UNIVERSITÉ DU QUÉBEC
INSTITUT NATIONAL DE LA RECHERCHE SCIENTIFIQUE
CENTRE EAU TERRE ENVIRONNEMENT

CONTRIBUTIONS AU DÉVELOPPEMENT DE LA CARACTÉRISATION DE L'HÉTÉROGÉNÉITÉ HYDRAULIQUE DES AQUIFÈRES

Par

Daniel PARADIS

Thèse présentée pour l'obtention du grade de *Philosophiae Doctor*, Ph.D.
en sciences de la terre

Jury d'évaluation

Examinateur interne et président du jury	Dr. Richard Martel INRS Centre Eau Terre Environnement
Examinateur externe	Dr. Geoffrey C. Bohling Kansas Geological Survey
Examinateur externe	Dr. Laurence R. Bentley University of Calgary
Directeur de recherche	Dr. René Lefebvre INRS Centre Eau Terre Environnement
Codirecteurs de recherche	Dr. Erwan Gloaguen INRS Centre Eau Terre Environnement
	Dr. Alfonso Rivera Commission géologique du Canada

PARADIS, Daniel (2014) *Contributions au développement de la caractérisation de l'hétérogénéité hydraulique des aquifères.* Thèse de doctorat, Université du Québec, Centre Eau Terre Environnement. 290 pages, 83 figures, 25 tables et 6 annexes (1 papier et 5 DVD), avec articles originaux (I-V).

RÉSUMÉ

Cette thèse comprend cinq articles originaux et un article synthèse traitant de la caractérisation hydraulique de l'hétérogénéité des aquifères granulaires. Cette thèse reconnait l'importance des données de caractérisation pour le développement d'une bonne compréhension des processus d'écoulement et de transport dans les aquifères. De nouveaux tests hydrauliques in situ pour caractériser en détails les propriétés hydrauliques des aquifères et leur corrélation non paramétrique avec des propriétés géophysiques indirectes sont utilisés pour définir l'hétérogénéité hydraulique à petite échelle. Le but de cette thèse est d'explorer le contenu en information de ces nouvelles méthodes et d'évaluer leur potentiel pour des applications sur le terrain dans des conditions aquifères hétérogènes et anisotropes. Les tests hydrauliques in situ réalisés pour la région d'étude sont des adaptations du débitmètre de puits et des essais de perméabilité à choc hydraulique couramment utilisés. Les tests avec débitmètre ont été adaptés pour les aquifères granulaires par l'aménagement de puits par enfoncement dotés de longues crépines en contact direct avec les sédiments, ce qui permet l'acquisition de profils verticaux de la conductivité hydraulique horizontale (K_h). Ce type de puits, sans sable filtrant, réduit les erreurs de mesures dues à la dissipation hydraulique et aux court-circuits hydrauliques associés à la présence du sable filtrant utilisé pour les puits conventionnels, ce qui est tout indiqué pour la réalisation de tests hydrauliques sur de petits intervalles verticaux dans les aquifères granulaires. Des tests d'interférence verticale à choc hydraulique utilisant un seul puits et la tomographie à choc hydraulique entre puits sont aussi proposés pour définir l'hétérogénéité de K_h, l'anisotropie de K (K_v/K_h) et l'emmagasinement spécifique (S_s) par l'analyse simultanée des perturbations hydrauliques transitoires générés par plusieurs tests à choc hydraulique effectués sur de petits intervalles. De plus, une approche non paramétrique par machine d'apprentissage de type RVM (*Relevance Vector Machine*) est proposée pour prédire K_h à partir de sondages au piézocône (CPT) couplés avec une sonde de teneur en eau et de résistivité électrique (SMR). La machine d'apprentissage est construite à partir de données d'entrainement colocalisées de K_h et de CPT/SMT mises à une échelle verticale commune de 15 cm. Pour acquérir un ensemble de données d'entrainement représentatif de l'aquifère littorale étudié, une approche d'acquisition des données est développée et appliquée avec des méthodes de caractérisation générales précédant les méthodes plus spécifiques. Ceci minimise le nombre de sites d'investigation tout en maximisant la signification des données directes et indirectes acquises. Diverses expérimentations in situ montrent que les estimations des propriétés hydrauliques obtenues avec les nouvelles méthodes sont cohérentes par rapport aux tests in situ et de laboratoire conventionnels. En particulier, le débitmètre et l'approche hydro-géophysique sont très efficaces pour estimer K_h, tandis que les tests d'interférence verticale et la tomographie à choc hydraulique donnent aussi des estimations fiables de K_h, mais en plus permettent d'évaluer K_v/K_h et S_s qui sont des paramètres difficiles à estimer autrement. Par ailleurs, la conversion de plus de 500 m de données CPT/SMR en estimations de propriétés hydrauliques et leur interpolation sur la région d'étude couvrant 12 km^2 révèlent les caractéristiques fondamentales de l'environnement de sédimentation littoral avec l'inter-digitation et les variations spatiales graduelles des propriétés hydrauliques. La contribution majeure de cette thèse est le renforcement des capacités de l'hydrogéologue à représenter de façon détaillée l'hétérogénéité des propriétés hydrauliques à des échelles pertinentes pour l'étude de l'écoulement et du transport dans les aquifères.

ABSTRACT

This thesis comprises five original papers and one synthesis paper dealing with hydraulic characterization of unconsolidated aquifer heterogeneity. This thesis recognizes the importance of characterization data to develop a sound understanding of processes associated with flow and transport in granular aquifers. Newly developed in-situ hydraulic tests designed to characterize aquifer hydraulic properties in details and their non-parametric correlation with indirect geophysical properties are then used to define aquifer hydraulic heterogeneities at fine scale. The aim of the thesis is to explore the information content of the new methods, and to assess their potential for field application under real heterogeneous and anisotropic aquifer conditions. The in-situ hydraulic tests carried out in this study are adaptations of flowmeter tests and commonly used slug tests. Flowmeter testing provides vertical profiles of horizontal hydraulic conductivity (K_h) following their adaptation for granular aquifers using direct-push wells with long screened sections in direct contact with sediments. Those wells reduce measurement bias associated with skin effects and hydraulic short-circuits of conventional wells using sand packs, and are thus tailored for hydraulic testing over small vertical intervals in unconsolidated aquifers. Vertical interference slug tests in a single direct-push well and tomographic slug tests between direct-push wells are also proposed to resolve heterogeneity in K_h, K anisotropy (or K_v/K_h) and specific storage (S_s) by the simultaneous analysis of the transient hydraulic perturbations generated by multiple slug tests performed over small intervals. Moreover, a non-parametric learning machine approach is proposed to predict K_h from cone penetrometer tests (CPT) coupled with a soil moisture and resistivity probe (SMR) using relevance vector machines (RVMs). The learning machine is built from a colocated training data set of K_h and CPT/SMR up-scaled measurements at a common vertical resolution of 15 cm. To acquire a representative training data set of the littoral aquifer found in the study area, a top down data acquisition process is developed and applied using general measurements prior to specific ones to minimize the number of investigated sites while maximizing the quality, extent and significance of the collected direct and indirect data. Field proof-of-concept studies show that estimations of hydraulic properties from the new methods are consistent with conventional field and laboratory tests. In particular, flowmeter tests and a hydro-geophysical approach can provide K_h estimates in a timely fashion, whereas vertical interference tests and tomographic slug tests can also provide K_h data, but also K anisotropy and S_s estimates that are difficult to obtain otherwise. Moreover, the conversion of more than 500 m of CPT/SMR data into hydraulic properties values and their interpolation over the 12 km^2 study area are consistent with the littoral depositional environment showing inter-digitation and spatial trends in hydraulic properties. The major contribution of this thesis is to enhance the capability of hydrogeologists to image heterogeneity in hydraulic properties in details at a scale relevant for flow and transport studies in aquifers.

Key words (10 mots max): Groundwater, Hydraulic Conductivity Anisotropy, Heterogeneity, Direct-Push, Flowmeter, Vertical Interference Slug Tests, Hydraulic Tomography, Learning Machine, Unconsolidated Aquifer, Hydrogeophysics

« *Il est bon d'avoir appris à ses dépends ce qu'on a besoin de savoir.* »

Hermann Hesse

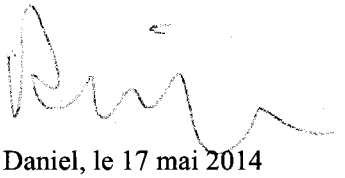
REMERCIEMENTS

Bien que la mise-à-terme d'un projet de thèse se veuille la démonstration d'un accomplissement personnel, la réalisation de mon projet thèse n'aurait pu voir le jour sans la présence, le support et le réconfort d'un bon nombre de proches.

Ma plus grande reconnaissance va à Yves Michaud qui a été le déclencheur et le plus cher supporteur de ce projet depuis son début. Je dois aussi ma reconnaissance à Donna Kirkwood et Alfonso Rivera qui ont facilité la réalisation de ce projet dans le cadre de mes fonctions professionnelles à la Commission géologique du Canada à Québec. Votre confiance et votre support, à vous tous, m'ont donnés l'inestimable latitude de développer toutes les idées présentées dans ce projet de recherche.

Ma gratitude va aussi à René Lefebvre qui a toujours été un mentor remarquable tout au long de mon parcours professionnel et académique, et dont j'apprécie particulièrement les plus grandes qualités: générosité et empathie. Je remercie aussi Erwan Gloaguen, qui arrivé à un moment charnière au début de cette thèse, a grandement inspiré son contenu, et qui au fil des années est devenus un précieux « *compagnon d'armes* ».

Je veux aussi remercier chaleureusement mes parents, Jacques et Nicole, pour leur présence et leur confiance lorsque ce fut nécessaire, ainsi que mes frangins, Nicolas et Élise, pour leur estime réciproque. Et finalement, merci à Anik, Anaïs et Thomas, qui sont une source inépuisable de bonheur.



Daniel, le 17 mai 2014

TABLE DES MATIÈRES

LISTE DES TABLEAUX

LISTE DES FIGURES

PREMIÈRE PARTIE : SYNTHÈSE

SYNTHÈSE

CONTRIBUTIONS AU DÉVELOPPEMENT DE LA CARACTÉRISATION DE L'HÉTÉROGÉNÉITÉ HYDRAULIQUE DES AQUIFÈRES

1 Introduction

La résolution de questions environnementales, telles que l'exploitation responsable des ressources en eau souterraine ou la réhabilitation de sites contaminés, exige la quantification des processus d'écoulement de l'eau souterraine et de transport des contaminants au moyen de modèles hydrogéologiques numériques (Anderson et Woessner 1992). Comme le montre la Figure S.1, la modélisation hydrogéologique se fait habituellement en plusieurs étapes, et comprend notamment la caractérisation du site, le développement d'un modèle conceptuel hydrogéologique, l'élaboration d'un modèle numérique, l'estimation des paramètres et l'évaluation de l'incertitude dans les prévisions (Neuman et Wieranga 2003). Les données de caractérisation et de suivi constituent le fondement de ce processus de modélisation. En effet, les données de caractérisation du site, qui permettent l'attribution de valeurs aux paramètres du modèle (ex., conductivité hydraulique, porosité, recharge), forment la base sur laquelle un ou plusieurs modèles conceptuels peuvent être postulés pour une région d'étude. Pour évaluer la capacité d'un modèle numérique à reproduire la dynamique réelle d'un système aquifère, il est également nécessaire de disposer de données de suivi qui fournissent des observations (ex., charges, concentrations, débit) des conditions hydrogéologiques réelles. Cependant, il existe généralement un vaste éventail de modèles et d'ensembles de paramètres qui peuvent mener à des simulations acceptables, même pour les sites faisant l'objet d'études intensives. Cela s'explique par l'impossibilité de décrire les systèmes aquifères de façon exhaustive, particulièrement l'hétérogénéité des propriétés hydrauliques des aquifères. Ainsi, la fiabilité des prédictions tirées d'un modèle hydrogéologique numérique est fortement liée au réalisme du modèle conceptuel qui peut être construit à partir de l'information hydrogéologique disponible.

Les études hydrogéologiques réalisées dans le but de comprendre l'écoulement de l'eau souterraine ou le transport de contaminants nécessitent des données représentatives sur la répartition spatiale hétérogène des propriétés hydrauliques des aquifères, comme la conductivité hydraulique (K) et la porosité (n) (Anderson *et al.* 1999; Eaton 2006). Pour les aquifères granulaires, l'hétérogénéité des propriétés hydrauliques est le résultat de processus géologiques complexes (Koltermann et Gorelick 1996). Cette hétérogénéité influence l'écoulement de l'eau souterraine et particulièrement les processus de transport de masse dans les aquifères à différentes échelles. En effet, l'hétérogénéité de K à grande échelle contrôle les voies d'écoulement de l'eau souterraine et des contaminants, tandis que les variations de K à petite comme à grande échelle modulent l'ampleur de la dispersion des contaminants. La représentation fine de l'hétérogénéité des propriétés hydrauliques au moyen de la définition d'hydrofaciès (HF) est généralement utile pour conceptualiser la complexité des systèmes aquifères (Anderson 1997; McKenna

et Poeter 1995; Ouellon *et al.* 2008; Frei *et al.* 2009). Un HF est une unité homogène, possiblement anisotrope, et distincte d'un point de vue de ses propriétés hydrauliques pour les besoins de la modélisation hydrogéologique (Anderson 1989). La caractérisation complète d'un système aquifère devrait permettre idéalement une représentation spatiale tridimensionnelle des HF et des variations de valeurs des propriétés hydrauliques dans chacun de ces HF.

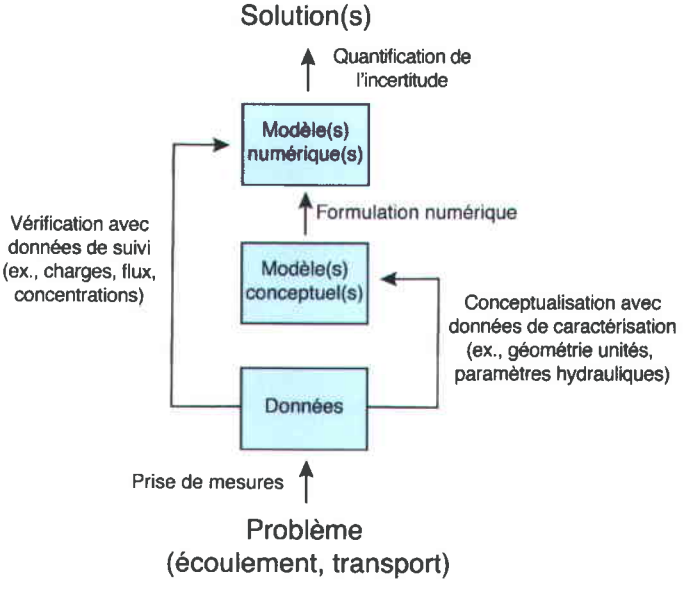


Figure S.1 Schéma illustrant le processus général de modélisation hydrogéologique pour la résolution d'un problème d'écoulement ou de transport.

Il peut être possible de concevoir un modèle hydrogéologique conceptuel pour un site à partir de données génériques de systèmes analogues ou à partir de mesures effectuées ailleurs sur des matériaux géologiques similaires. Toutefois, parce que chaque site est unique, la caractérisation locale devrait révéler les caractéristiques particulières du site. Donc, comme point de départ d'une bonne gestion d'un aquifère, l'échelle d'hétérogénéité pertinente à toute étude hydrogéologique devrait toujours être reconnue et caractérisée en conséquence à l'aide de mesures locales. Comme l'ont souligné de Marsily *et al.* (2005) et comme l'illustre la Figure S.2, le type de problème (ex., écoulement ou transport), l'échelle du site étudié, le milieu géologique, le type de réponse à donner (ex., ponctuelle ou diffuse), les méthodes de caractérisation disponibles et les coûts de l'acquisition des données doivent donc tous être considérés dans la résolution du problème.

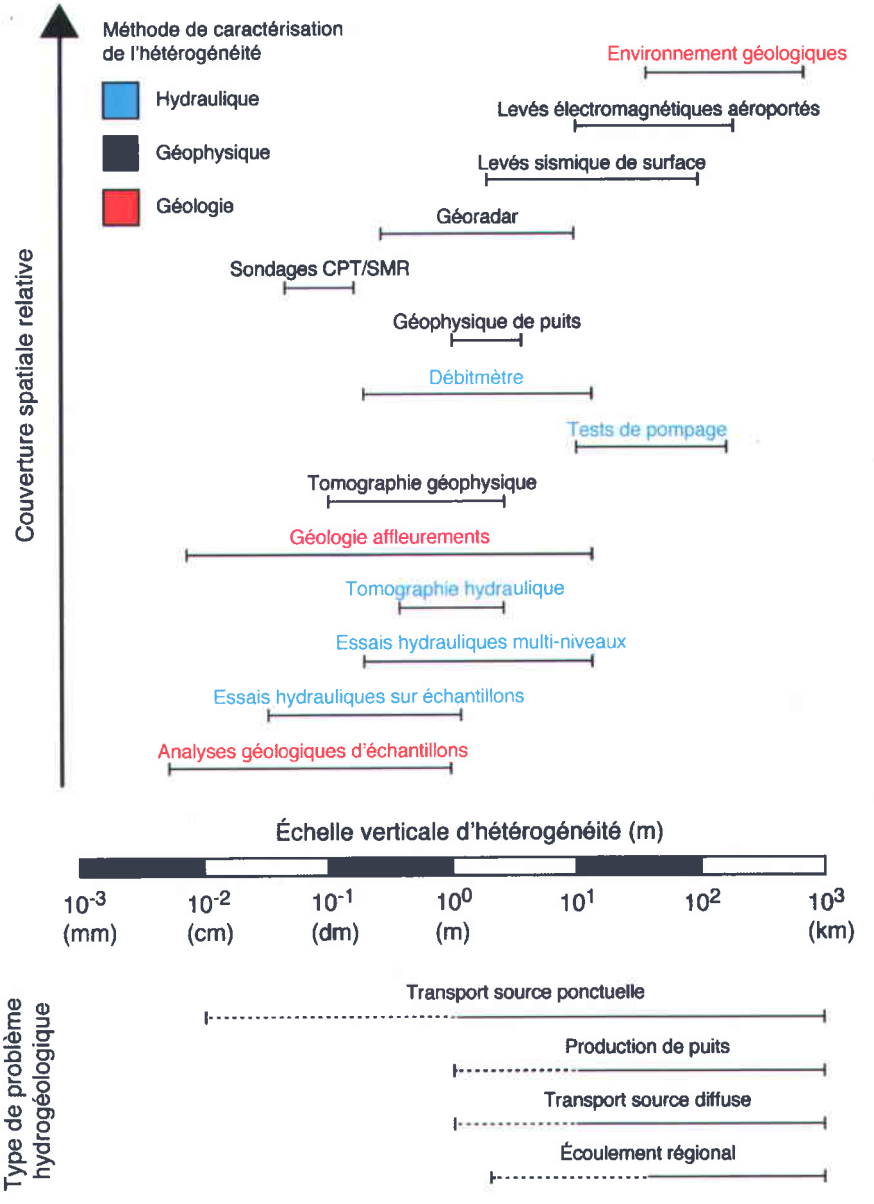


Figure S.2 La partie supérieure illustre les rapports relatifs entre l'échelle d'hétérogénéité d'un aquifère, l'échelle verticale de résolution et la couverture spatiale des méthodes de caractérisation. La partie inférieure montre l'échelle du problème d'écoulement ou de transport à considérer.

Les essais hydrauliques réalisés dans des puits sont habituellement des sources d'information fiables sur les propriétés hydrauliques des aquifères (Butler 2005). Cependant, les essais hydrauliques conventionnels, comme les tests de pompage, induisent principalement un écoulement horizontal et n'estiment par conséquent que la conductivité hydraulique horizontale (K_h). Une bonne compréhension de l'écoulement de l'eau souterraine et du transport de contaminants exige également la connaissance de l'anisotropie de K, c.-à-d. le rapport des conductivité hydraulique verticale K_v et horizontale K_h (K_v/K_h) et

de la porosité n. K_v et n peuvent être généralement obtenues à partir d'essais en laboratoire effectués sur des échantillons de sédiments, habituellement obtenus par carottage. Cependant, la durée habituellement prolongée des essais hydrauliques conventionnels et les difficultés expérimentales liées au prélèvement et à la manipulation d'échantillons de sédiments limitent souvent l'obtention de grandes quantités de mesures directes de K_h, K_v et n. Par conséquent, il est difficile d'acquérir suffisamment d'information et de détails sur la distribution spatiale des propriétés hydrauliques pour concevoir des modèles hydrogéologiques réalistes des sites à l'étude (Gelhar 1993; Anderson 1997; Butler 2005). La motivation de cette thèse est donc d'élaborer des approches de caractérisation qui permettraient une meilleure définition de l'hétérogénéité des propriétés hydrauliques des aquifères.

Dans cette synthèse, la contribution de la thèse à l'avancement des connaissances et la cohérence de la démarche scientifique dont les résultats sont divisés dans les différents articles sont illustrées au moyen de la présentation du cadre général qui a été appliqué pour la caractérisation des propriétés hydrauliques d'un aquifère granulaire effectuée pour l'étude de la migration des lixiviats d'un ancien site d'enfouissement sanitaire. La région visée par l'étude se trouve à St-Lambert-de-Lauzon, à 30 km au sud de Québec, et couvre une superficie de 12 km^2 autour de l'ancien site d'enfouissement (Tremblay *et al.* 2013). Les sédiments qui composent l'aquifère montrent des changements de texture transitionnels à petite et à grande échelle, reflétant le milieu littoral de leur sédimentation. Afin de définir le plus réalistement la complexité de l'aquifère étudié, l'accent a été mis sur l'acquisition de données géophysiques indirectes à haute résolution verticale de manière à acquérir un maximum d'information à moindre coût. Pour ce faire, des essais au piézocône (CPT : *Cone Penetration Tests*) combinés à une sonde permettant de mesurer la teneur en eau et la résistivité des sédiments (SMR : *Soil Moisture and Resitivity*) ont été utilisés pour répondre aux besoins de l'étude. L'avantage des données CPT/SMR pour la caractérisation hydrogéologique repose dans la résolution verticale à l'échelle du décimètre (Lunne *et al.* 1997; Schulmeister *et al.* 2003), qui ne peut être obtenue par les méthodes géophysiques de surface. En outre, l'importante couverture spatiale pouvant être obtenue par des sondages CPT/SMR comparativement à des essais hydrauliques dans des puits ou sur des carottes de sédiments permettent de déduire l'hétérogénéité des aquifères sur de grandes superficies (Lafuerza *et al.* 2005). Pour identifier les HF et estimer K_h à partir des données CPT/SMR, Paradis *et al.* (2011; 2014c) ont proposé une approche d'intégration des données hydro-géophysiques utilisant des machines d'apprentissage. Cette approche d'intégration représente le cœur du cadre de caractérisation de l'hétérogénéité des propriétés hydrauliques présentée dans cette synthèse et est aussi l'objet de l'Article V de la présente thèse.

Puisque l'approche d'intégration par machine d'apprentissage adoptée ici est basée sur l'établissement de relations empiriques propres à un site, un processus systématique d'acquisition des données a été suivi afin d'acquérir un ensemble de données représentatif de la région d'étude. Dans ce processus, l'application des méthodes de caractérisation plus générales précède l'application des méthodes plus spécifiques et le choix des sites de caractérisation pour les méthodes spécifiques est fait selon les caractéristiques des mesures obtenues avec les méthodes plus générales. Les méthodes de caractérisation générales offrent généralement une plus grande couverture spatiale et une résolution moindre, alors que les méthodes spécifiques permettent une plus grande résolution sur de plus petits volumes d'investigation. Ce processus qui vise à rationaliser l'acquisition des données est aussi utilisé couramment dans l'industrie pétrolière aux fins de prospection des hydrocarbures (Bradford et Babcock 2013).

À part la définition des relations empiriques pour prédire les HF et K_h à partir de données CPT/SMR, l'intégration des données comprend aussi la définition de relations empiriques qui lient K_v et n à K_h, et l'interpolation spatiale donnant la distribution spatiale de l'hétérogénéité des HF ainsi que de K_h, K_v et n sur la région d'étude. À l'appui du cadre général de caractérisation, des méthodes hydrauliques permettant de mesurer plus efficacement K_h et K_v ont également été développés pour cette thèse. En effet, l'utilisation du débitmètre de puits a été adaptée aux aquifères granulaires par Paradis *et al.* (2011) pour les mesures de K_h, tandis que des tests d'interférence verticale à chocs hydrauliques (Paradis et Lefebvre 2013) et la tomographique par chocs hydrauliques (Paradis *et al.* 2014a; 2014b) ont été développés pour l'estimation de K_v. Ces méthodes font l'objet des Articles I, II, III et IV, respectivement.

La contribution de cette thèse à l'avancement des connaissances se situe à deux niveaux. Premièrement, cette thèse contribue à l'amélioration de la pratique de la caractérisation hydrogéologique en :

- proposant et appliquant une approche systématique et rationnelle d'acquisition des données de base représentatives pour la définition de l'hétérogénéité des propriétés hydrauliques à l'échelle d'un sous-bassin (Synthèse et article de l'Annexe A);
- adaptant et développant des méthodes de mesures *in situ* des propriétés hydrauliques plus efficaces et plus sophistiquées (Articles I, II et IV);
- proposant, appliquant et démontrant le gain en efficacité d'une approche de caractérisation hydrogéologique fondée sur l'intégration de mesures hydro-géophysiques (Article V, Synthèse et article de l'Annexe A).

Deuxièmement, cette thèse contribue aussi à la compréhension des fondements physiques qui régissent l'écoulement dans les aquifères et des liens existant entre les propriétés hydrauliques et géophysiques des aquifères en :

- démontrant les possibilités et les limites des méthodes de caractérisation hydraulique par l'étude du contenu en information d'une analyse transitoire des tests de perméabilité (*slug tests*) réalisés en mode tomographique (Article III);
- démontrant les possibilités et les limites d'une approche de caractérisation hydro-géophysique fondée sur des mesures géophysiques multi-paramètres (Article V).

Cette synthèse présente tout d'abord la région d'étude ainsi qu'une description générale des travaux d'acquisition des données effectués pour la caractérisation de l'hétérogénéité des propriétés hydrauliques. L'approche d'intégration des données est ensuite présentée (ce qui comprend une brève description de la machine d'apprentissage et de l'approche d'interpolation) et illustrée pour une section représentative de l'aquifère. L'efficacité relative de l'approche hydro-géophysique proposée pour prédire K_h par rapport aux essais hydrauliques conventionnels est également un élément au cœur de la discussion. On retrouvera aussi en annexes (DVD) les mesures de base permettant de reproduire les analyses présentées dans cette synthèse et dans les articles originaux de la deuxième partie, et qui n'ont pas été intégrés au texte par soucis de concision. On retrouvera aussi à l'Annexe A la synthèse actuelle soumise pour publication en version anglaise abrégée.

2 Région d'étude et acquisition des données

2.1 Résumé de l'approche d'acquisition de données

La Figure S.3 illustre la succession générale des travaux d'acquisition des données suivis dans cette étude pour caractériser l'hétérogénéité des propriétés hydrauliques. La succession des travaux est systématique des plus générales vers les plus spécifiques et la localisation des mesures pour une méthode particulière dépend des réponses globales des méthodes appliquées précédemment. Les méthodes générales ont généralement une plus grande couverture spatiale et une plus faible résolution, tandis que les méthodes plus spécifiques sont appliquées sur de plus petits volumes d'investigations. De plus, les méthodes générales, comme l'analyse de la géologie régionale et les levés géoradar, fournissent des renseignements

indirects à propos des propriétés hydrauliques, tandis que les méthodes spécifiques comme les essais de perméabilité, fournissent des estimations directes. Cette succession des travaux d'acquisition de données vise à rationaliser l'acquisition des données et à maximiser la signification des mesures recueillies. Par exemple, la succession des travaux pour cette synthèse est centrée sur l'acquisition d'un ensemble de mesures hydrauliques et CPT/SMR représentatif de la région d'étude afin d'établir des relations empiriques fiables pour la prédiction des propriétés hydrauliques à partir des mesures géophysiques. Ainsi, des mesures CPT/SMR sont effectués dans un premier temps sur l'ensemble de la région d'étude afin de couvrir toute la plage possible de valeurs des paramètres géophysiques et des essais hydrauliques ciblés sont ensuite effectués avec parcimonie à des endroits stratégiques selon les réponses CPT/SMR obtenues. Les mesures hydrauliques ainsi acquissent bonifies les mesures CPT/SMR et leur nombre restreint permet de réduire considérablement le temps de caractérisation associé aux essais hydrauliques. De plus, parce que l'emplacement des sondages CPT/SMR est choisis selon l'étude de la géologie et les levés géoradar réalisés précédemment, les sondages peuvent fournir plus de détails pertinents à propos de l'hétérogénéité des propriétés hydrauliques par rapport à une approche conventionnelle utilisant seulement des essais hydrauliques dans des puits ou sur des carottes de sédiments. Les sous-sections suivantes présentent donc l'application de la démarche d'acquisition de données illustrée à la Figure S.3 en vue de la caractérisation de l'hétérogénéité des propriétés hydrauliques de l'aquifère de St-Lambert.

2.2 Géologie régionale et du site de St-Lambert

La démarche d'acquisition des données proposée a été appliquée à une région d'étude située à St-Lambert-de-Lauzon, à 30 km au sud de Québec, au Canada (Figures S.4a-b). Comme l'illustre la Figure S.4c, la région d'étude englobe un sous-bassin hydrographique de 12 km^2 entourant un site d'enfouissement sanitaire déclassé dans le but d'évaluer la migration et l'atténuation d'un panache de lixiviation. Les récepteurs potentiels du panache émis par l'ancienne décharge sont des ruisseaux qui délimitent la région d'étude au nord, à l'ouest et au sud (Tremblay *et al.* 2013). La limite à l'est correspond à une ligne de partage des eaux souterraines déterminée à l'aide de la carte piézométrique de la région d'étude.

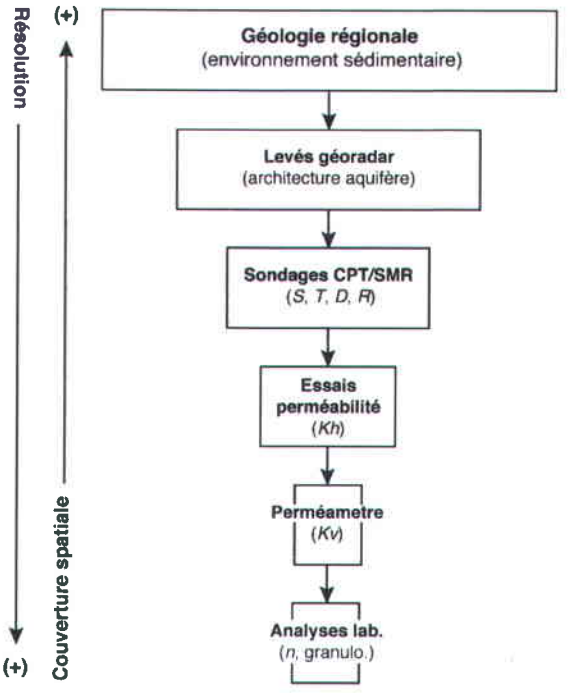


Figure S.3 Processus général d'acquisition séquentiel des données pour la caractérisation de l'hétérogénéité des propriétés hydrauliques de l'aquifère de St-Lambert (*S*: friction; *T*: résistance en pointe; *D*: constante diélectrique; *R*: résistivité électrique; K_h: conductivité hydraulique horizontale, K_v: conductivité hydraulique verticale; *n*: porosité totale).

L'une des premières étapes du processus de caractérisation de l'aquifère consistait à faire l'analyse des photos aériennes et des sédiments exposés dans les sablières afin de définir le milieu sédimentaire de la région d'étude. La connaissance du milieu sédimentaire est particulièrement utile pour apprécier l'architecture des sédiments et les tendances spatiales dans la variation de texture des sédiments, qui peuvent fournir des renseignements qualitatifs important à propos de l'hétérogénéité des propriétés hydrauliques (Heinz *et al.* 2003). Tel que décrit par Bolduc (2003), la géologie de surface de la région d'étude (Figure S.4b) est composée principalement de sables et de silts d'âge Quaternaire récent qui ont été déposés durant le retrait de la mer de Champlain, laquelle fut un bras de l'océan Atlantique occupant la vallée du St-Laurent au temps de la dernière déglaciation. Plus particulièrement, la sédimentation au site de St-Lambert a été contrôlée par des courants littoraux qui ont redéposés dans des environnements littoral et sous-littoral les sédiments provenant du paléo-delta de la rivière Chaudière. Ceci est indiqué à la Figure S.4c par la diminution de la granulométrie des sédiments suivant l'axe sud-ouest et par l'orientation sud-ouest nord-est de plusieurs crêtes de plages (Lamarche et Tremblay 2012). Les crêtes de plage et les barres de sable associés sont principalement composées de sable moyen à fin alors que des sédiments plus fins, tel que les silts, se retrouvent dans les zones de plus faible énergie entre les crêtes (Reading 1996). Les lithofaciès riche en silts sous les sables littoraux ont été déposés dans des

environnements sous-littoraux au front du littoral, tandis que ceux inter-stratifiés avec des sables (CPT24) ont été déposés dans les crêtes et les lagons aux arrières du littoral durant la régression marine.

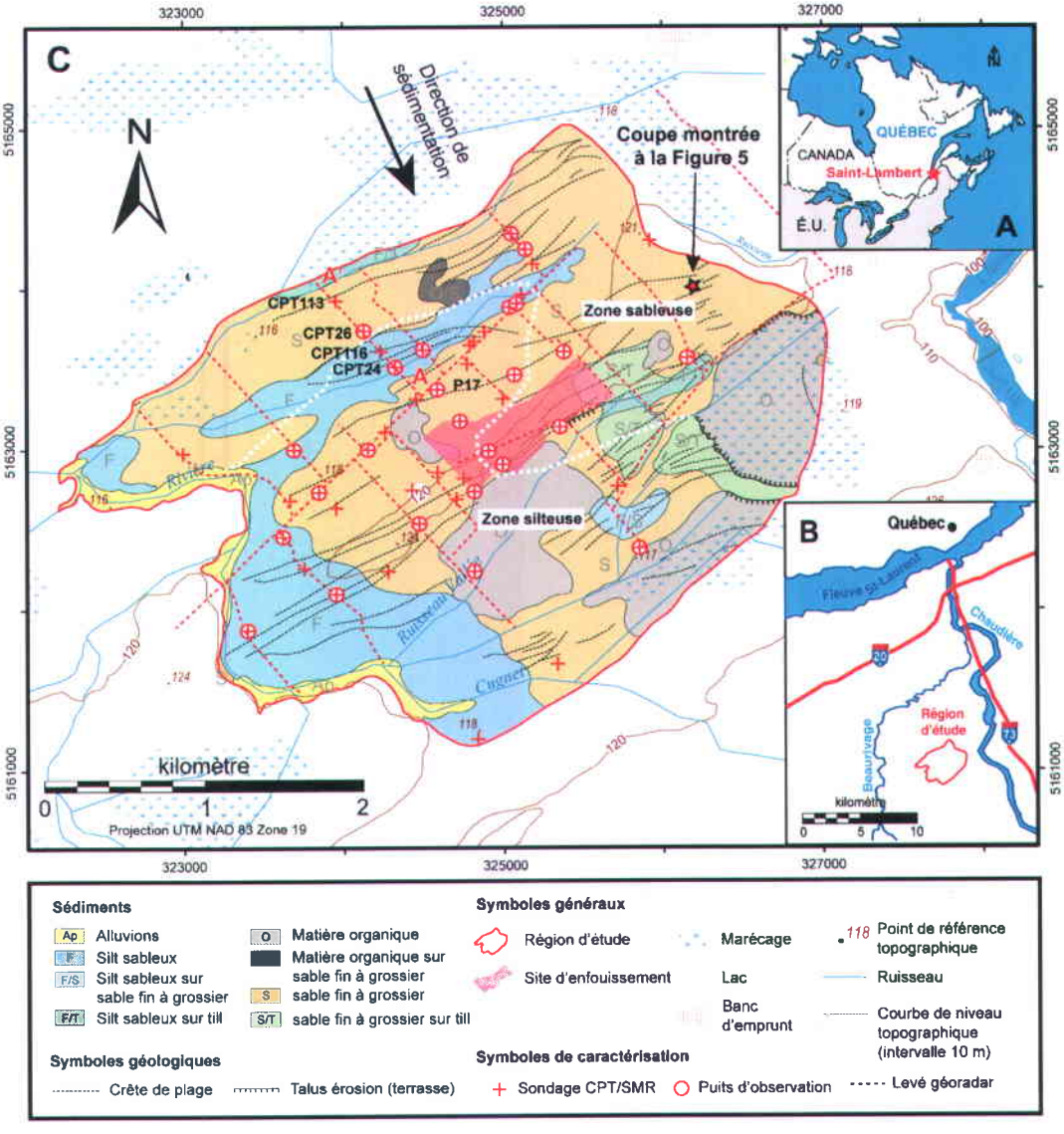


Figure S.4 Localisation générale de la région d'étude (a-b), avec la carte des sédiments Quaternaire de surface (c) pour le sous-bassin englobant l'ancien site d'enfouissement avec la localisation des levés géoradar, des sondages CPT/SMR et les puits d'observation utilisés pour la caractérisation de l'aquifère. La direction générale de déposition des sédiments composant l'aquifère est assumé perpendiculaire à l'orientation du paléo-rivage dans un environnement littoral. La ligne blanche pointillée représente la séparation entre la zone plus sableuse et la zone silteuse (tel que montré sur la section A-A' aux Figures S.4 et S.12). La carte Quaternaire a été modifiée de Lamarche et Tremblay (2012).

La granulométrie des sédiments de l'aquifère varie du sable fin jusqu'au silt très fin (silt argileux) avec des distributions granulométriques allant de mal à très mal trié. La granulométrie des sédiments et leur tri

est le résultat de deux processus sédimentaires, à savoir le transport par saltation et par suspension, qui se produisent simultanément dans un milieu littoral (Reineck et Singh 1980). Le degré de trie des sables et des silts au site de St-Lambert peut donc donner des valeurs de K relativement plus faibles comparativement à des sédiments beaucoup mieux triés, tel que retrouvé par exemple dans les environnements deltaïques. Comme le montre la Figure S.5, des transitions granulométriques marquées causés par des variations rapides d'énergie de sédimentation sont aussi souvent observés sur les sections de sédiments exposés dans les sablières. La superposition de couches de sable et de silt qui en résulte peut aussi créer des conditions aquifères semi-confinées en raison des contrastes de K associés aux différents sédiments. Dans chacune des couches, des transitions latérales marquées ou transitionnelles peuvent également être anticipés selon le niveau d'énergie ayant pu exister le long du littoral de la mer de Champlain.

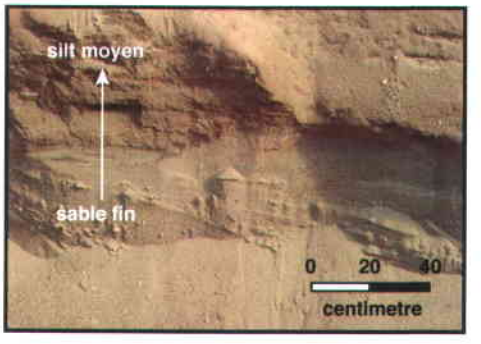


Figure S.5 Photo d'une coupe dans les sédiments montrant un contact transitionnel vertical allant d'un sable sec à un silt humide. La localisation de la coupe est montrée à la Figure S.2c.

2.3 Levés géoradar

Dans cette étude, le géoradar fut choisi comme outil d'exploration pour définir l'architecture de l'aquifère de la région d'étude en complément à la géologie régional. Le géoradar est une technique géophysique rapide et bien établie servant à l'acquisition non-destructive de données stratigraphiques des sédiments (Van Overmeeren 1998; Schmelzbach *et al.* 2011). Le géoradar utilise des ondes électromagnétiques à hautes fréquences transmises dans le sol, où elles se propagent en s'atténuant, et éventuellement sont réfractées et réfléchies à la surface du sol pour être mesurées par un receveur qui enregistre le signal électromagnétique en fonction du temps (Davis et Annan 1989). Les variations de la teneur en eau, de la fraction argileuse et de la granulométrie déterminent les coefficients de réflexion des différents matériaux. La résolution verticale des structures sédimentaires est de l'ordre d'un quart de longueur

d'onde qui correspond, pour les antennes de 100 MHz utilisée dans le cadre de cette étude, à environ 0.20 m (Tableau S.1).

Tableau S.1 Sommaire de l'acquisition des données de caractérisation (terrain et laboratoire).

Type de levé	Instrument	Configuration	Résolution verticale (m)	Nombre de levés	Longueur totale (m)
Géoradar	Pulse Echo	100 MHz antennes	0.2	20	21 000
CPT/SMR	Geotech 605D	Système acquisition Vertek	0.03 à 0.17	53	509
Puits	Geotech 605D avec tubage métal 7.72 cm (3 po) diamètre	Puits par enfoncement sans sable-filtrant (PVC 5.1 cm (2 po) diamètre)	n.a.	25	192 (longueur crépinée)
Sédiments	Geotech 605D avec Geotech Macro-Core Sampler	Tubage claire PETG de 1.5 m (3.8 cm (1.5 po) diamètre)	n.a.	75	95
Essais perméabilité (*slug test*) multi-niveaux (K_h)	Système initiation pneumatique avec capteur de pression	Système double obturateur	0.15	280	42
Essais perméamètre (K_v)	Perméamètre à charge variable avec capteur de pression	Avec le tubage original sans re-compaction des sédiments	0.15	59	9
Porosité (n)	Four à 50 °C (>7 jours)	Poids mouillé et sec	0.15	43	6.5
Distribution granulométrique	Fritsch Analysette 22	Réflexion laser	~0.01	59	0.6

Habituellement, les amplitudes des réflexions peuvent être corrélées aux contrastes de granulométries des sédiments et être utilisées pour déterminer les zones ayant des conditions hydrogéologiques différentes ou similaires. Jusqu'à 21 km de levés géoradar ont été acquis dans la région d'étude afin de déterminer la géométrie des structures internes de l'aquifère (Figure S.4c et Tableau S.1). Par exemple, la section géoradar de 100 MHz de la Figure S.6 illustre la stratification générale des sédiments observée dans la région de St-Lambert (voir l'emplacement de la section à la Figure S.4c). La Figure S.6 montre également les profils stratigraphiques obtenus à partir des sondages CPT/SMR, dont il est question dans la section suivante. Bien qu'en général, ces profils n'aient pas été disponibles au moment des levés géoradar, ils sont présentés ici pour appuyer la discussion.

Plusieurs observations peuvent être faites à partir de la Figure S.6. Premièrement, la stratification continue qui se prolonge sur plusieurs centaines de mètres, illustrée par les réflexions géoradar continues inclinées de quelques degrés (entre 3° et 5°) vers l'ancienne mer de Champlain (le fleuve Saint-Laurent d'aujourd'hui à la Figure S.4b), est caractéristique d'un milieu sédimentaire littoral (Van Overmeeren 1998). Plus précisément, une butte observée dans les réflexions géoradar situées à la distance de 300 m sur le profil géoradar de la Figure S.6 laisse présumer la présence d'une épaisse crête de plage qui agit comme un cordon de sable et divise la section géoradar en deux régions hydrogeéologiques distinctes. La région qui se trouve en avant du cordon de sable (à gauche, vers le point A') et le cordon lui-même montrent en effet une bonne résolution et une grande profondeur de pénétration des ondes électromagnétiques, qui sont liées aux sables (indiquées par les types de sédiments déduits des profiles CPT113 et CPT26) ayant une très faible teneur en matériaux conducteurs comme l'argile. La grande profondeur de pénétration dans cette zone (la zone sableuse de la Figure S.6) permet également la reconnaissance d'une couche de till recouvrant le substratum rocheux. En effet, les hyperboles observées sous les sédiments littoraux sont des caractéristiques géoradar typiques liées aux blocs qui se trouvent dans le till. Une couche de till, dont la profondeur varie de l'affleurement jusqu'à 22 m, est généralement présente partout dans la région d'étude et sa faible perméabilité limite l'échange d'eau souterraine entre les sédiments littoraux et le substratum rocheux. Au-delà du cordon de sable (à droite, vers le point A), les sédiments sont plus fins et montrent des couches alternant entre le sable et le silt, comme le laissent présumer les sédiments déduits des profils CPT116, CPT24 et P17. Les réflexions géoradar dans cette zone sont également diffuses à quelques endroits en raison de la présence de sédiments ayant une grande conductivité électrique, comme le silt argileux, qui limitent habituellement la pénétration du signal géoradar. En conséquence, la couche de till ne peut être identifiée dans cette zone parce que des sédiments fins limite la pénétration du signal géoradar dans cette séquence littorale. Des autres levés

géoradar effectués dans la région d'étude (non montré ici), une diminution générale de la profondeur de pénétration est également observée en avançant vers le sud-est (à droite, en s'éloignant du point A), ce qui laisse présumer que la granulométrie des sédiments décroit en s'éloignant du rivage (vers les terres) en raison d'une réduction de l'énergie des vagues et des courants. Enfin, le niveau de l'eau souterraine dans l'aquifère se trouve à environ 1 m sous la surface du sol, comme l'indique les niveaux d'eaux mesurés dans les puits (CPT26, CPT24 et P17). Ce niveau d'eau ne peut cependant être déterminée à partir du profil géoradar de la Figure S.6 en raison de sa proximité avec la surface du sol.

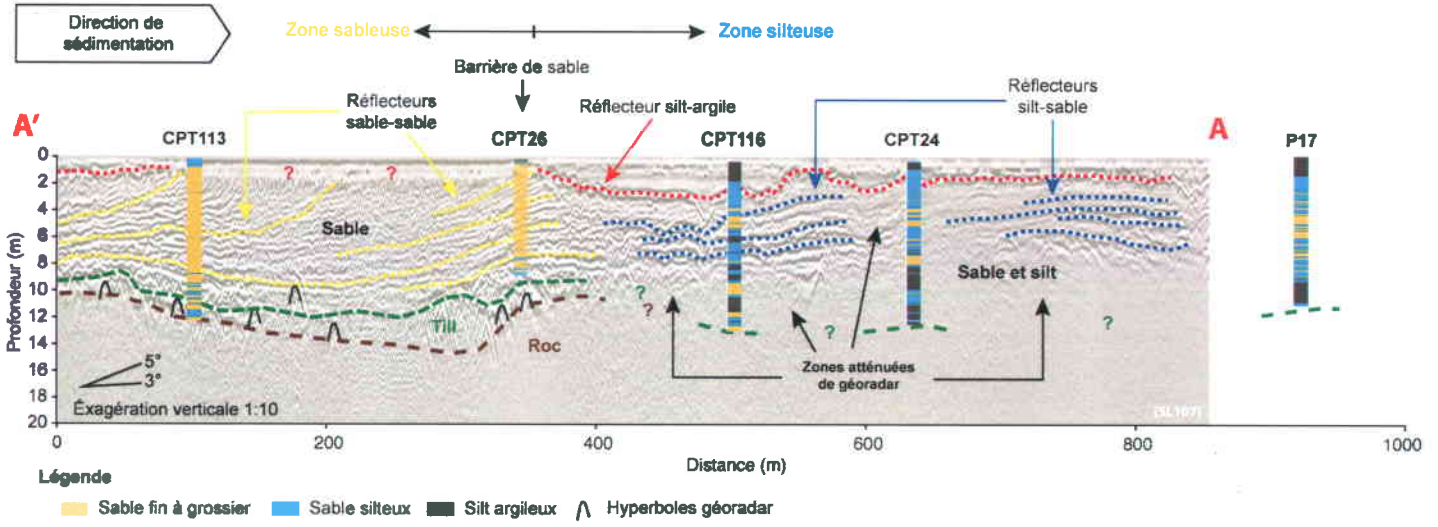


Figure S.6 Levé géoradar parallèle à la direction générale de sédimentation avec les profils de sédiments CPT113, CPT26, CPT116, CPT24 et P17 dérivés des essais au piézocône (*Cone Penetration Tests*, CPT) interprétés selon la classification proposée par Fellenius et Eslami (2000). La section géoradar et les sondages sont montrés à la Figure S.2c.

2.4 Sondages CPT/SMR

Selon le faciès géoradar reconnu précédemment, l'emplacement de 53 sondages CPT/SMR a été choisi (Figure S.4c et Tableau S.1) pour permettre une interprétation géologique approfondie des levés géoradar et pour acquérir les propriétés mécaniques et électriques des sédiments qui permettront d'établir des relations avec les mesures hydrauliques. Les sondages par enfoncement (CPT/SMR) ont été effectués à l'aide d'un système de sondage sur chenils Geotech 605-D muni d'un CPT comprenant un dispositif de mesure de la pression de l'eau interstitielle combiné à une sonde SMR (Figure S.7). Comme l'illustre la Figure S.8 (lignes rouges) pour l'emplacement du puits P17 (illustré à la Figure S.6), les sondages CPT/SMR permettent la mesure simultanée de deux propriétés mécaniques et de deux propriétés électriques des sédiments à des résolutions verticales variant entre 3 et 17 cm (Tableau S.1). Un cône de 15 cm^2 doté d'un bout conique à 60° a été utilisé conformément aux normes ASTM D3441 (ASTM 2000). Pour les sondages, le cône est enfoncé verticalement dans le sol à un rythme constant de 2 cm/s, bien que ce rythme doive être ralenti lorsque le pénétromètre atteint des couches compactes. À l'intérieur de la sonde CPT, deux jauges de force mesurent de façon indépendante la résistance à l'enfoncement (T) de la sonde et la friction latérale (S) le long du cône (Lunne *et al.* 1997). La résolution verticale de T et de S est de 4 cm et 17 cm, respectivement. Une jauge de pression est également utilisée dans le cône pour mesurer la pression d'eau interstitielle lorsque la sonde est enfoncée dans le sol. La pression interstitielle est un indicateur de présence d'argile et a été utilisée pour corriger les mesures de T. La sonde SMR est quant-à-elle composée de quatre électrodes branchées directement derrière le pénétromètre (Shinn *et al.* 1998). Les deux bagues intérieures espacées de 3 cm servent à mesurer la capacitance du sol. La sonde de capacitance utilise une fréquence de 100 MHz, réduisant ainsi les effets du type de sédiments sur la mesure. Lorsqu'elle pénètre le sol, la sonde de capacitance mesure les variations du signal à haute fréquence qui sont liées à la teneur en eau du sol et ces valeurs peuvent être convertis empiriquement en valeur de constante diélectrique (D). La mesure de la résistivité électrique (R) emploie les deux bagues extérieures de la sonde SMR séparées de 9 cm pour appliquer un courant et mesurer la chute de tension (configuration pôle-pôle). La sonde fonctionne à une fréquence de 1 000 Hz afin de réduire les effets de polarisation des sédiments.

Tel que mentionné précédemment, la Figure S.6 montre les types de sédiments déduits des mesures CPT pour illustrer de façon plus détaillée la stratigraphie généralement observée dans la région d'étude de St-Lambert. Les profils lithologiques sont obtenus par la conversion des propriétés mécaniques (T et S) des sédiments en utilisant la classification proposée par Fellenius et Eslami (2000). L'interprétation combinée

des sondages CPT/SMR et des levés géoradar a permis de reconnaitre deux contextes hydrogéologiques distincts dans la région d'étude, comme indiqué par la délimitation des zones sableuse et silteuse à la Figure S.6. La limite approximative entre les zones sableuse et silteuse pour toute la région d'étude est indiquée à la Figure S.4c (ligne blanche pointillée). La Figure S.4c montre également que les sondages CPT/SMR sont répartis sur l'ensemble de la région d'étude de manière à couvrir l'étendue complète des propriétés mécaniques et électriques. Tel que suggéré à la Figure S.6, les emplacements des sondages ont aussi été choisi pour caractériser les changements progressifs de texture des sédiments afin de mieux définir la continuité latérale de l'hétérogénéité de l'aquifère.

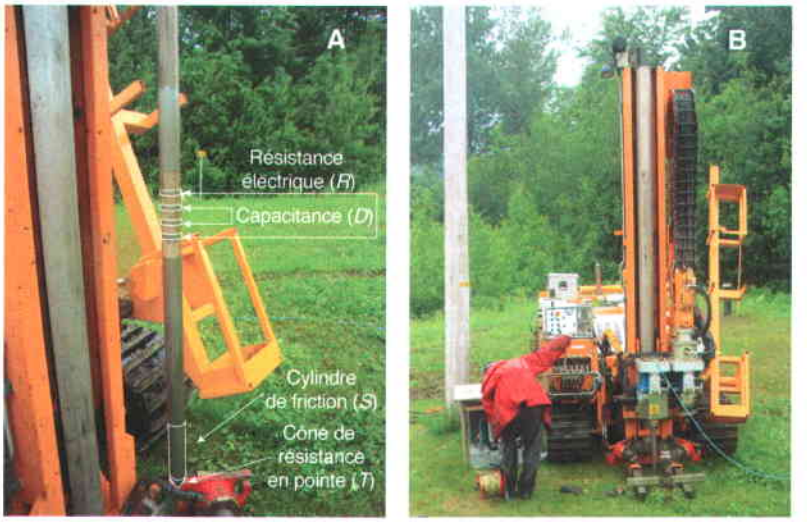


Figure S.7 Photos de l'équipement utilisé pour la réalisation des sondages CPT/SMR par enfoncement: (a) pénétromètre (paramètres *S* et *T*) et sonde SMR (paramètres *D* et *R*); et (b) système d'enfoncement motorisé sur chenilles (Geotech 605D).

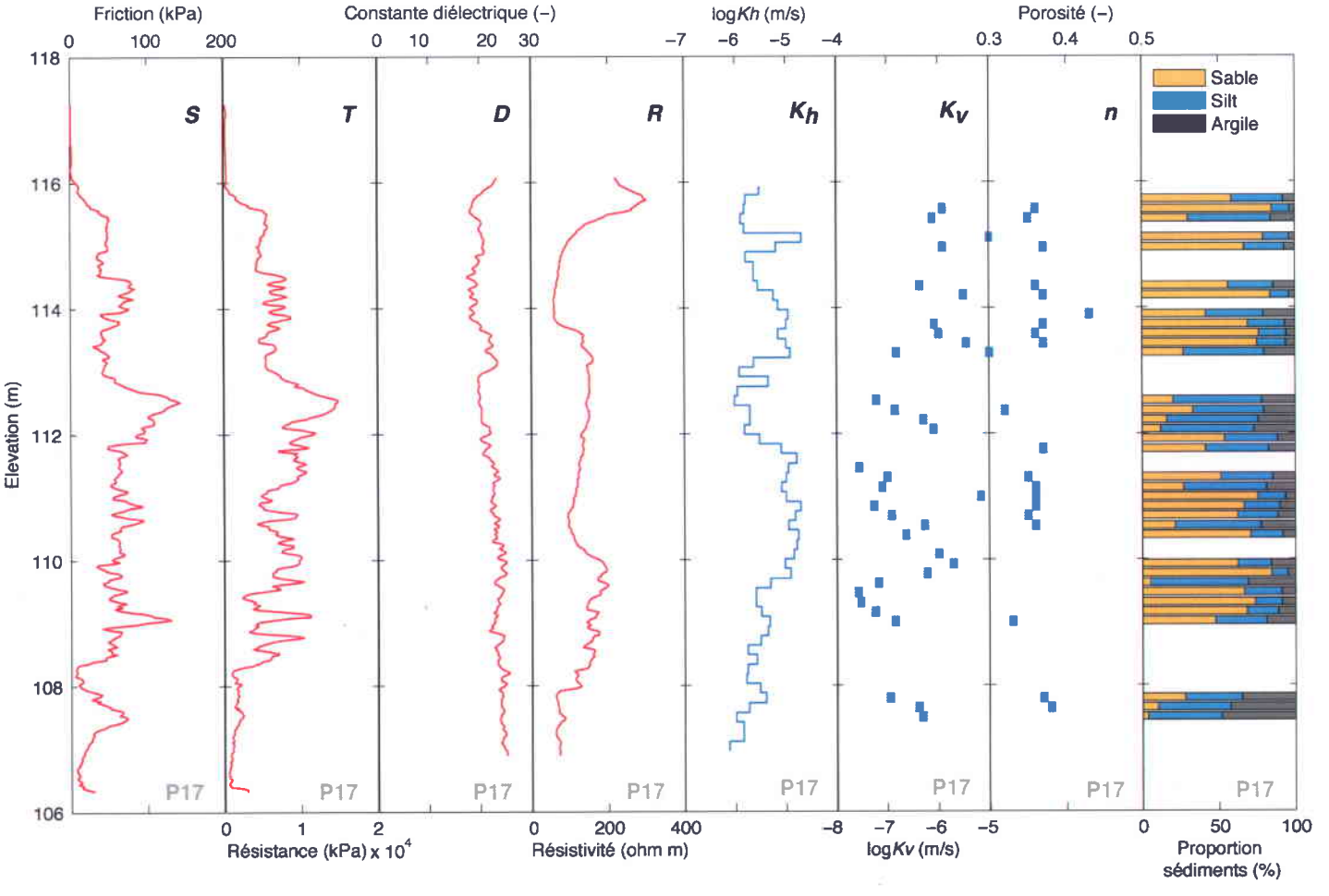


Figure S.8 Exemple de profils CPT/SMR (lignes rouge), de propriétés hydrauliques (lignes bleues) et lithologique obtenus au puits P17. Les paramètres CPT/SMR sont la friction (*S*), la résistance en pointe (*T*), la constante diélectrique (*D*) et la résistivité électrique (*R*). Les conductivités hydrauliques horizontale (K_h) et verticale (K_v) ont été obtenues respectivement des essais de perméabilité (*slug tests*) au puits P17 et des essais au perméamètre sur des échantillons de sédiments. La porosité (*n*) et la granulométrie ont été obtenues sur les mêmes échantillons de sédiments utilisés avec le perméamètre. Les proportions de sable, silt et argile sont basées sur les courbes granulométriques et les classes de Wentworth (1922). La localisation du puits P17 est montrée à la Figure S.2c.

2.5 Installation de puits par enfoncement

En s'appuyant sur les mesures CPT/SMR obtenues en temps réel durant les opérations de sondage, 25 des 53 emplacements de sondage ont été choisis pour installer des puits d'observation qui seront utilisés, entre autres, pour effectuer des essais hydrauliques (Figure S.4c et Tableau S.1). Chacun des puits a été installé dans le même trou créé par le sondages afin d'obtenir des données hydrauliques et géophysiques co-localisées, réduisant ainsi l'incertitude liées à la disparité des intervalles de mesures. Chaque puits d'observation a également été installé immédiatement à la suite de chaque sondage avec le même système d'enfoncement (Geotech 605D) afin de réduire au minimum le temps de mobilisation et de démobilisation de l'équipement.

L'installation de puits par enfoncement a suivi la procédure d'installation par crépine protégée proposée par ASTM (2004). Cette procédure d'installation n'exige pas l'utilisation de sable filtrant et la crépine du puits est en contact direct avec les sédiments. Ce type d'installation convient bien aux essais hydrauliques effectués sur de petits intervalles parce qu'il réduit la formation de court-circuit hydrauliques lors des tests. La procédure d'installation comprend, en premier lieu, le martelage d'un tubage métallique de 7.72 cm (3 po) de diamètre extérieur doté d'une pointe non récupérable sur sa base (Figure S.9). Cette pointe est utilisée pour éviter l'entrée de sédiments dans le tubage métallique lors de l'enfoncement pour la mise-en-place. Lorsque le tubage de métal atteint la profondeur désirée, un puits d'observation de 5.1 cm (2 po) de diamètre intérieur (tubes pleins et crépine de PVC) est ensuite inséré à l'intérieur du tubage de métal, bien appuyé sur la pointe non récupérable. Enfin, le tubage de métal est retiré jusqu'à la surface pour exposer le puits d'observation à la formation, tandis que la pointe non récupérable reste au fond du sondage. Notons que les diamètres du tubage de métal et du puits en PVC sont ajustés pour assurer un effondrement complet des sédiments sur le puits d'observation après le retrait du tubage métallique.

Pour la présente étude, des puits d'observation dotés de crépines couvrant la totalité de la zone saturée ont été installés pour permettre la mesure de profils continus des propriétés hydrauliques lors des essais hydrauliques. La longueur des crépines utilisés dans tous les puits varie de 4.6 à 16.8 m. Chacun des puits par enfoncement a également été développé de façon exhaustive après leur installation pour nettoyer la crépine et les sédiments environnants affecté lors des opérations d'installation (Butler 1998; Henebry et Robbins 2000). Chaque puits a donc été développé par pompage-pistonnage à des intervalles de 0.5 m à

l'aide d'une pompe inertielle de type Waterra dotée d'une bague de développement et ce jusqu'à ce que plus aucune turbidité ne soit observée dans l'eau pompée (Paradis *et al.* 2011).

Figure S.9 Photos d'une installation de puits par enfoncement: (a) tubage de PVC (blanc) inséré à l'intérieur du tubage de métal utilisé pour l'enfoncement à travers les sédiments; (b) pointe perdue placée à la base du tubage de métal pour éviter l'entrée des sédiments dans ce tubage; et (c) tubage de PVC installé avant le scellement du puits en surface (le tubage est plein dans sa partie supérieure et crépiné sur la majeure partie saturée en eau de l'aquifère). Voir aussi Paradis et al. (2011) pour les détails de l'installation des puits par enfoncement.

2.6 Échantillonnage des sédiments

En général, après chaque sondage CPT/SMR, les sédiments composant l'aquifère ont été échantillonnés à proximité des sondages (1 à 2 mètres). L'échantillonnage a été effectué à l'aide d'un échantillonneur à piston (échantillonneur Macro-Core de Geotech) permettant la récupération de sédiments à l'intérieur d'une gaine transparente en PETG de 38 mm de diamètre et de 1.52 m de longueur (Tableau S.1 et Figure S.10a). Cette méthode a principalement été choisie parce qu'elle préserve la stratification naturelle des sédiments, un élément fondamental pour l'analyse des structures sédimentaires et pour l'estimation de K_v en laboratoire avec des essais au perméamètre. Les échantillons ont été prélevés avec le même système par enfoncement (Geotech 605D) que celui utilisé pour les sondages CPT/SMR et les installations de puits. Ceci permet donc la réalisation d'une série d'opérations de terrain au même endroit en une seule mobilisation du système par enfoncement. Les intervalles d'échantillonnage ont été choisis en fonction

des réponses CPT/SMR obtenus en cours de sondage. Plus de 75 carottes ont été prélevés de façon sélective avec une récupération des sédiments estimée à 83 %.

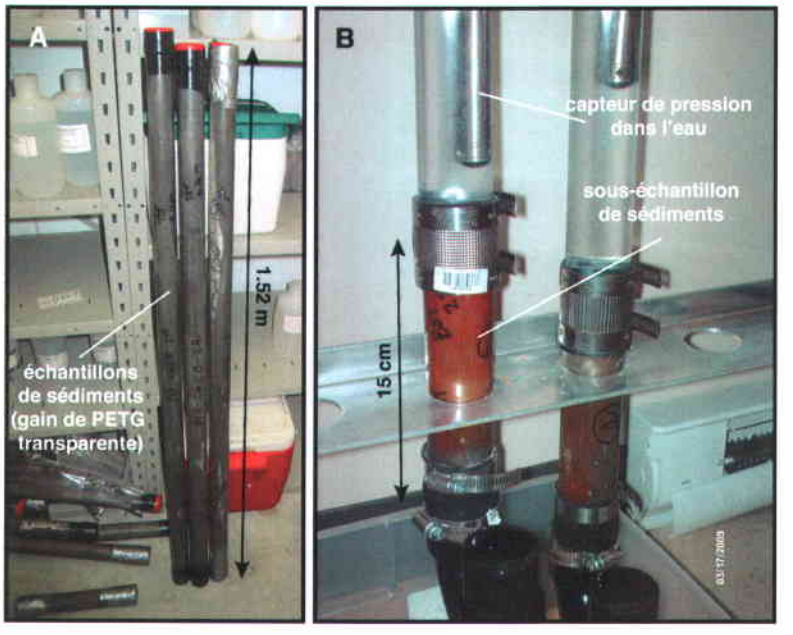


Figure S.10 (a) Carottes de sédiments de 1.5 m récupérées avec l'échantillonneur de sol à piston; (b) sous-échantillons de 15 cm placés dans des perméamètres à charge variable pour mesurer la conductivité hydraulique verticale.

2.7 Essais pour mesurer la conductivité hydraulique horizontale

2.7.1 Tests de perméabilité à choc hydraulique multi-niveaux

Les valeurs de K_h, telles que celles présentées à la Figure S.8, utilisées pour établir les relations avec les données CPT/SMR, ont été obtenues au moyen de tests de perméabilité (*slug tests*) réalisés à plusieurs niveaux dans les puits installés par enfoncement. Comme les puits sont dotés de crépines couvrant la totalité de l'aquifère, des profils continus de K_h peuvent ainsi être obtenus. Comme l'illustre la Figure S.11, une série de tests de perméabilité multi-niveaux nécessite l'utilisation d'obturateurs pour isoler successivement chacun des intervalles testés (Sellwood *et al.* 2005; Zemansky et McElwee 2005; Ross et McElwee 2007). Les chocs hydrauliques utilisés pour initier les tests ont été produits à l'aide d'un système pneumatique (Levy et Pannell, 1991), tel qu'illustré à la Figure S.10b, et les réponses hydrauliques correspondantes ont été interprétées en utilisant la méthode de Bouwer et Rice (1976). Une description plus détaillée du montage expérimental des essais, de leur réalisation et de leur analyse est fournie par Paradis *et al.* (2011). Les tests ont été effectués sur des intervalles de 15 cm afin de réduire

l'incertitude liée à la disparité des échelles entre les mesures hydrauliques et CPT/SMR. Aussi, des petits intervalles de mesures permettent, pour des aquifères relativement homogènes, d'approcher l'hypothèse d'isotropie pour K ($K_h \simeq K_v$), ce qui permet de simplifier la caractérisation de K. Un total de 280 intervalles choisis selon les mesures CPT/SMR obtenues précédemment a fait l'objet de tests de perméabilité (Tableau S.1).

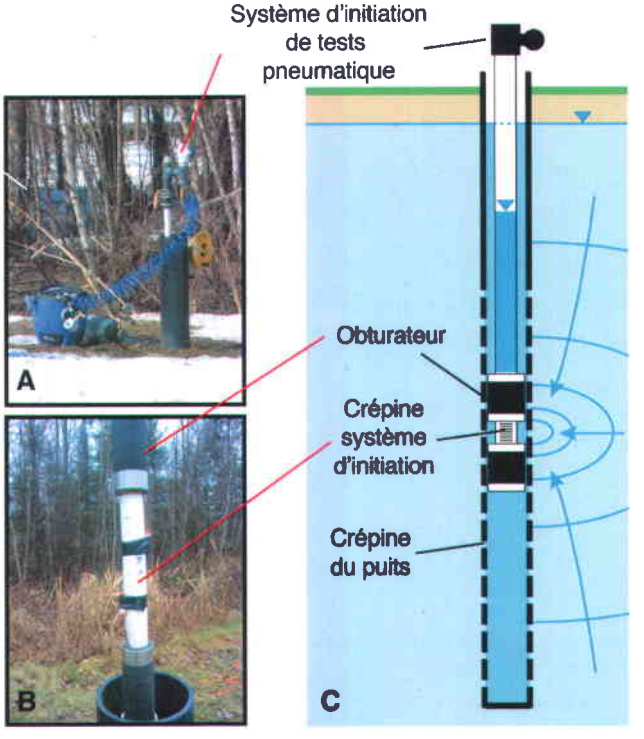


Figure S.11 (a) Système d'initiation pneumatique pour les tests de perméabilité (*slug tests*); (b) système à doubles obturateurs avec crépine pour isoler des intervalles spécifiques dans le puits d'observation; et (c) schéma d'un essai de perméabilité multi-niveaux.

2.7.2 Débitmètre de puits

Dans le cadre de cette thèse, Paradis *et al.* (2011) ont adapté l'utilisation du débitmètre de puits pour les formations granulaires afin d'améliorer l'efficacité de la caractérisation de K_h. Comme l'illustre la Figure S.12, un test de perméabilité au moyen d'un débitmètre comprend la mesure de l'apport d'eau souterrain à un puits à différentes profondeurs lorsque ce puits est pompé en régime permanent. Ceci permet d'obtenir un profil vertical de K_h selon l'hypothèse que K est proportionnel à l'apport d'eau (Hess 1986). Bien que le débitmètre soit couramment utilisé dans des puits ouverts pour analyser les formations

rocheuses consolidées, relativement peu d'études ont utilisé le débitmètre dans les formations granulaires à cause des problèmes reliés à la réalisation des essais dans des puits aménagés avec sable filtrant.

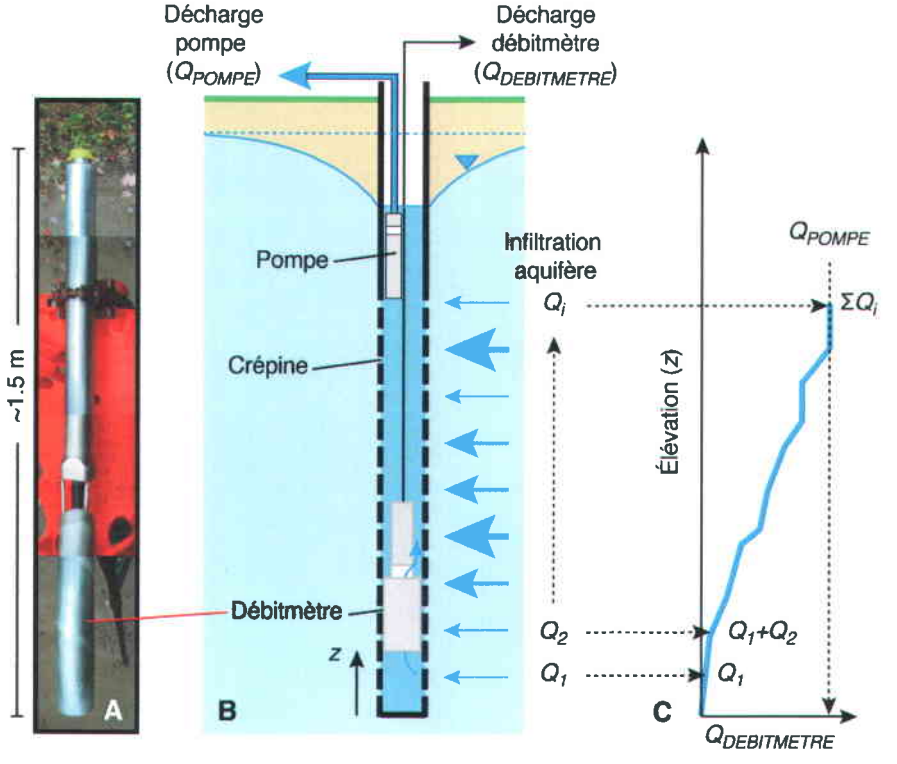


Figure S.12 (a) Photo d'un débitmètre de puits électromagnétique. Le centre du cylindre inférieur est ouvert à sa base et permet l'écoulement vertical de l'eau. Ce cylindre contient la bobine à induction pour la mesure du champ électromagnétique proportionnel à la vitesse d'écoulement de l'eau (Paradis *et al.* 2011). Le cylindre supérieur contient le système d'acquisition. (b) Schéma d'un essai avec débitmètre de puits. Le débitmètre est déplacé par paliers successifs pour l'obtention de la courbe cumulative du débit pompé dans le puits dans l'aquifère présenté en (c).

Dans leur étude, Paradis *et al.* (2011) ont décrit l'application du débitmètre dans des puits par enfoncement pour mesurer des profils de K_h et délimiter les HF dans les aquifères granulaires hétérogènes. La principale adaptation du débitmètre pour les aquifères granulaires est l'utilisation de puits par enfoncement sans sable filtrant et dotés de longues crépines, selon l'installation décrite précédemment, pour éviter les courts-circuits hydrauliques qui biaisent habituellement les tests de débitmètre dans les puits conventionnels (avec sable filtrant). L'effet de l'installation des puits par enfoncement sur les mesures de K_h dans les dépôts granulaires a tout d'abord été évalué en s'appuyant sur certaines études géotechniques antérieures indiquant que de telles installations perturbent minimalement les propriétés mécaniques des sédiments entourant la crépine. Par la suite, des profils de K_h de haute résolution (15 cm) obtenus dans des puits par enfoncement avec un débitmètre et au moyen d'essais de

perméabilité multi-niveaux ont été comparés pour des intervalles identiques. Pour 119 intervalles testés dans cinq puits différents, la différence de valeurs de logK_h obtenue à partir des deux méthodes est systématiquement inférieure à 10 % (Figure S.13). De plus, tel que discuté plus loin, un profil de K_h avec le débitmètre est obtenu en une fraction du temps nécessaire pour obtenir le même profil avec des tests de perméabilité multi-niveaux. Enfin, une approche graphique permettant l'interprétation des profils de débitmètre pour délimiter des zones homogènes de K_h est aussi proposée, fournissant ainsi une méthode en vertu de laquelle l'échelle et les contrastes de K dans les aquifères granulaires hétérogènes peuvent être représentés, ce qui facilite la représentation spatiale de l'hétérogénéité des aquifères.

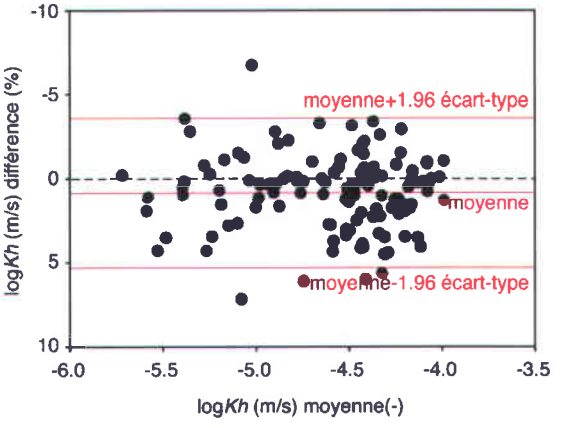


Figure S.13 Comparaison des mesures de conductivité hydraulique horizontale (K_h) obtenues avec le débitmètre et les essais de perméabilité (*slug tests*) multi-niveaux pour cinq puits à une résolution verticale de 15 cm. La différence dans les valeurs de logK_h obtenue par les deux méthodes est montrée par rapport à la moyenne des deux types de mesures à chacun des intervalles. Les intervalles de confiance de 95% (moyenne+1.96 écart-type) montrés supposent une distribution normale des différences.

2.8 Essais pour mesurer la conductivité hydraulique verticale

2.8.1 Perméamètre de laboratoire

Pour vérifier l'hypothèse d'isotropie de K à des intervalles de 15 cm, 59 échantillons de sédiments de 15 cm de longueur ont été choisis dans le but de mesurer K_v au laboratoire à l'aide d'essais avec un perméamètre (Figure S.7 et Tableau S.1). Pour réduire au minimum la perturbation des structures sédimentaires des échantillons de sol, un perméamètre de laboratoire à charge variable automatisé conçu pour loger directement la gaine de PETG de 38 mm de diamètre comprenant les échantillons de sédiments a été développé pour déterminer K_v, tel qu'illustré à la Figure S.10b. Chaque échantillon original de sol de 1.52 m de longueur a été subdivisé en sous-échantillons de 15 cm de longueur selon les intervalles soumis à des tests de perméabilité multi-niveaux. Des duplicata effectués sur plusieurs sous-

échantillons ont indiqué un niveau élevé de reproductibilité des valeurs de K_v et toutes ces valeurs ont été corrigées pour une température des eaux souterraines de 8 °C, représentative des conditions observées sur le terrain. Paradis et Lefebvre (2013) fournissent une description plus détaillée du montage expérimental et de la procédure d'analyse des essais. Comme le montre la Figure S.8, la comparaison des mesures de K_h avec celles de K_v au même profondeurs révèle que l'hypothèse d'isotropie à des intervalles de 15 cm ne tient habituellement pas, avec des valeurs de K_v jusqu'à deux ordres de grandeur plus faibles que K_h. Ceci indique que K_h et K_v devraient toutes deux être estimées pour chacun des intervalles de 15 cm afin de caractériser K pour la région d'étude et prédire adéquatement l'écoulement de l'eau souterraine et le transport de masse. Bien que les essais avec un perméamètre puissent être fiables pour estimer K_v dans des conditions favorables, les difficultés liées à la collecte et à la manipulation des échantillons de sédiments peuvent toutefois limiter la capacité à obtenir des estimations fiables de K_v dans certaines conditions plus difficiles, comme avec les sables grossiers (Stienstra et van Deen 1994). De plus, les essais effectués avec un perméamètre peuvent également prendre beaucoup de temps lorsque plusieurs mesures sont nécessaires (Klute et Dirksen 1986) et sont donc généralement restreints à un petit nombre d'échantillons.

2.8.2 Tests d'interférence verticale à choc hydraulique

Diverses méthodes de terrain permettent la mesure de K_h, mais K_v est rarement mesurée, faute de disponibilité de tests pratiques applicables sur le terrain. Dans le cadre de l'effort de caractérisation relié à cette thèse, Paradis et Lefebvre (2013) proposent des tests d'interférence verticale à choc hydraulique pour mesurer efficacement K_v sur le terrain. Ces tests sont une adaptation des tests d'interférence entre puits pour la mesure à un seul puits. Comme l'illustre la Figure S.14, un test d'interférence verticale est effectué dans un puits unique, entre un intervalle d'observation et un intervalle émetteur, isolé verticalement au moyen d'un montage composé de trois obturateurs pneumatiques. Un choc hydraulique est produit dans l'intervalle émetteur et les variations de charges hydrauliques qui en découlent sont enregistrées dans l'intervalle émetteur et d'observation.

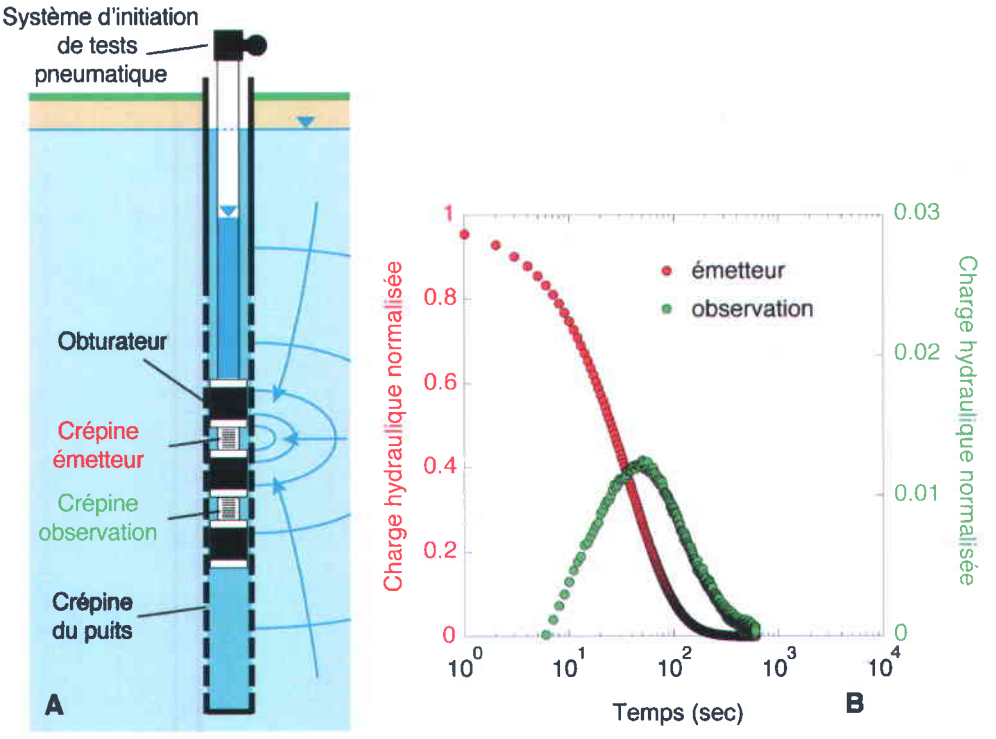


Figure S.14 (a) Schéma d'un essai d'interférence vertical à choc hydraulique. (b) Réponses hydrauliques types pour l'intervalle émetteur et l'intervalle d'observation (noter les échelles différentes pour ces deux intervalles).

Dans une étude sur le terrain visant à valider le principe du test proposé, 12 tests d'interférence verticale ont été effectués séquentiellement dans un puits aménagé dans un aquifère modérément hétérogène et très anisotrope composé de silts et de sables littoraux. Une méthode par enfoncement a été utilisée pour installer le puits sans sable filtrant et avec une longue crépine afin de permettre la mesure des propriétés hydrauliques *in situ* des sédiments et réduire au minimum les interférences entre les éléments de construction du puits et les essais hydrauliques. Les charges hydrauliques mesurées en fonction du temps dans l'intervalle émetteur et d'observation des douze essais ont été simultanément inversées numériquement avec l'algorithme de Bohling et Butler (2001) pour reconstruire les profils hétérogènes de K_h, de l'anisotropie de la conductivité hydraulique (K_v/K_h) et de l'emmagasinement spécifique (S_s). Les charges simulées résultant de l'inversion ont été comparées aux charges observées et les propriétés hydrauliques du modèle comparées aux valeurs de K_h et K_v obtenues par des tests de perméabilité multi-niveaux et par des essais en laboratoire avec perméamètre. Comme le montre la Figure S.15, le profil de K_v obtenu au moyen des tests d'interférence verticale suit un patron semblable avec la profondeur à celui obtenu au moyen d'un perméamètre sur des échantillons de sédiments recueillis sur les mêmes intervalles que les tests d'interférence. Cela démontre que les tests d'interférence verticale à choc hydraulique

peuvent fournir une base adéquate pour la caractérisation de K_h, tel que généralement obtenus des méthodes conventionnelles, et aussi de S_s et de K_v, ce qui est difficilement possible autrement.

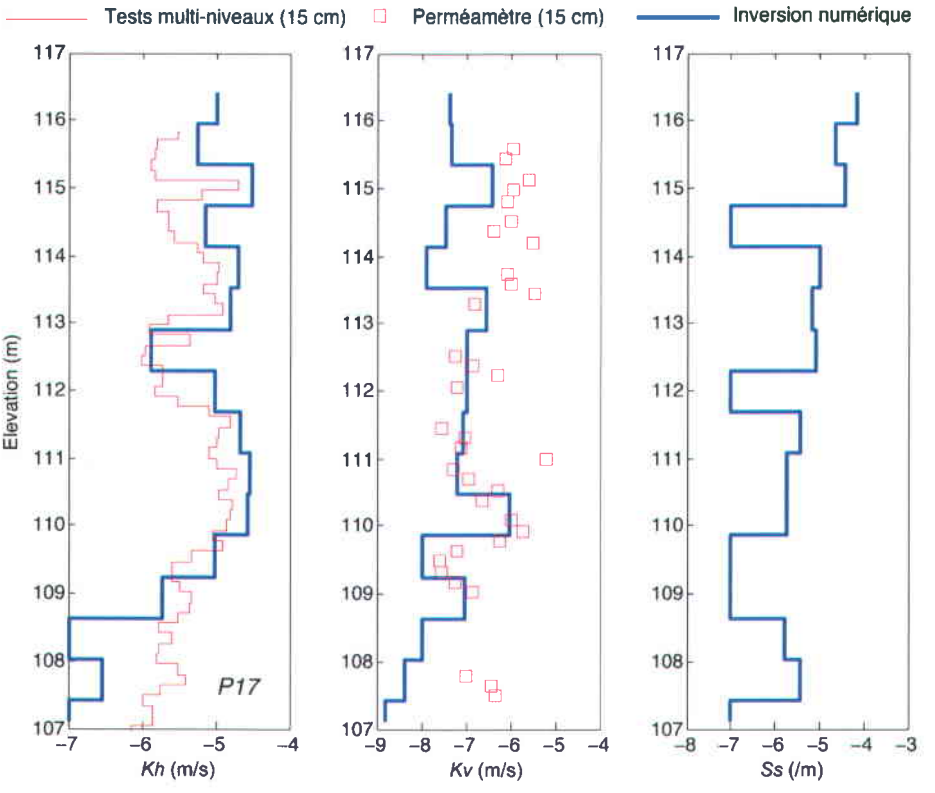


Figure S.15 Profils verticaux mesurés de K_h (essais de perméabilité multi-niveaux) et de K_v (perméamètre) comparés aux profils prédit par l'inversion numérique de plusieurs essais d'interférence vertical à choc hydraulique.

2.8.3 Tomographie à choc hydraulique

Un deuxième type d'essai pour estimer K_v, la tomographique à choc hydraulique, a également été élaboré par Paradis *et al.* (2014a; 2014b). Comme l'illustre la Figure S.16, la tomographique à choc hydraulique est une généralisation sur plusieurs puits des tests d'interférence verticale qui comprend la réalisation d'une série de tests de perméabilité à choc hydraulique entre obturateurs dans un puits émetteur avec l'enregistrement simultané des variations de charges hydrauliques dans l'intervalle émetteur et dans des intervalles d'observation aménagés dans un puits adjacent. La justification derrière la tomographie hydraulique est de déterminer les hétérogénéités locales entre les puits qui peuvent passer inaperçues à la suite d'essais sur un puits unique et de mieux représenter la variabilité spatiale des propriétés hydrauliques.

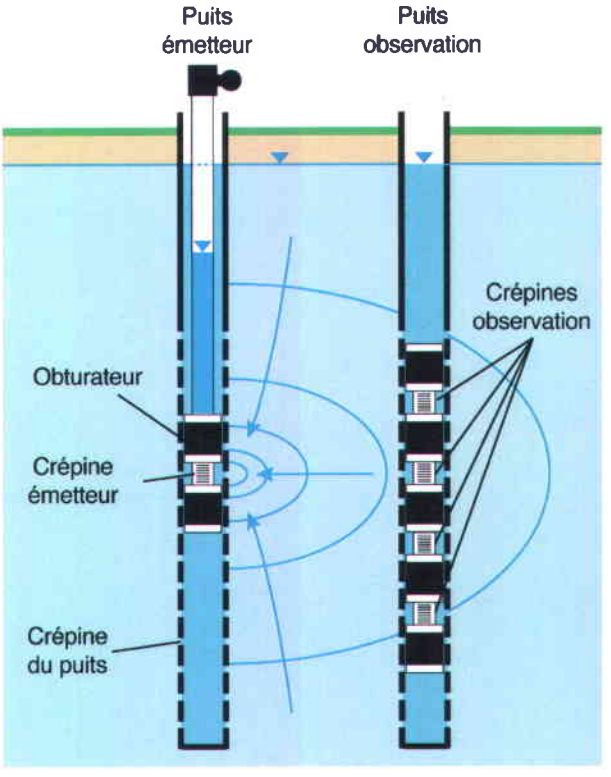


Figure S.16 Schéma d'un essai de tomographie à choc hydraulique entre deux puits. Les réponses hydrauliques à l'intervalle émetteur et dans les intervalles d'observations sont similaires aux courbes de la Figure S.14b.

L'étude fondamentale de Paradis *et al.* (2014a) visait à évaluer de manière quantitative le contenu en information des variations de charges hydrauliques associées à une analyse transitoire d'une expérimentation de tomographie à choc hydraulique et donc d'évaluer la capacité d'une telle expérimentation à résoudre l'hétérogénéité des propriétés hydrauliques dans un plan vertical entre deux puits. Une analyse de sensibilité et de résolution a donc été effectuée à l'aide de données synthétiques reflétant les caractéristiques de l'aquifère littoral de St-Lambert. Les résultats indiquent que la perturbation des charges induite par un test à choc hydraulique produit des patrons de sensibilité transitoire distincts dans les intervalles émetteur et d'observation, permettant ainsi la résolution indépendante de K_h, K_v/K_h et S_s dans le plan d'investigation et ne présente qu'une faible interférence des paramètres au-delà de ce plan (Figure S.17). Ceci est une importante généralisation par rapport aux approches de tomographie hydraulique actuellement proposées, puisqu'à ce jour la tomographie hydraulique n'avait été proposée que pour estimer K_h.

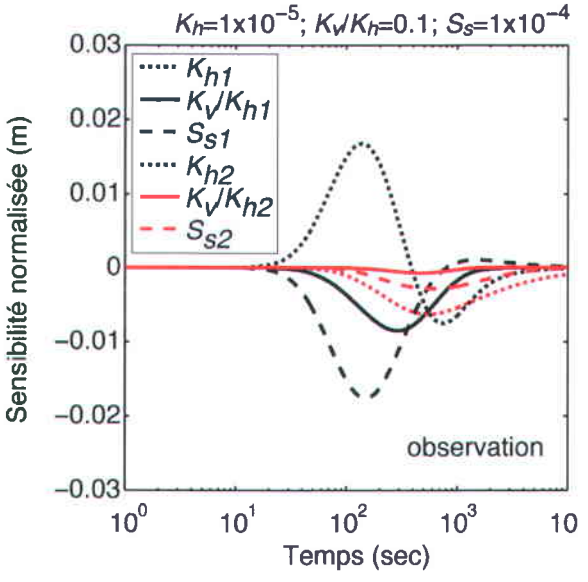


Figure S.17 Sensibilités normalisées de la charge hydraulique par rapport aux propriétés hydrauliques pour un intervalle d'observation dans un aquifère homogène et anisotrope suite à une perturbation hydraulique produite dans un intervalle émetteur. La configuration expérimentale est similaire à la Figure S.16 avec une distance entre le puits émetteur et d'observation d'approximativement 8 m. Les indices 1 et 2 réfèrent respectivement aux paramètres à l'intérieur et à l'extérieur de la région d'investigation comprise entre les puits émetteur et d'observation.

Cependant, la résolution latérale de chacun des paramètres hydrauliques est principalement centrée à proximité des puits émetteur et d'observation (Figure S.18). Bien que la résolution verticale puisse être augmentée en utilisant de plus petits intervalles émetteurs, la résolution latérale ne peut être augmentée en utilisant seulement les charges hydrauliques. En effet, le caractère dispersif de l'équation décrivant la propagation de la charge hydraulique dans l'aquifère et la géométrie de l'expérimentation sont des contraintes importantes pour la résolution latérale, même en utilisant un ensemble dense d'intervalles émetteurs et d'observation. Ces contraintes ne sont pas dues à l'amplitude de la perturbation hydraulique qui peut être générée, mais plutôt à la corrélation élevée qui existe entre les paramètres hydrauliques dans le plan d'investigation. Ainsi, une analyse transitoire de la tomographie à choc hydraulique permet uniquement la caractérisation des propriétés hydrauliques hétérogènes à une résolution relativement grossière. Pour une définition à plus haute résolution, de l'information supplémentaire doit être intégrée dans le processus d'inversion numérique.

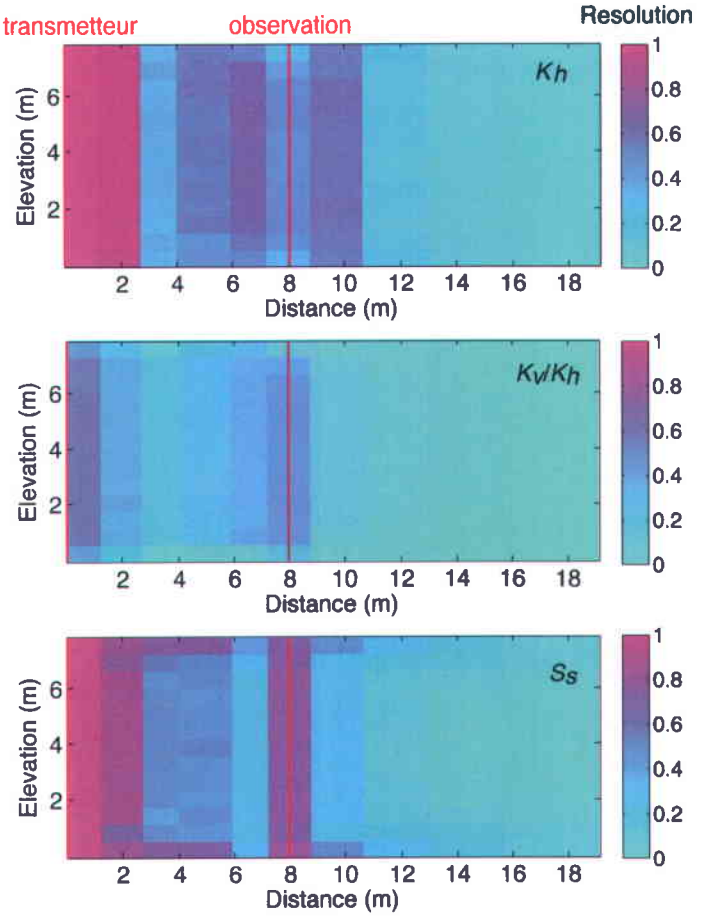


Figure S.18 Éléments de la diagonale de la matrice de résolution associés avec une analyse transitoire des réponses hydrauliques pour les intervalles émetteurs et d'observations de 13 essais de perméabilité (*slug tests*). L'analyse est basée sur la troncation d'une décomposition en valeurs singulières avec une erreur relative dans l'estimation des paramètres de 1%. Le niveau de bruit utilisé est de $1x10^{-5}$ m et correspond au niveau calculé avec l'instrumentation utilisée sur le terrain.

L'objectif de l'étude de Paradis *et al.* (2014b) consistait quant-à-elle à évaluer le potentiel de la tomographie à choc hydraulique, comme l'ont démontré Paradis *et al.* (2014a), dans des conditions réelles de terrain telle que celle rencontrées pour l'aquifère littoral hétérogène et anisotrope de St-Lambert. L'algorithme de Bohling et Butler (2001) a été utilisé pour l'inversion numérique de la tomographie effectué entre deux puits comprenant 12 tests à choc hydraulique pour un total de 58 intervalles d'enregistrement des variations de charges hydrauliques. L'exactitude des tomogrammes de K_h, K_v/K_h et S_s a été vérifiée au moyen de la simulation d'essais hydrauliques multi-échelles, non utilisés pour l'inversion, et avec des profils de K_h et de K_v provenant d'autres essais de terrain et de laboratoire. Comme l'illustre la Figure S.19, les estimations de K_h et de K_v obtenus avec la tomographie se comparent bien aux autres mesures. De plus, les différences entre les charges observées et prédites à partir d'essais multi-échelles indépendants montrent peu d'erreurs dans les estimations des paramètres hydrauliques.

Cette étude a donc démontré la capacité de la tomographie à choc hydraulique à définir des champs hétérogènes de K_h, K_v/K_h et S_s entre puits pour des conditions de terrain hétérogènes et anisotropes. Cette approche élargit donc les possibilités des méthodes hydrauliques en fournissant des renseignements à propos de l'hétérogénéité de K_h, habituellement déterminée par les méthodes conventionnelles, mais aussi à propos de l'hétérogénéité de K_v/K_h et S_s, ce qui est rarement possible avec les méthodes conventionnelles.

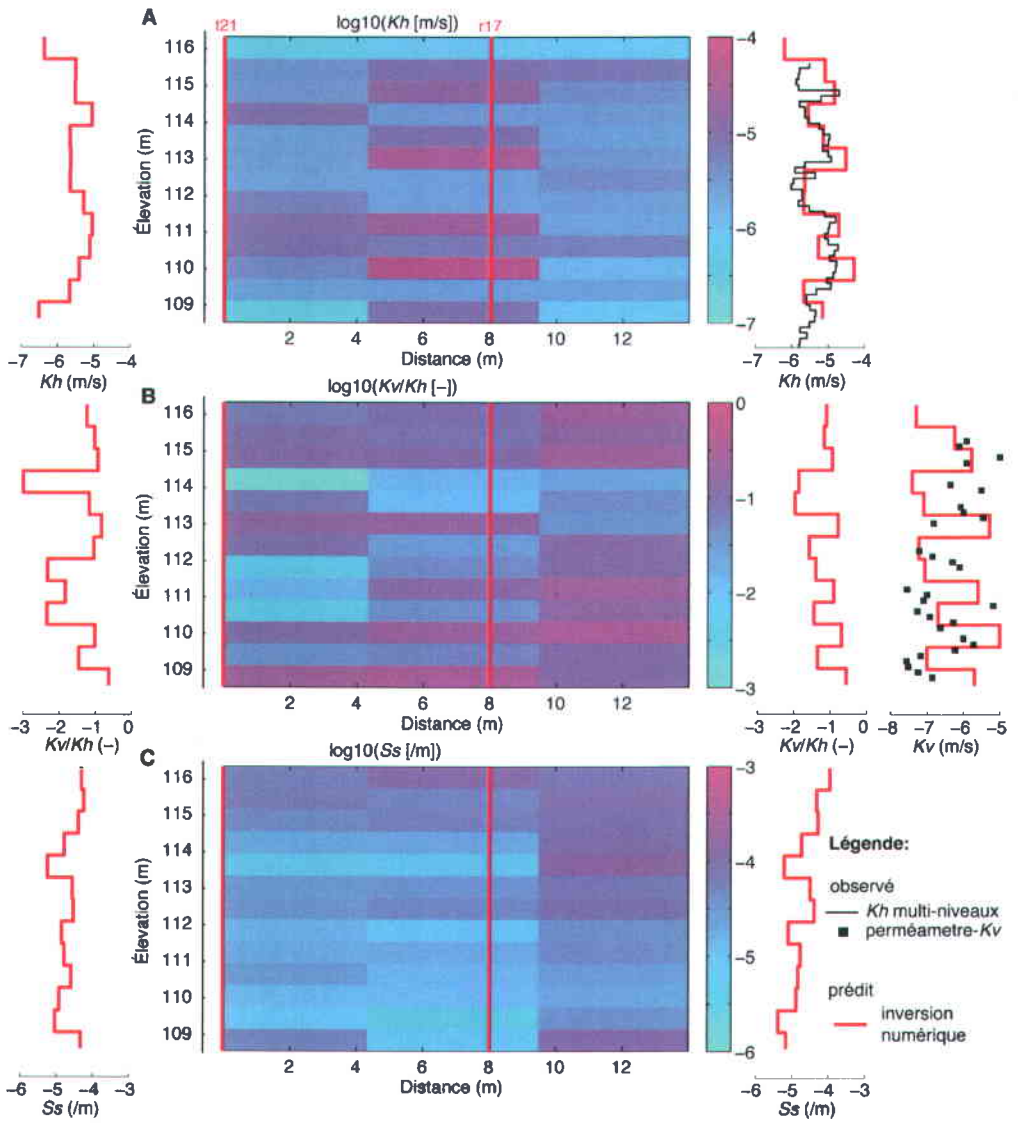


Figure S.19 Tomogrammes des paramètres hydrauliques résultant de l'inversion de tomographique t21r17 : (a) conductivité hydraulique horizontale (K_h); (b) anisotropie de la conductivité hydraulique (K_v/K_h); et (c) emmagasinement spécifique (S_s). Cette figure montre aussi les profils des propriétés hydrauliques prédit par l'inversion pour les puits P21 et P17 ainsi que les profils de K_h et K_v obtenus respectivement des essais de perméabilité (*slug tests*) multi-niveaux et des tests avec perméamètre sur des échantillons de sédiments.

2.9 Mesures de porosité et analyses granulométriques en laboratoire

À la suite des mesures de K_v en laboratoire avec un perméamètre sur des carottes de sédiments, des sous-échantillons ont été récupérés et saturés afin de mesurer la porosité totale (n) des sédiments (Figure S.8). La n a été estimée en soustrayant les poids sec et humide de chaque sous-échantillon de 15 cm (Tableau S.1). Pour obtenir le poids sec, les sous-échantillons ont été séchés dans un four à 50 °C pendant plus de 7 jours. Le nombre total de sous-échantillons dont n a fait l'objet de mesures a été de 43. Les sous-échantillons ont également été soumis (Figure S.8 et Tableau S.1) à une description géologique et à l'analyse granulométrique par diffusion de la lumière laser (Vitton et Sadler 1997) afin de fournir des renseignements géologiques complémentaires aux mesures hydrauliques et géophysiques.

3 Intégration des données

3.1 Aperçu de la méthode d'intégration des données

La méthodologie générale adoptée dans cette étude pour l'intégration des données suit le schéma illustré à la Figure S.20. Les travaux d'acquisition de données présenté à la Figure S.3 mène à la collecte d'un ensemble de données représentatif des conditions hydrogéologiques du site à l'étude (voir l'encadré intitulé « Données d'entrainement » à la Figure S.20). Cet ensemble de données comprend principalement des mesures hydrauliques, comprenant K_h, K_v et n, et les mesures géophysiques des paramètres S, T, D et R. Notons que lors de l'acquisition un effort particulier a été investi pour l'obtention d'un ensemble de données hydro-géophysiques co-localisé, c'est-à-dire avec des intervalles de mesures pour chacun des paramètres aussi prêt possible l'un de l'autre. Cet ensemble de données d'entrainement est utilisé pour définir les modèles d'HF, les relations hydro-géophysiques et les relations hydrauliques à l'aide d'une approche par machine d'apprentissage (voir l'encadré intitulé « Définition des relations » à la Figure S.20). Soulignons que la définition des modèles d'HF et des relations hydro-géophysiques entre K_h et les données CPT/SMR sont décrites de façon détaillée par Paradis *et al.* (2011; 2014c) et seuls les principaux résultats sont résumés dans cette synthèse. Une fois les modèles d'HF et les relations hydro-géophysiques définis et validés, les données CPT/SMR (pour lesquelles aucune mesure hydraulique n'est disponible) sont utilisées pour prédire les HF et K_h, tandis que les relations hydrauliques sont utilisées pour estimer K_v et n à partir des valeurs de K_h (voir l'encadré intitulé « Vérification et prédiction » à la Figure S.20). Enfin, les distributions spatiales des HF ainsi que de K_h, K_v et n dans la zone d'étude sont obtenues par interpolation (voir l'encadré intitulé « Interpolation » à la Figure S.20). Dans le reste de cette section,

l'application du processus d'intégration des données de la Figure S.20 est illustré pour la section A-A' de la Figure S.6. Soulignons que pour les besoins d'illustration de cette synthèse, les résultats du processus d'intégration des données ne sont illustrés que pour la section A-A', mais que la représentation de l'hétérogénéité des propriétés hydrauliques pourrait être fait pour l'ensemble de la région d'étude puisque les données d'entrainement ont été acquise de manière à être représentatives de l'ensemble de cette région.

