

Titre: Estimation des propriétés hydrauliques d'un aquifère par simulation séquentielle gaussienne bayésienne
Title: sequential gaussian bayesian estimation of hydraulic properties of an aquifer by simulation

Auteur: Camille Dubreuil-Boisclair
Author:

Date: 2009

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Dubreuil-Boisclair, C. (2009). Estimation des propriétés hydrauliques d'un aquifère par simulation séquentielle gaussienne bayésienne [Mémoire de maîtrise, École Polytechnique de Montréal]. PolyPublie.
Citation: <https://publications.polymtl.ca/8418/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/8418/>
PolyPublie URL:

Directeurs de recherche: Denis Marcotte, Michel Chouteau, & Erwan Gloaguen
Advisors:

Programme: Non spécifié
Program:

UNIVERSITÉ DE MONTRÉAL

ESTIMATION DES PROPRIÉTÉS HYDRAULIQUES D'UN AQUIFÈRE PAR
SIMULATION SÉQUENTIELLE GAUSSIENNE BAYÉSIENNE

CAMILLE DUBREUIL-BOISCLAIR
DÉPARTEMENT DES GÉNIES CIVIL, GÉOLOGIQUE ET DES MINES
ÉCOLE POLYTECHNIQUE DE MONTRÉAL

MÉMOIRE PRÉSENTÉ EN VUE DE L'OBTENTION
DU DIPLÔME DE MAÎTRISE ÈS SCIENCES APPLIQUÉES
(GÉNIE MINÉRAL)
AVRIL 2009



Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 978-0-494-53900-2
Our file Notre référence
ISBN: 978-0-494-53900-2

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

UNIVERSITÉ DE MONTRÉAL

ÉCOLE POLYTECHNIQUE DE MONTRÉAL

Ce mémoire intitulé:

ESTIMATION DES PROPRIÉTÉS HYDRAULIQUES D'UN AQUIFÈRE PAR
SIMULATION SÉQUENTIELLE GAUSSIENNE BAYÉSIENNE

présenté par: DUBREUIL-BOISCLAIR Camille
en vue de l'obtention du diplôme de: Maîtrise ès sciences appliquées
a été dûment accepté par le jury d'examen constitué de:

M. CHAPUIS Robert P., D.Sc.A., président

M. MARCOTTE Denis, Ph.D., membre et directeur de recherche

M. CHOUTEAU Michel, Ph.D., membre et codirecteur de recherche

M. GLOAGUEN Erwan, Ph.D., membre et codirecteur de recherche

M. GIROUX Bernard, Ph.D., membre

REMERCIEMENTS

Je tiens d'abord à remercier Denis Marcotte pour son encadrement juste comme il faut, sa compétence, sa compréhension et sa grande confiance.

Je remercie également Erwan Gloaguen pour son dévouement, ses innombrables explications, pour ses encouragements mais surtout, pour son indéfectible soutien tant académique que personnel.

Merci à Michel Chouteau pour m'avoir fait découvrir la géophysique et le monde de la recherche.

Un grand merci à Bernard Giroux pour son inestimable support en informatique et pour m'avoir toujours accueilli dans son bureau avec le sourire même si c'était la troisième fois de la journée !

Un grand merci également à mon amie et colocataire de bureau au tapis brun et aux murs couleur peau, Martine Rivest, pour son précieux soutien dans toutes les sphères de la vie quotidienne.

Vielen Dank mein Andreas, dass Du mich jeden Morgen mit einer schönen Nachricht aufweckst.

Je remercie mes parents pour leur amour et ma petite soeur Emma, simplement d'être là.

RÉSUMÉ

L'estimation de la conductivité hydraulique d'un aquifère est une étape cruciale de la caractérisation hydrogéologique d'un site. En effet, ce paramètre façonne le parcours de l'eau souterraine et par conséquent, celui de plusieurs types de contaminants. Ainsi, l'étude d'un panache de contamination passe obligatoirement par une bonne estimation de la conductivité hydraulique à l'échelle appropriée. Toutefois, ce paramètre hydraulique présente généralement une grande variabilité spatiale que les méthodes hydrogéologiques peinent à estimer sur une grille fine comprise entre deux forages.

Ainsi, ce projet vise à estimer la continuité spatiale de la conductivité hydraulique, dans des milieux hétérogènes, à une échelle décimétrique comprise entre deux forages, par intégration de données radar entre forages ainsi que de conductivités électriques, permittivités et conductivités hydrauliques le long de forages. La méthode s'appuie sur les relations *in situ* existantes entre les propriétés physiques, mesurées par géophysique et les propriétés hydrauliques du milieu étudié. Ces différents types de données sont combinés à l'aide d'une approche stochastique bayésienne permettant d'inclure les conductivités hydrauliques préalablement connues ainsi que les relations pétrophysiques du milieu afin de produire différents modèles de conductivités hydrauliques. L'analyse d'un grand nombre de simulations permet de quantifier la variabilité des valeurs estimées.

La méthode a d'abord été validée à l'aide d'un modèle synthétique hétérogène de porosité. À partir de ce modèle, des données radar en forage ont été modélisées numériquement, à partir desquelles des tomogrammes de vitesse et d'atténuation ont été générés. L'extraction des données le long de deux forages fictifs a permis de formuler une relation entre la porosité, la vitesse et l'atténuation. Cette rela-

tion est traduite sous forme d'une densité de probabilité à l'aide de la méthode des noyaux. Ensuite, pour chacun des pixels du plan entre les forages, un krigage simple des porosités préalablement connues permet de fournir une première valeur de ce paramètre. Cette valeur est ensuite mise à jour en y incluant l'information provenant des données géophysiques, par le biais de la vraisemblance, afin de restreindre la variabilité de l'estimation *a priori*. Cette méthode est appelée simulation séquentielle gaussienne selon une approche bayésienne.

Les résultats issus du modèle synthétique montrent que l'approche permet d'améliorer substantiellement l'estimation de la continuité spatiale de la porosité et la texture des estimés par rapport aux méthodes courantes comme le krigage simple, l'utilisation de la formule de Topp ou la simulation séquentielle gaussienne (SGS). De plus, la variance des estimés est grandement inférieure à celle obtenue par SGS ou par simulations sans *a priori*. Notons toutefois que ces résultats sont similaires à ceux obtenus par cosimulation dans le cas où la relation entre les paramètres physiques et la porosité est plutôt linéaire.

Ces résultats étant concluants, la méthode a été testée avec un cas réel. Le site à l'étude se trouve dans la municipalité de Saint-Lambert-de-Lauzon, dans une ancienne décharge de déchets d'origines diverses. La simulation bayésienne a été appliquée à des données acquises en forages provenant d'un piézocône, muni d'une sonde de résistivité et d'une sonde capacitive, d'essais de perméabilité *in situ* à niveaux multiples et d'une acquisition tomographique radar. Cette approche a permis d'obtenir des réalisations de la conductivité hydraulique du site, sur une grille fine (~ 15 cm) comprise entre deux forages distants de 9 m. Les caractéristiques des différentes lithologies possibles identifiées grâce aux simulations concordent étroitement avec le modèle géologique de ce site, connu à ce jour.

ABSTRACT

Hydraulic conductivity is generally considered as the most important parameter of a hydrogeological or environmental site characterization. Indeed, it shapes the course of groundwater flow and consequently that of many contaminants. However, hydraulic conductivity generally presents a large spatial variability which classical hydrogeological methods (e.g. pumping test, tracer test,...) fails to identify at a decimetric scale between two boreholes.

Consequently, this project aims to estimate the hydraulic conductivity of heterogeneous medias, at decimetric scale between two boreholes, by integrating ground penetrating radar (GPR) tomographic data, electrical conductivity, permittivity as well as hydraulic conductivity logs. To achieve this, the sequential gaussian simulation algorithm, in a bayesian framework, is proposed. This method is based on the determination of an *in situ* relationship between the physical and the hydraulic properties of the studied area. This relationship is then applied to the GPR tomographies to « update » a prior probability given by the previous known hydraulic conductivity data. Moreover, the stochastic framework of this method allows to produce many realizations of the hydraulic conductivity field, which can be analyzed to quantify the variability of the estimations.

The proposed method is first tested with a synthetic heterogeneous porosity model. From this model, GPR velocity and attenuation tomographic data are generated (Giroux et Chouteau, 2008a). The data along two fictive boreholes allows to formulate a relationship between porosity, GPR velocity and attenuation which is converted into a 3D probability density function using the kernel probability density function (*pdf*) estimator. For each pixel included in the plane between the boreholes, simple kriging of the previously known porosity data is used to estimate

the prior probability. This latter is then updated by the information contained in the geophysical data, through the likelihood function.

The results obtained from the synthetic model show that this approach considerably improves the estimation of the spatial continuity of the porosity compared to common estimation methods like simple kriging, Topp's formula or standard sequential gaussian simulation. Furthermore, the variance of the estimations is dramatically smaller than that obtained by standard SGS or by bayesian simulation without a prior estimate. Nevertheless, in the case where the relationship between the variables is quasi-linear, such as this one, the results are very comparable with the one obtained from cosimulation.

The previous results being conclusive, the method is tested with a real data set. The studied area is located in the municipality of St-Lambert-de-Lauzon, Quebec, on an old landfill where wastes from many sources were buried. The bayesian simulation algorithm was applied to data obtained from 1) a GPR tomographic acquisition, 2) a CPTu probe equipped with a resistivity and a capacitance probe and 3) a multi-level slug test. This approach allows to generate different hydraulic conductivity realizations at a decimetric scale, on a grid between the two boreholes. The different lithologies identified by the simulation are in good agreement with the geological model of the studied area, confirming the validity of the proposed approach.

TABLE DES MATIÈRES

REMERCIEMENTS	iv
RÉSUMÉ	v
ABSTRACT	vii
TABLE DES MATIÈRES	ix
LISTE DES FIGURES	xii
LISTE DES NOTATIONS ET DES SYMBOLES	xvi
LISTE DES TABLEAUX	xviii
CHAPITRE 1 INTRODUCTION	1
1.1 Objectif	2
1.2 Plan de travail	3
CHAPITRE 2 REVUE DE LITTÉRATURE	4
2.1 Estimation directe	4
2.2 Inversion conjointe et inversion contrainte	6
2.2.1 Approche par zonation	7
2.2.2 Approche structurale	8
2.2.3 Approche stochastique	9
2.3 Estimation par intégration	11
2.3.1 Analyse par classification non-hiéarchique	11
2.3.2 Approche bayésienne	13
CHAPITRE 3 REVUE THÉORIQUE	16
3.1 Notions de base d'hydrogéologie	16

3.1.1	Loi de Darcy	16
3.1.2	Écoulement en milieu saturé	18
3.2	Relations entre les propriétés hydrauliques et électriques du sol	19
3.2.1	Propriétés hydrauliques	21
3.2.1.1	Porosité et surface spécifique	21
3.2.1.2	Conductivité hydraulique et perméabilité	24
3.2.2	Propriétés électriques	27
3.2.3	Modèles de conductivité électrique	29
3.2.4	Modèles de permittivité	32
3.3	Géophysique en forage	35
3.3.1	Géoradar (GPR)	35
3.3.1.1	Ondes électromagnétiques et sols - Survol	36
3.3.1.2	Appareillage et modes d'acquisition	39
3.3.1.3	Tomographie	41
3.3.2	Piezocône multi-capteurs (CPTu)	46
3.3.3	Essai de perméabilité <i>in situ</i> à niveaux multiples	48
3.3.3.1	Acquisition et traitement des données	48
3.4	Outils statistiques	50
3.4.1	Géostatistique	50
3.4.1.1	Variogramme	50
3.4.1.2	Krigeage simple	54
3.4.1.3	Cokrigeage simple	56
3.4.2	Théorème de Bayes	57
3.4.2.1	Densité de probabilité - Estimateur de noyau	59
3.4.3	Simulations conditionnelles	62
3.4.3.1	Simulation séquentielle gaussienne	63
3.4.3.2	Cosimulation	64
3.4.3.3	Simulation séquentielle gaussienne bayésienne	65

CHAPITRE 4 MÉTHODOLOGIE	67
4.1 Approche proposée	67
4.2 Déroulement de la simulation	69
4.3 Validation de la méthode	71
CHAPITRE 5 RÉSULTATS - MODÈLE SYNTHÉTIQUE	73
5.1 Données et paramètres de pré-simulation	73
5.1.1 Modèles et données géophysiques simulés	73
5.1.2 Densité de probabilité conjointe	77
5.1.3 Modèle de covariance de la porosité	79
5.2 Simulations	80
5.3 Comparaison des simulations	83
5.4 Comparaison des variances	87
5.5 Comparaison des coefficients de corrélation	89
CHAPITRE 6 RÉSULTATS - SITE DE ST-LAMBERT-DE-LAUZON . .	91
6.1 Description du site à l'étude	91
6.2 Données et paramètres de pré-simulation	93
6.2.1 Données géophysiques	94
6.2.2 Données hydrogéologiques	95
6.2.3 Modèle de covariance des \log_{10} de la conductivité hydraulique	97
6.2.4 Densité de probabilité conjointe	98
6.3 Simulations	99
CHAPITRE 7 DISCUSSION	104
CHAPITRE 8 CONCLUSION	109
RÉFÉRENCES	110

LISTE DES FIGURES

FIG. 3.1	Variation de la porosité d'un sol selon la pourcentage de particules fines qu'il contient pour des processus de compactage identiques (Tiré de Morin, 2006)	23
FIG. 3.2	Variations des parties réelles et imaginaires de la permittivité en fonction de la fréquence normalisée à la fréquence de transition ($\sim 3 \times 10^4$ Hz) (Adapté de Grolier et al., 1991) . . .	33
FIG. 3.3	Propagation d'une onde électromagnétique dans un milieu à faibles pertes électriques : la forme du signal reste constante et l'amplitude décroît de manière exponentielle (Adapté de Annan, 2005)	37
FIG. 3.4	Acquisition en mode réflexion en (a) et acquisition en mode transmission en (b) Tx : émetteur, Rx : récepteur . . .	40
FIG. 3.5	Schéma d'une acquisition en tomographie radar. Tx : émetteur à la position x_i , Rx : récepteur à la position x_j . . .	41
FIG. 3.6	Schéma des pointés du temps et de l'amplitude dite « pic à pic » d'une onde provenant d'un géoradar en forage . . .	42
FIG. 3.7	Illustration schématique d'un trajet de rai à travers un milieu hétérogène . . .	43
FIG. 3.8	Schématisation d'un essai <i>in situ</i> à niveaux multiples . . .	49
FIG. 3.9	Variogramme expérimental (x), modèle ajusté (trait) et paramètres géostatistiques . . .	52
FIG. 4.1	Résumé schématique de la méthode proposée (Inspiré de Gastaldi et al., 1998) . . .	69
FIG. 5.1	Modèle de porosité compris entre deux forages . . .	74
FIG. 5.2	Histogramme du modèle de porosité . . .	74

FIG. 5.3	Modélisation (a) des vitesses (m/ns) et (b) des atténuations (np/m) radar obtenues à partir du modèle synthétique de porosité	75
FIG. 5.4	Porosités extraites le long des forages	76
FIG. 5.5	Vitesses (a) et atténuations (b) extraites le long des forages	77
FIG. 5.6	Graphique des vitesses (m/ns) (a) et des atténuations (np/m) (b) le long des forages selon la porosité	78
FIG. 5.7	Densité de probabilité conjointe porosité-vitesse-atténuation obtenue par la méthode des noyaux	79
FIG. 5.8	Ajustement d'un modèle exponentiel sur les observations de porosités le long des forages	80
FIG. 5.9	Modèle de porosité (a), exemples de simulations séquentielles gaussiennes par approche bayésienne (b), (c) et (d)	81
FIG. 5.10	Histogrammes du modèle en (a) et des simulations de la figure 5.9 dans le même ordre	82
FIG. 5.11	Modèle (a) et porosité calculée par la méthode séquentielle bayésienne (moyenne de 100 simulations) (b), par la formule de Topp et al. (1980) (c) et par krigeage simple (d)	84
FIG. 5.12	Porosité calculée par la méthode séquentielle bayésienne (a), la méthode séquentielle gaussienne (classique) (b), la méthode bayésienne sans <i>a priori</i> (c) et la cosimulation (d)	86
FIG. 5.13	Variance de la porosité calculée à partir de cent simulations pour la méthode bayésienne (a), la méthode séquentielle gaussienne (classique) (b), la méthode bayésienne sans <i>a priori</i> (c) et la cosimulation (d)	88

FIG. 5.14	Coefficients de corrélation des cent réalisations pour la méthode bayésienne en rouge, la cosimulation en jaune, la méthode séquentielle gaussienne classique en bleu et bayésienne sans <i>a priori</i> en vert	89
FIG. 6.1	Photographie de la position des puits sur le site à l'étude.	92
FIG. 6.2	Position des puits sur le site à l'étude.	93
FIG. 6.3	(a) Vitesses (m/ns) et (b) atténuations (np/m) radar obtenues entre les puits P19 et P21	94
FIG. 6.4	(a) Constante diélectrique et (b) conductivité électrique (S/m) calculées à partir des tomogrammes de vitesse et d'atténuation	95
FIG. 6.5	Profil du logarithme (base 10) des conductivités hydrauliques, obtenus par essai de perméabilité <i>in situ</i> à niveaux multiples, pour le puits P17	96
FIG. 6.6	Ajustement d'un modèle exponentiel sur le \log_{10} de la conductivité hydraulique le long de P17	97
FIG. 6.7	(a) Graphique de la constante diélectrique et (b) de la conductivité électrique le long de plusieurs forages selon la conductivité hydraulique	98
FIG. 6.8	Densité de probabilité conjointe	99
FIG. 6.9	Réalisations de la conductivité hydraulique (\log_{10}), entre les puits P19 et P21, obtenue par simulation séquentielle gaussienne par approche bayésienne	100
FIG. 6.10	Conductivité hydraulique (\log_{10} m/s) calculée par la méthode (a) de krigeage simple, (b) séquentielle gaussienne et (c) bayésienne sans <i>a priori</i>	101

LISTE DES NOTATIONS ET DES SYMBOLES

K	Conductivité hydraulique [m/s]
h	Charge hydraulique [m], distance entre deux points de mesure, largeur de fenêtre dans un noyau.
θ	Teneur en eau [%]
ϕ	Porosité [%]
S_s	Surface spécifique [m^2/Kg]
μ_f	Viscosité dynamique [Pa·s]
EM	Électromagnétique
GPR	<i>Ground Penetrating Radar</i>
BFT	<i>Borehole flowmeter test</i>
σ	Conductivité électrique [S/m]
σ_w	Conductivité électrique de l'eau [S/m]
ϵ	Permittivité électrique [F/m]
ϵ_0	Permittivité du vide [$8.89 \times 10^{-12} F/m$]
ϵ_r	Permittivité relative [sans dimension]
μ	Perméabilité magnétique [H/m]
μ_0	Perméabilité magnétique du vide [$4\pi \times 10^{-7} H/m$]
ω	Vitesse angulaire [$2\pi f$ rad/s]
f	Fréquence [Hz]
A_r	Amplitude radar à une distance r de la source [mV]
A_0	Amplitude radar à la source [mV]
α	Atténuation de l'onde radar [Np/m]
v	Vitesse de l'onde radar [m/ns]
β	Phase [rad]
c	Vitesse de la lumière [0.3 m/ns]

(*)	Forme complexe
(')	Partie réelle
(")	Partie imaginaire
Q	Débit [m^3/s]
Var, σ^2	Variance
Cov	Covariance
m	Moyenne
E	Espérance

LISTE DES TABLEAUX

CHAPITRE 1

INTRODUCTION

L'eau douce est l'une des plus grandes richesses de la terre. Elle est au coeur de l'écosystème planétaire et est vitale pour l'être humain. Actuellement, environ 70% de la réserve mondiale d'eau douce se retrouve sous forme de glace ou de neige. Le 30% restant est enfoui dans le sous-sol sous forme d'eau souterraine. En comparaison, le volume d'eau douce s'écoulant par les lacs et les rivières représente une portion négligeable de l'ensemble.

La subsurface est donc une zone qui doit être préservée avec soin. Or, il en est souvent autrement. Dans les pays industrialisés, l'abondance d'une eau propre et à faible coût contribue à véhiculer la fausse perception que cette ressource est inépuisable. De plus, dans la plupart des pays, cette zone accueille des quantités faramineuses de déchets de toutes sortes. Ces deux facteurs entraînent l'épuisement graduel et la contamination de cette ressource. Une gestion durable et concertée à l'échelle planétaire de l'eau souterraine est un défi politique majeur qui interpelle également la communauté scientifique. Conséquemment, une meilleure caractérisation de la subsurface est un des domaines d'étude qui s'inscrit dans ce cadre global.

La subsurface est une zone géologique présentant généralement une grande variabilité. En effet, les propriétés hydrauliques d'un sol peuvent varier de plusieurs ordres de grandeur sur de très courtes distances (Lesme et Friedman, 2005). Ces variations façonnent le parcours de l'eau souterraine aussi bien que celui des contaminants.

Dans la pratique courante, la caractérisation hydrogéologique d'un site se fait au moyen d'essais tels que l'essai de pompage, de traceur, *in situ* ou l'analyse granulométrique. Toutefois, ces derniers ne permettent pas de déterminer les propriétés hydrauliques à une échelle assez fine pour prédire avec précision le mode de transport d'un contaminant. Cette lacune des méthodes conventionnelles s'ajoute au fait qu'elles sont généralement coûteuses en temps et en forages.

Ainsi, depuis une dizaine d'années, une nouvelle branche scientifique, appelée *hydrogéophysique*, se penche sur les possibilités d'intégrer la géophysique au processus de caractérisation hydrogéologique de la subsurface. En effet, des méthodes géophysiques comme le géoradar et l'électrique de surface ou en configuration tomographique fournissent une information de haute résolution sur les propriétés physiques d'un milieu. Cependant, la conversion des paramètres physiques en paramètres hydrauliques ainsi que l'intégration et l'unification de l'échelle des différents types de données sont des exemples de difficultés auxquelles cette branche fait face. L'approche statistique constitue l'un des moyens d'aborder ces obstacles.

Ce mémoire tente donc de faire un petit pas vers une meilleure compréhension de la subsurface en testant une approche stochastique intégrant géophysique et hydrogéologie.

1.1 Objectif

L'objectif général de ce projet de maîtrise est de développer une méthode simple et applicable « *in situ* » permettant d'estimer la conductivité hydraulique d'un aquifère à une échelle fine (~ 15 cm) sur un plan compris entre deux forages distants d'au plus 10 m. L'approche vise à tester la validité d'une méthode permettant de combiner deux techniques de plus en plus utilisées en hydrogéophysique : des

méthodes géophysiques de haute résolution telles que le géoradar en forage et le piézocône multi-capteurs, et des méthodes hydrogéologiques telles que l'essai de perméabilité *in situ* à niveaux multiples.

Plus précisément, les travaux visent à formuler un lien statistique *in situ* entre les mesures de conductivité électrique, de permittivité et de conductivité hydraulique recueillies le long de forages et d'appliquer cette relation à la zone entre ces forages grâce aux mesures de radar géologique en mode tomographique. Cette combinaison statistique de données de différentes provenances, mais toutes à la même échelle, est réalisée à l'aide de la simulation séquentielle gaussienne par approche bayésienne.

1.2 Plan de travail

Le mémoire comprend 8 chapitres. Le deuxième chapitre passe en revue les différentes approches hydrogéophysiques publiées à ce jour ayant comme objectif l'estimation du champ de conductivité hydraulique à l'aide de la géophysique. Les notions hydrogéologiques, géophysiques et statistiques mises en oeuvre dans ce projet sont succinctement décrites dans le chapitre 3. Le chapitre 4 se concentre sur le mode opératoire de l'approche proposée. Les chapitres 5 et 6 présentent les résultats obtenus avec un modèle synthétique d'abord et avec un jeu de données réelles ensuite. Une discussion sur les différents résultats est présentée au chapitre 7. Finalement, le chapitre 8 fait la synthèse des éléments avancés dans ce travail et conclut par des recommandations pour de futurs travaux.

CHAPITRE 2

REVUE DE LITTÉRATURE

Ce chapitre passe en revue les différentes approches actuelles visant à estimer les paramètres hydrogéologiques d'un aquifère à l'aide de la géophysique. Celles-ci se classent globalement en trois groupes : la méthode par estimation directe, souvent appelée *direct mapping* en anglais, l'inversion conjointe et la méthode par intégration de données géophysiques et hydrogéologiques. Cette classification englobe plutôt bien l'ensemble des publications dans le domaine, quoique plusieurs d'entre elles n'entrent pas précisément dans une des catégories.

Une revue de ces différentes approches a aussi été réalisée par Slater (2007), Linde et al. (2006b) et Hubbard et Rubin (2005).

2.1 Estimation directe

La méthode par estimation directe se définit comme la transformation d'un modèle géophysique en un modèle hydrogéologique à l'aide d'une relation pétrophysique, existante ou déterminée *in situ* (Linde et al., 2006b). C'est la seule méthode possible lorsque l'on ne dispose d'aucune donnée hydrogéologique. Les résultats ainsi obtenus sont plutôt qualitatifs que quantitatifs en raison de la fiabilité relative des relations pétrophysiques comme celle d'Archie (1942) ou celle de Topp et al. (1980) (voir section 3.2), de l'échelle à laquelle elles sont appliquées, de l'hétérogénéité du site étudié (Lesme et Friedman, 2005) et de la résolution limitée de la reconstruction des images géophysiques par inversion.

Un exemple représentatif où ce type d'approche donne de bons résultats est l'estimation de la teneur en eau du sol à partir de données tomographiques obtenues par géoradar (Binley et al., 2001; Alumbaugh et al., 2002). Brièvement, cette méthode géophysique consiste à placer un émetteur d'ondes aux fréquences radar dans un forage et un récepteur dans un autre, situé à une distance de quelques mètres. L'émetteur et le récepteur sont ensuite déplacés de telle sorte à couvrir toute la surface entre les forages et, à chacune des positions, le temps de parcours de l'onde (ainsi que son amplitude) sont enregistrés par le récepteur. Les vitesses sont ensuite déterminées par inversion des temps de parcours (Giroux et al., 2007; Gloaguen et al., 2007a,b) qui, à leur tour, peuvent être converties en valeurs de constante diélectrique (voir section 3.3.1) et finalement en teneur en eau par des équations telles que celles de Topp ou de CRIM. Alumbaugh et al. (2002) ont réalisé une étude portant sur la validité et la précision de l'estimation de la teneur en eau, dans la zone vadose, à partir de données radar en tomographie et d'une relation pétrophysique établie *in situ*. Ils ont conclu que l'estimation de la teneur en eau par tomographie radar engendrait une erreur quadratique (*root mean square*) de 2.0 à 3.1%.

Plus récemment, Hördt et al. (2007) ont obtenu des résultats intéressants pour estimer la conductivité hydraulique de sédiments non consolidés à partir de données de polarisation provoquée (PP). Ceux-ci se sont basés, entre autre, sur les travaux en laboratoire de Slater et Lesme (2002) et de Börner et al. (1996) qui se consa- craient à établir des relations empiriques entre la conductivité hydraulique et la conductivité électrique complexe, correspondant à celle mesurée par la méthode PP.

Selon certains auteurs (Linde et al., 2006c; Kowalsky et al., 2006; Day-Lewis et al., 2005), l'approche directe présente certaines lacunes. L'une d'elles provient du fait que la relation entre les paramètres géophysiques et les propriétés hydrauliques

est considérée stationnaire sur tout le domaine d'étude, indépendante des erreurs d'acquisition et de la méthode d'inversion utilisée. Or, ces conditions sont rarement rencontrées puisque les relations pétrophysiques tendent à varier en fonction de l'hétérogénéité du sol. De plus, le tomogramme est fortement influencé par la méthode d'inversion adoptée, par la qualité des données ainsi que par l'opérateur de régularisation qui tend généralement à « lisser » les structures géologiques.

Finalement, bien que cette approche soit relativement rapide, les recherches se tournent maintenant davantage vers l'intégration de données hydrogéologiques aux données géophysiques. Ceci est le cas des deux autres approches présentées dans les sections suivantes.

2.2 Inversion conjointe et inversion contrainte

Une approche prometteuse, mais encore embryonnaire, est l'inversion conjointe. On entend par inversion conjointe que des données hydrogéologiques sont incluses lors de l'inversion géophysique (ou vice versa) ou encore que des données provenant de deux méthodes géophysiques différentes ont été conjointement inversées. Cette méthode vise à minimiser l'impact des erreurs engendrées par la prise de mesures et par le processus d'inversion (Linde et al., 2006b).

Selon Hubbard et Rubin (2005), même s'il est évident qu'il est profitable à l'inversion géophysique d'être contrainte par des données hydrogéologiques, une inversion conjointe réelle est rarement réalisée en pratique. Ceci est dû au fait que les différents types de données doivent être sensibles aux mêmes propriétés physiques, permettant ainsi de générer un modèle physique minimisant les écarts entre les deux jeux de données. Or, les données géophysiques et hydrogéologiques sont en général sensibles à différents paramètres physiques (Slater, 2007). Surmonter cet

obstacle constitue l'un des grands défis de l'hydrogéophysique.

Par ailleurs, des approches basées sur les méthodes d'optimisation heuristique (recuit simulé, réseau de neurones,...) permettent de fusionner plusieurs types de données dans des fonctions objectives presque sans contrainte. Cependant, le temps de calcul et la pondération des fonctions objectives rendent leur utilisation difficile.

Cette section présente donc quelques exemples récents d'inversion conjointe réalisées dans le but d'estimer des paramètres hydrauliques.

