

Université du Québec  
Institut national de la recherche scientifique  
Centre Eau Terre Environnement

**ASSIMILATION DE DONNÉES MULTIPLES AVEC ENSEMBLE  
SMOOTHER DANS LE CADRE DE MODÉLISATIONS  
HYDROGÉOLOGIQUES.**

Par

Thomas Béraud

Thèse présentée pour l'obtention du grade de  
*Docteur en Sciences*, PhD.  
en sciences de la terre

**Jury d'évaluation**

Responsable du programme	Bernard Giroux INRS
Jury externe	Frédéric Nguyen Université de Liège
Jury externe	Philippe Renard Université de Neuchâtel
Jury interne	John Molson Université Laval
Directeur de recherche	Erwan Gloaguen INRS
Co-Directeur de recherche	Maxime Claprood INRS



# Remerciements

Je voudrais tout d'abord remercier mon directeur de thèse Erwan et mon codirecteur Maxime pour m'avoir fait confiance pendant ces années, m'avoir épaulé tout au long de mes recherches et pour les vendredis après-midi été comme hiver.

Un merci également à Envisol pour m'avoir fait découvrir le Québec, sa culture, son climat et m'avoir donné envie d'y retourner pour ces dernières années.

Une mention pour Martin qui m'a convaincu de la bonne humeur de cette région du monde et m'a appris la géophysique à travers les terrains, été comme hiver québécois.

Je tiens également à remercier Guillaume Caumon, Paul Cupillard et Modeste Irakarama pour m'avoir donné le goût de la recherche pendant mon Master en géologie numérique à l'École Nationale Supérieure de Géologie à Nancy.

À tous les étudiants rencontrés ici à l'INRS qui ont rendu les repas du midi sympathique, en particulier Ceci et Mathieu.

Aux camarades de longue date, Arnaud, Florian et Tom malgré la distance et le décalage horaire.

À René Coty.

Un grand merci à Bastien le coloc des études sup' pour nos nombreuses années de jeux et la bonne humeur permanente sur Discord. Tu m'auras aidé à surmonter bien des épreuves et appris que gagner ce n'est pas le plus important (Un jour le top 1 promis).

Ma famille, pour son soutien pendant toutes ces années loin d'eux malgré des explications toujours trop approximatives sur ce que je faisais.

Et enfin à ma blonde, Leenane, qui m'a toujours donné de la motivation quand elle manquait.

Cette thèse est dédiée à la mémoire de ma maman, Geneviève partie trop tôt pour m'avoir vu devenir ingénieur ou docteur, et qui en aurait été très fière.



# Résumé

L'objectif de ce doctorat a été de développer une méthode numérique pour calibrer des modèles d'écoulement hydrogéologique à partir d'observations temporelles. Les méthodes d'ensemble ont été retenues pour cette tâche. Les méthodes d'ensemble reposent sur un ensemble de modèles équiprobables prédisant des paramètres physiques. Les paramètres d'états sont mis à jour en étudiant la covariance de ces paramètres dans les modèles de l'ensemble et les simulations issues de chaque modèle. L'ensemble smoother développé dans le cadre de ce travail repose sur un nouveau schéma d'assimilation. Chaque puits est assimilé individuellement, pour chaque type de données observé. De plus, une analyse locale est définie autour de chacun de ces puits pour mettre à jour les paramètres d'états des modèles numériques. Cette méthode a été testée sur un modèle synthétique représentant un aquifère côtier avec une intrusion saline. Dans ce modèle, 3 puits suivaient les concentrations en sel et les charges hydrauliques. Ces puits ont été assimilés avec la nouvelle méthode d'assimilation d'ensemble pour mettre à jour 68940 valeurs de perméabilités. Le résultat était une réduction moyenne de la RMSE de 42% pour les charges hydrauliques et de 63% pour les concentrations en sel entre les modèles de l'ensemble initial et final. La méthode a ensuite été testée sur un cas réel plus complexe, avec 1094771 éléments et 24 puits d'observation de charges hydrauliques. Dans ce cas, l'amélioration moyenne de la RMSE avec cette nouvelle méthode d'assimilation était d'environ 18%. Sans la nouvelle méthode développée dans ce doctorat, les ensembles de modèles divergeaient à la première itération de l'assimilation, avec une RMSE moyenne doublant ou triplant. Pour conclure, il est possible d'appliquer les méthodes d'ensemble smoother à des problématiques complexes hydrogéologiques à condition d'utiliser un schéma d'analyse locale et des assimilations puits par puits.

**Mots-clés** Ensemble Smoother; Assimilation; Analyse Locale; Localisation; Hydrogéologie; Modèles; Incertitudes



# Abstract

The objective of this PhD was to develop a numerical method to calibrate hydrogeological flow models from transient observations. Ensemble methods were chosen for this task which are based on a set of equiprobable models predicting physical parameters. The parameters are updated by studying the covariance of these parameters in the ensemble models and the simulated results from each model. The ensemble smoother developed in this work is based on a new assimilation scheme. Each well is assimilated individually, for each type of observed data. Moreover, a local analysis is defined around each of these wells to update the state parameters of the numerical models. This method has been tested on a synthetic model representing a coastal aquifer with saline intrusion. In this model, 3 wells followed salt concentrations and hydraulic heads. These wells were assimilated with the new ensemble assimilation method to update 68940 permeability values. The result was an average reduction in the RMSE error of 42% for hydraulic heads and 63% for salt concentrations between the initial and final ensemble. The method was then tested on a more complex real case, with 1094771 elements and 24 observation wells. The wells only tracked the hydraulic heads. In this case, the average improvement of the RMSE with this new assimilation method was about 18%. Without the newly developed method in this PhD, models diverged at the first iteration of the assimilation, with an average RMSE doubling or tripling. In conclusion, it is possible to apply ensemble smoother methods to complex hydrogeological problems with a local analysis scheme and well-by-well assimilations.

**Keywords** Ensemble Smoother; Assimilation; Local Analysis; Localisation; Hydrogeology; Model; Uncertainty



# Table des matières

<b>Remerciements</b>	<b>iii</b>
<b>Résumé</b>	<b>v</b>
<b>Abstract</b>	<b>vii</b>
<b>Table des matières</b>	<b>ix</b>
<b>Liste des figures</b>	<b>xiii</b>
<b>Liste des tableaux</b>	<b>xxi</b>
<b>Introduction</b>	<b>1</b>
<b>Contributions</b>	<b>3</b>
<b>1 Revue bibliographique</b>	<b>5</b>
1.1 Assimilation de données . . . . .	5
1.2 Historique . . . . .	6
1.3 Les différentes familles d'assimilation . . . . .	7
1.4 Localisation, Analyse Locale et Inflation . . . . .	10
1.4.1 Localisation . . . . .	10
1.4.2 Analyse Locale . . . . .	11
1.4.3 Inflation . . . . .	12
<b>2 Théorie</b>	<b>13</b>
2.1 Introduction à la modélisation . . . . .	13
2.2 Méthodes d'ensemble . . . . .	14
<b>3 Méthodologie</b>	<b>25</b>
3.1 Ratio d'information . . . . .	25
3.2 Schéma général . . . . .	26
3.2.1 Modélisation conceptuelle . . . . .	27
3.2.2 Modélisation géostatistique . . . . .	28
3.2.3 Modélisation hydrogéologique . . . . .	28
3.2.4 Méthodes d'assimilation . . . . .	29
3.3 Nouvelle méthode de Localisation et d'Analyse locale . . . . .	29
3.3.1 Localisation . . . . .	29
3.3.2 Schémas de Localisation et d'Analyse locale . . . . .	34

3.3.3	Implémentation de l'ellipsoïde d'Analyse Locale . . . . .	39
3.4	Code développé . . . . .	45
<b>4</b>	<b>Modèle synthétique</b>	<b>47</b>
4.1	Modèle géologique synthétique . . . . .	47
4.1.1	Modèle 3D synthétique d'aquifère côtier . . . . .	51
4.1.2	Assimilation . . . . .	57
4.2	Résultats . . . . .	58
4.2.1	RMSE sur les concentrations en sel . . . . .	59
4.2.2	RMSE sur les charges hydrauliques . . . . .	61
4.2.3	Amélioration moyenne en pourcent des RMSE . . . . .	63
4.2.4	Erreur moyenne des champs de conductivité hydraulique . . . . .	67
4.2.5	Réalisations individuelles dans l'ensemble . . . . .	68
4.3	Discussion sur le modèle synthétique . . . . .	71
<b>5</b>	<b>Modèle Réel : Les lagunes de Mercier</b>	<b>73</b>
5.1	Objectifs du modèle . . . . .	73
5.2	Présentation du site . . . . .	74
5.2.1	Localisation du site . . . . .	74
5.2.2	Contexte géologique . . . . .	75
5.2.3	Carte géologique . . . . .	79
5.2.4	Lignes sismiques . . . . .	81
5.3	Modèle 3D des lagunes de Mercier . . . . .	82
5.3.1	Modèle conceptuel . . . . .	83
5.3.2	Modèle géostatistique . . . . .	85
5.4	Modèle hydrogéologique . . . . .	87
5.4.1	Maillage 2D FEFLOW . . . . .	88
5.4.2	Grille d'écoulement 3D FEFLOW . . . . .	89
5.4.3	Propriétés hydrauliques initiales . . . . .	91
5.4.4	Conditions aux frontières . . . . .	91
5.4.5	Recharge . . . . .	92
5.4.6	Calibration du modèle régional permanent . . . . .	93
5.5	Définition de l'ensemble initial du processus d'assimilation . . . . .	94
5.5.1	Définition de l'ensemble initial pour le processus d'assimilation . . . . .	97
5.6	Essais de pompages . . . . .	104
5.6.1	Essai de pompage numéro 1 . . . . .	104
5.6.2	Essai de pompage numéro 2 . . . . .	106
5.7	Résultats . . . . .	109
5.7.1	Évaluation des performances des schémas de localisation envisagés . . . . .	110
5.7.2	Évaluation des performances des schémas d'analyse locale retenus . . . . .	115
5.8	Comparaison avec la méthode PESTPP-IES existante . . . . .	129
5.9	Discussion . . . . .	135
	<b>Conclusion</b>	<b>139</b>
	<b>Liste des abréviations</b>	<b>141</b>
	<b>Références</b>	<b>141</b>

<b>A</b>	<b>Courbes individuelles de charges hydrauliques pour le schéma 500/20 amplifié</b>	<b>147</b>
<b>B</b>	<b>Simulations numériques du cas synthétique</b>	<b>161</b>



# Liste des figures

1.1	Schéma de la mise à jour dans l'ensemble Kalman filter. En horizontal l'axe du temps avec les mesures temporelles effectuées à différents temps. En vertical l'axe des mises à jour. La flèche bleue représente la simulation de l'ensemble avec l'apport d'information provenant des données (flèches rouges). La flèche verte indique la mise à jour de l'ensemble grâce à l'EnKF. L'ensemble indiqué en bleu est modifié à chaque nouvelle mesure. Modifié d'après Evensen (2009a). . . . .	7
1.2	Schéma de la mise à jour dans l'ensemble Kalman Smoother. En horizontal l'axe du temps avec les mesures temporelles effectuées à différents temps. En vertical l'axe des mises à jour. La flèche bleue représente la simulation de l'ensemble avec l'apport d'information provenant des données (flèches rouges). L'ensemble indiqué en bleu est modifié à chaque nouvelle mesure. La flèche rose représente l'ensemble mis à jour avec une propagation de l'information jusqu'au temps initial de la simulation. Modifié d'après Evensen (2009a). . . . .	8
1.3	Schéma de la mise à jour dans l'ensemble smoother. En horizontal l'axe du temps avec les mesures temporelles effectuées à différents temps. En vertical l'axe des mises à jour. La flèche bleue représente la simulation de l'ensemble avec l'apport d'information provenant des données (flèches rouges). La flèche rose représente l'ensemble mis à jour. La flèche verte indique la mise à jour unique de l'ensemble. Modifié d'après Evensen (2009a). . . . .	8
2.1	n <sup>e</sup> grille de porosité obtenue par simulation géostatistique pour une grille en 2 dimensions contenant 20 cases. . . . .	15
2.2	5 réalisations de porosité obtenue par simulation pour une grille en 2 dimensions contenant 20 cases. . . . .	16
2.3	Porosité moyenne calculée sur les 5 réalisations de porosité obtenue sur la Figure 2.2.	18
2.4	5 grilles d'anomalies obtenues depuis les simulations pour une grille en 2 dimensions contenant 20 cases. . . . .	19
2.5	Matrice de covariance des 5 modèles de porosité. . . . .	20
2.6	Schéma d'un cycle complet d'assimilation de données d'un filtre de Kalman d'ensemble. En bleu le posterior provenant de l'étape d'analyse 2.8, en vert foncé le simulateur équivalent à H, en orange le prior provenant des simulations d'écoulement et en vert clair la vraisemblance des données mesurées. Le modèle réel est assumé étant présent dans la même distribution que l'ensemble. Modifié depuis (Raanes, 2015b). .	23

3.1	Le schéma peut être décomposé en deux cycles. Le premier cycle avec les flèches vertes indique les étapes classiques de modélisation hydrogéologique. Les flèches bleues traduisent le cycle d'assimilation apportée par les méthodes d'ensembles. Les ellipses rouges indiquent les zones mises à jour dans les modèles de paramètres d'états. Ces ellipses sont définis grâce à l'analyse locale. . . . .	27
3.2	Valeurs simulées dans un modèle en orange et valeurs mesurées en vert. Deux puits sont suivis pour deux types d'informations, des concentrations et des charges hydrauliques. . . . .	30
3.3	Anomalies résultantes de la différence entre données simulées et mesurées. . . . .	31
3.4	Anomalies de concentration. . . . .	31
3.5	Anomalies de charges hydrauliques. . . . .	32
3.6	Anomalies de concentrations sur les puits 1 et 2. . . . .	33
3.7	Anomalies de charges hydrauliques sur les puits 1 et 2. . . . .	33
3.8	Schémas d'assimilation de données où toutes les anomalies sont assimilées en une seule fois, donnant une mise à jour unique. . . . .	35
3.9	Schémas d'assimilation de données où les charges hydrauliques et la concentration sont assimilées séparément, donnant deux mises à jour pour chaque cellule. . . . .	36
3.10	Schémas d'assimilation de données où toutes informations sont assimilées successivement, une pour chaque puits et type de donnée. Chaque cellule du modèle est mise à jour par la contribution de ces 4 assimilations. . . . .	37
3.11	Schémas d'assimilation de données où toutes informations sont assimilées successivement, une pour chaque puits et type de donnée. Uniquement les cellules du voisinage des points d'observations sont mises à jour. . . . .	38
3.12	Ellipsoïde de poids défini autour d'un point d'observation donné dans une grille d'éléments non structurés. Le centre de l'ellipsoïde est défini en (50, 50, 10). La portée maximale est mise à 50 m, la portée moyenne à 20 m et la verticale à 10 m. L'azimut de l'axe majeur est défini à un angle de 75° avec le nord et la fonction de poids choisis est une fonction gaussienne. . . . .	44
3.13	Sections horizontale et verticale de l'ellipsoïde de poids défini autour d'un point d'observation donné dans une grille d'éléments non structurés. Le centre de l'ellipsoïde est défini en (50, 50, 10). La portée maximale est mise à 50 m, la portée moyenne à 20 m et la verticale à 10 m. L'azimut de l'axe majeur est défini à un angle de 75° avec le nord et la fonction de poids choisis est une fonction gaussienne. . . . .	44
3.14	Les flèches vertes indiquent l'import de données. Les flèches rouges indiquent l'export de données. Les flèches grises traduisent l'appel d'un code par un autre code. . . . .	45
4.1	Puits virtuels dans le modèle synthétique. . . . .	49
4.2	Histogramme de conductivité hydraulique observé dans les 5 puits virtuels du modèle synthétique 4.1a. . . . .	50
4.3	Variogrammes calculés et les paires de points utilisées comme support. . . . .	50
4.4	Position des 3 types de puits dans le modèle synthétique. Les points orange représentent les puits d'observations pour suivre les charges hydrauliques et les concentrations tout au long des simulations. Les pentagones rouges sont les puits de pompage de la Figure 4.1b. Les triforces vertes sont les puits de références utiliser pour étudier le champ de perméabilité. . . . .	52
4.5	Section en deux dimensions le long de l'axe $x$ des conditions aux limites du modèle FEFLOW. . . . .	53

4.6	Concentration en sel maximale observée dans les 5 puits du modèle de référence. La plage de donnée en bleu correspond à la fenêtre de calibration. Les données avec un fond vert correspondent aux données de validation. La position de ces puits est donnée en vue du dessus sur la Figure 4.4. . . . . .	56
4.7	RMSE des concentrations en sel pour les 8 méthodes à la fin de la 4 <sup>e</sup> assimilation. Les barres bleues correspondent aux données dans la fenêtre temporelle de calibration, les barres vertes correspondent aux données dans la fenêtre temporelle de validation. 0 : Ensemble initial sans aucune assimilation. 1 : Assimilation brute avec toutes les observations assimilées en une fois. 2 : Assimilation par type de données (tous les puits ensemble), sans analyse locale. 3 : Assimilation puits par puits, charge hydraulique seulement, sans analyse locale. 4 : Assimilation puits par puits, concentration seulement, sans analyse locale. 5 : Assimilation puits par puits, type de données par type de données, sans analyse locale. 6 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale 100m et 20m en vertical. 7 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale 50m et 10m en vertical. 8 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale 25m et 5m en vertical. . . . . .	60
4.8	RMSE sur les charges hydrauliques pour les 8 méthodes à la fin de la 4 <sup>e</sup> assimilation. Les barres bleues correspondent aux données dans la fenêtre temporelle de calibration, les barres vertes correspondent aux données dans la fenêtre temporelle de validation. 0 : Ensemble initial sans aucune assimilation. 1 : Assimilation brute avec toutes les observations assimilées en une fois. 2 : Assimilation par type de données (tous les puits ensemble), sans analyse locale. 3 : Assimilation puits par puits, charge hydraulique seulement, sans analyse locale. 4 : Assimilation puits par puits, concentration seulement, sans analyse locale. 5 : Assimilation puits par puits, type de données par type de données, sans analyse locale. 6 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale 100 m et 20 m en vertical. 7 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale 50 m et 10 m en vertical. 8 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale 25 m et 5 m en vertical. . . . . .	62
4.9	Amélioration moyenne en pourcentage de la RMSE par rapport à l'ensemble initial de calibration. L'évolution de la RMSE des concentrations est donnée en bleu foncée, celle des charges hydrauliques est donnée est cyan. La moyenne d'évolution globale de la RMSE est indiquée en rouge. . . . . .	64
4.10	Amélioration moyenne en pourcentage de la RMSE par rapport à l'ensemble initial de validation. L'évolution de la RMSE des concentrations est donnée en bleu foncée, celle des charges hydrauliques est donnée est cyan. La moyenne d'évolution globale de la RMSE est indiquée en rouge. . . . . .	65
4.11	Amélioration moyenne en pourcentage de la RMSE par rapport à l'ensemble initial. L'évolution de la RMSE des concentrations est donnée en bleu foncée, celle des charges hydrauliques est donnée est cyan. La moyenne d'évolution globale de la RMSE est indiquée en rouge. . . . . .	66
4.12	Amélioration en pourcentage de l'erreur moyenne des champs de conductivité hydraulique pour les 8 méthodes à chaque itération d'assimilation par rapport au champ de référence. . . . . .	67

4.13	Réalisations individuelles de charges hydrauliques en bleue pour l'ensemble initial, en orange pour l'ensemble final assimilé. . . . .	69
4.14	Réalisations individuelles de concentrations en bleue pour l'ensemble initial, en orange pour l'ensemble final assimilé. La ligne verticale représente la séparation entre données de calibration et de validation. . . . .	70
5.1	Localisation des anciennes lagunes de Mercier. Fond de carte OpenStreetMap du sud de Montréal distribué sous licence CC BY-SA. Fond de carte de l'Amérique du Nord sous License Creative Commons réalisé par Q-rieux. . . . .	75
5.2	Coupe schématique régionale simplifiée de la géologie du Bas-Saint-Laurent. La coupe suit un axe nord-ouest sud-est comme précisé dans l'encart supérieur gauche de la figure. Modifié de <a href="http://www2.ggl.ulaval.ca/personnel/bourque/intro.pt/planete_terre.html">http://www2.ggl.ulaval.ca/personnel/bourque/intro.pt/planete_terre.html</a> (12/03/2019) et Brun & Brisebois (1994). . . . .	76
5.3	Carte géologique et structurale du substratum rocheux extraite de la base de données du Système d'Information GÉOMinière du Québec (SIGEOM). . . . .	77
5.4	Schéma de l'histoire géologique quaternaire simplifiée sous la ville de Mercier. Encarts dans l'ordre chronologique du haut vers le bas. Substratum rocheux Cambrio-Ordovicien, mise en place des glaciers et naissance du till par abrasion du roc, apparition des dépôts fluvio-glaciaires, fonte des glaciers et déposition d'argiles dans la Mer de Champlain. . . . .	78
5.5	Carte géologique des dépôts meubles du quaternaire autour des anciennes lagunes de Mercier. Modifié d'après Tremblay (2008). . . . .	80
5.6	a) Lignes de sismiques réalisées sous les anciennes lagunes de Mercer et l'incinérateur selon un profil quasiment orienté ouest-est; b) Interpretation géologique du profil sismique contraint avec 12 puits. Les deux profils sont issus des travaux de Tremblay <i>et al.</i> (2010). . . . .	81
5.7	Horizons sismiques interprétés lors de la campagne d'acquisition de 2019 sur le site de Mercier. Exagération verticale 4x. . . . .	82
5.8	Schéma du contexte régional du site des anciennes lagunes de Mercier ((Lefebvre, 2007)). . . . .	84
5.9	Modèle conceptuel du site des anciennes lagunes de Mercier représentant les unités géologiques et l'écoulement de l'eau souterraine ((Lefebvre, 2007)). . . . .	84
5.10	Vue 3D des unités hydrostratigraphiques dans le modèle conceptuel régional. Exagération verticale de la vue régionale de 40x, exagération verticale de la vue locale de 20x. La carte de fond est la géologie de surface, avec le réseau hydrologique. En rouge le roc, en vert les tills, en jaune le coeur de l'esker, en orange les flancs de l'esker et en cyan les argiles marines. . . . .	86
5.11	Vue 3D des unités hydrostratigraphiques au niveau du site de l'UTES, avec une exagération verticale de 5. La carte de fond est la géologie de surface. Les puits de l'UTES PW-113, PW-112, et PW-111 sont présentés en blanc et noir de gauche à droite, respectivement. En rouge le roc, en vert les tills, en jaune le coeur de l'esker, en orange les flancs de l'esker et en cyan les argiles marines. . . . .	87
5.12	Représentation 2D de la grille de modélisation. Le réseau routier est indiqué en magenta. Les lignes bleues sont les sections verticales EW3 et SN3 présentées à la figure suivante. . . . .	89

5.13	Section verticale (EW3 sur la 5.12) d'ouest (gauche) vers l'est (droite) du modèle hydrogéologique 3D, montrant la représentation numérique des unités hydrostratigraphiques avec (haut) et sans (bas) la discrétisation latérale et verticale des éléments. Exagération verticale de 80x. De bas en haut : roc sain(marron), roc fracturé (rouge), till (vert), coeur de l'esker (jaune), esker étendu (orange), argile (cyan) et remblais organique (gris). . . . .	90
5.14	Section verticale (SN3 sur la 5.12) du sud (gauche) ou nord (droite) du modèle hydrogéologique 3D, montrant la représentation numérique des unités hydrostratigraphiques avec (haut) et sans (bas) la discrétisation latérale et verticale des éléments. Exagération verticale de 30x. De bas en haut : roc sain(marron), roc fracturé (rouge), till (vert), coeur de l'esker (jaune), esker étendu (orange), argile (cyan) et remblais organique (gris). . . . .	91
5.15	Distribution spatiale de la recharge efficace assignée à la première couche du modèle hydrogéologique. Le réseau routier est indiqué en magenta. . . . .	93
5.16	Distribution spatiale régionale des niveaux d'eau à calibrer en régime permanent sur le modèle hydrogéologique de Ville Mercier. . . . .	94
5.17	À gauche, les puits de mesures de la conductivité hydraulique dans les trois unités modélisées. À droite, les puits d'observation et de pompage qui seront utilisés lors du processus d'assimilation. Les limites du site de l'UTES sont reprises des Figures 5.15 et 5.16. . . . .	95
5.18	Modèle FEFLOW de perméabilité dans une couche du roc. L'insert géostatistique local présente des variabilités spatiales lissées dans le modèle régional. Le réseau routier est indiqué en magenta. . . . .	97
5.19	Variogrammes modélisés pour l'unité géologique du coeur de l'esker. . . . .	99
5.20	À gauche, puits de mesures utilisés pour les simulations géostatistiques. À droite, une réalisation de perméabilité pour l'unité du coeur de l'esker. . . . .	99
5.21	Variogrammes modélisés pour l'unité géologique des bordures de l'esker. . . . .	100
5.22	À gauche, puits de mesures utilisés pour les simulations géostatistiques. À droite, une réalisation de perméabilité pour l'unité des bordures de l'esker. . . . .	101
5.23	Variogrammes modélisés pour l'unité géologique du roc. . . . .	102
5.24	À gauche, puits de mesures utilisés pour les simulations géostatistiques. À droite, une réalisation de résidus de perméabilité pour l'unité de roc. . . . .	103
5.25	Position des puits d'observations et de pompage pour l'essai de pompage 1. . . . .	104
5.26	Débits prélevés par les puits de pompage PW112 et PW113 lors de l'essai de pompage 1. . . . .	105
5.27	Charges hydrauliques mesurées sur 4 puits de l'essai de pompage 1. . . . .	106
5.28	Position des puits d'observations et de pompage pour l'essai de pompage 2. . . . .	107
5.29	Débits prélevés par le puits de pompage PW8701 lors de l'essai de pompage 2. . . . .	108
5.30	Charges hydrauliques mesurées sur 4 puits de l'essai de pompage 12. . . . .	109
5.31	Évolution des distributions moyennes de perméabilités entre l'ensemble initial et 4 schémas d'assimilations après une itération. L'histogramme bleu sur les 4 sous-figures correspond à la distribution moyenne de l'ensemble initial. . . . .	111
5.32	Évolution d'une distribution individuelle de perméabilités entre l'ensemble initial et 4 schémas d'assimilations après une itération. L'histogramme bleu sur les 4 sous-figures correspond à la distribution de l'ensemble initial. . . . .	112
5.33	Erreur moyenne au carré (RMSE) des charges hydrauliques au droit des 12 puits de l'essai de pompage 1 après une itération d'assimilation. Chaque couleur correspond à un schéma d'assimilation différent. . . . .	113

5.34	Erreur moyenne au carré (RMSE) des charges hydrauliques au droit des 13 puits de l'essai de pompage 2 après une itération d'assimilation. Chaque couleur correspond à un schéma d'assimilation différent. . . . .	114
5.35	. . . . .	116
5.36	. . . . .	116
5.37	RMSE des charges hydrauliques à la fin des 4 schémas d'assimilation pour les puits de l'essai de pompage 1. La RMSE de l'ensemble initial est rappelée pour évaluer la performance. . . . .	118
5.38	RMSE des charges hydrauliques à la fin des 4 schémas d'assimilation pour les puits de l'essai de pompage 2. La RMSE de l'ensemble initial est rappelée pour évaluer la performance. . . . .	119
5.39	RMSE des charges hydrauliques à chaque itération du processus d'assimilation pour les puits de l'essai de pompage 1 et le schéma 250/15. . . . .	120
5.40	RMSE des charges hydrauliques à chaque itération du processus d'assimilation pour les puits de l'essai de pompage 2 et le schéma 250/15. . . . .	120
5.41	RMSE des charges hydrauliques à chaque itération du processus d'assimilation pour les puits de l'essai de pompage 1 et le schéma 250/15 amplifié. . . . .	121
5.42	RMSE des charges hydrauliques à chaque itération du processus d'assimilation pour les puits de l'essai de pompage 2 et le schéma 250/15 amplifié. . . . .	122
5.43	RMSE des charges hydrauliques à chaque itération du processus d'assimilation pour les puits de l'essai de pompage 1 et le schéma 500/20. . . . .	123
5.44	RMSE des charges hydrauliques à chaque itération du processus d'assimilation pour les puits de l'essai de pompage 2 et le schéma 500/20. . . . .	124
5.45	RMSE des charges hydrauliques à chaque itération du processus d'assimilation pour les puits de l'essai de pompage 1 et le schéma 500/20 amplifié. . . . .	125
5.46	RMSE des charges hydrauliques à chaque itération du processus d'assimilation pour les puits de l'essai de pompage 2 et le schéma 500/20 amplifié. . . . .	126
5.47	RMSE des charges hydrauliques à chaque itération du processus d'assimilation pour les puits de l'essai de pompage 1 et le schéma 1000/30 amplifié. . . . .	127
5.48	RMSE des charges hydrauliques à chaque itération du processus d'assimilation pour les puits de l'essai de pompage 2 et le schéma 1000/30 amplifié. . . . .	128
5.49	Charges hydrauliques relatives par rapport à la valeur initiale au début de l'essai. Mesure en rouge, ensemble initial en orange, ensemble final de la méthode 500/20 4x en bleu, et ensemble final PESTPP-IES en vert. . . . .	132
5.50	Charges hydrauliques relatives par rapport à la valeur initiale au début de l'essai. Mesure en rouge, ensemble initial en orange, ensemble final de la méthode 500/20 4x en bleu, et ensemble final PESTPP-IES en vert. . . . .	132
5.51	Charges hydrauliques relatives par rapport à la valeur initiale au début de l'essai. Mesure en rouge, ensemble initial en orange, ensemble final de la méthode 500/20 4x en bleu, et ensemble final PESTPP-IES en vert. . . . .	133
5.52	Charges hydrauliques relatives par rapport à la valeur initiale au début de l'essai. Mesure en rouge, ensemble initial en orange, ensemble final de la méthode 500/20 4x en bleu, et ensemble final PESTPP-IES en vert. . . . .	133
5.53	Charges hydrauliques relatives par rapport à la valeur initiale au début de l'essai. Mesure en rouge, ensemble initial en orange, ensemble final de la méthode 500/20 4x en bleu, et ensemble final PESTPP-IES en vert. . . . .	134

5.54 Charges hydrauliques relatives par rapport à la valeur initiale au début de l'essai. Mesure en rouge, ensemble initial en orange, ensemble final de la méthode 500/20 4x en bleu, et ensemble final PESTPP-IES en vert. . . . .	135
--	-----



# Liste des tableaux

3.1	Ratio d'information entre le nombre de paramètres d'états et le nombre d'observations dans la littérature. . . . .	26
4.1	Paramètres utilisés pour construire le champ de perméabilité de référence dans FGEN92, développé par (Robin <i>et al.</i> , 1993). . . . .	48
4.2	Paramètres de variogrammes utilisés pour générer la Figure 4.3. . . . .	50
4.3	Crépines des puits utilisées pour suivre la concentration maximale en sel dans les puits d'observations et de contrôles. . . . .	55
5.1	Synthèse de l'évolution des RMSE après la dernière assimilation par rapport à l'ensemble initial pour chacun des schémas étudiés. . . . .	129
5.2	Synthèse de l'évolution des RMSE relatives en pour cent par rapport à l'ensemble initial. Comparaison entre la méthode développée dans ce doctorat et la méthode PESTPP-IES. . . . .	130

# Introduction

Les sites contaminés prennent aujourd’hui une place de plus en plus importante à l’heure où la conscience écologique se développe et l’étalement urbain vient à manquer de place. Face à ces problématiques environnementales, économiques et sociales, il ne sera plus possible de laisser d’anciens sites industriels à l’abandon sans les restaurer pour leur donner une seconde vie. L’ampleur de la tâche à accomplir est visible sur le site Repère GTC (gestion des terrains contaminés) du Ministère de l’Environnement et de la Lutte contre les Changements Climatiques (MELCC) du Québec en visionnant le nombre de terrains contaminés présents.

Les modèles hydrogéologiques utilisés pour gérer les aquifères sont généralement déterministes. Peu de données étant disponibles, les incertitudes ne sont pas étudiées en profondeur, et un modèle est ajusté au mieux sur les données disponibles. Or avec ce faible nombre de données, il serait d’autant plus important de créer des modèles stochastiques pour évaluer les incertitudes. Dans les autres domaines des géosciences, comme dans les modélisations climatiques et l’exploration pétrolière, les méthodes d’ensembles sont utilisées pour répondre à ce problème. Même si la quantité de données sera proportionnellement plus réduite en environnement et apportera son lot de complexité, les méthodes d’ensemble peuvent être adaptées pour simuler les processus physiques d’écoulement des eaux souterraines. La méthode retenue sera inspirée de celle développée dans le domaine de l’exploration pétrolière pour mettre à jour des modèles géologiques du sous-sol à l’aide de données temporelles. Une meilleure compréhension du sous-sol et des incertitudes associés aux modèles permettra des décisions plus éclairées dans le futur pour la gestion de ce site contaminé.

L’objectif de ce doctorat sera de produire une méthode d’assimilation de données pour mettre à jour des modèles d’écoulement des eaux souterraines. Les méthodes d’ensembles reposent sur des ensembles de modèles équiprobables. Cet ensemble de modèles permet d’évaluer les incertitudes

sur le sous-sol, et donne une plage d'observations plausible plutôt qu'une solution unique. Cette nouvelle méthode d'assimilation sera validée dans un premier temps sur un cas synthétique. Ce cas synthétique sera un analogue d'aquifère côtier, avec une intrusion saline. Dans un second temps, la nouvelle méthode sera appliquée à un ancien site industriel avec une géologie complexe.

Les données nécessaires à la réalisation du cas réel ont été obtenues grâce à une collaboration entre l'Institut National de la Recherche Scientifique (INRS) de Québec et le Ministère de l'Environnement et de la Lutte contre les Changements Climatiques (MELCC). Ce partenariat a permis l'accès à un site suivi depuis une trentaine d'années par le MELCC. La quantité d'information disponible pour le site des anciennes lagunes de Mercier est exceptionnelle pour une étude environnementale, ce qui en fait un jeu de donnée idéal pour ce doctorat. La quantité de puits disponibles au droit de la zone d'étude n'a rien à envier à des projets habituellement réservés aux domaines minier et pétrolier. Cette densité d'information va permettre de tester de nouvelles méthodes pour définir des modèles hydrogéologiques complexes.

Le document débutera par une revue bibliographique pour définir l'état des connaissances sur les méthodes d'assimilation d'ensembles. Que ce soit dans le cadre d'application hydrogéologique ou non. Une synthèse des défauts de ces méthodes sera réalisée ainsi qu'une présentation des solutions pour les atténuer. La théorie des méthodes d'ensembles sera expliquée en détail dans le chapitre suivant. Cette explication permettra de visualiser le processus d'assimilation et les équations seront données pour assister le lecteur dans sa compréhension.

Les nouveaux apports réalisés dans le cadre de ce doctorat seront présentés dans la partie méthodologie. Ces apports seront par la suite testés de plusieurs manières. Dans un premier temps, une étude complète de ces nouvelles méthodes sera réalisée sur un cas synthétique simulant un aquifère côtier. Les méthodes les plus efficaces seront enfin appliquées sur le cas réel.

# Contributions

- Présentation d'une affiche au Venice International University Graduate Seminar. Le thème du séminaire était Hydrogeophysical inversion and data assimilation for the characterization and monitoring of coastal aquifers;
- Examen pré-doctoral avec présentation d'un premier exemple d'assimilation en 2 dimensions;
- Développement d'une nouvelle méthode d'analyse locale et de localisation pour la méthode d'assimilation multiple développée par Emerick & Reynolds (2013);
- Création d'un essai synthétique complet, des données au modèle d'écoulement pour tester la nouvelle méthode;
- Création d'un modèle géostatistique réel à partir des travaux de Claprood *et al.* (2022). Confirmation des performances de la méthode développée sur le modèle réel plus complexe;
- Article scientifique en soumission pour la revue Applied Computing and Geosciences présentant la nouvelle méthode d'analyse locale sur un cas synthétique;
- Publication des codes numériques commentés et documentés en Open Source sur [https://github.com/Thomas-Beraud/FEFLOW\\_MDA\\_MPMW.github.io](https://github.com/Thomas-Beraud/FEFLOW_MDA_MPMW.github.io);
- Présentation orale à l'International Association for Mathematical Geosciences 2022 dans la session "Filter and Smoothers. Filters or Smoothers?";
- Présentation orale à l'atelier sur les aquifères côtiers de Rennes en 2022.