3.2 Données hydro-géophysiques d'entrainement

Pour être comparées adéquatement aux mesures de K_h et pour définir des modèles d'HF et des relations hydro-géophysiques fiables (Isaaks et Srivastava 1989), les données CPT/SMR ont tout d'abord été ré-échantillonnées sur une grille régulière et ensuite été remises à la même échelle verticale que les mesures de K_h. En effet, bien que les mesures CPT/SMR aient été prises à un intervalle de temps régulier, la vitesse de pénétration des sondages n'est pas toujours constante. Les mesures CPT/SMR ont donc été ré-échantillonnées sur une grille régulière de 2 cm par intégration trapézoïdale (Davis 1973). Ensuite, l'uniformisation à 15 cm des échelles verticales de mesures de chacun des paramètres CPT/SMR (Tableau S.1) a été réalisée à l'aide d'une moyenne mobile. Le nombre de mesures utilisées pour faire la moyenne mobile a été déterminé selon l'échelle verticale originale de chaque paramètre CPT/SMR.

Les statistiques descriptives des paramètres hydrauliques et CPT/SMR pour les données d'entrainement sont présentées au Tableau S.2. Étant donné que la plage de valeurs des paramètres peut varier de quelques ordres de grandeur pour la plupart des paramètres, une transformation logarithmique a été appliquée pour rendre leur répartition plus normale, comme l'illustrent les histogrammes des Figures S.21 et S.22. La plage de valeurs de chacun des paramètres est semblable, sauf celle du paramètre D et celle de n, qui ont des distributions passablement plus étroites que les autres paramètres.

En ce qui concerne les paramètres hydrauliques, K_h et K_v varient de deux ordres de grandeur, tandis que n varie de moins d'un quart d'ordre de grandeur, tel que couramment observé (Castany 1963; Bear 1979; Freeze et Cherry 1979). Ceci suggère que la caractérisation de K est plus importante que celle de n parce que l'étendue des variations de K est en général plus grande que celle de n, et parce que K agit de manière à changer à la fois la magnitude et la direction du transport advectif, contrairement à n, qui ne change que la magnitude (Koltermann et Gorelick 1996).

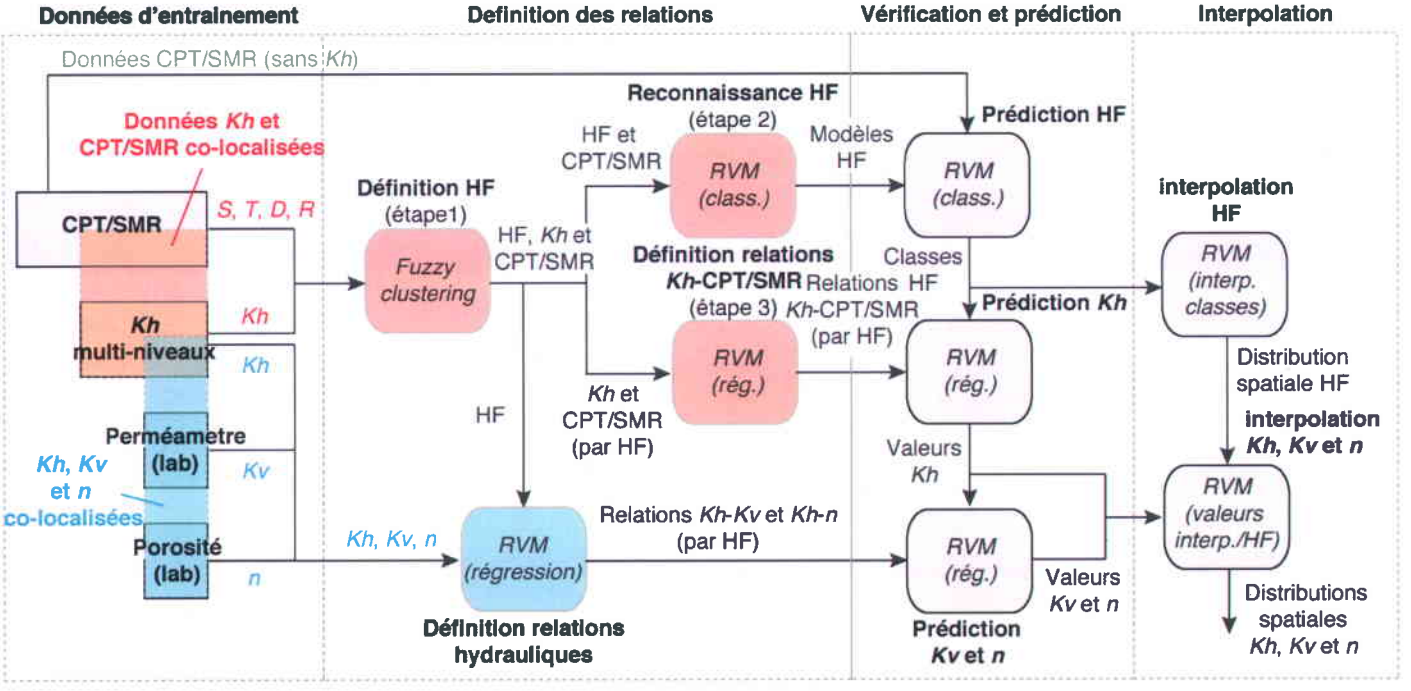


Figure S.20 Processus général d'intégration des données pour la définition de l'hétérogénéité des propriétés hydrauliques, tel qu'appliqué au site de St-Lambert. Les données CPT/SMR sans mesures de K_h sont utilisées pour la prédiction de K_h ou pour vérifier la précision des relations définies en utilisant une fraction (20%) des données d'entrainement. Les acronymes sont : *S* pour friction, *T* pour résistance en pointe, *D* pour constante diélectrique, *R* pour résistivité électrique, K_h pour conductivité hydraulique horizontale K_v pour conductivité hydraulique verticale, *n* pour porosité, HF pour hydrofaciès et RVM pour Relevant Vector Machine. Abréviations: class. (classification), rég. (régression), interp. (interpolation).

Tableau S.2 Statistiques descriptives des données pour les paramètres hydrauliques (K_h: conductivité hydraulique horizontale (m/s); K_v: conductivité hydraulique verticale (m/s); n: porosité totale (fraction du volume)) et CPT/SMR (S: friction (kPa); T: pression en pointe (kPa); D: constante diélectrique (-); R: résistivité électrique (ohm·m)) utilisés comme données d'entrainement.

Paramètre	Nombre	Moyenne	Médiane	Min.	Max.	Plage	Écart-type	Skewness	Kurtosis
log$S$	280	1.74	1.76	0.15	2.51	2.36	0.31	-0.96	2.94
log$T$	280	3.94	4.00	2.41	4.42	2.01	0.29	-1.44	3.63
log$D$	280	1.39	1.39	1.24	1.53	0.29	0.06	0.36	0.85
log$R$	280	2.16	2.16	1.61	3.08	1.47	0.25	1.09	3.25
logK_h	280	-5.04	-5.05	-6.24	-3.92	2.32	0.57	0.09	-1.10
logK_v	59	-6.40	-6.39	-7.56	-4.84	2.72	0.71	0.40	-0.62
log$n$	43	-0.45	-0.45	-0.53	-0.37	0.16	0.03	-0.09	0.95

Tableau S.3 Matrice de corrélation des données pour les paramètres hydrauliques (K_h: conductivité hydraulique horizontale (m/s); K_v: conductivité hydraulique verticale (m/s); n: porosité totale (fraction du volume)) et CPT/SMR (S: friction (kPa); T: pression en pointe (kPa); D: constante diélectrique (-); R: résistivité électrique (ohm·m)) utilisés comme données d'entrainement.

Paramètre	log$S$	log$T$	log$D$	log$R$	logK_h	logK_v	log$n$
log$S$	1.00						
log$T$	0.70	1.00					
log$D$	-0.02	0.16	1.00				
log$R$	0.06	0.40	0.21	1.00			
logK_h	-0.02	0.29	0.33	0.52	1.00		
logK_v	-0.27	-0.02	-0.09	0.20	0.37	1.00	
log$n$	-0.24	-0.19	-0.25	-0.29	0.30	0.59	1.00

Les graphiques présentés aux Figures S.21 et S.22 ainsi que la matrice de corrélation présentée au Tableau S.3 illustrent les relations qui existent entre tous les paramètres hydrauliques et géophysiques. Comme l'indique le Tableau S.3, les corrélations entre K_h et chacun des paramètres CPT/SMR sont faibles à modérées, sauf dans le cas de S pour lequel la corrélation est nulle. Pour les trois paramètres hydrauliques entre eux, les corrélations sont modérées et varient entre 0.37 et 0.59. Cette absence de fortes corrélations entre les divers paramètres hydro-géophysiques laisse donc présumer que la définition des relations pour prédire l'information hydraulique à partir de mesures CPT/SMR peut être une tâche difficile en raison de la non-linéarité des relations liants l'ensemble des données hydro-géophysiques. Dans la prochaine section, une approche d'intégration des données utilisant des machines d'apprentissage est proposée pour relever ce défi.

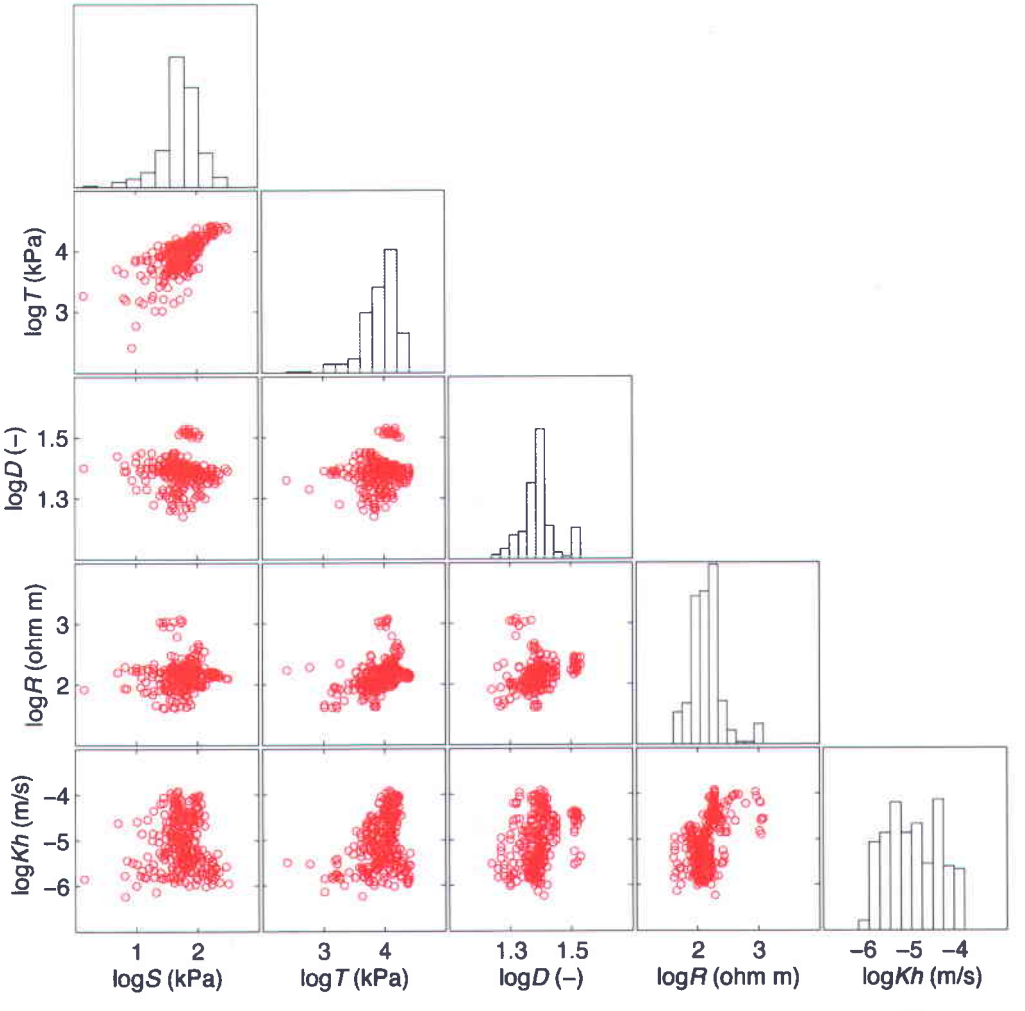


Figure S.21 Matrice de dispersion avec histogrammes pour la conductivité hydraulique horizontale (K_h) et les paramètres CPT/SMR (S: friction; T: résistance en pointe; D: constante diélectrique; R: résistivité électrique) pour les données d'entrainement. Le nombre d'intervalles co-localisés est de 280 (80% des données sont utilisées pour l'entrainement de la machine d'apprentissage et 20% pour sa vérification; voir Figure S.20).

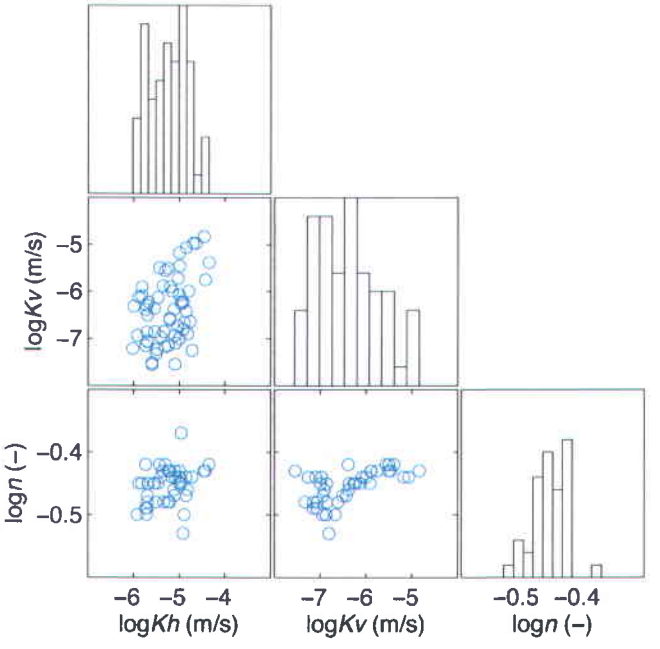


Figure S.22 Matrice de dispersion avec histogrammes pour la conductivité hydraulique horizontale (K_h), la conductivité hydraulique verticale (K_v) et la porosité totale (n) pour les données d'entrainement. Le nombre d'intervalles co-localisés est de 59 pour K_h et Kv, et de 43 pour n.

3.3 Définition des hydrofaciès

Pour composer avec la non-linéarité des données d'entrainement, une approche par machine d'apprentissage a été adoptée. L'entrainement de la machine d'apprentissage, illustrée de façon schématique par les carrés rouges de la Figure S.20, est une procédure en trois étapes qui utilise des algorithmes de classification (non-supervisée et supervisée) et de régression (supervisée) pour concevoir des prédicteurs pour les HF et K_h à partir des données CPT/SMR. La première étape du processus d'entrainement est la définition des HF au moyen d'un algorithme de classification non-supervisé (Dunn 1973; Bezdek 1981; Gustafson et Kessel 1979) en utilisant les mesures co-localisées de K_h et de CPT/SMR. La principale justification pour la classification non-supervisé est que divers matériaux géologiques peuvent avoir des comportements hydro-géophysiques différents et le regroupement statistique permet l'amalgame des caractéristiques hydro-géophysiques semblables sans aucune connaissance géologique préalable. Aux fins de la présente étude, seules les données de K_h sont intégrées dans le processus de classification parce que la quantité moindre de mesures de K_v et de n limiterait l'utilisation de toute les mesures de K_h disponibles. Bien que des relations indépendantes pour K_v et pour n pourraient être élaborées de la même manière que pour K_h, une intégration *a posteriori* de ces paramètres est plutôt proposée, tel qu'illustré à la Figure S.20. En raison du caractère non-unique qui

existe habituellement entre les paramètres hydrauliques et géophysiques (Rubin et Hubbard 2005), l'exactitude de l'information hydrauliques prédite à partir des mesures CPT/SMR peut souffrir d'une perte d'information qui est inhérente à la sensibilité des paramètres géophysiques relativement aux variations des paramètres hydrauliques d'intérêt. Ainsi, une procédure systématique a été appliquée pour la recherche des paramètres CPT/SMR qui sont les plus sensibles aux changements de valeurs de K_h, de manière à maximiser la qualité des prédictions. Le chevauchement des HF dans l'espace géophysique a été défini comme critère de sélection pour la recherche de la combinaison de paramètres CPT/SMR avec la meilleure capacité prédictive. Le nombre d'HF et le type de distance de similitude ont également été incorporés dans cette procédure de recherche. Comme l'illustre la Figure S.23, les résultats de la procédure de recherche montrent que la capacité prédictive des paramètres CPT/SMR augmente généralement avec le nombre de paramètres géophysique utilisés, ce qui laisse présumer que le caractère non-unique entre K_h et les paramètres CPT/SMR est mieux résolu avec un plus grand nombre de paramètres. Il s'agit là d'une caractéristique fondamentale des relations qui lient les paramètres hydrauliques et géophysiques.

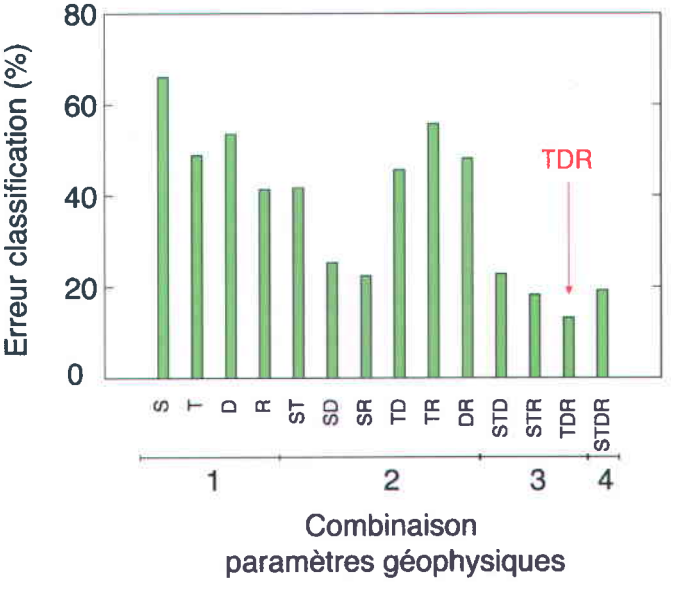


Figure S.23 Erreur de classification liée à la non-unicité des réponses hydrauliques et CPT/SMR pour différentes combinaisons de paramètres géophysiques et un nombre d'hydrofaciès de quatre. La flèche rouge indique que la combinaison de paramètres géophysique TDR a été retenue.

La combinaison des paramètres *T*, *D* et *R* avec 4 HF utilisant la distance de Mahalanobis (Gustafson et Kessel 1979) s'est avérée la classification avec la meilleure capacité prédictive. Les distributions des données à l'intérieur des HF pour chacun des paramètres géophysiques et de K_h issues de cette expérience

sont résumées aux Figures S.24a-d. Comme le montre ces figures, les médianes et la plage de valeurs pour chaque paramètre sont assez distinctes d'un HF à l'autre et ne comportent que quelques valeurs aberrantes. Les médianes de K_h augmentent de manière progressive de l'HF1 à l'HF4 et comportent de légers chevauchements d'un HF à l'autre qui pourraient être attribués à la complexité des réponses hydro-géophysiques et à la nature transitionnelle des sédiments littoraux. En outre, chaque HF présente des profils distincts pour K_h et pour les paramètres géophysiques telles qu'attendues des divers sédiments qui composent l'aquifère puisque ceux-ci peuvent présenter des réponses hydro-géophysiques différentes. Par exemple, l'HF4 a des valeurs médianes plus élevées pour les paramètres T, D, R et K_h, tandis que l'HF1 a des valeurs les plus basses pour les mêmes paramètres.

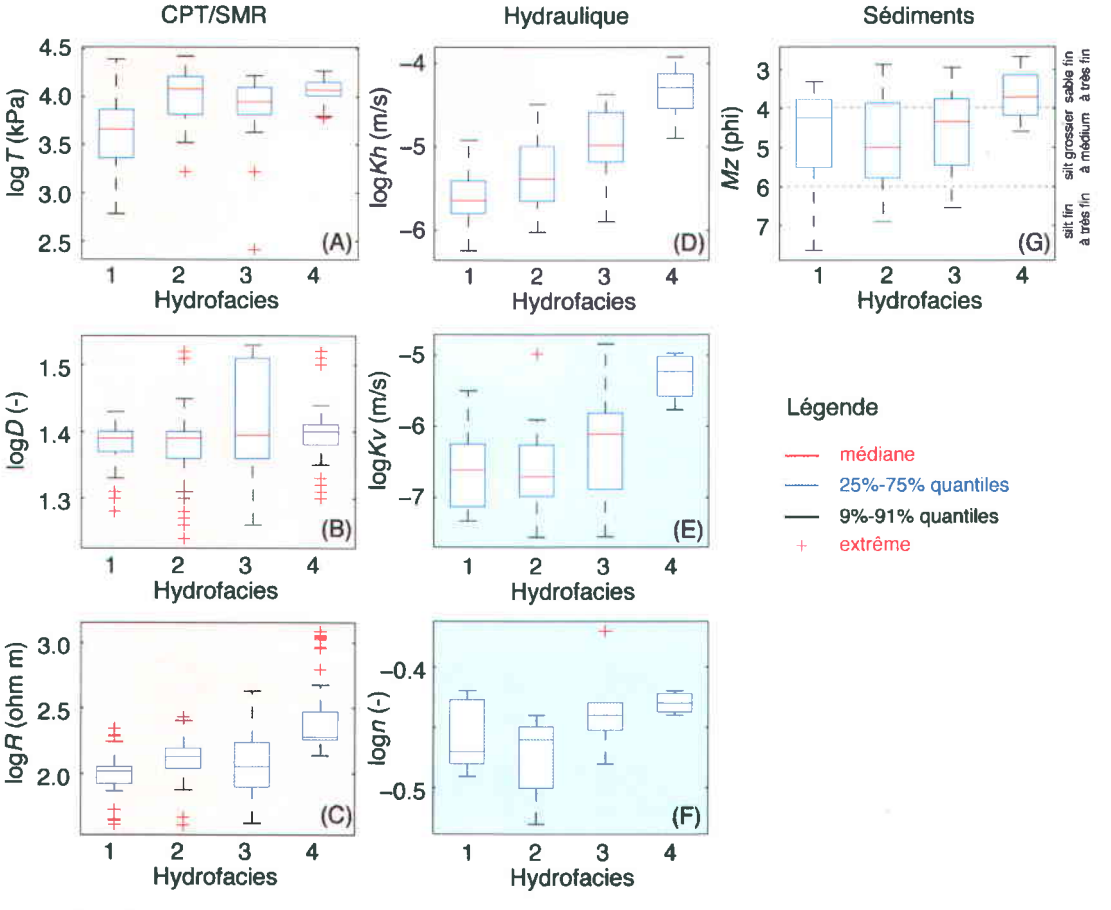


Figure S.24 Distributions des données pour chacun des quatre hydrofaciès résultant de la classification non supervisée (*clustering*) avec la combinaison de paramètres CPT/SMR donnant le meilleur potentiel prédictif (a-d). Cette classification utilise la résistance en pointe (T), la constante diélectrique (D) et la résistivité électrique (R) avec la conductivité hydraulique horizontale (K_h). Les distributions pour la conductivité hydraulique verticale (K_v) et la porosité (n) pour chacun des hydrofaciès sont le résultat d'une intégration *a posteriori* avec les données CPT/SMR (e-f). Les types de sédiments pour chacun des hydrofaciès sont basés sur les analyses granulométriques (g). La classification des sédiments suit la charte de Wentworth (1922): l'échelle phi est basée sur le logarithme en base 2 avec l'exposant négatif de la granulométrie en mm et la moyenne granulométrique des sédiments (Mz) évaluée à partir de (phi14+phi50+phi84)/3 d'après Krumbein et Sloss (1963). Une valeur de Mz faible correspond à des sédiments grossiers.

Les distributions de données pour K_v et n associé *a posteriori* à chacun des HF sont également présentées aux Figures S.24e-f. Basé sur la médiane, le quantile à 25 % et le quantile à 75 %, les HF pour K_v et n sont habituellement distincts l'un de l'autre et suivent essentiellement la même tendance que pour K_h, sauf dans le cas de l'HF1. La comparaison de la valeur médiane pour K_h avec celles de K_v pour chaque HF indique que l'anisotropie de K varie entre 0.05 et 0.2 pour l'HF1 et l'HF4, respectivement.

Enfin, des analyses granulométriques ont été réalisés pour fournir une interprétation géologique aux HF, comme l'illustre la Figure S.24g. Selon la tendance des valeurs minimales et maximales pour la granulométriques moyennes (Mz) des sédiments, les sédiments plus grossiers sont associés aux HF plus perméables et poreux. Il importe de noter que le volume de sédiments utilisés pour les analyses granulométriques est beaucoup plus petit que celui du sous-échantillon de 15 cm et ainsi les valeurs de Mz peuvent afficher une certaine fluctuation dans un sous-échantillon, ce qui peut expliquer la plus grande répartition de la granulométrie des sédiments pour l'HF1.

3.4 Définition des modèles d'hydrofaciès, des relations hydro-géophysiques et des relations hydrauliques

Bien que le processus de classification non-supervisé ait révélé des associations causales entre les données hydro-géophysiques, ce type de classification ne peut être utilisé pour faire des prédictions. Donc, les HF définis précédemment et les données appartenant à chacun des HF doivent être utilisés avec des algorithmes appropriés pour obtenir des modèles prédictifs, tel qu'illustré la Figure S.20. Deux types de modèle prédictif sont ici utilisés en parallèle. Premièrement, les données CPT/SMR pour chacun des HF avec les étiquettes des HF déterminés durant la classification non-supervisée sont utilisées pour entrainer un algorithme de classification supervisé de type RVM pour construire des modèles d'HF. Ces modèles d'HF permettront de faire la reconnaissance des HF basé sur les mesures CPT/SMR seulement. Deuxièmement, les mesures de K_h et les données CPT/SMR associés à chacun des HF sont utilisées pour définir des relations hydro-géophysiques propres à chaque HF à l'aide d'un algorithme de régression supervisé de type RVM. Le choix des RVMs est justifié par leur capacité à produire des relations ayant une bonne capacité de généralisation avec des ensembles de données restreints et fortement non-linéaires (Khalil *et al.* 2005; Camps-Valls *et al.* 2006; Samui 2007; Ghosh et Mujumdar 2008). L'idée principale derrière les RVMs est d'abord d'utiliser une fonction noyau pour linéariser un problème non-linéaire dans un espace de données de dimensions supérieures afin de faciliter la résolution du problème de régression ou de classification. Par la suite, le problème linéarisé est résout dans un cadre Bayésien avant d'être

finalement renvoyé dans l'espace de données non-linéaire original (Tipping 2001). Il est à noter que pour ce processus d'entrainement des RVM une approche de validation croisée utilisant l'ensemble des données d'entrainement a été suivi afin de sélectionner les paramètres optimaux des RVM (ex., fonction noyau et largeur du noyau). Les détails de cette procédure sont présentés par Paradis *et al.* (2014c).

Pour tester les performances des modèles et des relations obtenues, les données hydro-géophysiques d'entrainement ont été divisé en des ensembles d'entrainement et de vérification utilisant respectivement chacun 80 % et 20 % de l'ensemble des données d'entrainement. L'idée étant de simuler les performances de l'approche proposée tel qu'elle serait appliquée pour prédire les HF et K_h à partir des mesures CPT/SMR seulement (voir l'encadré intitulé « Vérification et prédiction » à la Figure S.20). Pour le processus de vérification, les vecteurs des RVMs pour la classification et la régression ont d'abord été obtenus en utilisant l'ensemble de données d'entrainement (80%) avec les paramètres d'ajustement optimaux obtenus au cours de l'entrainement. Les performances prédictives des RVMs ont ensuite été évaluées en utilisant l'ensemble de vérification. Pour éviter le biais dans le choix de l'ensemble de vérification, 100 ensembles d'entrainement et de vérification différents ont été choisies au hasard et appliquées tour à tour. Les résultats du processus de vérification avec les 100 ensembles montrent que le pourcentage médian d'HF mal classifiés à la suite de la classification supervisée avec la RVM est de 14 % (Figure S.25a). Étant donné que l'erreur de classification supervisé (Figure S.25a) est proche de l'erreur de classification obtenue lors de la classification non-supervisée pour cet ensemble de paramètres (Figure S.23), l'erreur de classification avec la RVM peut être attribuée au caractère non-unique qui existe entre K_h et les données CPT/SMR plutôt qu'au manque de performance de la machine. De plus, le coefficient linéaire médian de corrélation entre les valeurs sur le terrain et celles prédites de K_h obtenues par la régression supervisée avec la RVM est de 82 % (Figure S.25b).

Bien que les mesures de K_h soient habituellement faciles à acquérir au moyen d'essais hydrauliques conventionnels, chacun des HF devrait idéalement être également caractérisé par des valeurs de K_v et de n pour mieux comprendre l'écoulement et le transport de l'aquifère étudié. Ces mesures de K_v et de n sont toutefois difficiles à acquérir à cause des difficultés généralement rencontrées pour la collecte d'échantillons de sédiments et leur manipulation en vue d'éventuelles mesures en laboratoire, tel que précédemment discuté. Cela limite par conséquent l'acquisition de volumineux ensembles de données pour ces paramètres et leur intégration statistique direct avec l'ensemble de mesures de K_h beaucoup plus important. Pour ce faire, l'intégration des mesures de K_v et de n est faite *a posteriori,* après l'étape de classification non-supervisée, par association avec les mesures co-localisées de K_h, comme l'illustre la

Figure S.20. Des relations hydrauliques qui lient les mesures de K_v et de n à K_h ont donc été construites pour chacun des HF à l'aide de régression de type RVM et un total de huit relations sont donc disponibles pour prédire K_v et n à partir des données de K_h.

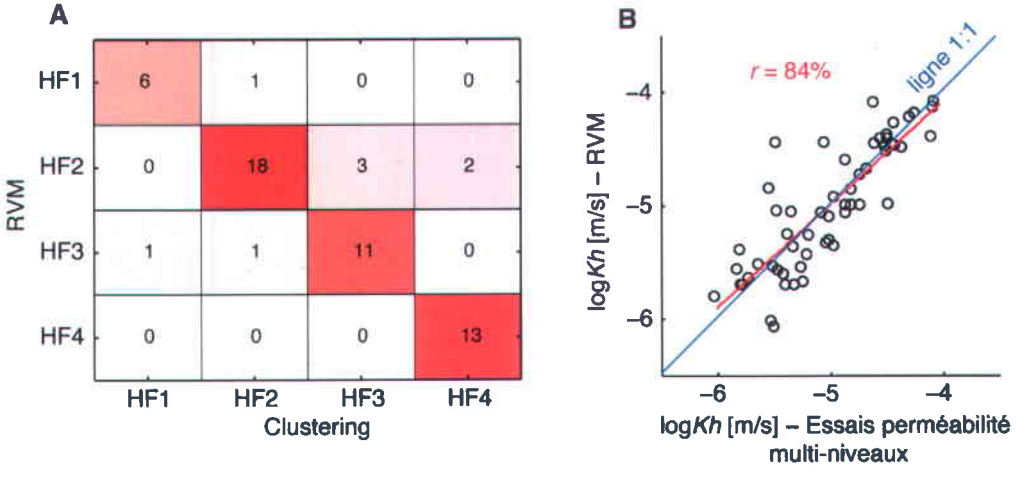


Figure S.25 (a) Matrice de confusion comparant, avec un ensemble de 56 données de vérification, la classification des hydrofaciès obtenue par la machine d'apprentissage (RVM) avec la classification originale effectuée par classification non supervisée (*clustering*). Les éléments de la diagonale indiquent les observations avec des classifications identiques pour les deux méthodes (HF1 6/7, 86%; HF2 18/20, 90%; HF3 11/14, 79%; HF4 13/15, 87%; dans l'ensemble 48/56, 86%). Les éléments hors de la diagonale correspondent aux observations mal classifiées par la classification RVM. (b) Comparaison de la conductivité hydraulique horizontale (K_h) mesurée par les essais à choc hydraulique multi-niveaux avec les estimations provenant de la régression RVM. Le même ensemble de données est utilisé pour (a) et (b).

3.5 Prédiction et distribution spatiale des hydrofaciès et des propriétés hydrauliques

Une fois la machine d'apprentissage entrainée de manière optimale, chaque observation CPT/SMR (vecteur de données pour S, T, D et R) disponible pour la région d'étude a été converti en information hydraulique, comme montré aux Figures S.26a-d pour chacun des sondages disponibles le long de la section géoradar A-A' de la Figure S.6. Toutes les données CPT/SMR ont été ré-échantillonnées et remises à une échelle commune de 15 cm avant d'appliquer la procédure de prédiction séquentielle pour estimer les HF, K_h, K_v et n, comme l'illustre la Figure S.20. Les paramètres CPT/SMR T, D et R avec 4 HF, comme discuté précédemment dans la section portant sur la classification non-supervisée, ont été utilisés pour les prédictions. Les résultats de la conversion montrent des profils verticaux distincts pour les HF et les paramètres hydrauliques pour les zones devant et derrière la barrière de sable, à l'instar de la tendance observée à la Figure S.6 avec les pseudo profils sédimentaires. De plus, les données CPT/SMR converties concordent en général assez bien avec les profils mesurés de K_h, K_v et n, comme l'illustrent les Figures S.26b-d pour l'emplacement P17. Prenez note que pour les besoins des prédictions à l'aide de la

machine d'apprentissage, la totalité des données d'entrainement disponibles a été utilisé pour extraire les vecteurs des RVMs avec les paramètres d'ajustement optimaux obtenus au cours de l'entrainement. Ceci permet donc de se servir de toute l'information disponible pour accroitre la capacité de généralisation des RVMs. Les Figures S.26a-d présentent également les distributions spatiales des HF ainsi que de K_h, K_v et n obtenues au moyen de l'interpolation des données CPT/SMR converties. Bien que plusieurs types d'interpolateurs existent, pour les besoins de cette thèse des interpolateurs ont été construits à partir de RVMs utilisant l'information hydrogéologique (HF, K_h, K_v ou n) provenant des sondages converties et les coordonnées spatiales des intervalles de mesures, comme l'ont proposé Smirnoff *et al.* (2008) pour l'interpolation utilisant la classification supervisée de type SVM. La Figure S.26a présente la distribution spatiale des HF le long de la section géoradar de la Figure S.6 obtenue à la suite de l'interpolation utilisant une RVM pour la classification. Cette interpolation montre nettement la transition entre la zone sablonneuse (HF1) et la zone silteuse (HF 2 à 4) avec la juxtaposition de l'HF1 plus grossier sur les HF plus fins, tel qu'on peut le prévoir pour un environnement littoral à cause de la dynamique des vagues et des marées.

Les Figures S.26b-d présentent également les résultats de l'interpolation de K_h, K_v ou n utilisant des RVMs pour la régression. Une machine d'apprentissage différente a été construite pour chacun des paramètres hydrauliques. Pour préserver les relations entre K_h, K_v et n observées pour l'ensemble des données d'entrainement (Figure S.22), les mêmes paramètres d'interpolation spatiale ont été utilisés pour tous les RVMs (classification et régression). Comme l'illustre la Figure S.26b, la zone sablonneuse est plus homogène et perméable que la zone silteuse. En outre, nous observons une diminution générale des valeurs de K_h plus nous nous approchons du rivage (de A vers A') comme nous pouvons le prévoir avec la réduction des énergies littorales lorsque nous nous déplaçons derrière le barrière de sable. Des observations semblables peuvent également être faites pour les Figures S.26c-d en ce qui concerne K_v et n, respectivement.

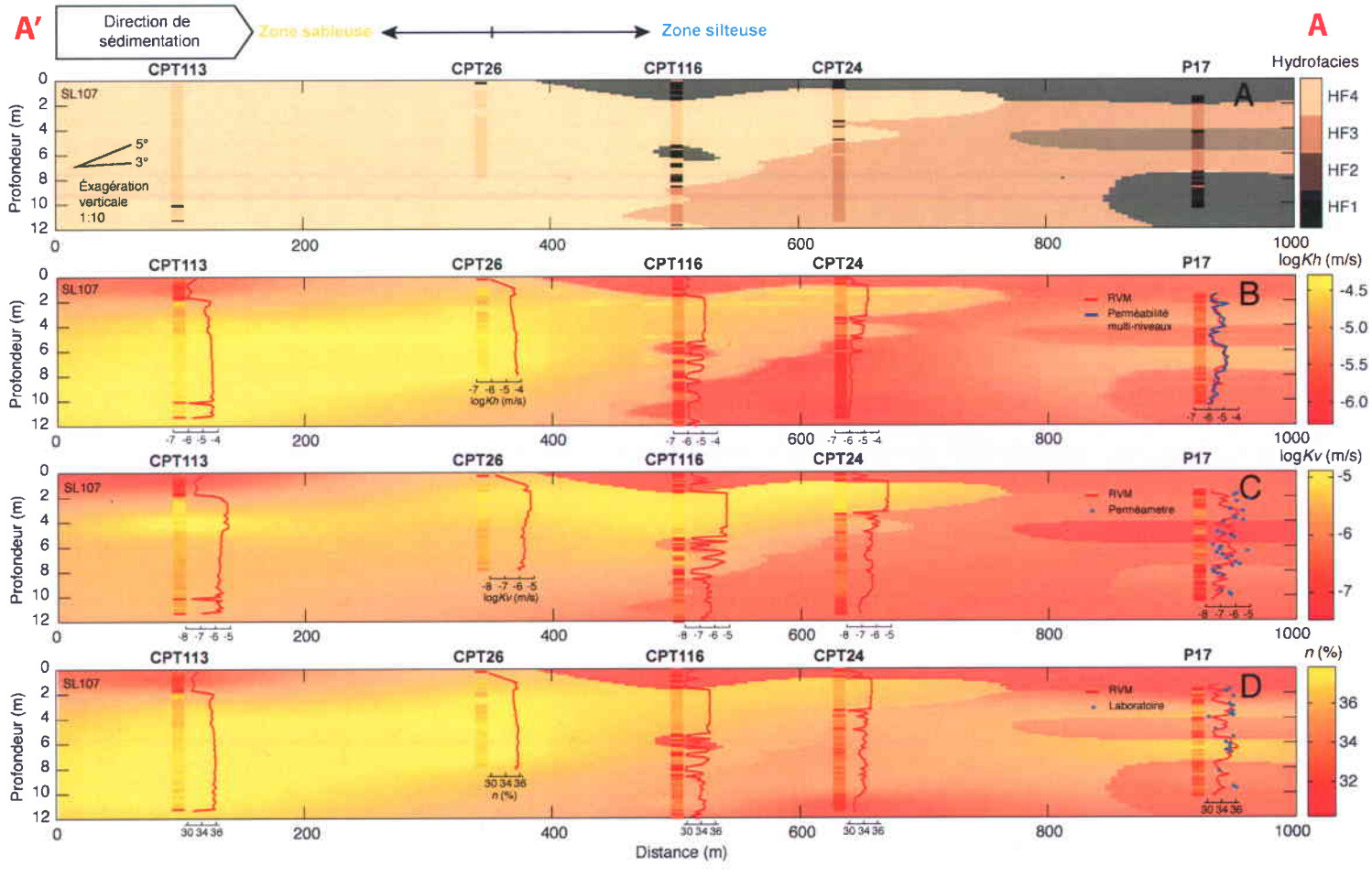


Figure S.26 Distributions spatiales (a) des hydrofaciès (HF); (b) de la conductivité hydraulique horizontale (K_h); (c) de la conductivité hydraulique verticale (K_v); et de la porosité (n) pour la section géoradar A-A' (localisation à la Figure S.4) selon sur l'interpolation de la reconnaissance des hydrofaciès et l'estimation des propriétés hydrauliques basée sur les profils CPT/SMR en utilisant l'approche de la Figure S.20. L'intensité des couleurs pour (a) est proportionnel à la médiane de K_h pour chacun des HF de la Figure S.24.

4 Discussion

4.1 Efficacité des méthodes de caractérisation de la conductivité hydraulique horizontale

Bien que les bénéfices de l'adoption du processus d'acquisition et d'intégration des données proposé puissent être difficiles à quantifier, dans cette section, nous comparons l'efficacité de l'approche d'intégration des données hydro-géophysiques pour estimer K_h à partir des données CPT/SMR, tel qu'appliqué dans l'effort de caractérisation présenté dans cette synthèse, avec divers essais hydrauliques conventionnels. Les essais hydrauliques comparés sont des tests à choc hydraulique multi-niveaux, des tests avec débitmètre de puits et des essais en laboratoire avec perméamètre réalisés sur des échantillons de sédiments. Ces méthodes ont été largement utilisées dans la région d'étude, que ce soit aux fins de production ou de recherche, ce qui permet de faire des estimations assez justes du temps requis pour faire l'acquisition et l'analyse des données. Aux fins de comparaison, une évaluation du temps nécessaire pour estimer K_h à des intervalles de 15 cm le long de 50 profils verticaux d'une longueur de 10 m (500 m au total) a été fait pour chacune des méthodes. Cette base de comparaison correspond aux caractéristiques types (nombre total, longueur moyenne et résolution verticale) des sondages CPT/SMR qui ont été effectués dans le cadre de cette étude (Tableau S.1). Selon les travaux de caractérisation effectués à l'aide des quatre méthodes choisies tout au long de cette étude, une estimation du temps requis pour l'acquisition et l'analyse des données a été faite, comme le résume le Tableau S.4 et l'illustre la Figure S.27. Il importe de noter que les résultats de cette analyse sont représentatifs des travaux effectués par notre équipe pour le site de St-Lambert et ces résultats pourraient varier selon le contexte géologique et l'expérience de l'équipe sur le terrain pour d'autre étude similaires.

Tableau S.4 Temps requis pour différentes méthodes afin d'acquérir 500 m de mesures de conductivité hydraulique horizontale (K_h) à des intervalles verticaux de 15 cm. Ne considère pas le temps de mobilisation et démobilisation, ni les tests de contrôle de qualité.

Méthode	Résolution verticale (m)	Temps par unité	Temps total pour 500 m de K_h à 15 cm (heure)	Commentaires
Sondages CPT/SMR	0.05 à 0.16	1.5 h par 10 m de sondage	483	Inclus installation puits (48 h), essais K_h multi-niveaux (280 h) dans 8 puits et analyse des essais (80 h) pour la définition des relations hydro-géophysiques
Débitmètre	0.15	7 h par puits de 10 m	700	Inclus temps pour atteindre le régime permanent du pompage et l'analyse des données. Temps total inclus l'installation et le développement de 8 puits.
Essais perméabilité (*slug tests*) multi-niveaux	0.15	0.5 h par intervalle de 0.15 m	3633	Assume 0.5 h par intervalle pour l'analyse des essais. Temps total inclus l'installation et le développement de 8 puits.
Perméamètre	0.15	1.5 h par intervalle de 0.15 m	5167	Inclus temps de préparation des échantillons et l'analyse.
		Travaux complémentaires		
Installation et développement de puits	n.a.	6 h par puits de 10 m	300	Inclus 1 h pour installation de puits et 5 h pour le développement du puits.
Échantillonnage de sédiments	n.a.	0.5 h par carotte de 1.5 m	167	Assume conditions homogènes à l'échelle de l'échantillon.

Tel qu'anticipé, les essais avec perméamètre effectués sur des échantillons de sédiments ont nécessité plus de temps que toute autre méthode (Figure S.27) en raison de la collecte des échantillons sur le terrain et de leur préparation en laboratoire (Tableau S.4). De plus, bien que des essais avec perméamètre puissent être effectués de façon continue selon toute la profondeur dans l'aquifère, en pratique la récupération des sédiments est souvent partielle et peut empêcher la réalisation d'essais pour certains intervalles. En effet, comme on peut le constater à la Figure S.8, le profil de K_v à l'emplacement P17 est discontinu parce que la récupération des sédiments était souvent incomplète ou l'état de récupération souvent non propices à la réalisation d'essais de perméabilité. Il importe de noter également que l'hypothèse d'isotropie de K (K_h=K_v) devrait être respectée pour obtenir des estimations valides de K_h à partir d'essais avec perméamètre parce que l'écoulement induit par ce type d'essai est habituellement perpendiculaire à la stratification naturelle des sédiments. Comme l'illustrent les profils de K_h et de K_v de la Figure S.8, des échantillons de sédiments plus petits que 15 cm auraient donc été nécessaires pour obtenir des mesures de perméamètre représentatives de K_h pour les sédiments de l'aquifère de St-Lambert.

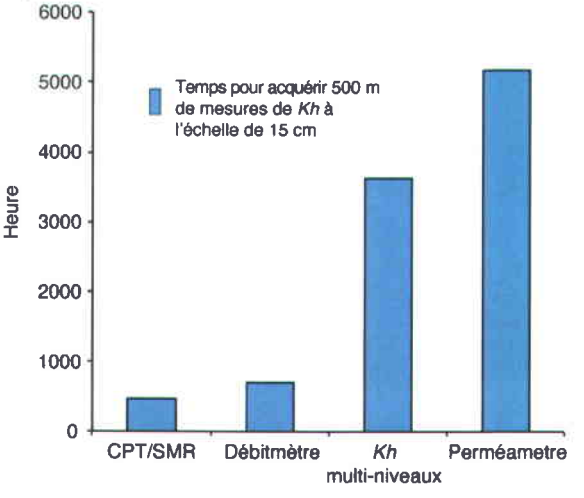


Figure S.27 Temps total pour acquérir 500 m de mesures de conductivité hydraulique horizontale (K_h) à l'échelle de 15 cm pour trois méthodes hydrauliques conventionnelles (débitmètre, essais de perméabilité et perméamètre de laboratoire) et l'approche hydro-géophysique basée sur l'estimation de K_h à partir des données CPT/SMR. Les détails des estimations de temps sont présentés au Tableau S.4.

Même si les tests à choc hydraulique multi-niveaux prennent environ 25 % moins de temps (Tableau S.4 et Figure S.27) et que des profils continus de K_h peuvent être obtenus beaucoup plus facilement qu'avec un perméamètre, cette méthode prend néanmoins beaucoup de temps. Une bonne partie du temps nécessaire pour effectuer les tests est passée à effectuer les essais hydrauliques eux-mêmes,

particulièrement pour les intervalles moins perméables, de même qu'à effectuer le développement des puits essentiel pour obtenir des estimations fiables de K_h (voir le Matériel supplémentaires de l'étude de Paradis *et al.* 2011). Nous remarquons que l'estimation du temps requis pour la réalisation des tests à choc hydraulique est quelque peu surestimée parce qu'elle assume que les puits ne sont utilisés qu'aux fins d'essais hydrauliques. Du point de vue plus général dans l'acquisition des données, ces puits peuvent servir à plusieurs fins (ex., géophysique de puits, échantillonnage géochimique) et le temps nécessaire à l'installation et à l'aménagement de puits doit être en fait réparti sur l'ensemble des opérations. Pour des estimations précises de K_h au moyen des tests à choc hydraulique sur de petits intervalles, les puits soumis à de tels essais devraient être préférablement aménagés sans sable-filtrant pour éviter les effets néfastes des court-circuit hydrauliques. S'il n'est pas possible de procéder de la sorte, les caractéristiques (ex., perméabilité, épaisseur) du sable-filtrant devraient au moins être déterminées et être prises en considération durant l'analyse des essais.

Comme l'illustre la Figure S.27, les tests avec débitmètre offrent une amélioration assez importante du temps d'acquisition des mesures de K_h. Après que le pompage du puits est atteint un régime stationnaire (ou quasi-stationnaire), les mesures de débits à l'aide d'un débitmètre ne prennent habituellement que quelques minutes par intervalle et un profil complet de K_h peut donc être obtenu assez rapidement (Tableau S.4). La valeur moyenne de K_h utilisée pour transposer les mesures du débit en valeurs individuelles de K_h le long du puits peut également être facilement obtenue des données de pompage elles-mêmes ou à partir d'un test à choc hydraulique réalisé indépendamment des mesures de débits. Cependant, la limite de sensibilité d'un débitmètre aux variations de débit peut empêcher l'estimation de K_h pour les intervalles avec de très faibles valeurs de K_h (débit très faible) et seulement des valeurs seuil peuvent être obtenues. De plus, l'important rabattement provoqué par le pompage dans des formations de très faible K_h peut dénoyer certains intervalles dans la partie supérieure de l'aquifère et ainsi compromettre l'estimation de K_h. Tout comme pour les tests à choc hydraulique, les tests avec débitmètre dans les aquifères granulaires nécessitent une installation de puits adéquate sans sable-filtrant et un développement intensif des puits.

Enfin, l'approche hydro-géophysique semble être celle qui est la plus efficace en termes de temps (Figure S.27). Il importe de noter que l'estimation du temps pour cette approche comprend le temps nécessaire à l'installation et à l'aménagement de 8 puits par enfoncement (incluant le développement) et à la réalisation de 280 tests à choc hydraulique sur des intervalles de 15 cm, tel qu'effectué dans cette étude. Contrairement aux tests hydrauliques dans les puits (tests à choc hydraulique et débitmètre), une fois que

les relations hydro-géophysiques ont été définies à l'aide d'un nombre suffisant de puits, aucun autre puits n'est nécessaire, et seuls des sondages CPT/SMR sont effectués. Ceci permet de réduire considérablement le temps total de caractérisation, puisque le temps nécessaire pour effectuer un sondage CPT/SMR, sans l'installation et l'aménagement de puits, est relativement court (Tableau S.4). La réalisation de tests avec débitmètre au lieu de tests à choc hydraulique pourraient également réduire considérablement le temps d'acquisition des données associés à la réalisation des essais hydrauliques pour l'approche hydro-géophysique. Cependant, l'intégration des données pour la définition des relations propres au site exige plus de temps que l'analyse des tests hydrauliques conventionnels et l'estimation du temps requis pour l'analyse des données pour l'approche hydro-géophysique présentée au Tableau S.4 suppose de toute évidence un certain degré de familiarisation avec les outils statistiques.

Dans un ordre d'idée différent, si les levés géoradar avait été utilisés au lieu des sondages CPT/SMR pour l'établissement des relations hydro-géophysiques, il aurait été possible d'augmenter considérablement l'efficacité de la caractérisation de l'aquifère, puisque pour cette étude, les 21 km de levés géoradar ont pris environ la moitié du temps nécessaire pour la réalisation des 53 sondages par enfoncement (Tableau S.1). Même si la résolution verticale des levés géoradar est plus grande que celle des sondages CPT/SMR, la meilleure continuité spatiale des levés géoradar permet de mieux définir l'architecture des hétérogénéités de l'aquifère. Cependant, la forte atténuation du signal géoradar observé sur un bon nombre de levés effectués dans la zone silteuse (Figure S.4) masque souvent la présence de lit sableux inter-digités avec les silts, ce qui limitait l'utilisation des données géoradar aux fins d'estimation des propriétés hydrauliques dans cette section de l'aquifère. Une autre limitation du géoradar, qui n'est pas exclusive à la région d'étude cette fois, est que les levés géoradar ne procurent que des renseignements à propos de la constante diélectrique, ce qui peut affaiblir la capacité prédictive d'une approche hydro-géophysique basée sur le géoradar en raison du nombre limité de paramètres géophysiques disponibles pour effectuer des corrélations avec les données hydrauliques. Ceci met en relief, avec comme appui la Figure S.23, qu'une approche hydro-géophysique basée sur des méthodes géophysiques de surface (ex., géoradar, sismique réflexion, levés électrique) ou aéroporté (ex., TDEM), devrait inclure des levés géophysique de plusieurs types sur la même section d'aquifère pour obtenir une meilleure conversion des données géophysiques en propriétés hydrauliques.

De plus, des tests d'interférence verticale à choc hydraulique (Paradis et Lefebvre 2013) pourraient également accroitre l'efficacité globale du processus de caractérisation. En effet, les tests d'interférence verticale nécessite à peu près le même équipement et le même temps de réalisation que des essais à choc

hydraulique multi-niveaux conventionnels et fournissent en plus des valeurs de K_h, des valeurs de K_v aussi. Donc, l'adoption de tests d'interférence verticale au lieu de tests à choc hydraulique multi-niveaux, tel qu'effectué dans cette étude, pourrait augmenter le nombre de mesures de K_v afin d'établir un meilleur ensemble de données d'entrainement et à remplacer les essais avec perméamètre sur des échantillons de sédiments. Cependant, le temps requis pour l'analyse des données issues des tests d'interférence verticale peut être plus long que dans le cas des tests à choc hydraulique multi-niveaux, surtout pour les profils hétérogènes qui nécessite d'utiliser l'inversion numérique pour l'analyse des données.

Finalement, la tomographie hydraulique tel que proposée par Paradis *et al.* (2014a; 2014b) pourrait également être utilisée pour remplacer les tests à choc hydraulique multi-niveaux dans des puits et les tests avec perméamètre en laboratoire pour estimer K_h et K_v. Cependant, la nécessité d'avoir au moins deux puits rapprochés pour réaliser la tomographie est surement une alternative moins attrayante qu'un test d'interférence verticale nécessitant seulement un puits. Cette technique est alors certainement plus intéressante pour acquérir l'hétérogénéité tridimensionnelle de K_h, K_v et S_s à moyenne échelle pour des sites relativement restreints. Néanmoins, pour des études hydrogéologiques régionales qui nécessitent habituellement des valeurs de propriétés hydrauliques sur des volumes importants d'aquifère, la tomographie hydraulique peut être un complément intéressant aux essais de pompage généralement utilisés pour ce genre d'études. En effet, une meilleure connaissance de l'hétérogénéité des propriétés hydraulique à proximité d'un puits d'essai, en particulier pour des conditions très hétérogènes, peut permettre de mieux comprendre la dynamique observée d'un essai de pompage ainsi que les caractéristiques de mise-à-l'échelle des propriétés hydrauliques de ce test. La tomographie hydraulique à choc hydraulique permet également d'obtenir K_v et S_s, ce que permettent généralement difficilement les essais de pompage. Notons qu'une couverture tomographique de faible à moyenne densité peut être obtenue dans un temps équivalent à celui d'un essai de pompage de duré moyenne.

4.2 Travaux complémentaires à cette thèse au site de St-Lambert

En concomitance avec cette étude sur la caractérisation de l'hétérogénéité des propriétés hydrauliques, d'autres travaux ont été effectués dans la région d'étude de St-Lambert. Pour renforcer davantage la confiance envers les distributions spatiales des propriétés hydrauliques interpolées pour la région d'étude, partiellement illustrées aux Figures S.27a-d, des travaux complémentaires ont été effectués par Brunet *et al.* (2012) pour contraindre la structure et les valeurs des propriétés hydrauliques des hétérogénéités avec des profils verticaux de charges hydrauliques mesurées dans des puits. En effet, il a été démontré que les

variations de charges hydrauliques mesurées à différentes profondeurs dans l'aquifère peuvent indiquer le degré de continuité latérale d'hétérogénéités observées dans des puits. Bien que divers algorithmes d'interpolation et de simulation géostatistique peuvent être utilisés selon le type d'hétérogénéités géologiques existant (Koltermann et Gorelick 1996; Falivene *et al.* 2007), le nombre de champs de propriétés hydrauliques pouvant être générés peut être très élevé, en particulier pour les sites avec peu de données, et plusieurs d'entre eux peuvent ne pas être nécessairement représentatifs de la région d'étude. Dans leur approche, Brunet *et al.* (2012) utilise un algorithme de simulation géostatistiques multi-points (Strebelle 2002) qui est combiné à des algorithmes d'inversion de points pilotes (Doherty *et al.* 2010) et de déformation graduelle (Le Ravalec et Mouche 2012) pour contraindre la distribution spatiale de K_h avec des charges hydrauliques mesurées à différentes profondeur dans l'aquifère. Cette approche qualifiée d'interpolation contrainte permet ainsi de sélectionner les champs de propriétés hydrauliques les plus plausibles.

En outre, des profils d'âge géochimique de l'eau souterraine ont été mesurés par Tremblay *et al.* (2012) dans les puits par enfoncement dotés de crépine couvrant la totalité de l'aquifère en vue de les intégrer dans un modèle numérique pour contraindre le champ de vitesse d'écoulement de l'eau souterraine dans l'aquifère de St-Lambert. Par exemple, les champs de K_h, K_v et n présentés à la Figure S.27a-d peuvent être utilisés dans un modèle numérique d'écoulement pour générer un champ de vitesse qui peut être comparé aux profils d'âge mesurés. À l'instar des charges hydrauliques mesurées à différentes profondeurs dans l'aquifère, la géométrie des hétérogénéités et les valeurs des propriétés hydrauliques peuvent être ajustés de manière à honorer les mesures d'âge. Soulignons que cette approche a déjà été mise en application avec succès pour d'autres sites (ex., Murphy *et al.* 2008).

Comme approche alternative à celle présentée dans cette thèse, Ruggeri *et al.* (2013) ont proposé l'intégration de levés électrique de surface (ERT : *Electrical Resistivity Tomography*) avec la résistivité électrique des sondages CPT/SMR et des mesures de K_h aux puits pour convertir en valeurs de K_h à haute résolution des sections d'ERT. Dans cette approche, les levés ERT procurant des mesures continues de résistivité électrique entre les puits sont utilisés comme grille de fond sur laquelle les mesures de K_h aux puits sont extrapolées, au lieu de recourir à des algorithmes d'interpolation conventionnelles qui utilise seulement l'information aux puits pour déduire des relations spatiales pour distribuer l'information hydrogéologique loin des puits. Jusqu'à ce jour, cette approche a été appliquée avec succès pour convertir un levé ERT effectué dans la zone sableuse (voir la Figure S.4c).

Enfin, Gloaguen *et al.* (2012) ont proposé l'intégration des valeurs de résistivité électrique et de constante diélectrique des sondages CPT/SMR avec les données géoradar pour faire la conversion des radargrammes en profondeur. Cette approche a été développée pour mieux estimer la profondeur des réflecteurs géoradar et pour éventuellement mieux utiliser les données géoradar pour l'estimation des propriétés hydrauliques.

5 Résumé et conclusions

Cette synthèse a présenté un cadre général d'acquisition et d'intégration des données développé et appliqué pour la caractérisation de l'hétérogénéité des propriétés hydrauliques à une échelle intermédiaire (quelques km^2), qui se veut une échelle adaptée à la modélisation numérique de l'écoulement de l'eau souterraine et du transport des contaminants en relation avec les zones sources de ces contaminants ainsi qu'à leur récepteurs. Ce cadre a été appliqué pour une zone d'étude de 12 km^2 en vue d'étudier la migration et l'atténuation naturelle d'un panache de lixiviat émis par un ancien site d'enfouissement sanitaire aménagé dans un aquifère granulaire littoral peu profond. L'approche de caractérisation est basée sur l'utilisation de données géophysiques indirectes à haute résolution et l'acquisition ciblée de données hydrauliques directes. Une campagne intensive de levés géoradar a été effectuée afin de fournir le contexte géologique général pour la sélection optimale des emplacements de sondages CPT/SMR et l'installation de puits par enfoncement dotés de longues crépines pour la réalisation de tests hydrauliques. L'approche de caractérisation s'appuie donc sur la conversion de données CPT/SMR en information hydraulique au moyen des relations hydro-géophysiques développées spécifiquement pour le site à l'étude à l'aide d'une approche par machine d'apprentissage. Les paramètres hydrauliques comprenaient la conductivité hydraulique horizontale (K_h), mesurées dans les puits par enfoncement par des tests de perméabilité avec une résolution verticale de 15 cm, et la conductivité hydraulique verticale (K_v) et la porosité (n) obtenues en laboratoire par des mesures réalisés sur des échantillons de sédiments. Les paramètres CPT/SMR comprenaient la résistance à la pénétration (T), la friction mécanique (S), la constante diélectrique (D) et la résistivité électrique (R). Toutes les mesures hydrauliques et CPT/SMR co-localisées ont été mises à une échelle verticale commune de 15 cm dans le but d'établir des relations statistiques. L'acquisition d'un ensemble de données d'entrainement hydro-géophysiques représentatif de l'éventail des sédiments observés dans la région d'étude était basée sur une approche de type générale-vers-spécifique. Spécifiquement, l'information sur la géologie régionale et les levés géoradar ont été

utilisés pour localiser les sondages CPT/SMR, les puits d'observation et les échantillons de sédiments de façon à couvrir la plage complète des réponses hydrauliques et géophysiques. Quatre hydrofaciès (HF) représentant les sédiments et ayant des relations hydro-géophysiques distinctes ont été définis à l'aide d'algorithmes de classification (non-supervisé et supervisé) et de régression (supervisé) pour permettre la prédiction des HF et de K_h à partir des données CPT/SMR. Le peu de mesures de K_v et de n obtenues en laboratoire ont été intégrées *a posteriori* avec les mesures de la K_h co-localisées pour définir pour chacun des HF des relations hydrauliques propres au site. La conversion des sondages CPT/SMR distribuées sur l'ensemble de la région d'étude a permis d'obtenir une banque de données importante sur HF, K_h, K_v et n afin de mieux définir l'hétérogénéité des propriétés hydrauliques de l'aquifère de St-Lambert. Un exemple d'interpolation de l'information hydraulique a été présenté pour une section géoradar pour laquelle plusieurs sondages CPT/SMR et données hydrauliques directes sont disponibles. Les principales constatations de cette étude sont résumées ci-dessous.

- Le processus d'acquisition de données de type générale-vers-spécifique permet une collecte de données rationnelle, systématique et progressive. Ce processus, élaboré et appliqué pour la région d'étude, a été pensé pour réduire au minimum le nombre de sites de mesures tout en optimisant la qualité, la portée et la signification des données directes et indirectes recueillies. Par exemple, la définition d'un modèle conceptuel hydrogéologique général au moyen de la géologie régionale, des levés géoradar et des sondages CPT/SMR a permis le repérage d'intervalles significatifs pour les tests de perméabilité multi-niveaux, ce qui a permis de réduire le nombre d'essais hydrauliques nécessitant beaucoup de temps de réalisation et a assuré de couvrir l'étendue complète des réponses géophysiques et hydrauliques observées pour la région d'étude. Cette approche d'acquisition de données a fourni un ensemble de données d'entrainement représentatif et complet pour l'entrainement de la machine d'apprentissage utilisée pour évaluer les propriétés hydrauliques de l'aquifère à partir des données CPT/SMR.

- Une approche hydro-géophysique pour caractériser les propriétés hydrauliques hétérogènes des aquifères est passablement plus efficace en termes de temps que les approches d'essais hydrauliques conventionnelles, qui peuvent rarement fournir la quantité de données nécessaires pour évaluer l'hétérogénéité d'un aquifère, surtout à une échelle intermédiaire de quelques km^2. Bien que scientifiquement il n'y ait aucune limite quant à la quantité de données qui peut être recueillie sur un aquifère, des facteurs économiques dictent cependant le plus souvent cette limite, qui dans bien des cas peut compromettre l'acquisition du niveau de données nécessaire pour résoudre adéquatement un

problème hydrogéologique. Par conséquent, une approche hydro-géophysique, tel que celle proposée dans cette thèse, peut contribuer à atteindre un certain équilibre entre la représentation adéquate de l'hétérogénéité et les contraintes économiques.

- Par le biais de relations hydro-géophysiques spécifiques au site, les données CPT/SMR contiennent beaucoup d'information à propos de l'hétérogénéité des propriétés hydrauliques de l'aquifère granulaire dans la région d'étude. En effet, l'interpolation des données de sondages converties en information hydraulique montre que les distributions spatiales des HF ainsi que de K_h, K_v et n reflètent l'architecture des sédiments de l'aquifère littoral, ce qui comprend l'inter-digitation et les changements spatiaux graduels.

- Des relations hydrauliques et hydro-géophysiques bien définies sont essentielles pour utiliser avec succès les données géophysiques pour la caractérisation des propriétés hydrauliques. Dans le cadre de cette étude, l'utilisation d'une sonde multi-paramètres, comme la sonde utilisée pour effectuer les sondages CPT/SMR, a permis de réduire le caractère non-unique des liens qui existent entre les propriétés hydrauliques et géophysiques en fournissant une série de paramètres géophysiques qui peuvent être corrélés aux divers paramètres hydrauliques. En effet, le succès de la prédiction des valeurs de propriétés hydrauliques à partir des mesures CPT/SMR augmente avec le nombre de paramètres géophysique utilisés. L'utilisation d'algorithmes tel que les RVMs a également permis de composer avec la non linéarité des relations hydro-géophysiques et contribué au développement de relations ayant de bonnes capacités de généralisation.

- La quantité et la qualité des données disponibles sont également importantes pour le succès de la caractérisation d'un aquifère et des efforts ont été consacrés au développement de méthodes hydrauliques efficaces et permettant de représenter les conditions de K. À cet égard, l'adaptation du débitmètre de puits pour les aquifères granulaires a été proposé pour permettre d'acquérir beaucoup plus rapidement des mesures de K_h qu'avec les tests hydrauliques conventionnels. De plus, la compréhension approfondie des principes physiques qui gouvernent l'écoulement de l'eau souterraine a aussi permis de développer les tests d'inférence verticale et la tomographie à choc hydraulique. Ces tests permettent de mesurer K_v lorsque l'hétérogénéité des sédiments à très petites échelles induit l'anisotropie de K à l'échelle de mesure des essais hydrauliques qui ont généralement des limites pratiques d'expérimentation ne permettant pas la mesure sur de très petits volumes. Des analyses de

sensibilité et de résolution des charges hydrauliques par rapport aux variations de propriétés hydrauliques ont ainsi permis de concevoir des dispositifs expérimentaux et des approches d'analyse permettant de généraliser les capacités de caractérisation des méthodes hydrauliques conventionnelles. Les données obtenues à l'aide de ces nouvelles méthodes hydrauliques, utilisées comme telle ou pour l'établissement de relations hydro-géophysiques, contribuent donc à donner des images plus complètes des paramètres qui contrôlent l'écoulement et le transport dans les aquifères.

- L'utilisation de puits par enfoncement aménagés sans sable filtrant et entièrement crépinés sur toute l'épaisseur saturée de l'aquifère a été un élément clé pour l'acquisition de mesures hydrauliques de qualité et le développement des nouvelles méthodes de caractérisation hydraulique proposées. D'abord, l'utilisation de puits sans sable filtrant permet la mesure directe des propriétés hydrauliques des sédiments en contact direct avec la crépine. Aussi, l'absence de sable-filtrant permet de réduire les effets de dissipation et d'interférence généralement associés à la présence de sable-filtrant qui altèrent souvent les mesures hydrauliques. Ceci, combiné à l'aménagement de puits entièrement crépinés, a aussi permis d'effectuer des tests de perméabilité (*slug tests*) de haute résolution (15 cm) et de fournir des profils continus de K_h pour caractériser tous les types d'hydrofaciès retrouvés dans la région d'étude. Dans le processus d'acquisition des données décrit dans la présente synthèse, les méthodes par enfoncement ont donc été essentielles pour fournir à la fois des données indirectes (CPT/SMR) et directes (K_h) à haute résolution nécessaires pour évaluer l'hétérogénéité des propriétés hydrauliques.

Compte tenu de la complexité des systèmes aquifères en général et de la difficulté à les représenter, le focus général de cette thèse était le développement de méthodes et d'approches pragmatiques de caractérisation de l'hétérogénéité des propriétés hydrauliques pouvant mener à une meilleure compréhension des processus d'écoulement et de transport dans les aquifères. Bien que les méthodes utilisées et les relations définies dans l'effort de caractérisation présenté soient propres au site de St-Lambert, le cadre d'acquisition et d'intégration des données est quant-à-lui général par rapport à ses principes de base. Ainsi, ce cadre pourrait être élargi à d'autres contextes hydrogéologiques en utilisant des méthodes de caractérisation hydraulique et géophysique similaires ou différents selon le cas. Bien que les méthodes et les approches proposées dans cette thèse puissent nécessiter un certain changement dans les pratiques et présentent un certain niveau de complexité dans la maitrise des outils, une des prémisses de cette thèse est que l'approche d'homogénéisation des conditions hydrogéologiques, qui a guidé bon nombre d'études hydrogéologiques, a peut être atteint sa limite. Des études détaillées basées sur des

données significatives (quantité et qualité) de caractérisation et de suivi sont nécessaires. À cet égard, à la lumière des résultats de caractérisation présentés pour le site d'étude de St-Lambert, ce site peut sembler être un cas particulier de complexité de l'hétérogénéité. Cependant, la caractérisation détaillée d'autres sites pourrait révéler que le site de St-Lambert est une règle plutôt qu'une exception et que pour faire face avec confiance aux enjeux environnementaux et sociétaux grandissant reliés aux ressources en eaux souterraines, un changement de paradigme est peut être nécessaire au niveau de la caractérisation des aquifères.

Références

American Society for Testing and Materials-ASTM (2000) Standard test method for performing electronic friction cone and piezocone penetration testing of soils. Annual book of ASTM standards. D5578-95 (Reapproved 2000)

American Society for Testing and Materials-ASTM (2004) D6724: Standard guide for installation of direct push ground water monitoring wells. ASTM International, 9 pp.

Anderson MP (1997) Characterization of geological heterogeneity. In: Dagan G, Neuman SP (ed) Stochastic Subsurface Hydrology, Cambridge University Press, Cambridge

Anderson MP, Aiken JS, Webb EK, Mickelson DM (1999) Sedimentology and hydrogeology of two braided stream deposits. Sediment Geol 129(3):187-199

Anderson MP, Woessner WW (1992) Applied Groundwater Modeling. Academic Press

Bear J (1979) Hydraulics of groundwater. McGraw-Hill, New York

Bezdek JC (1981) Pattern recognition with fuzzy objective function algoritms. Plenum Press, New York

Bolduc A (2003) Géologie des formations superficielles, Charny, Québec. Commission Géologique du Canada, Dossier public 1976, échelle 1/50000

Bouwer H, Rice RC (1976) A slug test method for determining hydraulic conductivity of unconfined aquifers with completely or partially penetrating wells. Water Resour Res 12(3):423-428

Bohling, GC, Butler JJJr (2001) Lr2dinv: A finite-difference model for inverse analysis of two-dimensional linear or radial groundwater flow. Comput Geosci, 27:1147-1156

Bradford JH, Babcock E (2013) The need to adapt the exploration model from the oil patch to contaminated-site characterization: A case from Hill AFB, Utah, USA. The Leading Edge, 32(7):750-756

Brunet P, Paradis D, Lefebvre R, Gloaguen E (2012) Using CPT and multi-level piezometric data to assess aquifer heterogeneity at the sub-watershed scale. 39th IAH Congress, Niagara Falls, Canada, September 16-21, Paper 844

Butler JJJr (1998) The Design, Performance, and Analysis of Slug Tests. Lewis Publishers, Boca Raton, Florida

Butler JJJr (2005) Hydrogeological methods for estimation of hydraulic conductivity. In: Rubin Y, Hubbard S (ed) Hydrogeophysics, Springer, New York, pp 23-58

Camps-Valls G, Gómez-Chova L, Muñoz-Marí J, Vila-Francés J, Amorós-López J, Calpe-Maravilla J (2006) Retrieval of oceanic chlorophyll concentration with relevance vector machines. Remote Sens Environ 105(1):23-33

Castany G (1963) Traité pratique des eaux souterraines. Dunod, Paris

Davis JC (1973) Statistics and data analysis in geology. Wiley, New York

Davis JL, Annan AP (1989) Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy. Geophys Prospect 37(5):531–551

de Marsily G, Delay F, Gonçalvès J, Renard P, Teles V, Violette S (2005) Dealing with spatial heterogeneity. Hydrogeol J 13:161-183

Doherty JE, Fienen MN, Hunt RJ (2010) Approaches to highly parameterized inversion: Pilot-point theory, guidelines, and research directions. U.S. Geol. Surv. Sci. Invest. Rep., 2010-5168, 36 pp.

Dunn JC (1973) A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. J Cybernetics 3:32-57

Eaton TT (2006) On the importance of geological heterogeneity for flow simulation. Sediment Geol 184:187-201

Falivene, O., Cabrera, L., Munoz, J.A., Arbués, P., Fernandez, O. et Saez, A. 2007. Statistical grid-based facies reconstruction and modelling for sedimentary bodies. Alluvial-palustrine and turbiditic examples. Geologica Acta, 5(3) 199-230

Fellenius BH, Eslami A (2000) Soil profile interpreted from CPTu data. "Year 2000 Geotechnics" Geotechnical Engineering Conference, Asian Institute of Technology, Bangkok, Thailand, November 27-30, 2000

Freeze RA, Cherry JA (1979) Groundwater. Prentice-Hall, Englewood Cliffs, New-Jersey

Frei S, Fleckenstein JH, Kollet SJ, Maxwel RM (2009) Patterns and dynamics of river–aquifer exchange with variably-saturated flow using a fully-coupled model. Journal of Hydrology 375, 383–393

Gelhar LW (1993) Stochastic subsurface hydrology. Prentice-Hall, Englewood Cliffs, New-Jersey

Ghosh S, Mujumdar P (2008) Statistical downscaling of GCM simulations to streamflow using relevance vector machine. Adv Water Resour 31:132-146

Gloaguen E, Lefebvre R, Ballard JM, Paradis D, Tremblay L, Michaud Y. (2012) Inference of the two dimensional GPR velocity field using collocated cokriging of direct push permittivity and conductivity logs and GPR profiles. J Appl Geophys 78:94-101

Gustafson DE, Kessel WC (1979) Fuzzy clustering with a fuzzy covariance matrix. Proceeding of IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes, San Diego, CA, 761-766

Heinz J, Kleineidam S, Teutsch G, Aigner T (2003) Heterogeneity patterns of Quaternary glaciofluvial gravel bodies (SW-Germany): application to hydrogeology. Sediment Geol 158:1-23

Henebry BJ, Robbins GA (2000) Reducing the influence of skin effects on hydraulic conductivity determinations in multilevel samplers installed with direct push methods. Ground Water 38(6):882-886

Hess AE (1986) Identifying hydraulically conductive fractures with a slow-velocity borehole flowmeter. Can Geotech J 23:69-78

Isaaks EH, Srivastava RM (1989) An introduction to applied geostatistics. Oxford University Press, New York

Khalili A, Almasri MN, McKee M, Kaluarachchi JJ (2005) Applicability of statistical learning algorithms in groundwater quality modeling. Water Resour Res 41:W05010, doi:10.1029/2004WR003608

Klute A, Dirksen C (1986) Hydraulic conductivity and diffusivity: laboratory methods. Methods of soil analysis. Part 1. In: Klute A (ed) Physical and Mineralogical Methods, Agronomy Monograph 9. American Society of Agronomy, pp 687-734

Koltermann CE, Gorelick SM (1996) Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches. Water Resour Res 32(9):2617-2658.

Krumbein WC, Sloss LL (1963) Stratigraphy and Sedimentation. 2nd edn. W.H. Freeman and Company, San Francisco

Kruseman GP, de Ridder NA (1990) Analysis and evaluation of pumping test data. ILRI Publishing, Netherlands

Lafuerza S, Canals M, Casamor JL, Devincenzi JM (2005) Characterization of deltaic sediment bodies based on in situ CPT/CPTU profiles: A case study on the Llobregat delta plain, Barcelona, Spain. Mar Geol 222-223:497-510

Lamarche L, Tremblay L (2012) Géologie des formations superficielles pour le site de St-Lambert-de-Lauzon, Québec, Unpublished research data

Lunne T, Robertson PK, Powell JJM (1997) Cone penetration testing in geotechnical practice. Spon Press, New York

Le Ravalec M, Mouche E (2012) Calibrating transmissivities from piezometric heads with the gradual deformation method: An application to the culebra dolomite unit at the waste isolation pilot plant (WIPP), New Mexico, USA. Journal of Hydrology, 472-473:1-13

Levy BS, Pannell L (1991) Evaluation of a pressure system for estimating in-situ hydraulic conductivity. Ground Water Management No. 5. In Proceedings 5th National Outdoor Action Conference, 31–45. Las Vegas, Nevada

Murphy S, Ouellon T, Ballard JM, Lefebvre R, Clark ID (2011) Tritium–helium groundwater age used to constrain a groundwater flow model of a valley-fill aquifer contaminated with trichloroethylene (Quebec, Canada). Hydrogeology Journal, 19(1):195-207

Neumann SP, Wierenga PJ (2003) A comprehensive strategy of hydrogeological modeling and uncertainty analysis for nuclear facilities and sites. U.S. Nuclear Regulatory Commission, Office of Nuclear Regulatory Research, Washington, DC 20555-0001

Ouellon T, Lefebvre R, Marcotte D, Boutin A, Blais V, Parent M (2008) Hydraulic conductivity heterogeneity of a local deltaic aquifer system from the kriged 3D distribution of hydrofacies from borehole logs, Valcatier, Canada. J Hydrol 351(1-2): 71-86

Paradis D, Lefebvre R (2013) Single-well interference slug tests to assess the vertical hydraulic conductivity of unconsolidated aquifers. J Hydrol 478(25):102-118

Paradis D, Lefebvre R, Morin RH, Gloaguen E (2011) Permeability profiles in granular aquifers using flowmeters in direct-push wells. Ground Water 49(4):534-547

Paradis D, Lefebvre R, Gloaguen E, Rivera A (in prep) Hydrofacies and hydraulic conductivity from direct-push data using relevant vector machines: motivation, algorithm and application. In preparation for Water Resour Res

Paradis D, Gloaguen E, Lefebvre R, Giroux B (submitted a) Sensitivity and resolution analysis of the information content of tomographic slug tests. Submitted to Water Resour Res September, 2013

Paradis D, Gloaguen E, Lefebvre R, Giroux B (submitted b) A fully-transient analysis of tomographic slug tests in a heterogeneous and anisotropic unconsolidated aquifer. Submitted to Water Resour Res September, 2013

Paradis D, Gloaguen E, Lefebvre R, Rivera A (2011) Tying geophysics to hydrogeology: a learning machine 1138 approach to characterize heterogeneous granular aquifers. Geohydro2011, Joint IAH-CNC, CANQUA 1139 and AHQ conference, Quebec City, Canada, August 28-31

Reading HG (1996) Sedimentary environments: Processes, facies and stratigraphy. 3rd edn. Blackwell Science, Oxford

Reineck HE, Singh IB (1980) Depositional sedimentary environments. 2nd edn. Springer-Verlag, New York

Robertson PK (1990) Soil classification using the cone penetration test. Can Geotech J 27(1):151-158

Ross HC, McElwee CD (2007) Multi-level slug tests to measure 3-D hydraulic conductivity distributions. Water Resour Res 16(1):67-79

Rubin Y, Hubbard S (2005) Hydrogeophysics. Springer, Netherlands

Ruggeri P, Irving J, Gloaguen E, Holliger K (2013) Regional-scale integration of multiresolution hydrological and geophysical data using a two-step Bayesian sequential simulation approach. Geophys J Int 194(1): 289-303

Samui P (2007) Seismic liquefaction potential assessment by using relevance vector machine. Earthq Eng Eng Vib 6(4):331-336

Schmelzbach C, Tronicke J, Dietrich P (2011) Three-dimensional hydrostratigraphic models from ground-penetrating radar and direct-push data. J Hydrol 398(3-4):235-245

Schulmeister M, Butler JJ, Healey J, Zheng L, Wysocki D, McCall G (2003) Direct-push electrical conductivity logging for high-resolution hydrostratigraphic characterization. Ground Water Monit R 23:52-62

Sellwood SM, Healey JM, Birks S, Butler JJJr (2005) Direct-push hydrostratigraphic profiling: Coupling electrical logging and slug tests. Ground Water 43(1):19-29

Shinn JD, Timian DA, Morey RM, Mitchell G, Antle CL, Hull R (1998) Development of a CPT deployed probe for in situ measurement of volumetric soil moisture content and electrical resistivity. Field Anal Chem Tech 2(2):103-110

Smirnoff A, Boisvert E, Paradis SJ (2008) Support vector machine for 3D modelling from sparse geological information of various origins. Comput Geosci 34(2):127-143

Stienstra P, van Deen JK (1994) Field data collection techniques - Unconventional sounding and sampling methods. In: Rengers N (ed) Engineering Geology of Quaternary Sediments. Balkema, pp 41-55

Strebelle S (2002) Conditional simulation of complex geological structures using multiple-point statistics. Mathematical Geology, 34(1):1-21

Tipping ME (2001) Sparse Bayesian learning and the relevance vector machine. J Mach Learn Res 1:211-244

Tremblay L, Lefebvre R, Cloutier V, Molson JW (2012) Value of geochemical data for understanding groundwater flow and natural attenuation of a leachate plume at a former landfill. 39th IAH Congress, Niagara Falls, Canada, September 16-21, Paper 317

Tremblay L, Lefebvre R, Paradis D, Gloaguen E (in press) Conceptual model of leachate migration in a granular aquifer from the integration of detailed characterization data. HJ-2012-2516 accepted for publication on August 3, 2013 by Hydrogeol J

Van Overmeeren RA (1998) Radar facies of unconsolidated sediments in The Netherlands: A radar stratigraphy interpretation method for hydrogeology. J Appl Geophys 40(1-3)

Vitton SJ, Sadler LY (1997) Particle-Size Analysis of Soils Using Laser Light Scattering and X-Ray Absorption Technology. Geotech Test J 20(1):1-11

Zemansky GM, McElwee CD (2005) High-resolution slug testing. Ground Water 43(2):222-230

DEUXIÈME PARTIE : ARTICLES

LISTE DES PUBLICATIONS ORIGINALES ET CONTRIBUTIONS DES AUTEURS

Cette thèse comprend cinq publications originales référencées comme suit:

Article I Paradis, Daniel, René Lefebvre, Roger H. Morin, et Erwan Gloaguen (2011) Permeability profiles in granular aquifers using flowmeters in direct-push wells. *Ground Water*, 49(4): 534-547.

Article II Paradis, Daniel et René Lefebvre (2013) Single-well interference slug tests to assess the vertical hydraulic conductivity of unconsolidated aquifers. *Journal of Hydrology*, 478: 102-118.

Article III Paradis, Daniel, Erwan Gloaguen, René Lefebvre et Bernard Giroux (-) Sensitivity and resolution analysis of the information content of tomographic slug tests. Accepté pour *Water Resources Research* le 25 novembre 2013; Version modifiée re-soumise le 8 mars 2014

Article IV Paradis, Daniel, Erwan Gloaguen, René Lefebvre et Bernard Giroux (-) Transient analysis of tomographic slug tests in a heterogeneous and anisotropic unconsolidated aquifer. En préparation pour *Water Resources Research.*

Article V Paradis, Daniel, René Lefebvre, Erwan Gloaguen, et Alfonso Rivera (-) Predicting hydrofacies and hydraulic conductivity from direct-push data using a data-driven relevance vector machine approach: motivations, algorithms and application. Accepté pour publication pour *Water Resources Research* le 7 mai 2014.

De plus, une version abrégée en anglais de la synthèse présentée à la Partie I a été soumise et acceptée pour publication et peut être consultée à l'Annexe A :

Article VI Paradis, Daniel, Laurie Tremblay, René Lefebvre, Erwan Gloaguen, Alfonso Rivera, Michel Parent, Jean-Marc Ballard, Yves Michaud and Patrick Brunet (2014) Field characterization and data integration to define the hydraulic heterogeneity of a shallow granular aquifer at a sub-watershed scale. *Environmental Earth Sciences*, doi: 10.1007/s12665-014-3318-2

Daniel Paradis a été l'auteur correspondant pour chacun de ces articles ainsi que le principal responsable de leur rédaction. Les co-auteurs des articles ont tous contribué par leurs commentaires aux idées originales présentées dans ces manuscrits et ont aussi activement participé à leur révision.

Les idées de l'adaptation du débitmètre pour les aquifères granulaires et de l'approche de définition des hydrofacies à partir des mesures brutes du débitmètre contenu dans l'article I sont des contributions de René Lefebvre. De plus, René Lefebvre a été le premier à souligner l'importance des mesures de conductivité hydraulique verticale pour la caractérisation hydrogéologique, ce qui a mené à la préparation des articles II, III et IV. René Lefebvre et Erwan Gloaguen sont aussi les instigateurs de l'idée originale de l'article V de

combiner les mesures hydrauliques et géophysiques pour la caractérisation hydrogéologique. Erwan Gloaguen a aussi contribué activement à l'avancement des idées pour le développement de la tomographie hydraulique pour les articles III et IV. Bien que tous les co-auteurs aient apportés une contribution scientifique importante pour la réalisation des articles I à V, Daniel Paradis est le principal responsable des choix méthodologiques et de l'ensemble du matériel contenu dans ces articles.

D'un point de vue technique, Daniel Paradis a bénéficié de la contribution de Jean-Marc Ballard et de René Lefebvre pour le développement de l'approche d'installation des puits par enfoncement qui ont été utilisés pour les tests hydrauliques des articles I-V. Les sondages CPT/SMR pour l'article V ont été réalisés par Jean-Marc Ballard avec la foreuse géotechnique de l'INRS sous la responsabilité de René Lefebvre. Le choix des emplacements des sondages CPT/SMR et des puits d'observation a été fait avec la collaboration de Laurie Tremblay, René Lefebvre et Jean-Marc Ballard. La réalisation des tests hydrauliques avec débitmètre de puits de l'article I a été sous la responsabilité de Roger Morin avec les sondes et le camion de diagraphies de la USGS-Denver. Daniel Paradis a cependant contribué à modifier certaines pratiques pour l'adaptation des mesures avec débitmètre pour les conditions particulières du site à l'étude. L'instrumentation pour les tests hydrauliques des articles I-V a été conçue par Daniel Paradis et les tests hydrauliques ont réalisés par Daniel Paradis ou ont été sous sa supervision.

Les codes numériques utilisés dans les articles II, III et IV pour la modélisation inverse ont été fournis par Geoffrey Bohling et ceux de l'article V pour la classification et la régression par machine d'apprentissage RVM par Mike Tipping. Erwan Gloaguen et Bernard Giroux ont aussi respectivement fournis les codes MATLAB pour la mise à l'échelle des mesures CPT/SMR de l'article V et pour la visualisation des mesures de tomographie des articles III et IV, respectivement. L'ensemble de tous les autres codes MATLAB pour la préparation, l'analyse et la visualisation des données des articles I-V a été développé entièrement par Daniel Paradis.

ARTICLE I

PERMEABILITY PROFILES IN GRANULAR AQUIFERS USING FLOWMETERS IN DIRECT-PUSH WELLS

Résumé

Pour bien comprendre les processus d'écoulement et de transport, la modélisation numérique des aquifères devrait être basée sur une distribution spatiale réaliste de la conductivité hydraulique (K). Cette propriété est cependant rarement définie adéquatement à cause de la difficulté des méthodes de caractérisation en général à acquérir suffisamment d'information. Pour des puits ouvert au roc, des mesures de débit sur de courts intervalles le long d'un puits en pompage à l'aide d'un débitmètre électromagnétique ont démontré leur efficacité pour définir des profils verticaux de K. Peu d'études ont cependant porté sur l'utilisation du débitmètre dans les aquifères granulaires, et les tests réalisés dans ces milieux avec des puits aménagés avec sable-filtrant ont montré que les mesures obtenues sont parfois erronées. Cet article présente une application du débitmètre dans des puits aménagés par enfoncement pour l'obtention de profils de K dans un aquifère granulaire hétérogène ayant des valeurs faibles à moyennes de K (10-6 to 10-4 m/s). L'effet de l'enfoncement des puits sur les mesures de K dans les aquifères granulaires est d'abord étudié et indique que de telles installations minimisent l'erreur sur l'estimation de K. Après l'installation et le développement de puits entièrement crépinés, ceux-ci sont ensuite utilisés pour valider des profils de K obtenus par des mesures de débitmètre à haute résolution (15 cm) avec des profils dérivés d'essai de perméabilité par choc hydraulique sur des intervalles identiques. Pour les 119 intervalles testés provenant de 5 puits différents, la différence dans le logarithme des valeurs de K obtenus par les deux méthodes est systématiquement en-dessous de 10%. Finalement, une approche graphique pour l'interprétation des profils de débitmètre est proposée pour définir des intervalles correspondant à des hydrofaciès distincts, donnant ainsi une méthode pour représenter les contrastes d'échelle et de magnitude de K dans les aquifères granulaires hétérogènes.

Citation: Paradis, Daniel, René Lefebvre, Roger H. Morin, and Erwan Gloaguen. 2011. Permeability profiles in granular aquifers using flowmeters in direct-push wells. *Ground Water*, 49(4): 534-547

Permeability Profiles in Granular Aquifers Using Flowmeters in Direct-Push Wells

Daniel Paradis[1,2*], René Lefebvre[2], Roger H. Morin[3] and Erwan Gloaguen[2]

1: Natural Resources Canada, Geological Survey of Canada, Québec, Canada

2: Institut national de la recherche scientifique, Centre Eau Terre Environnement, Québec, Canada

3: U.S. Geological Survey, Denver, Colorado, USA

* Corresponding author: Daniel.Paradis@RNCan-NRCan.gc.ca

Abstract

Numerical hydrogeological models should ideally be based on the spatial distribution of hydraulic conductivity (K), a property rarely defined on the basis of sufficient data due to lack of efficient characterization methods. Electromagnetic borehole flowmeter measurements during pumping in uncased wells can effectively provide a continuous vertical distribution of K in consolidated rocks. However, relatively few studies have used the flowmeter in screened wells penetrating unconsolidated aquifers, and tests conducted in gravel-packed wells have shown that flowmeter data may yield misleading results. This paper describes the practical application of flowmeter profiles in direct-push wells to measure K and delineate hydrofacies in heterogeneous unconsolidated aquifers having low-to-moderate K (10^{-6} to 10^{-4} m/s). The effect of direct-push well installation on K measurements in unconsolidated deposits is first assessed based on previous work indicating that such installations minimize disturbance to the aquifer fabric. The installation and development of long-screen wells are then used in a case study validating K profiles from flowmeter tests at high resolution intervals (15 cm) with K profiles derived from multi-level slug tests between packers at identical intervals. For 119 intervals tested in 5 different wells, the difference in log K values obtained from the two methods is consistently below 10%. Finally, a graphical approach to the interpretation of flowmeter profiles is proposed to delineate intervals corresponding to distinct hydrofacies, thus providing a method whereby both the scale and magnitude of K contrasts in heterogeneous unconsolidated aquifers may be represented.

Keywords: Borehole Flowmeter, Direct-Push Wells, Hydraulic Conductivity, Unconsolidated Aquifers, Heterogeneity.

1 Introduction

It is generally recognized that the hydraulic conductivity (K) heterogeneity of an aquifer controls groundwater flow and solute transport. Failure to adequately account for this heterogeneity is also known to be one of the main causes of remediation failures (de Marsily *et al.*, 2005). In heterogeneous aquifers, the spatial distribution of hydraulic properties may be complex and numerous direct and indirect estimates may be required to develop a realistic understanding of the hydrogeological system (Ouellon *et al.*, 2008). Although the spatial definition of aquifer heterogeneity has to rely in large part on the use of indirect data, often acquired from surface geophysical surveys (Rubin and Hubbard, 2005), numerous direct measurements are also required to define the properties of individual lithologic units (hydrofacies) having a distinctive range of hydraulic conductivities (Frei *et al.*, 2009). Therefore, a practical means of adequately quantifying the spatial distribution of hydraulic properties under the constraints of a limited amount of time and resources is essential as part of a hydrologic characterization program aimed at identifying and defining aquifer K heterogeneity.

Borehole flowmeter measurements obtained during pumping can efficiently determine the vertical distribution of K in fractured rocks (Morin *et al.*, 1988; Molz *et al.*, 1989; Hess *et al.*, 1992; Hanson and Nishikawa, 1996; Paillet, 1998; Crisman *et al.*, 2001). However, previous studies have suggested that head losses along the flowmeter can have important and detrimental effects on its use to determine K profiles (Boman *et al.*, 1997; Dinwiddie *et al.*, 1999; Ruud *et al.*, 1999). These head losses are exacerbated when a packer or diverter is used to force water in a well through the flowmeter. The major effects of head loss are: (1) possible bypass flow around the packer if the well is gravel packed (Dinwiddie *et al.*, 1999); (2) flow redistribution with meter position (Dinwiddie *et al.*, 1999), and (3) hydraulic isolation in the vicinity of the pump caused by flow behind casing and through the gravel pack above and below the pump in an unconfined aquifer (Boman *et al.*, 1997).

In an attempt to evaluate and reduce head losses, Arnold and Molz (2000) showed that the use of a 2.54-cm diameter electromagnetic (EM) flowmeter reduced head losses by a factor of 16 when compared with a 1.27-cm diameter probe. Ruud *et al.* (1999) and Arnold and Molz (2000) also proposed the use of an EM flowmeter without a packer or flow diverter (centralized configuration). In this adaptation, flow is allowed to bypass the flowmeter along the well screen. Clemo (2010) found that changes in inflow to the wellbore caused by blank casing joints could also affect the ratio of flow through and around the flowmeter.

Mathematical corrections that account for head losses have been proposed by Zlotnik and Zurbuchen (2003) for a packer configuration, and by Clemo (2010) for a centralized configuration. However, further refinement is needed in both cases to develop a practical model that is reliable for field use under a variety of flow conditions and casing sizes.

In consolidated aquifers, flowmeter measurements are generally conducted in open boreholes; in unconsolidated aquifers, these measurements have to be made in screened wells (Morin, 2006). Installation of screened wells with conventional methods (i.e. rotary or auger drilling) involves filling the annulus between the screen and borehole wall with a sand or gravel pack. Because a sand or gravel pack is comprised of material different from the adjacent geological unit, its presence causes either a reduction or an increase in K surrounding the screened interval of a well compared to that of the host material. For example, Young (1998) recommended the use of gravel packs rather than back-fill material for wells installed in heterogeneous granular aquifers to avoid a reduction in K around the most permeable material of the formation. However, it has been recognized by many authors (Rehfeldt *et al.*, 1992; Boman *et al.*, 1997; Dinwiddie *et al.*, 1999; Ruud *et al.*, 1999) that flowmeter measurements conducted in gravel-packed wells may yield misleading results due to the combined effect of head losses produced by the flowmeter and an annulus of high-permeability material surrounding the screen that allows some flow to bypass the instrument. Partly due to the potential impact of the gravel pack typically used in conventional well installation, flowmeter applications in unconsolidated aquifers have been restricted mainly to the characterization of dominant flow pathways (Young, 1995) or coarse-grained materials (i.e. Hess *et al.*, 1992; Morin, 2006; Li *et al.*, 2008).

The objective of this study is to effectively extend the application of the borehole flowmeter to the characterization of granular aquifers. Continuous, high vertical-resolution K profiles are generated that are deemed better suited to the description of K heterogeneity than are the point measurements provided by slug tests, the most common source of K data (Butler, 1998). As discussed in this paper, direct-push well installation without sand pack was chosen as a means of 1) obtaining continuous long screens, 2) minimizing head losses by selecting a well diameter only slightly larger than the flowmeter to ensure a tight fit and eliminate the necessity for a packer or a diverter, 3) avoiding flow bypass related to sand packs, and 4) minimizing the disturbance to the surrounding formation. This study also investigates the validity of flowmeter measurements by comparison against slug tests.

The selected study area is composed of an unconsolidated heterogeneous aquifer exhibiting a low-to-moderate K range (10^{-6} to 10^{-4} m/s); this type of hydrogeologic environment has not commonly been characterized with borehole flowmeters. High vertical resolution (15 cm) K profiles derived from both flowmeter measurements and multi-level slug tests between packers provided 119 K measurements over duplicate intervals in 5 different wells. In the study area of St-Lambert (Figure I.1a), Quebec, Canada, these 5 wells and their tested intervals were selected to provide a representative sampling of K across the range of hydrofacies determined to be typical for this region based on data profiles from Cone Penetration Tests (CPTu) (Lunne *et al.*, 1997) and Soil Moisture Resistivity (SMR) measurements (Shinn *et al.*, 1998). The potential impact of skin effects related to direct-push well installation was assessed in the field using slug interference tests. Finally, a graphical interpretation method using flowmeter profiles is introduced for the purpose of defining K values across depth intervals corresponding to distinct hydrofacies. This method also proved to be useful in enhancing the precision of K values determined across low-permeability intervals where individual flowmeter measurements may not have the resolution to quantitatively resolve the differential flow and the corresponding K value from an adjacent interval.

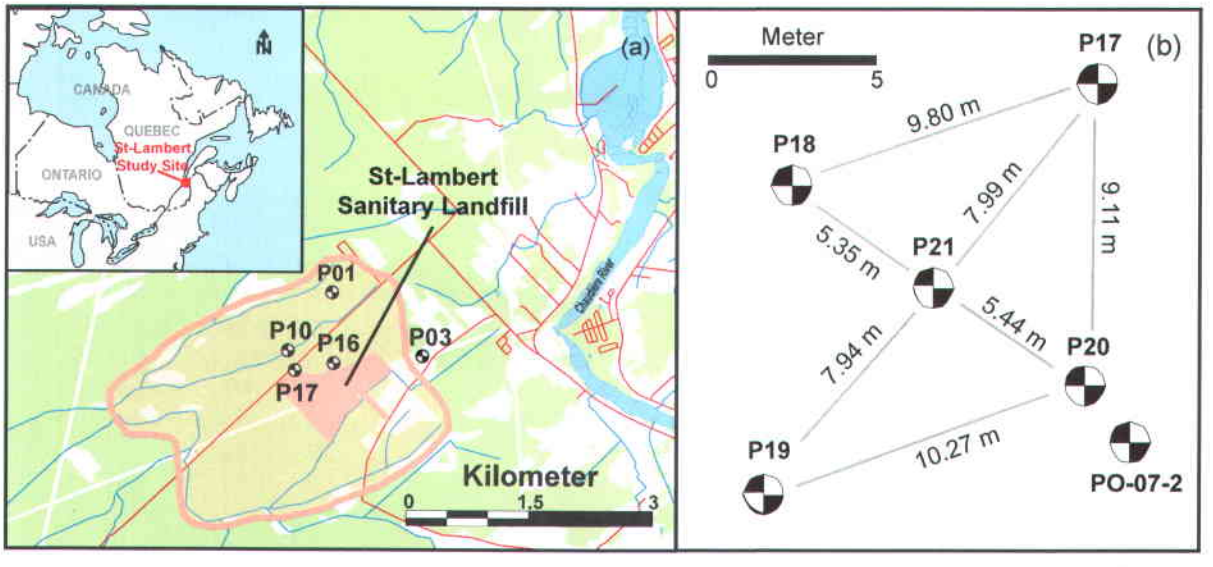


Figure I.1 (a) Location of the St-Lambert study area. The boundary of the study area corresponds to the sub-watershed surrounding the St-Lambert sanitary landfill. Only the locations of the wells used in this study are shown (P01, P03, P10, P16 and P17). (b) Location of well P17 and nearby observation wells that were used for multi-well pumping and interference tests.

2 Methods

2.1 Site Description

The study area is located 40 km south of Quebec City, Canada and encompasses a 12 km^2 sub-watershed surrounding a sanitary landfill (Figure I.1a). The aquifer is composed primarily of surficial Quaternary sediments 5-to-20-m thick that were deposited and reworked by the Champlain Sea (Bolduc, 2003). These sediments consist mostly of fine-to-medium sand, but sediments can range from coarse sand to clayey silt. Based on regional geological data and more than 25 CPTu/SMR soundings, the hydrostratigraphy of the aquifer is controlled by the distinctive structure of a spit that formed in a littoral environment. This feature is defined as a narrow finger-like ridge of sand formed by the longshore movement of sediment at the mouths of estuaries that extends from land into open water. This depositional environment leads to the presence of a wide range of sediment types.

Vertical profiles of K and hydraulic head measured as part of this study indicate generally semi-confined conditions, the result of alternating sand and silt layers related to the formation of the spit (Figure I.2). The water table is 1 to 2 m below ground surface and, from bottom to top, sediments are associated with three distinctive chronostratigraphic phases (Figure I.2). The first phase is the deposition of marine clayey silt, which occurred when the Champlain Sea level was at an elevation of approximately 180 to 200 m asl (above sea level); the present elevation of the study area is approximately 120 m asl. The second phase corresponds to the formation of the spit itself, when sea level was roughly at the present-day elevation of the site. In this sequence, sediments coarsen upward (silt to sand) to a maximum and then reverse and fine upward (sand to silt). This general sedimentary sequence was observed in all CPTu/SMR soundings of the study area. Finally, phase 3 is represented by a spatially discontinuous clayey silt unit at the top of the sediment sequence, which is related to the eventual retreat of the Champlain Sea.

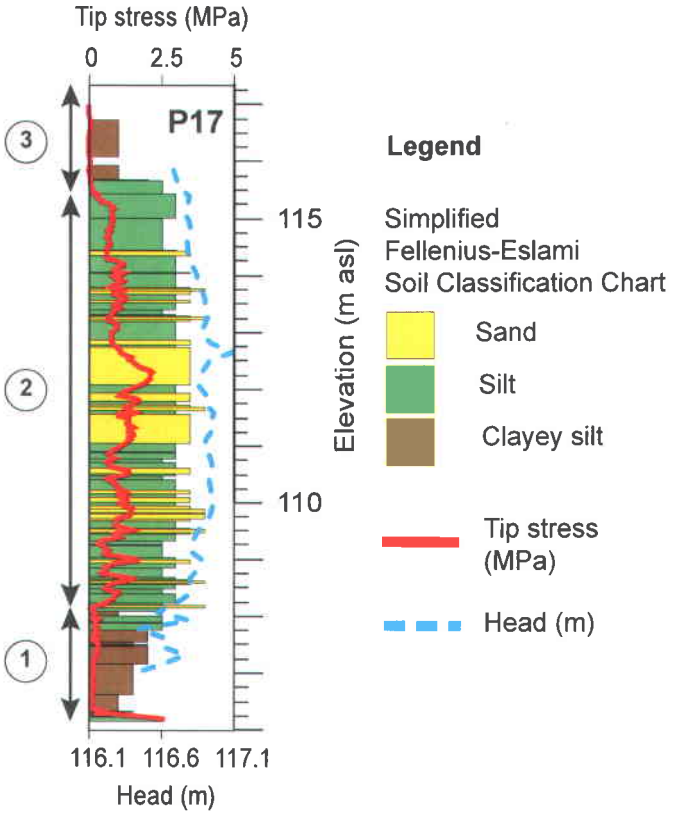


Figure I.2 Hydrostratigraphy of the St-Lambert site as illustrated by the well P17 cone penetration test (CPT). Soil textures are defined according to the Fellenius and Eslami (2000) CPT soil classification chart. Also shown are tip stress measured with the CPT, and static hydraulic head recorded prior to slug tests conducted between packers. Numbers to the left refer to sediments deposited during three chronostratigraphic phases described in the text.