2.2.1 Approche par zonation

Sous l'hypothèse que l'écoulement souterrain est principalement dominé par les hétérogénéités de grande échelle, Hyndman et al. (1994) ont réalisé une inversion conjointe 2D de données synthétiques sismiques et d'un essai de traceur. Cette méthode permet de délimiter les grandes zones lithologiques à partir des vitesses sismiques obtenues par inversion tomographique et d'associer, à chacune d'elle, une conductivité hydraulique déterminée par l'essai de traceur. Ceci est fait selon un processus itératif par l'algorithme qu'ils ont appelé « *Split inversion methodology* » (SIM) qui permet de déterminer les zones correspondant le mieux à plusieurs types de données (ici, vitesses sismiques et concentration d'un traceur). Cette approche, reprise plus tard par Hyndman et Gorelick (1996) pour un cas 3D réel, semble représenter plutôt bien les hétérogénéités ainsi que les variations de conductivités hydrauliques entre deux forages. Par la suite, Hyndman et al. (2000) ont développé une méthode permettant d'établir une relation entre les vitesses sismiques et la conductivité hydraulique à une échelle « locale ». L'approche est basée sur la minimisation de l'erreur entre les données de concentrations réelles de traceur et celles simulées à partir d'un champ de conductivités hydrauliques, celui-ci déterminé par

les vitesses sismiques et une relation linéaire variable entre ces deux paramètres.

Une approche similaire à celle de Hyndman et al. (2000) a été présentée par Linde et al. (2006c) à partir de données synthétiques. Seulement, ce sont des tomodigraphies radar de vitesse qui ont permis de délimiter les zones de même lithologie. Une relation pétrophysique distincte, déterminée au cours de l'inversion hydrogéologique, est ensuite assignée à chacune de ces zones. Selon Slater (2007), la validité d'une relation pétrophysique à une telle échelle est discutable. À ce propos, plusieurs auteurs (Moysey et al., 2005; Day-Lewis et al., 2005; Renard, 1997) se sont penchés sur le problème récurrent du changement d'échelle. Ceux-ci ont questionné la représentativité statistique d'appliquer à une échelle « locale » une relation pétrophysique déterminée à petite échelle. Différentes méthodes de changement d'échelle sont donc proposées par ces auteurs.

2.2.2 Approche structurale

Une seconde approche, dite « structurale » (Haber et Oldenburg, 1997; Gallardo et Meju, 2003, 2004; Linde et al., 2006a), permet la réalisation d'une inversion conjointe de deux méthodes géophysiques sous l'hypothèse que chacun des modèles sous-jacents présente la même structure et que les changements de propriétés physiques s'opèrent à une même position dans la même direction. L'inversion est résolue grâce à la minimisation d'une fonction objective quantifiant les différences de structure des deux modèles, tout en respectant les contraintes de chacun des jeux de données. Ce problème non linéaire est résolu itérativement (Haber et Oldenburg, 1997). Conséquemment, cette méthode ne requiert pas de relation pétrophysique prédéterminée.

Haber et Oldenburg (1997) ont d'abord testé un tel algorithme sur des données synthétiques et conclu que l'inversion conjointe ainsi réalisée est supérieure à l'inversion individuelle de chacun des deux jeux de données. Plus récemment, Gallardo et Meju (2004) ont réalisé une inversion conjointe 2D de données de résistivité DC et de sismique réfraction afin de mieux caractériser la subsurface. Ceux-ci aboutissent aux mêmes conclusions que Haber et Oldenburg (1997). Bien que ces derniers n'aient pas tenté d'estimer des paramètres hydrauliques, la méthode d'inversion conjointe présente, selon Slater (2007), un grand potentiel.

En se basant sur les travaux précédemment mentionnés, Linde et al. (2006a) ont réalisé une inversion conjointe 3D des temps de parcours obtenus par géoradar et de résistivité électrique en tomographie. Tout en visant à améliorer le modèle géophysique résultant de cette inversion, ils ont développé une méthode permettant de calculer des opérateurs de régularisation stochastiques, basés sur des modèles géostatistiques stationnaires qui eux sont estimés à partir de diagraphies EM, de conductivité électrique et de données radar en mode « ZOP » (émetteur et récepteur placés à la même élévation dans les deux forages). Les avancées proposées ont été appliquées à des données provenant de la zone vadose d'un aquifère sablonneux. À l'aide de diagrammes de dispersion, les auteurs sont parvenus à délimiter quatre zones ayant des propriétés électrophysiques semblables. Cette zonation n'aurait pas été possible uniquement par inversion simple de chacune des méthodes. En utilisant les relations pétrophysiques classiques, ils ont déterminé la perméabilité effective et la teneur en eau du site.

2.2.3 Approche stochastique

D'autres travaux importants en inversion conjointe hydrogéophysique sont ceux de Kowalsky et al. (2004) et Kowalsky et al. (2005). Ceux-ci proposent un algorithme

d'inversion permettant d'estimer la distribution des propriétés hydrauliques et de prédire le comportement de l'écoulement dans la zone vadose à l'aide de mesures géophysiques (géoradar) et de mesures hydrogéologiques recueillies en régime transitoire.

L'algorithme d'inversion est basé sur l'hypothèse que la perméabilité est le seul paramètre hydraulique non uniforme. On la considère comme une fonction aléatoire de distribution log-normale dont la corrélation spatiale est connue. Par la méthode du « maximum *a posteriori* » (MAP) et celle des points pilotes, la distribution de la perméabilité estimée est conditionnelle aux mesures de perméabilité, respecte les corrélations spatiales spécifiques et concorde avec les données hydrogéologiques et géophysiques. La méthode a d'abord été testée sur des données synthétiques (Kowalsky et al., 2004) en combinant des temps de parcours obtenus par radar en configuration « ZOP » (l'émetteur et le transmetteur sont maintenus à la même élévation dans chacun des trous de forage) et des mesures de perméabilité.

Les résultats de cet exemple synthétique ont montré que l'ajout de données radar dans l'inversion réduit grandement l'incertitude sur la prédiction des propriétés hydrauliques du sol dans la zone vadose. Par contre, pour des conditions réelles, les hypothèses qui sous-tendent l'inversion ne sont pas toujours valides (Linde et al., 2006b). La même approche a été appliquée à des données réelles par Kowalsky et al. (2005) lors d'un essai d'infiltration. Les résultats ont été comparés à des mesures provenant d'une sonde neutron, recueillies au même moment. Les conclusions tirées sont les mêmes que celles précédemment mentionnées. Un des atouts majeur de cette méthode est la possibilité de travailler directement avec les temps de parcours ce qui supprime l'étape de l'inversion tomographique individuelle, souvent source d'erreur.

2.3 Estimation par intégration

Cette dernière catégorie se réfère aux cas où l'inversion des données géophysiques est réalisée indépendamment de l'inversion des données hydrogéologiques. Ces deux types de données sont ensuite combinés à l'aide d'une approche géostatistique ou stochastique. Les propriétés hydrauliques sont alors estimées par interpolation des données hydrogéologiques disponibles en tenant compte de leur incertitude, des relations pétrophysiques existantes et de toute autre information connue. La méthode proposée dans ce mémoire entre dans cette catégorie. C'est pourquoi les quelques travaux présentant une telle approche seront abordés plus en détails que dans les sections précédentes.

2.3.1 Analyse par classification non-hiérarchique

L'analyse par classification non-hiérarchique est une méthode statistique qui consiste à définir, à partir d'un ensemble de données, des sous-groupes comportant chacun des caractéristiques distinctes. En hydrogéologie, cette méthode peut permettre d'identifier différentes unités lithologiques.

Une telle approche a été adoptée par Gloaguen et al. (2007b). En premier lieu, ceux-ci présentent un algorithme d'inversion tomographique radar dont le rayonnement est modélisé sous forme de rais courbes. Cet algorithme est basé sur la modélisation de la covariance entre les temps (et les amplitudes) déterminée à partir des données réelles. Les covariances calculées ainsi que les données réelles et toutes autres données additionnelles provenant de mesures le long des forages, par exemple, sont cokrigées. Des simulations conditionnelles sont également réalisées. Ces travaux montrent que l'utilisation de contraintes le long des forages diminue substantiellement l'incertitude sur les tomographies de vitesse et d'atténuation.

En deuxième lieu, les champs de vitesse et d'atténuation sont convertis en conductivités électriques et permittivités grâce aux relations pétrophysiques classiques. Une méthode de classification non-hierarchique a permis de distinguer deux familles lithologiques.

Tronicke et Holliger (2005) ont étudié l'apport des simulations stochastiques conditionnelles, basées sur la méthode du recuit simulé, afin d'estimer un champ de porosité par intégration de différents types de données géophysiques et hydrogéologiques. Un modèle hétérogène de porosité est d'abord créé à partir duquel des données synthétiques de diagraphies neutron et de géoradar en forage sont générées. Ces données sont ensuite combinées par méthode géostatistique et des simulations sont réalisées. La comparaison de ces simulations au modèle montre que cette approche reproduit bien l'hétérogénéité de la porosité.

Dans un même ordre d'idée, Paasche et al. (2006) montrent le potentiel de l'analyse par classification « *fuzzy c-means* » (FCM) pour intégrer des données radar et sismiques en forages, des diagraphies neutron et des essais *in situ* provenant d'un aquifère alluvionnaire. Les résultats indiquent que cette approche permet de bien délimiter les principales unités sédimentaires et produit des estimés plausibles de conductivités hydrauliques dans les zones non échantillonnées. Suite à ces résultats concluants, Paasche et al. (2006) ont présenté un algorithme, basé sur l'analyse par classification « *fuzzy c-means* », permettant d'intégrer plusieurs types de données géophysiques 2D et hydrogéologiques afin d'identifier les principales lithologies et les conductivités hydrauliques qui leur sont associées.

2.3.2 Approche bayésienne

L'approche dite bayésienne est une approche stochastique basée sur le théorème de Bayes. Brièvement, cette approche consiste à combiner l'information préalablement connue du paramètre recherché avec toutes autres données reliées de façon linéaire ou non avec cette variable. La force de cette méthode stochastique réside dans sa flexibilité. En effet, elle permet d'intégrer différents types de données tout en tenant compte des liens non-linéaires existant entre eux. Elle permet aussi d'inclure toute information préalablement connue sur la variable à estimer. C'est précisément le caractère non-linéaire des relations entre les propriétés physiques et hydrologiques du sol ainsi que la faible densité des points de mesures qui a incité certains auteurs (Chen et al., 2001; Hubbard et al., 2001; Gastaldi et al., 1998) à opter pour ce type d'approche. Pour ces mêmes raisons, c'est aussi l'approche qui a été adoptée dans ce mémoire.

Dans le cadre d'un grand projet mené conjointement par l'Université de Californie et le Laboratoire de recherche *Lawrence Berkeley*, une grande quantité de données géophysiques et hydrogéologiques a été relevée au site « *South Oyster* » situé sur la péninsule est de l'état de Virginie. Parmi les nombreuses publications issues de ce projet, deux d'entre elles présentent les résultats de l'approche bayésienne pour estimer les propriétés hydrogéologiques du site.

Chen et al. (2001) proposent une méthode d'estimation de la distribution spatiale des conductivités hydrauliques d'un aquifère par la combinaison de mesures géophysiques entre forages et de conductivités hydrauliques le long de ces forages. Un des objectifs poursuivi par les auteurs est de mesurer l'apport des données géophysiques pour réduire l'incertitude liée à l'estimation des conductivités hydrauliques. Les données géophysiques en question sont des vitesses et atténuations radar ainsi que des vitesses sismiques, toutes trois obtenues par tomographies entre

forages.

C'est donc dans un cadre bayésien, où chacune des variables est considérée comme une variable aléatoire, que les conductivités hydrauliques conditionnelles aux conductivités hydrauliques déjà connues et aux données géophysiques, ont été obtenues. La particularité de cette approche est l'élaboration d'un modèle de régression linéaire normal permettant d'étudier systématiquement les relations entre chacun des paramètres pour former la fonction appelée *vraisemblance*. Les estimations appelées *a priori* sont calculées à partir des conductivités hydrauliques déjà connues et sont ensuite mises à jour à l'aide des relations établies avec les données géophysiques.

Les résultats montrent que l'ajout de données géophysiques améliore l'estimation des conductivités hydrauliques. Cet apport est particulièrement important loin des zones où des données hydrauliques sont déjà connues. De plus, ceux-ci concluent que les vitesses radar et sismiques sont plus corrélées avec la conductivité hydraulique que l'atténuation radar.

La même approche a été reprise par Hubbard et al. (2001) en ajoutant, aux données précédentes, des données de radar de surface, de pénétromètre à cône et de débitmètre électromagnétique. Ces recherches ont permis une interprétation de la stratigraphie et une estimation haute résolution du champ de conductivité hydraulique. Ce dernier a ensuite été utilisé afin de contraindre un modèle numérique de transport de bactéries. Suite à la validation du modèle par un essai de traceur, les auteurs concluent que la géophysique apporte une information supplémentaire valide et de haute résolution qui n'est pas accessible par les seules méthodes hydrogéologiques.

Finalement, Gastaldi et al. (1998) ont utilisé une approche bayésienne pour estimer

l'épaisseur d'un réservoir à partir de données sismiques 3D. Des simulations non-linéaires ont été préférées aux relations linéaires qui, selon les auteurs, sont trop imprécises dans des cas hétérogènes. En plus des données sismiques réelles provenant de 15 puits, un modèle analogue a été créé afin de générer un plus grand nombre de données. C'est grâce au lien existant entre les isochrones sismiques du réservoir, les isochrones sismiques de la stratigraphie et l'épaisseur du réservoir que ce dernier paramètre a pu être estimé. Selon les auteurs, la relation multivariable non-linéaire unissant ces trois paramètres peut être décrite par la densité de probabilité conjointe, déterminée par la méthode non-paramétrique de noyau. En insérant cette dernière à des simulations séquentielles gaussiennes, les auteurs ont produit un grand nombre de réalisations de l'épaisseur du réservoir contraintes par les données sismiques. En comparant les résultats à cinq autres puits, leur travaux montrent que la méthode améliore grandement l'estimation de l'épaisseur du réservoir par rapport aux méthodes traditionnelles comme le krigeage ou les méthodes linéaires comme le cokrigeage. Une méthodologie très similaire à celle-ci a été appliquée dans ce mémoire.

CHAPITRE 3

REVUE THÉORIQUE

Ce chapitre introduit les différentes notions théoriques essentielles à la réalisation du travail présenté dans ce mémoire. La première section présente quelques concepts hydrogéologiques de base, suivis des relations pétrophysiques utiles entre les propriétés électriques et hydrogéologiques du sol. La section suivante porte sur les différentes méthodes géophysiques impliquées dans ce projet. On y retrouve une explication sommaire des principes physiques régissant chacune de ces méthodes géophysiques ainsi qu'une description plus approfondie du mode d'acquisition et de traitement des données. Les principales notions géostatistiques entourant la simulation séquentielle selon une approche bayésienne sont finalement présentées.

3.1 Notions de base d'hydrogéologie

3.1.1 Loi de Darcy

La loi de Darcy est un des concepts fondamentaux de l'hydrogéologie. Pour un écoulement unidirectionnel, elle exprime le débit par unité de temps d'un fluide s'écoulant à travers une matrice poreuse de section A et de longueur L (éq. 3.1). Il est important de mentionner que cette relation n'est valide que sous l'hypothèse d'un écoulement laminaire (nombre de Reynolds < 10).

$$Q = KA \frac{\Delta h}{L} , \quad (3.1)$$

où Q est le débit mesuré en m^3/s , K , la conductivité hydraulique en m/s et h , la charge hydraulique en m. La différence de charge par unité de longueur, $(\frac{\Delta h}{L})$, représente le gradient hydraulique. On suppose que l'écoulement se produit toujours dans le sens contraire à celui-ci, c'est à dire d'un point où la charge est plus élevée vers un point où elle est plus faible.

La formulation précédente, où K est un scalaire, s'applique uniquement à des milieux isotropes, c'est-à-dire des milieux où la dynamique de l'écoulement est la même dans toutes les directions. Lorsque ceci n'est pas vérifié, le milieu est alors anisotrope et la loi de Darcy s'exprime sous forme vectorielle. Dans ce cas, la conductivité hydraulique devient tensorielle et la loi de Darcy s'énonce par l'équation suivante, si l'on suppose que les directions principales d'anisotropie coïncident avec les axes x , y et z

$$\begin{bmatrix} q_{xx} \\ q_{yy} \\ q_{zz} \end{bmatrix} = - \begin{bmatrix} K_{xx} & 0 & 0 \\ 0 & K_{yy} & 0 \\ 0 & 0 & K_{zz} \end{bmatrix} \begin{bmatrix} \partial h / \partial x \\ \partial h / \partial y \\ \partial h / \partial z \end{bmatrix}. \quad (3.2)$$

La conductivité hydraulique saturée, K , exprime la facilité avec laquelle le fluide traverse le milieu poreux. Elle dépend à la fois des propriétés du milieu et des propriétés du fluide. C'est donc le paramètre principal dans la caractérisation d'un écoulement. Elle se définit par la formule suivante pour un milieu isotrope

$$K_s = k \frac{\rho_w g}{\mu_f}, \quad (3.3)$$

où k est la perméabilité (m^2), ρ_w , la masse volumique du fluide (kg/m^3), g , la constante gravitationnelle (m/s^2) et μ_f , la viscosité dynamique du fluide ($\text{Pa}\cdot\text{s}$).

Ainsi, plus le fluide est visqueux, plus la conductivité hydraulique saturée est faible. La perméabilité est le paramètre décrivant les propriétés des pores entre les grains solides du milieu. Ce terme sera abordé plus en détail dans les sections suivantes.

3.1.2 Écoulement en milieu saturé

La dynamique de l'écoulement de l'eau souterraine est régie par l'équation de Darcy mais obéit aussi à la loi de conservation de masse. Ainsi, la masse d'eau entrant dans un milieu moins la masse d'eau sortant doit être égale à la variation dans le temps de l'eau y étant emmagasinée. Cette dernière expression, combinée à la loi de Darcy se traduit par l'équation 3.4, pour un milieu anisotrope, saturé et en régime transitoire (Domenico et Schwartz, 1997)

$$\nabla \cdot (\mathbf{K} \nabla h) = S_s \frac{\partial h}{\partial t}, \quad (3.4)$$

où \mathbf{K} (m/s) est le tenseur des conductivités hydrauliques :

$$\begin{bmatrix} K_{xx} & K_{xy} & K_{xz} \\ K_{yx} & K_{yy} & K_{yz} \\ K_{zx} & K_{zy} & K_{zz} \end{bmatrix}, \quad (3.5)$$

h est la charge hydraulique (m), S_s est le coefficient d'emmagasinement spécifique (sans unité), q représente un terme source¹, x , y et z (m) sont les coordonnées spatiales orientées selon la règle de la main gauche² et t est le temps (s). La solution de l'équation 3.4 décrit la variation de la charge hydraulique, en tout

¹associé aux fuites ou à une source ponctuelle

²lorsque z augmente avec la profondeur

point et à tout moment.

À des fins de simplification, le milieu est parfois considéré isotrope. Dans ce cas, le tenseur de conductivité hydraulique se réduit à une constante et l'équation 3.4 s'exprime sous la forme de l'équation de diffusion. Finalement, si le régime d'écoulement est permanent, c'est-à-dire que la charge ne varie pas dans temps ($\frac{\partial h}{\partial t} = 0$), l'équation permettant d'obtenir la charge en tout point du domaine, se réduit à l'équation de Laplace (Domenico et Schwartz, 1997)

$$\nabla^2 h = 0 . \quad (3.6)$$

L'obtention du champ de charges hydrauliques à partir des équations mentionnées requiert une connaissance des conditions aux frontières ainsi que de la charge hydraulique (ou du tenseur), dans le cas d'un régime permanent. Pour un régime transitoire, le coefficient d'emmagasinement ainsi qu'une solution initiale doivent aussi être connus (Bear, 1979).

3.2 Relations entre les propriétés hydrauliques et électriques du sol

La détermination de la continuité spatiale des paramètres hydrauliques, à une échelle d'environ 15 cm pour une zone comprise entre deux forages, n'est pas aisée disposant seulement de données hydrogéologiques traditionnelles (essai de pompage, essai *in situ*, essai de traceur,...). En effet, les propriétés hydrauliques peuvent varier de plusieurs ordres de grandeur sur une très petite distance, rendant l'estimation imprécise (Lesme et Friedman, 2005).

La géophysique tente de combler cette lacune en apportant des mesures de haute

précision et de haute densité spatiale des propriétés physiques du sol. C'est donc un outil intéressant à incorporer dans une étude hydrogéologique menée à l'échelle locale, c'est-à-dire des zones de quelques centaines de mètres carrés de surface par une dizaine de mètres de profondeur. Outre sa précision, la géophysique possède aussi l'avantage d'être très peu destructive et de ne pas perturber le régime d'écoulement.

Cependant, les méthodes géophysiques ne mesurent jamais directement les propriétés hydrogéologiques mais bien les propriétés physiques du sol. Donc, des relations empiriques ou statistiques doivent être établies afin d'incorporer adéquatement des données géophysiques à une étude hydrogéologique.

Les principales méthodes géophysiques utilisées en hydrogéophysique sont le géoradar, la résistivité électrique (ER), la polarisation provoquée (PP), l'induction électromagnétique (EMI) et la réflectométrie dans le domaine temporel (TDR) et plus récemment par les sondages par résonance magnétique. Ces méthodes électromagnétiques se distinguent par la gamme de fréquences à laquelle elles opèrent, allant du courant continu direct (DC) (très basse fréquence) pour la résistivité électrique, à plus de 1 GHz pour le géoradar (Lesme et Friedman, 2005). Les propriétés électriques mesurées par ces différentes méthodes peuvent apporter de l'information sur les contacts géologiques, la composition du sol et de l'eau interstitielle et sur sa teneur en eau. Dans ce projet, on utilise le géoradar et la résistivité électrique.

La section suivante décrit les principales relations pétrophysiques existantes à ce jour entre les propriétés hydrogéologiques et les propriétés électriques du sol. Le principal problème de ces relations est leur non unicité (Rubin et Hubbard, 2005). Cet obstacle est présentement contourné par la combinaison de deux ou plusieurs méthodes géophysiques et hydrogéologiques et par le recours à des méthodes géostatistiques dans l'analyse des données. Après un survol des propriétés hydrau-

liques et électriques des milieux poreux, les modèles de permittivité et de conductivité électrique qui leur sont associés sont présentés. Nous nous limitons à l'analyse en milieu saturé.

3.2.1 Propriétés hydrauliques

3.2.1.1 Porosité et surface spécifique

En général, un sol se compose de trois phases : la phase solide (grains), la phase liquide (eau) et la phase gazeuse (air). Le volume occupé par la phase liquide et la phase gazeuse constitue le volume des « vides » ou pores.

La *porosité* est donc un paramètre essentiel dans la description d'un sol. Celle-ci se traduit par l'équation 3.7 et se définit comme le rapport entre le volume des vides et le volume total d'un échantillon (Grolier et al., 1991). C'est une valeur sans unité mais souvent exprimée en pourcentage qui prend la forme

$$\phi = \frac{V_v}{V_t} . \quad (3.7)$$

Lorsque le volume occupé par le gaz est totalement remplacé par le liquide, on dit que le milieu est saturé en liquide. Ce sont les pores interconnectés qui permettent à l'eau de s'écouler. On distingue la porosité totale (éq. 3.7), de la *porosité effective* correspondant seulement aux pores et interstices interconnectés permettant le passage de l'eau. Même si cette distinction est importante, elle se révèle difficile à faire concrètement. Ainsi, tout au long du mémoire, le terme porosité sera utilisé plutôt que porosité effective. Le tableau (3.1) présente des valeurs types de porosité totale de quelques sols non-consolidés.

Tab. 3.1: Gamme de valeurs de la porosité pour quelques sols non-consolidés (Extrait de Domenico et Schwartz (1997))

Type de sol	Porosité (%)
Gravier grossier	24-36
Gravier fin	25-38
Sable grossier	31-46
Sable fin	26-53
Silt	34-61
Argile	34-60

Selon Schön (1998), les facteurs influençant la porosité sont :

- les propriétés géométriques des grains et des pores (arrangement spatial, forme des grains, également) ;
- le processus de sédimentation et le degré de cimentation ;
- la profondeur et la pression.

Ainsi, pour des processus de compactage identiques³, des études expérimentales (Marion et al., 1992; Fraser, 1935) ont montré que la porosité d'un sol poreux granulaire varie en fonction de la fraction de particules fines qu'il contient (Figure 3.1). La porosité minimale est atteinte lorsque le volume des particules fines égale la porosité du matériau le plus grossier.

Il est important de noter que la porosité est, en théorie, indépendante de la taille des grains. Cependant, les grains de très petite taille comme ceux des argiles génèrent, par leur composition, des forces de surface qui « retiennent » l'eau, provoquant ainsi une augmentation de la porosité lorsque la taille des grains diminue

³pour d'autres modes de préparation et/ou de compactage, on obtient une courbe différente

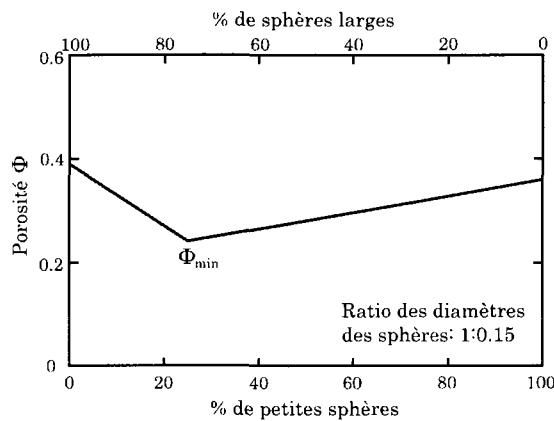


Fig. 3.1: Variation de la porosité d'un sol selon la pourcentage de particules fines qu'il contient pour des processus de compactage identiques (Tiré de Morin, 2006)

(Schön, 1998).

La *surface spécifique* est un autre paramètre contrôlant grandement l'écoulement de l'eau dans les sols. Celle-ci se définit comme la surface totale d'une masse M_s de particules divisée par cette masse (Mernoud, 2006). Ce paramètre est généralement exprimé en m^2/kg mais parfois aussi en m^2/m^3 selon qu'il soit multiplié par la masse volumique du grain solide, ou non.

Selon Schön (1998), la surface spécifique dépend principalement de la taille des grains et de leur forme ainsi que de la microstructure et de la morphologie de l'interface entre la matrice et les vides. Elle tend à augmenter lorsque la taille des grains diminue. Pour une même taille de grain, la sphéricité et l'arrondissement des grains influencent aussi inversement la surface spécifique. C'est une valeur qui est, pour le moment, principalement accessible par des prédictions à partir d'essais simples en laboratoire (Chapuis et Aubertin, 2003) mais qui présente tout de même un grand intérêt pour des travaux à plus grande échelle en raison du rôle prépondérant qu'elle joue dans la détermination de la perméabilité. Des travaux récents indiquent que la surface spécifique peut être estimée directement sur le ter-

rain à l'aide de la PP spectrale (Kirsch, 2006). Ces travaux, bien que préliminaires, sont encourageants pour le futur.

3.2.1.2 Conductivité hydraulique et perméabilité

Tel que mentionné à la section 3.1.1, la conductivité hydraulique décrit la facilité avec laquelle un fluide s'écoule à travers une matrice poreuse. C'est un paramètre clé de toute caractérisation hydrogéologique d'un site qui dépend à la fois des propriétés géométriques du milieu et des propriétés du fluide qui y circule. Elle peut varier de plusieurs ordres de grandeur sur de très courtes distances et peut aussi varier considérablement selon l'échelle à laquelle on la mesure. En effet, selon Dallaire (2004), la conductivité hydraulique mesurée par un essai de pompage peut être, pour un même site, jusqu'à 6 fois plus élevée que celle mesurée au laboratoire quand on a un mauvais modèle du terrain. Le tableau 3.2 présente quelques valeurs représentatives de conductivités hydrauliques pour différents sols homogènes.

Tab. 3.2: Valeurs représentatives de conductivité hydraulique (Extrait de Domenico et Schwartz (1997))

Type de sol	Conductivité hydraulique (m/s)
Gravier	3×10^{-4} à 3×10^{-2}
Sable grossier	9×10^{-7} à 6×10^{-3}
Sable moyen	9×10^{-7} à 5×10^{-4}
Sable fin	2×10^{-7} à 2×10^{-4}
Silt	1×10^{-9} à 2×10^{-5}
Till	1×10^{-12} à 2×10^{-6}
Argile	1×10^{-11} à $4,7 \times 10^{-9}$

La conductivité hydraulique, pour un milieu isotrope et saturé, s'exprime par la formule 3.3.

Sous l'hypothèse que les variables liées à l'eau souterraine (viscosité, masse volumique, salinité) demeurent plutôt constantes à l'échelle étudiée, soit quelques dizaines de mètres carrés tout au plus, le terme critique dans la détermination de la conductivité hydraulique devient la *perméabilité*. Celle-ci s'exprime généralement en m^2 mais parfois aussi en Darcy ($1\text{d} = 0.9869 \times 10^{-12} \text{ m}^2$).

Selon Schön (1998), la perméabilité dépend essentiellement de :

- de la taille des pores et de leur distribution ;
- la porosité ;
- de la forme des pores, de la morphologie de l'interface des grains et de la surface spécifique ;
- de l'arrangement spatial des pores (connectivité et tortuosité).

Les principales tendances existantes entre la perméabilité et les variables mentionnées ci-haut sont :

- la perméabilité augmente avec l'augmentation de la taille des grains ;
- la perméabilité augmente avec l'augmentation de la porosité ;
- la perméabilité diminue avec l'augmentation du compactage et de la cimentation ;

Ces tendances, valables seulement pour des sols qui ne sont pas endommagés mécaniquement, hydrauliquement ou thermiquement, sont examinées plus en détails au chapitre 5.

Il existe plusieurs formulations empiriques ou semi-empiriques permettant d'estimer la perméabilité (ou la conductivité hydraulique) d'un sol à partir de ses caractéristiques géométriques (Hazen, 1892; Kozeny, 1927; Carman, 1939; Chauvet, 1999). Parmi celles-ci, l'équation la plus connue (sans être la plus utilisée en

raison de la complexité des variables qui la composent) est celle dite de « Kozeny-Carman »⁴.

La formule de Kozeny-Carman apparaît sous différentes formes selon l'auteur qui la cite. L'équation 3.8 présente l'une d'elle, tirée de Carrier (2003).

$$k = \left(\frac{1}{C_{KC}} \right) \left(\frac{1}{S_s^2} \right) \left(\frac{\phi^3}{(1-\phi)^2} \right), \quad (3.8)$$

où C_{KC} est un coefficient empirique, S_s est la surface spécifique en (m^2/m^3) et ϕ est la porosité.

