighborhood of each cell of the subspace (noted N). This distribution is set up among an inner circle of "activator" "matrix" neighbor cells and an outer circle of "inhibitor" "matrix" neighbor cells (Figure 1). An "activator" "matrix" neighbor will tend to transform a given cell of the subspace (in grey in Figure 1) in a state "matrix," while an "inhibitor" "matrix" neighbor will tend to transform this cell in a state "conduit." These circles are also radially split in eight weighting sectors for a higher weight distribution possibility. At a CA time step t_{CA} , a global transition rule compares,

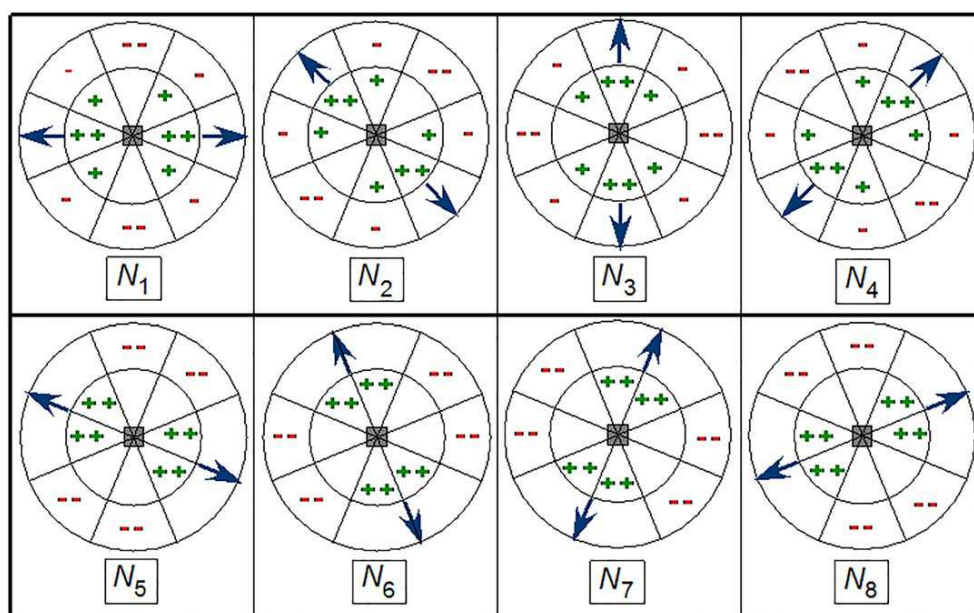


Figure 1. Scheme of the eight different weighting distributions N possibilities to parameterize the CA subspaces. Each distribution defines a different direction for the conduit-state generation shown by the arrows. The dual radius neighborhood is described here for a given cell in grey (the other cells are not shown for a reason of readability). In the configurations $N_i, i \in [1, 4]$, the circles are defined by an inner circle of radius 2 cells and an outer circle of radius 6 cells, and in the configurations $N_i, i \in [5, 8]$, the circles are defined by an inner circle of radius 4 cells and an outer circle of radius 5 cells. The neighbor cells of the greyed cell are split in eight internal “activator” weighting sectors and eight external “inhibitor” weighting sectors represented by the two radially split circles. A neighbor cell in state matrix can be associated (given its position in the neighborhood) to a positive weight “+ +” which is twice higher than a “+” weight, or to a negative weight “- -” which is twice higher than a “-” weight, or to a null weight in the empty sectors and beyond the neighborhood.

alternatively for each cell of the subspace, the cells in state “matrix” in the “activator” and “inhibitor” sectors of its neighborhood. For example, for a given cell, if the sum of “matrix”-state weights in its activator sector is higher than the “matrix”-state weights in its inhibitor sector, then this cell will become “matrix” in the next time step $t_{CA} + 1$, and in the opposite case this cell will become “conduit.” In the CADI algorithm, we configured eight different weighting distribution possibilities $N_i, i \in [1, 8]$ which define eight different directions of propagation for the conduit in a CA subspace (see the eight configurations in Figure 1). After several successive CA time steps transitions, a subspace ϕ will converge to a stable geometry noted $\hat{\phi}$ which depends on the weighting distribution chosen for N .

The eight configurations aforementioned are considered as the different possibilities for the CA subspace parameterization during the structural optimization of the inversion process. Thus, in the CADI method a converged CA subspace $\hat{\phi}(N, \beta)$ is parameterized by one of the eight weighted neighborhood configuration $N_i, i \in [1, 8]$ as structural parameter and by the transmissivity values $\beta = [\beta_{matrix}, \beta_{conduit}]$ as property parameter. The global partitioned model composed of all converged CA subspaces $\hat{\phi}_i, i \in [1, m_{CA}]$ will be referred to as $\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)$ with \mathbf{P}_N a m_{CA} vector of the different structural parameters piloting each CA subspace and \mathbf{P}_β a $2m_{CA}$ vector of the $\beta_{conduit}$ and β_{matrix} transmissivity values in each CA subspace of the model.

Initially in the CA temporal process, the whole model starts in an entire matrix state with only a single cell in state conduit. In the first CA time step, the conduit appears in the subspace of this initial cell, and as it arrives to the limit of this subspace it will potentially enter new subspaces (by local symmetry at the boundary between these subspaces) with another generation direction. Thus, from a unique conduit cell, the different subspaces permit the generation of a complex conduit network model at the end of the convergence of all CA (Figure 2).

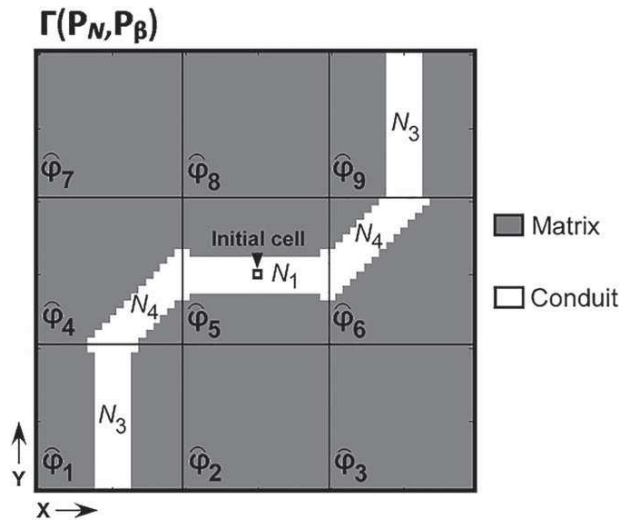


Figure 2. Presentation of a model in the CADI algorithm. Here the model is partitioned in nine subspaces controlled by CA. The model is parameterized by a structural parameter \mathbf{P}_N (here $\mathbf{P}_N(5)=N_1$; $\mathbf{P}_N(4)=\mathbf{P}_N(6)=N_4$, and $\mathbf{P}_N(1)=\mathbf{P}_N(9)=N_3$ (see Figure 1)) and a property values parameter \mathbf{P}_β (here every subspace is defined by the same β but it could vary in each subspace). Initially the whole model is considered as matrix, except an initial conduit cell. Within the CA time process, the conduit is generated from this initial cell and propagates through the model depending on the subspaces structural parameters until it reaches a global converged geometry.

2.2. Inverse Problem

In the Bayesian framework, the inverse problem retrieves a model matching the observed data and respecting the priori information (Tarantola & Valette, 1982). The optimization of the unknown parameters (here structural and property parameters) can be achieved by using an iterative deterministic algorithm that minimizes sequentially the following objective functions:

$$\Psi_{structure}(\mathbf{P}_N) = \frac{1}{2} (\mathbf{d}_{obs} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)))^T \mathbf{C}_d^{-1} (\mathbf{d}_{obs} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta))) + \frac{1}{2} (\mathbf{P}_{N,prior} - \mathbf{P}_N)^T \mathbf{C}_{P_N}^{-1} (\mathbf{P}_{N,prior} - \mathbf{P}_N) \quad (1)$$

$$\Psi_{properties}(\mathbf{P}_\beta) = \frac{1}{2} (\mathbf{d}_{obs} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)))^T \mathbf{C}_d^{-1} (\mathbf{d}_{obs} - f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta))) + \frac{1}{2} (\mathbf{P}_{\beta,prior} - \mathbf{P}_\beta)^T \mathbf{C}_{P_\beta}^{-1} (\mathbf{P}_{\beta,prior} - \mathbf{P}_\beta) \quad (2)$$

with \mathbf{d}_{obs} the vector of $n \times I$ observed data (such as hydraulic responses from pumping tests), $\mathbf{P}_{N,prior}$ and $\mathbf{P}_{\beta,prior}$ the a priori information to constrain the inversion of the structural and property parameters, \mathbf{C}_d a $(n \times n)$ covariance matrix of uncertainties on data, and \mathbf{C}_{P_N} a $(m_{CA} \times m_{CA})$ and \mathbf{C}_{P_β} a $(2m_{CA} \times 2m_{CA})$ covariance matrices of uncertainties on prior parameters. $f(\Gamma(\mathbf{P}_N, \mathbf{P}_\beta))$ denotes the nonlinear forward problem operator that links the hydraulic head data and the transmissivity field. $\Gamma(\mathbf{P}_N, \mathbf{P}_\beta)$ is the spatial partition of the transmissivity model Γ that is parameterized by the CA via the structural parameters \mathbf{P}_N and its property parameters \mathbf{P}_β .

The inversion process is conducted sequentially. First, the parameters \mathbf{P}_N and \mathbf{P}_β are initialized with reasonably chosen structural directions and transmissivity values for each subspace of the model. Then, the structural parameter \mathbf{P}_N is iteratively estimated with the fixed initial transmissivity model \mathbf{P}_β . Afterward, once the structural geometry is optimized, the spatial distribution of the transmissivity parameters \mathbf{P}_β is reconstructed considering this optimized structural geometry.

2.3. Optimization and Uncertainties Estimation

The optimization process begins with a sensitivity analysis: a local “One-factor-At-the-Time” (OAT) perturbation method, according to the classification in Pianosi et al. (2016). The structural sensitivity analysis considers the eight different neighborhood configurations (Figure 1) as structural parameters possibilities in each CA subspace in order to modify the conduit network and minimize the difference between the modeled data to the observed data. This sensitivity analysis establishes a $(8 \times m_{CA})$ sensitivity matrix \mathbf{S} .

At a k th iteration, for a modification in a CA subspace j by testing a configuration N_i , the element (i, j) of the matrix \mathbf{S} is calculated as:

$$\mathbf{S}^k(i, j) = \left(\mathbf{d}_{obs} - f \left(\Gamma \left(\mathbf{P}_N^k \Big|_{\mathbf{P}_N^k(j)=N_i}, \mathbf{P}_\beta \right) \right) \right)^T \mathbf{C}_d^{-1} \left(\mathbf{d}_{obs} - f \left(\Gamma \left(\mathbf{P}_N^k \Big|_{\mathbf{P}_N^k(j)=N_i}, \mathbf{P}_\beta \right) \right) \right) + \frac{1}{2} (\mathbf{P}_{N,prior}(j) - N_i)^T \mathbf{C}_{P_N}^{-1} (\mathbf{P}_{N,prior}(j) - N_i) \quad (3)$$

where $f(\Gamma(\mathbf{P}_N^k \Big|_{\mathbf{P}_N^k(j)=N_i}, \mathbf{P}_\beta))$ represents the data modeled through the local subspace parameter perturbation and $\mathbf{P}_{N,prior}(j) - N_i$ represents the gap between the local prior direction and the perturbation direction.

The new structural parameter \mathbf{P}_N^{k+1} is updated with the minimal value found in the matrix \mathbf{S} at the index $(i, j)_{\min}$ (where i represents the new configuration N_i for the subspace j). The updated parameter is built as an unique local improvement: $\mathbf{P}_N^{k+1} = \mathbf{P}_N^k$ expect for $\mathbf{P}_N^{k+1}(j_{\min}) = N_{i_{\min}}$. This update will generate a new structure for the iteration $k+1$. The same sensitivity analysis is repeated at each iteration until there is no more possible improvements in the structure.

Then, the uncertainties of the inverted structure are estimated from the last iteration sensitivity matrix, the prior uncertainties, and the value of the objective function:

$$\mathbf{C}_{\mathbf{P}_N}^{post}(j) = \left(\frac{1}{8} \sum_{i=1}^8 \hat{\mathbf{S}}(i,j) - \Psi_{structure}^{post} + \mathbf{C}_{\mathbf{P}_N}^{-1}(j,j) \right)^{-1} \quad (4)$$

with $\mathbf{C}_{\mathbf{P}_N}^{post}(j)$ the structural uncertainty for the subspace j , $\hat{\mathbf{S}}$ the last iteration sensitivity matrix and $\Psi_{structure}^{post}$ the value of the objective function after optimization. If a subspace conduit is well-constrained its value $\mathbf{C}_{\mathbf{P}_N}^{post}$ should be low.

After the structural optimization, a second optimization is led on the subspaces conduit and matrix transmissivity values \mathbf{P}_β , for the previously inverted structure. This optimization is an iterative process using a finite difference sensitivity analysis. The $(n \times 2m_{CA})$ Jacobian sensitivity matrix is defined for its index (i,j) as:

$$\mathbf{J}(i,j) = \left. \frac{\partial f_i}{\partial \mathbf{P}_\beta} \right|_{\mathbf{P}_\beta(j) = \mathbf{P}_\beta(j) + \Delta \mathbf{P}_\beta} \quad (5)$$

with $\Delta \mathbf{P}_\beta$ the finite difference step, f_i the forward problem variation on a data i for a variation on $\mathbf{P}_\beta(j)$.

At an iteration k , the updated values \mathbf{P}_β^{k+1} are calculated from a Newton linearization:

$$\begin{aligned} \mathbf{P}_\beta^{k+1} = & \mathbf{P}_\beta^k + \left((\mathbf{J}^k)^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{J}^k + \mathbf{C}_{\mathbf{P}_\beta}^{-1} \right)^{-1} \cdot (\mathbf{J}^k)^T \cdot \mathbf{C}_d^{-1} \cdot (\mathbf{d}_{obs} - f(\mathbf{P}_N, \mathbf{P}_\beta^k)) \\ & + \mathbf{C}_{\mathbf{P}_\beta}^{-1} \cdot (\mathbf{P}_{\beta,prior} - \mathbf{P}_\beta^k) \end{aligned} \quad (6)$$

This iterative process continues until an acceptable convergence of the objective function is achieved.

Then, the uncertainties of the property values are estimated from the last iteration Jacobian matrix \mathbf{J}^{post} with the posterior covariance matrix $\mathbf{C}_{\mathbf{P}_\beta}^{post}$:

$$\mathbf{C}_{\mathbf{P}_\beta}^{post} = \left((\mathbf{J}^{post})^T \cdot \mathbf{C}_d^{-1} \cdot \mathbf{J}^{post} + \mathbf{C}_{\mathbf{P}_\beta}^{-1} \right)^{-1} \quad (7)$$

The diagonal entries of this matrix give the variances of the property values. The square root values of these entries represent the standard deviation of the uncertainties on the estimated transmissivity field.

3. Application

3.1. Site Presentation

The Terrieu experimental site is located in the North of Montpellier, Southern France (Figure 3a). The site has been performed for the hydrodynamic studies of the Lez aquifer that is mainly composed of Early Cretaceous and Late Jurassic limestones. This site is one of the sites of the French research network SO Karst which was developed to monitor the karstic aquifers in France (Jourde et al., 2011; www.sokarst.org).

The Terrieu site sits on a local monocline structure that trends NE-SW and dips at about 15–20° toward NW. The surface area of the experimental site is about 2,400 m² (40 m × 60 m). Detailed fracture mappings, conducted at the ground surface of the site and nearby outcrops, have indicated that two major fracture sets (trending ENE-WSW and NW-SE, respectively) are present in the study area (Jazayeri Noushabadi, 2009; Wang et al., 2016).

Twenty-two boreholes have been drilled at the site to study the local-scale hydraulic behavior of the aquifer (Figure 3b). These boreholes are vertical and have a mean total drilled depth of 55 m. Downhole videos recorded in some boreholes have shown that the upper 30–40 m of the drilled formations are largely comprised of thin-layered, marly, early Cretaceous limestones while the lower part mainly consists of purer, massive, and nonaquifer late Jurassic limestones. The early Cretaceous limestone has a low permeability therefore it forms a confined upper boundary for the aquifer existing at the interface between these two units. A number of well-developed karstic conduits, with apertures up to 50 cm, have also been identified on downhole video logs. These karstic conduits were found to be present at a depth between 35 and 40 m (Jazayeri Noushabadi et al., 2011; Wang et al., 2016) at the interface of the aforementioned two units. The observed local orientation of the karstic conduits is indicated as green lines in Figure 3c. The extent of the

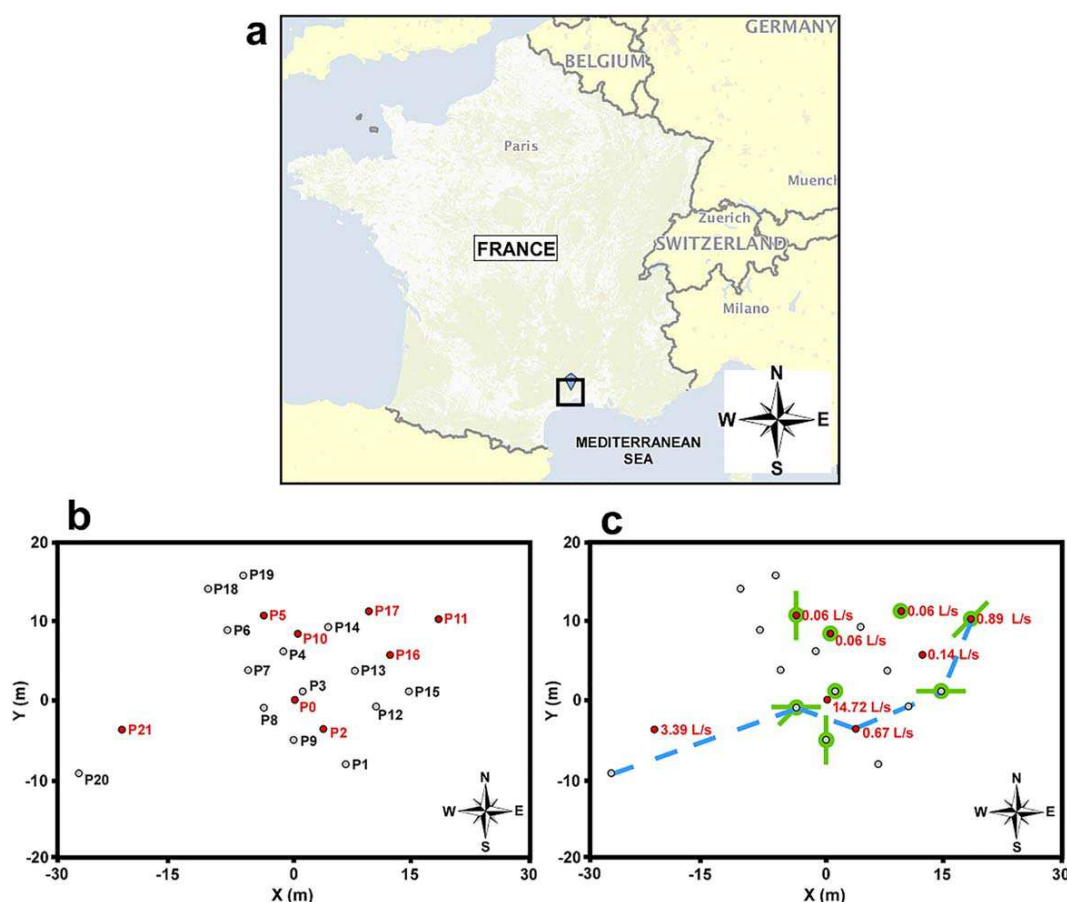


Figure 3. (a) Map indicating the location of the experimental site. The black square indicates the location of the Lez aquifer in which the Terrieu site is included. (b) Distribution of 22 boreholes of the Terrieu experimental site. The red dots indicate the boreholes where the pumping tests were performed while the grey dots indicate the measurement boreholes. (c) Pumping rates (red captions). Inferred principal flow path connectivity (blue dotted lines) and local karstic conduits (green lines) based on downhole videos, well logs, and packer tests. The orientation of the green lines indicates the orientation of local karstic features observed on downhole videos. A green dot indicates that no karstic features were seen in this borehole.

lower rock unit is unknown due to limited drilled depth (maximum of 60 m). Well logs (temperature and electrical conductivity) and straddle packer tests have shown that a preferential flow path (blue line in Figure 3c) exists along the major bedding plane corresponding to the interface between the two major rock units (Dausse, 2015; Jourde et al., 2002). All the observations from downhole videos, well logs, and packer tests have led to the conceptual model that a network of interconnected karstic conduits developed along an important bedding plane comprises the main flow paths of the experimental site (Wang et al., 2016).

Eight cross-hole pumping tests, in the form of hydraulic tomography, have been performed at the experimental site. The applied flow rate of each pumping test ranges from 0.2 to 53 m³/h depending on the well productivity and whether the well is connected to a high-permeability feature (Figure 3c). The drawdown of water table level generated by the pumping tests did not reach to the depth of the important bedding plane where main flow occurs; this means that the karstic network was saturated during the pumping tests. All tests reached steady state. The field-scale hydraulic tomography yielded a total number of 168 draw-down steady state measurements, which is the main data set of the present work.

This set of drawdown responses represents the observed data used in the inversion process, while the inferred preferential flow path and local conduits direction information presented in Figure 3c are not taken into account in the inversion but will be used for evaluating the effectiveness of the inversion results.

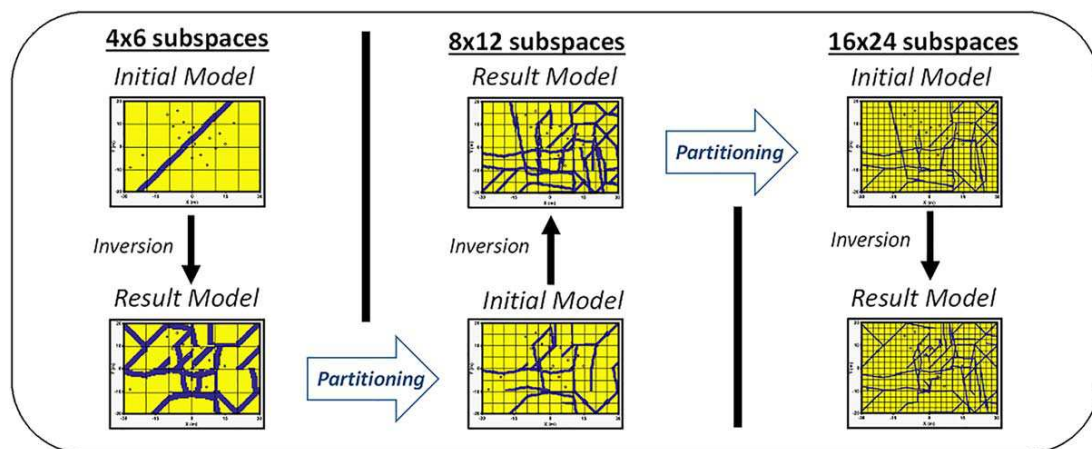


Figure 4. Schematic showing the sequential series of inversions led to obtain the final flow network model. The initial model was partitioned with 4×6 subspaces for its inversion. The inverted flow network model was then used as a new initial model for an inversion with 8×12 subspaces. The same operation was repeated on last time so that our final flow network has a partitioning of 16×24 subspaces.

3.2. Modeling Method

We have applied the CADI method to image the hydraulic transmissivity distribution at the Terrieu site from the joint inversion of the set of 168 steady state drawdown data. A two-dimensional equivalent porous medium parameterization was adopted to model the domain. The porous medium was distributed in two states: bedding plane fissured matrix (CADI matrix state) or karstic conduit (CADI conduit state). We adopted an evolving meshing, that is, for each iteration, refined preferentially at the boundaries of the conduits. This avoids an over-meshing of the model, especially in the matrix zones and, thus, reduces the computation time of the forward problem solver. The forward problem (i.e., steady state diffusion equation in saturated porous media) was solved using a commercial software (COMSOL Multiphysics). The inversion process was implemented as a MATLAB code that is connected to the COMSOL solvers via a local server. The local flow network model was enclosed in a large regional buffer zone ($1,000 \times 1,000 \text{ m}^2$) to reduce the influence of the boundaries conditions.

Based on a multiscale inversion method (Grimstadt et al., 2003), the flow network model was sequentially partitioned during the inversion process (see Figure 4).

The different inversion parameters chosen for the final inversion (16×24 CA subspaces) are presented in Table 1. The property value parameter β was chosen to define the exponent of the transmissivity ($\log_{10}(T)$). Thus, the prior standard deviation also applies to the exponent ($T \times 10^{\pm \sigma_T}$). The initial transmissivities values were chosen according to the values found during the field characterizations done by Jazayeri Noushabadi (2009) and Dausse (2015). We did not make any prior assumption on the structure directions in \mathbf{P}_N and \mathbf{C}_{P_N} .

Table 1
List of the Inversion Parameter Values Chosen for the Final Inversion (16×24 Subspaces)

Parameter	Value
Model dimension	$X = [-30, 30] \text{ m}; Y = [-20, 20] \text{ m}$
Partitioning/total cells amount ($Y \times X$)	$16 \times 24 / 320 \times 480$
Conduit aperture	50 cm
Data uncertainties	$\mathbf{C}_d = \sigma_{data}^2 \times \text{Id}(n); \sigma_{data}^2 = 0.01 \text{ m}$
\log_{10} transmissivity prior standard deviation	$\mathbf{C}_{P_N} = \sigma_T^2 \times \text{Id}(2m_{CA})$
Initial \log_{10} transmissivities	$\sigma_T^2 = 0.001$ for the conduits; $\sigma_T^2 = 1$ for the matrix $\beta_{conduit} = -1 (10^{-1} \text{ m}^2/\text{s}); \beta_{matrix} = -8 (10^{-8} \text{ m}^2/\text{s})$ $\beta_{regional} = -2 (10^{-2} \text{ m}^2/\text{s})$
Finite difference step	$\Delta \mathbf{P}_\beta = 10^{-4}$
Observation data (\mathbf{d}_{obs})	168 steady state drawdown responses

We began the inversion with a model partitioned in a large discretization 4×6 CA subspaces. In the flow network model, the karstic conduit was generated from an initial fracture cell placed at the P8 borehole (Figure 4). This choice was made based on field hydrogeological knowledge that P8 has the highest probability of intersecting the main flow path. The initial network was a simple conduit with a direction choice based on the principal flow path direction and with an aperture of 2 m. The remaining parts of the model were assigned with a transmissivity of matrix. As the structural geometry for this initial conduit was optimized, we added a second conduit in the model, orthogonally to the principal flow path and intersecting P8, and we led the structural inversion with these two principal conduits directions. When the inversion process found a 4×6 subspaces result model, we refined the obtained solution to a model partitioned in 8×12 subspaces (a CA subspace was divided into four new ones with the same parameters) and used it as initial model for a more precise inversion solution with conduits with an aperture of 1 m. Finally, we partitioned once more this new 8×12 solution into a 16×24 subspaces model which was once again used as the initial model for a final inversion solution with conduits with an aperture of 50 cm. This sequential modeling reduces the inversion time and allows interpretation of the importance of the flow paths found according to their emergence at different scales during the inversion process and thus their influence on the convergence of the objective function.

3.3. Results and Discussion

The simulated hydraulic heads from the final emergent flow network model show a high correlation to the field measurements (Figure 5a). It can be seen that a higher degree of difference between the simulated and measured hydraulic head exists for small drawdowns. This may be due to the fact that the small hydraulic drawdowns are more sensible to microfractures that were embedded in the equivalent porous matrix in the current model. The resulting karstic flow network model from the inversion process is presented in Figure 5b. The hydraulic transmissivity of the bedding plane fissured matrix was only slightly perturbed (i.e., a small deviation from the initial value of $10^{-8} \text{ m}^2/\text{s}$) during the entire inversion process (Figure 5b), the drawdown data were essentially reproduced only by the karstic network geometry. The emergent flow network shows a high consistency with field observations shown in Figure 3c in terms of connectivity between boreholes (Figure 5c) and local direction of karstic conduits (Figure 5d). The simulated drawdown maps from the resultant flow network model are presented in Figure 6.

The simulated karstic network structure in the Terrieu experimental site investigated through the hydraulic tomography can be schematized using the emergent flow network model as shown in Figure 7. This schema conceptualizes the site in three dimensions taking into account the slope of the bedding plane.

The maps of the simulated drawdowns (Figure 6) using the final flow network model highlight the high degree of heterogeneity of the experimental site. The steady state drawdown cone of each pumping test is highly irregular, and the shape of the drawdown cone can have very different geometry in different pumping tests. Three types of behaviors can be distinguished: (i) pumping in the nonkarstified bedding plane matrix (P2, P10, P17) associated with a low pumping rate and a small influence zone of drawdown localized around the pumping borehole; (ii) pumping in low-productivity conduits (P5, P11, P16) associated with a low pumping rate and a large influence zone of small drawdowns covering the entire field; and (iii) pumping in high-productivity conduits (P0, P21) associated with a high pumping rate and a large influence zone of drawdowns impacting the whole field. In our model, pumping in the matrix only induces a response in the vicinity of the pumped well, while pumping in the conduit network induces responses impacting significantly a much larger impact area (delineated by dotted white lines in Figure 6). This shows that the global flows in the model are controlled by the karstic conduits, which can be linked to the real behavior of the site in which flows follow the important discontinuities (Jourde et al., 2002). The modeled steady state drawdowns were compared to the field observations made by Jazayeri Noushabadi (2009) and Dausse (2015), and they appear to be similar to the estimated drawdowns from the field pumping tests.

To verify the results, we have compared the inverted flow network to the mapped connectivity and local direction knowledges gained from slug tests and downhole videos (Figure 3c). This information was not taken into account during the inversion process, but they permit to assess the effectiveness of the CADI approach with the steady state drawdown responses.

The observed connectivity described in Figure 3c between P20, P8, P2, P12, P15, and P11 can be reproduced by the inverted flow network model (Figure 5c). However, the drawdown map for P2 (Figure 6)

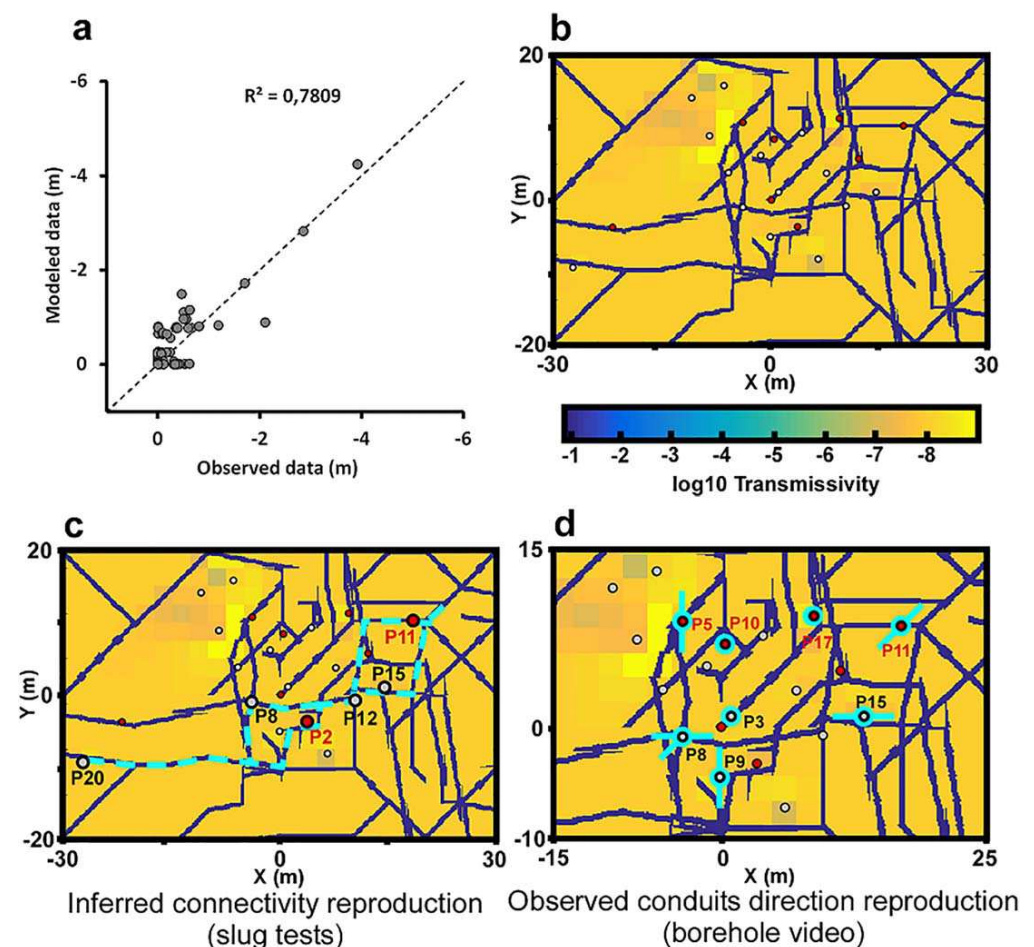


Figure 5. (a) Comparison of the observed drawdowns to the drawdowns modeled by the inverted flow model. (b) Resultant model of the inversion modeling showing the heterogeneous distribution of the transmissivities. (c) Comparison of the result model with the known preferential flow path connectivity (interpreted in the model in dotted blue lines). (d) Superimposition of the known local conduits direction (shown as blue lines) presented in Figure 3c.

shows that it is not directly connected to the main flow network in the model. We noticed that the establishment of this connectivity appeared quickly during the multiscale structural inversion (Figure 4). It indicates the importance of the connectivity between these boreholes toward the reproduction of the global drawdown distribution on the site.