2.2 Direct-push well installation and development

Direct-push monitoring well installation is a rapid and economical procedure. Like conventional wells, these wells provide representative groundwater samples and a means of obtaining geophysical and hydrogeological measurements (USEPA, 2005). Direct-push installations also offer advantages over conventional installations in unconsolidated formations: they reduce disturbance to the surrounding formation, eliminate drill cuttings, and avoid gravel-packs so that sediments are in direct contact with the screen. In this study, the wells used for K measurements were installed with a direct-push rig (Geotech 605D) following the protected screen standard technique (ASTM, 2004). Figure I.S1a illustrates the installation steps and Figure I.S1b depicts the tubing cross-sectional dimensions. A 76-mm (3-in OD, outside diameter) metal casing with an expendable bottom point is first hammered into the ground to the desired depth. Then 52-mm (2-in ID, inside diameter) (60-mm OD) fully-screened PVC tubing is inserted inside the metal casing before the outer metal casing is withdrawn. The width of the annulus between the

OD of the metal casing (corresponding to the borehole diameter) and the OD of the PVC screen is approximately 8 mm. The slot aperture in the PVC screen was selected to be 0.024 cm (0.001 in) based on the lower range of sediment particle-size distributions sampled in the study area. The inside diameter of the PVC tubing (52-mm ID) was selected to allow a tight fit between it and the flowmeter, thus eliminating the need for a flow diverter or packer. Special care was taken when installing the wells to ensure they were vertical and straight. Because of the small size tolerance between casing and tool, a slight bend in the casing or deviation from vertical was found to prevent the flowmeter from entering the well. Inclination sensors on the direct-push rig facilitated the installation of vertical wells.

In addition to proper installation, it is critical that direct-push wells be developed thoroughly to obtain representative measurements of hydraulic properties (Butler, 1998; Henebry and Robbins, 2000). For example, a poorly developed well can underestimate K due to the presence of fine sediments clogging the screen or formation. Consequently, an aggressive development program is necessary to correct for the inherent disturbance that occurs during the installation process (e.g. sediments smearing). In this study, the development was realised by pumping-surging operations performed with an inertial pump (Waterra) equipped with a foot valve and a surge block having a diameter slightly smaller than the inside diameter of the screen tubing. The foot valve was moved systematically along the screen at 0.5 m intervals and each interval was pumped until turbidity was no longer observed in the discharged water. The evolution of K at different stages of well development was also monitored for selected wells and intervals to assess the effectiveness of well development operations (Figure I.S2a and 1.S2b).

2.3 Assessment of skin effects from interference tests

Drilling operations and/or well construction can change the K of the material surrounding a borehole and produce either a reduction (low-K skin) or an increase (high-K skin) in K. However, the direct-push well installation method offers a potential advantage over conventional well-installation methods in unconsolidated aquifers because it may reduce disturbance to the formation and eliminate drill cuttings and gravel-pack. Nevertheless, it is important to both assess the representativeness of K values measured in direct-push wells and determine whether a skin effect from sediment deformation exists.

For the identification of well skin, Butler and Healey (1996) proposed a hydraulic experiment where repeat slug tests with varying effective radius of the well casing are performed in conjunction with a ratio method

derived from Hvorslev (1951) technique. However, this approach only distinguishes low-K skin when the difference in skin to formation K exceeds more than an order of magnitude.

To evaluate skin produced by direct-push well installation, interference slug tests were performed between closely spaced direct-push wells (see Figure I.1b for well locations). Results were examined using the analytical solution of Hyder *et al.* (1994), which yields estimates of both the aquifer and the skin material K's, as well as the specific storage (S_s). For these tests, one well was used to induce hydraulic perturbation and the other was used as an observation well. Slug tests were performed using a pneumatic method and hydraulic responses were monitored in both wells. Two kinds of interference tests were conducted: (1) on fully screened wells, and (2) between isolated intervals (60.96 cm) with injection and observation intervals at the same vertical position.

2.4 Electromagnetic borehole flowmeter

The borehole flowmeter measures vertical flow in a well at various depths to determine the inflow to the well across specific intervals. This instrument has been used in numerous studies to identify permeable intervals and measure the vertical distribution of horizontal K surrounding the well (Hess, 1986; Morin *et al.*, 1988; Molz *et al.*, 1989; Rehfeldt *et al.*, 1992; Paillet, 1998). Data analysis presumes the aquifer to be perfectly layered so that the K of each tested interval is proportional to the measured additional flow coming from that interval. The product of the interval thickness and its K value is a fraction of the average transmissivity (T) of the entire well (Molz *et al.*, 1989).

For this study, the vertical flow measurements at each station were made with an electromagnetic (EM) flowmeter (Century Geophysical Corp.; Model 9721). During testing, the perturbation created by the borehole fluid passing across a magnetic field generated inside the probe is proportional to the average vertical velocity of the water (Young and Pearson, 1995). The flowmeter is 1.4 m long with a sensor-housing outside diameter (OD) of 51 mm, and the measurement barrel dimensions are 42 mm OD and 25.4 mm inside diameter (ID) with a trumpet-shaped entryway. Among the different types of flowmeters available, the EM flowmeter was chosen for its low detection limit, its large dynamic range of measurement (0.05 to 40 L/min), and its relative ease of operation. Impeller flowmeters or heat-pulse flowmeters may prove to be more appropriate for other types of applications.

Generally, when conducting flowmeter tests in large-diameter wells, a diverter or a packer is attached to the flowmeter to force all flow through the instrument's sensor. This arrangement may produce a sharp change in fluid velocity that creates head losses and affects accuracy (Boman *et al.* 1997; Dinwiddie *et al.* 1999; Ruud *et al.* 1999). To avoid this situation, the diameter of the slotted PVC tubing used for the wells (52 mm ID) was deliberately chosen to allow the flowmeter to fit inside it with only a small clearance to minimize flow bypassing the tool and head losses due to velocity disturbances.

Applying the flow nozzle equation (Foley, 1997) to the flowmeter wellbore hydraulics, the head losses are proportional to the difference between the inverse of the fourth power of the radius of the flowmeter and the inverse of the fourth power of the radius of the well. Thus, for a 25.4-mm ID flowmeter used inside a 52-mm ID well, the head loss across the flowmeter associated with the maximum flow rate employed in this study (19 L/min) is approximately 1.9 mm. Therefore, we assume head loss to be negligible, with fluid making a smooth and laminar transition as it flows through the sensor. Probe calibration was also performed in the field using the same PVC tubing; the correlation between volumetric flow measured at the surface and that measured downhole with the flowmeter was accurate and linear across the range of pumping rates used for this study.

Before pumping each well, measurements were made to confirm the state of ambient flow. Within the resolution specifications of the EM probe, no evidence of ambient flow was found at the test site. Flowmeter measurements during pumping were made every 15.24 cm (6 in). For brevity, these intervals are heretofore referred to simply as being 15 cm long. Pumping was performed using a centrifugal pump at volumetric flow rates ranging between 4 to 19 L/min. The testing strategy involved selecting a pumping rate sufficient to produce measureable flow above the detection limit of the tool, while also minimizing drawdown in the well to preserve a saturated upper section of the aquifer and retain a longer interval accessible to flowmeter measurements. Flow measurements were obtained once steady-state conditions were approached. Pumping rate and hydraulic head were continuously monitored to ensure stable conditions throughout the entire flowmeter operation. Also, pumping rates and, consequently, borehole fluid velocities were slow enough to avoid turbulent conditions. Running flowmeter tests in a 10-m long screen at 15-cm intervals took approximately 4 hours.

2.5 Multilevel slug tests

When a slug test is conducted between packers that isolate (or straddle) a particular screened interval of a well, it is usually referred to as a multilevel slug test (e.g., Sellwood *et al.*, 2005; Zemansky and McElwee, 2005; Ross and McElwee, 2007). For this study, inflatable packers were fabricated over 25.4-mm ID PVC tubing, and threads on the tubing allowed the use of variable screen lengths between packers. An air line connected to an air compressor at the surface was attached to the packers to inflate them to the desired pressure. The dual-packer assembly was also connected by a 25.4-mm ID PVC riser pipe to the surface, within which water levels were monitored during the slug test. A rigid tape was attached near the top to accurately locate the depth of the straddled interval.

Slug tests were performed using a pneumatic method to induce an initial lowering of the water level (Levy and Pannell, 1991). For this purpose, a wellhead assembly was attached to the top of the riser pipe. This device contained an airtight adapter that allowed a transducer cable to pass and a ball valve to rapidly release pressure. An air compressor was also connected to the wellhead assembly to increase air pressure in the riser. A precision digital air-pressure gauge was used to accurately set the desired initial hydraulic head change for the slug test and to verify air pressure stabilization before initiating the slug test.

Multilevel slug tests were conducted at 15-cm intervals located at elevations coincident with the intervals tested by the EM flowmeter. Varying and successive head changes were imposed at some intervals for quality control and to verify data repeatability (Butler *et al.*, 1996). Because tested wells were extensively developed, intervals were randomly selected for quality control. Slug tests performed in a 10-m long screen at 15-cm intervals took approximately 40 hours. A total of 227 K estimates were obtained from multilevel slug tests.

The K values were estimated from slug tests using the interpretation method of Bouwer and Rice (1976), a method applicable to data collected in partially or completely penetrating wells under either unconfined or confined conditions. The ratio of the screen length to the screen radius used in this study is within the range of geometric conditions allowed by the method. This Bouwer and Rice (1976) straight-line method is also valid within the recommended head change range (0.20 to 0.30 m) proposed by Butler (1996) to produce reasonable estimates of K for varying values of aquifer storage.

3 Results and discussion

3.1 Potential impact of skin effects

To assess the potential impact of skin effects on estimates of K, interference slug tests were performed between fully screened direct-push wells (Figure I.1b) and between isolated intervals. Figure I.S3 illustrates an example of the curve-fitting exercise used to estimate aquifer and skin hydraulic parameters. Initially, aquifer hydraulic properties were adjusted based on the observation well response (both aquifer and skin hydraulic parameters being the same). Proper matching of the injection well response then defines the K skin parameters. In this example, S_s values used for aquifer and skin were the same. Observation well response is less sensitive than that of the injection well to skin properties, although both well responses are sensitive to aquifer properties. This distinction improved our ability to match both the aquifer and the skin hydraulic properties.

Table I.1 provides results of adjusted aquifer and skin hydraulic properties applying the solution of Hyder *et al.* (1994) for different pairs of wells or intervals. For all the tested wells and intervals, the skin is systematically slightly more permeable than the aquifer, and the aquifer-skin K ratio ranges from 0.36 to 0.86. For the interference tests between intervals, a good match between theoretical and observed curves is achieved if S_s is assumed to be constant. However, this assumption yields less satisfactory results for interference tests conducted on fully screened wells and S_s aquifer values were increased from initial estimates. This discrepancy may be explained by the fact that no packer was used in the observation wells. Although the diameter of the wells is small (2.54 cm), wellbore storage might have affected the response in the observation well. Physically, wellbore storage in the observation well will create a delay and a reduction in the amplitude of the hydraulic response that may be interpreted as a high S_s value for the aquifer. Nevertheless, a small increase in K around wells may be observed from both kinds of interference tests.

These results are consistent with those of Ngan-Tillard *et al.* (2005) who applied high spatial resolution X-Ray computed tomography scans to visualise the local deformation patterns caused by cone penetration in calibration chamber filled in with fine grained sand. Tomography results showed that a dilating zone having the shape of a sphere was formed ahead of the cone tip before being penetrated by the cone. Friction along the cone shaft did not generate visible density changes. The final dilating zone consisted of several coaxial cylinders of different density, all having a rounded tip. The dilated zone in direct contact with the cone has

an external diameter of up to 2.6 times the cone diameter. It corresponds to an average increase in porosity of 4.4%. Within this zone, a thin zone characterised by a very high dilation can be discerned. Its increase in porosity is 7.2%. Around these zones, a zone affected by a light dilation (2.3% porosity change) is found and its boundaries are approximately 5 times the cone diameter.

If K is assumed to be related to sand porosity, then an increase in porosity of 5% for fine sand may result in a K increase of approximately 15 to 50% [refer to Figure I.8 of Koltermann and Gorelick (1995) for fine sand mixed with clay or silt at 0% of weight fraction of fines]. That porosity change corresponds roughly to the average porosity increase observed by Ngan-Tillard *et al.* (2005) for the dilated zone. For example, for a K value of 7×10^{-6} m/s for natural aquifer material, the anticipated high-K skin created around the screen with a direct-push well could range from 8×10^{-6} to 1×10^{-5} m/s. Consequently, this change in K around the direct-push well is small compared to the K of the gravel pack (in the range of 1×10^{-3} m/s to 1×10^{-2} m/s) that is generally used with conventional wells (Driscoll, 1986). Whether this increase is the result of dilating material during direct-push well installation or well development operations cannot be resolved. Nevertheless, the impact of skin effects on K measurements in direct-push well is expected to be small. Ngan-Tillard *et al.* (2005) also shown that increasing sand density results in a significant increase in the extent of the dilation zone and the change in porosity that occurs around the cone. Then both the extent and magnitude of the high-K skin around a pushed tubing can be expected to change following sediment density.

Table I.1 Hydraulic properties of aquifer and skin estimated from interference slug tests performed on wells and isolated intervals (wells shown on Figure I.1b). Screen length for well P17 is 10.61 m and is 7.62 m for all other wells. Interval lengths for injection and observation wells are 61 cm and 30 cm, respectively. When present, number following well name refers to the depth of the top of the screen interval.

Injection Well	Observation Well	Aquifer		Skin		Ratio K_a/K_s (-)
		K_a (m/s)	Ss_a (/m)	K_s (m/s)	Ss_s (/m)	
Interval						
P21:3.97	P17:4.43	6.38×10^{-6}	2.03×10^{-5}	1.11×10^{-5}	2.03×10^{-5}	0.57
P21:4.58	P17:4.43	3.92×10^{-6}	2.21×10^{-5}	9.41×10^{-6}	2.21×10^{-5}	0.83
P21:5.80	P17:5.95	8.35×10^{-6}	1.79×10^{-5}	1.10×10^{-5}	1.79×10^{-5}	0.76
P21:6.41	P17:6.71	8.74×10^{-6}	1.83×10^{-5}	2.40×10^{-5}	1.83×10^{-5}	0.36
P21:7.02	P17:7.47	9.58×10^{-6}	2.50×10^{-5}	1.74×10^{-5}	2.50×10^{-5}	0.55
P21:7.63	P17:7.47	5.03×10^{-6}	1.21×10^{-5}	1.20×10^{-5}	1.21×10^{-5}	0.42
Average		7.00×10^{-6}	1.93×10^{-5}	1.42×10^{-5}	1.93×10^{-5}	0.58
Full Screen						
P21	P17	7.77×10^{-6}	2.75×10^{-5}	1.17×10^{-5}	6.21×10^{-6}	0.66
P21	P18	6.76×10^{-6}	3.85×10^{-5}	1.07×10^{-5}	1.10×10^{-5}	0.63
P21	P19	7.42×10^{-6}	3.08×10^{-5}	1.07×10^{-5}	1.10×10^{-5}	0.69
P21	P20	6.46×10^{-6}	3.78×10^{-5}	1.80×10^{-5}	6.21×10^{-6}	0.36
Average		7.10×10^{-6}	3.37×10^{-5}	1.28×10^{-5}	8.61×10^{-6}	0.59
P18	P17	9.76×10^{-6}	3.53×10^{-5}	1.14×10^{-5}	1.80×10^{-5}	0.86
P18	P18	9.76×10^{-6}	2.33×10^{-5}	1.14×10^{-5}	1.01×10^{-5}	0.86
P18	P19	9.76×10^{-6}	3.19×10^{-5}	1.14×10^{-5}	1.01×10^{-5}	0.86
P18	P21	9.02×10^{-6}	4.82×10^{-5}	1.14×10^{-5}	1.60×10^{-5}	0.79
Average		9.58×10^{-6}	3.47×10^{-5}	1.14×10^{-5}	1.36×10^{-5}	0.84

To further assess the impact of the high-K skin on K aquifer estimates, results obtained from various analytical methods applied to the response of a single well are compared to those obtained when an observation well is available to estimate skin hydraulic properties. Table I.S1 and Figure I.S4 display values of K estimated from conventional slug-test interpretation methods (Hvorslev, 1951; Bouwer and Rice, 1976) as well as the method of Hyder et al. (1994) without and with skin taken into account. Neglecting high-K skin results in an average overestimation of aquifer K of approximately 1.5 to 1.8. This overestimate is in general agreement with the work of Hyder et al. (1994) and indicates that, for an aquifer/skin ratio of 0.15 to 0.5 (a K increase of 15 to 50%), the resulting K obtained by a slug test with the Hvorslev (1951) method will be slightly overestimated [see Figure I.7a of Hyder et al. (1994) for a skin radius twice the radius of the well]. Interestingly, these authors also report that, for an aquifer/skin ratio of 5 (instead of 0.5), the resulting K estimate without an observation well would be roughly 10 times less (instead of slightly >1). Thus, the formation of a high-K skin around a well has less impact on the estimate of K than a low-K skin having a similar magnitude of aquifer/skin K ratio.

3.2 Estimation of average *K* from flowmeter data analysis

To estimate a K profile along a well, flowmeter data analysis requires a value of the average K (or total transmissivity) over the entire well. This value is usually obtained from an independent hydraulic test, such as a slug test or a single well pumping test (drawdown or recovery interpretation). However, it is also possible to obtain the average K value of the well concurrently with flowmeter operations, as the general procedure for flowmeter measurements requires steady-state (or quasi steady-state) conditions to be established prior to obtaining flow data. During initial transient conditions, pumping rate and drawdown records can be used to estimate the average K of the well, instead of relying on an independent hydraulic test. Moreover, the availability of individual and average K values estimated at the same time and under the same conditions is more likely to provide a representative K profile from flowmeter measurements than the use of an independent test performed under different conditions (e.g., variations in the total saturated thickness of the aquifer).

Young (1998) suggested the Cooper-Jacob straight-line (CJSL) method (Kruseman and de Ridder, 1990) be used to estimate the average K for proper analysis of flowmeter data. However, in a heterogeneous aquifer having low-to-moderate K, it may be difficult to estimate this property from short-duration pumping tests. To illustrate this point, Figure I.S5a presents the drawdown and temperature measured during a flowmeter test in well P18. Well P18 is located close to well P17 (Figure I.1b) and was used only to estimate average

K from flowmeter data. Flowmeter results from well P18 were not included in the comparative analysis with slug test data due to their similarity with those from adjacent well P17. Matching the Theis solution to early drawdown measurements suggests the presence of leakage because observed late drawdowns are lower than those predicted by the Theis solution and the value of S_s required to yield a reasonable match is greater than 0.7. This response indicates that more water is reaching the well cone of influence than is predicted for a confined aquifer. A generally improved match to observed drawdowns is obtained with the Moench (1985) solution for leakage from an aquifer top. Hydraulic conductivity estimated by this method for well P18 is 1.3×10^{-5} m/s (see Table I.2), a magnitude consistent with K values averaging from 9.58×10^{-6} to 1.14×10^{-5} m/s determined by interference tests (see Table I.1). Similar indications of leakage are observed for all the wells tested throughout the study area (not shown). As no stream is within a radius of less than 100 m from any tested well, leakage originating from this source cannot explain the observed behaviour.

Table I.2 Aquifer parameters derived from the interpretation of drawdown data measured in well P18 (Figure I.S5) during pumping related to flowmeter tests. Results are interpreted using the solutions of Moench (1985), Theis, and Cooper-Jacob (Kruseman and de Ridder, 1990), where T is transmissivity, K is hydraulic conductivity, S_s is specific storage, and parameters B' and b' are related to leakage properties and follow the definition of Hantush (1960).

Type Curve	T (m^2/s)	K (m/s)	S (-)	$1/B'$ (m^{-1})	β'/r (m^{-1})
Moench (1985) (Leaky) Case 1-Top	9.8×10^{-5}	1.3×10^{-5}	8.3×10^{-4}	0.0086	0.23
Theis (Confined) Early Data Fit	5.3×10^{-5}	7.0×10^{-6}	0.74	-	-
Theis (Confined) Same parameters as Moench	9.8×10^{-5}	1.3×10^{-5}	8.3×10^{-4}	-	-
Cooper-Jacob straight line (CJSL)	1.5×10^{-4}	2.1×10^{-5}	3.5×10^{-5}	-	-

Temperatures of pumped water recorded as a function of time during flowmeter testing (Figure I.S5a) and the static temperature profile measured before pumping (Figure I.S5b) further support the occurrence of aquifer leakage. Pumped water temperature measured at well P18 during the flowmeter test increased from 9.4 to 10.6 °C (Figure I.S5a). This rise is consistent with leakage from the aquifer top because the

temperature profile measured several days before pumping showed that the top of the aquifer was warmer (approaching 15 °C) than at the depth where drawdown and temperature were directly measured (Figure I.S5b). This suggests that leakage originating from the warm sand at the aquifer top gradually infiltrated through the underlying silt layer, thus leading to the observed temperature increase at the logger position. Temperature data indicate that steady-state conditions that balance flow from the aquifer with leakage from the top are eventually reached at the end of the test.

In the case of a heterogeneous aquifer, such as the one being considered in this study, choosing the correct analytical model is critical to the estimation of average K. For the example above, the application of the CJSL method would result in an overestimation of K (2.1×10^{-5} instead of 1.3×10^{-5} m/s, see Table I.2). Applying an average K based on an incorrect analytical and conceptual model can result in the correct relative vertical distribution of K but the incorrect magnitude of K (Hanson and Nishikawa, 1996). Aquifer conditions should thus be established to ensure the selection of an appropriate well-test solution. In addition, monitoring the temporal change in water temperature during pumping may provide an effective diagnostic indicator of leaky conditions in shallow heterogeneous aquifers.

A possible cause for concern in this analysis is the validity of the assumption of quasi-horizontal flow when using the interpretation method of Molz *et al.* (1989) with leaky conditions. As discussed by Molz and Melville (1996), the assumption of strictly horizontal flow applies only in the immediate vicinity of the well, not throughout the aquifer. Aquifer leakage, which is assumed by Moench (1985) to occur mainly in the vertical direction, does not disturb the predominantly horizontal flow in evidence throughout the tested aquifer even though it does affect the rate and distribution of drawdown.

3.3 Comparison of flowmeter and slug-test data

To demonstrate the usefulness of the EM flowmeter for characterizing K variability in unconsolidated aquifers, high-resolution profiles of K obtained from EM flowmeter tests were compared with corresponding profiles determined from multilevel slug tests. Five wells were selected for this purpose and a total of 123 intervals having a constant vertical resolution of 15 cm were tested by both methods (Table I.3). Wells were selected such that tested intervals represented a wide distribution of K. For the flowmeter method, the tested interval corresponded to the distance between two successive vertical displacements of the probe; for the multilevel slug method, the tested interval was the inner distance between the ends of the two packers inflated against the screen. To avoid a systematic bias in the comparison of discrete K values

derived from each method separately (next section), the average K value used in the analysis was the one determined from slug tests.

Table I.3 Well descriptions and average aquifer parameters used for comparison of flowmeter and multilevel slug tests.

Data Set: 15- cm Intervals with K Measurements from Both Methods			Multilevel Slug Tests	
Well	Number of Intervals	Total Length (m)	T (m^2/sec)	K (m/sec)
P01	28	4.42	2.5×10^{-4}	5.7×10^{-5}
P03	41	6.55	3.6×10^{-4}	5.5×10^{-5}
P10	16	2.59	2.2×10^{-5}	8.4×10^{-6}
P16	12	1.98	6.8×10^{-5}	3.4×10^{-5}
P17	22	4.27	3.3×10^{-5}	7.7×10^{-6}

Descriptive statistics of the logarithm of K (log K) measured by multilevel slug tests and by EM flowmeter tests are presented in Table I.4. The table lists data obtained from all intervals tested over the study area and estimated by both methods, as well as the overlapping intervals used for direct comparison of results. For the log K distribution that includes all data, the range of log K values obtained from slug tests is wider than that from the flowmeter tests. The maximum values in the range are approximately the same for both methods, whereas the minimum value is higher for flowmeter tests. This lower limit for the flowmeter is restricted by the resolution of the tool used for this study under the pumping rates employed; it corresponds to a minimum detectable value for K of approximately 1.6×10^{-6} m/s for a 15-cm measurement interval. Multilevel slug tests are not constrained by such a limitation. However, slug test duration (corresponding to drawdown recovery) with the configuration used in this study could persist for over 20 min when K values were as low as 1.6×10^{-6} m/s. A flowmeter interpretation scheme is proposed later to provide representative K values for low permeability intervals that vertically span numerous 15-cm flowmeter measurements. For the data subset used for comparison of methods, the mean values of log K for both techniques are slightly different (Table I.4), even though the average K value for the entire well (derived from multilevel slug tests) was used to determine interval K values from flowmeter data analysis. The difference in the mean is related

to the occurrence of null measurements (below detection limit) for the flowmeter, which were not used in the comparison.

Figure I.3a shows the differences between log K values derived from the multilevel slug tests and the EM flowmeter tests against their mean. According to Bland and Altman (1999), this type of graph is an effective way of displaying data when: (1) the range of variation in the measurements is large in comparison with the differences between the methods, and (2) there is an increase in variability of the differences as the magnitude of the values increases. Since slug test and flowmeter measurements are both subject to uncertainty, the mean K values computed from the two methods are used as a best estimate. Figure I.3a and Table I.4 are used to assess the "equivalent" nature of slug tests and flowmeter measurements.

For 123 intervals tested in five wells, the mean difference in log K is 0.0077 with 95% limits of agreement between 0.44 and -0.43 as shown in Figure I.3a. Thus flowmeter and slug estimates of K are similar on average despite computing an average K that includes the sections with null measurements. The differences between log K values obtained from flowmeter and slug tests also form a normal distribution, as revealed by a kurtosis-skewness distance probability equal to 0.58 (Table I.4). The systematic error can also be determined by the rank correlation between the absolute differences and the average in Figure I.4a. The Spearman rank correlation coefficient is 0.14, meaning that the error is statistically random.

Figure I.4a further illustrates the similarity between flowmeter and slug test methods (coefficient of determination R^2 = 0.78). Consequently, due to the high correlation between both methods and the reasonably large sample size, the confidence intervals for the limits of agreement are narrow. For example, by converting the difference between log K on the vertical axis of Figure I.3a to a percentage of their average (Bland and Altman, 1999), the 95% limits of agreement become 9.6 and -9.4% (with a mean of 0.17%).

Finally, to further demonstrate the relation between flowmeter and slug test measurements, the profiles of K obtained by both methods at well P03 are compared in Figure I.3b. These data depict similar trends, though flowmeter measurements generally oscillate around the slug test values and display a greater variability. This behaviour may simply be a manifestation of the probe design and calibration specifications. Because vertical measurement intervals are so thin (15 cm), the individual flow rates determined for each of these are

generally small compared to the broad resolution range of the probe. For example, individual flow rates measured for this study fell between 0.05 to 0.73 L/min (with an average of 0.31 L/min), whereas the accuracy of the flowmeter is 0.02 L/min. Larger tested intervals would presumably produce a smoother profile.

Table I.4 Descriptive statistics of log *K* (m/s) data derived from multilevel slug tests and borehole flowmeter measurements. Statistics are derived from all tested intervals in the study area as well as from common intervals tested by both methods. Also shown are statistics related to the differences in log *K* between slug tests and borehole flowmeter measurements. The *K* values from flowmeter and slug tests were estimated following Molz *et al.* (1989) and Bouwer and Rice (1976), respectively.

Method	Number of tests	Mean	SD	Range	Max	Min	Skewn.	Kurt.	Kurtosis-skewness distance (K-S Prob.)[b]
All data used for aquifer characterization									
Slug Test	227	-4.75	0.54	2.09	-3.92	-6.02	-0.62	-0.63	0.11(<0.001)
Flowmeter	141	-4.67	0.42	1.95	-3.79	-5.79	-0.70	-0.18	0.10(<0.001)
Subset used for comparison									
Slug Test	123	-4.58	0.43	1.80	-3.93	-5.73	-0.72	-0.41	0.13(<0.001)
Flowmeter	123[a]	-4.59	0.46	1.95	-3.79	-5.74	-0.67	-0.23	0.102(0.003)
Difference (Slug - Flowmeter)									
Difference	123	0.0077	0.22	1.39	0.71	-0.67	-0.0018	0.58	0.058(0.38)

a: comparison intervals for the flowmeter include 7 null values not considered in the statistical analyses.
b: K-S prob is for kurtosis-skewness probability.
SD: Standard-deviation
Skewn.: Skewness
Kurt.: Kurtosis

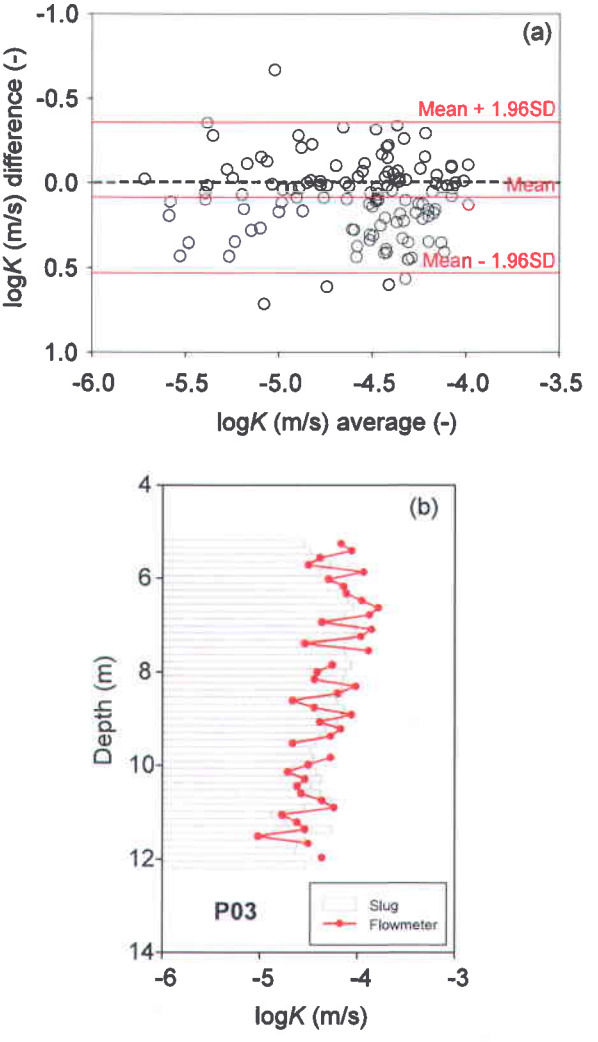


Figure I.3 Comparison of *K* data measured with the flowmeter and multilevel slug tests: (a) In five wells at a 15-cm vertical resolution: the difference in log *K* values obtained from the two measurements is shown against the average of the two measurements. The 95% confidence intervals are shown assuming a normal distribution (± 1.96 SD, SD being the standard deviation). (b) log *K* profile at well P03.

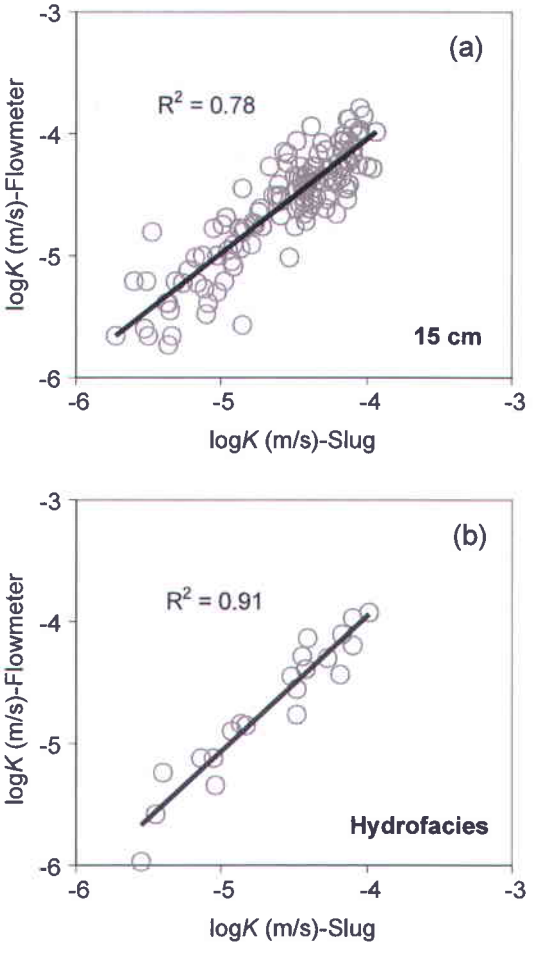


Figure I.4 Comparison of *K* data obtained from multilevel slug tests and flowmeter tests: (a) at 15-cm vertical intervals and (b) for selected intervals corresponding to specific hydrofacies (see Figure I.5).

3.4 Hydrofacies definition from a cumulative flow curve

Our results demonstrate that the flowmeter can effectively measure *K* at a vertical resolution as small as 15 cm. However, this degree of resolution may not always be viable or even necessary, especially when a hydraulic-testing program encompasses broader objectives such as the integration of field data into a regional numerical model. In addition, the spatial correlation among high-resolution *K* data may be difficult to establish when well density over the study area is low. Under such conditions, it would be beneficial to identify and distinguish larger vertical intervals having relatively homogeneous *K* (corresponding to hydrofacies) before interpolation or zonation between wells is attempted to define the hydrostratigraphy. From this perspective, the following method is proposed to distinguish individual hydrofacies using borehole flowmeter measurements (Figures 1.5a to 1.5d):

- Convert the cumulative flow values directly measured with the flowmeter into a curve of cumulative flow percentage. This is done by normalizing flow values by the sum of all values over the screen length independently for each well;

- Define hydrofacies from marked breaks in the slope of the cumulative percentage flow curve. This is done graphically and each line between breaks corresponds to an individual hydrofacies unit. Because changes in flow rate slope are related to variations in K, a break in the slope corresponds to a change in hydrofacies identified by its distinct K;

- Calculate the particular K for each hydrofacies from the difference in flow percentage at each break following the method of Molz *et al.* (1989). Care must be taken to use the appropriate average K, and the same method should be used for all wells. This ensures a uniform scaling of K values and avoids misinterpretation when comparing wells.

An example of this graphical technique is illustrated in Figure I.5a, where straight lines are drawn between each break in the cumulative flow curve. For the profile shown, five hydrofacies are delineated and numbered. The slope of the straight line associated with each hydrofacies unit is related to K, and an average K value is calculated from the difference in flow values at each break. The resulting hydrofacies and their K values are presented in Figure I.5b. In Figures 1.5c and 1.5d, this method is applied once again to well P17, but using a vertical resolution of 60.96 cm (24 inches) instead of 15 cm. In this lower-resolution profile, contrasts in K are less pronounced and the aquifer is reduced to three hydrofacies.

Hence, it is important to acquire borehole flowmeter measurements at a high enough vertical resolution to delineate individual hydrofacies units. However, this approach requires a compromise between the scale of the heterogeneity and the flow detection limit of the probe. If the vertical intervals defined in the testing program are too large, the spatial variability in K will not be adequately represented. For instance, for the range of K observed across the study area, a resolution of 15 cm proved to be an efficient choice. For more homogeneous aquifers (i.e., fewer breaks in the cumulative percentage flow curve of Figure I.5), larger intervals may be satisfactory and could be used to reduce the testing time. For less permeable aquifers, a testing program with larger intervals could be implemented to avoid too many measurements below the

detection limit of the flowmeter. Operationally, it takes much less time (by a factor of roughly 5 to 8) with the flowmeter method to perform the same measurements as with the multilevel slug test method in sediments having a low-to- moderate range of K as encountered at this study site.

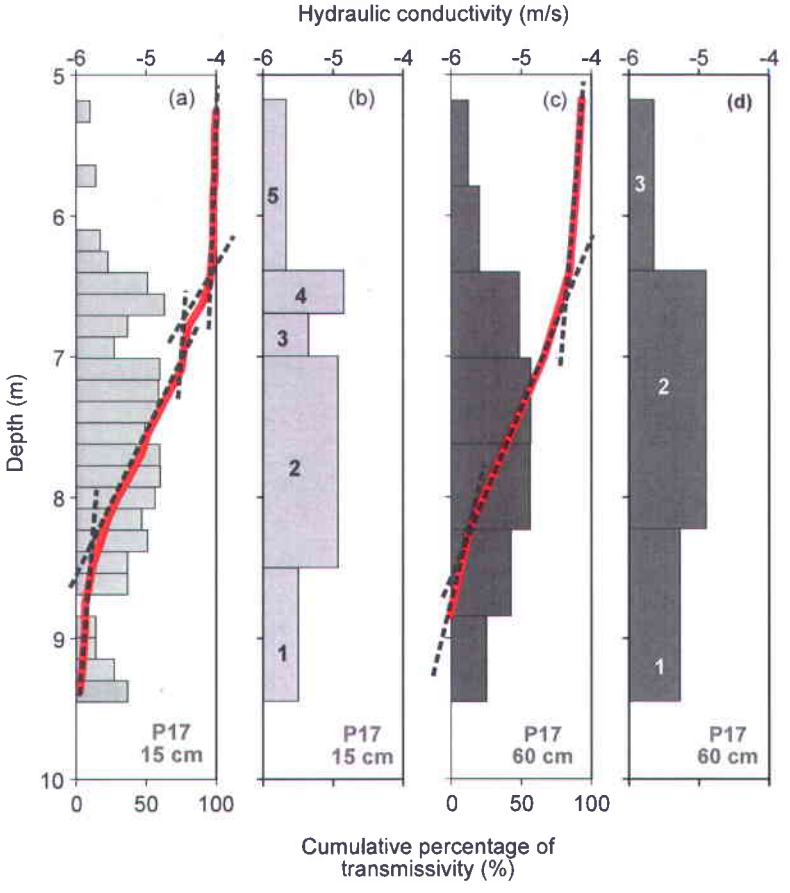


Figure I.5 Application of the cumulative flow-curve interpretation method. The figure shows calculated *K* (grey blocks) and cumulative transmissivity profiles (equivalent to cumulative flow, red line) using two vertical resolutions of flowmeter measurements at well P17: (a) *K* estimates for intervals at a vertical resolution of 15 cm (b) *K* estimates for intervals corresponding to hydrofacies units based on intervals with a steady slope of the cumulative flow curve shown in (a); (c) *K* estimates for intervals at a vertical resolution of 61 cm; (d) *K* estimates for intervals corresponding to hydrofacies units based on intervals with a steady slope of the cumulative flow curve shown in (c). Black dashed lines are the constant-slope sections of the cumulative flow curves over intervals for which *K* values are estimated.

As an alternative to calculating K for each hydrofacies, the average of individual 15-cm measurements can be used. However, using the cumulative curve to estimate hydrofacies K is particularly useful when null K values are obtained due to the flowmeter sensitivity range. Null values are averaged into the cumulative curve and can thus provide a representative K value for low-permeability hydrofacies as long as thicknesses span a few measurement intervals. This approach of using entire hydrofacies intervals to estimate K yields

more representative values than those based on individual 15-cm intervals; these latter estimates will overestimate K because low K values (non detected flow) cannot be determined and are thus not used in computing the average. However, when there are no null values, using the average of the individual intervals or using entire hydrofacies intervals should be roughly equivalent because the average error between K values estimated from flowmeter and slug tests is randomly distributed as shown on Figure I.3a. As expected, using the difference in flow to calculate K for each hydrofacies provides a more smoothed estimate of K than that derived for each individual tested intervals. This is illustrated in Figures 1.4a and 1.4b, where flowmeter results are compared to those obtained from slug tests between packers; R^2 for K values associated with hydrofacies is 0.91 whereas R^2 determined from 15-cm intervals is 0.78. For hydrologic studies concerned with general flow behavior, average values obtained for each hydrofacies may be sufficient to understand the groundwater system; for studies examining transport, the variation of K within each hydrofacies as provided by the individual estimates may be more relevant.

4 Conclusions

The objective of this study was to develop a practical and verifiable approach for the application of the borehole flowmeter in the hydrogeologic characterization of heterogeneous unconsolidated aquifers. Important conclusions and observations resulting from this effort include the following:

- To be useful, K measurements derived from the borehole flowmeter or multilevel slug tests must initially rely on proper well installation and extensive well development. In unconsolidated aquifers, wells installed by the direct-push technique should be favored over conventional methods (i.e., hollow-stem auger, mud-rotary) that use backfilled material (natural or gravel-pack). The direct-push method minimizes disturbance to the aquifer fabric and is more suitable for hydrogeologic studies that employ downhole measurements. However, care should be taken in installing direct-push wells to ensure a firm contact between the sediments and the screen. For example, in this study the distance between the outer casing radius and the outer casing screen was designed to be as thin as possible. Moreover, this kind of installation may be more appropriate in non-cohesive soils (e.g. gravel, sand) than in cohesive soils, where the hole created to accommodate the well may not collapse. As a consequence, voids may remain around the screen that will create a water reservoir around the screen that will favored hydraulic short circuits between tested intervals.

- The borehole flowmeter is a practical tool for quantifying the spatial variability of K in unconsolidated sediments. Values of K measured with a borehole flowmeter are found to be equivalent to those obtained from multilevel slug tests. This equivalence remains as long as flow velocity is above the flowmeter detection limit. Sufficiently fast velocities can be obtained if the material is relatively permeable ($> 10^{-6}$ m/s as in this study) or by increasing the well diameter in lower permeability material. Furthermore, this method is more time and cost efficient compared to slug tests performed at the same vertical resolution, provided the detection limit of the probe is not exceeded.

A proposed graphical-interpretation method is introduced that is based on the cumulative transmissivity curve constructed from flowmeter data. This technique is useful in distinguishing among individual hydrofacies and delineating the units at scales (1) that are computationally manageable for regional aquifer modeling and (2) that effectively represent the relative K contrasts of heterogeneous unconsolidated aquifers. However, to avoid the loss of hydrogeologic spatial details, borehole flowmeter measurements should be carried out at a relatively high sampling resolution (near 15 cm for this study). The choice of vertical resolution should be based on a balance between the detection limit of the borehole flowmeter and anticipated degree of aquifer heterogeneity.

The testing strategy of obtaining borehole flowmeter measurements in direct-push wells provides a reliable and cost-effective means of delineating hydrofacies in unconsolidated heterogeneous aquifers. Moreover, the results of this study extend the range of applicability of the flowmeter, conventionally restricted to fractured-rock and sand-and-gravel aquifers, to aquifers having low-to-medium K.

Acknowledgments

Constructive comments and references provided by Fred Molz helped improve a draft version of this manuscript. The authors would like to acknowledge the important technical support provided by Dave Martin and Barbara Corland, and are grateful to Christine Rivard, Geoff Bohling, Cynthia Dinwiddie and Carl Keller for insightful comments that improved the original manuscript. We also sincerely thank Yves Michaud, Alfonso Rivera, Donna Kirkwood, Stéphane Chalifoux and Gérard Goyette for their sustained support. This project was promoted under the Memorandum of Cooperation (Annex 18) between the USGS and Natural Resources Canada and was supported by the Geological Survey of Canada, under the Groundwater Resources Inventory Program, by the Régie intermunicipale de gestion des déchets des

Chutes-de-la-Chaudière, and by NSERC Discovery Grants held by R.L. and E.G. This is ESS contribution number 20090447. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. or Canadian Governments.

References

Arnold, K.B. and F.J. Molz. 2000. In-well hydraulics of the electromagnetic borehole flowmeter: Further studies. Ground Water Monitoring and Remediation 20, no. 1: 52-55.

American Society for Testing and Material (ASTM). 2004. D6724: Standard Guide for Installation of Direct Push Ground Water Monitoring Wells. ASTM International, 9 pp.

Bland, J.M. and D.G. Altman. 1999. Measuring agreement in method comparison studies. Statistical Methods in Medical Research 8: 135-160.

Bolduc, A. 2003. Géologie des formations superficielles, Charny, Québec. Commission Géologique du Canada, Dossier public 1976, échelle 1/50000.

Boman, G.K., F.J. Molz, and K.D. Boone. 1997. Borehole flowmeter application in fluvial sediments: methodology, results, and assessment. Ground Water 35, no.3: 443-450.

Bouwer, H. and R.C. Rice. 1976. A slug test method for determining hydraulic conductivity of unconfined aquifers with completely or partially penetrating wells. Water Resources Research 12, no.3: 423-428.

Butler, J.J. Jr. 1996. Slug tests in site characterization: Some practical considerations. Environmental Geosciences 3, no. 3: 154-163.

Butler, J.J. Jr. 1998. The Design, Performance, and Analysis of Slug Tests. Lewis Publishers. 252 pp.

Butler, J.J. Jr. and J.M. Healey. 1996. Detection of low-permeability well skins using slug tests. Kansas Geological Survey Open-File Report 96-20, 17 pp.

Butler, J.J., C.D. McElwee, and W. Liu. 1996. Improving the quality of parameter estimates obtained from slug tests. Ground Water 34, no. 3: 480-490.

Clemo, T. 2010. Coupled aquifer-borehole simulation. Ground Water 48, no. 1: 68-78.

Crisman, S.A., F.J. Molz, D.L. Dunn, and F.C. Sappington. 2001. Application procedures for the electromagnetic borehole flowmeter in shallow unconfined aquifers. Ground Water Monitoring and Remediation 21, no. 4: 96-100.

de Marsily, G., F. Delay, J. Gonçavès, P. Renard, V. Teles, and S. Violette. 2005. Dealing with spatial heterogeneity. Hydrogeology Journal 13: 161-183.

Dinwiddie, L.C., N.A. Foley, and F.J. Molz. 1999. In-well hydraulics of the electromagnetic borehole flowmeter. Ground Water 37, no. 2: 305-315.

Driscoll, F.G., 1986. Groundwater and Wells. Johnson Filtration Systems Inc., St.Paul, Minnesota, 1021 pp.

Fellenius, B.H. and A. Eslami. 2000. Soil profile interpreted from CPTu data. "Year 2000 Geotechnics" Geotechnical Engineering Conference, Asian Institute of Technology, Bangkok, Thailand, November 27-30, 2000, 18 pp.

Foley, N.A. 1997. Pressure distribution around an electromagnetic borehole flowmeter in an artificial well. Masters thesis, Environmental, Engineering & Science Department, Clemson University, Anderson, SC.

Frei, S., J.H. Fleckenstein, S.J. Kollet, and R.M. Maxwell, 2009. Patterns and dynamics of river-aquifer exchange with variably-saturated flow using a fully-coupled model. Journal of Hydrology 375: 383-393.

Hanson, R.T. and T. Nishikawa. 1996. Combined use of flowmeter and time-drawdown data to estimate hydraulic conductivities in layered aquifer systems. Ground Water 34, no. 1: 84-94.

Hantush, M. S., 1960. Modification of the theory of leaky aquifers. Journal of Geophysical Research 65, no. 11, 3713-3725.

Henebry, B.J. and G.A. Robbins. 2000. Reducing the Influence of Skin Effects on Hydraulic Conductivity Determinations in Multilevel Samplers Installed with Direct Push Methods. Ground Water 38, no. 6: 882-886.

Hess, A.E. 1986. Identifying hydraulically conductive fractures with a slow-velocity borehole flowmeter. Canadian Geotechnical Journal 23: 69-78.

Hess, K.M., S.H. Wolf, and M.A. Celia. 1992. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts, 3, Hydraulic conductivity and calculated macrodispersivities. Water Resources Research 28, no. 8: 2011-2027.

Hvorslev, M.J. 1951. Time Lag and Soil Permeability in Ground-Water Observations. Bull. No. 36, Waterways Exper. Sta. Corps of Engrs, U.S. Army, Vicksburg, Mississippi, p. 1-50.

Hyder, Z, J.J. Jr. Butler, C.D. McElwee, and W. Liu. 1994. Slug tests in partially penetrating wells. Water Resources Research 30, no. 11: 2945-2957.

Koltermann, C.E. and S.M. Gorelick, 1995. Fractional packing models for hydraulic conductivity derived from sediment mixtures. Water Resources Research 31, no. 12: 32837-3297.

Kruseman, G.P. and N.A. de Ridder. 1990. Analysis and evaluation of pumping test data, 2nd ed. ILRI Publication 47, International Institute for Land Reclamation and Improvement, Wageningen, The Netherlands, 367 pp.

Levy, B.S. and L. Pannell. 1991. Evaluation of a pressure system for estimating in-situ hydraulic conductivity. Ground Water Management No. 5, Proceedings 5th National Outdoor Action Conference, Las Vegas, NV, p. 31-45.

Li, W., A. Englert, O.A. Cirpka, and H. Vereecken. 2008. Three-dimensional geostatistical inversion of flowmeter and pumping test data. Ground Water 46, no. 2: 193-201.

Lunne, T., P.K. Robertson, and J.J.M. Powell. 1997. Cone penetration testing in geotechnical practice. Spon Press, Taylor and Francis Group, New York, 312 pp.

Moench, A.F. 1985. Transient flow to a large-diameter well in an aquifer with storative semiconfining layers. Water Resources Research 21, no. 8: 1121-1131.

Molz, F.J. and J.G. Melville. 1996. Comment on "Combined Use of Flowmeter and Time-Drawdown Data to Estimate Hydraulic Conductivities in Layered Aquifer Systems" , Ground Water 34, no. 5: 770.

Molz, F.J., R.H. Morin, A.E. Hess, J.G. Melville, and O. Guven. 1989. The impeller meter for measuring aquifer permeability variations - evaluation and comparison with other tests. Water Resources Research 25, no. 7: 1677-1683.

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.

Morin, R.H., A.E. Hess, and F.L. Paillet. 1988. Determining the distribution of hydraulic conductivity in a fractured limestone aquifer by simultaneous injection and geophysical logging. Ground Water 2, no. 6: 587-595.

Ngan-Tillard, D.J.M., X.H. Cheng, J. van Nes, and P.L.J. Zitha. 2005. Application of x-ray computed tomography to cone penetration tests in sands. Proceedings of the Geo-Frontiers 2005, Austin, Texas, January 24-26, 2005, ASCE Geotechnical Special Publications Nos. 130-142.

Ouellon, T., R. Lefebvre, D. Marcotte, A. Boutin, V. Blais, and M. Parent. 2008. Hydraulic conductivity heterogeneity of a local deltaic aquifer system from the kriged 3D distribution of hydrofacies from borehole logs, Valcatier, Canada. Journal of Hydrology 351, nos. 1-2: 71-86.

Paillet, F.L. 1998. Flow modeling and permeability estimation using borehole flow logs in heterogeneous fractured formations. Water Resources Research 34, no. 5: 997-1010.

Rehfeldt, K.R., J.M. Boggs, and L.W. Gelhar. 1992. Field study of dispersion in a heterogeneous aquifer. 3: Geostatistical analysis of hydraulic conductivity. Water Resources Research 28, no. 12: 3309-3324.

Ross, H.C. and C.D. McElwee. 2007. Multi-level slug tests to measure 3-D hydraulic conductivity distributions. Water Resources Research 16, no. 1: 67-79.

Rubin, Y. and S. Hubbard. 2005. Hydrogeophysics. Springer, Water Science and Technology Library, Vol. 50, 523 pp.

Ruud, N.C., Z.J. Kabala, and F.J. Molz. 1999. Evaluation of flowmeter-head loss effects in the flowmeter test. Journal of Hydrology 224, no. 1: 55-63.

Sellwood, S.M., J.M. Healey, S. Birks, and J.J. Jr. Butler. 2005. Direct-push hydrostratigraphic profiling: Coupling electrical logging and slug tests. Ground Water 43, no. 1: 19-29.

Shinn, J.D., D.A. Timian, R.M. Morey, G. Mitchell, C.L. Antle, R. Hull. 1998. Development of a CPT deployed probe for in situ measurement of volumetric soil moisture content and electrical resistivity. Field Analytical Chemistry and Technology 2, no. 2: 103-110.

United States Environmental Protection Agency (USEPA). 2005. Groundwater sampling and monitoring with direct push technologies. Solid Waste and Emergency Response, EPA 540/R-40/005, 67 pp.

Young, S.C. and H.S. Pearson. 1995. The electromagnetic borehole flowmeter: description and application. Ground Water Monitoring and Remediation 15, no. 4: 138-147.

Young, S.C. 1995. Characterization of high-K pathways by borehole flowmeter and tracer tests. Ground Water 33, no. 2: 311-318.

Young, S.C. 1998. Impacts of positive skin effects on borehole flowmeter tests in a heterogeneous granular aquifer. Ground Water 36, no. 1: 67-75.

Zemansky, G.M. and C.D. McElwee. 2005. High-resolution slug testing. Ground Water 43, no. 2: 222-230.

Zlotnik, V.A. and B.R. Zurbuchen. 2003. Estimation of hydraulic conductivity from borehole flowmeter tests considering head losses. Journal of Hydrology 281, nos. 1-2: 115-128.

Supporting information

Additional Supporting Information may be found in the online version of this article:

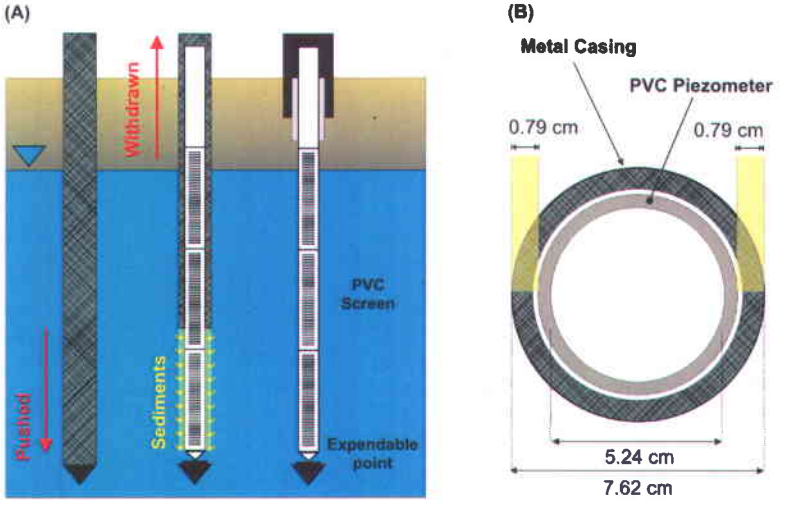
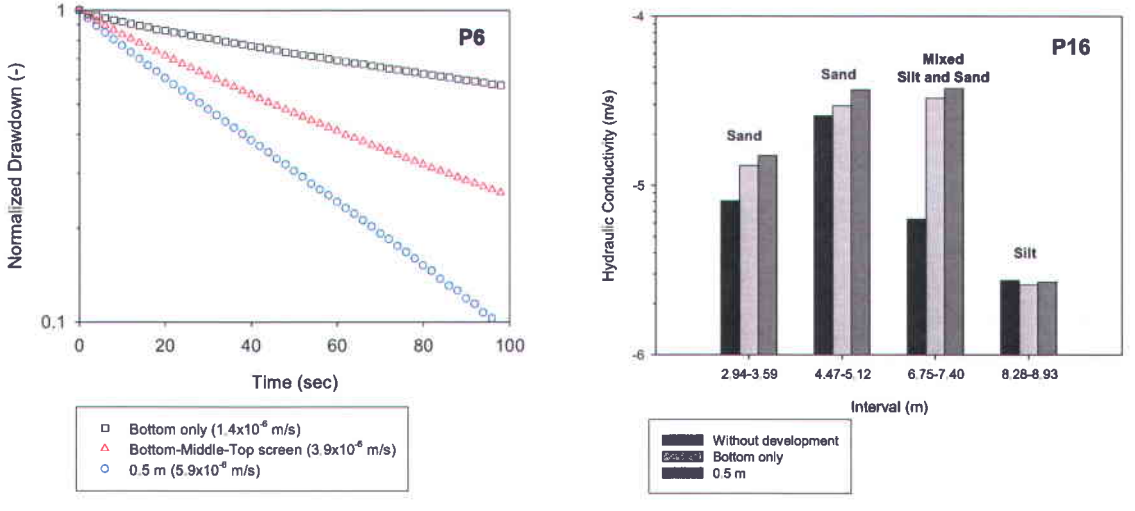


Figure I.S1 Direct-push well installation without gravel-pack. (a) Three-step installation procedure: metal casing hammered into the ground; PVC tubing slotted over saturated zone inserted inside the metal casing before its withdrawal; protection tubing placed around the sealed top of blank PVC tubing. (b) Casing and tubing dimensions showing the approximately 8-mm-wide annular space between the PVC tubing and borehole wall to be filled by sediments.



a) b)

Figure I.S2 Hydraulic response to well development following different pumping-surging configurations. (a) Normalized drawdown as a function of time for slug tests performed at observation well P6 following different pumping-surging configurations. (b) Apparent hydraulic conductivity evolution for selected discrete intervals within observation well P16.

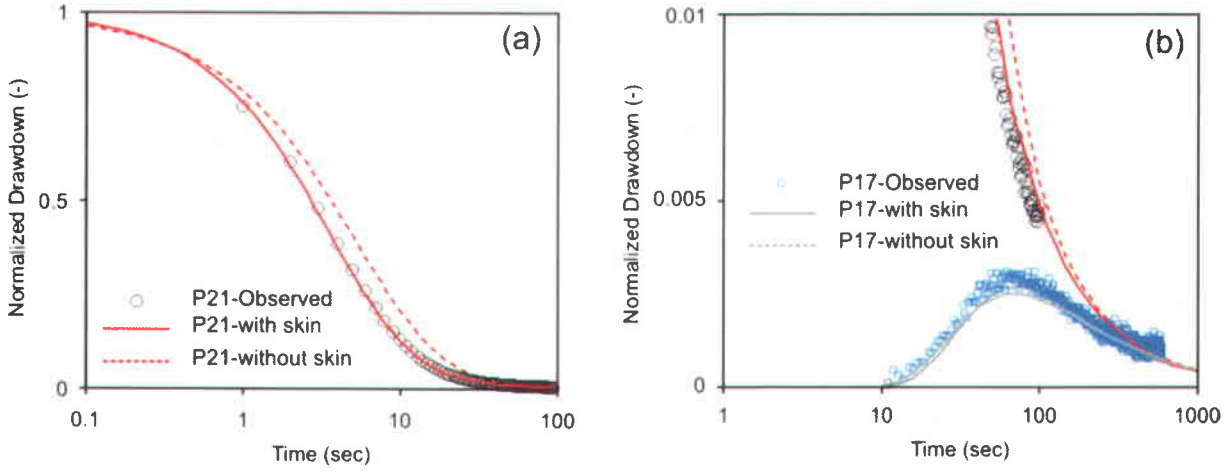


Figure I.S3 (a) Observed drawdown during interference slug test at injection well P21 and (b) expanded view of late-time drawdown at well P21 and observation well P17. Symbols refer to measurements and lines to adjustments to the solution of Hyder *et al.* (1994). Derivative curves (not shown) were also used to improve curve fitting.

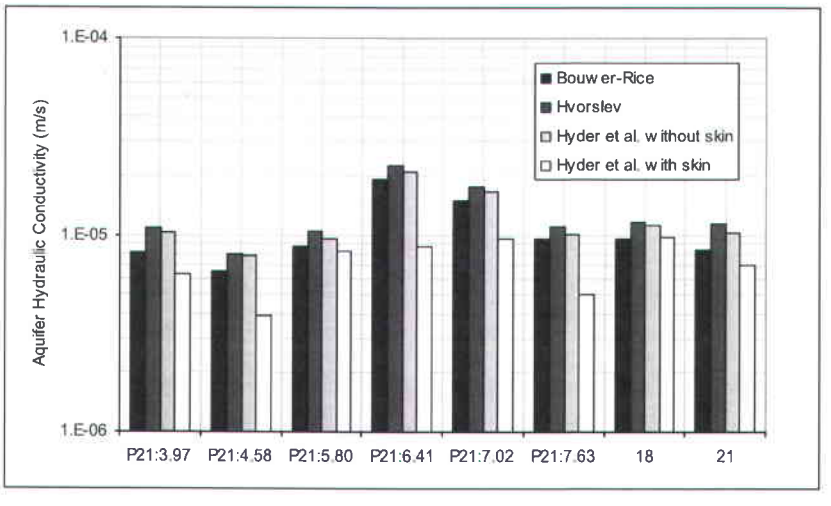
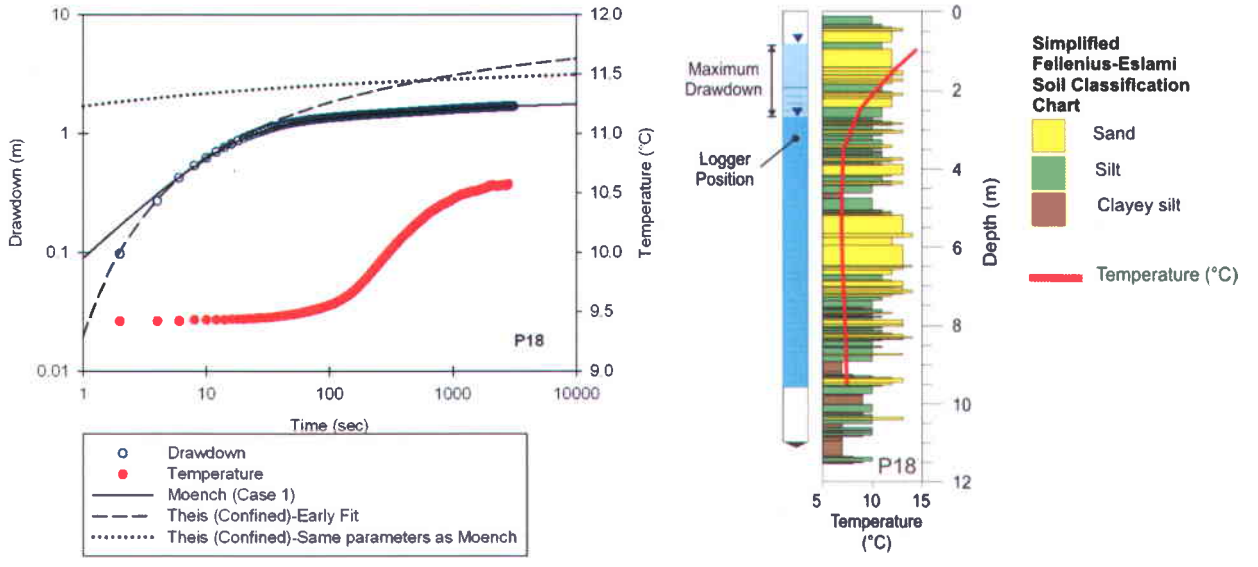


Figure I.S4 Hydraulic conductivity derived from slug tests using four analytical solutions.



a) b)

Figure I.S5 a) Drawdown and groundwater temperature measured during a flowmeter test in well P18. Early-time drawdown data are matched to the Theis solution for confined aquifers and the entire data record is matched to the solution of Moench (1995) that considers leakage from the aquifer top. For comparison, the Theis solution is also shown using the hydraulic parameters determined from the Moench solution (Table I.2). b) CPT-derived hydrostratigraphy, borehole fluid temperature profile, and well installation for well P18 (see Figure I.1b for location).

Table I.S1 Hydraulic conductivity values derived from slug tests using four different analytical solutions. Hydraulic conductivities estimated for each interval from Hyder *et al.* (1994) with skin are reported directly from Table I.1. *K* values for full screen are averages for each injection-observation combination presented in Table I.1. When present, number following well name refers to the depth of the top of the screen interval.

Well or Interval	Hydraulic Conductivity (m/s)			
	Bouwer and Rice (1976)	Hvorslev (1951)	Hyder *et al.* (1994) without skin	Hyder *et al.* (1994) with skin
	Interval			
P21:3.97	8.20×10^{-6}	1.08×10^{-5}	1.03×10^{-5}	6.38×10^{-6}
P21:4.58	6.54×10^{-6}	7.95×10^{-6}	7.83×10^{-6}	6.82×10^{-6}
P21:5.80	8.72×10^{-6}	1.04×10^{-5}	9.58×10^{-6}	9.58×10^{-6}
P21:6.41	1.92×10^{-5}	2.26×10^{-5}	2.09×10^{-5}	1.10×10^{-5}
P21:7.02	1.51×10^{-5}	1.77×10^{-5}	1.66×10^{-5}	7.97×10^{-6}
P21:7.63	9.62×10^{-6}	1.11×10^{-5}	1.01×10^{-5}	5.03×10^{-6}
	Full Screen			
P18	9.64×10^{-6}	1.17×10^{-5}	1.12×10^{-5}	9.58×10^{-6}
P21	8.44×10^{-6}	1.15×10^{-5}	1.02×10^{-5}	7.10×10^{-6}

ARTICLE II

SINGLE-WELL INTERFERENCE SLUG TESTS TO ASSESS THE VERTICAL HYDRAULIC CONDUCTIVITY OF UNCONSOLIDATED AQUIFERS

Résumé

La connaissance de la conductivité hydraulique et de son anisotropie pour les matériaux composant les aquifères est essentielle pour une bonne compréhension de l'écoulement de l'eau souterraine et du transport de contaminants. Plusieurs méthodes in situ existent pour déterminer la composante horizontale de la conductivité hydraulique (K_h), mais la conductivité hydraulique verticale (K_v) est cependant rarement mesurée faute de méthodes appropriées. Cet article propose des tests d'interférence verticaux par choc hydraulique, une adaptation des tests d'interférence entre puits pour un seul puits, pour la mesure in situ de K_v. Chacun des tests est réalisé entre un intervalle d'injection et d'observation isolés verticalement le long d'un puits par trois obturateurs. Un choc hydraulique est d'abord produit dans l'intervalle d'injection et les rabattements induits dans les intervalles d'injection et d'observation sont ensuite enregistrés. Pour une démonstration du concept, 12 tests d'interférence verticaux ont été réalisés séquentiellement le long d'un puits entièrement crépiné aménagé dans un aquifère modérément hétérogène et très anisotrope composé de silt et de sable littoraux. Le puits utilisé a été aménagé par enfoncement de façon à ce que les sédiments en place recouvrent la zone crépinée et en évitant ainsi l'utilisation de sable filtrant. Les puits aménagés par enfoncement permettent donc la mesure directe des propriétés des sédiments en place avec un minimum d'interférence avec les éléments de construction du puits. Les rabattements mesurés dans les intervalles d'injection et d'observation des tests ont été analysés par modélisation numérique inverse pour reconstruire les profils hétérogènes de K_h, K_v et d'emmagasinement spécifique. Les résultats ont été vérifiés en comparant les rabattements mesurés avec ceux prédits par la modélisation ainsi qu'avec des mesures in situ et de laboratoire de K_h et K_v. Les résultats indiquent que le profil de K_v obtenu par tests d'interférence verticaux est similaire au profil obtenu de mesures en laboratoire à l'aide d'un perméamètre sur des échantillons de sédiments pris dans les mêmes intervalles que les tests d'interférence. Ceci démontre donc que les tests d'interférence verticaux par choc hydraulique sont efficaces pour la mesure in situ de K_v à l'échelle du puits.

Citation: Paradis, Daniel and René Lefebvre. 2013. Single-well interference slug tests to assess the vertical hydraulic conductivity of unconsolidated aquifers. *Journal of Hydrology*, 478: 102–118

Single-Well Interference Slug Tests to Assess the Vertical Hydraulic Conductivity of Unconsolidated Aquifers

Daniel Paradis[1,2*] and René Lefebvre[2]

[1] Geological Survey of Canada, 490 rue de la Couronne, Quebec City, Canada G1K 9A9

[2] Institut National de la Recherche Scientifique, Centre Eau Terre Environnement (INRS-ETE), 490 rue de la Couronne, Quebec City, Canada G1K 9A9

*Corresponding author:

Daniel Paradis

Geological Survey of Canada, 490 rue de la Couronne, Quebec City, Canada G1K 9A9

Tel : (418) 654-3713 Fax : (418) 654-2604 E-mail: dparadis@nrcan.gc.ca

Abstract

Meaningful understanding of flow and solute transport in general requires the knowledge of hydraulic conductivity and its anisotropy. Various field methods allow the measurement of the horizontal component (K_h), but vertical hydraulic conductivity (K_v) is rarely measured, for lack of practical field tests. This paper proposes vertical interference slug tests, an adaptation of inter-well interference slug tests to a single well, for the efficient field measurement of K_v. The test is carried out in a single well between a stress and an observation interval that are vertically isolated with a three-packer assembly. An instantaneous pressure pulse is induced in the stress interval and resulting drawdowns are recorded in both the stress and the observation intervals. In a proof-of-concept field study, 12 vertical interference tests were carried out sequentially along a fully-screened well across a moderately heterogeneous and highly anisotropic aquifer made up of littoral silts and sands. A direct-push method was used to install the well, which was completed without sand-pack to allow the natural collapse of sediments in the thin annular space around the screen. Direct-push wells allow the measurement of in situ hydraulic properties of sediments and minimize well construction interferences with hydraulic tests. Drawdowns measured in stress and observation intervals of multiple tests were simultaneously inverted numerically to reconstruct heterogeneous profiles of K_h, hydraulic conductivity anisotropy (K_v/K_h), and specific storage (S_s). Results were validated by comparison of observed versus predicted drawdowns and with field and laboratory measurements of K_h and K_v made along the tested well. Results indicate that the profile of K_v values obtained with vertical interference slug tests follows a similar pattern with depth than the profile with lab measurements made with a permeameter on soil samples collected in the same intervals as the interference tests, which demonstrates that vertical interference slug tests could provide an efficient method for the field measurement of well-scale K_v values.

Keywords: Vertical hydraulic conductivity, Interference slug test, Heterogeneity, Anisotropy, Inverse modeling, Direct-push well

1 Introduction

Vertical hydraulic conductivity, K_v, values are commonly obtained from laboratory permeameters on vertical soil samples collected in the field (Wenzel, 1942). Permeameter tests can be carried out either on the original sample, if relatively undisturbed, or on a repacked sample. The difficulties in the experimental procedure related to soil sample collection and manipulation for certain geological units might, however, limit the capability to obtain estimations of hydraulic parameters from laboratory measurements (Stienstra and van Deen, 1994). Field tests involving undisturbed volumes of investigation would thus provide a more suitable basis for the estimation of both K_h and K_v, even though under suitable conditions permeameter tests can represent a reliable method to estimate K_v.

Permeameter tests may also be time consuming when many measurements are needed (Klute and Dirksen, 1986). Permeameter tests are thus mostly restricted to a small number of samples and seldom used extensively as part of routine hydrogeological characterizations; its application being mostly restricted to scientific demonstrations (Johnson *et al.* 2005; Sudicky *et al.*, 2010). Furthermore, in the case of permeameter tests, they only provide K_v estimates for the shallow unsaturated parts of aquifer that are accessible at soil surface. Thus, there is a need for an efficient K_v measurement method that can apply in observation wells that reach deeper units.

The restriction imposed by soil sample collection and manipulation for laboratory measurements often requires the use of field tests, instead, to estimate hydraulic parameters. Kabala (1993) proposed the dipole-flow test for the direct measurement of both K_h and K_v. In the dipole-flow test, two chambers are isolated in a well with inflatable packers. During the test, water is pumped at a constant rate from one chamber to the other, which generates a circulation of water in the aquifer next to the well. Pressure transducers are installed in the chambers to monitor head changes. Steady state and transient analysis of the dipole-test have been used to estimate aquifer properties (Kabala, 1993; Xiang and Kabala, 1997; Zlotnick and Ledder, 1996; Zlotnick and Zurbuchen, 1998; Zlotnick *et al.*, 2001). However, the solution for both K_h and K_v only using head changes is non-unique. To resolve this non-uniqueness, Hvilshoj *et al.* (2000) proposed an inverse numerical multilayer model constrained with pressure responses in the dipole and in adjacent piezometers. Sutton *et al.* (2000) also proposed to combine a steady-state dipole flow test with the use of a tracer to measure the travel time of the circulation loop between the two chambers. For larger scale evaluation of K_v, Goltz *et al.* (2008) proposed to extend the dipole-flow test to chambers isolated into distant wells. Tandem circulation wells tests are carried out with or without tracers.

Although the dipole test has become a proven method, for most hydrogeological studies the logistics involved in such a test restricts it from becoming a routinely used characterization method. In the petroleum industry, vertical pulse interference tests are used to estimate K_v (Burns, 1969; Hirasaki, 1974; Onur *et al.*, 2004; Sheng, 2009). In a vertical pulse interference test, the stress interval and the observation point must be vertically separated in a single borehole. Inflatable packers can achieve such a separation. The stress interval is a partially penetrating screen where a constant pumping rate is maintained for the duration of the test. The transient pressure data at the stress interval and at the observation point are recorded for the analysis. However, most analytical solutions available for the analysis of vertical pulse interference tests are based on superposition of point source or line source solutions (Burns, 1969; Hirasaki, 1974; Sheng, 2009). Point and line source solutions are not valid for early-test times as the wellbore storage is neglected (Satter *et al.*, 2007). As wellbore storage acts to negatively influence test results, the pumping time has to be long enough to dissipate the wellbore storage effects. In shallow aquifers, middle to long-term pumping may be subject to outer boundary conditions, which may lead to erroneous interpretation of the test (Kruseman and de Ridder, 1990). As used in the petroleum industry, vertical pulse interference tests are thus not suitable to be used for routine hydrogeological characterization.

As an effective alternative to existing laboratory and field tests, this paper proposes a field method to measure K_v, a hydraulic parameter that is difficult to obtain in practice with existing methods. The proposed vertical interference slug test is an adaptation to a single well arrangement of the inter-well interference slug test (Audouin and Bodin, 2007; Belitz and Dripps, 1999; Brauchler *et al.*, 2007; Brauchler *et al.*, 2010; Liu and Butler, 1995; Novakovski, 1989; Spane, 1996; Spane *et al.*, 1996). The proposed test involves a conventional slug test carried out between two isolated screened intervals of a single well using a three-packer assembly. An instantaneous hydraulic pulse is initiated within a stress interval and pressure responses are measured in both the stress and the observation intervals. Vertical interference slug tests are deemed effective because they extend the characterization potential of multilevel slug tests, which measure essentially K_h, to also measure K_v and S_s using a similar field-testing effort. Only an observation interval, for instance, has to be added below a conventional multilevel slug test assembly.

Testing in direct-push wells without sand packs around screens is also proposed to reduce problems related to well construction on hydraulic responses. Direct-push installations offer advantages over conventional installations in unconsolidated formations: they reduce disturbance to the formation,

eliminate drill cuttings and avoid sand packs, sediments being instead in direct contact with the screen. This installation method also facilitates the installation of wells that are fully screened across the entire aquifer thickness, thus providing continuous profiles of hydraulic properties or geochemical groundwater conditions (ASTM, 2010; USEPA, 2005).

Two alternative field-testing and data interpretation approaches are proposed depending on the level of heterogeneity of the tested aquifer vertical interval: 1) based on assumptions of hydraulic conductivity anisotropy ($K_h \neq K_v$) but vertical homogeneity of the tested aquifer interval, a curve matching procedure can be used to estimate K_h and K_v as well as specific storage (S_s) for individual tests; and 2) when heterogeneous aquifer conditions are considered, the same parameters (K_h, K_v and S_s) can be derived from the sequential testing of multiple intervals along a well and the simultaneous numerical inversion of all data obtained from these multiple tests.

This paper reports the results of a field-based proof of concept study aiming to assess the capability of single-well vertical interference slug tests to provide representative measurements of in situ hydraulic parameters, especially K_v, at the well-scale. The paper is organized as follows. In Section 2, a sensitivity analysis is done with an analytical model to understand the effects of aquifer properties and field test conditions on vertical interference slug test responses. Insights gained from the sensitivity analysis will help to design vertical interference slug tests to produce the best dataset. Hydraulic short-circuits and skin effects around a well that may affect vertical interference slug tests are discussed in Section 3, so we may identify these effects in the field data and correct it for a more accurate interpretation. Section 4 describes the methods used and results obtained from a field application of the proposed test in a heterogeneous and anisotropic unconsolidated aquifer. In order to assess the validity of results, Section 4 compares observed drawdowns to predicted values obtained from inverse modeling. Validation of results is also based on comparisons of estimates of K_h and K_v obtained from vertical interference slug tests with measurements of K_h obtained from high-resolution multilevel slug tests and laboratory permeameter K_v measurements on soil samples. The potential application of vertical interference slug tests is discussed in Section 5, and the key findings of this study are summarized in Section 6.

2 Sensitivity analysis of vertical interference slug tests

To assess the effects of aquifer parameters on the responses of vertical interference slug tests, a sensitivity analysis was carried out with a semi-analytical model. The sensitivity study aims to assess if these changes in responses hold the potential to resolve the effects of each individual hydraulic properties, and thus provide a means to measure these properties based on test data. A sensitivity analysis on key test and wellbore parameters is also presented to help in the design of vertical interference slug tests.

For the sensitivity analysis, we used the Kansas Geological Survey (KGS) model developed for the interpretation of cross-well slug tests under both confined and unconfined conditions (Butler, 1995; Hyder *et al.*, 1994; Liu and Butler, 1995). The phenomenon of interest is the head h response, as a function of radial distance from the well r, elevation z and time t, produced by an instantaneous pressure disturbance into the screened or open section of a well. The KGS model accounts for K_h, K_v/K_h and S_s of the aquifer, wellbore storage in the stress well, and the presence of a well skin around the stress well. As illustrated on Figure II.1, the KGS model presumes that the well skin of radius r_{sk} extends over the full thickness of the aquifer. The skin has transmissive and storage properties that may differ from the formation as a whole. Hydraulic parameters are assumed uniform within both the skin and the aquifer, although the components of hydraulic conductivity (K_v and K_h) may differ.

Equations (1)-(9) of Hyder *et al.*, (1994) describe the flow conditions of the system considered for the KGS model. The partial differential equation representing the flow of groundwater in response to an instantaneous change in water level at a well screened in an aquifer is the same for both the skin and the aquifer and can be written as

$$\frac{\partial^2 h_i}{\partial r^2}+\frac{1}{r}\frac{\partial h_i}{\partial r}+\left(\frac{K_{zi}}{K_{ri}}\right)\frac{\partial^2 h_i}{\partial z^2}=\left(\frac{S_{si}}{K_{ri}}\right)\frac{\partial h_i}{\partial t} \qquad (1)$$

where h_i is the head in zone i [L]; S_{si} is the specific storage of zone i [$1/L$]; K_{zi} and K_{hi} are the vertical and horizontal (or radial) components, respectively, of the hydraulic conductivity of zone i [L/T]; t is time [T]; r is radial direction [L]; z is vertical direction (depth) [L] with z=0 at the top of the aquifer and increases downward [L]; i is the designator, for $r \le r_{sk}$, i=1, and for $r \ge r_{sk}$, i=2; and r_{sk} is the outer radius of the skin [L].

The initial conditions involving equal heads h_1 and h_2 in zones 1 and 2 whose depths z are between 0 and the total thickness of the aquifer can be written as

$$h_1(r,z,0)=h_2(r,z,0)=0 \qquad r_{ws}<r<\infty \qquad 0<z<b \tag{2}$$

$$H(0)=H_0 \tag{3}$$

where r_{ws} is the stress interval radius $[L]$, b is aquifer thickness $[L]$, H is level of water in the well $[L]$, and H_0 is the static water column height above the observation interval, equal to the level of water in the well at $t=0$, $[L]$.

The outer boundary conditions are the following:

$$h_2(\infty,z,t)=0 \qquad t>0 \qquad 0\le z\le b \tag{4}$$

$$\frac{\partial h_i(r,0,t)}{\partial z}=\frac{\partial h_i(r,b,t)}{\partial z}=0 \qquad r_{ws}<r<\infty \qquad t>0 \tag{5}$$

whereas the inner boundary conditions at the stress interval are the following:

$$\frac{1}{L_s}\int_{d_s}^{d_s+L_s} h_1(r_{ws},z,t)dz=H(t) \qquad t>0 \tag{6}$$

$$2\pi r_w K_{h1}\frac{\partial h_1(r_w,z,t)}{\partial r}= \begin{cases} 0 & z<d_s, z>d_s+L_s \qquad t>0 \\ \dfrac{\pi r_c^{\,2}}{L_s}\dfrac{dH(t)}{dt} & d_s>z<d_s+L_s \qquad t>0 \end{cases} \tag{7}$$

where d_s is distance from the top of the aquifer to the top of the stress screen $[L]$; L_s is stress interval length $[L]$; r_c is radius of well casing$[L]$.

In order to ensure continuity of flow between the skin and the formation, auxiliary conditions at the skin-formation boundary ($r=r_{sk}$) must also be met:

$$h_1(r_{sk},z,t) = h_2(r_{sk},z,t) \qquad 0 \le z \le b \qquad t > 0 \tag{8}$$

$$K_{h1}\frac{\partial h_1(r_{sk},z,t)}{\partial r} = K_{h2}\frac{\partial h_2(r_{sk},z,t)}{\partial r} \qquad 0 \le z \le b \qquad t > 0 \tag{9}$$

Note that the KGS model assumes that wellbore storage effects can be neglected in the observation interval. To obtain test data with minimal wellbore storage effects, the observation interval should be straddled with packers (Sageev, 1986). If need be, wellbore storage in the observation interval could also be considered using the superposition approach of Novakowski (1989) and Tongpenyai and Raghavan (1981).

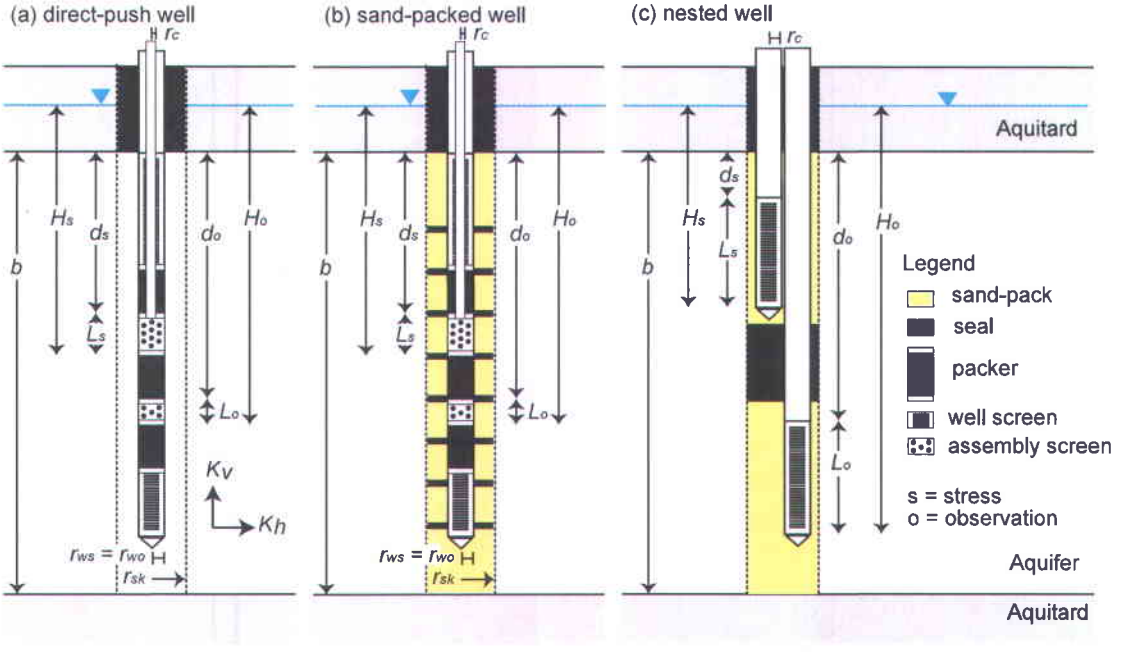


Figure II.1 Well arrangement and definition of parameters for a vertical interference slug test in a hypothetical confined aquifer: (a) between two intervals isolated using a three-packer assembly in a direct-push well fully screened across the aquifer (as used in the field study reported in this paper); (b) the same configuration as (a) but in a sand-packed well with inter-spaced bentonite seals in the sand pack to prevent hydraulic short-circuiting; and (c) between two piezometers of a nested well. Symbols used for parameters are defined in Table II.1.

For homogeneous and anisotropic aquifer conditions, the KGS model can thus be used to simulate the responses of a single-well vertical interference slug test for the three types of test configurations illustrated in Figures II.1a-c. For this sensitivity analysis, the KGS model is used to simulate conditions representative of a direct-push installation with fully-screened well (Figure II.1a). Further details about

test configuration are provided in Section 3.3. Test parameters (and their symbols) applying to these conditions are illustrated in Figure II.1a and defined in Table II.1. For the fully-screened configuration, a three-packer assembly was used to isolate a stress (transmitter) interval, in which an instantaneous head change was induced, from an observation (receiver) interval. Following that pulse, changes in heads with time were recorded in both the stress and observation intervals. Further details about field operations are provided in Section 4.

Table II.1 Aquifer and borehole parameters used for the base case in the sensitivity analysis with the KGS model. The vertical distance (corresponding to depth) is increasing downward and its origin (zero value) is the top of the aquifer (Figure II.1).

Aquifer properties	
Horizontal hydraulic conductivity (K_h)	1×10^{-5} m/s
Specific storage (S_s)	1×10^{-4} m^{-1}
Hydraulic conductivity anisotropy (K_v/K_h)	0.1
Aquifer thickness (b)	12 m
Stress interval parameters	
Static water column height (H_s)	4.56 m
Depth to top of interval (d_s)	3.95 m
Interval length (L_s)	0.61 m
Inside radius of well casing (r_c)	0.0127 m
Inside radius of screen (r_{ws})	0.0254 m
Well coordinates (R_s)	0 m
Observation interval parameters	
Static water column height (H_o)	5.48 m
Depth to top of interval (d_o)	5.18 m
Interval length (L_o)	0.30 m
Inside radius of screen (r_{wo})	0.0254 m
Well coordinates (R_o)	0.0254 m

2.1 Aquifer parameters effects

The shape and amplitude of the hydraulic responses in both the stress and observation intervals during a vertical interference slug test is determined by diffusivity. The diffusivity is the ratio of K (K_h and K_v) over S_s. To better understand the role of each component of the diffusivity on pressure pulse propagation, the values of individual hydraulic parameters were varied one at a time, while holding all other properties and parameters constant. The base case used for the sensitivity analysis is for the aquifer parameters, test well/aquifer configuration listed in Table II.1 and shown in Figure II.1a. Simulations were made for the base case and compared to simulations with lower and higher values of K_h, S_s and K_v/K_h. Semi-log plots of dimensionless drawdown (H/H_o) versus time presented in Figure II.2 illustrate the results of the sensitivity analysis. As expected, varying hydraulic parameters causes significant changes in the shape and amplitude of the predicted hydraulic responses in both the stress and the observation intervals.

The influence of K_h on the slug interference response in both the stress and the observation interval is illustrated in Figure II.2a. As shown, K_h significantly controls the timing of slug test head responses in both the stress and observation intervals. The stress interval exhibits the typical progressive drawdown recovery of slug tests, with faster recovery for higher K_h. The dimensionless drawdown response in the observation interval has the form of a pulse that is more delayed in time for lower K_h. For a constant value of S_s, the relationship between K_h and drawdown response time is thus inversely proportional, with higher K_h associated with faster propagation of the head pulse. Horizontal hydraulic conductivity, however, exerts no effect on shape and amplitude of the drawdown response in the stress and observation intervals.

In contrast, Figure II.2b shows that S_s significantly influences the amplitude and shape of the response in the observation interval. Specific storage also affects the arrival time of the initial head response and the peak amplitude of the pulse. Lower values of S_s are associated with larger head responses and faster head propagation. In contrast to the large variations of responses in the observation interval, the responses in the stress interval only show a slight delay at early-test times for differing values of S_s, whereas late-test times are not affected by changes in S_s for the test assembly used in the sensitivity analysis.

Figure II.2c shows the effects of K_v/K_h on slug test response in the stress and observation intervals. As for S_s, changes in K_v/K_h exert a pronounced effect on the amplitude and shape of the head response in the

observation interval. Lower values of K_v/K_h are associated with smaller head responses and slower head propagation. Hydraulic conductivity anisotropy also affects the arrival time of the head pulse peak. Comparison of Figures II.2b and 2.2c for the observation interval shows that K_v/K_h effects are quite similar to those caused by S_s; the effect of increasing K_v/K_h by one order of magnitude is similar to the effect obtained from decreasing S_s also by one order of magnitude. Figure II.2c shows that K_v/K_h also slightly affects the response in the stress interval in a way that differs from the effect of S_s. Figure II.2b shows that S_s mostly influences drawdown for early-test times, whereas K_v/K_h effects are exhibited at late-test times. The distinct effect of S_s and K_v/K_h on the drawdown response obtained in the stress interval thus offers a means of distinguishing differing values of these two aquifer properties.

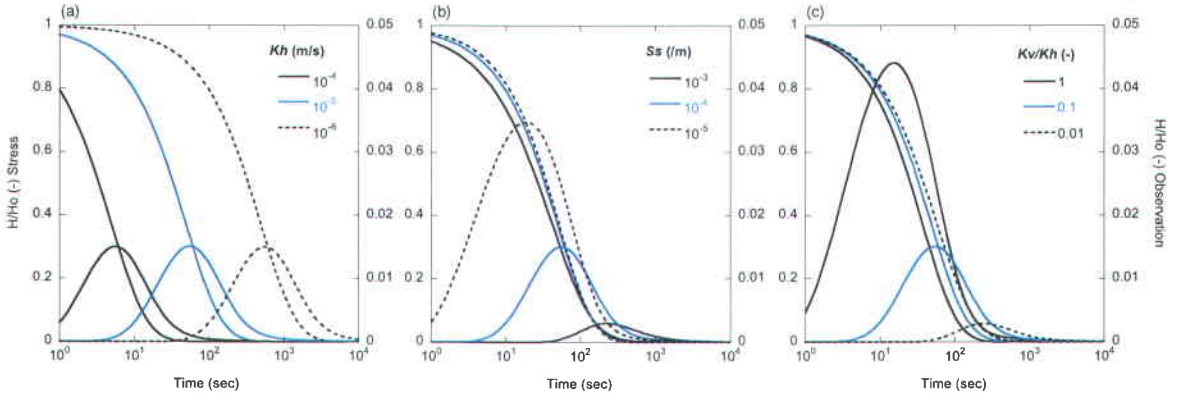


Figure II.2 Dimensionless drawdowns H/H_o versus time in the stress (S-shaped curves) and observation (bell-shaped curves) intervals considering the influence of hydraulic parameters on vertical interference slug test responses: (a) effect of horizontal hydraulic conductivity (K_h); (b) effect of specific storage (S_s); and (c) effect of hydraulic conductivity anisotropy (K_v/K_h). On these graphs and those of Figures II.3, II.4 and II.5, the plain line, bold line and dashed line represent the responses for the base case, for a high value and for a low value relative to the base case, respectively. The scale of H/H_o for the observation interval is not the same as the scale used for the stress interval in (a) and (b), as well as in Figures II.3, II.4, II.5, II.6, II.9 and II.13.

2.2 Test and wellbore parameters

A sensitivity analysis was also carried out to evaluate the effects of selected test and wellbore parameters. The base case used for the comparison of responses is the same as the one presented in Section 2.1 for the sensitivity analysis of aquifer parameters (Table II.1). Parameters considered in this sensitivity analysis included 1) stress interval wellbore storage, 2) stress and observation interval distance, and 3) skin horizontal hydraulic conductivity (K_h').

Figure II.3a shows test responses in the stress and observation intervals for different wellbore storage values obtained by varying the radius of the stress well casing (r_c), while holding constant the radius (r_{ws})

and the length of the stress screen (L_s). As discussed by Brauchler *et al.* (2007), Figure II.3a shows that large wellbore storage values, which correspond to a larger volume of water injected or withdrawn for the slug test, greatly increase the peak amplitude and duration of the head response in the observation interval. In practice, the peak amplitude in the observation interval may be quite small for conditions of high hydraulic conductivity anisotropy, which may limit the accuracy of the measured response. In order to obtain an accurately measurable head response in the observation interval, the volume of injected water (or wellbore storage) should be as large as possible, which may be obtained by using either a large stress well casing radius or a high initial water displacement.

Figure II.3b shows the effects of the distance between stress and observation intervals on the hydraulic response in these intervals. A shorter distance between stress and observation intervals leads to earlier head peak arrival time and longer duration head response having higher amplitude. The maximum distance to ensure accurate head measurements in the observation interval is difficult to quantify because of the combined effects of hydraulic properties, the magnitude of initial water displacement, and the sensitivity of the pressure transducers used. For the range of aquifer properties used in this sensitivity analysis, a stress-observation distance of up to 1 m with an initial water displacement of 1 m is expected to produce a maximum head displacement in the observation interval of approximately 1 cm, which is considered sufficiently accurate for quantitative interpretation considering common pressure transducer accuracy.

As indicated in Figure II.3c, vertical interference slug test responses in the observation interval are more sensitive to low-K skin than high-K skin. For this sensitivity analysis, K_v/K_h and S_s for the skin and the aquifer were assumed to be the same. A high-K skin slightly reduces the peak arrival time and increases the peak amplitude in the observation interval. Low-K skin effects are in the opposite direction but with more pronounced effects for the same magnitude of change in K_h'. These simulations indicate what could be observed for interference slug tests carried out in wells having various levels of development. A high-K skin could result from an over-developed well around which most of the fine material is removed by the development, whereas a low-K skin could result from a partially clogged screen. Figure II.3c thus shows that over-development would not significantly affect a vertical interference slug test response, whereas poor development would seriously deteriorate the response if skin effects cannot be modeled in the data inversion. The high-K skin may also correspond to the effect of a sand pack. In this perspective, Figure II.3c suggests that the installation of a relatively thin sand pack around the screen of a conventional well would provide interference slug test data leading only to a slightly over-estimated K_v. For instance the simulation with a K_h' one order of magnitude higher than the K_h of the formation and a

skin radius r_{sk} twice the radius of the well r_w, will produce a change in K_v of less than 10%, holding S_s constant.

2.3 Considerations for test analysis and design

As previously shown, S_s and K_v/K_h have similar effects on the shape and amplitude of the head response obtained from an interference slug test in the observation interval. For instance, for a given value of K_h, various combinations of S_s and K_v/K_h values can provide nearly identical head responses in the observation interval. To fully discuss the implications of results shown in Figures II.2b and 2.2c, Figure II.4 shows results obtained from combinations of S_s and K_v/K_h values leading to the same responses in the observation interval. Contrary to the identical observation interval responses, the stress interval responses, which have the same overall shape, are however translated on the time axis for the same combinations of S_s and K_v/K_h and the same value of K_h. This translation of head responses in the stress interval is due to the fact that S_s and K_v/K_h act at different times on the stress interval response; S_s acting on early-test times (Figure II.2b), whereas K_v/K_h affects mostly late-test times (Figure II.2c). Then to distinguish unique combinations of K_h, S_s and K_v/K_h with vertical interference slug tests, it is necessary to integrate the measurements obtained from both the stress and observation intervals into the interpretation.

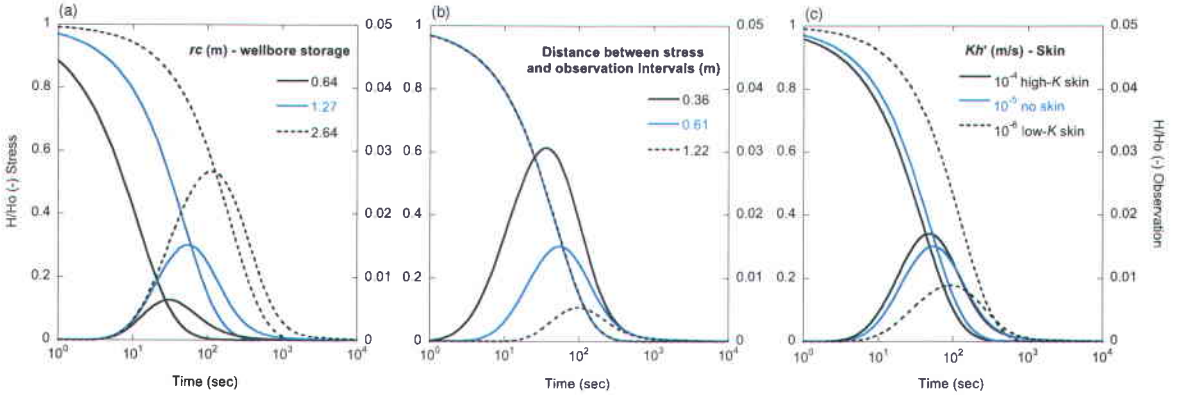


Figure II.3 Influence of various borehole-aquifer test configurations on vertical interference slug test responses in the stress and observation intervals: (a) effect of wellbore storage; (b) effect of vertical distance between stress and observation intervals; and (c) effect of skin horizontal hydraulic conductivity (K_h') - the radius of the skin is two times the radius of the screen ($r_{sk} = 2r_{wo} = 2r_{ws}$) (anisotropic conditions for the aquifer and the skin were assumed to be identical, i.e. K_v'/K_h'=0.1).

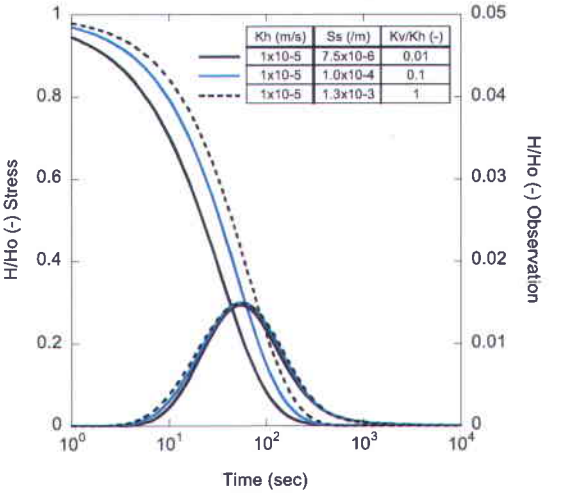


Figure II.4 Sensitivity analysis of responses for various combinations of specific storage (S_s) and hydraulic conductivity anisotropy (K_v/K_h) for a fixed value of horizontal hydraulic conductivity (K_h). This analysis shows that hydraulic properties for vertical interference slug test can be resolved by integrating the responses obtained in the observation and stress intervals in the interpretation.

The Figure II.5 illustrates the effects of the stress interval aspect ratio on the vertical interference slug test responses. The aspect ratio is defined as the screen length (L_s) over screen radius (r_{ws}). The aspect ratio and K_v/K_h appear in the same dimensionless parameter in the KGS model. A smaller aspect ratio leads to a greater proportion of slug-induced vertical flow in a homogeneous formation. By varying the aspect ratios and K_v/K_h values, the Figure II.5 shows that smaller aspect ratios (smaller stress interval length) lead to larger K_v/K_h effects on the stress interval response. A small aspect ratio should thus be used to obtain test data facilitating the estimation of K_v from vertical interference slug tests.

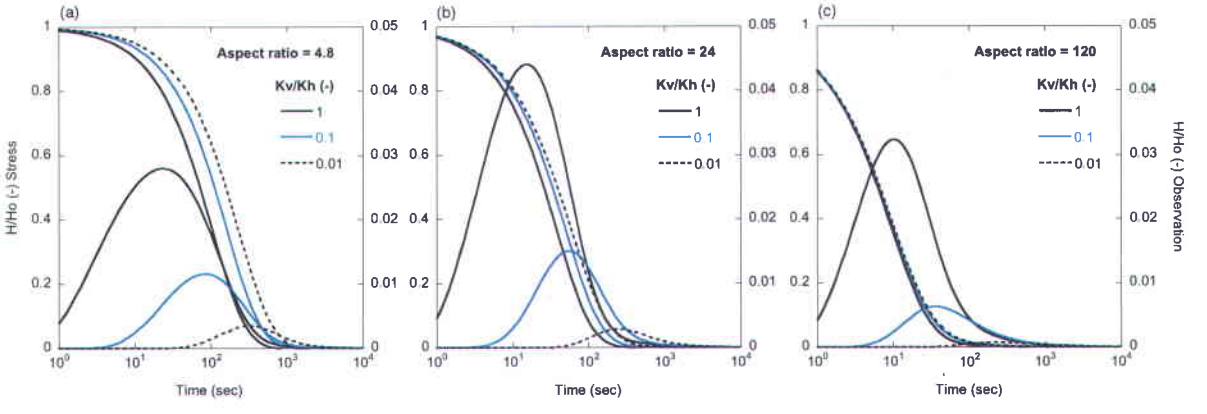


Figure II.5 Sensitivity analysis of responses for various combinations of hydraulic aspect ratio (L_s/r_{ws}): (a) 4.8; (b) 24; and (c) 120. For a screen interval radius of 2.54 cm, those aspect ratios correspond to stress interval lengths of 0.12 m, 0.61 m and 3.05 m, respectively. An aspect ratio of 24 was used for the sensitivity analysis of Figures II.2, II.3 and II.4 as well as for the field application reported in this paper.

3 Hydraulic short-circuit effects

Multiple tests over different parts of an aquifer in the same well can greatly improve aquifer characterization. Multiples tests can also provide indications as to whether test responses are characteristic of the aquifer or are affected by poor well installation. Skin effects that can alter the responses of vertical interference slug tests were illustrated in Section 2.2.

Another kind of well construction bias that can affect hydraulic responses, in particular with vertical interference tests carried out in direct-push wells, are hydraulic short-circuits. For this study, hydraulic short-circuits are defined as preferential flow paths that have been created by the incomplete collapsing of sediments around a screen during direct-push well installation and development. Two kinds of hydraulic short-circuits were observed during field experiment reported in this paper: static short-circuits (SSC) and dynamic short-circuits (DSC). "Static" refers to the fact that the transmissivity of the short-circuit is constant throughout the duration of the test, as opposed to "dynamic", for which the transmissivity is decreasing during the course of the test.

It should be mentioned that DSC could be modeled only with a porous-elastic model, which is beyond the scope of this study. Short-circuit effects are instead illustrated with vertical interference slug test datasets from the forthcoming field program reported in this study, so that such effects can be recognized and accounted for during the analysis (type-curve matching or numerical inversion).

3.1 Static hydraulic short-circuits (SSC)

From a hydraulic viewpoint, a SSC may have similar effects as high-K skin (Figure II.2c). However, a SSC may have greater effects on hydraulic responses because preferential flow paths are not necessary filled with sediments and thus constitute a direct connection between the stress and observation intervals.

As shown in Figure II.6a, a major SSC may be detected when the response in the observation interval have a similar shape than the theoretical curve but a much higher than expected amplitude response. For this test, the amplitude of the response in the observation interval is approximately half the amplitude recorded in the stress interval (H/H_o=0.5). This is far more than the amplitude recorded for other tests in the same well ($H/H_o \approx$ 0.01-0.03). As an additional insight on the occurrence of a SSC, the hydraulic response in the observation interval could not be reproduced with the KGS model without modeling a

skin with low S_s and high K. As the real skin hydraulic parameters are unknown for this case, it was not possible to estimate aquifer parameters from this test. The more cohesive clayey silt present at the base of the well (Figure II.7a) may have prevented the collapse of the formation on the well screen, thus explaining the presence of such a large SSC.