Cette équation est basée sur le modèle comparant un milieu poreux à des tubes capillaires tortueux dans lesquels un fluide visqueux s'écoule sous un régime laminaire. Le paramètre émergeant de ce modèle s'appelle la *tortuosité* et se définit comme le rapport entre la longueur d'un tube capillaire et la longueur de l'échantillon (Grolier et al., 1991). Malgré la simplicité du concept, la tortuosité reste une valeur plutôt abstraite à déterminer en réalité. Elle est donc « incluse » dans un coefficient empirique qui comprend aussi les effets liés à la sphéricité des grains. En supposant que les grains sont des sphères parfaites, le coefficient C_{KC} est estimé à 5 (Carrier, 2003).

Par ailleurs, il existe d'autres équations empiriques non pas basées sur la surface spécifique, mais sur une valeur représentative du diamètres des grains (d_5 , d_{10}). Or, le diamètre des grains présente une variance beaucoup plus faible que la surface spécifique, ces équations semblent plus appropriées à grande échelle. Lors d'une

⁴à noter que le comte von Kozeny et le baron von Carman n'ont jamais publié ensemble. Ce sont des chercheurs du domaine pétrolier qui ont assemblé des morceaux des théories de Kozeny et de Carman pour bâtir une théorie qu'ils ont appelée l'équation de « Kozeny-Carman » (R.P. Chapuis, 2009, communication personnelle)

revue des différentes équations empiriques dépendantes de la porosité et de la taille des grains, Vukovic et Soro (1992) recommande l'utilisation de l'équation de Zunker pour des sols à grains moyens et de l'équation de Zamarin pour des sols sablonneux à grains plus grossiers. Cette dernière revêt la forme

$$k = \beta_{ZN} \left(\frac{\phi^3}{(1 - \phi)^2} \right) d_e^2 , \quad (3.9)$$

où β_{ZN} est un coefficient empirique valent en moyenne 8.2×10^{-3} , ϕ est la porosité et d_e est une valeur du diamètre représentatif des grains.

La formule de Zamarin, tout comme celle de Kozeny-Carman et des autres qui ne sont pas mentionnées ici, prédisent plutôt bien la perméabilité lorsque chacun des paramètres entrant dans la formulation (porosité, surface spécifique, tortuosité, d_m ou d_{10} , etc.) sont déterminés avec précision, si possible au laboratoire. Or, cette pratique s'avère impossible à réaliser à une échelle de quelques mètres carrés. Conséquemment, ces paramètres doivent être « estimés », générant ainsi une erreur importante d'autant plus qu'elle s'ajoute à l'imprécision des mesures elles-mêmes. La solution consiste donc à se tourner vers d'autres approches, dont l'approche probabiliste. Les notions entourant cette dernière approche sont abordées à la section 3.4.

3.2.2 Propriétés électriques

Les propriétés électriques des milieux poreux saturés sont davantage contrôlées par celles du fluide interstitiel, généralement de l'eau, que par celles des minéraux eux-mêmes (Linde, 2005). Ceci est dû au fait que l'eau est un électrolyte et donc qu'elle possède une certaine quantité de charges libres et de charges mobiles permettant

le passage du courant.

Les propriétés électriques des matériaux se caractérisent par la *conduction électrique*, liée à la migration de charges libres sous l'application d'un champ électrique et par la *polarisation diélectrique*, liée à la séparation de charges de signes opposés sous l'application d'un champ électrique (Schön, 1998)⁵. Les paramètres associés à ces deux phénomènes sont la *conductivité électrique*⁶(σ), exprimée en S/m et la *permittivité diélectrique* (ϵ), exprimée en F/m . Dans un milieu anisotrope, σ et ϵ ne sont plus des scalaires mais des tenseurs.

Lorsqu'un champ électrique est appliqué à un matériau, les charges (libres ou liées) ne répondent pas instantanément à ce champ ce qui explique la nécessité d'exprimer la conductivité électrique et la permittivité sous forme complexe, c'est-à-dire, comportant une partie réelle ('') et une partie imaginaire ('") et on les note

$$\sigma^* = \sigma' + i\sigma'' , \quad (3.10)$$

$$\epsilon^* = \epsilon' - i\epsilon'' . \quad (3.11)$$

En géophysique, il est plus courant de faire référence à la *permittivité relative* (ϵ_r) qu'à la permittivité. Ce paramètre se définit comme le rapport entre la permittivité et celle du vide (Lesme et Friedman, 2005)

⁵La perméabilité magnétique est aussi une propriété des matériaux géologiques mais celle-ci, étant généralement négligeable, n'est pas abordée dans la présente section.

⁶la conductivité électrique est aussi souvent exprimée par son inverse, la résistivité $\rho = 1/\sigma$ [$\Omega - m$].

$$\epsilon_r = \frac{\epsilon'}{\epsilon_0} + \frac{\sigma''}{\omega\epsilon_0} \quad (3.12)$$

où ϵ_0 est la permittivité du vide ($8.89 \times 10^{-12} F/m$).

À de très basses fréquences avoisinant le zéro, la partie réelle de la conductivité électrique peut être mesurée. Inversement, à de très hautes fréquences, c'est la partie réelle de la permittivité qui domine (Schön, 1998). C'est pourquoi un courant direct, considéré comme ayant une fréquence presque nulle, est généralement utilisé pour mesurer la conductivité électrique et que le géoradar, opérant à des fréquences variant entre 1 MHz et 1000 MHz est utilisé pour mesurer la permittivité relative. Une description plus en profondeur de ces techniques est présentée à la section 3.3.

Il existe plusieurs relations pétrophysiques empiriques ou semi-empiriques reliant les propriétés physiques (conductivité électrique, permittivité relative, polarisation induite, etc.) à des propriétés hydrogéologiques (teneur en eau, perméabilité, porosité, etc.). Les deux sections suivantes relèvent quelques-uns des modèles les plus couramment utilisés.

3.2.3 Modèles de conductivité électrique

La formule d'Archie (Archie, 1942) est une relation pétrophysique empirique reliant la conductivité électrique d'un milieu poreux à sa porosité et à son degré de saturation. Cette relation met en évidence un lien de proportionnalité entre la conductivité effective du milieu saturé (σ) et la conductivité électrique de l'eau contenue dans les pores (σ_w)

$$\sigma = \frac{1}{F}\sigma_w . \quad (3.13)$$

Le terme F est appelé *facteur de formation* et reflète la microgéométrie du matériau (Lesme et Friedman, 2005). Ce facteur est aussi lié à la porosité (exprimée ici par le symbole ϕ pour éviter toute confusion avec l'exposant « n » apparaissant dans l'équation 3.15) par un exposant empirique m appelé *facteur de cimentation*, soit

$$F = \phi^{-m} . \quad (3.14)$$

Selon Archie (1942), ce facteur augmente avec la compaction du sol et se situe autour de 1.3 pour des sables non consolidés. Selon Jackson et al. (1978), le facteur de cimentation augmente lorsque la sphéricité des grains diminue alors que les variations dans la taille des grains n'ont que peu d'effet (Lesme et Friedman, 2005).

Dans sa forme la plus générale (éq. 3.15), la formule d'Archie exprime la conductivité électrique d'un matériau résistant contenant un fluide conducteur en fonction de sa porosité

$$\sigma = \sigma_w \phi^m . \quad (3.15)$$

Une des limites de ce modèle est de ne prendre en compte que la porosité totale et non la porosité effective. Or, dans un sol non consolidé, les pores sont rarement tous interconnectés, entraînant, dans ce cas, une surestimation de la conductivité électrique réelle du sol. Par ailleurs, la conduction de surface qui peut survenir dans un sol contenant des particules fines ou encore dans un sol saturé par une solution peu conductrice, n'est pas pris en compte dans l'équation 3.15. Ces effets peuvent mener à sous-estimer la conductivité effective réelle du sol (Lesme et Friedman, 2005). De plus, la conduction de surface a tendance à rendre la relation entre (σ)

et (σ_w) non linéaire (Schön, 1998). Suite à ces constatations, un nouveau modèle, basé sur celui d'Archie, a été proposé afin de tenir de compte de la conduction de surface, que l'on écrit

$$\sigma = \sigma_w \phi^m + \sigma_{surface} , \quad (3.16)$$

où $\sigma_{surface}$ peut être déterminée par la prolongation, jusqu'à l'ordonnée à l'origine, du graphique de (σ) vs (σ_w) (Lesme et Friedman, 2005) ou encore par des formules empiriques basées sur un diamètre représentatif des grains. Selon Schön (1998), la conduction de surface est principalement contrôlée par la taille des grains et la capacité d'échange cationique.

D'autres modèles ont été développés (Waxman et Smith, 1968; Pride, 1994; Revil, 1999) afin de prendre en compte la conduction de surface. Cependant, ces modèles font intervenir des termes reliés à la double-couche de Helmholtz. Ces derniers sont complexes à déterminer à l'échelle d'un site de plusieurs mètres cubes.

Les principales conclusions sur la relation entre la conductivité électrique⁷ d'un sol poreux saturé et sa composition sont les suivantes (Lesme et Friedman, 2005) :

- la conductivité électrique d'un sol poreux saturé est principalement contrôlée par la conductivité électrique du fluide qui le sature. Ainsi, la concentration en sels dissous, la mobilité et la charge des ions en solution ainsi que la température de l'eau détermineront la facilité avec laquelle un courant traverse le milieu ;
- la forme des grains et des pores ainsi que le taux de compaction, contrôlant la porosité, ont aussi un rôle déterminant dans la réponse électrique d'un sol ;

⁷sa partie réelle, indépendante de la fréquence

- la présence de particules fines, particulièrement des argiles, provoque l'apparition d'une conduction de surface, augmentant ainsi considérablement la conductivité d'un sol. Ce phénomène est principalement contrôlé par la surface spécifique des grains et augmente avec celle-ci mais de façon non linéaire par rapport à la conduction électrolytique (Schön, 1998).

Les relations précédentes ne font pas intervenir la conductivité hydraulique. Actuellement, il n'existe aucun modèle robuste permettant de calculer la conductivité hydraulique à partir de la conductivité électrique. Ceci s'explique, en partie, par l'influence des particules fines sur la porosité et sur la conductivité électrique. En effet, Purvance et Andricevic (2000) ont montré qu'en l'absence de particules fines, les conductivités hydraulique et électrique sont positivement corrélées puisqu'ils augmentent tous deux avec la porosité. Par contre, en présence de particules fines, la corrélation a tendance à s'inverser.

La section suivante traite de modèles reliant la permittivité relative (ϵ_r) et les propriétés hydrauliques d'un sol saturé.

3.2.4 Modèles de permittivité

Une attention particulière est aussi portée à la permittivité puisque c'est un paramètre très sensible à la présence d'eau. Tel que mentionné précédemment, la permittivité dépend de la fréquence du champ électrique appliqué. Toutefois, comme le montre la figure 3.2, lorsqu'une certaine fréquence est atteinte (relaxation), la partie réelle de la permittivité reste plutôt constante et la partie imaginaire devient quasiment nulle⁸. Ce phénomène se produit à plusieurs fréquences et selon divers

⁸L'impact de cette approximation sur la permittivité mesurée en radar est analysée en détail dans Giroux et Chouteau (2008b)

mécanismes. La bande fréquences radar pour lequel il se produit est environ de 50 MHz à 2 GHz. L'influence de la permittivité sur la propagation des ondes radar dans le sol est examinée plus en détail à la section 3.3.

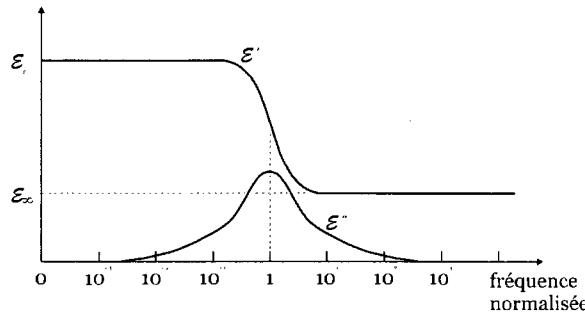


Fig. 3.2: Variations des parties réelles et imaginaires de la permittivité en fonction de la fréquence normalisée à la fréquence de transition ($\sim 3 \times 10^4$ Hz) (Adapté de Grolier et al., 1991)

Pour la suite du document, les constantes diélectriques mentionnées seront celles mesurées dans la bande radar [50 MHz - 2 GHz].

La permittivité relative est un paramètre influencé principalement par la teneur en eau du sol. Cette corrélation s'explique par le fort contraste existant entre la permittivité relative des sols secs ou saturés ($\epsilon_r \approx 4 - 30$) et la permittivité relative de l'eau ($\epsilon_{rw} \approx 80$). Conséquemment, plus la teneur en eau augmente, plus la permittivité relative du mélange augmente. Il est intéressant de mentionner que, contrairement à la conductivité électrique, la permittivité n'est que très peu influencée par la concentration ionique de l'eau (Olhoeft, 1981).

Les deux modèles les plus couramment utilisés pour calculer la permittivité relative d'un milieu poreux sont le modèle appelé « *complex refractive index model* » ou CRIM (Roth et al., 1990) et le modèle empirique de Topp et al. (1980).

Le modèle de CRIM est basé sur l'hypothèse que les ondes électromagnétiques traversent chacun des constituants du sol (eau, sol, air), en série⁹ (Lesme et Friedman, 2005). Pour un milieu saturé, donc comprenant deux phases, ce modèle s'énonce par l'équation suivante

$$\sqrt{\epsilon_{ref}} = \phi\sqrt{\epsilon_{rw}} + (1 - \phi)\sqrt{\epsilon_{rs}} , \quad (3.17)$$

où ϵ_{ref} est la permittivité relative effective du milieu, ϵ_{rw} est la permittivité relative de l'eau ($\epsilon_{rw} \approx 80$), ϵ_{rs} est la permittivité relative des grains solides et ϕ est la porosité. La permittivité relative de la matrice solide varie environ entre 5 et 20. Binley (2007) ont ajusté cette formule afin de tenir compte de la diminution de ϵ_{rw} lorsque la température ou la salinité augmentent (Lesme et Friedman, 2005).

Un second modèle de prédiction de la permittivité diélectrique « maximale » est le modèle de *Maxwell-Garnet*. Il s'énonce par l'équation

$$\epsilon_{ref} = \epsilon_{rw} + 3(1 - \phi)\epsilon_{rw} \frac{\epsilon_{rs} - \epsilon_{rw}}{\epsilon_{rs} + 2\epsilon_{rw} - (1 - \phi)(\epsilon_{rs} - \epsilon_{rw})} . \quad (3.18)$$

Finalement, un dernier modèle principalement utilisé pour des sols non saturés mais qui s'applique aussi à saturation est la formule empirique de Topp (Éq. 3.19) :

$$\epsilon_{ref} = 3.03 + 9.3\theta + 146\theta^2 - 76.7\theta^3 \quad (3.19)$$

où θ est la teneur en eau.

⁹i.e. le temps nécessaire à l'onde pour traverser un milieu poreux équivaut à la somme des temps pour passer à travers chacune des phases

Friedman (1998) a comparé le comportement de cette dernière formule avec la formule « CRIM » mentionnée plus haut appliquée à différents sols. Les résultats montrent que la formule de Topp est plus robuste pour des sols grossiers à moyens. Par contre, la formule présente une moins bonne estimation de la permittivité relative en présence de particules plus fines (Lesme et Friedman, 2005). Il existe d'autres relations plus complexes permettant l'estimation de la permittivité. Cependant, au delà des forces et des faiblesses de chacun des modèles, une erreur potentiellement plus grande réside dans l'estimation de la permittivité relative de la phase solide, principalement pour de grands volumes de sol.

Comme pour la conductivité électrique, aucune relation empirique ou théorique n'existe à ce jour entre la permittivité et la conductivité hydraulique.

3.3 Géophysique en forage

Cette section passe en revue les principes physiques ainsi que les modes d'acquisition et de traitement des données des différentes méthodes géophysiques utilisées lors de ce projet. Ces méthodes sont : le géoradar en forage, la diagraphie électrique et l'essai de perméabilité *in situ* à niveaux variables (*Multilevel slug test*). Contrairement aux deux premières méthodes, l'essai *in situ* se classe davantage parmi les méthodes hydrogéologiques que géophysiques. Il est toutefois considéré ici comme une méthode géophysique au sens large du terme.

3.3.1 Géoradar (GPR)

Le géoradar ou *Ground penetrating radar* (GPR) est une méthode géophysique qui utilise la propagation des ondes de fréquences radio afin de déduire les propriétés

électriques des sols. Dans le domaine environnemental, l'intérêt voué à cet appareil ne cesse de croître dû au fait que son signal est principalement dominé par la présence ou l'absence d'eau. Ses modes d'acquisition multiples, non-destructifs ainsi que la finesse de l'échelle spatiale qu'il peut généralement atteindre, font du géoradar un instrument de plus en plus employé pour la caractérisation géologique et hydrogéologique de la subsurface.

Cette section débute donc par un survol des principes électromagnétiques sur lesquels s'appuient la méthode. S'en suivra une description de l'appareil et des principaux modes d'acquisition. Enfin, l'algorithme d'inversion tomographique utilisé lors de ce projet est présenté.

3.3.1.1 Ondes électromagnétiques et sols - Survol

Tel que mentionné précédemment, le géoradar est un appareil permettant d'investiguer les propriétés électriques du sol en utilisant comme source des ondes électromagnétiques (EM) aux fréquences comprises entre 1 et 1000 MHz. Les fondements de cette méthode géophysique sont issus de la théorie de l'électromagnétisme développée il y a deux siècles par Maxwell. La propagation des ondes EM à travers un milieu peut être décrite par la combinaison des équations de Maxwell, décrivant mathématiquement le lien unissant les paramètres électriques et magnétiques, et des équations constitutives, décrivant le comportement des matériaux face à un champ électromagnétique.

Brièvement, les propriétés clés d'un champ d'ondes EM sont la vitesse de phase v , l'atténuation α et l'impédance électromagnétique Z . Aux basses fréquences, le champ électromagnétique se diffuse dans le matériau en fonction de celles-ci. Ce sont les fréquences utilisées par les méthodes d'induction EM. Par contre, aux

hautes fréquences, c'est-à-dire aux fréquences couvertes par le géoradar, le champ électromagnétique se propage davantage comme une onde à travers le matériau et, pour ϵ , μ et σ constants, la vitesse et l'atténuation deviennent indépendantes de la fréquence. À ces fréquences, dans un milieu homogène à faibles pertes, la propagation des ondes radar se fait avec conservation de la phase et avec une atténuation des amplitudes de type exponentielle décroissante, comme le montre la figure 3.3 (Annan, 2005).

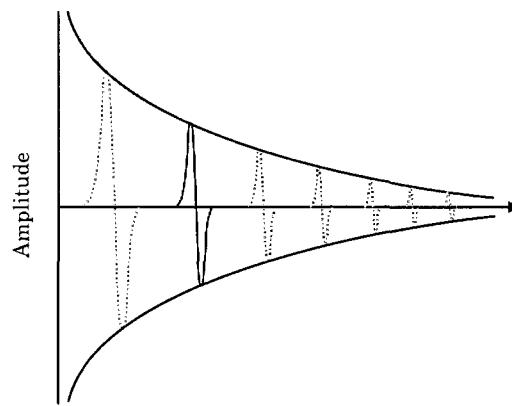


Fig. 3.3: Propagation d'une onde électromagnétique dans un milieu à faibles pertes électriques : la forme du signal reste constante et l'amplitude décroît de manière exponentielle (Adapté de Annan, 2005)

Il y a trois paramètres physiques qui contrôlent la propagation des ondes électromagnétiques à travers un milieu, soit : la perméabilité magnétique μ (capacité d'un milieu à générer un flux magnétique sous l'influence d'un champ magnétique), la permittivité diélectrique ϵ (capacité du sol à se polariser sous l'action d'un champ électrique) et la conductivité électrique σ (disponibilité et mobilité des charges) (Gloaguen, 2004). Ces deux dernières propriétés sont discutées plus en détail dans la section 3.2.2. Quant à la perméabilité magnétique, puisque l'eau et la majorité des minéraux formant les sols sont non ferromagnétiques, leur perméabilité

magnétique est considérée comme étant égale à celle du vide ($\mu_0 = 4\pi \times 10^{-7}$ H/m).

Ainsi, dans un milieu non dispersif, de perméabilité μ_0 et où la conductivité électrique est faible, la vitesse v et l'atténuation α des ondes EM s'expriment par les équations suivantes

$$\alpha = \frac{\omega}{c} \left[\frac{1}{2} \left(\sqrt{\epsilon_r^2 + \left(\frac{\sigma}{\omega \epsilon_0} \right)^2} - \epsilon_r \right) \right]^{1/2}, \quad (3.20)$$

$$\beta = \frac{\omega}{c} \left[\frac{1}{2} \left(\sqrt{\epsilon_r^2 + \left(\frac{\sigma}{\omega \epsilon_0} \right)^2} + \epsilon_r \right) \right]^{1/2}, \quad (3.21)$$

où c est la vitesse de la lumière en m/s. La relation entre la vitesse v de propagation de l'onde et sa phase β est donnée par la formule 3.22

$$v = \frac{\omega}{\beta}. \quad (3.22)$$

C'est la conductivité électrique qui est principalement responsable de l'atténuation des ondes. Dans un milieu composé essentiellement de sable et de gravier, la conductivité électrique est majoritairement due aux ions en solution dans l'eau interstitielle ; ainsi, plus l'eau est conductrice, plus le milieu est atténuant. De plus, la présence de silt ou d'argile provoque une augmentation considérable de la surface spécifique des grains, accentuant substantiellement l'effet de la conduction de surface (phénomène de la double couche). Un tel milieu devient donc très atténuant et limite grandement la profondeur d'investigation du géoradar.

Lors d'un levé radar, l'instrument mesure, en fonction du temps, le champ électrique causé par le passage de l'onde EM. De cette mesure, on peut extraire le temps de parcours de même que l'amplitude des ondes transmises, réfractées ou réfléchies entre un émetteur et un récepteur. Selon le mode d'acquisition des données, le géoradar permet de retrouver soit l'emplacement des principales structures géologiques ou autres (niveau de la nappe souterraine, stratifications, canalisations...), soit les champs de vitesse et d'atténuation. Les deux sections suivantes traitent de ce sujet.

3.3.1.2 Appareillage et modes d'acquisition

Bien qu'il existe plusieurs types de géoradars, le principe général de cet appareil est toujours le même : un émetteur génère une onde électromagnétique qui réagit avec le milieu et un récepteur enregistre le temps de parcours et l'amplitude du signal réfléchi ou transmis.

Le type d'antenne émettrice le plus couramment utilisé est l'antenne *impulsionnelle* où un signal à large bande et de courte durée est créé en appliquant un courant aux bornes d'un dipôle de l'antenne (Gloaguen, 2004). Il existe un autre type d'antenne, dite à *ondes synthétiques*, mais celui-ci ne sera pas abordé ici. L'appareil utilisé lors de cette étude est un radar Malå Geoscience, fabriqué en Suède. C'est un système en forage pour lequel des antennes de type impulsionnel ont été conçues. La fréquence centrale d'émission des antennes utilisées est de 100 MHz.

Les géoradars peuvent être divisés en deux catégories : les radars de surface et les radars en forage où au moins une des antennes est insérée dans un puits. Les radars de surface fonctionnent en mode *réflexion* alors que les radars en forage peuvent fonctionner en mode réflexion mais aussi en mode *transmission*. La différence entre

ces deux modes d'acquisition est présentée aux figures 3.4 (a) et (b).

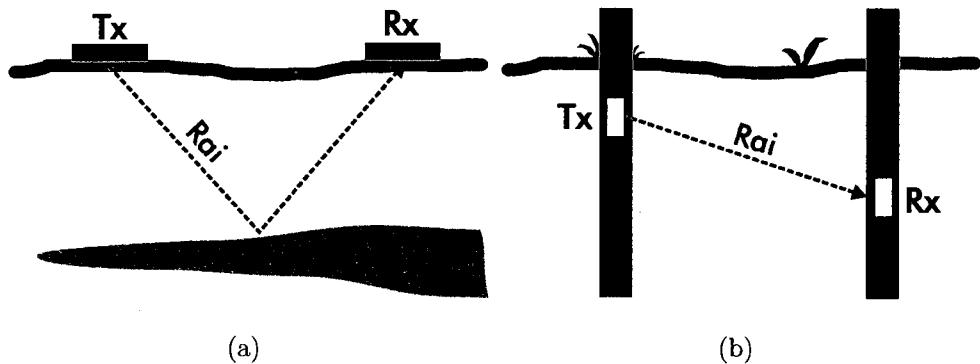


Fig. 3.4: Acquisition en mode réflexion en (a) et acquisition en mode transmission en (b) Tx : émetteur, Rx : récepteur

Le mode réflexion, généralement employé en surface, permet essentiellement de situer l'interface entre des couches ou d'autres structures présentant un contraste de permittivité électrique avec le milieu adjacent. Le mode transmission, quant à lui, est généralement associé à l'acquisition tomographique. Celle-ci consiste à placer l'antenne émettrice dans un forage et l'antenne réceptrice dans un autre. Les positions des deux antennes varient de telle sorte à couvrir adéquatement le plan entre les forages. Pour chacune des paires de positions émetteur-récepteur, le temps de parcours de l'onde directe ainsi que son amplitude sont enregistrés en fonction du temps (Gloaguen, 2004). Le concept d'un mode d'acquisition tomographique est présenté à la figure 3.5.

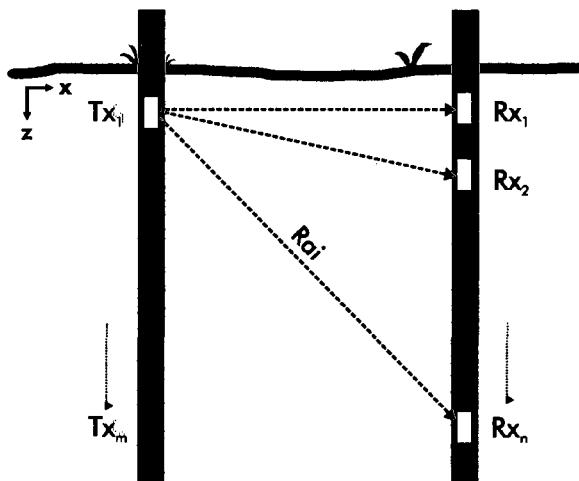


Fig. 3.5: Schéma d'une acquisition en tomographie radar. Tx : émetteur à la position x_i , Rx : récepteur à la position x_j

3.3.1.3 Tomographie

La tomographie est une méthode d'inversion. Cela signifie qu'à partir d'observations de certains paramètres, on tente de reconstruire le modèle physique qui lui est associé. Dans le cas de la tomographie radar, l'inversion vise à obtenir, d'une part, le champ de vitesses à partir des temps de parcours et d'autre part, le champ d'atténuation à partir des amplitudes. On appelle ces inversions respectivement tomographie de vitesse et tomographie d'atténuation. L'inversion est un domaine complexe. C'est pourquoi, seulement les grandes lignes de l'algorithme utilisé seront présentées. Cet algorithme, implémenté sous *Matlab*®, se nomme bh-tomo et a été développé par Giroux et al. (2007).

Avant d'entrer dans le processus d'inversion, une étape préliminaire consiste à « pointer » les temps correspondant aux premières arrivées du signal ainsi que les

extrema des amplitudes de celles-ci. Un exemple de ces pointés est présenté à la figure 3.6.

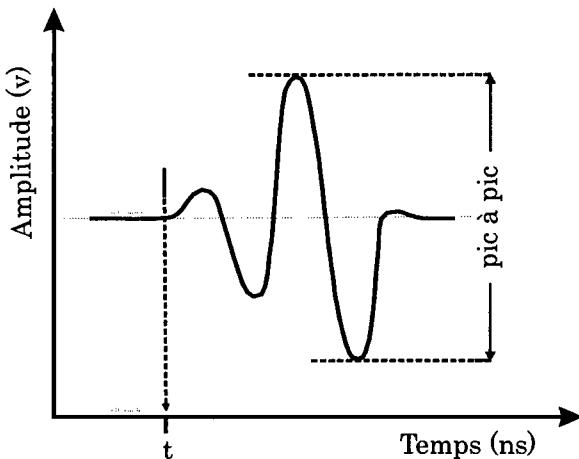


Fig. 3.6: Schéma des pointés du temps et de l'amplitude dite « pic à pic » d'une onde provenant d'un géoradar en forage

Une série de tirs aériens réalisés à différentes distances permet de trouver le temps t_0 correspondant au temps de référence. Ce dernier correspondant au prolongement de la droite des distances en fonction du temps. Une correction analogue doit également être faite pour les amplitudes.

Tomographie de vitesse

Tout d'abord, si l'on pose l'hypothèse que la fréquence de l'onde tend vers l'infini, il est possible de modéliser la trajectoire de celle-ci sous forme d'un rai. La complexité de cette modélisation réside dans le fait que, à l'instar d'un courant électrique, ces rais tendent à emprunter le chemin qui minimise le temps de parcours d'un point à un autre (principe de Fermat). Ainsi, à moins d'être dans un milieu homogène, ces trajectoires seront courbées. L'algorithme implémenté afin de réaliser cette modélisation se nomme « méthode des graphes ». Celle-ci consiste es-

sentiellelement à représenter le plan par un réseau de noeuds reliés entre eux par une distance proportionnelle au temps de parcours de l'onde le long de cette connexion (Gloaguen, 2004). La figure 3.7 résume bien l'approche. Il est important de mentionner que le fait de modéliser des rais courbes augmente, en théorie, la résolution spatiale. Toutefois, si le milieu présente de forts contrastes, les rais convergent vers les zones rapides et il peut devenir impossible d'estimer les vitesses dans les zones lentes, faute de rais (Gloaguen, 2004).