Regarding the knowledge about local conduit direction (Figure 5d), in the inverted model, P8 and P15 are in fact intersecting conduits following the same local direction as observed from downhole videos. In the cases of P5 and P11, where conduits have been observed in the field, the inverted flow network map shows that they intersect preferential flow paths, but not following exactly the same direction as field observations. From the drawdown maps (Figure 6), we can see that simulated pumping in P10 and P17 in our model behave like pumping in the matrix; thus, P10 and P17 are not connected to the karstic network. This is consistent with the information highlighted by camera observations. Only two inverted local flow structures do not reproduce downhole videos observations: P3 seems to be connected to the karstic network; however, no karstic conduits were identified on its downhole video. On the contrary, P9 seems to be located in the matrix in the inverted model, while a conduit has been observed in its borehole video.

In general, the inverted flow network model reconstructs most of the knowledges that we have about the site, even if this information was not used in the inversion process. Thus, it seems that the drawdown data set alone provides sufficient information to reproduce the true network connectivity between boreholes.

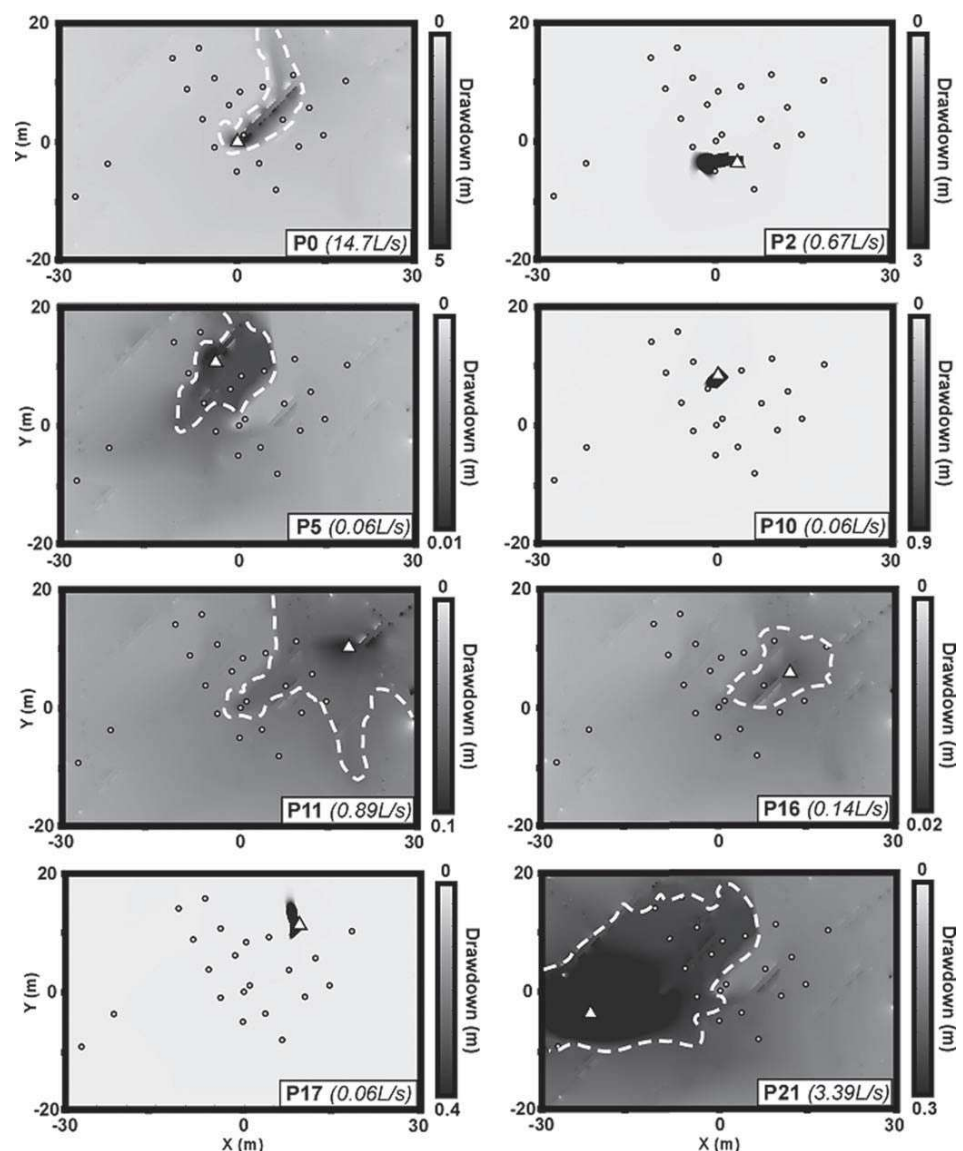


Figure 6. Maps of hydraulic drawdowns calculated from the result flow network model. The drawdowns are shown for each of the pumping wells (white triangles) used for the hydraulic tomography (the pumping rate is indicated in each figure). The drawdowns can have very different forms depending on the localization of the borehole in a conduit or in the matrix, highlighting the heterogeneity of the model. Pumping in the matrix (P2, P10, P17) results in a very local drawdown, while pumping in a conduit (P0, P5, P11, P16, P21) produces a more global drawdown in the whole model (in these cases, the area the most impacted by the pumping is delimited by white dotted lines).

Figure 8 presents uncertainty maps, which reflects the reliability of the inversion results, and the areas of the model that are covered by the information given from the different local responses in boreholes. The posterior structural uncertainty map indicates that the structure uncertainty is lower in the center of the model domain than in the periphery. The highest uncertainty in the inverted flow structures (in orange/yellow in Figure 8) occurs consistently in the regions beyond the one constrained by the borehole pattern, where no hydraulic information is available. The most reliable parts of the structure (in blue/green) are located in the middle of the domain, where the borehole pattern is dense, and therefore provides more hydraulic information. The map of the posterior standard deviations of the \log_{10} transmissivity shows that the matrix transmissivity value is well-constrained only near some boreholes (e.g., P0, P2, P4, P9, P10, P13,

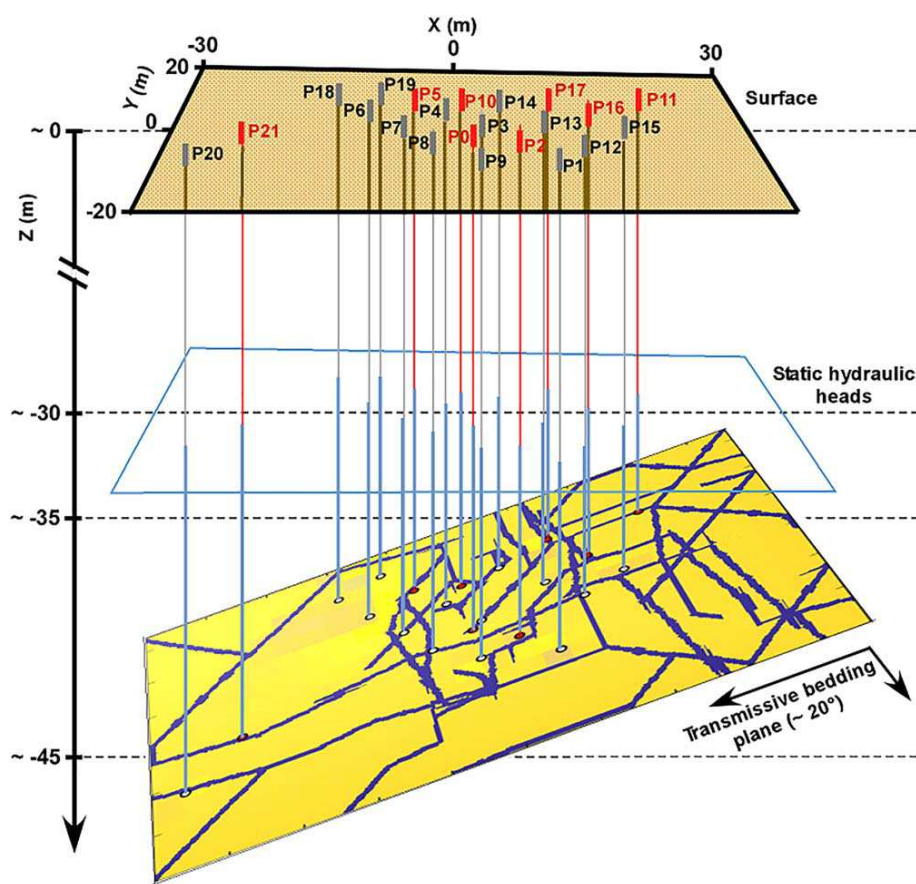


Figure 7. Schematic representation of the modeled karstic structure at the Terrieu experimental site, considering the geological information, the hydraulic tomography investigation, and the flow network produced by inversion with the CADI method. The red lines indicate the boreholes where the pumping tests were performed, while the grey lines indicate the measurement boreholes.

and P17). Thus, the pumping tests do not permit a characterization of the bedding plane transmissivity very far from the boreholes.

Figure 9 shows maps of flow velocities for pumping in P0 and P21, the two most productive boreholes, intersected by karstic conduits. The distribution of the modeled flow velocities during each pumping test are controlled by the high-transmissivity karstic features in the model. It appears clearly that the water pumped in these boreholes is mobilized in the karstic network, while the velocity of the water in the bedding plane matrix is low. Even if the CADI method allows a certain imagery of the flow paths that honors the observed hydraulic connectivity to be determined, it may not produce a geomorphologically realistic model. This is caused by the nonunique nature of inverse problem. To provide more realistic karst networks, the hydraulic data should be inverted jointly with other sources of data coming from geophysical and tracer tests or/and constrained by geomorphological and geological information. By comparing the velocity maps for P0 and for P21, we observed that a larger pumping rate (P0) induces an increase of the flow velocities around the borehole, but it also generates a mobilization of water from conduits located further from the pumping borehole, and thus, characterization of a larger area. On the contrary, a lower pumping rate (P21) will more specifically characterize the karstic conduits near the borehole. This implies that the use of transient modeling of a variable pumping rate to characterize both small scale and large scale karstic network would be useful. The velocities maps also indicate that the North-South oriented conduit in the center of the model ($X = -5$ m) is less solicited by the pumping in P0 and P21. This same conduit is, however, characterized more specifically by a pumping in P5, which would indicate a dual flow direction in the karstic network of the model (N-S and E-W).

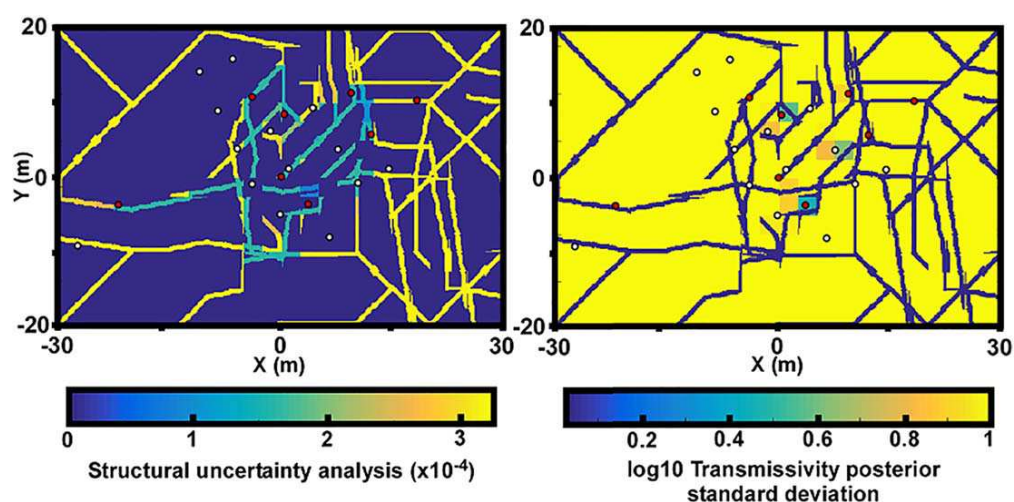


Figure 8. The map of the network structural uncertainties (left) shows that the network geometry is well-constrained especially in a zone between each borehole in the center of the model, and compared to the map of transmissivities standard deviation (right), the hydraulic data permitted to constrain more the conduits position than the matrix.

In Figure 10, the inversion result produced by the CADI method is compared to the one produced by the Sparse Nonlinear OPTimizer (SNOPT) method at the same field site and with the same data set (Wang et al., 2016). The SNOPT method is a classical efficient algorithm for nonlinear large-scale inverse problems. The model of Wang et al. (2016) is composed of a grid of squared cells of $0.5 \times 0.5 \text{ m}^2$; each cell is assigned with a transmissivity value, which is optimized in the inversion process. The optimization of the transmissivity values was constrained by upper and lower limits (10^{-1} and $10^{-8} \text{ m}^2/\text{s}$, respectively).

While the CADI method, associated to the multiscale process, can be initialized with a simple model to achieve an inverted flow network consistent with field observations, the SNOPT method required some connectivity information included in the initial model to converge to a coherent model. Overall, the SNOPT and CADI methods generate similar results in terms of medium morphology, where a highly conductive conduit network is embedded in a background matrix. The SNOPT method let more freedom for the optimization of both conduit and matrix transmissivities instead of primarily constraining the flows in a network as done by the CADI method. Thus, it allows establishing a smoother transient of transmissivity from flow paths to the

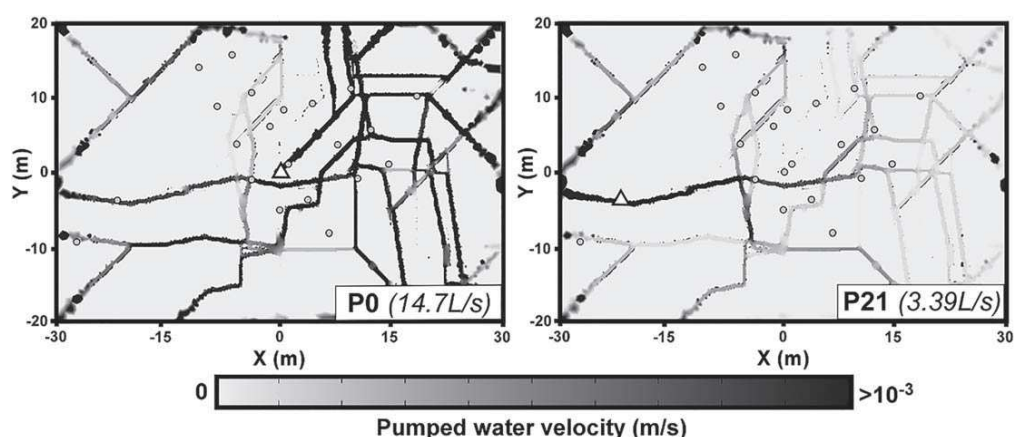


Figure 9. Maps of the pumped water velocities calculated by the result model for a pumping in borehole P0 and in borehole P21 (the two most productive pumping). The pumping boreholes are indicated by white triangles. For a reason a better readability of the low velocities, the scale has been fixed on a maximal velocity of 10^{-3} m/s , thus in the blackest zones, the velocity can be higher than this value (up to 10^{-2} m/s near the pumping point for P0).

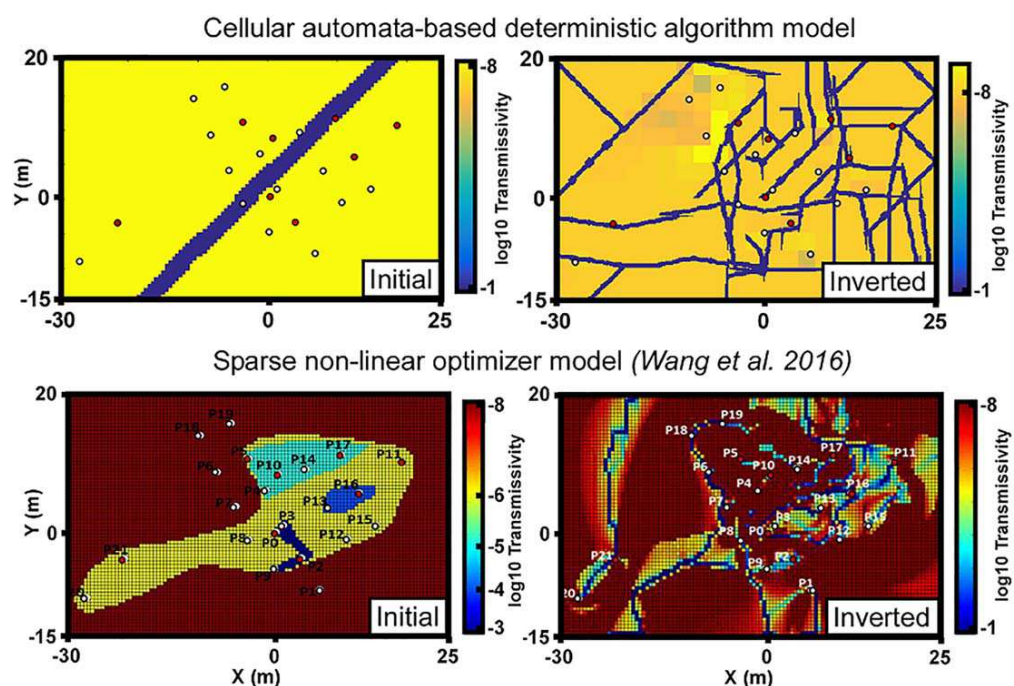


Figure 10. Comparison of the inversion result produced by the CADI method and by the SNOPT method (Wang et al., 2016) at the same scale of the Terrieu field site and with same hydraulic data set. The initial models are shown on the left and the inverted models are presented on the right.

background medium. This is important, in some cases, to reproduce the diffusive behavior of local microfractures, which cannot be modeled by the channelized flow in conduits. For this reason, the SNOPT result permits a better reproduction of the measured drawdowns ($R^2 \sim 0.9$ of SNOPT compared to the $R^2 \sim 0.78$ of CADI). However, the result by the CADI method is more consistent with the field knowledge of the local conduit orientation and borehole connectivity presented in Figure 3c. The constraints imposed in the CADI method also present the advantage to represent a network in an equivalent porous media result to mimic the hydraulic behaviors in the karstic and fractured environment, thus allowing an easier interpretation of the preferential flow paths positioning and of the connectivity between boreholes.

Although the combination of the CADI method and steady state hydraulic tomography technique appears effective in the identification of flow networks in karstic fields due to its ability of constraining the flows in a network, some limits can be highlighted.

First, we point out that even though the inverted flow paths near the boreholes show a high consistency to field observations, it is possible that some of the inverted conduits may represent the assemble hydraulic behavior of highly fractured areas, while in reality conduits may not be present in those areas. In the same idea, the network in the model is generated with a constant aperture, and thus does not permit to identify the flow paths associated to developed conduits to those associated to fractured zones. To achieve a more realistic network configuration, this method needs to be further developed to include a variable aperture model.

Second, the use of steady state head data may also limit the identification of the preferential paths. In fact, a pumping reaching the steady state in a karstic field mobilizes water in both the conduit network and the fractured rock matrix. The latter tends to blur the hydraulic responses of the conduit network. Using transient data could be useful to improve the inversion result and to reduce the nonuniqueness of the inversion. However, transient inversions may increase significantly the computational demand. A more efficient solution may be to use harmonic pumping (e.g., Rabinovich et al., 2015; Soueid Ahmed et al., 2016). The harmonic data permit to perform inversions in the frequency domain and may highlight the influence of flow from specific components of the system, and therefore would achieve a better characterization of the conduit network.

4. Conclusion

The use of the CADI method for a site scale karstic field characterization shows its capabilities in identifying the preferential flow network. The inverted distributed karstic flow network model reproduces the observed data while maintaining the realism of a highly heterogeneous flows distribution. The constraints imposed for the inversion process result in a model that localizes the conduits of the karstic network, which in turn control the main drawdowns direction.

The use of hydraulic tomography data in the inversion permits the characterization of the highly heterogeneous discrete conduit network. However, this requires a large number of boreholes, and, especially boreholes intersecting the karstic network. Results of the flow velocities in P0 and P21 show that pumping in boreholes intersecting the karstic network with different pumping rate could characterize the network at different scales (locally or more globally). This promotes the use, in the future, of the recently developed harmonic pumping technique at the Terrieu experimental site to identify more precisely the geometry of the karstic network with less pumping tests.

Some necessary conditions were met in the case of the Terrieu experimental site to successfully apply the CADI method. The karstic network at this site was mainly constrained within a bedding plane, which permits a two-dimensional modeling. Also, the subsurface flow in the Terrieu site is constrained mainly by the karstic conduits, while the matrix has a very low transmissivity. This permitted the model to start from an assumption of a two domain site (matrix and conduit), distinguished by a high contrast in transmissivity. In the future, the current method will be further developed to deal with transient hydraulic data and/or to use harmonic pumping responses, which are more sensitive to the karst connectivity, for improving the characterization of the geometry of the conduit networks.

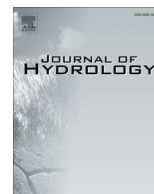
Acknowledgments

We thank the Normandy region for providing financial support for the PhD of Pierre Fischer. We would like to thank three anonymous reviewers for their relevant comments and corrections which permitted to significantly improve the quality of this article. The hydraulic data collection and hydrogeological characterization of the Terrieu experimental site presented in this work are the property of the French research network SO Karst and are freely accessible on demand to a person in charge of the network (see www.sokarst.org for more information about the SO Karst institution and contacts).

References

- Abusaada, M., & Sauter, M. (2013). Studying the flow dynamics of a karst aquifer system with an equivalent porous medium model. *Ground Water*, 51(4), 641–650.
- Arfib, B., & Charlier, J.-B. (2016). Insights into saline intrusion and freshwater resources in coastal karstic aquifers using a lumped Rainfall-Discharge-Salinity model (the Port-Miou brackish spring, SE France). *Journal of Hydrology*, 540, 148–161.
- Bohling, G. C., Zhan, X., Butler, J. J. Jr., & Zheng, L. (2002). Steady shape analysis of tomographic pumping tests for characterization of aquifer heterogeneities. *Water Resources Research*, 38(12), 1324. <https://doi.org/10.1029/2001WR001176>
- Cardiff, M., Barrash, W., Kitanidis, P. K., Malama, B., Revil, A., Straface, S., & Rizzo, E. (2009). A potential-based inversion of unconfined steady-state hydraulic tomography. *Ground Water*, 47(2), 259–270.
- Cardiff, M., & Kitanidis, P. K. (2009). Bayesian inversion for facies detection: An extensible level set framework. *Water Resources Research*, 45: W10416. <https://doi.org/10.1029/2008WR007675>
- Castagna, M., Becker, M. W., & Bellin, A. (2011). Joint estimation of transmissivity and storativity in a bedrock fracture. *Water Resources Research*, 47, W09504. <https://doi.org/10.1029/2010WR009262>
- Dausse, A. (2015). *Facteurs d'échelle dans la hiérarchisation des écoulements au sein d'un aquifère karstique: Analyse multi-échelles des propriétés hydrodynamiques et de transport de l'aquifère de Lez* (PhD thesis), French: Université de Montpellier.
- Dreiss, S. J. (1982). Linear kernels for Karst aquifers. *Water Resources Research*, 18, 865–876.
- Fischer, P., Jardani, A., & Lecoq, N. (2017b). A cellular automata-based deterministic inversion algorithm for the characterization of linear structural heterogeneities. *Water Resources Research*, 53, 2016–2034. <https://doi.org/10.1002/2016WR019572>
- Fischer, P., Jardani, A., Soueid Ahmed, A., Abbas, M., Wang, X., Jourde, H., & Lecoq, N. (2017a). Application of large-scale inversion algorithms to hydraulic tomography in an alluvial aquifer. *Ground Water*, 55, 208–218.
- Ghasemizadeh, R., Hellweger, F., Butscher, C., Padilla, I., Vesper, D., Field, M., & Alshawabkeh, A. (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*, 20, 1441–1464.
- Grimstadt, A.-A., Mannseth, T., Naevdal, G., & Urkedal, H. (2003). Adaptive multiscale permeability estimation. *Computers & Geosciences*, 7(1), 1–25.
- Hartmann, A., Goldscheider, N., Wagener, T., Lange, J., & Weiler, M. (2014a). Karst water resources in a changing world: Review of hydrological modeling approaches. *Reviews of Geophysics*, 52, 218–242.
- Hartmann, A., Mudarra, M., Andreo, B., Marin, A., Wagener, T., & Lange, J. (2014b). Modeling spatiotemporal impacts of hydroclimatic extremes on groundwater recharge at a Mediterranean karst aquifer. *Water Resources Research*, 50, 6507–6521. <https://doi.org/10.1002/2014WR015685>
- Hoeksema, R. J., & Kitanidis, P. K. (1984). An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resources Research*, 20, 1003–1020.
- Jaquet, O., Siegel, P., Klubertanz, G., & Benabderrhamane, H. (2004). Stochastic discrete model of karstic networks. *Advances in Water Resources*, 27, 751–760.
- Jazayeri Noushabadi, M. R. (2009). *Characterization of relationship between fracture network and flow-path network in fractured and karstic reservoirs: Numerical modeling and field investigation (Lez aquifer, Southern France)* (PhD thesis). France: Université de Montpellier.
- Jazayeri Noushabadi, M. R., Jourde, H., & Massonnat, G. (2011). Influence of the observation scale on permeability estimation at local and regional scales through well tests in a fractured and karstic aquifer (Lez aquifer, Southern France). *Journal of Hydrology*, 403, 321–336.

- Jourde, H., Batiot-Guilhe, C., Bailly-Comte, V., Bicalho, C., Blanc, M., Borrell, V., . . . Van-Exter, S. (2011). The MEDCYSS observatory, a multi scale observatory of flood dynamics and hydrodynamics in karst (Mediterranean border Southern France). In N. Lambrakis, G. Stournaras, & K. Katsanou (Eds.), *Advances in the research of aquatic environment. Environmental earth sciences*. Berlin, Heidelberg: Springer.
- Jourde, H., Cornaton, F., Pistre, S., & Bidaux, P. (2002). Flow behavior in a dual fracture network. *Journal of Hydrology*, 266, 99–119.
- Kong-A-Siou, L., Johanne, A., Borrell Estupina, V., & Pistre, S. (2015). Neural networks for karst groundwater management: Case of the Lez spring (Southern France). *Environmental Earth Sciences*, 74, 7617–7632.
- Kordilla, J., Sauter, M., Reimann, T., & Geyer, T. (2012). Simulation of saturated and unsaturated flow in karst systems at catchment scale using a double continuum approach. *Hydrology of Earth System Sciences*, 16, 3909–3923.
- Kovacs, A., & Sauter, M. (2007). Modelling karst hydrodynamics. In N. Goldscheider & D. Drew (Eds.), *Methods in karst hydrogeology* (pp. 65–91). London, UK: Taylor & Francis/Balkema.
- Labat, D., Ababou, R., & Mangin, A. (1999). Linear and nonlinear input/output models for karstic springflow and flood prediction at different time scales. *Stochastic Environmental Research and Risk Assessment*, 13, 337–364.
- Ladouche, B., Marechal, J.-C., & Dorfliger, N. (2014). Semi-distributed lumped model of a karst system under active management. *Journal of Hydrology*, 509, 215–230.
- Lee, J., & Kitanidis, P. K. (2013). Bayesian inversion with total variation prior for discrete geologic structure identification. *Water Resources Research*, 49, 7658–7669. <https://doi.org/10.1002/2012WR013431>
- Lee, J., & Kitanidis, P. K. (2014). Large-scale hydraulic tomography and joint inversion of head and tracer data using the principal component geostatistical approach (PCGA). *Water Resources Research*, 50, 5410–5427. <https://doi.org/10.1002/2014WR015483>
- Long, A. J., & Derickson, R. G. (1999). Linear systems analysis in a karst aquifer. *Journal of Hydrology*, 219, 206–217.
- Pianosi, F., Beven, K., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., & Wagener, T. (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, 79, 214–232.
- Rabinovich, A., Barrash, W., Cardiff, M., Hochstetler, D. L., Bakhos, T., Dagan, G., & Kitanidis, P. K. (2015). Frequency dependent hydraulic properties estimated from oscillatory pumping tests in an unconfined aquifer. *Journal of Hydrology*, 531, 2–16.
- Saller, S. P., Ronayne, M. J., & Long, A. J. (2013). Comparison of a karst groundwater model with and without discrete conduit flow. *Hydrogeology Journal*, 21, 1555–1566.
- Sharmeen, R., Illman, W. A., Berg, S. J., Yeh, T.-C. J., Park, Y.-J., Sudicky, E. A., & Ando K. (2012). Transient hydraulic tomography in a fractured dolostone: Laboratory rock block experiments. *Water Resources Research*, 48. <https://doi.org/10.1029/2012WR012216>
- Soueid Ahmed, A., Jardani, A., Revil, A., & Dupont, J. P. (2016). Joint inversion of hydraulic head and self-potential data associated with harmonic pumping tests. *Water Resources Research*, 52, 6769–6791. <https://doi.org/10.1002/2016WR019058>
- Tarantola, A., & Valette, B. (1982). Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics and Space Physics*, 20(2), 219–232.
- Von Neumann, J., & Burks, A. W. (1966). *Theory of self-reproducing automata*. Champaign, IL: University of Illinois Press.
- Wang, X., Jardani, A., & Jourde, H. (2017). A hybrid inverse method for hydraulic tomography in fractured and karstic media. *Journal of Hydrology*, 551, 29–46.
- Wang, X., Jardani, A., Jourde, H., Lonergan, L., Cosgrove, J., Gosselin, O., & Massonat, G. (2016). Characterisation of the transmissivity field of a fractured and karstic aquifer, Southern France. *Advances in Water Resources*, 87, 106–121.
- White, W. B. (2002). Karst hydrology: Recent developments and open questions. *Engineering Geology*, 65, 85–105.
- Yeh, T.-C. J., & Liu, S. (2000). Hydraulic tomography: Development of a new aquifer test method. *Water Resources Research*, 36, 2095–2105.
- Zhu, J., & Yeh, T.-C. J. (2005). Characterization of aquifer heterogeneity using transient hydraulic tomography. *Water Resources Research*, 41, W07028. <https://doi.org/10.1029/2004WR003790>
- Zimmerman, R. W., Chen, G., Hadgu, T., & Bodvarsson, G. S. (1993). A numerical dual-porosity model with semianalytical treatment of fracture/matrix flow. *Water Resources Research*, 29, 2127–2137.



Research papers

Hydraulic analysis of harmonic pumping tests in frequency and time domains for identifying the conduits networks in a karstic aquifer



P. Fischer^{a,*}, A. Jardani^a, M. Cardiff^b, N. Lecoq^a, H. Jourde^c

^a Normandie Univ, UNIROUEN, UNICAEN, CNRS, M2C, 76000 Rouen, France

^b Department of Geoscience, University of Wisconsin-Madison, Madison, WI, USA

^c Hydrosiences Montpellier (HSM), Univ. Montpellier, CNRS, IRD, Montpellier, France

ARTICLE INFO

Article history:

Received 24 September 2017

Received in revised form 28 February 2018

Accepted 4 March 2018

Available online 6 March 2018

This manuscript was handled by P. Kitanidis, Editor-in-Chief, with the assistance of Jean-Raynald de Dreuzy, Associate Editor

Keywords:

Oscillatory signal

Heterogeneity

Characterization

Aquifer

Frequency domain

Karst

ABSTRACT

In a karstic field, the flow paths are very complex as they globally follow the conduit network. The responses generated from an investigation in this type of aquifer can be spatially highly variable. Therefore, the aim of the investigation in this case is to define a degree of connectivity between points of the field, in order to understand these flow paths. Harmonic pumping tests represent a possible investigation method for characterizing the subsurface flow of groundwater. They have several advantages compared to a constant-rate pumping (more signal possibilities, ease of extracting the signal in the responses and possibility of closed loop investigation). We show in this work that interpreting the responses from a harmonic pumping test is very useful for delineating a degree of connectivity between measurement points. We have firstly studied the amplitude and phase offset of responses from a harmonic pumping test in a theoretical synthetic modeling case in order to define a qualitative interpretation method in the time and frequency domains. Three different type of responses have been separated: a conduit connectivity response, a matrix connectivity, and a dual connectivity (response of a point in the matrix, but close to a conduit). We have then applied this method to measured responses at a field research site. Our interpretation method permits a quick and easy reconstruction of the main flow paths, and the whole set of field responses appear to give a similar range of responses to those seen in the theoretical synthetic case.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Characterization of the hydrodynamic properties of porous or fractured aquifers is a common challenge in many areas including: exploitation and protection of water resources, oil production, geothermal energy extraction, civil engineering, remediation engineering and storage of radioactive waste. Among many approaches (tracer tests, slug tests, geophysical investigations, etc.), this characterization can be accomplished using pumping tests, in which the values of the transmissivity and storativity parameters are derived from the analysis of the hydraulic responses to the aquifer stimulation (Butler, 2005). However, in field conditions, many noise sources are unmanageable hydraulic contributions that can contaminate the signal induced by pumping, such as: aquifer recharge, river-aquifer interactions, evapotranspiration by plants, tidal fluctuations, or unplanned/unknown pumping at nearby water supply boreholes. For example, in unconfined aquifers

evapotranspiration occurring during pumping tests may result in a spurious “signal” (see, e.g., Cardiff et al., 2009).