# Chapitre 1

## Revue bibliographique

### 1.1 Assimilation de données

Les méthodes d'assimilation de données sont utiles pour estimer les paramètres d'état d'un système physique grâce à des observations et à un ensemble de modèles de paramètres. L'ensemble initial de modèles contient toutes les connaissances et les incertitudes associées au problème. En géoscience, l'ensemble va typiquement pouvoir contenir la géologie, les données géophysiques et pétrophysiques, les connaissances scientifiques, les modèles analogues, etc. Chacun des modèles individuels honorera les données mesurées (perméabilités, porosité, affleurement géologique, etc.) grâce à par exemple des simulations géostatistiques (Goovaerts, 1997). Evensen (1994) décrit les méthodes d'assimilation dans le cadre de systèmes océaniques, avec l'état initial des systèmes contenant forcément des erreurs. L'ensemble initial aura un impact sur les prédictions, même si le même modèle déterministe est utilisé pour simuler l'ensemble. Aucun modèle de l'ensemble ne devrait être meilleur ou pire qu'un autre, chacun étant équiprobable. Les méthodes d'assimilation peuvent uniquement être utilisées dans le cadre d'un problème pour lequel la physique est connue et que des solutions numériques existent pour la simuler.

La variabilité des paramètres dans les modèles de l'ensemble sera plus forte loin des points d'observation. L'ensemble de modèles permet d'approximer la matrice de covariance entre les paramètres des modèles (Raanes, 2015b). Pour éviter les fausses corrélations spatiales entre paramètres,

et s’approcher de la matrice réelle de covariance, un nombre suffisant de modèles sera nécessaire dans l’ensemble.

Le nombre de modèles requis pour l’assimilation de gros modèles géoscientifiques n’est pas clairement défini, Evensen (1997), Anderson & Anderson (1999), Mitchell *et al.* (2002), Houtekamer *et al.* (2005) et Altaf *et al.* (2014) indiquent que l’ensemble doit contenir entre 10 et 100 modèles. Certains auteurs recommandent plusieurs centaines de modèles, comme Emerick & Reynolds (2013), Zhou *et al.* (2011), Raanes (2015b), voire un millier pour Evensen (1997) et Zovi *et al.* (2017). La capacité de prédiction de l’ensemble devrait augmenter avec le nombre de modèle dans l’ensemble.

## 1.2 Historique

Historiquement, le filtre de Kalman a été introduit par Kalman (1960) comme une méthode d’assimilation de données pour résoudre des problèmes linéaires bruités. En 1994, Evensen (1994) développa le filtre de Kalman d’Ensemble (EnKF pour Ensemble Kalman Filter) pour gérer des systèmes non linéaires. Les premières applications étaient pour les prédictions météorologiques et la modélisation des courants océaniques. Par la suite, de nombreux autres domaines ont profité de ces développements. Dong *et al.* (2006) utilisa les EnKF pour assimiler des données sismiques 4D dans une modélisation de réservoir. Il nota que les méthodes classiques d’adjoint étaient plus restrictives par leur nécessité d’accès au simulateur de réservoir, i.e. au code source, pour calculer les gradients. Les EnKF utilisent la moyenne et la covariance inférées depuis l’ensemble pour estimer les perméabilités et covariances réelles. Il est donc possible d’assimiler l’ensemble de modèles sans avoir accès au code propriétaire du simulateur utilisé. En contrepartie, la taille de l’ensemble doit être suffisante pour estimer la moyenne et la covariance, tout en explorant dimensionnalité de l’ensemble. Johns & Mandel (2008) a appliqué les EnKF pour de modèles de propagation de feu. De nombreuses variations des méthodes d’ensembles sont développées, Evensen (2009b) en résume quelques-unes dans sa publication. Les différences principales résident dans la manière d’incorporer les données à chaque pas de temps (Filter) ou de manière unique (Ensemble Smoother), ainsi que dans la manière de calculer numériquement les solutions. Crestani *et al.* (2013) a utilisé les EnKF et Ensemble Smoother, pour des essais de traceur en 2D. Bauer *et al.* (2015) utilisait les méthodes d’ensemble pour de la prédiction météorologique sur des centaines de variables, sur des grilles de milliards de points avec uniquement 100 modèles dans l’ensemble. De nos jours, l’European Centre

for Medium-Range Weather Forecasts (ECMWF) réalise des assimilations de données sur  $10^7$  observations par jour sur des grilles de 650 millions de points. Ji *et al.* (2018) applique les EnKF pour prédire les mouvements de fumée dans les tunnels en cas d'incendie. Dans son article, Defforge *et al.* (2019) utilise un ES pour simuler les vents atmosphériques à une échelle locale à travers un champ d'éolienne. Leur ensemble présentait des améliorations avec uniquement 5 modèles. Wang *et al.* (2019) a estimé la topographie du fond marin en 2 dimensions grâce à une méthode d'assimilation d'ensemble. La qualité de l'assimilation de l'ensemble dépend de la distribution spatiale des observations dans l'espace. La qualité augmentera avec la couverture du domaine. Tous ces exemples démontrent l'efficacité et la versatilité des méthodes d'ensembles dans de nombreux domaines.

### 1.3 Les différentes familles d'assimilation

Plusieurs types de méthodes d'assimilation ont été utilisées dans tous les exemples précédents : Ensemble Kalman Filter (EnKF), Ensemble Kalman Smoother (EnKS) et Ensemble Smoother (ES). L'EnKF consiste en une mise à jour du modèle à chaque nouvelle donnée temporelle, puis une propagation de la simulation de ce nouveau modèle jusqu'à la prochaine mesure temporelle.

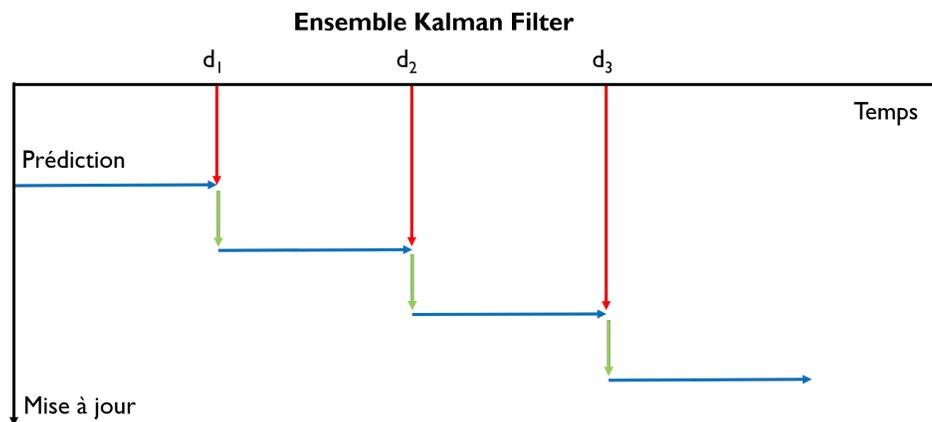


Figure 1.1 – Schéma de la mise à jour dans l'ensemble Kalman filter. En horizontal l'axe du temps avec les mesures temporelles effectuées à différents temps. En vertical l'axe des mises à jour. La flèche bleue représente la simulation de l'ensemble avec l'apport d'information provenant des données (flèches rouges). La flèche verte indique la mise à jour de l'ensemble grâce à l'EnKF. L'ensemble indiqué en bleu est modifié à chaque nouvelle mesure. Modifié d'après Evensen (2009a).

L'EnKS quant à lui, correspond à une mise à jour du modèle à l'instant  $t$  puis à une propagation depuis le temps initial  $t_0$  jusqu'à la prochaine mesure non assimilée. Chaque mise à jour implique donc le redémarrage depuis le temps initial.

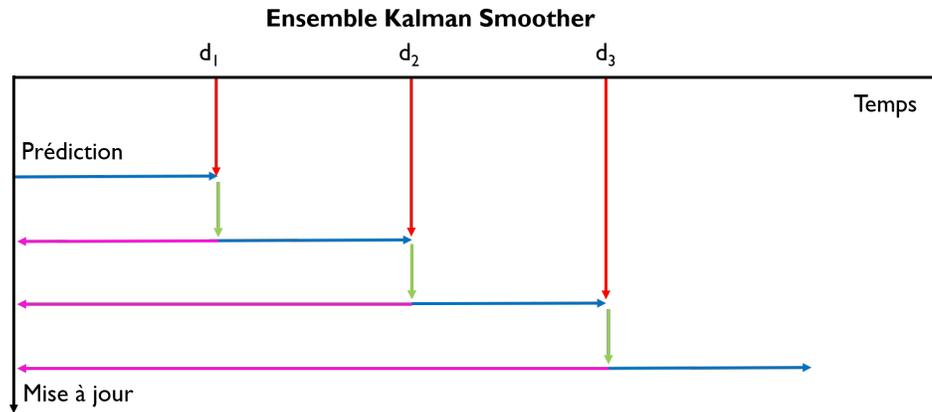


Figure 1.2 – Schéma de la mise à jour dans l'ensemble Kalman Smoother. En horizontal l'axe du temps avec les mesures temporelles effectuées à différents temps. En vertical l'axe des mises à jour. La flèche bleue représente la simulation de l'ensemble avec l'apport d'information provenant des données (flèches rouges). L'ensemble indiqué en bleu est modifié à chaque nouvelle mesure. La flèche rose représente l'ensemble mis à jour avec une propagation de l'information jusqu'au temps initial de la simulation. Modifié d'après Evensen (2009a).

La variante de l'ES, pour Ensemble Smoother également définie par Evensen (2009a), permet de faire une seule mise à jour pour l'ensemble des données temporelles. La correction va être évaluée en assimilant toutes les données en une seule fois. La simulation est conduite du temps initial  $t_0$  jusqu'à la dernière donnée temporelle mesurée en continu.

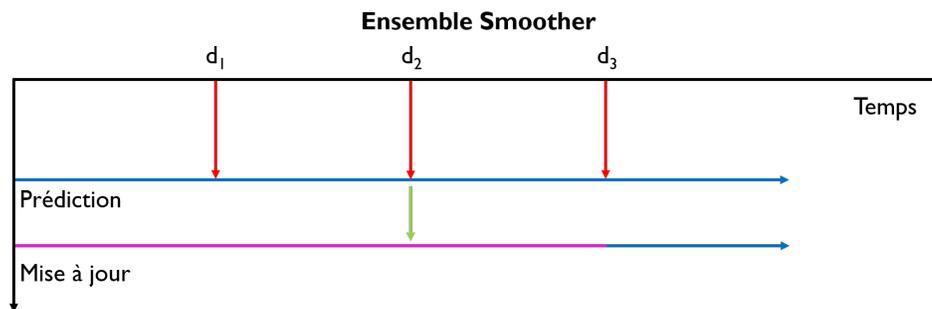


Figure 1.3 – Schéma de la mise à jour dans l'ensemble smoother. En horizontal l'axe du temps avec les mesures temporelles effectuées à différents temps. En vertical l'axe des mises à jour. La flèche bleue représente la simulation de l'ensemble avec l'apport d'information provenant des données (flèches rouges). La flèche rose représente l'ensemble mis à jour. La flèche verte indique la mise à jour unique de l'ensemble. Modifié d'après Evensen (2009a).

Evensen (2004), Chen & Oliver (2013) et Crestani *et al.* (2013) indiquent que les EnKF sont plus performants que les ES, dans le cas de courants océaniques pour Evensen (2004) et pour des essais de traceurs pour Crestani *et al.* (2013). Pour les deux auteurs, cette surperformance était due aux mises à jour successives des EnKF qui gardaient l'ensemble cohérent. Dans le même temps, Sakov *et al.* (2010), Skjervheim *et al.* (2011) et Li *et al.* (2018) notaient une importante réduction dans les temps de simulation en utilisant les ES par rapport aux EnKF. Ce gain de temps provient du fait que les simulateurs n'ont pas besoin d'être arrêtés et redémarrés à chaque pas de temps. Dans son article, Emerick & Reynolds (2012) introduit l'idée d'assimilation multiple ou Multiple Data Assimilation (MDA). Les MDA sont faciles à implémenter dès que l'on possède déjà des EnKF ou ES fonctionnelles. D'après l'auteur, 4 assimilations successives, ou MDA 4x, semblent être le meilleur compromis entre l'amélioration de l'erreur moyenne au carré (RMSE) de l'ensemble et le nombre d'assimilations. Chen & Oliver (2013) notait que dans le cas de système fortement non linéaire, les ES ne performait pas très bien. Ce problème a été résolu en utilisant les assimilations de données multiples. Les meilleurs résultats obtenus par les auteurs ont été atteints en utilisant une méthode appelée Levenberg–Marquardt-Ensemble Randomized Maximum Likelihood (LM-EnRML), qui est également de la famille des méthodes itératives. L'algorithme de Levenberg-Marquardt permet de déterminer la direction et l'intensité de la mise à jour des modèles. Les ES itératifs reposent sur l'idée de simuler tous les modèles du temps initial au final, assimiler l'ensemble des observations, puis resimuler tout l'ensemble à nouveau, et ceci pour un nombre prédéfini d'itérations. Les méthodes LM-EnRML produisent de meilleurs résultats que les ES MDA avec 16 itérations Chen & Oliver (2013) grâce à des régularisations / TSVD itératives. Cette méthode est donc numériquement plus complexe. Cependant, dans le cas de simulations numériquement très lourdes, les auteurs ont montré que 7 itérations de LM-EnRML étaient nécessaires pour atteindre les mêmes performances que les ES MDA avec 4 itérations. Emerick & Reynolds (2013) démontra que le ES-MDA 4x fonctionnait bien sur des problématiques de réservoirs pétroliers. Les ES-MDA rentrent également dans la catégorie des méthodes itératives, avec plusieurs petites corrections dans les paramètres de l'ensemble, contre une large correction dans le cas des méthodes Gauss-Newton. D'après la littérature, les méthodes ES-MDA 4x produisent de meilleurs RMSE que les EnMRL, et une meilleure correspondance avec les données que les EnKF pour un coût de calcul comparable. Enfin, Raanes *et al.* (2019) a démontré une faible différence dans la performance entre les méthodes ES-MDA et En-RML/IEnKS. La méthode ES-MDA numériquement et informatiquement plus simple performe quasiment aussi bien que les méthodes plus complexes En-RML et IEnKS.

À la suite de cette revue de littérature, le choix d'implémenter une méthode ES-MDA 4x au regard des bénéfiques coûts de calcul / performance a été fait.

## 1.4 Localisation, Analyse Locale et Inflation

Malgré les bons résultats relevés dans la littérature, certains problèmes sont inhérents à ces méthodes. Evensen (2009a) indique qu'avec de très larges modèles, il est impossible d'obtenir autant de modèles que de paramètres à estimer. Cette incertitude est représentée par un ensemble de modèles, plutôt qu'une matrice de covariance explicite entre paramètres. La taille de l'ensemble limite l'espace des solutions et peut donc introduire de fausses corrélations entre paramètres (spurious correlation). Les méthodes de localisations ont pour but de réduire l'impact des observations loin des points de mesures. L'analyse locale quant à elle sert à augmenter l'espace des solutions pour l'ensemble tout en réduisant les fausses corrélations entre paramètres. Les fausses corrélations entre mesures et paramètres des modèles mènent à de petites mises à jour sans sens physique dans les modèles de l'ensemble. Ces mises à jour purement mathématique vont réduire la variance dans l'ensemble, ce qui peut dans des cas extrêmes, réduire l'ensemble à des modèles identiques. Deux méthodes principales existent pour réduire ces fausses corrélations, la localisation de la covariance et l'analyse locale. Sakov & Bertino (2011) indiqua que les deux méthodes devraient donner les mêmes résultats. Le choix de la méthode sera basé sur l'efficacité d'implémentation pour un problème donné.

### 1.4.1 Localisation

La localisation de la covariance permet de réduire numériquement les valeurs dans la matrice de covariance pour une distance numérique donnée. La localisation permet artificiellement de réduire la taille du domaine d'influence pour chaque observation, grâce à la réduction de la matrice de covariance. Avec un nombre limité d'ensembles, la matrice de covariance définie sera sous-optimale (rank deficient). La localisation pourra être réalisée sur la matrice de covariance grâce à la multiplication de cette dernière par une fonction de corrélation. Hunt *et al.* (2007) et Chen & Oliver (2013) indiquent que les méthodes de localisation sont souvent nécessaires pour obtenir de bonnes performances dans l'assimilation, particulièrement avec beaucoup de données et un petit ensemble

de modèles. Les méthodes d'assimilation permettent de réduire le temps d'assimilation et améliorer les résultats. La localisation est plus facile à implémenter avec les ES que les EnKF, car les ES utilisent les paramètres d'états initiaux plutôt que des états dynamiques en cours de simulation.

### 1.4.2 Analyse Locale

L'analyse locale permet de prendre uniquement en compte un sous-ensemble d'observations ou de paramètres, ce qui réduit l'impact de la faible taille de l'ensemble. L'analyse locale permet d'approximer l'erreur de covariance dans un sous-domaine du modèle. La discontinuité entre sous-domaines est adressée par Hunt *et al.* (2007) en augmentant la covariance proche de chaque bordure de sous-domaine. L'analyse locale est bien adaptée à la parallélisation et est un bon choix pour les systèmes de grande échelle. Les sous-domaines locaux peuvent être définis autour de chaque point d'observation pour réduire physiquement les corrélations entre observations éloignées. Carrassi *et al.* (2018) indique qu'à la fin du processus d'assimilation, ES et EnKF auront assimiler la même quantité d'information. La localisation de domaine cherche à uniquement mettre à jour les paramètres sur certains éléments de la grille. Le nombre d'analyses locales à réaliser va augmenter linéairement avec le nombre de paramètres. Avec de nombreux paramètres, l'assimilation pourrait prendre trop de temps. Chen & Oliver (2017) avance que l'analyse locale a été particulièrement utilisée dans les prédictions météo et les réservoirs pétroliers. L'idée principale derrière l'analyse locale est de décomposer le modèle complet en plusieurs sous-modèles pour l'assimilation. La distance critique d'assimilation définit la région locale, cette taille peut varier de quelques éléments du modèle au modèle entier. Dans le cas du modèle entier, l'assimilation sera équivalente à celle du modèle sans localisation. Janjić *et al.* (2011) explique que si toutes les observations sont utilisées dans des sous-domaines disjoints, alors l'assimilation sera équivalente à une assimilation sans analyse locale. Il est possible de donner un poids à chaque observation en fonction de la distance au paramètre à mettre à jour.

Dans la littérature, les analyses locales consistent à mettre à jour chaque point de la grille avec un sous-ensemble des observations comme décrit par Sakov & Bertino (2011) et Carrassi *et al.* (2018). Janjić *et al.* (2011) explique que la localisation par domaine consiste à mettre à jour un domaine avec certaines observations basées sur la distance, en une seule assimilation. Une limitation existe avec cette approche, Carrassi *et al.* (2018) remarquait qu'avec l'augmentation du nombre de paramètres

dans le domaine, le nombre d'analyses locales allait augmenter linéairement avec ce nombre. Cette augmentation finit par atteindre un seuil rendant l'assimilation impossible. Les méthodes d'analyses locales convergent généralement plus vite que la localisation seule lorsque la quantité de données est largement supérieure au nombre de modèles.

### 1.4.3 Inflation

Suite aux multiples assimilations successives dans les MDA, (Emerick & Reynolds, 2013) note que le bruit sur les données doit être augmenté par un facteur égal au nombre d'assimilations prévues. Sans inflation, l'assimilation répétée pourrait faire s'effondrer la variance de l'ensemble, réduisant ce dernier à quasiment un modèle unique. (Chen & Oliver, 2013) ajoute que l'augmentation de la variance sur les erreurs de mesures semble améliorer la convergence et la RMSE des MDA.

De nombreuses autres applications existent pour les méthodes d'assimilation d'ensemble. Elles ne seront pas toutes détaillées, car elles sont très bien expliquées dans les papiers de leurs auteurs respectifs. Pour les applications retenues dans ce doctorat, il a été décidé de se baser sur les travaux de (Emerick & Reynolds, 2013) et (Raanes, 2015a). L'objectif est d'évaluer la possibilité d'application d'une méthode ES-MDA 4x pour des modélisations hydrogéologiques avec un nombre de données limitées. L'apport de ce doctorat sera majoritairement sur les méthodes de localisation et d'analyses locales pour gérer la forte disparité entre les nombres d'observations disponible et la complexité des modèles à mettre à jour.

# Chapitre 2

## Théorie

### 2.1 Introduction à la modélisation

Pour aider le lecteur à comprendre les méthodes d'ensemble, une explication rapide des principes de la modélisation numérique va être effectuée.

Un modèle numérique a pour but de reproduire de manière la plus fidèle possible, une partie du monde réel. Ce modèle est généralement nécessaire pour répondre à une problématique spécifique. Il peut s'agir par exemple d'un modèle de réservoir pétrolier pour prédire la production future, ou bien un modèle d'aquifère pour planifier la disponibilité en eau tout au long de l'année.

Ces modèles reposent sur des paramètres d'états, ces paramètres représentent les propriétés physiques du sous-sol. Cela peut correspondre à la perméabilité, la porosité, la teneur en eau, la saturation de la phase non aqueuse, etc. Lors de la conception du modèle, ces données sont obtenues en réalisant par exemple des forages, des essais de pompages ou bien de la géophysique. Le problème avec ces données, c'est qu'elles ne sont disponibles qu'en certains points de l'espace et du temps où elles ont été obtenues.

Il faut donc réaliser un modèle en 3 dimensions à partir de certains points dans l'espace. Une manière commune de gérer ce problème est de réaliser des modèles géostatistiques. Sans rentrer dans les détails, ces modèles vont corréliser les statistiques des données et leur distribution spatiales. Il sera donc possible d'estimer les valeurs des paramètres d'états dans tout le modèle, en dehors

des points de données. Ces simulations de paramètres d'états contiendront une incertitude dans le modèle 3D. Cette incertitude augmentera en s'éloignant des points de données acquis dans le monde réel. Chaque modèle simulé donnera une représentation possible du sous-sol à partir des données disponibles.

Pour évaluer la qualité d'un modèle obtenu par simulation, il est possible de simuler un processus physique à l'aide de ce dernier et de le comparer au processus réel. Dans le cas d'un modèle d'aquifère, des mesures sont disponibles dans le monde réel, il s'agit souvent de niveau d'eau dans le sous-sol, mais il pourrait également s'agir de concentration en espèces chimiques dans l'eau. Pour évaluer la qualité d'un modèle du sous-sol, il suffit de comparer les simulations dans le modèle par rapport aux mesures du monde réel.

## 2.2 Méthodes d'ensemble

L'objectif des méthodes d'assimilation de données est d'approximer les paramètres d'états d'un modèle à partir d'observations, de mesures et d'un ensemble de modèles. L'ensemble de modèles va contenir toute l'information à priori disponible, les connaissances géologiques, les mesures géophysiques, des modèles analogues, etc. Chacun de ces modèles va individuellement honorer les paramètres d'états mesurés sur site (perméabilité, porosité, etc.) grâce à par exemple, des simulations géostatistiques (Goovaerts, 1997).

Cette variabilité est présente entre chacun des modèles lorsque les parties étudiées se situent loin des données (forages, sismiques, affleurements, etc.). C'est cette variabilité, synonyme d'incertitude, qui va être évaluée dans l'ensemble des modèles. Il faut cependant avoir un nombre de modèles suffisamment important par rapport au nombre de paramètres à étudier pour pouvoir représenter les directions d'incertitudes (matrices de covariances) (Raanes, 2015b). La taille de l'ensemble pour des problèmes de modélisation géoscientifique est souvent supérieure de l'ordre de la centaine de modèles (Emerick & Reynolds, 2013) et (Raanes, 2015b).

Pour l'explication du principe des filtres d'ensemble de Kalman, un exemple va être expliqué comme s'il s'agissait d'un petit modèle hydrogéologique en 2 dimensions. Ce modèle numérique serait basé uniquement sur une grille simple de porosité en 2 dimensions. Dans un cas réel, plusieurs paramètres pourront être estimés en parallèle (porosité, perméabilité, etc.). Chacune des grilles de

l'ensemble contient une simulation de porosité obtenue grâce à l'incorporation des informations à priori. Un exemple de simulation obtenu est donné sur la Figure 2.1, ici il s'agit de la simulation  $n$  sur un nombre de simulations  $N$ . Sur cette grille, 20 paramètres de porosité  $\phi$  sont simulés.

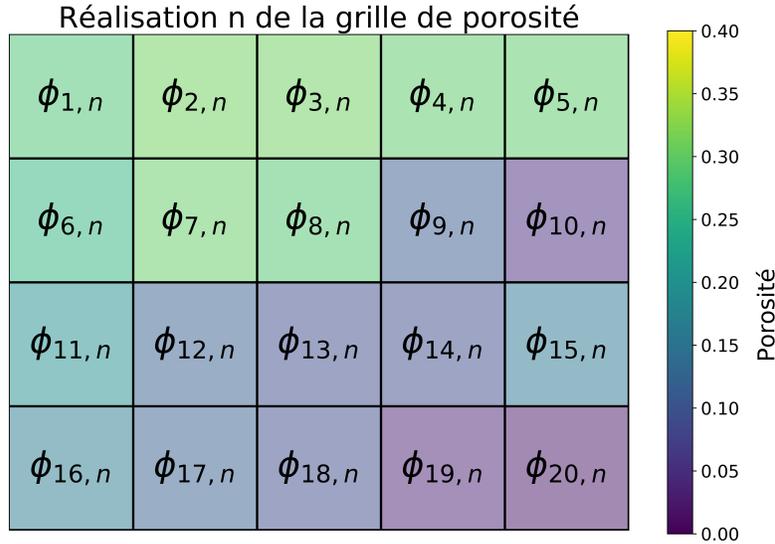


Figure 2.1 –  $n^{\text{e}}$  grille de porosité obtenue par simulation géostatistique pour une grille en 2 dimensions contenant 20 cases.

Le nombre de réalisations utilisées pour analyser les statistiques de l'ensemble est noté  $N$ . La dimension de  $N$  permet de plus ou moins bien représenter la variabilité et les statistiques de la variable étudiée (Raanes, 2015b). Chaque grille est déroulée dans un vecteur à 1 dimension noté  $x_n$  contenant l'ensemble des paramètres. L'ensemble des simulations est regroupé dans une matrice de dimension  $m \times N$ . Le nombre de paramètres à estimer, ici égal au nombre de cases de la grille est noté  $m$ .

L'ensemble est noté  $E$  :

$$E = [x_1, \dots, x_n, \dots, x_N] \in \mathbb{R}^{m \times N} \quad (2.1)$$

Grâce aux simulations géostatistiques, il est possible de générer de multiples réalisations respectant les données mesurées et les variogrammes imposés. Les différentes méthodes utilisées historique-

ment pour ces simulations sont décrites en détail par (Goovaerts, 1997). Ces multiples réalisations équiprobables sont la matière principale pour les filtres d'ensemble de Kalman. Toutes ces simulations donnent des grilles en deux dimensions dans le cadre de l'exemple donné sur la Figure 2.2. Chaque grille présente des variations de porosité inhérentes aux simulations spatiales des paramètres d'états entre les points de données.

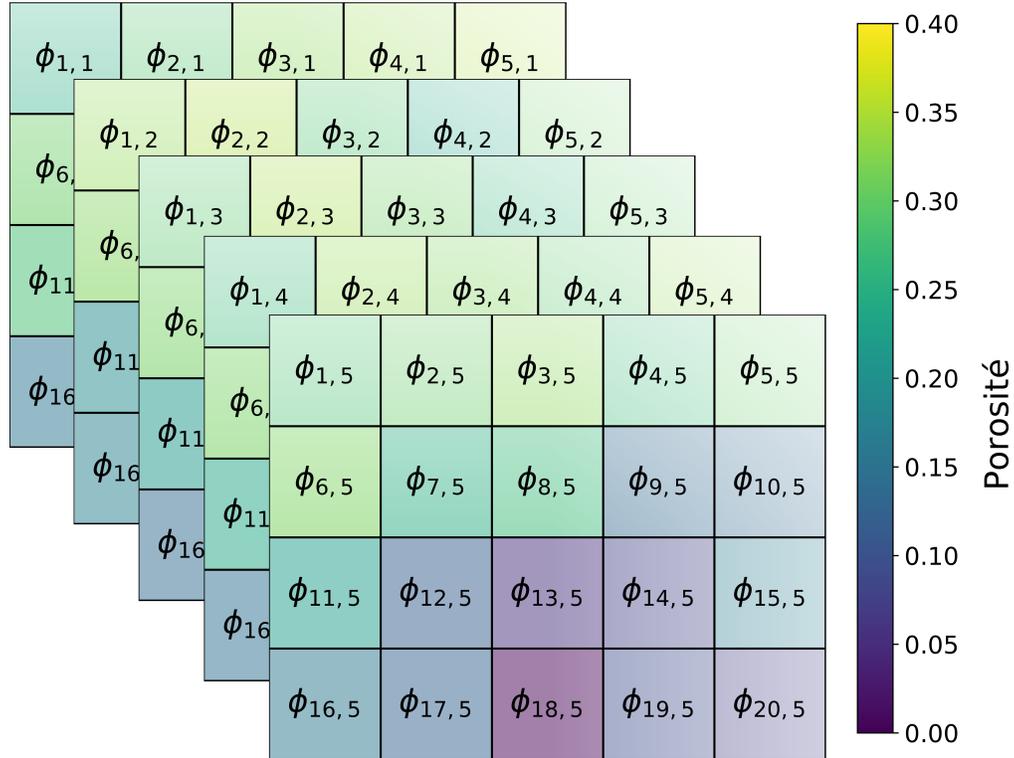


Figure 2.2 – 5 réalisations de porosité obtenue par simulation pour une grille en 2 dimensions contenant 20 cases.

En pratique, les porosités de la première réalisation  $n = 1$  stockées dans le vecteur  $x_1$  vont donc aller de  $\varphi_{1,1}$  à  $\varphi_{m,1}$ . Le modèle  $n$  ira donc quand à lui de  $\varphi_{1,n}$  à  $\varphi_{m,n}$ . Toutes les réalisations sont stockées dans une grande matrice contenant l'ensemble  $E$  :

$$E = \begin{bmatrix} \varphi_{1,1} & \cdots & \varphi_{1,N} \\ \vdots & \ddots & \vdots \\ \varphi_{m,1} & \cdots & \varphi_{m,N} \end{bmatrix}$$

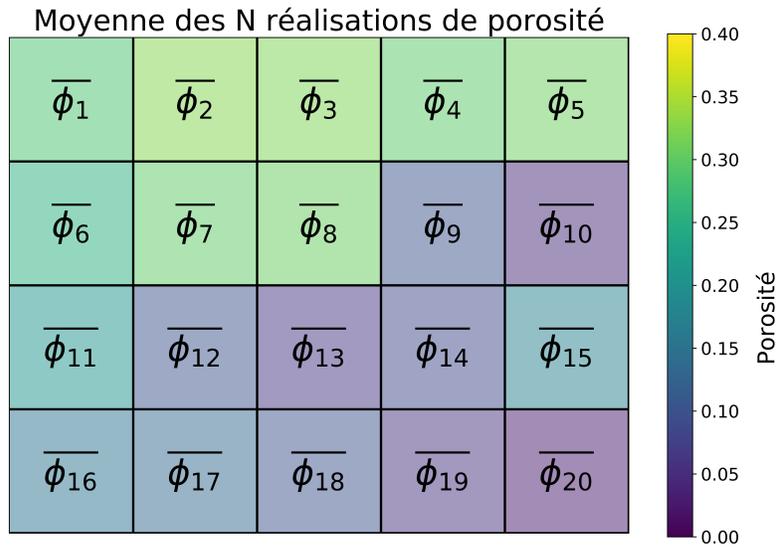
Les notations reprises ici sont issues des travaux de (Raanes, 2015b). L'ensemble  $E$  étant constitué il est maintenant possible d'approximer la moyenne et la covariance à partir de ce nombre fini de réalisations. Ces valeurs sont notées avec une barre supérieure pour les différencier de la moyenne et de la covariance réelle qui sont inconnues.

$$\bar{x} = \frac{1}{N} \sum_{n=1}^N x_n, \quad \bar{P} = \frac{1}{N-1} \sum_{n=1}^N (x_n - \bar{x})(x_n - \bar{x})^T. \quad (2.2)$$

$x_n$  est simplement le vecteur de la moyenne par case du modèle, il s'agit donc de l'espérance de porosité pour une case donnée du modèle grâce à l'ensemble des  $N$  simulations réalisées.

$$\bar{x} = \begin{bmatrix} \frac{\varphi_{1,1} + \varphi_{1,2} + \dots + \varphi_{1,N}}{N} \\ \vdots \\ \frac{\varphi_{m,1} + \varphi_{m,2} + \dots + \varphi_{m,N}}{N} \end{bmatrix}$$

Cela revient donc à additionner la valeur de porosité de chaque simulation puis de la diviser par le nombre de simulations. Le résultat est donc une grille moyenne de porosité donnée sur la Figure 2.3. Cette moyenne est nécessaire pour l'étape suivante consistant à mesurer les déviations individuelles de chaque modèle par rapport à la moyenne des modèles.



**Figure 2.3** – Porosité moyenne calculée sur les 5 réalisations de porosité obtenue sur la Figure 2.2.

Une fois le vecteur moyen  $\bar{x}$  calculé, il est possible de définir la matrice d'anomalies  $A$ . Cette matrice de dimension  $\mathbb{R}^{m \times N}$  mesure l'écart entre chaque case du modèle et la moyenne pour cette case :

$$A = [x_1 - \bar{x}, \dots, x_n - \bar{x}, \dots, x_N - \bar{x}] \in \mathbb{R}^{m \times N}, \quad (2.3)$$

$$A = \begin{bmatrix} \varphi_{1,1} - \bar{x}_1 & \varphi_{1,2} - \bar{x}_1 & \dots & \varphi_{1,N} - \bar{x}_1 \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{m,1} - \bar{x}_m & \varphi_{m,2} - \bar{x}_m & \dots & \varphi_{m,N} - \bar{x}_m \end{bmatrix}$$

Dans l'exemple visuel donné pour assister le raisonnement mathématique, 5 grilles d'anomalies peuvent être calculées. Chaque réalisation individuelle de porosité sur la Figure 2.2 est soustraite à la grille de porosité moyenne de la Figure 2.3. Il en résulte des anomalies négatives et positives autour de la moyenne comme indiqué sur la Figure 2.4.

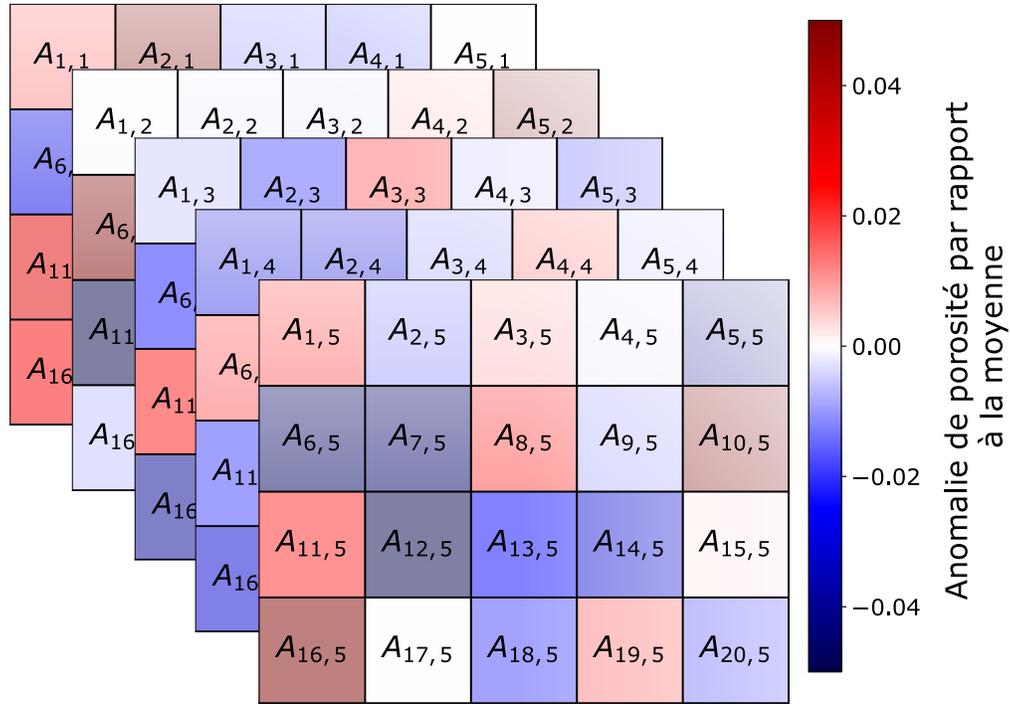


Figure 2.4 – 5 grilles d’anomalies obtenues depuis les simulations pour une grille en 2 dimensions contenant 20 cases.

Cette matrice d’anomalies permet de calculer la covariance entre les différents modèles. La covariance revient à étudier si les anomalies de paramètres, dans le cas de l’exemple, les anomalies de porosité ont tendance à évoluer dans une même direction ou à l’opposé.