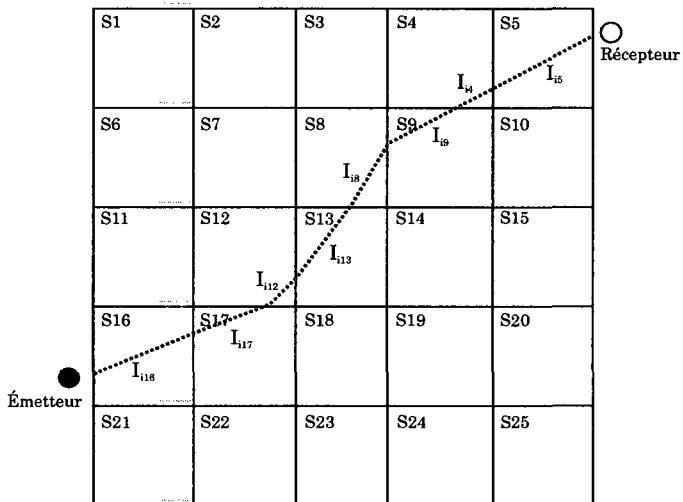


Fig. 3.7: Illustration schématique d'un trajet de rai à travers un milieu hétérogène

En quelques mots, la tomographie de vitesse se résume à la résolution du système linéaire matriciel :

$$Ls=t \ , \quad (3.23)$$

où s est le vecteur contenant les np valeurs de lenteur (inverse de la vitesse) à déterminer, t , le vecteur des nt temps observés et L , la matrice de taille $nt \times np$ contenant les longueurs de chacun des rais pour chacune des cellules. Si l'on

considère que l'on a nt temps observés pour l'ensemble des couples émetteur-récepteur et un plan discrétré en np cellules de lenteur constante, on obtient le système matriciel suivant

$$\mathbf{s} = \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_{np} \end{pmatrix}, \quad \mathbf{t} = \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_{nt} \end{pmatrix} \quad \text{et} \quad \mathbf{L} = \begin{pmatrix} l_{11} & l_{12} & \dots & l_{1np} \\ l_{21} & l_{22} & \dots & l_{2np} \\ \vdots & \vdots & \ddots & \vdots \\ l_{nt1} & l_{nt2} & \dots & l_{ntnp} \end{pmatrix}.$$

Puisque le parcours des ondes dépend des propriétés du milieu traversé et que celles-ci sont inconnues *a priori*, l'inversion est non-linéaire. Ainsi, seule \mathbf{t} est connue et la tâche consiste à déterminer \mathbf{L} et \mathbf{s} . Malheureusement, comme il arrive fréquemment en géophysique, ce problème est dit « mal posé », ce qui signifie que la matrice \mathbf{L} n'est pas directement inversible. Le système doit donc être régularisé afin de le rendre stable.

L'algorithme utilisé fait appel à une approche stochastique afin de régulariser l'inversion. Celui-ci est basé sur le principe que si deux paramètres sont reliés de façon linéaire, comme le sont la lenteur et les temps de parcours, leur covariance l'est aussi. À partir de cet énoncé, il est donc possible de modéliser la covariance des lenteurs à partir de la covariance expérimentale des temps de parcours mesurés. La matrice de covariance entre les données est donc donnée par

$$\text{cov}(\mathbf{t}, \mathbf{t}) = \mathbf{L} \text{cov}(\mathbf{s}, \mathbf{s}) \mathbf{L}^T + \mathbf{C}_0, \quad (3.24)$$

où $\text{cov}(\mathbf{s}, \mathbf{s})$ est la matrice de covariance des lenteurs et \mathbf{C}_0 est une matrice diagonale

représentant l'erreur sur les données.

La détermination de ces matrices de covariance permet de réaliser un cokrigage (Chilès et Delfiner, 1999) du champ de lenteur à partir des temps mesurés. Un avantage majeur de cette approche est de permettre l'ajout de lenteurs connues venant contraindre le processus d'inversion.

Puisque le trajet des rais n'est pas connu *a priori*, un trajet en rais droits fait office de modèle initial. Le cokrigage est ensuite effectué et, par comparaison des résidus des temps modélisés et des temps calculés, les trajets sont mis à jour et utilisés dans l'itération suivante. Ce processus itératif se termine lorsque la différence entre les temps mesurés et calculés, entre deux itérations successives, atteint un seuil suffisamment petit.

Afin de palier à l'effet de lissage engendré par le cokrigage, des simulations sont aussi réalisées, permettant de visualiser les différentes valeurs que peut prendre le champ de vitesse en fonction du modèle de covariance retenu et des données de temps pointées. L'algorithme de simulation implémenté est FFT-MA (*Fast Fourier Transform Moving Average*) qui permet de générer des simulations non conditionnelles de processus gaussiens stationnaires. Le conditionnement aux temps observés est ensuite réalisé par cokrigage.

Il est important de mentionner que, contrairement aux autres méthodes tomographiques (LSQR, SIRT, CG), celle-ci a l'avantage de tenir compte de la structure spatiale du champ de lenteur ainsi que de la corrélation entre les temps de parcours.

Une fois les trajets de chacun des rais connus (matrice L), la tomographie d'atténuation se résout de façon similaire à la tomographie de vitesse¹⁰.

¹⁰Ces algorithmes d'inversion (vitesse et atténuation) ainsi que ceux des simulations sont décrits en détail dans Gloaguen (2004)

Les tomographies de vitesse et d'atténuation permettent finalement de calculer la permittivité et la conductivité électrique en chacun des points de la grille grâce aux équations suivantes, déduites à partir des équations 3.20 et 3.21

$$\epsilon_r = \frac{\left(\frac{1}{v}\right)^2 - \left(\frac{\alpha}{\omega}\right)^2}{\mu_0 \epsilon_0}, \quad (3.25)$$

$$\sigma = \frac{2\alpha}{\mu_0 \epsilon_0}. \quad (3.26)$$

3.3.2 Piézocône multi-capteurs (CPTu)

Le piézocône multi-capteurs est un outil de sondage statique offrant une mesure instantanée *in situ* de plusieurs paramètres géotechniques. Il est composé d'une sonde munie de plusieurs capteurs électroniques, qui est enfoncee dans le sol à une vitesse constante d'environ 1 ou 2 cm/s. Les paramètres pouvant être mesurés par cet appareil sont :

- la résistance en pointe ;
- la friction ;
- la pression interstitielle ;
- l'inclinaison ;
- la température ;
- la résistivité ;
- la vitesse séismique,
- la permittivité.

Ces mesures sont prises tous les 2-3 cm, ce qui donne un profil stratigraphique et mécanique complet le long du forage investigué. C'est un appareil de pointe

de plus en plus apprécié dans l'industrie pour la rapidité et la précision de ses mesures. Parmi l'ensemble des mesures énoncées ci-haut, seules la résistivité et la permittivité sont incluses dans ce projet. Le mode de fonctionnement des capteurs permettant de prendre ces mesures est abordé plus en détail dans ce qui suit.

Sonde de résistivité

Le principe du capteur permettant de mesurer la conductivité électrique est simple. Il est composé de quatre électrodes, alignées verticalement, dont deux qui injectent un courant DC et deux autres qui mesurent la différence de potentiel générée. La loi d'Ohm ainsi que la géométrie des électrodes permettent de calculer la résistivité apparente en un point situé au centre des électrodes. La profondeur à laquelle le courant pénètre dépend de la résistivité du matériau.

La conductivité électrique mesurée dépend de la teneur en eau du sol, de la température de la formation et de la salinité de l'eau interstitielle. Ainsi, lorsque la porosité et les propriétés de l'eau souterraine demeurent relativement constantes, les contrastes de résistivité sont alors dus à un changement dans la fraction argileuse (Butler, 2005).

Sonde de permittivité

La permittivité relative d'un milieu se mesure à l'aide d'un capteur capacitif. Ce dernier est généralement composé de deux électrodes en forme d'anneau et d'un oscillateur. Le sol adjacent au capteur agit comme matériau diélectrique, formant ainsi un circuit fermé avec les électrodes, équivalent à un circuit RLC. Lors de l'enfoncement du dispositif dans le sol, un champ électrique oscillant est d'abord généré entre les deux anneaux. Puis, à l'approche du sol adjacent, la capacitance de couplage du condensateur est perturbée. La fréquence doit ensuite être ajustée pour que le système entre en résonance, ce qui permet de déterminer à capacitance du sol à proximité du capteur et ensuite, sa teneur en eau à l'aide de courbes de

calibration.

La sonde CPTu utilisée lors des présents travaux est de marque Geoprobe.

3.3.3 Essai de perméabilité *in situ* à niveaux multiples

L'essai de perméabilité *in situ* est un essai couramment utilisé en hydrogéologie afin de déterminer la conductivité hydraulique le long d'un puits. Contrairement à l'essai de pompage classique, ce test fournit un estimé beaucoup plus local, comme son nom l'indique, de la conductivité hydraulique d'un aquifère. Toutefois, les essais traditionnels ne donnent qu'une seule valeur moyenne de conductivité hydraulique pour la portion crépinée d'un puits ce qui ne permet pas d'observer les variations verticales de ce paramètre. Or, la compréhension d'un écoulement passe par la connaissance de la variabilité spatiale de la conductivité hydraulique. C'est dans le but de palier à ce problème que l'essai de perméabilité *in situ* à niveaux multiples a été conçu. Ce dernier s'opère exactement comme l'essai standard (à niveau variable) mais à plusieurs niveaux de l'aquifère. La simplicité et la rapidité de son exécution rend cet essai plus intéressant que d'autres essais tels l'essai de traceur ou de pompage. Le principe hydraulique de cet essai ainsi que le mode de traitement des données sont présentés à la section suivante.

3.3.3.1 Acquisition et traitement des données

Le principe hydraulique d'un essai *in situ* classique et d'un essai à niveaux multiples est exactement le même. Celui-ci consiste à imposer de façon quasi instantanée une variation dans la charge hydraulique dans le puits et à mesurer le temps nécessaire pour que la charge initiale se rétablisse. Lors d'un essai à niveaux multiples, cette procédure est exécutée à différents niveaux le long du puits, ceux-ci maintenus

isolés du reste de l'aquifère par des « coussins gonflables » (*packers* en anglais). Un exemple schématique d'un essai *in situ* à niveaux multiples (MLST) est présenté à la figure 3.8.

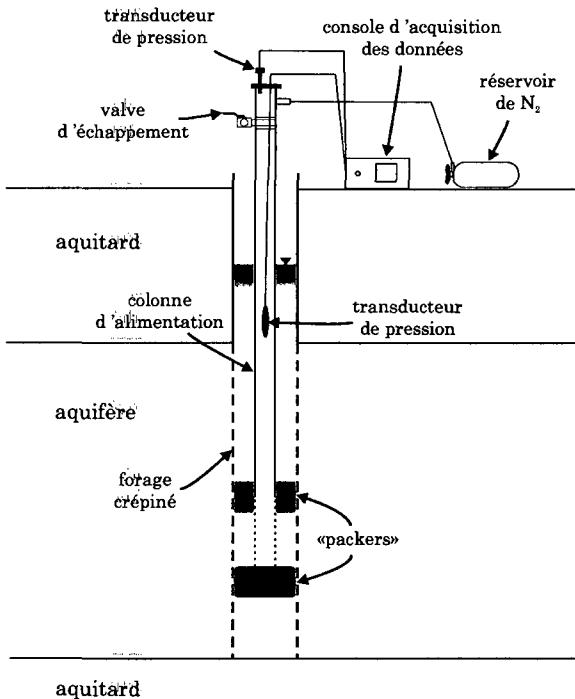


Fig. 3.8: Schématisation d'un essai *in situ* à niveaux multiples

Comme le montre la figure 3.8, l'essai *in situ* aussi bien que l'essai à niveaux variables nécessitent l'installation d'un puits crepiné tout le long de l'aquifère. Bien sûr, dans le cas d'un site contaminé, il est primordial de s'assurer à ne pas propager une contamination de surface en profondeur.

Pour chacun des niveaux, une courbe du temps en fonction du logarithme naturel de la différence de charge est tracée. Théoriquement, une droite est obtenue dont la pente est proportionnelle à la conductivité hydraulique recherchée. La constante de proportionnalité, dans ce cas appelé « coefficient de forme », est ensuite calculée

selon les dimensions du puits (Hvorslev, 1951), ce qui permet ensuite de déduire la conductivité hydraulique. La norme canadienne pour de tels essais mentionne que le niveau piézométrique local doit aussi être mesuré pour chacun des niveaux.

3.4 Outils statistiques

Cette section présente les principaux outils statistiques nécessaires à la réalisation de ce projet. On y retrouve donc le cœur théorique de l'approche proposée. Les notions géostatistiques essentielles à l'élaboration de la méthode sont d'abord exposées suivies du théorème de Bayes et des concepts qui le sous-tendent. Enfin, le chapitre se termine par la revue théorique des simulations séquentielles par approche bayésienne.

3.4.1 Géostatistique

Cette section présente la notion de variogramme et les méthodes d'interpolation que sont le krigage et le cokrigage.

3.4.1.1 Variogramme

Le variogramme permet de décrire le lien spatial existant entre des observations d'un phénomène. Celui-ci est basé sur l'hypothèse que deux observations rapprochées se ressemblent davantage en moyenne que deux observations éloignées.

Le variogramme théorique est défini par la formule

$$\gamma(h) = \frac{1}{2}E[(Z(x) - Z(x + h))^2] , \quad (3.27)$$

où x est le vecteur des coordonnées de chacune des observations $Z(x)$ en 1, 2 ou 3 dimensions et h le vecteur des distances entre chacun de ces points. Si le phénomène étudié présente une anisotropie géométrique, ce vecteur h est défini en distance et en directions.

D'après la formule 3.27, on remarque que le variogramme ne dépend pas des positions exactes de chacune des observations mais bien uniquement de la distance qui les sépare. Cette propriété du variogramme découle de l'hypothèse de *stationnarité de second ordre*¹¹ qui implique que les deux premiers moments (espérance et covariance) de la fonction aléatoire Z existent et soient constants sous l'effet d'une translation¹² (Chilès et Delfiner, 1999).

Puisque l'on ne dispose généralement que de quelques observations et non de toutes les observations possibles d'un phénomène, le variogramme peut être estimé par l'expression

$$\gamma(h, \theta) = \frac{1}{2N(h, \theta)} \sum_{i=1}^{N(h, \theta)} [(Z(x_i) - Z(x_i + h))^2], \quad (3.28)$$

où $Z(x_i)$ est l'observation au point x_i et $N(h, \theta)$ représente le nombre de paires d'observations séparées de h , dans la direction θ .

Une fois le variogramme expérimental obtenu, une fonction analytique, ou modèle, est ajustée avec le plus de précision possible, comme le montre la figure 3.9. Lorsqu'un plateau est atteint, comme dans ce cas-ci, les principales caractéristiques extraites sont

¹¹Il existe une autre hypothèse moins restrictive que la dernière, appelée hypothèse *intrinsèque*, qui permet à la moyenne d'avoir une dérive linéaire mais ce cas ne sera pas abordé dans ce qui suit.

¹²Cette hypothèse s'énonce : $E[Z(x)] = m$ et $Cov(Z(x), Z(x+h)) = C(h)$, m étant la moyenne et $C(h)$, la covariance.

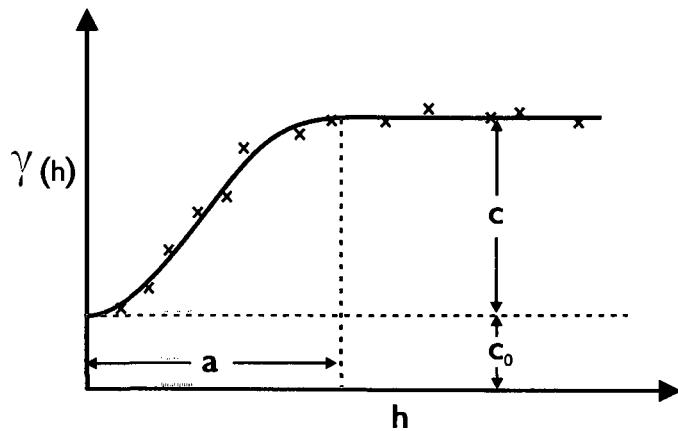


Fig. 3.9: Variogramme expérimental (x), modèle ajusté (trait) et paramètres géostatistiques

- la **portée** (a) : représente la distance à partir de laquelle deux observations ne sont plus corrélées en moyenne ;
- le **palier** (σ^2) : représente la variance de la variable aléatoire, c'est-à-dire l'écart quadratique moyen entre deux observations. On note que la variance totale est souvent le résultat de la somme de deux composantes : la variance à grande échelle (C) et la variance à petite échelle (C_0) ;
- l'**effet pépite** (C_0) : représente les variations à très courte échelle (i.e. $h \rightarrow 0$). Celles-ci peuvent être dues à plusieurs phénomènes comme la présence de micro-structures ayant une portée plus petite que la taille du support, des structures ayant une portée plus petite que la distance entre les observations ou encore à des erreurs de mesure ou de positionnement (Chilès et Delfiner, 1999). On peut détecter la présence d'un effet pépite lorsque le variogramme est non nul à l'origine, comme le montre la figure 3.9.

La présence d'un palier permet aussi de déterminer la covariance ($C(h)$), nécessaire

lors du krigage, entre deux observations séparées d'une distance h , grâce à la relation suivante

$$C(h) = \sigma^2 - \gamma(h) . \quad (3.29)$$

Il existe plusieurs modèles dont les plus courants sont : l'effet pépite et les modèles sphérique, gaussien et exponentiel. Le choix du modèle dépend de la nature du phénomène étudié. Pour des phénomènes présentant une très grande variabilité et donc peu de continuité, comme c'est généralement le cas de la conductivité hydraulique, le modèle sphérique ou exponentiel, combiné à un effet pépite, sont généralement plus adéquats.

Dans le cas où l'on souhaite étudier simultanément la structure spatiale d'un phénomène à l'aide de plusieurs fonctions aléatoires, on a alors recours au variogramme (ou au covariogramme) croisé. Sous sa forme théorique, celui-ci se généralise directement du variogramme simple par l'expression suivante (Chilès et Delfiner, 1999)

$$\gamma_{ij}(h) = \frac{1}{2} E [(Z_i(x) - Z_i(x + h))(Z_j(x) - Z_j(x + h))] , \quad (3.30)$$

où Z_i et Z_j sont deux fonctions aléatoires conjointement intrinsèques, ce qui signifie qu'elles obéissent aux conditions suivantes :

$$\begin{cases} E [Z_i(x) - Z_i(x + h)] = 0 & \text{pour } i = 1, \dots, p \\ Cov [Z_i(x) - Z_i(x + h), Z_j(x) - Z_j(x + h)] = 2\gamma_{ij} & \text{pour } i, j = 1, \dots, p \end{cases}$$

Lorsque l'hypothèse plus restrictive de stationnarité conjointe du second ordre est respectée, il est alors possible d'exprimer la relation entre le variogramme croisé et le covariogramme croisé par

$$\gamma_{ij}(h) = C_{ij}(0) - \frac{1}{2} [C_{ij}(-h) + C_{ij}(+h)] \quad \text{pour } i, j = 1, \dots, p . \quad (3.31)$$

L'équation 3.31 montre que le variogramme croisé ne contient que la partie paire du covariogramme croisé, ce qui se traduit par une perte d'information. Une autre limitation est que son estimation demande l'échantillonnage des variables i et j aux mêmes points, ce qui n'est pas le cas de la covariance croisée. Pour cette raison, on préférera souvent cette dernière.

Dans le cas expérimental à deux variables aléatoires, le variogramme croisé entre chacune des observations situées aux mêmes points, se définit par

$$\gamma_{ij}(h) = \frac{1}{2N_{ij}(h)} \sum_{i,j=1}^{N_{ij}(h)} [(Z_i(x_\beta) - Z_i(x_\alpha)) (Z_j(x_\beta) - Z_j(x_\alpha))] , \quad (3.32)$$

où $N_{ij}(h)$ est le nombre de paires (x_β, x_α) espacées de h (Chilès et Delfiner, 1999).

La modélisation des variogrammes et des variogrammes croisés expérimentaux est une étape difficile mais essentielle à la réalisation d'un cokrigeage (voir section 3.4.1.3).

3.4.1.2 Krigeage simple

Le krigeage est une méthode d'estimation linéaire et sans biais qui se base sur la structure spatiale des observations connues, déterminée à l'aide du variogramme

(voir section précédente). Les valeurs estimées sont obtenues par une combinaison linéaire des valeurs observées. Bien qu'il existe plusieurs types de krigeage (simple, ordinaire, universel, ...), seul le krigeage simple est abordé ici.

Le krigeage simple se limite aux cas stationnaires où l'on connaît la moyenne. Ainsi, l'estimateur qui lui est associé se traduit par

$$Z_0^* = m + \sum_{i=1}^n \lambda_i (Z_i - m) , \quad (3.33)$$

où Z_0^* est l'estimation au point x_0 , m est la moyenne, λ_i est le poids associé à l'observation située en x_i et Z_i est l'observation au point x_i .

L'estimation passe donc par la détermination des poids de krigeage associés à chacune des observations connues tout en minimisant la variance d'estimation σ_e^2 qui s'énonce

$$\sigma_e^2 = Var[Z_0] + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j Cov[Z_i, Z_j] - 2 \sum_{i=1}^n \lambda_i Cov[Z_0, Z_i] , \quad (3.34)$$

où $Var[Z_0]$ est la variance de la variable à estimer et $Cov[Z_i, Z_j]$ est la covariance entre les observations i et j et $Cov[Z_0, Z_i]$, la covariance entre la variable à estimer et l'observation au point i .

Le système de krigeage simple (éq. 3.35) est obtenu en posant les dérivées partielles de σ_e^2 , par rapport à chacun des λ_i , égales à zéro

$$\sum_{j=1}^n \lambda_j Cov[Z_i, Z_j] = Cov[Z_0, Z_i] \quad \text{pour } i = 1..n . \quad (3.35)$$

La variance de krigage devient alors

$$\sigma_{ks}^2 = \text{Var}[Z_0] - \sum_{i=1}^n \lambda_i \text{Cov}[Z_0, Z_i]. \quad (3.36)$$

Tel que mentionné précédemment, le krigage simple ne peut s'effectuer que lorsque le variogramme présente un palier, ce qui permet de déterminer les covariances nécessaires à la résolution du système 3.35. Une fois les poids de krigage obtenus, la valeur estimée est calculée à l'aide de l'équation 3.33.

Il est important de mentionner que le krigage est un interpolateur exact, c'est-à-dire qu'en un point où une valeur est connue, l'estimation sera toujours égale à cette valeur (Chilès et Delfiner, 1999).

3.4.1.3 Cokrigage simple

Le cokrigage s'apparente au krigage en ce sens que c'est aussi un estimateur linéaire et sans biais. Toutefois, il ne se limite pas qu'à une seule variable mais incorpore dans son estimation l'information contenue dans d'autres variables qu'on appelle, *variables secondaires*. Ainsi, si ces dernières sont bien corrélées avec la variable principale, il en résultera une estimation plus précise que celle obtenue par krigage. Encore une fois, il existe plusieurs types de cokrigage mais on se concentrera ici sur le cokrigage simple.

Comme pour le krigage simple, la particularité du cokrigage simple est de connaître les moyennes de chacune des variables. La condition de non biais est obtenue simplement en soustrayant celles-ci aux observations afin de travailler uniquement avec les résidus de moyenne nulle.

Dans le cas où l'on a une variable principale $Z(x)$, de moyenne m_z et une variable secondaire $Y(x)$, de moyenne m_y , l'estimateur prend la forme

$$Z_0^* = m_z + \sum_{i=1}^{nz} \lambda_i (Z_i - m_z) + \sum_{i=1}^{ny} \alpha_i (Y_i - m_y) . \quad (3.37)$$

Selon le même principe de minimisation de la variance d'estimation, on obtient le système de cokrigeage suivant

$$\sum_{j=1}^{nz} \lambda_j Cov[Z_i, Z_j] + \sum_{j=1}^{ny} \alpha_j Cov[Z_i, Y_j] = Cov[Z_0, Z_i] \quad \text{pour } i = 1..nz , \quad (3.38)$$

$$\sum_{j=1}^{nz} \lambda_j Cov[Y_i, Z_j] + \sum_{j=1}^{ny} \alpha_j Cov[Y_i, Y_j] = Cov[Z_0, Y_i] \quad \text{pour } i = 1..ny . \quad (3.39)$$

Et la variance d'estimation s'écrit

$$\sigma_{ck}^2 = Var[Z_0] - \sum_{i=1}^{nz} \lambda_i Cov[Z_0, Z_i] - \sum_{i=1}^{ny} \alpha_i Cov[Z_0, Y_i] . \quad (3.40)$$

La complexité du cokrigeage réside dans la détermination de modèles admissibles pour chacune des covariances simples, $Cov[Z_i, Z_j]$ et $Cov[Y_i, Y_j]$, et croisées, $Cov[Y_i, Z_j]$.

3.4.2 Théorème de Bayes

Le théorème de Bayes est introduit ici puisqu'il est au cœur de l'approche proposée dans ce mémoire. Développé par le révérend anglais Thomas Bayes et publié en

1763, ce théorème a été conçu afin de pouvoir inclure dans le calcul d'une probabilité conditionnelle toutes autres probabilités provenant d'observations ayant un lien avec l'événement étudié.

Par définition, la probabilité conditionnelle qu'un événement A se produise étant donné un second événement B est (Walpole et al., 2002) :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)P(B|A)}{P(B)} \quad (3.41)$$

où $P(A \cap B)$ représente la probabilité conjointe que les événements A et B se produisent simultanément. Cette notion est abordée plus en détail à la section suivante (3.4.2.1). Dans le terme de droite de l'égalité, $P(A)$ est appelée la probabilité *non-conditionnelle ou marginale* de l'événement A.

Issu de la théorie des probabilités conditionnelles (éq. 3.41), le théorème de Bayes s'énonce comme suit :

Soit A_1, \dots, A_k , k événements mutuellement exclusifs (ne pouvant pas se produire simultanément) et B_1, \dots, B_m , m événements observables mutuellement exclusifs, alors

$$P(A_i|B_j) = \frac{P(B_j|A_i) P(A_i)}{\sum_{j=1}^k P(B_j|A_j) P(A_j)}, \quad (3.42)$$

où les probabilités $P(A_1), \dots, P(A_k)$ sont appelées probabilités *a priori* et contiennent l'information statistique préalablement connue sur les phénomènes étudiés. Les probabilités $P(B_j|A_i)$ sont appelées *fonctions de vraisemblance* (ou *likelihoods* en anglais) et décrivent les paramètres A_i d'une loi statistique en fonction des observations B_j connues. Finalement, les résultats $P(A_i|B_j)$ se nomment probabilités *a posteriori*, qui signifie l'actualisation de l'*a priori* une fois les événements

B_j observés.

L'estimation de la fonction de vraisemblance et de la distribution *a priori* sont les deux étapes critiques de l'application du théorème de Bayes. Les méthodes employées pour y arriver sont détaillées dans les sections suivantes.

3.4.2.1 Densité de probabilité - Estimateur de noyau

L'approche bayésienne adoptée dans ce projet nécessite la détermination de la densité de probabilité multivariable. L'estimation par noyau (*Kernel density estimation*) est une des méthodes non-paramétriques permettant de déterminer la densité de probabilité multivariable.

Le concept de l'estimateur par noyau s'apparente grandement à celui de l'histogramme mais de façon continue et lisse. En effet, plutôt que d'estimer la densité de probabilité d'un point x en additionnant le nombre d'observations incluses dans une « boîte » de largeur fixe h , la méthode de noyau remplace la boîte par une fonction en forme de « cloche » centrée en x , de variance h . Ainsi, plus une observation se rapproche de ce point, plus sa contribution est importante. La densité de probabilité finale est obtenue par la somme de chacune de ces contributions que l'on normalise ensuite pour obtenir une fonction de densité. Dans un cas « multivariable », cette fonction sera obtenue par la multiplication des densités de probabilité correspondant à chacune des variables aléatoires.

L'estimateur de noyau pour une seule variable aléatoire se définit comme suit (Zuber, 1999) : soit x_1, x_2, \dots, x_n, n observations d'une variable aléatoire, alors l'estimateur par noyau permettant de déterminer la densité de probabilité au point x s'écrit

$$\hat{f}_h(x) = \frac{1}{n \cdot h} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right), \quad (3.43)$$

où K est appelée le noyau et h est une « fenêtre » contrôlant le degré de lissage de l'estimation.

Le noyau est une fonction de pondération dont l'intégrale est égale à l'unité. Parmi les noyaux les plus courants (uniforme, triangle, epanechnikov, quadratique, cubique, gaussien, et circulaire), c'est le noyau gaussien (moyenne nulle et variance de un) qui a été retenu puisque l'on est dans un contexte gaussien. Celui-ci s'énonce par la formule suivante

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \quad (3.44)$$

Il est possible de généraliser l'équation 3.43 à d dimensions. Dans ce cas, l'estimateur de noyau devient

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_1 \dots h_d} K\left(\frac{x_1 - x_{i1}}{h_1}, \dots, \frac{x_d - x_{id}}{h_d}\right), \quad (3.45)$$

où h_1, \dots, h_d sont les « fenêtres » pour chacune des dimensions.

La solution la plus simple permettant de combiner chacun des noyaux k associés à une dimension (noyau univariable) est de les multiplier entre eux afin de déterminer la densité de probabilité conjointe des d variables aléatoires,

$$K(x) = K(x_1) \cdot \dots \cdot K(x_d). \quad (3.46)$$

De là, on tire l'estimateur de noyau

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n \left\{ \prod_{j=1}^d \frac{1}{h_j} K \left(\frac{x_j - x_{ij}}{h_j} \right) \right\} . \quad (3.47)$$

L'étape délicate dans la détermination d'une fonction de densité est l'estimation des largeurs de fenêtre (h). Plus la fenêtre est petite, plus la fonction varie jusqu'à provoquer l'apparition de détails artificiels non désirés. Inversement, plus la fenêtre est grande, plus la fonction est lissée jusqu'à effacer les caractéristiques du noyau. En théorie, la fenêtre idéale est celle qui minimise l'erreur quadratique moyenne intégrée (EQMI) de l'estimateur, c'est-à-dire la somme des erreurs quadratiques moyennes (EQM) pour chacune des valeurs de x . Formellement, cela s'écrit

$$EQM_x(\hat{f}) = E \left[(\hat{f}(x) - f(x))^2 \right] , \quad (3.48)$$

$$EQMI(\hat{f}) = \int EQM_x(\hat{f}) dx . \quad (3.49)$$

Cependant, dans le cas où l'on utilise un noyau gaussien et où la vraie fonction de densité est aussi gaussienne de variance σ^2 , le largeur de fenêtre qui permet de minimiser l'EQMI est donnée par (Silverman, 1986)

$$h_d = 0.9 A n^{-0.2} , \quad (3.50)$$

où $A = \text{minimum}(\text{écart-type}, \text{écart interquartile}/1.34)$.