To address this difficulty, harmonic pumping tests have been suggested as an efficient approach to characterize the hydraulic properties, by making the hydraulic signal exploitable even for low signal amplitudes and noises corruptions (Cardiff and Barrash, 2015). The oscillations in the hydraulic responses caused by a harmonic excitation with a known frequency can, in fact, be more easily extracted from the ambient noises, by applying filtering techniques (Bakhos et al., 2014), than the hydraulic responses generated by a constant rate pumping. Harmonic pumping tests also offer the possibility of avoiding the non-linear regimes of groundwater flow by controlling the characteristics of the periodic excitation. This excitation can be conducted through a pumping-reinjection system (Rasmussen et al., 2003), or even without any pumping or injection, by only using a moving mass (Guiltinan and Becker, 2015). Harmonic pulse tests by using cyclic injection rates have also been conducted by Sun et al. (2015, 2016) in a deep subsurface aquifer for leakage detection. Harmonic pumping tests have been especially investigated for their abilities

* Corresponding author.

E-mail address: pierre.fischer1@univ-rouen.fr (P. Fischer).

for characterizing aquifers properties. The analysis and interpretation of the harmonic data for a characterization are most often realized with analytical solutions (Renner and Messar, 2006; Dagan and Rabinovich, 2014; Rabinovich et al., 2015). Among these attempts, Black and Kipp (1981) proposed solutions for analytical models, under the homogeneity assumption, to derive the average transmissivity and storativity parameters from harmonic test analysis. Rasmussen et al. (2003) applied an analytical model expressed in frequency form to interpret the piezometric data generated from harmonic pumping tests in a porous contaminated aquifer, in order to estimate the average values of the hydraulic properties. The numerical models have also been used particularly in the frequency domain instead of the time domain in order to reduce the computing time and to take into account the spatial heterogeneity of the hydraulic properties (Black and Kipp, 1981; Cardiff et al., 2013). Modeling and imaging potential of harmonic pumping tests data were also already addressed in the literature for the reconstruction of the spatial variability of hydraulic properties (Lavenue and de Marsily, 2001; Cardiff et al., 2013; Zhou et al., 2016; Soueid Ahmed et al., 2016).

Some works have focused more specifically on the capacity of harmonic pumping for characterizing fractured aquifers. Renner and Messar (2006) applied harmonic pumping tests on a fractured sandstone aquifer to deduce the average values of transmissivity and storativity fields using a cyclical hydraulic excitation by alternating pumping, no flow and injection periods. At the same study site Maineult et al. (2008) used the self-potential method to monitor, remotely at the ground surface, the piezometric fluctuations caused by these harmonic excitations. More recently, Guiltinan and Becker (2015) similarly conducted periodic slug tests on isolated fractures without any extraction or injection, only by oscillating a slug at different depths of the water column in the well, to characterize the hydraulic connectivity of the fracture using analysis of phase shift and attenuation of the signal with an analytical model. Schuite et al. (2017) used tilt data recorded at the ground surface to follow the oscillatory deformations induced by harmonic pulses performed in a fractured aquifer.

Among the works previously cited and dealing with this subject, none have sought to model the impact of spatial high contrasts of hydraulic property variations on oscillatory testing. In karstic aquifers, where low-transmissivity host rocks are directly adjacent to sparsely-distributed, high-transmissivity fractures and conduits, such a scenario is clearly present and leads to flow paths particularly constrained. Wells in karstic aquifers therefore may be expected to have drastically different responses depending on whether they are located on or near a fracture or a conduit. In fact, the main flow paths in subsurface karstic field follow the conduit and fracture network and one would need to have an idea of its positioning for understanding the flow behavior (White, 2002; Saller et al., 2013). Commonly, karstic fields are described by their connectivity, whether the flows between wells follows a highly conductive path or not. Jazayeri Noushabadi et al. (2011) defined three types of hydrodynamic response to pumping and pulse tests, as a function of the degree of conductivity of the flow path network in the investigated karstic aquifer. The authors assumed that a short time lag and a high amplitude hydrodynamic response (conduit type hydrodynamic response) was the consequence of a high permeability and high connectivity between the observation well and the main flow path network (karst conduits that generate the large scale permeability of the aquifer), while ‘fracture type’ and ‘matrix type’ hydrodynamic responses were related to both a lower permeability and lower connectivity between the observation wells and the main flow path network. To be more consistent with previous works, this connectivity must be defined in term of property distribution as a ‘static’ or ‘topological’ connectivity, and in term of physical flow/transport processes as a ‘dynamical’

connectivity (Renard and Allard, 2013; Tyukhova and Willmann, 2016).

In this work, we propose an oscillatory signal analysis method, based on a synthetic and simplified model, for qualitatively interpreting hydraulic responses of a karstic field to a harmonic pumping test performed at different frequencies. This method consists in interpreting the ‘dynamical’ connectivity information from the responses in terms of ‘topological’ connectivity within the karstic field. In the first section of the manuscript, we present the mathematical framework used to simulate numerically, in both temporal and frequency domains, the groundwater flow responses to a sinusoidal excitation in a fractured aquifer. These numerical approaches are applied in the second section to a hypothetical karstic aquifer characterized by the presence of a simple karstic network. In the third section, we perform signal analysis by comparing point measurements of spatial amplitude decay and phase shift values with respect to the source of the oscillating signal at the pumping well to determine a conduit (conduit network), dual (fissure, conduit proximity) or matrix flow connectivity. Finally, we apply the same analysis method on real field data acquired on a karstic field located in Southern France. This signal analysis method permits an easy, fast and coherent interpretation of the preferential flow paths’ location on this site.

2. Theoretical background

In this article we use several terms such as harmonic, oscillatory, periodic, or sinusoidal signals, which all refer to the same idea of a signal defined by amplitude, a mean value and a period, repeating over time.

The harmonic pumping signal Q used in this manuscript is mathematically defined as:

$$Q(t) = -Q_A \cos(\omega t) + Q_m \quad (1)$$

where t is the time (s), Q is the time-dependent flow rate signal (m^3/s), Q_A is the amplitude of the oscillatory portion of the flowrate signal (m^3/s), Q_m is the flow rate signal mean value (m^3/s), and $\omega = \frac{2\pi}{T}$ is the angular frequency where T denotes the period of the pumping signal (s). We note that this flow formulation reproduces only the extraction of water with a flow rate sinusoidally fluctuating around the mean value Q_m . This flow rate signal can also be rewritten in a complex form:

$$Q(t) = -Q_{\text{osc}}(t) + Q_m \quad (2)$$

with $Q_{\text{osc}}(t) = \text{Re}(Q_A e^{i\omega t})$, and i representing the imaginary unit.

In order to describe the hydraulic drawdown responses of a confined karstic aquifer in 2D to a harmonic pumping signal, we represent the aquifer in a model domain Ω with a coupled discrete-continuum concept (for more details about the coupled discrete-continuum modeling, see Teutsch, 1993; Liedl et al., 2003). We combine Darcy’s law and the law of conservation of mass in a 2D matrix domain (intact rock) containing a 1D discrete conduits domain, which gives us the following partial differential equations, both defined in the Cartesian coordinates dimension:

$$\begin{cases} S_{S,\text{mat}} \frac{\partial h}{\partial t} - \nabla \cdot (K_{\text{mat}} \nabla h) = \frac{-Q(t)}{V_{\text{el}}} \delta(x - x_s) & \text{in the 2D matrix} \\ S_{S,\text{cond}} \frac{\partial h}{\partial t} - \nabla_T \cdot (K_{\text{cond}} \nabla_T h) = \frac{-Q(t)}{V_{\text{el}}} \delta(x - x_s) & \text{in the 1D conduits} \end{cases} \quad (3)$$

where h represents the hydraulic drawdown which links the 1D and the 2D domains (m), Q is the harmonic pumping rate (m^3/s) in an elementary volume V_{el} (m^3) of the matrix or the fractures, $\delta(x - x_s)$ represents the Dirac distribution where x_s denotes the coordinates of the pumping well that can be positioned in the matrix domain or in a conduit. $S_{S,\text{mat}}$ and $S_{S,\text{cond}}$ are the specific

storages in the matrix and the conduits (m^{-1}), K_{mat} and K_{cond} are the conductivities in the matrix and the conduits (m/s), and ∇_T is the tangential gradient operator for the hydraulic equation in the conduits described as discrete elements at the internal boundaries of the domain.

Both governing equations can be solved numerically by following no-drawdowns initial and boundary conditions:

$$\begin{cases} h(x, y, t) = 0 & \forall (x, y) \in \Omega & \text{when } t = 0 \\ h(x, y, t) = 0 & \forall t \geq 0 & \text{when } (x, y) \in \Omega_{boundary} \end{cases} \quad (4)$$

A sufficiently large equivalent porous media buffer zone encloses the model, in order to reduce the effects of the boundaries $\Omega_{boundary}$ on the flows within the model area of interest. The nature of the excitation signal applied at the pumping borehole creates hydraulic responses $h(t)$ composed of a sinusoidal signature $h_{osc.}(t)$ and a non-sinusoidal signature $h_{lin.}(t)$, which must be eliminated to deal only with the harmonic component.

$$h(x, y, t) = h_{osc.}(x, y, t) + h_{lin.}(x, y, t) \quad (5)$$

The hydraulic periodic signature $h_{osc.}$ can be expressed in a complex formulation:

$$h_{osc.}(x, y, t) = Re(H_\omega(x, y)e^{i\omega t}) \quad (6)$$

where H_ω is a complex number representing the wave phasor at the given frequency.

Therefore we can compute numerically the harmonic responses due to the sinusoidal signal of the pumping rate by a reformulation of the time domain groundwater equation in a frequency domain (Black and Kipp, 1981; Cardiff et al., 2013):

$$\begin{cases} i\omega S_{S,mat} H_\omega - \nabla \cdot (K_{mat} \nabla H_\omega) = \frac{Q_A}{V_{el}} \delta(x - x_s) & \text{in the 2D matrix} \\ i\omega S_{S,cond} H_\omega - \nabla_T \cdot (K_{cond} \nabla_T H_\omega) = \frac{Q_A}{V_{el}} \delta(x - x_s) & \text{in the 1D conduits} \end{cases} \quad (7)$$

with H_ω the complex field variable which links the 1D and the 2D domains and describes the harmonic signal responses spatially.

In the frequency domain the initial and boundary conditions become:

$$\begin{cases} H_\omega(x, y) = 0 & \forall (x, y) \in \Omega & \text{initially} \\ H_\omega(x, y) = 0 & \text{when } (x, y) \in \Omega_{boundary} \end{cases} \quad (8)$$

The amplitude and phase offset values of the response to the harmonic pumping signal at a given position (x, y) are calculated from the complex variable H_ω value:

$$\begin{aligned} \text{Amplitude: } A(x, y) &= \sqrt{(Re H_\omega(x, y))^2 + (Im H_\omega(x, y))^2} \text{ in m,} \\ \text{Phase offset: } \Phi(x, y) &= \frac{180}{\pi} \text{atan2}(-Im H_\omega(x, y), Re H_\omega(x, y)) \text{ in }^\circ \end{aligned} \quad (9)$$

with Re the real part of H_ω , Im the imaginary part of H_ω , and atan2 the two-arguments inverse tangent function in radian mode (rad).

Then the time-dependent periodic response can be reconstructed as:

$$h_{osc.}(x, y, t) = A(x, y) \cos(\omega t - \Phi(x, y) \frac{\pi}{180}) \quad (10)$$

Thus the frequency domain solver permits to reproduce periodic time-dependent responses while avoiding the use of a time domain solver.

3. Synthetic application

3.1. Model structure

In order to interpret drawdown responses to harmonic pumping tests in a real karstic aquifer (Terrieu, Southern France), we have studied the spatial drawdown responses from a synthetic case model of dimensions 50 m \times 50 m. This synthetic case model was built in consideration of our current understanding of the Terrieu karstic field (presented more specifically in Section 4.1) and our pumping signals during investigations at the field site. A simple karstic network is represented in the model as a 1D discrete geometry in a 2D continuum homogeneous matrix (Fig. 1).

The homogeneous matrix is associated with a 10^{-6} m/s conductivity value and a 10^{-4} m $^{-1}$ specific storage value. The karstic network is composed of a large conduit with a 0.1 m/s equivalent hydraulic conductivity and two thin conduits with a 0.01 m/s equivalent hydraulic conductivity. The whole conduit network is associated with a 10^{-8} m $^{-1}$ specific storage.

Eight different boreholes have been placed in this model at different strategic positions (Table 1). Three boreholes are located in the large conduit (LC: P5, P6, P8), two in the thin conduit (TC: P2, P3), two in the matrix but near to the conduit network (M-NC: P1, P4) and one in the matrix, distant from the conduit network (M: P7). We applied a 5 min period harmonic pumping signal (as defined in Eq. (1)) in each of these boreholes alternatively, while measuring the drawdown responses in the seven other boreholes. Different values of pumping amplitudes and mean flow rates were chosen according to the positioning of the pumping borehole (in a conduit or the matrix) in order to simulate a difference of productivity at each location.

The synthetic case model is enclosed in a 1,000 m \times 1,000 m buffer zone with a global 10^{-3} m/s conductivity value and a 10^{-4} m $^{-1}$ specific storage value. The boundaries of the buffer zone are associated with a no-drawdown condition. Thus the buffer zone reduces boundaries effects on drawdowns simulated within the central area of the model.

The governing equation (Eqs. (3) and (7)) for the simulation in the model were solved with the software COMSOL Multiphysics using a finite element method on a triangular adaptive grid (with

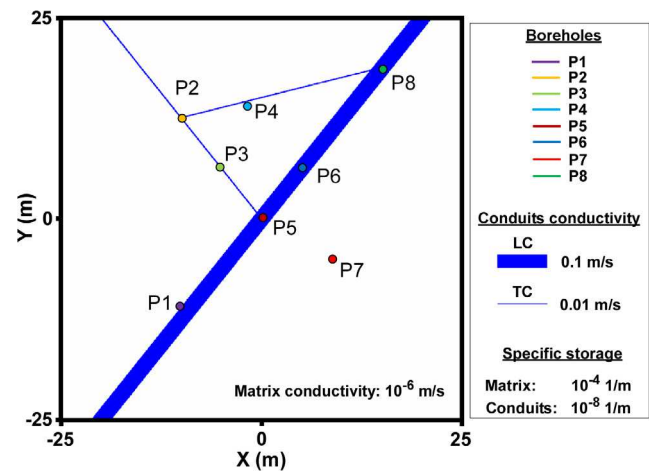


Fig. 1. The theoretical synthetic case used to study the responses of a harmonic pumping in a karstic field. A karstic network (in blue) composed of a large conduit (LC) and two thin conduits (TC) crosses a homogeneous matrix (in white). All conduits are 1-D features in the model, but shown with conductivity-weighted thicknesses for clarity. Eight different boreholes are positioned in the model and represent pumping or measurement points. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Coordinates, position and pumping signal parameters for the eight boreholes. For the positioning M = Matrix, TC = Thin Conduit, LC = Large Conduit and NC = Near Conduit. The pumping signal parameters are the amplitude (Q_A) and the mean flow rate (Q_m) (see Eq. (1)).

	P1	P2	P3	P4	P5	P6	P7	P8
X; Y (m)	−10; −11	−10; 12.5	−5; 6.25	−2; 14	0; 0	5; 6.25	9; −5	15; 18.75
Position	M-NC (1 m from LC)	TC	TC	M-NC (50 cm from TC)	LC	LC	M	LC
Q_A/Q_m (L/s)	0.5/1	2/4	2/4	0.5/1	2/4	2/4	0.5/1	2/4

a mesh refinement around the model discrete structures: the linear conduits and the boreholes points presented in Fig. 1) considering the initial and boundary conditions described in Eqs. (4) and (8).

3.2. Modeling in time domain

We firstly solved the synthetic case model in the time domain (Eq. (3)). The top graphic in Fig. 2 shows the time domain drawdown response in all boreholes for a harmonic pumping in P3 (in

a thin conduit). Except for P7, we notice a periodic signal in the responses of each borehole. Moreover, past the first signal period, the drawdown responses can be represented as the sum of a linear drawdown (extracted by linear regression) and a purely oscillatory signal (Fig. 2). This oscillatory signal can be parameterized, for each borehole, by a value of amplitude and a value of phase offset. In the particular case of P7, its oscillatory signal is almost null (<1 mm), thus its drawdown response is composed almost exclusively of a linear signal.

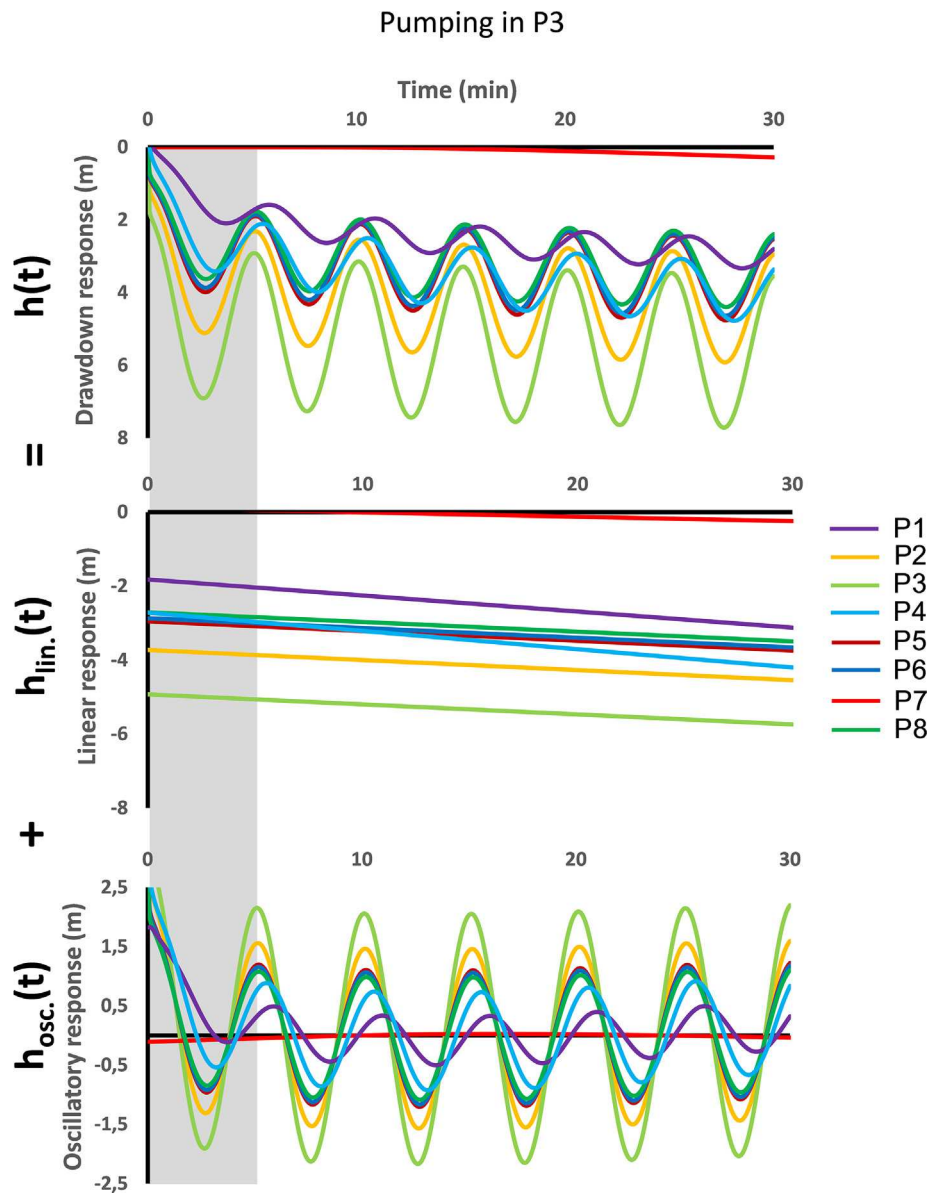


Fig. 2. Drawdown responses h in each borehole to a harmonic pumping in P3 in a time domain simulation. If the greyed portion of the time series is not considered, these drawdown responses can be described as the sum of a linear signal h_{lin} and a purely oscillatory signal h_{osc} .

From the observations made in Fig. 2, the drawdown response at a given position, after the first signal period, can be mathematically approximated as the sum of a linear function and an oscillatory function:

$$h(t) \simeq \underbrace{-at - h_0}_{h_{lin.}(t)} + \underbrace{A \cos\left(\frac{2\pi}{T}t - \Phi \frac{\pi}{180}\right)}_{h_{osc.}(t)} + H_0 \quad (11)$$

with h (m) the time domain drawdown response, $h_{lin.}$ (m) the linear part of the response with its slope a (m/min) and its intercept values h_0 (m), $h_{osc.}$ (m) the oscillatory part of the response with A (m), Φ (°) and T (min) its amplitude, phase offset and period values, and H_0 (m) the initial water level (in our synthetic case, this value is considered as 0).

For the interpretation of the responses we will be more specifically interested in the oscillatory response $h_{osc.}$.

3.3. Modeling in frequency domain

3.3.1. Frequency/time domains comparison

We also solved the synthetic case in the frequency domain (Eq. (7)). In this way, instead of spatial drawdown values, we calculate the spatial variation of amplitude and phase offset values in the response signal. We can then reconstruct for a given position the oscillatory signal $h_{osc.}$ in the drawdown response from these amplitude and phase offset values as described in Eq. (10). By comparing the reconstructed frequency domain signals to the time domain oscillatory ones, we show that they are almost identical (see the example for pumping in P3 in Fig. 3 and Appendix 1), except for the first signal period (see Fig. 2), in which the time domain signals have not reached a stationary behavior (because of the pumping signal oscillating around a non-null mean rate).

We can simulate the same oscillatory responses signals to a harmonic pumping test either with a time domain solver, or with a frequency domain solver (if we put aside the first period where the time domain responses have not reach their stationary behavior, and thus are not reproducible in the frequency domain). However, in our case, the frequency-domain solver is more useful, as it provides directly the spatial distribution of the oscillatory responses amplitude and phase offset values that interest us for the interpretation part. There is no need of signal decomposition as in the time domain simulations. The frequency domain solver is also faster than the time domain one as we don't have to simulate different time steps (for the simulations on the presented synthetic case the frequency domain solver was approximately 120 times faster than the time domain one).

Therefore, for the interpretation of the responses oscillatory signals $h_{osc.}$, we have used the frequency domain simulations, and more specifically the responses' amplitude and phase offset values at the position of the different boreholes.

3.3.2. Analysis of the harmonic hydraulic responses

In order to interpret the spatial responses to harmonic pumping at a point in the karstic synthetic case, we have specifically studied the amplitude and phase offset values in these oscillatory signal responses. We have observed that, depending on the pumping borehole location, the responses could be highly variable. From Fig. 4 we first state that the oscillatory signal amplitude response is not proportional to the distance between the measurement points and the pumping point, as it would be expected in a purely homogeneous aquifer. In our synthetic case, the spatial responses were controlled by the degree of connectivity between the pumping point, the flow path induced by the pumping, and the

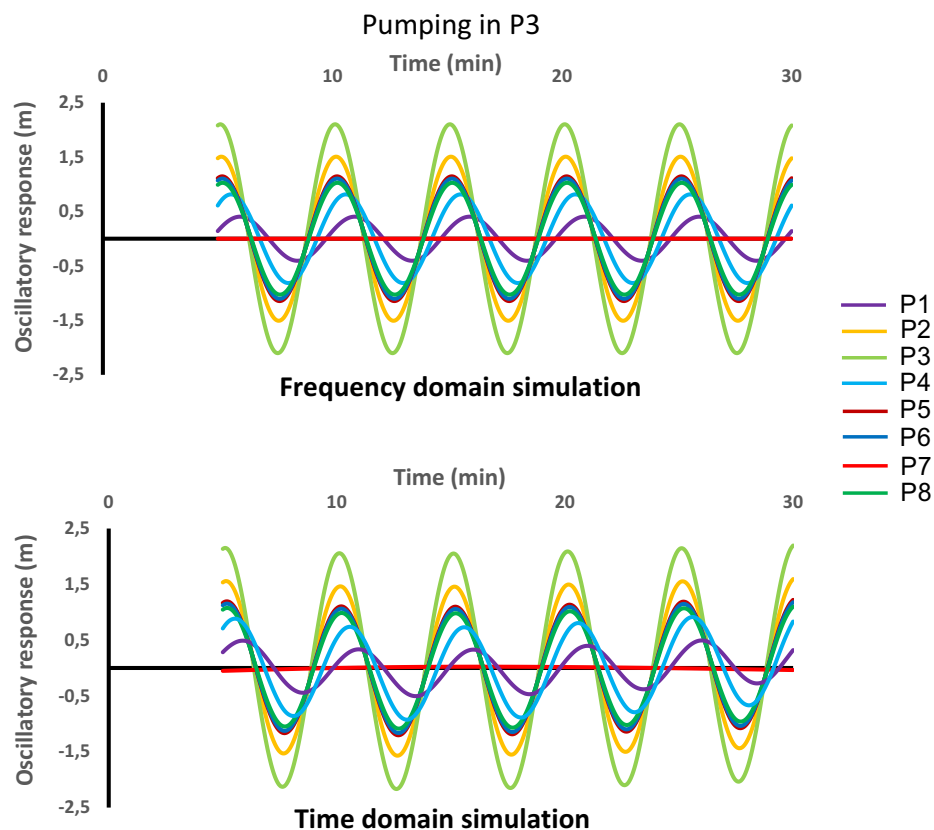


Fig. 3. Oscillatory signals responses in each borehole for a harmonic pumping in P3, for a frequency domain simulation and a time domain simulation (avoiding the first signal period). One sees that these signals are almost the same for the two simulations.

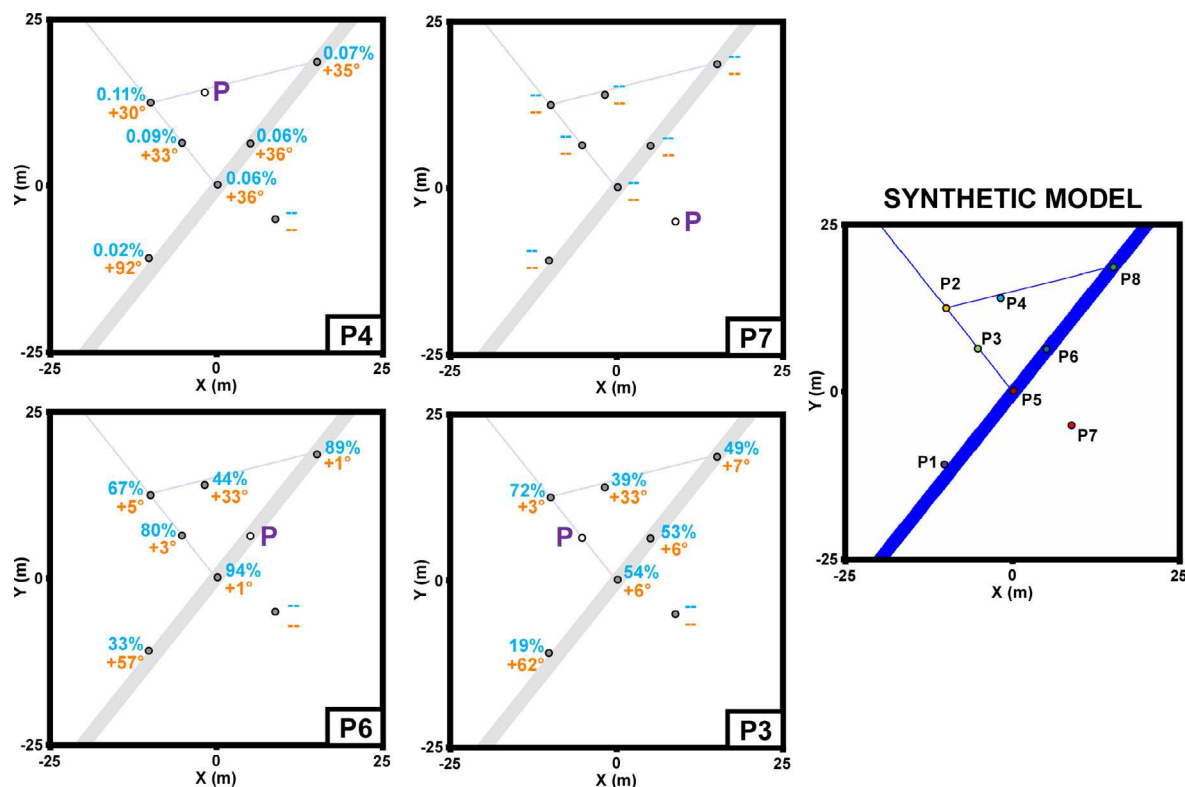


Fig. 4. Relative amplitude (% in blue) and relative phase offset ($^{\circ}$ in orange) values in the oscillatory responses in each borehole for different harmonic pumping locations (P4, P7, P6, P3). A dash represents an absence of oscillatory response (<1 mm). The pumping location is indicated by 'P' and its drawdown oscillatory signal is considered as a 100% amplitude signal with a 0° phase offset. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

measurement point. In particular cases, if the pumping borehole is positioned in the matrix, no oscillatory responses can be seen in the other boreholes (see case P7 in Fig. 4).

Four representative examples of response are presented in Fig. 4, with the complete table of responses amplitudes and phases offsets for these cases presented in Appendix 2. From the analysis of the 8 different pumping cases in the synthetic model, if we pump in or near to a conduit, three degree of connectivity can be interpreted by comparing the spatial oscillatory responses relatively to the pumping signal:

- *Amplitude responses associated with a low phase shift* (see P2, P3, P5, P8 responses for a pumping in P3 or P6 in Fig. 4): the pumping point is in a conduit and the measurement point has a *conduit connection* to it through a flow path in the karstic network,
- *Amplitude responses associated with a high phase shift*: the measurement point and the pumping point have a *dual connection*. A prevailing part of the flow path follows the karstic network and another part is in the matrix. The response phase offset value is proportional to the matrix flow path importance. Then, either the pumping point is in the matrix near to a karstic network (see P2, P3, P5, P8 responses for a pumping in P4 in Fig. 4), or the measurement point is in the matrix near to the network (see P1 and P4 responses for a pumping in P3 or P6), or both with a higher phase offset (see P1 response for a pumping in P4),
- *Negligible amplitude responses (almost no oscillatory signal)* (see P7 responses in Fig. 4): the prevailing part of the flow path between the measurement point and the pumping point is located in the matrix, it generates a *matrix connection* response.

Further information can be interpreted from the relative amplitude value of the measured signals for the responses with the same

phase offset. If the flow path follows a thin conduit in the network, the amplitude of the signal will decrease along the flow path away from the pumping point (see the responses in the thin conduit for the pumping case P6). The rate of decay in the responses amplitude becomes less important when the signal reaches more conductive conduits of the network (see the amplitude decay of the responses in the large conduit for the pumping case P3). The decay rate is inversely proportional to the conductivity of the conduits.

The choice of the harmonic pumping period is important for a good interpretation of the oscillatory responses. We show, in Fig. 5, the spatial responses' signal differences for harmonic pumping in P6 and P3 with a same signal amplitude ($Q_A = 2$ L/s) but two different periods (1–5 min).

The main effect of changing the period duration affects specifically the measurement boreholes that have a dual connection to the pumping borehole, while the other boreholes will have little consequent response differences. Decreasing the period from 5 min to 1 min will produce a relative decrease of the amplitude and a relative phase offset increase in the response signal of the boreholes with a dual conduit/matrix connection. Globally, this would tend to bring a 'dual connectivity' response closer to a 'matrix connectivity' response. This can be clearly seen in the oscillatory responses reconstructed from Eq. (10) (Fig. 6). For a 1 min pumping period, there is no significant signal changes for the boreholes in the conduits, but the oscillatory signals for the boreholes in the matrix tend to disappear compared to their responses to a 5 min pumping period signal (especially for P1 that becomes closer to the P7 'matrix connectivity' signal).