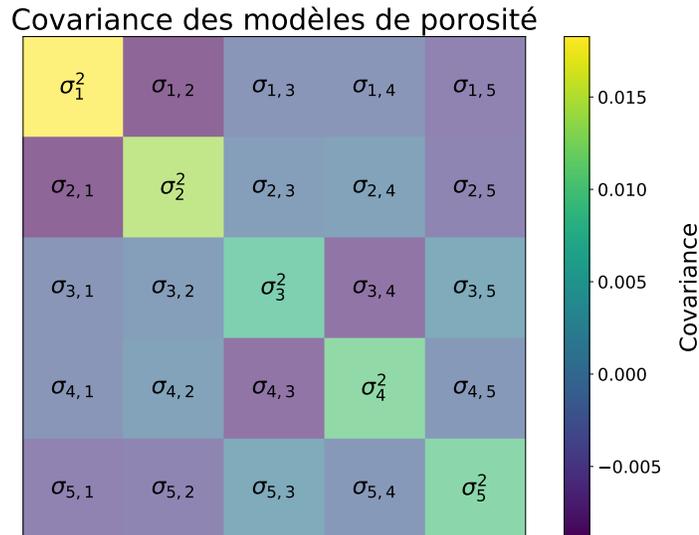
La matrice  $\bar{P}$  est la matrice de covariance qui traduit quelles cases du modèle ont tendance à varier ensemble. La dimension de  $\bar{P} \in \mathbb{R}^{N \times N}$ .

Les anomalies étant les écarts à la moyenne, il est possible de calculer la matrice de covariance  $\bar{P}$  directement à partir de la matrice  $A$  :

$$\bar{P} = \frac{AA^T}{N - 1} \tag{2.4}$$

$$\bar{P} = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} & \dots & \sigma_{1,m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{m,1} & \sigma_{m,2} & \dots & \sigma_m^2 \end{bmatrix}$$

La matrice de covariance  $P$  de la Figure 2.5 traduit les corrélations entre modèles. Une covariance de 0 indique que les modèles ne présentent pas de corrélation ou très faibles et locales. Une covariance positive indique que deux modèles ont des régions présentant des anomalies communes. Enfin, une corrélation négative indique des anomalies opposées. Cette matrice sera utile pour savoir quelles parties du modèle seront mises à jour pour mieux prédire les données observées.



**Figure 2.5** – Matrice de covariance des 5 modèles de porosité.

Les méthodes d'assimilation d'ensembles reposent sur une étape de prédiction et une étape d'analyse. L'étape de prédiction correspond à la simulation de valeurs physiques dans l'ensemble des modèles paramètres d'états. Si l'on considère un exemple d'aquifère, chacun des modèles de porosité de la Figure 2.2 sera simulé. Le résultat serait donc 20 simulations de niveaux d'eau. Une fois les simulations réalisées, l'étape de prédiction est terminée. L'étape suivante est donc l'étape

d'analyse. Lors de l'analyse, les différences entre les observations de niveaux d'eau dans les modèles simulés et les mesures réelles vont servir à mettre à jour les paramètres d'états.

À partir des paramètres d'états du modèle, il va être possible de réaliser une simulation d'écoulement et évaluer les paramètres dynamiques du modèle comme la charge piézométrique.  $N$  simulations d'écoulement vont être réalisées et de la même manière, les anomalies de charges hydrauliques vont être observées par rapport à la moyenne. Dans la littérature, Evensen (1994) définit une fonction d'observation  $H$  pour faire le lien entre les paramètres d'états  $m$  et les observations. Cette matrice permet de passer d'un espace à  $m$  paramètres à l'espace des observations de dimensions  $p$  (nombre de mesure). En pratique, cette matrice  $H$  n'est pas calculée, il s'agit d'un simulateur de processus physique. Pour des problématiques hydrogéologiques, cette matrice  $H$  est inconnue et est remplacée par un logiciel comme FEFLOW qui prend un champ de paramètres en entrée (porosité, perméabilité, etc.) et donne des charges hydrauliques ou des concentrations en sorties. Ces observations, de niveaux d'eau par exemple, seront stockées dans une matrice d'observations d'observation  $y \in \mathbb{R}^{p \times N}$ . Avec  $\nu$  les valeurs observées dans les modèles :

$$y = \begin{bmatrix} \nu_{1,1} & \cdots & \nu_{1,N} \\ \vdots & \ddots & \vdots \\ \nu_{p,1} & \cdots & \nu_{p,N} \end{bmatrix}$$

Cette matrice regroupera donc une ou plusieurs observations par modèle d'écoulement. Ces observations permettront de calculer la moyenne des observations, la matrice d'anomalies d'observations  $Y$ .

$$Y = [\nu_1 - \bar{\nu}, \dots, \nu_n - \bar{\nu}, \dots, \nu_N - \bar{\nu}] \in \mathbb{R}^{p \times N}, \quad (2.5)$$

$$Y = \begin{bmatrix} \nu_{1,1} - \bar{\nu}_1 & \nu_{1,2} - \bar{\nu}_1 & \cdots & \nu_{1,N} - \bar{\nu}_1 \\ \vdots & \vdots & \ddots & \vdots \\ \nu_{p,1} - \bar{\nu}_p & \nu_{p,2} - \bar{\nu}_p & \cdots & \nu_{p,N} - \bar{\nu}_p \end{bmatrix}$$

Dans le cadre de modélisation d'écoulement en réservoir, il est possible de négliger les incertitudes de la simulation (les simulations d'écoulement sont des fonctions stables par rapport aux propriétés des roches) (Emerick & Reynolds, 2013).

Les données mesurées souffrent d'erreurs de mesures. Pour en tenir compte, le vecteur de données observées  $D_{obs}$  est obtenu en supposant des erreurs de mesures gaussiennes autour de la valeur mesurée sur le terrain. En ajoutant de petites erreurs sur les mesures, l'ensemble tendra à ajuster les observations sur les mesures plus ou moins une petite erreur. Cette erreur peut également être vue comme un critère de tolérance entre les observations et les mesures. Ainsi Emerick & Reynolds (2013) définit ce vecteur comme tel :

$$D_{obs} \sim \mathcal{N}(D_{mesurés}, R). \quad (2.6)$$

La matrice  $R$  est construite à partir de la covariance sur l'erreur des données mesurées.

La matrice de gain de Kalman notée  $\bar{K}$  permet de mesurer l'apport d'information provenant du modèle ou des données.

$$\bar{K} = AY^T(YY^T + R)^{-1} \quad (2.7)$$

$YY^T$  est la covariance entre les anomalies d'observations depuis le modèle (charge hydraulique simulée),  $R$  est la matrice de covariance traduisant l'incertitude sur les mesures effectuées sur site. Enfin  $AY^T$  est la matrice de covariance reliant les anomalies de paramètres d'états du modèle et les anomalies d'observations. La dimension de  $\bar{K}$  est de  $m \times p$ . Cette dimension traduit l'apport d'informations depuis la dimension réduite des mesures  $p$  vers le nombre de paramètres d'états  $m$ .

L'étape finale des méthodes d'assimilation de données est la mise à jour des modèles de paramètres d'états.  $E^f$  est l'ensemble des modèles de paramètres initiaux ou issus de l'étape précédente. Cet ensemble de modèles va être mis à jour en observant l'innovation, c'est-à-dire la différence entre les données observées  $D_{obs}$  et les valeurs simulées  $y$ . Ces valeurs simulées pourront être par exemple des charges hydrauliques ou bien des concentrations calculées dans les  $N$  modèles. Dans la littérature, les valeurs simulées  $y$  sont obtenues grâce à une fonction complexe  $h()$  qui dépend des paramètres d'états contenus dans l'ensemble  $E^f$ .  $h(E^f)$  est en pratique obtenue par un logiciel

de simulation de processus physique, comme FEFLOW dans le cas de problèmes de simulations d'écoulement hydraulique. La fonction  $h()$  représente le processus complet effectué par un simulateur d'écoulement pour prédire l'état du système dans le futur depuis les paramètres d'états connus actuellement.

$$E^a = E^f + \bar{K} \{D_{obs} - h(E^f)\} = E^f + \bar{K} \{D_{obs} - y\} \quad (2.8)$$

L'innovation va avoir plus ou moins d'impacts sur certaines régions et modèles en fonction de la valeur du gain de Kalman noté  $\bar{K}$ . Lorsque la covariance est forte entre mesures et paramètres, l'innovation aura un impact plus élevé sur les paramètres d'états de l'ensemble.

Les nouvelles valeurs de paramètres sont stockées dans l'ensemble  $E^a$ , pour ensemble analysé. À l'étape suivante, au moment d'incorporer de nouvelles mesures, l'ensemble analysé deviendra l'ensemble prédit  $E^f$  et une nouvelle itération d'assimilation pourra avoir lieu. La Figure 2.6 illustre ce cycle d'assimilation d'après les travaux de Raanes (2015b).

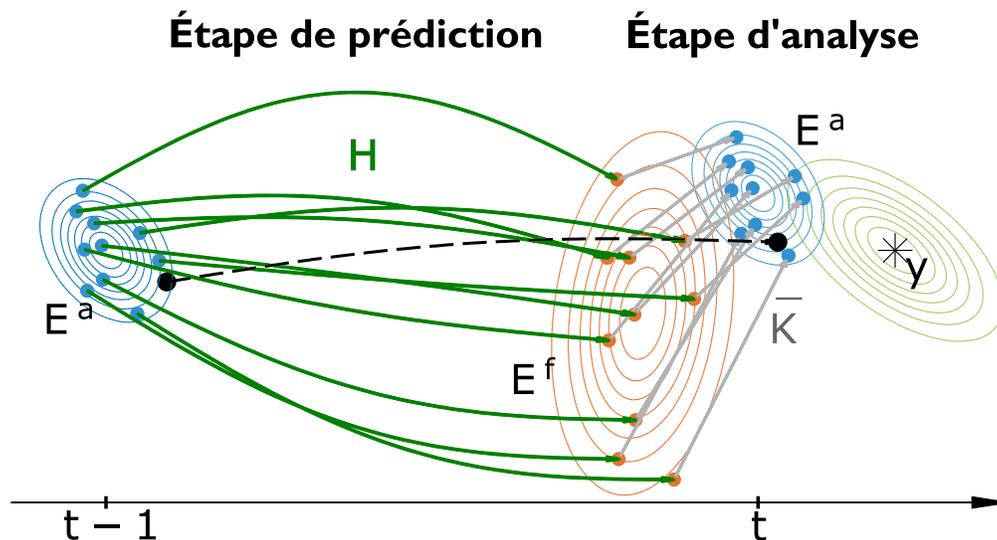


Figure 2.6 – Schéma d'un cycle complet d'assimilation de données d'un filtre de Kalman d'ensemble. En bleu le posterior provenant de l'étape d'analyse 2.8, en vert foncé le simulateur équivalent à  $H$ , en orange le prior provenant des simulations d'écoulement et en vert clair la vraisemblance des données mesurées. Le modèle réel est assumé étant présent dans la même distribution que l'ensemble. Modifié depuis (Raanes, 2015b).

L'explication réalisée dans cette partie repose sur le principe de l'EnKF de (Evensen, 2009a) avec les notations mathématiques issues de (Raanes, 2015b). La Figure 2.6 indique une assimilation à chaque pas de temps de la simulation numérique. Ceci est cohérent avec les méthodes d'EnKF.

Dans le cas des ES, comme indiqué sur la Figure 1.3, l'assimilation ne se fait plus à chaque pas de temps de la simulation, mais se fait une fois à la fin de la simulation complète. Avec les ES, la différence se fera sur la matrice d'observation  $y$ . Celle-ci contiendra toutes les données temporelles concaténées.

# Chapitre 3

## Méthodologie

La méthode d'ensemble retenue est celle de l'assimilation multiple ES-MDA 4x de Emerick & Reynolds (2013). Il s'agit de 4 assimilations ES à la suite, avec une simulation complète de l'écoulement hydraulique pour tous les modèles de l'ensemble à chaque itération. Une nouvelle méthode d'analyse locale et de localisation va être développée à partir de cette méthode d'assimilation pour gérer le faible nombre de données disponibles pour mettre à jour les paramètres d'états. Ce faible rapport d'information est illustré par le ratio d'information défini ci-dessous.

### 3.1 Ratio d'information

Pour évaluer la quantité d'information nécessaire pour utiliser les méthodes d'assimilation d'ensemble, un ratio d'information a été défini. Ce ratio est le rapport entre le nombre de paramètres d'états et le nombre d'observations. Il s'agit donc d'une sorte de rapport spatial ou volumétrique pour voir combien d'éléments sont mis à jour par observation dans la littérature. Cette information semble cruciale dans le cas d'étude environnementale où le nombre de puits d'observation est souvent très réduit. Les ratios calculés dans le Tableau 3.1 sont plus indicatifs qu'exhaustifs, car certains éléments de modèles peuvent contenir plusieurs paramètres d'états en fonction des auteurs. L'idée générale est d'évaluer la quantité d'observations nécessaire pour mettre à jour l'ensemble de modèles. Le ratio maximum observé dans la littérature était de 5500. Le but de ce manuscrit sera donc d'étudier la viabilité de la méthode pour des applications environnementales avec un ratio d'in-

**Tableau 3.1 – Ratio d’information entre le nombre de paramètres d’états et le nombre d’observations dans la littérature.**

Auteur(s)	Géométrie	Paramètres	Observations	Ratio
(Dong <i>et al.</i> , 2006)	2D	256	5	51
(Zhou <i>et al.</i> , 2011)	3D	8000	20-80	100-400
(Emerick & Reynolds, 2012)	2D	3600	7	514
(Emerick & Reynolds, 2013)	3D	4840	40	121
(?)	2D	3297	885	4
(Chen & Oliver, 2013)	3D	165000	30	5500
(Zovi <i>et al.</i> , 2017)	3D	11939	-	50-750
(Li <i>et al.</i> , 2018)	2D	8000	110	73
(Wang <i>et al.</i> , 2019)	2D	1800	24	90
Cas synthétique	3D	68940	6	11490
Cas réel	3D	1094771	24	45616

formation plus élevé, 11490 pour le cas synthétique de ce doctorat puis 45616 pour le cas réel. Le cas synthétique aura 68490 éléments FEFLOW avec 3 points d’observation mesurant chacun la charge hydraulique et la concentration en sel. Le cas réel présenté à la fin de ce manuscrit aura lui 1094771 éléments FEFLOW pour 24 points d’observations suivant la charge hydraulique uniquement.

## 3.2 Schéma général

Le schéma utilisé repose sur deux cycles distincts. Dans un premier temps une modélisation géostatistique et hydrogéologique classique, puis dans un second temps une partie assimilation. Un schéma est présentée à la Figure 3.1. Lors de ce doctorat, les nouveaux développements se sont concentrés sur l’analyse locale lors du processus d’assimilation. Ce focus est nécessaire pour gérer le fort ratio d’information attendu dans le cadre d’étude hydrogéologique. Chacune des étapes du schéma général sera rapidement présentée dans les paragraphes suivants.

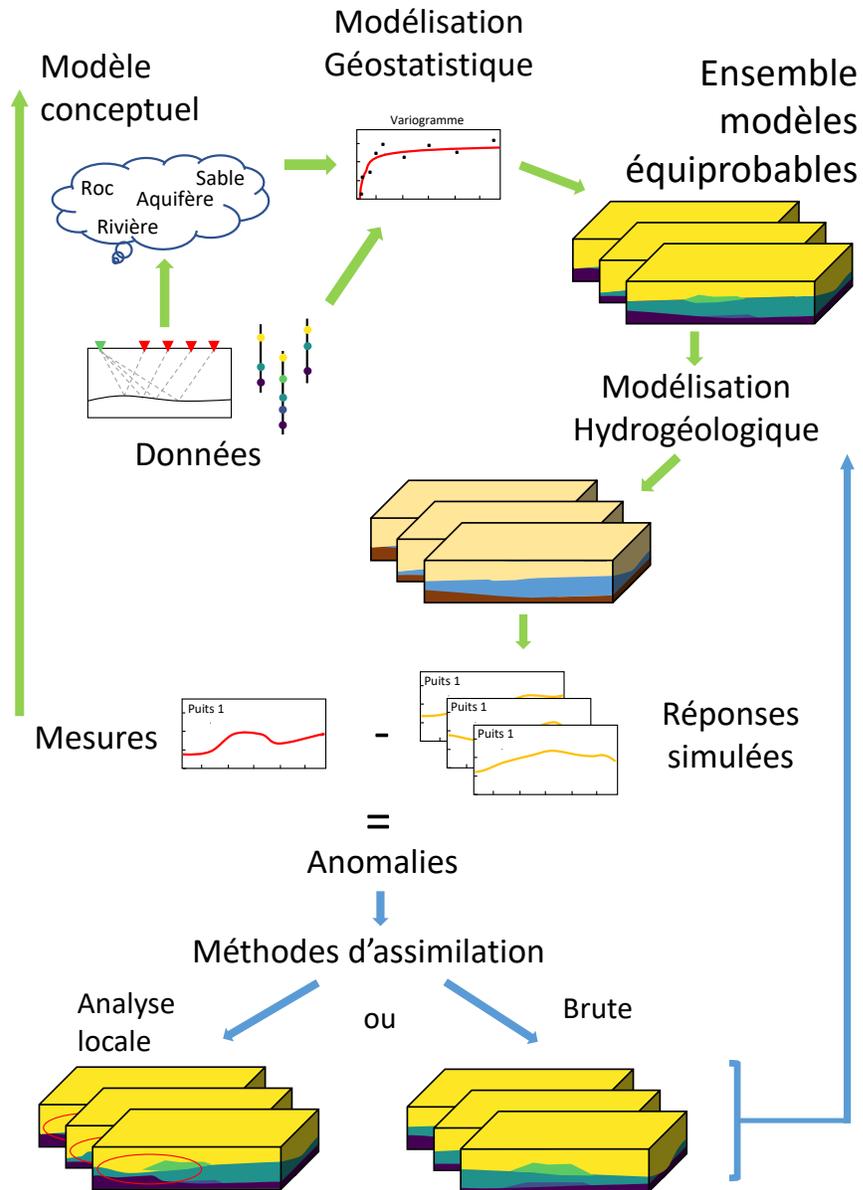


Figure 3.1 – Le schéma peut être décomposé en deux cycles. Le premier cycle avec les flèches vertes indique les étapes classiques de modélisation hydrogéologique. Les flèches bleues traduisent le cycle d'assimilation apportée par les méthodes d'ensembles. Les ellipses rouges indiquent les zones mises à jour dans les modèles de paramètres d'états. Ces ellipses sont définis grâce à l'analyse locale.

### 3.2.1 Modélisation conceptuelle

Dans un premier temps, un modèle conceptuel doit être établi. Il doit s'agir du premier raisonnement avant de penser à comment construire le modèle. Le modèle conceptuel doit intégrer les données disponibles (perméabilités, porosité, géophysique, etc.), les mesures (charges hydrau-

liques, concentrations, etc.) et toutes les connaissances disponibles (géologie, hydrologie, études précédentes, etc.) à travers des hypothèses. Ces hypothèses sont utiles pour discriminer les modèles qui correspondent aux mesures de terrain ou non.

### 3.2.2 Modélisation géostatistique

Une fois les données regroupées, un modèle géostatistique est bâti. Il est bâti grâce à un géomodeleur comme SGEMS ou GOCAD. L'analyse exploratoire détaillée des caractéristiques statistiques des différents jeux de données (perméabilité, porosité, etc.), et une analyse des corrélations spatiales est conduite à l'aide d'un géomodeleur ou bien de code python externe. Une fois ces informations connues, des méthodes de simulations géostatistiques sont utilisées pour produire un ensemble de modèles équiprobables. Chacun des modèles produits honore les données imposées, les variations apparaissent progressivement en s'éloignant des données. Dans le cas synthétique, des simulations gaussiennes séquentielles (SGS) ont été réalisées à l'aide du logiciel SGEMS. Dans le cas réel, des bandes tournantes séquentielles (TBS) au nombre de 2400 ont été utilisées avec le logiciel ISATIS.

### 3.2.3 Modélisation hydrogéologique

Les modèles hydrogéologiques sont bâtis en tenant compte des conditions aux limites disponibles. Dans le cas synthétique, elles peuvent être arbitraires. Dans le cas réel, ces limites sont normalement basées sur des limites naturelles où cela est possible. Il s'agira de rivières, bassin versant, limites géologiques, etc. Les valeurs de recharges et de pompages régionaux sont imposées pour correspondre aux valeurs publiées et suivies. Ces modèles sont construits en intégrant les données issues des géomodeleurs. Une attention particulière est portée à l'intégrité du transfert des propriétés depuis les géomodeleurs vers le logiciel d'écoulement des eaux souterraines FEFLOW. Dans le cas du modèle synthétique, les charges hydrauliques et concentrations en sel sont suivies dans un modèle d'aquifère côtier simplifié. Le cas réel cherche à reproduire les charges hydrauliques sur un modèle multi échelle complexe.

Ces étapes sont nécessaires pour obtenir les données d'entrées pour le développement de ce doctorat, mais les méthodes développées peuvent s'appliquer à n'importe quel problème d'assimilation de données.

### 3.2.4 Méthodes d'assimilation

À partir des observations obtenues lors des étapes précédentes, les anomalies sont calculées en calculant les différences entre mesures et simulations dans les modèles de l'ensemble. Ces anomalies sont les données d'entrées nécessaires au processus d'assimilation décrit en détail dans le chapitre Méthodes d'ensemble.

La mise à jour des modèles de paramètres peut se faire de plusieurs manières. Lors de ce doctorat, la méthode de référence issue de la littérature est la méthode ES-MDA 4x détaillée par Emerick & Reynolds (2013), sans schéma d'analyse. Il s'agit de la méthode brute sur la Figure 3.1. Tous les éléments du modèle sont mis à jour, indépendamment de leurs distances aux points d'observation.

La nouvelle méthode développée repose sur un schéma d'analyse local afin de mettre uniquement à jour les paramètres proches des points d'observations. Une explication détaillée de cette nouvelle méthode est donnée dans les chapitres qui suivent.

## 3.3 Nouvelle méthode de Localisation et d'Analyse locale

### 3.3.1 Localisation

Les méthodes d'assimilation d'ensemble souffrent d'un problème conceptuel lorsque les données assimilées proviennent de différentes sources et sont de différents ordres de grandeur. L'idée est de trouver un moyen d'assimiler des données différentes en utilisant par exemple, des charges hydrauliques et des concentrations en sel. Pour expliquer le raisonnement et la solution, un exemple théorique va suivre. Cet exemple est constitué de deux puits, sur lesquels des charges hydrauliques et des concentrations sont suivies. Lors du suivie, un puits de pompage est mis en route avant d'être arrêté pour permettre un retour à l'état initial. La Figure 3.2 reprend les valeurs qui seraient mesurées sur le terrain et les valeurs simulées qui seraient issues de modèles numériques. Sans technique d'assimilation particulière ou de normalisation, de l'information sera perdu. La Figure 3.2 indique ce problème avec une échelle verticale volontairement identique pour différent types de données.

Dans cet exemple, les concentrations en sel varient de 0 à 700 mg/l tandis que les charges hydrauliques restent très proches de 20 m. Les valeurs mesurées et simulées sont très proches pour

les charges hydrauliques et les concentrations sur le puits 1. Les concentrations au puits 2 présentent une grande disparité entre les valeurs simulées et mesurées.

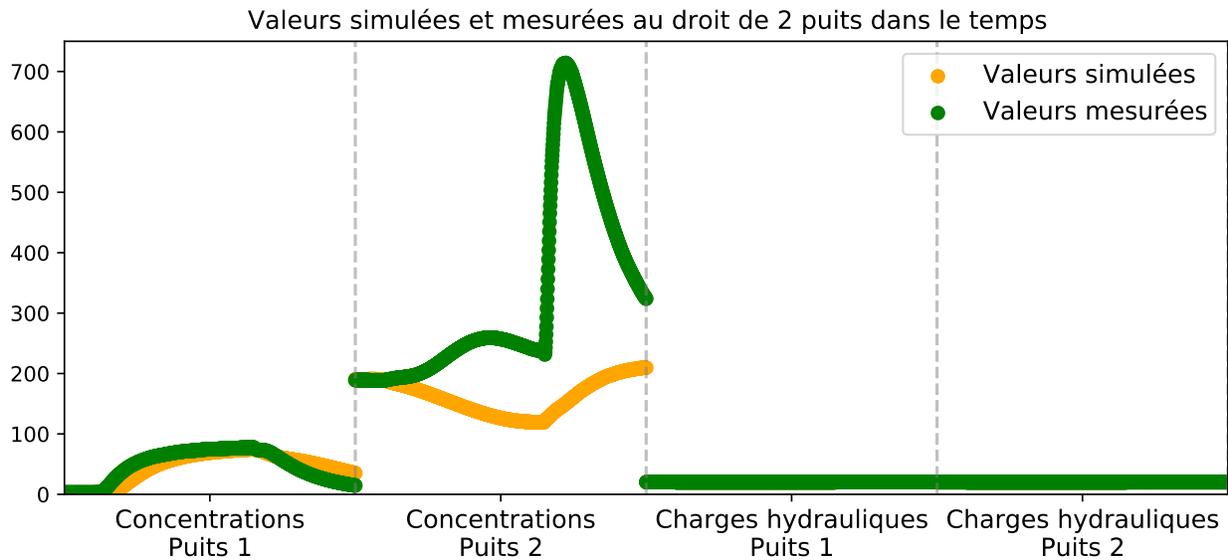


Figure 3.2 – Valeurs simulées dans un modèle en orange et valeurs mesurées en vert. Deux puits sont suivis pour deux types d'informations, des concentrations et des charges hydrauliques.

Les anomalies sont définies comme la différence entre les données mesurées en verte et les données simulées en orange sur la Figure 3.2. L'amplitude de ces anomalies varie fortement en fonction des types de données, et des modèles de l'ensemble. Dans l'exemple de la Figure 3.3, les anomalies sur les charges hydrauliques sont proches de 0.1 m, les anomalies sur les concentrations en sel varient de 0 g/l à 600 mg/l. Ces différences d'échelles sont problématiques pour une assimilation de toutes les données en une seule fois. L'information provenant des anomalies de concentration écraserait l'information apportée par les anomalies de charges hydrauliques, la mise à jour des paramètres du modèle se ferait quasi exclusivement en fonction de la concentration. Cette différence d'amplitude est visible sur la Figure 3.3.

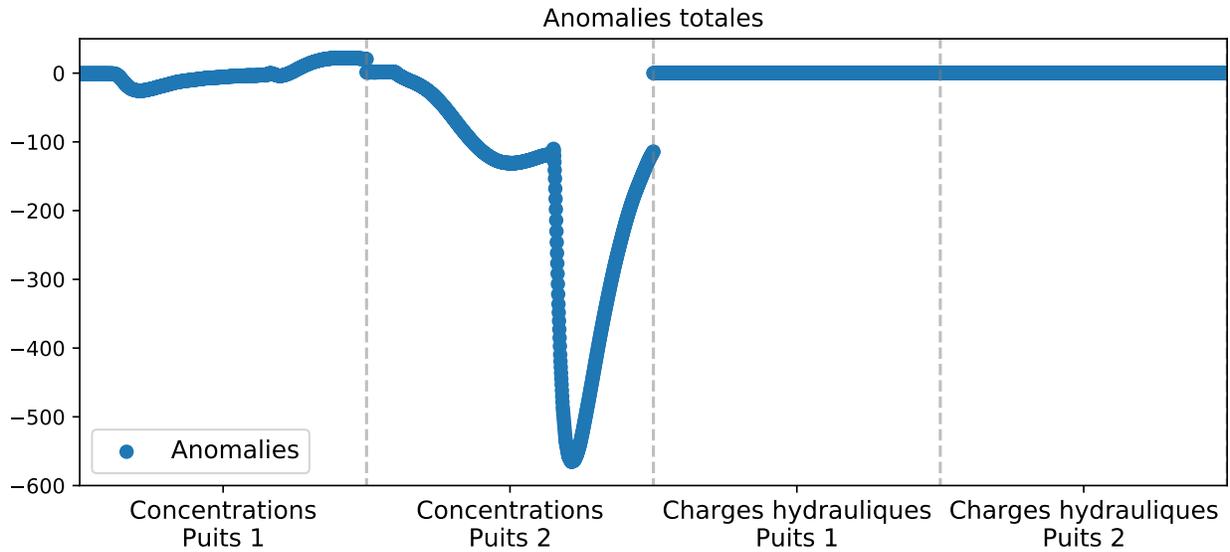


Figure 3.3 – Anomalies résultantes de la différence entre données simulées et mesurées.

Ces courbes représenteraient les résultats d'un modèle pris au hasard dans un ensemble de modèles. Lorsque les anomalies sont séparées par types de données, l'amplitude relative des anomalies est plus faible, permettant une meilleure définition. La Figure 3.4 donne les anomalies de concentration uniquement. La Figure 3.5 décrit les anomalies de charges hydrauliques uniquement. Hors, dans le cas des concentrations, même s'il s'agit d'un même type de donnée, les anomalies présentent une grande différence d'amplitude en fonction des puits.

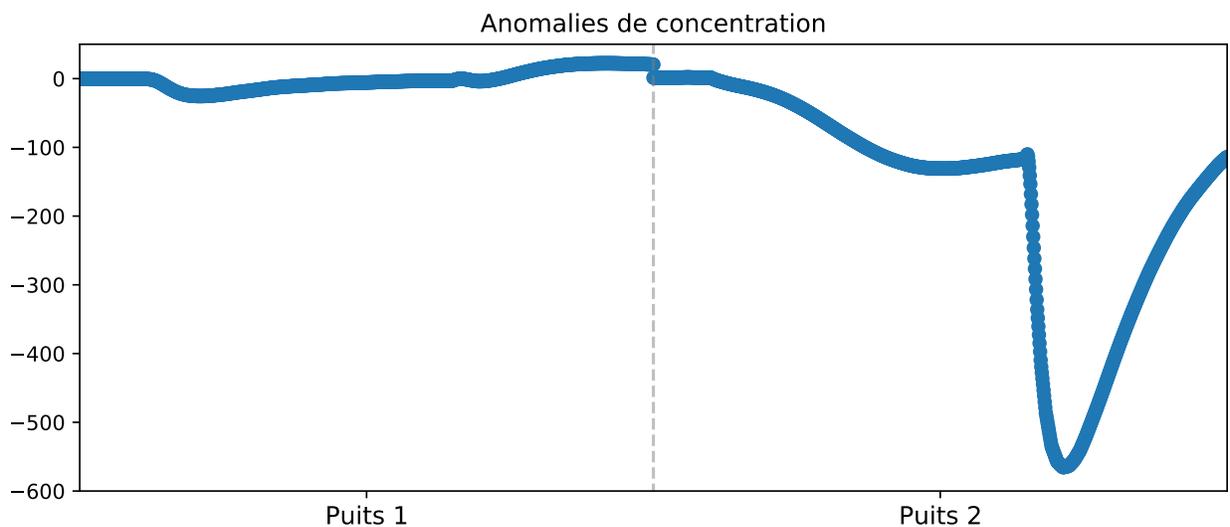
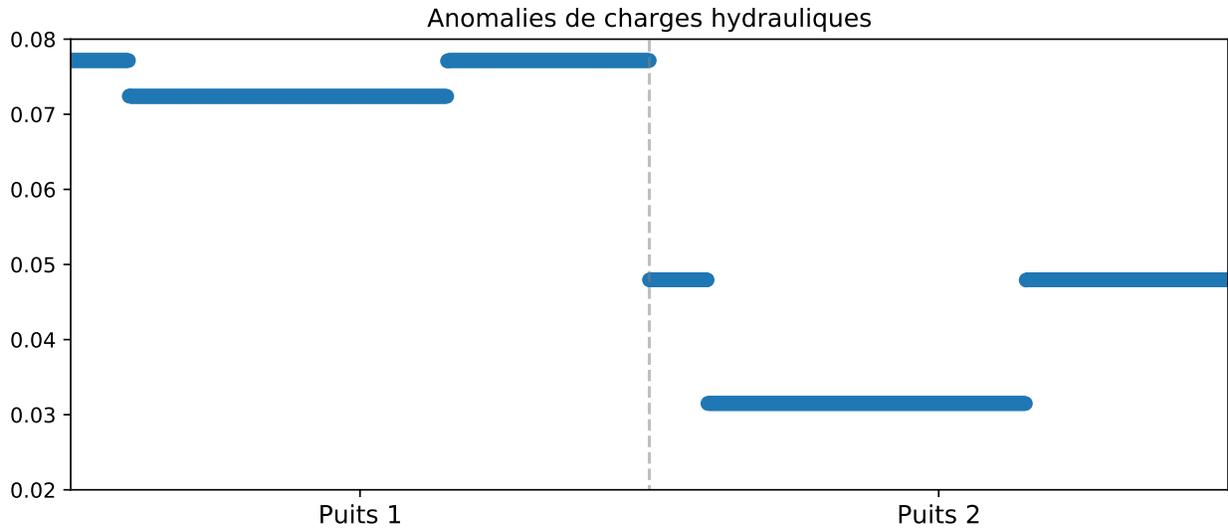


Figure 3.4 – Anomalies de concentration.



**Figure 3.5 – Anomalies de charges hydrauliques.**

Les anomalies de charges hydrauliques de la Figure 3.5 sont plutôt stables dans le temps, s'établissant immédiatement à une nouvelle valeur après application d'un débit de pompage. Les anomalies de concentrations sont dynamiques, et répondent progressivement aux modifications du débit de pompage. Ramener ces deux types de données à une même moyenne écraserait tout de même de l'information. Une assimilation de données, type de donnée par type de donnée, devrait permettre une assimilation plus précise que d'assimiler toutes les données en une seule fois.

Il est possible d'aller encore plus loin, en séparant également chaque puits pour avoir une anomalie relative la plus précise possible. L'anomalie de concentration du puits 1 de la Figure 3.6 était écrasée par celle du puits 2. La Figure 3.7 présente une faible amélioration de la définition de l'anomalie des charges hydrauliques.

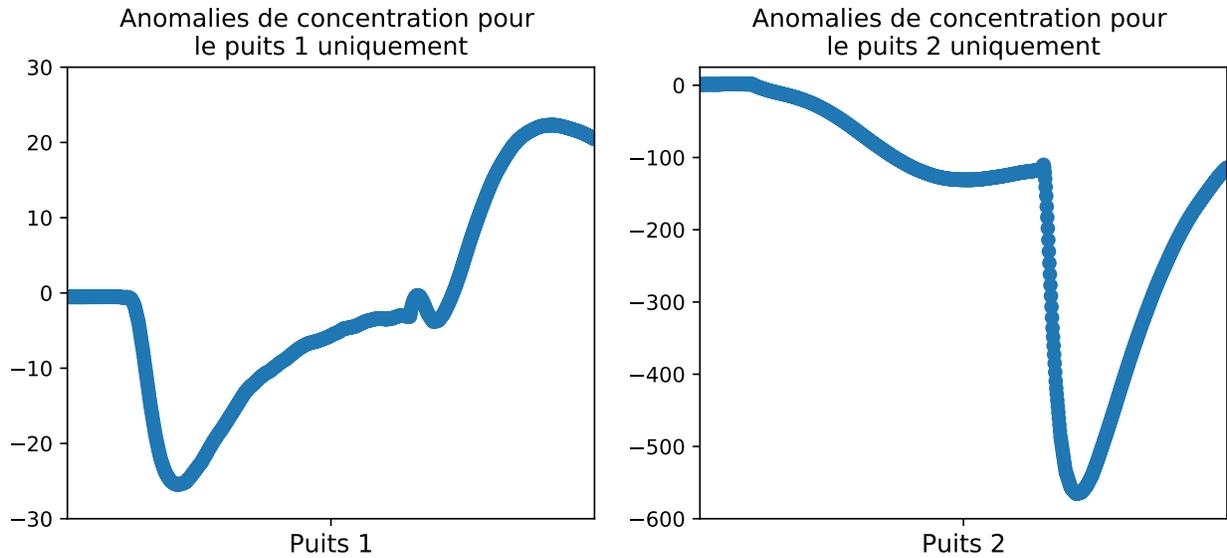


Figure 3.6 – Anomalies de concentrations sur les puits 1 et 2.