Il existe plusieurs formules analytiques simples pour estimer *a priori* la largeur de

fenêtre idéale. Cependant, puisqu'on se limite ici à un cas à trois variables, une méthode visuelle itérative est aussi adéquate.

3.4.3 Simulations conditionnelles

Pour l'estimation de variables très hétérogènes comme la porosité ou la conductivité hydraulique, les simulations sont des outils statistiques intéressants permettant de mieux quantifier la variabilité spatiale du phénomène que le krigage ou le cokrigage. En effet, contrairement à ces deux dernières méthodes d'estimation, les simulations ont la propriété de respecter l'histogramme ainsi que le variogramme des données (Goovaerts, 1997). Ainsi, il est possible de générer une infinité de simulations, appelées *réalisations*, toutes statistiquement équivalentes et indépendantes les unes des autres (Chilès et Delfiner, 1999).

Les méthodes de simulations peuvent être divisées en deux catégories : les méthodes gaussiennes et non-gaussiennes. De plus, si l'algorithme de simulation respecte les valeurs observées, qu'on appelle alors *données conditionnantes*, les simulations sont dites *conditionnelles*. Dans le cas contraire, ce sont des simulations non-conditionnelles. Les algorithmes de simulation présentés dans ce mémoire, soit la simulation séquentielle gaussienne (SGS), la cosimulation et la simulation séquentielle gaussienne par approche bayésienne, sont, toutes les trois, des simulations gaussiennes conditionnelles.

3.4.3.1 Simulation séquentielle gaussienne

Toute simulation gaussienne et conditionnelle a comme point de départ N valeurs conditionnantes, normales¹³.

La simulation séquentielle gaussienne s'opère en quatre étapes (Goovaerts, 1997) :

1. La première étape consiste à définir un chemin aléatoire visitant, une seule fois, chacun des noeuds à simuler.
2. À chacun des noeuds, l'espérance et la variance conditionnelles de la variable aléatoire sont déterminées par krigeage simple (voir section 3.4.1.2) des points conditionnans et des points précédemment simulés. (Le krigeage simple est utilisé car, étant donné que la variable aléatoire est gaussienne, la moyenne et la variance ainsi obtenues sont équivalentes à l'espérance conditionnelle et à la variance conditionnelle de la variable aléatoire (Goovaerts, 1997)

$$E\{Z(x_1)|Z(x_2)\} = m + C_{12}C_{22}^{-1}(Z(x_2) - m) ,$$

$$Var\{Z(x_1)|Z(x_2)\} = C_{11} - C_{12}C_{22}^{-1}C_{21} ,$$

où $Z(x)$ est un vecteur multivariable de distribution gaussienne, m est la moyenne connue et où C_{ij} sont les covariances entre les observations aux points x_1 et x_2 .)

3. On tire aléatoirement une valeur de la fonction de densité cumulative (*cdf*) obtenue par krigeage et on ajoute cette nouvelle valeur aux noeuds déjà simulés.

¹³Si les données ne sont pas gaussiennes, il est possible de leur faire subir une transformation pour qu'elles s'en approchent. Cependant, la transformation gaussienne de la loi marginale ne garantit pas que la distribution conjointe soit aussi gaussienne.

4. On visite le noeud suivant en répétant les étapes 2 et 3 jusqu'à ce que tous les noeuds de la grille aient été visités.

Une infinité de simulations peut être réalisée en modifiant le chemin aléatoire et en tirant toujours aléatoirement une valeur de la *cdf* obtenue par krigeage simple.

Le temps de calcul est souvent la faiblesse de cette méthode lorsque, vers la fin de l'algorithme, il y a beaucoup de noeuds inclus dans le krigeage. Une des façons de pallier à ce problème est de limiter le krigeage à un voisinage restreint de noeuds autour du noeud à simuler (grâce à l'effet d'écran). La méthode devient alors seulement approximative.

3.4.3.2 Cosimulation

L'algorithme SGS, tel que présenté précédemment, permet de ne simuler qu'une seule variable. Or, il est possible de disposer de plusieurs types de données (hydrogéologiques, géophysiques, chimiques,...) reliées au même phénomène. Dans le cas où chacune de ces variables montre un comportement gaussien et où le lien les unissant est plutôt linéaire, des simulations conditionnelles peuvent être réalisées par cosimulation.

La cosimulation est le prolongement multivariable directe de la méthode SGS en ce sens que les étapes à suivre sont identiques à la seule différence que le krigeage simple est remplacé par un cokrigeage simple (voir section 3.4.1.3). Ainsi, à chacun des noeuds, l'espérance et la variance conditionnelles sont calculées par cokrigeage à l'aide des variogrammes simples et croisées de chacune des variables. La complexité de la méthode réside principalement dans la détermination de ces derniers.

La cosimulation est une méthode efficace lorsque la relation entre les variables

secondaires et la variable principale tend à être linéaire. Dans le cas contraire, une autre méthode de simulation est sans doute préférable. De plus, à l'instar de la méthode SGS, la cosimulation peut devenir extrêmement coûteuse en temps de calcul.

3.4.3.3 Simulation séquentielle gaussienne bayésienne

Selon Gastaldi et al. (1998), les relations linéaires sont parfois trop limitatives pour la prédiction de certains phénomènes.

Il convient de réécrire la formule de Bayes sous une forme moins abstraite où k est la variable principale (ici, la conductivité hydraulique) et σ et ϵ sont les variables secondaires (ici, la conductivité électrique et la permittivité) desquelles on possède des observations à chacun des points où l'on souhaite simuler k

$$\underbrace{P(k|\sigma, \epsilon_r)}_{a \text{ posteriori}} = \frac{1}{P(\sigma, \epsilon_r)} \cdot \underbrace{P(\sigma, \epsilon_r|k)}_{\text{vraisemblance}} \cdot \underbrace{P(k)}_{a \text{ priori}}. \quad (3.51)$$

Ainsi, la simulation par approche bayésienne se déroule en cinq étapes (la densité de probabilité conjointe étant d'abord obtenue par la méthode des noyaux présentée à la section 3.4.2.1) :

1. la première étape consiste à définir un chemin aléatoire visitant, une seule fois, chacun des noeuds à simuler ;
2. ensuite, à chacun des noeuds :
 - (a) détermination de la distribution *a priori*

À l'instar de la méthode SGS, l'espérance et la variance conditionnelles de la variable principale sont déterminées par krigeage simple

des noeuds conditionnans et des noeuds précédemment simulés ;

(b) détermination de la *vraisemblance*

La vraisemblance est ensuite obtenue à partir de la densité de probabilité conjointe des différentes variables, considérant qu'au noeud à simuler, les variables secondaires sont connues ;

(c) détermination de la distribution *a posteriori*

Ces deux distributions, évaluées aux mêmes noeuds, sont ensuite multipliées et le résultat est divisé par la probabilité conjointe $P(\sigma, \epsilon)$. Cette dernière probabilité est obtenue par la sommation des densités de la distribution *a posteriori* non-normalisée ;

(d) de cette distribution *a posteriori*, une valeur du paramètre (k) est tirée aléatoirement et ajoutée à la grille des valeurs simulées ;

3. cette procédure est répétée jusqu'à ce que tous les noeuds de la grille (ou autre configuration de noeuds) aient été visités.

La force de cet algorithme réside d'abord dans l'*a priori*, qui permet de tenir compte des observations que l'on a de la variable principale. Deuxièmement, la vraisemblance, qui s'obtient à l'aide de la densité de probabilité conjointe, permet d'inclure dans l'estimation toutes autres variables secondaires connues. De plus, une densité de probabilité conjointe calculée par la méthode des noyaux ne nécessite pas la détermination de relations numériques entre chacune des variables. Finalement, les relations entre celles-ci peuvent être non-linéaires.

Bien que l'approche bayésienne soit utilisée dans une multitude de domaines, à notre connaissance, il semble qu'aucun autre auteur n'ait appliqué la méthodologie présentée par Gastaldi et al. (1998) à l'hydrogéophysique.

CHAPITRE 4

MÉTHODOLOGIE

Ce chapitre présente la méthode proposée dans ce mémoire. Celle-ci est d'abord résumée dans un schéma où chacune des étapes est brièvement décrite, puis, les différentes approches employées pour valider la méthode sont exposées.

4.1 Approche proposée

L'objectif est de développer une procédure *in situ* permettant d'estimer la conductivité hydraulique d'un site, à une échelle fine (~ 15 cm) sur un plan compris entre deux forages, par la combinaison de données géophysiques et hydrogéologiques en forages. Pour ce faire, la méthode de simulation stochastique séquentielle gaussienne par approche bayésienne, décrite à la section 3.4.3.3, a été adoptée. Plusieurs aspects de ce problème hydrogéophysique viennent appuyer ce choix, entre autres, la non linéarité des relations pétrophysiques, la grande variabilité des paramètres hydrauliques à une telle échelle ainsi que la formulation particulière de l'approche bayésienne permettant d'inclure dans la simulation toute information de conductivité hydraulique préalablement connue. Cette approche est d'abord testée sur un modèle synthétique de porosité pour ensuite être appliquée à un cas réel.

Bien que la méthode permette d'incorporer plusieurs types de données tant géophysiques qu'hydrogéologiques, ce projet se limite à deux sources de données :

1. hydrogéologique :
 - mesures de la conductivité hydraulique le long de forages, à une échelle

d'environ 15 cm, à l'aide d'un essai *in situ* à niveaux multiples (voir section 3.3.3) ;

2. géophysique :

- mesures de la conductivité électrique et de la permittivité le long de forages, à une échelle d'environ 2 à 3 cm, à l'aide d'une sonde CPTu (voir section 3.3.2) ;
- mesures de la vitesse et de l'atténuation des ondes radar entre forages, à une échelle d'environ 30 cm, à l'aide d'une acquisition tomographique radar (voir section 3.3.1). Ces deux paramètres sont ensuite transformés en conductivité électrique et permittivité.

4.2 Déroulement de la simulation

La figure 4.1 présente une synthèse schématique d'une simulation par approche bayésienne. Chacune des cinq étapes est ensuite expliquée sommairement.

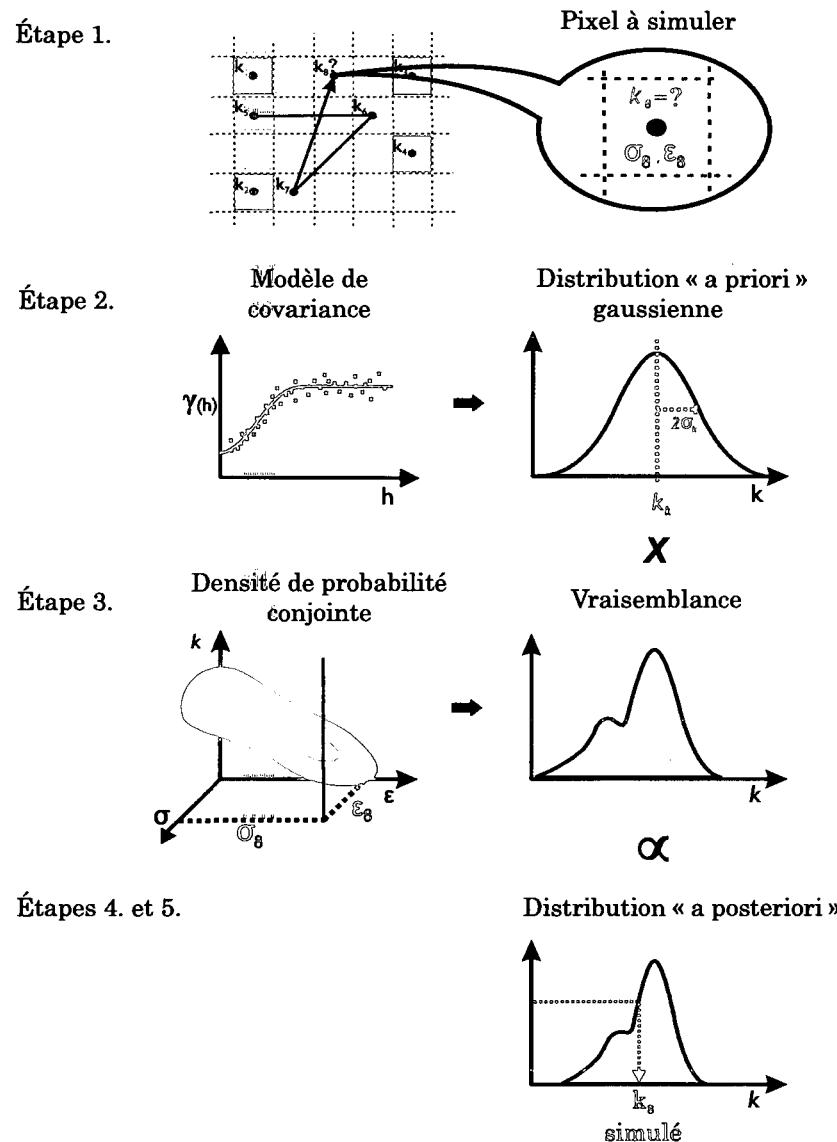


Fig. 4.1: Résumé schématique de la méthode proposée (Inspiré de Gastaldi et al., 1998)

Étape 1. Définition d'un chemin aléatoire

La première étape consiste à définir un chemin aléatoire visitant, une seule fois, chacun des points à simuler. Les noeuds en rouge correspondent aux données de conductivité hydraulique préalablement connues et ceux en jaune correspondent aux noeuds déjà simulés.

Les étapes suivantes visent à simuler une valeur de conductivité hydraulique au noeud encore indéterminé (#8 sur la figure 4.1).

Étape 2. Détermination de la distribution *a priori*

La distribution appelée *a priori* correspond à la probabilité $P(K|K_1, K_2, \dots, K_7)$, où K_1, K_2, \dots, K_7 représentent les valeurs de conductivité hydraulique connues jusqu'ici. Ce paramètre ayant un comportement gaussien, la distribution recherchée l'est aussi. Dans un tel cas, la détermination de cette première distribution se fait par krigeage simple. Ce dernier est réalisé à l'aide de la fonction *cokri* (Marcotte, 1991), implémentée en Matlab[©].

Étape 3. Détermination de la *vraisemblance*

Au noeud considéré ici, la vraisemblance se définit comme $P(\sigma_8, \epsilon_{r8}|K)$ et correspond à la probabilité d'avoir simultanément les valeurs de conductivité électrique σ_8 et de permittivité ϵ_{r8} , connues au noeud à simuler grâce à la tomographie radar, ceci pour toutes valeurs de K .

La détermination de cette distribution nécessite la connaissance de la densité de probabilité conjointe entre la conductivité hydraulique, la conductivité électrique et la permittivité. Tel que mentionné précédemment, cette distribution conjointe est calculée à l'aide de la méthode des noyaux présentée à la section 3.4.2.1. Étant donné le caractère gaussien du paramètre recherché, c'est un noyau gaussien qui a été adopté. Les largeurs de bande associées à

cette méthode ont été ajustées en se basant sur la formule empirique proposée par Silverman (1986).

Étape 4. Détermination de la distribution *a posteriori*

La distribution finale est appelée *a posteriori*. Elle est définie par la probabilité $P(K|\sigma_8, \epsilon_{r8}, K_1, \dots, K_7)$, qui correspond à la distribution des valeurs de K étant donnée les paramètres géophysiques à ce noeud ainsi que les valeurs de K déjà connues ou simulées. Cette distribution s'obtient par la multiplication de l'*a priori* et de la vraisemblance, évalués aux mêmes points. Ce résultat est ensuite divisé par la probabilité conjointe $P(\sigma, \epsilon_r)$ afin d'obtenir une distribution ayant une aire sous la courbe égale à 1.

Étape 5. Sélection aléatoire d'une valeur de K

De cette distribution *a posteriori*, une valeur de K est tirée aléatoirement et ajoutée à la grille des valeurs simulées.

Cette procédure est répétée jusqu'à ce que tous les noeuds de la grille (ou autre configuration de noeuds) aient été visités. Après avoir simulé un certain nombre de noeuds, on doit restreindre le voisinage afin de conserver une taille raisonnable du système de krigeage simple. La méthode, en ce sens, est approximative tout comme la simulation séquentielle gaussienne classique.

4.3 Validation de la méthode

Afin de valider la robustesse de la méthode proposée, celle-ci a d'abord été testée sur un modèle synthétique de porosité. Bien que, dans ce cas, le paramètre recherché ne soit pas la conductivité hydraulique, le résultat obtenu grâce à ce modèle

offre l'avantage de pouvoir être comparé à d'autres résultats obtenus à l'aide de méthodes d'estimation ou de simulation différentes. Ainsi, les méthodes auxquelles la simulation bayésienne est comparée sont

1. méthodes d'estimation directe :
 - (a) krigeage simple ;
 - (b) porosité obtenue par la formule de Topp et al. (1980) (seulement possible pour la porosité) ;
2. méthodes de simulation :
 - (a) simulation séquentielle gaussienne ;
 - (b) cosimulation ;
 - (c) simulation séquentielle sans *a priori*.

Lorsque l'on dispose d'un grand nombre de réalisations pour une même méthode, la variance de la valeur à chacun des pixels permet de repérer les zones moins bien connues et les zones plus précises. Cependant, une information précise peut être biaisée. C'est pourquoi, une comparaison avec le modèle de base est aussi faite à l'aide du coefficient de corrélation qui lui mesure le degré de similitude de la forme simulée avec le modèle de base.

Une bonne méthode devrait montrer la meilleure corrélation possible avec le modèle. De plus, la variance de chacun des pixels calculée à partir d'un grand nombre de simulations bayésiennes devrait être la plus faible possible, sous condition toutefois de bien reproduire l'incertitude du problème. On ne peut malheureusement pas vérifier cette dernière condition à partir d'une seule image et d'un seul échantillon.

CHAPITRE 5

RÉSULTATS - MODÈLE SYNTHÉTIQUE

5.1 Données et paramètres de pré-simulation

5.1.1 Modèles et données géophysiques simulés

Avant de travailler avec des données réelles, la méthode a d'abord été testée à l'aide d'un modèle synthétique de porosité (fig. 5.1). Celui-ci représente une coupe 2D de 7 mètres de largeur et de 9 mètres de profondeur d'un sol non-consolidé, qu'on suppose saturé, possiblement composé de silt, de sable et de gravier. Ce modèle provient d'une image numérique d'un dépôt de sable où une valeur de porosité a été associée à chacun des niveaux de gris. Celle-ci varie de 0.08 à 0.53¹ et est répartie de façon hétérogène quoiqu'il soit possible de distinguer certaines structures sub-horizontales.

On suppose qu'un premier forage est situé à $x = 0$ m et un second à $x = 6$ m². C'est un modèle ayant une variance de 0.0015 et un comportement plutôt gaussien, comme le montre la figure 5.2 suivante.

¹on note toutefois que l'intervalle correspondant à 5% et 95% de la distribution est 0.19 et 0.31

²L'emplacement de ces forages est représenté, sur chacune des figures, par deux lignes noires.

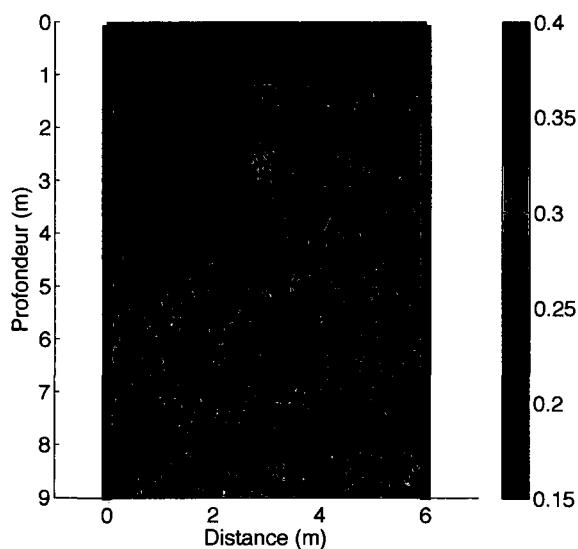


Fig. 5.1: Modèle de porosité compris entre deux forages

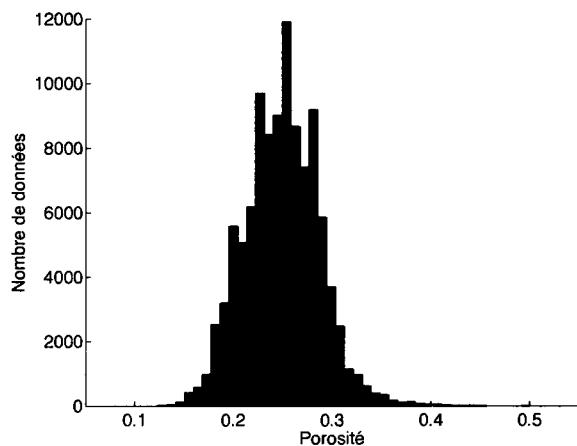


Fig. 5.2: Histogramme du modèle de porosité

À partir du modèle de porosité, un algorithme développé par Giroux et Chouteau (2008a) a permis de générer des modèles de permittivité et de conductivité électriques. Le logiciel MoRad (Giroux et Chouteau, 2004) a ensuite permis la modélisation des traces radar synthétiques. Le pointé des premières arrivées ainsi

que des amplitudes pic-à-pic des premières arrivées a été effectué à l'aide du logiciel bh-tomo (Giroux et al., 2007). Par la suite, les tomographies stochastiques de vitesse et d'atténuation, présentées à la figure 5.3, ont été calculées (Gloaguen, 2004)³.

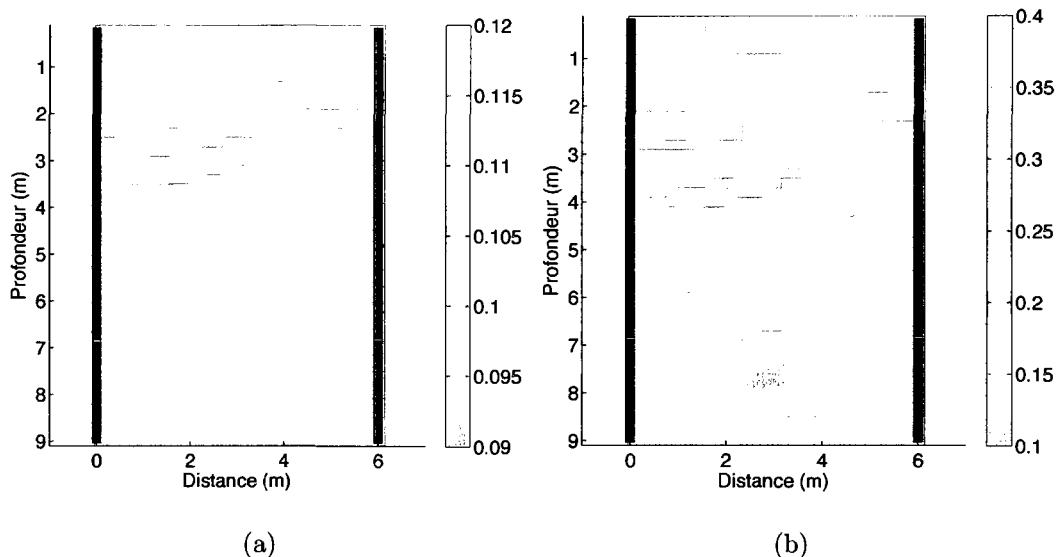


Fig. 5.3: Modélisation (a) des vitesses (m/ns) et (b) des atténuations (np/m) radar obtenues à partir du modèle synthétique de porosité

Ces tomogrammes mettent en évidence la présence de couches à haute vitesse en superficie, à une profondeur de 3 m ainsi qu'à une profondeur de 4.5 m. Ces deux dernières couches sont légèrement inclinées. De plus, on note la présence d'une couche horizontale de basse vitesse à environ 8.5 m de profondeur sous laquelle se trouve une dernière couche de plus haute vitesse. Les mêmes structures sont présentes sur le tomogramme d'atténuation. Il est toutefois important de mentionner qu'un problème lié aux ondes guidées tend à surestimer, par endroits, la valeur de l'atténuation. En effet, de forts contrastes de conductivité électrique

³sans inclure de contraintes le long des forages

tendent à canaliser les ondes le long des couches les moins atténuantes (Ellefsen, 1999). Dans un cas réel, ce problème peut être contourné lorsque l'on dispose de valeurs de conductivité électrique le long des forages adjacents à la tomographie, comme on le verra à la section 6.

À partir du modèle de porosité et des tomogrammes synthétiques, les valeurs le long des forages ont été extraites. Dans un cas réel, celles-ci auraient été disponibles à l'aide des sondes de conductivité et de permittivité électriques et de porosité.

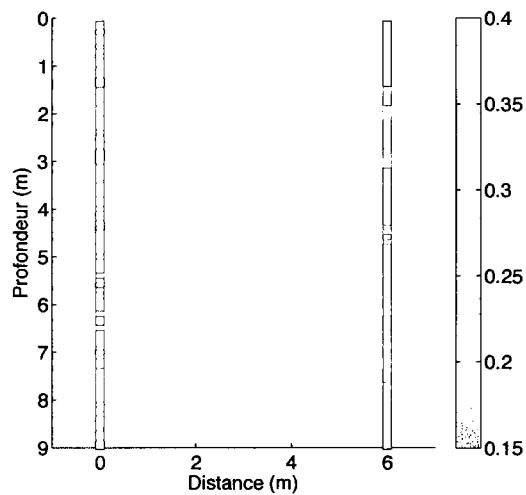


Fig. 5.4: Porosités extraites le long des forages

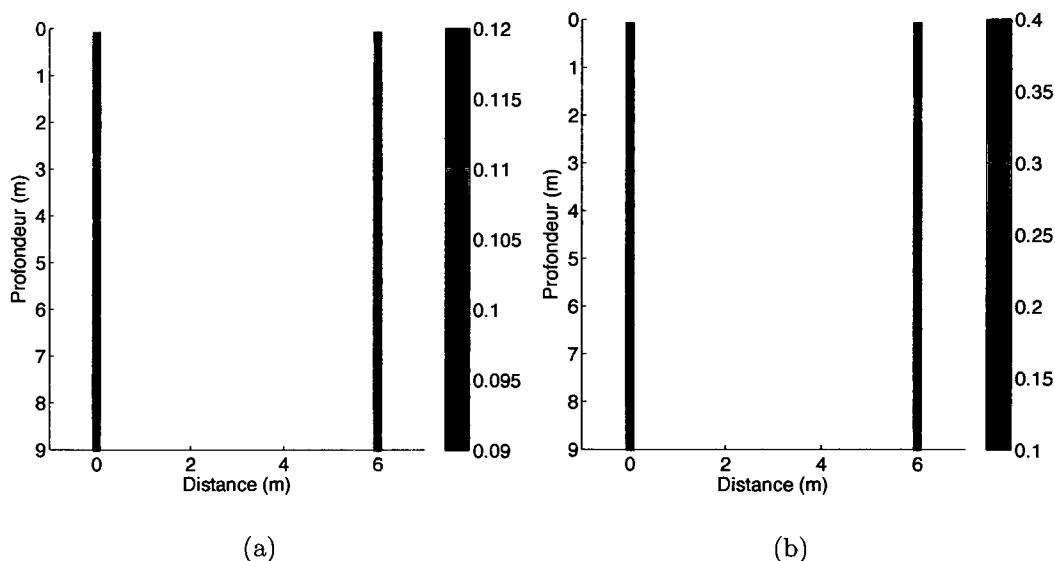


Fig. 5.5: Vitesses (a) et atténuations (b) extraites le long des forages

La variance des porosités le long des forages est de 0.0013, valeur plus faible que celle du modèle complet. Cette « perte d'information », liée à l'échantillonnage, peut mener à une sous-estimation de la variance des simulation séquentielle gaussienne classique et bayésienne.

5.1.2 Densité de probabilité conjointe

Grâce aux valeurs de porosité, de vitesse et d'atténuation connues aux mêmes points le long des forages, une densité de probabilité conjointe à trois variables a été calculée par la méthode des noyaux (éq. 3.47). Les relations entre les deux paramètres géophysiques et la porosité sont présentées aux graphiques 5.6 (a) et (b).

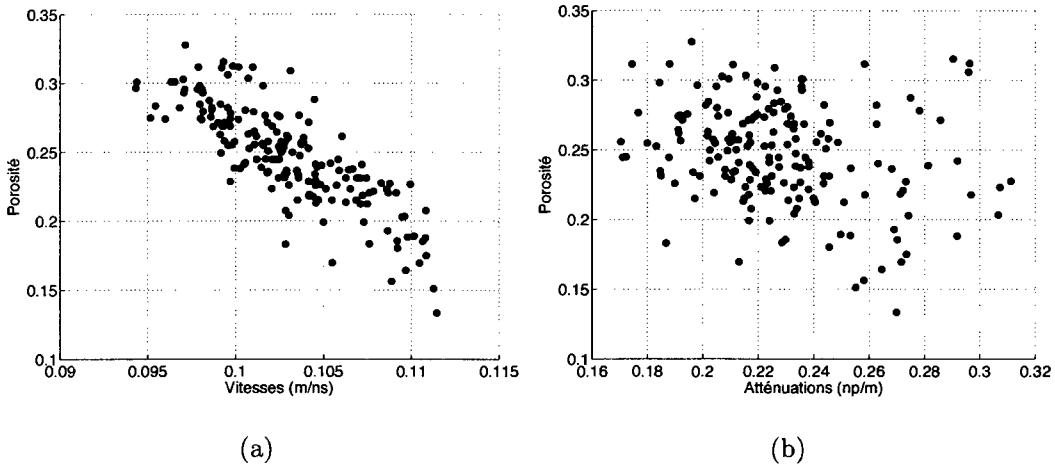


Fig. 5.6: Graphique des vitesses (m/ns) (a) et des atténuations (np/m) (b) le long des forages selon la porosité

On remarque que, plus la vitesse augmente, plus la porosité diminue et ce, de façon plutôt linéaire. Étant donnée la forte influence de la permittivité sur la vitesse des ondes, ce résultat est en accord avec les formules empiriques analytiques comme celle de Topp et al. (1980) reliant la permittivité et la porosité (dans un cas saturé). Par contre, on remarque que la corrélation entre la porosité et l'atténuation est plus faible. Ce résultat est attribuable au fait que la tomographie d'atténuation est généralement moins précise que la tomographie de vitesse. Cette perte de précision s'explique, entre autre, par la forte influence des contrastes de propriétés aux abords des forages sur le diagramme de rayonnement des antennes radar et par l'erreur qu'occasionne une mauvaise approximation du parcours des rays sur l'atténuation. Du plus, les forts contrastes de propriétés peuvent entraîner un problème d'ondes guidées, influençant aussi la précision de l'atténuation.