From the previous observation, we have studied the evolution of the relative amplitude and phase offset responses more specifically in the points with a dual connection to the network (P1, P4) for increasing periods (Table 2). It appears that the most important factor impacting the responses in these points is their distances to

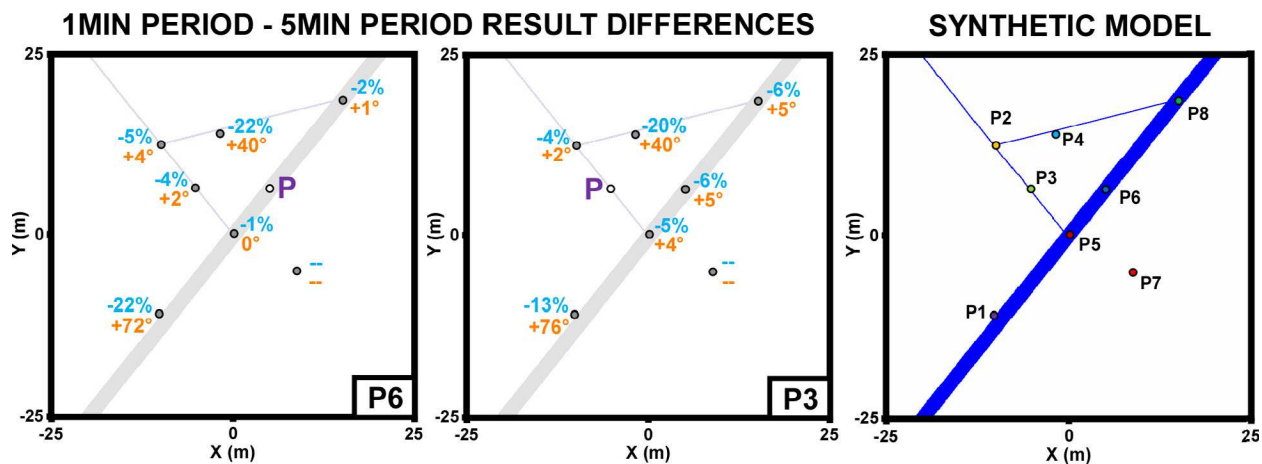


Fig. 5. Differences in relative amplitude (in blue, in %) and relative phase offset (in orange, in °) values in the oscillatory responses by decreasing from a 5 min period signal to a 1 min period signal for two different harmonic pumping locations (P6, P3). A dash represents an absence of oscillatory response (<1 mm). The pumping location is indicated by 'P'. The main signal differences appear for the boreholes located in the matrix, near to a conduit (P1, P4) (dual connection). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

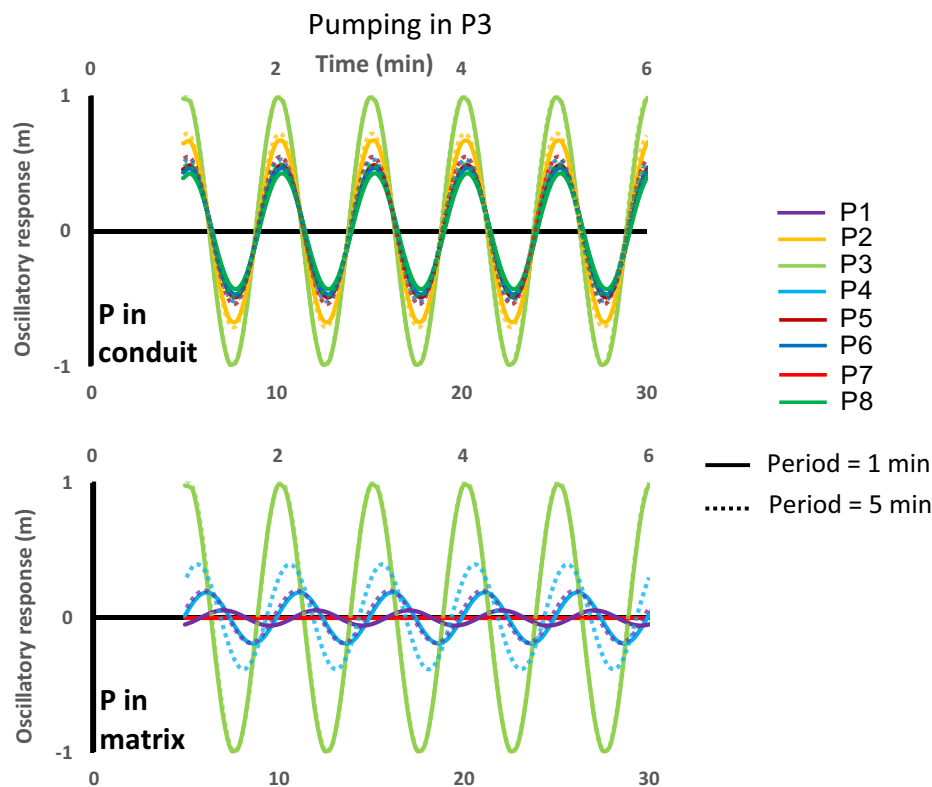


Fig. 6. Comparison of the oscillatory relative responses for a harmonic pumping in P3 for a 1 min period signal during 6 min (full line) and a 5 min period signal during 30 min (dotted line). The measurement boreholes have been separated regarding their location: in a conduit (P2, P5, P6, P8) or in the matrix (P1, P4, P7). The main signal differences appear for the boreholes located in the matrix, near to a conduit (P1, P4) (dual connection).

the network. Independently from the importance of the conduit in which the pumping is performed, the more a measurement point is distant from the network the more its phase offset response will evolve with a period change (see in Table 1 P1, distant from 1 m, compared to P4, distant from 50 cm). On the contrary, the relative amplitude of the response seems to be related to the productivity of the pumping location, but is therefore less interesting for delineating the position of the measurement point. Globally, for a good characterization of the conduits positioning, it is important to choose at least two different periods in order to compare the evo-

lution of the phase offset in the responses. These periods should be sufficiently high to avoid the risks related to a too low amplitude response (unreadable response) or a phase offset value exceeding one cycle (see Table 2) that may lead to incorrect interpretations.

To summarize, when pumping in the network, studying the amplitude of the responses provides some information about the conductivity of the conduits along the flow path, while studying the phase offset of the responses permits to characterize the degree of connectivity between the measurement points. A low period pumping (high frequency) tends to highlight more specifically

Table 2

Table of the relative amplitude and phase offset values in the oscillatory responses of P1 (1 m away from the network) and P4 (50 cm away from the network) to harmonic pumping in P3 and P6 and for increasing signal periods. In this table Amp. = Amplitude, P.O. = Phase Offset, TC = Thin Conduit, LC = Large Conduit, NTC = Near Thin Conduit and NLC = Near Large Conduit. Values in parentheses signify phase offsets greater than one cycle ($>360^\circ$).

Pump. point	Meas. point		Period						
			10 s	30 s	1 min	2 min	5 min	10 min	30 min
P3 (TC)	P1 (NLC)	Amp.	2%	3%	6%	11%	19%	26%	37%
		P.O.	(-14°)	($+128^\circ$)	$+138^\circ$	$+96^\circ$	$+62^\circ$	$+45^\circ$	$+27^\circ$
	P4 (NTC)	Amp.	4%	11%	19%	28%	39%	46%	55%
		P.O.	($+79^\circ$)	$+107^\circ$	$+73^\circ$	$+52^\circ$	$+33^\circ$	$+24^\circ$	$+16^\circ$
P6 (LC)	P1 (NLC)	Amp.	4%	6%	11%	19%	33%	44%	57%
		P.O.	(-20°)	($+124^\circ$)	$+129^\circ$	$+90^\circ$	$+57^\circ$	$+40^\circ$	$+23^\circ$
	P4 (NTC)	Amp.	5%	13%	22%	32%	44%	53%	63%
		P.O.	($+51^\circ$)	$+106^\circ$	$+73^\circ$	$+51^\circ$	$+33^\circ$	$+24^\circ$	$+15^\circ$

the boreholes directly connected to the pumping point by conduits (network flow propagation). A high period pumping (low frequency) tends to generate responses in boreholes with a conduit or a dual connectivity to the pumping point (part of network and part of diffusive flows propagation).

In order to locate only the boreholes directly connected through the conduit network, a high frequency pumping is necessary. But adding in combination responses from a lower frequency permits the identification of boreholes close to this network, which is useful information for imaging the karstic network arrangement.

3.3.3. Spatial analysis on simulation maps

The previous interpretation of the oscillatory signal responses to pumping signal with a 5 min period can be generalized in maps

of spatial distribution of the amplitude (Fig. 7) and phase offset (Fig. 8) responses.

The case of a pumping in P7 (in the matrix) in Fig. 7 shows a typical response map that would be expected from a homogeneous aquifer with an amplitude response decreasing with the distance to the pumping point forming a circle. In the case of a pumping in P4 (in the matrix), this response circle reaches the karstic network and the signal can propagate in the conduits with a subdued amplitude. If the pumping point is directly located in a conduit of the network (P6 and P3), the oscillatory signal propagates uniquely through the flows of the conduits, highlighting the karstic network. A linear decrease of the signal amplitude is visible in the thin conduits along the flow path, but in the most conductive conduits, the signal easily propagates at a site scale without attenuation.

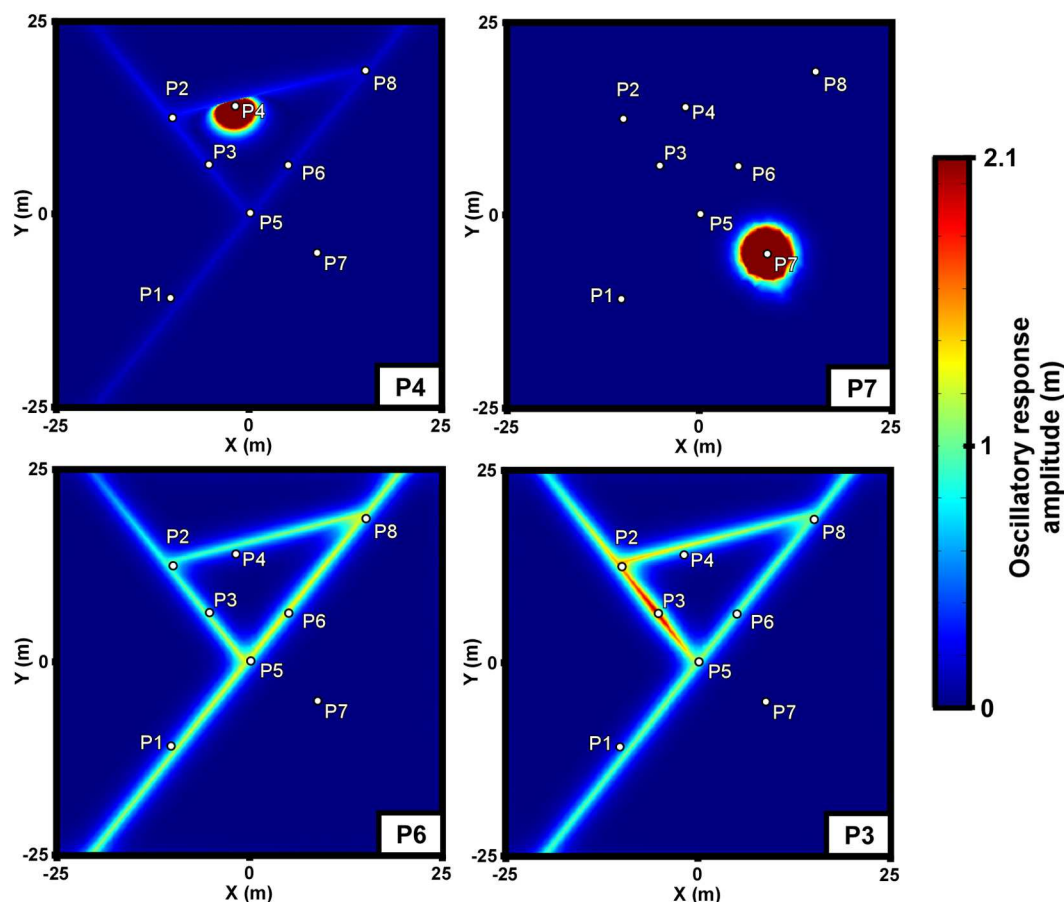


Fig. 7. Maps of distribution of the amplitude value in the responses to a harmonic pumping signal with a 5 min period at different locations: in the matrix near a conduit (P4), in the matrix (P7), in a large conduit (P6), in a thin conduit (P3).

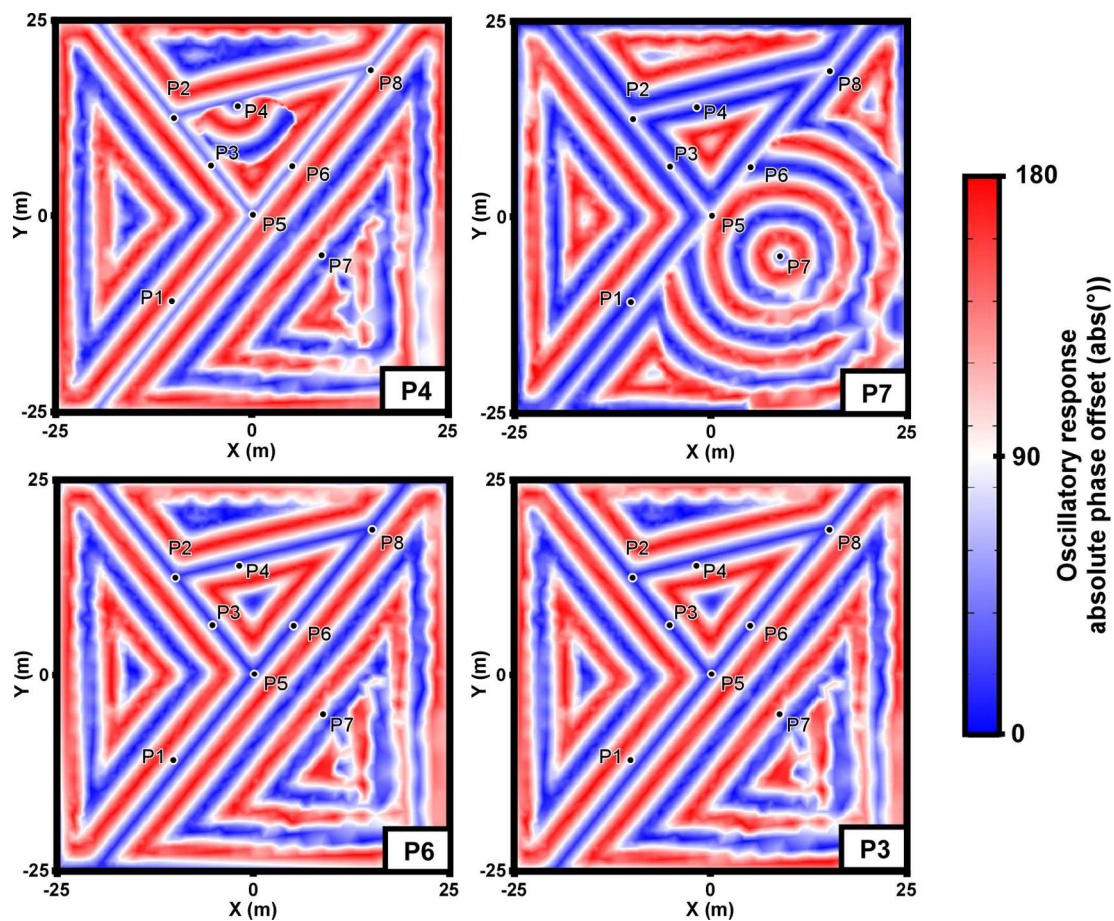


Fig. 8. Maps of distribution of the phase offset value in the responses to a harmonic pumping signal with a 5 min period at different locations: in the matrix near a conduit (P4), in the matrix (P7), in a large conduit (P6), in a thin conduit (P3).

Concerning the spatial phase offset distribution in Fig. 8, the case of a pumping in P7 in the matrix highlights what would be expected from a homogeneous aquifer. The phase offset value varies in function of the distance to the pumping point, forming a circle in the figure. As this circle reaches the network, its homogeneous behavior gets stopped, but at this position the signal has already lost its whole amplitude (see P7 in Fig. 7). The same 'homogeneous behavior' can also be seen locally in the case of a pumping in P4 (in the matrix), but the phase offset variation gets quickly 'controlled' by the conduit's disposition. The phase offset remains then roughly constant along the network geometry, with a value dependent on the pumping point's distance to the network. For a direct pumping in P7 in the network (P6, P3), independently of the conduits conductivities (as long as these conductivities are significantly higher than the matrix), there is no significant phase shifting along the conduits of the network. For all cases, once the signal has reached the karstic network, its phase shifting value will stay constant along the conduits, but it will increase rapidly in flow paths orthogonal to the conduit (conduit to matrix flows).

The effect of the signal period, already described in the previous part, can be generalized by using these same maps (Fig. 9).

Fig. 9 shows that a shorter period (high frequency) highlights more specifically the conduit network: the signal amplitude quickly decreases and the phase offset quickly increases when the signal enters the matrix. Thus, the propagation of the signal through diffusive flows in the matrix is almost absent. On the contrary a longer period signal can propagate at longer distances along matrix diffusive flows and with less phase shifting, which permits the boreholes in the matrix to respond. By comparing the cases of

P6 in the figure, for a 5 min period the diffusive behavior of this signal seems to 'blur' the karstic network. In fact, the responses associated to boreholes located near the conduits become undistinguishable from the responses of boreholes located directly in the conduits, and thus a precise localization of the conduit will become more difficult. Concretely this behavior is seen for P1 or P4: their responses are close to the ones of in the conduits network for a 5 min period, but their responses for a 1 min period clearly show that they are located in the matrix.

4. Example of harmonic pumping investigation

4.1. Site presentation

We now apply the lessons learned from our analysis of synthetic data trends to a true field case. We have performed an oscillatory pumping test on the well-studied Terrieu karstic field (approximately 2,500 m²) near to Montpellier, in Southern France. The oscillations in the pumping rate were controlled by a programmable electrical device based on a dimmer, linked to the pump. This site has been studied in two recent thesis (Jazayeri Noushabadi, 2009; Dausse, 2015) and several articles (Jourde et al., 2002; Jazayeri Noushabadi et al., 2011; Wang et al., 2016; Wang et al., 2017). The property values used in the synthetic model are inspired from conduits, fractures, matrix and regional property estimations issued from investigations presented in Dausse (2015). The Terrieu karstic field is part of the MEDCYSS observatory (Jourde et al., 2011), and an experimental site of the French National Karst Observatory (SNO Karst – www.sokarst.org).

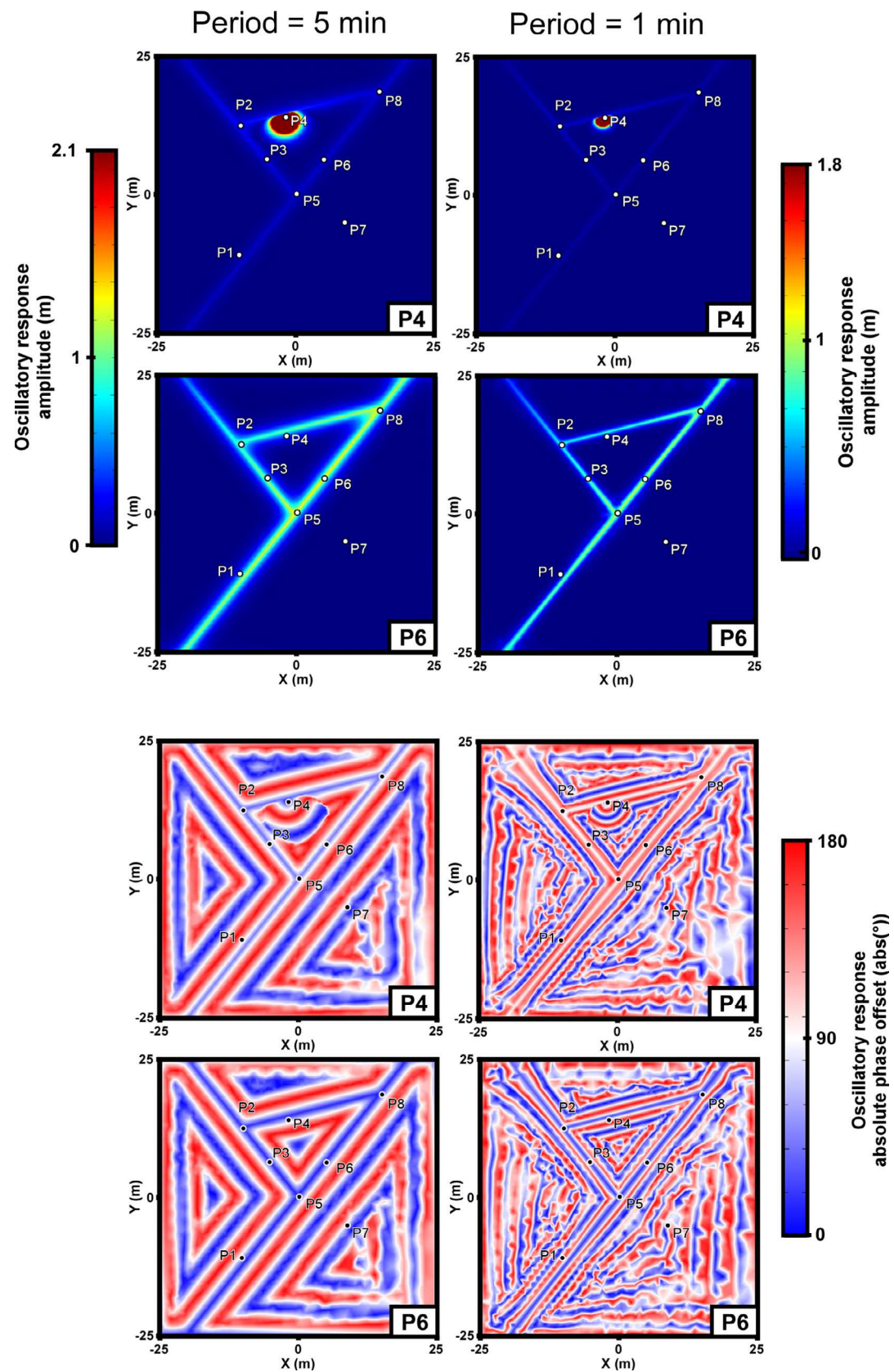


Fig. 9. Comparative maps of distribution of the amplitude and absolute phase offset values in the responses to a harmonic pumping at two different locations (in the matrix near a conduit (P4), in a large conduit (P6)) for a 5 min period (left) and 1 min period (right) signal.

This site is investigated through 22 boreholes (Fig. 10) and lies on a confined aquifer. Observed conduits through downhole videos (Jazayeri Noushabadi et al., 2011), located between 35 m and 45 m under the surface, have been generated at a sloped and fractured interface between marly Cretaceous and massive Jurassic limestones. Both units have very low permeability, which permits to consider the karstic aquifer to be confined. At the time of the field investigation karst features located at the sloped interface were fully saturated. These conduits have an aperture that can reach 20–50 cm. Previous field investigations (packer tests, temperature

and electrical conductivity logging) have permitted to highlight a preferential flow path (see blue line in Fig. 10) within this karstified interface crossing several boreholes (Jazayeri Noushabadi, 2009; Dausse, 2015). This preferential flow path could be the consequence of a network of conduits directly connecting these points.

A pumping investigation was conducted with an electronic automata device connected to the pump and generating oscillations in the pumping rate. A harmonic pumping test was led for 30 min on the borehole P15 with a period of 5 min with flow rates that varied between 3.2 m³/h and 7.4 m³/h. In other words, the signal from this pumping can be represented as a constant-rate pumping test of magnitude 5.3 m³/h (Q_m) convolved with an oscillatory (net zero) pumping signal with period 5 min and amplitude 2.1 m³/h (Q_A). The signal drawdown responses were measured in 12 other boreholes, additionally to a measurement in the pumping borehole itself.

4.2. Example of typical responses

The upper graph in Fig. 11 shows the drawdown measurement in the pumping borehole (P15), and example of measured responses in three other boreholes (P2, P9 and P10). These three boreholes are approximately at the same distance from the pumping point.

This graph shows that the field measured responses have the same behavior as the theoretical ones, with additional noise. The drawdown responses, if we pass over the first pumping period, can be approximately decomposed as an addition of a purely linear signal and a purely oscillatory signal of the form of Eq. (10) (shown in the lower graph of Fig. 11). The amplitude and phase offset analysis of the oscillatory signal of the three chosen measured responses examples also show that we have the same type of responses in this field case as seen in the theoretical case:

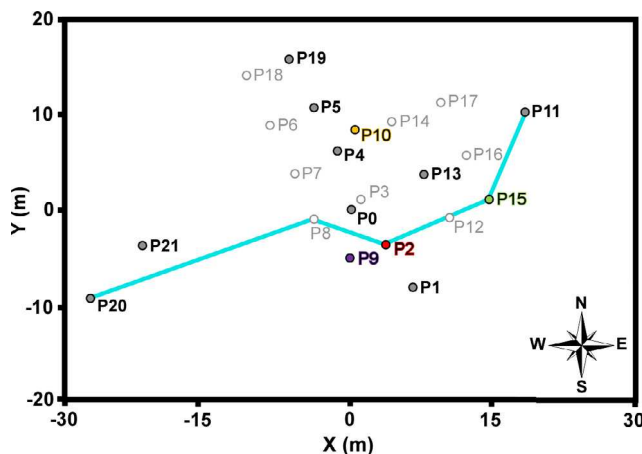


Fig. 10. Boreholes locations on the Terrieu site. The colors for P2, P9, P10 and P15 refer to the colors used to designate these boreholes in Fig. 11. The blue line indicates a conduit connectivity assessed from previous investigations (Dausse, 2015; Wang et al., 2016). The boreholes in light grey were not measured during the harmonic pumping test. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

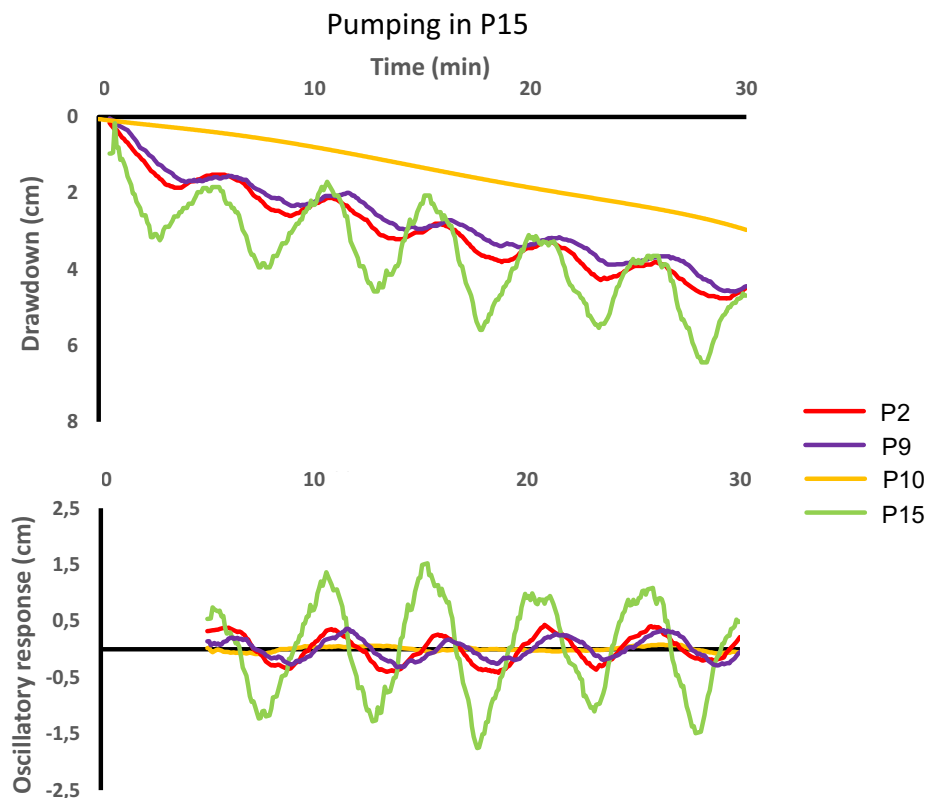


Fig. 11. Example of different type of responses registered during the 5 min period harmonic pumping test in P15 on the Terrieu site. The top graph shows the complete responses and the bottom graph shows the purely oscillatory responses after having subtracted the linear signal.

responses with measurable amplitude and a low phase offset (P2), responses with similar amplitude but a high phase offset (P9), and responses that contain no measurable oscillatory component (P10). Therefore, it seems acceptable to test the same interpretation that we made on the theoretical case for the field data.

4.3. Interpretation of the responses

We decomposed the entire set of measured drawdowns (following Eq. (11)) in order to keep only their oscillatory signal (the oscillatory responses in each borehole are presented in Fig. 12). We have then fit these oscillatory signals to function of the form of Eq. (10) with amplitude and phase offset as variable parameters.

As for the interpretation of the amplitude and phase offset values from the frequency domain modeling of the theoretical case in Part 3.3.2, we have produced in Fig. 13 a map of the measured amplitude and phase offset values from the oscillatory responses to pumping at a point within the karstic network.

The relative amplitude values of the responses vary between 0% (no oscillatory response) and 35%, and the relative phase offset values vary between $+30^\circ$ and $+140^\circ$ (relatively to the pumping signal). From downhole observations and productivity estimations on the pumping well P15 (Jazayeri Noushabadi et al., 2011), we know that this well is located in a conduit of the network.

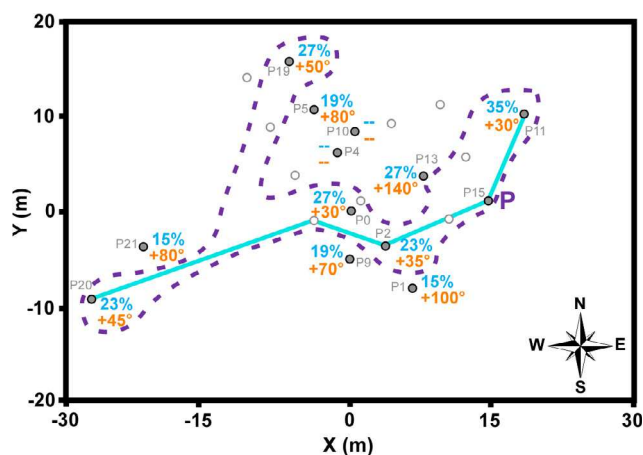


Fig. 13. Example of a possible conduits network (inside the zone delineated by violet dotted boundaries) interpreted from the boreholes connectivity by applying the same analysis than in the synthetic case. The captions represent the relative amplitude (in blue, in %) and relative phase offset (in orange, in $^\circ$) values in the oscillatory responses in each measured borehole. A dash represents an absence of oscillatory response (<1 mm). The pumping location is indicated by 'P'. The blue line indicates a conduit connectivity known from previous investigations (Jazayeri Noushabadi, 2009; Dausse, 2015). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

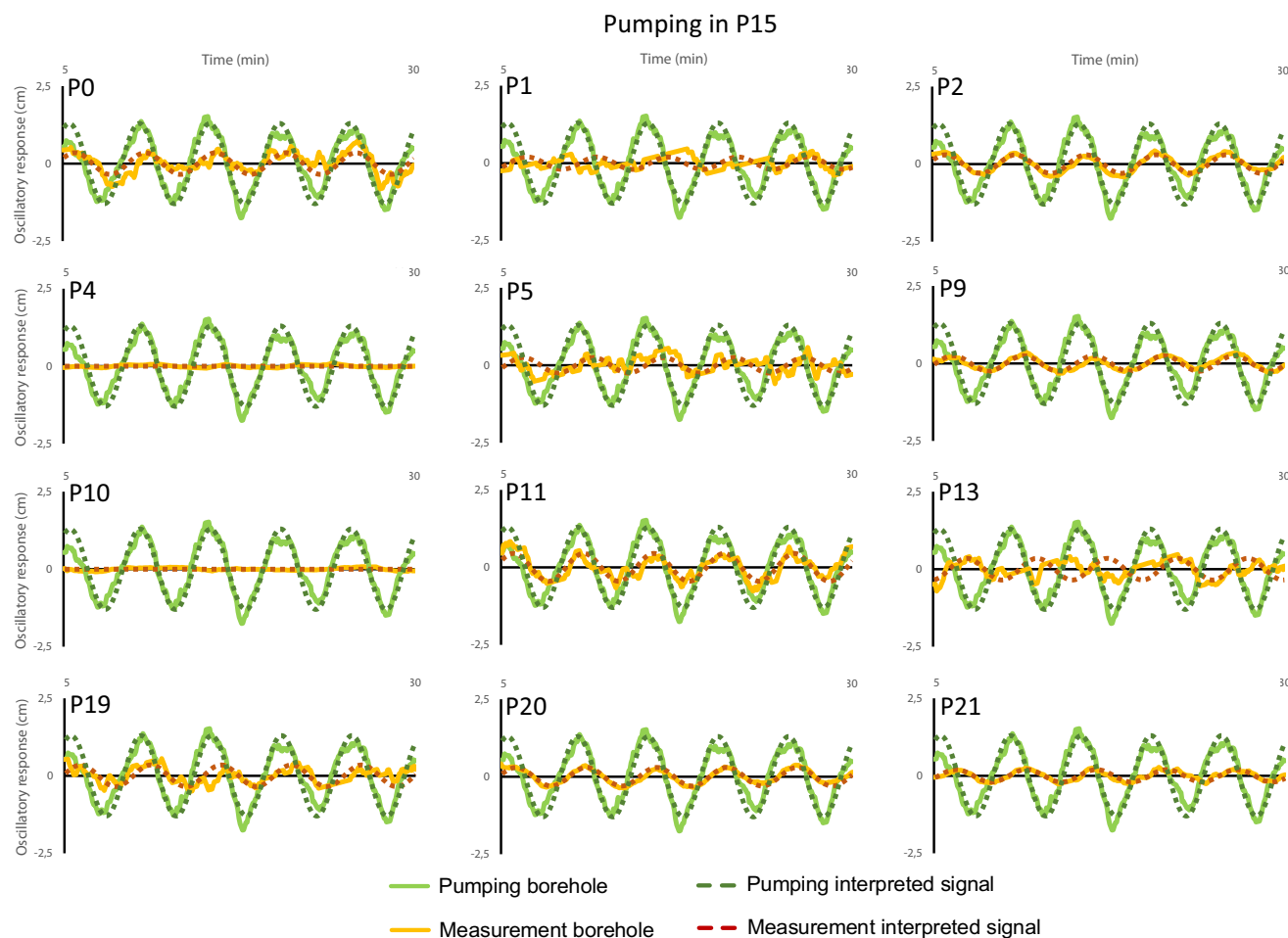


Fig. 12. Registered oscillatory responses for each measurement borehole compared to the $T = 5$ min period pumping borehole signal (full lines) and the interpreted signals for an equation form of Eq. (10) with variables amplitude and phase offset values (dotted lines).