It should be noted that a smaller SSC could also affect vertical interference tests and that their presence may be harder to diagnose. As for a high-K skin, a K_v over-estimation would be obtained from data affected by the presence of a SSC.

3.2 Dynamic hydraulic short-circuits (DSC)

Unlike SSC, the transmissivity of a DSC is pressure-dependent of the stress response and thus decreases with time after the initiation of the head change at the beginning of a vertical interference slug test. This is the most common type of test bias observed in the study presented in this paper. It is hypothesized that even though sediments might have completely collapsed around the screen during direct-push well installation and development, the sediments in contact with the well screen might have lost some of their stiffness. Sediments such as clayey silt may have been compressed over their limit of elasticity by the casing used to install the direct-push well. As a result, when a large head change is applied in the stress interval, the sediment in contact with the screen are then more easily deformed and thus allow the head difference to propagate along the well. As the head change in the stress interval decreases, sediments gradually recover their original shape and the pressure propagation along the well is reduced accordingly.

A dataset showing a typical DSC response (observation in Figure II.6b) was analyzed in more details to explain the effects of a DSC on field data. The effects of a DSC can be recognized by the deviation of the hydraulic response for early-test times, compared to the theoretical response that would be obtained without any DSC effects (full red line in Figure II.6b). This theoretical curve was obtained with the KGS model based on a match to observed peak amplitude and late-test times data (see Section 4.4.2). Measurements for early-test times are above the theoretical response (H/H_o of 0.066 instead of 0). Specifically, observed drawdowns in the observation interval decreased for the first 3 seconds, stabilized for 2 seconds and then increased until the peak amplitude was reached. This suggests that a DSC affects the response in the observation interval only for the first 3 seconds and that over this period both DSC and aquifer responses are superposed to produce the observed response.

To better understand the dynamics of a DSC, modeling was done with the KGS model of the hydraulic response produced by a high-K skin (or SSC) having an equivalent effect on the initial drawdown measured (at 1 sec) in the observation interval (long-dashed line in Figure II.6b). The high-K skin response was modeled by adjusting the skin hydraulic parameters to fit the first drawdown points observed in the observation interval, while keeping the aquifer parameters the same as those estimated in Section 4.4.2 (or plain red line in Figure II.6b). Measured drawdowns in the observation interval are actually well below the calculated response that an equivalent high-K skin would produce. It appears that we cannot reproduce the characteristic shape measured in the observation interval (decreasing then increasing drawdowns), which is affected by DSC, by assuming a constant skin transmissivity.

Since a DSC cannot be modeled with the KGS model, the DSC response was empirically represented. As it is expected that the variation in transmissivity of a DSC is pressure-dependent of the stress interval response, an exponential decay function was thus used to model the drawdowns induced by the DSC. Two exponential models were fitted (Figure II.6c): the exponential decay was fitted with the first 3 seconds of the observation interval response, which are only affected by the DSC (*DSC3sec* in Figure II.6b); the stress interval response for the entire test duration was rescaled to the first drawdown measured in the observation interval (DSCend in Figure II.6b). For scenario DSC3sec the head effect of the DSC is completely dissipated after 96 seconds, whereas for scenario DSCend the DSC effects last for the entire duration of the test (700 sec). The latter scenario may be considered as a worst-case scenario.

To illustrate the DSC effects on K_v estimates, the observed drawdowns in the observation interval were corrected for the two pressure scenarios (Figure II.6b). The drawdown induced by the DSC was subtracted from the observed drawdown in the observation interval. With the scenario DSC3sec, only early-test times are affected without changing significantly the peak amplitude and late-test times drawdown. For scenario DSCend, peak amplitude is somewhat reduced as well as drawdown for late-test times. For this scenario, the K_v that was initially estimated at 1.2×10^{-7} m/s with the original data is now reduced to 9.5×10^{-8} m/s with corrected data (DSC3end). It can be concluded that for the magnitude of DSC observed in this study, we may expect a slight over-estimation (26%) of K_v with vertical interference slug tests. In addition, the implication for test analysis is that measured drawdowns in observation intervals can be corrected, as previously done, or early-test time data affected by DSC can be ignored in the analysis.

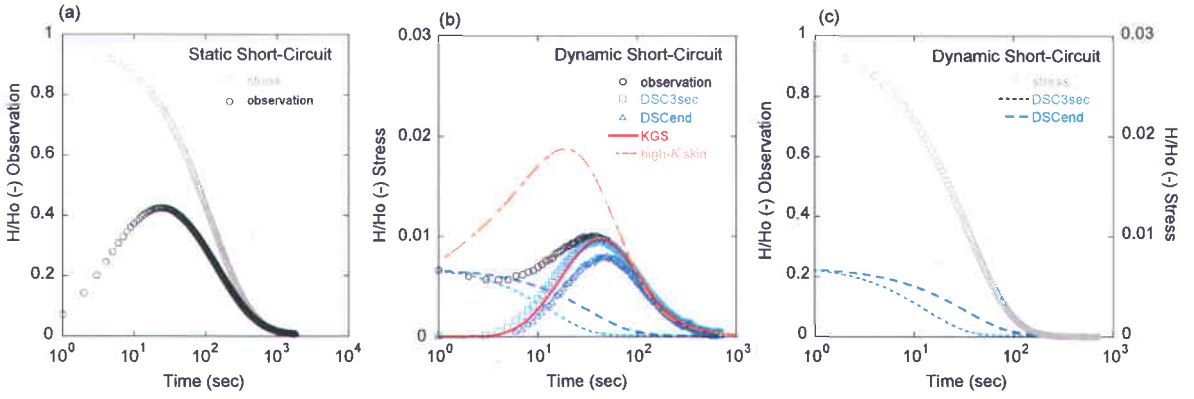


Figure II.6 (a) Example of a major static short-circuit (SSC) detected in the observation interval as see by the very high amplitude response. (b-c) Sensitivity analysis to simulate the effects of a dynamic hydraulic short-circuit (DSC) in the observation interval. Figure (b) shows the observed dimensionless drawdowns H/H_o in the stress interval (t8) and presumed exponential decays of the initial H/H_o values in the observation interval (*r8*), which are related to a DSC, based on the observed decay over the first 3 seconds (DSC3sec) and a worst-case scenario with decay lasting until the end of the 700 seconds observation period (DSCend). These calculations are also reproduced in Figure (c), which shows H/H_o in the observation interval for several cases: the actual measurements (*r8*), measurements corrected using the calculated DSC exponential decays (DSC3sec and DSCend), and simulated responses using the KGS model without DSC effect (KGS) and using a constant high-*K* skin that does not dissipate dynamically.

3.3 Considerations for well design

A vertical interference slug test can be conducted using several different test assembly configurations (Figures II.1a-c). Figure II.1a illustrates tests carried out in the same fully-screened well with a three-packer assembly used to isolate two intervals. The section between the upper two packers is used as the stress interval, while the section between the lower two packers is the observation interval. This assembly can then be used to test a single homogeneous interval or moved sequentially vertically down the well to test the entire screen length. In non-cohesive soils (e.g., gravel, sand), fully-screened direct-push wells (Paradis *et al.*, 2010) or hammer-driven wells (Morin *et al.*, 1988) installed without sand-packs are best suited to measure natural hydraulic properties of the formations. During well installation, however, care should be taken to minimize the risk of hydraulic short-circuits. Open boreholes in rock formations may also be tested with such a test configuration.

For cohesive or semi-cohesive soils (e.g. clayey silt), where the natural collapse of the sediments onto the screen well is uncertain, a sand-packed well (Figure II.1b) with inter-spaced bentonite seals placed to prevent short-circuiting may be a better alternative in such formations (Butler et al., 1994; Ptak and Teutsch, 1994). To obtain continuous profiles, the locations of the seals should be carefully selected to isolate each stress interval from the observation interval along successive testing positions. An alternative

might be to test only homogeneous intervals along the well. In this case, seals should be used to separate the homogeneous intervals and for each interval another seal should be used to avoid short-circuits.

Depicted in Figure II.1c is another configuration that can test a single interval between two isolated piezometers in a nested well. With this arrangement, the bentonite seal between the screens prevents hydraulic short-circuits. In nested wells, either piezometer can be used as a stress or an observation interval. This configuration is restricted to cases for which the stress and the observation intervals are in the same homogeneous and anisotropic layer. If long screens are installed, a packer should be used to reduce the length of the stress interval in order to get a small aspect ratio.

It is not recommended to conduct vertical interference slug tests in conventional wells without any special considerations for the installation of the sand pack. The radius and the permeability of the sand pack must be designed properly to avoid the sand pack from completely hiding the hydraulic response of the aquifer. Sand pack parameters can then be integrated into the model (analytical or numerical) to estimate aquifer parameters. In addition to reducing the risk of creating a skin around the well screen, well development operations should make sure to unclog the screen, but over-development that can remove too much fine sediment should be avoided.

4 Field test examples

4.1 Study site

To demonstrate the potential use of vertical interference slug tests to estimate K_v, a field study was designed and carried out to 1) apply the field testing procedure, 2) generate data to be interpreted using two approaches depending on the level of heterogeneity of the tested interval, and 3) validate hydraulic parameter estimates derived from interference slug tests by a comparison with results derived from a high resolution K_h depth profile measured with multilevel slug tests and K_v laboratory permeameter measurements using soil samples. The field test was conducted in a direct-push well installed into an aquifer composed primarily of surficial Quaternary sediments that were deposited and reworked by the Champlain Sea (Bolduc, 2003). The aquifer is located in Saint-Lambert-de-Lauzon along the St. Lawrence River, 40 km south of Quebec City, Quebec, Canada (Paradis *et al.*, 2010). Considering regional geological data and more than 25 cone penetration tests (CPT) with soil moisture resistivity (SMR), the hydrostratigraphy of the aquifer was found to be controlled by the sedimentary architecture of

a spit that formed in a littoral environment. At this site, the spit has a range of grain textures from fine sand to clayey silt (Figure II.7a). Sediment samples often show rapid transitions in grain-size over less than a decimeter. Vertical profiles of hydraulic conductivity and hydraulic heads indicate semi-confined conditions, which result from alternating sand and silt layers. The water table is generally 1 to 2 m below the ground surface. A till layer overlying bedrock and varying in depth from surface outcrop to 22 m acts as a hydraulic barrier at the base of the aquifer.

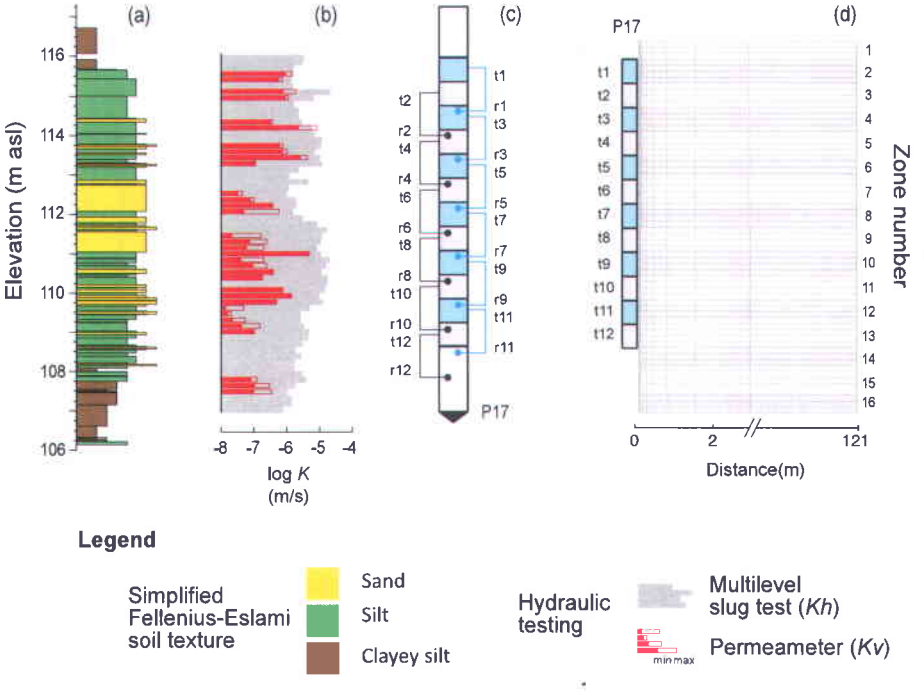


Figure II.7 (a) Hydrostratigraphy of the test well shown by cone penetration test (CPT) data translated into soil textures (according to the Fellenius and Eslami (2000) CPT soil classification chart). (b) Hydraulic conductivity data at a vertical resolution of 15 cm measured by multilevel slug tests (K_h) and permeameter tests on soil samples (K_v) over the same intervals. (c) Locations of tested intervals (61 cm) by vertical interference slug tests. (d) Numerical grid used for data inversion. In the test location numbers, *t* stands for transmitter (stress interval) and *r* stands for receiver (observation interval).

4.2 Direct-push well and sediments core sampling

A single test well was installed using a direct-push rig and the protected screen standard technique (ASTM, 2004). After well installation and development, the sediment was continuously sampled at a location 1 m away from the well using a piston-rod sampler hammered with a direct-push rig (Paradis *et al.*, 2010). This method allowed for the recovery of 8 cores over the thickness of the aquifer (38 mm-diameter and 1.52 m-long inside a transparent polyethylene teraphthalate (PETG) liner, 83% overall recovery).

4.3 Multilevel slug tests and laboratory permeameter tests

Multilevel slug tests were used to establish a K_h profile along the well. Multilevel slug tests involve the use of packers to isolate a screened interval of a well to conduct a slug test (e.g., Butler, 1998; Ross and McElwee, 2007; Zemansky and McElwee, 2005). A total of 60 K_h estimates were obtained from multilevel slug tests at 15 cm interval (Figure II.7b). Paradis *et al.* (2010) provide a more a detailed description of the test assembly and procedure.

As proposed by Johnson *et al.* (2005), an automated falling head laboratory permeameter was set up by installing pressure transducers at the base of multiple falling head devices (reservoir) and recording changes in hydraulic head with time with the pressure transducers. To estimate K_v values of soil samples, permeameters were designed to accommodate the 38 mm-diameter PETG liner containing soil samples with varying length. This was done to reduce sample manipulation in order to minimize disturbance of original sediments. Each 1.52 m-long sample were subdivided into 15 cm-long subsamples. The subdivision was coincident with the test intervals used in the multilevel slug tests and vertical interference slug tests.

Subsamples were then compacted at a pressure corresponding to the sampling depth with a hydraulic press. Occasional voids in the subsamples were locally repacked to avoid a hydraulic short-circuit along the sample sleeve. Each falling head device was placed over a subsample saturated with water at room temperature to initiate a test. The rate of decline of the water level in the reservoir was then used to calculate K_v using a solution based on Darcy's law (Fetter, 2001). The total number of subsamples tested with falling head permeameter was 34 (Figure II.7b). Duplicate tests (2 to 3) that were conducted on numerous subsamples indicated a high degree of reproducibility (coefficient of determination of 78 %). All K_v values were temperature-corrected to a groundwater temperature of 8 °C representative of field conditions (Sudicky *et al.*, 2010).

4.4 Vertical interference slug tests

For this study, inflatable packers were built over 2.54 cm inside-diameter polyvinyl chloride (PVC) tubing (Figure II.8a). Threads on the PVC tubing allowed the use of variable screen lengths between packers. An airline was attached to the packers and connected at the surface to an air compressor to inflate the packers to the desired pressure (510-620 kPa). The dual-packer assembly was also connected

by a 2.54 cm inside-diameter PVC riser pipe to the surface, within which water levels were monitored with a pressure transducer during the slug test. A rigid tape was attached near the top to accurately locate the position of the straddled interval. The arrangement of screens used for this study is shown in Figure II.8. The screens corresponding to the stress and observation intervals are 61 cm and 30 cm-long, respectively, while the separation between the bottom of the stress interval and the top of the observation interval is 61 cm. Vertical interference slug tests were made at 61 cm intervals at elevations coincident with the intervals tested by the multilevel slug tests and the permeameter tests.

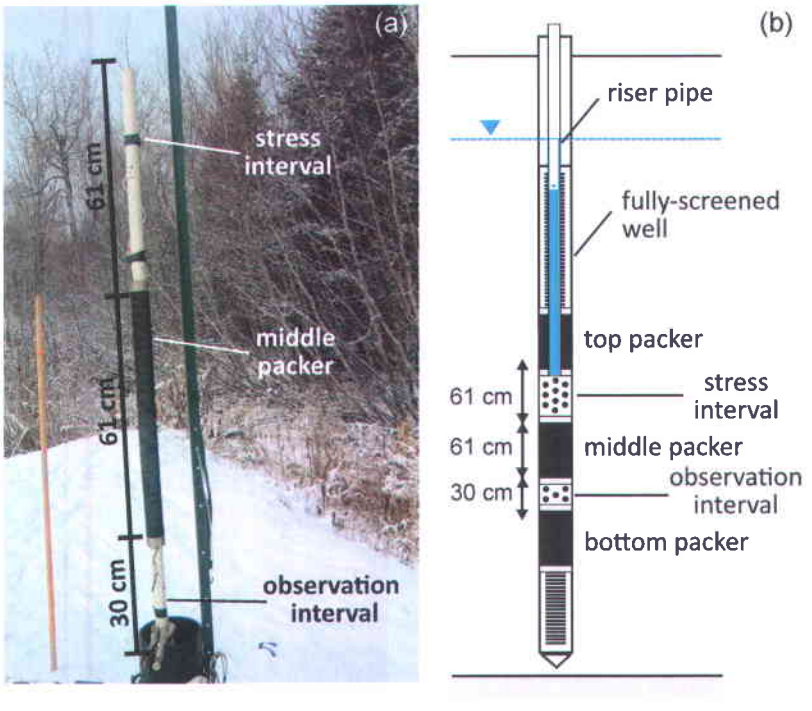


Figure II.8 Picture (a) and (b) schematic representation of the vertical interference slug test assembly used for the field experiment.

Once packers were inflated, to facilitate test interpretation, a trend-monitoring period was continued until stabilization of pressure transducers (pressure interval 0-10 m; minimum accuracy and resolution 0.5 cm and 0.00006 m, respectively) located within the stress and the observation intervals as well as the air pressure applied to the packers.

After this period, slug tests were carried out using a pneumatic method to induce the initial lowering of the water level (Levy and Pannell, 1991). For this purpose, a wellhead assembly was attached to the top of the riser pipe that contained an airtight adapter that allowed a transducer cable to pass and a ball valve

for an instantaneous release of pressure. An air compressor was also connected to the wellhead assembly to increase air pressure in the riser. A calibrated digital air pressure gauge was used to accurately set the desired initial hydraulic head change for the slug test and to verify air pressure stabilization before initiating the slug test. For each test, the initial drawdown was as high as practical to produce a response in the observation interval that is larger than the head resolution of the pressure transducer. The initial head induced for all tests with this device ranged from 1.06 to 7.78 m (Table II.2).

Table II.2 Summary of vertical interference slug tests carried out along the well. Prefix *t* is for transmitter (stress interval) whereas *r* is for receiver (observation interval). *: values for data recorded in the observation intervals. The scaling factor is used in the Levenberg-Marquart algorithm to account for the discrepancy between stress and observation head responses.

Test #	Initial head in stress interval (m)	Maximum observed *H/Ho* (-)*	Maximum observed amplitude (m)*	Scaling factor (stress/obs.)	Early time data interval removed (second)*
t1r1	1.06	0.005	0.005	0.1/0.001	0
t2r2	1.49	0.002	0.003	0.1/0.001	0
t3r3	1.15	0.002	0.003	0.1/0.001	0
t4r4	2.79	0.012	0.034	0.1/0.005	0
t5r5	3.59	0.028	0.099	0.1/0.005	0
t6r6	3.65	0.004	0.015	0.1/0.005	70
t7r7	4.25	0.011	0.045	0.1/0.005	25
t8r8	4.83	0.010	0.049	0.1/0.005	25
t9r9	4.44	0.032	0.142	0.1/0.005	10
t10r10	6.85	0.021	0.144	0.1/0.005	8
t11r11	7.05	0.066	0.468	0.1/0.005	20
t12r12	7.78	0.425	3.306	-	-

A total of 12 vertical interference slug tests were carried out to measure K_v along the well (Figure II.7c). Vertical interference slug tests were conducted systematically at successive 61 cm interval to obtain data over the entire well screen length. Figure II.9 shows recorded data for the 12 tests whose duration was a maximum of 45 minutes. As summarized in Table II.2, drawdowns measured in observation intervals were often quite small ranging from 0.30 to 46 cm. This emphasizes the importance of designing a slug test device able to induce a large stress on the aquifer combined with the need to use very sensitive pressure transducers. As with any slug test, varying and repeated head changes were imposed for quality control and to verify repeatability (Butler *et al.*, 1996).

Testing of a single well by vertical interference slug tests can be carried out at any depth increment without extracting the test assembly from the tested well. Therefore, testing time for a specific well depends on testing time for a single interval, which includes several steps: repositioning the test assembly, inflation of packers, waiting for pressure stabilization within tested intervals by following real-time transducer measurements, running tests with several hydraulic heads, and deflation of packers. The duration of the various steps involved in vertical interference slug tests can vary, but in moderate hydraulic conductivity silts and sands, a single interval can be tested in less than 60 minutes. This time decreases for more permeable formations. For a measure of hydraulic properties over a different scale, the test assembly can be modified after removing the assembly outside of the well. Modification, under field conditions, can usually be accomplished within a few minutes by simply exchanging the middle packer.

4.4.1 Analysis by inverse modeling for heterogeneous cases

To handle the heterogeneous and anisotropic conditions observed along the study well, we analyzed vertical interference slug test data using the lr2dinv (linear and radial 2D inversion) code developed by Bohling and Butler (2001). The program consists of a two-dimensional finite-difference flow model coupled with a Marquardt-Levenberg regularization of a least-squares estimation of hydraulic parameters. The code uses a logarithmic transformation of radial coordinates to transform the radial flow problem into Cartesian coordinates (axis symmetric). The linear problem is then solved on a rectangular grid, with regular grid increments in the vertical and transformed radial directions. The basic hydraulic parameters used in the model are K_h, K_v/K_h and S_s. A separate zonation of the model grid can be specified for each parameter to simulate complex aquifer conditions, and each parameter can be defined as known or unknown on a zone-by-zone basis. The code allows for simultaneous analysis of head responses from any number of tests, each represented by different sets of initial and boundary conditions. In addition, the program considers wellbore storage and the placement of packers in wells.

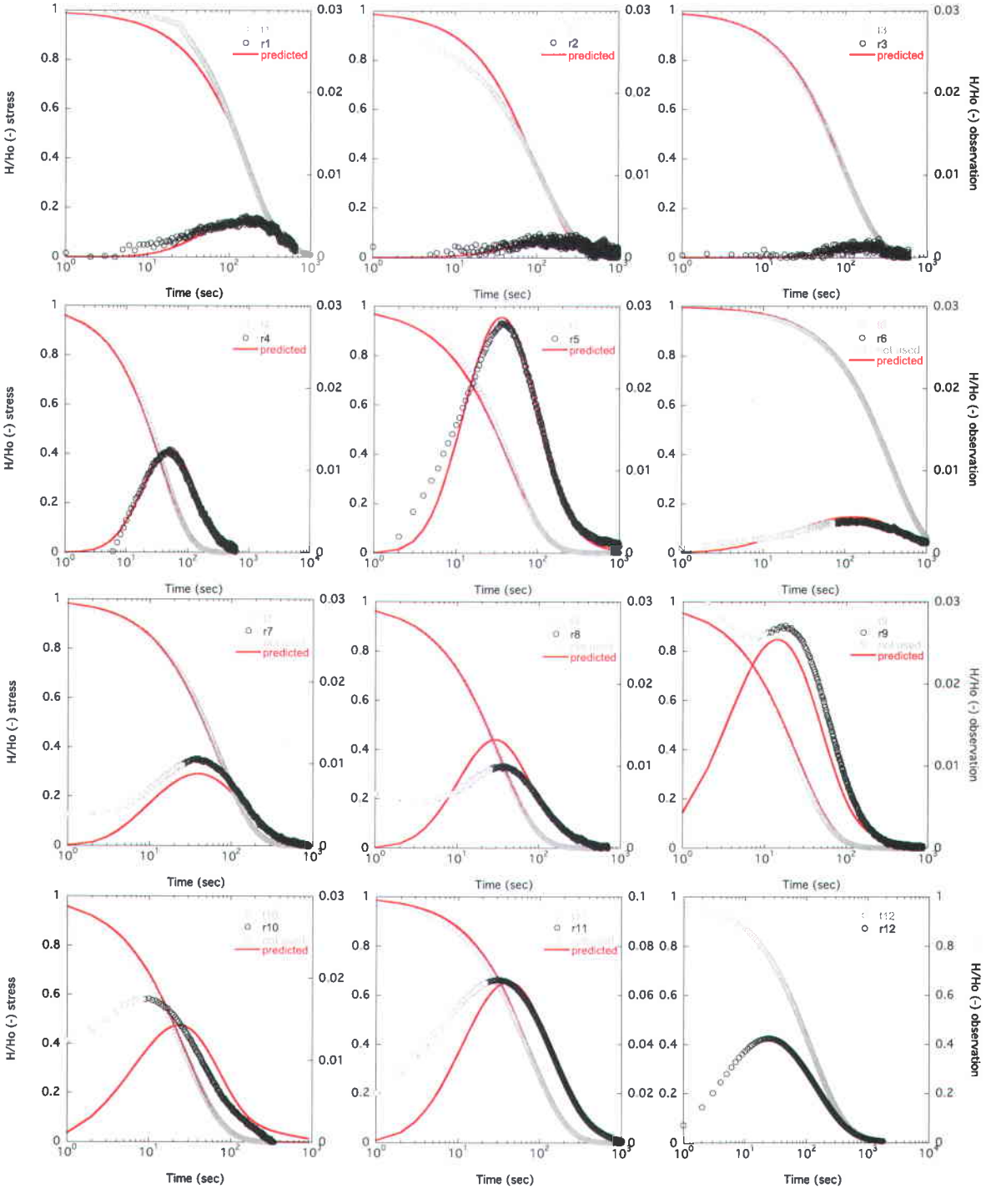


Figure II.9 Vertical interference slug test responses measured (symbols) and predicted by inversion (curves). Graphs show each of the 12 stress interval responses (grey circles) along with the responses from observation intervals (black circles). Grey squares with a circle inside correspond to data that were not used in the inversion due to apparent DSC effects (see Figure II.6). Scales of observation interval responses are not the same as those of other tests for *t11* and *t12* (larger scale).

1.1.1.1 Numerical grid

The domain used for the inversion was composed of variable-size rectangular elements and measured 9.45 m × 121.5 m (Figure II.7d). A simulation grid with 14 cells of dimensions ranging from approximately 0.021 m to 55 m was used along the transformed radial axis and 62 cells of 0.1524 m thickness were used along the vertical axis. The grid was adjusted to fit the locations of the packers used during the field experiment reported in this paper.

The grid was thereafter subdivided into 16 layers with a thickness of 61 cm, which correspond to the length of the stress interval used for the tests (Figure II.7d). The comparison of the first two stress interval responses (*t1r1* and *t2r2*) with previous slug tests carried out earlier for another study suggests that a low-K skin is present. The first column of cells adjacent to stress intervals t1 and t2 were then simulated as a skin with imposed values of 1×10^{-6} m/s, 1 and 1×10^{-6} m^{-1} for K_h, K_v/K_h and S_s, respectively. Values of hydraulic parameters of these 16 layers were not constrained during the inversion except for the cells skin.

The discretization used here places the outer boundary of the model about 90 m (at node location) from the tested well, so that this boundary has negligible effect on transient head simulations within the time frame of the slug tests and on the corresponding simulations. The imposed boundary conditions were no flow for the outer and the bottom boundaries and constant head for the top boundary (unconfined conditions). The initial head condition was assumed to be at steady-state prior to each slug test and a dimensionless drawdown initial value of 1 was applied at the stress interval topmost cell to simulate the slug test. Packers used to isolate the stress and the observation intervals were modeled as infinitely impermeable (K_h=K_v=1×10^{-20} m/s).

To account for the influence of wellbore storage on slug test responses, wellbore storage effects were simulated for each stress interval (inner boundary). This was accomplished by specifying large values of K_h and K_v for the stress interval cells, a S_s value of 1 for the upper cell and a S_s and a value of 0 for the remaining cells. Large values of hydraulic conductivity were used to indicate an infinitely conducting wellbore. The S_s value of 1 was used to indicate that the hydraulic stress, which is applied to the topmost cell, would translate directly into the wellbore, whereas the S_s value of 0 was used to represent the relative incompressibility of water. For the observation intervals, no wellbore storage effects were simulated because all observation intervals were straddled with packers (Novakowski, 1989).

1.1.1.2 Inverted dataset

Eleven of the twelve vertical interference slug test datasets presented in Figure II.9 were inverted to produce heterogeneous profile of K_h, K_v and S_s along the well (stress and observation intervals for each test shown in Figure II.7b). Test *t12r12* was not used in the inversion due to a strong evidence of the presence of a hydraulic short-circuit acting for the entire duration of the test. This test was carried out in the lower portion of the aquifer where clayey silt is mostly present. The cohesiveness of that sediment might have prevented the collapse of the formation on this section of the direct-push well (Figure II.7a). For each of the 11 slug tests, both stress and observation responses were simultaneously inverted. To account for the scale discrepancy between stress and observation responses, a weighted regression was carried out. The scaling factors of the Levenberg-Marquardt algorithm associated with each stress and observation responses were then adjusted according to the relative amplitude of the various responses (Table II.2).

All original 1 sec-interval measurements were used for the inversion, except early measurements affected by DSC (Table II.2 and Figure II.9). It should be noted that a DSC could not be modeled with the current code, which explains why early data had to be removed. A total number of 23670 points were then inverted.

1.1.1.3 Results

Predicted drawdowns from the inversion are compared to observed values in Figure II.9 as well as final K_h, K_v and S_s profile along the well in Figures II.10a-c, respectively. The inverse model was run for the first eleven slug tests listed in Table II.2 until the convergence criterion was met, which occurs when the difference between predicted and observed normalized drawdowns is less than 1×10^{-3}. All test data were included in the inversion simultaneously and started with uniform K_h, K_v/K_h and S_s fields (K_h=1×10^{-5} m/s, K_v/K_h=0.01 and S_s=1×10^{-5} m^{-1}). No constraint was imposed on the inversion, except the low-K skin adjacent to stress intervals *t1* and *t2* that was simulated as an isotropic layer with K_h and S_s values lower and higher than the values of the formation, respectively ($K_h' = K_v' = 1\times10^{-6}$ m/s and S_s'=1×10^{-3} m^{-1}). For instance, the clogging of the formation by sediment in place reduces the K of the formation, whereas the non-compact nature of the skin allows the storage of more water under a unit change in hydraulic head. Lower K' value or higher S_s' value of the skin did not significantly change the hydraulic responses. The duration of the simulation was 4000 s to give enough time to reproduce all field tests. The inversion was

done on a OSX (64-bit) operating system running on an Intel Quad Core i7 processor operating at 2.2 GHz with 8 GB of RAM. The total computational time for inverting the data was 201 min.

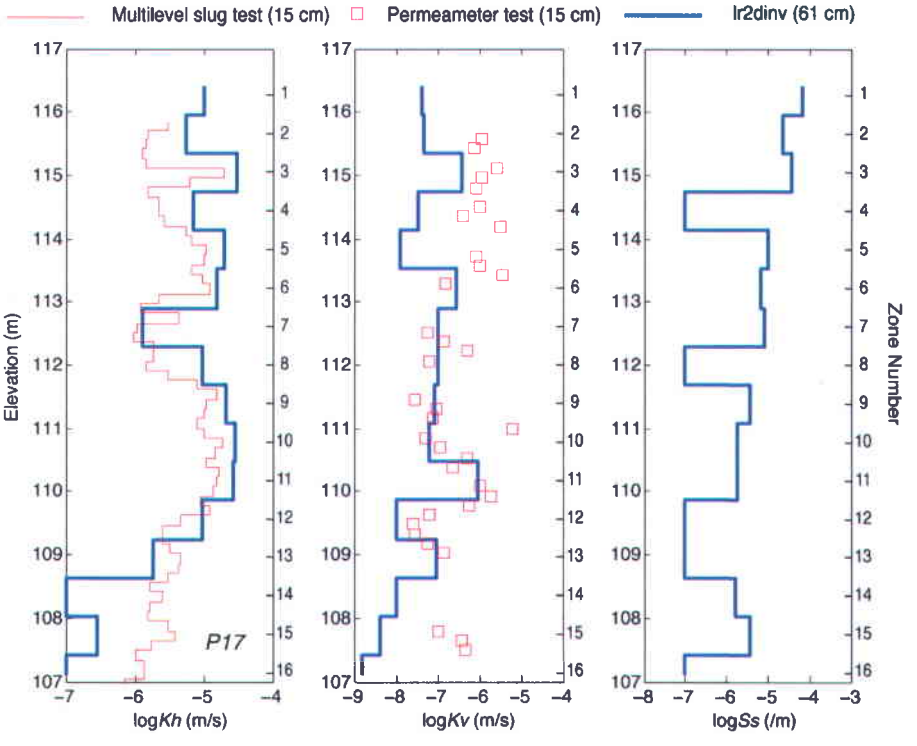


Figure II.10 Vertical profiles of observed logK_h (from slug tests) and logK_v (from cores) along the well with corresponding profiles predicted by the simultaneous numerical inversion of multiple vertical interference slug tests.

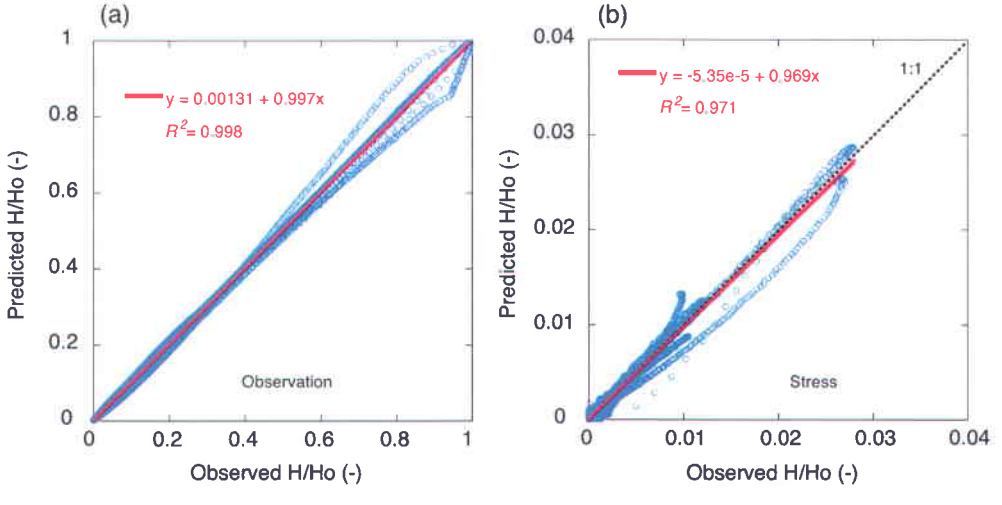


Figure II.11 Scatter plot of predicted versus observed dimensionless drawdowns (H/H_o) for the 12 vertical interference slug tests. Separate scatter plots are shown for the stress intervals (a) and the observation intervals (b). The dotted line is a reference 1:1 line indicating a perfect match and the solid red line is the linear regression of data.

To examine the quality of the numerical inversion, predicted and observed drawdowns are first compared for the 11 inverted tests in all stress and observation intervals (Figures II.11a and II.11b, respectively). Examination of Figures II.11a and II.11b reveals that the data tends to closely fit the perfect fit 1:1 line with a slope (m) of 0.997 and 0.969 and a R^2 of 0.998 and 0.971 for all stress and observation intervals, respectively. As presented in Figures II.11a and II.11b, the inversion captures the general behavior of the drawdown data and only a few tests present large scatter and bias (Figure II.9; Table II.3). For instance, the slope m between observed and predicted drawdowns for individual tests are ranging from 0.954 to 1.072 (stress) and 0.370 to 1.238 (observation), whereas R^2 values go from 0.998 to 1.000 (stress) and 0.741 to 0.999 (observation) (Table II.3). The fit for stress intervals is generally better than for observation intervals because observation responses are more sensitive to well construction (Figure II.3c), S_s (Figure II.2b) and K_v/K_h (Figure II.2c).

The validity of predicted K_h and K_v values was assessed by comparing K_h and K_v profiles along the well to available K_h values from multilevel slug tests and K_v measurements from permeameter tests (Figure II.10a and II.4). The K_h profile obtained from vertical interference slug tests follows the same general trend as K_h values measured with multilevel slug tests. Values for zone number 1 and 13 to 16 should be excluded for this comparison because they were not tested. K_h values from interference tests measured over a screen length of 61 cm are shifted toward higher K_h values obtained with multilevel slug tests at 15 cm interval. This is in accordance with the principle of parallel flow, which implies that upscaled K_h values are weighted toward the most permeable layers (Freeze and Cherry, 1979). Nevertheless, predicted K_h values are still systematically slightly higher than multilevel slug test values. This bias is explained by the fact that K_h values obtained with multilevel slug tests were estimated with the Bouwer and Rice (1976) method assuming isotropic conditions. For instance, neglecting K_v/K_h ratio generally results in K_h underestimation that is proportional to the K_v/K_h ratio (Zlotnik, 1994).

The predicted K_v values by inversion presents also a profile similar to the measured K_v values with permeameter tests (Figure II.10b and Table II.4). The predicted values tend toward smaller K_v values obtained with permeameter tests, which follows the serial flow principle. In opposite to the parallel flow principle, upscaled K_v values are weighted toward lower K_v values. Excluding zones 1, 2 and 14 to 16 that were not constrained vertically by interference tests, the only discrepancies with permeameter tests are for zones 4 and 5. For these zones, predicted values are more than one order of magnitude lower than laboratory measurements. Two hypotheses have been explored to explain those differences. First, the widening of the hole during direct-push operations to install the well might have favored the creation of a

skin around the screen. This might have occurred especially in the first meter of the saturated zone where sediments from upper layers might have filled the void around the screen. In addition, the development operations might have been not efficient for the first two meters due to the dewatering of the well caused by pumping. However, all skin modeling attempts failed to improve the fit between predicted and laboratory K_v values for zones number 4 and 5. The only skin modeling that was efficient was for zones number 1 to 3, where the presence of low-K skin was evident.

Table II.3 Validation statistics for observed and predicted drawdowns for each vertical interference slug test used in the inversion. R^2 and m stand for coefficient of determination and slope of the linear regression curve, respectively. The root-mean-square error (*RMS*) is presented in percentage of the maximum head recorded in the respective stress interval.

Test	Stress Interval			Observation Interval		
	R^2	m	*RMS* (%)	R^2	m	*RMS* (%)
t1r1	0.999	0.954	1.58	0.949	0.961	0.036
t2r2	0.998	1.072	1.77	0.773	0.697	0.047
t3r3	1.000	1.005	0.39	0.741	0.370	0.044
t4r4	1.000	0.972	0.52	0.994	1.009	0.056
t5r5	1.000	0.961	0.69	0.999	1.052	0.072
t6r6	1.000	1.006	0.53	0.995	1.238	0.032
t7r7	1.000	0.984	0.47	0.999	0.814	0.053
t8r8	1.000	1.008	0.20	0.986	1.129	0.065
t9r9	0.999	1.045	0.70	0.993	0.819	0.120
t10r10	1.000	1.032	0.46	0.974	0.845	0.200
t11r11	1.000	1.034	0.62	0.999	0.994	0.077
t12r12	-	-	-	-	-	-
All Tests	0.998	1.002	0.88	0.986	0.981	0.082

The other explanation is from the experimental procedure related to permeameter tests in the laboratory, especially those related to samples collection and manipulation. The major challenge with permeameter

tests on minimally disturbed samples (or non-recompacted samples) is to collect cores that do not contain any voids or preferential flow paths. The two most important factors that control the quality of core recovery are sediment cohesiveness and grain size. Generally, loose sediments with large grain size, such as coarse sand and gravel, are not favorable to the collection of good cores. Specifically for this study, the quality of recovery for the core collected between elevations 114 to 116 m was very poor. It is expected that a coarse sand layer of about 30 cm at elevation 115 m affected the quality of the core. This sand layer, which is indicated by high K_h values in Figure II.10, was only partially recovered and left a hole in the core. Many permeameter tests on subsamples of this core were thus rejected because K_v values were much higher than K_h values estimated with multilevel slug tests on the same intervals. Despite K_v values for other subsamples that were not rejected for this core because they appeared to be valid, we may suspect by association that K_v values between elevations 114 to 116 m are overestimated.

For the S_s profile, there are no independent measurements to compare with. The modeled profile however shows a general decrease with depth (Figure II.10c and Table II.4). This can be related to an increase of the overburden load with depth, which reduces the elastic property of sediments and thus their specific storage capacity.

Finally, examination of Figure II.12 reveals that the dimensionless variances in the estimates of K_h are consistently lower than those for K_v/K_h and S_s, indicating that there is greater confidence in the estimates of K_h relative to the estimates of the other two hydraulic properties. The dimensionless variances represent the uncertainties relative to the estimated parameter values. Dimensionless variances are scaled by the parameter estimates themselves and thus are not influenced by the differing magnitudes of the estimates (Bohling and Butler, 2001). All variance profiles show similar general trend with important variance increase in the uppermost and lowest zones, where no interference tests were done. Excluding the shallowest and deepest zones (zone number 1 and 14 to 16), the average of the logarithm of the dimensionless variance for K_h is -1.10, whereas it is 0.09 and 0.62 for K_v/K_h and S_s, respectively.

Moreover, K_v/K_h and S_s variances have similar magnitude, but the S_s variance profile presents some zones (4 and 8) with values that depart from the main trend. Those zones correspond to the same zones depicted in Figure II.10c, where S_s values are much higher than the average value. As a result, there is thus less confidence in estimated S_s values for those zones. The similarity in K_v/K_h and S_s variance profiles is related to the fact that K_v/K_h and S_s have similar effects on observation responses in vertical interference slug tests (Figures II.2b and II.2c). Moreover, K_v/K_h and S_s variances are higher for the few

last interference tests (zones 11 to 13), which also correspond to the tests with the poorest match between predicted and observed drawdowns (Figure II.9).

Table II.4 Summary of hydraulic tests used to estimate vertical hydraulic conductivity (K_v). Interval length for vertical interference slug test is 61 cm. Permeameter and multilevel slug tests hydraulic conductivity is upscaled to 61 cm.

Test	Lab[1]	Multilevel Slug Test		Vertical Interference Slug Test		
	K_v (m/s)	K_h (m/s)	K_v (m/s)	K_h (m/s)	K_v (m/s)	S_s (m^{-1})
t1r1	9.4×10^{-7}	1.8×10^{-6}	1.7×10^{-6}	5.7×10^{-6}	4.5×10^{-8}	2.2×10^{-5}
t2r2	1.1×10^{-6}	7.4×10^{-6}	2.6×10^{-6}	3.0×10^{-5}	4.0×10^{-7}	3.9×10^{-5}
t3r3	7.5×10^{-7}	3.2×10^{-6}	2.8×10^{-6}	7.1×10^{-6}	3.6×10^{-8}	1.0×10^{-7}
t4r4	9.1×10^{-7}	8.9×10^{-6}	8.4×10^{-6}	2.1×10^{-5}	1.3×10^{-8}	1.1×10^{-5}
t5r5	2.9×10^{-7}	6.2×10^{-6}	2.7×10^{-6}	1.6×10^{-5}	3.0×10^{-7}	7.0×10^{-6}
t6r6	8.3×10^{-8}	2.1×10^{-6}	1.5×10^{-6}	1.3×10^{-6}	1.1×10^{-7}	8.5×10^{-6}
t7r7	6.1×10^{-7}	3.6×10^{-6}	2.4×10^{-6}	9.9×10^{-6}	1.0×10^{-7}	1.0×10^{-7}
t8r8	5.1×10^{-8}	1.1×10^{-5}	1.1×10^{-5}	2.1×10^{-5}	8.5×10^{-8}	3.6×10^{-6}
t9r9	1.4×10^{-7}	1.4×10^{-5}	1.3×10^{-5}	2.8×10^{-5}	6.6×10^{-8}	1.8×10^{-6}
t10r10	5.0×10^{-7}	1.4×10^{-5}	1.3×10^{-5}	2.7×10^{-5}	9.4×10^{-7}	1.8×10^{-6}
t11r11	4.6×10^{-8}	5.5×10^{-6}	3.7×10^{-6}	9.5×10^{-6}	1.1×10^{-8}	1.0×10^{-7}
t12r12	8.1×10^{-8}	3.8×10^{-6}	3.7×10^{-6}	1.9×10^{-6}	9.3×10^{-8}	1.0×10^{-7}

1: Laboratory data from permeameter measurements on soil samples.

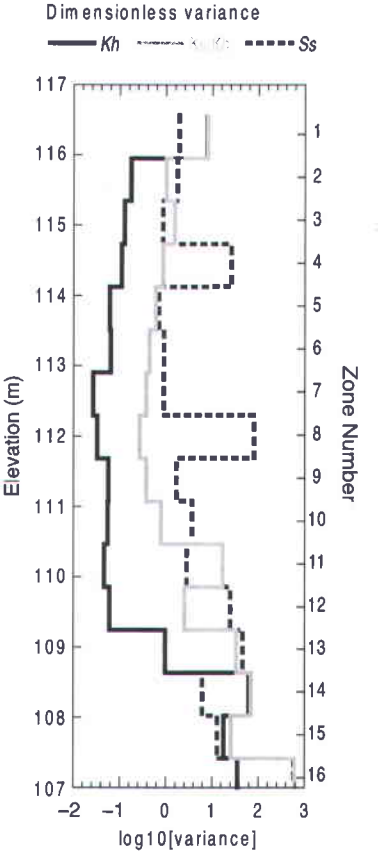


Figure II.12 Vertical profiles of dimensionless variance along the well for hydraulic parameters predicted by the inversion: (a) horizontal hydraulic conductivity (K_h); (b) hydraulic conductivity anisotropy (K_v/K_h); and (c) specific storage (S_s).

4.4.2 Analysis by type-curve matching for homogeneous cases

For homogeneous and anisotropic aquifers, a type-curve matching procedure with the KGS model can be used to interpret vertical slug test data. The test analysis procedure involves the simultaneous match of both stress and observation interval responses to resolve K_v/K_h, S_s and K_h. Aquifer parameters for the entire test assembly length (from the top of the stress interval to the bottom of the observation interval) must be reasonably assumed homogeneous, else both stress and observation responses could not be matched together. Such analysis is thus restricted to homogeneous and anisotropic cases.

This procedure was applied to the interpretation of data from test *t8r8*, which is a relatively homogeneous but anisotropic interval along the well (Figures II.7 and II.10). The match between observed head responses and the KGS model is depicted in Figure II.13. It is important to note that skin hydraulic properties are not modeled in this analysis and that the material around the screen is considered

undisturbed. For this test, the early-test times dimensionless head data in the observation interval suggest the influence of a high-K skin or a DSC. As the dynamic short-circuit speeds up mainly early-test times, the data match was done for late-test times data measured from and beyond the peak amplitude of dimensionless head in the observation interval. The value of K_v/K_h obtained for interval t8r8 was 0.0061 for an estimated K_v value of 1.2×10^{-7} m/s. In comparison, the K_v value obtained by inversion for this interval (zone number 9) was 8.5×10^{-8} m/s.

Type-curves for K_v/K_h values one order of magnitude above and below the best match are also shown on Figure II.13 to indicate the sensitivity of hydraulic responses to K_v/K_h. As discussed in the sensitivity analysis, we observe that K_v/K_h mainly influences late-test times in the stress interval and modify the slope of the curve without changing the response at early-test times. A small modification of this slope results in a large change in the amplitude and arrival time of the dimensionless head in the observation interval. These observations combined with a complete and an accurate measurement of hydraulic responses in both the stress and the observation intervals as well as the careful design of a small aspect ratio for the stress interval allow the resolution of distinct effects and values for K_v/K_h, S_s and K_h.

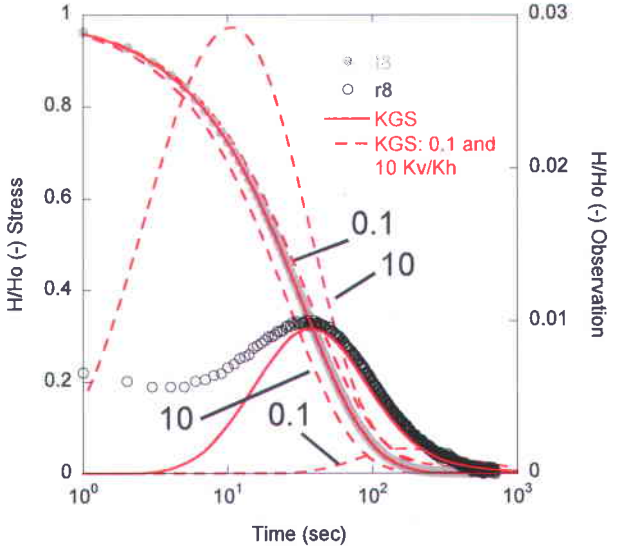


Figure II.13 Observed vertical interference slug test response in the stress (*t8*) and observation (*r8*) intervals compared to the data match (solid lines) obtained with the KGS model for test *t8r8* assuming homogeneous aquifer conditions. Dashed lines are KGS model responses obtained using K_v/K_h values that are 10 times lower and higher than the best match value (0.0061).

4.4.3 Equivalent vertical hydraulic conductivity from multilevel slug tests

A simple procedure to estimate K_v from field measurements is by invoking the serial flow principle on K_h values obtained from high-resolution multilevel hydraulic tests carried out along a single well (Freeze and Cherry, 1979). As an example, we calculated an equivalent K_v for the multilevel slug tests carried out at 15 cm interval (Table II.4). It is assumed that 15 cm interval from multilevel slug tests are isotropic and homogeneous at this scale. We thus applied a harmonic mean on the four (when available) 15 cm intervals comprised within each 61 cm interval of the vertical interference slug tests. The comparison in Table II.4 of rescaled multilevel slug tests with K_v values with permeameter and vertical interference tests shows a discrepancy of up to two orders of magnitude in K_v estimates. This is in agreement with the depositional environment of the study area, where fine centimeter-scale transitions in grain-size were often observed in sediment samples. For instance, as multilevel slug tests induce predominantly horizontal flow patterns, the presence of thin horizontal low hydraulic conductivity layers does not significantly reduce K_h (parallel flow), whereas these layers essentially control the flow in the vertical direction (serial flow) and then lead to low K_v values for permeameter and vertical interference slug tests.

To be useful for K_v estimation, the vertical resolution of the multilevel slug tests should have been as small as the heterogeneity of the sediments (approximately 1-10 cm). For the study site, this situation would not have been practical with multilevel slug tests. The HRK tool proposed by Liu *et al.* (2009) could be however an interesting alternative. This tool, which is the coupling of a direct-push permeameter (Butler *et al.*, 2007; Lowry *et al.*, 1999) with a direct-push injection logger (Dietrich *et al.*, 2008), has the potential to measure K_h at a vertical resolution of approximately 1.5 cm.

5 Summary and conclusions

This paper explored the use of vertical interference slug tests for the efficient measurement of K_v at the field scale. The test is carried out in a single well between a stress and an observation interval that are isolated vertically with a three-packer assembly. Similar to inter-well interference slug tests, an instantaneous pressure pulse is induced in the stress interval and resulting drawdowns are recorded in both the stress and the observation intervals. In a field proof-of-concept study, twelve vertical interference tests were carried out sequentially along a fully-screened direct-push well in a moderately heterogeneous and highly anisotropic silts and sands littoral aquifer. The direct-push well was completed without sand-pack by the natural collapse of sediment around the screen in order to minimize well

construction interferences with the tests. The simultaneous numerical inversion of multiple tests was done to reconstruct heterogeneous profiles of K_h, K_v/K_h and S_s. For a homogeneous section of the aquifer a semi-analytical model was also applied. Results of the numerical inversion and the semi-analytical model were compared with permeameter tests carried out in the laboratory to measure K_v on sediment samples and with high-resolution multilevel slug tests to measure K_h at a fine scale. The key findings of this study are summarized below:

- Vertical interference slug test responses measured in both a stress and an observation interval contain all the information needed to resolve the values of K_h, K_v/K_h and S_s. The design of a stress interval with a small aspect ratio is however essential as well as the acquisition of high-quality drawdown data.
- Vertical interference slug test responses measured between isolated intervals along the same well may be affected by well construction. Well construction effects, such as hydraulic short-circuits, can however be detected and taken into account during the analysis. Direct-push wells, such as the one used in this study, seem suitable to carry out vertical interference slug tests in comparison to less well-adapted sand-packed wells.
- In a moderately heterogeneous and highly anisotropic aquifer, it is possible to analyze vertical interference slug tests using inverse modeling that accounts for wellbore storage effects. In such an environment, however, a sequential field-testing scheme should be followed where stress and observation intervals overlap. The simultaneous inversion of multiple tests yields K_h, K_v/K_h and S_s values at discrete intervals along the well, as well as their uncertainties.
- For cases for which the aquifer can be considered homogeneous at the scale of the vertical interference test assembly, a type-curve matching procedure with a semi-analytical model can also be used to resolve K_h, K_v/K_h and S_s values.
- Vertical interference slug tests hold the potential for K_v measurement under really challenging field conditions. Results of this study indicate that K_v values obtained with vertical interference slug tests are in agreement with permeameter tests conducted for the same tested intervals. Such a direct measurement is far more interesting than the calculation of an equivalent K_v from high-resolution multilevel slug tests assuming isotropic conditions at vertical resolution as low as 15 cm. Such a calculation for the tested well did not show a good correlation with both vertical interference slug tests and permeameter data, suggesting that K_v/K_h in the study aquifer can be quite high, even over a relatively small vertical scale.

In a practical sense, the vertical interference slug test is an extension of the single well multilevel slug test. It requires only the addition of an observation interval at the bottom of a multilevel slug test assembly. Vertical interference slug tests are however an improvement over conventional multilevel slug tests because they can provide not only K_h but also K_v/K_h and S_s from a single test, and thus contribute to provide the data required for meaningful understanding of flow and transport in aquifers.

Acknowledgements

Sincere thanks to J. Butler for his comments that helped improve a draft version of this manuscript as well as to G. Bohling for providing a modified version of lr2dinv. The authors would like to acknowledge the important technical support provided by D. Martin, C. DeLouis and J.-M. Ballard and recognize the invaluable in-depth reviews and constructive comments made by two anonymous reviewers. The first author would like sincerely thanking Y. Michaud, A. Rivera and D. Kirkwood for their support. This study was supported by the Geological Survey of Canada under the Groundwater Resources Inventory Program, by the Régie intermunicipale de gestion des déchets des Chutes-de-la-Chaudière and by a NSERC Discovery Grant held by R.L. This is Earth Science Sector contribution number 20100189.

References

ASTM. 2004. D6724: Standard Guide for Installation of Direct Push Ground Water Monitoring Wells. ASTM International, 9 p.

ASTM. 2010. D6725-04: Standard Practice for Direct Push Installation of Prepacked Screen Monitoring Wells in Unconsolidated Aquifers. ASTM International, 15 p.

Audouin, O., J. Bodin. 2007. Cross-borehole slug test analysis in a fractured limestone aquifer. Journal of Hydrology, 348(3-4): 510-523.

Barry, F., D. Ophori, J. Hoffman, and R. Canace. 2009. Groundwater flow and capture zone analysis of the Central Passaic River Basin, New Jersey. Environmental Geology, 56: 1593-1603.

Belitz, K., and W. Dripps. 1999. Cross-well testing in unconfined aquifers: A case study from the Sleepers River Watershed, Vermont. Ground Water, 37(3): 438-447.

Bolduc, A. 2003. Géologie des formations superficielles, Charny, Québec. Commission Géologique du Canada, Dossier public 1976, 1/50000 map sheet.

Brauchler, R., J. Cheng, M. Everette, B. Johnson, P. Dietrich, R. Liedl, M. Sauter. 2007. An inversion strategy for hydraulic tomography: Coupling travel time and amplitude inversion. Journal of Hydrology, 345(3-4): 184-198.

Brauchler, R., R. Hu, T. Vogt, D. Halbouni, T. Heinrichs, T. Ptak, M. Sauter. 2010. Cross-well slug interference tests: An effective characterization method for resolving aquifer heterogeneity. Journal of Hydrology, 384(1-2): 33-45.

Burns, W.A. 1969. New single-well test for determining vertical permeability. Journal of Petroleum Technology, 246: 743-752

Butler Jr., J.J., 1995. Slug tests with observation wells; extension of Hyder *et al.* (1994) Solution to case of head in an observation well. Kansas Geological Survey, Open-file Report no. 95–43, 14 pp.

Butler, J.J. Jr. 1998. The Design, Performance, and Analysis of Slug Tests. Lewis Publishers. 252 pp.

Butler, Jr. J.J., G.C. Bohling, Z. Hyder, C.D. McElwee. 1994. The use of slug tests to describe vertical variations in hydraulic conductivity. Journal of Hydrology, 156 (1-4) 137-162.

Butler, J.J., and McElwee, C.D., and Liu, W. 1996. Improving the quality of parameter estimates obtained from slug tests. Ground Water, 34(3) 480-490.

Butler, J. J., Jr., P. Dietrich, V. Wittig, and T. Christy. 2007. Characterizing hydraulic conductivity with the direct-push permeameter. Ground Water, 45(4), 409-419.

Dietrich, P., J. J. Butler Jr., and K. Faiß. 2008. A rapid method for hydraulic profiling in unconsolidated formations. Ground Water, 46(2), 323-328.

Falta, R.W., N. Basu, and P.S. Rao. 2005. Assessing impacts of partial mass depletion in DNAPL source zones: II. Coupling sources strength functions to plume evolution. Journal of Contaminant Hydrology, 79: 45-66.

Fellenius, B.H., and A. Eslami. 2000. Soil profile interpreted from CPTu data. Year 2000 Geotechnics, Geotechnical Engineering Conference, Asian Institute of Technology, Bangkok, Thailand, November 27-30, 2000, 18 p.

Fetter, C.W. Jr. 2001. Applied Hydrogeology. 4th edition, Prentice-Hall, Upper Saddle River, NJ, 691 p.

Freeze, R.A. and J.A. Cherry. 1979. Groundwater. Prentice-Hall, Inc. Englewood Cliffs, NJ. 604 p.

Freeze, A. and Witherspoon, P.A. 1967. Theoretical analysis of regional flow: 2. Effect of water-table configuration and subsurface permeability variation. Water Resources Research, 3(2) 623-634.

Gerber, R. E., and K. Howard. 2000, Recharge through a regional till aquitard: three-dimensional flow model water balance approach. Ground Water, 38(3): 410-422.

Goltz, M.N., J. Huang, M.E. Close, and M.J. Flintoft, L. Pang. 2008. Use of tandem circulation wells to measure hydraulic conductivity without groundwater extraction. Journal of Contaminant Hydrology, 100(3-4): 127-136.

Hart, D.J., K.R. Bradbury and D.T. Feinstein. 2006. The vertical hydraulic conductivity of an aquitard at two spatial scales. Ground Water, 44(2): 201-211.

Hirasaki, G.J. 1974. Pulse tests and other early transient pressure analyses for in-situ estimation of vertical permeability. Transactions of the AIME, 257: 75-90.

Hvilshoj, S., K.H. Jensen, and B. Madsen. 2000. Single-well dipole flow tests: Parameter estimation and field testing. Ground Water, 38(1): 53-62.

Hyder, Z., J.J. Jr. Butler, C.D. McElwee and W. Liu. 1994. Slug test in partially penetrating wells, Water Resources Research, 30(8): 2945-2957.

Hyder, Z., and J.J Jr. Butler. 1995. Slug tests in unconfined formations: An assessment of the Bouwer and Rice technique. Ground Water, 33(1): 16-22.

Johnson, D.O., F.J. Arriaga, and B. Lowery. 2005. Automation of a falling head permeameter for rapid determination of hydraulic conductivity of multiple samples. Soil Science Society of America Journal, 69(3): 828-833.

Kabala, Z.J. 1993. The dipole-flow test: a new single-borehole test for aquifer characterization. Water Resources Research, 29(1): 99-107.

Klute, A. and C. Dirksen. 1986. Hydraulic conductivity and diffusivity: laboratory methods. Methods of soil analysis. Part 1. In: Physical and Mineralogical Methods. ed. by A. Klute, Agronomy Monograph 9, American Society of Agronomy, 687-734.

Kruseman, G.P. and N.A. de Ridder. 1990. Analysis and evaluation of pumping test data, 2nd ed. ILRI Publication, 377 p.

Levy, B.S., and L. Pannell. 1991. Evaluation of a pressure system for estimating in-situ hydraulic conductivity. Ground Water Management No. 5, Proceedings of the 5th National Outdoor Action Conference, Las Vegas, NV, 31-45.

Liu, W.Z., and J.J. Jr. Butler. 1995. The KGS model for slug tests in partially penetrating wells. Version 3.0, Lawrence, Kansas: Kansas Geological Survey Comput. Ser. Rep. 95-1.

Liu, G., J. J. Butler Jr., G.C. Bohling, E. Reboulet, S. Knobbe, and D.W. Hyndman. 2009. A new method for high-resolution characterization of hydraulic conductivity. Water Resources Research, 45, W08202.

Lowry, W., N. Mason, V. Chipman, K. Kisiel, and J. Stockton 1999. In-situ permeability measurements with direct push techniques: Phase II topical report. Rep. SEASF-TR-98-207, 102 pp., Fed. Energy Technol. Cent., Dep. of Educ., Morgantown, W. Va.

Morin, R.H., D.R. LeBlanc, and W.E. Teasdale. 1988. A statistical evaluation of formation disturbance produced by well-casing installation methods. Ground Water, 26(2): 207-217.

Novakowski, K.S. 1989. Analysis of pulse interference tests. Water Resources Research, 25(11): 2377-2387.

Onur, M., P.S. Hegeman, and F.J. Kuchuk. 2004. Pressure–pressure convolution analysis of multiprobe and packer-probe wireline formation tester data. SPE Reservoir Engineering and Evaluation, 351-364.

Paradis, D., R. Lefebvre, R.H. Morin, E. Gloaguen. 2010. Permeability profiles in granular aquifers using flowmeters in direct-push wells. Ground Water, 49(4): 534-547, doi: 10.1111/j.1745-6584.2010.00761.x

Ptak, T. and G. Teutsch. 1994. Forced and natural gradient tracer tests in a highly heterogeneous porous aquifer: Instrumentation and measurements. Journal of Hydrology, 159: 79-104.

Riva, M., L. Guadagnini, A. Guadagnini, T. Ptak and E. Martac. 2006. Probabilistic study of well capture zones distribution at the Lauswiesen field site. Journal of Contaminant Hydrology, 88 (1-2): 92-118.

Ross, H.C., and C.D. McElwee. 2007. Multi-level slug tests to measure 3-D hydraulic conductivity distributions. Natural Resources Research, 16(1): 67-79.

Rushton, K.R., M. Owen, and L.M. Tomlinson. 1992. The water resources of the Great Oolite aquifer in the Thames Basin. Journal of Hydrology, 32(1-4): 225-248.

Sageev, A. 1986. Slug test analysis. Water Resources Research, 22(8): 1323-1333.

Satter, A., G. Iqbal and J.L. Buchwalter. 2007. Practical Enhanced Reservoir Engineering. PennWell Corporation, 706 p.

Sellwood, S.M., J.M. Healey, S. Birks, and J.J. Jr. Butler. 2005. Direct-push hydrostratigraphic profiling: Coupling electrical logging and slug tests. Ground Water, 43(1): 19-29.

Sheng, J.J. 2009. A new technique to determine horizontal and vertical permeabilities. Transport in Porous Media, 77: 507-527.

Spane, F.A. Jr. 1996. Applicability of slug interference tests for hydraulic characterization of unconfined aquifers: (1) analytical assessment. Ground Water, 34(1): 66-74.

Spane, F.A. Jr., P.D. Thorne and L.C. Swanson. 1996. Applicability of slug interference tests for hydraulic characterization of unconfined aquifer: (2) field test examples. Ground Water, 34(5): 925-933.

Stienstra, P. and J.K. van Deen. 1994. Field data collection techniques – Unconventional sounding and sampling methods. In: Engineering Geology of Quaternary Sediments. ed. by N. Rengers, Balkema, 41-55.

Sudicky, E.A., W.A. Illman, I.K. Goltz, J.J. Adams, and R.G. McLaren. 2010. Heterogeneity in hydraulic conductivity and its role on the macroscale transport of a solute plume: From measurements to a practical application of stochastic flow and transport theory, Water Resources Research, 46, W01508, doi:10.1029/2008WR007558..

Sutton, D.J., Z.J. Kabala, D.E. Schaad, and N.C. Ruud. 2000. The dipole-flow test with a tracer: a new single-borehole tracer test for aquifer characterization. Journal of Contaminant Hydrology, 44: 71-101.

Tongpenyai, Y. and R. Raghaven. 1981. The effect of the wellbore storage and the skin on interference test data. Journal of Petroleum Technology, 33: 151-160.

Toth J. 1963. A theoretical analysis of groundwater flow in small drainage basins. Journal of Geophysical Research, 68: 4785-4812.

USEPA. 2005. Groundwater sampling and monitoring with direct push technologies. Office of Solid Waste and Emergency Response, Washington, DC 20460, EPA 540/R-04/005, 67 p.

Wenzel, L.K. 1942. Methods for determining permeability of water-bearing materials, with special reference to discharging-well methods, with a section on direct laboratory methods and bibliography on permeability and laminar flow, by V.C. Fishel. Water Supply and Irrigation Paper no. 887. U.S. Geological Survey, Reston, VA.

Wu, J., C. Zheng, and C.C. Chien. 2005. Cost-effective sampling network design for contaminant plume monitoring under general hydrogeological conditions. Journal of Contaminant Hydrology, 77: 41–65.

Xiang, J. and Z.J. Kabala. 1997. Performance of the steady-state dipole-flow test in layered aquifers. Hydrological Processes, 11(12): 1595-1605.

Zemansky, G.M., and C.D. McElwee. 2005. High-resolution slug testing. Ground Water, 43(2): 222-230.

Zlotnik, V.A. 1994. Interpretation of slug and packers tests in anisotropic aquifers. Ground Water, 32(5): 761-767.

Zlotnik, V.A. 1997. Effects of anisotropy on the capture zone of a partially penetrating well. Ground Water, 35(5): 842-847.

Zlotnik, V.A., G. Ledder. 1996. Theory of dipole flow in uniform anisotropic aquifers. Water Resources Research, 32(4): 1119-1128.

Zlotnik, V.A., and B.R. Zurbuchen. 1998. Dipole probe: Design and field applications of a single-borehole device for measurements of vertical variations of hydraulic conductivity. Ground Water, 36(6): 884-893.

Zlotnik, V.A., B.R. Zurbuchen, and T. Ptak. 2001. The steady-state dipole-flow test for characterization of hydraulic conductivity statistics in a highly permeable aquifer: Horkheimer Insel site, Germany. Ground Wate 39(4): 504-51

ARTICLE III

SENSITIVITY AND RESOLUTION ANALYSIS OF THE INFORMATION CONTENT OF TOMOGRAPHIC SLUG TESTS

Résumé

Il a été démontré que la propagation des charges hydrauliques en réponse à un essai de perméabilité à choc hydraulique est sensible aux variations de propriétés hydrauliques (K_h, K_v/K_h et S_s) des matériaux à travers lesquels cette perturbation hydraulique se propage. La tomographie par choc hydraulique comprend la réalisation d'une série d'essais de perméabilité à choc hydraulique dans des intervalles isolés d'un puits avec l'enregistrement simultané des charges hydrauliques dans l'intervalle émetteur et dans des intervalles d'observation isolés entre obturateurs dans un puits adjacent. Cette étude a pour but d'évaluer quantitativement le contenu en information d'une analyse transitoire de tomographie par choc hydraulique par rapport aux propriétés hydrauliques, et donc d'estimer sa capacité à résoudre l'hétérogénéité des propriétés hydrauliques dans un plan vertical entre deux puits. Une analyse de sensibilité et de résolution de premier ordre a donc été réalisée en utilisant des données synthétiques reflétant les conditions d'un aquifère littoral avec une importante anisotropie de K. Les résultats indiquent que la perturbation hydraulique induite par un choc hydraulique produit des patrons transitoires de sensibilité distincts pour les intervalles émetteur et d'observation. Ceci permet la résolution indépendante de K_h, K_v/K_h et S_s à l'intérieur du plan d'investigation, avec une faible interférence des paramètres situés à l'extérieur de ce plan. La résolution spatiale effective des paramètres hydrauliques peut aussi être considérablement améliorée en utilisant à la fois les réponses des intervalles émetteurs et d'observations dans l'analyse. L'utilisant de la réponse des émetteurs réduit la corrélation entre les paramètres, ce qui serait difficile d'obtenir autrement, même en utilisant un réseau dense d'intervalles émetteurs et d'observation. Une analyse transitoire de tomographie par choc hydraulique permet donc d'imager l'hétérogénéité des propriétés hydrauliques à une bonne résolution en plus de considérer l'anisotropie de K, pour ainsi permettre de caractériser efficacement les propriétés hydrauliques contrôlant l'écoulement et le transport de masse dans les aquifères.

Accepté pour *Water Resources Research* le 25 novembre 2013; Version modifiée re-soumise le 8 mars 2014

Sensitivity and resolution analysis of the information content of tomographic slug tests

Daniel Paradis[1,2,*], Erwan Gloaguen[2], René Lefebvre[2] and Bernard Giroux[2]

1: Geological Survey of Canada, 490 rue de la Couronne, Quebec City, Canada G1K 9A9

2 : Institut national de la recherche scientifique, Centre Eau Terre Environnement (INRS-ETE), 490 rue de la Couronne, Quebec City, Canada G1K 9A9

*Corresponding author: Daniel Paradis

Phone: (418) 654-3713 Fax: (418) 654-2604 E-mail: dparadis@nrcan.gc.ca

Abstract

Hydraulic head propagation in response to a slug test has been shown to be sensitive to variations in hydraulic properties (K_h, K_v/K_h and S_s) of the material through which the head perturbation propagates. Tomographic slug tests involve a series of packer slug tests carried out in stressed intervals of a borehole and the simultaneous recording of heads in the stressed interval and in observation intervals between packers in an adjacent borehole. This study aims to quantitatively assess the information content about hydraulic properties associated with the transient analysis of tomographic slug tests, and thus their capability to resolve heterogeneous hydraulic properties in a vertical plane between two boreholes. A first-order sensitivity and resolution analysis is thus carried out using synthetic data reflecting a known littoral aquifer with strong K anisotropy. Results indicate that the head perturbation induced by a slug test produces distinct transient sensitivity patterns in stressed and observation intervals, thus allowing the independent resolution of K_h, K_v/K_h and S_s within the plane of investigation, with a weak interference from parameters radially outward from the observation well. Effective spatial resolution can be improved using both stressed and observation interval responses in the analysis. Stressed interval responses reduce the correlation between parameters, which would be difficult to achieve otherwise, even with a dense network of stressed and observation intervals. Based on a homogeneous parameter distribution, the estimated sensitivity matrix indicates that the inversion of tomographic slug tests shows a high potential to resolve heterogeneous fields of K_h, K_v/K_h and S_s.

Keywords: Hydraulic tomography, Slug tests, Aquifer characterization, Heterogeneity, Anisotropy, Jacobian matrix

1 Introduction

It is widely recognized that knowledge of the spatial distribution of aquifer and aquitard hydraulic properties within an aquifer system is essential to the understanding of its dynamics, which can provide the basis for sound groundwater management. Indeed, comprehensive aquifer characterization should ideally provide a three-dimensional model of the architecture of the aquifer system units and their respective hydraulic properties, including hydraulic conductivity (K), K anisotropy (ratio of vertical and horizontal K, K_v/K_h) and specific storage (S_s) [*Anderson*, 1989]. Hydraulic tomography, which is essentially the simultaneous analysis of inter-well hydraulic responses, measured at discrete and isolated intervals, to multiple hydraulic tests (pumping or slug tests), is increasingly recognized as a mean of imaging heterogeneity in hydraulic properties, such as preferential flow paths or impermeable barriers that control flow and transport in aquifers. A comprehensive summary of previously published studies dedicated to hydraulic tomography was recently made by *Cardiff and Barrash,* [2011].

Pumping tests are the most commonly used approach for test initiation of hydraulic tomography experiments.. Tomographic pumping tests drawdown data are either analyzed under steady state or transient flow conditions. For tomographic experiments using traditional constant-rate pumping tests, drawdown sensitivities to K and S_s in different regions of an aquifer have been shown to be fairly correlated, which may lead to difficulties in identifying hydraulic properties independently and in spatially resolving those properties [*Bohling*, 2009]. It is the sensitivity to a given hydraulic property at a given time and location, defined as the ratio of the change in drawdown to a change in its property value, and its magnitude along with its correlation with sensitivity to properties at surrounding locations that indicates the ability of a pumping test to estimate a property value from drawdown data [*Vasco et al.*, 1997]. Meanwhile, it has been recognized that varying the pumping rate scheme (magnitude and frequency) of a pumping test can increase parameter resolutions by increasing sensitivity magnitudes while simultaneously constraining sensitivity correlations [*Butler and McElwee*, 1990]. Such variable pumping rate approach takes advantage of the transient behavior of sensitivities that are generated, which allows hydraulic properties in a given region of an aquifer to be better distinguished from properties in other regions. It should be noted that the oscillatory pumping approach recently proposed by *Cardiff et al.* [2013] might also fulfill a similar function.

A slug test involves a sudden change in the volume of water (or hydraulic head) in a well, by adding or removing water through a stressed interval, which causes a distance and time decaying hydraulic head perturbation into an aquifer [*Cheng et al.*, 2009]. The sensitivity behavior related to a slug test could thus also be potentially exploited to increase parameters resolution. Indeed, slug tests exhibit transient sensitivity behaviors to hydraulic properties in any regions of an aquifer interrogated by the test because of the time decaying head perturbation [*McElwee et al.*, 1995a; 1995b]. While hydraulic tomography based on slug tests have been extensively studied using the concept of hydraulic travel time [*Brauchler et al.*, 2007; 2010 and 2011; *Hu et al.*, 2011] to estimates heterogeneity in hydraulic diffusivity (the ratio K/S_s), tomographic slug tests data have never been studied through a sensitivity analysis to assess their resolution potential for defining heterogeneity in hydraulic properties. Moreover, previous studies on hydraulic tomography have essentially focused on identification of heterogeneity in K_h, and to a lesser extend in S_s. While aquifers could exhibit hydraulic conductivity anisotropy at various scales, and may influence K_h and S_s estimations if not taken into account [e.g., *Hyder et al.*, 1994; *Zlotnik*, 1994; *Butler*, 1995; *Spane*, 1996; *Belitz and Dripps*, 1999; *Butler and Zhan*, 2004; *Paradis and Lefebvre*, 2013], its explicit characterization using hydraulic tomography has never been the object of an analysis.

The main objective of this study is thus to assess the information content associated with a transient analysis of tomographic slug tests data to estimate heterogeneity in K_h, K_v/K_h and S_s. As widely used to assess geophysical inversion results [e.g., *Menke*, 2012; *Aster*, 2005], this study uses the sensitivity and resolution matrices to quantify the potential for tomographic slug tests data to resolve heterogeneity in K_h, K_v/K_h and S_s between wells. While sensitivity analysis is used to understand the fundamental physics underlying a particular hydraulic experiment, resolution analysis can be used to describe how well the model can be reproduced when solving the inverse problem of that particular experiment [*Vasco et al.*, 1997]. Resolution analysis is an extension of the sensitivity analysis and combines sensitivities from a number of head measurements and for a number of properties into a single measure related to each hydraulic property. The resolution matrix describes the degree of averaging contained in the estimate of a parameter at a particular location and it indicates whether a given parameter can be resolved or not for this particular experiment. In well hydraulic studies, the sensitivity and resolution matrices have been used by *Vasco et al.* [1997] and *Clemo et al.* [2003] to examine the ability of single or multiple fully penetrating pumping tests with multiple observation wells to resolve two-dimensional areal variations in transmissivity, and by *Bohling* [2009] to compare the sensitivity and resolution associated with transient and steady-shape approaches to the analysis of a set of tomographic pumping tests. Although the purpose of this study is not to compare the various approaches used to carry out hydraulic tomography, a similar

analysis of the information content as the one used by *Bohling* [2009] is adopted here to provide a common basis of comparison with previous studies. We note also that our resolution analysis does not consider any form of regularization, as the purpose of this study is to illustrate the information contained in head data alone, without the information that could be provided by other measurements.

The sensitivity and resolution analysis presented in this study are characteristic of the experimental configuration and general aquifer characteristics presented in a field implementation of tomographic slug tests in a littoral aquifer that shows evidence of strong K anisotropy at small-scale [*Paradis et al.*, 2014], which are described in Section 2.1. Then in Section 2.2, we present the radial flow model used to simulate various synthetic experiments and to derive a corresponding sensitivity matrix using small perturbations of a homogeneous and anisotropic model. This model allows the simulation of wellbore storage effects and placement of packers in the stressed well. Also, the approach used for resolution analysis, through singular value decomposition (SVD) of sensitivity matrix, is detailed in Section 2.3. In Section 3.1, we verify that the head perturbation induced by a slug test produces distinct transient sensitivity patterns in stressed and observation intervals that would allow the independent identification of K_h, K_v/K_h and S_s within the region of investigation (ROI) encompassed by a stressed well and an observation well, and whether the influence of hydraulic properties beyond the observation well can be isolated. We then illustrate in Section 3.2 through a spatial mapping of temporal sensitivities for a single slug test that parameter correlations may influence the resolution potential, while in Section 3.3 we demonstrate using a resolution analysis how tomographic slug tests can reduce parameter correlations with the ROI. Then, in Section 3.4, we investigate the influence of various factors that might be controlled to increase the resolution potential of tomographic slug tests. Moreover, in complement to previous analysis, Section 3.5 demonstrates for simple heterogeneous and anisotropic models that the non-applicability of the principle of reciprocity for slug tests is an intrinsic consequence of the pulsed mode of test initiation. Finally, we conclude and summarize our results in Section 4.

2 Methodology

2.1 Experimental configuration and general aquifer characteristics

The sensitivity and resolution analysis presented in this paper reflects the experimental configuration and general aquifer characteristics of a field implementation of tomographic slug tests reported by *Paradis et al.* [2014]. The present paper does not aim to exactly reproduce the test configuration or the hydraulic property fields themselves, but instead use similar test design and mean physical property values as the basis for various synthetic experiments to assess the information content of a transient analysis of tomographic slug tests.

The first step of the methodology consists in building synthetic models that are similar to the case study reported by *Paradis et al.* [2014]. That study involved a set of tomographic slug tests performed in a moderately permeable and anisotropic littoral aquifer [*Paradis et al.*, 2011; *Paradis and Lefebvre*, 2013; *Paradis et al.*, 2014]. The tomography field experiment involved 12 slug tests performed within 0.61 m long stressed intervals straddled by packers along two fully screened direct-push wells separated by a distance of 7.98 m. During the slug tests, head responses were recorded using a network of observation pressure transducers also located within 0.30 m long intervals straddled by packers. Other measurements were also made at that site, notably multilevel slug tests and lab permeameter tests on vertical sediment samples [*Paradis and Lefebvre*, 2013; *Paradis et al.*, 2014]. Values of K_h based on multilevel slug tests done at 15 cm intervals ranged from 7×10^{-7} to 2×10^{-5} m/s using the *Bouwer and Rice* [1976] method, whereas K_v estimates from permeameter tests at the same scale ranged from 2.7×10^{-8} to 1.0×10^{-5} m/s. The comparison of K_h and K_v values for the same intervals provides estimates of K_v/K_h between 0.0025 and 0.83. The bulk average K_h and K_v/K_h are about 1×10^{-5} m/s and 0.1, respectively. Moreover, considering the results of the inversion of the field data, a bulk average S_s of 1×10^{-4} m^{-1} will be used in the synthetic experiments. On that basis, Table III.1 summarizes the parameters used for the aquifer model as well as test characteristics used for numerical experiments. The upper part of the aquifer consists of fine sediments that have been shown to act as a leaky confining layer during pumping tests [*Paradis et al.*, 2011]. While in the field implementation the aquifer was modeled as an unconfined aquifer, for the sake of the synthetic experiments confined conditions will be assumed to get symmetric lower and upper boundary conditions (Figure III.1a) and then avoid confusion in the interpretation of the synthetic results.

Table III.1 Parameters used for the aquifer model and test characteristics used for numerical experiments.

Aquifer Parameters	
Horizontal hydraulic conductivity (K_h)	1×10^{-5} m/s
Specific storage (S_s)	1×10^{-4} m^{-1}
Hydraulic conductivity anisotropy (K_v/K_h)	0.1
Aquifer thickness	8 m
General Test and Well Parameters	
Initial head displacement (H_0)	4.5 m
Stress interval length	0.61 m
Inside radius of stress interval casing	0.0254 m
Inside radius of stress screen	0.0254 m
Observation interval length	0.305 m
Inside radius of observation screen	0.0254 m
Inter-well distance	7.98 m

2.2 Radial flow model

The computer program lr2dinv was used for the numerical simulations needed to carry out this study. The lr2dinv numerical simulator is a two-dimensional radial-vertical finite difference flow model developed by *Bohling and Butler* [2001]. Under conditions of radial symmetry, meaning the absence of angular variations in both hydraulic properties and boundary conditions, the flow to a partially penetrating well (stressed interval) in response to an instantaneous change in water level in a confined aquifer of infinite areal extent is described by:

$$\frac{1}{r}\frac{\partial}{\partial r}\left(rK_r\frac{\partial h}{\partial r}\right)+\frac{\partial}{\partial z}\left(K_z\frac{\partial h}{\partial z}\right)=S_s\frac{\partial h}{\partial t} \tag{1}$$

where h is the head [L], S_s is the specific storage [1/L], K_r and K_z are respectively the hydraulic conductivity in the radial (or horizontal- K_h) and vertical directions, t is time[T], r is the radial coordinate [L], and z is the vertical coordinate (positive upward from a zero reference at the base of the aquifer) [L].

The inner boundary of the model is at the stressed well radius r_w [L] and the initial conditions are given by

$$h(r,z,0)=0 \quad r_w < r < r_m \quad 0 < z < b \tag{2}$$

$$H(0)=H_0 \tag{3}$$

where r_m is the outer boundary of the model located far away from the stressed well in order not to interfere with it, b [L] is aquifer thickness, H is level of water in the well [L], and H_0 is the static water column height in the well where the stressed interval is located, which is equal to the initial level of water in the well at t=0, [L].

The outer boundary conditions are the following:

$$h(r=r_m,z,t)=0 \quad t>0 \quad 0 \le z \le b \tag{4}$$

$$\frac{\partial h(r,0,t)}{\partial z}=\frac{\partial h(r,b,t)}{\partial z}=0 \quad t>0 \quad r_w \le r \le r_m \tag{5}$$

and the inner boundary conditions at the stressed interval is given by

$$\frac{1}{L}\int_{d}^{d-L} h(r_w,z,t)dz = H(t) \qquad t>0 \tag{6}$$

$$2\pi r_w K_r \frac{\partial h(r_w,z,t)}{\partial r} = \begin{cases} 0 & z<d, z>d-L \quad t>0 \\ \dfrac{\pi r_c^2}{L}\dfrac{dH(t)}{dt} & d>z<d-L \quad t>0 \end{cases} \tag{7}$$

where d is the distance from the top of the aquifer to the bottom of the stressed interval [L], L is the stressed screen length [L], and r_c is the radius of well casing [L].

The program lr2dinv also allows the representation of the wellbore into the model domain, enabling the simulation of wellbore storage effects and placement of packer intervals in the well. Wellbore storage is defined as the relative volume change ΔV per unit head change Δh in the wellbore ($W = \Delta V/V \cdot \Delta h$). Wellbore processes are approximated using Darcy's Law, with open sections of the wellbore represented as high-K regions and packers as low-K regions. To simulate wellbore effects, one column of cells is used to represent the region inside the wellbore, from the wellbore radius, r_w, to the inner radius of the model grid, $r_{min}=r_w \cdot \exp(\Delta r')$, where $\Delta r'$ is the constant spacing in the transformed radial direction. Most cells in the wellbore should be assigned a value of S_s equal to 0, representing the fact that water is essentially incompressible. However, the top cell in the wellbore (or the top in any screened interval) should have a value of S_s equal to 1, which the program will convert to the value

$$S_s = \frac{1}{\Delta z}\left(\frac{r_w^2}{r_w^2 - r_{\min}^2}\right) = \frac{1}{\Delta z}\left(\frac{1}{1-\exp(-2\Delta r')}\right) \tag{8}$$

to account for the vertical thickness of the cell and compensate for the exclusion of the inner portion of the wellbore (between $r=0$ and $r=r_{min}$) from the model domain. For slug tests, the initial head in the top cell should be set to the initial displacement for the test, H_0, with all other initial heads in the model set to zero. The storage properties of the wellbore are set to represent wellbore storage effects, which is a key factor needed to be considered for a meaningful interpretation of inversion results with tomographic slug tests. Wellbore storage affects the hydraulic response in an observation interval in two ways. First, for a

given distance between the stressed and observation intervals, wellbore storage causes a delay in the time at which the pressure change is observed. Second, the response amplitude grows with the wellbore storage if the head change remains the same [*Prats and Scott*, 1975; *Spane*, 1996; *Brauchler et al.*, 2007; *Paradis and Lefebvre*, 2013]. Wellbore storage effects (time delay and amplitude change) are also a function of S_s, and for an aquifer within which S_s is expected to vary, simulation of wellbore storage as allowed by lr2dinv greatly improves the accuracy of the inverted tomograms. Note that the lr2dinv assumes that wellbore storage effects can be neglected in the observation interval. To obtain test data with minimal wellbore storage effects, the observation interval should be straddled with packers [*Sageev*, 1986].

The model uses a logarithmic transform of the radial coordinate, $r'=\ln(r/r_w)$, to transform the radial flow problem into an equivalent Cartesian problem in (r', z) space [*Butler and McElwee*, 1995; *Bohling and Butler*, 2001]. This study used a simulation grid with 43 cells of dimension $\Delta r'=0.19492$ along the transformed radial axis and 26 cells of dimension $\Delta z=0.3048$ m (1 foot) along the vertical axis (Figure III.1a). In the physical space, the radial location of the grid node with radial index i is $r_w \cdot \exp((i-0.5)\cdot\Delta r')$ and the location of the outer face of the corresponding cell is $r_w \cdot \exp(i\cdot\Delta r')$. The scale of the logarithmic transform was adjusted to fit the radial location of grid nodes to the location of the wells with a well spacing identical to the field implementation [*Paradis et al.*, 2014]. The exponentially increasing grid size in the radial direction improves the detailed description near the stressed well and provides a coarser representation with increasing distance from the stressed well, representing the fundamental sensitivity behavior of radial flow [*Butler*, 1990]. This type of grid also provides an efficient means to place the outer boundary far from the stressed well. The discretization used places the zero-head outer boundary of the model about 111 m from the stressed well, so that this boundary has negligible effects on simulated slug test heads at the stressed and observation wells.

For the sensitivity and resolution analysis presented in this paper, the simulation grid was divided into 143 parameter cells of 0.61 m in height that corresponds to the length of the stressed interval, as shown in Figure III.1b. Given the logarithmic change in cell dimensions in the radial direction, the cells of the simulation grid were merged to obtain inversion cells width of approximately equal size within the region encompassing the stressed and observation wells, referred to as the region of investigation (ROI) (region 1 in Figure III.1b). The average width of the cells within the ROI is 1.46 m, which means that 6

columns of cells separate the stressed and observation wells. Hereafter the term "parameter grid" is used to refer to this zonation of the simulation grid.

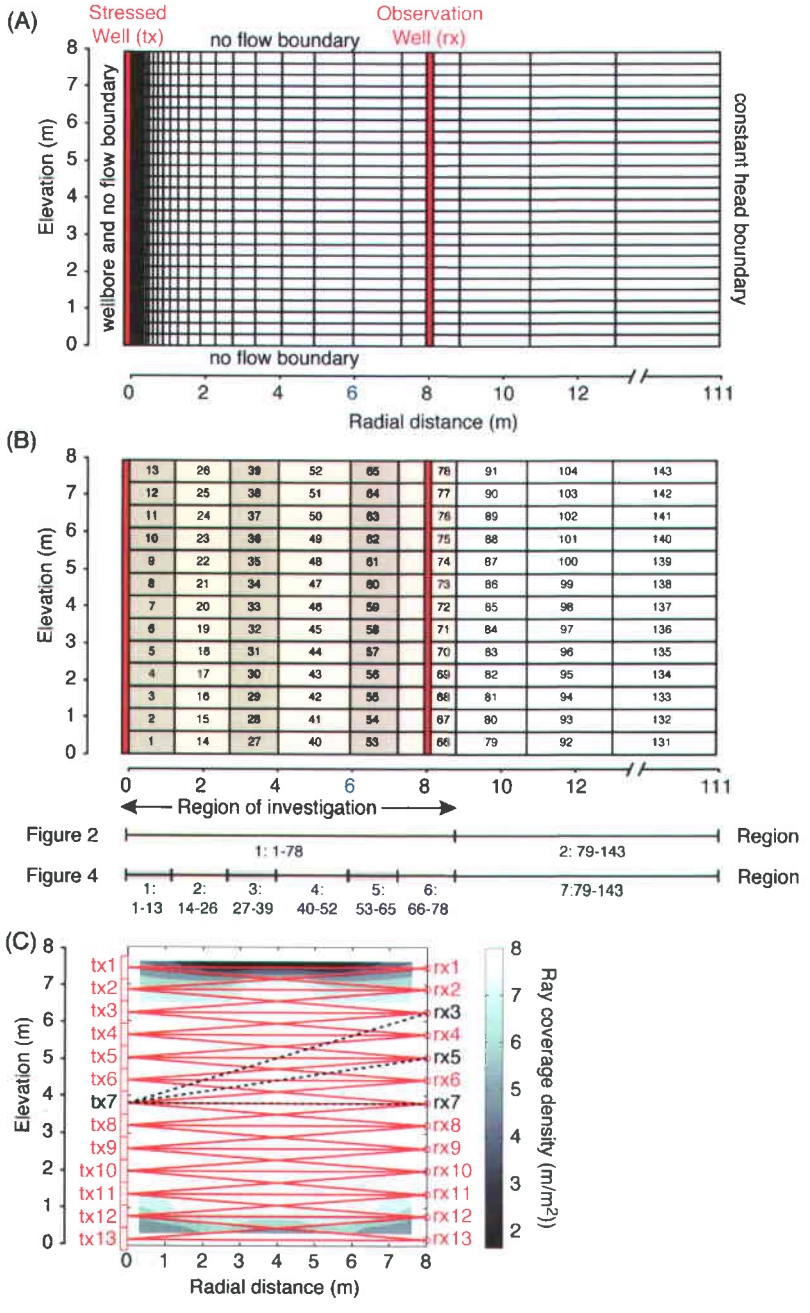


Figure III.1 (a) Numerical model grid and boundary conditions used for all synthetic experiments. (b) Parameter grid with 143 cells and zonation used for simulations in Figures III.2 and III.4. (c) Locations of the stressed and observation intervals (red plain lines) with ray coverage density for the synthetic tomographic experiment with 13 slug tests and 37 observation intervals (base case) presented in Figures III.5a-b. Table III.2 presents the correlation matrix for sensitivity to hydraulic properties with respect to the tx7 stressed interval and the rx7 observation interval shown in Figure III.1c

(black dotted line). Figure III.2 shows the normalized head sensitivities to hydraulic properties for the tx7 stressed interval and observation intervals at different angles relative to rx7 (rx7, rx5 and rx3).

2.3 Sensitivity matrix

In this paper, the sensitivity analysis of the transient analysis of tomographic slug tests is based on the Jacobian (sensitivity) matrix. The sensitivity matrix relates a change in a parameter to a corresponding change in the heads and each J_{ij} element in the sensitivity matrix represents the sensitivity of the head at a given time and location, h_i, to one of the hydraulic properties, p_j:

$$J_{ij} = \frac{\partial h_i}{\partial p_j} \tag{9}$$

The head index i runs over all observation times and locations for all tests, and p_j represents either the $\ln(K_h)$, $\ln(K_v/K_h)$ or $\ln(S_s)$ values associated with each of the cells in the grid. Thus, the sensitivity matrix represents a linear approximation of the behavior of the flow model nearby the parameters used in the simulations [*Vasco et al.*, 1997] and serves as an approximation in the solution steps of most nonlinear groundwater gradient inverse algorithms [*Aster et al.*, 2005].

In this work, the sensitivity matrix is estimated using a numerical derivative technique, i.e. it is computed from the simulation of heads for various test configurations using average uniform values for K_h, K_v/K_h and S_s as the base hydraulic property values. The sensitivity matrix elements are constructed by a sequence of groundwater flow simulations using lr2dinv, one simulation per grid cell, in which each hydraulic property in a single cell is slightly perturbed from the base hydraulic property values and the differences in heads are noted.

The sensitivity used in this study is built from the normalized sensitivities, or sensitivities to logarithmic hydraulic property values, defined as follows:

$$K_h \frac{dh}{dK_h} = \frac{dh}{dK_h / K_h} = \frac{dh}{d(\ln K_h)} \tag{10a}$$

$$(K_v / K_h) \frac{dh}{d(K_v / K_h)} = \frac{dh}{d(K_v / K_h)/(K_v / K_h)} = \frac{dh}{d(\ln(K_v / K_h))} \tag{10b}$$

$$S_s \frac{dh}{dS_s} = \frac{dh}{dS_s / S_s} = \frac{dh}{d(\ln S_s)} \tag{10c}$$

This normalization brings the sensitivities of the different parameters to an equal level for a change of head in response to a unit relative change in the parameter value or unit change in the natural log of the parameter value [*Bohling and Butler*, 2001].

2.4 Singular value decomposition (SVD)

The problem being ill-posed, the tomographic slug test inversion is regularized using a singular value decomposition (SVD) of the sensitivity matrix [*Aster et al.* [2005]; *Bohling*, 2009; *Jimenéz et al.*, 2013]. The problem considers m observed data and n model parameters. The singular value decomposition of the $m \times n$ sensitivity matrix, **J**, yields

$$\mathbf{J} = \mathbf{U}\mathbf{S}\mathbf{V}^T \tag{11}$$

where **U** is a $m \times m$ orthogonal matrix with columns that are unit basis vectors spanning the data space, R^m. **V** is a $n \times n$ orthogonal matrix with columns that are unit basis vectors spanning the model space, R^n. **S** is a $m \times n$ diagonal matrix with nonnegative diagonal singular values of **J** on the diagonal.

The singular values along the diagonal of **S** are arranged in order of decreasing magnitude, and the columns of **U** and **V** are arranged in the corresponding order. The columns of **V** represent linear

combinations of the model parameters, and the leading columns, corresponding with the largest singular values, represent the linear combinations that are most strongly resolved by the data.

The SVD can also be used to compute an inverse of **J**, called the Moore-Penrose pseudoinverse, which is given by:

$$\mathbf{J}^{\dagger} = \mathbf{V}_p \mathbf{S}_p^{-1} \mathbf{U}_p^T \tag{12}$$

and where only the p largest singular values and vectors, that correspond to the most strongly resolved parameters, have been retained. This pseudoinverse can be used in computing an $n \times 1$ vector of estimated parameters, $\mathbf{m}_{\dagger}$, from an $m \times 1$ vector of observed data, **d**, as

$$\mathbf{m}_{\dagger} = \mathbf{J}^{\dagger}\mathbf{d} \tag{13}$$

To characterize the conditioning of an inverse problem it is useful to analyze the singular value spectrum, which is the range of singular values. A least-squares problem is said to be rank-deficient if there is a clear distinction between the nonzero and zero singular values and the rank of **J**, which is the number of nonzero singular values, is less than the number of unknown parameters n. Rank-deficient problems have non-unique solutions, i.e. there is an infinite number of acceptable parameter vectors that will fulfill the inversion criteria. A number of unknown parameters n larger than the number of independent data m is characteristic of rank-deficient problems. Rank-deficient problems using SVD can be solved after truncating ($p<n$) the zero singular values, but a model of limited resolution will be produced. Even for noise-free data, it will not be possible to recover the true model **m** in a rank-deficient problem.

In the case of a sensitivity matrix with the number of independent data m exceeding the number of unknown parameters n, the rank of **J** will be limited by the number of parameters and there will be at most n nonzero singular values. Using $p=n$ in the computation of the generalized inverse will yield an estimated parameter vector $\mathbf{m}_{\dagger}$ that is close to the exact solution for this problem. Retaining a large number of singular values and vectors in the generalized inverse will allow a better fit to the observed

data and a better apparent resolution of the detail in the estimated parameters. In many inverse problems, however, singular values decay gradually toward zero and do not show an obvious jump between nonzero and zero singular values. The ratio of the largest to the smallest singular value, referred as the condition number, indicates the degree of ill-conditioning. With ill-conditioned problems the terms associated with smaller singular values will tend to amplify the noise in the data, increasing the variance in the parameter estimates. Regularization involving the truncation ($p<n$) of the smaller singular values may thus be necessary to produce a model with a lower variance, but this leads to a loss of resolution and introduces bias in the estimated parameters (variance-resolution tradeoff). Note that the problems related to ill-conditioning and bias-variance tradeoff are inherent to inverse problems and the SVD analysis allows to elucidate these problems.

2.5 Model resolution matrix

The generalized inverse always gives a solution, $\mathbf{m}_\dagger$, but it is essential to investigate how faithful this representation of the model representing the spatial distribution of parameter values is likely to be of the "true" model [*Vasco et al.*, 1997]. The concept of model resolution is an important way to characterize the bias of the generalized inverse solution. If the true model is represented by **m** and the corresponding true data vector is represented by **d**=**Jm**, then the parameter vector estimated from application of the pseudoinverse (13) is:

$$\mathbf{m}_\dagger = \mathbf{J}^\dagger \mathbf{J}\mathbf{m} \tag{14}$$

The matrix multiplying the true model is the model resolution matrix:

$$\mathbf{R}_m = \mathbf{J}^\dagger \mathbf{J} = \mathbf{V}_p \mathbf{V}_p^T \tag{15}$$

where the elements of matrix $\mathbf{R}_m$ are the averaging coefficients. For instance, R_{ij} denotes the contribution of the j^{th} parameter to the estimate of the i^{th} parameter. The resolution matrix is an array whose diagonal elements correspond to each grid cell of the model domain. The values of diagonal elements ($i=j$) are between 0 and 1. A value of zero means that a given parameter cannot be resolved using the available

observation data, and a value of one means it can be resolved perfectly. The off-diagonal elements ($i \neq j$) reflect the influence of other elements on resolving the elements on the diagonal. Thus a value of zero in all the ij^{th} (j not equal to i) elements means that the parameter in the j^{th} element has no effect on determining the i^{th} element and, consequently, no spatial averaging would be present in the estimate.

It should be noted that the resolution matrices $\mathbf{R}_m$ do not depend on specific data or models, but are exclusively properties of the sensitivity matrix **J** [*Aster et al.*, 2005]. They reflect the physics and geometry of a hydraulic experiment, and thus provide a powerful tool to assess the information content of tomographic slug tests.

3 Results and discussion

3.1 Temporal sensitivities for a single slug test

To provide basic insight into how each hydraulic property can be resolved by the transient analysis of tomographic slug tests, this section describes the results of the computation of temporal sensitivities for K_h, K_v/K_h and S_s based on the homogeneous and anisotropic base model of aquifer parameters and test characteristics described in Table III.1. To start this analysis, we distinguish the region within the ROI from the region beyond the observation well that represents the aquifer outside the ROI (regions 1 and 2 in Figure III.1b, respectively). These two regions are considered to assess how hydraulic properties can be resolved within the ROI with the lowest interference from parameters outside the ROI. Given the radial discretization of the numerical grid, the cells centered on the observation well were included in the ROI. Figure III.2 shows head sensitivities to K_h, K_v/K_h and S_s for the region within and outside the ROI over the entire duration of a slug test. The sensitivities were computed with lr2dinv [*Bohling and Butler*, 2001] in a forward modeling mode with an initial head displacement in the stressed interval of 4.5 m. The plotted sensitivities are normalized by the corresponding hydraulic property values as described by Equations 10a-c to compare the relative importance of each property. The sign of the sensitivity value indicates whether the resulting change in head is proportional (positive value) or inversely proportional (negative value) to a unit change in property value.

Figures III.2a-b present the normalized head sensitivities for a stressed and an observation interval, respectively, that are aligned along the same horizontal path (tx7 and rx7 shown in Figure III.1c). It can be seen that sensitivities to all hydraulic properties for the stressed interval are larger than sensitivities for the observation interval. Despite the difference in scale of the two graphs, we note that sensitivities to K_v/K_{h1} and S_{s1} for the stressed interval (Figure III.2a) are significant with respect to sensitivities for the observation interval (Figure III.2b). In fact, sensitivities are proportional to the flow rate that contributes to the head response at a particular point of observation (stressed or observation interval) [e.g., *McElwee and Yukler* 1978]. The total flow rate at a point of observation is the sum of the flow (K_h and K_v) and storage (S_s) components. Maximum sensitivities are thus found near the stressed interval where both horizontal and vertical hydraulic gradients and the rate of change in heads are the largest.

Figure III.2a illustrates that the head response in the stressed interval is clearly more sensitive to K_{h1} than to K_v/K_{h1} or S_{s1} with high sensitivity correlations between those properties (Table III.2). Thus, only the K_{h1} component can be realistically identified using only stressed interval data [*McElwee et al.* 1995a]. Note that the stressed interval response is sensitive to K_v/K_{h1} here because a small aspect ratio (screen length over screen radius) that increases the proportion of slug-induced vertical flow is used for the experiment and that sensitivity to K_v/K_{h1} generally decreases with increasing aspect ratio [*Paradis and Lefebvre*, 2013]. Although sensitivity correlations presented in Table III.2 for properties within the ROI are moderately anti-correlated with properties outside the ROI, sensitivity amplitudes for those properties are very low and do not affect significantly the stressed interval response.