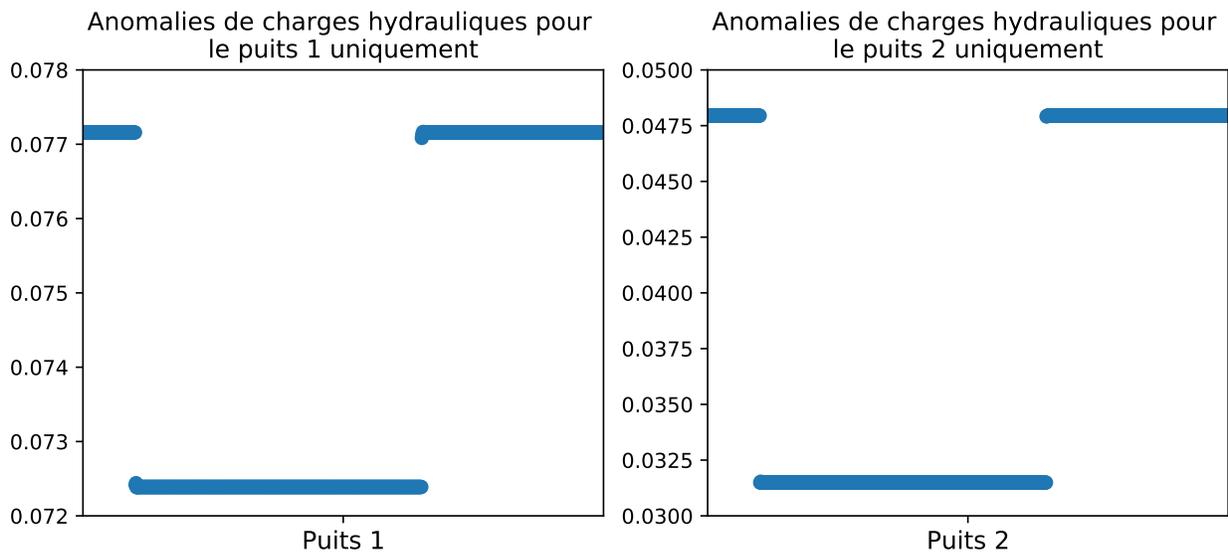


Figure 3.7 – Anomalies de charges hydrauliques sur les puits 1 et 2.

En séparant chaque type de données et chaque puits, les anomalies sont définies avec une échelle relative la plus précise possible. Une anomalie importante sur un type de donnée ne devrait pas influencer sur la quantité d'information exploitable sur les autres observations. En séparant les assimilations pour chaque puits et chaque type de données, l'objectif est de mettre à jour l'ensemble de paramètres d'état en sacrifiant le minimum d'information provenant des anomalies d'observa-

tions. Il s'agit donc d'une sorte de localisation. Au lieu de réduire numériquement les matrices de covariances comme classiquement dans la littérature, il est préférable de la faire grâce à une approche physique basée sur la distance. Cette dernière semble plus logique en se basant sur chaque observation individuellement.

### **3.3.2 Schémas de Localisation et d'Analyse locale**

Les méthodes de localisation sont majoritairement ajoutées pour atténuer les fausses corrélations dans les matrices de covariance et stabiliser numériquement l'inversion. Le choix du schéma d'atténuation peut être compliqué si ce dernier n'est pas basé sur une distance physiquement mesurable. Le moyen de coupler l'analyse locale et la localisation avec une distance perceptible dans l'espace serait donc de définir un voisinage d'observation variable autour de chaque point d'observation.

Plusieurs schémas d'assimilation ont été définis pour tester différentes hypothèses.

#### **Assimilation brute**

L'assimilation brute correspond à l'assimilation de l'ensemble de données disponibles en une étape. Tous les éléments sont mis à jour en une assimilation englobant toute l'information. Ceci correspondrait à l'assimilation de toutes les anomalies présentes sur la Figure 3.3.

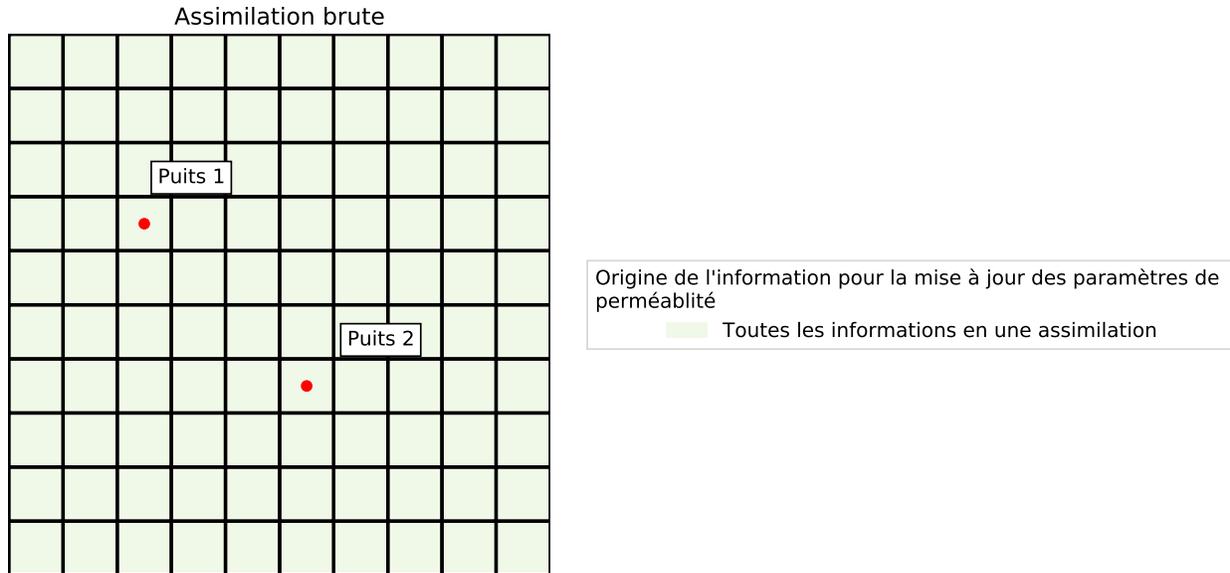
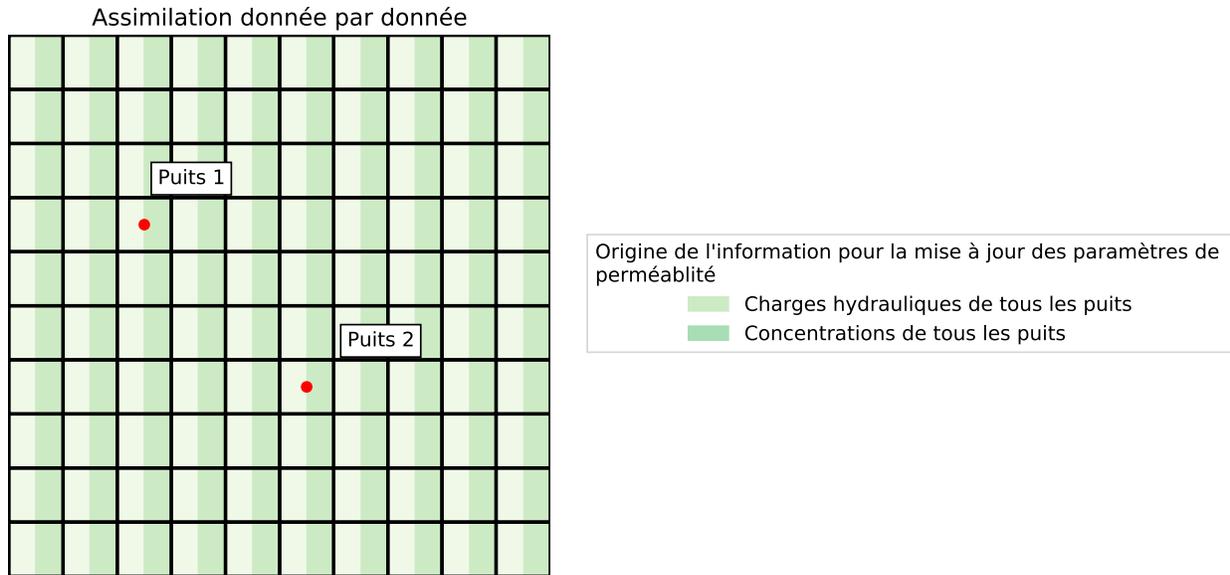


Figure 3.8 – Schémas d’assimilation de données où toutes les anomalies sont assimilées en une seule fois, donnant une mise à jour unique.

La Figure 3.8 schématise cette mise à jour pour tous les paramètres de perméabilités du modèle de l'exemple. Tous les paramètres sont mis à jour à l'issue d'une assimilation.

### Assimilation par type de données

L'assimilation par type de données revient à effectuer une assimilation de tous les puits par type de donnée. Tous les éléments sont mis à jour avec une combinaison d'information provenant des anomalies de charges hydrauliques et des anomalies de concentration assimilées séparément. Cela reviendrait à assimiler les anomalies des Figures 3.4 et 3.5 séparément.

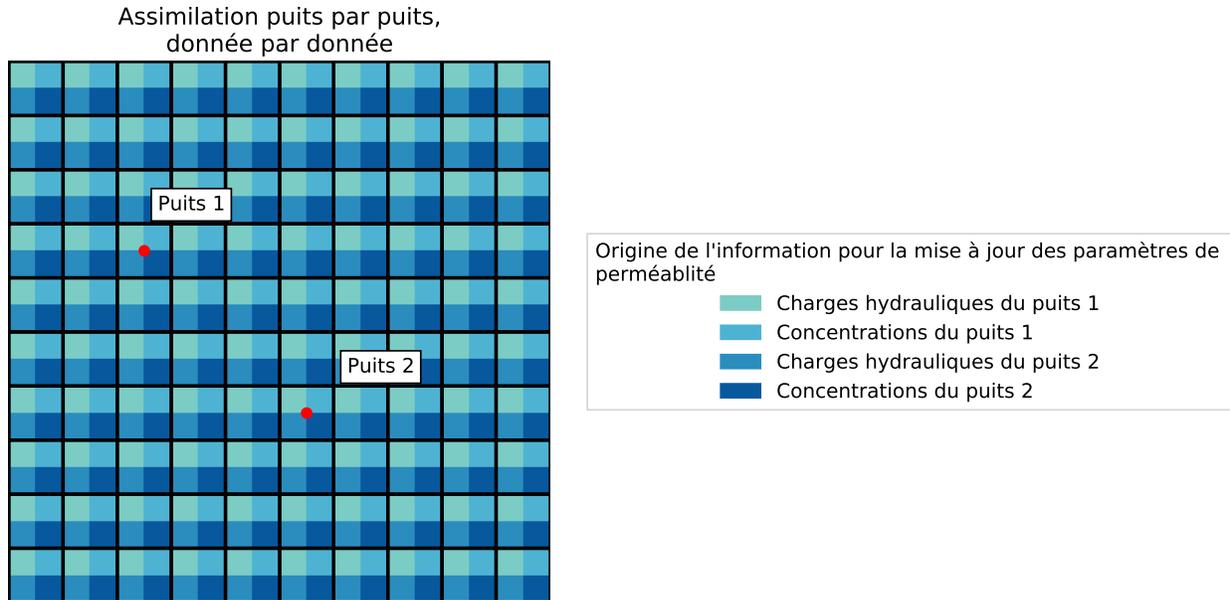


**Figure 3.9 – Schémas d’assimilation de données où les charges hydrauliques et la concentration sont assimilées séparément, donnant deux mises à jour pour chaque cellule.**

La Figure 3.9 schématise cette mise à jour pour tous les paramètres de perméabilités du modèle de l'exemple. Tous les paramètres sont mis à jour à l'issue de deux assimilations. Ces derniers seront mis à jour à moitié par l'information provenant de anomalies de charges hydrauliques et à moitié de l'information provenant des anomalies de concentrations.

### **Assimilation par type de données et par puits**

L'assimilation par type de données et par puits revient à utiliser chaque anomalie individuellement. Chaque élément du modèle est mis à jour par 4 assimilations, une par type de données et par puits. Cette méthode correspond à l'assimilation individuelle des anomalies des Figures 3.6 et 3.7.

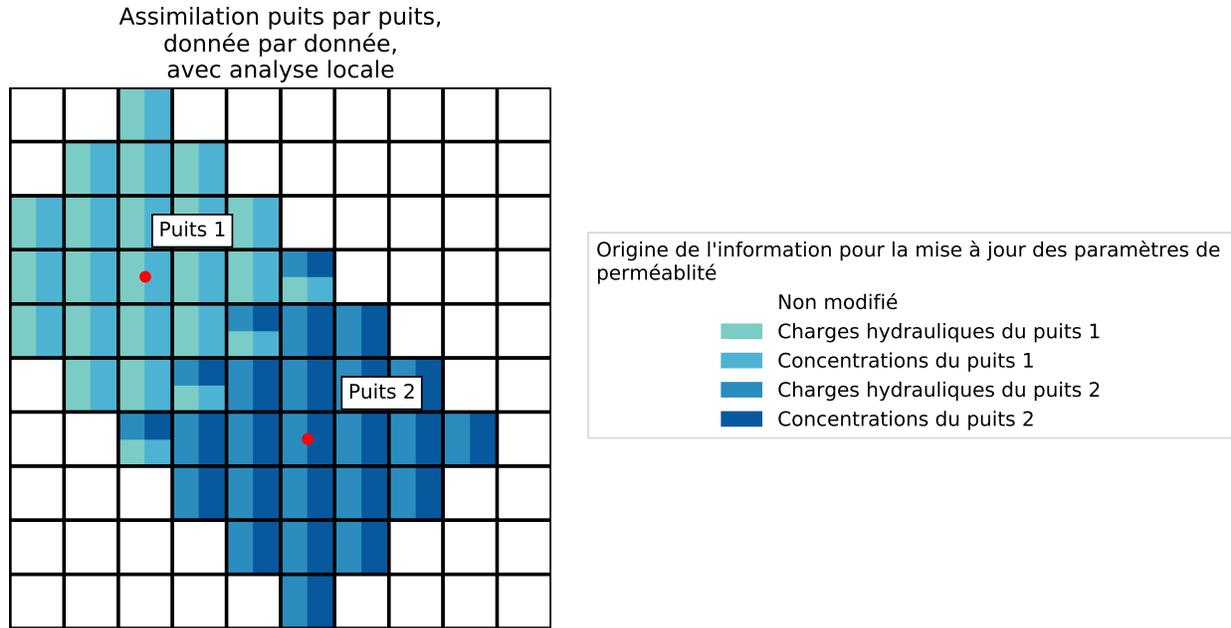


**Figure 3.10 – Schémas d’assimilation de données où toutes informations sont assimilées successivement, une pour chaque puits et type de donnée. Chaque cellule du modèle est mise à jour par la contribution de ces 4 assimilations.**

La Figure 3.10 schématise cette mise à jour pour tous les paramètres de perméabilités du modèle de l'exemple. Tous les paramètres sont mis à jour à l'issue de quatre assimilations. Ils seront mis à jour à un quart par l'information provenant de anomalies de charges hydrauliques du puits 1, à un quart du puits 2, à un quart des anomalies des concentrations du puits 1 et à un quart du puits 2.

### Assimilation par type de données et par puits avec analyse locale

La méthode combine l'analyse locale et la localisation. Pour se faire, chaque type de donnée et chaque puits est assimilé individuellement. Chacune de ces assimilations va uniquement mettre à jour les paramètres dans un voisinage défini autour de chaque puits d'observation. Cette méthode requiert également 4 assimilations. Lorsque plusieurs points d'observation sont assez proches l'un de l'autre, l'information provenant des plusieurs de ces points sera utilisée.



**Figure 3.11 – Schémas d’assimilation de données où toutes informations sont assimilées successivement, une pour chaque puits et type de donnée. Uniquement les cellules du voisinage des points d’observations sont mises à jour.**

La Figure 3.11 schématise la mise à jour pour les paramètres de perméabilités du modèle de l'exemple présent dans un voisinage proche des points d'observations. Les paramètres sont mis à jour à l'issue de quatre assimilations. Ils seront mis à jour à un quart par l'information provenant de anomalies de charges hydrauliques du puits 1, à un quart du puits 2, à un quart des anomalies des concentrations du puits 1 et à un quart du puits 2.

Dans le cas de l'analyse locale centrée autour de chaque point d'observation, le modèle est uniquement mis à jour dans un voisinage défini par l'utilisateur et non plus globalement. Il est maintenant possible de mettre à l'échelle l'amplitude de la mise à jour en fonction de la distance de l'observation. Plus l'élément du modèle est éloigné du point d'observation, moins la mise à jour devrait être importante, traduisant une diminution dans le lien physique. Il serait également possible de définir un poids additionnel par puits, afin de refléter l'a priori sur la qualité des données ou le type de données. Des mesures géophysiques pourraient avoir un poids plus faible qu'une mesure directe du sous-sol par exemple.

La contribution à la mise à jour de chaque paramètre d'états par chaque puits individuel et par chaque type de données est définie comme suit :

$$\text{màj paramètre final} = \frac{1}{n_{\text{données}}} \sum_{i=1}^{n_{\text{données}}} \left( \frac{1}{n_{\text{puits}}} \sum_{j=1}^{n_{\text{puits}}} (\text{poids puits } [j] * \text{màj analyse locale } [i, j]) \right); \quad (3.1)$$

$$\sum_{j=1}^{n_{\text{puits}}} (\text{poids puits } [j]) = 1. \quad (3.2)$$

Dans les cas synthétique et réel présentés plus bas dans ce document, le poids individuel de chaque puits était uniforme. Aucune importance supplémentaire n'a été assignée à un puits par rapport à un autre. Mais cette possibilité existe si la crédence est variable en fonction des points observations.

Dans le cas d'une assimilation globale de toutes les cellules du modèle, le vecteur de mise à jour aura la même taille que celui de l'ensemble des paramètres. Dans le cas des voisinages variables, chaque mise à jour résultant d'une assimilation est stockée dans un dictionnaire. Ce dictionnaire contient l'index de l'élément dans le modèle global, la mise à jour et son poids lié à la distance du point d'observation assimilé. Le poids calculé basé sur la distance du point d'observation est décrit dans le paragraphe suivant. Pour accélérer le processus d'assimilation, tous les indices d'éléments qui seront mis à jour autour des points d'observations sont précalculés. Ces indices sont uniques pour tout le processus d'assimilation multiple, et varient uniquement en fonction de la définition du voisinage par l'utilisateur. De cette manière, le processus d'assimilation mettra uniquement à jour les éléments présents dans ses voisinages.

### 3.3.3 Implémentation de l'ellipsoïde d'Analyse Locale

Pour la méthode développée dans ce document, les données seront assimilées observation par observation, afin de minimiser les fausses corrélations. L'amplitude de la mise à jour diminuera pour atteindre 0 sur les bordures du voisinage autour de chaque point d'observation. Chaque sous-domaine sera centré sur un point d'observation. Par rapport aux méthodes de mises à jour locales citées plus tôt, le nombre d'assimilations sera limité aux nombres de points d'observations avec la nouvelle méthode. Ce nombre étant très petit par rapport aux nombres de paramètres, il ne devrait pas y avoir de limites au niveau du nombre d'assimilations à effectuer. Finalement, avec la

nouvelle méthode, il sera possible de tenir compte d'anomalies ayant des magnitudes très différentes. Ceci n'était pas possible en mettant à jour un paramètre en assimilant plusieurs anomalies à la fois. L'assimilation brute consiste à assimiler toutes les données provenant des observations en une assimilation. Cette approche est la plus simple, mais elle présente une limitation. La régularisation de la matrice de covariance peut écraser des informations provenant de différents puits. Comme vu précédemment, différentes amplitudes sont possibles en fonction des données utilisées. Pour atténuer ce risque, il a été décidé d'assimiler chaque point d'observation individuellement. Avec cette méthode, l'amplitude des observations et des anomalies est relativement consistante au niveau d'un même puits. À la fin du processus d'assimilation, la moyenne de la mise à jour provenant de l'ensemble des points d'observations, comme définis dans l'Équation 3.1, est prise.

De plus, l'idée serait de diminuer l'importance d'une observation lorsque cette dernière se situe loin des paramètres à mettre à jour. Pour apporter cette possibilité à la méthode, l'inspiration est venue du concept de portée et de variance défini par (Goovaerts, 1997) dans le cadre des géostatistiques. Un ellipsoïde de poids a été créé autour d'un point d'observation. Cet ellipsoïde repose sur un azimut, un pendage et 3 portées définissant les dimensions des axes principaux de l'ellipsoïde. Comme dans le cas des variogrammes, la distribution des poids suivra des formes linéaire, gaussienne ou exponentielle. Le poids d'observation est défini à 1 au centre de l'ellipsoïde, puis il décroît pour atteindre 0 aux bordures de l'ellipsoïde. Il est possible de spécifier une valeur non nulle pour tout de même mettre à jour les paramètres en dehors de l'ellipsoïde. Ceci est analogue à l'effet pépite présent dans les variogrammes définis par (Chilès & Delfiner, 1999).

Chaque point d'observation a son propre ellipsoïde de poids défini autour de lui même. Il est possible d'utiliser de l'information à priori sur la qualité ou l'incertitude des données en diminuant le poids maximum dans l'ellipsoïde à moins que 1. L'information provenant de l'assimilation de ce point aura donc une importance plus faible par rapport aux autres points assimilés. Les voisinages sont donc définis par 3 portées le long des axes et 2 angles. Chaque portée d'ellipsoïde donne la distance dans le modèle où les paramètres seront mis à jour autour du point d'observation. Le plan horizontal est défini par deux axes et l'angle d'azimut. Cet angle est l'angle trigonométrique avec l'axe  $x$ . Le plan vertical est défini par l'angle de pendage avec le plan horizontal. Par analogie avec les géostatistiques et les variogrammes, une fonction de poids le long des 3 axes de l'ellipsoïde a été définie. L'équation d'une ellipse est rappelée ci-dessous :

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1. \quad (3.3)$$

Avec  $a, b, c$  les 3 axes de l'ellipsoïde alignés avec les axes  $x, y, z$  et le centre positionné en  $(0, 0, 0)$ .

La première étape est de transformer les coordonnées de chaque point du modèle dans l'ellipsoïde en coordonnées locales centrées en  $(0, 0, 0)$  et alignés sur les axes  $x, y, z$ . Il faut pour cela effectuer une rotation autour de l'axe  $y$  pour annuler l'effet du pendage  $\varphi$ . Puis une rotation autour de l'axe  $z$  pour annuler l'angle d'azimut  $\theta$ . À ce point l'ellipsoïde est aligné sur les axes locaux et centré en  $(0, 0, 0)$ . Il est donc possible d'utiliser la forme simplifiée de l'Équation 3.3. La matrice de coordonnées locales  $A$  est définie comme telle :

$$A = \begin{bmatrix} x - x_c \\ y - y_c \\ z - z_c \end{bmatrix}, \quad (3.4)$$

avec  $x, y, z$  les coordonnées des éléments du modèle à l'intérieur du voisinage et  $x_c, y_c, z_c$  les coordonnées du centre de l'ellipsoïde. L'étape suivante est donc la rotation autour de l'axe  $y$  pour annuler le pendage  $\varphi$  :

$$B = \begin{bmatrix} \cos(\varphi) & 0 & \sin(\varphi) \\ 0 & 1 & 0 \\ -\sin(\varphi) & 0 & \cos(\varphi) \end{bmatrix}, \quad (3.5)$$

$$C = B \times A^T. \quad (3.6)$$

La matrice  $C$  représente les coordonnées locales des points de l'ellipsoïde projeté sur le plan horizontal.

Après cette étape, les coordonnées locales sont encore pivotées autour de l'axe  $z$  par l'angle d'azimut  $\theta$ . La dernière rotation est effectuée pour finir l'alignement sur les axes locaux :

$$D = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) & 0 \\ \sin(-\varphi) & \cos(-\varphi) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3.7)$$

$$E = D \times C. \quad (3.8)$$

La matrice  $E$  représente les coordonnées locales des points de l'ellipsoïde projetées sur les axes locaux  $x, y, z$ .

La méthode choisie pour avoir une continuité dans le poids, avec 3 différentes portées, est d'utiliser une même fonction de poids et une combinaison linéaire de 3 portées différentes.

Depuis l'Équation 3.3, la somme des poids doit être égale à 1. Le maximum de l'ellipsoïde le long de chaque axe sera atteint aux portées maximales définies par l'utilisateur pour son ellipsoïde de voisinage. 3 scalaires compris entre  $[0, 1]$  ont été définis:

$$u = \left( \frac{E[0]}{\text{max range}} \right)^2, v = \left( \frac{E[1]}{\text{med range}} \right)^2, w = \left( \frac{E[2]}{\text{min range}} \right)^2. \quad (3.9)$$

3 modèles de variogrammes ont été définis depuis (Chilès & Delfiner, 1999) pour servir de fonctions de poids. 4 paramètres sont donnés à ces fonctions,  $n$  l'effet pépité,  $s$  le plateau,  $r$  la portée et  $h$  le pas étudié.

Modèle exponentiel :

$$\gamma(h) = (s - n)(1 - \exp(-3h/r)) + n \quad (3.10)$$

Modèle sphérique :

Si  $h < r$  :

$$\gamma(h) = (s - n) \left( \left( \frac{3h}{2r} - \frac{h^3}{2r^3} \right) \right) \quad (3.11)$$

Sinon :

$$\gamma(h) = s \quad (3.12)$$

Modèle gaussien :

$$\gamma(h) = (s - n) \left( 1 - \exp \left( -\frac{3h^2}{r^2} \right) \right) + n \quad (3.13)$$

Avec ces 3 modèles, il est possible de calculer un poids d'observation qui répond aux besoins pour chaque point dans l'ellipsoïde où  $u + v + w \leq 1$ .

$$\begin{aligned} \text{Poids d'observation} &= s - u \times \text{fonction}(s, n, \text{max range}, E[0]) \\ &\quad - v \times \text{fonction}(s, n, \text{med range}, E[1]) \\ &\quad - w \times \text{fonction}(s, n, \text{min range}, E[2]). \end{aligned} \quad (3.14)$$

Avec une **fonction** qui correspond à un des 3 modèles de variogramme décrits dans les Équations 3.10, 3.11, 3.12 et 3.13. Un ellipsoïde de poids d'observations est donné sur les Figures 3.12 et 3.13 pour une grille non structurée d'éléments d'un modèle FEFLOW. Pour alléger la lecture, seuls les éléments avec un poids strictement supérieur à 0 sont affichés.

Ellipse de poids autour d'un point d'observation centré en (50,50,10)

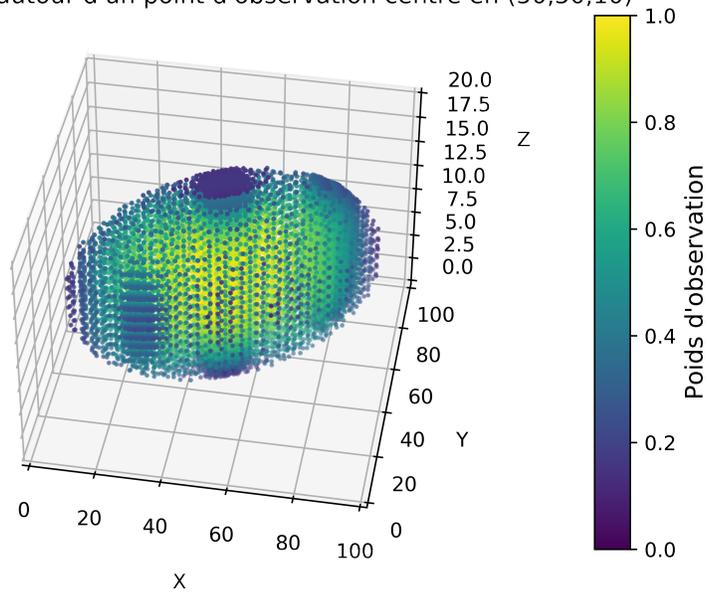


Figure 3.12 – Ellipsoïde de poids défini autour d'un point d'observation donné dans une grille d'éléments non structurés. Le centre de l'ellipsoïde est défini en (50, 50, 10). La portée maximale est mise à 50 m, la portée moyenne à 20 m et la verticale à 10 m. L'azimut de l'axe majeur est défini à un angle de  $75^\circ$  avec le nord et la fonction de poids choisie est une fonction gaussienne.

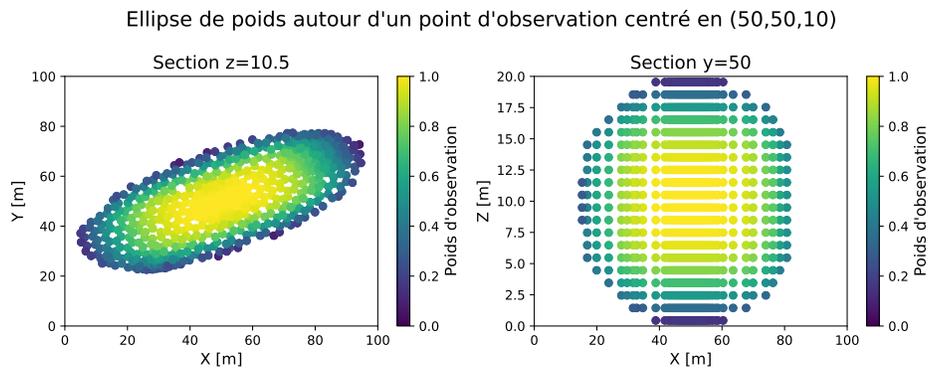


Figure 3.13 – Sections horizontale et verticale de l'ellipsoïde de poids défini autour d'un point d'observation donné dans une grille d'éléments non structurés. Le centre de l'ellipsoïde est défini en (50, 50, 10). La portée maximale est mise à 50 m, la portée moyenne à 20 m et la verticale à 10 m. L'azimut de l'axe majeur est défini à un angle de  $75^\circ$  avec le nord et la fonction de poids choisie est une fonction gaussienne.

L'efficacité de cette nouvelle méthode d'analyse locale sera étudiée dans les parties suivantes du manuscrit. L'assimilation de toutes les données en une seule assimilation sera comparée à une assimilation pour chaque type de données puis à une assimilation pour chaque puits et chaque type de données. Enfin, l'impact des différentes distances dans l'analyse locale sera évalué.

### 3.4 Code développé

L'ensemble du code développé lors de doctorat est disponible en open source sur Github.

[https://github.com/Thomas-Beraud/FEFLOW\\_MDA\\_MPMW.github.io](https://github.com/Thomas-Beraud/FEFLOW_MDA_MPMW.github.io)

Le code a été développé en s'articulant autour du logiciel d'écoulement FEFLOW. Cependant, les méthodes d'assimilation, les classes de gestion des données et les méthodes d'analyses locales sont entièrement utilisables pour n'importe quel simulateur de processus physique. Un premier essai en 2D était interfacé avec le logiciel Saltflow Molson & Frind (2022). Il suffira à l'utilisateur d'extraire les observations et les paramètres d'états envoyés au simulateur.

Un schéma symbolique de l'organisation du code développé est donné sur la Figure 3.14. Le fichier `Import_Data.py` est modulaire de par sa programmation objet pour permettre la gestion de différents simulateurs. Le fichier `Localization_Methods.py` devra être modifié en fonction du format de coordonnées des paramètres d'état dans le cas d'un simulateur différent.

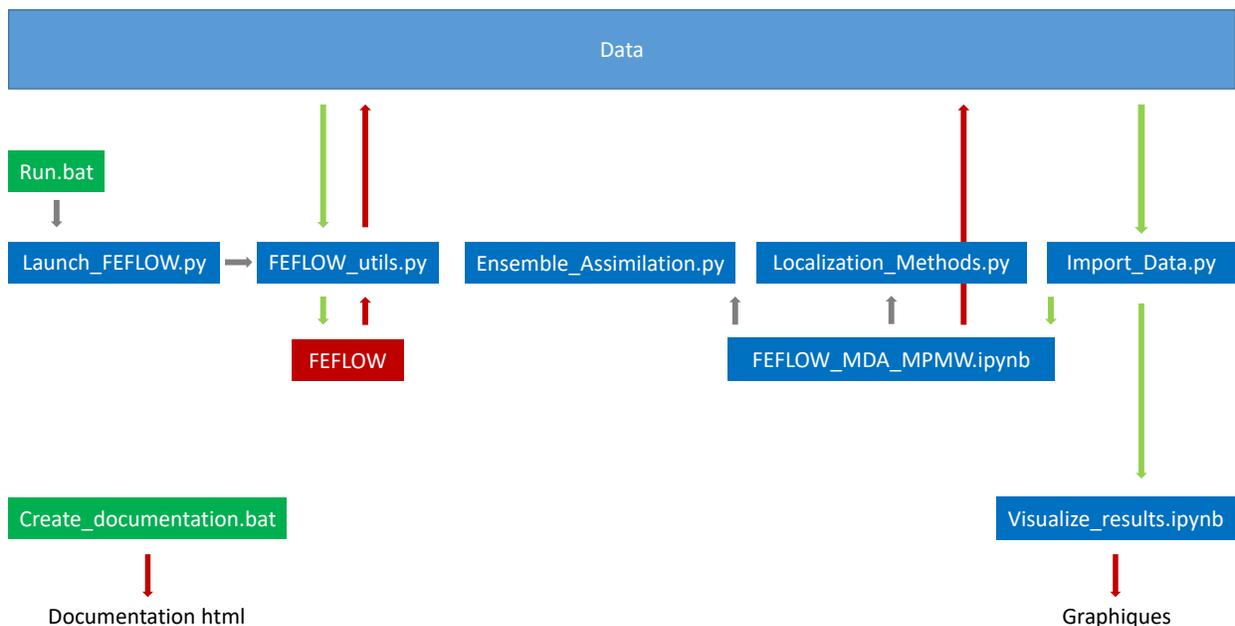


Figure 3.14 – Les flèches vertes indiquent l'import de données. Les flèches rouges indiquent l'export de données. Les flèches grises traduisent l'appel d'un code par un autre code.



# Chapitre 4

## Modèle synthétique

Pour tester les différentes hypothèses avancées sur la localisation et l'analyse locale, un modèle synthétique a été construit. Ce dernier permettra d'évaluer les performances des schémas d'assimilation au niveau des mesures et observations, mais également sur le champs de perméabilité. Dans la réalité, le champ de perméabilité réel est inconnu.

Dans un premier temps les hypothèses sur la géologie du modèle, le modèle conceptuel et les hypothèses hydrogéologiques seront détaillées. Dans un second temps, l'ensemble des schémas d'assimilation qui seront évalués, ainsi que leurs objectifs respectifs, seront expliqués.

Enfin, une analyse des différents résultats au travers des RMSE (Erreurs Moyennes au Carré) sur les charges hydrauliques, les concentrations et sur les champs de perméabilités sera conduite.

### 4.1 Modèle géologique synthétique

Pour essayer de reproduire une étude de cas réel, un modèle géologique du sous-sol a été créé comme modèle de référence. L'idée est de définir un modèle reproduisant un aquifère côtier. Ce modèle sera majoritairement sableux, avec une couche argileuse horizontale en son centre. Ce modèle de référence ne sera jamais directement utilisé pour le processus d'assimilation. Il sera utilisé pour comparer les modèles de perméabilités obtenus à la fin des méthodes d'assimilations testées. Le modèle fait 100 m par 100 m , dans le plan horizontal et fait 20 m d'épaisseur. Dans ce modèle

géologique, 5 puits virtuels ont été positionnés. Les perméabilités observées entre l'intersection des puits et du modèle ont servi de données d'entrées pour la modélisation géostatistique. Aucune erreur de mesure n'a été ajoutée sur ces mesures de perméabilités, ce qui est un cas idéal impossible à atteindre dans la réalité. Le champ de perméabilité synthétique en 3D a été généré à l'aide du programme FGEN92 développé par (Robin *et al.*, 1993). Les paramètres d'entrées utilisés sont donnés dans le Tableau 4.1. La couche centrale argileuse a été pensée avec une portée différente pour introduire une région non-stationnaire dans le modèle.