By analyzing the phase offset value, we note that the lower phase offset values are around $+30^\circ/+50^\circ$, and that several points have this same phase shift. This would indicate the presence of a network of conduits directly connecting these boreholes ('conduit' connectivity: P0, P2, P11, P19, P20) to the pumping point P15 (a possible network of conduits is proposed, within the zone delineated by the violet dotted boundaries, in Fig. 13). Then, the other boreholes ('dual' connectivity: P1, P5, P9, P21; 'matrix' connectivity: P4, P10) would have coherent responses toward this conduit network, similar to those seen in the study of the synthetic case ('dual' or 'matrix' connectivity). The more a borehole is distant to this possible conduit network, the more its response signal relatively to the pumping signal has a low amplitude and a high phase shift (for the more distant boreholes P4 and P10, the oscillatory signal disappear which indicate a 'matrix' connectivity). Only P13 would present an incoherent response (with both a high amplitude response and a high phase shift), but its true signal was too noisy to permit a good amplitude and phase offset parameters fitting (see P13 in Fig. 12). This incoherent result can be attributed to a bad signal measurement. Except this point, the other 11 measured responses on the field are coherent with the behaviors interpreted in the theoretical study case.

The conduit connectivity highlighted from the method presented in this article is coherent with the direct connectivity already highlighted from previous investigations (blue line in Fig. 13), which tend to strengthen the validity of this interpretation. Furthermore, P19, which was not found as connected in the previous investigation, appears to be directly connected in this work. However, this connectivity information would require a supplementary validation in the next field campaigns.

We have finally made quantitative estimation on the property values in the field, relatively to each response, in order to compare these estimations to the previously found degree of connectivity. These estimations were made by finding, for each response, an equivalent homogeneous model (K_{eq} and $S_{s,eq}$) reproducing the response amplitude and phase offset. The optimization of K_{eq} and $S_{s,eq}$ in each case was performed through a classical least square minimization criterion inversion. It appears that P0, P2, P11, P20 (interpreted as connected through conduits), and P13 require high K_{eq} values (10^{-2} – 10 m/s) and low $S_{s,eq}$ (10^{-8} – 10^{-7} m $^{-1}$). P5, P9, P21 (interpreted as dual connectivity responses), and P19 (interpreted as directly connected) require high K_{eq} values (10^{-3} – 10^{-2} m/s) and high $S_{s,eq}$ (10^{-4} – 10^{-3} m $^{-1}$). Finally, P1 (interpreted as dual connectivity response), and P4 and P10 (interpreted as a prevailing matrix connection) require low K_{eq} values (10^{-6} m/s) and high $S_{s,eq}$ (10^{-4} m $^{-1}$). Globally three categories of estimated equivalent properties can also be highlighted and would correspond to the previously described degree of connectivity: conduit connectivity (high K_{eq} , low $S_{s,eq}$), dual connectivity (high K_{eq} , high $S_{s,eq}$), and matrix connectivity (low K_{eq} , high $S_{s,eq}$). In this case only P19 appears to be rather connected in a 'dual' manner and P1 rather as a matrix point, which nuance the interpretation of some responses at the limit of the defined categories.

From these results, we believe that the qualitative method of interpretation of harmonic signal responses developed in this work can facilitate interpreting the degree of connectivity of karstic field responses to an oscillatory signal, especially in order to characterize and localize the conduits network through the boreholes connectivity. This interpretation can be led manually as shown in this field example. However the next step would be to use this interpretation in an inverse modeling approach (associated to a frequency domain modeling and by considering the amplitude and phase offset values of the observed and simulated responses). This approach becomes, in fact, very useful when the amount and complexity of responses to interpret is important. The Cellular

Automata-based Deterministic Inversion developed in Fischer et al. (2017) would provide in this case interesting results, as it would permit to generate both adequate conduits network and property distributions (conductivity and specific storage) for the reproduction of the responses, and thus go further in the quantitative analysis of the harmonic data in karstic fields.

5. Discussion and conclusion

We have studied the responses to harmonic pumping in a theoretical synthetic study, by applying a time domain solver and a frequency domain solver. We have firstly demonstrated that the harmonic result from the frequency domain simulation was very similar to the time domain oscillatory part, and thus, as the frequency domain solver is much faster, it is more useful for the simulation of periodic responses.

From an analysis of the amplitude and the phase offset of the response relative to the pumping signal and its positioning in the model, we have proposed a global method for qualitatively interpreting a degree of hydraulic connectivity between each borehole. The amplitude and phase offset values permit to distinguish either a conduit connectivity between boreholes (flow path in the conduit network), or a dual connectivity (flow path partly in the matrix for boreholes near to a conduit but not directly in it), or a matrix connectivity (majority of flow associated with flow in the matrix). By modifying the period of the pumping signal, we can dissociate more precisely the conduit connectivity and the connectivity and obtain some information about the distance of a measurement point relatively to the karstic network. A high frequency signal will more specifically highlight the conduit flows, while a low frequency signal will give more importance to the matrix diffusive flows.

In previous works Renner and Messar (2006) and Guiltinan and Becker (2015) used analytical solutions to show that, in a fractured aquifer, increasing the pumping period decreased the estimates of effective hydraulic diffusivity, due to the increase of the values of estimated storativity (while the transmissivity estimations remained almost unchanged), which was associated to a mobilization of the surrounding fracture void spaces. In our karstic model this observation would be related to the behavior of a well in the matrix near to a conduit, where the signal can better propagate within the high storage that represents the matrix when the pumping period is increased. As it has been noticed by Rabinovich et al. (2015) with the use of a heterogeneity model for a porous aquifer, flows (by extension the harmonic signal) will preferentially propagate in the most conductive media, especially at lower pumping period, which is also what we observe in our karstic model between conduits and matrix flows, when pumping at different periods. The results we present in this article between amplitude ratio and phase shifting and the degree of connectivity of boreholes show equivalent relations (in a more accentuated way) than those presented by Guiltinan and Becker (2015) in a fractured aquifer. The harmonic signal will arrive with an important attenuation and phase shift between areas badly connected and, on the contrary, almost unchanged when fractures (or in our case a conduit) connect two boreholes.

In order to test the validity of our signal analysis method, we have confronted the same interpretation for a set of responses from a true karstic field to a harmonic pumping. The interpretation permitted to delineate the main flow paths easily and quickly by satisfying all measurement. This method could be generalized for a manual interpretation of a set of field responses. The benefits brought here by the harmonic tests compared to the other connectivity investigations done on the same site (packers, temperature and electrical conductivity) and to constant-rate pumping can be

found in the easier extraction of the signal in the responses (with filtering techniques), even when responses are noised, and from the possibility to simulate the responses with a modeling in a frequency domain, much quicker than the time domain. Furthermore, the possibility of simulating the amplitude and phase offset values in the responses using a frequency domain modeling permits a more affordable application of this interpretation method in an inversion approach.

We have presented in the field part only the results from one pumping test associated to one period, selected from a campaign of pumping tests in which several different boreholes and different period values were tested. In order to interpret simultaneously all responses from all harmonic pumping tests, we have to use an inverse modeling. This technique is already widely used to characterize and quantify the heterogeneity in fractured and karstic fields, by interpreting the responses to constant-rate pumping hydraulic tomography (Hao et al., 2008; Illman et al., 2009; Castagna et al., 2011; Illman, 2014; Zha et al., 2016; Wang et al., 2017). In a future work, we plan to focus more specifically our discussion on the quantitative analysis of the harmonic pumping responses in a karstic field, by associating them to a conduit network hydraulic imagery, based on an inverse modeling approach and its ability to reproduce the complete set of responses with a given distribution of properties.

Conflict of interest

None.

Acknowledgments

We thank the region Normandy for financially supporting the PhD of Pierre Fischer. We would like to thank the four anonymous reviewers for their relevant comments and propositions which permitted to significantly improve the quality of this article.

Appendix

Appendix 1

Oscillatory response amplitude and phase offset values at the position of the different boreholes for a harmonic pumping in P3. We have estimated these values from a frequency domain simulation and from a time domain simulation (avoiding the first signal period). One sees that these values are almost the same for the two simulations.

	Frequency domain simulation		Time domain simulation	
	Amplitude (m)	Phase offset (°)	Amplitude (m)	Phase offset (°)
P1	0.4	70	0.4	72
P2	1.51	11	1.51	12
P3	2.1	8	2.1	9
P4	0.81	41	0.81	42
P5	1.15	14	1.15	15
P6	1.1	14	1.11	15
P7	0	–	–	–
P8	1.03	15	1.03	15

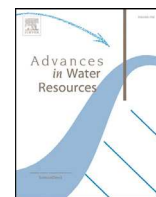
Appendix 2

Table of the real and relative amplitude and phase offset values in the oscillatory responses to different harmonic pumping locations and for two different signal periods. For each pumping case, the boreholes are sort by increasing distance to the pumping point. A dash represents an absence of oscillatory response (<1 mm).

Pumping Point	Borehole	Distance (m)	Period = 5 min				Period = 1 min			
			Amplitude (m)	Phase offset (°)	Relat. amplitude (%)	Relat. phase offset (°)	Amplitude (m)	Phase offset (°)	Relat. amplitude (%)	Relat. phase offset (°)
P3	P3	0	2.1	8	100	0	1.81	11	100	0
	P2	8	1.51	11	72	3	1.23	16	68	5
	P5	8	1.15	14	54	6	0.89	21	49	10
	P4	8	0.81	41	39	33	0.34	84	19	73
	P6	10	1.1	14	53	6	0.85	22	47	11
	P1	18	0.4	70	19	62	0.1	149	6	138
	P7	18	–	–	–	–	–	–	–	–
	P8	24	1.03	15	49	7	0.78	23	43	12
P4	P4	0	236.31	9	100	0	191.12	20	100	0
	P2	8	0.26	39	0.11	30	0.11	80	0.06	60
	P3	8	0.2	41	0.09	33	0.09	84	0.05	64
	P6	10	0.15	45	0.06	36	0.06	90	0.03	69
	P5	14	0.15	45	0.06	36	0.06	90	0.03	70
	P8	18	0.16	44	0.07	35	0.16	88	0.08	68
	P7	22	–	–	–	–	–	–	–	–
	P1	26	0.05	101	0.02	92	0.06	–142	0.03	–162
P6	P6	0	1.38	12	100	0	1.12	17	100	0
	P5	8	1.3	12	94	1	1.04	18	93	1
	P3	10	1.1	14	80	3	0.85	22	76	5
	P4	10	0.61	45	44	33	0.25	90	22	73
	P7	12	–	–	–	–	–	–	–	–
	P8	16	1.23	13	89	1	0.97	19	87	2
	P2	16	0.93	17	67	5	0.69	25	62	9
	P1	23	0.46	68	33	57	0.12	146	11	129
P7	P7	0	174.03	22	100	0	110.86	40	100	0
	P5	10	–	–	–	–	–	–	–	–
	P6	12	–	–	–	–	–	–	–	–
	P3	18	–	–	–	–	–	–	–	–
	P1	20	–	–	–	–	–	–	–	–
	P4	22	–	–	–	–	–	–	–	–
	P8	24	–	–	–	–	–	–	–	–
	P2	26	–	–	–	–	–	–	–	–

References

- Bakhos, T., Cardiff, M., Barrash, W., Kitanidis, P.K., 2014. Data processing for oscillatory pumping tests. *J. Hydrol.* 511, 310–319.
- Black, J.H., Kipp Jr., K.L., 1981. Determination of hydrogeological parameters using sinusoidal pressure tests: a theoretical appraisal. *Water Resour. Res.* 17 (3), 686–692.
- Butler, J.J., 2005. Hydrogeological methods for estimation of spatial variations in hydraulic conductivity. In: Rubin, Y., Hubbard, S.S. (Eds.), *Hydrogeophysics*. Water Science and Technology Library. Springer, Dordrecht.
- Cardiff, M., Barrash, W., Kitanidis, P.K., Malama, B., Revil, A., Straface, S., Rizzo, E., 2009. A potential-based inversion of unconfined steady-state hydraulic tomography. *Ground Water* 47, 259–270.
- Cardiff, M., Bakhos, T., Kitanidis, P.K., Barrash, W., 2013. Aquifer heterogeneity characterization with oscillatory pumping: Sensitivity analysis and imaging potential. *Water Resour. Res.* 49, 5395–5410.
- Cardiff, M., Barrash, W., 2015. Analytical and semi-analytical tools for the design of oscillatory pumping tests. *Groundwater* 53 (6), 896–907.
- Castagna, M., Becker, M.W., Bellin, A., 2011. Joint estimation of transmissivity and storativity in a bedrock fracture. *Water Resour. Res.* 47, W09504. <https://doi.org/10.1029/2010WR009262>.
- Dagan, G., Rabinovich, A., 2014. Oscillatory pumping wells in phreatic, compressible, and homogeneous aquifers. *Water Resour. Res.* 50 (No.8), 7058–7066.
- Dausse, A., 2015. Facteurs d'échelle dans la hiérarchisation des écoulements au sein d'un aquifère karstique: Analyse multi-échelles des propriétés hydrodynamiques et de transport de l'aquifère du Lez. PhD Thesis, Université de Montpellier. French.
- Fischer, P., Jardani, A., Lecoq, N., 2017. A cellular automata-based deterministic inversion algorithm for the characterization of linear structural heterogeneities. *Water Resour. Res.* 53, 2016–2034.
- Guiltinan, E., Becker, M.W., 2015. Measuring well hydraulic connectivity in fractured bedrock using periodic slug tests. *J. Hydrol.* 521, 100–107.
- Hao, Y., Yeh, T.C.J., Xiang, J., Illman, W.A., Ando, K., Hsu, K.C., Lee, C.H., 2008. Hydraulic tomography for detecting fracture zone connectivity. *Ground Water* 46, 183–192.
- Illman, W.A., 2014. Hydraulic tomography offers improved imaging of heterogeneity in fractured rocks. *Groundwater* 52, 659–684.
- Illman, W.A., Liu, X., Takeuchi, S., Yeh, T.C.J., Ando, K., Saegusa, H., 2009. Hydraulic tomography in fractured granite: mizunami underground research site, Japan. *Water Resour. Res.* 45, W01406. <https://doi.org/10.1029/2007WR006715>.
- Jazayeri Noushabadi, M.R., 2009. Characterization of relationship between fracture network and flow-path network in fractured and karstic reservoirs: Numerical modeling and field investigation (Lez aquifer, Southern France). PhD Thesis, Université de Montpellier. English.
- Jazayeri Noushabadi, M.R., Jourde, H., Massonnat, G., 2011. Influence of the observation scale on permeability estimation at local and regional scales through well tests in a fractured and karstic aquifer (Lez aquifer, Southern France). *J. Hydrol.* 403, 321–336.
- Jourde, H., Cornaton, F., Pistre, S., Bidaux, P., 2002. Flow behavior in a dual fracture network. *J. Hydrol.* 266, 99–119.
- Jourde, H., Batiot-Guilhe, C., Bailly-Comte, V., Bicalho, C., Blanc, M., Borrell, V., Bouvier, C., Boyer, J.F., Brunet, P., Cousteau, M., Dieulin, C., Gayard, E., Guinot, V., Hernandez, F., Kong-A-Siou, L., Johannet, A., Leonardi, V., Mazzilli, N., Marchand, P., Patris, N., Pistre, S., Seidel, J.L., Taupin, J.D., Van-Exter, S., 2011. The MEDYCYSS observatory, a multi scale observatory of flood dynamics and hydrodynamics in karst (Mediterranean border Southern France). In: Lambrakis, N., Stournaras, G., Katsanou, K. (Eds.), *Advances in the research of aquatic environment*. Environmental Earth Sciences. Springer, Berlin, Heidelberg.
- Lavenue, M., de Marsily, G., 2001. Three-dimensional interference test interpretation in a fractured aquifer using the pilot point inverse method. *Water Resour. Res.* 37 (11), 2659–2675.
- Liedl, R., Sauter, M., Huckinghaus, D., Clemens, T., Teutsch, G., 2003. Simulation of the development of karst aquifers using a coupled continuum pipe flow model. *Water Resour. Res.* 39, 1057.
- Maineult, A., Strobach, E., Renner, J., 2008. Self-potential signals induced by periodic pumping tests. *J. Geophys. Res.* 113, B01203. <https://doi.org/10.1029/2007JB005193>.
- Rabinovich, A., Barrash, W., Cardiff, M., Hochstetler, D.L., Bakhos, T., Dagan, G., Kitanidis, P.K., 2015. Frequency dependent hydraulic properties estimated from oscillatory pumping tests in an unconfined aquifer. *J. Hydrol.* 531, 2–16.
- Rasmussen, T.C., Haborak, K.G., Young, M.H., 2003. Estimating aquifer hydraulic properties using sinusoidal pumping at the Savannah River site, South California, USA. *Hydrogeol. J.* 11, 466–482.
- Renard, P., Allard, D., 2013. Connectivity metrics for subsurface flow and transport. *Adv. Water Resour.* 51, 168–196.
- Renner, J., Messar, M., 2006. Periodic pumping tests. *Geophys. J. Int.* 167, 479–493.
- Saller, A.P., Ronayne, M.J., Long, A.J., 2013. Comparison of a karst groundwater model with and without discrete conduit flow. *Hydrogeol. J.* 21, 1555–1566.
- Schuite, J., Longuevergne, L., Bour, O., Guihéneuf, N., Becker, M.W., Cole, M., Burbey, T.J., Lavenant, N., Boudin, F., 2017. Combining periodic hydraulic tests and surface tilt measurements to explore in situ fracture hydromechanics. *J. Geophys. Res.* Solid Earth 122, 6046–6066.
- Soueid Ahmed, A., Jardani, A., Revil, A., Dupont, J.P., 2016. Joint inversion of hydraulic head and self-potential data associated with harmonic pumping tests. *Water Resour. Res.* 52 (9), 6769–6791.
- Sun, A.Y., Lu, J., Hovorka, S., 2015. A harmonic pulse testing method for leakage detection in deep subsurface storage formations. *Water Resour. Res.* 51, 4263–4281.
- Sun, A.Y., Lu, J., Freifeld, B.M., Hovorka, S.D., Islam, A., 2016. Using pulse testing for leakage detection in carbon storage reservoirs: a field demonstration. *Int. J. Greenhouse Gas Control* 46, 215–227.
- Teutsch, G., 1993. An extended double-porosity concept as a practical modeling approach for a karstified terrain. *Hydrogeol. Process. Karst Terranes* 207, 281–292.
- Tyukhova, A.R., Willmann, M., 2016. Connectivity metrics based on the path of smallest resistance. *Adv. Water Resour.* 88, 14–20.
- Wang, X., Jardani, A., Jourde, H., Lonergan, L., Cosgrove, J., Gosselin, O., Massonnat, G., 2016. Characterisation of the transmissivity field of a fractured and karstic aquifer, Southern France. *Adv. Water Resour.* 87, 106–121.
- Wang, X., Jardani, A., Jourde, H., 2017. A hybrid inverse method for hydraulic tomography in fractured and karstic media. *J. Hydrol.* 551, 29–46.
- White, W.B., 2002. Karst hydrology: recent developments and open questions. *Eng. Geol.* 65, 85–105.
- Zha, Y., Yeh, T.C.J., Illman, W.A., Tanaka, T., Bruines, P., Onoe, H., Saegusa, H., Mao, D., Takeuchi, S., Wen, J.-C., 2016. An application of hydraulic tomography to a large-scale fractured granite site, Mizunami, Japan. *Groundwater* 54, 793–804.
- Zhou, Y., Lim, D., Cupola, F., Cardiff, M., 2016. Aquifer imaging with pressure waves - Evaluation of low-impact characterization through sandbox experiments. *Water Resour. Res.* 52, 2141–2156.



Harmonic pumping tomography applied to image the hydraulic properties and interpret the connectivity of a karstic and fractured aquifer (Lez aquifer, France)

P. Fischer^{a,*}, A. Jardani^a, H. Jourde^b, M. Cardiff^c, X. Wang^b, S. Chedeville^b, N. Lecoq^a

^a Normandie Univ, UNIROUEN, UNICAEN, CNRS, M2C, 76000 Rouen, France

^b Laboratoire Hydrosiences, Université de Montpellier, CNRS, 34000 Montpellier, France

^c Department of Geoscience, University of Wisconsin-Madison, Madison, WI, USA

ARTICLE INFO

Keywords:

Hydraulic tomography
Karst
Oscillatory signal
Connectivity
Modeling
Conduit network

ABSTRACT

In this work, we present a novel method to interpret, at a field scale, the preferential flows generated by harmonic pumping tests, in which the pumped flowrate varies according to a sinusoidal function with a given period. The experimental protocol relies on the application of harmonic pumping tests in a karstic field located near to Montpellier (Southern France) at 4 different boreholes, each time with a shorter and a longer period, and the analysis of the hydraulic responses recorded at the 13 observation wells. A qualitative analysis of the oscillatory component in the hydraulic responses, in term of amplitude decay and phase lag, permitted to propose a preliminary model of degree of connectivity between the boreholes, through the network of conduits. Then, a quantitative interpretation of the harmonic responses was applied to image the spatial heterogeneity of the hydraulic properties (hydraulic conductivity and storage coefficient) by using a deterministic inverse algorithm called CADI. This algorithm is based on an equivalent porous medium concept and parameterized by a Cellular Automata approach in order to provide a realistic reconstruction of the karstic network. This algorithm is linked to the groundwater flow equation, reformulated in frequency domain, to simulate the amplitudes and phase shifts responses to the harmonic pumping tests. The inverse process was successfully applied on the dataset collected with both periods, in a separate and joint way. The results obtained allowed for a discussion on the efficiency of the harmonic pumping tomography for the characterization of the karstic structures.

1. Introduction

The protection and the management of the water resources involve the identification of the preferential flow paths in the ground. Therefore, one needs to characterize the spatial distribution of the hydraulic properties in the field subsurface. A common way to assess the hydraulic properties of a field, such as conductivity and specific storage, is the analysis of the drawdown responses to a pumping test from which local or average properties can be inferred from analytical equations that relate the hydraulic response to the hydraulic properties (Butler, 2005).

However, in the case of karstic aquifers, the assessment of the hydraulic properties is challenging (White, 2002; Hartmann et al., 2014) as the hydraulic properties in this type of aquifer can vary by several orders of magnitude within a short distance (Wang et al., 2016). This makes the characterization of the karstic fields very complex. To face this difficulty, it is then necessary to interpret the responses of the field by taking into account the positioning of the conduits network, which

determines the preferential flow paths (Kovacs, 2003; Ghasemizadeh et al., 2012; Saller et al., 2013).

The hydraulic tomography concept has been widely applied to map the spatial variability of hydraulic properties, in both type of aquifers (porous and fractured), by performing a joint interpretation of hydraulic data recorded simultaneously at several wells, as responses to extraction/injection of water (Yeh and Liu, 2000; Bohling et al., 2002; Zhu and Yeh, 2005; Yeh and Lee, 2007; Cardiff et al., 2009a; Castagna et al., 2011; Berg and Illman, 2013; Cardiff et al., 2013a; Zha et al., 2015; Zha et al., 2016; Zha et al., 2017). This approach relies on a numerical technique (such as finite difference, finite element and finite volume) to solve the groundwater flow equation, and the inverse process to reconstruct the heterogeneity of the hydraulic conductivities and the storage properties by fitting the piezometric responses. The inversion process usually provides a non-unique solution which can produce an ambiguous interpretation of the hydraulic data. To overcome this issue, a prior information on the distribution of the properties can be used to constrain and guide the inversion to a more realistic solution (Carrera and Neuman, 1986). In the case of aquifers with a low heterogeneity, the geostatistical constraints remain the most simple and efficient way to find accurate solutions (Hoeksema and Kitanidis, 1984; Kitanidis, 1995; Fischer et al., 2017a). In the context of fractured and

* Corresponding author.

E-mail address: pierre.fischer1@univ-rouen.fr (P. Fischer).

karstic aquifers, the definition of the a priori model, or even the strategy for the numerical groundwater flow simulation (which can be performed by using various techniques such as equivalent porous media or discrete fractures networks), remain the subject of open debates among hydrogeologists. In fact, for a successful interpretation of hydraulic responses of karstic aquifers, the ‘classical’ geostatistical inversion method would require a dense network of measurement and a significant resolution of model parameterization because of the high contrasts existing in the distribution of the spatial properties. Recently, several inversion methods have been developed for characterizing karst networks. One way is to directly incorporate a discrete geometry within a background model using a discrete-continuum forward model (Teutsch, 1993; Liedl et al., 2003; de Rooij et al., 2013). In this case, the parameterization of the inverse problems usually relies on a stochastic generation of discrete fracture networks that are conditioned to statistical (Li et al., 2014; Le Coz et al., 2017), mechanical (Jaquet et al., 2004; Bonneau et al., 2013), or structural data (Pardo-Iguzquiza et al., 2012; Collon et al., 2017). Another way is based on a deterministic optimization of the geometry of discrete networks (Borghi et al., 2016; Fischer et al., 2018a).

Previous studies have shown that equivalent porous media models are able to reproduce the hydraulic flows in karstic aquifers at a kilometer scale (Larocque et al., 1999; Abusaada and Sauter, 2013) or a decametric scale (Wang et al., 2016). However, if the scale of investigation is too small, this type of model can become unreliable for the characterization of the properties of fractured rocks, extremely contrasted and structured at a small scale (Illman, 2014). Although the classical geostatistical inverse approaches were originally proposed for inversion of hydraulic fields, they can be made adaptive to discrete geometries with special treatments to the prior model (e.g. the total variation prior model, Lee and Kitanidis, 2013), or using an iterative procedure (e.g. the sequential successive linear estimator, Ni and Yeh, 2008; Hao et al., 2008; Illman et al., 2009; Sharmeen et al., 2012). Other methods for inversion of complex discrete structures involve introducing constraints of a priori knowledge to the inverse model using a guided image (Hale 2009; Soueid Ahmed et al., 2015), a training image (Lochbühler et al., 2015), a probability perturbation (Caers and Hoffman, 2006), a transition probability distribution (Wang et al., 2017), a multi-scale resolution (Ackerer and Delay, 2010), a level-set method (Lu and Robinson, 2006; Cardiff and Kitanidis, 2009b), or based on cellular automata (Fischer et al., 2017b).

Apart from these challenges in modeling techniques, a further difficulty in karst aquifer characterization arises from characteristics of hydraulic tests. Due to the integration nature of pressure diffusion, the steady state responses of long-term constant-rate pumping tests in a karst aquifer represent the combined effect of the multiple media (conduits, fissures, and matrix) rather than specific components. Although the interpretation of transient responses may provide additional information about the relative importance of each aquifer components, the inclusion of such data in a modeling in the time domain is also computationally demanding. Recently, harmonic pumping tests have been introduced as an alternative for the identification of the networks of preferential groundwater flows. Harmonic characterization designates an investigation in which an oscillatory/sinusoidal signal is used to disturb the water level of an aquifer. Different ways to produce such signals have already been proposed: a pumping-reinjecting system (Rasmussen et al., 2003; Renner and Messar, 2006), a moving mass at the water table interface (Gultinan and Becker, 2015), or a controlled pumping system (Lavenue and de Marsily, 2001). Then, the response signals among the aquifer contain an oscillatory part (characterized by an amplitude and a phase offset values) that can be easily isolated from the ambient noise (Bakhos et al., 2014; Cardiff and Barrash, 2015). Harmonic characterization has already been successfully applied to a large range of complex cases such as contaminated aquifer (Rasmussen et al., 2003), leakage detection (Sun et al., 2015), or fractured aquifers (Renner and Messar, 2006; Maineult et al., 2008; Gultinan and Becker, 2015). The theoretical aspects of the application of harmonic pumping to karstic

aquifers have also been developed in Fischer et al. (2018b). The imagery potential of harmonic investigations has been studied for mapping the distribution of hydraulic properties in heterogeneous aquifers with models solved in the time domain (Lavenue and de Marsily, 2001; Soueid Ahmed et al., 2016) or in the frequency domain (Cardiff et al., 2013b; Zhou et al., 2016).

In this article we will present a field characterization of karst network based on a harmonic pumping tomography. Hydraulic data were obtained from the Terrieu experimental site located in Montpellier, in Southern France. At the same site, results of hydraulic tomography, in which hydraulic responses were generated with constant-rate pumping tests, have already been presented and discussed in our previous works (Wang et al., 2016, 2017; Fischer et al., 2017c). In this new work, we rely our analysis on a set of responses to harmonic pumping tests with different oscillation periods, to characterize the karst network. We describe in Section 2 the experimental study site, the harmonic pumping investigation led on it, and the processing made on the measured field responses for the later interpretation. Then, in Section 3 we briefly introduce the numerical model setup and the Cellular Automata-based Deterministic Inversion (CADI) algorithm. Further details of our inverse algorithm can be found in Fischer et al., (2017b). In Section 4 we present the inversion results obtained with the CADI method at the Terrieu field site and the efficiency of the method in reproducing the observed hydraulic responses. Finally, Section 5 presents a discussion of the effect of the harmonic signal period on the inversion results.

2. Field investigation

2.1. Experimental site presentation

The Terrieu experimental site is located ~15 km in north of the town of Montpellier in southern France. The site consists of 22 vertical boreholes drilled within a surface area of approximately 2500 m² (40 × 60 m) and permits the study of karstic flows at a local scale (Fig. 1). As a part of the network of the French Karst Observatory (SNO Karst, www.sokarst.org) and the Medycys network (Jourde et al., 2011), the site has been used as a field laboratory for testing new field hydraulic methods and tools developed for the characterization of karstic aquifers (Jourde et al., 2002; Jazayeri Noushabadi, 2009; Jazayeri Noushabadi et al., 2011; Dausse, 2015; Wang et al., 2016; Wang et al., 2017; Fischer et al., 2017c).

The geological logs collected from the different boreholes shows that the subsurface of the field is composed of 35–45 m of thin-layered marly Cretaceous limestones, deposited on pure and massive Jurassic limestones. The interface between these two units is a sloped monocline bedding plane dipping at 20° toward Nord-West (Wang et al., 2016).

The Terrieu field is located in the Lez regional aquifer. Temperature and electrical conductivity measurements, and packer tests in boreholes presented in previous works (Jazayeri Noushabadi, 2009; Dausse, 2015) have shown the existence of preferential flow paths (shown in Fig. 1) along the bedding plane between the Cretaceous and Jurassic limestones. Downhole videos in the boreholes show, that, at this interface, karstic conduits with aperture up to 50 cm are present (Jazayeri Noushabadi et al., 2011).

The massive Jurassic limestones are non-aquifer and the Cretaceous limestones have a low permeability. This causes the aquifer to be confined at the interface between these two layers, in the bedding plane. A network of karstic conduits has developed preferentially on this bedding plane, and controls the fluid circulation within the aquifer.