La figure 5.7 présente des coupes à travers la densité de probabilité conjointe. Les zones en rouge signifient une haute densité de probabilité.

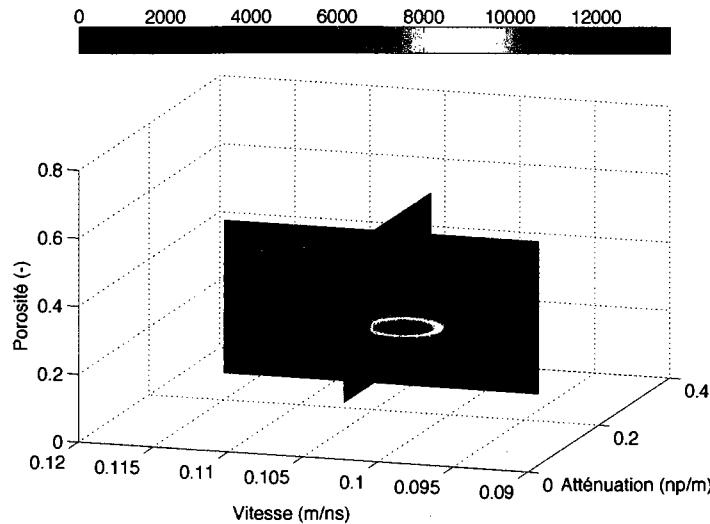


Fig. 5.7: Densité de probabilité conjointe porosité-vitesse-atténuation obtenue par la méthode des noyaux

Les largeurs de fenêtre adoptées sont de 0.0012 pour la vitesse, de 0.0085 pour l'atténuation et de 0.0115 pour la porosité. Elles ont été calculées à l'aide de la formule 3.50, développée par Silverman (1986).

5.1.3 Modèle de covariance de la porosité

Tel que mentionné précédemment, le krigeage simple effectué lors de la simulation bayésienne nécessite l'ajustement d'un modèle de covariance sur les porosités connues le long des forages. L'ajustement, présenté à la figure 5.8, a été réalisé à l'aide d'un modèle exponentiel ayant une portée en z de 0.65 m, un effet pépite de 0.0004 et un palier de 0.0015. Puisqu'il n'est pas possible de déterminer la portée dans la direction des x avec les seules données entre forages, celle-ci est jugée équivalente à la portée de 7 m modélisée lors de l'inversion tomographique. Cette pratique, couramment employée en géostatistique, est basée sur le l'hypothèse que

les deux variables présentent la même structure.

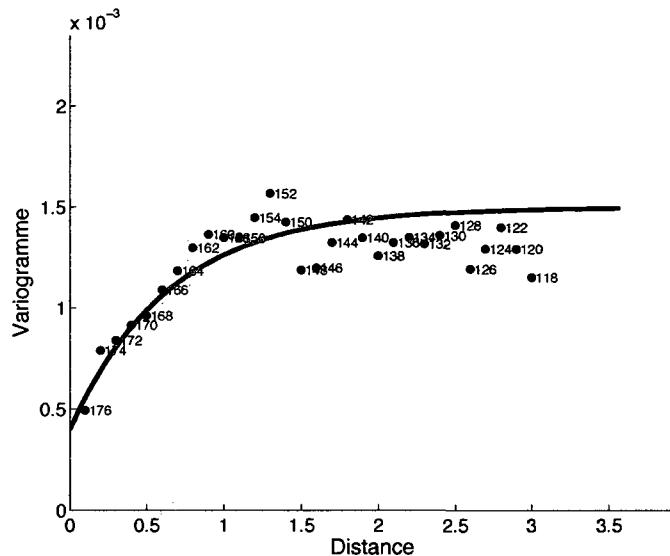


Fig. 5.8: Ajustement d'un modèle exponentiel sur les observations de porosités le long des forages

5.2 Simulations

Les figures 5.9 (b), (c) et (d) sont des exemples, choisis aléatoirement, de résultats obtenus par simulation séquentielle gaussienne par approche bayésienne parmi un ensemble de 100 réalisations. En comparant ceux-ci au modèle (fig. 5.9 (a)), on remarque que les principales structures sont toujours représentées et les contrastes de porosité, respectés. De plus, les simulations montrent une texture assez similaire à celle du modèle, excluant les valeurs extrêmes.

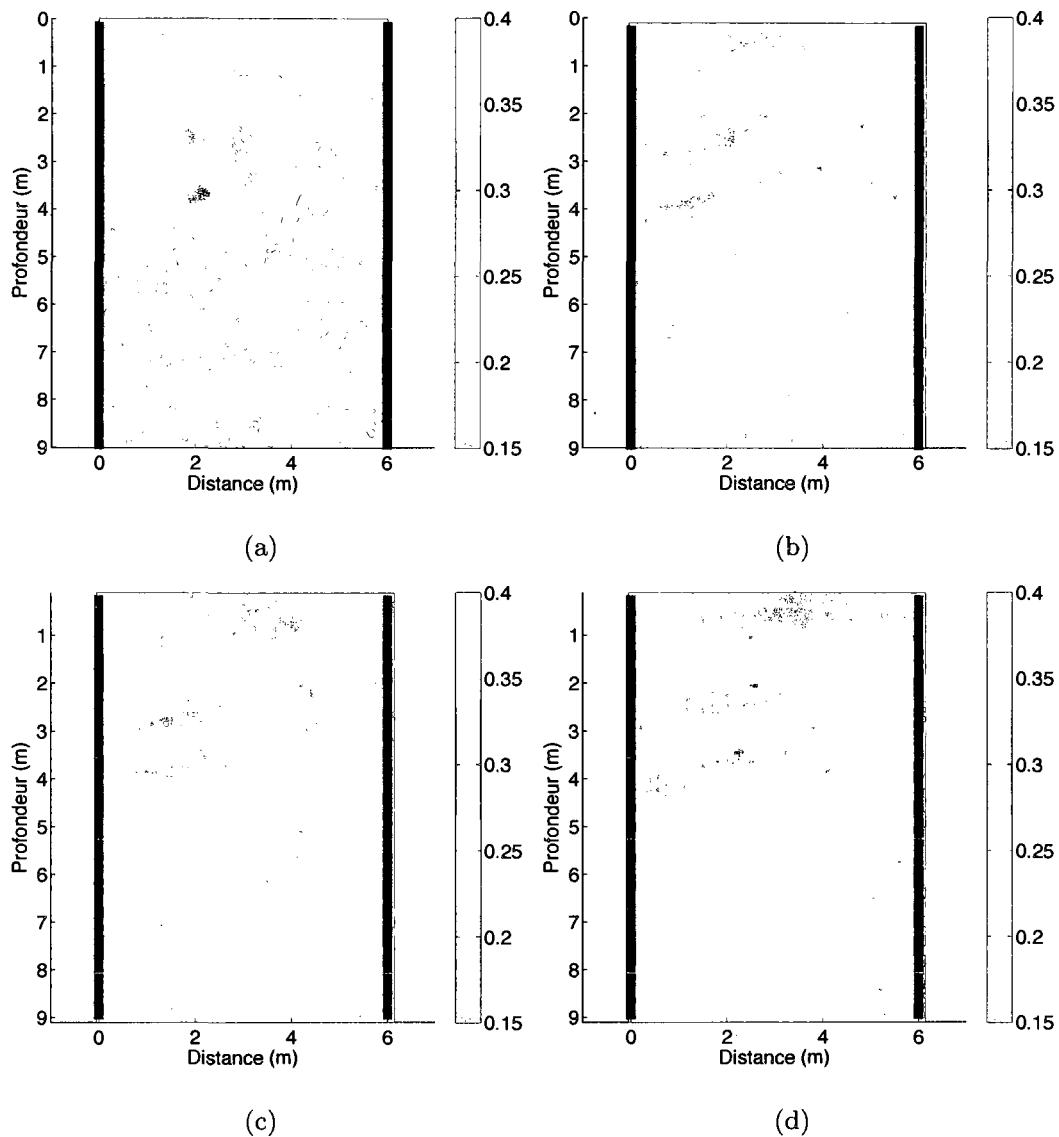


Fig. 5.9: Modèle de porosité (a), exemples de simulations séquentielles gaussiennes par approche bayésienne (b), (c) et (d)

La figure 5.10 présente les histogrammes du modèle et des simulations. Les variances de ces simulations sont autour de 0.0009, ce qui se compare à la variance de 0.0015 du variogramme. Ceci s'explique par l'information conditionnante supplémentaire fournie par les données géophysique. De plus, une seule tomogra-

phie géophysique a été utilisée, limitant ainsi la variabilité. Bien que la variance du modèle soit plus grande que celle des simulations, la forme gaussienne que présente les données du modèle est bien reproduite pour chacune des simulations, à quelques petites variations près.

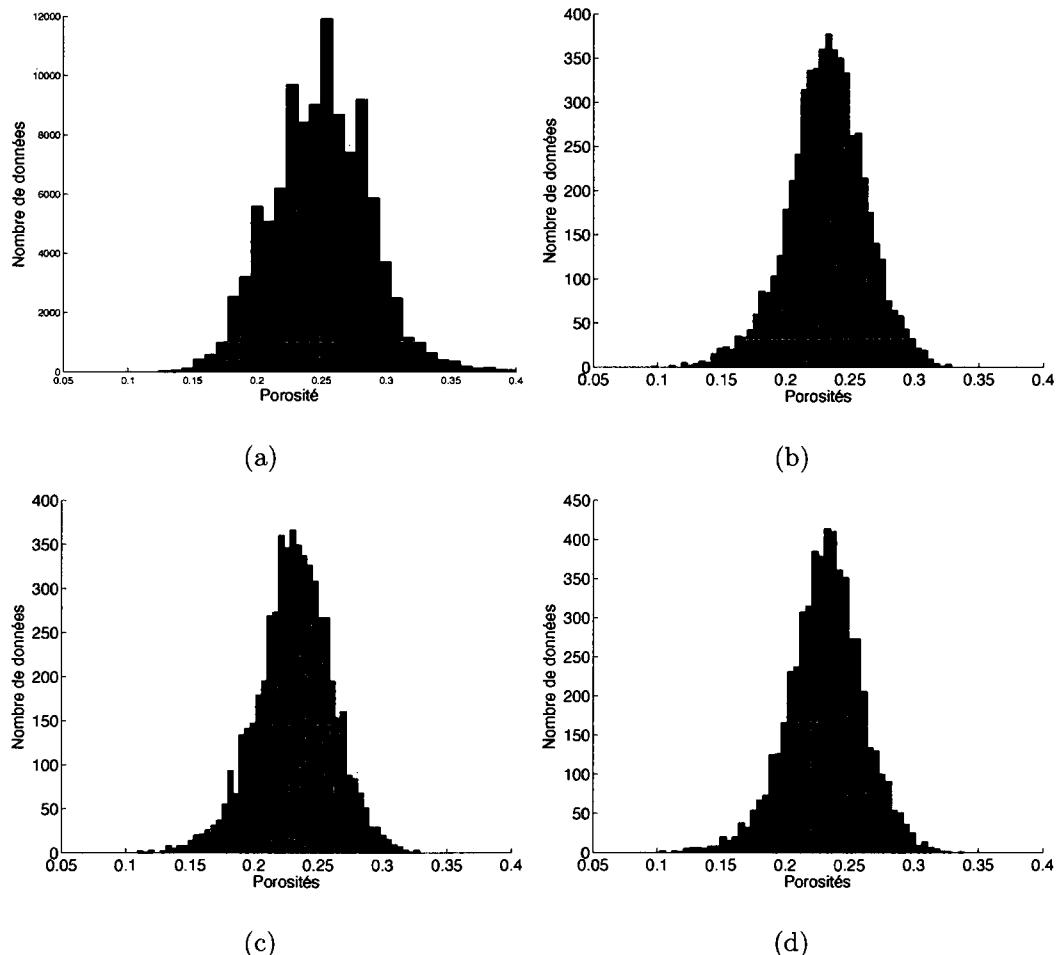


Fig. 5.10: Histogrammes du modèle en (a) et des simulations de la figure 5.9 dans le même ordre

Les coefficients de corrélation de ces trois simulations avec le modèle sont de 0.44, 0.43 et 0.40 respectivement.

5.3 Comparaison des simulations

La figure 5.11 montre les résultats de l'approche proposée comparés à ceux calculés à partir des méthodes directes (voir section 4.3). Le résultat obtenu à l'aide de la formule de Topp appliquée au tomogramme de permittivité, permet d'obtenir le meilleur coefficient de corrélation (0.57), comparativement au krigeage utilisant la porosité le long des forages (0.43) et à la corrélation moyenne de cent simulations bayésiennes (0.55). On remarque également que les principales structures lithologiques sont présentes et sont aux bons endroits. Toutefois, les méthodes directes produisent des résultats très lisses qui ne sont pas représentatifs de la texture réelle d'un sol hétérogène. De plus, ces méthodes ne fournissent qu'un estimé de la porosité ce qui ne permet pas de quantifier la variabilité sur l'estimation.

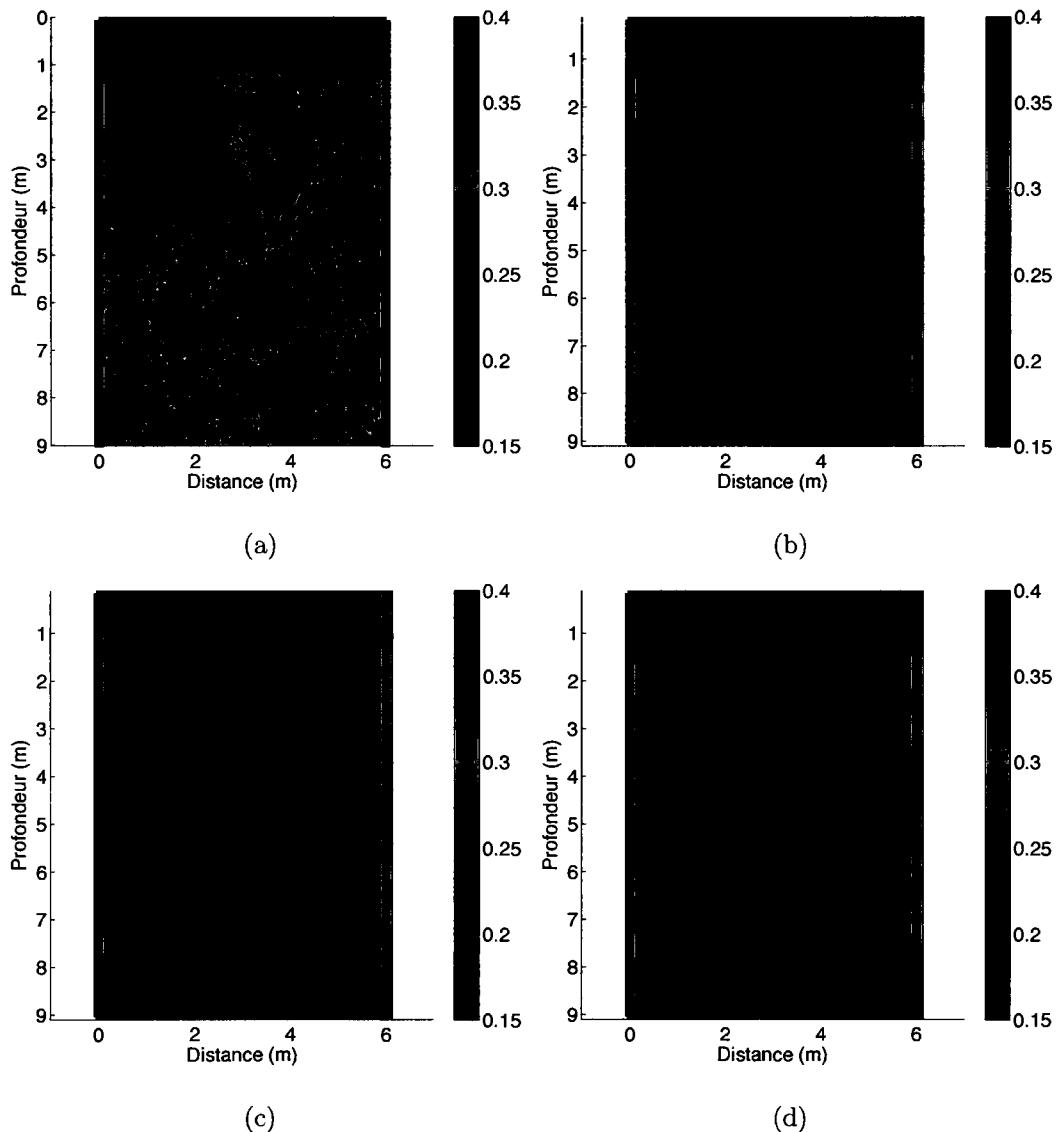


Fig. 5.11: Modèle (a) et porosité calculée par la méthode séquentielle bayésienne (moyenne de 100 simulations) (b), par la formule de Topp et al. (1980) (c) et par krigeage simple (d)

La figure 5.12 montre une réalisation (choisie aléatoirement) de la porosité calculée à partir de différentes méthodes de simulation. On note que les simulations reproduisent bien la texture du modèle. On remarque également que la simulation

bayésienne (a) et la cosimulation (d) sont nettement supérieures aux deux autres méthodes en ce qui a trait à la continuité et au positionnement des différentes lithologies. Par exemple, les deux couches inclinées, de porosité élevée et situées à 2 et 3.5 m. de profondeur, sont clairement représentées par ces deux méthodes.

Par ailleurs, en comparant les différentes simulations au modèle vrai, on s'aperçoit qu'aucune ne montre autant de variabilité que ce dernier. En effet, les valeurs extrêmes ne peuvent être générées par les simulations si cette variabilité n'est pas représentée par les données hydrogéologiques et géophysiques le long des forages.

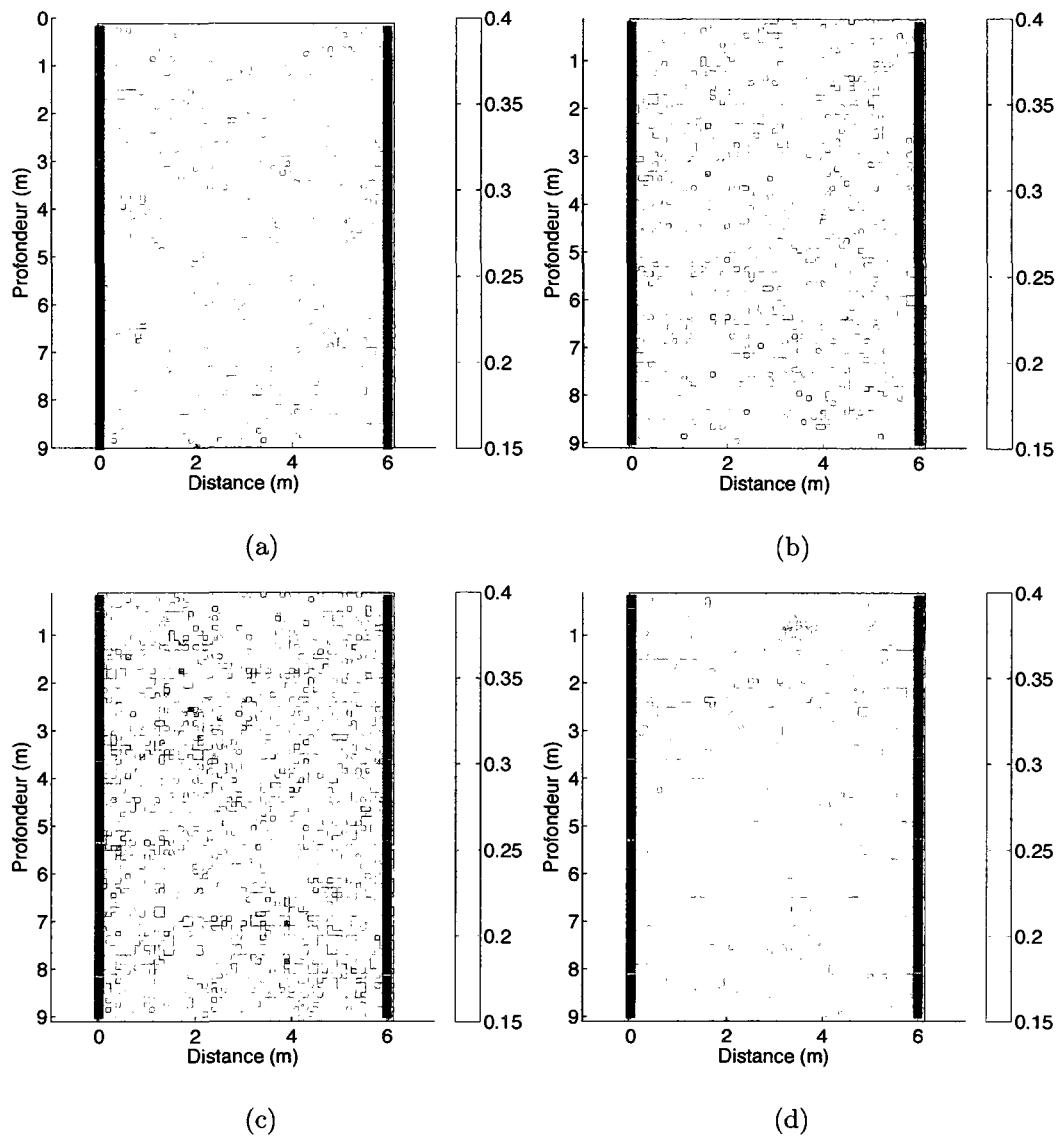


Fig. 5.12: Porosité calculée par la méthode séquentielle bayésienne (a), la méthode séquentielle gaussienne (classique) (b), la méthode bayésienne sans *a priori* (c) et la cosimulation (d)

Un modèle linéaire de corégionalisation à deux composantes (exponentiel + effet pépite) est adopté. Celui-ci a une portée de 1.5 m en z et 7 m en x . Les paliers de la composante exponentielle sont (ordre des variables : porosité (-), vitesse (m/ns) 2

et atténuation (np/m)² :

$$\begin{pmatrix} 1.2 & -0.12 & -0.25 \\ -0.12 & 0.017 & 0.12 \\ -0.25 & 0.12 & 2 \end{pmatrix} \times 10^{-3}$$

et l'effet pépite est :

$$\begin{pmatrix} 0.3 \times 10^{-3} & 0 & 0 \\ 0 & 0.1 \times 10^{-6} & 0 \\ 0 & 0 & 2 \times 10^{-6} \end{pmatrix}$$

On note que l'effet pépite sur la porosité présenté ici est légèrement différent de celui modélisé pour le krigeage simple. Ce petit changement a été réalisé afin que la matrice de l'exponentiel soit définie positive et donc, que le modèle soit valide.

5.4 Comparaison des variances

La carte des variances chacune des méthodes, calculées à partir des cent réalisations, sont présentées à la figure 5.13. Ceci permet de visualiser rapidement la répartition de l'incertitude pour chacun des pixels. On remarque aisément que la simulation séquentielle par approche bayésienne (a) ainsi que la cosimulation (d) ont des variances beaucoup plus faibles que la simulation séquentielle gaussienne classique (b) et que la simulation bayésienne sans estimation *a priori* (c). Ceci s'explique par le fait que, lors de la cosimulation et de la simulation bayésienne, on tient compte à la fois de l'information géophysique et de la corrélation spatiale ; avec la simulation classique, on n'inclut pas l'information géophysique et avec la simulation

bayésienne sans *a priori*, on n'inclut pas la structure spatiale de la porosité.

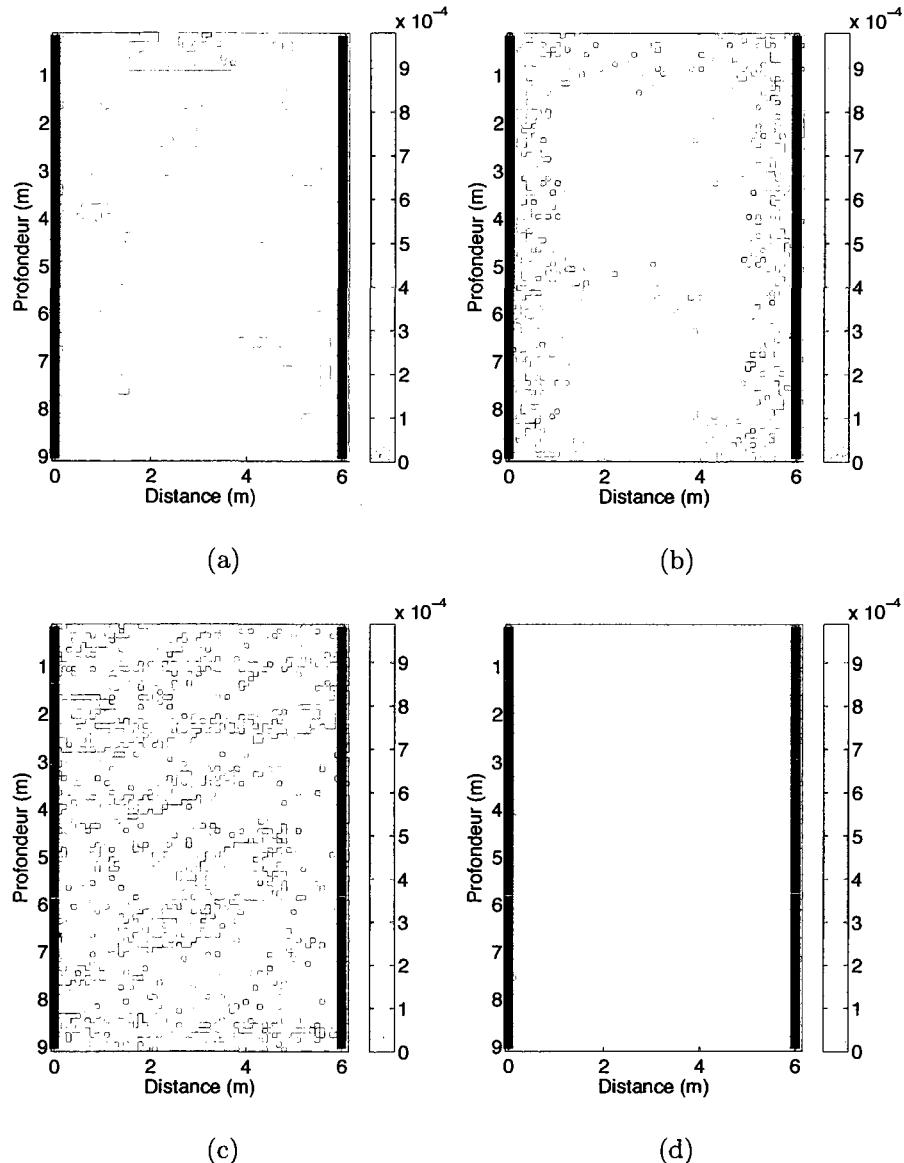


Fig. 5.13: Variance de la porosité calculée à partir de cent simulations pour la méthode bayésienne (a), la méthode séquentielle gaussienne (classique) (b), la méthode bayésienne sans *a priori* (c) et la cosimulation (d)

5.5 Comparaison des coefficients de corrélation

Le coefficient de corrélation a été calculé pour chacune des cent réalisations de chacune des méthodes. Les résultats obtenus sont présentés à la figure 5.14.

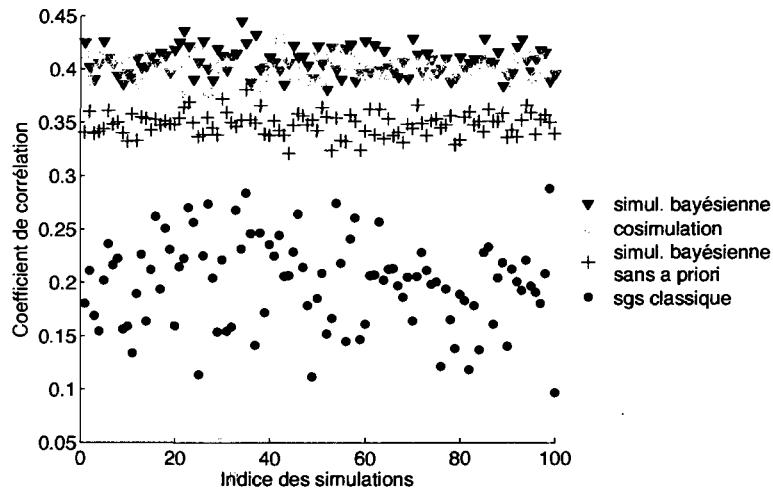


Fig. 5.14: Coefficients de corrélation des cent réalisations pour la méthode bayésienne en rouge, la cosimulation en jaune, la méthode séquentielle gaussienne classique en bleu et bayésienne sans *a priori* en vert

Ce graphique permet de faire une distinction marquée entre les méthodes combinant les données géophysiques et hydrogéologiques et celles qui n'utilisent qu'un des deux types de données. De ce fait, les coefficients de corrélation de la simulation par approche bayésienne (rouge) ainsi que la cosimulation (jaune) sont nettement supérieurs à ceux de la simulation séquentielle gaussienne classique et de la simulation bayésienne sans *a priori*.