**Tableau 4.1 – Paramètres utilisés pour construire le champ de perméabilité de référence dans FGEN92, développé par (Robin *et al.*, 1993).**

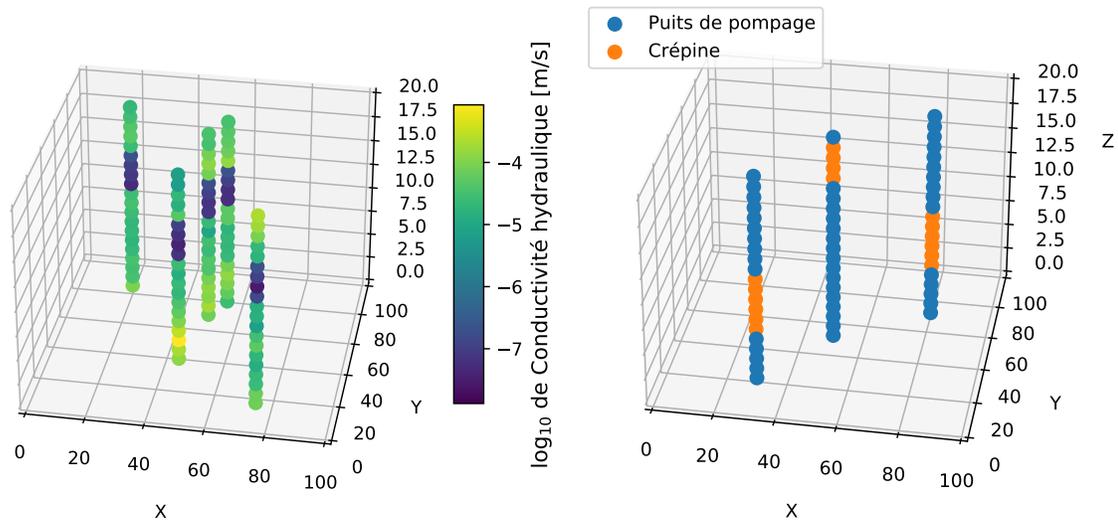
<b>Région</b>	<b>ln moyen perméabilité</b>	<b>ln Std perméabilité</b>	<b>Portée hor. <math>x</math> en <math>m</math></b>	<b>Portée hor. <math>y</math> en <math>m</math></b>	<b>Portée ver. <math>z</math> en <math>m</math></b>
Sableuse	-10	1.2	15	10	3
Couche argileuse	-16	1.2	30	30	3

Les paramètres FGEN92 utilisés pour le modèle général sableux sont des portées horizontales de 15 m, 10 m et une portée verticale de 3 m. La conductivité hydraulique moyenne est de  $4.5 * 10^{-5}$  m/s et la variance de  $8.1 * 10^{-6} m^2/s^2$  dans les parties supérieure et inférieure de la couche centrale argileuse. Dans le milieu du modèle, une couche argileuse horizontale de plus faible conductivité est définie entre 10 m et 13 m. La conductivité hydraulique moyenne est de  $1.1 * 10^{-7}$  m/s et la variance de  $1.8 * 10^{-7} m^2/s^2$ .

Pour reproduire un cas réel, des biais d'échantillonnage réel ont été reproduits dans le modèle synthétique.

Comme décrit par (Goovaerts, 1997), le pas d'échantillonnage devrait varier pour étudier la corrélation spatiale. Le positionnement des puits a été réalisé à des distances variables les uns des autres. La quantité de puits a été limitée dans le modèle. Cette décision provient de l'analyse d'études environnementales où le budget était très limité, et donc le nombre de forages. Le résultat de cet échantillonnage est donné sur la Figure 4.1a, avec la position des 5 puits virtuels.

À ce point de la modélisation, la méthode d'échantillonnage a déjà introduit un biais dans les données. Un des puits virtuels est tombé dans une zone fortement perméable en profondeur, comme indiqué sur la Figure 4.1a. Ceci aura pour impact de surestimer la perméabilité dans tous les modèles



(a) 5 puits virtuels forés dans le modèle géologique synthétique 3D obtenu avec le programme FGEN92 développé par (Robin *et al.*, 1993).

(b) Position de 3 puits de pompage dans le modèle tridimensionnel. Les points orange représentent la zone crépinée utilisée pour le pompage.

**Figure 4.1 – Puits virtuels dans le modèle synthétique.**

de perméabilité de l'ensemble autour de ce point. Aucun autre point de contrôle n'étant présent dans la partie sud-ouest du modèle, cette surestimation pourra s'étendre dans cette direction en fonction de la portée des variogrammes choisis. Les histogrammes de conductivité hydraulique calculés sur les puits échantillonnés sont donnés sur la Figure 4.2. Les deux familles de conductivité hydraulique sont toujours bien visibles. Au regard de cette information, deux régions ont été créées pour les simulations géostatistiques.

À cause du faible nombre de points présent dans la région de faible conductivité, le variogramme défini dans la région de haute conductivité a été réutilisé. Une hypothèse de continuité dans la portée de corrélation spatiale a été faite. Ce manque d'information reflète souvent la problématique rencontrée dans les études environnementales, avec malheureusement souvent peu d'information. Ce manque peut souvent être lié au faible budget de caractérisation disponible. Les variogrammes montrés sur la Figure 4.3 ne sont pas parfaits, mais ils sont représentatifs de certains variogrammes avec peu de données. Le variogramme horizontal omnidirectionnel a été obtenu avec un pas de 2 m et une tolérance de 1 m. Le variogramme vertical a été obtenu avec un pas de 1 m et une tolérance de 0.5 m. Le variogramme horizontal de la Figure 4.3a a été choisi pour favoriser les distances les plus courtes. Après 25 m, le modèle respecte la variance verticale et la variance moyenne observée entre

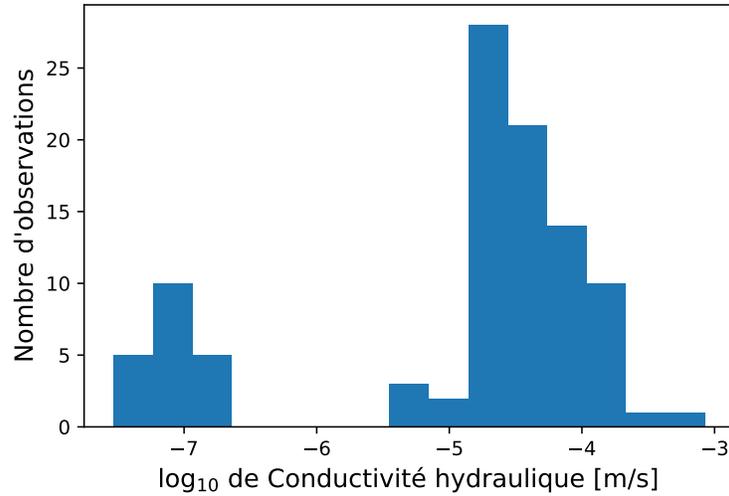


Figure 4.2 – Histogramme de conductivité hydraulique observé dans les 5 puits virtuels du modèle synthétique 4.1a.

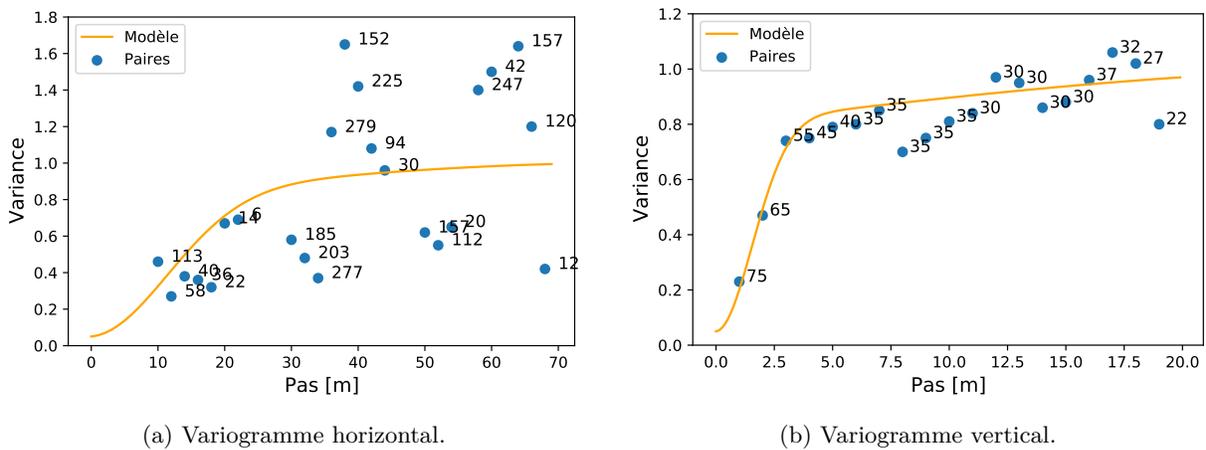


Figure 4.3 – Variogrammes calculés et les paires de points utilisées comme support.

les principales paires de points. Ces variogrammes ont été calculés après transformation gaussienne des perméabilités.

Les variogrammes affichés sur la Figure 4.3 sont composés de 3 structures élémentaires résumées dans le tableau suivant :

Tableau 4.2 – Paramètres de variogrammes utilisés pour générer la Figure 4.3.

Type	Plateau	Portée horizontale $m$	Portée verticale $m$
Effet pépite	0.05	-	-
Gaussien	0.75	28	3.8
Sphérique	0.2	80	30

À cause du manque d'échantillons de perméabilité, la corrélation spatiale inférée depuis les variogrammes ne correspond plus exactement aux paramètres imposés en entrées du modèle synthétique. L'anisotropie horizontale dans le plan  $(x, y)$  a été perdue, suite à l'utilisation d'un modèle omnidirectionnel. La plus longue corrélation spatiale présente dans la couche argileuse a également été perdue. Pour construire les variogrammes, un modèle omnidirectionnel a dû être choisi pour les deux couches. Le variogramme final a une portée horizontale de 28 m et une portée verticale de 3.8 m. Une tendance à plus longue portée a aussi été ajoutée. Ces tendances sont le résultat du sous-échantillonnage de la distribution de perméabilité par rapport à la distribution initiale du Tableau 4.1.

Pour prendre en compte ces incertitudes, des simulations séquentielles gaussiennes ont été effectuées pour laisser varier les champs de perméabilité loin des points d'observation tout en conservant une structure spatiale réaliste (Goovaerts, 1997). Les champs de perméabilité ont été construits grâce au logiciel SGEMS (Wu *et al.*, 2009). 100 simulations de perméabilité ont été créées pour construire l'ensemble de paramètres d'état de départ. Les 5 puits de la figure 4.1a imposent les données brutes aux modèles. Une région horizontale a été définie autour des faibles perméabilités entre 10 m et 13 m, comme identifié sur l'histogramme de la Figure 4.2. Les variogrammes et histogrammes simulés furent validés dans SGEMS pour assurer la qualité des simulations (même portée, variance et graphiques. Q-Q P-P).

À la fin de cette étape, 100 réalisations équiprobables de champ de perméabilité ont été obtenues à partir des 5 puits virtuels forés dans le modèle synthétique de référence du sous-sol. Dans la section suivante, le modèle d'aquifère côtier utilisé dans les simulations d'écoulement du cas synthétique sera présenté.

#### 4.1.1 Modèle 3D synthétique d'aquifère côtier

Le cas synthétique a été construit pour reproduire un aquifère côtier avec une intrusion saline sur un côté du modèle. La durée totale de la simulation fut de 2000 jours. Pour assurer la stabilité numérique des simulations, tous les modèles ont d'abord été roulés en régime permanent, pour atteindre un équilibre entre l'eau fraîche et salée avant d'être perturbée en régime transitoire. 3 puits de pompes étaient actifs entre 200 et 1300 jours pour introduire des variations dans l'équilibre des modèles. La position de ces puits de pompage est donnée sur la Figure 4.4 avec des pentagones

rouges. 2 puits avaient leurs crépines sous la couche imperméable, un puits l'avait au-dessus. Les crépines sont rappelées sur la Figure 4.1b.

3 types de puits sont présents dans le modèle présenté sur la Figure 4.4 :

- Puits de pompage : Introduit une modification dans les charges hydrauliques et la concentration en sel après l'établissement du régime permanent;
- Puits de référence : Puits virtuellement foré dans le modèle de référence pour étudier la perméabilité. Analogue au puits de caractérisation dans les études environnementales où les slugs tests sont réalisés. Ces puits ont été utilisés pour construire l'ensemble de paramètres d'états initiaux;
- Puits d'observations : Puits de suivi des charges hydrauliques et des concentrations tout au long des simulations. Uniquement les puits d'observations sont nommés sur la Figure 4.4 car ce sont ces derniers qui vont être simulés pour les processus d'écoulement.

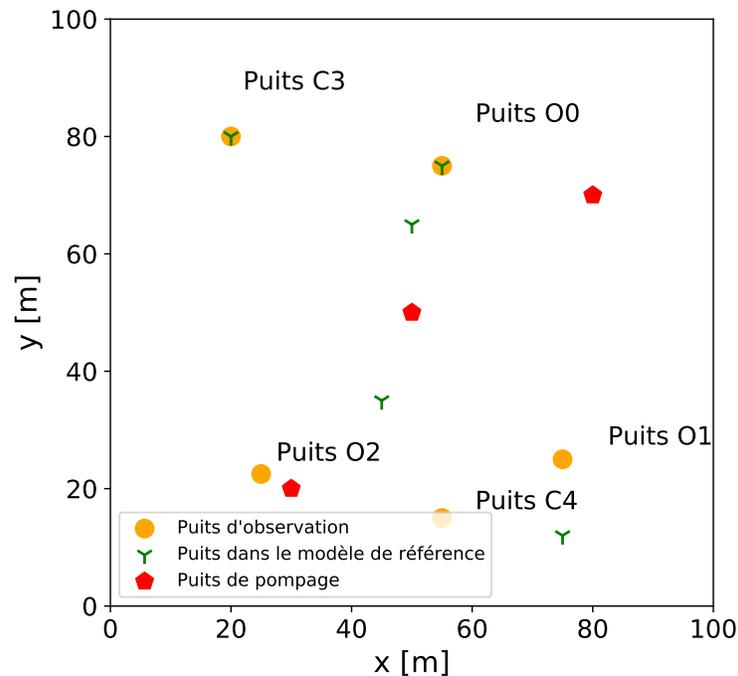
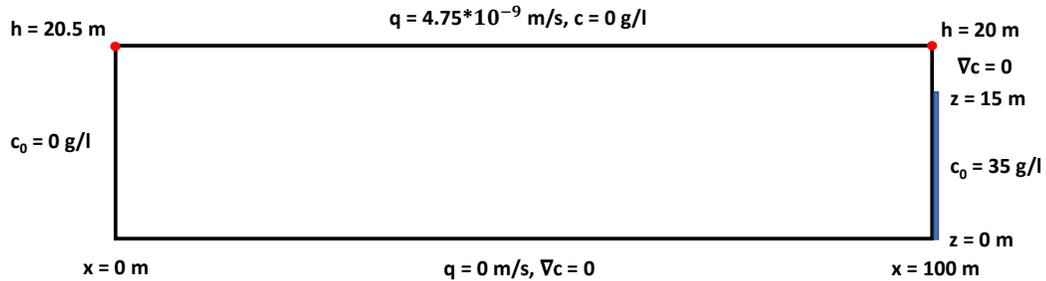


Figure 4.4 – Position des 3 types de puits dans le modèle synthétique. Les points orange représentent les puits d'observation pour suivre les charges hydrauliques et les concentrations tout au long des simulations. Les pentagones rouges sont les puits de pompage de la Figure 4.1b. Les triforces vertes sont les puits de références utiliser pour étudier le champ de perméabilité.



**Figure 4.5** – Section en deux dimensions le long de l’axe  $x$  des conditions aux limites du modèle FEFLOW.

Pour évaluer la capacité d’assimilation de la méthode d’ensemble, plusieurs combinaisons d’observation, de pompage et de puits de référence furent établies. Comme indiqué sur la Figure 4.4, certains puits d’observation et de référence sont les mêmes. Les autres puits d’observation sont loin de tout puits de référence et plus ou moins proche des puits de pompage. Cette distribution de puits a été choisie afin de mesurer la performance de la méthode ES-MDA 4x en différentes conditions. La simulation d’écoulement et de concentration a été réalisée avec le logiciel FEFLOW. Les paramètres imposés au modèle conceptuel FEFLOW sont donnés dans la section suivante.

### Paramètres du modèle d’écoulement souterrain dans FEFLOW

Une charge hydraulique de 20.5 m a été imposée sur la bordure gauche du modèle et 20 m sur la bordure droite. Une recharge de  $4.75 \times 10^{-9}$  m/s ou 150 mm/an a été définie sur le haut du modèle. Une arrivée d’eau salée avec une concentration de 35 g/l est fixée sur la bordure droite du domaine entre 0 m et 15 m. Une condition d’imperméabilité était imposée sur la limite inférieure du modèle. Enfin, de l’eau avec une concentration saline de 0 g/l a été imposé sur les bordures supérieures et gauches du modèle. Un résumé graphique de toutes ces contraintes et des conditions initiales est donné sur la Figure 4.5.

Une porosité de 0.3 et un coefficient de diffusion de  $10^{-9}$   $m^2/s$  sans décroissance ou retard ont été imposés dans l’ensemble du modèle. La dispersivité longitudinale était de 5 m et la dispersivité transversale verticale de 0.5 m. Le débit de pompage pendant l’essai de pompage était égal à 40  $m^3/j$  pour la somme des crépines de tous les puits.

Tous ces paramètres n'ont pas été choisis pour reproduire quelconque aquifère côtier réel. Ils ont été choisis pour assurer une stabilité numérique dans les simulations et valider ou non les méthodes d'assimilation testées. La discrétisation spatio-temporelle du modèle a été basée sur les nombres de Peclet et Courant respectivement. Le nombre de Peclet a été imposé pour rester en dessous de 2 afin de minimiser la dispersion numérique. Le nombre de Peclet est le ratio entre le taux de transport advectif et le taux de transport dispersif.

$$P_e = \frac{\text{taux de transport advectif}}{\text{taux de transport dispersif}}, P_e \leq 2, P_e = \frac{v\Delta x}{\alpha_L v} = \frac{\Delta x}{\alpha_L}. \quad (4.1)$$

Avec  $v$  la vitesse d'écoulement,  $\Delta x$  le pas de la grille et  $\alpha_L$  la dispersivité longitudinale. En accord avec l'équation précédente, la dimension des éléments de la grille a été choisie tel que :

$$\Delta x \leq 2\alpha_L, \Delta x \leq 10m. \quad (4.2)$$

La discrétisation spatiale devrait être inférieure à 10 m pour éviter la dispersion numérique dans les simulations. En accord avec ce résultat, la taille maximum des éléments dans FEFLOW a été fixée à 6.4 m dans le plan horizontal et 1 m dans la direction verticale  $z$ .

Pour éviter les oscillations numériques autour de la solution, le nombre de Courant devrait être plus petit que la moitié du nombre de Peclet. La vitesse maximum de la vitesse d'écoulement souterraine était autour de  $1 * 10^{-5}$  m/s dans le modèle de référence pendant le pompage, et 10 fois plus petit sans pompage:

$$C_r = \frac{v\Delta t}{\Delta x}, C_r \leq \frac{P_e}{2}, \Delta t \leq \frac{\Delta x}{v}, \quad (4.3)$$

$$\Delta t \leq 10 \text{ jours} \quad (4.4)$$

Le pas de temps adaptatif est disponible dans FEFLOW pour assurer la convergence numérique de la solution à chaque pas de temps, mais un pas de temps maximum a été imposé à 10 jours en accord avec l'équation 4.4. Dans la suite du processus d'assimilation, tous les temps ont été rééchan-

**Tableau 4.3 – Crépines des puits utilisées pour suivre la concentration maximale en sel dans les puits d’observations et de contrôles.**

Puits	Élévation minimum	Élévation maximale
O0	6	10
O1	6	10
O2	11	15
C3	5	9
C4	9	13

tillonnés pour tous observer la même charge ou concentration au même moment. Une interpolation linéaire entre les pas de temps issus des simulations FEFLOW a été effectuée.

La concentration en sel et les charges hydrauliques étaient suivies sur 5 puits. 3 de ces puits ont été utilisés pour l’assimilation. Les 2 puits restants servant de puits de contrôle pour évaluer la qualité de l’assimilation loin de points de mise à jour. Sur la Figure 4.4, les puits de contrôle sont nommés CX et les puits d’observations OX. Les puits de contrôle C3 et d’observation O0 sont situés sur des puits de référence du modèle. Sur ces des puits, les charges et la concentration devraient être proches de celle du modèle de référence. La concentration maximum de sel était suivie sur ces 5 puits. L’élévation des crépines de ces 5 puits est donnée dans le Tableau 4.3.

FEFLOW a été utilisé en régime permanent dans un premier temps pour que chaque modèle de perméabilité atteigne son équilibre. Ensuite, seulement, les modèles ont été lancés en régime transitoire avec le pompage. La concentration en sel était stable avant le démarrage du pompage pour l’ensemble des modèles. On peut voir cette stabilité sur le modèle de référence dans la Figure 4.6. La concentration maximale en sel pour le modèle de référence varie entre  $1.0 * 10^{-2}$  mg/l et  $7.5 * 10^2$  mg/l. Cette amplitude sera un challenge pour le processus d’assimilation à venir. Les anomalies provenant des hautes concentrations en sel pouvant effacer de l’information provenant des puits à plus faible concentration.

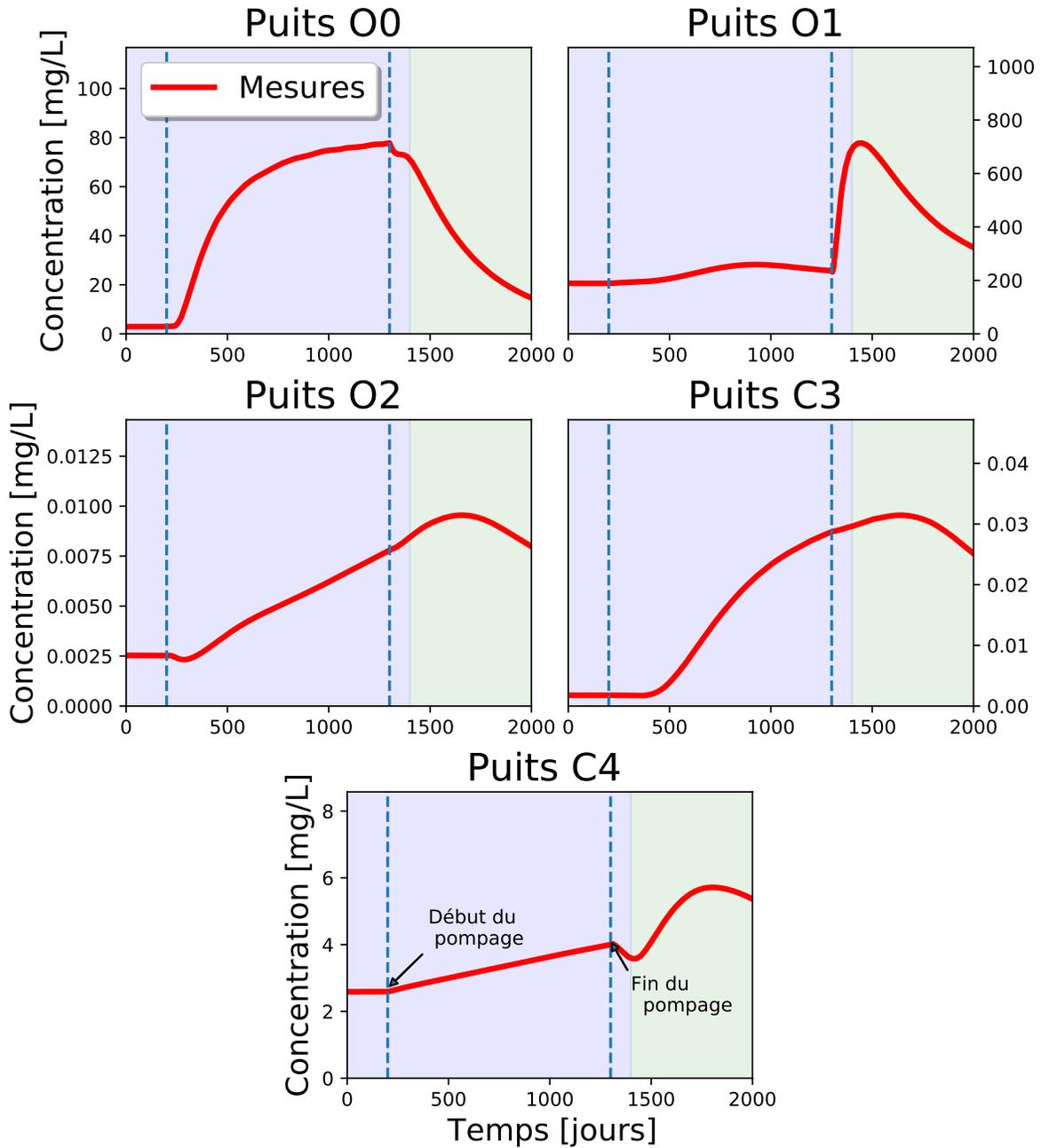


Figure 4.6 – Concentration en sel maximale observée dans les 5 puits du modèle de référence. La plage de donnée en bleu correspond à la fenêtre de calibration. Les données avec un fond vert correspondent aux données de validation. La position de ces puits est donnée en vue du dessus sur la Figure 4.4.

### 4.1.2 Assimilation

3 puits ont été utilisés pour assimiler les modèles de perméabilité en étudiant les charges hydrauliques et les concentrations entre le modèle de référence et les modèles de l'ensemble. 2 puits étaient également suivis pour suivre l'évolution des charges et des concentrations dans l'ensemble plus loin de points assimilés.

Pour tester l'efficacité des différents schémas de localisation et d'analyse locale, 8 schémas ont été testés :

- 0 : Ensemble initial sans aucune assimilation;
- 1 : Assimilation brute avec toutes les observations assimilées en une fois;
- 2 : Assimilation par type de données (tous les puits ensemble), sans analyse locale;
- 3 : Assimilation puits par puits, charge hydraulique seulement, sans analyse locale;
- 4 : Assimilation puits par puits, concentration seulement, sans analyse locale;
- 5 : Assimilation puits par puits, type de données par type de données, sans analyse locale;
- 6 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale  $100m$  et  $20m$  en vertical;
- 7 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale  $50m$  et  $10m$  en vertical;
- 8 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale  $25m$  et  $5m$  en vertical;

Les différents schémas d'analyses locales pour l'affectation de poids ont été choisis en lien avec les portées de variogramme et la dimension du modèle. L'analyse locale la plus grande reflète la dimension du modèle 3D. L'idée avec cette grande portée est de mettre à jour toutes les cellules du modèle, mais avec un poids qui décroît en s'éloignant de points d'observations. Les deux autres portées choisies sont des multiples des portées des variogrammes.  $50/10$  est presque le double des portées de variogrammes tandis que le schéma  $25/5$  correspond environ à la portée des variogrammes. Ces 3 portées différentes dans l'analyse locale ont été testées pour évaluer l'impact de la distance autour des points d'observations dans la méthode ES-MDA 4x. Les points en dehors de la portée d'analyse locale conserveront leurs incertitudes, avec des variations uniquement dues aux variations issues des simulations géostatistiques.

Chaque schéma testé évaluera une hypothèse sur la méthode d'analyse locale utilisée. L'assimilation brute en une seule fois servira de point de référence. Des schémas avec et sans analyse locale, mais puits par puits, seront utilisés pour évaluer l'apport de la méthode basée sur des assimilations puits par puits et type de donnée par type de donnée. Enfin, la combinaison d'assimilation puits par puits, donnée par donnée avec des analyses locales de différentes portées sera étudiée. Chacun des schémas testés s'appuie sur le processus brut ES-MDA 4x décrit par (Emerick & Reynolds, 2013).

Dans la section suivante, les données de calibration et de validation vont être présentées. Uniquement une sous-partie des données disponibles sur les points d'observations a été utilisée pour mettre à jour les ensembles. Cela a permis d'évaluer la performance de l'assimilation sur les données calibrées utilisées par l'ES MDA 4x et sur les données de validation jamais assimilées par la méthode. De plus, les puits de contrôles permettront d'évaluer la performance de la méthode sur des puits jamais vus par la méthode. Les puits de contrôles sont également séparés en données de calibration et de validation, mais ces labels sont uniquement liés au temps des observations. Les méthodes d'assimilation ont utilisé les données temporelles de  $t = 0$  jour à  $t = 1400$  jours, ce qui représente 70% des données disponibles sur les points d'observations retenus. La fenêtre de données de calibration est en bleue sur la Figure 4.4, les données de validation sont quant à elle en verte. Il a été choisi volontairement de ne pas assimiler toutes les données des puits pour pouvoir évaluer la performance de la méthode de manière plus robuste, sur des données non assimilées.

## 4.2 Résultats

Pour évaluer les performances de la méthode d'assimilation ES-MDA 4x, une évaluation de l'erreur moyenne au carré (RMSE) pour la concentration en sel et pour les charges entre les modèles de l'ensemble et le modèle de référence va être conduite. La RMSE au niveau des champs de perméabilité entre les modèles de l'ensemble et celui de référence sera également suivie, ceci est uniquement possible avec un cas synthétique. L'observation de ces erreurs devrait permettre d'étudier les performances pour tous les schémas d'analyses. L'évolution combinée permettra d'évaluer si les mises à jour ne sont pas contraintes par les observations, mais mettent également à jour la perméabilité dans la direction du modèle de référence de perméabilité. Tous les graphiques d'erreur affichent de manière séparée l'erreur de calibration et de validation. Les puits de calibration utilisés pour l'assimilation étaient les puits O1, O2 et O3. Les données de calibration sur les puits de contrôle C3

et *C4* n'ont pas été utilisées pour l'assimilation, ils correspondent uniquement aux mêmes indices temporels. Enfin, les données de validation sont présentes sur tous les puits pour évaluer la qualité des assimilations sur les puits assimilés ou non.

### 4.2.1 RMSE sur les concentrations en sel

La Figure 4.7 montre la RMSE sur les concentrations au niveau des 5 puits, au niveau des données de calibration et de validation. L'ensemble initial représente le résultat des simulations issues directement des modèles géostatistiques sans aucune assimilation. Ces simulations étant basées uniquement sur les données des puits de la Figure 4.1a.

Tous les modèles suivants ont été assimilés avec la méthode ES-MDA 4x décrite par Emerick & Reynolds (2013). 4 assimilations successives ont été réalisées à chaque fois, avec une simulation complète d'écoulement et de simulation de transport de contaminants. Les différents schémas testés sont basés sur la manière de calculer les anomalies durant le processus d'assimilation et la portée de la mise à jour. Toutes les méthodes, à l'exception de la portée de mise à jour 30/5m, ont été capables de réduire la RMSE sur les données calibrées par rapport à l'ensemble initial. Ces résultats démontrent la robustesse des méthodes définies par Evensen (2003) et Emerick & Reynolds (2013). Toutes les assimilations ont tendance à surajuster les observations (overfitting), la RMSE et la variance étant toujours plus faible sur les données calibrées que celle de validation. La RMSE médiane la plus faible pour le puits *O0* est de  $9.3 * 10^0$  pour le schéma 4 contre une RMSE médiane initiale de  $1.7 * 10^0$ . La RMSE médiane la plus faible sur le puits *O1* est de  $1.2 * 10^2$  pour le schéma 4 contre une RMSE médiane initiale de  $1.4 * 10^2$ . Pour le puits *O2* la RMSE médiane est à nouveau la plus faible pour le schéma 4 avec  $7 * 10^{-3}$  contre  $1.3 * 10^{-2}$  initialement. La RMSE médiane la plus faible sur le puits *C3* est de  $3.4 * 10^{-3}$  pour le schéma 3 contre une RMSE médiane initiale de  $3.8 * 10^{-3}$ . Enfin, le puits *C4* a une RMSE médiane minimale de  $9.0 * 10^{-1}$  pour une RMSE médiane initiale de  $1.7 * 10^0$  pour le schéma 5. La meilleure méthode basée uniquement sur la réduction de la RMSE sur les concentrations, sans tenir compte de la RMSE sur les charges hydrauliques, est le schéma d'assimilation 4 sur les données de calibration. Le schéma 4 correspond à l'assimilation de la concentration seule puits par puits. En tenant compte des données de validation, le schéma le plus efficace est le schéma d'analyse locale 3. La variance des ensembles est réduite sur les puits *O1*, *O2* et *C4*. Les puits *O0* et *C3* ont été que légèrement améliorés, ceci pourrait être lié

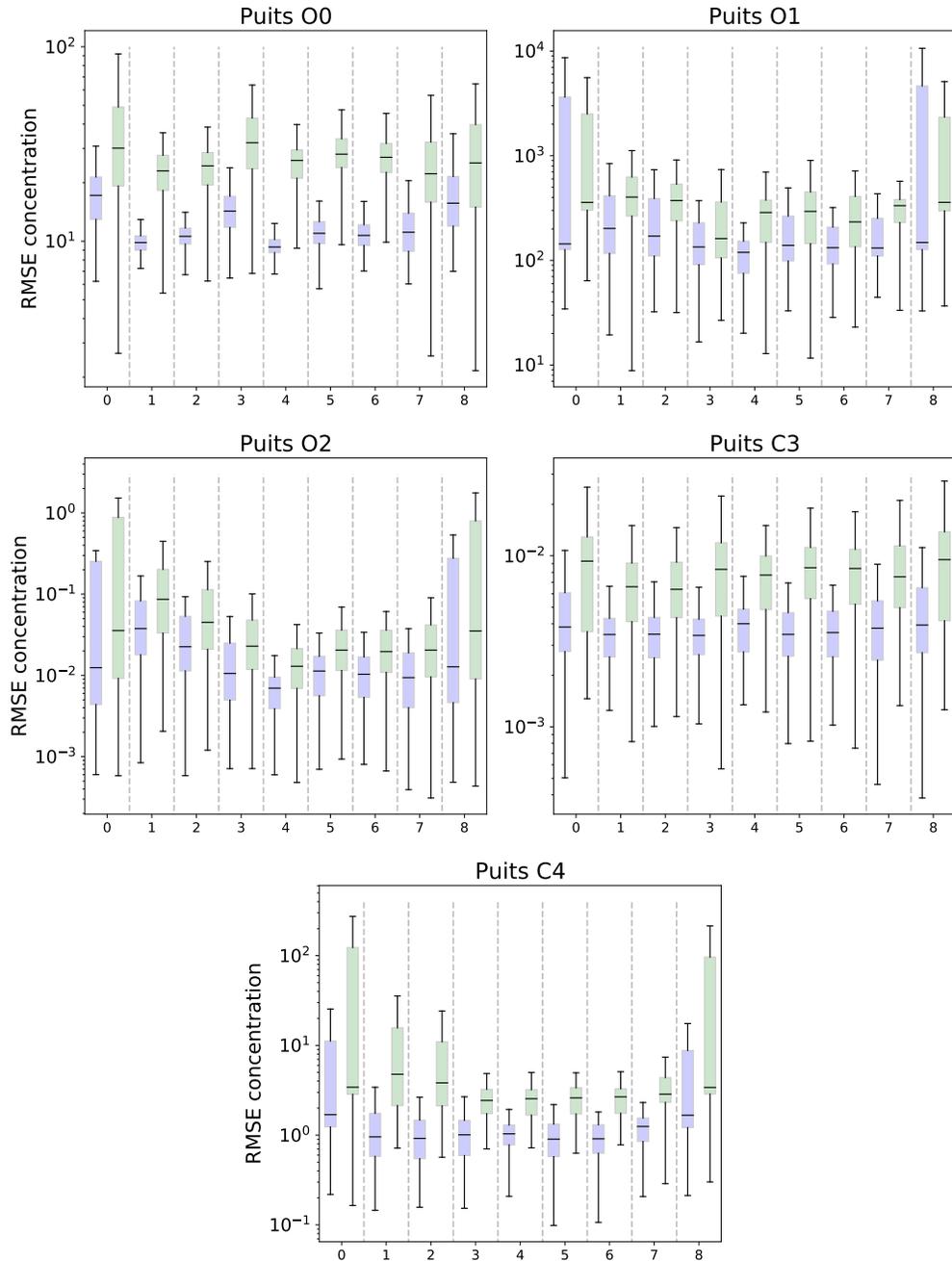


Figure 4.7 – RMSE des concentrations en sel pour les 8 méthodes à la fin de la 4<sup>e</sup> assimilation. Les barres bleues correspondent aux données dans la fenêtre temporelle de calibration, les barres vertes correspondent aux données dans la fenêtre temporelle de validation. 0 : Ensemble initial sans aucune assimilation. 1 : Assimilation brute avec toutes les observations assimilées en une fois. 2 : Assimilation par type de données (tous les puits ensemble), sans analyse locale. 3 : Assimilation puits par puits, charge hydraulique seulement, sans analyse locale. 4 : Assimilation puits par puits, concentration seulement, sans analyse locale. 5 : Assimilation puits par puits, type de données par type de données, sans analyse locale. 6 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale 100m et 20m en vertical. 7 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale 50m et 10m en vertical. 8 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale 25m et 5m en vertical.

au fait que les puits  $O0$  et  $C3$  étaient imposés dans les simulations gaussiennes séquentielles. Très peu de variations étaient présentes dans les ensembles à ces endroits, ce qui donnait un ensemble artificiellement contraint restant proche des valeurs initiales. On pourra également noter que les valeurs initiales de concentration des puits  $O0$  et  $C3$  était très proche des valeurs de référence comme indiqué sur la Figure 4.14. Cette proximité rendant toute amélioration relative plus compliquée. Le puits  $O2$  a la plus faible concentration des 3 puits assimilés, sur ce puits l'apport de l'assimilation puits par puits est supérieur à celle de l'assimilation globale. Un résultat encourageant est que l'assimilation des charges hydrauliques seules (schéma 3) permet tout de même d'améliorer la RMSE sur les concentrations. Ce résultat est très important pour les études environnementales où les concentrations sont très peu suivies, mais pourtant si cruciales.