The curves depicted in Figure III.2b for the observation interval (rx7 shown on Figure III.1c) indicate that sensitivities for properties within the ROI have similar amplitudes with different behaviors. Sensitivity curves for K_{h1} and S_{s1} have a fairly strong anti-correlation (Table III.2), which may results in difficulty in resolving those properties using only observation interval data. The maximum sensitivity for properties within the ROI actually occurs before (K_h and S_s) or at (K_v/K_h) the peak of the normalized head response recorded in the observation interval (Figure III.2b). At later times, when the peak of the head perturbation is moving beyond the observation well, the sensitivity for properties outside the ROI reach their maximum amplitude and sensitivity to K_{h2}, K_v/K_{h2} and S_{s2} are thus delayed in time with respect to K_{h1}, K_v/K_{h1} and S_{s1} (Figure III.2b). The sensitivity to K_{h2} and S_{s2} with respect to sensitivity to properties within the ROI indicate a low to moderate influence on the observation interval response, whereas the very low sensitivity to K_v/K_{h2} indicates that this property cannot be identified from head data. Indeed, the finite duration of a slug test and the attenuation of the head perturbation that travels in the aquifer indeed

constrain the growth in sensitivities for hydraulic properties outside the ROI. Although the strong sensitivity correlation between K_{h2} and S_{s2} indicates that only their combined effects can be assessed, their weak to moderate correlations with respect to sensitivity to properties within the ROI suggest that K_{h2} and S_{s2} can be isolated from K_{h1}, K_v/K_{h1} and S_{s1} (Table III.2). While hydraulic properties within the ROI cannot be completely resolved using either stressed or observation interval response alone, it appears from Figures III.2a-b and Table III.1 that K_{h1}, K_v/K_{h1} and S_{s1} can be better identified using both stressed and observation interval responses [*McElwee et al.*, 1995b]. That is the combined use of different head responses contributes to reduce correlations among the properties.

Table III.2 Correlation matrix for sensitivity to hydraulic properties for regions within (subscript 1) and outside (subscript 2) the region of investigation (ROI) with respect to stressed (tx7) and observation (rx7) interval. The ROI is defined in Figure III.1b and location of tx7 and rx7 are shown in Figure III.1c.

Stressed Interval (tx7)						
	Within the region of investigation			**Outside the region of investigation**		
Properties	K_{h1}	K_v/K_{h1}	S_{s1}	K_{h2}	K_v/K_{h2}	S_{s2}
K_{h1}	1	0.98	0.86	-0.64	-0.36	-0.52
K_v/K_{h1}	0.98	1	0.73	-0.60	-0.31	-0.49
S_{s1}	0.86	0.73	1	-0.62	-0.41	-0.53
K_{h2}	-0.64	-0.60	-0.62	1	0.82	0.98
K_v/K_{h2}	-0.36	-0.31	-0.41	0.82	1	0.88
S_{s2}	-0.52	-0.49	-0.53	0.98	0.88	1
Observation Interval (rx7)						
	Within the region of investigation			**Outside the region of investigation**		
Properties	K_{h1}	K_v/K_{h1}	S_{s1}	K_{h2}	K_v/K_{h2}	S_{s2}
K_{h1}	1	-0.12	-0.84	0.45	0.41	0.38
K_v/K_{h1}	-0.12	1	0.64	0.68	0.82	0.72
S_{s1}	-0.84	0.64	1	0.03	0.15	0.11
K_{h2}	0.45	0.68	0.03	1	0.89	0.99

K_v/K_{h2}	0.41	0.82	0.15	0.89	1	0.89
S_{s2}	0.38	0.72	0.11	0.99	0.89	1

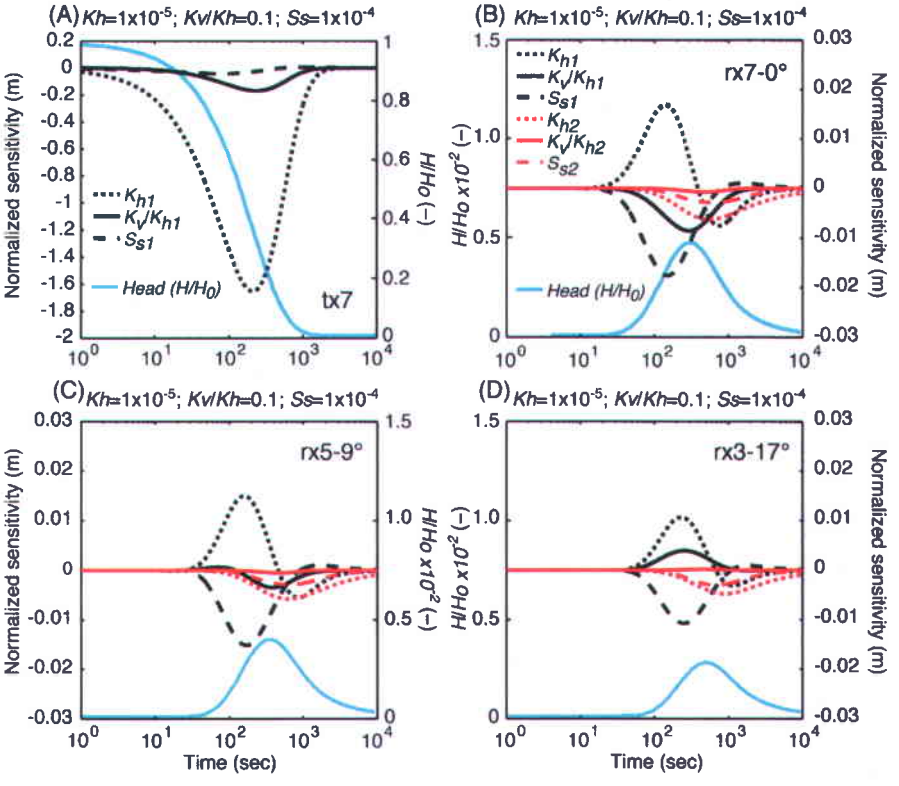


Figure III.2 (a) Normalized head sensitivities to hydraulic properties (K_h, K_v/K_h and S_s) for a stressed interval; and (b, c and d) observation intervals located at different angles from the stressed interval based on a homogeneous and anisotropic aquifer model (Figure III.1b and Table III.1). Locations of the intervals (tx7, rx7, rx5 and rx3) are indicated in Figure III.1c (black dotted lines). Note the difference in scale for stressed and observation intervals. Subscripts 1 and 2 refer to parameters within and outside the region of investigation (ROI), respectively. Stressed interval sensitivities for properties outside the ROI are not significant and are thus not shown in Figure III.2a. Corresponding normalized head (H/H_0) responses are also shown (plain blue lines).

The sensitivity analysis is also extended to the effects of the angle between the locations of the stressed and observation intervals on the observation interval response. Figures III.2b-c show head sensitivities to properties within and outside the ROI for various observation interval angles (rx7, rx5 and rx3 in Figure III.1c). Sensitivity amplitudes for K_{h1} and S_{s1} generally decrease for larger angles, although they retain relatively similar shapes, whereas sensitivity curves for K_{h2}, K_v/K_{h2} (practically nil) and S_{s2} remain almost unchanged. The relative influence of properties outside the ROI over K_{h1} and S_{s1} is therefore slightly increased. Most importantly, Figures III.2b-d show that the sensitivity curve for K_v/K_{h1} is strongly affected by observation interval location. Indeed, sensitivities range from positive to negative values depending on the location of the observation interval in the flow field. We note that for isotropic conditions (K_v=K_h) the sensitivity to all properties do not change with observation interval location. The increases in travel path lengths at the angles used (9° and 17°) are actually too small (<5%) to observe any significant head attenuation. Thus from a practical viewpoint, because sensitivity to K_v/K_{h1} is

maximal at low angle and may decrease rapidly to zero for anisotropic conditions, it may be sufficient to only monitor the propagation of the head perturbation with observation intervals located at low angles with respect to the source.

In summary, the front of the head perturbation, which begins with the first arrival time and ends with the peak of the head amplitude recorded in the observation interval, contains the majority of the information about hydraulic properties within the ROI. The different temporal behaviors of the sensitivity curves suggest that information contained in stressed and observation interval responses of slug tests can be used to independently identified K_h, K_v/K_h and S_s within the ROI without much interference from hydraulic properties beyond the observation well. It should be noted that the head perturbation induced by a slug test produces transient sensitivity patterns that are fundamental to independently identify hydraulic properties. *Butler and McElwee* [1990] demonstrated that only data collected during periods of changing sensitivities provide information about hydraulic properties. Thus, unlike sensitivity curves for pumping tests that reach constant sensitivity values for hydraulic conductivity components (K_h and K_v/K_h) once the cone of depression moves beyond the point of observation (e.g., *Butler and McElwee,* 1990; *Bohling,* 2009; *Mao et al.,* 2013; *Sun et al.,* 2013), the sensitivity curves induced by slug test for a particular location within the ROI vary for all hydraulic properties over the entire duration of a test. That is, a larger time window of opportunity is offered by slug tests to identify hydraulic properties within the ROI. Moreover, the resolution potential for S_s can be strongly lowered using pumping tests because wellbore storage effects contribute to attenuate sensitivity to S_s, especially on drawdowns recorded in stressed well. Finally, hydraulic properties outside the ROI can have a much more negative impact on the identification of properties within the ROI using pumping tests because the increasing cone of influence over time (in opposite to slug tests that last for a short duration), may increase the correlation between properties within and outside the ROI.

3.2 Spatial sensitivities for a single slug test

To further illustrate the fundamental aspects of tomographic slug tests, a spatial sensitivity analysis was carried out on the hydraulic properties for a single slug test for the homogeneous and anisotropic aquifer represented by the properties in Table III.1. A forward simulation was run to simulate a slug test using the base case values (Table III.1) and the sensitivities were evaluated for all cells of the 143-parameter grid (Figure III.1b), considering all time steps for a stressed interval (tx7) and an observation interval (rx7) separately.

Figures III.3a-b show the root-mean-square (*RMS*) normalized sensitivity of heads to K_h, K_v/K_h and S_s for each cell of the parameter grid and for the entire duration of the test (10000 s). The RMS normalized sensitivity to K_h for parameter grid cell j is given by:

$$RMS_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Kh\frac{dh_i}{dKh}\right)^2} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{dh_i}{d\ln Kh}\right)^2} \tag{16}$$

and similarly for K_v/K_h and S_s. The summation is over the entire set of n simulated heads at the observation point (stressed or observation interval) and for all times over the test. The RMS normalized sensitivity is an integrated measure of the sensitivity of the simulated heads at the observation point and for all times to a small relative change in the given hydraulic property in a given cell. The root mean square of sensitivity values for each cell emphasizes the relative variation of the sensitivities associated with each properties and observation point and it is these relative variations that determine the resolution potential.

For the interpretation of Figures III.3a-b, it is first important to note the difference in magnitude between the sensitivities for the stressed and the observation intervals. As previously noted in Figures III.2a-b, maximum sensitivities are for the stressed interval where the flow rates are the largest. Sensitivities for the observation interval would then have presented larger values for an observation well located closer to the stressed interval or using a higher initial head to initiate the test. The comparison of Figure III.3a to Figure III.3b shows that the overall spatial sensitivity patterns for the stressed and observation intervals are different. Sensitivities to K_h, K_v/K_h and S_s for the stressed interval are focused mainly near the stressed interval, whereas sensitivities for the observation interval are more evenly distributed over the ROI. Also, the specific sensitivity patterns for each hydraulic property are somewhat different according to the flow pattern induced by slug tests performed in stressed intervals of small aspect ratio. Indeed, K_h and S_s larger sensitivities are focused along the horizontal line that joins stressed and the observation intervals, where horizontal head gradients and temporal variation in head gradients are maximal, respectively. For K_v/K_h, the larger sensitivity values are rather located on both sides (over and under) of the stressed and observation intervals where vertical head gradients are more important.

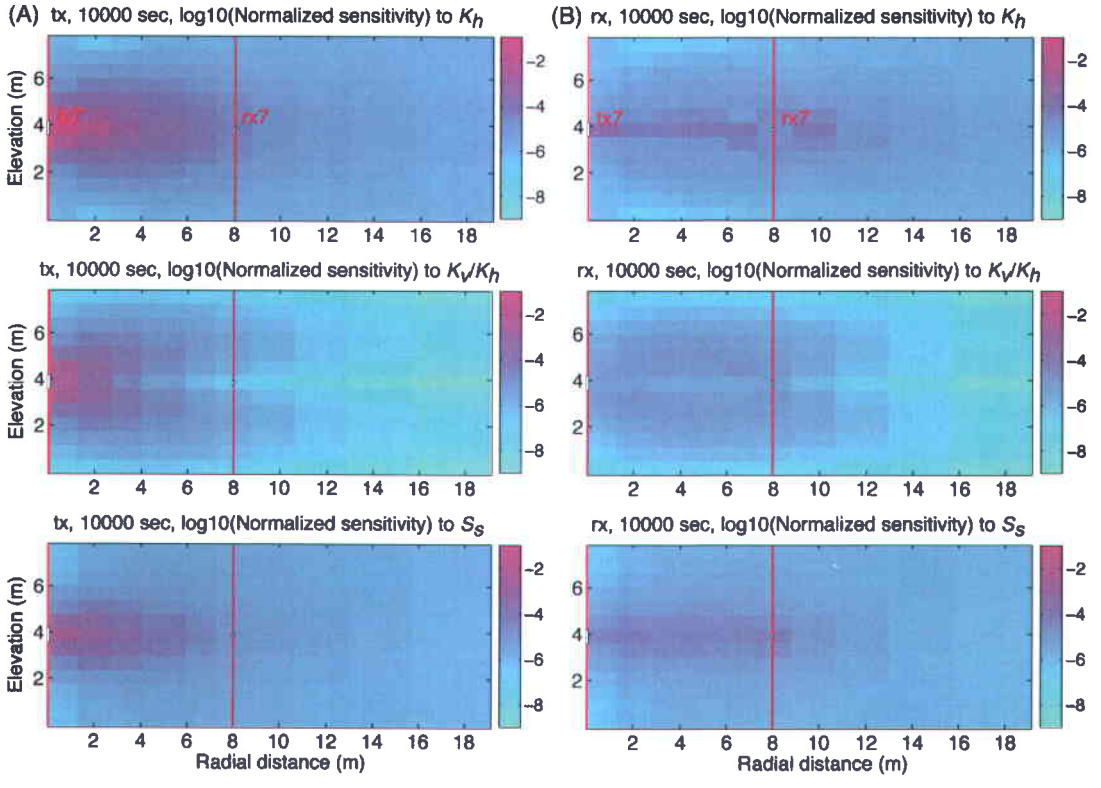


Figure III.3 Spatial distributions of root-mean-square normalized sensitivity to K_h, K_v/K_h and S_s over a single slug test for: (a) a stressed interval (tx7); and (b) an observation interval (rx7). Locations of the intervals (tx7, and rx7) are indicated in Figure III.1c (black dotted line).

Figures III.4a-b show temporal head sensitivity to K_h, K_v/K_h and S_s for six regions (column of cells) within the ROI that we distinguished from the region outside the ROI, as indicated in Figure III.1b. Sensitivities for the 143-parameter grid in Figures III.3a-b were summed over each region at each time step, as previously done in Figures III.2a-d. As suggested by Figures III.3a-b, we observe from Figure III.4b that while the observation interval response is sensitive to all regions within the ROI, the regions that influence the stressed interval response are rather focused near the stressed interval with a decreasing influence away from this well (note however the difference in scale between Figure III.4a and III.4b). Thus, the benefits of using both stressed and observation interval to resolve correlations between properties, as previously discussed with Figures III.2a-b, decrease with increasing distance from the stressed interval as the sensitivities for the stressed interval vanish. Moreover, Figures III.4a-b show that there are fairly strong correlations among sensitivity curves for adjacent regions in the middle of the ROI. This suggests that it may be difficult to separate the contribution on the head responses of each parameter within the ROI, particularly when considering normally noisy head data. Note that the same sensitivity correlations can be observed for parameters outside the ROI (not shown). In the next sections, we will

investigate how slug tests carried out in a tomographic format can cope with the different spatial patterns of sensitivity for properties and observation points (stressed and observation intervals) and can alleviate correlations between parameters to resolve K_h, K_v/K_h and S_s throughout all the ROI.

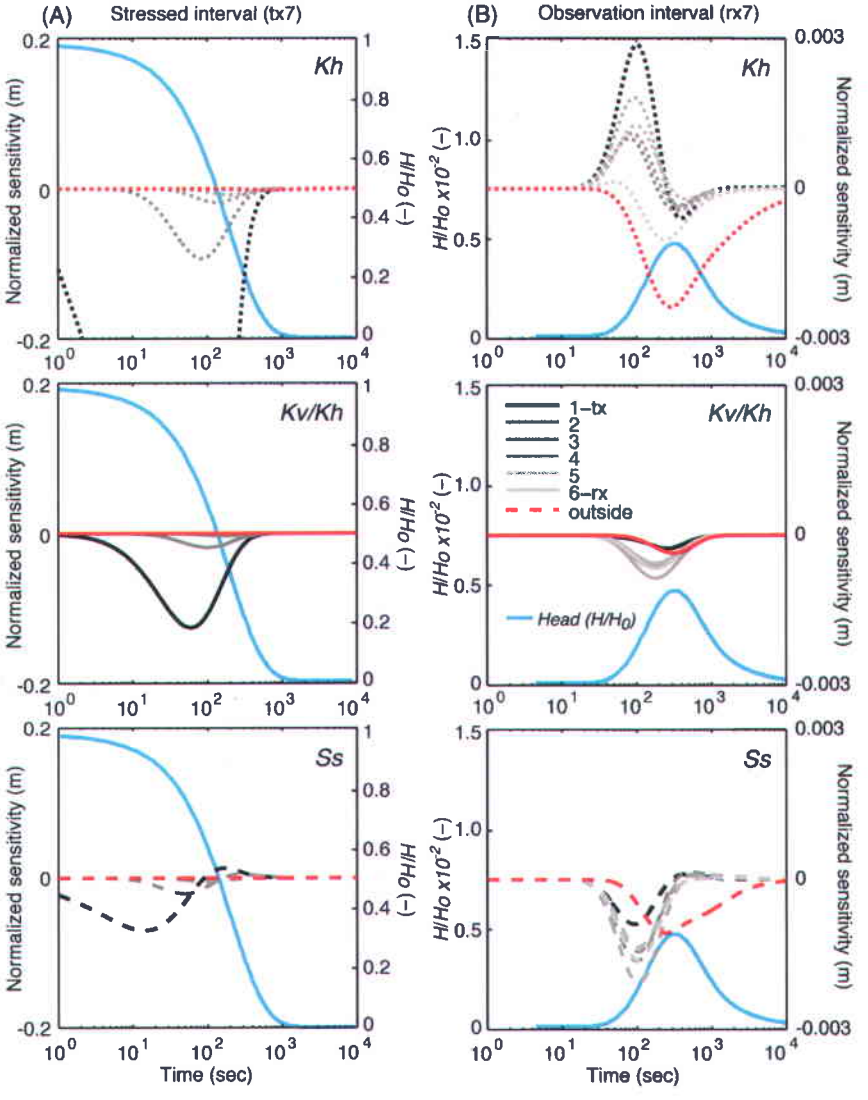


Figure III.4 (a) Normalized head sensitivities to hydraulic properties (K_h, K_v/K_h and S_s) for a stressed interval (tx7); and (b) an observation interval (rx7) based on a homogeneous and anisotropic aquifer model (Figure III.1b and Table III.1). Locations of the intervals (tx7 and rx7) are indicated in Figure III.1c (black dotted line). Note the difference in scale for stressed and observation intervals. Subscripts 1 to 7 are defined in Figure III.1b. Corresponding normalized head (H/H_0) responses are also shown (plain blue lines).

3.3 Spatial sensitivity and resolution of a tomographic experiment

This section examines the spatial sensitivity and resolution associated with a transient analysis of the head responses obtained from tomographic slug tests. For that purpose, a tomographic experiment is numerically simulated, with 13 slug tests along the stressed well with 3 observation intervals (except for the upper and lower most stressed intervals that use 2 observation intervals) distributed along the observation well to provide relatively uniform spatial coverage of the aquifer, as illustrated in Figure III.1c. A total of 37 observation interval responses and 13 stressed interval responses are thus used for the sensitivity and resolution analysis to K_h, K_v/K_h and S_s, which is considered a reasonable test configuration (moderate spatial coverage density) that might be practically obtained in the field. The present analysis is carried out with a homogeneous and anisotropic model using all available stressed and observation interval responses of the tomographic experiment, as suggested by sensitivity patterns shown in Figures III.2-4, in order to fully assess the resolution potential of a transient analysis of tomographic slug tests to resolve K_h, K_v/K_h and S_s within the ROI. The analysis is based on the sensitivity matrix computed from a simulation of the head responses for each test, using the base property values presented in Table III.1 and the 143-parameter grid illustrated in Figure III.1b. The sensitivity of the head are then computed using relations 10a-c for all times and all locations of the 13 stressed intervals and 37 observation intervals for all hydraulic properties (K_h, K_v/K_h or S_s), individually, associated with each cell of the parameter grid. The resolution analysis is based on a truncated SVD of the sensitivity matrix, as previously described.

Figure III.5a shows the RMS normalized sensitivity of head observations to K_h, K_v/K_h and S_s values over all 13 tests in the 143-parameter grid for a 10000 s time frame. As expected from Figures III.3a-b for a single test, a transient analysis of tomographic slug tests using stressed and observation interval responses focuses the sensitivity more strongly on cells near the stressed well, with gradually decreasing sensitivities farther from the stressed well. Sensitivities are also found to be homogeneously distributed along the vertical direction as expected from the relatively constant ray coverage density illustrated in Figure III.1c.

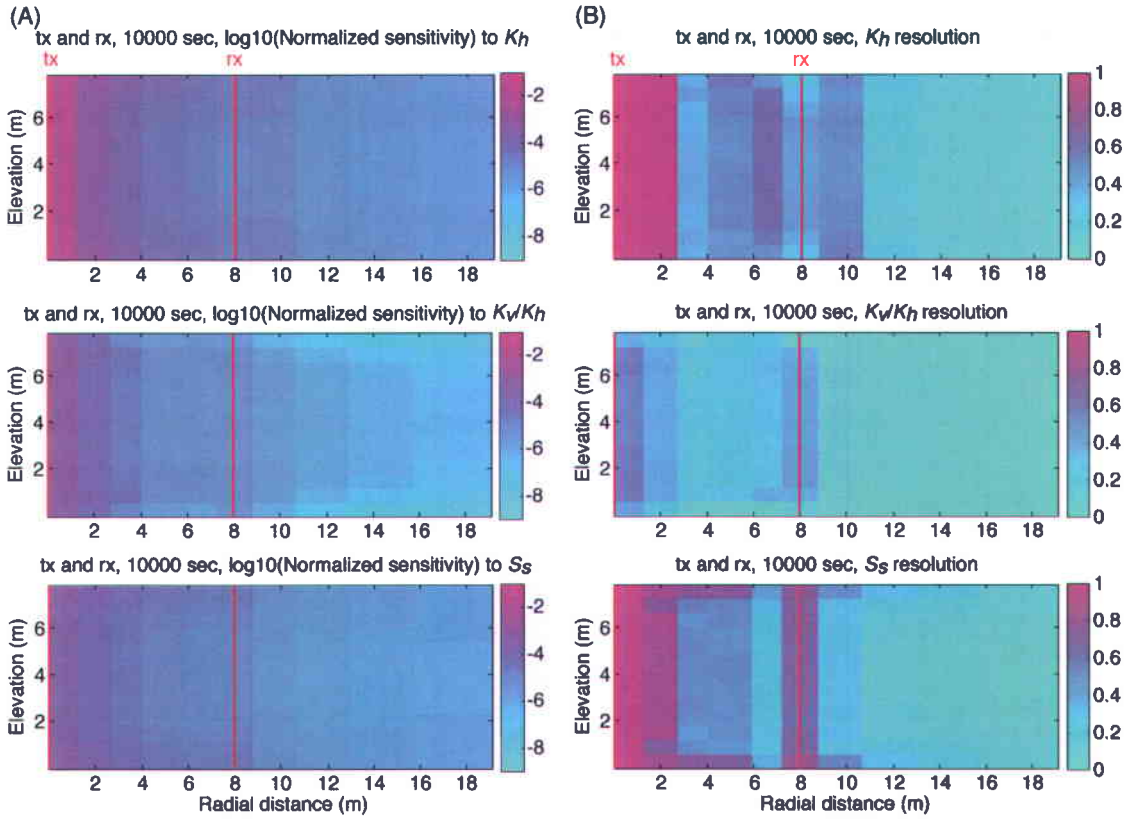


Figure III.5 (a) Spatial distributions of root-mean-square normalized sensitivity to K_h, K_v/K_h and S_s; and (b) diagonal elements of resolution matrix (Figure III.7) for K_h, K_v/K_h and S_s associated with the analysis of the 10000-sec head records for the synthetic tomographic experiment with 13 slug tests and 37 observation intervals (base case) shown in Figure III.1c. The resolution values are based on truncated singular value decomposition (SVD) inversion with a relative parameter error of 1% using the 143-parameter grid shown in Figure III.1b. Simulation parameters and average resolution values for hydraulic properties within the region of investigation (ROI) are presented in Table III.3.

Figure III.6a shows the singular value spectra for the sensitivity matrix of the 143-parameter grid that is composed of 429 singular values, corresponding to the number of hydraulic properties (3: K_h, K_v/K_h and S_s) times the number of cells (143). A singular value spectrum is a graph of the singular values in descending order. A larger singular value means a better resolution potential for those parameters. The bump in the first 13 singular values corresponds to K_h values for the 13 cells closer to the stressed intervals where the sensitivities for those parameters are much higher than in any other cells, as depicted by the sensitivity patterns in Figures III.5a-b. The slope of the spectrum also indicates the ability to accurately estimate the parameter combinations associated with successive singular values. Indeed, the independent information contained in head data exceeds or equals the number of unknown parameters, as indicated in Table III.3 by the rank of the sensitivity matrix, which equals the number of parameters. This means that with noise-free head data the true model of the hydraulic properties can be fully recovered, which obviously does not represent real field conditions. The gradual decay in singular values toward

very small singular values illustrated in Figure III.6a and the corresponding large condition number of $1.9x10^{11}$ (Table III.3) both suggest that the tomographic experiment is severely ill-conditioned, which means that the experiment is very sensitive to noise. That is, parameters that correspond to the smaller singular values can be hardly resolved because they have very small influence on the head responses or their sensitivities are highly correlated. As it will be shown later, most of the smaller singular values correspond to parameters beyond the observation well.

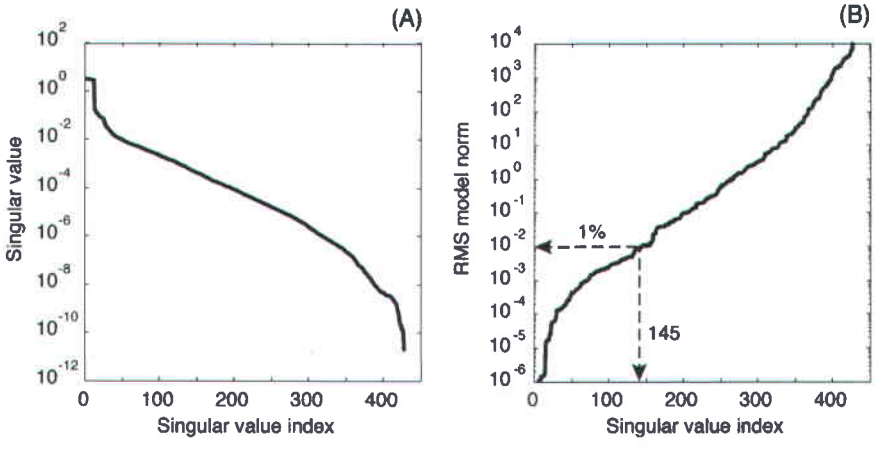


Figure III.6 (a) Singular value spectra, and (b) root-mean-square model norm associated with truncated singular value decomposition (SVD) inversion of white noise (error standard deviation of $2x10^{-4}$ m) head vector associated with the analysis of the 10000-sec head records for the base case synthetic tomographic experiment.

Table III.3 Simulation parameters and average resolution for hydraulic properties within the region of investigation (ROI) for various scenarios. SD stands for standard-deviation.

Scenario	Noise SD (m)	Initial head (m)	Cell size (m)	Inter-well distance (m)	Grid cell number	Unknown parameters	Rank	Singular values retained	Condition number	Average resolution within ROI (-)		
										K_h	K_v/K_h	S_s
Base case (37txrx)	2E-4	4.5	1.46	7.98	143	429	429	145	1.9E+11	0.63	0.30	0.62
Noise-free	2E-8	4.5	1.46	7.98	143	429	429	366	1.9E+11	0.99	0.99	1.00
H0=9m	2E-4	9.0	1.46	7.98	143	429	429	161	9.5E+10	0.64	0.34	0.66
H0=2.25m	2E-4	2.25	1.46	7.98	143	429	429	131	8.8E+10	0.56	0.26	0.55
H0=1.13m	2E-4	1.13	1.46	7.98	143	429	429	96	9.8E+10	0.44	0.14	0.43
Distance=6.56m	2E-4	4.5	1.46	6.56	143	429	429	143	2.5E+11	0.70	0.38	0.72
Distance=5.40m	2E-4	4.5	1.46	5.40	143	429	428	147	3.0E+11	0.83	0.50	0.86
Distance=3.65m	2E-4	4.5	1.46	3.65	143	429	427	161	3.4E+11	0.98	0.80	0.99

Scenario	Noise SD (m)	Initial head (m)	Cell size (m)	Inter-well distance (m)	Grid cell number	Unknown parameters	Rank	Singular values retained	Condition number	Average resolution within ROI (-)		
										K_h	K_v/K_h	S_s
Distance=2.03m	2E-4	4.5	1.46	2.03	143	429	427	153	3.7E+11	0.99	0.98	0.99
Column=4	2E-4	4.5	2.19	7.98	117	351	351	131	1.9E+11	0.78	0.50	0.80
Column=3	2E-4	4.5	2.92	7.98	104	312	312	124	1.9E+11	0.91	0.71	0.92
Column=2	2E-4	4.5	4.38	7.98	91	273	273	102	1.9E+11	0.96	0.91	0.97
59txrx	2E-4	4.5	1.46	7.98	143	429	429	155	1.5E+11	0.66	0.34	0.65
37txrx&13tx^{-1}	2E-4	4.5	1.46	7.98	143	429	429	195	1.2E+11	0.81	0.36	0.75
37txrx&13txrx^{-1}	2E-4	4.5	1.46	7.98	143	429	429	217	1.1E+11	0.89	0.44	0.82
37txrx&37txrx^{-1}	2E-4	4.5	1.46	7.98	143	429	429	246	1.1E+11	0.93	0.62	0.88
37txrx&37txrx^{-1} resolution	2E-4	4.5	2.92	7.98	104	312	312	171	1.2E+11	0.99	0.96	0.99

Using all singular values in the computation of the generalized inverse with noise-free head data will yield an estimated model of the hydraulic properties, given by relation 13, which is close to the exact solution. In order to provide a realistic estimate of parameter resolution under realistic field conditions with actual synthetic data, we define the noise amplification behavior associated with the tomographic experiment and use it to truncate the singular value spectra to a predefined noise level. Figure III.6b presents a plot of the RMS model norm obtained when a truncated SVD inversion of the sensitivity matrix is applied to a white noise data residual vector versus the number of singular values retained in the inversion. The head data residual vector is Gaussian with a mean of zero and a standard deviation of $2x10^{-4}$ m, which was evaluated from a residual analysis of head data recorded during the field implementation of a tomographic slug tests experiment [*Paradis et al.*, 2014]. The RMS model norm is a measure of the noise amplification error as described by *Vogel* [2002], which is based on a particular realization of the residual vector and treating the sensitivity matrix as a linear approximation of the flow behavior in the vicinity of the base property values [*Clemo et al.*, 2003; *Bohling*, 2009]. Note that various realizations of the residual vector showed similar model norm results. Thus Figure III.6b (black line) depicts the noise amplification behavior associated with the tomographic experiment previously described for K_h, K_v/K_h and S_s using the 143-parameter grid. With more terms added to the truncated SVD, the magnitude of the error increases, reflecting the increasing amplification of noise associated with the increasing demand for parameter resolution. For this analysis, we used a predefined relative parameter error of 1% to truncate the singular value spectra (Figure III.6b).

Figure III.7 shows the resolution matrix associated with the analysis of the 10000-s records from the 13 slug tests based on truncated SVD inversion of the sensitivity matrix using 145 singular values, which corresponds to a relative parameter error of 1% (Figure III.6b). Thus, only the number of singular values that amplify the assumed head noise into average relative error of 1% in the parameter estimates is retained. The resolution matrix is 439x439, with upper left, center and lower right 143x143 matrices associated with the K_h, K_v/K_h and S_s values of the 143-parameter grid, respectively. The upper middle, upper right and middle right matrices and their corresponding middle left, lower left and lower middle mirror image matrices describe dependence among the K_h, K_v/K_h and S_s values. Each diagonal element of the 439x439 matrix describes the degree of resolution of the K_h, K_v/K_h or S_s value in a particular model cell and ranges from 1 for a perfectly resolved parameter to 0 for a completely unresolved parameter. Off-diagonal elements of the resolution matrix describe the extent of blurring of a particular parameter estimate due to correlations with other parameters during the inversion process [*Aster et al.*, 2005]. Resolution values lower than 0.5 indicate that the parameter estimate for a cell is influenced more

strongly by other parameter values than it is by the corresponding parameter of the cell itself. Higher resolution values are thus essentially focused on the diagonal of the resolution matrix for cells encompassed by the ROI, as illustrated in Figure III.7. Small resolution values are also spread along two off-diagonal stripes running parallel to the diagonal, suggesting a correlation between particular parameters.

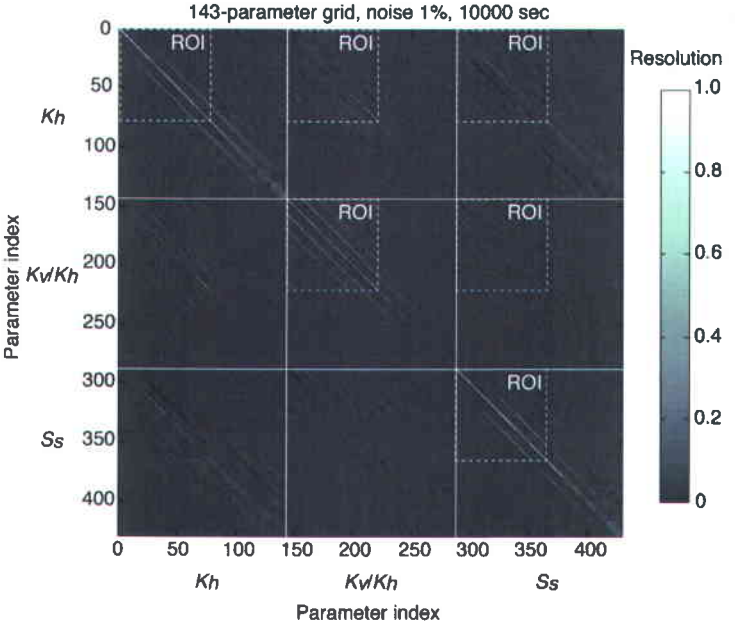


Figure III.7 Resolution matrix associated with the analysis of the 10000-sec records from the base case synthetic tomographic experiment based on truncated SVD inversion of the sensitivity matrix with a relative parameter error of 1%. Full white lines superposed on the resolution matrix separates the different sub-matrices corresponding to each hydraulic parameter and dashed white boxes encompass parameters within the region of investigation (ROI).

To further examine the resolution behavior of the tomographic experiment, Figure III.5b shows the diagonal elements of the resolution matrix at the corresponding parameter grid cells [*Aster et al.*, 2005]. As expected from the resolution matrix, resolution values for K_h, K_v/K_h and S_s are higher within the ROI, but they are mostly focused on the stressed and observations wells with better resolution near the stressed well. This results in a weak resolution potential for K_h, K_v/K_h and S_s within the ROI with average resolution values for this region ranging from 0.30 to 0.62 (Table III.3). We note also the fairly lower resolution potential for K_v/K_h in comparison to K_h and S_s, as expected from Figures III.2b-d that show relatively smaller sensitivity to K_v/K_h. This is expected to be the result of the test initiation method that induces predominantly horizontal flow, even using stressed intervals with a small aspect ratio. For cells within the ROI with low-resolution values, the inspection of the resolution matrix indicates that the

resolution potential for those parameters is spread to adjacent cells that are horizontally aligned in the grid. The off-diagonal stripes pointed out in Figure III.7 are indeed exactly offset by a value of 13 parameters with respect to the diagonal, which corresponds to the number of vertical cells in the grid (see Figure III.1b for the numbering of the parameter grid cells). This indicates that for each incompletely resolved cell within the ROI, only the combined effects of this cell and the cells on its sides can be effectively resolved from the inversion because the sensitivity to individual hydraulic properties for those cells are highly correlated together, as it was also previously suggested by Figures III.4a-b. It should be noted that retaining more terms in the truncated SVD inversion will lead to an increase in average resolution for parameters within the ROI, but the price paid for this resolution will be an increase in the variance, as expressed by the noise amplification curve associated with the tomographic experiment in Figure III.6b. In addition, Figure III.5b shows that it is not possible to extract information about K_h, K_v/K_h and S_s outside the ROI, except for K_h just beyond the observation well. The inexistent resolution potential for parameters outside the ROI is explained by the combined facts that sensitivities outside the ROI are relatively small with respect to sensitivities within the ROI, with decreasing values away from the observation well (Figure III.5a), and the strong correlations between adjacent parameters for this region, as deduced from the resolution matrix in Figure III.7. Very low-resolution values on the diagonals of the upper right and middle right matrices suggest however that K_h, K_v/K_h and S_s are not significantly correlated together everywhere over the modeled domain. Thus, this analysis indicates that even using information contained in the head records of a moderately dense network of stressed and observation intervals, the effects of the hydraulic properties in several adjacent cells on head responses can be inseparable (high correlation) or their effects in certain regions can be hindered (low sensitivity). Estimates of the parameters are thus weighted averages of the true model parameters in the vicinity of the poorly resolved cells and only their combined effects can be thoroughly resolved at the scale of the parameter grid cell size used for this analysis. Note that a similar analysis could also be carried out using varying vertical cell dimensions. But, using vertical cells smaller than the length of the stressed intervals would result in too many parameters with respect to the information contained in the stressed interval responses. That is, an increase in parameter variance and a decrease in parameter resolution would be observed with smaller vertical cell dimensions.

Moreover, comparing Figures III.5a and III.5b reveals that although head responses induced by multiple slug tests are sensitive to hydraulic properties throughout the whole ROI and, to a lesser extent, at a fair distance away beyond the observation well, the resolution potential for K_h, K_v/K_h and S_s is essentially focused on the stressed and observation wells. This is consistent with previous sensitivity and resolution

analysis using pumping tests [e.g., *Vasco et al.*, 1997; *Clemo et al.*, 2003; *Bohling,* 2009]. Thus, the governing groundwater flow equation that describes hydraulic head propagation into an aquifer and the geometry of the tomographic experiment itself are major limitations to spatial resolutions of hydraulic properties, regardless of the test initiation method (pulse versus constant rate). This sensitivity and resolution analysis clearly illustrates that head sensitivity to parameters is not a condition sufficient to ensure the resolution of hydraulic properties in an aquifer, and correlations between parameters must be taken into account [*Aster et al.*, 2005]. That contradicts the recent claims by *Huang et al.* [2011] and *Sun et al.* [2013], which largely stems from their synthetic numerical inversion with noise-free data. Note that the truncated SVD inversion of the sensitivity matrix for our previous synthetic tomographic experiment with noise-free head data obviously leads to a resolution close to unity for each hydraulic property (K_h, K_v/K_h and S_s) within the ROI (Table III.4).

3.4 Analysis of factors affecting resolution of tomographic slug tests

This section explores the effects of various experimental parameters that can be considered in the design of a tomographic experiment in order to maximize the resolution potential of an experiment. As pointed out in the previous sections, the factors affecting the resolution of a tomographic experiment are basically the head measurement error, the absolute magnitude of the sensitivities and the correlation between the parameters (sensitivity correlation). Thus by means of resolution analysis, we varied different experimental parameters to illustrate their impact on resolution factors and corresponding hydraulic property resolution. For each scenario, the resolution of parameters within the ROI is calculated and compared to the previous tomographic experiment in Section 3.3 (base case scenario) for discussion. As previously done, resolution calculation is based on truncated SVD inversion of the sensitivity matrix with a relative parameter error of 1%. For all scenarios the vertical cell dimension is also kept constant at 0.61 m, i.e. the length of the stressed intervals.

The relative magnitude of the sensitivity compared to the measurement error (signal-to-noise ratio) is a critical consideration for parameter resolution. If the signal-to-noise ratio is small, it is difficult to separate the real signal from the noise caused by measurement errors and thus obtain reliable parameter estimates. For instance, Figure III.4a-b shows how the sensitivity curves for adjacent regions can be close together. Although a noise-free inversion could recover exact parameter values, adding noise to head measurements would obviously make it difficult to separate the respective contributions of each

parameter. Table III.3 shows that using noise-free head data (white noise with a standard deviation of $2x10^{-8}$ m) results in perfectly resolved hydraulic parameters within the ROI.

Although the noise level is dependent on the technology available to record head measurements, the signal-to-noise ratio can also be improved by increasing the initial head of the slug used to initiate tests in order to increase the relative magnitude of the sensitivity compared to the measurement error. Figure III.8a and Table III.3 show the evolution of the average resolution for K_h, K_v/K_h and S_s within the ROI with respect to the initial head. Those curves show that up to a certain initial head, the resolution is not improved significantly. Our interpretation is that when the head measurements in response to a hydraulic perturbation are large enough to be differenced from the noise, the resolution potential is thus only controlled by sensitivity correlation, which is inherent to the actual tomographic experiment (e.g., governing equation, source form, tests configuration).

We saw in Figures III.4a-b that sensitivity curves for parameters in the middle of the ROI were highly correlated, which resulted in low resolution values for this region, as illustrated in Figure III.5b. Meanwhile, as estimates of the parameters are weighted averages of the true model parameters in the vicinity of the incompletely resolved cells (Figure III.7), an inversion can be regularized through discretization to lower sensitivity correlation, and provide parameter estimates with higher resolution [*Aster et al.*, 2005]. Regularization through discretization can be achieved by reducing the number of cells between the stressed and observation wells, either by increasing the lateral size of the parameter grid cells or by decreasing the distance between wells while keeping the size of the 143-parameter grid cells constant. With the former strategy, we reduce the cell resolution of the experiment, whereas with the latter an increase in the cell resolution is achieved at the expense of a loss in spatial coverage (cell-coverage tradeoff).

Figure III.8b and Table III.3 show the evolution of the average resolution for K_h, K_v/K_h and S_s within the ROI with respect to the average lateral cell size. For this analysis, the horizontally adjacent cells of the 143-parameter grid (Figure III.1) were merged together to separate the ROI in two, three or four parts, while cell dimensions outside the ROI were kept as in the 143-parameter grid cells. The resolution curves for K_h, K_v/K_h and S_s show that the resolution for each property increases with increasing cell size, as expected from the resolution matrix in Figure III.7. Maximum resolution is achieved when two columns of cells separate the wells, for corresponding cell size of lateral dimension of 4.38 m. However, while the

selection of a coarse grid for inversion increases the individual resolution within each cell, it may also induce a bias in the estimated parameters. This bias is generally expressed by the smoothing of the estimated parameters. For aquifers showing gradual variations in hydraulic properties, this averaging procedure could be acceptable, but for highly heterogeneous distributions with sharp changes in property values the effects of this discretization should be evaluated carefully. This is also valid for vertical heterogeneity assessment.

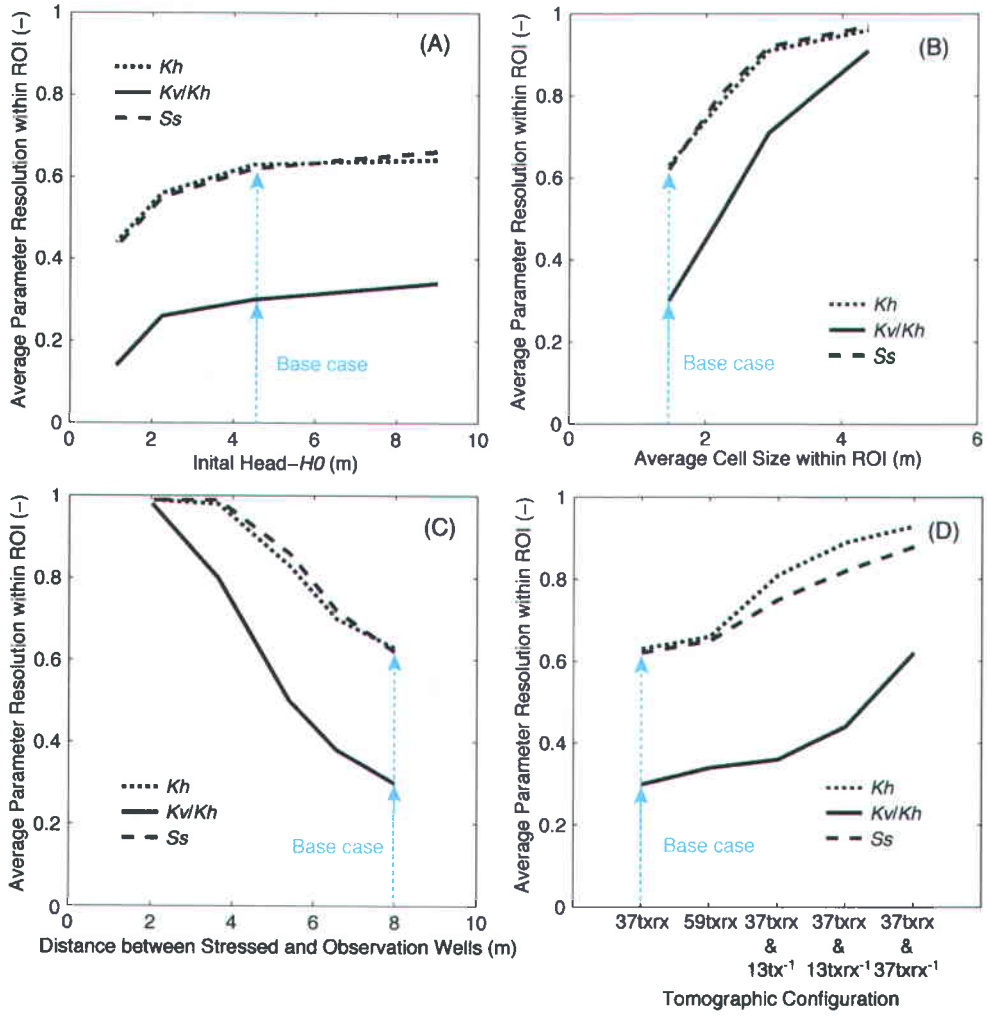


Figure III.8 Average parameter resolution within the region of investigation (ROI) for various scenarios of: (a) initial head (H_0) used to initiate slug tests; (b) cell size; (c) distance between stressed and observation wells; and (d) tomographic configuration using different sets of stressed and observation intervals. Results are for an analysis of the 10000-sec records based on truncated SVD inversion of the sensitivity matrix associated to respective scenario with a relative parameter error of 1%. Simulation parameters and average resolution values for hydraulic properties within the region of investigation (ROI) for each scenario are compiled in Table III.3. Blue arrows point to results corresponding to conditions used for the base case.

Figure III.8c and Table III.3 show the evolution of the average resolution for K_h, K_v/K_h and S_s within the ROI with respect to the distance between the stressed and observation wells. For this analysis, the distance between the stressed and observation wells was gradually decreased, while keeping the size of the 143-parameter grid cell constant. The resolution curves in Figure III.8c show that the resolution potential increases with decreasing distance between stressed and observation wells, and maximum resolution is achieved when two column of cells separate the wells, for corresponding lateral cell size of 1.46 m and an inter-well distance of 2.03 m. The comparison of Figures III.8b and III.8c thus reveals that the resolution potential is effectively tied to the sensitivity matrix with resolution focused on the stressed and the observation wells, regardless of the size of the cells. This means that the major limitation of lateral resolution of tomographic slug tests is not the radius of influence of a slug test, but instead the sensitivity correlation between parameters. When designing a tomographic experiment, one should thus consider the tradeoff between cell resolution and spatial coverage with respect to the scale of aquifer heterogeneity considered. For instance, highly heterogeneous aquifers will require closely spaced wells for higher cell resolution in order to reduce bias in hydraulic property estimates, whereas reasonable estimates can be achieved with more distant wells (coarser cells) within aquifers with more gradual changes in properties.

As shown in Figures III.2a-d, sensitivity curves for stressed and observation intervals are different, which suggests that reciprocal tests where roles of stressed and observation wells are interchanged can be used to reduce sensitivity correlation between parameters and thus increase parameter resolution within the ROI. To evaluate the potential of resolution of a tomographic experiment using reciprocal tests, we combined sensitivity values of two independent simulations originating at different wells, to simulate the interchanged roles of stressed and observation wells, because the actual radial flow model does not allow simultaneous simulation of reciprocal tests. The sensitivities evaluated for every cell, time and stressed and observation intervals of each reciprocal simulation are then combined on a common 143-parameter grid with sensitivity values adjusted to account for the change in cell width and symmetry of the sensitivities around wells. The resolution matrix is thereafter computed from this combined sensitivity matrix.

For this analysis, Figure III.8d and Table III.3 present average resolution for K_h, K_v/K_h and S_s within the ROI for different configurations of stressed and observation intervals, including the base case configuration (37txrx) illustrated in Figure III.1c. Three different configurations using reciprocal tests are considered: (1) the base case with 13 reciprocal slug tests using only stressed interval responses

(37txrx&13tx^{-1}), (2) the base case with 13 reciprocal slug tests using stressed interval responses and only one observation interval response per test (37txrx&13txrx^{-1}), and (3) the base case with its reciprocal mirror image (37txrx&37txrx^{-1}). We also consider a configuration without reciprocal tests similar to the base case configuration, but using a total of 59 observation intervals spread over a larger range of observation angles (59txrx).

As shown in Figure III.8d, using 59 observation intervals (59txrx) instead of 37 (37txrx) only slightly increases the resolution of hydraulic properties within the ROI. In contrast, the reciprocal configuration (37txrx&13txrx^{-1}) that uses approximately the same additional number of hydraulic responses provides a much better the resolution potential (37txrx & 13txrx^{-1} and 59txrx use 26 and 22 additional intervals, respectively). Using only 13 reciprocal stressed interval responses (37txrx&13tx^{-1}) is even slightly better. Obviously, using the base case with its reciprocal mirror image (37txrx&37txrx^{-1}) considerably increases the resolution within the ROI, especially for K_v/K_h. We evaluated that using this configuration, a cell resolution of 2.92 m (three column of cells) can be achieved with corresponding hydraulic property resolutions over 0.95 for K_h, K_v/K_h and S_s within the ROI ("37txrx&37txrx^{-1} resolution" in Table III.3). This clearly demonstrates that reciprocal tests are far more informative than observation intervals added to existing stressed intervals. That is, reciprocal tests better contribute to decrease the level of correlation between parameters within the ROI, which is achieved to a lesser extent with additional observation intervals because of the higher redundancy in the sensitivity curves and the general decrease in sensitivity amplitudes with increasing observation interval angles.

3.5 Principle of reciprocity for slug tests

To further understand the information contained in stressed and observation interval responses, the implications of the principle of reciprocity for tomographic slug tests are investigated. If the principle of reciprocity applies to tomographic slug tests, a change in the direction of slug testing (switching stressed and observation interval locations) would not affect observation interval responses, regardless of the degree of aquifer heterogeneity [*McKinley et al.*, 1968; *Bruggeman*, 1972; *Falade*, 1981; *Barker*, 1991]. Figures III.9a and III.9b illustrate heterogeneous (layered) aquifer models and test configurations used for the reciprocity analysis, where interchanging the roles of a pair of stressed-observation intervals simulates reciprocal tests. For instance, the two reciprocal responses are obtained from a first simulation with the stressed interval at Well 1 and the observation interval at Well 2 and the second simulation with the stressed and observation intervals at Well 2 and Well 1, respectively. Moreover, two different

heterogeneous aquifer models are used to simulate symmetrical and asymmetrical test configurations relative to heterogeneity. Symmetrical and asymmetrical test configuration means that the heterogeneity patterns seen by the traveling head perturbation for each of the two reciprocal tests are mirror or else flipped images, respectively. For both cases, the heterogeneity of the model is changed by varying one property at a time while holding the two other properties constant, as presented in Table III.4. It should be noted that for reciprocal simulations two grids with different origins but with identical representation of the heterogeneity are used to take into account the grid axisymmetry. For instance, a first grid with finer cells at Well 1 is used for the simulations with tx11 (red path in Figure III.9) and a second grid originating at Well 2 for simulations with tx3 (blue path in Figure III.9). We note also that the stressed and observation interval locations are symmetrical to the upper and lower boundaries. In this paper, heterogeneous models are used only to demonstrate the applicability of the principle of reciprocity with slug tests.

Figures III.9a and III.9b present results of reciprocal simulations for symmetrical and asymmetrical cases, respectively, with varying one hydraulic property value for the region 2 at the time (Table III.4). Obviously, head responses for the symmetrical case are identical whether the test is initiated in Well 1 or Well 2 (Figure III.9a). For the asymmetrical case, however, the observation interval responses are clearly different for the two reciprocal tests, in particular with respect to K_h heterogeneity (Figure III.9b). Thus, observation interval responses resulting from reciprocal slug tests cannot be interchanged for test configurations asymmetrical to K_h heterogeneity, which refutes the principle of reciprocity.

The principle of reciprocity is not applicable for slug tests with test configurations asymmetrical to heterogeneity because sensitivity patterns for the two reciprocal tests are different. As previously noted, sensitivities are functions of the flow rate pattern induced by the slug test, which is controlled by both the hydraulic properties and the intensity of the source. Obviously, the intensity of the source for a slug test varies with time following head variations in the stressed interval (Figure III.2a), which is maximal for the initial time and gradually decrease to zero at the end of the test. For a slug test, the rate of head decreases and the total duration of a test are both controlled by hydraulic properties of material near the stressed interval (Figure III.3a), which differ for the two reciprocal tests. Consequently, the underlying sensitivity patterns that shape reciprocal observation interval responses are also different and depend on the spatial distribution of aquifer heterogeneity between the stressed and observation intervals.

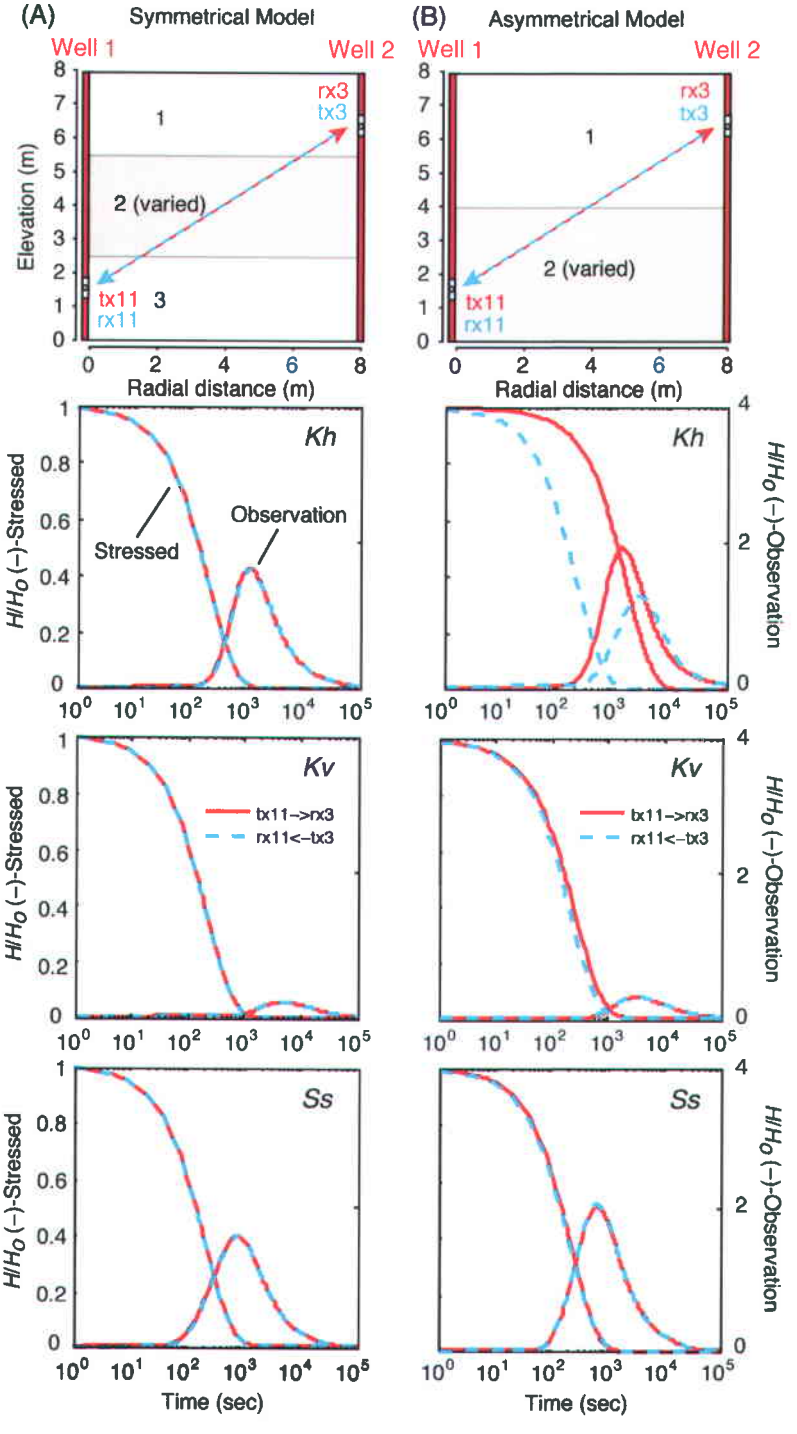


Figure III.9 Normalized head (H/H_0) responses for reciprocal slug tests for: (a) a symmetrical; and (b) asymmetrical test configuration relative to heterogeneity in K_h, K_v and S_s. Alternatively changing values of K_h, K_v/K_h and S_s in the different layers are compiled in Table III.4.

Table III.4 Hydraulic parameters used for the simulations to illustrate the principle of reciprocity with slug tests.

Simulation	Region	K_h (m/s)	K_v (m/s)	S_s (m^{-1})
Symmetrical and asymmetrical: K_h change	2	$1x10^{-6}$	$1x10^{-6}$	$1x10^{-4}$
Symmetrical and asymmetrical: K_v change	2	$1x10^{-5}$	$1x10^{-7}$	$1x10^{-4}$
Symmetrical and asymmetrical: S_s change	2	$1x10^{-5}$	$1x10^{-6}$	$1x10^{-5}$
For all simulations	1 and 3	$1x10^{-5}$	$1x10^{-6}$	$1x10^{-4}$

It should be noted that the principle of reciprocity does not hold also for reciprocal constant rate pumping tests with significant wellbore storage [*Deng*, 1996]. For a pumping test, the total pumping rate is the summation of the aquifer pumping rate and the wellbore storage pumping rate [*Park and Zhan*, 2002]. The wellbore storage supplies most of the initial pumped water, and initially equals to the total pumping rate and gradually decreases to zero when pumping continues. In contrast, the aquifer pumping rate is initially zero and gradually approaches the total pumping rate when pumping continues. Thus, because the aquifer pumping rate is controlled by the hydraulic properties of material near the stressed pumping well, similar to slug tests, reciprocal pumping tests in asymmetrical heterogeneous aquifers exhibit different sensitivity patterns. The principle of reciprocity is applicable for pumping tests regardless of the heterogeneity only when it can be assumed that the total pumping rate is instantaneously applied to the aquifer (line source assumption assuming no wellbore storage pumping rate). Note that the applicability of the principle of reciprocity for reciprocal pumping tests, as discussed above with wellbore storage and line source, have been verified numerically with the heterogeneous models depicted in Figures III.9a-b. The applicability of the principle of reciprocity is then related to wellbore storage that acts differently for pumping and slug tests.

Thus, the observation interval response contains information about hydraulic properties between stressed and observation intervals, as depicted in Figure III.3b, but also about the intensity and shape of the source that produced the head perturbation recorded in the observation interval, which are controlled by material near the stressed interval (Figure III.3a). Consequently, the non-applicability of the principle of

reciprocity for slug tests requires that both stressed and observation interval responses be used together for tomographic slug test analysis in order to isolate the influence of the stressed interval response on the observation interval response. We note that independent analysis of each reciprocal test using both stressed and observation interval responses provide the same average diffusivity value, regardless of the heterogeneity [*Streltsova*, 1988]. That may have implications for tomographic slug tests analysis based on ray tracing techniques using only observation interval responses [e.g., *Brauchler et al.*, 2007, 2010 and 2011; *Hu et al.*, 2011], where observation interval responses may need to be transformed with respect to the shape of the stressed interval responses.

4 Summary and conclusions

Through a sensitivity and resolution analysis, this study examined the information content of a transient analysis of tomographic slug test data for characterizing K_h, K_v/K_h and S_s heterogeneities for a synthetic case based on the known characteristics of a littoral aquifer showing K anisotropy at the scale of the experiment. The analysis used a radial flow model enabling the simulation of wellbore storage effects to represents hydraulic head propagation induced by slug tests in the plane encompassing the stressed and observation wells. The first-order sensitivity analysis presented involved small perturbations from a homogeneous and anisotropic model, which serves as a linear approximation of the effects of heterogeneity on hydraulic head responses recorded in stressed and observation intervals used for tomography. Resolution analysis was carried out using singular value decomposition of the sensitivity matrix, as a convenient way of analyzing least-squares inverse problems, in the absence of an explicit regularization term and using a noise level representative of field conditions. Various synthetic experiments were assessed to show how head measurement error, sensitivity magnitude and sensitivity correlation are affecting resolution potential of tomographic slug tests, and simple heterogeneous models were also used to verify the applicability of the principle of reciprocity for slug tests. The main conclusions of this study can be stated as follows:

- A transient analysis of tomographic slug test data is capable of resolving independently K_h, K_v/K_h and S_s within the region of investigation (ROI) encompassed by the stressed and observation wells. This resolution is better achieved with the combined use of stressed and observation interval data because of 1) the generally weak correlation of the sensitivity curves for hydraulic properties within the ROI, as expressed by their different temporal behaviors, and 2) the non-applicability of the principle of

reciprocity to slug tests for heterogeneous aquifers, which leads to a dependence of observation interval responses to the direction of testing, that can only be resolved with the use of stressed interval responses. The resolution potential for K_v/K_h is however generally lower than for K_h and S_s likely due to the test initiation method that induces predominantly horizontal flow despite using stressed intervals with a small aspect ratio.

- Hydraulic properties outside the ROI exert only a limited influence on the resolution of K_h, K_v/K_h and S_s within the ROI because of 1) the attenuation over space and time of the slug test response, which reduces the relative sensitivities for parameters beyond the observation well with respect to sensitivities for parameters within the ROI, and 2) the distinct behavior of sensitivities for parameters within and outside the ROI during passage of the head perturbation over the observation intervals. In other words, parameters outside the ROI either have no influence on the observation responses (K_v/K_h) or their effects can be isolated (K_h and S_s) using the entire head records in the analysis.
- The resolution potential of tomographic slug tests for K_h, K_v/K_h and S_s with a noise level representative of field conditions is essentially focused on the stressed and observation wells as a result of the high correlation between parameters within or outside the ROI. In particular for the ROI, hydraulic head responses are sensitive to each individual parameter, but only their global behavior can be resolved. Although reciprocal tests, where the role of stressed and observations intervals are interchanged, greatly reduces the correlation between parameters in adjacent regions of the aquifer, not much can be done in practice to increase the lateral resolution using head data alone, even using a dense network of stressed and observation intervals.

While transient sensitivity patterns induced by slug tests could provide different resolution potential than pumping tests due to different behaviors of respective sensitivity curves, the same general limitation of the governing groundwater flow equation applies on the lateral resolution of hydraulic properties between wells [*Vasco et al.*, 1997; *Clemo et al.*, 2003; *Bohling*, 2009]. Despite this limitation, the analysis of tomographic slug test data could be viewed as a mean to estimate local heterogeneity of hydraulic properties [*Vasco et al.*, 1997], where the inversion can be regularized through grid discretization. Considering that only a few column of cells between stressed and observation wells can be effectively resolved, characterization of K_h, K_v/K_h and S_s with tomographic slug tests data is thus a tradeoff between cell resolution (small cells with closely spaced wells) and spatial coverage (coarse cells with distant wells), where the scale of an experiment need to be adapted to the degree of heterogeneity of an aquifer to reduce bias in parameter estimates. For many hydrogeological investigations, estimates of local heterogeneity in K_h, K_v/K_h and S_s could however provide the necessary level of information needed.

Meanwhile, the anisotropic framework offered by a transient analysis of tomographic slug tests, which potentially could also be extended to tomographic pumping tests, provides greater confidence in the estimated model parameters [*Babuska and Cara*, 1991]. Even if the result is that no anisotropy is required, one can have greater confidence in the isotropic model thus obtained if the analysis method is set in terms of an anisotropic framework. This knowledge about anisotropy could also be used to further constrain geostatistical parameters required by commonly used geostatistically based inversion approaches [e.g., *Yeh et al.*, 1995, 1996; *Kitanidis and Vomvoris*, 1983; *Kitanidis*, 1995], as inappropriate use of geostatistical parameters could induce considerable errors in hydraulic property estimates [*Cardiff and Barrash*, 2011].

Although this study investigated the resolution for a two-dimensional wedge of an aquifer with a radial flow model, results of this study are deemed to reflect the fundamentals of tomographic slug tests. Under real field conditions, radially asymmetric heterogeneity relative to the stressed well may induce significant angular variations influx toward the stressed well and bias head responses measured in the plane between the stressed and observation wells. For the sensitivity and resolution analysis presented in this paper, such situation is obviously not a concern because small perturbations of parameters from a homogeneous and anisotropic model can be considered radially symmetric [*Bohling*, 2009]. Moreover, such representation of an aquifer with a radial flow model is considered a fair approximation for layered sequence of sediments with predominantly vertical rather than lateral hydraulic properties variations. Indeed, sensitivity to hydraulic properties for slug tests are heavily focused near the stressed well (Figures III.3a-b), with a sharp decrease in sensitivity over time (Figure III.2a) and distance (Figure III.3a). Thus, sensitivity patterns for K_h, K_v/K_h and S_s around the stressed and observation wells in a 2D areal view are expected to be approximately symmetric with respect to the stressed well, with maximum influence of asymmetric heterogeneity on the stressed interval response at early times when sensitivities are maximal. The 2D areal sensitivity patterns for slug tests using stressed and observation intervals in the analysis is likely different from sensitivity patterns for pumping tests with stressed and observation intervals, which are asymmetrical around pumping wells [e.g., *Oliver*, 1993; *Leven and Dietrich*, 2006], due to the difference in the form of the source used to initiate the respective tests (pulsed versus constant rate). Then as long as lateral hydraulic properties variations, especially K_h, are gradual around the stressed well, head responses measured in the plane between the stressed and observation wells can be considered fairly unbiased and the use of a two-dimensional radial flow model fully justified.

This study used a sensitivity and resolution analysis assuming a homogeneous and anisotropic aquifer model and further work will have to be carried out to address the resolution potential of tomographic slug tests related to the geometry of the heterogeneities (e.g., continuous or discontinuous layers), their contrast in hydraulic property values (e.g., conduit or barrier flow), as well as on the appropriate experimental design to capture those features. A field-based application of a transient analysis of tomographic slug tests would also be the next logical step to further demonstrate the potential of the approach.

Acknowledgements

This study was supported by the Geological Survey of Canada as part of the Groundwater Resources Inventory program, by the Régie intermunicipale de gestion des déchets des Chutes-de-la-Chaudière and by NSERC Discovery Grants held by R.L. and E.G. This is ESS contribution number 20090447.

References

Anderson, M.P. (1989), Hydrogeologic facies models to delineate large-scale spatial trends in glacial and glaciofluvial sediments, Geol. Soc. Am. Bull. 101, 501–511.

Aster, R. C., B. Borchers, and C. H. Thurber (2005), Parameter Estimation and Inverse Problems, 301 pp., Elsevier, Amsterdam.

Babuska, V., and M. Cara (1991), Seismic anisotropy in the earth, 217 pp. Kluwer Academic Publishers, Boston.

Barker, J. A. (1991), The reciprocity principle and an analytical solution for Darcian flow in a network, Water Resour. Res., 27 (5), 743-746.

Belitz, K., and W. Dripps (1999), Cross-well slug testing in unconfined aquifers: a case study from the sleepers river watershed, Vermont, Ground Water, 37 (3), 438-447.

Berg, S. J., and W. A. Illman (2011), Three-dimensional transient hydraulic tomography in a highly heterogeneous glaciofluvial aquifer-aquitard system, Water Resour. Res., 47, W10507, 10.1029/2011WR010616.

Bohling, G. C. (2008), Information fusion in regularized inversion of tomographic pumping tests, in Quantitative Information for Hydrological Sciences, Stud. Comput. Intell., 79, 137–162, Sringer, Berlin.

Bohling, G. C. (2009), Sensitivity and resolution of tomographic pumping tests in an alluvial aquifer, Water Resour. Res., 45, W02420, 10.1029/2008WR007249.

Bohling, G. C., and J. J., Butler Jr. (2001), Lr2dinv: A finite-difference model for inverse analysis of two-dimensional linear or radial groundwater flow, Comput. Geosci., 27, 1147-1156.

Bouwer, H., and R.C. Rice (1976) A slug test method for determining hydraulic conductivity of unconfined aquifers with completely or partially penetrating wells. Water Resour. Res., 12(3): 423-428.

Brauchler, R., J. Cheng, M. Everette, B. Johnson, P. Dietrich, R. Liedl, and M. Sauter (2007), An inversion strategy for hydraulic tomography: Coupling travel time and amplitude inversion, J. Hydrol., 345, 184-198, 10.1016/j.jhydrol.2007.08.011.

Brauchler, R., R. Hu, P. Dietrich, and M. Sauter (2011), A field assessment of high-resolution aquifer characterization based on hydraulic travel time and hydraulic attenuation tomography, Water Resour. Res., 47, W03503, 10.1029/2010WR009635.

Brauchler, R., R. Hu, T. Vogt, D. Halbouni, T. Heinrichs, T. Ptak, and M. Sauter (2010), Cross-well slug interference tests: An effective characterization method for resolving aquifer heterogeneity, J. Hydrol., 384, 33-45, 10.1016/j.jhydrol.2010.01.004.

Bruggeman, G. A. (1972), The reciprocity principle in flow through heterogeneous porous media, In Fundamentals of Transport Phenomena in Porous Media, Developments in soil science, 2, 136-149 pp., Elsevier, Amsterdam.

Butler Jr., J. J. (1990), The role of pumping tests in site characterization: Some theoretical considerations, Ground Water, 28(3), 394-402, 10.1111/j.1745-6584.1990.tb02269.x.

Butler Jr., J. J. (1995), Slug tests with observation wells; extension of Hyder et al. (1994) Solution to case of head in an observation well, Kansas Geological Survey, Open-file Report no. 95–43, 14 pp.

Butler Jr., J. J., and C. D. McElwee (1990), Variable-rate pumping tests for radially symmetric nonuniform aquifers, Water Resour. Res., 26(2), 291– 306.

Butler Jr., J. J., and C. D. McElwee (1995), Well-testing methodology for characterizing heterogeneities in alluvial-aquifer systems: Final technical report, Kans. Geol. Surv. Open File Report no. 75-95.

Butler Jr., J. J., and X. Zhan (2004), Hydraulic tests in highly permeable aquifers, Water Resour. Res., 40, W12402, 10.1029/2003WR002998.

Cardiff, M., and W. Barrash (2011), 3-D transient hydraulic tomography in unconfined aquifers with fast drainage response, Water Resour. Res., 47, W12518, doi:10.1029/2010WR010367.

Carrera, J., A. Alcolea, A. Medina, J. Hidalgo, and L. J. Slooten (2005), Inverse problem in hydrogeology, Hydrogeol. J., 13, 206-222, 10.1007/s10040-004-0404-7.

Cardiff, M., T. Bakhos, P. K. Kitanidis, and W. Barrash (2013), Aquifer heterogeneity characterization with oscillatory pumping: Sensitivity analysis and imaging potential, Water Resour. Res., 49: 5395-5410, doi:10.1002/wrcr.20356.

Cheng, J., R. Brauchler, and E. Everett (2009), Comparison of early and late travel times of pressure pulses induced by multilevel slug tests, Eng. Appl. Comput. Fluid Mech., 3(4), 514–529.

Clemo, T., P. Michaels, and R. M. Lehman (2003), Transmissivity resolution obtained from the inversion of transient and pseudo-steady drawdown measurements, in Proceedings of MODFLOW and More 2003 Understanding Through Modeling, pp. 629-633, Int. Ground Water Model. Cent., Golden Colo.

Deng, X. (1996), Description of heterogeneous reservoirs using pressure and tracer data, Ph.D. dissertation, Stanford University, 170 pp.

Falade, G. K. (1981), Analysis of the Reciprocity Concept in a Porous Medium, Water Resour. Res., 17(4), 918-920.

Fienen, M. N., T. Clemo, and P. K. Kitanids (2008), An interactive Bayesian geostatistical inverse protocol for hydraulic tomography, Water Resour. Res., 44, W00B01, 10.1029/2007WR006730.

Hu, R., R. Brauchler, M. Herold, P. Bayer (2011), Hydraulic tomography analog outcrop study: Combining travel time and steady shape inversion, J. Hydrol., 409 (1-2), 350-362, 10.1016/j.jhydrol.2011.08.031.

Huang, S.-Y., J.-C. Wen, T.-C. J. Yeh, W. Lu, H.-L. Juan, C.-M. Tseng, J.-H. Lee, and K.-C. Chang (2011), Robustness of joint interpretation of sequential pumping tests: Numerical and field experiments, Water Resour. Res., 47, W10530, doi:10.1029/ 2011WR010698.

Hyder, Z., J. J. Butler Jr., C. D. McElwee, and W. Liu (1994), Slug test in partially penetrating wells, Water Resour. Res., 30(8), 2945-2957.

Jimenéz, S., R. Brauchler, and P. Bayer (2013), A new sequential procedure for hydraulic tomographic inversion, Adv. Water Resour., 62:59-70.

Kitanidis, P. K. (1995), Quasi-linear geostatistical theory for inversing, Water Resour. Res., 31(10), 2411–2419.

Kitanidis, P. K., and E. Vomvoris (1983), A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimen- sional simulations, Water Resour. Res., 19(3), 677–690.

Leven, C., and P. Dietrich (2006), What information can we get from pumping tests?-Comparing pumping test configurations using sensitivity coefficients, J. Hydrol., 319(1-4), 199-215, doi:10.1016/ j.jhydrol.2005.06.030.

Mao, D., T.-C. J. Yeh, L. Wan, C.-H. Lee, K.-C. Hsu, J.-C. Wen, and W. Lu (2013), Cross-correlation analysis and information content of observed heads during pumping in unconfined aquifers, Water Resour. Res., 49, doi:10.1002/wrcr.20066.

McElwee, C. D., and M. A. Yukler (1978), Sensitivity of groundwater models with respect to variations in transmissivity and storage, Water Resour. Res., 14(3), 451–459, 10.1029/WR014i003p00451.

McElwee, C.D., G.C. Bohling, and J. J. Butler Jr. (1995a), Sensitivity analysis of slug tests I: The slugged well, J. Hydrol., 164, 53-67.

McElwee, C.D., J. J. Butler Jr., and G. C. Bohling (1995b) Sensitivity analysis of slug tests II: Observation wells, J. Hydrol., 164, 69-87.

McKinley, R. M, S. Vela, and L. A. Carlton (1968), A field application of pulse-testing for detailed reservoir description, J. Petrol. Tech., 20 (3), 313-321.

Menke, W. (2012), Geophysical Data Analysis: Discrete Inverse Theory, 3rd ed., 293 pp., Academic Press.

Oliver, D. S. (1993), The influence of nonuniform transmissivity and storativity on drawdown, Water Resour. Res., 29(1), 169-178, doi:10.1029/92WR02061

Paradis, D., and R. Lefebvre (2013), Single-well interference slug tests to assess the vertical hydraulic conductivity of unconsolidated aquifers, J. Hydrol., 478(25), 102-118, 10.1016/j.jhydrol.2012.11.047.

Paradis, D., L. Tremblay, R. Lefebvre, E. Gloaguen, A. Rivera, M. Parent, J.-M. Ballard, Y. Michaud, and P. Brunet (2014), Field characterization and data integration to define the hydraulic heterogeneity of a shallow granular aquifer at a sub-watershed scale. Accepted for publication in Environ. Earth Sci., February 1st, 2014.

Paradis, D., R. Lefebvre, R. H. Morin, and E. Gloaguen (2011), Permeability profiles in granular aquifers using flowmeters in direct-push wells, Ground Water, 49, 534–547, 10.1111/j.1745-6584.2010.00761.x.

Park, E., and H. Zhan (2002), Hydraulics of a finite-diameter horizontal well with wellbore storage and skin effect, Adv. Water Resour., 25(4), 389-400, 10.1016/S0309-1708(02)00011-8.

Prats, M., J. B. Scott (1975), Effect of wellbore storage on pulse-test pressure response, J. Petrol. Tech., 707–709.

Rubin, Y., and S. S. Hubbard (Eds.) (2005), Hydrogeophysics, 523 pp., Springer, Dordrecht, Netherlands.

Sageev, A. (1986), Slug test analysis. Water Resour. Res., 22(8), 1323-1333.

Spane Jr., F.A. (1996), Applicability of slug interference tests for hydraulic characterization of unconfined aquifers: (1) analytical assessment, Ground Water, 34(1), 66-74.

Streltsova, T. D. (1988), Well Testing in Heterogeneous Formations, John Wiley & Sons, 413 p.

Sun, R., T.-C. J. Yeh, D. Mao, M. Jin, W. Lu, and Y. Hao (2013), A temporal sampling strategy for hydraulic tomography analysis, Water Resour. Res., 49, 3881-3896, doi:10.1002/wrcr.20337.

Vasco, D. W., A. Datta-Gupta, and J. C. S. Long (1997), Resolution and uncertainty in hydrologic characterization, Water Resour. Res., 33(3), 379-397, 10.1029/96WR03301.

Vogel, C. R. (2002), Computational Methods for Inverse Problems, 183 pp., Soc. for Ind. and Appl. Math., Philadelphia, Pa.

Yeh, T.-C. J., A. Gutjahr, and M. Jin (1995), An iterative cokriging-like technique for groundwater flow modeling, Ground Water, 33(1), 33–41.

Yeh, T.-C. J., M. Jin, and S. Hanna (1996), An iterative stochastic inverse method: Conditional effective transmissivity and hydraulic head fields, Water Resour. Res., 32(1), 85–92.

Zlotnik, V. (1994), Interpretation of Slug and Packer Tests in Anisotropic Aquifers. Ground Water, 32:761-766, doi: 10.1111/j.1745-6584.1994.tb00917.x

ARTICLE IV

TRANSIENT ANALYSIS OF TOMOGRAPHIC SLUG TESTS IN A HETEROGENEOUS AND ANISOTROPIC UNCONSOLIDATED AQUIFER

Résumé

La caractérisation des propriétés hydrauliques des aquifères devrait idéalement donner la distribution spatiale de la conductivité hydraulique (K), de l'anisotropie de K (rapport de K vertical et horizontal, K_v/K_h) et de l'emmagasinement spécifique (S_s). *Paradis et al.* [2014] a démontré qu'une analyse transitoire de tomographie par choc hydraulique peut imager l'hétérogénéité de K_h, K_v/K_h et S_s à une échelle de quelques mètres. La tomographie par choc hydraulique comprend la réalisation d'une série d'essais de perméabilité à choc hydraulique dans un puits émetteur avec l'enregistrement simultané des charges hydrauliques dans les intervalles émetteurs et dans des intervalles d'observations isolés dans un puits à proximité. Cet article présente l'analyse des données obtenues au cours d'un essai de terrain réalisé dans un aquifère de silts et de sables littoraux modérément hétérogène et très anisotrope. L'objectif est d'évaluer les résultats d'une analyse transitoire de l'essai de tomographie par choc hydraulique. L'algorithme de *Bohling et Butler* [2001] est utilisé pour l'inversion numérique transitoire des mesures de charges hydrauliques prisent dans plusieurs intervalles émetteurs et d'observations, en considérant l'emmagasinement de puits. La justesse des tomogrammes de K_h, K_v/K_h et S_s est vérifiée par la simulation de tests hydrauliques à l'échelle globale et locale, non utilisés dans l'inversion, ainsi qu'avec des profils de K_h et K_v obtenus de tests indépendants de terrain et de laboratoire. Les résultats indiquent que les estimations de K_h et K_v par tomographie se comparent bien aux autres types de tests. De plus, les différences entre les charges observées et prédites pour les tests multi-échelles résultent en seulement de petites erreurs d'estimation des paramètres hydrauliques. Cette étude démontre donc la capacité de la tomographie par choc hydraulique à définir des champs hétérogènes de K_h, K_v/K_h et S_s entre puits.

En préparation pour *Water Resources Research*

Transient analysis of tomographic slug tests in a heterogeneous and anisotropic unconsolidated aquifer

Daniel Paradis[1,2,*], Erwan Gloaguen[2], René Lefebvre[2] and Bernard Giroux[2]

1: Geological Survey of Canada, 490 rue de la Couronne, Quebec City, Canada G1K 9A9

2 : Institut national de la recherche scientifique, Centre Eau Terre Environnement (INRS-ETE), 490 rue de la Couronne, Quebec City, Canada G1K 9A9

*Corresponding author: Daniel Paradis

Phone: (418) 654-3713 Fax: (418) 654-2604 E-mail: dparadis@nrcan.gc.ca

Abstract

Hydraulic aquifer characterization should ideally provide the spatial distribution of hydraulic conductivity (K), K anisotropy (ratio of vertical and horizontal K, K_v/K_h) and specific storage (S_s) *Paradis et al.* [2014] have shown that a transient analysis of inter-well tomographic slug tests could assess heterogeneities in K_h, K_v/K_h and S_s at the scale of a few meters assuming that S_s is the sole storage parameter affecting slug test head responses. Tomographic slug tests involve a series of packer slug tests carried out in a well and the simultaneous recording of heads in the stressed interval and in observation intervals in an adjacent well. This paper presents the processing of data obtained from a field test carried out in a moderately heterogeneous and highly anisotropic silts and sands aquifer. The objective is to assess the results of a transient analysis of the tomographic slug tests. The algorithm of *Bohling and Butler* [2001] is used for the transient inversion of heads recorded in multiple stressed and observation intervals, considering wellbore storage. The accuracy of K_h, K_v/K_h and S_s tomograms is verified by simulation of large-scale and local hydraulic tests, not used in inversion, and with profiles of K_h and K_v from other field and laboratory tests. K_h and K_v estimates from tomographic slug tests compare well with other measurements. Furthermore, differences between observed and predicted heads from independent multi-scale tests (inter-well and vertical interference slug tests) show small error in hydraulic parameter estimates. This study thus demonstrates the capability of tomographic slug tests to define heterogeneous fields of K_h, K_v/K_h and S_s between wells.

Keywords: Hydraulic tomography, Slug tests, Heterogeneity, Anisotropy, Inverse modeling, Aquifer characterization

1 Introduction

Hydrogeological investigations, carried out either to understand groundwater flow or contaminant transport, require representative data on the heterogeneous spatial distribution of hydraulic properties in aquifers [*de Marsily et al.*, 2005]. The geometry and texture of geological materials in aquifers vary naturally through space as a result of complex geological processes, leading to heterogeneity in hydraulic properties that greatly influence groundwater flow and contaminant transport at different scales [*Koltermann and Gorelick*, 1996]. Indeed, the large-scale heterogeneity, particularly in hydraulic conductivity (K), controls the paths of groundwater and contaminants, while both large and small-scale variations control the magnitude of contaminant dispersion [e.g., *Sudicky*, 1986; *LeBlanc et al.*, 1991]. Hydraulic tomography has been identified by many investigators [e.g., *Neumann*, 1987; *Tosaka et al.*, 1993; *Bohling*, 1993; *Zhu and Yeh*, 2005 *Bohling et al.*, 2007; *Illman et al.*, 2008; *Brauchler et al.*, 2011; *Berg and Illman*, 2011; *Cardiff et al.*, 2012] as a promising alternative to more conventional hydraulic tests to image the heterogeneous spatial distribution of hydraulic properties in aquifers at field-scale [*Butler*, 2005].

Hydraulic tomography consists in the simultaneous analysis of inter-well hydraulic responses at various observation points to multiple hydraulic tests (pumping or slug tests). Considering the sensitivity and resolution analysis provided by *Vasco et al.* [1997], *Clemo et al.* [2003] and *Bohling* [2009] for vertical averaged and tomographic pumping tests as well as by *Paradis et al.* [2014] for tomographic slug tests, a tomographic analysis of drawdown or head data proved to hold the potential for imaging aquifer heterogeneities at a scale of relevance for groundwater flow and large-scale contaminant transport investigations. However, the groundwater flow governing equation describing aquifer hydraulics and the geometry of tomographic experiments are both major limitations to the resolution in hydraulic properties that can be achieved between wells. In fact, the high sensitivity correlation for parameters located between stressed and observation interval limits the horizontal resolution because of predominantly horizontal flow between vertical wells. This is exacerbated in real field conditions with the presence of noise in the head data. Although the vertical spatial resolution of an aquifer can be increased using smaller stressed intervals, not much can be done to increase the horizontal resolution using hydraulic responses alone. It follows that hydraulic tomography is limited to the estimation of medium-scale (few meters) averages of hydraulic properties, where the horizontal cell-size will depend mainly on the distance between wells. Such estimates can nevertheless provide the necessary level of information for many hydrogeological investigations that require images of the global architecture of an aquifer.

Previous field investigations using hydraulic tomography mainly focused on K estimates, assuming isotropy in K (see the literature review in *Cardiff and Barrash,* 2011). For tomographic experiments at sites where the assumption of isotropy does not apply at the scale of the experiment (e.g. length of the stressed interval > vertical variations in sediments grain-size), K anisotropy (ratio of vertical and horizontal K, K_v/K_h) has to be considered to ensure a meaningful interpretation of tomographic data. For instance, such conditions can occur in aquifers where fine layering in sediments may induce large anisotropy in K. Also, the knowledge of specific storage (S_s) may be useful to determine the hydraulic diffusivity and flow system response times through aquifer systems (aquifer and aquitards units) in order to assess impact of various hydraulic stresses on aquifers [*Heath*, 1983; *Bredehoeft*, 2011]. Moreover, there are few accurate and reliable in situ methods to estimate S_s [*Smith et al.*, 2013] even from pumping tests due to deleterious wellbore storage effects [*Kruseman and de Ridder*, 1990]. Considering the sensitivity and resolution analysis provided by *Paradis et al.* [2014], a transient analysis of tomographic slug test data would be a promising approach for obtaining such information on heterogeneities in K_h, K_v/K_h and S_s.

The purpose of this paper is thus to assess the quality of the information provided by a transient analysis of tomographic slug test data. Although the analysis presented by *Paradis et al.* [2014] shows that such an analysis has the potential to provide aquifer heterogeneity in K_h, K_v/K_h and S_s, there is a need to assess the performance of this approach using data acquired with the constraints of real field conditions. We address this need here by comparing the results of the inversion of a tomographic experiment performed in moderately heterogeneous and highly anisotropic littoral aquifer with estimates obtained using other hydraulic tests and by the simulation of various hydraulic tests not used in the inversion. Previous assessments of hydraulic tomography have been done on field data for tomographic pumping tests by *Bohling et al.* [2007], *Berg and Illman* [2011] and *Cardiff et al.* [2012] as well by *Brauchler et al.* [2011] for tomographic slug tests analyzed using a ray tracing approach. To our knowledge this paper presents the first assessment of a transient analysis of tomographic slug tests using field data, as well as the first tomographic analysis that considers K anisotropy.

This paper first presents a general description of the study site and the tomographic experimental design used in this study. The inverse method employed for the transient analysis of tomographic slug tests is then presented, with a description of the tomographic data set. Results of the tomography analysis are then compared with K_h and K_v estimates obtained from other field and laboratory tests, and validated using independent large-scale and local hydraulic tests.

2 Site description

This study was carried out at the Saint-Lambert-de-Lauzon Research Site (SLLRS) located 40 km south of Quebec City, Quebec, Canada (Figure IV.1a). The SLLRS encompasses a 12 km^2 sub-watershed area surrounding a decommissioned sanitary landfill emitting a landfill leachate plume (Figure IV.1b). The SLLRS was used to develop an integrated aquifer characterization approach [*Tremblay et al.*, 2013]. Considering the regional geological context, as well as GPR surveys and cone penetration tests (CPT), the hydrostratigraphy of the granular aquifer is found to have been controlled by longshore currents, tidal fluctuations and wave actions that occurred in a littoral depositional environment. The surficial geology of the study area is composed primarily of Quaternary sediments that were deposited and reworked during the retreat of the Champlain Sea [*Bolduc*, 2003]. As depicted in Figure IV.1b, many beach ridges, generally composed of sand, were formed by wave energy of the Champlain Sea [*Lamarche and Tremblay*, 2012]. Between the ridges, areas of lower energy were present and finer sediments were deposited.

The 100 MHz GPR section and the sediment profiles in Figure IV.2 depict the general stratification of sediments found at the SLLRS (see section and profile locations in Figure IV.1b). Sediment profiles are obtained from cone penetrometer tests (CPT) where mechanical properties (tip and sleeve stresses) of sediments are converted into sediment textures based on the CPT classification chart proposed by *Fellenius and Eslami* [2000]. It should be noted that converted sediment profiles are used as proxy of the real sediment textures and they are thus more useful for relative comparison over the study site than to estimate absolute sediment classes. Several observations can be made from Figure IV.2. First, the systematic stratification extending several hundred of meters shown by continuous GPR reflections dipping a few degrees (between 3° to 5°) towards the sea, which is a characteristic feature of a littoral depositional environment [*Van Overmeeren*, 1998]. Moreover, a mound in GPR reflections located at the distance of 500 m on the GPR profile of Figure IV.2 suggests a sand barrier that separates the GPR section in two distinct zones. The zone in front of the beach ridge (left, toward A') and the barrier itself is mostly composed of sand (see sediment profiles CPT113 and CPT26) that was deposited as a result of current reduction likely due to a change in bedrock topography. The lower reflection in Figure IV.2 is associated with the contact between the littoral sediments and a till layer overlying bedrock. Small hyperbolas below this contact are typical GPR characteristics related to blocks found in till. Beyond the barrier (right, toward A), sediments are finer and show alternating sand and silt layers, as suggested by profiles P10 and P17. GPR reflections within this region are also diffuse at a few spots as a result of more

electrically conductive sediments, such clayey silt, which generally limits GPR signal penetration. The water table is approximately 1 m below ground surface, as indicated by water level in wells, but cannot be identified from the GPR profile.

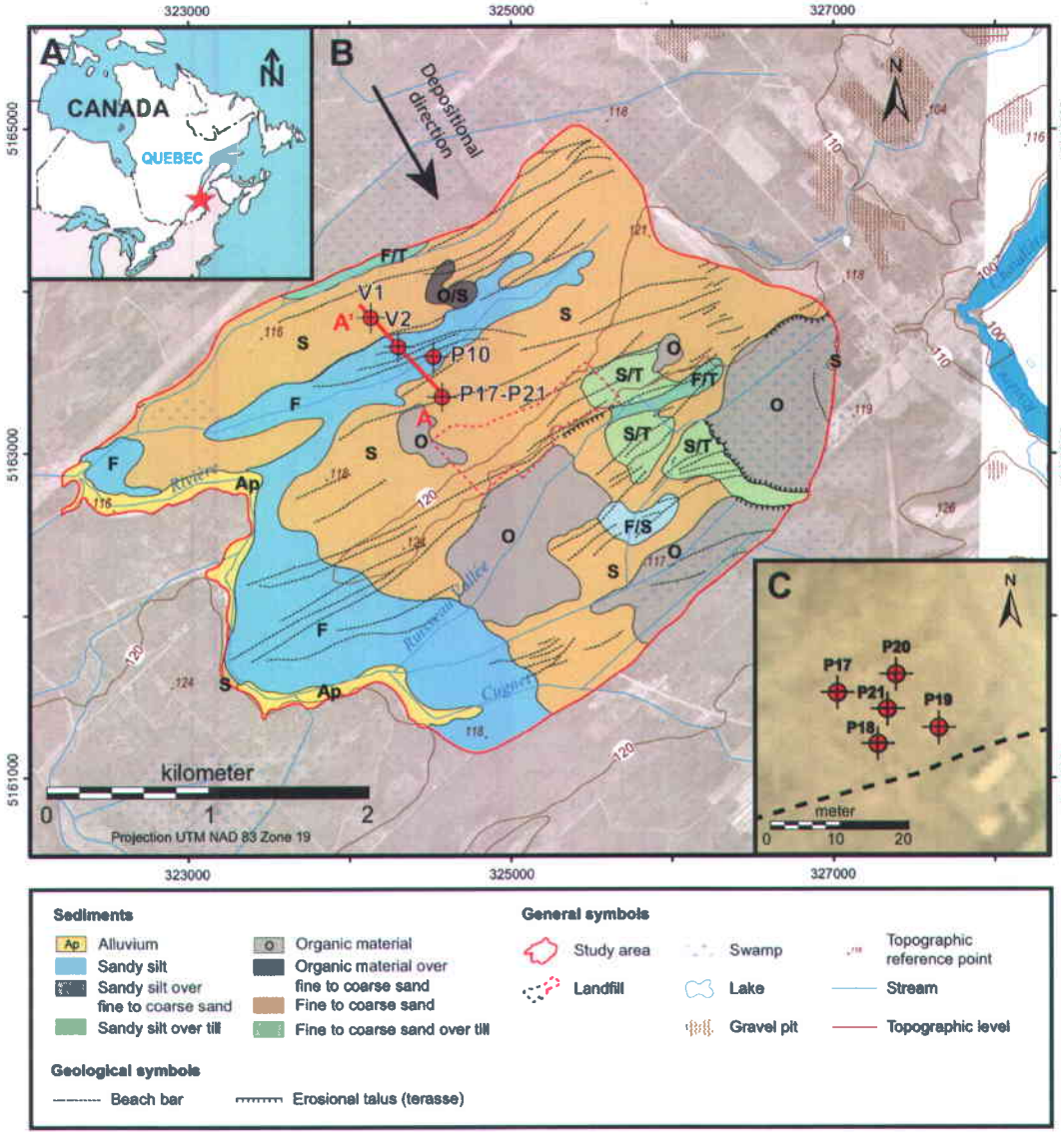


Figure IV.1 General location of the study site (a), with the Quaternary sediments map for the sub-watershed enclosing the former sanitary landfill (b), and the location of the wells used for the tomographic experiment (c). The Quaternary map is modified from Lamarche and Tremblay [2012].

The depositional conditions of sediments (Figure IV.2) was also based on tip stress data measured with CPT. Tip stress is a measure of sediments compaction that is associated with grain-size and sorting [*Lunne et al.*, 1997]. Tip stress generally increases with grain-size and poorly-sorted sediments. Relative

variations in tip stress can thus be used to infer depositional conditions. For profiles located within the region beyond the beach ridge, four layers that may be associated with the same stage of deposition are recognized. Interestingly, K_h values measured from multilevel slug tests at 15 cm intervals show similar profiles for wells P10 and P17 that are more than 200 m away from each other, as illustrated in Figure IV.2. More details on multilevel slug tests procedure are provided by *Paradis et al.* [2011].

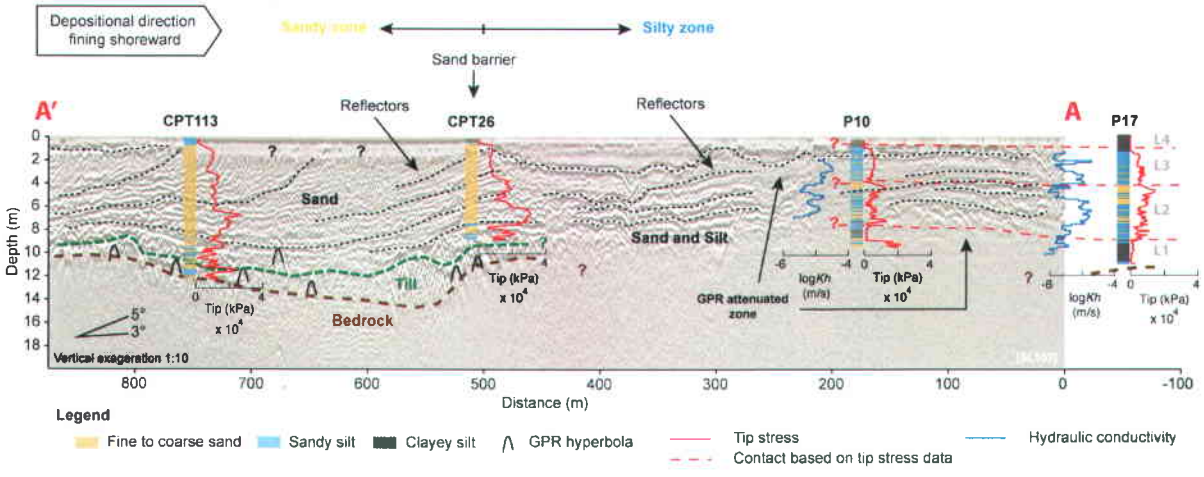


Figure IV.2 GPR section aligned parallel with the general littoral depositional direction along with profiles of sediments texture derived from cone penetrometer test data (tip and sleeve stresses) and tip stress data for wells V1, V2, P10 and P17. Horizontal hydraulic conductivity (*K*) obtained from multilevel slug tests is also presented for wells in which these measurements were made. GPR section and well locations are shown in Figure IV.1b.

Figure IV.3 presents the sequence of sediments derived from CPT data and tip stress profiles along wells used for the tomographic experiment (see Figure IV.1c for well locations). The tomographic plot is located 50 m away from the end of the GPR section and the sediments architecture is thought to be an extension of the region beyond the sand barrier as already depicted in Figure IV.2. Although there are some differences in tip stress data between each well, the same general layering observed in Figure IV.2 can be recognized. The major difference is in the sand proportion derived from CPT data that suggests a lateral increase in sediments grain-size. The large variations in tip stress over small vertical intervals also suggest fine variations in grain-size (layering) that may induce *K* anisotropy at larger scale.

The comparison of K_v measured on sediment samples using permeameter tests with K_h measured in the field with multilevel slug tests provide evidence of *K* anisotropy at SLLRS, as previously suggested by tip stress variations. To obtain K_v by laboratory testing, 8 core samples were collected over the entire aquifer thickness close to well P17. A piston-rod operated sampler driven by a direct-push rig allowed the

recovery of samples inside a PETG liner (38 mm diameter and 1.52 m long). A falling head laboratory permeameter was designed to directly accommodate the PETG liner containing sediment subsamples of 15 cm long to minimize disturbance of the original sedimentary structures and hence K_v estimates. More details on sampling and testing procedures are provided by *Paradis and Lefebvre* [2013]. Results of the permeameter tests presented in Figure IV.3 show that K_v ranges from 2.7×10^{-8} to 1.0×10^{-5} m/s and the comparison with K_h values available at the same elevation suggests that the K_v/K_h ratio may vary from 0.0025 to 0.83. This K anisotropy for a sample size of 15 cm is coherent with small-scale variations in CPT data and the fine layering in grain-size observed in sediment samples.

In summary, the layered nature of sediments found at the SLLRS, as shown by the GPR section, the CPT profiles and the multilevel slug tests, all suggest that the sedimentary architecture is relatively continuous laterally across the study site. Although large-scale layers can be recognized over the study area, the fine layering in sediments grain-size induces strong K anisotropy, as revealed by the comparison of K_v and K_h values measured on the same intervals with permeameter and multilevel slug tests, respectively. Thus to ensure a meaningful understanding of flow and transport characteristics at the SLLRS, a hydraulic testing approach has to capture vertical and horizontal variations in hydraulic properties, including K anisotropy.

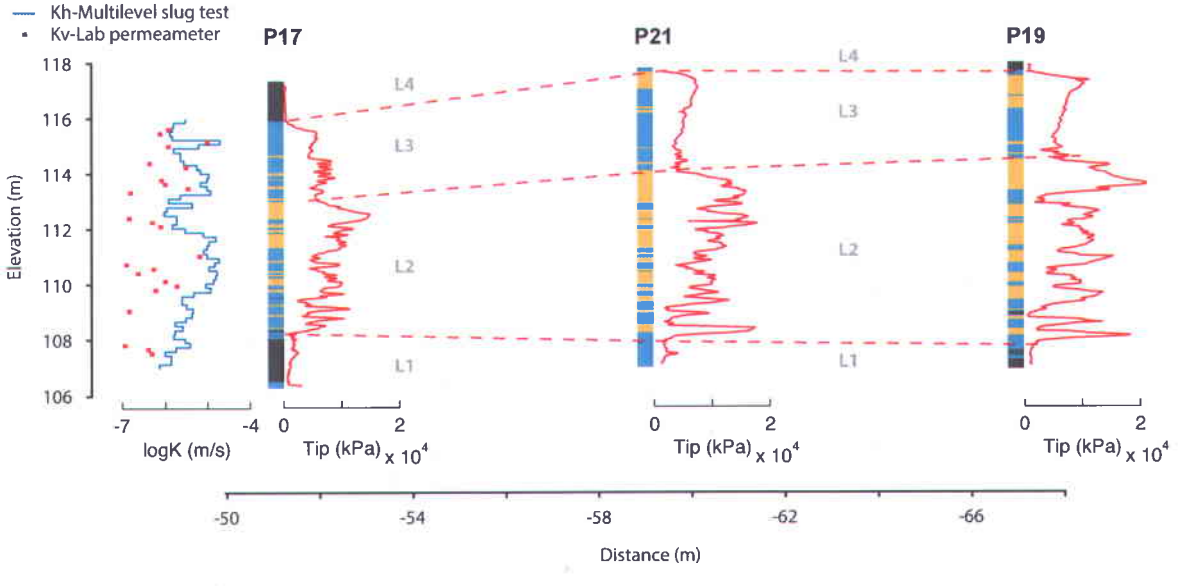


Figure IV.3 Sediment layers in the aquifer based on tip stress data (numbered from L1 to L4) along with profiles of sediment texture derived from cone penetrometer test data (tip and sleeve stresses) for wells P17, P21 and P19 at the tomography site (legend in Figure IV.2). Vertical (K_v) and horizontal (K_h) hydraulic conductivities obtained from laboratory permeameter and field multilevel slug tests, respectively, are also provided for well P17. Well locations are shown in Figure IV.1c.