Par contre, si l'on s'intéresse aux coefficients de corrélation calculés sur la moyenne des cent simulations, pour chacune des méthodes, on obtient les résultats suivants en ordre décroissant, incluant les méthodes directes :

Tab. 5.1: Classement des méthodes selon leur coefficient de corrélation.

1. Formule de Topp et al. (1980)	0.57
2. Simulation séquentielle bayésienne	0.55
3. Cosimulation	0.46
4. Krigage simple	0.43
5. Simulation séquentielle bayésienne sans <i>a priori</i>	0.4
6. Simulation séquentielle gaussienne classique	0.39

En excluant le résultat obtenu par la formule de Topp⁴, on peut conclure que la méthode proposée donne des résultats équivalents ou supérieurs aux autres méthodes de simulation. Examinons maintenant l'applicabilité de la méthode à un cas réel.

⁴cette formule permettant uniquement de calculer la porosité et que l'on s'intéresse davantage à la conductivité hydraulique qu'à la porosité

CHAPITRE 6

RÉSULTATS - SITE DE ST-LAMBERT-DE-LAUZON

6.1 Description du site à l'étude

Dans ce chapitre, la méthode proposée est appliquée à des données réelles. Le site à l'étude se situe à 35 km au sud-est de la ville de Québec, dans la municipalité de St-Lambert-de-Lauzon. Le site, d'une superficie de 12 km², est délimité à l'est par une ligne de partage des eaux de surface et souterraine et ailleurs par le réseau hydrique. En son centre se trouve un site d'enfouissement municipal, en opération depuis 1974. Jusqu'en 1997, les cellules d'enfouissement n'étaient pas recouvertes et ne contenaient pas de géo-membrane à leur base. Durant ces 24 années d'exploitation, 900 000 tonnes de déchets municipaux et industriels ont été entreposés directement sur un aquifère sableux non-confiné de 10 m d'épaisseur, reposant sur une unité d'argile et de silt (Tremblay et al., 2008). Depuis ce temps, l'eau de pluie et de fonte de neige traversant cette masse de déchets s'infiltra dans l'aquifère.

La zone d'étude est située directement sur le site d'enfouissement, en bordure d'une cellule. D'une superficie d'environ 400 m², ce site, hautement instrumenté, permet d'étudier chacun des aspects de la contamination. La figure 6.1 permet de visualiser une partie des installations, dans la zone où les mesures ont été acquises.

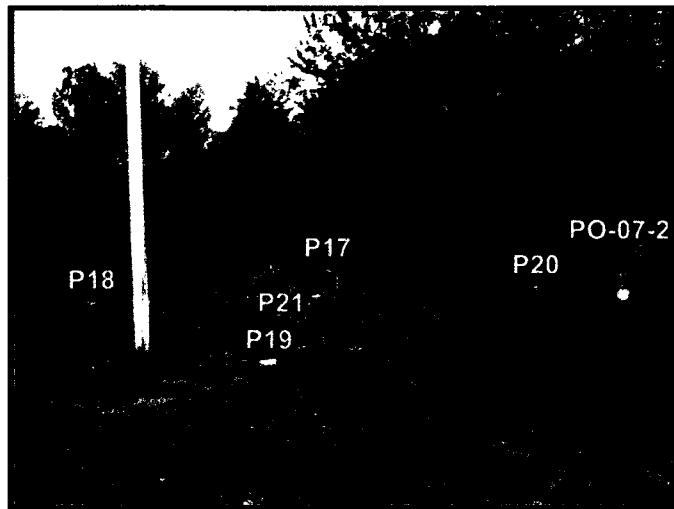


Fig. 6.1: Photographie de la position des puits sur le site à l'étude.

La stratification du site est connue grâce aux carottes de forage ainsi qu'aux différentes mesures recueillies jusqu'à présent (CPTu, essais *in situ*,...).

La figure 6.2 montre la disposition des puits. L'acquisition tomographique radar a été réalisée entre les puits P19 et P21, marqués en rouge, et le profil de conductivité hydraulique, le long du puits P17, marqué en jaune. Notons que l'on ne dispose pas de données hydrogéologiques adjacentes aux tomogrammes, ce qui augmente le degré d'incertitude sur l'estimation. Au moment de prendre les mesures, la nappe était à moins d'un mètre de la surface du sol. On considère donc que le milieu est saturé.

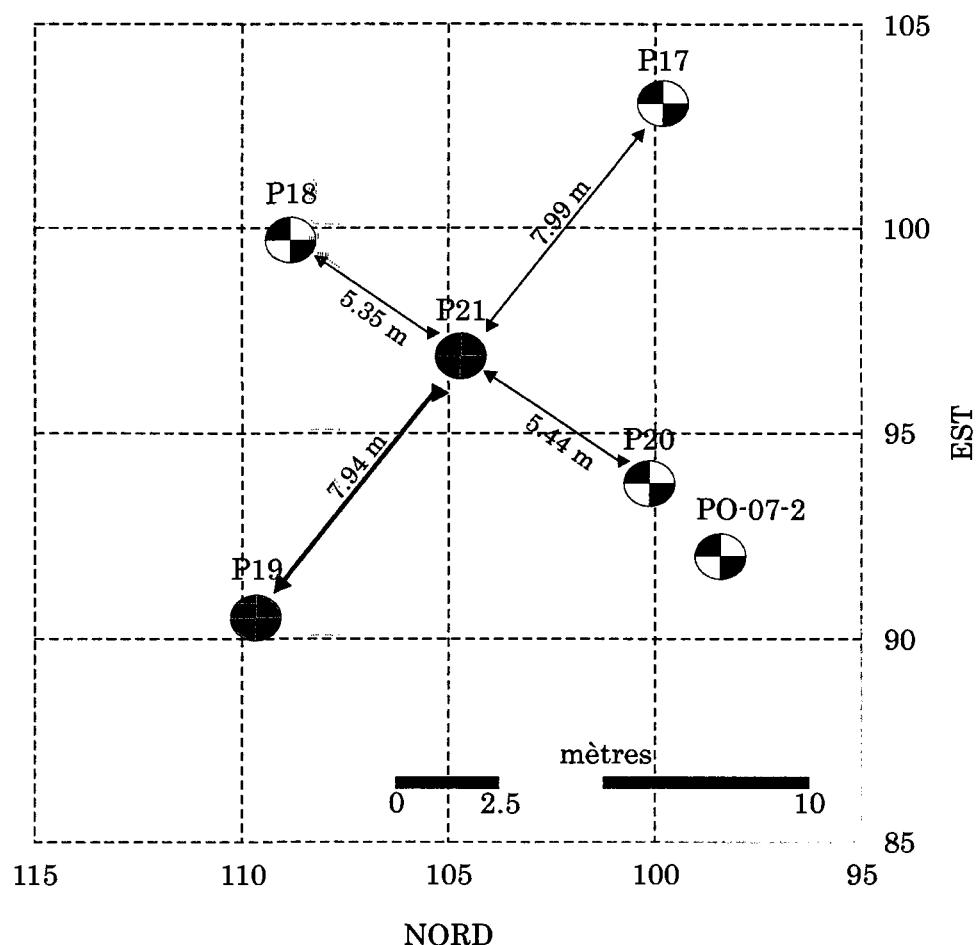


Fig. 6.2: Position des puits sur le site à l'étude.

6.2 Données et paramètres de pré-simulation

Cette section présente les différentes données dont on dispose ainsi que les étapes préalables à la réalisation des simulations.

6.2.1 Données géophysiques

Une acquisition de tomographie radar a été réalisée entre les forages P19 et P21 à l'aide d'un système *ProEx* et des antennes de 100 MHz, fabriqué par *Malå Geoscience*. Les temps de parcours ainsi que les amplitudes ont ensuite été inversés grâce à l'algorithme *bh_tomo* (Gloaguen et al., 2007b; Giroux et al., 2007). Afin de pouvoir réaliser le cokrigeage du champ de lenteur à partir des temps mesurés, un modèle exponentiel ayant une portée de 0.25 m en z , de 7 m en x , une variance de 0.9 et un effet pépite de 2 a été ajusté sur ces temps de parcours. Aucune contrainte le long des forage n'a été imposée lors de l'inversion.

Ces tomogrammes permettent d'identifier principalement trois couches sub-horizontales aux propriétés contrastées : une couche à faible vitesse et à atténuation élevée située entre 4 et 6 m et comprise entre deux couches de haute vitesse et de plus faible atténuation. La petite zone de très fortes vitesses située au sommet de l'image correspond à la zone non-saturée.

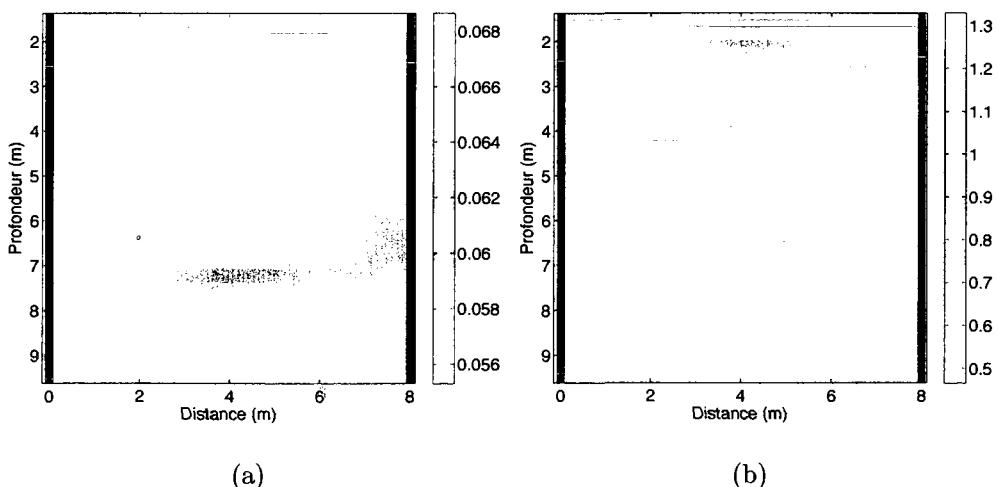


Fig. 6.3: (a) Vitesses (m/ns) et (b) atténuations (np/m) radar obtenues entre les puits P19 et P21

Les vitesses et les atténuations sont ensuite converties en permittivité relative et en conductivité électrique grâce aux équations 3.25 et 3.26, présentées à la section 3.3.1.

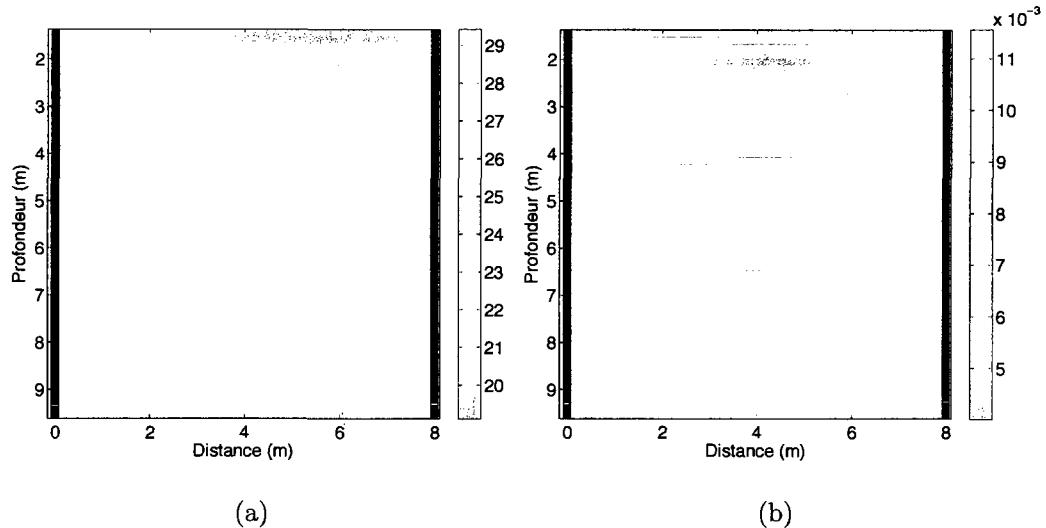


Fig. 6.4: (a) Constante diélectrique et (b) conductivité électrique (S/m) calculées à partir des tomogrammes de vitesse et d'atténuation

On remarque, en toute logique, que les principales structures identifiées précédemment sont reproduites par les tomogrammes de permittivité relative et de conductivité électrique. On note que les deux premiers mètres des tomogrammes présentent des valeurs extrêmes qui correspondent à la zone non-saturée. Cette zone n'est donc jamais prise en compte dans l'analyse des résultats.

6.2.2 Données hydrogéologiques

Le profil du logarithme (base 10) de la conductivité hydraulique mesuré au puits P17 est présenté à la figure 6.5. Ces mesures ont été acquises lors d'un essai de perméabilité *in situ* à niveaux multiples. Pour chacun des niveaux, un minimum

de trois essais à niveau variable ont été réalisés afin de s'assurer de la qualité du résultat. De plus, le niveau piézométrique local a été systématiquement mesuré afin de respecter la norme canadienne (Chapuis, 1995) et de récupérer le maximum de données possibles.

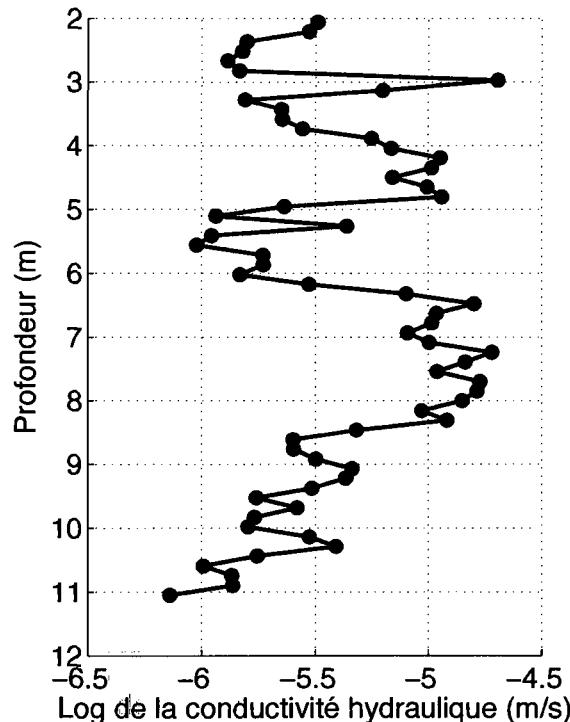


Fig. 6.5: Profil du logarithme (base 10) des conductivités hydrauliques, obtenus par essai de perméabilité *in situ* à niveaux multiples, pour le puits P17

Sur ce profil, on peut identifier trois zones distinctes (on ne s'intéresse ici qu'aux 9 premiers mètres puisque cela correspond à la profondeur des tomogrammes radar) :

1. entre 2 et 5 m, une conductivité hydraulique d'environ $10^{-4.8}$ m/s s'apparentant à un sable ;
2. entre 5 et 6.2 m, une conductivité hydraulique d'environ 10^{-6} m/s s'apparentant à un sable fin contenant une fraction de silt ;

3. entre 6.2 et 8.5 m, une conductivité hydraulique d'environ $10^{-4.6}$ m/s s'apparentant à un sable.

Ces zones concordent bien avec le modèle géologique élaboré jusqu'à ce jour sur ce site (Gloaguen, 2009, communication personnelle) ainsi qu'avec les contrastes observés sur les tomogrammes.

6.2.3 Modèle de covariance des \log_{10} de la conductivité hydraulique

Les mesures de conductivité hydraulique le long du puits P17 permettent ensuite de calculer un variogramme sur lequel un modèle est ajusté. Cette étape est nécessaire à la réalisation du krigage simple, réalisé pour chacun des pixels, lors de la simulation bayésienne (voir section 4).

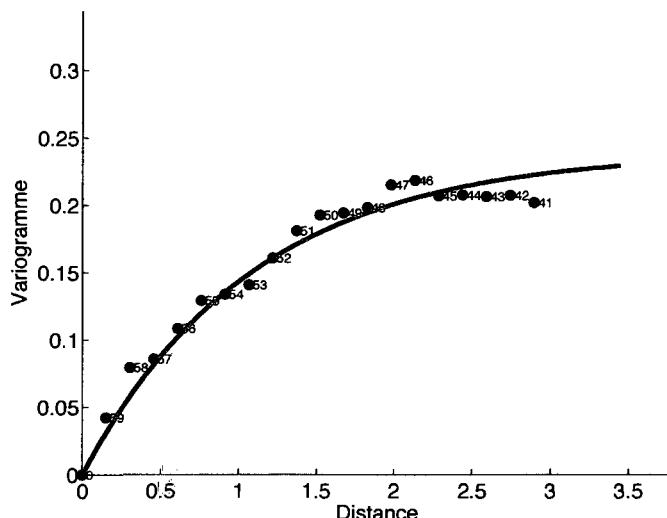


Fig. 6.6: Ajustement d'un modèle exponentiel sur le \log_{10} de la conductivité hydraulique le long de P17

Comme pour l'inversion tomographique, un modèle exponentiel a été choisi. Les paramètres d'ajustement sont : une variance de 0.24 (m/s)^2 , aucun effet pépite et

une portée verticale de 1.1m. La moyenne arithmétique du \log_{10} de la conductivités hydrauliques de ce puits est $1 \times 10^{-5.4}$ m/s.

6.2.4 Densité de probabilité conjointe

La densité de probabilité conjointe a pour but d'établir la relation *in situ* entre la constante diélectrique, la conductivité électrique et la conductivité hydraulique. Celle-ci a été obtenue grâce aux mesures de ces trois paramètres le long de forages (P1 à P17) situés dans les environs du site. La figure suivante présente les relations entre le logarithme de la conductivité hydraulique et les paramètres physiques.

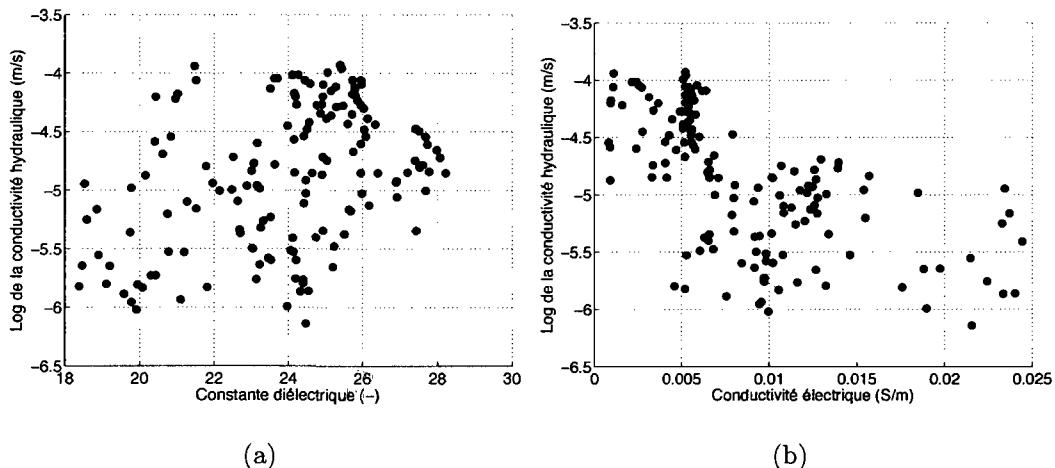


Fig. 6.7: (a) Graphique de la constante diélectrique et (b) de la conductivité électrique le long de plusieurs forages selon la conductivité hydraulique

La corrélation de la conductivité hydraulique avec la constante diélectrique est de 0.4 alors qu'avec la conductivité électrique, elle est de -0.62. Dans le cas de la conductivité électrique (fig. 6.7 (b)), cette corrélation serait sans doute plus élevée avec un modèle non-linéaire, ce qui justifie l'approche proposée qui ne requiert pas la détermination d'une relation numérique.

Conformément au protocole établi, la méthode des noyaux a été utilisée afin de calculer la densité de probabilité conjointe pour ces trois paramètres, comme le montre la figure 6.8.

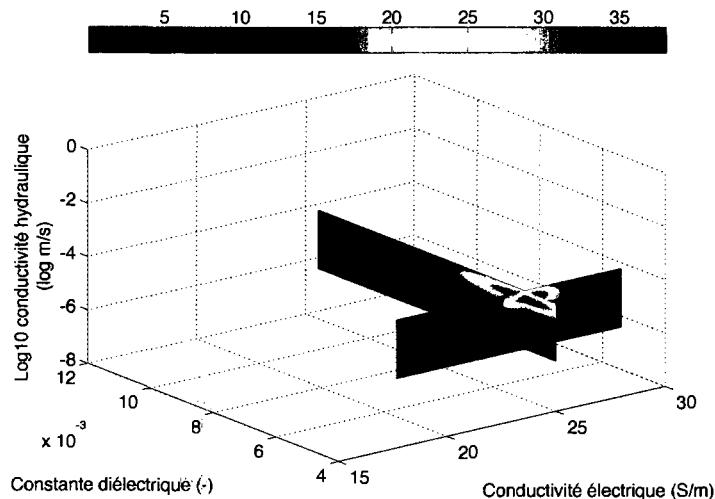


Fig. 6.8: Densité de probabilité conjointe

La largeur des fenêtres est de 0.0015 pour la conductivité électrique, de 0.77 pour la constante diélectrique et de 0.2 pour la conductivité hydraulique. À l'instar du cas synthétique, ces largeurs de fenêtre ont été calculées à l'aide la formule empirique de Silverman (1986).

6.3 Simulations

La figure 6.9 présente quatre exemples de simulations de la conductivité hydraulique obtenues grâce à l'approche bayésienne. Les valeurs sont déterminées sur un carré de maille 15 cm × 15 cm. On retrouve, sur chacune des simulations, la couche de silt de très faible conductivité hydraulique comprise entre deux couches perméables de conductivité hydraulique plus grande que 2×10^{-5} m/s.

Ainsi, ces résultats concordent bien avec la coupe stratigraphique connue à ce jour sur ce site.

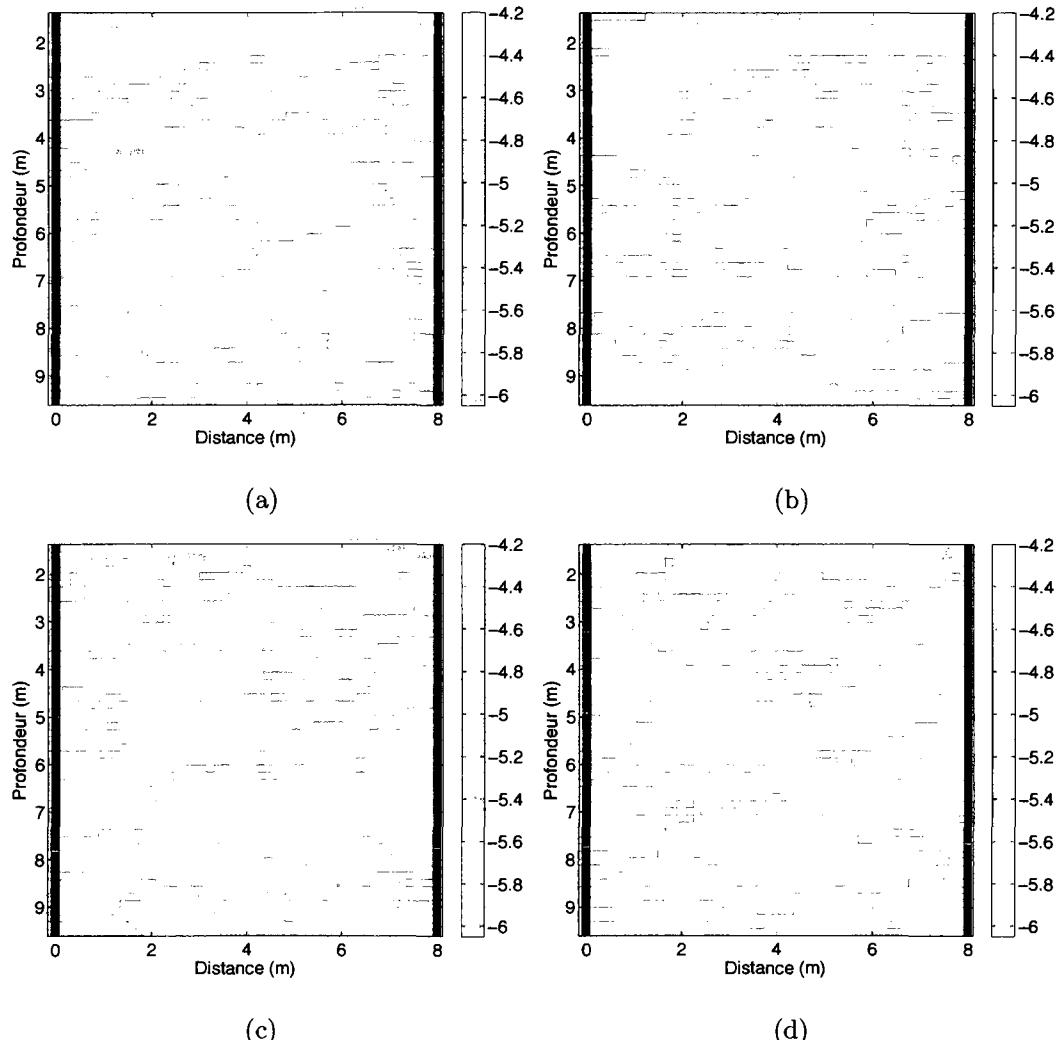


Fig. 6.9: Réalisations de la conductivité hydraulique (\log_{10}), entre les puits P19 et P21, obtenue par simulation séquentielle gaussienne par approche bayésienne

Si l'on compare les résultats précédents avec d'autres méthodes de simulation (voir figure 6.10), on remarque que les trois couches mentionnées précédemment sont à peine visibles sur ces trois résultats. La piétre qualité du krigeage et de la simulation

séquentielle gaussienne classique est due au manque de données de conductivité hydraulique à proximité du panneau étudié. Quant à la simulation bayésienne sans *a priori*, on explique sa très grande hétérogénéité par le fait que l'algorithme ne permet pas de forcer une certaine continuité spatiale puisque chaque pixel est tiré de la densité de probabilité conjointe indépendamment de ses voisins.

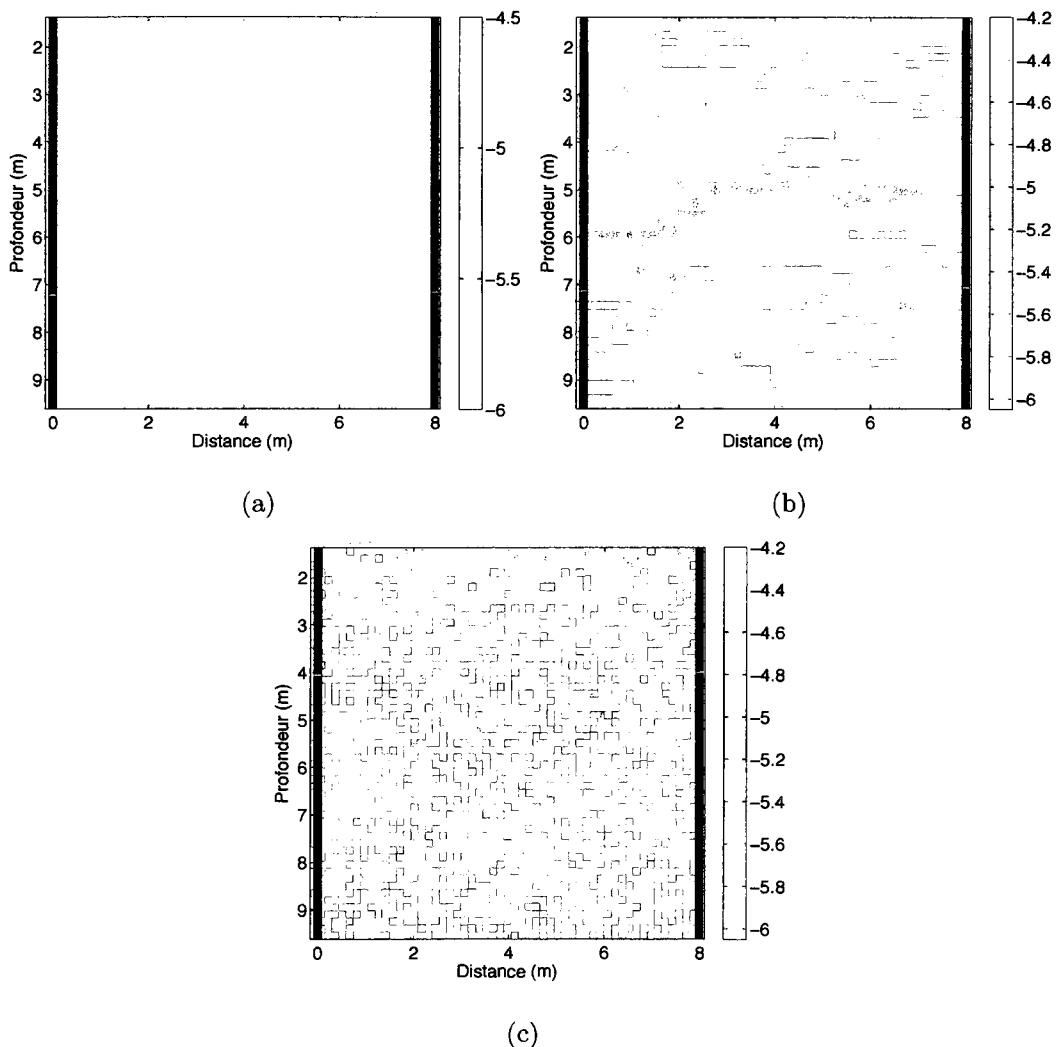


Fig. 6.10: Conductivité hydraulique (\log_{10} m/s) calculée par la méthode (a) de krigage simple, (b) séquentielle gaussienne et (c) bayésienne sans *a priori*

Par ailleurs, la variance de chacun des pixels, pour cent simulations, est représentée à la figure 6.11. L'examen de ce résultat permet d'identifier les zones plus incertaines pour chacune des méthodes. On remarque qu'une d'elles se situe au sommet de la coupe et à quelques endroits dans la couche la plus perméable à 8 m de profondeur. La variance très élevée située directement sous une zone à variance quasiment nulle, au sommet des deux simulations bayésiennes, est occasionnée par le fort contraste de vitesses des ondes près de la surface, en zone non-saturée, et par l'ordre dans lequel les pixels sont visités. En effet, si par hasard plusieurs pixels situés dans la zone non-saturée sont d'abord visités, les fortes valeurs à cet endroit se répercuteront sur la valeur simulée à un pixel en bas de 2 m. Ainsi, puisque la situation inverse peut aussi se produire (plusieurs pixels situés sous la zone non-saturée sont d'abord visités), cela entraîne une zone de grande variance à l'interface.