#### 4.2.2 RMSE sur les charges hydrauliques

Assimiler les charges hydrauliques uniquement, puits par puits, est la meilleure méthode pour réduire la RMSE sur les charges hydrauliques avec le schéma 3. Les meilleures méthodes pour réduire la RMSE sont à nouveau celles basées sur des assimilations puits par puits pour les points  $O0$ ,  $O1$  et  $O2$  comme montré sur la Figure 4.8. Le puits de contrôle  $C3$  n'a pas montré d'amélioration lors du processus d'assimilation, ceci pourrait toujours être lié au fait que  $C3$  était aussi un puits de référence dans le modèle synthétique. Finalement, la RMSE sur les charges hydrauliques sur le puits  $C4$  est améliorée sur tous les puits à l'exception de la méthode d'analyse locale  $25/5m$ . Les méthodes qui réduisent le plus la RMSE sur les charges hydrauliques en assimilant les charges et la concentration sont celles basées sur des assimilations puits par puits et type de données par type de données. L'assimilation de la concentration seule ne semble pas capable de réellement améliorer la RMSE sur les charges hydrauliques.

La RMSE médiane la plus faible pour le puits  $O0$  est de  $8.9 * 10^{-3}$  pour le schéma 3 contre une RMSE médiane initiale de  $3.5 * 10^{-2}$ . La RMSE médiane la plus faible sur le puits  $O1$  est de  $7.5 * 10^{-3}$  pour le schéma 3 contre une RMSE médiane initiale de  $3.4 * 10^{-2}$ . Pour le puits  $O2$  la RMSE médiane est la plus faible pour le schéma 3 avec  $7.1 * 10^{-3}$  contre  $5.3 * 10^{-2}$  initialement. La RMSE médiane la plus faible sur le puits  $C3$  est de  $1.1 * 10^{-2}$  pour le schéma 4 contre une RMSE médiane initiale de  $4.1 * 10^{-2}$ . Enfin, le puits  $C4$  a une RMSE médiane minimale de  $8.6 * 10^{-3}$  pour

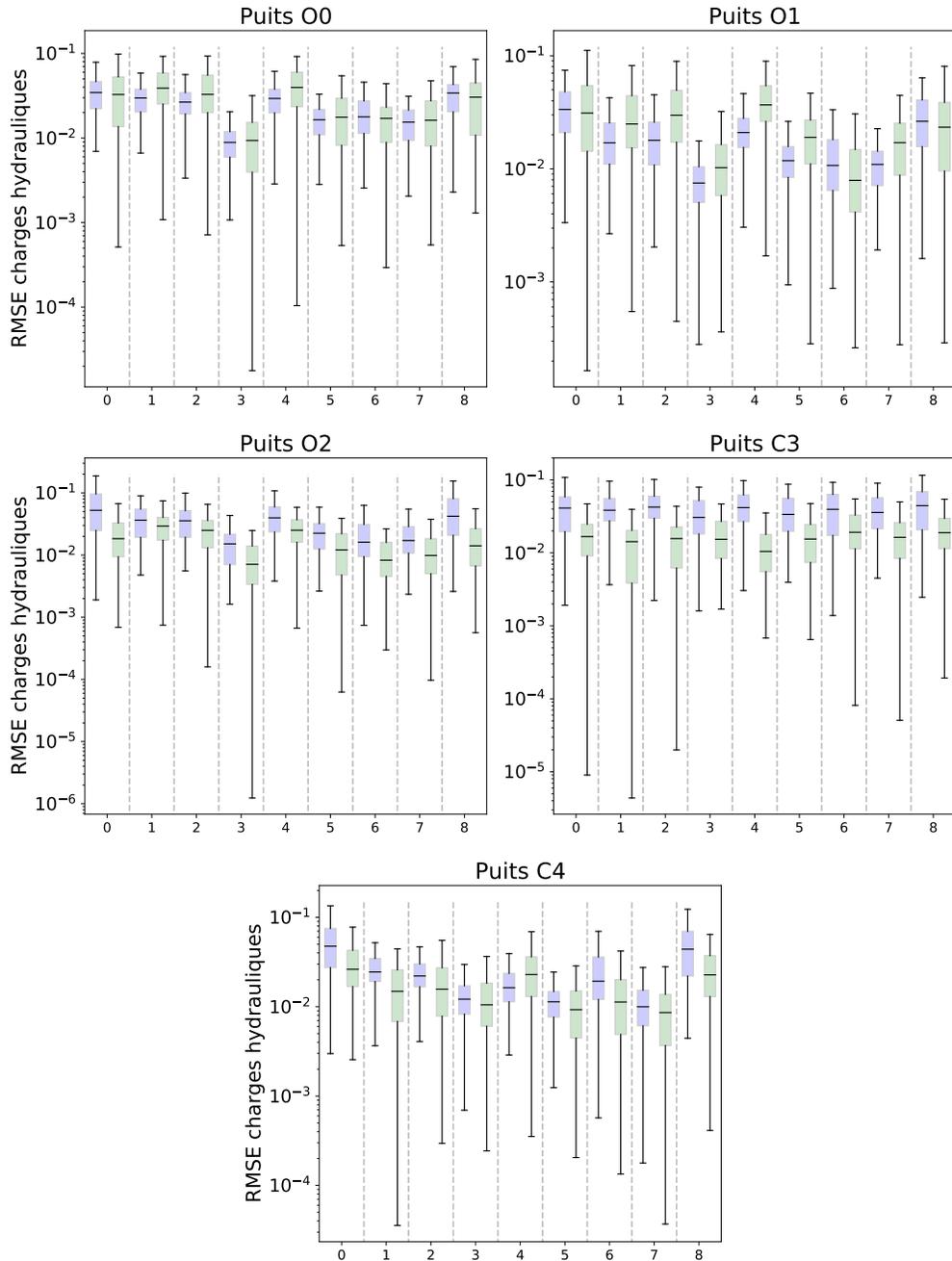


Figure 4.8 – RMSE sur les charges hydrauliques pour les 8 méthodes à la fin de la 4<sup>e</sup> assimilation. Les barres bleues correspondent aux données dans la fenêtre temporelle de calibration, les barres vertes correspondent aux données dans la fenêtre temporelle de validation. 0 : Ensemble initial sans aucune assimilation. 1 : Assimilation brute avec toutes les observations assimilées en une fois. 2 : Assimilation par type de données (tous les puits ensemble), sans analyse locale. 3 : Assimilation puits par puits, charge hydraulique seulement, sans analyse locale. 4 : Assimilation puits par puits, concentration seulement, sans analyse locale. 5 : Assimilation puits par puits, type de données par type de données, sans analyse locale. 6 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale 100 m et 20 m en vertical. 7 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale 50 m et 10 m en vertical. 8 : Assimilation puits par puits, type de données par type de données, analyse locale de portée horizontale 25 m et 5 m en vertical.

une RMSE médiane initiale de  $4.8 * 10^{-2}$  pour le schéma 7. Pour les charges hydrauliques, ce n'est plus nécessairement les données de calibration qui ont les RMSE les plus faibles.

L'amélioration absolue sur les charges hydrauliques est plus faible que sur les concentrations, l'ensemble initial étant beaucoup plus proche des valeurs de référence pour les charges. Les résultats sur les Figures 4.7 et 4.8 indiquent que la portée de l'analyse locale ne doit pas être trop petite. La portée 25/5 m est trop faible pour permettre une mise à jour suffisante des modèles de l'ensemble. Dans ce cas synthétique, l'analyse locale devrait au moins être capable de mettre à jour la bordure droite du domaine, laissant entrer plus ou moins de sel. Une autre manière de choisir une portée d'analyse locale est de prendre une portée qui permet à chaque élément d'être mise à jour au moins par un point d'observation. Une couverture multiple de chaque élément par plusieurs points d'observation reste idéale, il s'agit du meilleur moyen de réduire les incertitudes et les fausses corrélations. Pour évaluer les performances globales de toutes les méthodes, la RMSE sur les champs de perméabilité va être étudiée.

Toutes les méthodes sont efficaces sur les puits d'observation et de contrôle, comme l'indiquent les améliorations de la RMSE sur les deux jeux de puits. L'amélioration est plus faible sur les puits d'observation qui sont également des puits de référence. Cette plus faible amélioration provient d'une combinaison d'ensemble de perméabilité moins variable autour de ces points et de valeurs initiales simulées plus proche des valeurs de référence.

### 4.2.3 Amélioration moyenne en pourcent des RMSE

#### Données de calibration

Pour évaluer la performance globale du processus d'assimilation, la variation médiane en pourcent de la RMSE a été calculée par rapport à l'ensemble initial. Le pourcentage permettra de gérer les différences d'amplitudes entre les types de données.

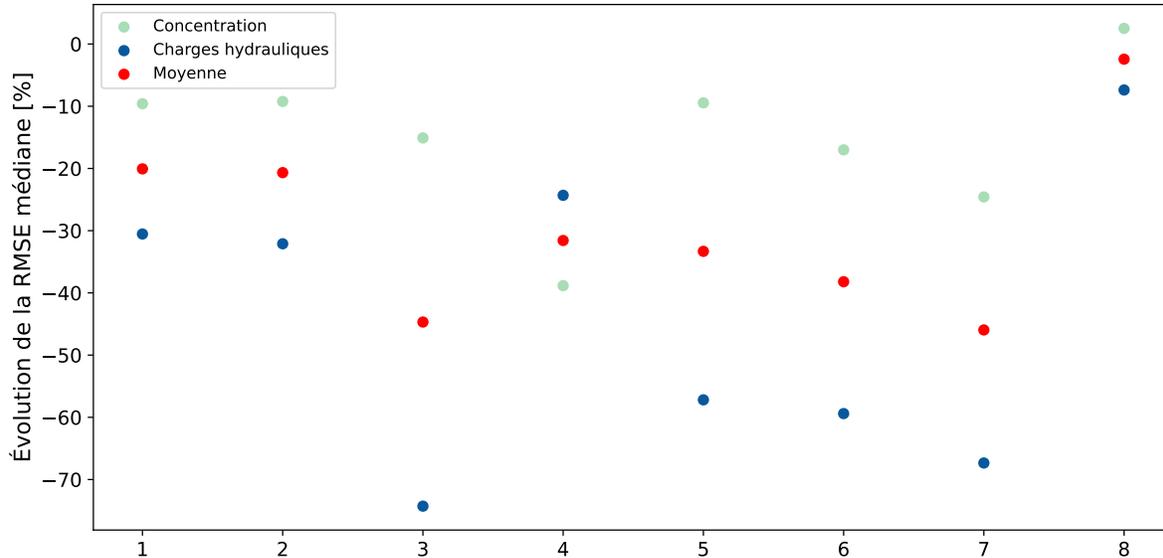


Figure 4.9 – Amélioration moyenne en pourcentage de la RMSE par rapport à l’ensemble initial de calibration. L’évolution de la RMSE des concentrations est donnée en bleue foncée, celle des charges hydrauliques est donnée est cyan. La moyenne d’évolution globale de la RMSE est indiquée en rouge.

Les points sur la Figure 4.9 et 4.10 sont les améliorations en pourcentage pour chaque schéma étudié. Les points bleus foncés représentent l’amélioration en pourcentage de la RMSE médiane sur les charges hydrauliques. Les points rouges traduisent l’amélioration moyenne de la RMSE médiane en pourcentage pour les deux types de données. Les points cyan donnent l’amélioration en pourcentage de la RMSE médiane sur les concentrations.

La Figure 4.9 indique l’amélioration en pourcentage pour les données calibrées seulement. Les meilleures méthodes d’assimilation sont à nouveau celles basées sur des assimilations puis par puits et type de donnée par type de donnée. Avec une réduction médiane de la RMSE de 43.1% pour les schémas 3 et 6 et de 45.3% pour le schéma 7. Sur tous les schémas à l’exception du 4 (assimilation de la concentration uniquement), la réduction de la RMSE médiane était plus importante sur les charges hydrauliques que sur les concentrations. L’analyse locale de portée 25/5 m, ou la portée de mise à jour est trop faible pour permettre l’assimilation, voit sa RMSE médiane se détériorer de 2.5%. L’assimilation des charges hydrauliques seule fonctionne mieux que l’assimilation des concentrations uniquement. Avec une amélioration moyenne de 43.1% contre une amélioration moyenne de 30.3%.

## Données de validation

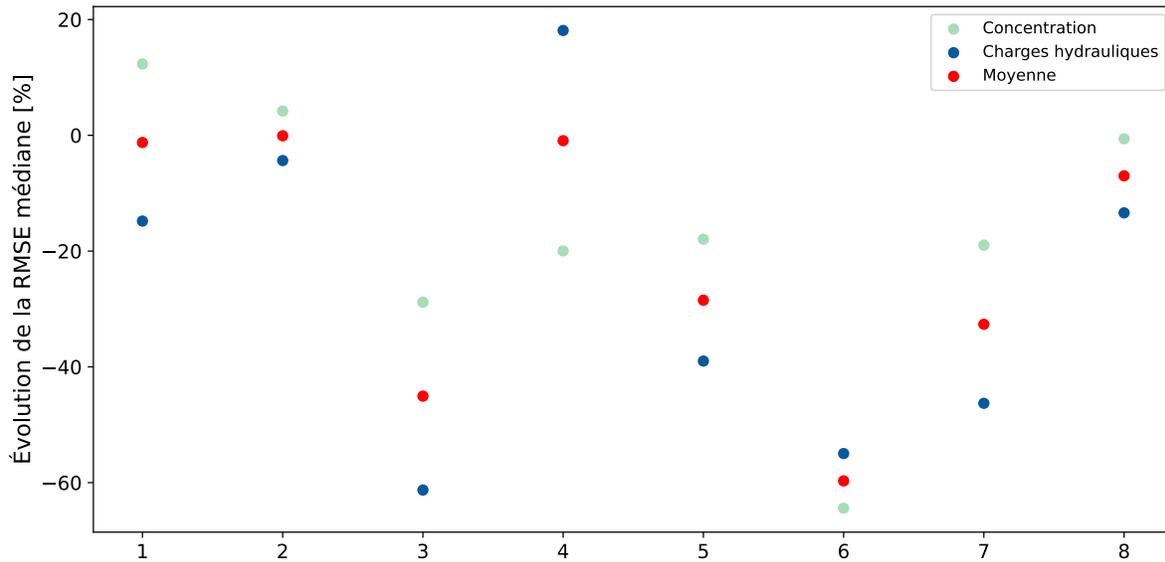


Figure 4.10 – Amélioration moyenne en pourcentage de la RMSE par rapport à l’ensemble initial de validation. L’évolution de la RMSE des concentrations est donnée en bleu foncée, celle des charges hydrauliques est donnée en cyan. La moyenne d’évolution globale de la RMSE est indiquée en rouge.

Sur la Figure 4.10, l’amélioration en pourcentage a été uniquement calculée sur les données de validation, celles-ci n’ayant jamais été utilisées par le processus d’assimilation. Sur ces données, le schéma ayant le plus réduit la RMSE médiane est le numéro 6 avec une réduction de 63.0%, suivi à nouveau par le numéro 3 avec 46.3%. Sur ces données, la meilleure méthode est l’assimilation puits par puits, type de donnée par type de donnée avec une analyse locale de 100/20 m. La différence avec les autres méthodes est que l’analyse locale a l’air de réduire le surajustement sur les données de calibration. Sans analyse locale, les éléments loin des points d’observations pourraient être surajustés sur de fausses corrélations pour mieux correspondre aux données observées, mais dans le cas de nouvelles observations les modèles seraient moins performants.

## Données complètes

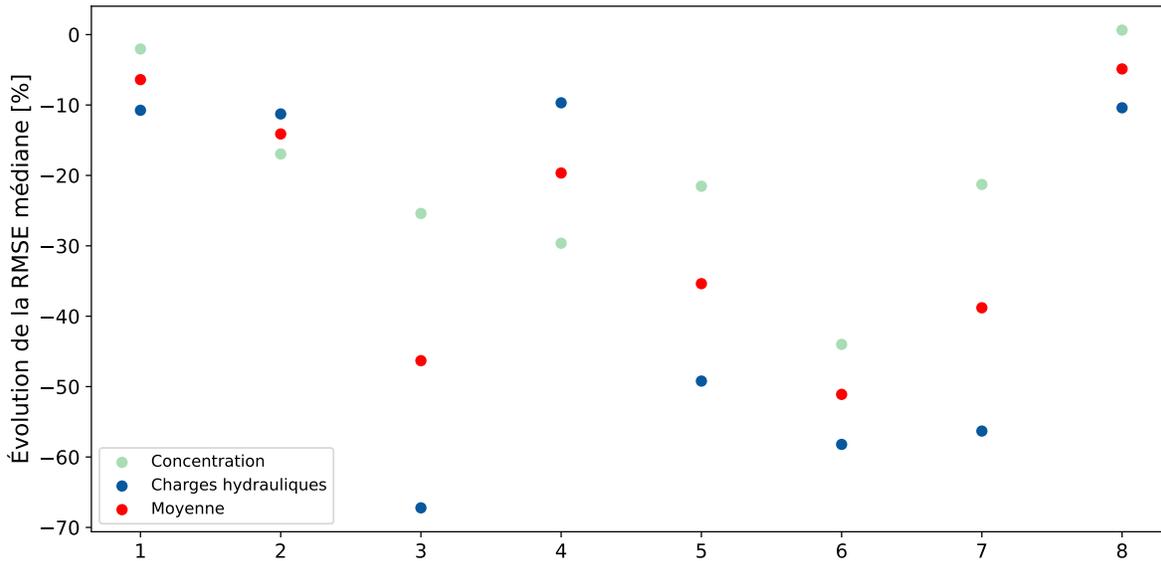


Figure 4.11 – Amélioration moyenne en pourcentage de la RMSE par rapport à l'ensemble initial. L'évolution de la RMSE des concentrations est donnée en bleu foncée, celle des charges hydrauliques est donnée est cyan. La moyenne d'évolution globale de la RMSE est indiquée en rouge.

L'amélioration complète sur les données de calibration et de validation est donnée sur la Figure 4.11. En regardant l'ensemble des données simulées, le schéma ayant la meilleure réduction médiane de RMSE est le schéma 6 avec une réduction moyenne de 53.1%, suivi par les schémas 3 (45.8%) et 7 (41.8%).

Les méthodes d'assimilation puits par puits et type de donnée par type de donnée donnent de meilleurs résultats que les schémas d'assimilation brute. L'assimilation de la concentration seule avec l'analyse locale 25/5 m est le pire des schémas dans les 8 méthodes testées. Les résultats de l'assimilation avec une portée de 25/5 m indiquent que l'assimilation est trop limitée spatialement dans les modèles pour avoir un impact. L'assimilation de la concentration seulement pourrait manquer d'information à longue distance qui est contenue dans les charges hydrauliques.

#### 4.2.4 Erreur moyenne des champs de conductivité hydraulique

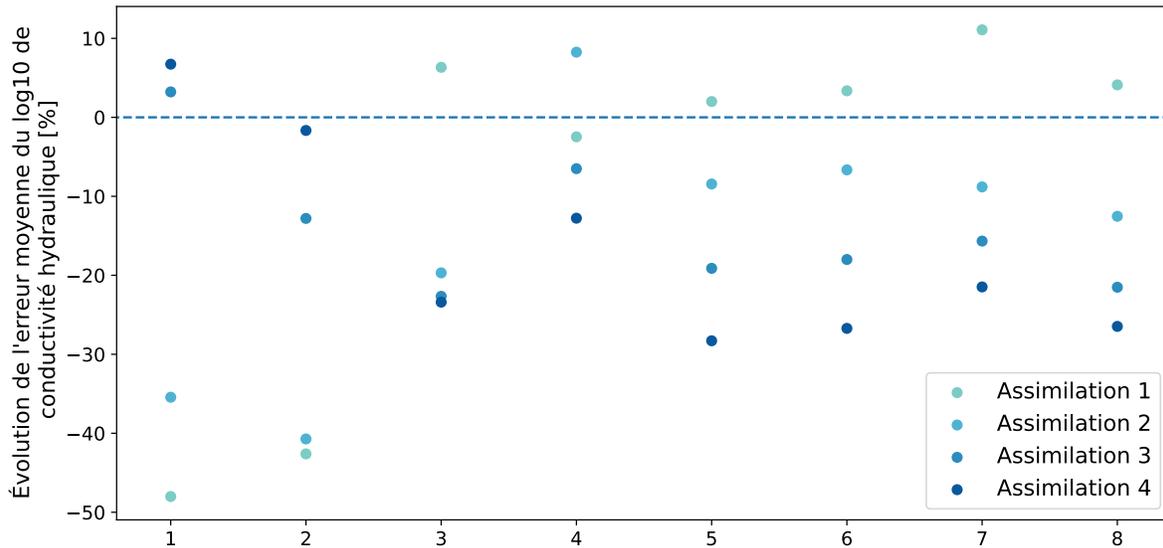


Figure 4.12 – Amélioration en pourcentage de l’erreur moyenne des champs de conductivité hydraulique pour les 8 méthodes à chaque itération d’assimilation par rapport au champ de référence.

Dans le cas d’un modèle synthétique, il est possible de comparer les modèles de conductivité hydraulique de l’ensemble tout au long du processus d’assimilation par rapport au modèle de référence. Dans la réalité, cette comparaison est impossible, le modèle réel du sous-sol pouvant seulement être inféré à partir de puits, essais de perméabilités ou mesures géophysiques. Pour reproduire ces complications dans le modèle synthétique, des biais d’échantillonnage ont été introduits en définissant uniquement les modèles à travers 5 puits, comme rappelés sur la Figure 4.1a. L’erreur moyenne sur la conductivité hydraulique s’améliore à chaque itération d’assimilation sur la Figure 4.12 dans le cas des assimilations puits par puits 3 à 8.

Les schémas 1 et 2 indiquent une meilleure performance pour la première itération d’assimilation. Cependant les conductivités hydraulique se dégradent par la suite à chaque nouvelle itération. De plus, ces améliorations ne se traduisent pas sur les RMSE de charge hydraulique et de concentration. On peut dénoter l’efficacité des nouvelles méthodes d’analyses locales, qui arrivent à mettre à jour la conductivité hydraulique de l’ensemble dans la direction du modèle de référence, tout en améliorant grandement les charges hydrauliques et les concentrations. Les schémas 3 et 6 permettent une réduction de l’erreur moyenne de conductivité hydraulique de 24% et de 27%.

En étudiant la RMSE sur l'ensemble des sections précédentes, il est possible de conclure que la méthode d'assimilation puits par puits et donnée par donnée permet d'améliorer la méthode d'assimilation brute proposée par (Emerick & Reynolds, 2013) sans analyse locale.

#### 4.2.5 Réalisations individuelles dans l'ensemble

Pour visualiser l'évolution des modèles individuels dans les ensembles, les simulations individuelles de concentrations et de charges hydrauliques initiales et finales dans le cadre de l'assimilation 100/20 m ont été affichées.

Les charges hydrauliques sur l'ensemble initial étaient très proches du modèle de référence, comme le montre la Figure 4.13. La moyenne de l'ensemble a été améliorée pendant le processus d'assimilation sur les puits *O0*, *O1*, *O2* et *C4*. La moyenne de l'ensemble au puits *C3* à quand a elle légèrement été dégradée. Dans tous les ensembles, une réduction de la variance a eu lieu entre l'ensemble initial et final.

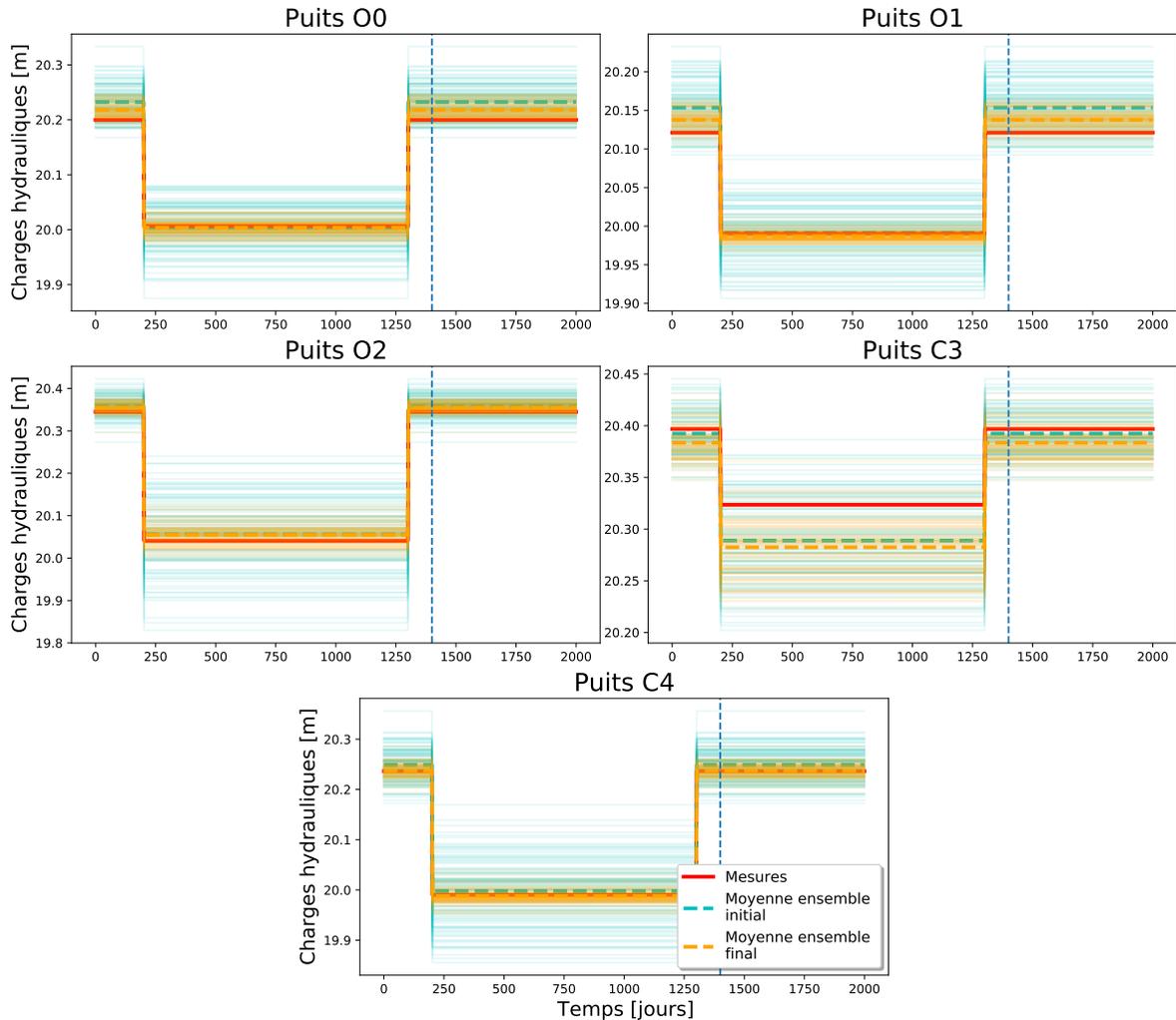


Figure 4.13 – Réalisations individuelles de charges hydrauliques en bleu pour l'ensemble initial, en orange pour l'ensemble final assimilé.

Pour les courbes de concentrations, l'échelle logarithmique a été tronquée dans la partie inférieure. Sans cela, la Figure 4.14 aurait été illisible à cause de quelques concentrations proches de 0 faisant tendre le logarithme vers  $-\infty$ . Le puits O0 montre le problème d'un ensemble initial trop contraint en ne contenant pas la concentration de référence. Pendant le processus d'assimilation, les membres de l'ensemble vont tenter de se rapprocher au maximum sans jamais y arriver. Les puits O1, O2 et C4 montrent des améliorations de l'ordre d'un à deux ordres de grandeur, que ce soit sur la moyenne ou sur la dispersion des individus. La Figure 4.14 permet de voir la capacité de la nouvelle méthode d'assimilation à gérer des données de valeurs très différentes. À nouveau, le puits

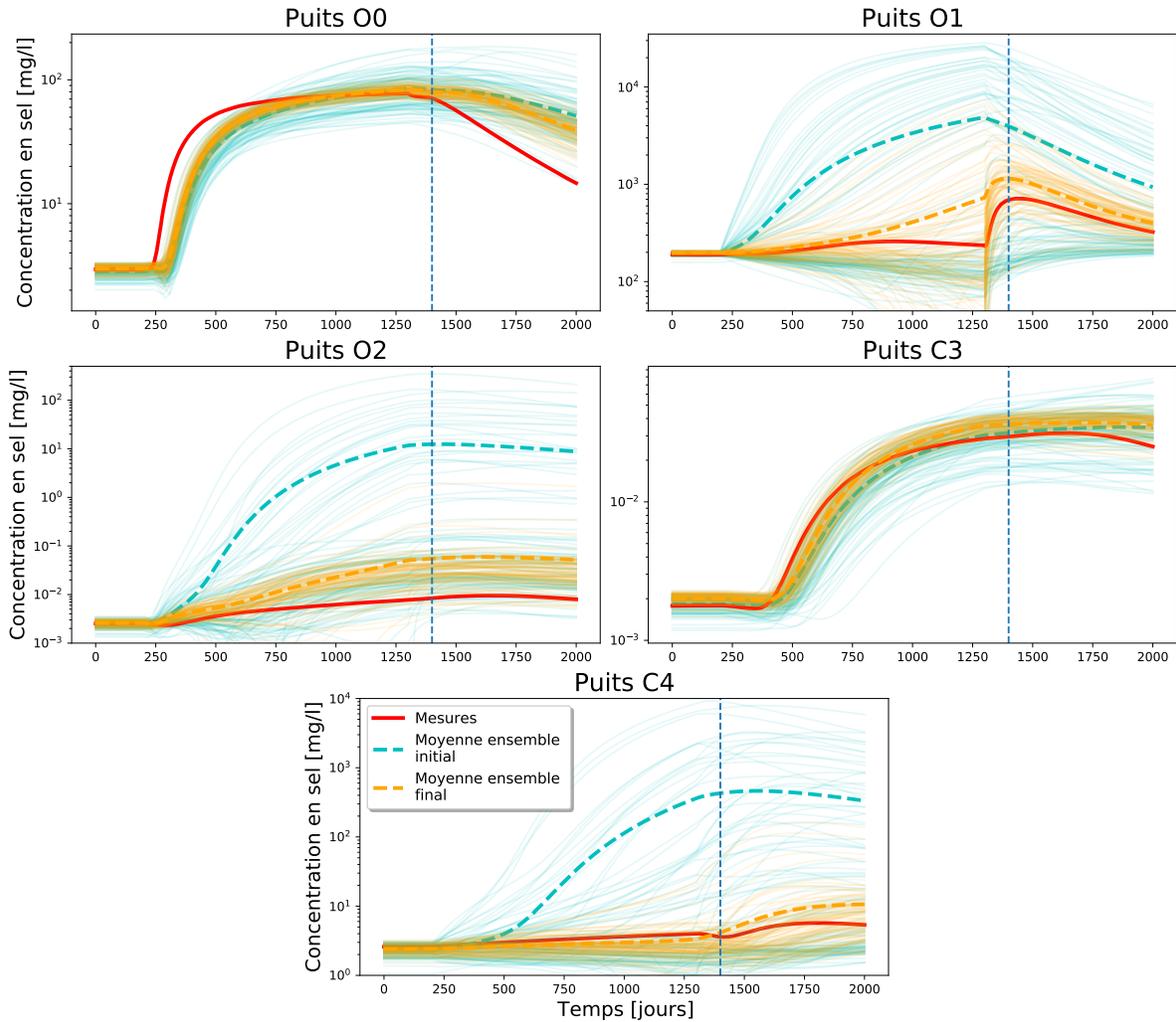


Figure 4.14 – Réalisations individuelles de concentrations en bleu pour l'ensemble initial, en orange pour l'ensemble final assimilé. La ligne verticale représente la séparation entre données de calibration et de validation.

*C3* montre une faible mise à jour, mais cela s'explique par un ensemble initial déjà très proche des concentrations de référence observées sur ce puits.

Les puits *O0* et *C3* étaient des puits de référence dans le modèle synthétique, leurs observations étaient déjà relativement proches de celle simulée dans le modèle de référence. L'amélioration moyenne était plus faible sur ces puits dans la Figure 4.11 à cause de cela. De plus, la variabilité réduite des modèles de l'ensemble autour de ces points n'a pas permis assez de variabilité pour l'exploration de l'espace des solutions. Ce qui a induit à ce qu'aucune réalisation individuelle ne corresponde aux valeurs recherchées sur ces deux puits comme le montre la Figure 4.14.

### 4.3 Discussion sur le modèle synthétique

Premièrement, la méthode proposée par Emerick & Reynolds (2013) a été capable d'améliorer les prédictions de l'ensemble et de réduire leurs variances dans le cas synthétique d'aquifère côtier. Utiliser un ES était plus rapide que de relancer de multiple EnKF en redémarrant FEFLOW à chaque pas de temps de simulation comme démontré par Skjervheim *et al.* (2011).

Toutes les interprétations numériques ont été réalisées sur les RMSE et les médianes de RMSE. Ces métriques sont sensibles aux valeurs extrêmes. Il faut donc également veiller à ce que la moyenne des ensembles soit correctement mise à jour, et non pas uniquement quelques modèles extrêmes. Cette confirmation visuelle a été obtenue sur les Figures 4.14 et 4.13. Il est possible de voir que la moyenne tend vers les mesures, en même temps que les modèles extrêmes sont ramenés vers la moyenne.

L'assimilation brute de toutes les données sans analyse locale ou localisation avait des problèmes pour gérer les données d'ordres de grandeur différentes. L'assimilation brute effaçait une partie de l'information provenant des données d'observation. L'analyse locale dans le cas synthétique améliore les modèles individuels de l'ensemble, comme décrit par Hunt *et al.* (2007), Sakov & Bertino (2011), Carrassi *et al.* (2018) et de nombreux autres auteurs. Avec des modèles constitués de nombreux paramètres d'états, le nombre de sous-domaines de localisation va définir trop important comme noté par Carrassi *et al.* (2018). Une manière d'éviter ce problème est donc de mettre à jour les paramètres des modèles dans tous les éléments autour d'un point d'observation. De cette manière-là, le nombre d'analyses locales à effectuer sera limité au nombre de points d'observations dans le modèle. Ce nombre de points d'observations sera toujours plus faible que le nombre de paramètres d'états dans le cadre de géomodèles complexes.

Les résultats ont montré que le pourcentage d'amélioration était plus faible sur les puits *O0* et *C3*. Ces points étant des puits de référence dans le modèle initial, les données simulées étaient déjà proches des données de référence. Une étude supplémentaire pourrait être conduite pour évaluer l'impact d'ajouter un bruit dans l'ellipsoïde d'analyse locale afin d'éviter l'effondrement de la variance à proximité de points de mesures. Les auteurs Evensen (2009a) et Carrassi *et al.* (2018) détaillent les effets de l'effondrement de la variance dans les ensembles.

L'amélioration moyenne sur des les données de calibration est proche entre les mises à jour globales et celles issues d'analyses locales. Ce comportement est normal lorsque les portées d'analyses locales sont grandes et englobent quasiment toutes les points d'informations. Les différences sont plus visibles sur les données de validation, ces dernières n'ayant pas été assimilées. Dans ce cas, les méthodes d'analyses locales améliorent la moyenne. La difficulté est donc de choisir la bonne portée d'analyse locale, comme le notait Chen & Oliver (2017). Lorsque la portée est trop faible, trop peu d'éléments peuvent être mis à jour, et les modèles n'arrivent pas à coller aux données de charges hydrauliques et de concentration, comme le montre la méthode 25/5 m. L'assimilation des charges hydrauliques seules est encourageante, due à la facilité d'obtenir ces informations par rapport à des concentrations.

Mais si plusieurs types de données sont disponibles, l'utilisation de toutes ces données améliorera les résultats de l'assimilation comme le montre les Figures 4.11 et 4.12.

Au final, les ensembles assimilés avec des analyses locales ont conservé de l'incertitude sans effondrement de la variance tout en améliorant les prédictions de charges hydrauliques, de concentration et en améliorant les champs de perméabilités.

Tous ces résultats ont validé la méthode développée pour un cas synthétique simplifié. Sans assimilation puits par puits, les réductions médianes de RMSE atteignaient environ 10 à 15%. Avec assimilation puits par puits et analyse locale cette réduction atteint 53.1%, et une meilleure mise à jour du champ de conductivité hydraulique.

La prochaine étape était donc de tester la méthode d'ES-MDA 4x avec analyse locale sur un cas réel. La prochaine section de ce document sera donc dédiée à la présentation du cas réel et aux résultats de la méthode sur ce dernier.

## Chapitre 5

# Modèle Réel : Les lagunes de Mercier

Pour tester la méthode, il a été possible d'accéder à un jeu de données complet pour une problématique réelle et d'actualité. Les anciennes lagunes de Mercier font l'objet d'une caractérisation extensive depuis plus de 40 ans pour suivre l'évolution de la contamination (BAPE, 1994). Pour présenter ce jeu de données, une description de la géologie du site va être réalisée. Dans un second temps, les choix pour la création du modèle conceptuel seront détaillés. Enfin, les résultats obtenus avec la méthode d'assimilation seront présentés. Les données et le modèle utilisés dans le cadre de ce projet reposent en partie sur les travaux de Claprood *et al.* (2022).