3 Tomographic experiment design

3.1 Well installations

The SLLRS was instrumented with a total of 5 wells arranged in a diamond pattern (one at each corner and one in the center) 15.9 m in the long axis and 10.8 m in the short axis (Figure IV.1c). Fully screened direct-push wells were installed over the entire aquifer thickness with a direct-push rig (Geotech 605D) following the protected screen standard technique [*ASTM*, 2004]. These wells were installed at the same locations as CPT soundings and screens were placed in direct contact with undisturbed sediments. Sand-pack was not used with well installation to reduce the effects of hydraulic short-circuits for observation intervals located along the stressed well [*Paradis and Lefebvre*, 2013]. The fully screened wells have a diameter of 5.1 cm with screen length ranging from 7.6 to 9.2 m. Wells were thoroughly developed to insure good hydraulic contact between the screens and sediments [*Paradis et al.*, 2011]. This well installation is a key step in the success of hydraulic tomography as it allows flexibility in the tomographic design with packers. For instance, stressed and observation intervals can be isolated with a packer spacing of any length into any well.

3.2 Data collection

In this paper, we only report on the tomographic experiment t21r17 carried out between wells P21 and P17 as stressed and observation well, respectively (Figures IV.1c and IV.3). For this experiment, a network of pressure transducers was installed within screen intervals isolated by packers to allow data collection during slug tests (Figure IV.4a). Inflatable packers were built of 25.4-mm ID PVC tubing, and threads on the tubing allowed the use of variable screen length between packers. A line connected to a nitrogen tank at the surface was attached to the packers to inflate them to the desired pressure; pressure regulators were used to keep the pressure constant for all tests. All observation intervals were isolated into the observation well by daisy chaining screens and packers. Observation intervals of 30 cm were used, to accommodate pressure transducers, separated by packers of 61 cm in length. All observation intervals were straddled to reduce wellbore storage effects.

Within the stressed well, a three-packer assembly was used to isolate the stressed interval and an observation interval underneath, as proposed by *Paradis and Lefebvre* [2013] for vertical interference slug tests. The stressed interval isolated by two packers at the top of the assembly was connected by a

25.4-mm ID PVC riser pipe to the surface (Figure IV.4b), within which water levels were monitored during each test to determine the end of the monitoring period. A stressed interval of 61 cm was used and the underneath 30-cm observation interval was separated by a packer of 61 cm in length. For each test of tomographic experiment t21r17, six intervals were thus monitored simultaneously: the stressed and the observation intervals in the stressed well P21, and four observation intervals in the adjacent well P17.

To reduce hydraulic head drift and noise caused by packer inflation and deflation, packers used to isolate intervals in the observation well were inflated at the beginning of the day and were left in place for entire duration of the tests. The download and transfer of all slug test data was only done at the end of the day. The stressed and observation intervals were monitored with 0-10 m pressure transducers (model LT 3001 Levelogger Gold, Solinst Canada Ltd.). To obtain good time-resolution recordings of hydraulic responses and to estimate noise level associated with the transducers the sampling rate was set to 1 sec.

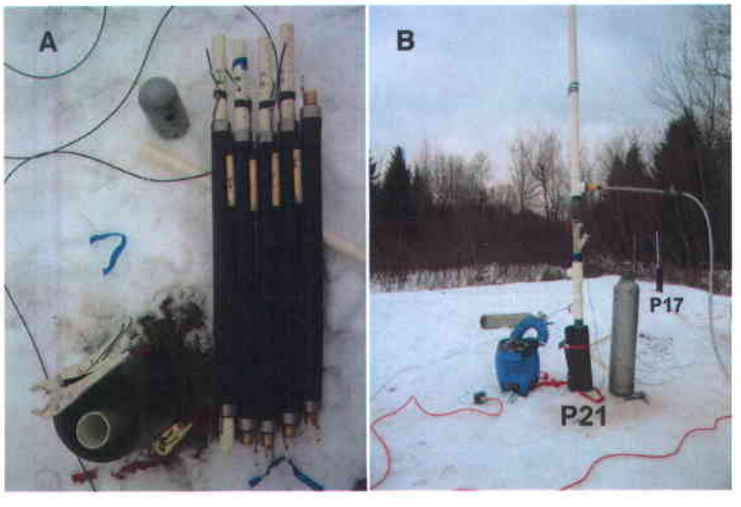


Figure IV.4 Slug test instrumentation for tomographic experiment: (a) packers, screens and pressure transducers used to monitor slug test responses; (b) water column (white PVC) used to initiate slug tests.

3.3 Tomographic slug tests description

Only for tomographic experiment t21r17, a total of twelve slug tests were carried out and monitored in the stressed well P21, while simultaneously recording hydraulic responses within four observation intervals straddled in well P17 and within one observation interval in well P21, as depicted in Figures IV.5a-b. Slug tests were done systematically at successive 61-cm intervals to stress all parts of the aquifer along well P21 and to better constrain the inversion. Slug tests were done using a water column to induce

an initial rise of the water level. For this purpose, a wellhead assembly was attached to the top of the riser pipe of the stressed interval (Figure IV.4b). This device contained a bottom-hole valve, which initially allows the filling of the device with water and the fast introduction of the water column into the stressed interval. The device also contained a watertight adapter that allowed a transducer cable to pass for real-time monitoring of the stressed interval. The initial head induced for all tests with this device ranged from 4.06 to 4.49 m (Table IV.1).

Table IV.1 Summary of slug tests carried out at SLLRS for tomographic experiment t21r17. Prefix "t" is for "transmitter" (stressed interval) whereas "r" is for "receiver" (observation interval).

Stressed interval	Initial head in stressed interval (m)	Maximum observed amplitude in observation well P17 (m)	Interval of maximum amplitude in observation well P17	Maximum observed amplitude in observation interval in well P21 (m)
t21-236	4.24	0.026	r17-214	0.016
t21-297	4.33	0.034	r17-305	0.010
t21-358	4.46	0.059	r17-396	0.112
t21-419	4.28	0.029	r17-488	0.191
t21-480	4.45	0.032	r17-488	0.065
t21-541a	4.46	0.029	r17-488	0.026
t21-541b	4.48	0.024	r17-580	-
t21-602	4.21	0.055	r17-671	0.087
t21-663	4.06	0.071	r17-671	0.074
t21-724	4.21	0.066	r17-762	0.226
t21-785	4.37	0.057	r17-762	-
t21-845	4.49	0.031	r17-854	-

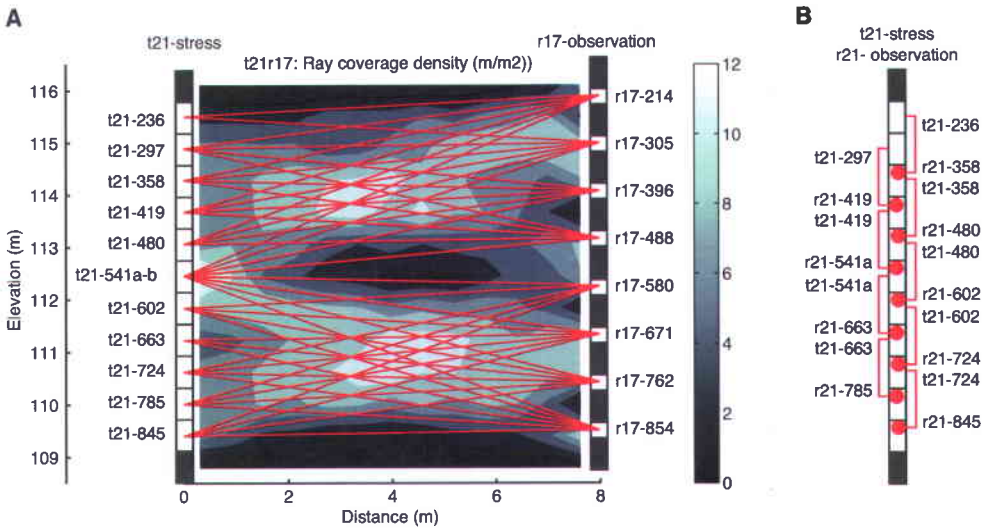


Figure IV.5 Coverage density (length of rays per area; m/m^2) between stressed and observation wells for tomographic experiment t21r17: (a) between wells P21 and P17; and (b) along well P21. The rays along well P21 in (b) are not used in the evaluation of coverage density in (a). White spaces indicate the lengths of open intervals between packers (stressed or observation interval). In the nomenclature used for stressed and observation interval numbers, "t" stands for "transmitter" and "r" for "receiver". The first two digits after the letter indicate the well number whereas the last three digits are top of the screen depth, in cm, from the top of the tubing.

Slug tests were recorded until the peak amplitude in the observation intervals had passed. This was done mainly to reduce test duration, and it is justified by the fact that data for times up to the first inflection point after peak amplitude contain sufficient information to determine hydraulic properties within and outside the region of investigation [*Paradis et al.*, 2014]. The slug tests duration ranged from 5 to 16 minutes, thus allowing multiple tests to be done in a single day.

The spacing between observation intervals was initially based on the thickness of the aquifer layers depicted by CPT data and multilevel slug tests at well P17 (Figure IV.3). The resulting coverage density depicted in Figure IV.5b is generally uniform, except for a zone around elevation 112.5 m where a defective packer at the end of the daisy chained assembly precluded the monitoring of heads in a fifth interval. Note that the coverage density in Figure IV.5b assumes straight ray paths and it is only presented to illustrates the configuration of stressed and observation intervals used for tomographic experiment t21r17. The angles between stressed and observation intervals (using screen centers as reference) ranged from 0° to 23°.

From a practical viewpoint, a hydraulic head perturbation can be generated from a slug test for a large range of aquifer K values in an aquifer, which does not limit tomographic testing to intervals with

moderate to high K as can be expected with the use of pumping. A head perturbation induced by a slug test is only delayed in time following K variations, so that K values do not affect the amplitude of the perturbation, as opposed to pumping tests that can desaturate small stressed intervals in low-K formations. However, one of the main challenges for a field implementation was to obtain a good signal-to-noise ratio. As summarized in Table IV.1, heads measured in observation intervals at well P17 were quite small and generally less than 8 cm. The amplitude for the observation interval placed under the stressed interval in well P21 were generally higher, but early head measurements may be subject to dynamic hydraulic short-circuits that should be recognized and taken into account during the analysis [*Paradis and Lefebvre*, 2013]. The high K anisotropy, that strongly attenuated the head perturbation in the vertical direction, and the moderate to high values of S_s both contributed to the small head variations recorded in observation intervals. This emphasizes the importance of designing a slug test device able to induce a large stress on the aquifer (Table IV.1) combined with the need to use very sensitive pressure transducers.

3.4 Inter-well interference slug test description

In complement to the previous tomographic experiment, an interference slug test between wells P21 and P17 using the entire screened section of each well was carried out. This test provides the average dynamic behavior of the aquifer under stress and can be used to constrain or to verify the results of the tomographic inversion. For this test, wells P21 and P17 were used as stressed and observation wells, respectively. The same device was used to initiate the test and the observation well was straddled at the top of the water table with a packer to reduce wellbore storage effects. The initial head in the stressed well was 2.99 m and the maximum amplitude recorded in the observation well was 0.023 m.

4 Data analysis methodology

In this study, we invert tomographic experiment t21r17 slug test data with a radial flow model. Observation interval responses recorded in the observation well, together with stressed interval responses, are used in the numerical inversion to estimate the spatial distribution of hydraulic properties between wells P21 and P17. The accuracy of the spatial distribution of K_h, K_v/K_h and S_s generated by the inversion of multiple slug tests is evaluated in the next section with observation interval responses recorded below the stressed interval in well P21 (vertical interference slug tests) and the inter-well interference slug test

conducted as part of the present tomographic effort. In this section, we first describe the inverse analysis adopted in this study.

4.1 Radial flow model

The program lr2dinv (linear and radial 2D inversion) developed by *Bohling and Butler* [2001] is used in this study to inverse the data from the tomographic slug tests experiment t21r17. lr2dinv is a two-dimensional radial-vertical finite difference flow model coupled with a Marquardt-Levenberg algorithm that can be used for simultaneous inverse analysis of head responses. Under conditions of radial symmetry, meaning the absence of angular variations in both the hydraulic properties and the boundary conditions, the flow to a partially penetrating stressed interval in response to an instantaneous change in water level in a confined aquifer of infinite areal extent is described by:

$$\frac{1}{r}\frac{\partial}{\partial r}\left(rK_r\frac{\partial h}{\partial r}\right)+\frac{\partial}{\partial z}\left(K_z\frac{\partial h}{\partial z}\right)=S_s\frac{\partial h}{\partial t} \tag{1}$$

where h=head [L], S_s=specific storage [1/L], K_r=hydraulic conductivity in the radial (or horizontal-K_h) direction [L/T], K_z= hydraulic conductivity in the vertical (or K_v) direction [L/T], t=time[T], r=radial coordinate [L], and z=vertical coordinate (positive upward from the base of the aquifer, which is used as a reference) [L].

The inner boundary of the model is at the stressed well radius r_w [L] and initial conditions given by:

$$h(r,z,0)=0 \quad r_w<r<\infty \quad 0<z<b \tag{2}$$

$$H(0)=H_0 \tag{3}$$

where b [L] is aquifer thickness, H is the level of water in the well [L], and H_0 is the static water column height above the stressed interval, equal to the level of water in the well at t=0, [L].

The outer boundary conditions are the following:

$$h(r=\infty,z,t)=0 \quad t>0 \quad 0\le z\le b \tag{4}$$

$$\frac{\partial h(r,0,t)}{\partial z}=\frac{\partial h(r,b,t)}{\partial z}=0 \quad t>0 \quad r_w\le r\le\infty \tag{5}$$

and the inner boundary conditions at the stressed interval are given by:

$$\frac{1}{L}\int_{d}^{d-L} h(r_w,z,t)dz=H(t) \qquad t>0 \tag{6}$$

$$2\pi r_w K_r\frac{\partial h(r_w,z,t)}{\partial r}=\begin{cases} 0 & z<d, z>d+L \quad t>0 \\ \dfrac{\pi r_c^2}{L}\dfrac{dH(t)}{dt} & d>z<d+L \quad t>0 \end{cases} \tag{7}$$

where d is the distance from the top of the aquifer to the top of the stressed interval [L]; L is stressed screen length [L]; r_c is the radius of well casing [L]. The imposed boundary conditions for this study were no flow for the bottom boundary and constant head for the top boundary (unconfined conditions), as illustrated in Figure IV.6a.

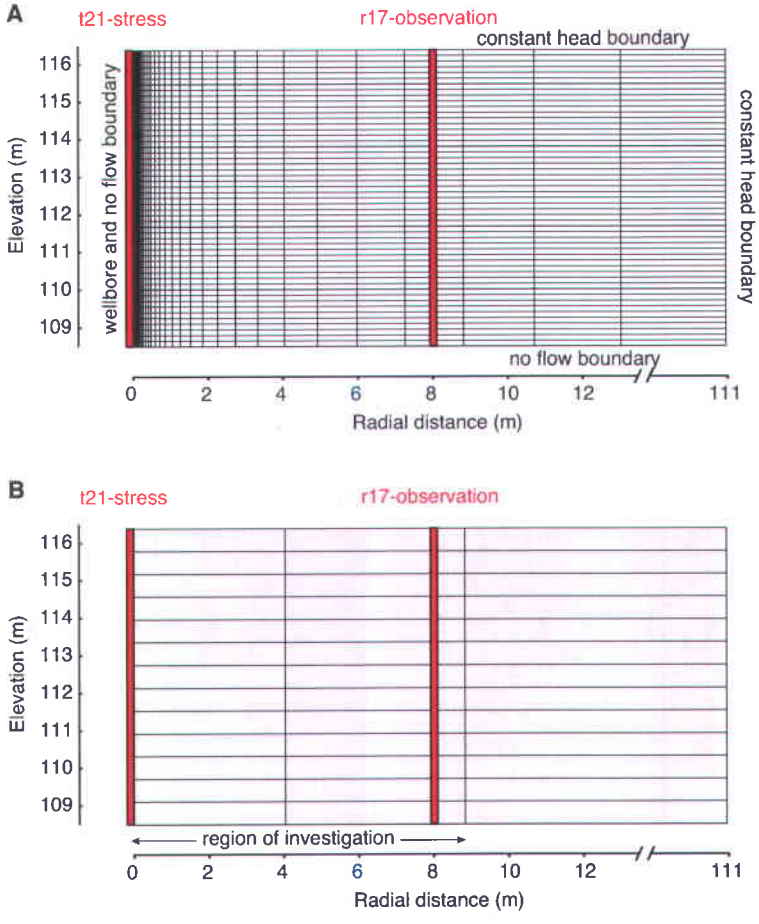


Figure IV.6 Simulation (a) and parameter (b) grids for the inverse analysis of tomographic experiment t21r17.

The program lr2dinv also allows incorporating the wellbore into the model domain, thus representing the effects of wellbore storage and the placement of packers in the stressed well. Wellbore storage is defined as the relative volume change ΔV per unit head change Δh in the wellbore ($W=\Delta V/(V\cdot\Delta h)$). Wellbore processes are approximated using Darcy's Law, with open sections of the wellbore represented as high-K regions ($K_h=1\times10^5$ m/s and $K_v=1\times10^8$ m/s) and packers as low-K regions ($K_h=K_v=1\times10^{-12}$ m/s). To simulate the wellbore, one column of cells is used to represent the region inside the wellbore, from the wellbore radius, r_w, to the inner radius of the model grid, $r_{min}=r_w\exp(\Delta r')$, where $\Delta r'$ is the constant spacing in the transformed radial direction. Most cells in the wellbore should be assigned a S_s value of 0, representing the fact that water is essentially incompressible. However, the top cell in the wellbore (or the top in any screened interval) should have a S_s value of 1 to indicates that the hydraulic stress would translate directly into the wellbore and the program will convert to the value

$$S_s = \frac{1}{\Delta z}\left(\frac{r_w^2}{r_w^2 - r_{\min}^2}\right) = \frac{1}{\Delta z}\left(\frac{1}{1-\exp(-2\Delta r')}\right) \tag{8}$$

to account for the vertical thickness of the cell and compensate for the exclusion of the inner portion of the wellbore (between r=0 and r=r_{min}) from the model domain. To simulate each slug test, the initial head in the top cell was set to the initial displacement used to initiate the test, H_0, with all other initial heads in the model set to zero assuming therein steady-state conditions prior to each test. The storage properties of the wellbore are set to represent wellbore storage effects, which is a key factor to obtain a meaningful interpretation of inversion results with tomographic slug tests. Wellbore storage affects the hydraulic response in an observation interval in two ways. First, for a given distance between the stressed and observation interval, wellbore storage causes a delay in the time at which the pressure change is observed. Second, the response amplitude grows with the wellbore storage if the initial head change remains the same [*Prats and Scott*, 1975; *Spane*, 1996; *Brauchler et al.*, 2007, 2011; *Paradis and Lefebvre*, 2013]. Wellbore storage effects (time delay and amplitude change) are also relative to S_s values and then for an aquifer where S_s is expected to vary, simulation of wellbore storage as provided by lr2dinv will greatly improve the accuracy of inverted tomograms. The lr2dinv model assumes that wellbore storage effects can be neglected in the observation interval. To obtain test data with minimal wellbore storage effects, the observation intervals were straddled with packers in the field.

The lr2dinv model uses a logarithmic transform of the radial coordinate, r'=ln(r/r_w), to transform the radial flow problem into an equivalent Cartesian problem in (r', z) space [*Butler and McElwee*, 1995; *Bohling and Butler*, 2001]. This study used a simulation grid with 43 cells of dimension $\Delta r'$=0.19492 along the transformed radial axis and 52 cells of dimension Δz=0.1524 m along the vertical axis (Figure IV.6a). In the physical space, the radial location of the grid node with radial index i is $r_w$exp((i-0.5) $\Delta r'$) and the location of the outer face of the corresponding cell is $r_w$exp($i\Delta r'$). The scale of the logarithmic transform was adjusted to fit the radial location of the grid nodes to the location of the wells. The exponentially telescoping grid in the radial direction provides an easy means to place the outer boundary far from the stressed well. The discretization used places the zero-head outer boundary of the model about 111 m from the stressed well, so that this boundary has negligible effects on simulated slug test heads at the stressed and the observation wells.

4.2 Parameter grid

To handle the issue of non-uniqueness when solving the inverse problem, we regularized the numerical inversion by the so-called discretization approach [*Aster et al.*, 2005]. With the lr2dinv model, a separate zonation of the simulation grid can be specified for each hydraulic property, cells of the simulation grid are merged to define a parameter grid, shown in Figure IV.6b, that is expected to correspond to the effective spatial resolution of the tomographic experiment t21r17. The cell size for the parameter grid is based on the sensitivity and resolution analysis provided by *Paradis et al.* [2014] for a synthetic tomographic experiment based on similar field conditions as the ones used in this study. In that study, it was found that the region of investigation encompassed by the stressed and observation wells, can be adequately resolved for K_h, K_v/K_h and S_s with an effective cell width that corresponds to half of the region of investigation when using only one stressed well and one observation well in the analysis. The vertical resolution of cells was set to 61 cm, which corresponds to the length of the stressed interval used for slug tests. The parameter grid used for the inversion is then composed of 39 cells with 13 cells in the vertical and three cells along the horizontal direction: two for the region of investigation and one for the region beyond the observation well. As distinct zonation for regions within and outside the region of investigation was done to take into account the influence of the two regions. Following the analysis provided by *Paradis et al.* [2014], K_h and S_s values for cells just outside the observation well can potentially be resolved from a transient analysis of tomographic slug tests, whereas K_v/K_h values for the entire region beyond the observation well have practically no influence on heads and thus cannot be resolved. Thus, for this study, we concentrate our analysis only for the region encompassed by the stressed and observation wells because of the larger uncertainty in estimates for parameters outside the region of investigation. This discretization is thought to be sufficient to image the vertical and lateral heterogeneity in hydraulic properties that is expected from available information. Using finer parameter grid cells would of course lead to over-parametrization because there is insufficient information to constrain the inversion at a smaller grid scale. The same parameter grid was used for K_h, K_v/K_h and S_s and no other form of regularization was applied to the inverse analysis.

4.3 Inverted data set

All the 12 slug tests of tomographic experiment t21r17 were analyzed simultaneously to produce heterogeneous distributions of K_h, K_v/K_h and S_s in the panel between wells P21 and P17. For each test, the stressed and four observation interval responses were simultaneously inverted, except for test t21-236 where two observation responses were removed from the dataset due to defective connections that created

too much noise (Figure IV.5a). A total of 12 stressed interval responses and 46 observation interval responses were then available for this experiment and analyzed. Using stressed interval responses in the inverse analysis is important in order to increase parameter resolutions near the stressed intervals and also to reduce uncertainties in property estimates from the shape of the observation interval responses due to the non-applicability of the principle of reciprocity for slug tests, as shown by *Paradis et al.* [2014].

Original measurements were subsampled using a 10-sec interval to reduce computational burden while remaining well below the Nyquist sampling criterion. It was found to be crucial to keep the shape of the head response in the analysis because maximum sensitivity to each hydraulic property occurs at different time periods during the test [*Paradis et al.*, 2014]. The total number of data points used for this inversion was 3408. Observed head data shown in Figure IV.7 are raw responses, as measured in the field, without any kind of filtering.

The noise level associated with head measurements was estimated from a residual analysis of head records. For this analysis, sections of head records that are nearly linear were extracted and the statistics on the deviation (residual) of the recorded heads with respect to a linear fitted model on the observations were computed. The value of the standard-deviation of the error is $1x10^{-4}$ m, which is well below values of recorded heads in observation intervals (Table IV.1 and Figure IV.7). Weighted inversion that accounts for the noise level was not used because all head measurements have similar noise level regardless of the absolute amplitude in head responses and it is thus similar to an unweighted inversion. Instead, a weighted regression was done to account for the scale discrepancy between stressed and observation head responses. The scaling factors of the Levenberg-Marquardt algorithm associated with each stressed response and their corresponding observation interval responses were then adjusted according to the relative amplitude of the stressed and observation interval responses.

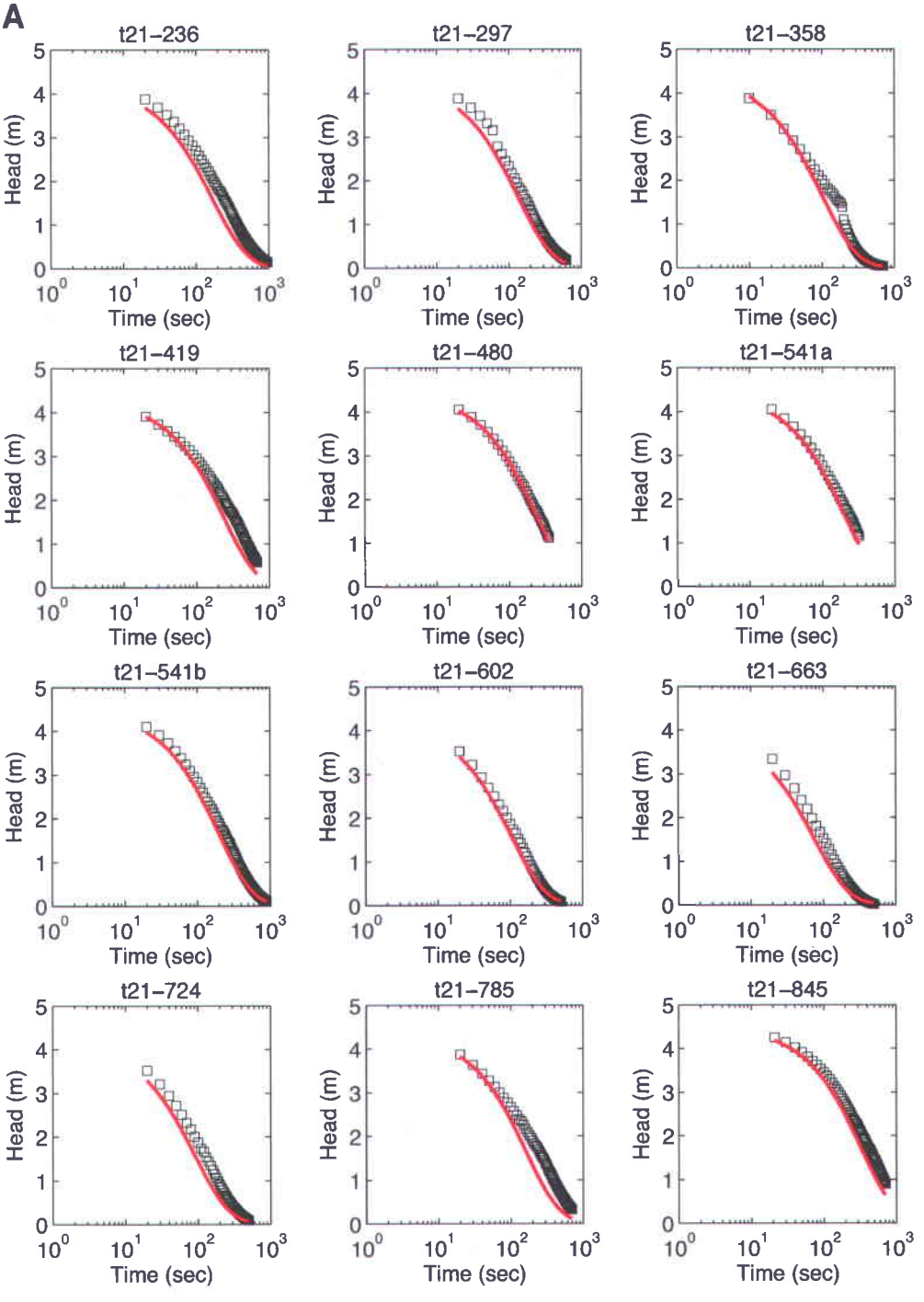


Figure IV.7 Observed (symbols) and simulated (red curves) head responses for tomographic experiment t21r17. Graphs (a) show each of the 12 stressed interval responses (black squares) along with (b) the responses from observation intervals (black circles) following the slug test in the indicated stressed interval.

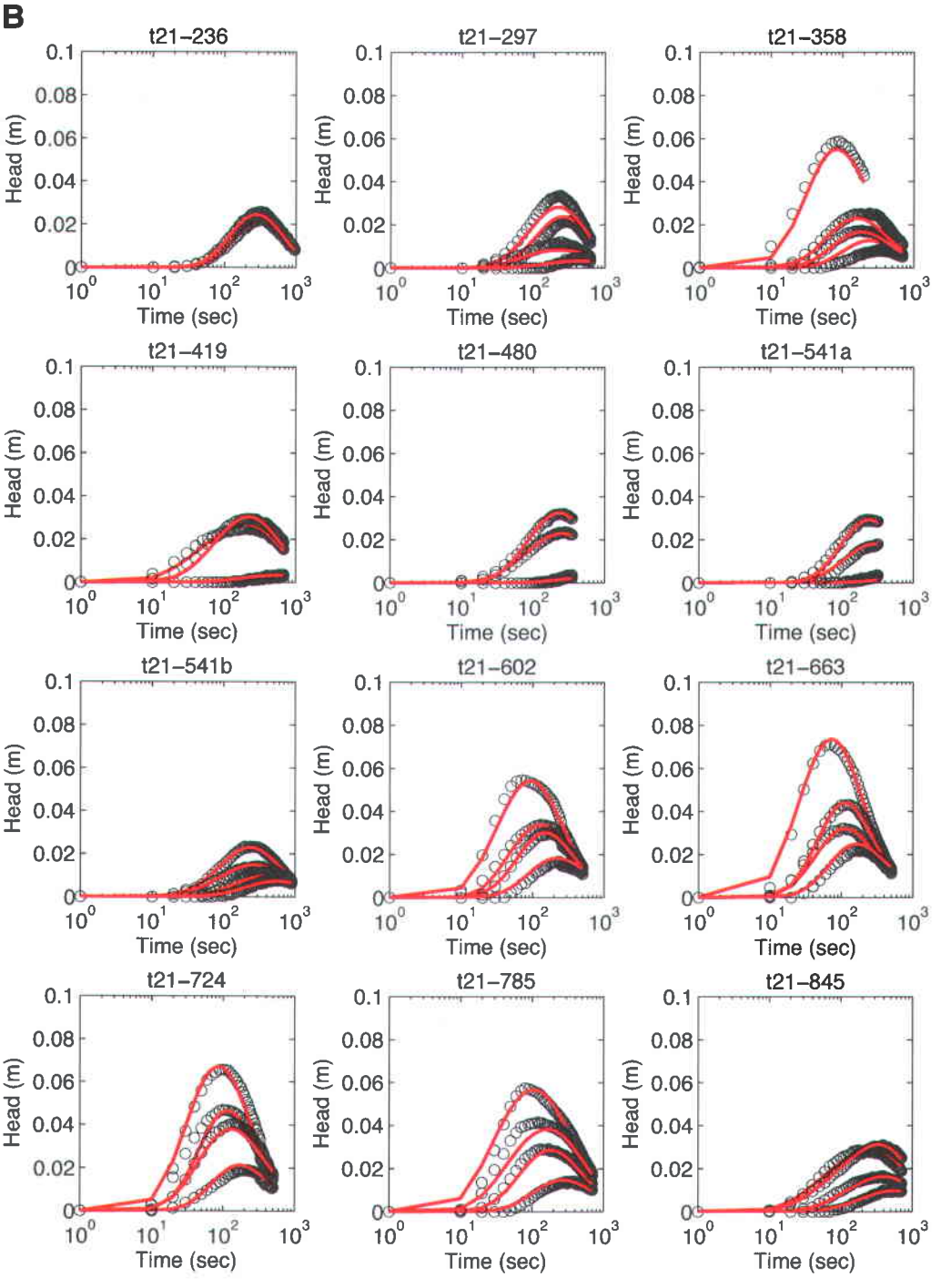


Figure IV.7 Continued

5 Transient hydraulic tomography: analysis of results

The inversion of the head data was run for all slug tests listed in Table IV.1 and the iterative estimation process terminated either when the relative difference between two consecutive unknown parameter estimates (**p**) was less than 1×10^{-3} or the relative difference between two successive values of the following chi-squared error was lower than 1×10^{-8}:

$$\chi^2 = \sum_{i=1}^{n}\left(\frac{h_i - f_i(\mathbf{p})}{\sigma_i}\right)^2 \tag{9}$$

where h_i is the i^{th} head observation and $f_i(\mathbf{p})$ is the corresponding simulated head, based on the current vector of unknown parameters, **p**, and s_i is the scaling factor associated with the i^{th} observation [*Bohling and Butler*, 2001]. All test data were included in the inversion simultaneously and the inversion started with uniform K_h, K_v/K_h and S_s fields (K_h=1×10^{-5} m/s, K_v/K_h=0.1 and S_s=1×10^{-4} m^{-1}) with no constraint imposed on the inversion. The duration of the hydraulic test considered in the simulation was 1000 sec to give enough time to reproduce the peak amplitudes recorded in all observation intervals. A linear biconjugate gradient scheme was used for solving the system of equations in the finite difference model. In the presence of strong vertical to horizontal anisotropy, this preconditionner is known to speed up convergence. The inversion of the tomographic experiment t21r17 was done on a OSX (64-bit) operating system running on an Intel Quad Core i7 processor running at 2.2 GHz with 16 GB of RAM. The total computation time for data inversion was 1235 min. Simulated heads for each test are compared to observed values in Figure IV.7 and validation statistics are presented in Table IV.2.

Analysis of simulated and observed heads for each slug test in Figures IV.7a-b reveals that the inversion captures fairly well the general behavior of head responses obtained during the tests. Indeed, first arrival times and maximum amplitudes of simulated observation interval responses mostly coincide with field observations (Figure IV.7b). Only a small delay in time is observed between simulated and observed stressed interval responses, as depicted by simulated responses slightly ahead of observed curves for few tests (Figure IV.7a), which may suggest a small overestimation of K_h values near well P21. The slope of the linear regression between observed and simulated heads (m) are 0.977 and 0.960 for all stressed and observation intervals (Table IV.2), respectively, which indicate that simulated heads overall slightly

underestimate observed heads. Nevertheless, coefficient of correlation (r) and m values close to unity between simulated and observed heads for almost all tests indicate that there is no systematic deviation in head estimates. Moreover, root-mean-square error (*RMSE*) values in percentage of the maximum head recorded in stressed or observation intervals indicate small errors for both stressed and observation intervals with *RMSE* values ranging from 1.8% to 9.9% for all tests.

Table IV.2 Validation statistics for observed and simulated heads for each slug test used in the inversion of tomographic experiments t21r17: "*r*" is the coefficient of correlation, "*m*" is the slope of the linear regression and "*RMSE*" is the root-mean-square error. Statistics for stressed and observation intervals are presented separately. RMS in percentage is calculated relative to the maximum head measured in stressed or observation interval according to Table IV.1.

Test	Stressed interval			Observation intervals		
	r	*m*	*RMSE* (m / %)	*r*	*m*	*RMSE* (m / %)
t21-236	0.991	0.936	0.318 / 7.5	0.942	0.919	0.0023 / 8.8
t21-297	0.997	0.944	0.207 / 4.8	0.973	0.913	0.0023 / 6.8
t21-358	0.986	0.940	0.219 / 4.9	0.975	0.854	0.0031 / 5.3
t21-419	0.994	1.100	0.365 / 8.5	0.994	0.979	0.0012 / 4.1
t21-480	0.999	1.037	0.078 / 1.8	0.998	1.017	0.0007 / 2.2
t21-541a	0.999	1.064	0.173 / 3.9	0.994	1.013	0.0012 / 4.1
t21-541b	0.997	0.964	0.160 / 3.6	0.987	0.966	0.0008 / 3.3
t21-602	0.996	0.954	0.132 / 3.1	0.992	0.931	0.0017 / 3.1
t21-663	0.992	0.903	0.210 / 5.2	0.994	0.994	0.0015 / 2.1
t21-724	0.991	0.933	0.269 / 6.4	0.987	0.922	0.0026 / 3.9
t21-785	0.987	1.041	0.432 / 9.9	0.988	0.950	0.0022 / 3.9
t21-845	0.998	1.080	0.301 / 6.7	0.995	1.009	0.0010 / 3.2
All tests	0.989	0.977	0.270 / 6.2	0.989	0.960	0.0019 / 4.4

The final K_h, K_v/K_h and S_s tomograms obtained from data inversion of tomographic experiment t21r17 are shown in Figures IV.8a-c, respectively. Inverted parameter values for outside the region of investigation are presented but our analysis focuses only on the region of investigation as previously discussed in Section 4.2. Although the spatial distribution for each hydraulic property is generally heterogeneous across the region of investigation, some layering between high and low K_h values can be recognized in the tomogram of Figure IV.8a. For instance, two dipping layers of higher K_h values (around elevations 111 m and 114 m) are interspersed by a layer of lower K_h in the middle of the tomogram. The dipping of those layers is consistent with the architecture of the geological environment revealed by GPR and CPT surveys in Figures IV.2 and IV.3. K_h values within each layer vary laterally and resulting vertical variations in K_h values along well P21 are smoother than for well P17, as depicted by simulated profiles in Figure IV.8a. The simulated K_h profile along well P17 also shows a similar trend as measured values from 15-cm interval multilevel slug tests. The small peak in K_h values around elevation 115 m is reproduced by the inversion and this layer, embedded in sediments of lower K_h values, seems to be discontinuous in space because it does not appear in the profile of well P21.

Layering for K_v/K_h is less obvious than for K_h and the spatial variation in K_v/K_h values depicted in Figure IV.8b is more heterogeneous. Nevertheless, the simulated K_v profile along well P17 compares well with K_v measurements from laboratory permeameter tests, as shown for the K_v profile in Figure IV.8b. A direct quantitative comparison between the two approaches is however difficult because of the discrepancy in scale between numerical cells (61 cm in height by 4.23 m width) and soil samples (15 cm in height by 3.8 cm in radius). Conceptually, if we assume that the fine layering in sediment samples is continuous across the larger inversion grid cells, inverted K_v values should thus be close to the harmonic average of K_v values measured on soil samples contained in each cell, according to the series flow principle for perfectly layered media [*Freeze and Cherry*, 1979]. On the other hand, if the fine layering in sediments, in particular low-K layers that impede vertical flow, is discontinuous within larger inversion cells, the resulting average K_v values over the cells will be higher than the harmonic average and will increase following the scale of the layering according to general means for layered media [*Abramowitz and Stegun*, 1972; *Ouellon et al.*, 2008].

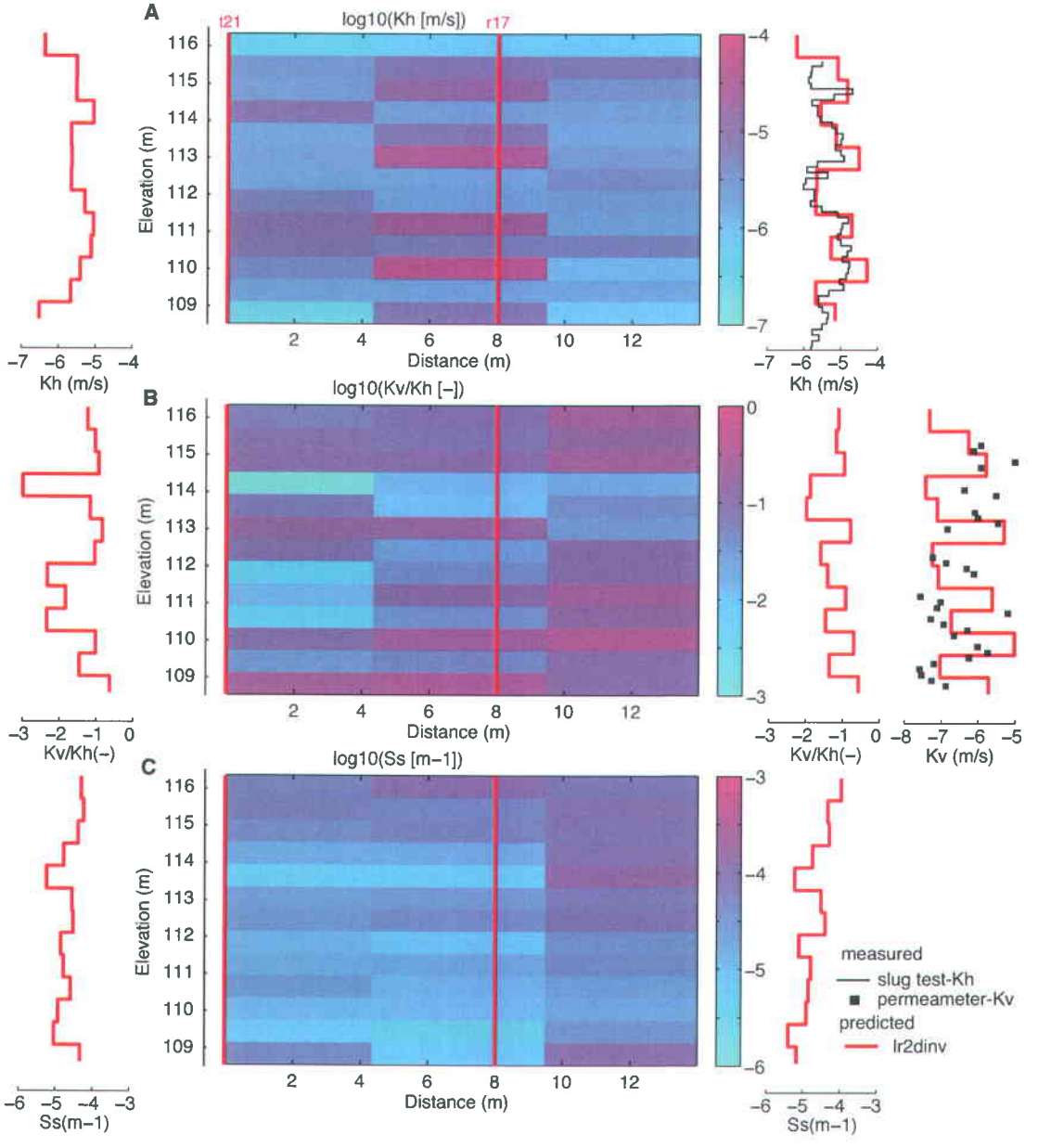


Figure IV.8 Tomograms of hydraulic parameter values from the inversion of tomographic experiment t21r17: (a) horizontal hydraulic conductivity (K_h); (b) hydraulic conductivity anisotropy (K_v/K_h); and (c) specific storage (S_s). Also depicted is the estimated vertical profile of hydraulic parameters along wells P21 and P17 with K_h values from 15-cm multilevel slug tests and K_v values from laboratory permeameter tests on 15-cm soil samples along well P17.

The tomogram and profiles for S_s in Figure IV.8c show a general decrease in S_s values with depth with relatively similar profiles along wells P21 and P17. There are no measured S_s values to compare with, but the simulated vertical trend can be related to an increase of the overburden load with depth. Conceptually, a heavier overburden load would reduce the elastic property of sediments and then their specific storage capacity. A similar behavior was observed by *Liu et al.* [2007] for sandbox experiments.

Figures IV.9a-c present dimensionless covariance values for K_h, K_v/K_h and S_s estimates depicted in Figures IV.8a-c, respectively, at corresponding cell location. Dimensionless covariance is evaluated using the root-mean square of the diagonal elements of the covariance matrix, which is approximated from the Jacobian matrix for the final parameters of the inversion. The dimensionless variances represent the uncertainties relative to the estimated parameter values. Dimensionless variances are scaled by the parameter estimates themselves and thus are not influenced by the differing magnitudes of the estimates [*Bohling and Butler*, 2001], which facilitate the comparison between different hydraulic properties. Overall, the uncertainty estimates are fairly comparable with K_h, K_v/K_h or S_s within the region of investigation, except for K_h cells near the stressed well. Dimensionless variances for K_h are indeed lower for those cells, indicating that there is greater confidence in the estimate of K_h in that region. This is due to the much higher sensitivity of stressed interval responses to K_h variations for cells near the stressed well, as demonstrated by *Paradis et al.* [2014]. Moreover, examination of dimensionless covariance patterns in Figure IV.9a-c does not reveal any correlation with coverage density of tomographic slug tests (Figure IV.5), especially for the zone in the middle of the grid where the density is lower.

In summary, results of the inversion of the tomographic experiment t21r17 show good agreement with observed data (e.g., heads, K_h and K_v data) and the littoral depositional environment (e.g., lateral trend in K_h, strong K anisotropy and decreasing S_s with depth). Although geophysical surveys (Figures IV.2 and IV.3) suggested a distinct vertical layering across the study area, estimates from tomography showed that hydraulic property variations within each geophysical-based layer are heterogeneous, which suggests complex relationships between sediments, geophysical signal and corresponding hydraulic properties. The resulting heterogeneity in hydraulic properties within the region of investigation can thus be considered representative of actual site conditions even using a coarse discretization of the aquifer. In the next section, to further strengthen this assessment, tomograms of hydraulic properties are further tested against independent hydraulic tests that have not been used in the previous inverse analysis.

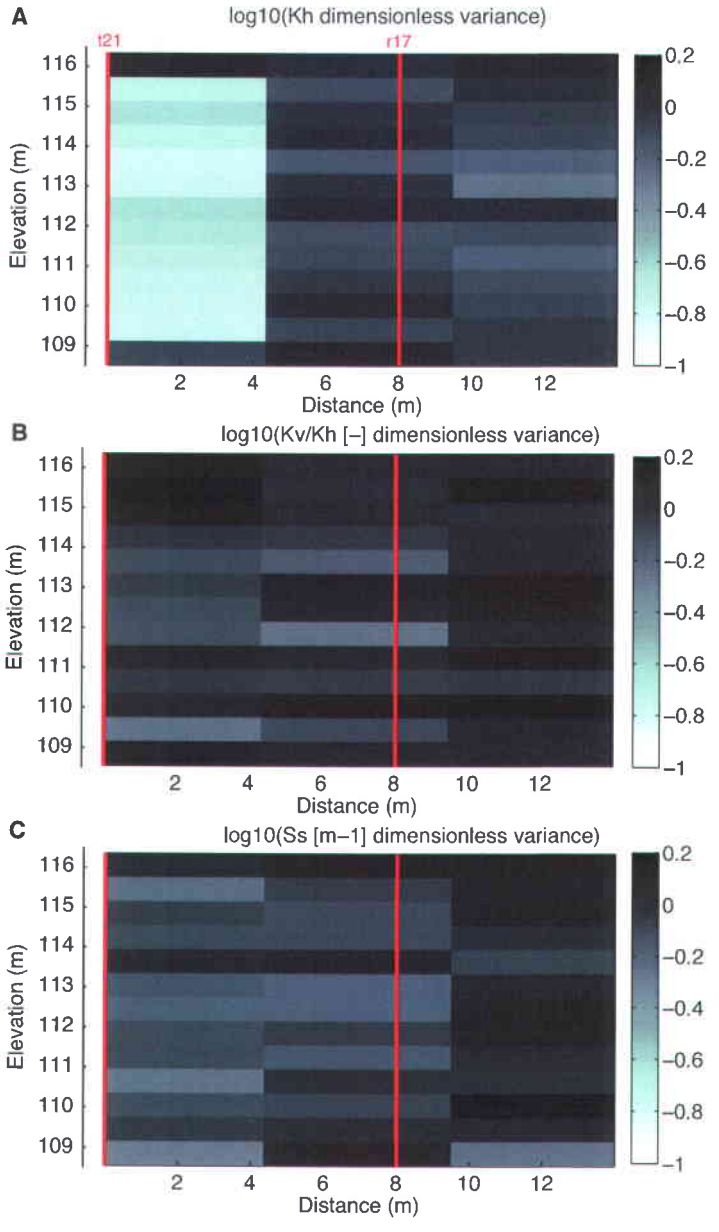


Figure IV.9 Spatial distributions of the logarithm of dimensionless variance from the inversion of tomographic experiment t21r17 for: (a) horizontal hydraulic conductivity (K_h); (b) hydraulic conductivity anisotropy (K_v/K_h); and (c) specific storage (S_s).

6 Verification of tomography results

A relatively coarse parameter grid discretization was used to minimize over-parameterization and simulated heads and hydraulic properties appear reasonable estimates of actual field conditions based on

available information. However, this does not mean that the estimated parameters can be accepted as accurate. In order to gain greater confidence in the representative nature of properties provided by the inverse analysis, a verification analysis is done using simulated K_h, K_v/K_h and S_s tomograms obtained from the inversion with hydraulic tests not used in the inversion. Two different types of hydraulic tests are used for the verification: (1) one inter-well interference slug test between wells P21 and P17 (Section 3.4); and (2) nine vertical interference slug tests done along stressed well P21 (Section 3.3). The former test, carried out over the entire well screen length, assesses the representative nature of simulated average hydraulic properties over the entire region of investigation, whereas the later tests assess parameter estimates, in particular K anisotropy, along discrete vertical intervals along the stressed well. While those tests could have been integrated in the inverse analysis to further constrain parameter inversion, they are used here as verification tools. Hence, head responses for the various verification tests are simulated with the lr2dinv model in forward mode using simulated K_h, K_v/K_h and S_s tomograms obtained from the previous inverse analysis described in Section 5 as aquifer properties. To assess the performance of the simulations, predicted head responses are compared to head observations and the error in parameter estimates in the inversion is inferred by matching predicted and observed head responses using an equivalent homogeneous and anisotropic analytical model.

6.1 Parameters verification with an inter-well interference slug test

Figure IV.10 shows the predicted head response for the inter-well interference slug test carried out using well P21 as the stressed well and well P17 as the observation well. For this test, the entire screen section of the two wells was used either as stressed or observation intervals, as described in Section 3.4, so the global behavior of the aquifer under a head perturbation can be assessed. As seen in Figure IV.10, the predicted head responses for both stressed and observation wells are fairly close to observations. More specifically, the predicted stressed and observation well responses are both slightly ahead of the observed responses and predicted head amplitudes at the observation well are a little underestimated. Consequently, r and m values are 0.99 and 0.93 for the stressed well, and 0.92 and 0.74 for the observation well, respectively, as presented in Table IV.3.

In order to quantify the error in hydraulic property estimates computed in the inversion, a type-curve matching procedure is used, assuming an equivalent homogeneous and anisotropic aquifer. The semi-analytical KGS model [*Butler*, 1995; *Hyder et al.*, 1994; *Liu and Butler*, 1995] is used here for this exercise. The test analysis procedure for the inter-well test involves the simultaneous match of both

stressed and observation well responses for the observed heads, first, and then for the predicted head responses. For each hydraulic property, the difference in parameter values for matched observed and predicted responses is then used to make an estimation of the error for this property. The error in parameter estimate for $\log K_h$, is given by:

$$E_{\log Kh} = \frac{100(\log K_h^{obs} - \log K_h^{pred})}{\log K_h^{obs}} \tag{10}$$

and similarly for $\log K_v/K_h$ and $\log S_s$. A negative value of the error indicates an underestimation of the parameter, and an overestimation for a positive value.

Results of this analysis presented in Table IV.3 for the inter-well test reveal that the error in $\log K_h$ and $\log S_s$ estimates obtained from the inversion are 5% and 1%, respectively, which indicates that average K_h and S_s values for the tested aquifer are fairly well estimated by the inversion. We note also that average K_v/K_h value cannot be estimated by the inter-well test because wells are screened over almost all the aquifer and thus the induced flow is predominantly horizontal.

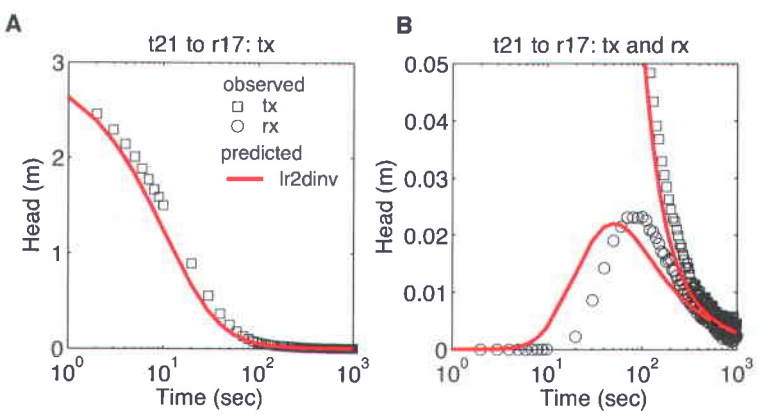


Figure IV.10 Observed (symbols) and predicted (red curves) head responses for the inter-well interference slug test used for the verification of hydraulic properties derived from tomographic experiment t21r17. Graph (a) shows the stressed well response (black squares) alone, whereas graph (b) illustrates both stressed (black squares) and observation (black circles) well responses for a zoomed view on the observation well response.

Table IV.3 Summary statistics between simulated and observed heads and hydraulic parameter error for each validation test used for the verification of the inverse analysis of tomographic experiment t21r17: "*r*" is the coefficient of correlation, "*m*" is the slope of the linear regression and "*RMSE*" is the root-mean-square error. Note that statistics for stressed interval responses of vertical interference slug tests are identical to those shown in Table IV.2 for corresponding tests because the same stressed interval was used for both tomographic and vertical interference slug tests.

Validation test	Head						Parameter error (%)		
	Stressed interval			Observation interval					
	r	*m*	*RMSE* (m / %)	*r*	*m*	*RMSE* (m / %)	$\log K_h$	$\log K_v/K_h$	$\log S_s$
Inter-wells interference slug test									
t21 to r17	0.99	0.93	0.063 / 2	0.92	0.74	0.003 / 10	5	-	1
Vertical interference slug tests									
t21-236 to r21-358	0.99	0.93	0.318 / 8	0.90	3.27	0.008 / 51	1	21	-9
t21-297 to r21-419	0.99	0.94	0.207 / 5	0.64	0.47	0.002 / 20	6	-3	2
t21-358 to r21-480	0.98	0.94	0.219 / 5	0.64	0.65	0.022 / 20	-1	-23	7
t21-419 to r21-541a	0.99	1.10	0.365 / 9	0.85	0.87	0.027 / 14	-4	-29	8
t21-480 to r21-602	0.99	1.03	0.078 / 2	0.77	0.93	0.017 / 26	-2	-17	5
t21-541a to r21-663	0.99	1.06	0.173 / 4	0.98	1.80	0.007 / 30	2	-14	5
t21-602 to r21-724	0.99	0.95	0.132 / 3	0.30	0.56	0.037 / 43	-5	-33	10
t21-663 to r21-785	0.99	0.90	0.210 / 5	0.76	0.53	0.020 / 28	1	-35	10
t21-724 to r21-845	0.99	0.93	0.269 / 6	0.99	0.57	0.087 / 38	6	-81	20
All vertical tests	0.99	0.97	0.252	0,91	0.66	0.027	-	-	-

6.2 Parameters verification with vertical interference slug tests

While an inter-well test provides an appreciation of the global behavior of the aquifer under stress, vertical interference slug tests can indicate how heterogeneity in K_h, K_v/K_h and S_s are well estimated by the inverse analysis at specific locations into the aquifer. Figure IV.11 shows the predicted observation interval response for each of the nine vertical interference slug tests carried out along well P21. We note that stressed interval responses are not shown in Figure IV.11 but they are identical to the responses presented in Figure IV.8. In fact, the stressed interval responses are the same for both vertical interference and tomographic slug tests because they were already used in the inverse analysis of tomographic data. Only observation interval responses of vertical interference tests were not used in the inversion. Although visual inspection of individual test depicted in Figure IV.11 reveals that predicted heads for some tests moderately depart from observations, the overall r and m statistics for all tests are relatively good with values of 0.91 and 0.66, respectively (Table IV.3). Moreover, vertical interference tests are very sensitive and a small modification in hydraulic properties generally results in a large change in the amplitude and arrival time of the head response in the observation interval [*Paradis and Lefebvre*, 2013]. Indeed, errors in parameter estimates presented in Table IV.3 for each vertical interference test are below 6%, 35% and 10% for $\log K_h$, $\log K_v/K_h$ and $\log S_s$, respectively, except for the last test (t21-724 to r21-845) that shows larger errors.

It should be noted that the error estimation procedure for vertical tests was slightly modified relative to the procedure used for the inter-well test. In fact, in heterogeneous profiles it is not possible to match both stressed and observation intervals assuming homogeneous and anisotropic conditions, and both K_v/K_h and S_s variations have similar influence on the observation interval response [*Paradis and Lefebvre*, 2013]. Thus, in order to evaluate error in parameter estimates for vertical tests measured observation interval response were first matched, then the predicted responses were obtained by varying K_h and S_s while holding K_v/K_h the same to estimate K_h and S_s. The same procedure was repeated by varying K_h and K_v/K_h while holding S_s constant to estimate K_h and K_v/K_h. The resulting estimates of errors are obviously gross approximations and errors reported in Table IV.3 for K_v/K_h and S_s are then expected to be overestimated. Overall, the differences between observed and predicted head responses from vertical interference slug tests result in fairly small error in K_h, K_v/K_h and S_s estimates.

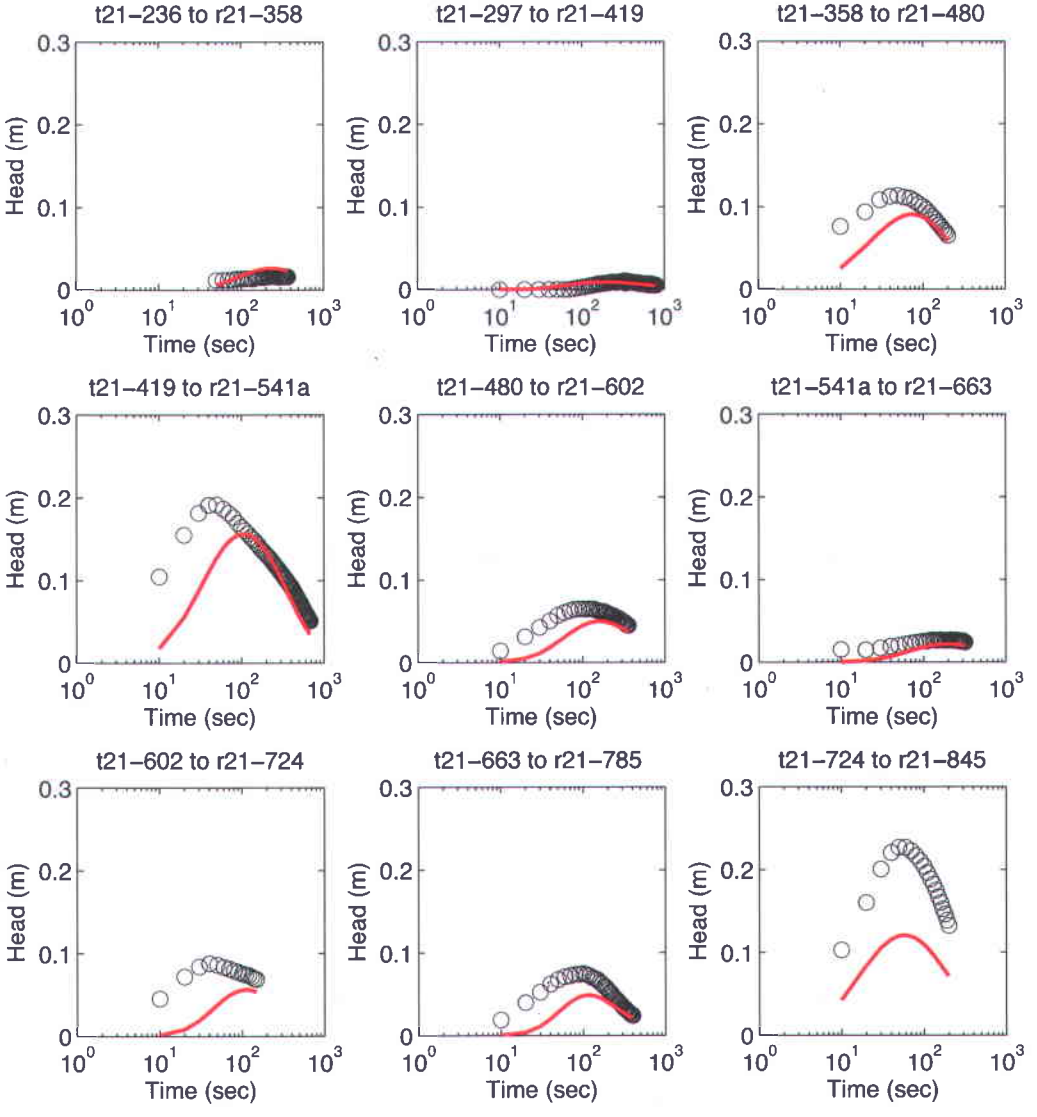


Figure IV.11 Observed (symbols) and predicted (red curves) head responses for vertical interference slug tests [Paradis and Lefebvre, 2013] used for the verification of tomographic experiment t21r17. Stressed interval responses are not shown for those tests but they are identical to head responses depicted in Figure IV.7 for corresponding tests.

7 Conclusions

The main objective of this study was to assess the performance of a transient analysis of tomographic slug test data in estimating heterogeneity in K_h, K_v/K_h and S_s in real field conditions. To achieve this goal, results of the inversion of a tomographic experiment performed in moderately heterogeneous and highly anisotropic silt and sand littoral aquifer were compared to estimates obtained using other hydraulic tests at the same site. The simulation of independent large-scale and single-well hydraulic tests not used in the

inversion tested the representativity of the estimates. According to the results, the main findings of this study can be summarized as follows.

- Simulated heads were obtained from the simultaneous inversion of the transient head responses recorded into multiple stressed and observation intervals with a two-dimensional radial finite difference flow model considering K_h, K_v/K_h, S_s and wellbore storage effects. These simulated heads closely match observations.

- The K_h and K_v estimates obtained from tomography compare well with those of high-resolution multilevel slug tests and laboratory permeameter tests on small soil samples, despite the difficulty to compare K_v values from each method due to scale disparity. The structure of the heterogeneity in K_h (dipping and lateral variations) is also coherent with the architecture of the geological environment revealed by geophysical surveys. Moreover, estimates of S_s from tomographic slug tests exhibit physically plausible vertically decreasing S_s profiles.

- K_h, K_v/K_h and S_s models estimated from the inverse analysis of tomographic data are used to compute the predicted transient head responses from the simulation of an inter-well interference slug test and nine vertical interference slug tests, which are shown to be in close agreement with field observations. That is, global and local aquifer dynamics are thus adequately represented by inverted K_h, K_v/K_h and S_s tomograms, which is critical for a good understanding of flow and transport in aquifers. This demonstrates the robustness of the tomographic slug tests approach.

In addition, in this study we used a flow simulation model that assumes radial symmetry to investigate both vertical and lateral hydraulic property variations in the plane encompassing the stressed and observation wells. A radially symmetric model should provide a reasonable estimate of flow behavior in this plane as long as lateral variations in the hydraulic properties do not induce significant angular variations in flux toward the well, in which case the narrow angular wedge encompassing the investigated plane can be treated in isolation as a two-dimensional radial-vertical problem bounded by zero-flux boundaries in the angular direction [*Bohling*, 2009]. That is, considering the heterogeneous nature of the study site, the use of a radially symmetric model could be questioned. However, considering the results of the verification analysis, it is suggested that the hydraulic behavior of the tested aquifer is reasonably well

represented by a radial-flow model. Both small head variations recorded in observation intervals and strongly focused flux near stressed intervals may explain why the present inverse analysis seems to be not affected by significant angular variations in flux (see *Paradis et al.*, 2014).

Furthermore, the fact that the inversion scheme supports anisotropy provides greater confidence in the estimated model parameters and may explain the success of the verification process. Considering the very small-scale (<15 cm) heterogeneity of the study site, it would have been practically impossible to design a tomographic experiment that satisfies the assumption of isotropy at the scale of the grid cells used for the inverse analysis. For that case, the problem of non-uniqueness in parameter estimates would have been obviously amplified because of the oversimplification of the anisotropic problem into an isotropic one. Thus, by considering K anisotropy, the inversion provides more flexibility in the design of the field survey. Even if the results of the inverse analysis reveal no anisotropy, we believe that greater confidence in the parameter estimates is gained by considering anisotropy.

Finally, this study demonstrated that a transient analysis of tomographic slug test data was able to capture the key features of the study aquifer, namely K_h, K_v/K_h and S_s heterogeneities, on a scale relevant for groundwater flow and contaminant transport investigations. Future work will focus on the integration of additional wells in the inverse analysis, in particular reciprocal tests between pair of stressed and observation wells, to increase the effective spatial resolution of hydraulic properties, as shown by *Paradis et al.* [2014]. A three-dimensional observation interval network would also have to be considered for a more accurate representation of aquifer heterogeneities, as already implemented by *Berg and Illman* [2011] and *Cardiff et al.* [2012] with tomographic pumping tests to image the spatial distribution of K_h.

Acknowledgements

The authors would like to acknowledge the important technical support provided by J.-M. Ballard, R. Lessard, É. Paquet, L. Fleury, M. Sauvageau and C. DeLouis. Sincere thanks to H. Russell and N. Benoit for their comments that helped improve an early version of this manuscript. We are also grateful to Y. Michaud, A. Rivera and D. Kirkwood of the Geological Survey of Canada (GSC). This study was supported by the GSC as part of the Groundwater Resources Inventory program, by the Régie intermunicipale de gestion des déchets des Chutes-de-la-Chaudière and by NSERC Discovery Grants

held by R.L. and E.G. This is ESS contribution number 20090447. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the Canadian Government.

References

Abramowitz, M., and I.A. Stegun (1972) Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables, 10th ed. Applied Math. Series US National Bureau of Standards, Washington, 1046pp.

American Society for Testing and Material (ASTM) (2004), D6724: Standard Guide for Installation of Direct Push Ground Water Monitoring Wells, ASTM International, pp. 9.

Aster, R. C., B. Borchers, and C.H. Thurber (2005), Parameter Estimation and Inverse Problems, 301 pp., Elsevier, Amsterdam.

Berg, S.J., and W.A. Illman (2011), Three-dimensional transient hydraulic tomography in a highly heterogeneous glaciofluvial aquifer-aquitard system, Water Resour. Res., 47, W10507, doi: 10.1029/2011WR010616.

Bolduc, A. (2003), Géologie des formations superficielles, Charny, Québec, Commission Géologique du Canada, Dossier public 1976, échelle 1/50000, Ottawa, Ont., Canada.

Bohling, G.C. (1993), Hydraulic tomography in two-dimensional, steady-state groundwater flow, Eos Trans. AGU, 74, 141.

Bohling, G.C., and J.J., Butler Jr. (2001), lr2dinv: A finite-difference model for inverse analysis of two-dimensional linear or radial groundwater flow, Comput. Geosci., 27(10), 1147-1156.

Bohling, G.C. (2009), Sensitivity and resolution of tomographic pumping tests in an alluvial aquifer, Water Resour. Res., 45, W02420, 10.1029/2008WR007249.

Bohling, G.C., J.J., Jr. Butler, X. Zhan, and M.D. Knoll (2007), A Field Assessment of the value of steady-shape hydraulic tomography for characterization of aquifer heterogeneities, Water Resour. Res., 43(5), W05430, doi:10.1029/2006WR004932.

Brauchler, R., J. Cheng, M. Everette, B. Johnson, P. Dietrich, R. Liedl, and M. Sauter (2007), An inversion strategy for hydraulic tomography: Coupling travel time and amplitude inversion, J. Hydrol., 345, 184-198, doi:10.1016/j.jhydrol.2007.08.011.

Brauchler, R., R. Hu, P. Dietrich, and M. Sauter (2011), A field assessment of high-resolution aquifer characterization based on hydraulic travel time and hydraulic attenuation tomography, Water Resour. Res., 47, W03503, doi:10.1029/2010WR009635.

Butler Jr., J.J. (1995) Slug tests with observation wells; extension of Hyder et al.

(1994) Solution to case of head in an observation well, Kansas Geological Survey, Open-file Report No. 95-43, p. 14.

Butler Jr., J.J. (2005), Hydrogeological methods for estimation of hydraulic conductivity, edited by Y. Rubin and S. Hubbard, in Hydrogeophysics, pp. 23-58, Springer, New York.

Butler Jr., J.J., and C.D. McElwee (1995), Well-testing methodologies for characterizing heterogeneities in alluvial-aquifer systems: final technical report, Open File Report 95-75, pp. 320, Kansas Geological Survey, Lawrence, Kansas.

Cardiff, M., and W. Barrash (2011), 3-D transient hydraulic tomography in unconfined aquifers with fast drainage response, Water Resour. Res., 47, W12518, doi:10.1029/2010WR010367.

Cardiff, M., W. Barrash, and P.K. Kitanidis (2012), A field proof-of-concept of aquifer imaging using 3-D transient hydraulic tomography with modular, temporarily-emplaced equipment, Water Resour. Res., 48: W05531.

Clemo, T., P. Michaels, and R. M. Lehman (2003), Transmissivity resolution obtained from the inversion of transient and pseudo-steady drawdown measurements, in Proceedings of MODFLOW and More 2003 Understanding Through Modeling, pp. 629-633, Int. Ground Water Model. Cent., Golden Colo.

De Marsily, G., F. Delay, J. Gonçavès, P. Renard, V. Teles, and S. Violette (2005), Dealing with spatial heterogeneity. Hydrogeol. J., 13, 161-183.

Fellenius, B.H., and A. Eslami (2000), Soil profile interpreted from CPTu data, Year 2000 Geotechnics, Geotechnical Engineering Conference, Asian Institute of Technology, Bangkok, Thailand.

Freeze, R.A. and J.A. Cherry (1979), Groundwater. Prentice-Hall, Inc. Englewood Cliffs, NJ. 604 p.

Heath, R.C. (1983), Basic ground-water hydrology, U.S. Geological Survey Water-Supply Paper 2220, p 8-9.

Hyder, Z., J.J. Butler Jr., C.D. McElwee, W. Liu (1994) Slug test in partially

penetrating wells, Water Resour. Res., 30(8), 2945-2957.

Illman, W.A., A.J. Craig, and X. Liu (2008), Practical issues in imaging hydraulic conductivity through hydraulic tomography, Ground Water, 46(1), 120–132, 10.1111/j.1745-6584.2007.00374.x.

Koltermann, C.E. and S.M. Gorelick, (1996), Heterogeneity in sedimentary deposits: A review of structure imitating, process imitating, and descriptive approaches, Water Resour. Res., 32: doi: 10.1029/96WR00025.

Lamarche, L., and L. Tremblay (2012), Géologie des formations superficielles pour le site de St-Lambert-de-Lauzon, Québec, Unpublished research data.

LeBlanc, D.R., S.P. Garabedian, K.M. Hess, L.W. Gelhar, R.D. Quadri, K.G. Stollenwerk, and W.W. Wood (1991), Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts, 1, Experimental design and observed tracer movement, Water Resour. Res., 27, 895-910.

Liu, X., W.A. Illman, A.J. Craig, J. Zhu, and T.-C.J. Yeh (2007), Laboratory sandbox validation of transient hydraulic tomography, Water Resour. Res., 43, W05404, 10.1029/2006WR005144.

Liu, W.Z., and J.J. Butler Jr., (1995) The KGS model for slug tests in partially penetrating wells, Version 3.0, Lawrence, Kansas: Kansas Geological Survey Comput. Ser. Rep. 95-1.

Lunne, T., P.K. Robertson, and J.J.M. Powell (1997), Cone penetration testing in geotechnical practice, Spon 1075 Press, Taylor and Francis Group, New York.

Neuman, S.P. (1987), Stochastic continuum representation of fractured rock permeability as an alternative to the REV and fracture network concepts, pp. 533-561, Rock Mechanics: Proceedings of the 28th U.S. Symposium, Tucson, Arizona, edited by I.W. Farmer, J.J.K. Daemen, C.S. Desai, C.E. Glass, and S.P. Neuman, A.A. Balkema, Rotterdam, 1240 p.

Ouellon, T., R. Lefebvre, D. Marcotte, A. Boutin, V. Blais, and M. Parent (2008), Hydraulic conductivity heterogeneity of a local deltaic aquifer system from the kriged 3D distribution of hydrofacies from borehole logs, Valcatier, Canada. J. of Hydrology, 351 (1-2), 71-86.

Paradis, D., and R. Lefebvre (2013), Single-well interference slug tests to assess the vertical hydraulic conductivity of unconsolidated aquifers, J. Hydrol., 478(25), 102-118, doi: 10.1016/j.jhydrol.2012.11.047.

Paradis, D., R. Lefebvre, R.H. Morin, and E. Gloaguen (2011), Permeability profiles in granular aquifers using flowmeters in direct-push wells, Ground Water, 49, 534–547, doi: 10.1111/j.1745-6584.2010.00761.x.

Paradis, D., E. Gloaguen, R. Lefebvre, and B. Giroux (2014), Tomographic slug tests information content, accepted pour publication pour Water Resources Research, 25 novembre 2013.

Prats, M., J.B. Scott (1975), Effect of wellbore storage on pulse-test pressure response, J. Petrol. Tech., 707–709.

Spane Jr., F.A. (1996), Applicability of slug interference tests for hydraulic characterization of unconfined aquifers: (1) analytical assessment, Ground Water, 34(1), 66-74.

Sudicky, E.A. (1986), A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process, Water Resour. Res., 22, 2069-2082.

Tosaka, H., K. Masumoto, and K. Kojima (1993), Hydropulse tomography for identifying 3-D permeability distribution in High Level Radioactive Waste Management, Proceedings of the Fourth Annual International Conference of the ASCE, pp. 955–959, Am. Soc. Civ. Eng., Reston, Va.

Tremblay, L., R. Lefebvre, D. Paradis, and E. Gloaguen, in print. Conceptual model of leachate migration in a granular aquifer from the integration of detailed characterization data. HJ-2012-2516 accepted for publication on August 3, 2013 by *Hydrogeology Journal.*

Van Overmeeren, R.A. (1998), Radar facies of unconsolidated sediments in The Netherlands: A radar stratigraphy interpretation method for hydrogeology, J. Appl. Geophys., 40(1-3), doi: 10.1016/S0926-9851(97)00033-5.

Vasco, D.W., A. Datta-Gupta, and J.C.S. Long (1997), Resolution and uncertainty in hydrologic characterization, Water Resour. Res., 33(3), 379-397, 10.1029/96WR03301.

Wu, C.-M., T.-C. J. Yeh, J. Zhu, T.H. Lee, N.-S. Hsu, C.-H. Chen, and A.F. Sancho (2005), Traditional analysis of aquifer tests: Comparing apples to oranges?, Water Resour. Res., 41, W09402, doi: 10.1029/2004WR003717.

Zhu, J., and T.-C. J. Yeh (2005), Characterization of aquifer heterogeneity using transient hydraulic tomography, Water Resour. Res., 41, W07028, 10.1029/2004WR003790.

ARTICLE V

PREDICTING HYDROFACIES AND HYDRAULIC CONDUCTIVITY FROM DIRECT-PUSH DATA USING A DATA-DRIVEN RELEVANCE VECTOR MACHINE APPROACH: MOTIVATIONS, ALGORITHMS AND APPLICATION

Résumé

L'hétérogénéité de la conductivité hydraulique (K) exerce un contrôle important sur l'écoulement de l'eau souterraine et le transport des contaminants. La distribution spatiale de K peut être représentée par des mesures géophysiques indirectes en utilisant des relations empiriques ou statistiques liant les mesures géophysiques et K. Cet article présente une approche non paramétrique par machines d'apprentissage RVM (*relevance vector machine*) pour prédire K à partir de sondages au piézocône (CPT) couplés avec une sonde de teneur en eau et de résistivité électrique du sol (SMR). L'approche par machine d'apprentissage est démontrée par une application dans un aquifère granulaire hétérogène couvrant un sous-bassin versant de 12 km^2. Avec les données de la région d'étude, les méthodes conventionnelles de régression ne permettent pas de prédire K de manière satisfaisante à cause de la très forte non linéarité des relations liant K et les paramètres CPT/SMR. L'approche proposée nécessite d'abord l'application d'un algorithme de groupage flou (*fuzzy clustering*) pour définir les hydrofacies (HF) sur la base des données CPT/SMR et de K, pour ensuite appliquer des machines d'apprentissage RVM pour la classification des HF et la prédiction de K pour chacun des HF sur la base des données CPT/SMR seulement. La machine d'apprentissage a été développée à partir d'un ensemble de données d'entrainement colocalisées de K mesures par des essais de perméabilité à choc hydraulique et des mesures CPT/SMR mise à une échelle verticale commune de 15 cm. Après l'entrainement, la machine d'apprentissage a été testée à l'aveugle en utilisant des données non utilisées lors du processus d'entrainement. Les résultats montrent que la reconnaissance des HF et la prédiction de K avec la machine d'apprentissage sont cohérents avec les observations. L'utilisation combinée de données CPT/SMR et de machines d'apprentissage RVM s'avère donc une approche robuste et efficace, qui pourrait être généralisable pour la caractérisation des aquifères granulaire en général.

Accepté pour *Water Resources Research* le 7 mai 2014

Predicting hydrofacies and hydraulic conductivity from direct-push data using a data-driven relevance vector machine approach: motivations, algorithms and application

Daniel Paradis[1,2,*], René Lefebvre[2], Erwan Gloaguen[2], and Alfonso Rivera[1]

[1] Geological Survey of Canada, 490 rue de la Couronne, Quebec City, Canada G1K 9A9

[2] Institut National de la Recherche Scientifique, Centre Eau Terre Environnement (INRS-ETE), 490 rue de la Couronne, Quebec City, Canada G1K 9A9

*Corresponding author: Daniel Paradis

Geological Survey of Canada, 490 rue de la Couronne, Quebec City, Canada G1K 9A9

Phone : (418) 654-3713 Fax : (418) 654-2604 E-mail: dparadis@nrcan.gc.ca

Abstract

The heterogeneity of hydraulic conductivity (K) exerts a major control on groundwater flow and solute transport. The heterogeneous spatial distribution of K can be imaged through indirect geophysical measurements using empirical or statistical relations between indirect geophysical data and K. This paper presents a non-parametric learning machine approach to predict aquifer K from cone penetrometer tests (CPT) coupled with a soil moisture and resistivity probe (SMR) using relevance vector machines (RVMs). The learning machine approach is demonstrated with an application to a heterogeneous unconsolidated littoral aquifer in a12 km^2 sub-watershed. In that study area, conventional regression methods could not successfully predict K due to strong nonlinearity in relations between K and CPT/SMR parameters. Our approach involved fuzzy clustering to define hydrofacies (HF) on the basis of CPT/SMR and K data prior to the use of RVM for HF classification and K prediction for each HF on the basis of CPT-SMR data. The learning machine was built from a colocated training data set of K measurements from slug tests in wells and CPT/SMR data up-scaled at a common vertical resolution of 15 cm. After training, the learning machine was blind tested using data that was withheld from the training process. Results show that HF recognition and K predictions from the learning machine are consistent with observations. The combined use of CPT/SMR data and RVM-based learning machine proved to be powerful and efficient and it appears as a generally applicable approach for K heterogeneity characterization for unconsolidated aquifers.

Keywords: Aquifer Characterization, Hydrogeophysics, Heterogeneity, Hydraulic conductivity, Direct-push, Relevance vector machine

1 Introduction

Knowing the heterogeneous spatial distribution of hydraulic conductivity (K) in aquifers is a prerequisite to tackle groundwater flow and transport problems. Indeed, since K may vary over several orders of magnitude and impacts both the magnitude and direction of advective transport, the primary focus of aquifer characterization is generally on the measurement of K [*Koltermann and Gorelick*, 1996]. Aquifer K is mostly measured using hydraulic tests carried out in wells (e.g., slug tests, pumping tests). Although such tests are generally reliable sources of data about K, they are however costly and time consuming. Such measurements are thus usually available only from a few wells and at a too low spatial resolution to adequately define K heterogeneities at the scale needed for most practical groundwater flow and mass transport studies [*Butler*, 2005].

Due to these limitations of conventional hydraulic characterization, hydro-geophysics is increasingly recognized as an effective alternative to better spatially distributes controls on hydraulic properties, which requires the translation of indirect geophysical data into hydraulic properties [*Rubin and Hubbard*, 2005]. The value of using geophysical data for hydrogeological characterization lies in the extensive spatial coverage generally offered by geophysical methods, which may be helpful to provide spatial continuity in K heterogeneities. Reliable predictions in K from geophysical data should however be based on sound relations between hydraulic and geophysical data, which are usually subject to a large degree of uncertainty under field conditions [*Chen et al.,* 2001]. The major problem with the integration of hydro-geophysical data is non-uniqueness. Typical causes of non-uniqueness are the scale and the resolution disparity between hydraulic and geophysical measurements and the uncertainty associated with field data acquisition and interpretation. Another fundamental aspect with non-uniqueness is the degree of sensitivity between hydraulic and geophysical parameters, which for a particular geological material may result in a fairly weak correlation. That problem is exacerbated under heterogeneous field conditions where sensitivities may vary for different geological materials and thus preclude reliable estimations of hydraulic properties from geophysical data. Thus, the overall motivation of this work is the need to develop efficient and robust aquifer characterization and data analysis approaches that can provide more information about K: higher number of control points, relatively fine vertical resolution, continuous vertical profiles, based on repeatable physical measurements. Such larger high quality dataset are needed

to define K heterogeneity using geostatistical interpolation schemes (estimation or simulation) in order to develop more realistic numerical groundwater flow and solute transport models [*Anderson*, 1997].

This paper explores the potential of using cone penetrometer tests (CPT) coupled with a soil moisture and resistivity probe (SMR) for estimating hydrofacies (HF) and K in unconsolidated aquifers. The paper is focused on the assessment of the usefulness of CPT/SMR soundings for HF and K estimation and on the integration of hydraulic and geophysical data through a learning machine approach based on relevance vector machines (RVMs). CPT/SMR is a multi-parameter probe that simultaneously provides vertical profiles of mechanical (tip stress, sleeve stress, pore pressure) and electrical (dielectric constant, bulk electrical resistivity) parameters of sediments. Thus, due to the number of simultaneously measured geophysical parameters and their similar volumes of investigation, CPT/SMR soundings have the potential to reduce non-uniqueness between geophysical measurements and K. From a practical viewpoint, the value of using CPT/SMR data for aquifer characterization lies in the vertical decimeter-scale resolution offered by this direct-push technique (*Lunne et al.*, 1997; *Schulmeister et al.*, 2003), which cannot be obtain by surface-based geophysical methods. Moreover, the relatively higher number of continuous vertical profiles that can be obtain by direct-push soundings in comparison to wells or core-based hydraulic tests allows the definition of aquifer heterogeneities over large investigation areas (*Lafuerza et al.*, 2005; *Paradis et al.*, 2014). CPT soundings for geological applications have been mostly used to deduce sediment texture from mechanical parameters [e.g., *Robertson*, 1990; *Fellenius and Eslami*, 2000]. *Farrar* [1996] also proposed a chart to evaluate K from sediment texture, but it only provides order-of-magnitude K estimates and it does not make full use of electrical parameters provided by the SMR probe.

Our objective is thus to develop a general approach to define site-specific relations to reliably predict K from CPT/SMR data as an extension of the spatial coverage of conventional hydraulic testing. Hydraulic testing is still needed to establish a relationship between K and CPT/SMR parameters, but it only needs to cover the representative range of materials in a study area, not to extensively cover the area. Also, to facilitate the spatial interpretation of K heterogeneity over a study area and to allow a better integration with geological depositional models, we also wish to use direct-push data to define hydrofacies (HF) [e.g., *Anderson*, 1997; *Koltermann and Gorelick*, 1996; *Ouellon et al.*, 2008; *Paradis et al.*, 2014]. A HF is a homogeneous unit that is hydrogeologically meaningful for the purposes of flow and transport modeling [*Anderson*, 1989] and it is defined here as a distinct unit in terms of K distribution. In this paper, the definition of relations to predict HF and K from CPT/SMR data is made through a non-

parametric learning machine approach because of the complex nonlinear relations that generally exist between hydro-geophysical parameters [e.g., *Mohaghegh et al.*, 1997; *Lee and Datta-Gupta*, 1999; *Dubois et al.*, 2007]. Learning machines do not assume a rigid functional form, they rely on the available data to build up a model of the system, and no a priori assumptions on parameter relations are made [*Mitchell*, 1997].

In this context, different models and architectures of artificial neural networks (ANNs) have been commonly considered and applied successfully for the prediction of lithofacies [*Chen and Rubin*, 2003; *Dubois et al.*, 2007] or hydraulic parameters in petroleum reservoirs [*Mohaghegh et al.*, 1997; *Wong et al.*, 1998; *Lee and Datta-Gupta*, 1999; *Shokir et al.*, 2006; *Al-Anazi et al.*, 2009; *Elshafei and Hamada*, 2009; *Kharrat et al.*, 2009] from cross-hole or borehole geophysics data. However, despite their potential effectiveness, ANNs present some important drawbacks: (i) design and training often results in a complex, time-consuming task, in which many parameters must be tuned; (ii) minimization of the training errors can lead to poor generalization performance; and (iii) performance can be degraded when working with small datasets.

A promising alternative to ANNs is the support vector machine (SVM), which was developed to solve both the classification and the regression problems [*Vapnik*, 1995; 1998]. Rather than following the empirical risk minimization of ANNs, where a function of the training errors is minimized, SVM follows the structural risk minimization, by which an upper bound of generalization performance is maximized. In addition, SVM can efficiently control model complexity through a regularization term, SVM can work with high dimensional input spaces, and deal with sparse datasets. The main idea behind SVM is to perform a linear regression in a high dimension feature space, through a kernel function, which returns a nonlinear regression in the original input space. The SVM is a non-parametric, regularized and nonlinear regression tool, which has yielded good results for the prediction of lithofacies, permeability and porosity of petroleum reservoirs from borehole geophysics data and in alleviating the aforementioned problems of ANNs [*Al-Anazi and Gates*, 2010a; 2010b].

The rationale for selecting relevance vector machine (RVM) approaches for this study over ANNs and SVM is that many studies have shown that RVM performs better than either ANNs or SVM in many applications for accuracy and sparsity of the solution [*Khalil et al.*, 2005; *Camps-Valls et al.*, 2006; *Samui*, 2007; *Ghosh and Mujumdar*, 2008]. A RVM is a Bayesian extension of the SVM to solve

nonlinear classification and regression models. Along with its ability to produces relations with good generalization capability with sparse and nonlinear data, which is typical in most geosciences applications, RVM produces probabilistic outputs that capture uncertainty in the predictions. RVM is also easier to apply since it has no regularization term needing to be adjusted [*Tipping*, 2001].

The remainder of this paper is organized as follows. Section 2 describes the study area and the training dataset used to develop classification and regression models. Section 3 outlines the learning machine approach and provides a description of the main algorithms, which includes fuzzy clustering and RVMs for classification and regression. Section 4 presents the application of the learning machine to the training dataset to recognize HF and estimate K. Conclusions about the key findings of this study are listed in Section 5.

2 Field characterization and training dataset

Since learning machines are based on empirical data, the collection of a representative training dataset for a given study area is fundamental to establish meaningful relations between hydraulic and geophysical parameters. *Paradis et al.* [2014] described the general data acquisition approach that was followed for the characterization of hydraulic properties of the study area, which includes the collection of the training dataset used in this paper. Data acquisition followed a topdown approach in which the characterization methods with larger spatial coverage were first applied, followed by more local methods. In particular, the definition of a general hydrogeological conceptual model [*Tremblay et al.*, 2013] through regional geology, GPR surveys and CPT/SMR soundings allowed the targeting of specific locations for well installations and K testing in order to cover the whole range of hydro-geophysical responses observed over the study area. In this study, it is those K measurements with corresponding colocated CPT/SMR data that compose the training dataset used to develop hydro-geophysical relations. In this section, we thus describe the study area, provide a summary of the data acquisition process for CPT/SMR and K data, and describe the hydro-geophysical training dataset.

2.1 Saint-Lambert study area

The proposed methodology was developed and applied in relation with a study carried out in St-Lambert-de-Lauzon, located 30 km south of Quebec City, Canada (Figures 1a-b). As illustrated in Figure 1c, the

study area encompasses a 12-km^2 sub-watershed area surrounding a decommissioned sanitary landfill where an assessment of the migration of a leachate plume was underway [*Tremblay et al.*, 2013].

As reported by *Bolduc* [2003], the surficial sediments of the study area (Figure 1c) consist primarily of Late Quaternary sandy and silty sediments that were deposited in the receding Champlain Sea, which was an arm of the Atlantic Ocean that had invaded the St.Lawrence Valley at the time of the last deglaciation. More specifically, deposition at the St-Lambert site was controlled mainly by longshore currents that redeposited in littoral and sublittoral settings the sediments supplied to the Chaudière River paleodelta. This is indicated in Figure 1c by the southwestward fining of the littoral sediments in conjunction with the southwest-northeast trend of the beach ridges [*Lamarche and Tremblay*, 2012]. These ridges and the associated nearshore bars are mostly composed of medium to fine sand while the intervening troughs composed of finer, silty sediments.

Sediment size for the aquifer ranges from fine sand to very fine silt (clayey silt) with poor to very poor grain-size sorting. Rapid transitions in sediments size and sorting caused by rapidly varying depositional energy are also generally observed in surface exposures. The resulting superposition of sand and silt layers may then create semi-confined aquifer conditions due to their respective K contrast. Lateral intrastratal transitions in these littoral and sublittoral sediments are expected as a result of changing energy levels along Champlain Sea shorelines. A till layer that varies in depth from surface outcrop to 22 m is present and its low permeability generally limits water exchanges between the Quaternary sediments and the underlying regional fractured sedimentary rock aquifer. The water table is shallow, being found 1 to 2 m below ground surface.

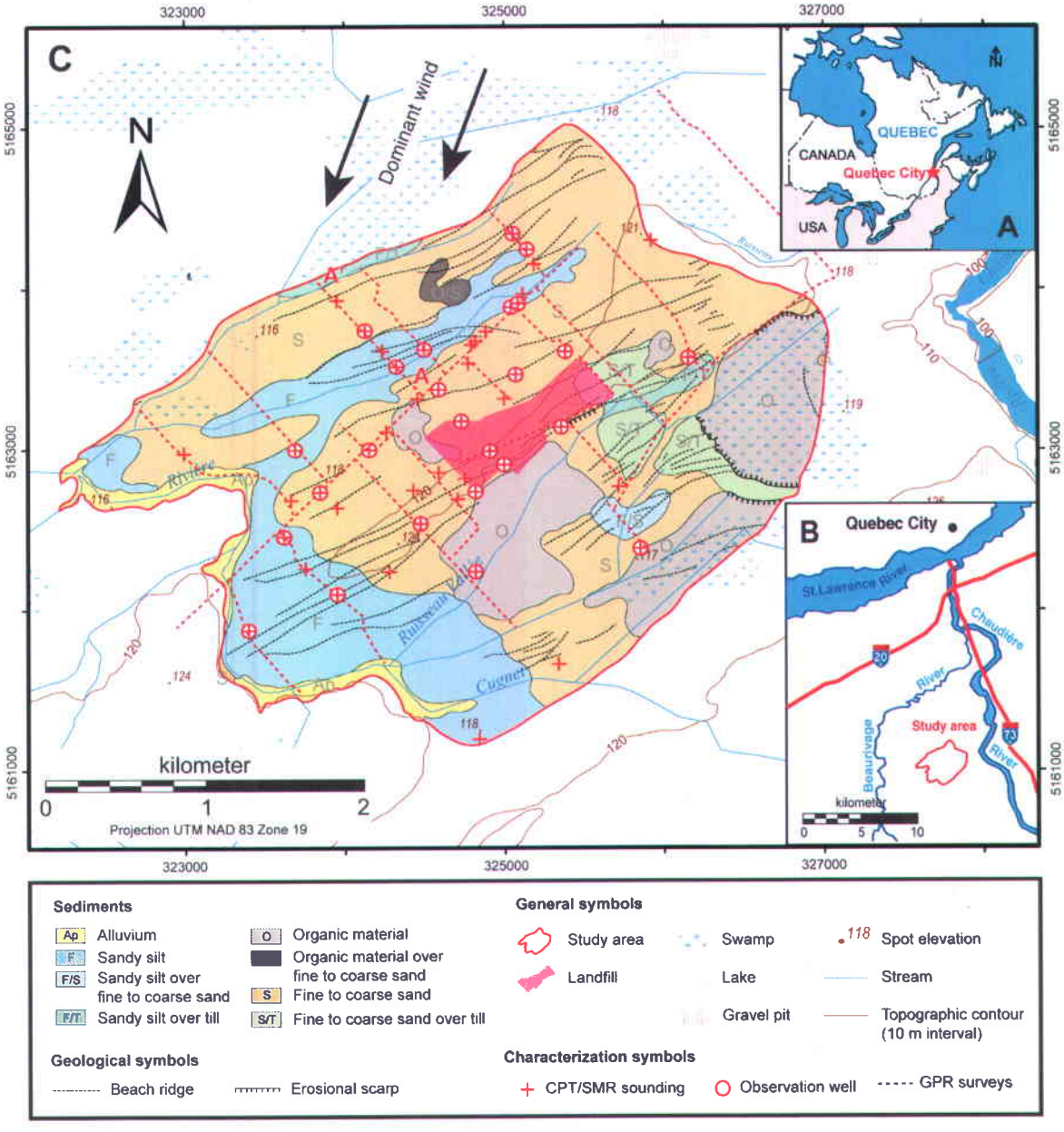


Figure V.1 General location of the St-Lambert study area (a-b), with the Quaternary sediments map (c) for the sub-watershed surrounding the decommissioned sanitary landfill showing the locations of direct-push soundings and observation wells used for aquifer characterization. The main depositional direction of sediments making up the granular aquifer is assumed to have been oblique to the orientation of the paleo-shore in a littoral environment. The Quaternary map was modified from *Lamarche and Tremblay* (2012).

2.2 Geophysical measurements from CPT/SMR soundings

According to the regional surficial sediments geology and GPR surveys, the location of 53 CPT/SMR soundings were selected (Figure 1c) to provide further geological interpretation and to acquire mechanical and electrical properties of sediments to establish relations with hydraulic measurements [*Paradis et al.*, 2014]. Direct-push soundings were carried out using a Geotech 605-D rig equipped with a CPT system including pore pressure measurement combined with a SMR probe. CPT and SMR probes allow the simultaneous measurement of three mechanical and two electrical properties of sediments, respectively, at vertical resolutions that range from 3 to 17 cm (Table 1). A 15-cm^2 penetrometer cone with a 60° conical tip was used according to ASTM D3441 standards [*ASTM*, 2000]. The penetrometer is advanced vertically into the soil at a constant rate of 2 cm/s, though this rate must be reduced when compact layers are met. Inside the probe, two load cells independently measure the vertical stress against the conical tip and the side friction along the sleeve [*Lunne et al.*, 1997]. The vertical resolution of the tip stress (T) and sleeve stress (S) with the CPT probe are 4 cm and 17 cm, respectively. A pressure transducer in the cone is also used to measure the pore water pressure as the probe is pushed into the ground. Pore pressure is an indicator of the presence of clay and was used to correct T data. However, the pore pressure measured by the CPT tool was not used for this study as it is hard to quantitatively relate to hydraulic properties; this parameter is thus not further discussed in this paper. The SMR probe is composed of four electrodes that are connected directly behind the penetrometer [*Shinn et al.*, 1998]. The inner two rings are used to measure soil capacitance and the spacing between the two rings is 3 cm. The soil moisture probe operates at 100 MHz, thereby reducing the effects of the electrical conductivity of the soil on the measured dielectric constant. The instrument measures shifts in the resonance frequency signal as it passes through the soil that may be related empirically to soil moisture content or the dielectric constant (D). The bulk electrical resistivity (R) measurement employs the outer two rings of the SMR probe, that are spaced 9 cm apart, to apply the current and to measure the voltage drop (pole-pole configuration). The probe operates at a frequency of 1000 Hz to avoid soil polarization effects.

Table V.1 Original and transformed vertical resolutions and vertical support of measurements for direct-push parameters and hydraulic conductivity.

Parameter	Symbol	Original		Transformed	
		Vertical resolution (cm)	Support (cm)	Vertical resolution (cm)	Support (cm)
Geophysical parameters (CPT/SMR)					
Mechanical resistance - Tip stress	T	2.6 +/- 3.6	4	2	16
Mechanical friction - Sleeve stress	S	2.6 +/- 3.6	17	2	17
Dielectric constant	D	2.6 +/- 3.6	3	2	15
Bulk DC electrical resistivity	R	2.6 +/- 3.6	9	2	15
Hydraulic parameter					
Hydraulic conductivity	K	15	15	15	15

2.3 Direct-push well installation for colocated hydraulic and geophysical data

Based on CPT/SMR data obtained in real-time during sounding operations, 25 of the 53 direct-push soundings were converted into observation wells to carry out hydraulic tests (Figure 1c). Each observation well was installed into the same hole created by the sounding to obtain colocated hydraulic and direct-push data and thus reduce uncertainty in data analysis related to disparity in interval measurements. Each observation well was also installed immediately following the CPT/SMR sounding with the same direct-push rig (Geotech 605D) to reduce mobilization operation of the rig. The observation wells were installed using the protected screen installation procedure [*ASTM*, 2004], as described by *Paradis et al.* [2014]. This direct-push well installation procedure does not require the use of a sand-pack and the observation well screen is in direct contact with sediments, which is more suitable for hydraulic tests carried out over small intervals because it reduces skin effects on test data. Observation wells that are fully screened across the saturated zone were also installed to provide continuous profiles of K using multilevel slug tests. Screen length for all wells varied between 4.57 to 16.76 m. Direct-push wells were also thoroughly developed after their installation to unclog screens and surrounding sediments related to well installation operations [*Butler*, 1998; *Henebry and Robbins*, 2000]. Each well was then

developed using a pumping-surging operation at 0.5-m intervals until turbidity was no longer observed in the discharged water [*Paradis et al.*, 2011].

2.4 Hydraulic conductivity measurement from multilevel slug tests

K values used to establish relations with CPT/SMR data were obtained by multilevel slug tests in 8 of the fully screened direct-push wells. A multilevel slug test involves the use of packers to isolate a screened interval of a well to conduct a slug test [*Sellwood et al.*, 2005; *Zemansky and McElwee*, 2005; *Ross and McElwee*, 2007]. Slug tests were performed using a pneumatic method to induce an initial lowering of the water level [*Levy and Pannell*, 1991] and hydraulic responses were interpreted using the *Bouwer and Rice* [1976] method. A more detailed description of the test assembly, field and test analysis procedure is provided by *Paradis et al.* [2011]. Multilevel slug tests were made over 15-cm vertical intervals to reduce uncertainty related to scale disparity between hydraulic and direct-push data. A total of 280 intervals were tested and selected according to the range and occurrence of CPT/SMR responses. As depicted in Figure 2, the intervals with available *K* data in these 280 tests cover quite well the ranges of mechanical (*T* and *S*) and electrical (*R* and *D*) measurements of the CPT/SMR in the study area.

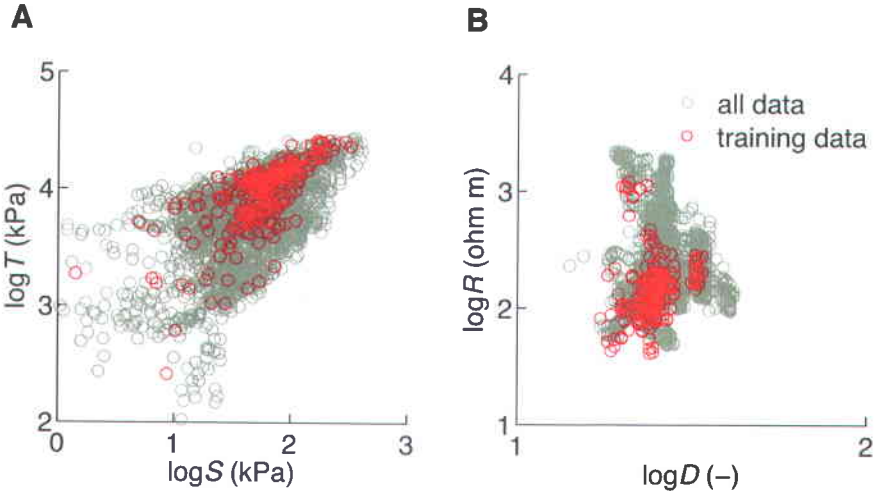


Figure V.2 Distribution of the training data set relative to the range of (a) mechanical: sleeve (*S*) and tip (*T*) stresses; and (b) electrical: dielectric constant (*D*) and resistivity (*R*) responses for the 16 direct-push soundings. A direct-push interval measurement of the training data set corresponds to a colocated measurement of hydraulic conductivity (*K*). Direct push measurements were transformed as summarized in Table V.1 so that their vertical resolution matches the one of *K* measurements. The number of such colocated hydro-geophysical measurements is 280. Parameter symbols are defined in Table V.1.

2.5 Data resampling and rescaling

Statistical techniques require that the different variables are measured at the same scale and on the same volume. However, this is not the case with CPT/SMR data where measurements are taken at a regular time interval but at a rate of penetration that is not necessarily constant. Therefore CPT/SMR data need to be interpolated on a regularly spaced grid from values measured at irregular intervals (Table 1). To accomplish this, we used a trapezoidal integration, a technique that considers all observations within an interval to estimate a single point. All observations on the irregular grid are first joined by a straight line and the value at the estimated point on the regular grid is calculated by summing the area under the original curve over the desired interval length [*Davis*, 1973]. CPT/SMR data were then resampled on a regular grid of 2 cm, a vertical resolution close to the original resolution and convenient to deal with a variety of measurement scales.

In addition, the vertical support of measurement for the different hydraulic and geophysical parameters are not identical (Table 1). In order to properly compare all measurements (T, S, R, D and K), the variations in their support need to be taken into account [*Isaaks and Srivastava*, 1989]. Hence, all the parameters with the smaller support were upscaled to the scale of the parameter with the larger support. Specifically, all CPT/SMR geophysical parameters were rescaled with a moving average to the same 15-cm interval that corresponds to the same K intervals over which hydraulic testing was carried out. Consequently, all hydro-geophysical measurements represent approximately the same vertical support of measurement.

2.6 Descriptive statistics of the hydro-geophysical training dataset

Since we are interested in defining relationships between K and CPT/SMR data, descriptive statistics presented here are for geophysical data available in the same intervals where K measurements are available, which together form the hydro-geophysical training dataset. Statistics for the hydro-geophysical training dataset are presented in Table 2 and histograms for each parameter are depicted in Figure 3. Since the range in parameter values for most parameter vary over a few orders of magnitude, a logarithmic transform was applied to make their distribution closer to a Gaussian distribution. Even though, histograms of the logarithm of geophysical parameter are all slightly asymmetric: negatively skewed for mechanical parameters ($\log S$ and $\log T$) and positively skewed for electrical parameters ($\log D$ and $\log R$). The distribution for $\log K$ is rather symmetrical and uniform (not normally distributed), which

suggests weak correlations with direct-push parameters that have different distributions, as indeed depicted by the scatter plots in Figure 3 and the correlation matrix shown in Table 3.