2.2. Harmonic pumping investigation

The main dataset used in this study was collected from an investigation using harmonic pumping tests performed at the Terrieu site. We have conducted pumping sequentially in four different boreholes while

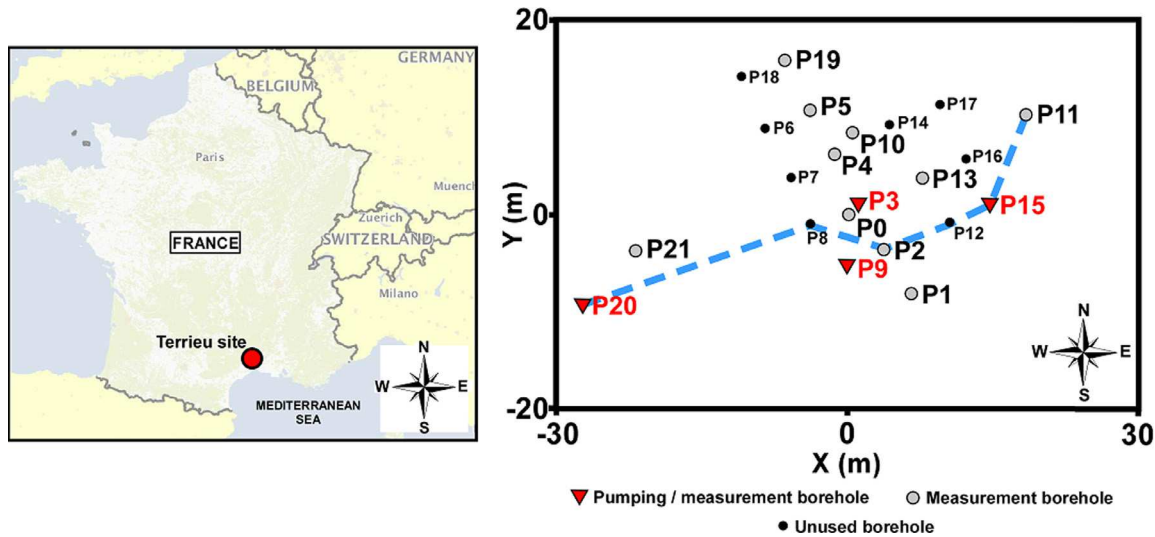


Fig. 1. Maps of localization of the Terrieu site in France (left) and well pattern on the site (right). Boreholes used as pumping and measurement points are indicated using red triangles, and boreholes used only as measurement points are indicated using grey circles. Boreholes indicated by solid black points were not used during the investigation. The blue dotted line delineates a preferential flow path identified by previous studies (Jazayeri Noushabadi, 2009; Dausse, 2015), which shows a connectivity between P2, P8, P11, P12, P15 and P20. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1

Harmonic pumping rates registered for each pumping point during the investigation. Q_A and Q_m refer to Eq. (1).

Pumping well:		P3	P9	P15	P20
Pumping rate (m^3/h)	Amp. Q_A	1	0.22	2.1	2.5
	Mean Q_m	4.1	0.35	5.3	3.8

recording the water-level responses in 13 selected observational boreholes (see Fig. 1). The water-level responses were also measured in the pumping wells P9, P15 and P20 but not in P3.

The static water level before the hydraulic investigation was at a depth of 20 m. The maximal drawdown generated by the pumping tests was 4 m. Therefore the karstic conduits (located at a depth of 35–45 m) were saturated during the entire duration of hydraulic tests.

The harmonic pumping tests were performed with a configurable electronic device, specially designed for this study by electronics engineers. This device controls a flow rate variator linked to the pump, which can generate a pumping signal with a sinusoidal shape around a mean value. The period and amplitude of the sinusoid signal can be configured with the device. The generated pumping rate can be described by:

$$Q(t) = Q_m - Q_A \cos(\omega t), \quad (1)$$

where Q is the output pumping signal (m^3/s), Q_m the mean pumping rate (m^3/s), Q_A the oscillatory signal amplitude (m^3/s), and $\omega = \frac{2\pi}{T}$ the pulsation (rad s^{-1}) with T the period (s).

Different signal amplitudes and mean values were independently applied in each different pumping borehole according to its productivity (see Table 1).

For each pumping location, two pumping tests with different periods ($T = 2$ min and $T = 5$ min) were conducted during 30 min (15 cycles for a 2 min period, 6 cycles for a 5 min period). Water-level variations were continuously measured with digital pressure sensors installed in the measurement wells.

Overall, this investigation permitted to record 104 drawdown curves (13 measurements for each 2 different periods of signal applied in each 4 pumping wells).

2.3. Data processing

In order to interpret the harmonic signal in the drawdown curves, we have performed the same signal decomposition as proposed in Fischer et al. (2018b). This decomposition consists in removing the linear part, induced by the mean pumping signal Q_m , from the drawdown curve (through a linear regression) to keep only the oscillatory response. This operation is feasible only after an early transient period (accordingly we truncate the first cycle of the recorded responses). As we show in Fig. 2 for the pumping in P15, the operation of removing the linear part is acceptable, as the resulting signals appear to be purely oscillatory. Some natural noises and vibrations induced by the pumping appear as high frequencies fluctuations in the oscillatory responses.

Mathematically the drawdown curves generated by the harmonic pumping tests can be approximated as a sum of a linear signal and an oscillatory signal applied on the initial water table level:

$$h(t) \simeq h_{lin.}(t) + h_{osc.}(t) + H_0$$

with $h_{lin.}(t) \simeq -at - h_0$ and $h_{osc.}(t) \simeq A \cos\left(\frac{2\pi}{T}t - \Phi \frac{\pi}{180}\right)$ (2)

where h represents the drawdown (m) over time, $h_{lin.}$ is a linear signal described by its slope a (m/s) and its intercept h_0 (m) (whose values can be retrieved by linear regression), $h_{osc.}$ is an oscillatory signal described by its amplitude A (m), its period T (s) and its phase shift Φ ($^\circ$), and H_0 represents the initial water table level (m) (in our case we considered $H_0 = 0$ m).

The linear signal $h_{lin.}$ can be easily estimated in a first approximation through a linear regression performed on each drawdown curve. After removing this linear trend, the amplitude and phase offset of the remaining signal of each borehole can be determined by a Fast Fourier Transform (FFT) on their oscillatory signals. The FFT permits to extract the main oscillatory components of a signal, to denoise it, and to interpret its parameters. Fig. 3 presents the FFT results for the oscillatory signals of three representative boreholes (P10, P11, P2) during a pumping in P15 with the two periods (2 min and 5 min). The interpretation results of amplitude and phase offset for the entire dataset are presented in Appendix Table A1.

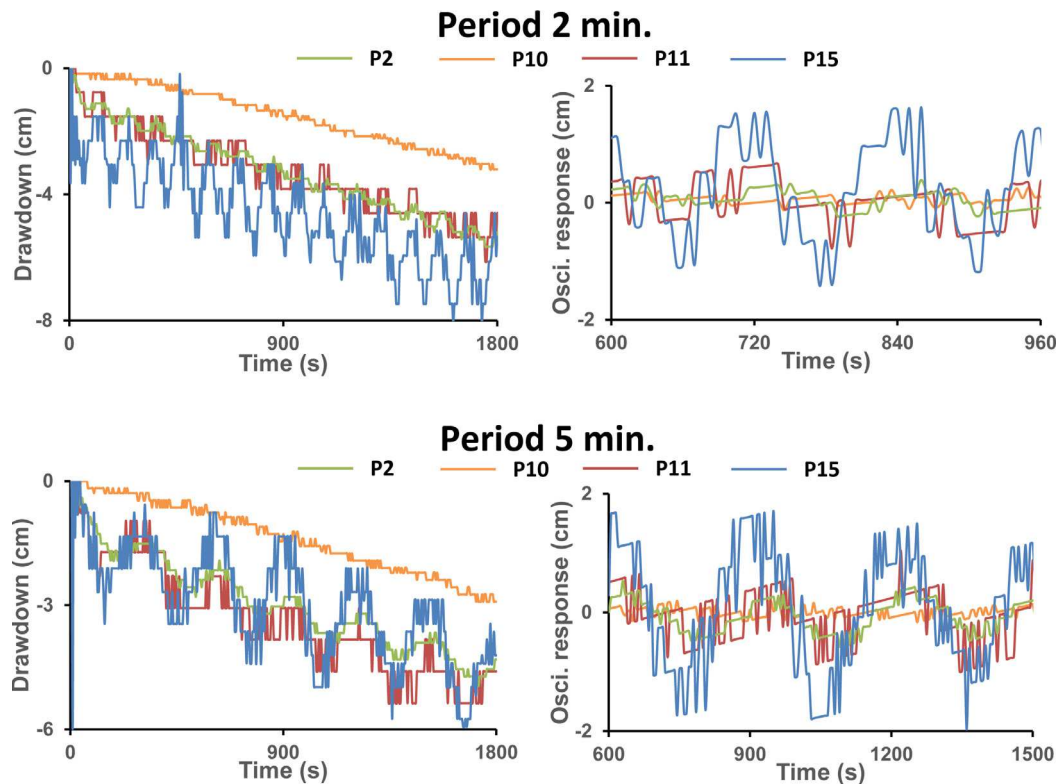


Fig. 2. Left: Measured drawdown curves for a selection of boreholes (P2, P10, P11, P15) during a pumping in P15 with a 2 min and a 5 min period. Right: Zoom-in view of three oscillation cycles after removing the linear part from the drawdown curves.

2.4. Preliminary analysis

The different responses of amplitude and phase offset interpreted in P10, P11 and P2 highlight three distinct flow behaviors (Fischer et al., 2018b). The responses in P10, having a negligible amplitude (< 1 mm) relatively to the pumping signal, which we interpreted as a negligible oscillatory response, is associated to a ‘matrix connectivity’ between the pumping and the observational well. In contrast, the response in P11 has a significant amplitude and an almost invariable phase relatively to the pumping signal for the two different periods. This behavior is associated to a ‘conduit connectivity’ response, meaning that P15 and P11 would be connected through a karstic conduit network. The response in P2 has a lower amplitude response than P11, and its phase offset relatively to the pumping signal increases as the pumping period decreases ($+71^\circ$ for a 2 min signal, $+38^\circ$ for a 5 min signal). This third behavior is associated to a ‘dual connectivity’ response, which corresponds to an inter-well connection either through fissures or when the observation borehole is located in the matrix but close to a conduit.

Following the method described in Fischer et al., (2018b) and through the integration of the amplitude and phase offset results interpreted for each pumping-observation well pair, it is possible to obtain a map of inter-well connectivity which contains qualitative information regarding the spatial distribution of the conduit network and the relative position of boreholes to the network. We can first link, on the map, the boreholes with a low phase shift relatively to the pumping signal, to represent a conduit connection. From this conduit connectivity we can then establish a possible conduit network, and then verify that the other boreholes responses would be adequate toward their position to the interpret network (dual connectivity with higher phase shift for boreholes close to the network and matrix connectivity with negligible responses for the others).

Possible connectivity maps interpreted with the responses to the harmonic pumping test in P15 for a period of 2 min. and 5 min. are proposed in Fig. 4.

The comparison between the two connectivity maps in Fig. 4 shows that the period of the harmonic pumping signal may have a slight impact on the connectivity interpretation. For example in the pumping test with a 2 min period P1 can be interpreted as connected to the pumping well through conduits, but not in the pumping test with a 5 min period. This implies that a change in the period of the pumping signal modifies the flow field induced by the pumping.

Furthermore, the manual interpretation is possible only when the amount of hydraulic data to deal with is limited (13 responses for each period in Fig. 4). Therefore, although such a qualitative analysis through manual interpretation of inter-well connectivity could provide some important guidance to hydrogeological investigation, such as indicating the general trend of the main conduits and relative inter-well connectivity, to obtain a quantitative hydrodynamics characterization and to integrate a larger amount of hydraulic measurements (104 responses from 4 different pumping locations with each time two different periods) an inverse modeling is required.

3. Modeling methodology

3.1. Forward problem and model parameterization

Inverse modeling involves the use of the forward problem in order to simulate, for a given model of hydraulic properties, the hydraulic responses. In this section we present a 2D model that represents the property field along the bedding plane interface, in which the karstic network has developed on the Terrieu site.

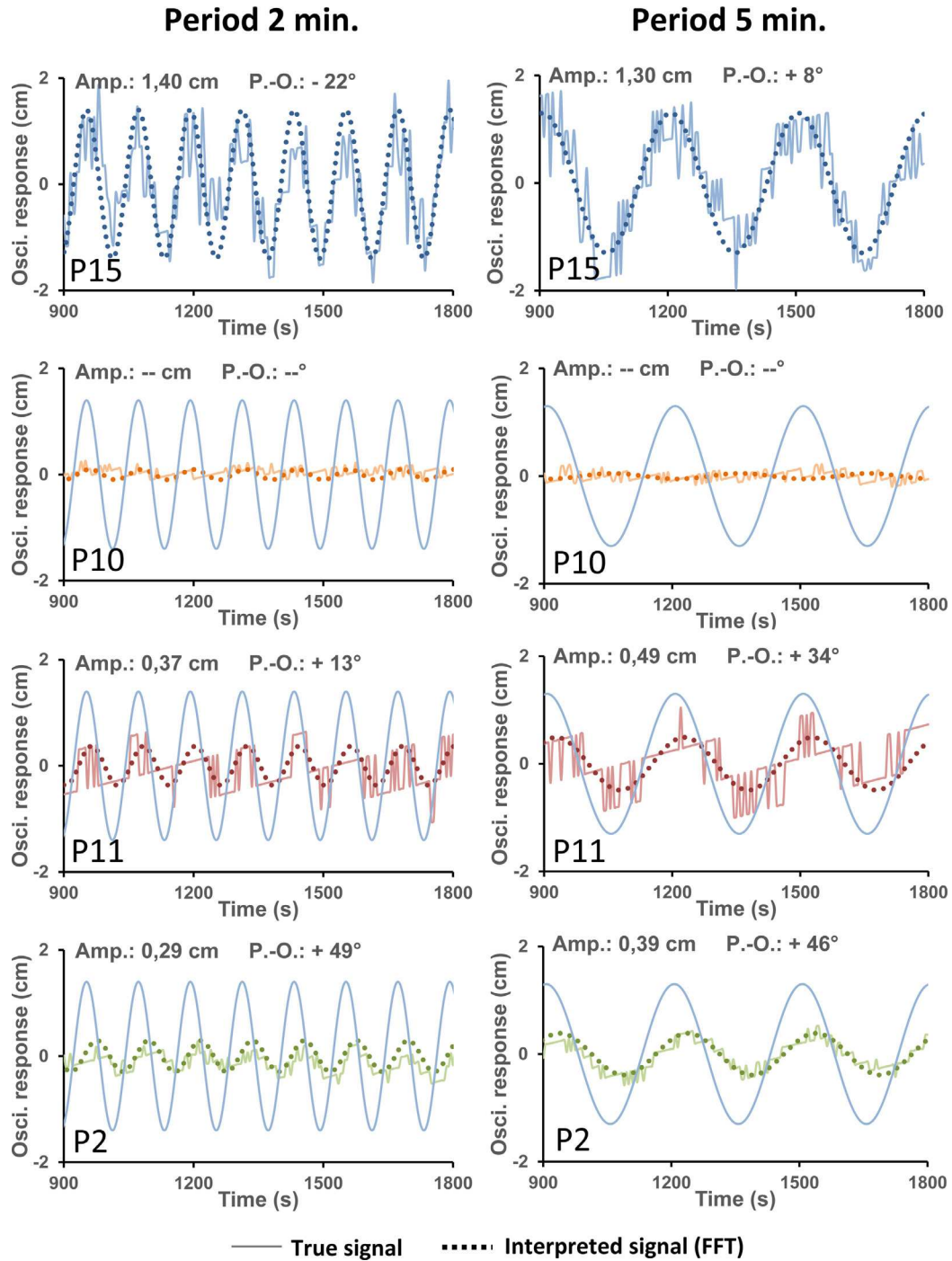


Fig. 3. Zoom-in on the oscillatory responses extracted from the drawdown measured in P2, P10, P11 and P15 during pumping tests in P15 with a 2 min (left) and a 5 min (right) signal periods and FFT results of the interpreted amplitude (Amp.) and phase offset (P.-O.) responses. Solid lines represent the measured signals, dotted lines represent the interpreted signals (h_{osc} in Eq. (2)) reconstructed from the amplitudes and phase offsets interpreted by FFT. For interpreted amplitudes smaller than 1 mm (for example here in P10), we considered the oscillatory responses to be negligible. The blue lines represent the interpreted pumping signals (P15) and are presented for each borehole for a better visualization of the interpreted phase offset responses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

As seen in the previous section, the variation of the piezometric level among the site, in response to a harmonic pumping, can be approximated by the sum of a linear drawdown and an oscillatory drawdown (Eq. 2). We simulate in the model only the oscillatory part of the drawdown responses h_{osc} . The inversion aims to reproduce the values of amplitude and phase offset of the oscillatory part in the measured responses. This oscillatory part can be described as a signal in the fre-

quency domain in the model:

$$h_{osc}(x, y, t) = \text{Re}(H_{\omega}(x, y)e^{i\omega t}). \quad (3)$$

with H_{ω} a complex parameter holding the amplitude and phase offset responses over space (x, y), Re the function returning the real part of a complex value, $\omega = \frac{2\pi}{T}$ the pulsation (rad s⁻¹) and i the imaginary unit.

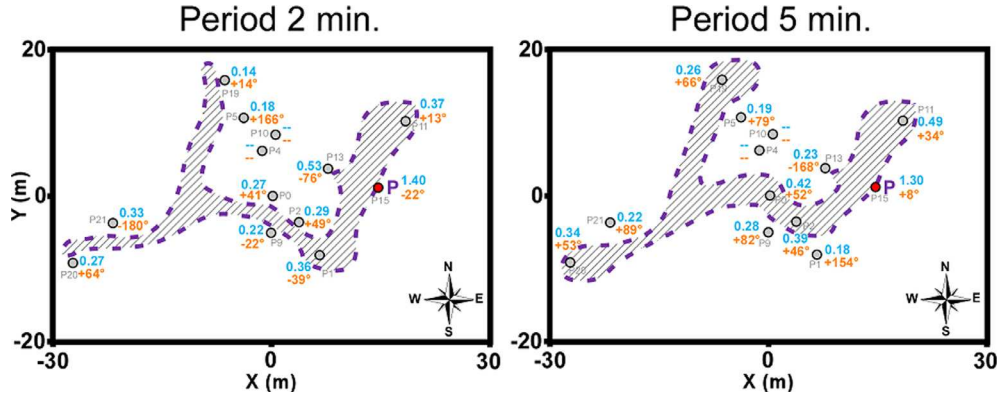


Fig. 4. Connectivity maps interpreted from the amplitude (in blue) and phase offset (in orange) responses to a pumping in P15 with a 2 min (left) and a 5 min (right) period of signal. The areas within the dotted lines delineate a possible area where boreholes are connected through a direct conduit connectivity. Dashes indicate negligible oscillatory responses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

This oscillatory feature of the hydraulic signal permits to rewrite the time domain form of the groundwater flow equation into a frequency form, in order to reduce the computation time of the forward problem.

In a 2D, porous, isotropic and saturated domain Γ the groundwater flow equation based on the Darcy's law in a frequency domain can be expressed as:

$$i\omega S_S H_\omega - \nabla \cdot (\mathbf{K} \cdot \nabla H_\omega) = \frac{Q_A}{V_{el}} \delta(x - x_s, y - y_s), \quad (4)$$

with S_S the specific storage distribution (m^{-1}), \mathbf{K} the conductivity distribution (m/s), Q_A the pumping amplitude (m^3/s), V_{el} an elementary volume of the finite element grid in the model, and $\delta(x - x_s, y - y_s)$ the Dirac distribution where x_s, y_s represents a pumping location. As the system is 2D, with a unit thickness, conductivity \mathbf{K} and transmissivity \mathbf{T} are of same value, as well as specific storage S_S and storativity S . In this study we considered Darcy's law to be acceptable for representing the flows generated in the karstic structures. In fact, it appears from the previous studies on the Terrieu site that the flows in the conduits have a low velocity, inducing a low Reynolds value, even for higher pumping rates than the ones used during this new investigation.

The initial and boundary conditions used for solving Eq. (4) are:

$$\begin{aligned} H_\omega(x, y) &= 0 \quad \forall (x, y) \in \Gamma \text{ as initial condition} \\ H_\omega(x, y) &= 0 \quad \text{when } (x, y) \in \Gamma_{bound.} \text{ as boundary condition.} \end{aligned} \quad (5)$$

The spatial distribution of the complex parameter H_ω permits the reconstruction of the oscillatory responses simulated among the model, through the calculation of their amplitude and phase offset values:

$$\begin{aligned} \text{Amplitude : } A(x, y) &= \sqrt{(\text{Re}H_\omega(x, y))^2 + (\text{Im}H_\omega(x, y))^2} \quad \text{in } m \\ \text{Phaseoffset : } \Phi(x, y) &= \frac{180}{\pi} \text{atan2}(-\text{Im}H_\omega(x, y), \text{Re}H_\omega(x, y)) \quad \text{in } ^\circ, \end{aligned} \quad (6)$$

where Re and Im are the functions returning the real and imaginary parts of a complex value, and atan2 is the function returning the inverse tangent value in radian mode from two arguments.

The simulated response signals are then reconstructed temporally and spatially:

$$h_{osc.}(x, y, t) = A(x, y) \cos\left(\omega t - \Phi(x, y) \frac{\pi}{180}\right). \quad (7)$$

In a karstic medium, the spatial response signals are very dependent to the highly heterogeneous distribution of the field properties \mathbf{T} and \mathbf{S} along the karstic conduits. Therefore this heterogeneity has to be taken into account in the distribution of these properties in the model in order

to simulate a realistic responses behavior. For this purpose we chose to apply as parameterization for our model and inverse problem the Cellular Automata-based Deterministic Inversion (CADI) method, developed and detailed in Fischer et al. (2017b). For a detailed description of the CADI method we refer the reader to Fischer et al. (2017b), as we will only briefly summarize the concept in this article.

The CADI method uses a particular parameterization of the property field in the model to generate linear structures (conduits) over a background (matrix). The field is composed of a grid of cells, each cell being assigned to a value of transmissivity and storativity. This grid of cells is divided in m_{CA} subspaces, each one being controlled by a cellular automaton piloting the part of the cells inside of its subspace (Fig. 5). The cellular automata concept is a mathematical tool which permits to generate structures within a grid with simple neighborhood and transition rules (Von Neumann and Burks, 1966). The cellular automata control the local direction of generation of the conduit in the different subspaces. They are piloted through eight different neighborhood definition $\mathbf{N}_i, i \in [1, 8]$ that permit to define eight different directions. One of these eight direction possibilities is assigned to each cellular automaton in the subspaces. The conduit network is generated by first assigning a state 'matrix' or 'conduit' to each cell. The whole cells are initially in state 'matrix', except an initial cell of the grid in state 'conduit' which designates the starting point and starting subspace for the generation of the conduits. The network of conduits then generates following the different local direction affected to each subspace it crosses. The generation ends once each part of the network has reached an end (either the limit of the model or a subspace in which the network has already generated).

Then property values are assigned to the cells depending on their state ('matrix' or 'conduit') and their localization (subspace). Each subspace defines locally a value for \mathbf{T}_{mat} and \mathbf{S}_{mat} for its cells in state 'matrix' and a value for \mathbf{T}_{cond} and \mathbf{S}_{cond} for its cells in state 'conduit'.

In order to be able to easily pilot and modify the configuration of the model through this parameterization, the structural directions and properties assigned to each subspace are defined in two parameter vectors: \mathbf{P}_N and \mathbf{P}_β . \mathbf{P}_N is a m_{CA} -vector containing the directions of generation $\mathbf{N}_i, i \in [1, 8]$ assigned to each subspace. Several independent networks can be generated in the same model with different directions parameters for each network. In this case \mathbf{P}_N becomes a $(m_{CA} \times \text{frac})$ matrix where **frac** represents the amount of independent networks in the model. Each column contains the subspaces directions for each network. \mathbf{P}_β is a $4m_{CA}$ -vector containing the \mathbf{T}_{mat} and \mathbf{S}_{mat} and the \mathbf{T}_{cond} and \mathbf{S}_{cond} values assigned to each subspace. In this way the parameters controlling the configuration of the model, and thus the model itself, can be easily and locally modified. \mathbf{P}_N and \mathbf{P}_β represent the parameters

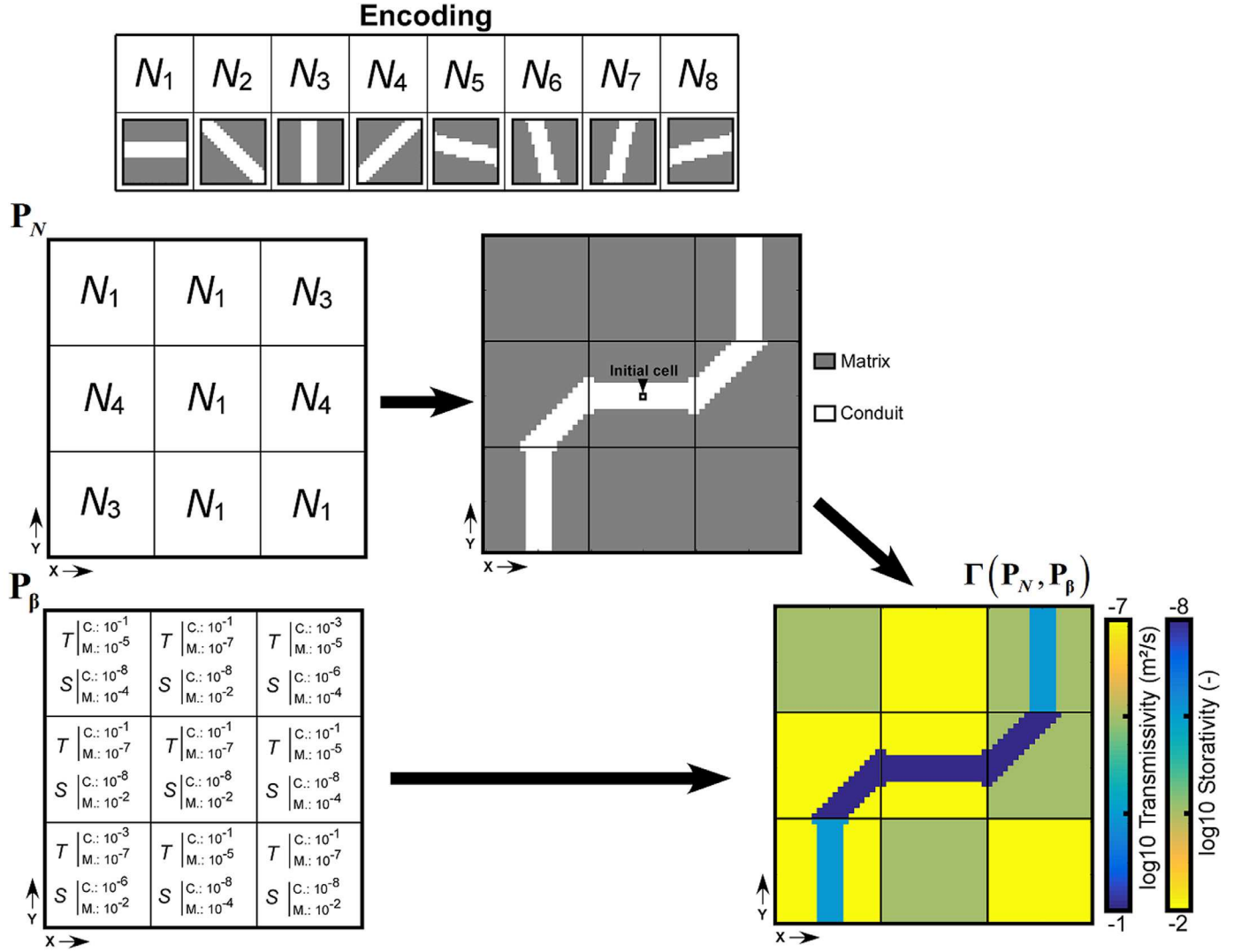


Fig. 5. Schema of the parameterization of a model with the CADI method. P_N contains the encoded (see Encoding) structural directions of generation associated to each subspace which permits to generate, from an initial 'conduit' cell, a network of conduits in the matrix. P_β contains the conduit (C) and matrix (M) transmissivity and storativity values associated to each subspace. $\Gamma(P_N, P_\beta)$ designates the model produced by applying the property values from P_β to the network generated from P_N .

to be optimized in the inverse problem in order to reproduce the observed data (amplitude and phase offset responses) through a suitable model.

3.2. Inverse problem

The inverse problem consists in retrieving the best values for the parameters contained in P_N and P_β regarding the minimization of the gap between the simulated data and the observed data (amplitude and phase offset responses at the measurement points for the different pumping tests). This inverse algorithm contains two steps in which we seek to minimize two objective functions sequentially, $\Psi_{structure}$ for the optimization of the structural parameter P_N and $\Psi_{properties}$ for the optimization of the property parameter P_β (Tarantola and Valette, 1982):

$$\Psi_{structure}(P_N) = \frac{1}{2} (d_{obs} - f(\Gamma(P_N, P_\beta)))^T C_d^{-1} (d_{obs} - f(\Gamma(P_N, P_\beta))) + \frac{1}{2} (P_{N,prior} - P_N)^T C_{P_N}^{-1} (P_{N,prior} - P_N), \quad (8)$$

$$\Psi_{properties}(P_\beta) = \frac{1}{2} (d_{obs} - f(\Gamma(P_N, P_\beta)))^T C_d^{-1} (d_{obs} - f(\Gamma(P_N, P_\beta))) + \frac{1}{2} (P_{\beta,prior} - P_\beta)^T C_{P_\beta}^{-1} (P_{\beta,prior} - P_\beta). \quad (9)$$

where d_{obs} is a n -vector containing the n measured responses, $f(\Gamma(P_N, P_\beta))$ is a n -vector containing the responses simulated with the model at the same positions than in d_{obs} , C_d is a $(n \times n)$ matrix of covariance on the data, $P_{N,prior}$ is a m_{CA} -vector holding *a priori* structural parameters for P_N , $P_{\beta,prior}$ is a $4m_{CA}$ -vector holding *a priori* property values for P_β , and C_{P_N} and C_{P_β} are $(m_{CA} \times m_{CA})$ and $(4m_{CA} \times 4m_{CA})$ matrices of prior covariance on the parameters P_N and P_β .

At the beginning of the inversion process, the variable parameters P_N and P_β are initialized with a priori conduit directions and property values in order to create the initial model. Then a sequential and deterministic process optimizes firstly the structural parameters in P_N (considering the initial properties in P_β as invariable), and then, in second step, the property parameters in P_β (considering the previously optimized P_N as invariable). Finally, after the optimization process, the posterior uncertainties on the structural and property parameters are estimated.

3.2.1. Optimization of the structural parameters

The optimization of the structural parameters in P_N is an iterative process in which a sensitivity matrix is computed, at each iteration step, to minimize the objective function in Eq. (8). This analysis requires the computation of a $(8 \times m_{CA})$ sensitivity matrix S . At a given iteration step k , each element (i, j) of the matrix is calculated as follow:

$$S^k(i, j) = \left(d_{obs} - f \left(\Gamma \left(P_N^k \Big|_{P_N^k(j)=N_i}, P_\beta \right) \right) \right)^T C_d^{-1} \times \left(d_{obs} - f \left(\Gamma \left(P_N^k \Big|_{P_N^k(j)=N_i}, P_\beta \right) \right) \right) + \frac{1}{2} (P_{N,prior}(j) - N_i)^T C_{P_N}^{-1} (P_{N,prior}(j) - N_i) \quad (10)$$

where $f(\Gamma(P_N^k \Big|_{P_N^k(j)=N_i}, P_\beta))$ represents the responses simulated with the modified direction N_i in the subspace $P_N^k(j)$, and $P_{N,prior}(j) - N_i$ denotes the angular gap between the modified direction N_i and the *a priori* direction $P_{N,prior}(j)$.

The coordinates (i_{min}, j_{min}) of the minimal value in the sensitivity matrix provide the subspace to be optimized (j_{min}) and the direction to apply (i_{min}) in order to minimize the objective function during this iteration. At the end of an iteration, the value of the structural objective function is recalculated. This iterative optimization ends when the objective function has reached a minimum (no more structural modification can decrease the objective function). The optimized parameters in P_N will then be considered as invariable during the optimization of P_β .

After the convergence of the objective function in the structural optimization, the uncertainties on the local directions of the geometry of the network can be estimated from the posterior structural covariance:

$$C_{P_N}^{post}(j) = \left(\frac{1}{8} \sum_{i=1}^8 S^{post}(i, j) - \Psi_{structure}^{post} + C_{P_N}^{-1}(j, j) \right)^{-1} \quad (11)$$

where $C_{P_N}^{post}(j)$ denotes the uncertainty associated to the direction of the subspace j , S^{post} is the sensitivity matrix of the last iteration, and $\Psi_{structure}^{post}$ is the value of the minimized objective function associated to the last iteration. The higher the uncertainty value associated to a subspace is, the more the direction of the subspace is uncertain. In the contrary a low value denotes a well constrained direction.