### 5.1 Objectifs du modèle

Actuellement, l'usine de traitement des eaux souterraines (UTES) pompe et traite les eaux impactées des anciennes lagunes.

Des nombreuses incertitudes sont présentes sous les lagunes, au droit des zones de connexion entre l'esker et le roc fracturé. L'esker et le roc ont des perméabilités variables en fonction des processus de déposition et de fracturation respectivement.

Dans un mandat d'évaluation des possibles méthodes de gestion de l'aquifère, le Ministère de l'Environnement et de la Lutte Contre les Changements Climatiques (MELCC) a mis à disposition de l'INRS des nombreuses données. Ces données étaient partiellement disponibles dans les

publications de Denis (1991), Poulin (1977), Pontlevoy *et al.* (2004), Lefebvre (2007) et Nastev & Lamontagne (2010). Elles contiennent des essais de pompage, des mesures de perméabilités, des données géophysiques, de nombreux forages, etc.

Les objectifs des modèles sont doubles. Dans un premier temps, le but est de construire un ensemble de modèles pour évaluer les incertitudes associées aux champs de perméabilités de l'esker et du roc. Dans un second temps, différents schémas d'acquisition seront testés pour améliorer cet ensemble en comparant les mesures d'essais de pompages avec les observations effectuées dans les modèles de l'ensemble.

Ces ensembles de modèles pourront être utilisés dans le futur pour évaluer différents scénarios de gestion du site, ainsi que les incertitudes de chaque scénario.

## 5.2 Présentation du site

### 5.2.1 Localisation du site

Les données utilisées lors de ce doctorat proviennent d'un site d'étude localisé au sud-est du Canada, au sud de la province de Québec et à proximité de la frontière américaine. Le terrain se situe dans les Basses Terres du Saint-Laurent à environ 20 km au sud-ouest de Montréal. Une carte de localisation est donnée sur la Figure 5.1. Les anciennes lagunes de la ville de Mercier, qui sont la source d'une ancienne contamination, sont situées sur un esker sablo-graveleux. C'est cette particularité géologique qui a permis l'écoulement de contaminants dans les eaux souterraines en profondeur. Une synthèse de la géologie régionale et locale est donnée dans les deux prochains paragraphes.

La zone étudiée est délimitée hydrauliquement par le fleuve Saint-Laurent au nord, la rivière Châteauguay à l'ouest et la rivière Esturgeon au sud.

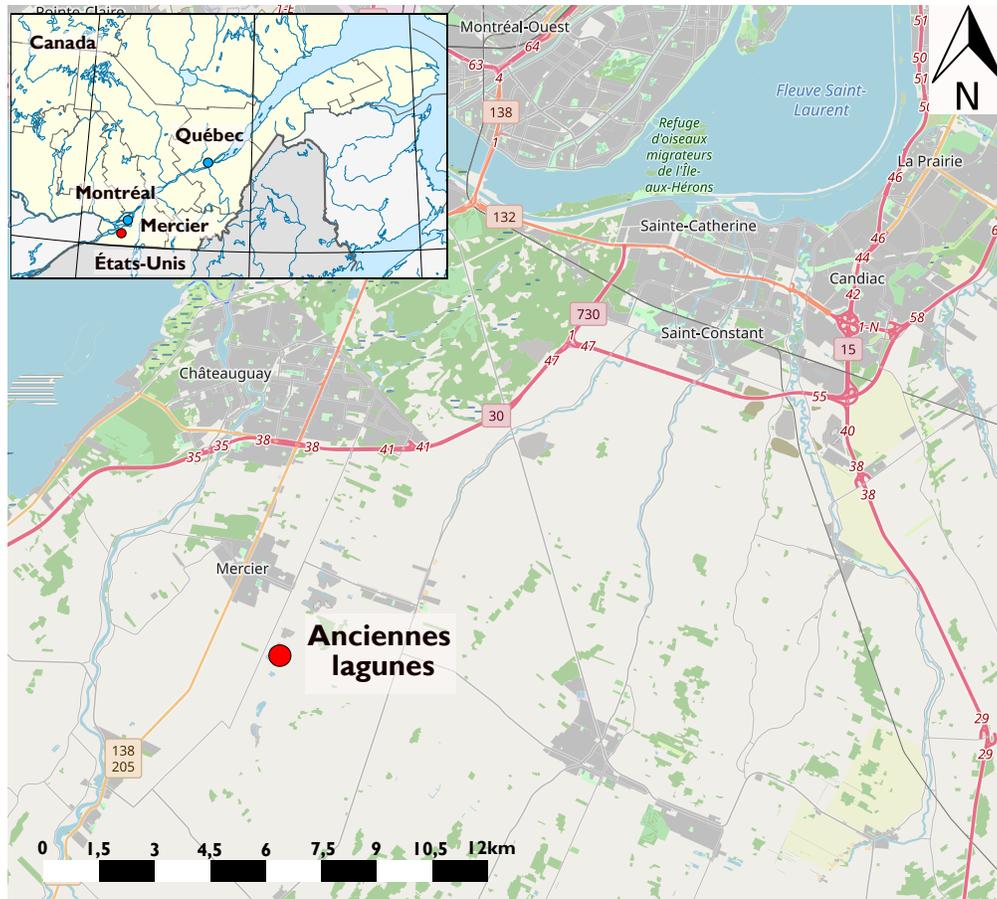


Figure 5.1 – Localisation des anciennes lagunes de Mercier. Fond de carte OpenStreetMap du sud de Montréal distribué sous licence CC BY-SA. Fond de carte de l'Amérique du Nord sous License Creative Commons réalisé par Q-rioux.

### 5.2.2 Contexte géologique

Au niveau régional, la plateforme du Saint-Laurent repose en discordance sur les roches du bouclier de Grenville datant du précambrien comme indiqué sur la Figure 5.2. Les Appalaches chevauchent la plateforme du Saint-Laurent par l'intermédiaire de la faille Logan (Brun & Brisebois, 1994). La série stratigraphique complète des roches sédimentaires des Basses Terres est de l'ordre de 1500 m à 3000 m d'épaisseur. Cela correspond à un cycle complet de transgression-régression. Les formations de la base et du sommet sont d'origine continentale tandis que le coeur de la pile sédimentaire est d'origine marine (Nastev & Lamontagne, 2010).

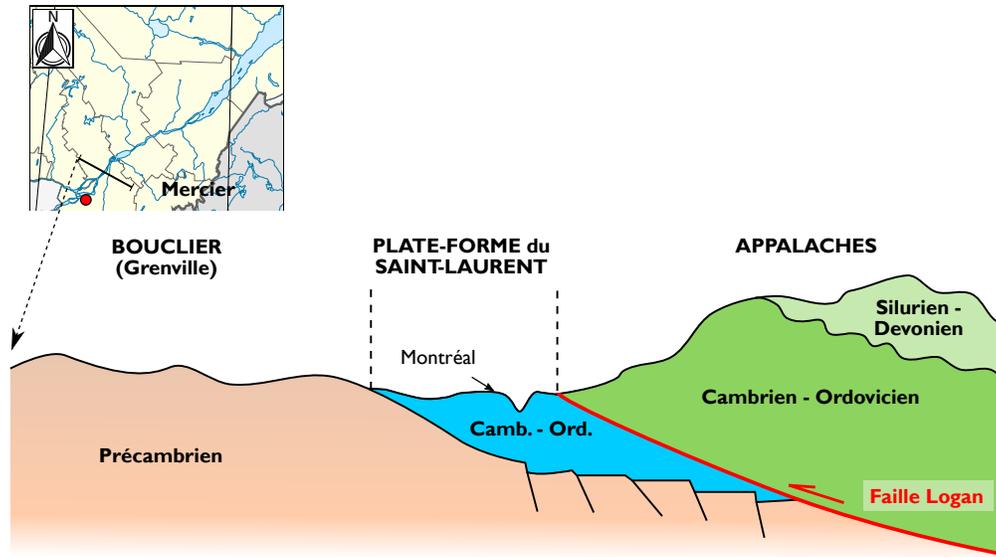


Figure 5.2 – Coupe schématique régionale simplifiée de la géologie du Bas-Saint-Laurent. La coupe suit un axe nord-ouest sud-est comme précisé dans l’encart supérieur gauche de la figure. Modifié de [http://www2.ggl.ulaval.ca/personnel/bourque/intro.pt/planete\\_terre.html](http://www2.ggl.ulaval.ca/personnel/bourque/intro.pt/planete_terre.html) (12/03/2019) et Brun & Brisebois (1994).

L’analyse de la carte géologique de la région de Mercier a permis de mettre en évidence les grandes structures géologiques. Le substratum rocheux est constitué des grès de Potsdam du Cambrien supérieur et des grès et dolomies du Beekmantown de l’Ordovicien inférieur. Le substratum rocheux en dessous de Mercier est parcouru par le synclinal de Sainte-Marie d’orientation générale sud-ouest nord-est. L’anticlinal d’Aubrey est présent à environ 6 km à l’est de Mercier comme indiqué sur la carte géologique en Figure 5.3. L’anticlinal majeur d’Oka-Beauharnois suit globalement une orientation nord-ouest sud-est, avec un cœur laissant apparaître des grès de Covey Hill du Cambrien. La faille normale de Havelock d’orientation nord-est sud-ouest se situe à l’est de Mercier à une dizaine de kilomètres, cette dernière ne devrait quasiment pas impacter l’écoulement souterrain des eaux. Deux autres familles de failles sont définies par Globensky (1986), un système est-ouest traduit par les failles de Delson, Dowler et Sainte-Anne-de-Bellevue et un système nord-ouest sud-est représenté par la faille de Saint-Régis.

Le sommet de l’anticlinal d’Aubrey se trouvant entre Mercier et la faille devrait agir comme une ligne de séparation des eaux souterraines. L’ensemble des roches sédimentaires cambro-ordoviciennes repose en concordance les unes sur les autres avec un léger pendage régional vers le nord-est (Nastev

& Lamontagne, 2010). La direction des couches serait comprise entre N10° et N125° avec un pendage relativement faible compris entre 1° et 5° vers le sud-est et le nord-ouest (Denis, 1991).

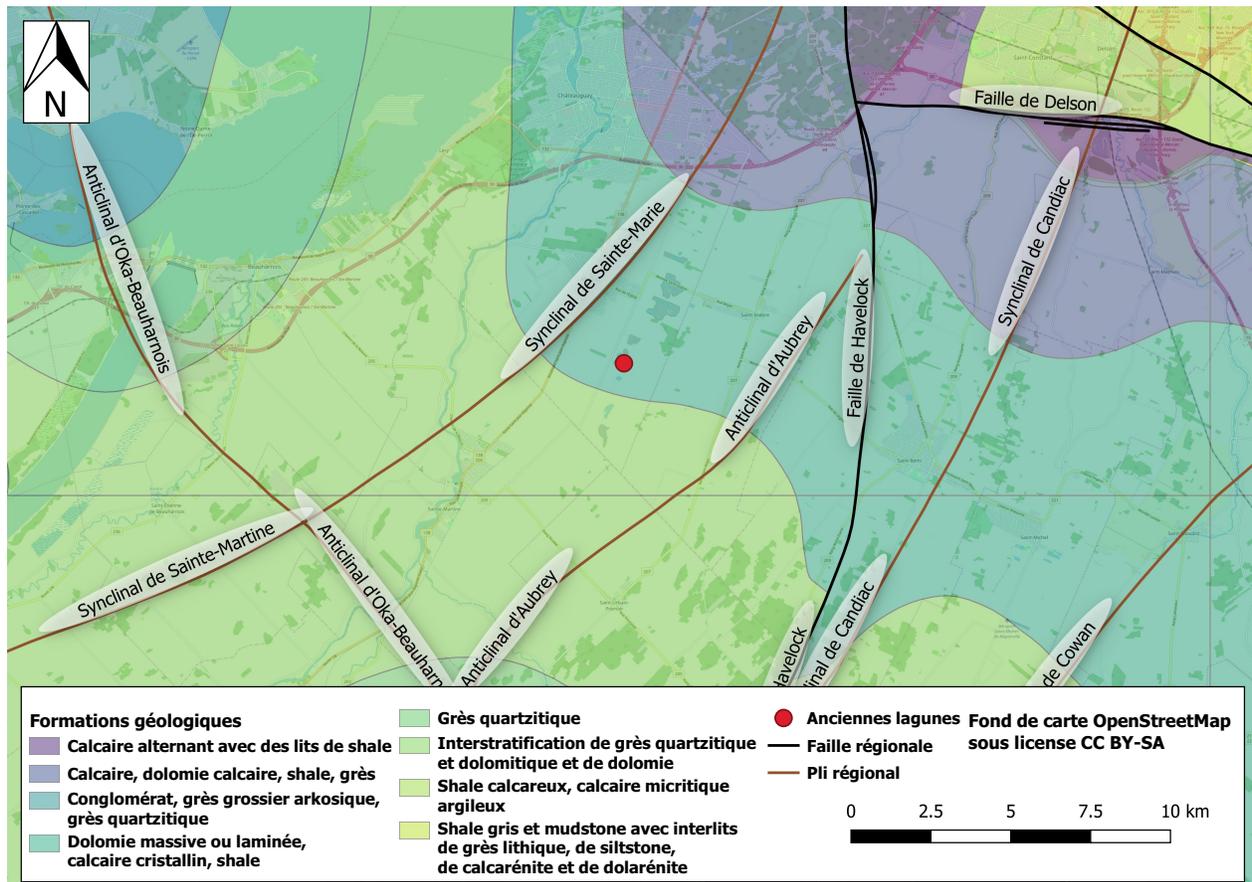


Figure 5.3 – Carte géologique et structurale du substratum rocheux extraite de la base de données du Système d’Information GÉOMinière du Québec (SIGEOM).

La Figure 5.4 résume l’histoire géologique locale. Le till, qui est en contact direct avec le roc à l’exception de quelques zones d’érosion, est le résultat de l’érosion glaciaire du Wisconsinien supérieur sur les roches Cambrio-Ordovicienne. Le glacier s’est déplacé vers le sud/sud-ouest, laissant des drumlins orientés dans cette direction à l’est des lagunes de Mercier. Certains drumlins ont été façonnés et réorientés lors des écoulements glaciaires de la déglaciation.

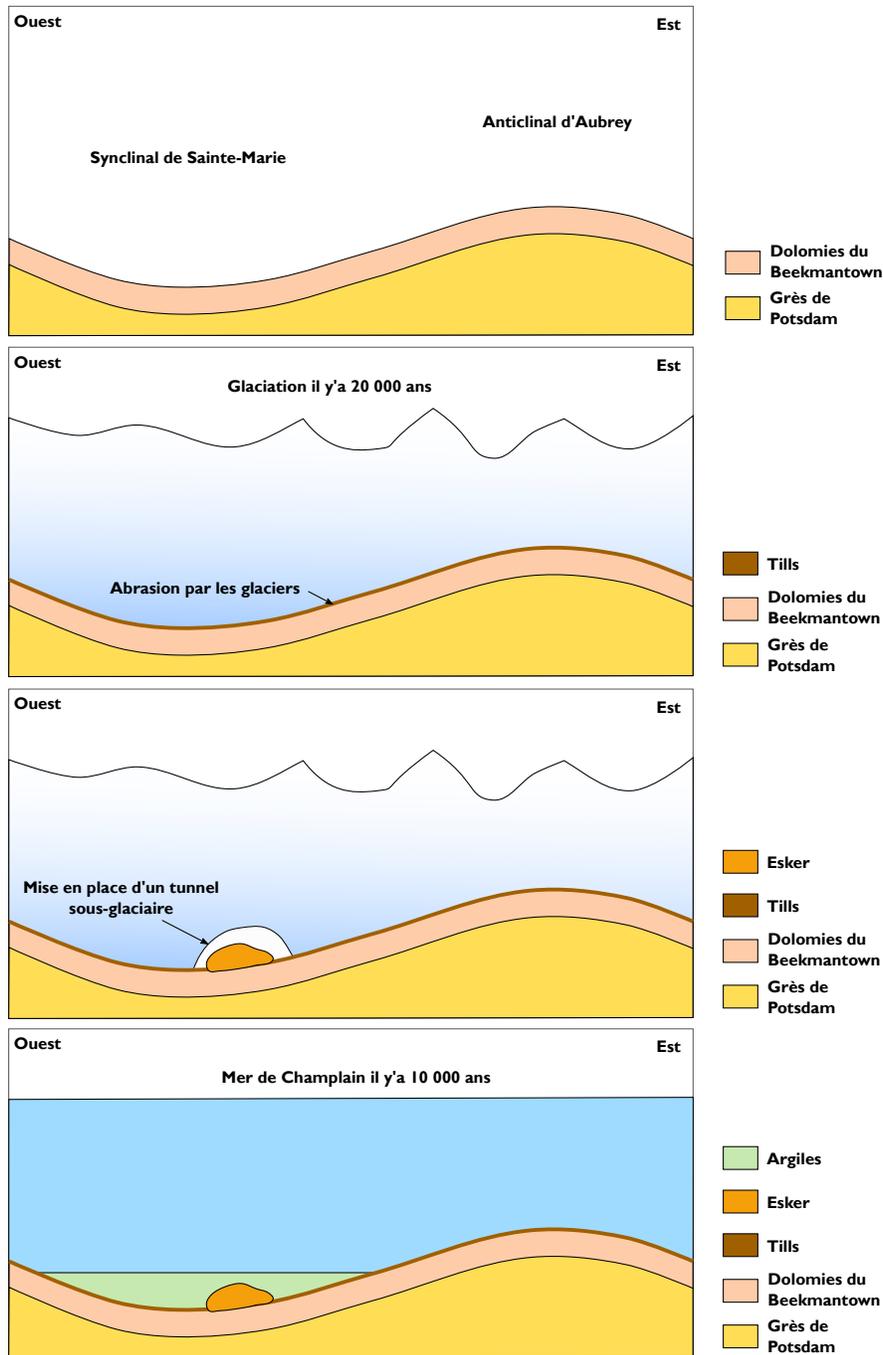


Figure 5.4 – Schéma de l'histoire géologique quaternaire simplifiée sous la ville de Mercier. Encarts dans l'ordre chronologique du haut vers le bas. Substratum rocheux Cambrio-Ordovicien, mise en place des glaciers et naissance du till par abrasion du roc, apparition des dépôts fluvio-glaciaires, fonte des glaciers et déposition d'argiles dans la Mer de Champlain.

Dans la phase de déglaciation, des dépôts fluvio-glaciaires ont été mis en place en conformité avec les tills sous-jacents. Dans l'esker, la séquence basale est constituée de sédiments sablo-graveleux

d'origine sous-glaciaire. Après le retrait des glaciers et l'invasion de la mer de Champlain, les sommets des drumlins et les rivages exposés à Covey Hill ont été lessivés par les vagues. Les tills et les silts remobilisés se sont ainsi déposés dans les bas topographiques en dessous de la limite d'action des vagues, créant les dépôts donnant les argiles de Champlain (Nastev & Lamontagne, 2010). Ces argiles de la Mer de Champlain sont présentes dans quasiment toute la région dans les bas topographiques en dessous de 50 m a.s.l. (above sea level). Toute la pile quaternaire a donc été modelée par les mouvements glaciaires ainsi que par la sédimentation marine et fluviale. La fin de déglaciation et le début de l'épisode de la Mer de Champlain ont eu lieu il y'a environ 11100 ans (Tremblay, 2008). Le niveau marin aurait atteint 160 m a.s.l. Une preuve de cette phase est la déposition de sédiments glaciolacustres et marins.

L'étude sismique menée par Tremblay *et al.* (2010) le long de l'esker de Mercier a révélé que la profondeur des sédiments meubles jusqu'au toit du roc est supérieure à 20 m et que des argiles marines saturées sont présentes en surfaces sur la quasi-totalité des lignes sismiques. Nastev & Lamontagne (2010) précise que l'épaisseur des sédiments meubles pouvait atteindre 45 m au coeur de la vallée de Châteauguay.

### 5.2.3 Carte géologique

Une carte géologique des formations du quaternaire a été réalisée par Tremblay (2008). Cette carte modifiée sur la Figure 5.5 permet de visualiser l'étendue de l'esker sous les anciennes lagunes. Comme précisée dans les parties précédentes, l'esker s'étend du nord-est vers le sud-ouest sur environ 11 km. De nombreuses gravières ont exploité les sables et graviers de l'esker. Elles se répartissent tout du long de la crête fluvio-glaciaire. Des drumlins sont également présents sur cette carte dans la même direction générale que l'esker. Il est également intéressant de constater la couverture de la région par des formations plutôt imperméables (silts, argiles et tills). La recharge en eaux souterraines autour de la ville de Mercier passera donc préférentiellement à travers les sables et graviers de l'esker comme les modélisations de Pontlevoy (2004), et Croteau *et al.* (2010) le confirment.

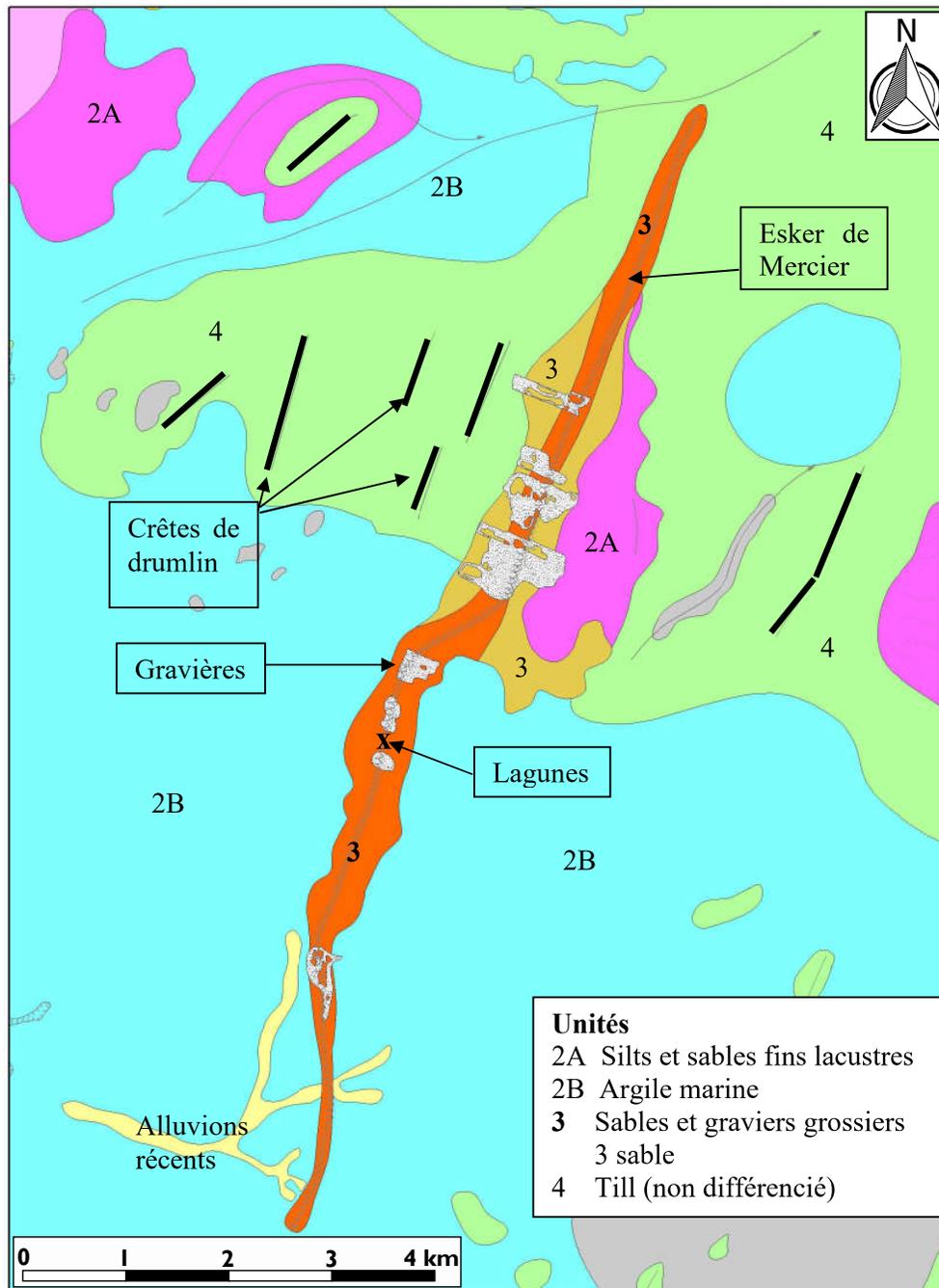


Figure 5.5 – Carte géologique des dépôts meubles du quaternaire autour des anciennes lagunes de Mercier. Modifié d'après Tremblay (2008).

### 5.2.4 Lignes sismiques

Les travaux de Tremblay *et al.* (2010) ont confirmés la présence et la forme de l'esker sous les lagunes, Figure 5.6. Une acquisition plus récente, présentée dans l'article de Claprood *et al.* (2022) indiquait la présence d'un esker secondaire au sud-est de l'esker principal.

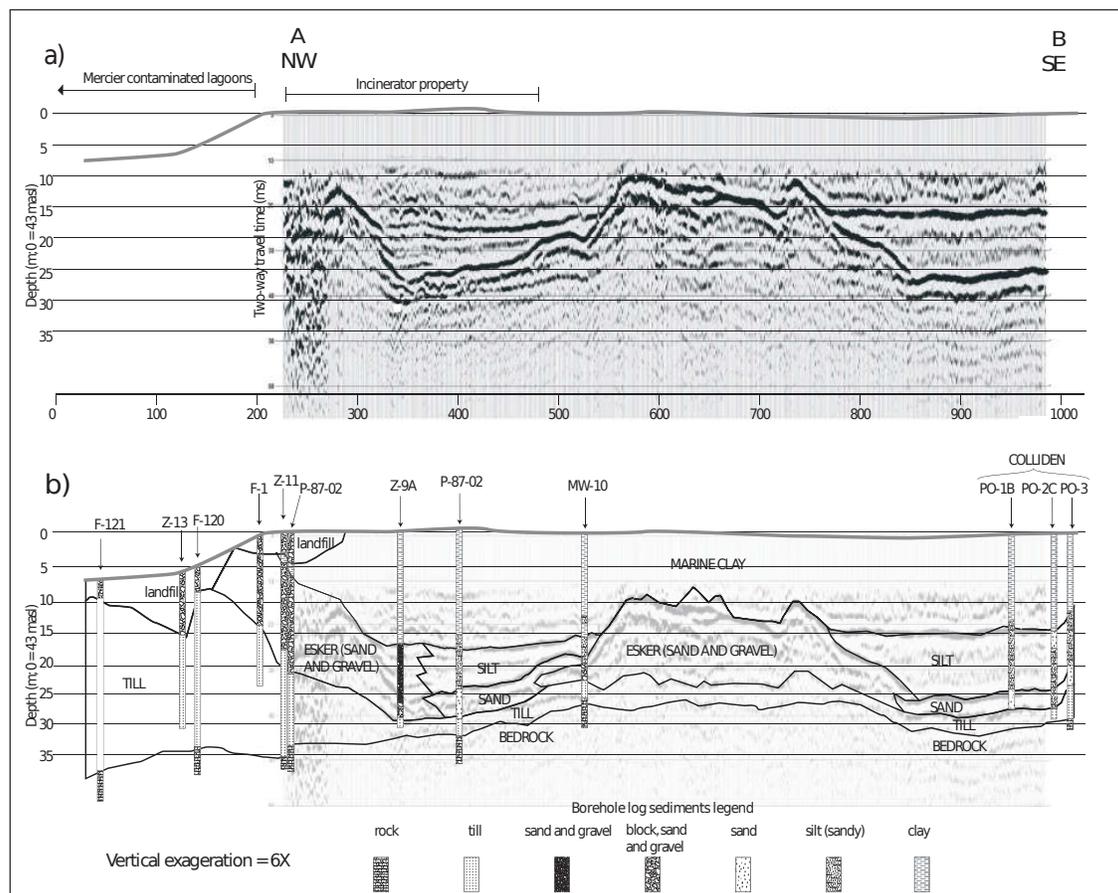


Figure 5.6 – a) Lignes de sismiques réalisées sous les anciennes lagunes de Mercer et l'incinérateur selon un profil quasiment orienté ouest-est; b) Interprétation géologique du profil sismique contraint avec 12 puits. Les deux profils sont issus des travaux de Tremblay *et al.* (2010).

La nouvelle campagne d'acquisition détaillée par Claprood *et al.* (2022) a permis de définir un modèle 3D des horizons géologiques majeurs au droit des lagunes de Mercier. Le traitement géophysique des 50 km de lignes sismiques acquises est précisé par Claprood *et al.* (2022). Le modèle interprété est donné sur la Figure 5.7. L'horizon cyan a été interprété comme étant le toit des argiles marines, le jaune correspondrait à une extension latérale sablo-silteuse et graveleuse de l'esker. Le marqueur orange indiquerait l'esker principal constitué majoritairement de sables et graviers. Enfin,

l'horizon rouge traduirait le haut du roc ou du till compacté sous la pile de sédiments quaternaire. L'interprétation des horizons par les géophysiciens a été appuyée par de nombreux forages présents dans la région d'intérêt.

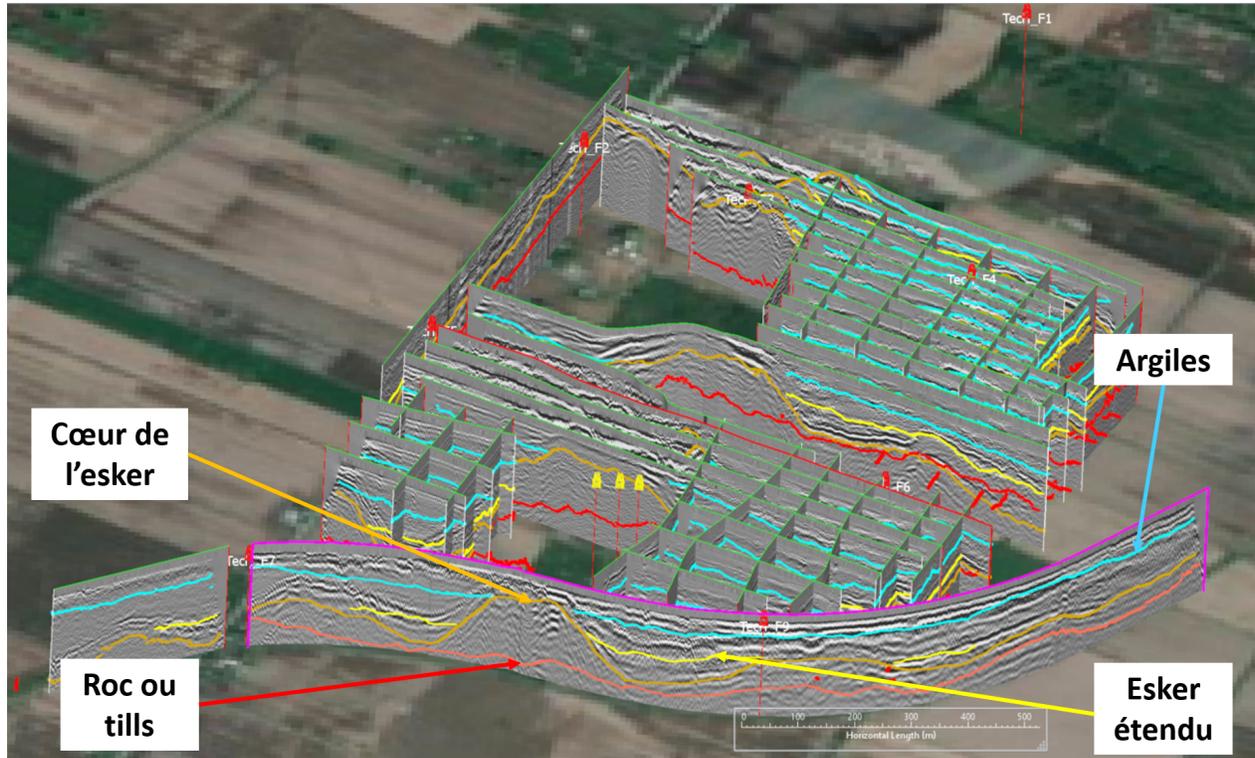


Figure 5.7 – Horizons sismiques interprétés lors de la campagne d'acquisition de 2019 sur le site de Mercier. Exagération verticale 4x.

### 5.3 Modèle 3D des lagunes de Mercier

Dans le cadre de ce doctorat, l'objectif ne sera pas de créer un modèle hydrogéologique à partir de zéro, mais d'améliorer un modèle tout en évaluant les incertitudes. Conventionnellement un seul modèle hydrogéologique est créé et ajusté sur les données mesurées ce qui rend l'évaluation des incertitudes très complexe. Le site de Mercier étudié a fait l'objet d'une caractérisation environnementale extensive avec près de 200 forages, 300 essais de perméabilités, des études géophysiques et des études de porosité. L'ensemble de ces informations vient d'être synthétisé et mis en forme pour servir de données conditionnantes dans les futures modélisations stochastiques. Cette synthèse de données est reprise en partie des travaux de Claproud *et al.* (2022). Les modèles conceptuel et hy-

drogéologique sont également issus de ces travaux, ils serviront de bases à la création de l'ensemble des modèles.

### 5.3.1 Modèle conceptuel

Le modèle conceptuel hydrogéologique, qui représente l'idée du sous-sol obtenue à partir des connaissances disponibles, a été bâti en se basant sur de précédents modèles définis entre autres par Pontlevoxy *et al.* (2004), Lefebvre (2007) et Claprood *et al.* (2022). Deux inconnues principales devront être étudiées grâce aux modèles à cause de leurs impacts potentiels sur l'écoulement des eaux souterraines:

- La position et l'emprise de l'esker;
- La disposition des zones de contacts entre le roc fracturé et l'esker.

Les anciennes lagunes de Mercier reposent sur un système aquifère constitué d'un roc fracturé recouvert par deux unités successives de till. Des sables et graviers d'origine fluvioglaciaire déposés par-dessus les tills constituent un deuxième aquifère superficiel de type esker, faisant quelques centaines de mètres de largeur et quelques kilomètres de longueur.

Bien que le site des anciennes lagunes de Mercier ne fasse que quelques centaines de mètres de longueur, il faut intégrer tous les éléments régionaux pouvant influencer l'écoulement des eaux souterraines. La limite Ouest du modèle hydrogéologique est la rivière Châteauguay qui draine la majorité des eaux de surface de la région, et qui est en contact hydraulique avec l'aquifère du roc fracturé au coin sud-ouest du modèle. La limite Nord du modèle est la limite du bassin versant de la Rivière Châteauguay, tandis que la limite Sud est la rivière des Fèves. La limite Est du modèle est la ligne approximative de partage des eaux selon les élévations du roc. La limite supérieure du modèle est la surface topographique, tandis que la limite inférieure du modèle est fixée à 40 mètres sous le toit du roc, afin de représenter l'écoulement des eaux souterraines dans le roc fracturé.

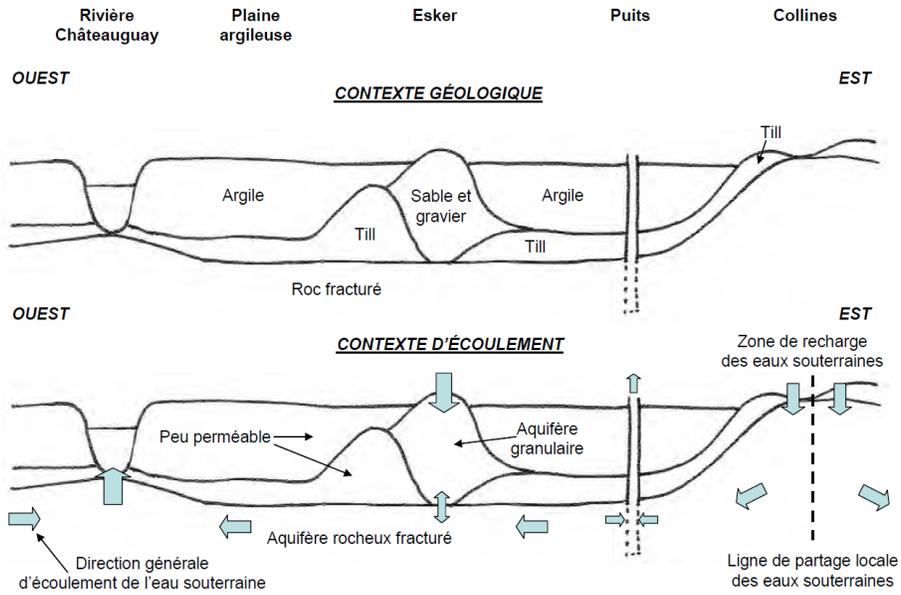


Figure 3.2. Illustration schématique du contexte régional du site des anciennes lagunes de Mercier (Voir le tableau 3.1 résumant les caractéristiques des composantes du contexte régional)

Figure 5.8 – Schéma du contexte régional du site des anciennes lagunes de Mercier ((Lefebvre, 2007)).