The correlation matrix in Table 3 shows no or low correlations between geophysical parameters together except for log$S$ and log$T$ where the correlation is relatively high (r=0.70) and may indicate redundancy in those two parameters. The correlations between log$K$ and geophysical parameters are generally significant but very low, except with log$S$ where the correlation is almost null. This indicates that the relations between K and direct-push parameters are not strait-forward because of the weak or non-linear relationships among the parameters that could exist. Standard-deviation for log$K$ is also at least twice the standard-deviation of geophysical parameters (even an order of magnitude greater with respect to log$D$), which suggests that K may be more sensitive to changes in sediment types than any direct-push parameter. Thus, the establishment of hydro-geophysical relations to predict K from direct-push data appears inherently difficult because of the nonlinearity in the hydro-geophysical data. In the next section, we describe the learning machine approach to handle this challenge.

Table V.2 Descriptive statistics for the distribution of measured values for direct-push parameters and hydraulic conductivity in the study area where they are both available. Parameter symbols are defined in Table V.1. The number of direct push measurements are for transformed values (Table V.1).

Parameter	Number	Mean	Median	Minimum	Maximum	Range	Standard Deviation	Skewness	Kurtosis
log$S$	280	1.74	1.76	0.15	2.51	2.36	0.31	-0.96	2.94
log$T$	280	3.94	4.00	2.41	4.42	2.01	0.29	-1.44	3.63
log$D$	280	1.39	1.39	1.24	1.53	0.29	0.06	0.36	0.85
log$R$	280	2.16	2.16	1.61	3.08	1.47	0.25	1.09	3.25
log$K$	280	-5.04	-5.05	-6.24	-3.92	2.32	0.57	0.09	-1.10

Table V.3 Correlation matrix showing the values of the correlation coefficient (r) for direct-push parameters and hydraulic conductivity. Parameter symbols are defined in Table V.1.

Parameter	log$S$	log$T$	log$D$	log$R$	log$K$
log$S$	1.00				
log$T$	0.70	1.00			
log$D$	-0.02	0.16	1.00		
log$R$	0.06	0.40	0.21	1.00	
log$K$	-0.02	0.29	0.33	0.52	1.00

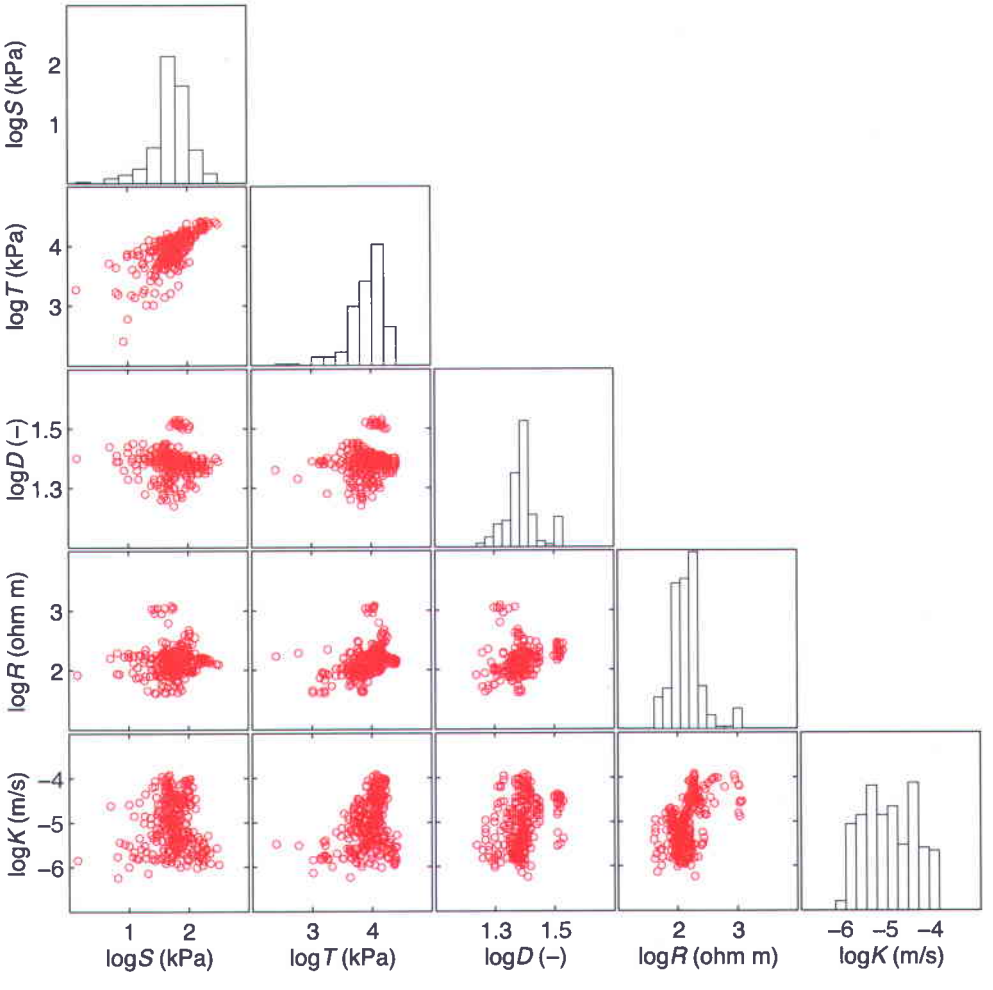


Figure V.3 Matrix scatter-plots with histograms for hydraulic conductivity (K) and CPT/SMR parameters (S: sleeve stress; T: tip stress; D: dielectric constant; R: electrical resistivity) for the hydro-geophysical training data set. The total number of colocated intervals is 280.

3 Outline of the learning machine approach

In this paper, the definition of relations to predict HF and K from CPT/SMR data is made through the training of a learning machine. This training is a three-step procedure, as proposed by *Lee and Datta-Gupta* [1999] for permeability prediction in complex petroleum reservoirs from borehole geophysics logs. This approach uses classification (unsupervised and supervised) and regression (supervised) algorithms that are applied with the training dataset to build good predictors for HF and K, as schematically illustrated in Figure 4. The first step in the training process is the definition of homogeneous groups using unsupervised fuzzy clustering with K and direct-push data of the training dataset. Thereafter those groups are referred to as hydrofacies (HF). Clustering involves the grouping of observations in such a way that observations in the same group (cluster) are more similar to each other than to those in other groups. The main rationale behind using clustering for this study is that various geological materials may have different hydro-geophysical behaviors and clustering allows the grouping of similar hydro-geophysical characteristics without any prior geological knowledge. This grouping can thus contribute to alleviate weak correlations among hydro-geophysical parameters and thus provide more accurate regression equations between K and direct-push data. The integration of K measurement in the clustering process is also helpful to define HFs that are hydrogeologically meaningful. Since clustering is only a tool applicable to make associations among datasets but cannot be used as a predictive tool, a RVM for classification is trained to recognize each HF (HF models) based on CPT/SMR responses, as defined by clustering. Finally, we develop a specific relation between K and direct-push (HG relation) data for each HF using RVMs for regression. Note that HF models and HG relations are defined independently (in parallel) using the same training dataset. The learning machine also includes a feedback path where the outputs obtained at any stage of the process can be reconsidered in a previous stage. Once the learning machine is optimally trained, the sequential predictive procedure illustrated in Figure 5 requires that for each set of new direct-push measurements (i.e., where no hydraulic tests have been carried out) HF is first identified using HF models, and then the associated hydro-geophysical relation is used to predict K. The remainder of this section presents the general algorithms used by the learning machine that involves fuzzy clustering and RVM for both regression and classification.

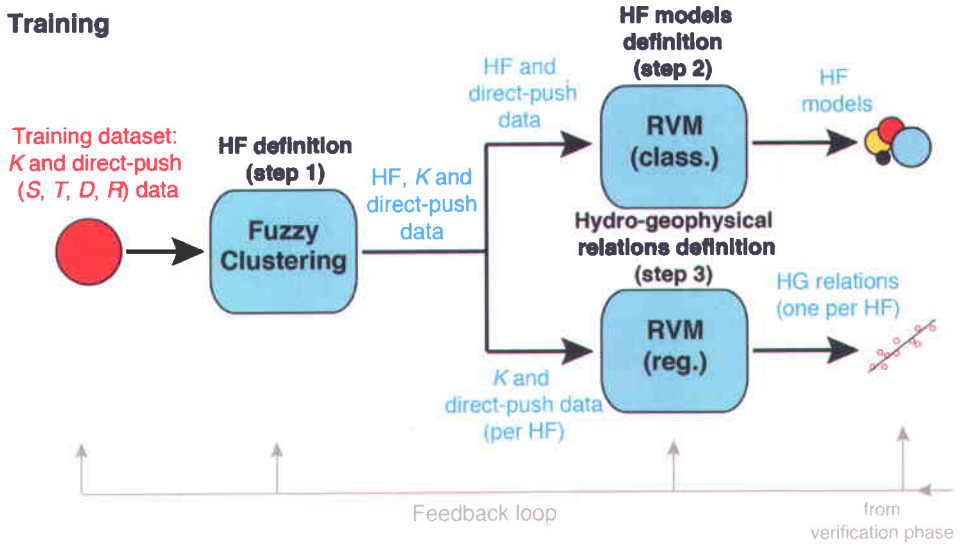


Figure V.4 General stages for the training of the learning machine to define hydrofacies (HF) models and hydro-geophysical relations from the hydro-geophysical training data set. RVM stands for relevance vector machine. Abbreviations: class. (classification), reg. (regression). Parameter symbols are defined in Table V.1. Note that HF models and HG relations are defined independently (in parallel).

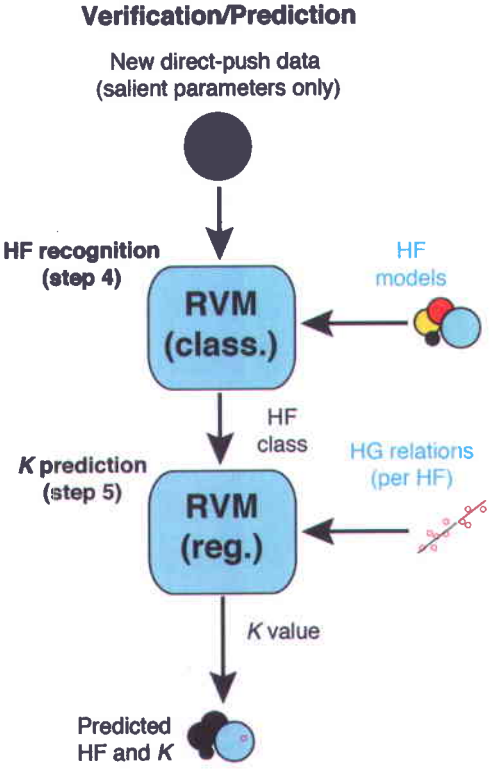


Figure V.5 General stages for the prediction of hydrofacies (HF) class and hydraulic conductivity (*K*) value from unseen direct-push data using the trained learning machine. This sequence is also used for testing the learning machine during the training phase. Parameter symbols are defined in Table V.1. Note that HF recognition and K prediction are carried out sequentially.

3.1 Fuzzy c-mean and Gustafson-Kessel clustering

The first computational stage for the training of the learning machine is to detect the underlying structure in the training dataset, upon which the learning machine is built on, in order to find associations among K measurements and the various direct-push data («HF definition» in Figure 4). There are many clustering methods available, and it is well known that the validity of the final clusters depends upon the method used to cluster the observations. Clustering methods can be classified according to whether the clusters are fuzzy (soft) or crisp (hard). Hard clustering means partitioning the data into specified mutually exclusive clusters whereas fuzzy clustering allows the observations to belong to several clusters simultaneously with different degrees of membership. Fuzzy representation is more natural than hard representation for hydro-geophysical datasets because the non-unique relationships that generally exist between hydraulic and geophysical parameters often result in overlapping HF.

The widely used fuzzy c-means (FCM) algorithm is here presented to illustrate the fuzzy clustering process used in this study. The FCM algorithm [*Dunn*, 1973; *Bezdek*, 1981] aims to find fuzzy partitioning of a given training set, by minimizing of the basic c-means objective function:

$$J(\mathbf{X};\mathbf{U},\mathbf{V}) = \sum_{i=1}^{c}\sum_{k=1}^{N} \mu_{ik}^{m} D_{ikA}^{2} \tag{1}$$

where $\mathbf{X} = [x_{kn}]$ is a $N \times n$ data matrix which contains N colocated observation intervals or simply observations, each having n parameters that correspond to K and direct-push parameters. $\mathbf{U} = [\mu_{ik}]$ is a $N \times c$ fuzzy partition matrix with c clusters, and represents cluster partial memberships of each observation x_k in $\mathbf{X}$. Fuzzy partition allows μ_{ik} to have real values in $[0,1]$. $m = (1,\infty)$ is the fuzziness weighting exponent, which determines the fuzziness of the resulting clusters. $\mathbf{V} = [v_c]$ is a vector of cluster centers, which have to be determined. $D_{ikA}^{2} = \|x_k - v_i\|_A^2 = (x_k - v_i)^T A (x_k - v_i)$ is a squared inner-product distance norm, which is a dissimilarity measure between the observations x_k and the centers v_i of each cluster.

The FCM algorithm uses the standard Euclidean distance norm with a common norm inducing matrix **A** for each cluster. The $n \times n$ norm inducing matrix **A** is a diagonal matrix that accounts for the variance of each parameter *n*. Hence the FCM algorithm induces hyperspherical clusters and can induce clusters with the same shape and orientation.

For a predefined number of clusters, the minimization of Eq. (1) that is carried out with the partition matrix and the centers identifies the structure in **X**. The generic optimization scheme involves a sequence of iterations, in which the values of the partition matrix is successively updated:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} \left(D_{ikA} / D_{jkA} \right)^{2/(m-1)}}, \quad 1 \le i \le c, \quad 1 \le k \le N \tag{2}$$

and the centers:

$$v_i = \frac{\sum_{k=1}^{N} \mu_{ik}^m x_k}{\sum_{k=1}^{N} \mu_{ik}^m}, \quad 1 \le i \le c \tag{3}$$

This iterative process ends when the difference between the fuzzy partition matrices in the following iterations is lower than a maximum termination tolerance value.

Other algorithms allow the definition of clusters having variable shapes or density. For instance, *Gustafson and Kessel* [1979] extended the standard FCM algorithm by employing an adaptive distance norm to detect clusters of different geometrical shapes. In this approach, each cluster has its own norm-inducing matrix A_i and D_{ikA}^2 in Eq. (1) is replaced by

$$D_{ikA_i}^2 = \left\| x_k - v_i \right\|_{A_i}^2 = \left(x_k - v_i \right)^T A_i \left(x_k - v_i \right),$$

which is the squared inner-product distance norm for the Gustafson and Kessel (GK) algorithm, wherein

$$A_i = \left[\rho_i \det\left(F_i\right)\right]^{1/n} F_i^{-1}$$

where ρ_i is the norm of A_i. This distance norm is known as the squared Mahalanobis distance. The matrices A_i are then used as optimization variables in the c-means functional, thus allowing each cluster to adapt the distance norm to the local topological structure of the data, in particular to the linear correlations among parameters. F_i is the fuzzy covariance matrix of the i-th cluster defined by:

$$F_i = \frac{\sum_{k=1}^{N} \mu_{ik}^m \left(x_k - v_i\right)\left(x_k - v_i\right)^T}{\sum_{k=1}^{N} \mu_{ik}^m}, \quad 1 \le i \le c \tag{4}$$

and it is updated along with the partition matrix **U** and the centers **V** in the iterative process leading to the minimization of Eq. (1).

3.2 Relevance vector machine for regression

With the HF labels and corresponding K, and the direct-push data resulting from the clustering process, predictive models and relations are independently defined, as illustrated in Figure 4. For the definition of hydro-geophysical relations, separate RVMs for regression are trained for each HF using corresponding K and direct-push data. The main idea behind RVM is to perform a linear regression in a high dimension feature space, through a kernel function, which returns a nonlinear regression in the original input space. The weights of the regression are solved into a Bayesian framework and can be used for regression and classification problems, as described below.

Consider a training dataset of N input-target pairs $\{g_k, K_k\}_{k=1}^N$ with the geophysical data vectors g_k as the input along with colocated K_k measurements as the target. We note that g_k can include many geophysical parameters, whereas K_k represents only one hydraulic parameter. For the prediction with many hydraulic parameters with the same set of geophysical data, separate regressions should be defined for each

hydraulic parameter of interest. For regression, the targets K_k are generally assumed to be some noisy realization of an underlying functional relationship $f(\mathbf{g})$ that we wish to estimate so that

$$\mathbf{K} = f(\mathbf{g};\mathbf{w}) + \varepsilon_k \tag{5}$$

where ε_k are the result of an additive noise process and $\mathbf{w}$ is a vector of adjustable weights. One interesting class of candidate functions for $f(\mathbf{g};\mathbf{w})$ is given by

$$f(\mathbf{g};\mathbf{w}) = \sum_{k=1}^{N} w_k \phi_k(\mathbf{g}) = \mathbf{w}^T \boldsymbol{\phi}(\mathbf{g}) \tag{6}$$

which represents a linearly-weighted sum of N nonlinear fixed basis functions denoted by $\boldsymbol{\phi}(\mathbf{g}) = \left(\phi_1(\mathbf{g}), \phi_2(\mathbf{g}), ..., \phi_N(\mathbf{g})\right)^T$. Models of the type of Eq. (6) are known as linear models since the function $f(\mathbf{g};\mathbf{w})$ is a linear function of the parameters $\mathbf{w} = \left(w_1, w_2, ..., w_N\right)^T$. For many practical regression problems, nonlinear basis functions are generally employed with Eq. (6) and, indeed, can be very complex and flexible if N is relatively large.

Classical treatment of a regression problem seeks a point estimate of the unknown weight vector $\mathbf{w}$. By contrast, in a Bayesian approach we characterize the uncertainty in $\mathbf{w}$ through a probability distribution $p(\mathbf{w})$. Observations of data points modify this distribution by virtues of Bayes' theorem, with the effect of the data being mediated through the likelihood function. In the context of RVM for regression, the class of candidate functions for $f(\mathbf{g};\mathbf{w})$ is of the form:

$$f(\mathbf{g};\mathbf{w}) = \sum_{k=1}^{N} w_k L(\mathbf{g}, \mathbf{g}_k) + w_0 \tag{7}$$

where $\mathbf{w}$ are the weights or relevance vector coefficients, w_0 is bias and $L(\mathbf{g},\mathbf{g}_k)$ is a kernel function. The kernel function is used to introduce nonlinearity in the mapping function. The kernel function defines a set of nonlinear fixed basis functions $\phi_k(\mathbf{g}) = L(\mathbf{g},\mathbf{g}_k)$ that is centered on each of the N training data points g_k. Selection of a suitable kernel type that best suits the data is an important step.

From a probabilistic viewpoint, the targets $\mathbf{K}$ in Eq. (5) are assumed to be independent and ε_k a Gaussian noise with zero mean and noise variance σ^2. Due to the assumption of independence of the $\mathbf{K}$, Eq. (5) can then be rewritten as the likelihood function

$$p(\mathbf{K}|\mathbf{w},\sigma^2) = (2\pi\sigma^2)^{-N/2}\exp\left\{-\frac{1}{2\sigma^2}\|\mathbf{K}-\Phi\mathbf{w}\|^2\right\} \tag{8}$$

where Φ is the design matrix that contains the response of all basis functions ϕ_k to the geophysical data g_k, with

$$\Phi = \left[\phi(g_1),\phi(g_2),...,\phi(g_N)\right]^T,$$

wherein,

$$\phi(g_k) = \left[1, L(g_k,g_1), L(g_k,g_2)..., L(g_k,g_N)\right]^T.$$

From classical statistics, the maximum likelihood estimation that allows point estimate of $\mathbf{w}$ and σ^2 from Eq. (5) generally leads to over-fitting due to large number of weight generally used for the regression. In the Bayesian approach proposed by *Tipping* [2001], a zero-mean Gaussian non-informative prior distribution is imposed over the model weights $\mathbf{w}$:

$$p(\mathbf{w}|\alpha) = \prod_{i=0}^{N} N(w_i|0,\alpha_i^{-1}) \tag{9}$$

with $\boldsymbol{\alpha}$ a vector of N+1 hyperparameters. Note that there is an individual hyperparameters associated independently with every weight, moderating the strength of the prior.

To complete the specification of this hierarchical prior, hyper-priors are defined over $\boldsymbol{\alpha}$, as well over the noise variance σ^2. These quantities are scale parameters, and suitable priors are Gamma distributions:

$$p(\alpha) = \prod_{i=0}^{N} Gamma(\alpha_i|a,b)$$

$$p(\beta) = Gamma(\beta|c,d)$$

with $\beta \equiv \sigma^{-2}$ and where

$$Gamma(\alpha|a,b) = \Gamma(a)^{-1} b^a \alpha^{a-1} e^{-b\alpha} \tag{10}$$

with $\Gamma(a)$ the gamma function. To make these priors non-informative, scale parameters (a, b, c and d) are set to zero.

The prior distribution expresses the uncertainty in $\mathbf{w}$. By virtue of Bayes' theorem this non-informative prior is later modified through the likelihood function to reflect the observed dataset of N input-target pairs. The prior is governed by a hyper-parameter α_k associated with each weight (iteratively estimated from the data), moderating the prior strength. The individual hyper-parameters control groups of weights and their associated basis functions $\varphi(\mathbf{g})$, which are associated with each geophysical parameter $\mathbf{g}$. The association of a hyper-parameter with each weight $\mathbf{w}$ is the key feature of the RVM that is responsible for its sparsity properties. Hence, the minimization of the number of hyper-parameters that control groups of weights, instead of each individual weight, leads to smooth regression models with less over-fitting.

Having defined the prior, Bayesian inference proceeds by computing, from Bayes' rule, the posterior over all unknowns given the data:

$$p(\mathbf{w},\alpha,\sigma^2|\mathbf{K}) = p(\mathbf{w}|\mathbf{K},\alpha,\sigma^2)p(\alpha,\sigma^2|\mathbf{K}) \tag{11}$$

The posterior distribution over the weights **w** that can be computed analytically is thus given by:

$$\begin{aligned} p(\mathbf{w}|\mathbf{K},\alpha,\sigma^2) &= \frac{p(\mathbf{K}|\mathbf{w},\sigma^2)p(\mathbf{w}|\alpha)}{p(\mathbf{K}|\alpha,\sigma^2)} \\ &= (2\pi)^{-(N+1)/2}|\Sigma|^{-1/2}\exp\left\{-\frac{1}{2}(\mathbf{w}-\mu)^{\mathrm{T}}\Sigma^{-1}(\mathbf{w}-\mu)\right\} \end{aligned} \tag{12}$$

where the posterior covariance and mean are respectively:

$$\boldsymbol{\Sigma} = (\mathbf{A}+\sigma^{-2}\boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1},$$

$$\mu = \sigma^{-2}\boldsymbol{\Sigma}\boldsymbol{\Phi}^T\mathbf{K}$$

with $\mathbf{A} = diag(\alpha_1,\alpha_2,...,\alpha_m)$.

The hyper-parameter posterior $p(\alpha,\sigma^2|\mathbf{K})$ is on the other hand represented by a delta-function at its mode (i.e., most probable values for **α** and σ^2), which leads to the approximation $p(\alpha,\sigma^2|\mathbf{K}) \propto p(\mathbf{t}|\alpha,\sigma^2)p(\alpha)p(\sigma^2)$. For the case of uniform hyper-priors (a=b=c=d=0), the training process becomes the search for most probable posterior hyper-parameter **α** and σ^2 by maximizing only the marginal likelihood

$$p(\mathbf{K}|\alpha,\sigma^2) = (2\pi)^{-N/2}\left|\sigma^2\mathbf{I}+\boldsymbol{\Phi}\mathbf{A}^{-1}\boldsymbol{\Phi}^T\right|^{1/2}\exp\left\{-\frac{1}{2}\mathbf{K}^T(\sigma^2\mathbf{I}+\boldsymbol{\Phi}\mathbf{A}^{-1}\boldsymbol{\Phi}^T)^{-1}\mathbf{K}\right\} \tag{13}$$

The **α** and σ^2 which maximize the marginal likelihood are then found iteratively by setting **α** and σ^2 to initial values, finding values for $\boldsymbol{\Sigma}$ and μ, using these to calculate new estimates **α** and σ^2 and repeating this process until a convergence criteria is met. In practice, during the iterative process it is found that

many of the α_i tend to infinity, which results in that several weights w_i become zero to achieve sparsity. The number of observations, RV ($RV < N$), that have nonzero weights are the relevance vectors.

Predictions are made on the basis of the posterior distribution over the weights $\mathbf{w}$, conditioned on maximizing values obtained at the convergence of the hyper-parameters estimation procedure. Having learned from the training values the hydraulic data $\mathbf{K}$, we now make a prediction of hydraulic parameter value at a new location $\mathbf{K}^*$ given new geophysical data $\mathbf{g}^*$, where the posterior from the training process is considered the prior for the prediction process. Therefore the new hydraulic data $\mathbf{K}^*$ for new geophysical data $\mathbf{g}^*$ are predicted according to

$$\mathbf{K}^* = \boldsymbol{\mu}^T \Phi(\mathbf{g}^*) \tag{14}$$

where $\boldsymbol{\mu}$ is a vector of the mean weights of the relevance vectors and $\Phi(\mathbf{g}^*)$ is the design matrix evaluated for the new geophysical data $\mathbf{g}^*$ with the best suited kernel to the data adjusted during the regression training process.

3.3 Relevance vector machine for classification

In parallel with the definition of hydro-geophysical relations, HF models are also developed with HF labels and direct-push data using a RVM for classification, as illustrated in Figure 4. RVM for classification follows an essentially identical framework as previously detailed for RVM regression, but using Bernoulli likelihood and a sigmoid link function to account for the change in the target quantities (continuous versus class data).

Considering a two-class problem with a dataset of N input-target pairs with the geophysical data vectors $\mathbf{g}$ and corresponding hydrofacies (class) **HF** with $HF_k \in \{0,1\}$. Applying the logistic sigmoid link function $\sigma(f) = 1/(1+\exp^{-f})$ to $f(\mathbf{g};\mathbf{w})$ and adopting the Bernoulli distribution for $P(HF|\mathbf{g})$, the likelihood function is expressed as

$$P(\mathbf{HF}|\mathbf{w}) = \prod_{k=1}^{N} \sigma\{f(\mathbf{g}_k;\mathbf{w})\}^{\mathbf{HF}_k} \left[1-\sigma\{f(\mathbf{g}_k;\mathbf{w})\}\right]^{1-\mathbf{HF}_k} \tag{15}$$

Note that there is no noise variance for the classification case. However, unlike the regression case, the weights cannot be evaluated analytically, and closed-form expressions for either the weight posterior $p(\mathbf{w}|\mathbf{HF},\alpha)$ or the marginal likelihood $P(\mathrm{HF}|\alpha)$ are denied. An approximation procedure based on Laplace's method has then to be used [*MacKay*, 1992].

For prediction, the probability of membership to one of the hydrofacies **HF*** for new geophysical data **g*** is thus given by

$$\mathbf{HF}^* = \sigma\left\{\boldsymbol{\mu}^T \Phi(\mathbf{g}^*)\right\} \tag{16}$$

where $\boldsymbol{\mu}$ is a vector of the mean weights of the relevance vectors and $\Phi(\mathbf{g}^*)$ is the design matrix evaluated for the new geophysical data **g*** with the best suited kernel to the data adjusted during the classification training process (Eq. 10).

For a model with more than two HFs, a multi-class classification approach for which a series of binary classifications is performed should be adopted. Two of the common methods for multi-class classification include the one-against-all and the all-against-all techniques. The all-against-all approach is adopted here because it generally produces better classification performance over the one-against-all approach [*Allwein et al.*, 2000; *Hsu and Lin*, 2002]. In the all-against-all approach, each HF is compared to each other HF [*Hastie and Tibshirani*, 1998] and a binary model is built to discriminate between each pair of HFs, while discarding the rest of the HFs. This requires building $c(c-1)/2$ binary HF models. When testing new geophysical data **g*** that do not have hydraulic information, a voting is performed among the various binary HF models and the HF with the maximum number of vote wins, and this HF label is assigned to the new geophysical data **g***.

4 Hydrofacies and hydraulic conductivity estimation from CPT/SMR data

In this section, we apply and test the learning machine approach proposed in Section 3 to the training dataset described in Section 2 to identify HF and predict K from CPT/SMR data for the St-Lambert site. The training and testing of the learning machine involves the following steps:

1. HF definition (Step 1 in Figure 4): We first select the most relevant CPT/SMR parameters to predict K and then extract HF structures in the training dataset using clustering algorithms. All observations of the training dataset are used in this procedure.
2. HF models definition (Step 2 in Figure 4): In this step we train a RVM for classification to recognize the HFs defined in Step 1 using HF labels and data of the most relevant CPT/SMR parameters. A 10-fold cross-validation procedure is followed to select a kernel type and adjust its window width. This process leads to the definition of site-specific HF models.
3. Hydro-geophysical relations definition (Step 3 in Figure 4): For each HF defined in Step 1, we train a RVM for regression using colocated K and direct-push data to define a specific hydro-geophysical relation for each HF. A leave-one-out (LOO) cross-validation procedure is followed here to select the kernel and window width of each regression equation.
4. Verification of HF and K predictions (Steps 4 and 5 in Figure 5): Finally, the performance of the trained learning machine to identify HFs and estimate K is assessed using the same sequential prediction procedure used to predict hydraulic information from new direct-push data that do not have hydraulic information. In order to provide statistics on the estimation errors for HF and K, the trained learning machine is applied several times with different randomly selected training and testing datasets.

More details on the application of the learning machine approach are given in the next sections.

4.1 Fuzzy clustering for hydrofacies definition

The first stage of the training process is the definition of HF with clustering (Step 1 in Figure 4). The clustering was carried out with the Matlab Fuzzy Clustering and Data Analysis Toolbox [*Balasko et al.*, 2005] with all the 280 colocated K and CPT/SMR measurements of the training dataset. To avoid the

largest-valued parameters to bias the clustering, the logarithm of the original data for each parameter were normalized to a common scale using their respective range. Every clustering simulation was also initialized with random seeds. Experiments were done with different seeds not reported here and those results were very similar to the ones given in this paper.

Due to the non-uniqueness that generally exists between hydraulic and geophysical parameters [*Rubin and Hubbard*, 2005], the accuracy of predicted hydraulic information from direct-push data may suffer from a loss of information. This loss of information causes an overlapping between HFs defined in the geophysical space that may preclude reliable estimates of hydraulic conductivity. It should be remembered that HFs are defined here in the hydro-geophysical space with K and direct-push data, but the prediction of HF and K is made in the geophysical space only using direct-push data. As suggested in Figure 5, an improper HF recognition during the prediction process will indeed result in the selection of the wrong hydro-geophysical relation to predict K. Thereby, a systematic procedure was applied to search for the combination of CPT/SMR parameters that produce the least HF overlapping in the geophysical space to obtain the best predictive capability.

Moreover, as an unsupervised learning technique, the clustering algorithm requires a predefined number of clusters. The correct choice of cluster numbers is often ambiguous, with interpretations depending on data structure and clustering purposes. Furthermore, the optimal choice of cluster numbers has to strike a balance between maximum compression of the data, using a single cluster, and maximum accuracy, by assigning each data point to its own cluster. One may estimate an optimal number of clusters using various statistical approaches or expert opinion. In this study the optimal number of HFs was included in the search procedure.

Similarity measured between two observations drawn from the same parameters space is also fundamental to the definition of a cluster. Similarity is usually measured by a distance function defined on pairs of observations. The distance measure must be chosen carefully following the data structure and the parameters characteristics (e.g., correlation). For instance, the FCM and GK clustering algorithms that use Euclidean and squared Mahalanobis distances, respectively, were also tested to find the similarity measure that minimizes HF overlapping.

To assess the degree of HF overlapping for each examined combination, we define the misclassification error (ME_{olap}) as the percentage of misclassified observations resulting from the projection of HFs from the hydro-geophysical to the geophysical space:

$$ME_{olap} = \frac{100}{N} \sum_{i=1}^{c} \sum_{k=1}^{N} (HF_{hg} \neq HF_{g})_{ik} \tag{17}$$

where c is the number of HF, N is the number of observations, and HF_{hg} and HF_g are HF labels in the hydro-geophysical and geophysical spaces, respectively. A lower value of ME_{olap} indicates less overlapping between HFs and a better recognition potential using direct-push data.

To calculate the ME_{olap}, we first extract coordinate centers of each HF clustered in the hydro-geophysical space. Then we project HF into the geophysical space by defining new HFs using only geophysical parameters and by imposing previously extracted HF centers to the new clustering. Note that we are not re-running the clustering here, only the coordinate centers of the direct-push parameters are used and those pertained to K are discarded. After the fuzzy memberships of each observation are transformed to integer numbers, by using the HF label with the maximum fuzzy membership, HF_{hg} and HF_g labels for each observation obtained in the two spaces are compared together. Observations with different HF labels are thus considered as incorrectly classified and the ME_{olap} is calculated using Eq. (17).

The results of the exhaustive search procedure for direct-push parameters, HF number and similarity measure are presented in Figures 6a-b. During this procedure, all combinations of geophysical parameters with various numbers of HFs were individually clustered using the FCM and GK algorithms to find HF structures with the lowest ME_{olap} value (less HF overlapping). A total of 168 combinations were thus examined with the number of HFs varied between 2 and 7, with K data included in all clustering experiments. The fuzziness weighting exponent m was fixed to a value of 2 for all experiments. Several observations can be made from Figures 6a-b. First, HF overlapping generally decreases with the number of geophysical parameters used to define HF, as expressed by decreasing ME_{olap} values. For instance, ME for a subset using only one geophysical parameter is up to 66% (e.g., SK with 4 HFs in Figure 6b), while HF overlapping is as low as 13% for subsets using three or four direct-push parameters (e.g., TDRK with 4 HFs in Figure 6b).

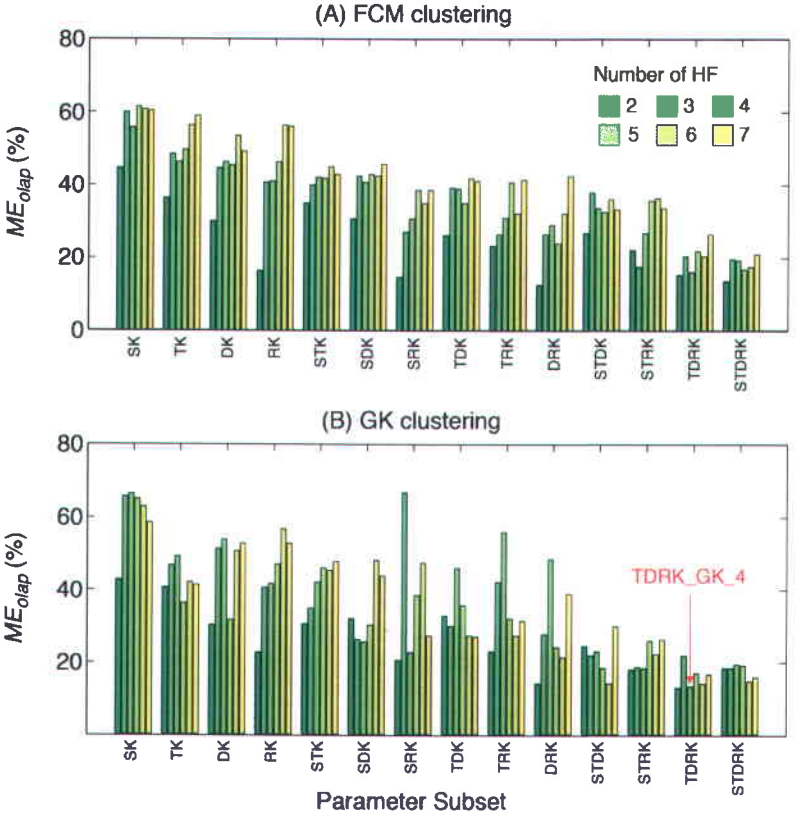


Figure V.6 Misclassification error (ME_{olap}) for the projection of the hydrofacies (HF) defined in the hydro-geophysical space by (a) fuzzy c-mean (FCM) and (b) Gustafson-Kessel (GK) clustering into the geophysical space for different numbers of HF and geophysical parameters set. The arrow indicates the GK clustering of TDRK data with 4 classes of HF that was selected for Step 1 of RVM training (see Figure V.4).

Figures 6a-b also show that the GK algorithm generally provides better performance than FCM clustering with an average ME_{olap} of 35% and 38%, respectively. In fact, the Euclidian distance that is used by the FCM algorithm generally works well for datasets with compact and isolated hypersherical clusters [*Mao and Jain*, 1996], which is unlike our dataset that is continuously distributed without sharp transitions in values. Also, strong linear correlation among parameters, such as between S and T in Table 3, are likely to distort distance measures in the Euclidian space. This distortion is however alleviated with squared Mahalanobis distance used in the GK algorithm, where different weights are assigned to each data based on their variances and co-variances linear correlations, and follows natural correlations among parameters within each HF.

According to Figures 6a-b, the clustering experiment that provides less HF overlapping in the geophysical space is the subset with direct-push parameters *T*, *D* and *R* using the GK algorithm with 4 HF. This subset is referred thereafter as the TDRK_GK_4 subset, which indicates the set of parameters (TDRK), the clustering algorithm (GK) and the number of HFs (4), respectively. We note that the same direct-push parameter subset with 2 or 6 HF with the GK algorithm also has a small ME_{olap} value. For the sake of hydrogeological interpretation, we selected the combination with 4 HFs because it strikes a balance between maximal compression (fewer HFs) and accuracy (many HFs) and roughly separates the range of *K* into 4 classes of half an order of magnitude each. Note that principal component analysis (PCA) re-expression of the dataset and features construction as product of original parameters were also tested [e.g., *Guyon and Elisseeff*, 2003]. Those experiments did not however provide lower ME_{olap} values than TDRK_GK_4, likely due to the loss of information caused by the filtering process and dimensionality reduction.

To further understand the link between *K* and direct-push parameters, we carried out a factor analysis using the hydro-geophysical dataset, as illustrated in Figure 7. In a factor analysis, the measured parameters depend on a smaller number of unobserved (latent) factors [*Davis*, 1973; *Harman*, 1976]. Since each factor may affect several variables in common, they are known as common factors. Each parameter is assumed to depend on a linear combination of the common factors, and the coefficients are known as loadings. For this study, the factor analysis was carried out using a maximum likelihood approach [*Jöreskog*, 1967].

Two important observations can be made from Figure 7. First, we see a high loading for mechanical parameters *S* and *T* on the factor 1 axis, whereas loading for *K* and electrical parameters *D* and *R* are more important on factor 2. This suggests that two independent factors are underlying hydro-geophysical relationships for the study training dataset. Based on the nature of each parameter, this suggests that factor 1 may be related to the structure of the sediments, while factor 2 relates to the pore structure. We note that according to the latent root criterion only two factors can be extracted from the training dataset and those factors explain 71% of the total variance. We should also note that factors rotated by a promax oblique scheme provided similar loading as the varimax orthogonal scheme illustrated in Figure 7, which indicates that the two main factors are fairly independent (orthogonal).

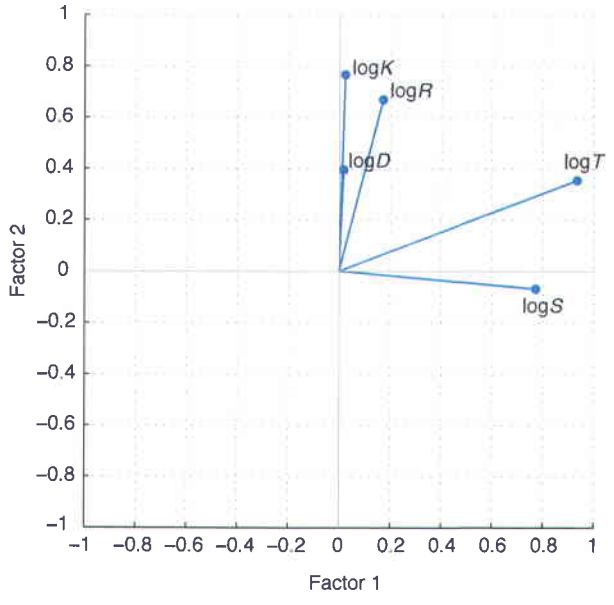


Figure V.7 Factor analysis for the hydro-geophysical training data set. The two-dimensional diagrams of directional vectors represents loadings on individual hydro-geophysical parameters projected against the two main factors rotated using the varimax scheme [*Harman*, 1976].

Another important observation from Figure 7 is that some CPT/SMR parameters are better suited to resolve *K*. For instance, high *T*, *D* and *R* loadings on axis 2 means that those parameters have a good potential of resolution of *K*, which also has a high loading on axis 2. For perfectly correlated geophysical parameters with *K*, only one geophysical parameter would be necessary to predict *K* from geophysical parameters. For perfectly correlated parameters the reduction in space dimension does not lead to overlapping. However, with imperfectly correlated parameters, and possibly heterogeneous correlation varying according to the geological materials, even a weak correlation will stretch the cloud of observations in the direction of the correlation. The larger the correlation, or anti-correlation, the stronger is the stretching. When using a large number of geophysical parameters, the distortion of the cloud will be larger and the clusters will be more distinct from each other and consequently less overlapping will result. Thus using more geophysical parameters with varying degrees of correlation with *K* increases the overall degree of correlation with *K* and thus the potential of geophysical parameters to resolve *K*. In Figure 7 we see that the only parameter that does not share variance with *K* is *S* (this can also be seen in the correlation matrix in Table 3), which explains why this parameter was not retained in the systematic search procedure with geophysical parameters. From Figure 6 we also observe that for the same number of parameters, subsets with S generally present the higher ME_{olap}.

Data distribution for each HF and parameter for the TDRK_GK_4 subset are also illustrated in Figure 8, where we assigned observations to the highest HF membership. The median and range of values for K and each retained direct-push parameter are fairly distinct between HFs with only a few outliers. Particularly, the median values for K gradually increase from HF1 to HF4, with slight overlaps between HFs that may be attributed to the complexity of the hydro-geophysical responses and to the transitional nature of the littoral depositional environment. Moreover, each HF presents distinct profiles of K and direct-push parameters as expected from the various sediments composing the aquifer that may present different hydro-geophysical responses. For instance, HF4 has the highest median values for T, D, R and K, whereas HF1 has the lowest values for the same parameters. As depicted in Table 4, clustering also results in better linear correlations between parameters with correlation values up to 0.90 between geophysical parameters together and up to 0.66 between K and geophysical parameters. This correlation information can indicate which geophysical parameters are best suited to predict K within each HF.

In summary, the number and the choice of geophysical parameters appear to have an important impact on the potential for HF recognition from CPT/SMR data. Our interpretation is that using more parameters allows a better distinction of the various HFs, which should reflect geological materials, by reducing non-uniqueness among hydraulic and geophysical parameters. And the process of selection of the number of HFs and similarity measure allows extracting those distinct HF structures in the training dataset.

Table V.4 Correlations matrix showing the values of the correlation coefficient (r) for direct-push parameters and hydraulic conductivity for each hydrofacies of the TDRK_GK_4 subset. Parameter symbols are defined in Table V.1.

Parameter	log$T$	log$D$	log$R$	log$K$
Hydrofacies 1				
log$T$	1.00			
log$D$	0.21	1.00		
log$R$	0.22	-0.44	1.00	
log$K$	0.22	0.23	0.01	1.00
Hydrofacies 2				
log$T$	1.00			
log$D$	-0.28	1.00		
log$R$	0.07	0.75	1.00	
log$K$	-0.60	0.48	0.18	1.00
Hydrofacies 3				
log$T$	1.00			
log$D$	0.78	1.00		
log$R$	0.77	0.90	1.00	
log$K$	0.57	0.66	0.42	1.00
Hydrofacies 4				
log$T$	1.00			
log$D$	-0.07	1.00		
log$R$	-0.10	-0.77	1.00	
log$K$	0.38	-0.01	-0.11	1.00

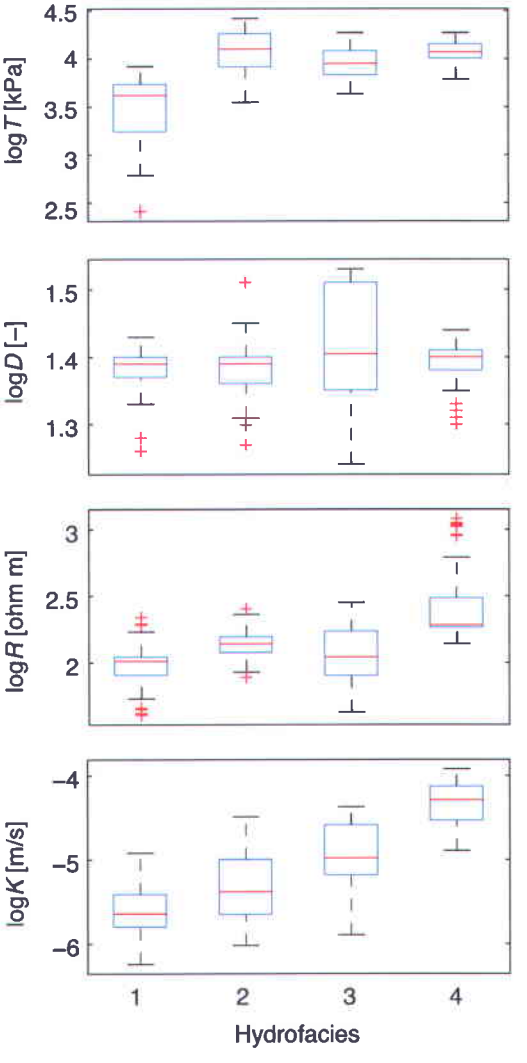


Figure V.8 Data distribution for each of the four hydrofacies resulting from the clustering experiment with the best predictive capability (TDRK_GK_4 subset; see Figure V.6). This clustering uses tip stress (*T*), dielectric constant (*D*) and resistivity (*R*) along with hydraulic conductivity (*K*).

4.2 Multi-class RVM training for hydrofacies models definition

To build predictive HF models for our study site, a multi-class RVM is trained to recognize HFs of the TDRK_GK_4 subset using data for parameter *T*, *D* and *R* (Step 2 in Figure 4). Note that HF data used for the training are integer numbers obtained from the transformation of fuzzy memberships resulting from clustering. Figures 9a-b show graphs used for the selection of the optimal kernel window length of the multi-class RVM using a Gaussian kernel function. To find the optimal kernel window length, the classifier performance (Figure 9a) and complexity (Figure 9b) are assessed for various kernel functions

and kernel window lengths using training and testing datasets. The classifier performance is defined here by the misclassification error (ME_{RVM}) that is evaluated using an equation similar to Eq. (17):

$$ME_{RVM} = \frac{100}{N}\sum_{i=1}^{c}\sum_{k=1}^{N}(HF_{clust} \neq HF_{RVM})_{ik} \tag{18}$$

where HF_{clust} and HF_{RVM} are HF labels from clustering and RVM classification, respectively. Note that HF_{clust} are integer numbers obtained from the transformation of fuzzy memberships, whereas HF_{RVM} are integer numbers resulting from the multi-class voting process. A lower value of ME_{RVM} indicates a better HF predictive capability of the RVM. The model structural complexity (sparsity) of the classifier is expressed by the total number of relevance vectors (RV) used by the RVM classifier:

$$RV = \sum_{i=1}^{k} RV_i \tag{19}$$

where RV_i is the number of the relevance vectors per HF i. A lower RV value produces a smoother solution.

Thus, for a given kernel function and a kernel window length, relevance vectors are first determined with the procedure in Section 3.3 using the training dataset. Then, HFs are predicted using the testing dataset with Eq. (16) and previous relevance vectors. Finally, performance and sparsity of the classifier are assessed using Eq. (18) and Eq. (19), respectively. This procedure was repeated for different kernel window lengths to produce Figures 9a-b with a Gauss kernel function. Note that we are using normalized parameters and kernel window widths for different parameters can thus be plotted on the same axis in Figures 9a-b. Different kernel functions can also be tested. The supervised classification with RVM was carried out with the SPARSEBAYES Matlab Toolbox [*Tipping and Faul*, 2003; *Tipping*, 2009] using Bernoulli likelihood.

To avoid bias in the selection of training and testing datasets, a 10-fold cross-validation procedure [*Geisser*, 1975] was followed to produce Figures 9a-b. Cross-validation with partial data splitting is

reported to be a robust procedure for model selection of classification problems [*Arlot and Celisse*, 2010]. With this procedure the entire training dataset is split randomly into ten groups of similar size and each group is used in turn as a testing set, while the other nine groups are used together to form a single training set. The average value of the ten experiments for each kernel window length of a given kernel function is then used to plot ME_{RVM} and RV curves. We note in Figures 9a-b that statistics for ME_{RVM} and RV are also provided for each of the six binary models used by the one-against-all approach for the training of the multi-class RVM.

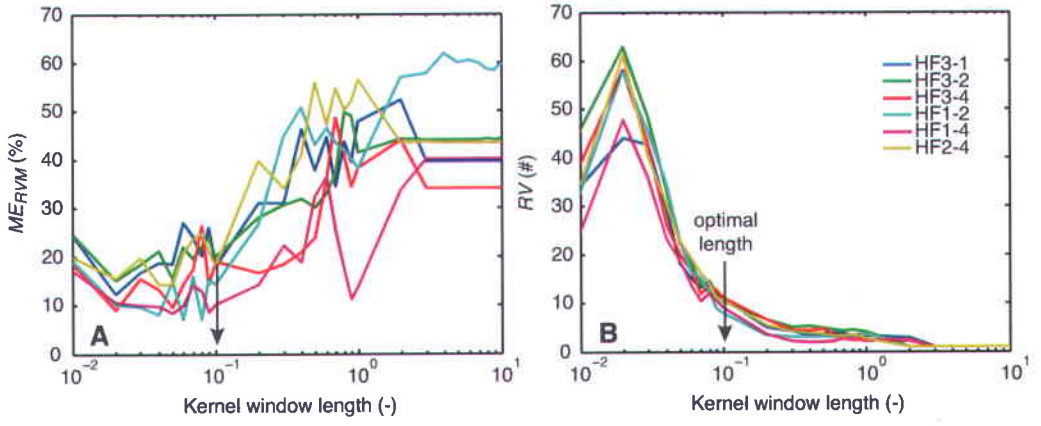


Figure V.9 Graph of (a) misclassification error (ME_{RVM}), and (b) number of relevance vector (RV) used for the selection of the optimal kernel window length of the relevance vector machine for classification. Each curve represents a pair of hydrofacies used for the all-against-all multi-class classification. Results are for a Gauss kernel function.

According to Figures 9a-b, we selected a unique kernel window length of 0.1 as the optimal value for best predictive capability of the RVM classifier with a Gaussian kernel function. Note that other kernel functions were also tested, but those experiments did not provide better predictive capabilities as verified by the procedure presented in Section 4.4. The selection of the kernel window length followed the elbow-criterion, to find the optimal number of relevance vectors that strike a balance between over-fitting and over-smoothing the testing data, while obtaining a RVM classifier with low ME_{RVM} value. For instance, using a large number of relevance vectors generally leads to good classifier performance with training data, but to poor generalization capability when used with testing data because the over-fitted model describes the random error instead of the underlying relationship [*Tetko et al.*, 1995]. Using very few relevance vectors leads to poor classifier performance with both training and testing data due to the over-smoothing of the underlying relationship.

4.3 RVM regression training for the definition of hydro-geophysical relations

Similar to the definition of HF models with RVM classification, kernel functions and kernel window lengths are tested to define hydro-geophysical relations with the best predictive capabilities using *K*, *T*, *D* and *R* data of the TDRK_GK_4 subset (Step 3 in Figure 4). As illustrated in Figures 10a-c for a Laplace kernel function, a hydro-geophysical relation is independently defined for each of the four HFs. Note that for developing hydro-geophysical relation for a given HF, we assign each observation (*K*, *T*, *D*, *R* data) to the HF with the maximum membership after the fuzzy memberships of each observation obtained from clustering are transformed to integer numbers. Reliability of the hydro-geophysical relations is defined in terms of goodness-of-fit statistics that also reflect the adequacy and significance of the predicted model. These key statistics are mean error (*Bias*) and root-mean-square error (*RMS*):

$$Bias = N^{-1}\sum_{i=1}^{N}\left(\mathrm{t}_i - \mathrm{t}_i^*\right) \tag{20}$$

$$RMS = \sqrt{N^{-1}\sum_{i=1}^{N}\left(\mathrm{t}_i - \mathrm{t}_i^*\right)^2} \tag{21}$$

Additionally, we used the number of relevance vector *RV* defined in Eq. (19) as an index of structural complexity of the RVM regressor.

For the training of RVM regressors a LOO cross-validation procedure [*Stone*, 1974] is followed, where each observation is successively left out from the entire training dataset and used for testing. This procedure is generally well suited for model selection of regression problems [*Arlot and Celisse*, 2010]. Thus, for a given kernel function and kernel window length, relevance vectors are determined with the procedure described in Section 3.2 using all available data except one observation, and *K* prediction is made using Eq. (14) with the observation left out. This process is repeated until all observations have been used as testing data and the average for all experiments is used to assess performance (Figures 10a-b) and sparsity (Figure 10c) of the RVM regression models. The supervised regression was carried out with the SPARSEBAYES Matlab Toolbox [*Tipping and Faul*, 2003; *Tipping*, 2009] using Gauss likelihood. According to Figures 10a-c, the optimal kernel window lengths for HF1 to HF4 are 0.4, 0.15,

0.4 and 0.5, respectively. Note that the Laplace kernel function used in Figures 10a-c provided the best predictive capabilities as verified by the procedure presented in Section 4.4.

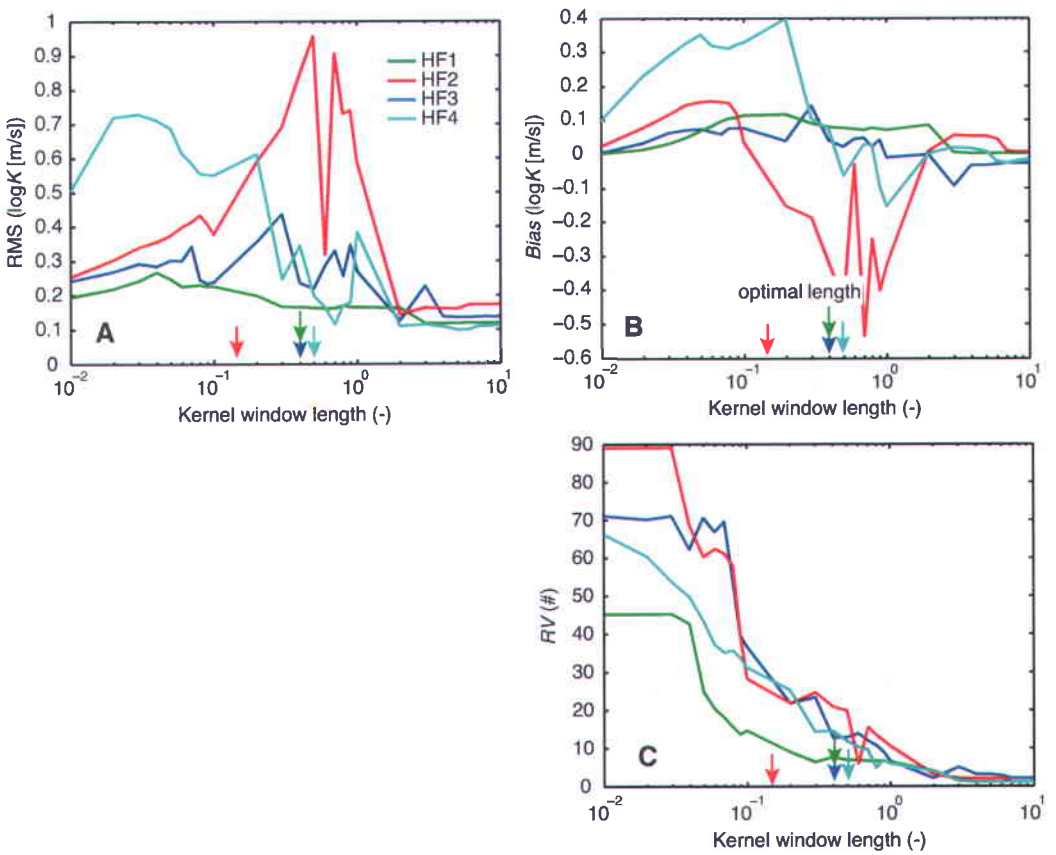


Figure V.10 Graph of (a) root-mean-square (*RMS*), (b) mean error (*Bias*), and (c) number of relevance vector (*RV*) used for the selection of the optimal kernel window length of each relevance vector machine for regression associated to the four hydrofacies of the TDRK_4GK subset. Results are for a Laplace kernel function.

4.4 Verification of hydrofacies and hydraulic conductivity predictions

In this section, we assess the error associated with the application of the previously trained learning machine to identify HF and predict K from CPT/SMR data for our study site. The verification process follows the same sequential steps for prediction illustrated in Figure 5 using training and testing datasets. First, the training data with parameters of the RVMs (kernel function and window length) are used to build HF models and hydro-geophysical relations. Then, for a given CPT/SMR observation (T, D, R for this example) of the testing set, HF is predicted using HF models and the hydro-geophysical relation corresponding to this HF is then applied to estimate K using the same CPT/SMR data. Predicted HF and K are then compared to known values of the testing set to assess performances. Note that predicted HFs

here are the result of the multi-class voting process. To assess classification and regression errors, we randomly selected 80% of the available colocated hydro-geophysical data and used it as a training set while the remaining 20% was used as a testing set. This procedure was repeated 100 times to provide error distributions associated with the selection of training and testing sets, as illustrated in Figure 11. The same RVM parameters (kernel function and window length) found in Sections 4.2 and 4.3 were used for all simulations.

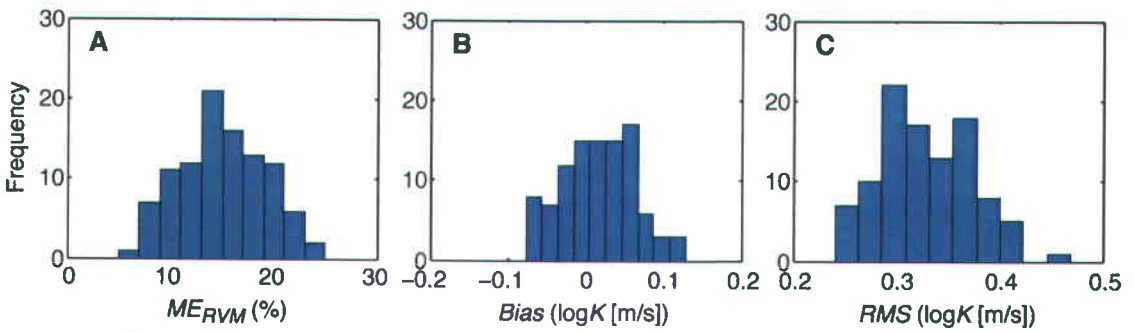


Figure V.11 Histograms of: (a) misclassification error (ME_{RVM}), (b) mean error (*Bias*), and (c) root-mean-square (*RMS*) for error assessment of hydrofacies (a) and hydraulic conductivity predictions (b-c).

According to the median ME_{RVM} value in Figure 11a, the error associated with HF recognition from CPT/SMR data is 14%. This value is similar to the ME_{olap} value (13%) obtained from clustering to evaluate the degree of HF over-lapping associated with non-uniqueness between K and CPT/SMR parameters. This means that the classification with the RVM is almost perfect, as expressed by non-zero ME_{RVM} value, and the obtained ME_{RVM} value is associated with non-uniqueness as discussed in Section 4.1. Figure 12a also presents a confusion matrix comparing the HF classification obtained by the RVM classifier to the original classification made by clustering for the simulation with the median ME_{RVM} value. The classification error is fairly well distributed over all HFs with ME_{RVM} for each HF ranging from 10% to 21%.

Figures 11b-c that assess regression performances, indicate that there is no significant bias in K estimates, as expressed by a median *Bias* value of 0.016. Also, the median *RMS* error for all simulations is 0.327, which represents approximately 14% of the total range in log$K$ values. According to the sequential procedure to estimate K (Figure 5), we note that the accuracy of predicted K depends on both the accuracy of the HF models and hydro-geophysical relations because the recognition of the wrong HF necessarily leads to the wrong hydro-geophysical relation. Figure 12b also presents a scatter plot

comparing log$K$ estimates obtained by the RVM regressors to the value obtained from slug tests for the same testing dataset used in Figure 12a. The correlation coefficient (r) between predicted and field log$K$ estimates is 84% and there is no bias in the estimate as the regression line overlaps the 1:1 perfect fit line.

Finally, Figures 13a-b and Table 5 show observed and predicted log$K$ distributions per HF, respectively, for the median testing dataset presented in Figure 12. Although the general trend in median log$K$ values between HFs is similar for both observed and predicted distributions, the range in predicted log$K$ for each HF is narrower than the observed ranges. This indicates that the learning machine smooth log$K$ estimations, which is inherent to any estimation process (classification and regression).

The actual error assessment use 80% of the available data for training of the learning machine and it is thus expected that when used to make prediction with 100% of the training data for field application (e.g., *Paradis et al.*, 2014), K estimates could be slightly better than values reported in this analysis. Meanwhile, the study of *Paradis et al.* [2011] that compared K estimates from flowmeter and multi-level slug tests, for 123 of the 280 intervals used in the present study, showed *RMS* and r values of 10% and 88%, respectively. Thus, the proposed indirect method to estimate K from CPT/SMR data compares fairly well with direct methods based on hydraulic tests.

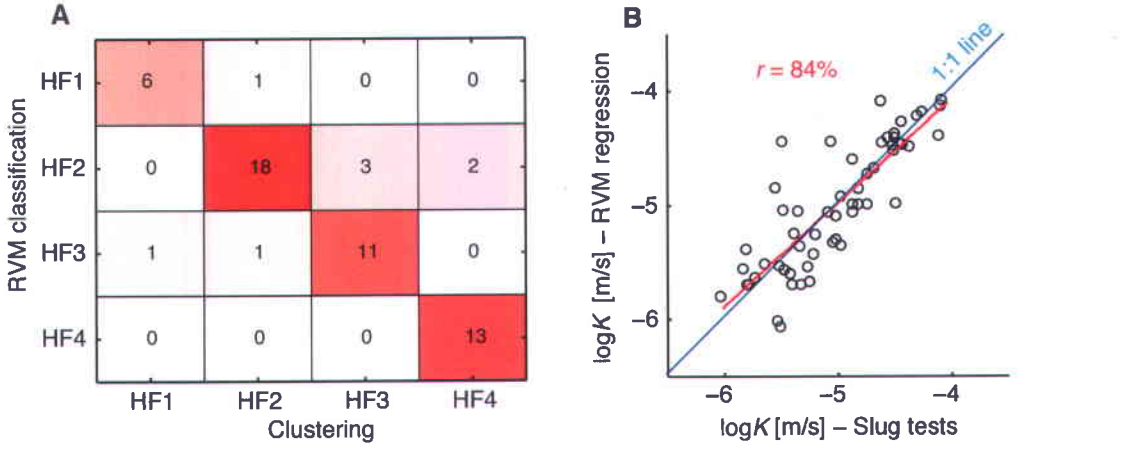


Figure V.12 (a) With a testing data set (n=56 observations), the confusion matrix compares the hydrofacies classification obtained by the relevance vector machine (RVM) classifier to the original classification made by fuzzy clustering (TDRK_GK_4 subset). This classification corresponds to the simulation using the median misclassification error illustrated in Figure V.11a. The diagonal indicates the observations for which both classifications are identical (HF1 6/7, 86%; HF2 18/20, 90%; HF3 11/14, 79%; HF4 13/15, 87%; overall 48/56, 86%). Off-diagonal observations were misclassified by RVM classification (HF1 1/7, 14%; HF2 2/20, 10%; HF3 3/14, 21%; HF4 2/15, 13%; overall 8/56, 14%). (b) Comparison of the logarithm of hydraulic conductivity (K) measured with multilevel slug tests with the estimation made using RVM regressors with the same testing data set as shown in Figure (a).

Table V.5 Logarithm of hydraulic conductivity (log*K*) statistics per hydrofacies and overall for the median testing data set presented in Figure V.12 resulting from the testing (error assessment) of the learning machine: (a) original log*K* data of the TDRK_GK_4 subset obtained with slug tests; and (b) predicted log*K* values from the learning machine following the approach illustrated in Figure V.5. Hydraulic conductivity distributions per hydrofacies are also illustrated in Figure V.13.

Hydrofacies	Number	Minimum	Maximum	Median	Standard-Deviation
Measured log*K* [m/s]-Slug tests					
1	7	-6.08	-5.33	-5.70	0.26
2	20	-5.81	-4.68	-5.32	0.35
3	14	-5.61	-4.46	-5.05	0.38
4	15	-4.60	-4.07	-4.39	0.16
Overall	56	-6.08	-4.07	-5.03	0.56
Predicted log*K* [m/s]-RVM					
1	7	-5.77	-5.23	-5.37	0.19
2	23	-6.02	-4.66	-5.39	0.40
3	13	-5.80	-4.41	-4.96	0.41
4	13	-4.85	-4.06	-4.42	0.24
Overall	56	-6.02	-4.06	-5.00	0.51

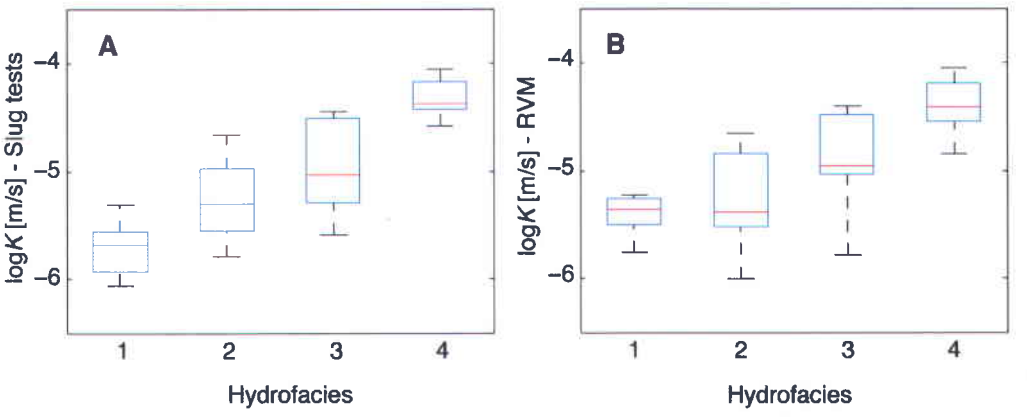


Figure V.13 Logarithm of hydraulic conductivity (log*K*) distribution per hydrofacies for the median testing data set presented in Figure V.12 resulting from the testing (error assessment) of the learning machine: (a) original hydraulic conductivity (log*K*) data of the TDRK_GK_4 subset obtained with slug tests; and (b) predicted log*K* values from the learning machine following the approach illustrated in Figure V.5.

5 Summary and conclusions

This paper presented a learning machine approach to define site-specific hydro-geophysical relationships in order to estimate aquifer hydraulic properties based solely on geophysical measurements. Specifically, we explored the use of CPT/SMR soundings data for K estimation using a statistical framework combining fuzzy clustering and relevance vector machines (RVMs). Hydro-geofacies (HFs) reflecting geological materials present within the studied aquifer were first extracted from a training dataset composed of K data measured in wells using 15-cm vertical resolution packer slug tests and CPT/SMR data that include resistance to penetration (tip stress, T), mechanical friction (sleeve stress, S), dielectric constant of bulk sediments (D) and DC electrical resistivity (R). All colocated hydraulic and CPT/SMR data were up-scaled to a common vertical resolution of 15 cm for the purpose of establishing hydro-geophysical relationships. RVMs for classification and regression were then trained independently using previous clustering data to define predictive HF models and hydro-geophysical relationships, respectively. Accuracy of HF and K estimates using the trained learning machine was assessed through a cross-validation procedure to evaluate the potential of CPT/SMR data for K estimation. Important conclusions and observations resulting from this study include the following:

Indirect CPT/SMR data hold the potential to estimate K at a high-vertical resolution even under real field conditions. The use of a multi-parameter probe, such as the CPT/SMR, reduces non-uniqueness between geophysical and hydraulic properties by providing a series of geophysical parameters that provide

complementary information on hydraulic data. Thus, using CPT/SMR data for aquifer characterization could be an efficient complement to conventional well-based hydraulic testing approach, as it contains enough information to infer K heterogeneities.

Even though up to four geophysical parameters from CPT/SMR were available to correlate with K for this study, it was not possible however to completely resolve the non-uniqueness in hydro-geophysical relations at our site, as expressed by the non-zero degree of over-lapping between HFs projected in the geophysical space. This remaining uncertainty is inherent to measurement and interpretation errors, but also to the fundamental nature of the relationships between geological materials and hydro-geophysical responses, which limits our capacity to fully represent hydraulic parameters through indirect approaches. This means that the choice and number of geophysical parameters to use as an indirect approach to characterize hydraulic properties for a particular site are crucial to ensure meaningful hydrogeological interpretation of geophysical measurements.

RVMs for classification and regression are effective to establish HF models and hydro-geophysical relationships with good generalization capabilities. Relations between hydraulic and geophysical parameters are generally nonlinear and difficult to define, as illustrated through the actual training dataset for the St-Lambert aquifer, and reliable statistical relationships are critical for successfully employing geophysical data for hydraulic properties characterization. In this study, HF models and hydro-geophysical relationships are defined through a kernel function that maps complex hydro-geophysical relations into a higher-dimensional space where classification and regression problems are easier to solve, which provides flexibility in fitting multivariate nonlinear relations. The division of the training dataset in HFs with distinct hydro-geophysical relations among the data also contributes to alleviate nonlinearity. From a practical perspective, for aquifer characterization over a relatively large study area, discrete representation of aquifer property heterogeneity with HFs is helpful to conceptualize the complexities present in most aquifers and to verify geological plausibility.

As non-parametric learning machines are based on empirical data, the selection of a representative hydro-geophysical training dataset is fundamental to establish meaningful hydro-geophysical relationships. In that respect, a systematic data acquisition approach should be adopted to provide a representative and complete training dataset for the training of the learning machine for a particular study area, as proposed

by *Paradis et al.* [2014] for the St-Lambert aquifer. Such an approach is needed to ensure the coverage of the entire range of hydraulic and geophysical responses observed over the study area.

A K heterogeneity characterization approach based on hydro-geophysical relations still has to involve hydraulic testing. Actually, such an approach has higher requirements on the type of K data needed, which leads to new ways to targeted hydraulic testing carried out in the perspective of providing K values over the observed range of geophysical responses in a given study area. Such a perspective and approach can contribute to a more efficient aquifer characterization process because less time consuming hydraulic tests are needed, but these tests have to meet more specific criteria, notably for high vertical resolution hydraulic testing suited for the definition of hydraulic heterogeneity. In that perspective, for the present study, we used a training dataset of K measurements from multi-level slug tests carried out in fully-screened direct-push wells at a vertical resolution of 15 cm. However, more efficient hydraulic testing approaches are still needed to acquire larger K training datasets (e.g., *Liu et al.*, 2009; *Paradis et al.*, 2011).

Finally, although the probing methods and relations defined in this paper are site-specific, the learning machine is general in its basic principles and could be extended to other hydrogeological contexts with different hydraulic and geophysical characterization tools.

Acknowledgements

The authors would like to acknowledge the important technical support provided by J.-M. Ballard and D. Martin as well as Y. Michaud and D. Kirkwood for their support. The Geological Survey of Canada (Groundwater Geoscience Program), the Régie intermunicipale de gestion des déchets des Chutes-de-la-Chaudière and NSERC Discovery Grants held by E.G. and R.L supported this study. This is an Earth Science Sector contribution number 20130456.

References

Al-Anazi, A., I. D. Gates, and J. Azaiez (2009), Fuzzy logic data-driven permeability prediction for heterogeneous reservoirs, in Proceedings of the Society of Petroleum Engineers of the

EUROPEC/EAGE Annual Conference and Exhibition, Amsterdam, The Netherlands, Paper SPE 121159, doi: 10.2118/121159-MS.

Al-Anazi, A., and I. D. Gates (2010a), A support vector machine algorithm to classify lithofacies and model permeability in heterogeneous reservoirs, *Eng. Geol.*, *114*, 267-277.

Al-Anazi, A., and I.D. Gates (2010b), On the capability of support vector machines to classify lithology from well logs, *Nat. Resour. Res.*, *19*, 125-139.

Allwein, E., R. Shapire, and Y. Singer (2000), Reducing multiclass to binary: A unifying approach for margin classifiers, *J. Mach. Learn. Res.*, *1*, 113-141.

American Society for Testing and Materials-ASTM (2000), D5578-95: Standard test method for performing electronic friction cone and piezocone penetration testing of soils, Annual book of ASTM standards.

American Society for Testing and Materials-ASTM (2004), D6724: Standard guide for installation of direct push ground water monitoring wells. ASTM International, 9 pp.

Anderson, M. P. (1989), Hydrogeologic facies models to delineate large-scale spatial trends in glacial and glaciofluvial sediments, *Geol. Soc. Am. Bull.*, *101*, 501-511.

Anderson, M. P. (1997), Characterization of geological heterogeneity, in: *Stochastic Subsurface Hydrology*, edited by G. Dagan and S. P. Neuman, Cambridge University Press, Cambridge.

Arlot, S., and A. Celisse (2010), A survey of cross-validation procedures for model selection, *Statistics Surveys*, *4*, 40-79.

Balasko, B., J. Abonyi, and B. Feil (2005), Fuzzy Clustering and Data Analysis Toolbox for Use with Matlab, University of Veszprem, Veszprem Hungary, Accessed: 29 Marsh 2011, http://www.mathworks.com/matlabcentral/fileexchange/7473

Bezdek, J. C. (1981), *Pattern recognition with fuzzy objective function algoritms*, Plenum Press, New York.

Bolduc, A. (2003), Géologie des formations superficielles, Charny, Québec, Commission géologique du Canada, Dossier public 1976, échelle 1/50000

Bouwer, H., R. C. Rice (1976), A slug test method for determining hydraulic conductivity of unconfined aquifers with completely or partially penetrating wells, *Water. Resour. Res.*, *12*, 423-428.

Butler, Jr. J. J. (1998), *The design, performance, and analysis of slug tests*, Lewis Publishers, Boca Raton, Florida.

Butler, Jr. J. J. (2005), Hydrogeological methods for estimation of hydraulic conductivity, in *Hydrogeophysics* edited by Y. Rubin, and S. Hubbard, pp. 23-58, Springer, New York.

Camps-Valls, G., L. Gómez-Chova, J. Muñoz-Marí, J. Vila-Francés, J. Amorós-López, and J. Calpe-Maravilla (2006), Retrieval of oceanic chlorophyll concentration with relevance vector machines, *Remote Sens. Environ.*, *105*, 23-33.

Chen, J., and Y. Rubin (2003), An effective Bayesian model for lithofacies estimation using geophysical data, *Water Resour. Res.*, *39*, 1118, doi:10.1029/2002WR001666, 5.

Chen, J., S. Hubbard, and Y. Rubin (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 Resour. Res.*, *37*, 1603-1613, doi:10.1029/2000WR900392.

Davis, J. C. (1973), *Statistics and data analysis in geology*, Wiley, New York

Dubois, M. K., G. C. Bohling, and S. Chakrabarti (2007), Comparison of four approaches to a rock facies classification problem, *Comput. Geosci.*, *33*, 599-617.

Dunn, J. C. (1973), A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, *J. Cybernetics*, *3*, 32-57.

Elshafei, M., G. Hamada (2009), Neural network identification of hydrocarbon potential of shaly sand reservoirs, *Journal of Petroleum Science and Technology*, *27*, 72-82.

Farrar, J. A. (1996), Research and standardization needs for direct push technology applied to environmental site characterization, in *Sampling Environmental Medias* edited by J. H. Morgan, pp. 93-107, ASTM Special Technical Publication 1282. American Society for Testing and Materials, Philadelphia, PA.

Fellenius, B. H., and A. Eslami (2000), Soil profile interpreted from CPTu data, presented at Year 2000 Geotechnics, Geotechnical Engineering Conference, Asian Institute of Technology, Bangkok, Thailand, November 27-30, 2000.

Geisser, S. (1975), The predictive sample reuse method with applications, *J. Amer. Statist. Assoc.*, *70*, 320-328.

Ghosh, S., and P. Mujumdar (2008), Statistical downscaling of GCM simulations to streamflow using relevance vector machine, *Adv. Water. Resour.*, *31*, 132-146.

Gustafson, D. E., and W. C. Kessel (1979), Fuzzy clustering with a fuzzy covariance matrix, Proceeding of IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes, San Diego, CA, pp. 761-766.

Guyon, I., and A. Elisseeff (2003), An introduction to variable and feature selection, *J. Mach. Learn. Res.*, *3*, 1157-1182.

Harman, H. H. (1976), *Modern Factor Analysis*. 3rd ed. Chicago: University of Chicago Press.

Hastie, T., and R. Tibshirani (1998), Classification by pairwise coupling, *Ann. Stat.*, *26*, 451-471.

Henebry, B. J., and G. A. Robbins (2000), Reducing the influence of skin effects on hydraulic conductivity determinations in multilevel samplers installed with direct push methods, *Ground Water*, *38*, 882-886.

Hsu, C. W., and C. J. Lin (2002), A comparison of methods for multi-class support vector machines, *IEEE Trans. Neural Netw.*, *13*, 415-425.

Isaaks, E. H., and R. M. Srivastava (1989), *An introduction to applied geostatistics*, Oxford University Press, New York.

Jöreskog, K. G. (1967), Some contributions to maximum likelihood factor analysis, *Psychometrika*, *32*, 443-482.

Khalili, A., M. N. Almasri, M. McKee, and J. J. Kaluarachchi (2005), Applicability of statistical learning algorithms in groundwater quality modeling, *Water. Resour. Res.*, 41:W05010, doi:10.1029/2004WR003608.

Kharrat, R., R. Mahdavi, H. Bagherpour, and S. Hejri (2009), Rock type and permeability prediction of a heterogeneous carbonate reservoir using artificial neural networks based on flow zone index approach, paper 120166 presented at SPE Middle East Oil and Gas Show and Conference, Bahrain, March 15-18.

Koltermann, C. E., and S. M. Gorelick (1996), Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches, *Water Resour. Res.*, *32*, 2617-2658.

Lafuerza, S., M. Canals, J. L. Casamor, and J. M. Devincenzi (2005), Characterization of deltaic sediment bodies based on in situ CPT/CPTU profiles: A case study on the Llobregat delta plain, Barcelona, Spain, *Mar. Geol.*, *222-223*, 497-510.

Lamarche, L., and L. Tremblay (2012), Géologie des formations superficielles pour le site de St-Lambert-de-Lauzon, Québec, in Tremblay, L. (2013). Caractérisation intégrée d'un aquifère granulaire pour l'évaluation des processus géochimiques influençant l'atténuation naturelle d'un panache de lixiviat. Ph.D. Thesis, Institut national de la recherche scientifique (INRS) p. 50.

Lee, S. H., and A. Datta-Gupta (1999), Electrofacies characterization and permeability predictions in carbonate reservoirs: Role of multivariate analysis and nonparametric regression, SPE Annual Technical Conference and Exhibition, 3-6 October 1999, Houston, Texas.

Levy, B. S., and L. Pannell (1991), Evaluation of a pressure system for estimating in-situ hydraulic conductivity, in Proceedings 5th National Outdoor Action Conference, NWWA, Dublin OH, pp. 131-146.

Liu, G., J. J. Butler Jr., G.C. Bohling, E. Reboulet, S. Knobbe, and D.W. Hyndman (2009) A new method for high-resolution characterization of hydraulic conductivity, *Water Resour Res*, *45*, W08202, doi:10.1029/2009WR008319.

Lunne, T., P. K. Robertson, and J. J. M. Powell (1997), *Cone penetration testing in geotechnical practice*, Spon Press, New York.

Mao, J., and A. Jain (1996), A self-organizing network for hyperellipsoidal clustering (HEC), *IEEE Trans. Neural Netw.*, *7*, 16-29.

MacKay, D. J. C. (1992), The evidence framework applied to classification networks, *Neural Comput.*, *4*, 720-736.

Mitchell, T. (1997), *Machine Learning*, McGraw Hill, New York.

Mohaghegh, S., B. Balan, and S. Ameri (1997), Permeability determination from well log data, *SPE Format. Eval.*, *12*, 170-174.

Ouellon, T., R. Lefebvre, D. Marcotte, A. Boutin, V. Blais, and M. Parent (2008), Hydraulic conductivity heterogeneity of a local deltaic aquifer system from the kriged 3D distribution of hydrofacies from borehole logs, Valcatier, Canada, *J. Hydrol.*, *351*, 71-86.

Paradis, D., R. Lefebvre, R. H. Morin, and E. Gloaguen (2011), Permeability profiles in granular aquifers using flowmeters in direct-push wells, *Ground Water*, *49*, 534-547.

Paradis, D., L. Tremblay, R. Lefebvre R, E. Gloaguen, A. Rivera, M. Parent, J.-M. Ballard, Y. Michaud, and P. Brunet (2014), Field characterization and data integration to define the hydraulic heterogeneity of a shallow granular aquifer at a sub-watershed scale, Accepted for publication in *Environ. Earth Sci.*, *February 2014*.

Robertson, P. K. (1990), Soil classification using the cone penetration test, *Can. Geotech. J.*, *27*, 151-158.

Ross, H. C., and C. D. McElwee (2007), Multi-level slug tests to measure 3-D hydraulic conductivity distributions, *Water Resour. Res.*, *16*, 67-79.

Rubin, Y., and S. Hubbard (2005), *Hydrogeophysics*, Springer, Netherlands.

Samui, P. (2007), Seismic liquefaction potential assessment by using relevance vector machine. *Earthq. Eng. Eng. Vib.*, *6*, 331-336.

Schulmeister, M., J. J. Butler Jr., J. Healey, L. Zheng, D. Wysocki, and G. McCall (2003), Direct-push electrical conductivity logging for high-resolution hydrostratigraphic characterization, *Ground Water Monit. R.*, *23*, 52-62.

Sellwood, S.M., J. M. Healey, S. Birks, and J. J. Butler Jr. (2005), Direct-push hydrostratigraphic profiling: Coupling electrical logging and slug tests, *Ground Water*, *43*, 19-29.

Shinn, J.D., D. A. Timian, R. M. Morey, G. Mitchell, C. L. Antle, and R. Hull (1998), Development of a CPT deployed probe for in situ measurement of volumetric soil moisture content and electrical resistivity, *Field Anal. Chem. Tech.*, 2, 103-110.

Shokir, E. M. El-M., A. Ateeq, and A. Al-Sughayer (2006), Permeability estimation from well log responses, *J. Can. Pet. Technol.*, *45*, 41-46.

Stone, M. (1974), Cross-validatory choice and assessment of statistical predictions, *J. Roy. Statist. Soc. Ser. B*, *36*, 111-147.

Tetko, I. V., D. J. Livingstone, and A. I. Luik (1995), Neural network studies : 1. Comparison of overfitting and overtraining, *J. Chem. Inf. Comput. Sci.*, *35*, 826-833.

Tipping, M. E. (2001), Sparse Bayesian learning and the relevance vector machine, *J. Mach. Learn. Res.*, *1*, 211-244

Tipping, M. E., and A. C. Faul (2003), Fast marginal likelihood maximisation for sparse Bayesian models, in Proceedings of the Ninth International Workshop on Artificial Intelligence and Statistics edited by C. M. Bishop and B. J. Frey, Key West, FL, Jan 3-6.

Tipping, M. E. (2009), SPARSEBAYES : An efficient Matlab implementation of the sparse Bayesian modelling algorithm (Version 2.0), accessed: December 2013, http://www.miketipping.com

Tremblay, L., R. Lefebvre, D. Paradis, and E. Gloaguen, (2013), Conceptual model of leachate migration in a granular aquifer derived from the integration of multi-source characterization data (St-Lambert, Canada). Published online December 6, 2013, *Hydrogeology Journal,* doi: 10.1007/s10040-013-1065-1.

Vapnik, V. (1995), *The nature of statistical learning theory*, Springer-Verlag, New York.

Vapnik, V. (1998), *Statistical learning theory*, John Wiley and Sons, New York.

Wong, P. M., D. J. Henderson, and L. J. Brooks (1998), Permeability determination using neural networks in the Ravva Field, Offshore India, *SPE Format. Eval.*, *1*, 99-104.

Zemansky, G. M., and C. D. McElwee (2005), High-resolution slug testing, *Ground Water*, *43*, 222-230.

ANNEXES

DESCRIPTION DES ANNEXES

Cette section comprend une série d'annexes fournie en support aux travaux présentés dans cette thèse. À l'Annexe A on retrouve la version anglaise abrégée de la synthèse de la première partie soumise et acceptée pour publication :

Annexe A **Synthèse (version anglaise)** **295**

Paradis, Daniel, Laurie Tremblay, René Lefebvre, Erwan Gloaguen, Alfonso Rivera, Michel Parent, Jean-Marc Ballard, Yves Michaud and Patrick Brunet (2014) Field characterization and data integration to define the hydraulic heterogeneity of a shallow granular aquifer at a sub-watershed scale. *Environmental Earth Sciences*, doi: 10.1007/s12665-014-3318-2

On retrouve aussi en version numérique (DVD) les mesures de terrain et de laboratoire permettant de reproduire les analyses présentées dans la synthèse et les articles originaux de cette thèse :

Annexe B **Données des tests à choc hydraulique multi-niveaux** **DVD**

Données utilisées pour la Synthèse et les Articles I et V

Annexe C **Données des tests avec débitmètre de puits** **DVD**

Données utilisées pour l'Article I

Annexe D **Données des tests d'interférence verticale à choc hydraulique** **DVD**

Données utilisées pour les Articles II et IV

Annexe E **Données de tomographie à choc hydraulique** **DVD**

Données utilisées pour l'Article IV

Annexe F **Données des sondages CPT/SMR brutes et mise-à-l'échelle** **DVD**

Données utilisées pour la Synthèse et l'Article V

Annexe G **Données d'entrainement lithologique, hydraulique et CPT/SMR co-localisées mise-à-l'échelle** **DVD**

Données utilisées pour la Synthèse et l'Article V

ANNEXE A

SYNTHÈSE (VERSION ANGLAISE)

Citation: Paradis, Daniel, Laurie Tremblay, René Lefebvre, Erwan Gloaguen, Alfonso Rivera, Michel Parent, Jean-Marc Ballard, Yves Michaud and Patrick Brunet (2014) Field characterization and data integration to define the hydraulic heterogeneity of a shallow granular aquifer at a sub-watershed scale. *Environmental Earth Sciences*, doi: 10.1007/s12665-014-3318-2

Field characterization and data integration to define the hydraulic heterogeneity of a shallow granular aquifer at a sub-watershed

Daniel Paradis[1,2,*], Laurie Tremblay[2], René Lefebvre[2], Erwan Gloaguen[2], Alfonso Rivera[1], Michel Parent[1], Jean-Marc Ballard[2], Yves Michaud[1] and Patrick Brunet[2]

[1] Geological Survey of Canada, 490 rue de la Couronne, Quebec City, Canada G1K 9A9

[2] Institut national de la recherche scientifique, Centre Eau Terre Environnement (INRS-ETE), 490 rue de la Couronne, Quebec City, Canada G1K 9A9

*Corresponding author: Daniel Paradis

Tel : (418) 654-3713 Fax : (418) 654-2604 E-mail: dparadis@nrcan.gc.ca

Abstract

Providing a sound basis for aquifer management or remediation requires that hydrogeological investigations carried out to understand groundwater flow and contaminant transport be based on representative data that capture the heterogeneous spatial distribution of aquifer hydraulic properties. This paper describes a general workflow allowing the characterization of the heterogeneity of the hydraulic properties of granular aquifers at an intermediate scale of a few km2. The workflow involves characterization and data integration steps that were applied on a 12 km2 study area encompassing a decommissioned landfill emitting a leachate plume and its main surface water receptors. The sediments composing the aquifer were deposited in a littoral-sublittoral environment and show evidence of small-scale transitional heterogeneities. Cone penetrometer tests (CPT) combined with soil moisture and electrical resistivity (SMR) measurements were thus used to identify and characterize spatial heterogeneities in hydraulic properties over the study area. Site-specific statistical relationships were needed to infer hydrofacies units and to estimate hydraulic properties from high-resolution CPT/SMR soundings distributed all over the study area. A learning machine approach was used due to the complex statistical relationships between colocated hydraulic and CPT/SMR data covering the full range of aquifer materials. Application of this workflow allowed the identification of hydrofacies units and the estimation of horizontal hydraulic conductivity, vertical hydraulic conductivity and porosity over the study area. The paper describes and discusses data acquisition and integration methodologies that can be adapted to different field situations, while making the aquifer characterization process more time-efficient and less labour-intensive.

Keywords: Heterogeneity, Hydrofacies, Hydraulic properties, Aquifer characterization, Direct-push, Hydrogeophysics, Data integration, Relevance vector machine (RVM)

1 INTRODUCTION

Hydrogeological investigations designed to assess groundwater flow or contaminant transport require representative data on the heterogeneous spatial distribution of hydraulic properties in aquifers, such as hydraulic conductivity (K) and porosity (n) (Anderson et al. 1999; Eaton 2006). In unconsolidated sediments, the heterogeneity of hydraulic properties results from complex geological processes (Koltermann and Gorelick 1996). This heterogeneity influences groundwater flow and especially mass transport processes in aquifers over various scales (e.g., Caputo et al. 2010; Dann et al. 2011; Güngör-Demirci and Aksoy 2011; Lu et al. 2011; Mastrocicco et al 2011; Blouin et al. 2013; Doro et al. 2013). Indeed, the large-scale heterogeneities in K control the paths of groundwater and contaminants, while both large and small-scale variations in K modulate the magnitude of contaminant dispersion. Thus, to provide a sound basis for aquifer management, the scale of heterogeneity relevant for a particular hydrogeological investigation should be recognized and characterized accordingly (e.g., Gräbe et al. 2013). In doing so, consideration has to be paid to the type of problem to be addressed (e.g., flow or transport), the scale of the investigated site, the medium, the type of answer to be provided (e.g., point or diffuse) and the cost of data collection (de Marsily et al. 2005).

Discrete representation of hydraulic property heterogeneity through the definition of hydrofacies (HF) unit is generally helpful to conceptualize the complexities present in most aquifers and to keep control on geological plausibility (Anderson 1997; Poeter and McKenna 1995; Ouellon et al. 2008; Frei et al. 2009). A HF unit is generally a functional unit that is hydrogeologically meaningful for the purposes of flow and transport modeling (Anderson 1989). Complete aquifer characterization would ideally yield a three-dimensional model representing the spatial distribution of hydraulic properties on the basis of the spatial arrangements of the HF units present in the aquifer system.

Field hydraulic tests are common and generally reliable sources of information on aquifer hydraulic properties (Butler 2005). Conventional hydraulic tests, such as pumping tests and slug tests, induce horizontal flow patterns and only estimate horizontal hydraulic conductivity (K_h). Complete understanding of groundwater flow and contaminant transport also requires the knowledge of K anisotropy, i.e. the ratio of vertical K_v and horizontal K_h hydraulic conductivities (or vertical hydraulic conductivity, K_v) and n. K_v and n can be obtained from laboratory tests carried out on soil samples that are usually obtained by coring during drilling operations. However, the typically long duration of conventional hydraulic tests and the difficulties

in the experimental procedure related to sample collection and manipulation generally limit our ability to obtain large amounts of direct data about K_h, K_v and n. As a result, there is rarely sufficient information at the needed level of detail to construct meaningful hydrogeological models (Gelhar 1993; Anderson 1997; Butler 2005). The motivation of this study is then to develop an in situ aquifer characterization approach that would enable time and cost effective methods for imaging the heterogeneity of hydraulic properties in granular aquifers.

In this paper, we present the general workflow that was applied for the characterization of hydraulic properties of a shallow unconfined aquifer for the study of leachate migration from a decommissioned sanitary landfill to its main receptors (Tremblay et al. 2013). The study covers a 12 km^2 area surrounding the landfill. Sediments composing the aquifer show evidence of small and large-scale transitional changes in sediment textures, reflecting their littoral to sublittoral depositional environment. To address the challenge posed by this aquifer for the definition of heterogeneity in the study area, the emphasis was placed on the acquisition of high vertical resolution indirect data that could be efficiently obtained at relatively low cost. The choice was made to use cone penetrometer testing (CPT) combined with a soil moisture and resistivity probe (SMR), as this met our requirements for defining the stratigraphy. CPT/SMR is a multi-parameter probe that provides vertical profiles of mechanical and electrical properties of sediments. The value of using CPT/SMR data for aquifer characterization lies in its vertical decimeter-scale resolution (Lunne et al. 1997; Schulmeister et al. 2003), which cannot be obtained by surface-based geophysical methods. Moreover, the relatively extensive spatial coverage that can be obtained by direct-push soundings in comparison to well or core-based hydraulic tests allows the imaging of aquifer heterogeneity over large areas (Lafuerza et al. 2005). To be used to estimate hydraulic properties, Paradis et al. (2011; 2014b) proposed a data integration approach based on learning machines to predict HF and K_h from CPT/SMR data. The estimation of hydraulic data from CPT/SMR data thus constitutes the core of the aquifer characterization approach presented in this paper. As the adopted learning machine approach is based on empirical site-specific relations between hydraulic and CPT/SMR data, we followed a top down data acquisition workflow to acquire a training dataset representative of the study area. In this process, general characterization methods with greater spatial coverage and less resolution are first applied followed by specific methods with higher resolution and smaller volumes of investigation to rationalize data collection, as commonly used in the oil industry for hydrocarbon exploration (Bradford and Babcock 2013). Besides the definition of hydro-geophysical relations to predict HF and K_h from direct-push data, the data integration approach involves the definition of relations that link sparse K_v and n data to K_h and spatial interpolation of HF, K_h, K_v and n over the study area to provide complete and meaningful fields of hydraulic properties. To the knowledge of the

authors, there are very few field studies reporting such a level of integration for aquifer heterogeneity characterization that spans the entire spectrum from adapted field data acquisition to high-end statistical data integration. The current study is also original in terms of its scale, as hydrogeophysical efforts to define aquifer heterogeneity are often restricted to small areas and also illustrates the potential of geophysical-based approach for efficient aquifer characterization over relatively large areas.

This paper first presents the study area with a general description of the data acquisition workflow used for the characterization of hydraulic property heterogeneity. The data integration approach is then presented, which includes a brief description of the learning machine approach used to predict hydraulic information from CPT/SMR data and the interpolation scheme used to obtain the spatial distribution of HF units and hydraulic properties over the study area and illustrated along a representative section of the aquifer. The relative time efficiency of the proposed approach to predict K from direct-push data in comparison to conventional hydraulic tests is also discussed.

2 Study area and data acquisition

2.1 Summary of the data acquisition approach

Figure 1 illustrates the general data acquisition workflow followed in this study to characterize heterogeneity of hydraulic properties. This workflow follows a top down approach initially using general methods with greater spatial coverage and less resolution, followed by more specific methods with higher resolution and smaller volumes of investigation according to the overall responses of previously applied methods. Also, general methods such as the regional geology analysis and GPR surveys provide indirect information about the stratigraphy, whereas specific methods such as slug tests provide direct estimates of K_h. This data acquisition workflow is intended to rationalize data collection and to maximize the significance of the information gathered. For instance, the proposed workflow mainly relies on the collection of a training dataset representative of the study area to establish relations between hydraulic and direct-push parameters, as proposed by Paradis et al. (2011; 2014b). Then hydraulic tests are carried out with parsimony only at strategic locations where changes in stratigraphy can be deduced from large-scale geophysical measurements in order to cover the range of material types. The approach thus relies on integration of direct data with more abundant indirect CPT/SMR data to estimate hydraulic properties. Since

CPT/SMR sounding locations are selected on the basis of previous geological and GPR surveys, they can provide more relevant details about heterogeneity of hydraulic properties than would be practically possible from well-based hydraulic tests alone. In the remainder of this section, we thus present the application of the data acquisition workflow, as illustrated in Figure 1, for field characterization of the test site.

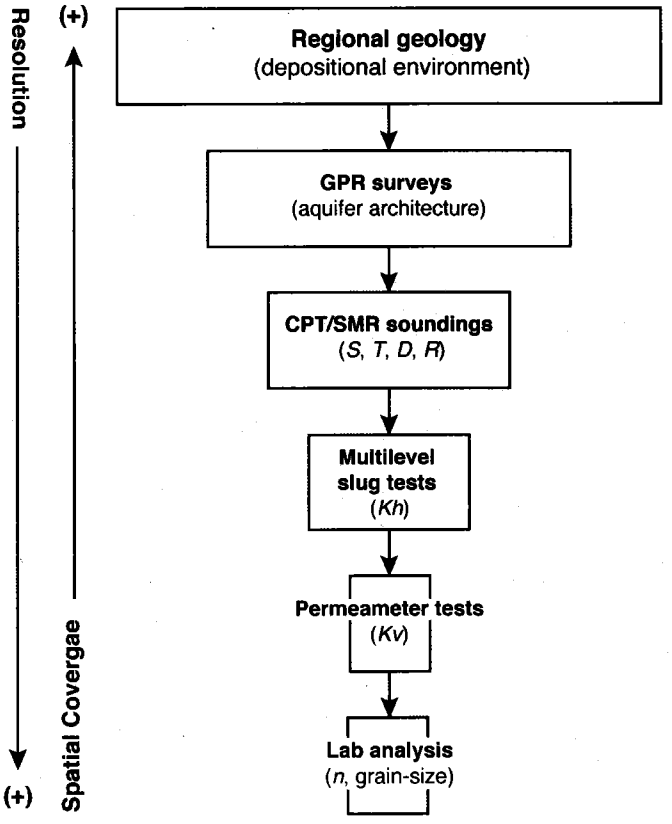


Figure 1 General sequential data acquisition workflow for the characterization of heterogeneities in hydraulic properties at the St-Lambert study area (*S*: sleeve stress; *T*: tip stress; *D*: dielectric constant; *R*: electrical resistivity; K_h: horizontal hydraulic conductivity, *Kv*: vertical hydraulic conductivity; *n*: total porosity).

2.2 St-Lambert site and regional geology

The proposed methodology was applied in a test area located in St-Lambert-de-Lauzon, 30 km south of Quebec City, Canada (Figures 2a-b). The St-Lambert study (Figure 2c) included a 12-km^2 sub-watershed surrounding a decommissioned sanitary landfill in order to assess the migration and natural attenuation of a leachate plume. The potential receptors of the plume emitted from the former landfill are streams delineating the north, west and south boundaries of the study area (Tremblay et al. 2013). The eastern limit corresponds to a groundwater divide identified from a potentiometric map.

As a first step in the aquifer characterization process, air photos and surficial sediment exposures were analyzed to identify the depositional environment and map the surficial geology of the study area (Lamarche and Tremblay 2012). The knowledge of the depositional environment is particularly helpful to gain insight on sediment architecture and spatial trends in sediment textures, which could provide qualitative information about the heterogeneity of hydraulic properties (Heinz et al. 2003). As reported by Bolduc (2003), the surficial sediments of the study area (Figure 2b) consist primarily of Late Quaternary sandy and silty sediments that were deposited in the receding Champlain Sea, which was an arm of the Atlantic Ocean that had invaded the St.Lawrence Valley at the time of the last deglaciation. More specifically, deposition at the St-Lambert site was controlled mainly by longshore currents that redeposited in littoral and sublittoral settings the sediments supplied to the Chaudière River paleodelta. This is indicated in Figure 2c by the southwestward fining of the littoral sediments in conjunction with the southwest-northeast trend of the beach ridges (Lamarche and Tremblay 2012). These ridges and the associated nearshore bars are mostly composed of medium to fine sand while the intervening troughs are composed of finer, silty sediments (Reading 1996). The silt-rich lithofacies found below littoral sands were deposited in sublittoral foreshore environments, while those interstratified with thin sand layers (CPT 24) were deposited in troughs and in lagoonal backshore settings during marine regression.

Sediments size for the aquifer ranges from fine sand to very fine silt (clayey silt) with poor to very poor grain-size sorting. Sediment size and sorting in sandy littoral environments result from the interaction of two main depositional processes, namely saltation and suspension transport occurring concurrently (Reineck and Singh 1980). The degree of sorting for sands and silts found at the St-Lambert site may thus result in relatively lower values of K in comparison to typically better sorted sediments. As shown in Figure 3, rapid transitions in sediments size and sorting caused by rapidly varying depositional energy are also generally observed in surface exposures. The resulting superposition of sand and silt layers may then create semi-confined aquifer conditions due to their respective K contrast. More gradual lateral intrastratal transitions in these littoral and sublittoral sediments are also expected as a result of changing energy levels along Champlain Sea shorelines.