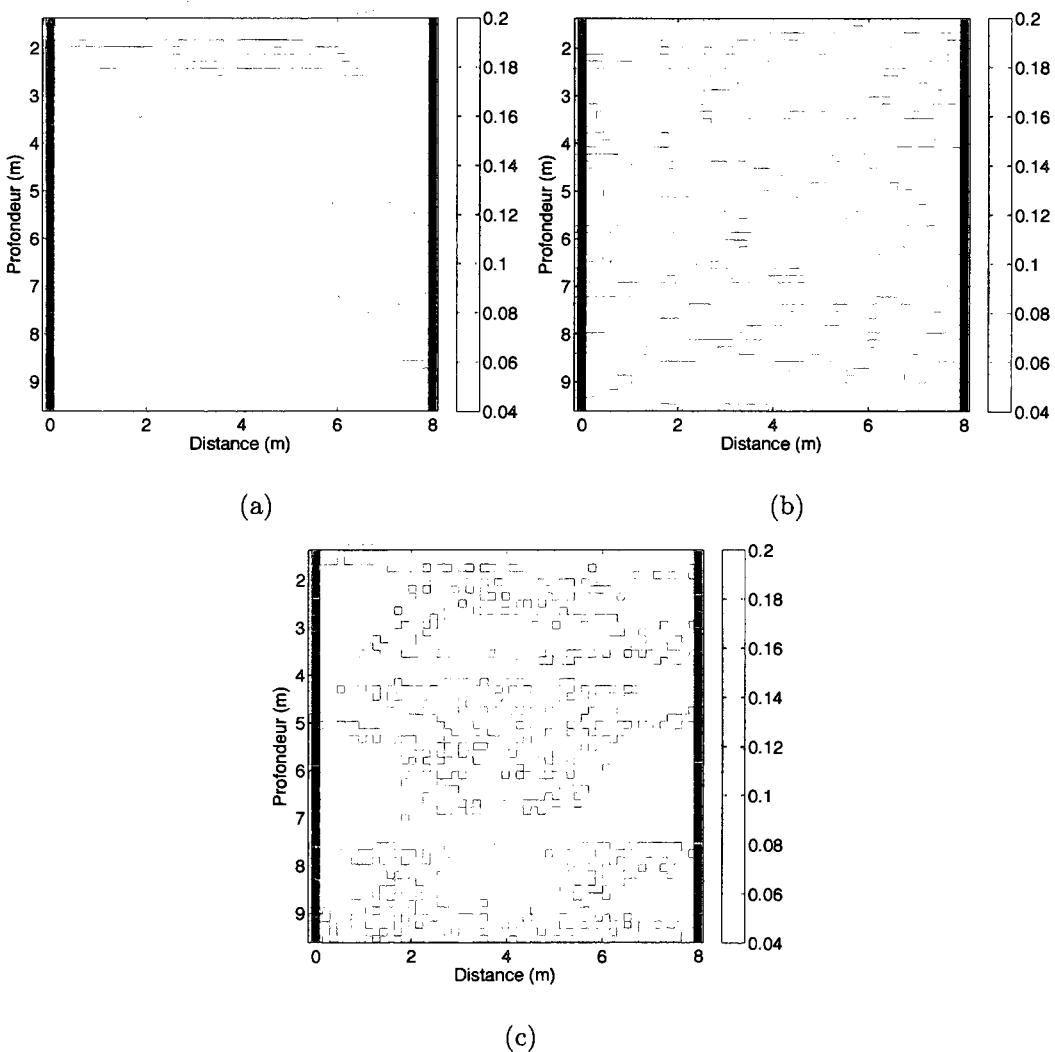


Fig. 6.11: Variance ($\log_{10} (\text{m/s})^2$) sur l'estimation de la conductivité hydraulique calculé avec cent simulations pour la méthode (a) bayésienne avec *a priori* (b), séquentielle gaussienne et (c) bayésienne sans *a priori*

CHAPITRE 7

DISCUSSION

À la lumière des résultats obtenus, l'algorithme de simulation séquentielle gaussienne par approche bayésienne permet de générer des réalisations de la conductivité hydraulique sur une grille 2D, à une échelle décimétrique. Cet algorithme est validé à l'aide d'un cas synthétique, puis appliqué à un cas réel. Ces deux cas ont permis de confirmer l'hypothèse que l'ajout de données géophysiques aux données hydrogéologiques précise substantiellement l'estimation des paramètres hydrauliques. À partir de ces résultats, les points forts et les points faibles de l'approche bayésienne ont pu être identifiés. Dans cette section, il est donc question d'analyser les forces et les limites de chacun des aspects de la méthode proposée, de la prise des mesures à l'obtention de réalisations.

Avantages

Un premier avantage de l'approche bayésienne est de permettre l'inclusion, dans la simulation, des données de conductivité hydraulique préalablement connues mais souvent imprécises, peu importe leur localisation dans le site, sous forme d'une première distribution appelée *a priori*. Celle-ci est calculée à l'aide d'un krigeage simple de la variable principale. Un second avantage de la méthode est l'intégration de cette approche paramétrique (ajustement d'un variogramme) à la flexibilité de l'approche non paramétrique de la méthode du noyau, qui permet de tenir compte de relations non linéaires et à dispersions variables.

La distribution *a priori* est « mise à jour » à l'aide de l'information apportée par les données géophysiques. Le lien entre les paramètres physiques et hydrauliques

est établi *in situ* et la méthode des noyaux permet de calculer la densité de probabilité conjointe de ceux-ci. Cette méthode a l'avantage de permettre l'inclusion de plusieurs variables, tant que celles-ci sont connues aux mêmes points, un des atouts par rapport à la simulation séquentielle gaussienne standard qui n'utilise que la variable principale.

Tel que mentionné précédemment, une des forces de la méthode des noyaux est de pouvoir supporter des relations non-linéaires entre les différentes variables. Ceci est pleinement justifié avec le cas réel où les relations établies entre la conductivité hydraulique et les paramètres physiques sont supposées non linéaires. En conséquence, l'approche proposée est avantageuse par rapport au cokrigeage¹, qui ne tient compte que du lien linéaire entre les variables. Par ailleurs, le cokrigage requiert l'ajustement de tous les variogrammes et variogrammes croisés, ce qui s'avère être une lourde tâche.

Par ailleurs, le fait d'établir une relation *in situ* entre les paramètres physiques et hydrauliques est probablement mieux adapté pour des sols hétérogènes que le recours à des formules pétrophysiques générales, restreintes à des limites de validité. De plus, contrairement aux méthodes de simulation, celles-ci ne permettent pas de mesurer l'incertitude sur les estimations.

Désavantages

Le krigeage simple requiert une estimation de la moyenne du paramètre recherché, ce qui peut être imprécis si l'on ne possède pas beaucoup de données. Une alternative serait de remplacer le krigeage simple par un krigeage ordinaire (Emery, 2008). De plus, la moyenne et la variance de krigeage, calculées au pixel à simuler, servent à générer la distribution *a priori*, de forme gaussienne. Cette étape suppose

¹avec la conductivité hydraulique comme variable principale et la constante diélectrique et la conductivité électrique comme variables secondaires

donc que le paramètre recherché présente un comportement gaussien, ce qui n'est pas toujours le cas. Ici, nous ferons l'hypothèse que la distribution de ce dernier est bien gaussienne. Il est possible de valider cette hypothèse en comparant les variogrammes des indicatrices associées à des quartiles complémentaires (e.g. 0.2 - 0.8, 0.1 - 0.9). Dans le cas gaussien, ces variogrammes devraient être identiques (Chilès et Delfiner, 1999; Marcotte et Gloaguen, 2008).

Le calcul de la densité de probabilité conjointe nécessite la détermination des largeurs de fenêtre. Lorsque l'on a plus de trois variables et qu'un examen visuel est impossible, cet exercice peut être délicat. De plus, la fonction d'interpolation servant à générer la probabilité de densité conjointe continue peut induire un biais se reflétant sur la vraisemblance (Rosenblatt, 1956). Ceci est d'autant plus vrai lorsque l'on dispose de peu de points de mesures.

Outre la quantité, la qualité des points de mesure du noyau est importante puisqu'elle peut influencer grandement la justesse des estimations. Lors de ce projet, la justesse de l'estimateur aurait pu être vérifiée en appliquant la méthode à plusieurs modèles synthétiques statistiquement équivalents. Malheureusement, cet exercice n'a pas été réalisé, faute de temps.

D'un côté pratique, les méthodes géophysiques et hydrogéologiques retenues comportent plus d'avantages que d'inconvénients :

- le piézocône multi-capteurs est un instrument de pointe, qui donne des mesures relativement précises et en continu, de plusieurs paramètres. Cet appareil est avantageux pour la vitesse à laquelle il opère et pour la quantité et la qualité des données qu'il mesure (Butler, 2005).
- l'acquisition tomographique radar présente l'avantage de mesurer simultanément

deux paramètres physiques en lien avec les propriétés hydrauliques du milieu. L'inversion de ces deux paramètres permet de déterminer la vitesse et l'atténuation des ondes radar sur une grille décimétrique entre les forages. Notons toutefois que les tomogrammes sont moins bien définis sur les bordures et au sommet, qu'au centre (Linde et al., 2006a). De plus, dans un sol composé d'une grande fraction argileuse, les ondes radar risquent d'être très atténuées en raison de la forte conductivité électrique de l'argile, ce qui constitue une limite à l'utilisation de cet appareil.

Une des caractéristiques de l'acquisition tomographique est de mieux reproduire les grandes structures du milieu que les variations à très petite échelle. Considérant que l'effet de pépite des tomographies radar est principalement occasionné par le processus d'inversion, on peut supposer que celui-ci n'est probablement pas corrélé avec les variations à petite échelle de la porosité. Dans cette étude, les tomogrammes résultant du cokrigeage ont été préférés aux simulations. Il aurait toutefois été intéressant de visualiser l'impact d'inclure les variations des tomographies simulées par la méthode FFT-MA (réalisées à l'aide du logiciel *bh-tomo*) sur les réalisations de porosité ou de conductivité hydraulique obtenues par l'approche proposée. Pour être valable, cette vérification aurait nécessité, par exemple, le calcul d'une réalisation bayésienne pour chacune des réalisations tomographiques.

- l'essai de perméabilité *in situ* à niveaux multiples est relativement rapide à mettre en oeuvre et assez précis pour repérer des lithologies d'une dizaine de centimètres, comportant des propriétés hydrauliques différentes. Toutefois, cet essai, comme les autres méthodes hydrogéologiques *in situ*, ne mesure réellement que la composante horizontale de ce paramètre. De plus, les mesures peuvent être biaisées par l'accumulation ou le drainage des particules fines à l'interface

entre le forage et le sol adjacent.

Il est important de mentionner que les mesures issues de ces différents appareils ne sont pas toutes à la même échelle lors de leur acquisition. Ainsi, une intégration des mesures les plus fines (prises ici par le piézocône et l'essai *in situ* à niveaux multiples), à une échelle plus grossière (tomographie radar) a été préalablement réalisée (Paradis et al., 2008) afin de tenir compte des différentes échelles.

Finalement, une faiblesse de la méthode est d'être tributaire d'une acquisition radar tomographique généralement réalisée en 2D alors qu'en réalité, la conductivité hydraulique est un phénomène tridimensionnel. La tomographie radar limite aussi la largeur de la zone étudiée à une dizaine de mètres, ce qui peut sembler faible par rapport à la taille d'un panache de contamination, par exemple. La combinaison 2D ou 3D de plusieurs panneaux est donc une avenue intéressante pour cette méthode. Une autre moyen pour générer un modèle 3D serait de recourir à la tomographie radar en 3D directement (Eppstein et Dougherty, 1998b,a). La simulation bayésienne pourrait donc être appliquée directement sur l'ensemble du bloc à condition d'avoir une valeur des paramètres géophysiques à tous les points où l'on souhaite simuler une valeur de conductivité hydraulique.

CHAPITRE 8

CONCLUSION

En somme, l'algorithme de simulation séquentielle gaussienne par approche bayésienne de données de tomographie radar et des paramètres hydrauliques, est une approche simple et applicable « *in situ* » permettant d'estimer la conductivité hydraulique d'un aquifère à une échelle fine (~ 15 cm), sur un plan compris entre deux forages. La robustesse de la méthode a été validée à l'aide d'un cas synthétique et d'un cas réel. Les résultats obtenus par cette approche sont concluants. Toutefois, divers travaux futurs sont envisageables.

D'abord, d'autres méthodes géophysiques et hydrauliques pourraient être intégrées aux simulations par le biais de la densité de probabilité conjointe. Des exemples de méthodes pertinentes seraient la tomographie électrique pour pallier aux effets des ondes guidées, la tomographie sismique pour mieux contraindre la tomographie radar ou la tomographie hydraulique pour une meilleure estimation *a priori* de la conductivité hydraulique. Il faut toutefois considérer que l'ajout d'autres méthodes exige une augmentation considérable du nombre de données afin d'avoir une estimation précise de la densité de probabilité conjointe.

Une avenue intéressante pour augmenter la quantité de données disponibles est le concept du modèle analogue. À l'instar de la méthode présentée par Gastaldi et al. (1998) et par Gloaguen et al. (2009), un modèle analogue de la coupe stratigraphique du site investigué peut permettre de générer des données géophysiques et hydrauliques synthétiques. Ces dernières peuvent ensuite être intégrées au calcul de la densité de probabilité conjointe afin de combler les zones non échantillonnées. Cette approche novatrice a comme avantage de préciser la relation entre les pa-

ramètres physiques et hydrauliques et donc, de fournir potentiellement un meilleur estimé de la conductivité hydraulique. Toutefois, une telle approche n'est possible que si l'on dispose de relations analytiques théoriques liant les variables géophysiques mesurées à la conductivité hydraulique. Ceci demeure un domaine actif de recherche.

De plus, il serait particulièrement intéressant d'extrapoler plusieurs panneaux 2D, compris entre deux forages, pour créer un modèle 3D de conductivité hydraulique connue à une échelle décimétrique. Un essai de traceur sur un site réel et modélisé à l'aide d'un simulateur d'écoulement pourrait ensuite permettre de valider l'interpolation. Un algorithme permettant de générer un tel modèle à l'aide de données géophysiques et hydrauliques améliorerait substantiellement les prédictions actuelles de transport de contaminants et permettrait de réaliser plus précisément des plans de réhabilitation de zones contaminées.

Ultimement, un modèle de conductivité hydraulique 3D, à une échelle décimétrique, pourrait aussi être généré par simulation bayésienne à partir d'une tomographie radar 3D.

RÉFÉRENCES

- Alumbaugh, D., Chang, P., Paprocki, L., Brainard, J., Glass, R. J., et Rautman, C. A. (2002). Estimating moisture contents in the vadose zone using cross-borehole ground penetrating radar : A study of accuracy and repeatability. *Water Resources Research*, **38**(12).
- Annan, A. (2005). *Ground-Penetrating Radar*, chapitre 11. Dans : Butler, K. Darwin, Near-surface geophysics, Society of Exploration Geophysicists.
- Archie, G. E. (1942). The electrical resistivity log as an aid in determining some reservoir characteristics. *Trans. AIME*, **146**, 54–62.
- Bear, J. (1979). *Hydraulics of groundwater*. McGraw-Hill Editions, London.
- Binley, A. (2007). *Hydrological - geophysical relationships*. Lancaster University.
- Binley, A., Winship, P., Middleton, R., Pokar, M., et West, L. J. (2001). High-resolution characterization of vadose zone dynamics using cross-borehole radar. *Water Resources Research*, **37**(11), 2639–2652.
- Butler, J. J. (2005). *Hydrogeological methods for estimation of hydraulic conductivity*, chapitre 2, pages 23–58. Dans : Rubin, Y., Hubbard, S., Hydrogeophysics, Springer.
- Carman, P. C. (1939). Permeability of saturated sands, soils and clays. *Journal of Agricultural Sciences*, **29**, 263–273.
- Carrier, D. W. (2003). Goodbye, Hazen ; Hello, Kozeny-Carman. *Journal of Geotechnical and Geoenvironmental Engineering*, **129**(11), 1054–1056.
- Chapuis, R. (1995). Controlling the quality of ground water parameters : some examples. *Canadian Geotechnical Journal*, **32**(1), 172–177.

- Chapuis, R. (1999). *Guide des essais de pompage et leurs interprétations*. Les publications du Québec, Ste-Foy.
- Chapuis, R. et Aubertin, M. (2003). On the use of Kozeny-Carman equation to predict the hydraulic conductivity of soils. *Canadian Geotechnical Journal*, **40**, 616–628.
- Chen, J., Hubbard, S., et Rubin, Y. (2001). Estimating the hydraulic conductivity at the South Oyster site from geophysical tomographic data using bayesian techniques based on the normal linear regression model. *Water Resources Research*, **37**(6), 1603–1613.
- Chilès, J.-P. et Delfiner, P. (1999). *Geostatistics*. Wiley Inter-science.
- Dallaire, V. (2004). Étude des effets d'échelle dans un aquifère alluvionnaire. Mémoire de maîtrise, École Polytechnique de Montréal.
- Day-Lewis, F., Singha, K., et Binley, A. (2005). Applying petrophysical models to radar travel time and electrical resistivity tomograms : Resolution-dependent limitations. *Journal of Geophysical Research*, **110**(B08206).
- Domenico, P. et Schwartz, F. (1997). *Physical and chemical hydrogeology*. John Wiley and Sons, New York, 2e édition.
- Ellefsen, K. J. (1999). Effects of layered sediments on guided wave in crosswell radar data. *Geophysics*, **64**(6), 1698–1707.
- Emery, X. (2008). Uncertainty modeling and spatial prediction by multi-gaussian kriging : Accounting for an unknown mean value. *Computers and Geosciences*, **34**(11), 1431–1442.
- Eppstein, M. J. et Dougherty, D. E. (1998a). Efficient 3-D data inversion : Soil characterization and moisture monitoring from crosswell GPR at a Vermont test site. *Water Resources Research*, **34**(8), 1889–1900.

- Eppstein, M. J. et Dougherty, D. E. (1998b). Optimal 3-D traveltome tomography. *Geophysics*, **63**(3), 1053–1061.
- Fraser, H. J. (1935). Experimental study of the porosity and permeability of clastic sediments. *Journal of Geology*, **43**(8), 910–1010.
- Friedman, S. P. (1998). A saturation degree-dependent composite spheres model for describing the effective dielectric constant of unsaturated porous media. *Water Resources Research*, **34**, 2949–2961.
- Gallardo, L. A. et Meju, M. A. (2003). Characterization of heterogeneous near-surface materials by joint 2D inversion of DC resistivity and seismic data. *Geophysical Research Letters*, **30**(13).
- Gallardo, L. A. et Meju, M. A. (2004). Joint two-dimensional DC resistivity and seismic travel time inversion with cross-gradients constraints. *Journal of Geophysical Research*, **109**(B3), B03311.
- Gastaldi, C., Roy, D., Doyen, P., et Den Boer, L. (1998). Using bayesian simulations to predict reservoir thickness under tuning conditions. *The Leading Edge*, pages 589–593.
- Giroux, B. et Chouteau, M. (2004). The effect of clay seams in borehole GPR attenuation tomography. In Slob, E. and Yarovoy, A., éditeurs, *Proceedings of the Tenth International Conference on Ground-Penetrating Radar*, Delft, The Netherlands.
- Giroux, B. et Chouteau, M. (2008a). A hydrogeophysical synthetic model generator. *Computers and Geoscience*, **34**(9), 1080–1092.
- Giroux, B. et Chouteau, M. (2008b). Validity of the low-loss approximation in the estimation of dielectric constant and water content from borehole radar data.

Proceedings of the 12th International Conference on Ground-Penetrating Radar, Birmingham, UK; June.

Giroux, B., Gloaguen, E., et Chouteau, M. (2007). Bh-Tomo - A Matlab borehole georadar 2D tomography package. *Computers and Geoscience*, **33**, 126–137.

Gloaguen, E. (2004). *Reconstruction des images bidimensionnelles de levés de radar géologique en forage*. Thèse de doctorat, École Polytechnique de Montréal.

Gloaguen, E., Giroux, B., et Marcotte, D. (2009). Two-dimensional geostatistical porosity simulation based on wavelet decomposition of GPR tomography image and model analogue.

Gloaguen, E., Giroux, B., Marcotte, D., et Dimitrakopoulos, R. (2007a). Pseudo-full-waveform inversion of borehole GPR data using stochastic tomography. *Geophysics*, **72**(5), J43–J51.

Gloaguen, E., Marcotte, D., Giroux, B., Dubreuil-B., C., Chouteau, M., et Aubertin, M. (2007b). Stochastic borehole radar velocity and attenuation tomographies using cokriging and cosimulation. *Journal of Applied Geophysics*, **62**, 141–157.

Goovaerts, P. (1997). *Geostatistics for Natural Resources Evaluation*. Applied geostatistics series. Oxford University Press.

Grolier, J., Fernandez, A., Hucher, M., et Riss, J. (1991). *Les propriétés physiques des roches - Théories et modèles*. Masson.

Haber, E. et Oldenburg, D. (1997). Joint inversion : A structural approach. *Inverse Problems*, **13**(1), 63–77.

Hazen, A. (1892). Some physical properties of sand and gravel with special reference to their use in filtration. Technical report, Massachussets State Board of Health, Boston.

- Hördt, A., Blaschek, R., Kemma, A., et Zisser, N. (2007). Hydraulic conductivity estimation from induced polarization data at the field scale - the Krauthausen case history. *Journal of Applied Geophysics*, **62**, 33–46.
- Hubbard, S., Chen, J., Peterson, J., Majer, E., Williams, K., Swift, D. J., Mailloux, B., et Rubin, Y. (2001). Hydrogeological characterization of the South Oyster bacterial transport site using geophysical data. *Water Resources Research*, **37**(10), 2431–2456.
- Hubbard, S. et Rubin, Y. (2005). *Introduction to Hydrogeophysics*, chapitre 1. Dans : Rubin, Y., Hubbard, S., *Hydrogeophysics*, Springer.
- Hvorslev, M. (1951). Time Lag and Soil Permeability in Ground-Water Observations. *U.S. Army Engineer Waterways Experiment Station*, pages 1–50.
- Hyndman, D. et Gorelick, M. (1996). Estimating lithologic and transport properties in three dimensions using seismic and tracer data : the Kesterton aquifer. *Water Resources Research*, **32**(9), 2659–2670.
- Hyndman, D., Harris, J. M., et Gorelick, M. (1994). Coupled seismic and tracer test inversion for aquifer property characterization. *Ground Water*, **30**(7), 1965–1977.
- Hyndman, D., Harris, J. M., et Gorelick, M. (2000). Inferring the relation between slowness and hydraulic conductivity in heterogeneous aquifers. *Water Resources Research*, **36**(8), 2121–2132.
- Jackson, P. D., Taylor-Smith, D., et Stanford, P. N. (1978). Resistivity-porosity-particles shape relationships for marine sands. *Geophysics*, **43**, 1250–1268.
- Kirsch, R. (2006). *Groundwater Geophysics*. Springer.

- Kowalsky, M., Chen, J., et Hubbard, S. (2006). Joint inversion of geophysical and hydrological data for improved subsurface characterization. *The Leading Edge*, pages 730–734.
- Kowalsky, M., Finsterle, S., Peterson, J., Hubbard, S., Rubin, Y., Majer, E., Ward, A., et Gee, G. (2005). Estimation of field-scale soil hydraulic and dielectric parameters through joint inversion of GPR and hydrological data. *Water Resources Research*, **41**(w11425).
- Kowalsky, M., Finsterle, S., et Rubin, Y. (2004). Estimating flow parameter distributions using ground-penetrating radar and hydrological measurements during transient flow in the vadose zone. *Advances in Water Resources*, **27**(6), 583–599.
- Kozeny, J. (1927). Über kapillare Leitung des Wasser im Boden. *Sitzungsberichte Wiener Akademie*, **136**(2a), 271–306.
- Lesme, D. P. et Friedman, S. P. (2005). *Relationships between the electrical and hydrogeological properties of rocks and soils*, chapitre 4, pages 87–128. Dans : Rubin, Y., Hubbard, S., Hydrogeophysics, Springer.
- Linde, N. (2005). *Characterization of hydrogeological media using electromagnetic geophysics*. Thèse de doctorat, Uppsala University.
- Linde, N., Binley, A., Tryggvason, A., Laust, B., et Revil, A. (2006a). Improved hydrogeophysical characterization using joint inversion of cross-hole electrical resistance and ground-penetrating radar traveltimes data. *Water Resources Research*, **42**(W12404).
- Linde, N., Chen, J., Kowalsky, M., et Hubbard, S. (2006b). *Hydrogeophysical parameter estimation approaches for field scale characterization*. Dans : Vereecken, H. et al. Applied Hydrogeophysics, Chap. 2.

- Linde, N., Finsterle, S., et Hubbard, S. (2006c). Inversion of tracer test data using tomographic constraints. *Water Resources Research*, **42**(W04410).
- Marcotte, D. (1991). Cokriging with Matlab. *Computers and Geoscience*, **17**(9), 1265–1280.
- Marcotte, D. et Gloaguen, E. (2008). A class of spatial multivariate models based on copulas.
- Marion, D., Nur, A., Yin, H., et Han, D. (1992). Compressional velocity and porosity in sand-clay mixtures. *Geophysics*, **57**(4), 554–563.
- Mernoud, A. (2006). *Cours de physique du sol - Propriétés de base du sol et de la phase liquide*. École Polytechnique fédérale de Lausanne.
- Morin, R. H. (2006). Negative correlation between porosity and hydraulic conductivity in sand-and-gravel aquifers at Cape Cod, Massachusetts, USA. *Journal of Hydrology*, **316**, 43–52.
- Moysey, S., Singha, K., et Knight, R. (2005). A framework for inferring field-scale rock physics relationships through numerical simulation. *Geophysical Research Letters*, **32**(L08304).
- Olhoeft, G. R. (1981). *Electrical properties of rocks*, pages 257–330. Dans : Y.S. Touloukian, Y.S. Judd and R.F. Roy, éditeurs, *Physical Properties of Rocks and Minerals*.
- Paasche, H., Tronicke, J., Holliger, K., Green, A. G., et Maurer, H. (2006). Integration de diverse physical-property models : Subsurface zonation and petrophysical parameter estimation based on fuzzy c-means cluster analyses. *Geophysics*, **71**(3), H33–H44.

- Paradis, D., Lefebvre, R., Gloaguen, E., Tremblay, L., , et Ballard, J. (2008). Multivariate integration of direct push (mechanical and electrical properties) and hydraulic conductivity for the definition of hydrofacies in unconsolidated sediments. In *Conférence présentée à GeoEdmonton*.
- Pride, S. R. (1994). Governing equations for the coupled electromagnetics and acoustic of porous media. *Physical Review B*, **50**(21), 15678–15696.
- Purvance, D. T. et Andricevic, R. (2000). On the electrical-hydraulic conductivity correlation in aquifers. *Water Resources Research*, **36**, 2905–2913.
- Renard, P. (1997). *Modélisation des écoulements en milieux poreux hétérogènes*. Thèse de doctorat, École des mines de Paris.
- Revil, A. (1999). Ionic diffusivity, electrical conductivity, membrane and thermoelectric potentials in colloids and granular porous media : a unified model. *Journal of Colloid and Interface Science*, **212**(2), 503–522.
- Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, **27**(3), 832–837.
- Roth, K. R., Schulin, R., Fluhler, H., et Attinger, W. (1990). Calibration of time domain reflectometry for water content measurement using a composite dielectric approach. *Water Resources Research*, **26**, 2267–2273.
- Rubin, Y. et Hubbard, S. (2005). *Stochastic forward and inverse modeling : the hydrogeophysical challenge*, chapitre 17, pages 487–512. Dans : Rubin, Y., Hubbard, S., *Hydrogeophysics*, Springer.
- Schön, J. H. (1998). *Physical properties of rocks - Fundamentals and Principles of Petrophysics*, volume 18. Elsevier.

- Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, New York.
- Slater, L. (2007). Near surface electrical characterization of hydraulic conductivity : from petrophysical properties to aquifer geometries - A review. *Surveys Geophys*, **28**, 169–197.
- Slater, L. et Lesme, D. P. (2002). Electrical-hydraulic relationships observed for unconsolidated sediments. *Water Resources Research*, **38**(10).
- Topp, G., Davis, J., et Annan, A. (1980). Electromagnetic determination of soil water content : mesurements in coaxial transmission lines. *Water Resources Research*, **16**(3), 574–582.
- Tremblay, L., Gloaguen, E., Lefebvre, R., et Ballard, J. (2008). Integration of geological, geochemical, and direct push methods for the detection of leachate plumes. In *Conférence présentée à GeoEdmonton*.
- Tronicke, J. et Holliger, K. (2005). Quantitative integration of hydrogeological data : Conditional geostatistical simulation for characterizing heterogeneous alluvional aquifers. *Geophysics*, **70**(3), H1–H10.
- Vukovic, M. et Soro, A. (1992). Determination of hydraulic conductivity of porous media from grainsize composition. *Water Resources Publications*, pages 1–83.
- Walpole, R., Myers, R., S., M., et Ye, K. (2002). *Probability and statistics for engineers and scientists*. Prentice Hall.
- Waxman, M. H. et Smith, L. J. M. (1968). Electrical conductivites in oil-bearing sands. *Journal of the Society of Petroleum Engineers*, **8**(107-122).
- Zuber, J. (1999). *Estimation non paramétrique d'une fonction de densité*, chapitre 7. École Polytechnique Fédérale de Lausanne.