3.2.2. Optimization of the property parameters

In the second step, we estimate the hydraulic properties in P_β through an iterative optimization process and by considering the previously optimized P_N as invariable. The process relies on a linearization of the objective function in Eq. (9), through the computation of the Jacobian matrix. At an iteration step k , the values of the properties in P_β are updated from the previous ones as follow:

$$P_\beta^{k+1} = P_\beta^k + \left((J^k)^T \cdot C_d^{-1} \cdot J^k + C_{P_\beta}^{-1} \right)^{-1} \cdot (J^k)^T \cdot C_d^{-1} \times \left(d_{obs} - f \left(\Gamma \left(P_N, P_\beta^k \right) \right) \right) + C_{P_\beta}^{-1} \cdot (P_{\beta,prior} - P_\beta^k) \quad (12)$$

where J is the $(n \times 4m_{CA})$ Jacobian matrix calculated with a finite difference method: $J(i, j) = \frac{\partial f_i}{\partial P_\beta} \Big|_{P_\beta(j)=P_\beta(j)+\Delta P_\beta}$ with ΔP_β a finite difference step.

The value of the properties objective function is recalculated at the end of each iteration. This iterative optimization ends when the objective function converges to a minimum value.

The uncertainties on the T and S values can then be estimated through the computation of the posterior properties covariance matrix. The square root values of the diagonal entries of this matrix represent the standard deviation associated to each property value:

$$C_{P_\beta}^{post} = \left((J^{post})^T \cdot C_d^{-1} \cdot J^{post} + C_{P_\beta}^{-1} \right)^{-1} \quad (13)$$

where $C_{P_\beta}^{post}(j, j)$ denotes the variance of the j th property parameter in P_β , and J^{post} is the Jacobian matrix of the last iteration.

Table 2

Parameters used for the inversion process.

Parameter	Value
Final partitioning	12 × 8
Final grid size	Δx: 0.25 m; Δy: 0.25 m
Final network thickness	1 m
T_{init}	10 ⁻¹ m ² /s for the conduits; 10 ⁻⁶ m ² /s for the matrix
S_{init}	10 ⁻⁸ for the conduits; 10 ⁻⁴ for the matrix
T_{buffer} ; S_{buffer}	10 ⁻² m ² /s; 10 ⁻³
σ_{data}^2	0.01 on amplitude values
	10 on phase offset values
$\sigma_{prop.}^2$	0.1 (applied on the exponent: 10 ^{β ± σ_{prop}})

4. Modeling application

4.1. Modeling strategy

We have applied the CADI method on the 2D parameterized model presented in the previous section, in order to find a structured property distribution able to reproduce the measured oscillatory responses (amplitudes and phase offsets) presented in the Section 2.3. We considered the oscillatory responses interpreted as negligible (amplitude < 1 mm) to be null for the inversion. We have coded the CADI algorithm with Matlab and used Comsol Multiphysics to solve the model in the frequency domain (see Eq. (5)). This resolution was led using a finite element method on a triangular mesh, refined around the boundaries between the matrix and the conduit network represented by the equivalent porous media properties distributed over the model. This mesh refinement is performed with the Comsol *mphimage2geom* function on the conduit network image generated by the cellular automata.

The distributed model is constructed as a 40 × 60 m² rectangle included in a 1000 × 1000 m² buffer zone. The external boundaries of the buffer zone are built with imposed Dirichlet boundary conditions as presented in Eq. (5). Thus, this zone permits to limit the effect of the boundary conditions on the parameterized model.

The values of the parameters chosen for the model parameterization and the inversion process are presented in Table 2. The initial values for the inversion were chosen accordingly to estimates from previous studies on the Terrieu site (Jazayeri Noushabadi, 2009; Dausse, 2015; Wang et al., 2016). In the inverse problem the properties values β in P_β were associated to the exponent of the transmissivity and the storativity. The initial standard deviation values on the data (σ_{data}) and on the property parameters (σ_T, σ_S) are used to construct the covariance matrices as diagonal matrices: $C_d = \sigma_{data}^2 \times Id(n)$ and $C_{P_\beta} = \sigma_{prop.}^2 \times Id(4m_{CA})$. No *a priori* information were considered for the structure local directions in $P_{N,prior}$ and C_{P_N} .

The inversion was led following a multi-scale method (Grimstadt et al., 2003), as described in Fig. 6. The multi-scale inversion consists in performing an inversion first for a coarse resolution of the model, and then use the inversion result as a new initial model with a higher resolution for a new inversion process. This permits to progressively reduce the size of the discretization cells for the property field during the inversion, which can be interesting for studies on heterogeneous fields with no prior information on the property distribution. This was already done with the CADI method in Fischer et al., (2017c).

The initial model has been constructed with two unidirectional conduits with uniform property values (see Table 2) as a coarse approximation of the manual estimation made in Fig. 4.

A first set of inversions were led by separating the 2 min and 5 min period responses, and with a 6 × 4 subspaces partitioning of the model (with a conduit thickness of 2 m). The results of these separate inversions have then been repartitioned in 12 × 8 subspaces models (with a conduit thickness of 1 m), which were used as initial models for a new inversions with the same datasets. This partitioning of the models per-

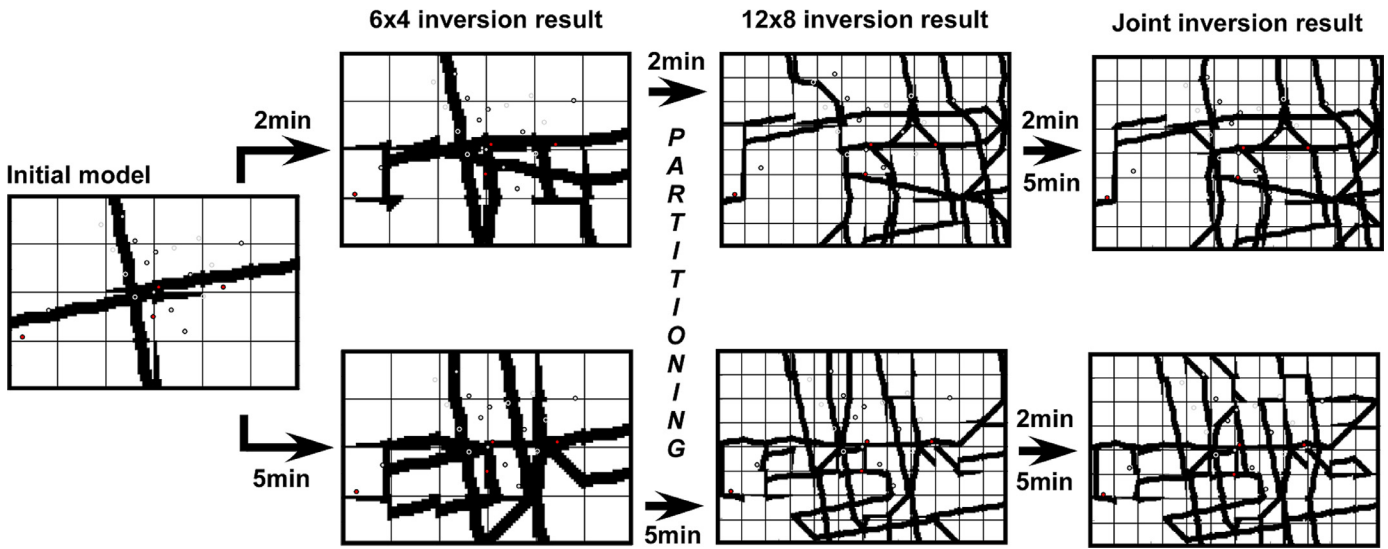


Fig. 6. Schematization of the complete multi-scale inversion process. Starting from an initial model, firsts inversions were led for a 6 × 4 partitioning (shown by the grid). The results were refined to 12 × 8 subspaces and used for new inversions. Finally, joint inversion were led starting from the results of the previous separate inversion.

Table 3
RMSEs on the amplitude (Amp.) and phase offset (P.-O.) values for the different inversion results. RMSEs values in brackets represent responses that were simulated through models generated for another period of signal (i.e. 5 min responses simulated with a model generated specifically for the 2 min responses and vice versa).

Results RMSEs		'2 min'	'5 min'	'2 min (+ 5 min)'	'5 min (+ 2 min)'
Amp.	2 min	1.1 cm	(11 cm)	5.4 cm	5.1 cm
	5 min	(9 cm)	0.5 cm	5 cm	5.3 cm
P.-O.	2 min	56 °	(112 °)	60 °	60 °
	5 min	(85 °)	66 °	67 °	66 °

mits to give more liberty to the inversion process, while starting from 'not too far' solutions, which is especially interesting for a deterministic process. The inversion results for the 12 × 8 models with 2 min and 5 min data separated will be presented in Section 4.2.

A final inversion process has consisted in starting from these 12 × 8 separate results, with the same partitioning, by adding the 5 min data to the 2 min inversion result and the 2 min data to the 5 min results, for joint inversions. The results of these joint inversions will be discussed in Section 4.3.

4.2. Modeling results

Fig. 9 shows the spatial distribution of hydraulic transmissivity and storativity inverted using the responses to the 2 min and 5 min periods, respectively. The comparison between the measured and simulated hydraulic responses is presented in Fig. 7 and Table 3. It can be seen that the simulated responses match the measured ones quite well.

The simulated responses to the P15 pumping appear to be slightly overestimated implying the existence of a very productive conduit in P15 that could not be simulated in the presented model. Otherwise, the responses proportions and behaviors for each pumping are well respected with the simulated signals. Overall, the Root Mean Square Error (RMSE) of the amplitude is 1.1 cm for the 2 min response signals, and 0.5 cm for the 5 min response signals. The RMSE of the phase offset is 56° for the 2 min response signals, and 66° for the 5 min response signals (see Table 3). The difficulties in reproducing the phase offset data with the CADI method may be contributed to the high degree of variation of the phase shift within the low transmissivity matrix (Fig. 8). A small displacement of a certain wellbore location by 1 m in the matrix

can modify the phase offset by a value of 90°. Thus, our phase offset RMSEs remain under the variations produced by a 1 m displacement on the field, which is acceptable at our scale.

The amplitude of the signal is also decaying very fast in the matrix (Fig. 8), thus the amplitude of a response is already a good information to characterize the proximity of a borehole with a conduit of the karst network. According to Fischer et al. (2018b) and the maps in Fig. 8 produced with the CADI method, the amplitude in the responses signals of a karstic aquifer permit to distinguish the boreholes in (or near) conduits from the ones in the matrix. On the other side, the phase offset response permit to characterize more precisely the distance of a responding borehole to conduits, as it varies orthogonally to the direction of the conduit and stays very low within the network.

One advantage of using the CADI method in this work is that the optimized conduits networks can be clearly distinguished from the matrix in the inverted fields in Fig. 9. If these optimized conduits networks represent only one possible geometry among other likely models, they permit to interpret the relative positioning of each borehole (in a conduit, close to a conduit, or in the matrix) and thus a degree of connectivity between them. The models of networks produced by inversions of pumping tests of different periods are very different, indicating that the both sets of responses provide different hydraulic information of the aquifer. The reconstructed network for the dataset of a period of 5 min is denser than that for a period of 2 min. The 5 min period dataset seems to carry information about karstic structures of different scales (conduits, fractures, fissures) around the measurement points, while the 2 min period dataset tends to characterize more specifically the most conductive karstic structures over the field scale. This hypothesis is supported by the maps of the amplitude in the models. Amplitude responses to a 2 min pumping signal in P15 quickly decrease around the borehole but remain visible in the coarse network over almost the entire field, while the ones to a 5 min pumping signal stay high in the dense karstic network around P15, and decrease beyond (Fig. 8).

Concerning the property values, it appears that the reproduction of the responses required more modifications on the transmissivity values than on the storativity, especially for the conduits. Some similarities appear between the results to both periods, for example the existence of conduits near P9 with lower transmissivities, which can then be assimilated more likely to large fissures rather than conduits. Also in both results the conduits around P15 have been associated to very high transmissivities (> 1 m²/s), which tends to indicate the existence of a very

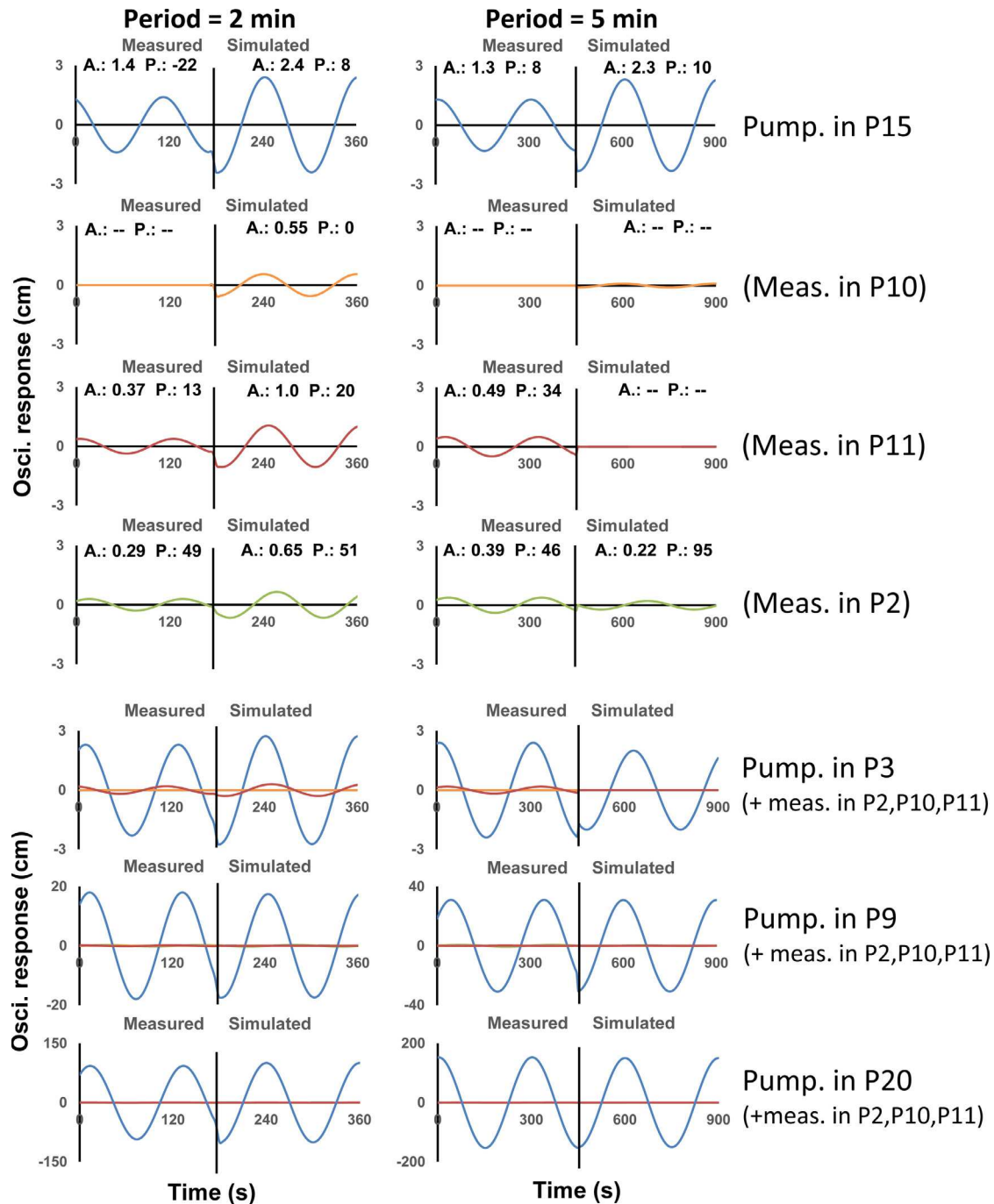


Fig. 7. Comparison of some measured and simulated (with the property distributions presented in Fig. 9) responses signals in observation points P2 (green), P10 (orange), P11 (red) and pumping points P3, P9, P15, P20 (each time in blue), for pumping signals with a 2 min (left) and a 5 min (right) period. In the case of the pumping in P3 we present in blue the signal in P0, located 1 m away from P3 (which was not measured during the investigation). For a better readability the responses are presented separately for a pumping in P15 with their amplitude (A. in cm) and their phase offset (P. in °) values. For the pumping in P3, P9 and P20 the responses are presented on a same graph. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

productive conduit at this position. This information could already been deduced by the fact that the responses induced by a pumping in P15 in the model were slightly overestimated (Fig. 7).

The fact that the two periods of signal lead to different solutions to the inverse problem, even by starting from a same initial model, tends to indicate that different periods of pumping signal induce different flow fields in the tested karst aquifer. In order to better understand the bene-

fits of each period in a harmonic pumping characterization, one would need to study the results of the joint inversions, led with the responses to both periods simultaneously.

4.3. Effect of the period of pumping signal on the inversion results

It appears, in fact (see Table 3), that the separate models can badly simulate the responses associated to the signal period not used in each

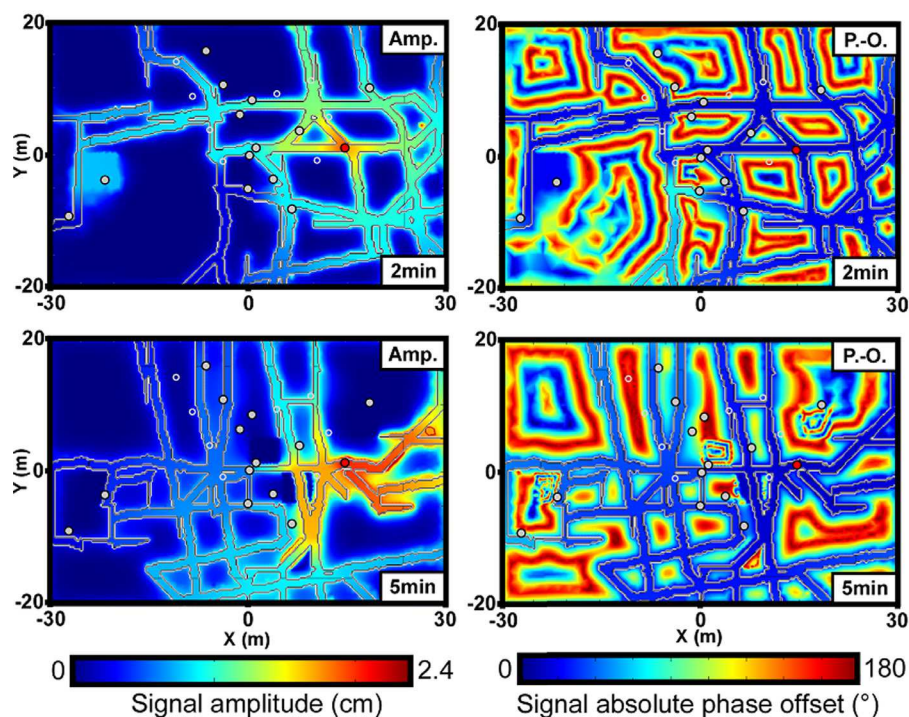


Fig. 8. Maps of simulated spatial amplitude (Amp.) and phase offset (P.-O.) with the models in Fig. 9 for a pumping in P15 with a signal period of 2 min and 5 min.

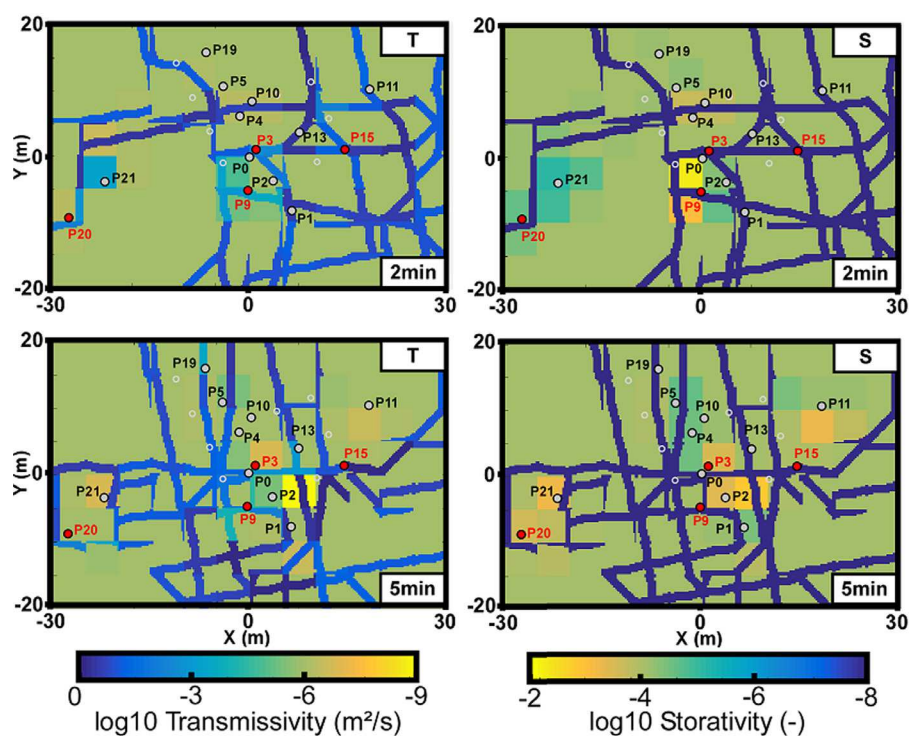


Fig. 9. Maps of the distributions of transmissivity (T) and storativity (S) found by separate inversions of the responses to periods of 2 min and 5 min.

inversion (i.e. 5 min responses simulated through the model generated with 2 min responses, and conversely), suggesting that each set of response contains different and complementary information for the characterization of the field. Therefore, new inversions were started from the results presented in Fig. 9 as initial models, by joining the mis-

ing responses to the ‘observed responses’ dataset in the inverse process (see the joint inversion results in Fig. 6). For a better understanding, we will mention as ‘2 min’ and ‘5 min’ separate results the model results in Fig. 9 produced from the inversion of the responses to the 2 min and 5 min periods separately. The models produced by inversions of the joint

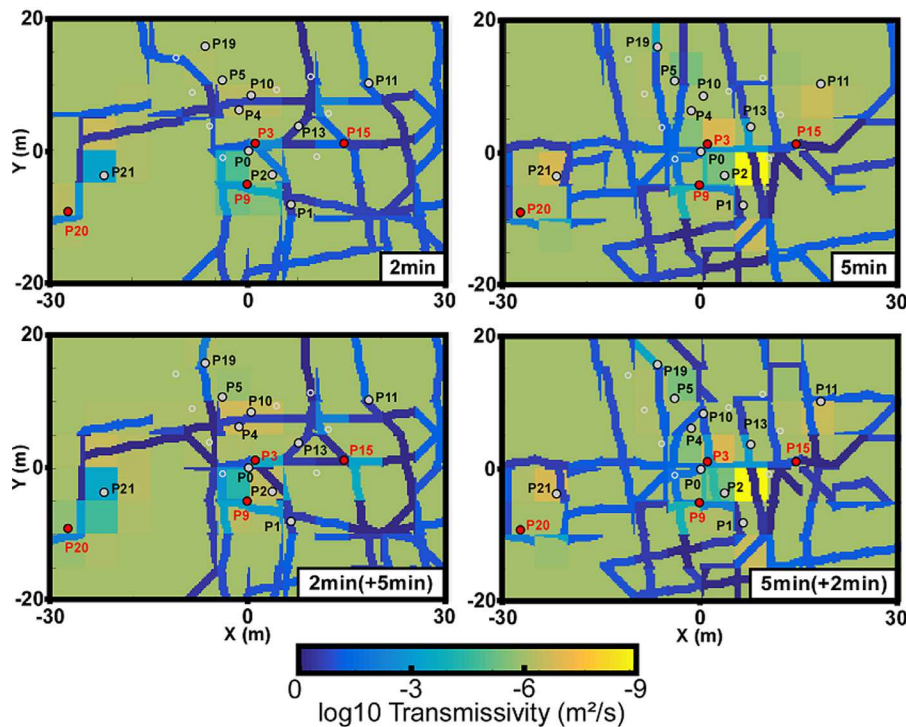


Fig. 10. Maps of the distributions of transmissivity found by inversions of the responses to the 2 min and 5 min periods, and joint inversions started with the 2 min result (2 min (+ 5 min)), and with the 5 min result (5 min (+ 2 min)).

datasets started with the ‘2 min’ model and with the ‘5 min’ model will be respectively mentioned as the ‘2 min (+ 5 min)’ and ‘5 min (+ 2 min)’ joint results. The results of the joint inversions are presented in Fig. 10.

The ‘2 min (+ 5 min)’ result, solution to the inversion started from the ‘2 min’ result, is very close to its initial model. The ‘5 min (+ 2 min)’ result, solution to the inversion started from the ‘5 min’ result, shows some modifications on the periphery of its network (P4, P5, P10, P11). Both joint solutions, however, do not permit a reproduction of the measured signals as good as the one generated by the ‘2 min’ and ‘5 min’ separate results (see Table 3).

Although the phase offsets RMSEs are almost the same for the joint results and the separate results, the amplitudes RMSEs are multiplied by 5–10 with the joint results. This shows that, even if the amplitude responses can be well reproduced by separate models for each period, they cannot be reproduced very well with a unique model. This tends to validate the hypothesis that different periods of pumping signal induce different flow fields in the aquifer, which need to be characterized separately. In fact, the CADI method is limited in its ability to represent a variation of aperture in the generated network, which can partly explain why the joint inversions are less good than the separate ones if the flows mobilized with each period occur in structures of different apertures.

However, even if the preferential flows change among the field for different pumping periods, the relative distance of each borehole toward the karstic network does not depend on the hydrodynamic but on the morphology of the karst structures, and thus their connectivity behavior should logically remain the same as the karst structure does not vary with the period of pumping. This is also what the reproduction of the phase offset values in the joint results would tend to indicate. In fact, according to Fig. 8, if the area of propagation of the amplitude response is dependent to the density of the conduit network at that scale, it is not true for its phase offset which remains null in the conduits regardless of the network geometry.

In Table 4 we present the interpretation of the position of each borehole relatively to the conduits network, with the data processing estimations, the separate modeling results and the joint modeling results.

First, it is interesting to point out that the estimations made manually in Fig. 4 for the P15 pumping match for 62% in term of position (in the conduit, close to a conduit, in the matrix) the separate modeling results ‘2 min’ and ‘5 min’. Taking account that the estimation were made with only 13 responses over the 52 available, it shows that the manual interpretation method described in Section 2.4 and in Fischer et al. (2018b) can already provide rather interesting and fast estimations.

According to Fig. 8, the amplitude and phase offset responses of an observable point very close to a conduit (< 0.5 m) are almost undistinguishable to the ones directly in a conduit. Therefore we consider these points to also have a conduit connectivity response in Table 4. Fig. 11 maps represent the interpreted connectivity of each borehole obtained from the joint inversion results. This figure also shows that these results reproduce the schema of connectivity of the preferential flow path established in Jazayeri Noushabadi (2009) and Dausse (2015) between P2, P11, P15, and P20 (see Fig. 1).

Taking into account the connectivity response, 57% of the boreholes show a similar behavior between the ‘2 min’ and the ‘5 min’ separate results. This degree of similarity increases to 93% when comparing the boreholes connectivity responses from the two joint results. The only behavior difference between the two results comes from P21, which appears as connected through the matrix in the ‘2 min (+ 5 min)’ result and as connected through conduits in the ‘5 min (+ 2 min)’ result. However the property distribution of the ‘2 min (+ 5 min)’ result in Fig. 10 shows that its matrix transmissivity near the P21 point has a value close to a conduit transmissivity. This permits to induce a conduit connectivity behavior for P21, even if the conduit is distant from the borehole. Therefore, we can consider that P21 should have a ‘conduit connection behavior’. The fact that the degree of similarity of the boreholes behavior has increased to almost 100% for two joint result clearly shows that the reproduction of the whole responses dataset requires to delineate a unique connectivity relation between each borehole, even if it is not sufficient to reproduce the amplitudes of the measured responses.

Table 4

Positioning or connectivity response of each borehole interpreted from the qualitative estimations (Fig. 4), the separate inversion results (Fig. 9), and the joint inversion results (Fig. 10).

	Period=2 min			Period=5 min			
	Est.	Sep.	Joint	Joint	Sep.	Est.	
P0	○	⊗	⊗	×	×	×	P0
P1	×	×	×	⊗	⊗	○	P1
P2	○	○	○	○	○	×	P2
P3	—	×	×	×	⊗	—	P3
P4	—	○	○	○	—	—	P4
P5	○	—	—	—	○	○	P5
P9	○	⊗	⊗	×	×	○	P9
P10	—	⊗	⊗	×	—	—	P10
P11	×	×	×	×	—	×	P11
P13	×	×	×	×	×	○	P13
P15	×	×	×	×	×	×	P15
P19	○	—	×	×	×	×	P19
P20	○	○	○	○	○	×	P20
P21	○	—	—	⊗	⊗	○	P21

Notation:

×: in a conduit = conduit connectivity response.

⊗: close to a conduit (<0.5 m) = conduit connectivity response.

○: close to a conduit (<2 m) = dual connectivity response.

—: in the matrix = matrix connectivity response.

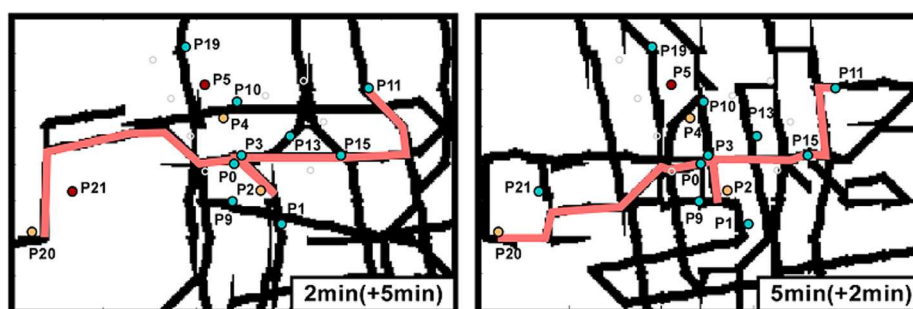


Fig. 11. Maps of the connectivity responses associated to each borehole from the networks (shown in background in black) inverted with the joints inversions. Boreholes in blue are associated to a conduit connectivity, in orange to a dual connectivity, and in red to a matrix connectivity. The red lines show flow paths in the models which show a same connectivity as the field preferential flow path highlighted in Jazayeri Noushabadi (2009) and Dausse (2015). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The remaining question is: which period of signal contains this information, as the separate results show only a 57% similarity? The conduits network almost didn't change during the joint inversion started from the '2 min' result, and in fact, the separate '2 min' result and the joint '2 min (+ 5 min)' result show a 93% similarity in their borehole behaviors, while the '5 min' result has a degree of similarity of 71% with '5 min (+ 2 min)'. Then, most of the borehole behaviors found in the joint results were already present in the result of the 2 min signal period, which indicates that the 2 min responses contain the most information about the connectivity of each borehole.

If the two sets of amplitude responses are not well reproduced by the joint inversion while they both delineate a similar type of response for each borehole, it shows that the characterized property distribution (*in extenso* the induced flow paths network) is different for each period. In fact, the responses to a 5 min period require a dense flow network to be reproduced, while the responses to a 2 min period require a much more dispersed network. The structural posterior uncertainty maps in Fig. 12 indicate that the network of the '2 min' result is very uncertain compared to the network of the '5 min' result. This indicates that the responses to a 5 min period contain more information on the localization of the flow paths around the measurement points than the responses to a 2 min period. These last ones seem to provide less precise information

on the localization of the flows between the boreholes among the field. When inverting jointly the responses from the two period, the networks in the joint results show overall lower uncertainties. Globally, the study of the structural uncertainties tends to indicate that, while, as seen before, responses to a the lower period contain more information on the type of connectivity of each borehole, the responses to the higher period contain more information on the position of the preferential flow paths around the boreholes. However, the fact that these both sets of responses cannot be well reproduced simultaneously also indicates that the flows highlighted in the '5 min' result do not exist with a period of 2 min. Therefore, while the lower period essentially mobilizes water from the most conductive karstic structures among the field, the dense flow field highlighted by the higher period can be assimilated to a mobilization of water also in less conductive karstic structures.