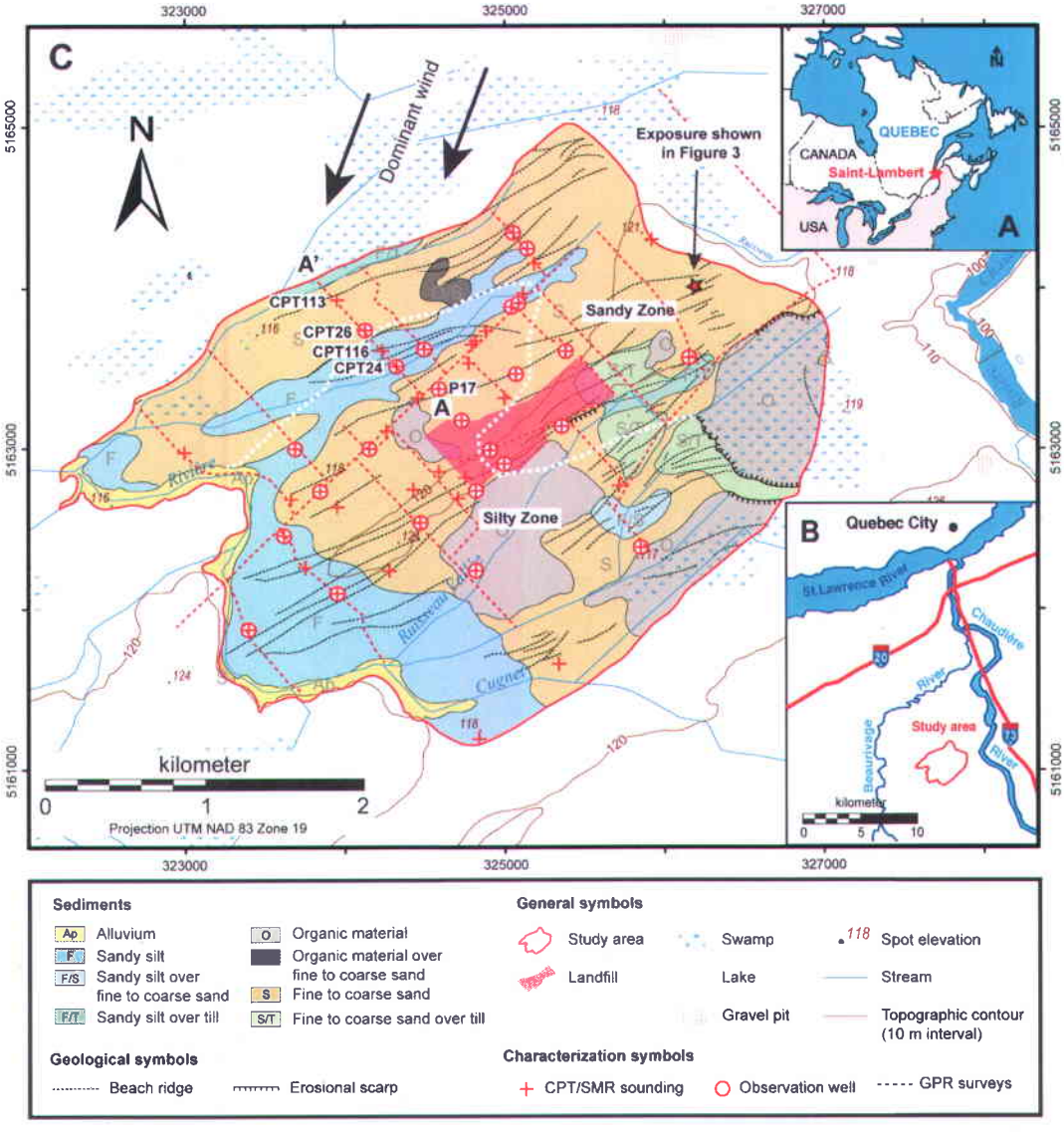


Figure 2 General location of the study area (a-b), with the Quaternary sediments map (c) for the sub-watershed enclosing the decommissioned sanitary landfill with the location of GPR surveys, direct-push soundings and observation wells used for aquifer characterization. The main depositional direction of sediments making up the granular aquifer is assumed to have been oblique (northeasterly winds at that time according to Filion 1987) to the orientation of the paleo-shore in a littoral environment. The white dashed line separates zones with relatively more sandy and more silty sediments (as shown on Section A-A' in Figures 4 and 12). The Quaternary map was modified from Lamarche and Tremblay (2012).

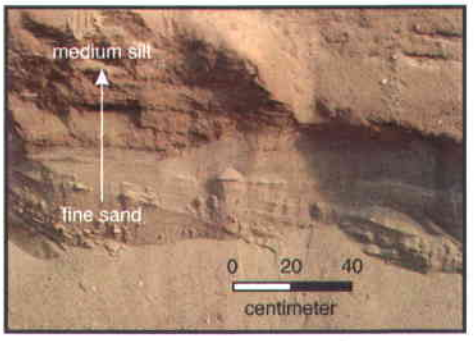


Figure 3 Photo of sediments showing rapid vertical transition dry sand beds upward into damp silt beds. The location of the exposure is shown in Figure 2c.

2.3 GPR surveys

Ground penetrating radar (GPR) surveys were chosen as a tool to characterize the aquifer architecture in complement to the regional geology. GPR is a fast, inexpensive and established geophysical technique for non-invasive imaging of stratigraphy of unconsolidated sediments (Van Overmeeren 1998; Schmelzbach et al. 2011). It uses high-frequency electromagnetic waves that are emitted in the ground, where they propagate and eventually are refracted or reflected to the surface where a receiver record the electro-magnetic signal as a function of time (Davis and Annan 1989). Changes in water content, clay fraction and grain-size distribution determine the reflection coefficients of the different materials. A GPR survey can resolve sedimentary structures on the order of a quarter of the signal wavelength, which for the 100 MHz antenna used for this study corresponds to approximately 0.2 m (Table 1).

Typically, reflection amplitudes can be correlated to the contrasts between different sediment grain sizes and can be used to identify zones with different hydrogeological conditions. Approximately 21 km of GPR surveys were acquired over the study area to identify bounding surfaces and the geometry of internal structures within the aquifer (Figure 2c and Table 1). As an example, the 100 MHz GPR section in Figure 4 depicts the general stratigraphic architecture of the littoral sediments in the Saint-Lambert study area (see section location in Figure 2c). Sediment profiles derived from CPT/SMR soundings, as discussed in the next section, are also shown in Figure 4. While those profiles were not generally all available at the time of the GPR surveys, they are presented here in support of observations and discussions.

Table 1 Summary of characterization data acquisition (field and laboratory).

Survey type	Instrument	Configuration	Vertical resolution (m)	Number of surveys	Total length (m)
GPR	Pulse Echo	100 MHz antenna	0.2	20	21 000
CPT/SMR	Geotech 605D	Vertek data acquisition system	0.03 to 0.17	53	509
Wells	Geotech 605D with 7.72 cm (3-inch) diameter metal casing	5.1 cm (2-inch) diameter PVC direct-push well without sand pack	n.a.	25	192 (screen length)
Sediment samples	Geotech 605D with Geotech Macro-Core Sampler	3.8 cm (1.5-inch) diameter and 1.5 m long clear PETG liner	n.a.	75	95
Multi-level slug tests (K_h)	Pneumatic test initiation system with pressure transducer	Dual-packer system	0.15	280	42
Permeameter tests (K_v)	Falling head permeameter with pressure transducer	Through original sediment cores without recompaction	0.15	59	9
Porosity (n)	Oven at 50° C (>7 days)	Wet and dry weights	0.15	43	6.5
Grain-size distribution	Fritsch Analysette 22	Laser light scattering	~0.01	59	0.6

Several observations can be made from Figure 4. First, the continuous stratification extending several hundred meters, shown by continuous GPR reflections dipping a few degrees (between 3° to 5°) towards the former sea (present-day St. Lawrence River in Figure 2b), is a characteristic feature of a littoral depositional environment (Van Overmeeren 1998). More specifically, a mound in GPR reflections located at the distance of 350 m on the GPR profile in Figure 4 suggests the presence of a thick beach ridge that acts as a sand bar or barrier and separates the GPR section into two distinct regions. The shoreface section lies to the left of the sand barrier (toward A') and the sand bar itself shows good GPR resolution and penetration depth, which are characteristic of sandy lithofacies (as also indicated by sediment types inferred from CPT113 and CPT26) with very low content of conductive material such as clay. The good GPR penetration depth in this zone (sandy zone in Figure 4) also allows the recognition of a till layer below the littoral sediments and overlying bedrock. Small hyperbolas below the littoral sediments are indeed typical GPR characteristics related to boulders or cobbles found in till. A till layer whose top varies in depth from 0 (surface outcrops) to 22 m is generally present above bedrock throughout most of the study area and its low permeability ($\sim 10^{-9}$ m/s) limits groundwater exchange between the littoral sediments and the underlying bedrock. To the right of sand barrier (toward A), sediments are finer-grained, consisting of interbedded sand and silt layers as suggested by sediment types inferred from profiles CPT116, CPT24 and P17. GPR reflections within this region are also blurred at a few spots as a result of more electrically conductive sediments, such clayey silt, which generally limits GPR signal penetration. Hence the basal till layer cannot be identified in these profiles because fine-grained sediments limit the penetration of the GPR signal through the littoral sequence. From other GPR surveys, we also observe a general decrease in penetration depth as we move toward the southeast (right, away from A), which indicates that sediments are fining shoreward (from sea to land) due to a reduction in depositional energy in this backshore setting. Finally, the water table is approximately 1 m below ground surface, as indicated by water level in wells (CPT26, CPT24 and P17), but cannot be identified from the GPR profile in Figure 4 due to its proximity to the surface.

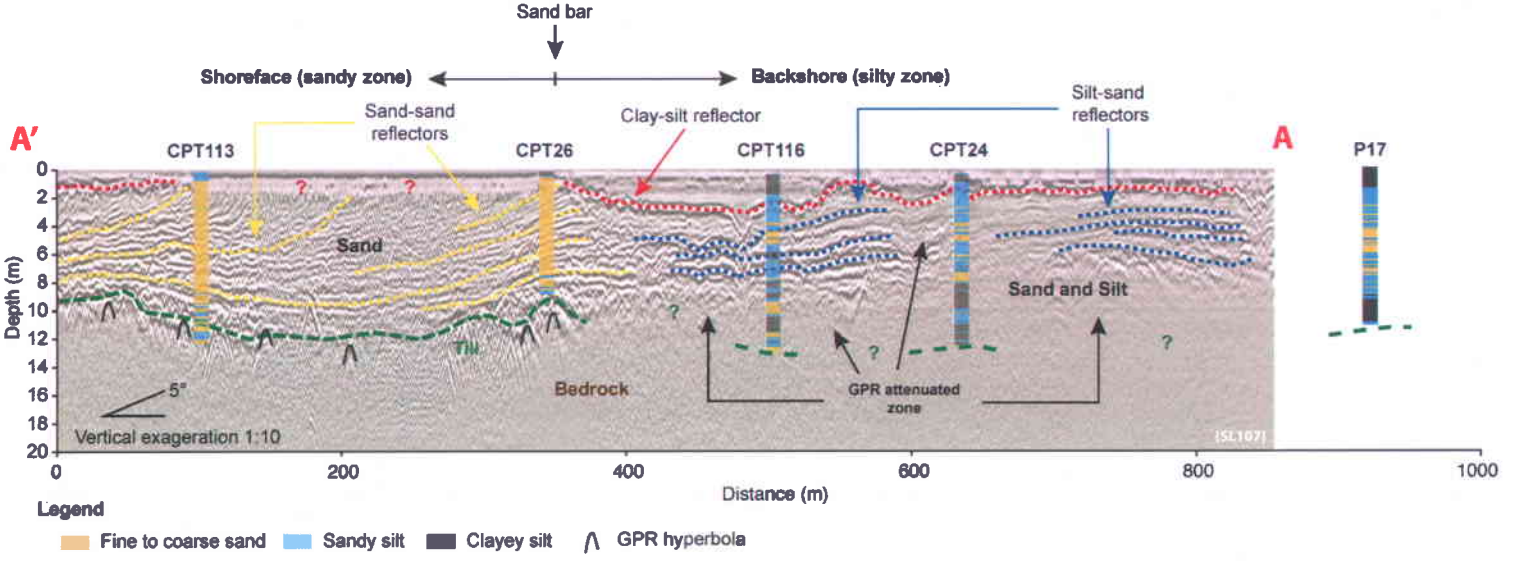


Figure 4 GPR section perpendicular to paleo-shoreline ridges along with profiles of sediment textures derived from cone penetrometer test data for soundings CPT113, CPT26, CPT116, CPT24 and P17 (based on the CPT interpretation chart of Fellenius and Eslami 2000). GPR section and sounding locations are shown in Figure 2c.

2.4 CPT/SMR soundings

On the basis of previously recognized GPR-facies, 53 CPT/SMR soundings were located (Figure 2c and Table 1) to allow further geological interpretation of the GPR surveys and to acquire mechanical and electrical properties of sediments and establish relations with hydraulic measurements. Direct-push soundings were carried out using a Geotech 605-D rig equipped with a CPT system including pore pressure measurement combined with a SMR probe. As illustrated in Figure 5 (red lines) for location P17 (shown in Figure 4), CPT/SMR probes allow the simultaneous measurement of two mechanical and two electrical properties of sediments at vertical resolutions that range from 3 to 17 cm (Table 1). A 15 cm^2 penetrometer cone with a 60° conical tip was used according to ASTM D5578-95 standards (ASTM 2000). The penetrometer is advanced vertically into the soil at a constant rate of 2 cm/s, though this rate can vary when compact layers are met. Inside the probe, two load cells independently measure the vertical stress against the conical tip and the side friction along the sleeve (Lunne et al. 1997). The vertical resolution of the tip stress (T) and sleeve stress (S) are 4 cm and 17 cm, respectively. A pressure transducer in the cone is also used to measure the pore water pressure as the probe is pushed into the ground. Pore pressure is an indicator of the presence of clay and was used to correct T data. The SMR probe is composed of four electrodes that are connected directly behind the penetrometer (Shinn et al. 1998). The inner two rings are used to measure soil capacitance and the spacing between the two rings is 3 cm. The soil moisture probe operates at 100 MHz, thereby reducing the effects of soil electrical conductivity on the measurement. The instrument measures shifts in the high frequency signal as it passes through the soil and this recorded signal is then empirically transformed into soil moisture content or dielectric constant (D) value. The electrical resistivity (R) measurement employs the outer two rings of the SMR probe, that are spaced 9 cm apart, to apply the current and measure the voltage drop (pole-pole configuration). The probe operates at a frequency of 1 KHz to avoid soil polarization effects.

As previously mentioned, Figure 4 shows, along with the GPR section, sediment types inferred from CPT data to further illustrate the general sediment sequence and architecture at the St-Lambert site. Pseudo-sediment profiles are obtained from CPT data using mechanical properties (T and S) of sediments based on the classification proposed by Fellenius and Eslami (2000). The combined interpretation of both CPT/SMR soundings and GPR surveys allowed the recognition of two distinct hydrogeological contexts over the study area, as indicated by the sandy and silty zones in Figure 4. The approximate limit between the sandy and silty zones is indicated in Figure 2c (white dotted line). Figure 2c also shows that CPT/SMR sounding

locations are fairly evenly distributed over the study area in order to cover the full range of mechanical and electrical properties and to map the gradual change in sediment textures and better define lateral continuity in aquifer material, as shown in Figure 4. Thus, indirect geological information obtained from pseudo-sediment textures derived from CPT data was used to guide preliminary site investigations before cores and hydraulic tests were available.

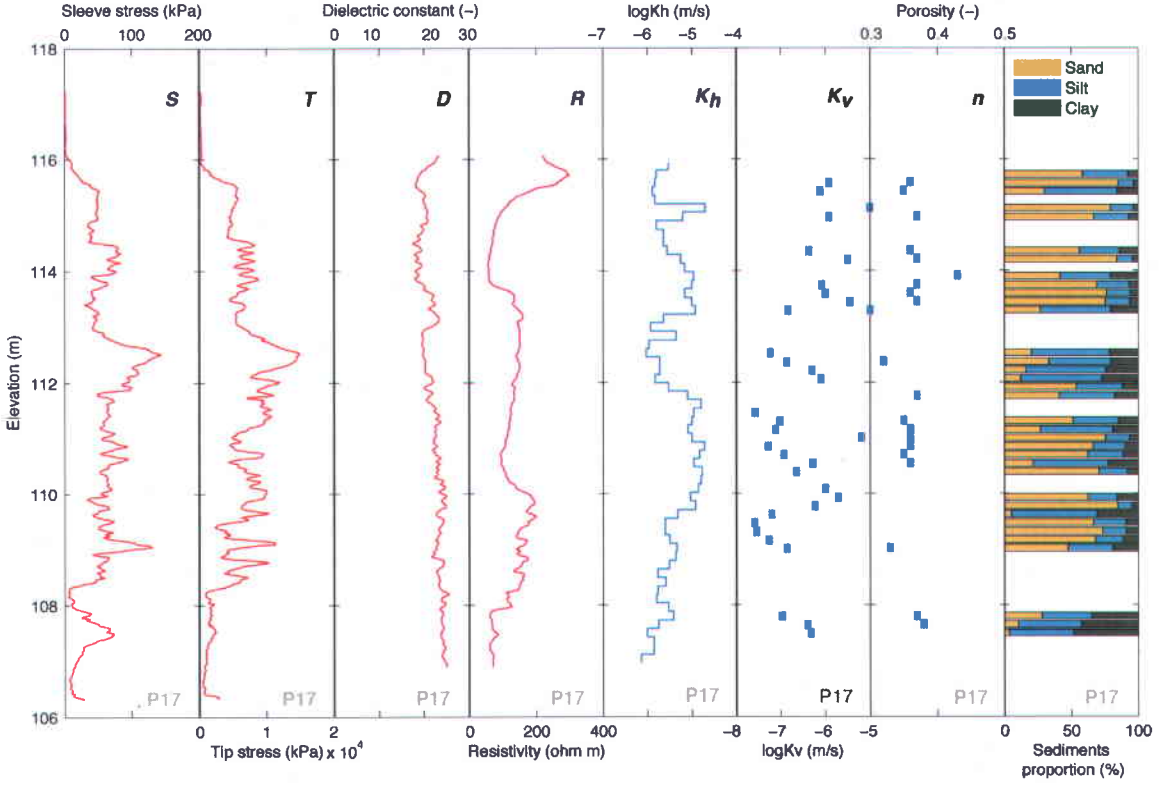


Figure 5 Example of CPT/SMR (red lines), hydraulic properties (blue line and squares) and lithological profiles obtained at location P17. CPT/SMR parameters are sleeve stress (*S*), tip stress (*T*), dielectric constant (*D*) and electrical resistivity (*R*). Horizontal (K_h) and vertical (K_v) hydraulic conductivities were obtained from multilevel slug tests in the P17 observation well and laboratory permeameter tests on sediment samples, respectively. Porosity (*n*) and grain sizes were obtained on the same sediment samples used for permeameter tests. Proportions of sand, silt and clay are based on grain size analyses using Wentworth (1922) grain size classes. P17 location is shown in Figure 2c.

2.5 Direct-push well installation

Based on CPT/SMR data obtained in real-time during sounding operations, 25 of the 53 direct-push sounding locations were selected for the installation of observation wells to be used, among other things, to carry out hydraulic tests (Figure 2c and Table 1). All wells were installed into the same hole created by the CPT/SMR sounding to obtain colocated hydraulic and direct-push data, which thus reduces the uncertainty

in data analysis related to disparities in interval measurements between direct and indirect data. Each observation well was also installed immediately following the CPT/SMR sounding with the same direct-push rig (Geotech 605D) to minimize mobilization and demobilization of the rig.

The installation of direct-push wells followed the protected screen installation procedure proposed by ASTM (2004). This well installation procedure does not require the use of sand-pack and the well screen is in direct contact with sediments, which is more suitable for hydraulic tests carried out over small intervals because it reduce skin effects on test data. The installation procedure involves, first, the hammering of a 7.72 cm OD (3 in., Geotech) metal casing with an expendable point at its base (Figure 6). The point is used to avoid the entry of sediments while hammering down the casing. When the metal casing is at the desired depth, the 5.1 cm ID (2 in) observation well (PVC tubing and screen) is inserted inside the metal casing, thus sitting on the expendable point. Finally, the metal casing is withdrawn to the surface to expose the observation well to the formation with the expendable point remaining at the bottom of the well. The hammered metal casing and PVC tubing diameters should have a tight fit to ensure a complete collapse of sediments on the observation well tubing after removal of the metal casing.

Figure 6 Photos of a direct-push well installation: (a) PVC tubing (white) inserted within the metal casing that was hammered down through sediments; (b) expendable point placed at the base of the metal casing to avoid sediments entrance; and (c) installed PVC tubing before surface grouting (the tubing is solid in its upper portion but slotted over most of the aquifer). See also Paradis et al. (2011) for details on direct push well installation, including their Supporting Information.

For this study, we installed fully screened observation wells across the whole saturated zone to provide continuous profiling of hydraulic properties with hydraulic tests. Screen length for all wells varied between 4.6 to 16.8 m. Direct-push wells were also thoroughly developed after their installation to unclog screens and surrounding sediments related to well installation operations (Butler 1998; Henebry and Robbins 2000). Each well was then developed by a pumping-surging operation at 0.5-m intervals using an inertial pump (Waterra) with a development ring until turbidity was no longer observed in the discharged water (Paradis et al. 2011).

2.6 Sediment sampling

Generally after each CPT/SMR sounding, sediments were also sampled at a location 1-2 m away from the location of the sounding. Sampling was carried out with a piston-rod operated sampler (Geotech Macro-Core Sampler) allowing the recovery of a 38 mm diameter and 1.52 m long sediment sample inside a PETG liner (Table 1). This method was chosen mainly because it preserves sediment stratification within the core, which is fundamental for the analysis of sediment structures and reliable estimates of K_v with laboratory permeameter tests. Samples were taken with the same direct-push rig (Geotech 605D) used for CPT/SMR soundings and well installation, which allows a series of direct-push operations at the same site in a single mobilization of the rig. To outline the process of a typical drilling sequence as described above, it usually took less than half a day per site, meaning two or more visited sites a day. Sample intervals were selected based on real-time CPT/SMR responses and more than 75 cores were selectively taken. Sediments recovery was estimated at 83% overall.

2.7 Multilevel slug tests in wells

K_h values used to establish relations with CPT/SMR data, an example of which is shown in Figure 5, were obtained by multilevel slug tests carried out into the fully screened direct-push wells. A multilevel slug test involves the use of packers to isolate a screened interval of a well to conduct a slug test (Sellwood et al. 2005; Zemansky and McElwee 2005; Ross and McElwee 2007). Slug tests were performed using a pneumatic method to induce an initial lowering of the water level (Levy and Pannell, 1991) and hydraulic responses were interpreted using the Bouwer and Rice (1976) method. A more detailed description of the test assembly, field and test analysis procedure is provided by Paradis et al. (2011). Multilevel slug tests were made at 15-cm intervals to reduce uncertainty related to scale disparity between hydraulic and direct-push data and to be as close as possible to the assumption of K isotropy. A total of 280 intervals were then

tested for K_h (Table 1). The whole set of tested intervals was also selected to have representative K_h data for the range and occurrence of CPT/SMR responses observed over the study area (Paradis et al., 2014b).

As part of this study, in order to develop a time-efficient hydraulic characterization approach, Paradis et al. (2011) adapted to unconsolidated formations the use of flowmeter tests that are commonly deployed in fractured media. A flowmeter test involves the measurement of groundwater inflow into a well at different depths during steady-state pumping conditions to obtain a vertical profile of K_h (Hess 1986). The main adaptation for flowmeter measurements in unconsolidated aquifers is the use of long-screen direct-push wells installed without sand-pack, as the installation was previously described, to avoid hydraulic short-circuits that generally bias flowmeter tests in sand-packed wells. The study of Paradis et al. (2011) shows that the difference in log K_h values obtained from flowmeter and multilevel slug tests for the same 15-cm intervals (n=128) is consistently below 10% with a flowmeter profile obtained in a fraction of the time needed to obtain the same profile with slug tests, as will be discussed later.

2.8 Laboratory permeameter tests

To verify the assumption of K isotropy over 15-cm intervals, 59 15-cm long sediment samples were selected to estimate K_v in the laboratory with permeameter tests (Figure 5 and Table 1). In order to minimize disturbance of original sediments, an automated falling head laboratory permeameter designed to directly accommodate the 38 mm diameter PETG liner containing sediment samples was set up to determine K_v values. Each 1.52 m long sample was subdivided into 15-cm long subsamples according to the tested intervals with multilevel slug tests. Duplicate tests carried out on a variety of subsamples indicated a high degree of reproducibility and all K_v values were temperature-corrected to a groundwater temperature of 8 °C representative of field conditions. Paradis and Lefebvre (2013) provide a more detailed description of the experimental assembly and the test analysis procedure.

As shown in Figure 5, the comparison of K_h with K_v measurements at similar depths reveals that the assumption of isotropy at 15-cm interval does not generally hold, with K anisotropy up to two orders of magnitude. In addition, K_h estimates using the Bouwer and Rice (1976) method as previously discussed neglected K anisotropy, which may result in an underestimation of K_h values (Zlotnik 1994). Thus discrepancy between K_v and K_h values could be larger than observed in Figure 5. This suggests that both K_h and K_v should be estimated for each 15-cm interval to provide representative conditions controlling flow

and mass transport over the study site. Although under suitable conditions permeameter tests can represent a reliable method to estimate K_v, the difficulties in the experimental procedure related to soil sample collection and manipulation may, however, limit the capability to obtain reliable estimations of K_v from laboratory measurements (Stienstra and van Deen 1994). Moreover, permeameter tests may also be time consuming when many measurements are needed (Klute and Dirksen 1986) and are thus mostly restricted to a small number of samples.

As part of this characterization effort, two new field tests involving undisturbed volumes of investigation were then developed to provide suitable basis for K_v estimations. The first test is a vertical interference slug test which involves a slug test carried out in a single well between a stress and an observation interval that are vertically isolated with a three-packer assembly (Paradis and Lefebvre 2013). An instantaneous pressure pulse is induced in the stress interval and resulting heads are recorded in both the stress and the observation intervals for the analysis. The second test consists of tomographic slug tests, a generalization of vertical interference slug tests to multiple wells, which indeed involves recording head responses in additional observation intervals in adjacent wells (Paradis 2014; Paradis et al. 2014a). For the two slug-based tests, a fully-transient analysis that makes use of the entire hydraulic head records is proposed to reconstruct heterogeneous profiles of K_h, K anisotropy (or K_v) and specific storage (S_s). In proof-of-concept field studies (Paradis and Lefebvre 2013; Paradis 2014), results indicate that profiles of K_v values obtained with vertical interference slug tests and tomographic slug tests follow similar patterns with depth that are similar to profiles obtained from laboratory permeameter tests on sediments samples.

2.9 Laboratory porosity measurements and grain-size analysis

After permeameter testing, subsamples were recovered and saturated to measure total porosity (n) (Figure 5). The n was estimated by subtracting wet and dry weights of each 15-cm subsample (Table 1). Subsamples were oven-dried at 50 °C for more than 7 days. The total number of subsamples tested for n was 43. Geological description and grain-size analysis by laser light scattering (Vitton and Sadler 1997) were also carried out on the subsamples (Figure 5 and Table 1) to provide geological information complementary to the hydraulic and geophysical measurements.

3 Data integration

3.1 Overview of the data integration approach

The general methodology adopted in this study for data integration follows the flow chart depicted in Figure 7. The application of the workflow for data acquisition presented in Figure 1 led to the collection of a training dataset representative of hydrogeological conditions in the St-Lambert study area (see the "Training DataSet" box in Figure 7). This training dataset provides mainly colocated hydraulic measurements of K_h, K_v, n and direct-push parameters S, T, D and R. This training dataset is used to define HF models, hydro-geophysical relations and hydraulic relations through the training of a learning machine (see the "Definition of Relations" box in Figure 7). The definition of HF models and hydro-geophysical relations between K_h and CPT/SMR data have been described in detail by Paradis et al. (2011; 2014b) and only the main results are summarized in this paper. Once HF models and hydro-geophysical relations are defined and validated, CPT/SMR data for which hydraulic data are not available are used to convert direct-push data into HF classes and K_h values, while hydraulic relations are used to estimate K_v and n from K_h data (see the "Testing and Prediction" box in Figure 7). Finally, the spatial distributions of HF, K_h, K_v and n over the study area are obtained through interpolation of observed and converted data (see the "Interpolation" box in Figure 7). In the remainder of this section, we thus present the application of the data integration workflow in Figure 7 for the section A-A' in Figure 4. In the present paper, the application is illustrated only for section A-A', but the application can be further extended as the training dataset presented in this section was acquired to be representative of the entire study area.

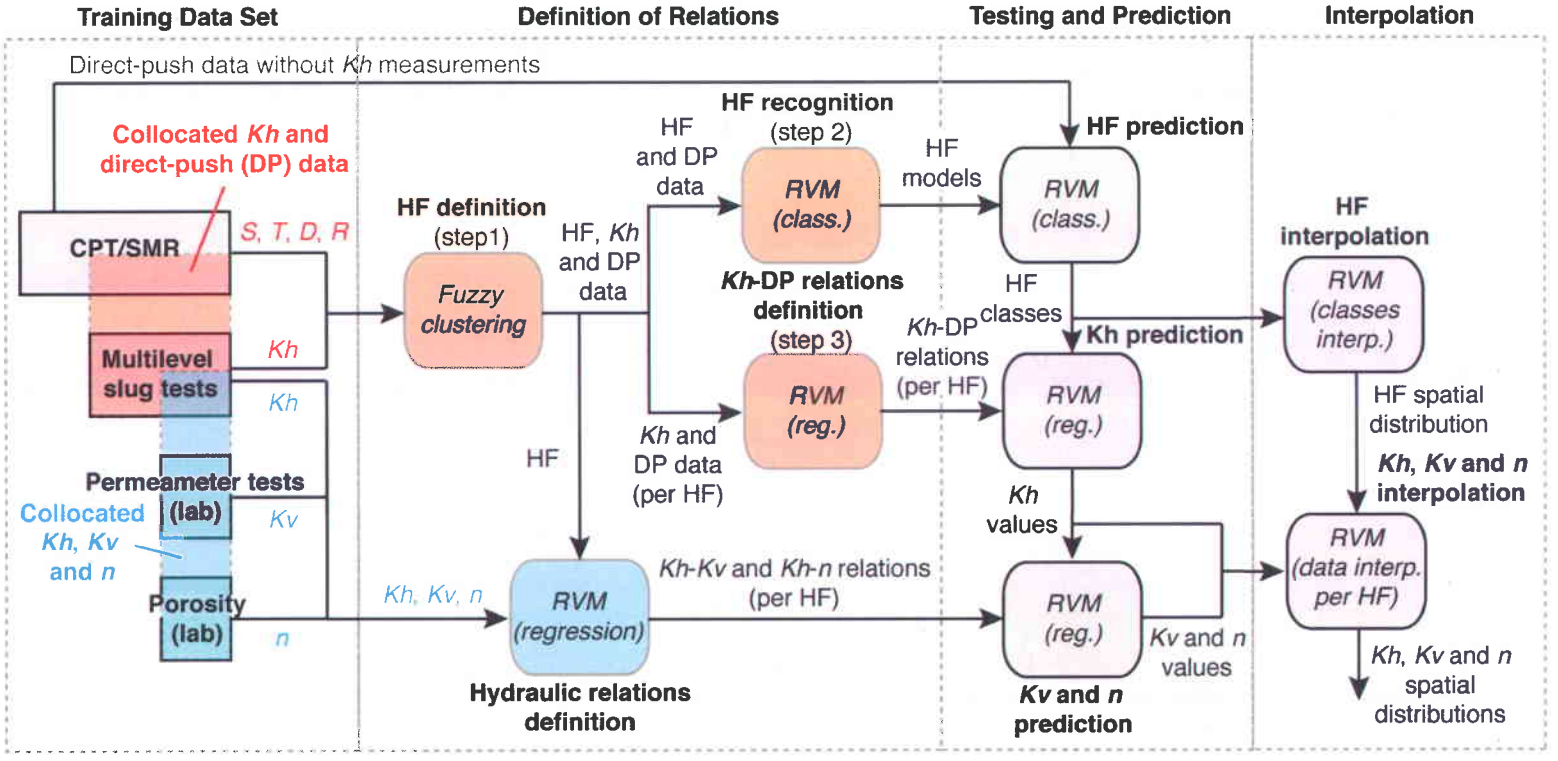


Figure 7 General data integration workflow for the definition of hydraulic heterogeneity, as applied to the St-Lambert site. The blue and the red rectangles in the "Training Data Set" dotted box indicate that the measurements are collocated. Direct-push data without K_h measurements are data used to make predictions or to test prediction accuracy using a fraction (20%) of the training data set. Where S is sleeve stress, T is tip stress, D is the dielectric constant, R is electrical resistivity, K_h is horizontal hydraulic conductivity, K_v is vertical hydraulic conductivity, n is porosity, HF is hydrofacies and RVM is Relevant Vector Machine. Abbreviations: class. (classification), reg. (regression), interp. (interpolation).

3.2 Training data set

To be properly statistically compared to K_h measurements and define reliable HF models and hydro-geophysical relations (Isaaks and Srivastava 1989), CPT/SMR data were first resampled to a regular grid and then rescaled to a common vertical spatial support of measurement with K_h data. Indeed, although CPT/SMR measurements were taken at a regular time interval, the rate of penetration is not always constant. CPT/SMR data were thus resampled on a regular grid of 2 cm using trapezoidal integration (Davis 1973). Then, to account for the different vertical support of measurements of each CPT/SMR parameter (Table 1), all direct-push parameters were upscaled to the same 15-cm intervals of K_h measurements using a moving average. The number of data points used in the moving average was adjusted according to the original length of the support of measurement of each CPT/SMR parameter. Descriptive statistics for hydraulic and CPT/SMR parameters of the training dataset are presented in Table 2. Because the range of values for most parameters varies over a few orders of magnitude, a logarithmic transformation was applied to normalize their distribution, as illustrated in histograms of Figures 8 and 9. The range of values for each colocated parameter is similar, except for D and n that have somewhat narrower distributions than the other parameters.

Table 2 Descriptive statistics for hydraulic (K_h: horizontal hydraulic conductivity (m/s); Kv: vertical hydraulic conductivity (m/s); n: total porosity (volume fraction)) and CPT/SMR parameters (S: sleeve stress (kPa); T: tip stress (kPa); D: dielectric constant (-); R: electrical resistivity (ohm·m)) used as training datasets.

Parameter	Nbr.	Mean	Median	Min.	Max.	Range	Standard Deviation	Skewness	Kurtosis
log$S$	280	1.74	1.76	0.15	2.51	2.36	0.31	-0.96	2.94
log$T$	280	3.94	4.00	2.41	4.42	2.01	0.29	-1.44	3.63
log$D$	280	1.39	1.39	1.24	1.53	0.29	0.06	0.36	0.85
log$R$	280	2.16	2.16	1.61	3.08	1.47	0.25	1.09	3.25
logK_h	280	-5.04	-5.05	-6.24	-3.92	2.32	0.57	0.09	-1.10
logK_v	59	-6.40	-6.39	-7.56	-4.84	2.72	0.71	0.40	-0.62
log$n$	43	-0.45	-0.45	-0.53	-0.37	0.16	0.03	-0.09	0.95

Hydraulic parameters in particular, K_h and K_v, vary over two orders of magnitude whereas n varies over less than a quarter of one order of magnitude, as commonly observed in many hydrogeological investigations (Castany1963; Bear 1979; Freeze and Cherry 1979). This shows that the primary focus of aquifer characterization should be on K because the range of K variation is generally larger than the range for n, and because K acts to change both the magnitude and direction of advective transport, as opposed to n which changes only the magnitude (Koltermann and Gorelick 1996).

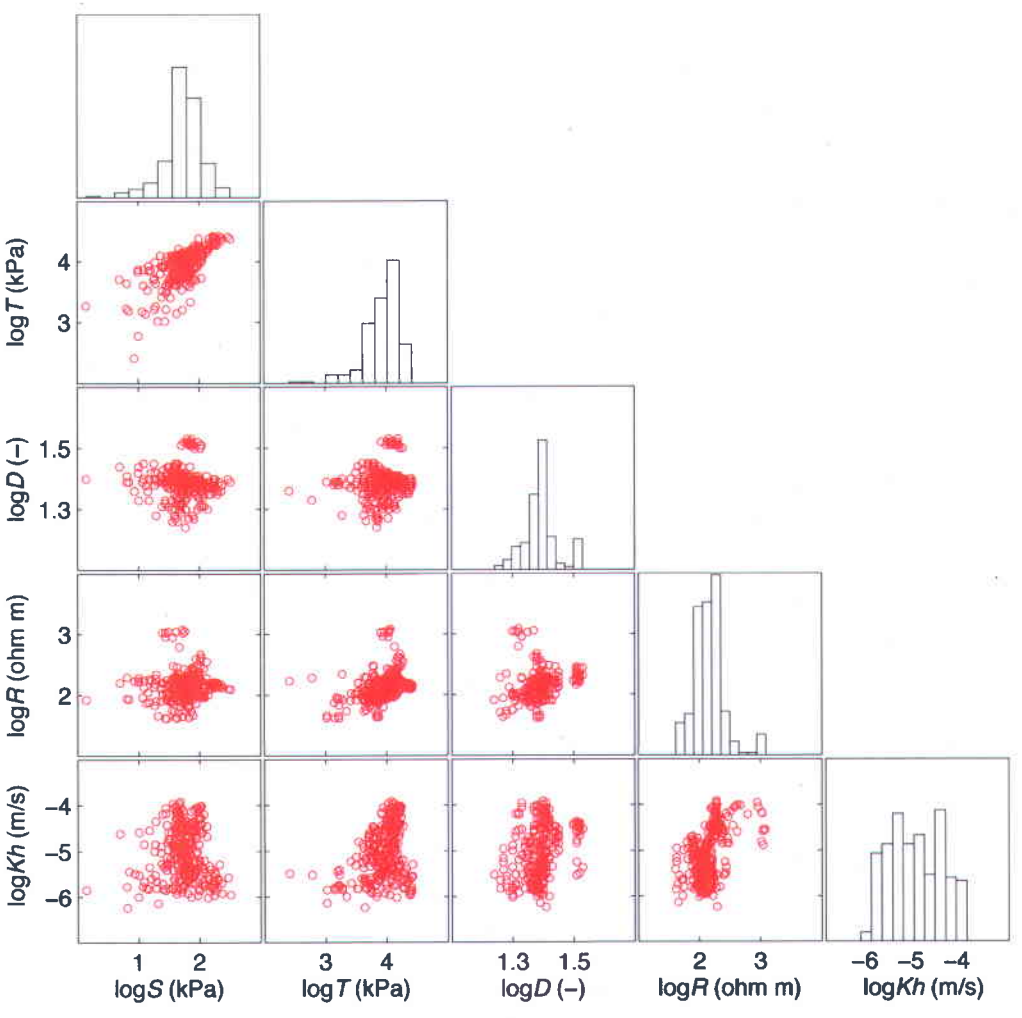


Figure 8 Matrix scatter-plots with histograms for horizontal hydraulic conductivity (K_h) and CPT/SMR parameters (*S*: sleeve stress; *T*: tip stress; *D*: dielectric constant; *R*: electrical resistivity) for the hydro-geophysical training data set. The total number of collocated intervals is 280 (80% of data points were used for Relevant Vector Machine training and 20% for testing; see Figure 7).

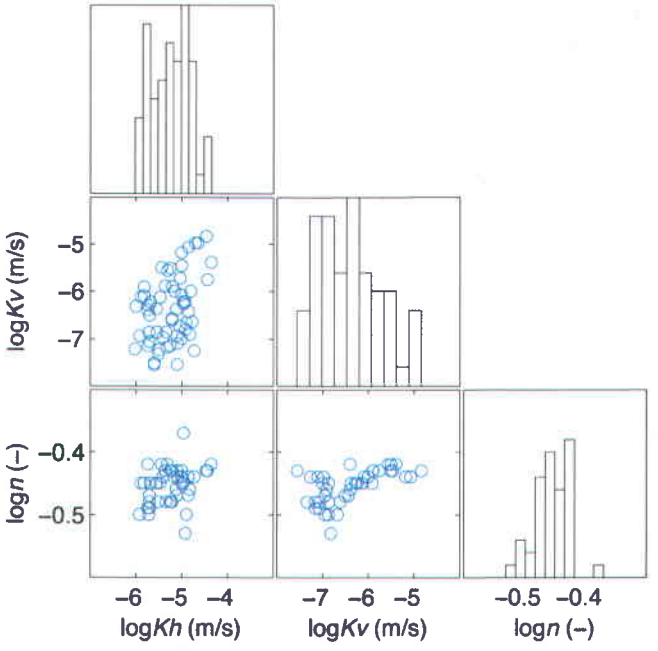


Figure 9 Matrix scatter-plots with histograms for horizontal hydraulic conductivity (K_h), vertical hydraulic conductivity (K_v) and total porosity (n) for the hydrogeological training data set. The total number of collocated intervals for K_h and K_v is 59, and 43 for n.

The scatter plots in Figures 8 and 9 also illustrate the relations between all hydraulic and CPT/SMR parameters with the corresponding correlation matrix presented in Table 3. As shown in Table 3, the correlations between K_h and each CPT/SMR parameters are low to moderate, except for S where the correlation is null. For the three hydrogeological parameters alone, the correlations are moderate and vary from 0.30 to 0.59. This lack of strong linear correlations between the various hydro-geophysical parameters suggests that the definition of relations to predict hydraulic information from CPT/SMR data is inherently difficult. In the next section, we will present the learning machine approach that was adopted to model non-stationary and heteroscedastic relations.

Table 3 Correlation matrix for hydraulic hydraulic (K_h: horizontal hydraulic conductivity (m/s); Kv: vertical hydraulic conductivity (m/s); n: total porosity (volume fraction)) and CPT/SMR parameters (S: sleeve stress (kPa); T: tip stress (kPa); D: dielectric constant (-); R: electrical resistivity (ohm·m)) used as training datasets.

Parameter	log$S$	log$T$	log$D$	log$R$	logK_h	logK_v	log$n$
log$S$	1.00						
log$T$	0.70	1.00					
log$D$	-0.02	0.16	1.00				
log$R$	0.06	0.40	0.21	1.00			
logK_h	-0.02	0.29	0.33	0.52	1.00		
logK_v	-0.27	-0.02	-0.09	0.20	0.37	1.00	
log$n$	-0.24	-0.19	-0.25	-0.29	0.30	0.59	1.00

3.3 Definition of hydrofacies

To handle complex relations that may be present in the hydro-geophysical training dataset, a learning machine approach was adopted. The training of the learning machine, as schematically illustrated by the red boxes in Figure 7, is a three-step procedure that uses classification (unsupervised and supervised) and regression (supervised) algorithms to build predictors for HF and K_h from direct-push data. The first step in the training process is the definition of HF units through unsupervised fuzzy clustering (Dunn 1973; Bezdek 1981; Gustafson and Kessel 1979) using colocated K_h and CPT/SMR data. The main rationale behind the use of clustering is that various geological materials may have discontinuous and non-stationary hydro-geophysical behaviors and clustering allows the grouping of similar hydro-geophysical characteristics without any prior geological knowledge. For this study, only K_h data are integrated in the clustering process because the smaller number of K_v and n measurements would restrict the use of the entire K_h dataset. Although independent relations for K_v and n with CPT/SMR data could be developed similarly as for K_h, an a posteriori integration of K_v and n data with colocated K_h measurements is proposed instead, as illustrated in Figure 7.

The accuracy of predicted hydraulic information from direct-push data depends largely on the degree of sensitivity of the geophysical parameters relative to the hydraulic parameter of interest. For geophysical parameters highly sensitive (perfectly correlated) to changes in hydraulic parameter values, there will be no ambiguity in hydraulic information estimates from noise-free geophysical data. When the degree of sensitivity between hydraulic and geophysical parameters is less than perfect and the data are noisy, the statistical relations between hydraulic and geophysical parameters become non-unique (Rubin and Hubbard 2005). The predictive capability of the geophysical tool to predict hydraulic information will also deteriorate because different sets of geophysical parameter values can lead to the same value of the hydraulic parameter of interest or a unique combination of geophysical parameter values can results in a range of hydraulic parameter values. As described in Paradis et al. (2014b), the overlapping of HF units in the geophysical space was used to express the non-uniqueness in the hydro-geophysical relations. A systematic procedure was then applied to search for the combination of CPT/SMR parameters that minimizes the overlap between the identified HF units in order to obtain the best predictive capability of the learning machine. The number of HF units and the similarity clustering measures were also included in this search procedure. The results of the systematic search procedure show that the predictive capability of CPT/SMR data generally increases with the number of direct-push parameters, suggesting that non-uniqueness between K_h and direct-push parameters is minimized with more parameters (Figure 10). The clustering with direct-push parameter T, D and R, with 4 HF units and using a squared Mahalanobis clustering distance (Gustafson and Kessel 1979) was the clustering experiment with the best predictive capability. Note that the Fellenius and Eslami (2000) classification in Figure 4 uses mechanical properties of CPT (S and T) to derive pseudo-sediment textures and using this classification to predict K_h would result in inaccurate estimates, as shown in Figure 10 with misclassification error over 40%. That justified the proposed learning machine to define a site-specific classification to predict hydraulic properties for the study site, taking advantages of all available CPT/SMR parameters.

The resulting data distributions for each HF unit are summarized in Figures 11a-d. As seen in those figures, the median and range of values for K_h and each direct-push parameter are fairly distinct between HF units with only a few outliers. In particular, the median values for K_h gradually increase from HF1 to HF4 with slight overlaps between HF units that may be attributed to the complexity of the hydro-geophysical responses and to the transitional nature of sediments in the littoral depositional environment. In addition, each HF unit presents distinct distributions of K_h and direct-push parameters as expected from the various sediments composing the aquifer that may have different hydro-geophysical responses. For instance, HF4 has higher median values for T, D, R and K_h, whereas HF1 has the lowest values for the same parameters.

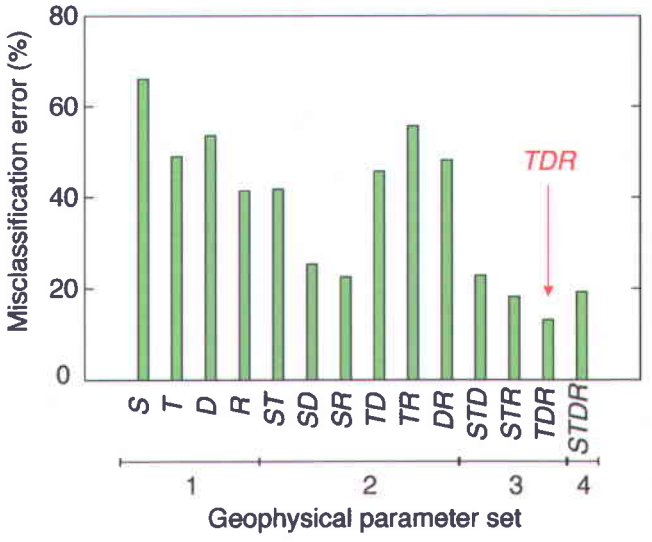


Figure 10 Misclassification error associated to the non-uniqueness between hydraulic and CPT/SMR parameters for all combination of geophysical parameters. The clustering for this figure is for 4 hydrofacies units using a squared Mahalanobis clustering distance. The red arrow indicates the geophysical parameter set used for this study.

Data distributions for K_v and n measurements associated a posteriori to each HF unit are also presented in Figures 11e-f. Based on median, 25% quantile and 75% quantile values, HF units for K_v and n are generally distinct from each other and follow essentially the same trend as for K_h, except for HF1. The comparison of median K_h and K_v values for each HF unit indicates that K anisotropy ranges from 0.05 to 0.2 for HF1 and HF4, respectively.

Finally, grain-size analyses were used to provide a geological meaning to the HF units, as illustrated in Figure 11g. According to the trend in minimal and maximal mean grain-sizes (M_z), coarser sediments are associated with more permeable and porous HF units. It should be noted that the volume of sediments used for the grain-size analysis is much smaller than the volume of the 15-cm subsample and then M_z values may show some fluctuation within a subsample. This may explain the large variations in M_z within each HF unit, as expressed by the spread of 9% quantile and 91% quantile values. Also, the small number of sediment samples for HF1 (n=4) may explain the deviation of M_z for this unit with respect to the trend observed for HF2 to HF4.

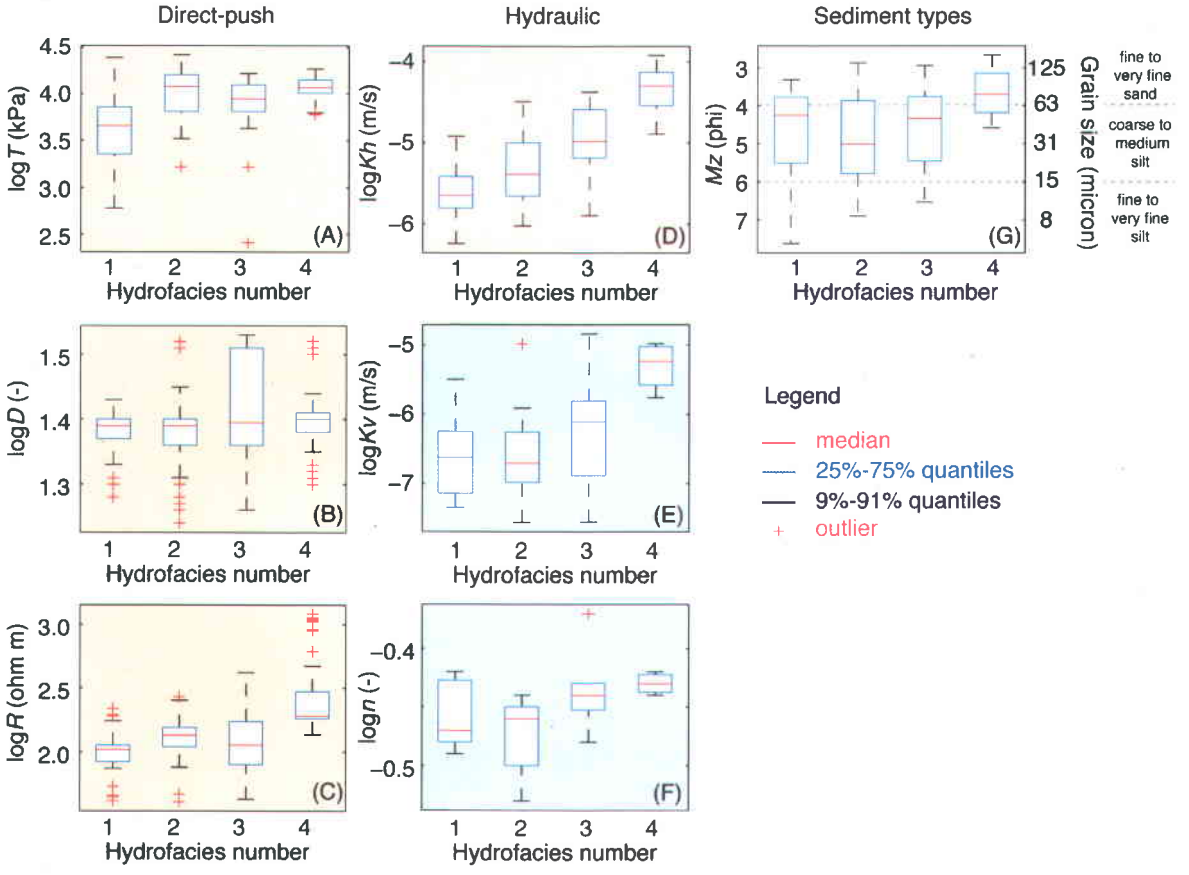


Figure 11 Data distribution for each of the four hydrofacies resulting from the clustering experiment with the best predictive capability (a-d). This clustering uses tip stress (T), dielectric constant (D) and resistivity (R) along with horizontal hydraulic conductivity (K_h). Distributions of vertical hydraulic conductivity (K_v) and porosity (n) for each hydrofacies are from a posteriori integration with collocated direct-push data (e-f). Sediment types of each hydrofacies are based on grain-size analyses (g). Sediments classification follows the Wentworth (1922) chart: the phi-scale is a base-2 logarithmic scale with the negative exponent of the grain size in mm and the mean grain-size of sediments (M_z) evaluated from (phi14+phi50+phi84)/3 according to Krumbein and Sloss (1963). A small M_z value corresponds to coarse sediments.

3.4 Definition of hydrofacies models, hydro-geophysical relations and hydraulic relations

Although the clustering process revealed associations among the hydro-geophysical data, unsupervised clustering cannot be used to make predictions. Thus, to obtain predictive models, previously defined HF units and their related hydro-geophysical data are then trained in a supervised manner, as illustrated in Figure 7. Two training processes are run in parallel. First, direct-push data for CPT/SMR parameters retained during the clustering process (T, D and R) along with corresponding HF labels are trained together using a relevance vector machine (RVM) for classification to be able to recognize HF units from direct-push

data (Figure 7, step 2). Second, K_h measurements and direct-push data associated with each HF unit (T, D and R) are used to define hydro-geophysical relations specific for each HF unit using a RVM for regression (Figure 7, Step 3). The rationale for selecting RVMs is their ability to produce relations with good generalization capability with sparse and nonlinear data (Khalil et al. 2005; Camps-Valls et al. 2006; Samui 2007; Ghosh and Mujumdar 2008). The main idea behind RVM is to use a kernel function to linearize a nonlinear problem into a higher dimensional data space and solve the linear problem within a Bayesian framework, which returns a nonlinear regression in the original data space (Tipping 2001).

To train the learning machine and test the obtained relations, the hydro-geophysical training dataset was split into training and testing sets using 80% and 20% of the data, respectively. The optimal model parameters (e.g., kernel type and kernel window length) for both RVM classification and RVM regression were retrieved using the training set through cross-validation and predictive performances were assessed using the testing set. The testing process followed the same procedure used to predict HF unit and K_h from CPT/SMR data, as illustrated in Figure 7, where for a particular vector of direct-push data the HF class is first identified using HF models, and then the associated hydro-geophysical relation is used to predict the corresponding K_h value using the same data vector. To avoid the bias in selecting the testing set, 100 different combinations of training and testing sets were randomly selected and applied in turns. Results of the verification process for the 100 training and testing sets have shown that the median percentage of misclassified HF unit resulting from the classification with the RVM is 14% (Figure 12a) and this misclassification error was attributed to the non-uniqueness that exists between K_h and direct-push data. Meanwhile, the median linear coefficient of correlation between field and predicted K_h values from the RVM for regression was 82% (Figure 12b).

Although K_h data are generally easy to acquire with conventional hydraulic tests to obtain a representative dataset for the study area, HF unit should ideally also be characterized by K_v and n values to be meaningful for flow and transport modeling. Those K_v and n measurements are however difficult to acquire because of the restriction generally imposed by sediment sample collection and manipulation for laboratory measurements, as previously discussed in the section on data acquisition. This therefore restricts the acquisition of large datasets for those parameters and their direct integration with the larger training dataset of K_h and direct-push data at the clustering stage would reduce the reliability of the defined HF units because of the reduced size of the training dataset. Then, it is proposed to integrate K_v and n data a posteriori after the clustering step by association with colocated K_h data, as illustrated in Figure 7. Hydraulic relations that link K_v and n measurements to K_h were thus developed for each HF unit using independent RVMs for

regression and a total of eight relations were available to predict K_v and n from K_h data. Note that since K_v is defined through a regression with K_h, K anisotropy for each HF unit as well as within each HF unit can be different and comparable with the colocated measurements.

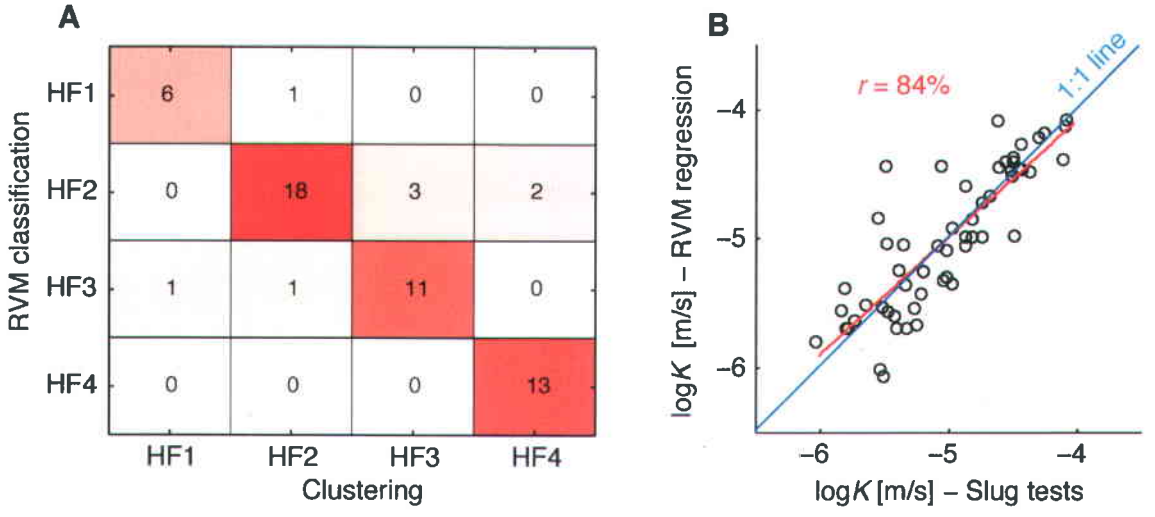


Figure 12 (a) Confusion matrix comparing, with a testing data set (n=56 observations), the HF classification obtained by the relevance vector machine (RVM) classification to the original classification made by fuzzy clustering. The diagonal indicates the observations for which both classifications are identical (HF1 6/7, 86%; HF2 18/20, 90%; HF3 11/14, 79%; HF4 13/15, 87%; overall 48/56, 86%). Off-diagonal observations were misclassified by RVM classification. (b) Comparison of horizontal hydraulic conductivity (K_h) measured with multilevel slug tests with the estimation made using RVM regression with the same testing data set as shown in Figure (a).

3.5 Prediction and spatial distribution of hydrofacies and hydraulic properties

The learning machine being optimally trained, each vector of direct-push data available for the study area was converted into hydraulic information as shown in Figures 13a-d for each of the direct-push soundings available along GPR section A-A'. Direct-push data were resampled and upscaled to a common 15-cm interval before applying the sequential predictive procedure to estimate HF, K_h, K_v and n, as illustrated in Figure 7. Direct-push parameters T, D and R with 4 HF units, as previously discussed in the section on clustering, were used for predictions. Results of the conversion show distinct vertical profiles of HF, K_h, K_v and n for the zones in front and behind the sand barrier, similar to the trend observed in Figure 4 with pseudo-sediment profiles. Also, the converted direct-push data are generally in fairly good agreement with measured profiles of K_h, K_v and n, as illustrated in Figures 13b-d for location P17. Note that for the purpose

of predictions with the learning machine the whole training dataset was used to make use of all available information and thus increase the generalization capabilities of the RVMs.

Figures 13a-d also present spatial distributions of HF, K_h, K_v and n obtained from the interpolation of converted direct-push data along soundings. For the purpose of this paper, spatial distributions of hydraulic information (HF, K_h, K_v and n) were obtained by interpolators built with RVMs, as proposed by Smirnoff et al. (2008) for lithological reconstruction with a similar supervised learning algorithm. The construction of the interpolators is similar as the previous process of HF models and hydro-geophysical relations definition, except that hydraulic information and spatial coordinates of observations are used instead of hydraulic and geophysical data. For instance, a RVM for classification is used to interpolate HF units, using converted HF information along soundings and corresponding spatial coordinates (CPT113, CPT26, CPT116, CPT24 and P17 in Figure 13), whereas a RVM for regression is built for each hydraulic property (K_h, K_v and n). Note that this RVM-based approach can be qualified as a geometrical interpolation (for classification or regression) technique and its intent, as used in this study, is to provide a visualization tool to map the information contained in the converted CPT/SMR data. Then for the sake of the visualization, large kernel window lengths were used to provide more generalized spatial representations with fewer details. As shown in Figures 13a-d, thin layers that likely have small spatial continuity are not mapped and information that dominates the data set is favoured (smoothed). As briefly discussed later, geostatistical simulation of hydraulic information to constrain the heterogeneity of the study area is addressed by Brunet et al. (2012).

Figure 13a presents the most probable spatial distribution of HF units along the GPR section shown in Figure 4 resulting from the integration of HF classes and corresponding spatial coordinates into a RVM for classification. From this interpolation, we clearly see the transition from the sandy zone (HF4) to the silty zone (HF1-3) with the juxtaposition of the coarser HF4 over the finer HF units as can be expected from wave and current dynamics in a littoral depositional environment. Figures 13b-d also present results of the interpolation for K_h, K_v and n, respectively. Similarly to HF interpolation, hydraulic properties with corresponding spatial coordinates were integrated into RVMs for regression. For K_h, K_v and n interpolations, hydraulic properties were interpolated individually for each HF unit with independent interpolation parameters. In order to preserve the relationships between K_h, K_v and n observed in the training dataset, the same RVM interpolation parameters were used for all parameters. As illustrated in Figure 13b, the sandy zone is more homogeneous and permeable than the silty zone. Moreover, we observe a general decrease in K_h values shoreward (from A' to A) as can be expected from a reduction of the littoral energies as we move behind the sand barrier. Similar observations can also be made in Figures 13c-d for K_v and n, respectively.

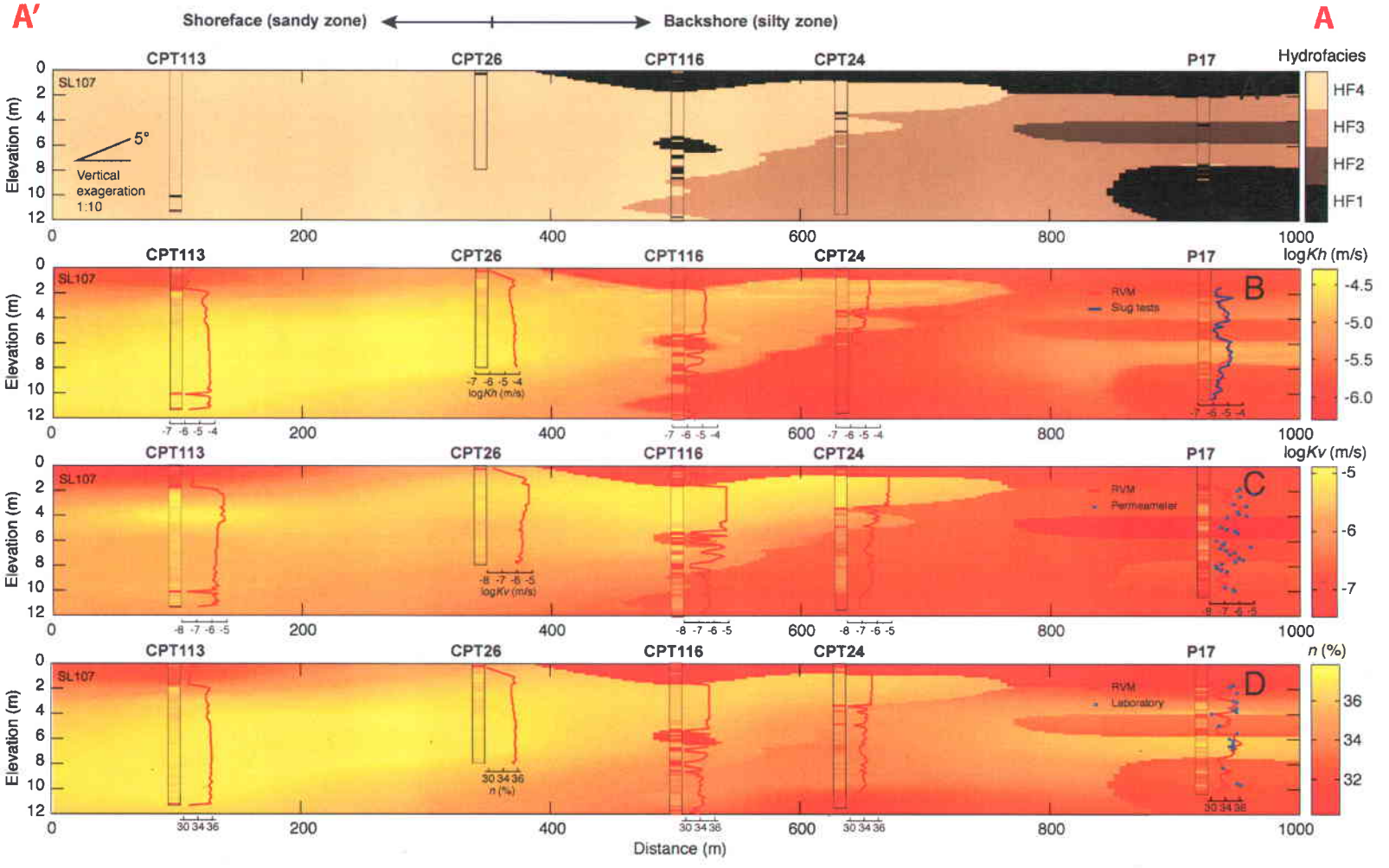


Figure 13 2D spatial distribution of (a) hydrofacies (HF) unit; (b) horizontal hydraulic conductivity (K_h); (c) vertical hydraulic conductivity (K_v); and (d) porosity (n) along GPR section A-A' (location in Figure 4) based on the interpolation of the recognized HF units and the estimated hydraulic properties based on vertical direct-push soundings using the approach shown on Figure 7. Colour intensities in (a) are proportional to median K_h values of each HF in Figure 11d.

4 Discussion

4.1 Time efficiency of Kh characterization methods

Although the benefits of the whole data acquisition and integration process may be difficult to quantify, it may be worthwhile comparing the time efficiency of the hydro-geophysical data integration approach to estimate K_h from direct-push data, as applied in the characterization effort presented in this paper, with various conventional hydraulic tests. The hydraulic tests are multilevel slug tests and flowmeter tests in wells and laboratory permeameter tests on soil samples. These methods have been extensively used over the study area, either for production or research purposes, which allows fair estimates of the time required for data acquisition and analysis. For comparison purposes, we evaluate the time required for each method to estimate K_h at 15-cm intervals along 50 10-m long vertical profiles (500 m in total), which corresponds to the approximate total number, average length and vertical resolution of CPT/SMR soundings that were carried in this study (Table 1). On the basis of the characterization work carried out with the four selected methods, we estimated of the time required for data acquisition and analysis, as summarized in Table 4 and illustrated in Figure 14. It should be noted that the results of this analysis are approximations that may vary according to the geological context and the experience of the field team.

As expected, permeameter tests carried out on sediment samples required more time than any other method (Figure 14) due to the collection in the field and subsequent laboratory preparation of the samples (Table 4). Moreover, although permeameter tests could be carried out continuously with depth within an aquifer, in practice partial sediment recovery may preclude the testing of certain intervals, as can be appreciated in Figure 5 where the K_v profile along P17 obtained from permeameter tests is discontinuous. It should be noted also that the assumption of K isotropy ($K_h = K_v$) should be met to obtain valid estimates of K_h values from permeameter tests because the flow induced with such tests is generally perpendicular to the naturally deposited sediments layers. As illustrated with profiles of K_h and K_v in Figure 5, sediment samples smaller than 15-cm would then have been required to get permeameter measurements representative of K_h for the study area. Alike permeameter tests, an approach based on the conversion into hydraulic information of lithology obtained from sediment observations (e.g., Klingbeil et al. 1999; Heinz et al. 2003; Kostic et al. 2005) would also be time consuming to characterize the aquifer at the level wanted for this analysis. Geological interpretation of sediment samples is nevertheless essential to understand the hydrogeological context and the degree of heterogeneity of the site.

Table 4 Time needed to acquire 500 m of horizontal hydraulic conductivity measurements (K_h) at 15-cm intervals for different methods. Does not consider mobilization and demobilization times nor duplicate and quality control tests and data analysis time.

Method	Vertical resolution (m)	Time per unit	Total time for 500 m of Kh measurements at 15-cm interval (hour)	Comments
CPT/SMR	0.05 to 0.16	1.5 h per 10-m sounding	483	Well installations (48 h), multilevel slug tests at 8 locations (280 h), data analysis for the definition of hydro-geophysical relationships (80 h).
Flowmeter tests	0.15	7 h per 10-m well	700	Time to reach steady state pumping rate and data analysis (50 h), well installation and development (300 h).
Multilevel slug tests	0.15	33 h per 10-m well	3600	Data analysis (1650 h), well installation and development (300 h).
Permeameter tests	0.15	99 h per 10-m well	5117	Includes samples preparation and data analysis. Sediment sampling (167 h).
Complementary work				
Well installation and development	n.a.	6 h per 10-m well	300	Includes 1 h for well installation and 5 h for well development.
Sediment sampling	n.a.	0.5 h per 1.5-m core	167	Assuming isotropic conditions at the scale of the sample.

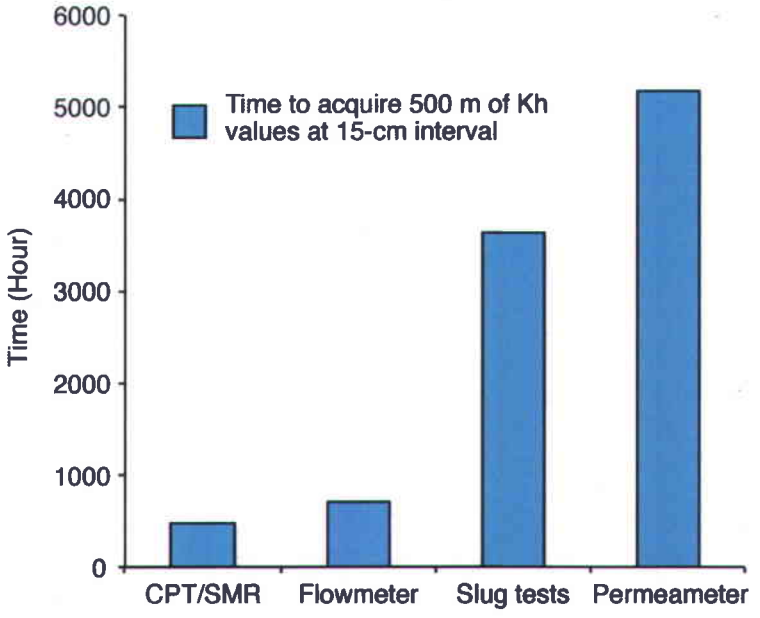


Figure 14 Total time to acquire 500 m of horizontal hydraulic conductivity (K_h) values at 15-cm interval for three conventional hydraulic testing methods (flowmeter tests, multilevel slug tests and laboratory permeameter test) actually measuring properties and estimation using the hydro-geophysical approach based on direct-push soundings. Details about estimations are provided in Table 4.

Though multilevel slug tests take about 25% less time than permeameter tests for the same testing interval length (Table 4 and Figure 14) and continuous profiles of K_h can be obtained more easily within long-screen observation wells, this method is nevertheless quite time consuming. Much of the time for multilevel slug tests is spent on hydraulic testing itself, in particular for less permeable intervals, as well as on extensive well development, which is essential to get reliable estimates of K_h (see supplementary materials in Paradis et al. 2011). We note that the time estimated for multilevel slug tests is somewhat overestimated because it considers that wells are used only for hydraulic testing purposes. In a broader data acquisition perspective for contaminant characterization those wells can serve multiple purposes (e.g., geochemical sampling) and the time for well installation and development would be distributed over all the operations. For accurate estimates of K_h with multilevel slug tests over small intervals, tested wells should be preferably installed without sand-pack to avoid the skin effects of the sand-pack. If that is not possible, the characteristics of the sand-pack should be at least taken into account during the test data analysis.

As illustrated in Figure 14, flowmeter tests are a fairly strong improvement in data acquisition time for K_h estimates. After reaching steady or quasi-steady state pumping rate in a well, flow measurements with a

flowmeter generally require only a few minutes per interval and thus a complete profile of K_h can be readily obtained (Table 4). The average K_h value to transpose flow measurements into specific K_h values can also be easily obtained from the pumping data itself or from a slug test performed over the entire length of the screen. However, the limit of sensitivity in flow changes of a flowmeter may hamper K_h estimations for low-K_h intervals and only threshold values can be obtained. Also, the large drawdown induced by the pumping in very low-K_h wells may preclude the testing of intervals at the top of the aquifer. As for multilevel slug tests, flowmeter tests in unconsolidated aquifers also require proper well installation without sand-pack and extensive well development to avoid hydraulic short-circuits.

Finally, the direct-push based approach relying on site-specific relationships appears as the more time efficient one (Figure 14). It should be noted that the time estimate for this approach includes the time required for the installation and development of 8 direct-push wells and for the testing of 280 intervals with multilevel slug tests, as done in this study. Unlike well-based methods, once the hydro-geophysical relationships are defined with enough wells no more wells are required and only direct-push soundings are carried out, which could considerably reduce the overall characterization time because the time to carry out a CPT/SMR sounding without well installation and development is relatively short (Table 4). Flowmeter tests instead of multilevel slug tests could also reduce the data acquisition time related to hydraulic testing. The tested intervals for K_h to build the training dataset should however be carefully selected to allow the establishment of hydro-geophysical relations that are representative of the study area. Also, data integration for the definition of site-specific relations requires a large amount of time and the time estimate for data analysis provided in Table 4 obviously assumes a certain degree of familiarity with the statistical tools.

For instance, if GPR data had been used instead of CPT/SMR data for the establishment of hydro-geophysical relations, it would have been possible to considerably increase aquifer characterization efficiency, as in this study it took approximately half the time to carry out the 21 km of GPR surveys in comparison with the 53 direct-push soundings (Table 1). Obviously the vertical resolution of the GPR surveys is larger than the vertical resolution of the CPT/SMR soundings, but for lateral aquifer heterogeneities that cannot be captured by direct-push soundings due to the spacing between each sounding by respect to the spatial scale of the heterogeneities, GPR surveys can then be valuable. For the study area, the strong attenuation of the GPR signal over much of the silty zone often hindered permeable sand layers under thick silt layers, which limited the use of GPR data for hydraulic properties estimation. Another limitation of the GPR method that may not be only specific to the study area is that GPR surveys only provide information about dielectric constant, which may weaken the predictive capacity of the GPR

method due to the limited number of geophysical parameters available to correlate with hydraulic data. This indicates that for surface-based geophysical methods in general, several geophysical survey methods along the same section can be necessary to get reliable conversion of geophysical data into hydraulic properties. Future work on hydro-geophysical data integration will focus on the use of GPR information either as an additional variable to CPT/SMR data at the wells or to guide the hydraulic information interpolation between wells.

Meanwhile, vertical interference slug tests carried out along a single well (Paradis and Lefebvre 2013) could also increase the efficiency of the overall characterization process. Indeed, vertical interference slug tests required about the same equipment and testing time as conventional multilevel slug tests and provide at the same time K_h and K_v values. Thus, the adoption of vertical interference slug tests instead of multilevel slug tests could increase the number of K_v measurements required to establish a better training dataset and replace permeameter tests on soil samples. However the time for data analysis of vertical interference slug tests may be longer than for multilevel slug tests, especially for heterogeneous profiles that may require numerical inversion to retrieve hydraulic properties.

Finally, hydraulic tomography as proposed by Paradis (2014) and Paradis et al. (2014a) could also be used as a complement to vertical interference slug tests to estimate both K_h and K_v values, but the need to have at least two wells closely spaced is surely more demanding operationally than a single-well test. Tomography may then be more relevant for the definition of medium-scale three-dimensional heterogeneity in K_h, K_v and S_s over a relatively small study site. For regional hydrogeological investigations, if multiple wells closely spaced are available, hydraulic tomography may nevertheless be an interesting replacement to conventional pumping tests. A low-density tomographic coverage can be obtained in an equivalent amount of time as an intermediate duration pumping test, but provides more information about hydraulic properties and their respective structures. It is noted that the knowledge of hydraulic property structures is essential even for regional hydrogeological investigations that generally require average values over large volume of aquifer, because of the bias induced by the averaging procedure for highly heterogeneous conditions (e.g., Wu et al. 2013). This knowledge needs to be then taken into account in the up-scaling process to obtain representative hydraulic property estimations over large volumes of aquifer.

4.2 Complementary work

In relation with this work, other complementary work was carried out at the St-Lambert study area. To further increase the confidence in the spatial distribution of hydraulic properties obtained over the study area, which are illustrated in Figures 13a-d, complementary work have been carried out by Brunet et al. (2012) to constrain the structure and hydraulic properties of the heterogeneities with vertical profiles of hydraulic heads measured in wells. Indeed, it has been shown that variations in hydraulic heads measured at different locations within an aquifer can indicate the degree of continuity of aquifer heterogeneities between wells. In their approach, a multi-points simulation scheme (Strebelle 2002) was combined with an inversion algorithm (e.g. pilot point, gradual deformation) method to constrain the K_h spatial distribution with measured a 3D hydraulic head distribution.

Moreover, geochemical and groundwater age profiles obtained along wells were measured by Tremblay et al. (2012) to be used to constrain the simulated velocity field in the aquifer by a numerical model. For instance, K_h, K_v and n fields obtained in Figure 13a-d can be used within a numerical groundwater flow and age transport model to generate groundwater residence times that can be compared to measured age profiles. Similarly to multilevel heads, both the heterogeneity structures and hydraulic properties can be tuned to match monitoring data. We note that this approach has been already successfully applied elsewhere (e.g., Murphy et al. 2008).

As an alternative to the approach presented in this paper, Ruggeri et al. (2013) proposed the integration of surface-based electrical resistivity tomography (ERT), direct-push resistivity probe and sparse K_h measurements to estimate high-resolution K_h distribution along ERT sections. In this approach, ERT surveys are indeed used as background to extrapolate away from wells K_h data, instead of relying on interpolation algorithms to distribute hydraulic information as done in this paper. To date this approach has been successfully applied for ERT surveys in the sandy zone (see Figure 2c).

Finally, to overcome the difficulties to infer the ground velocity models in processing GPR survey data, Gloaguen et al. (2012) proposed the integration of direct-push relative permittivity and electrical resistivity logs with GPR surveys data to convert the radargram time scale into depth. This approach was developed to better estimate GPR reflector depths and eventually make better use of GPR surveys for hydraulic properties estimations.

5 Summary and conclusions

This paper presented a workflow developed and applied for the characterization of hydraulic property heterogeneity at an intermediate scale (a few km^2), which is a relevant scale for the numerical modeling of groundwater flow and mass transport in relation with contaminant source zones and receptors. This workflow was applied to a 12 km^2 study area for the study of the migration and natural attenuation of a leachate plume emitted by a decommissioned sanitary landfill in a shallow littoral unconsolidated aquifer (Tremblay et al. 2013). The characterization approach is based on the use of high-resolution indirect hydrogeophysical data and targeted acquisition of direct hydraulic data. Extensive GPR surveys were used to provide the context for the selection of optimal locations for CPT/SMR soundings, and the installation of colocated fully-screened direct push observation wells for a subset of the locations to allow hydraulic tests. The approach then relies on the statistical conversion of CPT/SMR data into hydraulic information through the development of site-specific hydro-geophysical relationships using a learning machine approach. The hydrogeological parameters included horizontal hydraulic conductivity (K_h) data measured in observation wells using 15-cm vertical resolution packer slug tests, and vertical hydraulic conductivity (K_v) and porosity (n) data obtained from testing sediment samples in the laboratory. CPT/SMR parameters included resistance to penetration (tip stress, T), mechanical friction (sleeve stress, S), dielectric constant of bulk sediments (D) and DC electrical resistivity (R). All colocated hydraulic and direct-push data were upscaled to a common vertical resolution of 15 cm for the purposes of establishing relations between CPT/SMR measurements and K_h. The acquisition of a training dataset representative of the range of sediment types found in the study area to train the learning machine was based on a top down approach where information on the regional geology and GPR surveys were used to identify CPT/SMR sounding locations, select colocated observation well locations and sample sediments to cover the whole range of hydraulic and direct-push responses. Four hydrofacies (HF) units representing sediments with distinct hydraulic properties and hydro-geophysical relations allowing the prediction of K_h from CPT/SMR data were defined for each HF unit by the combination of fuzzy clustering and relevance vector machines for classification and regression, respectively. Sparse laboratory K_v and n data were integrated a posteriori with colocated K_h measurements and site-specific hydraulic relations defined for each HF unit. The conversion of available spatially distributed CPT/SMR data over the study area provided data on HF, K_h, K_v and n that can be used to define hydraulic property heterogeneity in the study area. An example of hydraulic property interpolation is provided for a GPR section in which multiple CPT/SMR soundings are available as well as direct hydraulic data.

Thus, this work recognizes the importance of site-specific characterization data to develop a sound understanding of processes associated with flow and transport in aquifers. Although a conceptual hydrogeological model can be developed for a specific site from generic data through analog systems or measurements taken elsewhere on similar geological materials, each site is unique and thus site-specific characterization is likely to provide representative characteristics of the site. Although the probing methods used and the relations defined in this paper are site-specific and may be not directly useable in other contexts, the general framework for data acquisition and data integration is general in its basic principles and could be readily generalized to other hydrogeological contexts with different hydraulic and geophysical characterization tools. In this regards, the key findings of this study are summarized below.

The coarse to fine data acquisition approach allows rational, systematic and progressive data collection. This procedure, as developed and applied in the study area, was designed to minimize the number of investigated sites while maximizing the quality, extent and significance of the collected direct and indirect data. For instance, the definition of a general hydrogeological conceptual model through regional geology, GPR surveys and CPT/SMR soundings allowed the targeting of specific intervals for multilevel slug tests, which reduced time consuming hydraulic testing, and ensured the coverage of the entire range of hydro-geophysical and hydraulic responses observed over the study area. This general framework to data acquisition provided a representative and complete training dataset for the learning machine approach used to assess hydraulic properties from indirect CPT/SMR data. Obviously, according to the scale and the hydrogeological system studied, any comprehensive combination of characterization tools can be used as proposed in this study.

CPT/SMR soundings appear as a promising tool to characterize heterogeneity in hydraulic property for shallow unconsolidated aquifers with sediment grain-size finer than gravel. The depth of investigation of direct-push tools generally depends on the applied load and mechanical properties of sediments. In sandy unconsolidated aquifers in the absence of coarse gravel or pebbles, depths of up to 40 m can be reached. Moreover, the use of a multi-parameter probe, such as the CPT/SMR, reduced non-uniqueness between geophysical and hydraulic properties by providing a series of geophysical parameters that provide complementary information on hydraulic data. However, for hydrogeological investigations where direct-push tools cannot be used, surface-based methods can replace CPT/SMR soundings. To obtain the same benefits of a multi-parameter probe with surface-based geophysical methods, several methods would have to be used in conjunction. For instance, mechanical properties of sediments can be obtained from compression and shear wave components of seismic reflection surveys similarly to tip (T) and sleeve

stresses (S) data of a CPT probe. The dielectric constant (D) and electrical resistivity (R) estimated with a SMR probe could also have their equivalent with GPR and ERT surveys. However, the choice of geophysical methods will always depend on the physical properties of the geological material investigated (e.g., Telford et al. 1990; Rubin and Hubbard 2005) and their respective sensitivity to hydraulic properties.

Reliable statistical relationships between hydraulic and hydro-geophysical parameters are critical for successfully employing geophysical data for hydraulic properties characterization. In this study, among the multi-parameter CPT/SMR probe used to reduce non-uniqueness between geophysical and hydraulic properties, the use of clustering to separate distinct hydro-geophysical responses and of RVMs to handle complex statistical relations contributed to the development of relationships with good generalization capabilities. For the littoral sediments aquifer found in the study area, gradual transitions in sediments posed a serious challenge to aquifer characterization. For instance, small changes in grain-size distribution that only show small changes in geophysical properties may induce large variations in hydraulic conductivity due to the change in grain-size that dominates the pore structure (e.g., Koltermann and Gorelick 1995). For environments with sharp contrasts in lithologies, with corresponding contrasts in both geophysical and hydraulic parameters, HF recognition is expected to be easier and hydraulic property estimation errors associated with HF recognition lower. Also, the need to use several geophysical parameters associated with hydro-geophysical parameters non-uniqueness could be alleviated. Nevertheless, CPT/SMR data through site-specific hydro-geophysical relations appear to contain enough information to infer the hydraulic property heterogeneity of the granular aquifer in the study area. Indeed, the interpolation of converted direct-push data along vertical profiles into hydraulic information shows that spatial distributions for HF, K_h, K_v and n reflect the sedimentary architecture of the littoral aquifer, which involves HF unit inter-digitation and spatial trends in hydraulic properties between and within HF. This suggests that the learning machine approach proposed here is robust and its use could be generalized to other characterization problems.

The amount of hydraulic data needed to build statistical relationships, and for hydrogeological characterization in general, is also important and efforts have to be dedicated to the development of more efficient hydraulic testing methods. In particular, methods that can provide reliable information about K_v and n, which are generally difficult to acquire, would be essential to provide a more complete picture of the parameters that control groundwater flow and contaminant transport. The number of data needed to build representative relationships or to cover the study area is however difficult to assess, and this concern would be guided through experience and knowledge of the local environment. The use of fully-screened observation wells installed without sand pack using a direct push method was a key component of our

characterization approach. These wells allowed high vertical resolution (15 cm) packer slug tests to be carried out and provide continuous profiles of K_h over the full range of materials found in the study area. Such wells also promote the use of less time consuming flowmeter surveys that can also provide high-resolution continuous vertical profiles of K_h (Paradis et al. 2011). Such wells are also needed to implement vertical interference tests (Paradis and Lefebvre 2013) and tomographic slug tests (Paradis 2014; Paradis et al. 2014a) that can provide more complete measurements of hydraulic properties (K_h, K_v and S_s). While hydraulic testing in open holes in consolidated formations is usually unrestricted by hydraulic short-circuits, direct-push wells installed with screens in direct contact with sediments offer a serious alternative to conventional wells, which are prone to hydraulic short-circuits due to the use of sand pack, for high-resolution hydraulic testing in unconsolidated aquifers. Note that direct-push hydraulic tests have been also proposed to avoid the limitations of conventional wells such as direct-push injection logging (Dietrich et al. 2008; Liu et al. 2009; Lessof et al. 2010), the hydraulic profiling tool (McCall et al. 2009; Köber et al. 2009), direct-push slug testing (Butler et al. 2002), and the direct-push permeameter (Butler et al. 2007).

A hydro-geophysical approach to characterize the heterogeneous hydraulic properties of aquifers is significantly more time efficient compared to conventional hydraulic testing approaches. The approach proposed takes advantage of the spatial coverage offered by CPT/SMR soundings, while minimizing the number of time consuming hydraulic tests. That is, for a similar characterization effort, much more hydraulic information could be obtained from a hydro-geophysical approach than from a conventional hydrogeological approach based exclusively on core descriptions and hydraulic tests. In keeping with this study, where we illustrated the benefits and the feasibility of implementing such a hydro-geophysical approach for a shallow unconsolidated aquifer at an intermediate scale of a few km^2, future work will be needed to adapt and further develop this type of approach for different hydrogeological contexts and scales. Considering the environmental and societal costs that may be due to the insufficient characterization of aquifer systems, those efforts would certainly be justified.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Marc Hinton, Geoffrey Bohling and an anonymous reviewer for their constructive review of this manuscript. The Geological Survey of Canada under the Groundwater Geoscience Program, the Régie intermunicipale de gestion des déchets des Chutes-de-la-Chaudière and

NSERC Discovery Grants held by E.G. and R.L supported this study. This is an Earth Science Sector contribution number 20130426.

REFERENCES

American Society for Testing and Materials-ASTM (2000) Standard test method for performing electronic friction cone and piezocone penetration testing of soils. Annual book of ASTM standards. D5578-95 (Reapproved 2000)

American Society for Testing and Materials-ASTM (2004) D6724: Standard guide for installation of direct push ground water monitoring wells. ASTM International, 9 pp.

Anderson MP (1989) Hydrogeologic facies models to delineate large-scale spatial trends in glacial and glaciofluvial sediments. Geol Soc Am Bull 101:501-511

Anderson MP (1997) Characterization of geological heterogeneity. In: Dagan G, Neuman SP (ed) Stochastic Subsurface Hydrology, Cambridge University Press, Cambridge

Anderson MP, Aiken JS, Webb EK, Mickelson DM (1999) Sedimentology and hydrogeology of two braided stream deposits. Sediment Geol 129(3):187-199

Bear J (1979) Hydraulics of groundwater. McGraw-Hill, New York

Bezdek JC (1981) Pattern recognition with fuzzy objective function algoritms. Plenum Press, New York

Blouin M, Martel R, Gloaguen E (2013) Accounting for Aquifer Heterogeneity from Geological Data to Management Tools. Ground Water 51:421-431

Bolduc A (2003) Géologie des formations superficielles, Charny, Québec. Commission géologique du Canada, Dossier public 1976, échelle 1/50000

Bouwer H, Rice RC (1976) A slug test method for determining hydraulic conductivity of unconfined aquifers with completely or partially penetrating wells. Water Resour Res 12(3):423-428

Bradford JH, Babcock E (2013) The need to adapt the exploration model from the oil patch to contaminated-site characterization: A case from Hill AFB, Utah, USA. The Leading Edge, 32(7):750-756

Brunet P, Paradis D, Lefebvre R, Gloaguen E (2012) Using CPT and multi-level piezometric data to assess aquifer heterogeneity at the sub-watershed scale. 39th IAH Congress, Niagara Falls, Canada, September 16-21, Paper 844

Butler JJ, Healey JM, McCall GW, Garnett EJ, Loheide II SP (2002) Hydraulic tests with direct-push equipment. Ground Water 40(1):25-36

Butler JJ, Dietrich P, Wittig V, Christy T (2007) Characterizing hydraulic conductivity with the direct-push permeameter. Ground Water 45(4):409-419

Butler JJJr (1998) The Design, Performance, and Analysis of Slug Tests. Lewis Publishers, Boca Raton, Florida

Butler JJJr (2005) Hydrogeological methods for estimation of hydraulic conductivity. In: Rubin Y, Hubbard S (ed) Hydrogeophysics, Springer, New York, pp 23-58

Camps-Valls G, Gómez-Chova L, Muñoz-Marí J, Vila-Francés J, Amorós-López J, Calpe-Maravilla J (2006) Retrieval of oceanic chlorophyll concentration with relevance vector machines. Remote Sens Environ 105(1):23-33

Caputo MC, De Carlo L, Masciopinto C, Nimmo JR (2010) Measurement of field-saturated hydraulic conductivity on fractured rock outcrops near Altamura (Southern Italy) with an adjustable large ring infiltrometer. Environ Earth Sci, 60(3):583-590, 10.1007/s12665-009-0198-y

Castany G (1963) Traité pratique des eaux souterraines. Dunod, Paris

Dann R, Turner M, Close M, Knackstedt M (2011) Multi-scale characterisation of coastal sand aquifer media for contaminant transport using X-ray computed tomography. Environ Earth Sci, 63(5):1125-1137 DOI, 10.1007/s12665-010-0788-8

Davis JC (1973) Statistics and data analysis in geology. Wiley, New York

Davis JL, Annan AP (1989) Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy. Geophys Prospect 37(5):531–551

de Marsily G, Delay F, Gonçalvès J, Renard P, Teles V, Violette S (2005) Dealing with spatial heterogeneity. Hydrogeol J 13:161-183

Dietrich P, Butler JJ, Faiss K (2008) A rapid method for hydraulic profiling in unconsolidated formations. Ground Water 46(2):323–328

Doro KO, Leven C, Cirpka OA (2013) Delineating subsurface heterogeneity at a loop of River Steinlach using geophysical and hydrogeological methods. Environ Earth Sci, 69(2):335-348, DOI: 10.1007/s12665-013-2316-0

Dunn JC (1973) A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. J Cybernetics 3:32-57

Eaton TT (2006) On the importance of geological heterogeneity for flow simulation. Sediment Geol 184:187-201

Fellenius BH, Eslami A (2000) Soil profile interpreted from CPTu data. "Year 2000 Geotechnics" Geotechnical Engineering Conference, Asian Institute of Technology, Bangkok, Thailand, November 27-30, 2000

Filion L (1987) Holocene development of parabolic dunes in the central St. Lawrence Lowland, Québec. Quaternary Res 28(2):196-209

Freeze RA, Cherry JA (1979) Groundwater. Prentice-Hall, Englewood Cliffs, New-Jersey

Frei S, Fleckenstein JH, Kollet SJ, Maxwel RM (2009) Patterns and dynamics of river–aquifer exchange with variably-saturated flow using a fully-coupled model. Journal of Hydrology 375, 383–393

Gelhar LW (1993) Stochastic subsurface hydrology. Prentice-Hall, Englewood Cliffs, New-Jersey

Ghosh S, Mujumdar P (2008) Statistical downscaling of GCM simulations to streamflow using relevance vector machine. Adv Water Resour 31:132-146

Gloaguen E, Lefebvre R, Ballard JM, Paradis D, Tremblay L, Michaud Y. (2012) Inference of the two dimensional GPR velocity field using colocated cokriging of direct push permittivity and conductivity logs and GPR profiles. J Appl Geophys 78:94-101

Grabe A, Rodiger T, Rink K, Fischer T, Sun F, Wang WQ, Siebert C, Kolditz O (2013) Numerical analysis of the groundwater regime in the western Dead Sea escarpment, Israel plus West Bank. Environ Earth Sci, 69(2):571-585, DOI: 10.1007/s12665-012-1795-8

Gungor-Demirci G, Aksoy A (2011) Variation in time-to-compliance for pump-and-treat remediation of mass transfer-limited aquifers with hydraulic conductivity heterogeneity. Environ Earth Sci, 63(6):1277-1288, DOI: 10.1007/s12665-010-0800-3

Gustafson DE, Kessel WC (1979) Fuzzy clustering with a fuzzy covariance matrix. Proceeding of IEEE Conference on Decision and Control including the 17th Symposium on Adaptive Processes, San Diego, CA, 761-766

Heinz J, Kleineidam S, Teutsch G, Aigner T (2003) Heterogeneity patterns of Quaternary glaciofluvial gravel bodies (SW-Germany): application to hydrogeology. Sediment Geol 158:1-23

Henebry BJ, Robbins GA (2000) Reducing the influence of skin effects on hydraulic conductivity determinations in multilevel samplers installed with direct push methods. Ground Water 38(6):882-886

Hess AE (1986) Identifying hydraulically conductive fractures with a slow-velocity borehole flowmeter. Can Geotech J 23:69-78

Isaaks EH, Srivastava RM (1989) An introduction to applied geostatistics. Oxford University Press, New York

Khalili A, Almasri MN, McKee M, Kaluarachchi JJ (2005) Applicability of statistical learning algorithms in groundwater quality modeling. Water Resour Res 41:W05010, doi:10.1029/2004WR003608

Klingbeil R, Kleineidam S, Asprion U, Aigner T, Teutsch G (1999) Relating lithofacies to hydrofacies: outcrop-based hydrogeological characterisation of Quaternary gravel deposits. Sediment Geol 129(3-4):299-310

Klute A, Dirksen C (1986) Hydraulic conductivity and diffusivity: laboratory methods. Methods of soil analysis. Part 1. In: Klute A (ed) Physical and Mineralogical Methods, Agronomy Monograph 9. American Society of Agronomy, pp 687-734

Köber R, Hornbruch G, Leven C, Tischer L, Grossmann J, Dietrich P, Weiss H, Dahmke A (2009) Evaluation of combined direct-push methods used for aquifer model generation. Ground Water 47(4):536-546

Koltermann CE, Gorelick SM (1995) Fractional packing model for hydraulic conductivity derived from sediment mixtures. Water Resour Res 31(12):3283-3297, doi:10.1029/95WR02020

Koltermann CE, Gorelick SM (1996) Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches. Water Resour Res 32(9):2617-2658.

Kostic B, Becht A, Aigner T (2005) 3-D sedimentary architecture of a Quaternary gravel delta (SW-Germany): Implications for hydrostratigraphy. Sediment Geol 181(3-4):147-171

Krumbein WC, Sloss LL (1963) Stratigraphy and Sedimentation. 2nd edn. W.H. Freeman and Company, San Francisco

Kruseman GP, de Ridder NA (1990) Analysis and evaluation of pumping test data. ILRI Publishing, Netherlands

Lafuerza S, Canals M, Casamor JL, Devincenzi JM (2005) Characterization of deltaic sediment bodies based on in situ CPT/CPTU profiles: A case study on the Llobregat delta plain, Barcelona, Spain. Mar Geol 222-223:497-510

Lamarche L, Tremblay L (2012) Géologie des formations superficielles pour le site de St-Lambert-de-Lauzon, Québec, In: Tremblay L (2013). Caractérisation intégrée d'un aquifère granulaire pour l'évaluation des processus géochimiques influençant l'atténuation naturelle d'un panache de lixiviat. Ph.D. Thesis, Institut national de la recherche scientifique (INRS) p 50

Lessoff SC, Schneidewind U, Leven C, Blum P, Dietrich P, Dagan G (2010) Spatial characterization of the hydraulic conductivity using direct-push injection logging. Water Resour Res 46(12):1-9

Levy BS, Pannell L (1991) Evaluation of a pressure system for estimating in-situ hydraulic conductivity. Ground Water Management No. 5. In Proceedings 5th National Outdoor Action Conference, 31–45. Las Vegas, Nevada

Liu G, Butler JJ, Bohling GC, Reboulet E, Knobbe S, Hyndman DW (2009) A new method for high-resolution characterization of hydraulic conductivity. Water Resour Res 45(8):1-6

Lu CP, Shu LC, Chen XH, Cheng C (2011) Parameter estimation for a karst aquifer with unknown thickness using the genetic algorithm method. Environ Earth Sci, 63(4):797-807, DOI: 10.1007/s12665-010-0751-8

Lunne T, Robertson PK, Powell JJM (1997) Cone penetration testing in geotechnical practice. Spon Press, New York

Mastrocicco M, Colombani N, Palpacelli S, Castaldelli G (2011) Large tank experiment on nitrate fate and transport: the role of permeability distribution. Environ Earth Sci, 63(5):903-914, DOI: 10.1007/s12665-010-0759-0

McCall W, Christy TM, Christopherson T, Issacs H (2009) Application of Direct Push Methods to Investigate Uranium Distribution in an Alluvial Aquifer. Ground Water Monit R 29(4):65-76

Murphy S, Ouellon T, Ballard JM, Lefebvre R, Clark ID (2011) Tritium–helium groundwater age used to constrain a groundwater flow model of a valley-fill aquifer contaminated with trichloroethylene (Quebec, Canada). Hydrogeology Journal, 19(1):195-207

Ouellon T, Lefebvre R, Marcotte D, Boutin A, Blais V, Parent M (2008) Hydraulic conductivity heterogeneity of a local deltaic aquifer system from the kriged 3D distribution of hydrofacies from borehole logs, Valcatier, Canada. J Hydrol 351(1-2): 71-86

Paradis D, (2014) Contributions au développement de la caractérisation de l'hétérogénéité hydraulique des aquifères. Thèse de doctorat, Université du Québec, Centre Eau Terre Environnement. 290 p.

Paradis D, Lefebvre R (2013) Single-well interference slug tests to assess the vertical hydraulic conductivity of unconsolidated aquifers. J Hydrol 478(25):102-118

Paradis D, Gloaguen E, Lefebvre R, Rivera A (2011) Tying geophysics to hydrogeology: a learning machine 1138 approach to characterize heterogeneous granular aquifers. Geohydro2011, Joint IAH-CNC, CANQUA 1139 and AHQ conference, Quebec City, Canada, August 28-31

Paradis D, Lefebvre R, Morin RH, Gloaguen E (2011) Permeability profiles in granular aquifers using flowmeters in direct-push wells. Ground Water 49(4):534-547

Paradis D, Gloaguen E, Lefebvre R, Giroux B (2014a) Sensitivity and resolution analysis of the information content of tomographic slug tests. Accepted for publication in Water Resour Res, November, 2013.

Paradis D, Lefebvre R, Gloaguen E, Rivera A (2014b) Hydrofacies and hydraulic conductivity from direct-push data using relevance vector machines: motivation, algorithm and application. Submitted to Water Resour Res, Marsh, 2014.

Poeter EP, McKenna SA (1995) Reducing uncertainty associated with ground-water flow and transport predictions. Ground Water 33:899-904

Reading HG (1996) Sedimentary environments: Processes, facies and stratigraphy. 3rd edn. Blackwell Science, Oxford

Reineck HE, Singh IB (1980) Depositional sedimentary environments. 2nd edn. Springer-Verlag, New York

Robertson PK (1990) Soil classification using the cone penetration test. Can Geotech J 27(1):151-158

Ross HC, McElwee CD (2007) Multi-level slug tests to measure 3-D hydraulic conductivity distributions. Water Resour Res 16(1):67-79

Rubin Y, Hubbard S (2005) Hydrogeophysics. Springer, Netherlands

Ruggeri P, Irving J, Gloaguen E, Holliger K (2013) Regional-scale integration of multiresolution hydrological and geophysical data using a two-step Bayesian sequential simulation approach. Geophys J Int 194(1): 289-303

Samui P (2007) Seismic liquefaction potential assessment by using relevance vector machine. Earthq Eng Eng Vib 6(4):331-336

Schmelzbach C, Tronicke J, Dietrich P (2011) Three-dimensional hydrostratigraphic models from ground-penetrating radar and direct-push data. J Hydrol 398(3-4):235-245

Schulmeister M, Butler JJ, Healey J, Zheng L, Wysocki D, McCall G (2003) Direct-push electrical conductivity logging for high-resolution hydrostratigraphic characterization. Ground Water Monit R 23:52-62

Sellwood SM, Healey JM, Birks S, Butler JJJr (2005) Direct-push hydrostratigraphic profiling: Coupling electrical logging and slug tests. Ground Water 43(1):19-29

Shinn JD, Timian DA, Morey RM, Mitchell G, Antle CL, Hull R (1998) Development of a CPT deployed probe for in situ measurement of volumetric soil moisture content and electrical resistivity. Field Anal Chem Tech 2(2):103-110

Smirnoff A, Boisvert E, Paradis SJ (2008) Support vector machine for 3D modelling from sparse geological information of various origins. Comput Geosci 34(2):127-143

Stienstra P, van Deen JK (1994) Field data collection techniques - Unconventional sounding and sampling methods. In: Rengers N (ed) Engineering Geology of Quaternary Sediments. Balkema, pp 41-55

Strebelle S (2002) Conditional simulation of complex geological structures using multiple-point statistics. Mathematical Geology, 34(1):1-21

Telford WM, Geldart LP, Sheriff RE (1990) Applied Geophysics (2nd edition). Cambridge University Press, pp. 770

Tipping ME (2001) Sparse Bayesian learning and the relevance vector machine. J Mach Learn Res 1:211-244

Tremblay L, Lefebvre R, Cloutier V, Molson JW (2012) Value of geochemical data for understanding groundwater flow and natural attenuation of a leachate plume at a former landfill. 39th IAH Congress, Niagara Falls, Canada, September 16-21, Paper 317

Tremblay L, Lefebvre R, Paradis D, Gloaguen E (2013) Conceptual model of leachate migration in a granular aquifer from the integration of detailed characterization data. Hydrogeol J, doi 10.1007/s10040-013-1065-1

Van Overmeeren RA (1998) Radar facies of unconsolidated sediments in The Netherlands: A radar stratigraphy interpretation method for hydrogeology. J Appl Geophys 40(1-3)

Vitton SJ, Sadler LY (1997) Particle-Size Analysis of Soils Using Laser Light Scattering and X-Ray Absorption Technology. Geotech Test J 20(1):1-11

Wentworth CK (1922) A scale of grade and class term for clastic sediments. J Geol 30:507-521

Wu SC, Tan YC, Chen CH, Lin HT (2013) Estimation of effective hydrogeological parameters by considering varying heterogeneity and pumping rates. Environ Earth Sci, 68(1):169-180, DOI: 10.1007/s12665-012-1727-7

Zemansky GM, McElwee CD (2005) High-resolution slug testing. Ground Water 43(2):222-230

Zlotnik V (1994) Interpretation of slug and packer tests in anisotropic aquifers. Ground Water 32(5):761-766