The maps of the posterior uncertainties on the transmissivity values in Fig. 13 tend to show that the responses to a 2 min period give more information on the transmissivity of conduits locally around the pumping points (average ± 0.1 on the transmissivity exponent) than the responses to a 5 min period (± 0.2 on the transmissivity exponent). The uncertainties on the storativity values remain high in the whole conduit network and for both periods (± 0.2 on the storativity exponent). Further from the pumping boreholes, both property values in the con-

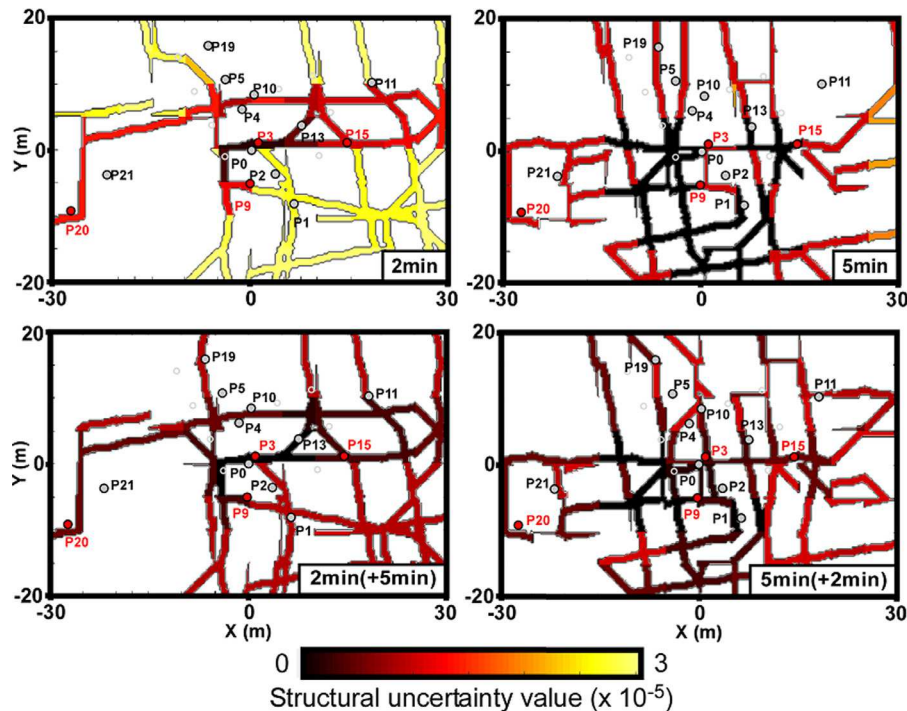


Fig. 12. Structural uncertainty values from the results found for separate inversions of the 2 min and 5 min responses, and joint inversions started with the 2 min result (2 min (+ 5 min)), and with the 5 min result (5 min (+ 2 min)).

duits network (transmissivity and storativity) globally remain uncertain (± 0.3 on the exponent), even in the joint results. The transmissivity and storativity values in the matrix are more constrained around the measurement points in the matrix. These information indicate that, for both periods, the characterization of the property values with oscillatory responses remains local, around the boreholes. Oscillatory responses provide more information on the global connectivity and the localization of preferential flows rather than on the property values of the matrix and the conduits and fissures in the aquifer.

Overall, regarding our results, it appears that the flow paths generated by periods of 2 min and a 5 min in the pumping signals in this karstic field are very different. It appears in fact that, at our field scale, higher frequency signals (here a period of 2 min) activate principally the most conductive flow paths over the field, mostly located in the conduits network. These frequencies permit to better characterize the distance between each borehole and the most conductive karstic structures. Therefore, they allow a better interpretation of a degree of connectivity between boreholes, through the network of karstic conduits. Lower frequency signals (here a period of 5 min) activate, at our field scale, both conductive and also less conductive structures. Therefore, these frequencies permit to better characterize the existence and the localization of networks of fractures and fissures around the boreholes.

5. Discussion

In Fischer et al. (2018b), the authors have described a qualitative method for interpreting inter-well connectivity from the responses to harmonic pumping tests in karstic aquifers, by categorizing the extracted oscillatory responses in three types (conduit, dual, matrix connectivity). In this work the method is further developed through the integration of a quantitative interpretation with an inverse algorithm, the CADI method, that can handle a large number of measured data simultaneously and generate complex distributions of properties. The integrated approach permits to produce spatial distributions of amplitude and phase offset responses consistent with those studied by Fischer et al. (2018b).

The comparison of inverted conduit networks from periods of 2 min and 5 min indicates different pumping frequency generate different flow fields. A higher frequency will permit to better characterize the flows in highly conductive structures, and the conduit connectivity at a field scale. This finding is consistent with the work of Rabinovich et al. (2015) where the authors show that the flow paths tend to follow the most conductive media especially at lower period. On the contrary, a lower frequency will permit to better characterize the flows in less conductive structures, and thus the localization of networks of smaller conduits, fractures and fissures. Each frequency of signal permits to generate responses holding different and complementary information on karst structures. There is therefore, according to our results, no 'best' choice of period for the characterization of a karstic field. This choice should be made accordingly to what structures one would most likely characterize. The important impact of the period of the pumping signal on the ratio of conduit/matrix flows has already been highlighted with a simplified study case in Fischer et al. (2018b), but our new work also shows a more complex role of the different structures (conduit, fracture, fissure, matrix) on the generation of different flow fields with the different frequencies of pumping signal.

To summarize, if at a regional scale one would imagine that the difference between a lower and a higher frequency pumping would principally concern the zone of influence in the aquifer (with a larger zone for a lower frequency), this difference implies more specific behaviors at a smaller site scale. In fact, at this scale, higher frequencies mobilize water essentially in the most conductive structures, while lower frequencies permit to reach also a mobilization of water in the less conductive structures and media. On another hand oscillatory responses do not provide precise information on the conductivity and specific storage values in the conduits and the matrix very far from the pumping point. However the same observation was made on the same field for the inversion of steady-state responses to eight pumping tests at constant rates in Fischer et al. (2017c). This previous article also showed that steady state data to constant rate pumping (comparable to an infinitely high oscillatory period) were more sensitive to fracture flows

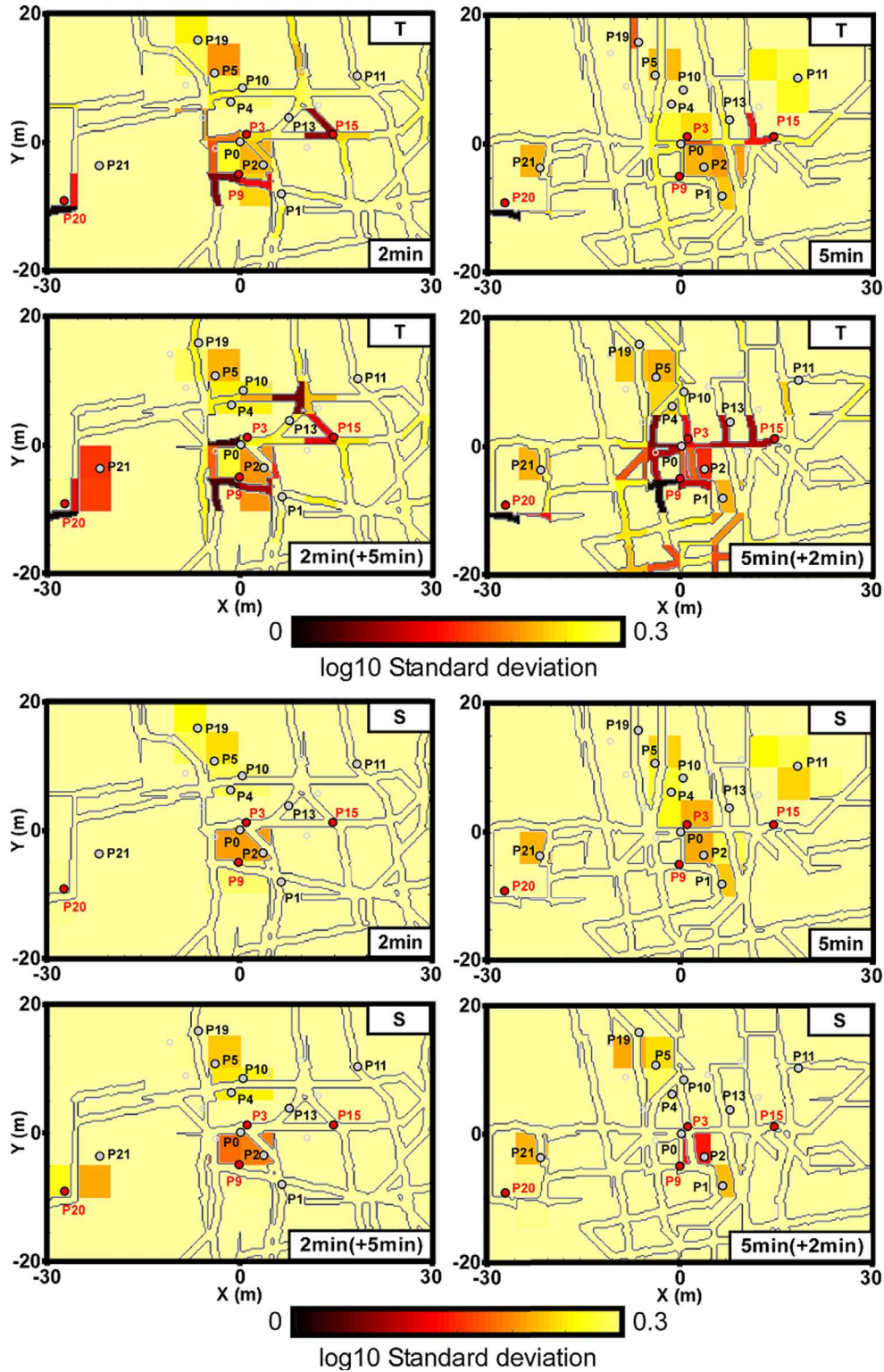


Fig. 13. Transmissivity (T) and storativity (S) standard deviation values of the results found for separate inversions of the 2 min and 5 min responses, and joint inversions started with the 2 min result (2 min (+ 5 min)), and with the 5 min result (5 min (+ 2 min)).

and required a dense inverted network to be reproduced, which is in agreement with our observation for a higher period of signal in this work. Furthermore, the uncertainty analysis from this previous article indicated that the constant rate data, as the data for a period of 5 min in the pumping signal in this work, permitted to better charac-

terize the flow structures in areas where we had measurement points. The inverted networks in the results from our new work and the ones in Fischer et al. (2017c) both reproduce the established connectivity of the Terrieu field (Jazayeri Noushabadi, 2009; Dausse, 2015), but in the case of the constant rate data, the inversions were led with responses to an

investigation of eight pumping points instead of four with the harmonic investigation.

The results of our new work, associated to the previous ones in Fischer et al. (2017c), tend to indicate that the steady state responses to a constant pumping rate would blur the connectivity associated to the most conductive conduits among the field, while better characterizing the karstic structures of different scales in the areas around the pumping wells. Therefore it would require several well distributed pumping points in order to characterize the whole karstic network. Responses to an oscillatory pumping rate, on contrary, allow an already good characterization of the karstic connectivity from a unique pumping in a borehole in the karstic network (as P15 in this work).

6. Conclusion

In this work we have extended the qualitative method presented in Fischer et al. (2018b), for interpreting a karstic network connectivity from the hydraulic responses to harmonic pumping tests, to a quantitative analysis by combining these responses with an inversion algorithm. The integrated approach is able to deal with a large set of data simultaneously and to construct structurally contrasted distributions of hydraulic properties conditioned to the measured tomographic harmonic pumping responses.

Our results show that tomographic harmonic pumping tests performed with different signal frequencies led to a characterization of different structures of the karstic network. Higher frequency signals tend to assist in interpreting a degree of connectivity between each borehole of the field and the most conductive structures, while lower frequency

signals are more useful in the localization of less conductive features, such as small fractures and fissures.

The CADI method, as imaging tool, shows limitations in its ability to represent complex structures of different aperture simultaneously, as already noticed in Fischer et al. (2017c), which can partly explain the less good results of the joint inversions compared to the separate inversions in this work. However we believe that the combination of the CADI method with tomographic harmonic pumping tests appears as a promising methodology for a quantitative characterization of the hydraulic properties and the hydraulic connectivity in karstic aquifers.

Conflict of interest

None

Acknowledgments

We would like to thank Todd C. Rasmussen and two anonymous reviewers for their relevant comments and propositions which permitted to significantly improve the quality of this article. We thank the Normandy region for providing financial support for the Ph.D. of Pierre Fischer. We also thank Michel Simon for his precious contribution during the field investigation. The hydraulic data collection and hydrogeological characterization of the Terrieu experimental site are part of work of the French research network SNO Karst.

Appendix A

Table A1

Interpreted amplitude and phase offset responses for both signal periods (2 min, 5 min) of each pumping point (P3, P9, P15, P20) and distances between the measurement points and the pumping points. A dash represents a negligible oscillatory response (amplitude lower than 1 mm). Dashes represent a negligible oscillatory response, considered as null for the inversion.

	P3					P9				
	Distance (m)	Amplitude (cm)		Phase offset (°)		Distance (m)	Amplitude (cm)		Phase offset (°)	
		2 min	5 min	2 min	5 min		2 min	5 min	2 min	5 min
P0	1.4	2.30	2.40	27	11	5.3	0.20	0.11	0	24
P1	10.6	0.17	0.17	−53	145	7.2	0.33	0.14	0	−110
P2	5.4	–	–	–	–	4.2	0.26	0.66	109	88
P4	5.7	–	–	–	–	11.4	–	–	–	–
P5	11	0.18	0.20	151	−126	16.4	0.22	0.23	143	82
P9	6.3	0.20	0.47	104	71	–	18	31	39	54
P10	7.5	–	–	–	–	13.7	–	–	–	–
P11	19.4	0.19	0.18	−20	41	24	0.17	0.12	0	−97
P13	6.9	0.33	0.22	40	85	11.7	0.18	0.14	143	0
P15	13.3	–	–	–	–	15.9	0.26	0.16	112	166
P19	16.7	0.13	0.11	−172	−25	22	0.27	0.11	110	−134
P20	30.1	–	–	–	–	27.3	–	–	–	–
P21	23.3	–	–	–	–	21.6	–	–	–	–
	P15					P20				
	Distance (m)	Amplitude (cm)		Phase offset (°)		Distance (m)	Amplitude (cm)		Phase offset (°)	
		2 min	5 min	2 min	5 min		2 min	5 min	2 min	5 min
P0	14.4	0.27	0.42	41	52	28.9	0.38	0.24	−7	68
P1	12.4	0.36	0.18	−39	154	33.6	0.37	0.15	6	−76
P2	11.7	0.29	0.39	49	46	31.5	0.24	0.26	3	64
P4	16.6	–	–	–	–	30.1	–	–	–	–
P5	20.7	0.18	0.19	166	79	30.9	0.25	0.14	17	135
P9	15.9	0.22	0.28	−22	82	27.3	0.26	0.20	14	114
P10	15.8	–	–	–	–	32.9	–	–	–	–
P11	9.9	0.37	0.49	13	34	49.5	0.33	0.26	11	68
P13	7.3	0.52	0.23	−76	−168	37.1	0.23	0.15	5	−9
P15	–	1.40	1.30	−22	8	43	0.46	0.35	3	56
P19	25.6	0.14	0.26	14	66	32.7	0.14	0.15	−78	84
P20	43	0.27	0.34	64	53	–	93	153	41	3
P21	36.5	0.33	0.22	−180	89	7.9	0.24	0.21	−8	98

References

- Abusaada, M., Sauter, M., 2013. Studying the flow dynamics of a karst aquifer system with an equivalent porous medium model. *Groundwater* 51 (4), 641–650.
- Ackerer, P., Delay, F., 2010. Inversion of a set of well-test interferences in a fractured limestone aquifer by using an automatic downscaling parameterization technique. *J. Hydrol.* 389, 42–56.
- Bakhos, T., Cardiff, M., Barrash, W., Kitanidis, P.K., 2014. Data processing for oscillatory pumping tests. *J. Hydrol.* 511, 310–319.
- Berg, S.J., Illman, W.A., 2013. Field study of subsurface heterogeneity with steady-state hydraulic tomography. *Groundwater* 51 (1), 29–40.
- Bohling, G.C., Zhan, X., Butler, J.J., Zheng, L., 2002. Steady shape analysis of tomographic pumping tests for characterization of aquifer heterogeneities. *Water Resour. Res.* 38, 1324.
- Bonneau, F., Henrion, V., Caumon, G., Renard, P., Sausse, J., 2013. A methodology for pseudo-genetic stochastic modeling of discrete fracture networks. *Comput. Geosci.* 56, 12–22.
- Borghi, A., Renard, P., Cornaton, F., 2016. Can one identify karst conduit networks geometry and properties from hydraulic and tracer test data? *Adv. Water Resour.* 90, 99–115.
- Butler, J.J., 2005. Hydrogeological methods for estimation of spatial variations in hydraulic conductivity. In: Rubin, Y., Hubbard, S.S. (Eds.), *Hydrogeophysics*. Water Science and Technology Library, 50. Springer, Dordrecht.
- Caers, J., Hoffman, T., 2006. The probability perturbation method: a new look at Bayesian inverse modeling. *Math. Geol.* 38 (1), 81–100.
- Cardiff, M., Barrash, W., Kitanidis, P.K., Malama, B., Revil, A., Straface, S., Rizzo, E., 2009a. A potential-based inversion of unconfined steady-state hydraulic tomography. *Groundwater* 47 (2), 259–270.
- Cardiff, M., Kitanidis, P.K., 2009b. Bayesian inversion for facies detection: an extensible level set framework. *Water Resour. Res.* 45, W10416. <https://doi.org/10.1029/2008WR007675>.
- Cardiff, M., Barrash, W., Kitanidis, P.K., 2013a. Hydraulic conductivity imaging from 3-D transient hydraulic tomography at several pumping/observation densities. *Water Resour. Res.* 49, 7311–7326.
- Cardiff, M., Bakhos, T., Kitanidis, P.K., Barrash, W., 2013b. Aquifer heterogeneity characterization with oscillatory pumping: Sensitivity analysis and imaging potential. *Water Resour. Res.* 49, 5395–5410.
- Cardiff, M., Barrash, W., 2015. Analytical and semi-analytical tools for the design of oscillatory pumping tests. *Groundwater* 53 (6), 896–907.
- Carrera, J., Neuman, S.P., 1986. Estimation of aquifer parameters under transient and steady state conditions: 1. Maximum likelihood method incorporating prior information. *Water Resour. Res.* 22 (2), 199–210.
- Castagna, M., Becker, M.W., Bellin, A., 2011. Joint estimation of transmissivity and storativity in a bedrock fracture. *Water Resour. Res.* 47, W09504. <https://doi.org/10.1029/2010WR009262>.
- Collon, P., Bernasconi, D., Vuilleumier, C., Renard, P., 2017. Statistical metrics for the characterization of karst network geometry and topology. *Geomorphology* 283, 122–142.
- Dausse, A., 2015. Facteurs d'échelle dans la hiérarchisation des écoulements au sein d'un aquifère karstique: analyse multi-échelles des propriétés hydrodynamiques et de transport de l'aquifère de Lez (Ph.D. thesis). Université de Montpellier, French.
- De Rooij, R., Perrochet, P., Graham, W., 2013. From rainfall to spring discharge: coupling conduit flow, subsurface matrix flow and surface flow in karst systems using a discrete-continuum model. *Adv. Water Resour.* 61, 29–41.
- Fischer, P., Jardani, A., Soueid Ahmed, A., Abbas, M., Wang, X., Jourde, H., Lecoq, N., 2017a. Application of large-scale inversion algorithms to hydraulic tomography in an alluvial aquifer. *Groundwater* 55, 208–218.
- Fischer, P., Jardani, A., Lecoq, N., 2017b. A cellular automata-based deterministic inversion algorithm for the characterization of linear structural heterogeneities. *Water Resour. Res.* 53, 2016–2034.
- Fischer, P., Jardani, A., Wang, X., Jourde, H., Lecoq, N., 2017c. Identifying flow networks in a karstified aquifer by application of the cellular automata-based deterministic inversion method (Lsez aquifer, France). *Water Resour. Res.* 53, 10508–10522.
- Fischer, P., Jardani, A., Lecoq, N., 2018a. Hydraulic tomography of discrete networks of conduits and fractures in a karstic aquifer by using a deterministic inversion algorithm. *Adv. Water Resour.* 112, 83–94.
- Fischer, P., Jardani, A., Cardiff, M., Jourde, H., Lecoq, N., 2018b. Hydraulic analysis of harmonic pumping tests in frequency and time domains for identifying the conduits networks in a karstic aquifer. *J. Hydrol.* 559, 1039–1053.
- Ghasemzadeh, R., Hellweger, F., Butscher, C., Padilla, I., Vesper, D., Field, M., Alshawabkeh, A., 2012. Review: groundwater flow and transport modeling of karst aquifers, with particular reference to the North Coast Limestone aquifer system of Puerto Rico. *Hydrogeol. J.* 20, 1441–1461.
- Grimstad, A.-A., Mannseth, T., Naevdal, G., Urkedal, H., 2003. Adaptive multiscale permeability estimation. *Comput. Geosci.* 7, 1–25.
- Guilinan, E., Becker, M.W., 2015. Measuring well hydraulic connectivity in fractured bedrock using periodic slug tests. *J. Hydrol.* 521, 100–107.
- Hale, D. (Ed.), 2009. Structure-oriented smoothing and semblance, Center for Wave Phenomena CWP 635. [Available at <http://inside.mines.edu/~dhale/papers/Hale09StructureOrientedSmoothingAndSemblance.pdf>].
- Hao, Y., Yeh, T.J., Xiang, J., Illman, W.A., Ando, K., Hsu, K., Lee, C., 2008. Hydraulic tomography for detecting fracture zone connectivity. *Ground Water* 46, 183–192.
- Hartmann, A., Goldscheider, N., Wagnen, T., Lange, J., Weiler, M., 2014. Karst water resources in a changing world: review of hydrological modeling approaches. *Rev. Geophys.* 52, 218–242.
- Hoeksema, R.J., Kitanidis, P.K., 1984. An application of the geostatistical approach of the inverse problem in two-dimensional groundwater modeling. *Water Resour. Res.* 20, 1003–1020.
- Illman, W.A., Liu, X., Takeuchi, S., Yeh, T.J., Ando, K., Saegusa, H., 2009. Hydraulic tomography in fractured granite: Mizunami underground research site, Japan. *Water Resour. Res.* 45. <https://doi.org/10.1029/2007WR006715>.
- Illman, W.A., 2014. Hydraulic tomography offers improved imaging of heterogeneity in fractured rocks. *Groundwater* 52, 659–684.
- Jaquet, O., Siegel, P., Klubertanz, G., Benabderrhamane, H., 2004. Stochastic discrete model of karstic networks. *Adv. Water Resour.* 27, 751–760.
- Jazayeri Noushabadi, M.R., 2009. Characterization of relationship between fracture network and flow-path network in fractured and karstic reservoirs: numerical modeling and field investigation (Lez aquifer, Southern France) (Ph.D. thesis). Université de Montpellier.
- Jazayeri Noushabadi, M.R., Jourde, H., Massonnat, G., 2011. Influence of the observation scale on permeability estimation at local and regional scales through well tests in a fractured and karstic aquifer (Lez aquifer, Southern France). *J. Hydrol.* 403, 321–336.
- Jourde, H., Cornaton, F., Pistre, S., Bidaux, P., 2002. Flow behavior in a dual fracture network. *J. Hydrol.* 266, 99–119.
- Jourde, H., et al., 2011. The MEDYCYSS observatory, a multi scale observatory of flood dynamics and hydrodynamics in karst (Mediterranean border Southern France). In: Lambrakis, N., Stournaras, G., Katsanou, K. (Eds.), *Advances in the Research of Aquatic Environment*. Environmental Earth Sciences. Springer, Berlin, Heidelberg.
- Kitanidis, P.K., 1995. Quasi-linear geostatistical theory for inverting. *Water Resour. Res.* 31 (10), 2411–2419.
- Kovacs, A., 2003. Estimation of conduits network geometry of a karst aquifer by the means of groundwater flow modeling (Bure, Switzerland). *Bol. Geol. y Min.* 114 (2), 183–192.
- Larocque, M., Banton, O., Ackerer, P., Razack, M., 1999. Determining karst transmissivities with inverse modeling and an equivalent porous media. *Ground Water* 37 (6), 897–903.
- Lavenue, M., de Marsily, G., 2001. Three-dimensional interference test interpretation in a fractured aquifer using the pilot point inverse method. *Water Resour. Res.* 37 (11), 2659–2675.
- Le Coz, M., Bodin, J., Renard, P., 2017. On the use of multiple-point statistics to improve groundwater flow modeling in karst aquifers: a case study from the hydrogeological experimental site of Poitiers, France. *J. Hydrol.* 545, 109–119.
- Lee, J., Kitanidis, P.K., 2013. Bayesian inversion with total variation prior for discrete geologic structure identification. *Water Resour. Res.* 49, 7658–7669.
- Li, Z.Y., Zhao, J.H., Qiao, X.H., Zhang, Y.X., 2014. An automated approach for conditioning discrete fracture network modelling to in situ measurements. *Aust. J. Earth Sci.* 61, 755–763.
- Liedl, R., Sauter, M., Hockinghaus, D., Clemens, T., Teutsch, G., 2003. Simulation of the development of karst aquifers using a coupled continuum pipe flow model. *Water Resour. Res.* 39, 1057.
- Lochbühler, T., Vrugt, J.A., Sadegh, M., Linde, N., 2015. Summary statistics from training images as prior information in probabilistic inversion. *Geophys. J. Int.* 201, 157–171.
- Lu, Z., Robinson, B.A., 2006. Parameter identification using the level set method. *Geophys. Res. Lett.* 33, L06404. <https://doi.org/10.1029/2005GL025541>.
- Maineult, A., Strobach, E., Renner, J., 2008. Self-potential signals induced by periodic pumping tests. *J. Geophys. Res.* 113, B01203. <https://doi.org/10.1029/2007JB005193>.
- Ni, C., Yeh, T.J., 2008. Stochastic inversion of pneumatic cross-hole tests and barometric pressure fluctuations in heterogeneous unsaturated formations. *Adv. Water Resour.* 31, 1708–1718.
- Pardo-Iguzquiza, E., Dowd, P.A., Xu, C., Duran-Valsero, J.J., 2012. Stochastic simulation of karst conduit networks. *Adv. Water Resour.* 35, 141–150.
- Rabinovich, A., Barrash, W., Cardiff, M., Hochstetler, D.L., Bakhos, T., Dagan, G., Kitanidis, P.K., 2015. Frequency dependent hydraulic properties estimated from oscillatory pumping tests in an unconfined aquifer. *J. Hydrol.* 531, 2–16.
- Rasmussen, T.C., Haborak, K.G., Young, M.H., 2003. Estimating aquifer hydraulic properties using sinusoidal pumping at the Savannah River site, South California, USA. *Hydrogeol. J.* 11, 466–482.
- Renner, J., Messar, M., 2006. Periodic pumping tests. *Geophys. J. Int.* 167, 479–493.
- Saller, A.P., Ronayne, M.J., Long, A.J., 2013. Comparison of a karst groundwater model with and without discrete conduit flow. *Hydrogeol. J.* 21, 1555–1566.
- Sharmeen, R., Illman, W.A., Berg, S.J., Yeh, T.J., Park, Y., Sudicky, E.A., Ando, K., 2012. Transient hydraulic tomography in a fractured dolostone: laboratory rock block experiments. *Water Resour. Res.* 48. <https://doi.org/10.1029/2012WR012216>.
- Soueid Ahmed, A., Zhou, J., Jardani, A., Revil, A., Dupont, J.P., 2015. Image-guided inversion in steady-state hydraulic tomography. *Adv. Water Resour.* 82, 83–97.
- Soueid Ahmed, A., Jardani, A., Revil, A., Dupont, J.P., 2016. Joint inversion of hydraulic head and self-potential data associated with harmonic pumping tests. *Water Resour. Res.* 52 (9), 6769–6791.
- Sun, A.Y., Lu, J., Hovorka, S., 2015. A harmonic pulse testing method for leakage detection in deep subsurface storage formations. *Water Resour. Res.* 51, 4263–4281.
- Tarantola, A., Valette, B., 1982. Generalized nonlinear inverse problems solved using the least squares criterion. *Rev. Geophys. Space Phys.* 20 (2), 219–232.
- Teutsch, G., 1993. An extended double-porosity concept as a practical modeling approach for a karstified terrain. *Hydrogeol. Process. Karst Terranes* 207, 281–292.
- Von Neumann, J., Burks, A.W., 1966. *Theory of Self-Reproducing Automata*. University of Illinois Press ISBN 0-598-37798-0.
- Wang, X., Jardani, A., Jourde, H., Lonergan, L., Cosgrove, J., Gosselin, O., Massonnat, G., 2016. Characterisation of the transmissivity field of a fractured and karstic aquifer, Southern France. *Adv. Water Resour.* 87, 106–121.
- Wang, X., Jardani, A., Jourde, H., 2017. A hybrid inverse method for hydraulic tomography in fractured and karstic media. *J. Hydrol.* 551, 29–46.

- White, W.B., 2002. Karst hydrology: recent developments and open questions. *Eng. Geol.* 65, 85–105.
- Yeh, T.J., Liu, S., 2000. Hydraulic tomography: development of a new aquifer test method. *Water Resour. Res.* 36, 2095–2105.
- Yeh, T.J., Lee, C., 2007. Time to change the way we collect and analyze data for aquifer characterization. *Groundwater* 45 (2), 116–118.
- Zha, Y., Yeh, T.J., Illman, W.A., Tanaka, T., Bruines, P., Onoe, H., Saegusa, H., 2015. What does hydraulic tomography tell us about fractured geological media? A field study and synthetic experiments. *J. Hydrol.* 531 (1), 17–30.
- Zha, Y., Yeh, T.J., Illman, W.A., Tanaka, T., Bruines, P., Onoe, H., Saegusa, H., Mao, D., Takeuchi, S., Wen, J.C., 2016. An application of hydraulic tomography to a large-scale fractured granite site, Mizunami, Japan. *Groundwater* 54, 793–804.
- Zha, Y., Yeh, T.J., Illman, W.A., Onoe, H., Mok, C.M.W., Wen, J.C., Huang, S.Y., Wang, W., 2017. Incorporating geologic information into hydraulic tomography: a general framework based on geostatistical approach. *Water Resour. Res.* 53. <https://doi.org/10.1002/2016WR019185>.
- Zhou, Y., Lim, D., Cupola, F., Cardiff, M., 2016. Aquifer imaging with pressure waves - Evaluation of low-impact characterization through sandbox experiments. *Water Resour. Res.* 52, 2141–2156.
- Zhu, J., Yeh, T.J., 2005. Characterization of aquifer heterogeneity using transient hydraulic tomography. *Water Resour. Res.* 41, W07028.

Résumé

Ce manuscrit de thèse présente une nouvelle approche pour caractériser qualitativement et quantitativement la localisation et les propriétés des structures dans un aquifère fracturé et karstique à l'échelle décamétrique. Cette approche est basée sur une tomographie hydraulique menée à partir de réponses à une investigation de pompages et interprétée avec des méthodes d'inversions adaptées à la complexité des systèmes karstiques. L'approche est appliquée sur un site karstique d'étude expérimental en France, une première fois avec des signaux de pompage constants, et une deuxième fois avec des signaux de pompage harmoniques. Dans les deux cas, l'investigation a fourni des réponses de niveaux d'eau de nappe mesurés pendant des pompages alternés à différentes positions. L'interprétation quantitative de ces jeux de réponses consiste à les reproduire par un modèle avec un champ de propriété réaliste adéquat généré par inversion. Les méthodes d'inversions proposées dans ce manuscrit permettent de reconstruire un champ de propriétés hydrauliques réaliste en représentant les structures karstiques soit par un réseau généré par automates cellulaires, soit par un réseau discrétisé. Les résultats d'interprétations obtenus sur le site d'étude expérimental permettent d'imager les structures karstiques sur une carte et de « lire » leur localisation. De plus, les résultats obtenus avec les réponses à des pompages harmoniques tendent à montrer le rôle de la fréquence du signal sur les informations portées par les réponses. En effet, les fréquences plus élevées caractérisent mieux les structures les plus conductrices, alors que les fréquences plus faibles mobilisent des écoulements également dans des structures karstiques moins conductrices.

Mots-clés : Caractérisation hydrogéologique, Aquifère karstique, Tomographie hydraulique, Pompage harmonique, Modélisation, Problème inverse.

Abstract

This thesis manuscript presents a novel approach to characterize qualitatively and quantitatively the structures localization and properties in a fractured and karstic aquifer at a decametric scale. This approach relies on a hydraulic tomography led from responses to a pumping investigation and interpreted with inversion methods adapted to the complexity of karstic systems. The approach is applied on a karstic experimental study site in France, a first time with constant pumping signals, and a second time with harmonic pumping signals. In both applications, the investigation resulted in groundwater level responses measured during alternated pumping tests at different locations. The quantitative interpretation of these sets of responses consists in reproducing these responses through a model with an adequate realistic property field generated by inversion. The inversion methods proposed in this manuscript permit to reconstruct a realistic hydraulic property field by representing the karstic structures either through a network generated by cellular automata, or through a discretized network. The interpretation results obtained on the experimental study site permit to image the karstic structures on a map and to 'read' their localization. Furthermore, the results obtained with the responses to harmonic pumping tests tend to show the role of the signal frequency on the information carried by the responses. In fact, higher frequencies better characterize the most conductive structures, while lower frequencies mobilize flows also in less conductive karstic structures.

Keywords: Hydrogeological characterization, Karst aquifer, Hydraulic tomography, Harmonic pumping, Modeling, Inverse problem.