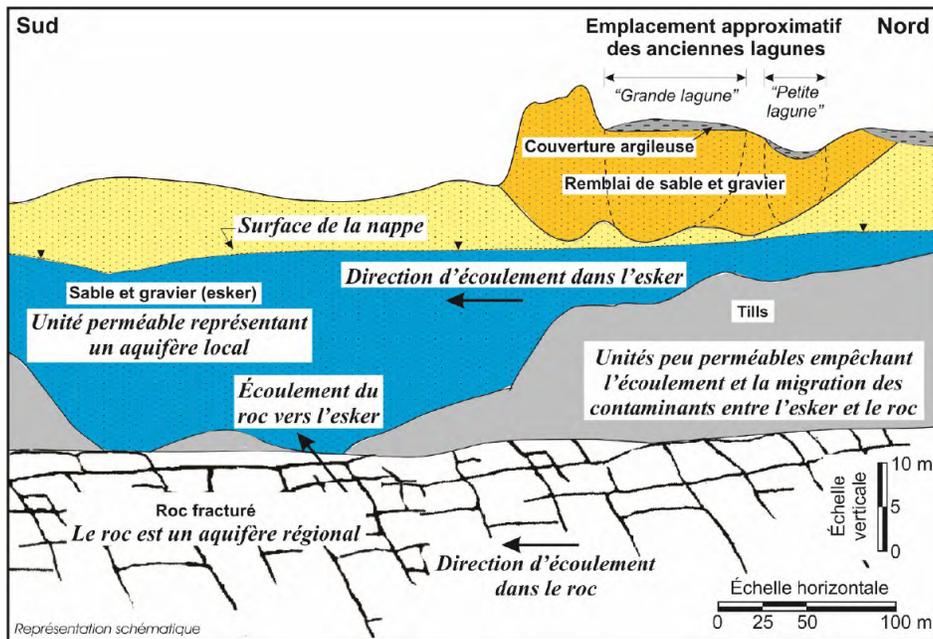


Figure 3.4a. Modèle conceptuel du site des anciennes lagunes de Mercier – Unités géologiques et écoulement

Figure 5.9 – Modèle conceptuel du site des anciennes lagunes de Mercier représentant les unités géologiques et l'écoulement de l'eau souterraine ((Lefebvre, 2007)).

La synthèse des publications précédentes et des données géologiques disponibles a révélé un milieu très hétérogène, avec des fenêtres d'écoulement localisées entre l'esker et le roc. Ces potentielles fenêtres d'écoulements ont été schématisées par Lefebvre (2007) dans ses travaux et reprises sur la Figure 5.9.

### 5.3.2 Modèle géostatistique

La variabilité dans les fenêtres entre les faciès perméables de l'esker et les parties fracturées du roc nécessite une étude poussée. L'étude de la continuité de l'horizon de till ou de son érosion sera primordiale pour la simulation de l'écoulement des eaux souterraines. Différentes méthodes étaient considérées pour définir les contacts entre l'aquifère semi-captif du roc et l'aquifère libre de l'esker. Ces contacts sont primordiaux pour définir les chemins d'écoulement des contaminants. Une première approche implicite en faisant varier les poids des différents paramètres de manières stochastiques en utilisant une approche dérivée de GemPy (De la Varga *et al.* (2019) basée sur les travaux de Calcagno *et al.* (2008) et Caumon *et al.* (2013)) a été envisagée. Cette approche permettrait de créer physiquement différentes surfaces s'intersectant pour représenter la variabilité d'épaisseur de till entre le roc et l'esker. La deuxième approche, reposerait sur une variation des propriétés de conductivité hydraulique dans la zone d'incertitude pour simuler numériquement des fenêtres de contact. La deuxième méthode présentait l'avantage de nécessiter un seul maillage, évitant la complexité de réaliser de nombreux modèles dans FEFLOW. Elle permettait également à l'utilisateur de faire varier son nombre de modèles équiprobables sans avoir à s'inquiéter du temps pour les porter sous FEFLOW. La deuxième méthode a donc été retenue.

La même réflexion a été tenue pour les autres horizons géologiques. Un krigeage pour les surfaces géologiques a été choisi et des simulations stochastiques pour peupler les propriétés pétrophysiques de chacune des unités. Le degré de confiance dans chacune des réalisations sera traduit par les paramètres des variogrammes choisis pour la modélisation géostatistique. La portée des variogrammes donnera la distance de corrélation entre les données et le plateau indiquera la variance entre les données non corrélées (Goovaerts, 1997). Enfin, l'effet pépite indiquera d'éventuelles erreurs de mesures ou bien des variations à plus courte distance que le pas d'échantillonnage (Journel & Huijbregts, 1978).

Une interpolation des marqueurs sismiques grâce à la méthode Discrete Smooth Interpolation (DSI) dans le géomodeleur GOCAD Mallet (1992) a été réalisée pour avoir une contrainte spatiale continue. Ces surfaces sismiques ont servi de dérive externe pour les krigeage d'horizons identifiés sur les profils sismiques (roc et esker principalement) (Galli & Meunier, 1987).

Le modèle 3D régional est issu des travaux de Claprood *et al.* (2022). Il combine le modèle conceptuel et toutes les données disponibles à ce jour. Le modèle des Figures 5.10 et 5.11 a été bâti grâce au géomodeleur GOCAD. Le modèle est constitué de 85 millions de cellules de taille respective 25 m x 25 m x 1 m. Le modèle régional obtenu autour des lagunes de Mercier correspond au modèle local fourni par le MELCC.

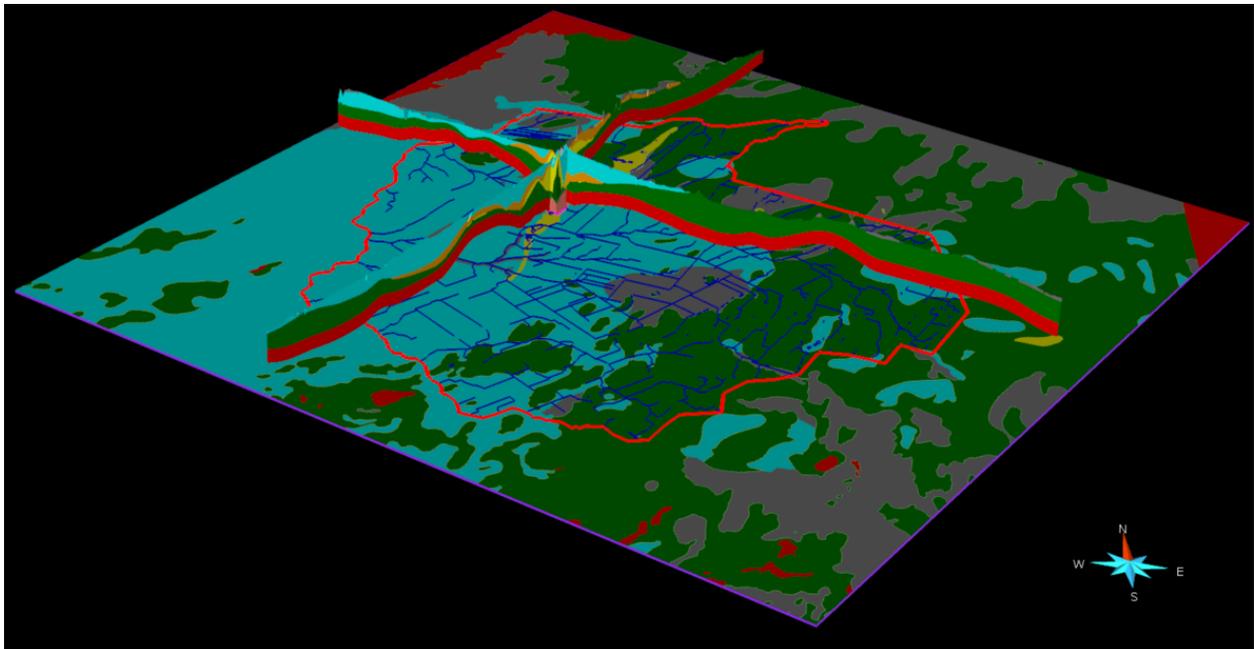


Figure 5.10 – Vue 3D des unités hydrostratigraphiques dans le modèle conceptuel régional. Exagération verticale de la vue régionale de 40x, exagération verticale de la vue locale de 20x. La carte de fond est la géologie de surface, avec le réseau hydrologique. En rouge le roc, en vert les tills, en jaune le coeur de l'esker, en orange les flancs de l'esker et en cyan les argiles marines.

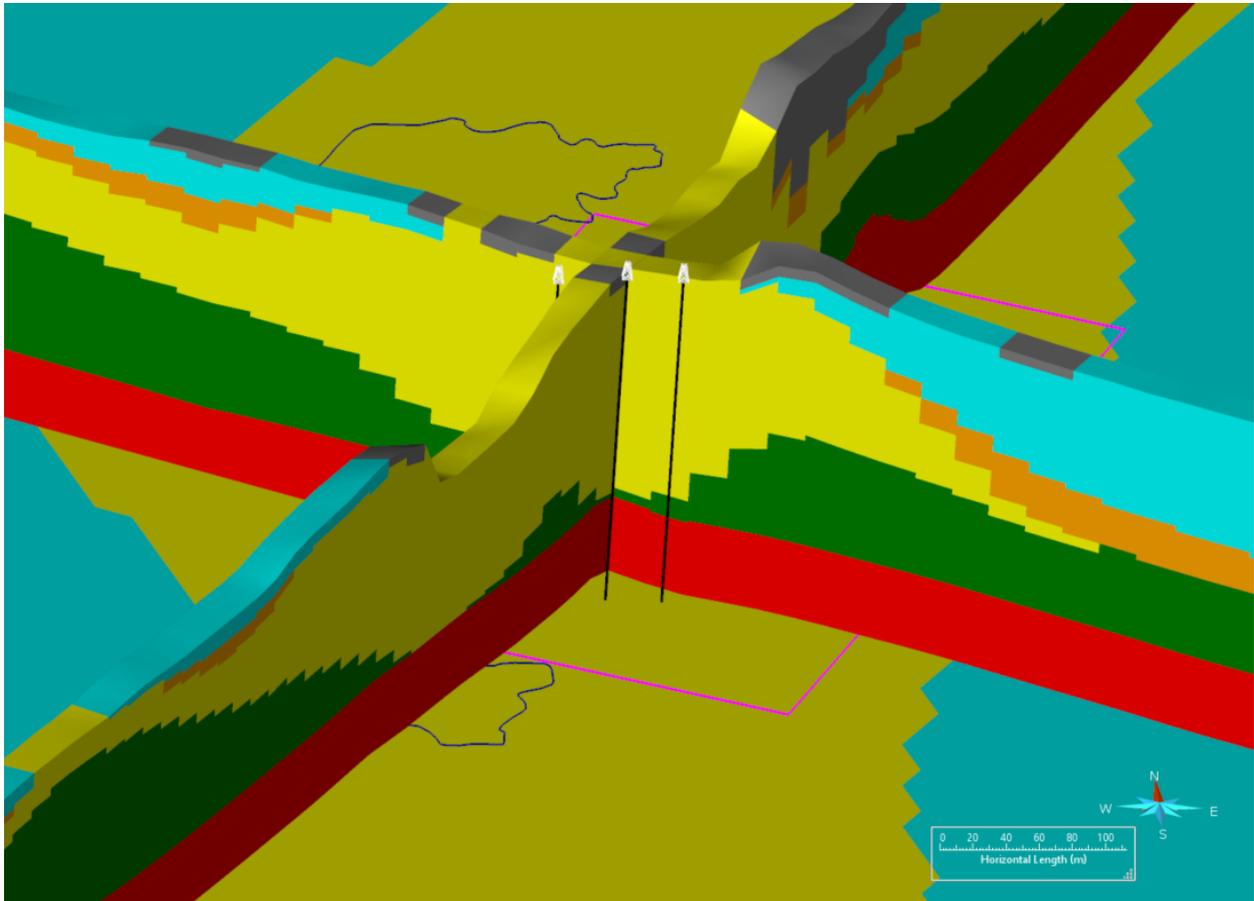


Figure 5.11 – Vue 3D des unités hydrostratigraphiques au niveau du site de l'UTES, avec une exagération verticale de 5. La carte de fond est la géologie de surface. Les puits de l'UTES PW-113, PW-112, et PW-111 sont présentés en blanc et noir de gauche à droite, respectivement. En rouge le roc, en vert les tills, en jaune le coeur de l'esker, en orange les flancs de l'esker et en cyan les argiles marines.

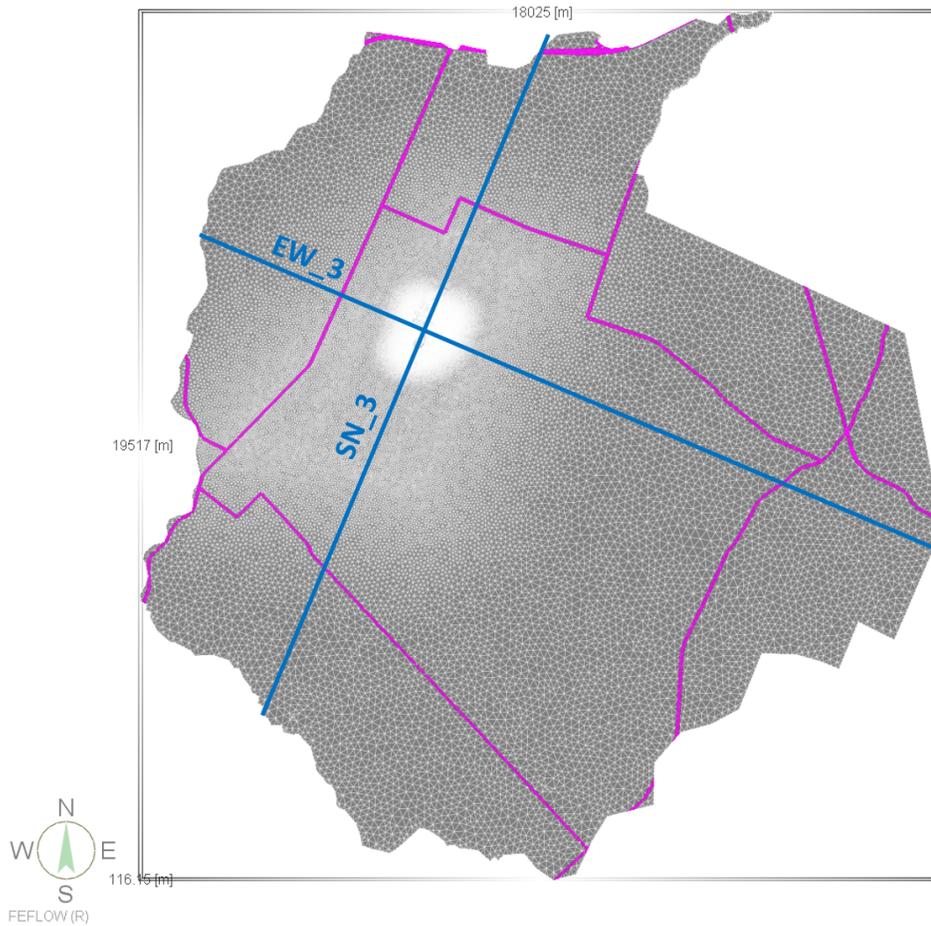
La modélisation géostatistique locale sera centrée autour des lagunes et des puits présentés sur la Figure 5.11.

## 5.4 Modèle hydrogéologique

Le modèle hydrogéologique régional a été calibré lors des travaux de Claprod *et al.* (2022). Ce modèle a été repris, avec pour objectif de le raffiner localement. Les valeurs de paramètres d'états n'étant plus régionaux, mais variables pour chaque élément autour des lagunes de Mercier.

### 5.4.1 Maillage 2D FEFLOW

La construction de la grille d'écoulement dans Feflow se fait en utilisant l'algorithme « Triangle » pour générer des éléments triangulaires en deux dimensions (2D) tout en permettant différents niveaux de raffinement souhaités autour d'éléments à considérer dans le modèle hydrogéologique. La grille a été construite afin de représenter parfaitement les éléments linéaires importants du modèle hydrogéologique tels que les cours d'eau et les limites latérales de certaines unités hydrostratigraphiques. Un niveau de raffinement a été appliqué pour obtenir des cellules d'au maximum 100 m autour des limites du modèle et des cours d'eau principaux. Une grille irrégulière de 44 points de contrôles, incluant les 3 puits de pompage de l'UTES, a été utilisée au niveau local afin de permettre une discrétisation fine des éléments (taille des éléments souhaitée de 5 m) au site des anciennes lagunes de Mercier. La grille 2D finale 5.12 comporte 57 257 nœuds (113 569 éléments) par couche. La taille des éléments varie entre 5 m et 150 m.



**Figure 5.12 – Représentation 2D de la grille de modélisation. Le réseau routier est indiqué en magenta. Les lignes bleues sont les sections verticales EW3 et SN3 présentées à la figure suivante.**

#### 5.4.2 Grille d'écoulement 3D FEFLOW

La grille 2D est extrudée en trois dimensions (3D) afin de représenter les variations verticales des unités hydrostratigraphiques ainsi que les hétérogénéités verticales des propriétés hydrauliques. Une couche de 0.5 m d'épaisseur est imposée à la surface afin d'avoir un contrôle distinct sur la géologie de surface et assurer une recharge plus efficace au modèle. Deux couches minces sont ajoutées à la base de l'unité d'Argile afin de bien représenter les différences de pressions et de charges entre les sédiments fins et peu conducteurs de l'argile et les sédiments très conducteurs de l'esker. Une surface est fixée à l'élévation du toit de l'unité Esker en tous points du modèle afin de représenter adéquatement cette unité. L'unité Esker est subdivisée en 20 couches d'épaisseur proportionnelle afin de représenter, dans une étape ultérieure au projet, toutes les hétérogénéités verticales de la

conductivité hydraulique dans cette unité. Afin d'éviter la création de couches très minces pouvant générer des instabilités numériques, la discrétisation se fait également sur l'unité de Till au niveau régional. L'élévation du toit du roc est également fixée afin d'être adéquatement représentée dans le modèle. Le roc est représenté par un total de 9 couches, pour permettre une représentation adéquate de l'hétérogénéité verticale de la conductivité hydraulique dans l'étape ultérieure de modélisation. De haut en bas, les 9 couches du roc ont les épaisseurs suivantes :

- 5 couches avec  $dz = 2$  m;
- 2 couches avec  $dz = 5$  m;
- 2 couches avec  $dz = 10$  m.

Le modèle hydrogéologique régional 3D contient 35 couches (36 surfaces), pour un total de 2 061 252 nœuds (3 974 915 éléments). La Figure 5.13 et la Figure 5.14 montrent deux sections verticales traversant le modèle régional 3D.

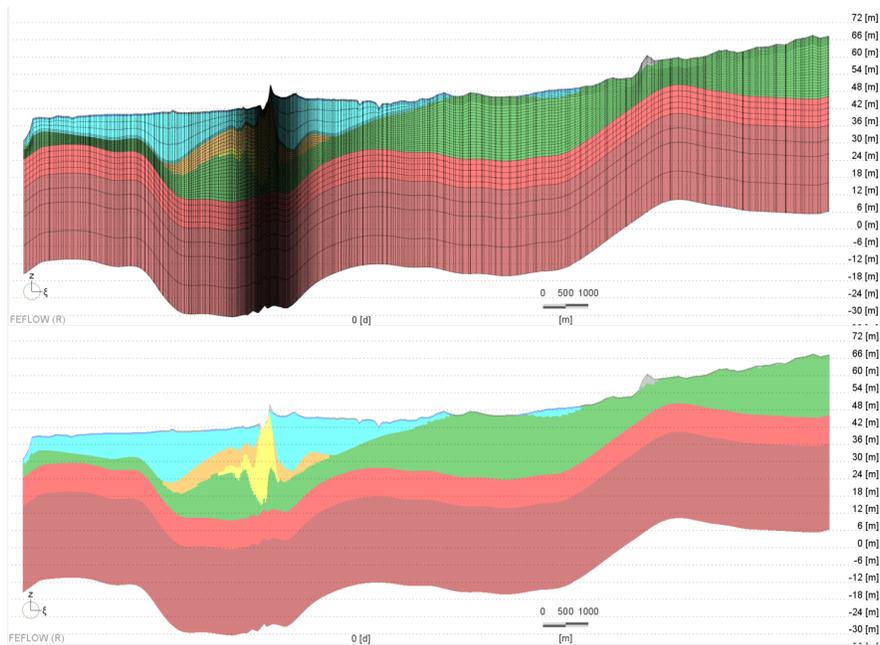


Figure 5.13 – Section verticale (EW3 sur la 5.12) d'ouest (gauche) vers l'est (droite) du modèle hydrogéologique 3D, montrant la représentation numérique des unités hydrostratigraphiques avec (haut) et sans (bas) la discrétisation latérale et verticale des éléments. Exagération verticale de 80x. De bas en haut : roc sain(marron), roc fracturé (rouge), till (vert), coeur de l'esker (jaune), esker étendu (orange), argile (cyan) et remblais organique (gris).

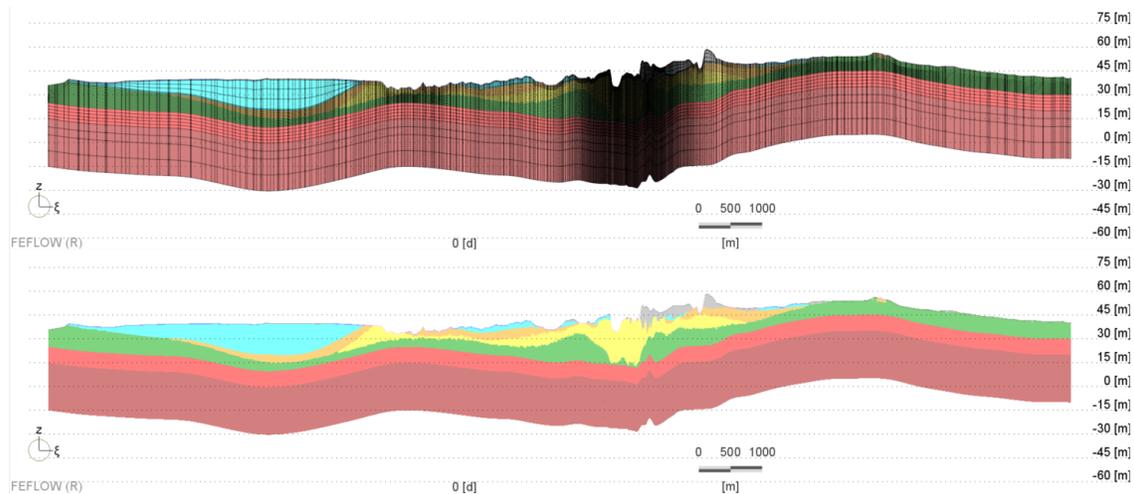


Figure 5.14 – Section verticale (SN3 sur la 5.12) du sud (gauche) ou nord (droite) du modèle hydrogéologique 3D, montrant la représentation numérique des unités hydrostratigraphiques avec (haut) et sans (bas) la discrétisation latérale et verticale des éléments. Exagération verticale de 30x. De bas en haut : roc sain(marron), roc fracturé (rouge), till (vert), coeur de l'esker (jaune), esker étendu (orange), argile (cyan) et remblais organique (gris).

### 5.4.3 Propriétés hydrauliques initiales

Le modèle régional de perméabilité a été initialement calibré à partir des données compilées à partir des études précédentes et de nouveaux essais effectués par le MELCC. Une valeur unique représentative a été assignée par unité du modèle. Ces valeurs ont été ajustées par la suite afin d'honorer les valeurs de charges hydrauliques lues aux piézomètres, régionalement et localement. Le roc est traité comme un milieu poreux équivalent (sans fractures discrètes).

### 5.4.4 Conditions aux frontières

La rivière Châteauguay située à la limite ouest du modèle est le principal élément drainant l'eau de surface de la région. La rivière l'Esturgeon est le principal cours d'eau situé à l'intérieur des limites du modèle, passant au sud du site des anciennes lagunes de Mercier. Les autres cours d'eau situés à l'intérieur des limites du modèle sont pour la plupart des canaux d'irrigation non permanents. Seuls les cours d'eau reconnus et les canaux d'irrigation plus importants sont inclus dans le modèle comme conditions aux frontières.

Une condition aux frontières de type « charge constante » égale à l'élévation de la surface topographique est assignée aux cours d'eau principaux à l'intérieur du domaine de modélisation (Figure 5.12).

Une charge constante est assignée au roc à la limite sud-est du modèle, variant de 48 masl au sud à 60 masl à l'est dans les environs de St-Rémi. Une charge constante à 30 masl est assignée au roc à la limite sud-ouest du modèle, au niveau de la rivière Châteauguay. Une charge constante de 34 masl est assignée au sud-ouest du modèle, où l'unité Esker est modélisée comme étant en connexion avec la rivière Châteauguay.

#### 5.4.5 Recharge

La distribution spatiale de la recharge provient du modèle développé par Croteau *et al.* (2010), effectué avec HELP, révisée pour tenir compte de l'utilisation d'eau dans la région par les municipalités et certains grands utilisateurs industriels (Figure 5.15). Cette recharge fluctue en fonction de la géologie de surface, et varie de 0 mm/an à 120 mm/a régionalement.

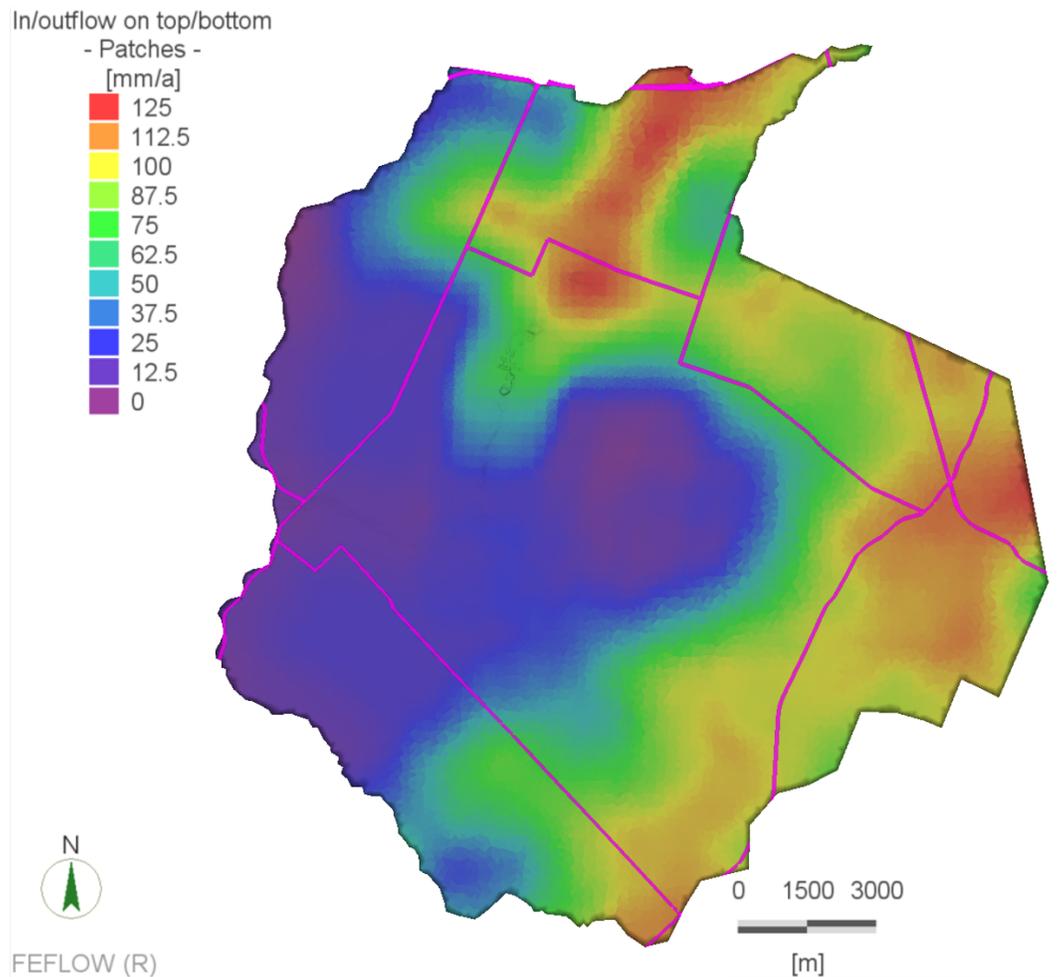


Figure 5.15 – Distribution spatiale de la recharge efficace assignée à la première couche du modèle hydrogéologique. Le réseau routier est indiqué en magenta.

#### 5.4.6 Calibration du modèle régional permanent

La calibration du modèle déterministe a été réalisée par le logiciel PEST en calibrant une valeur de conductivité hydraulique par unité hydrostratigraphique. Cette calibration initiale a permis de déterminer les valeurs de conductivité hydrauliques à assigner au niveau régional.

La calibration du modèle régional déterministe est faite en régime permanent seulement. Le modèle régional est défini en mode confiné avec un niveau de saturation de 100% sur tout le domaine de modélisation. Bien que ces hypothèses ne soient pas représentatives des conditions locales où la présence de l'esker en surface rende le modèle partiellement saturé et non confiné, ces conditions sont justifiables au niveau régional pour calibrer les valeurs de charges hydrauliques représentant

l'écoulement régional en régime permanent. La majorité du modèle est couvert par des tills et argiles imperméables en surface. Les données utilisées pour la calibration sont données sur la Figure 5.16. Une fois le modèle régional calibré, la définition de l'ensemble de modèles initiaux pour le processus d'assimilation a été possible.

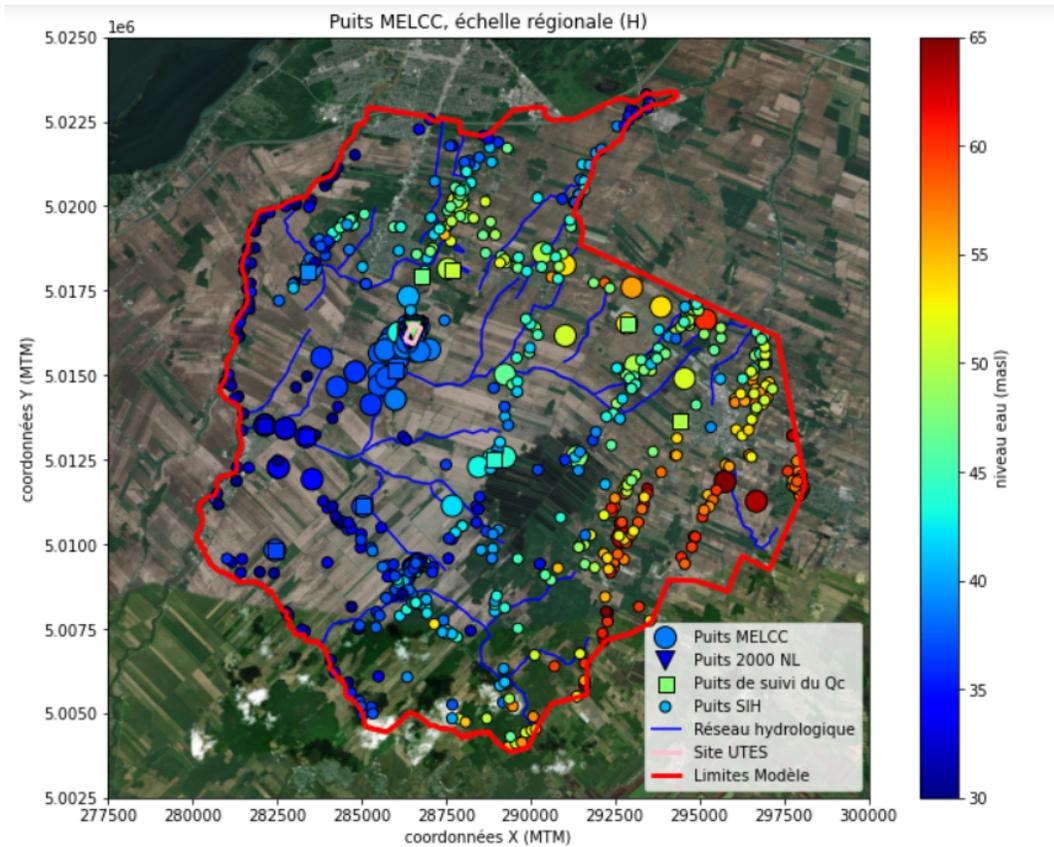


Figure 5.16 – Distribution spatiale régionale des niveaux d'eau à calibrer en régime permanent sur le modèle hydrogéologique de Ville Mercier.

## 5.5 Définition de l'ensemble initial du processus d'assimilation

Le processus d'assimilation va reposer sur deux essais de pompage disponibles. Ces deux essais de pompage sont localisés au droit des anciennes lagunes de Mercier. Le choix a donc été fait d'uniquement mettre à jour les paramètres de perméabilité et les coefficients d'emmagasinement localement autour de ces lagunes. Mettre à jour tout le modèle n'aurait aucun sens physiquement au vu des distances entre observations et paramètres.

Après observation des données géologiques, du modèle conceptuel et des valeurs de perméabilité, il a été décidé de créer des simulations géostatistiques locales uniquement au sein des 2 unités de l'esker et du roc. Les tills et argiles ont gardé leurs valeurs calibrées issues du modèle régional, le but de ces deux géologies est d'introduire des zones relativement imperméables dans les modèles. Les remblais organiques de surface ont également conservé leurs valeurs régionales, leurs emprises spatiales étant supposées trop faibles pour avoir un impact suffisant sur la distribution des charges hydrauliques. Les 2 unités de l'esker et le roc sont simulés pour représenter les incertitudes du processus de déposition sédimentaire dans le cas de l'esker et de la fracturation dans le cas du roc.

À proximité de ces deux essais de pompage, le maximum de données de perméabilité provenant des multiples campagnes d'acquisition réalisées a été compilé. Des mesures de perméabilité ont été effectuées dans le roc, dans le coeur de l'esker et dans les bordures de l'esker. La Figure 5.17 indique la position de ces mesures à gauche, avec l'unité échantillonnée. Le graphique de droite indique l'ensemble des points d'observations (qui peuvent être ou ne pas être un puits d'échantillonnage) et les trois puits de pompages qui seront utilisés pour l'assimilation de données.

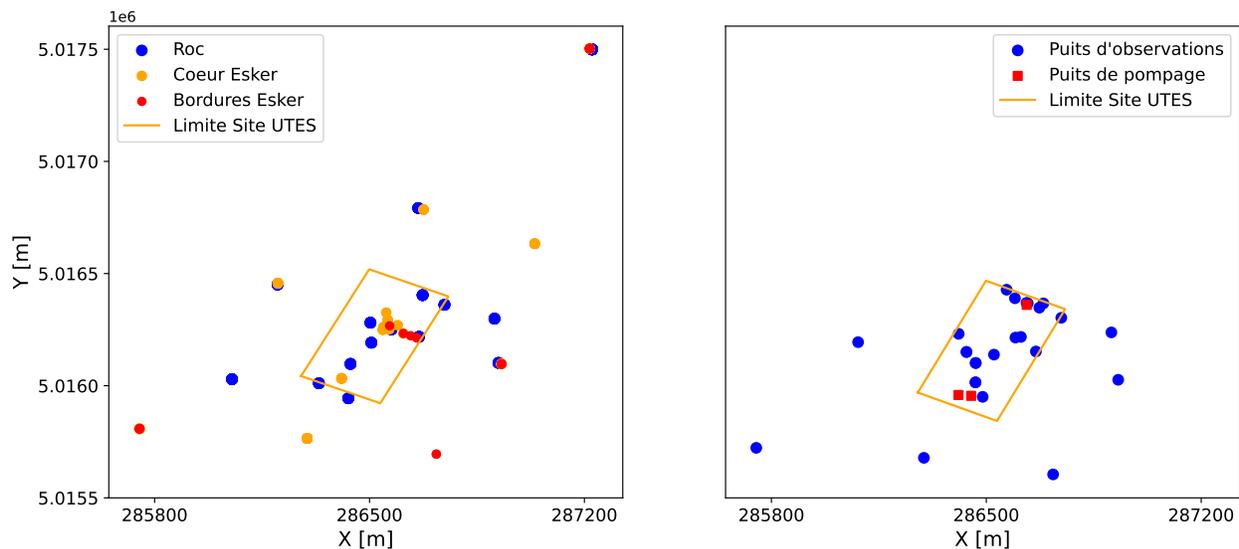
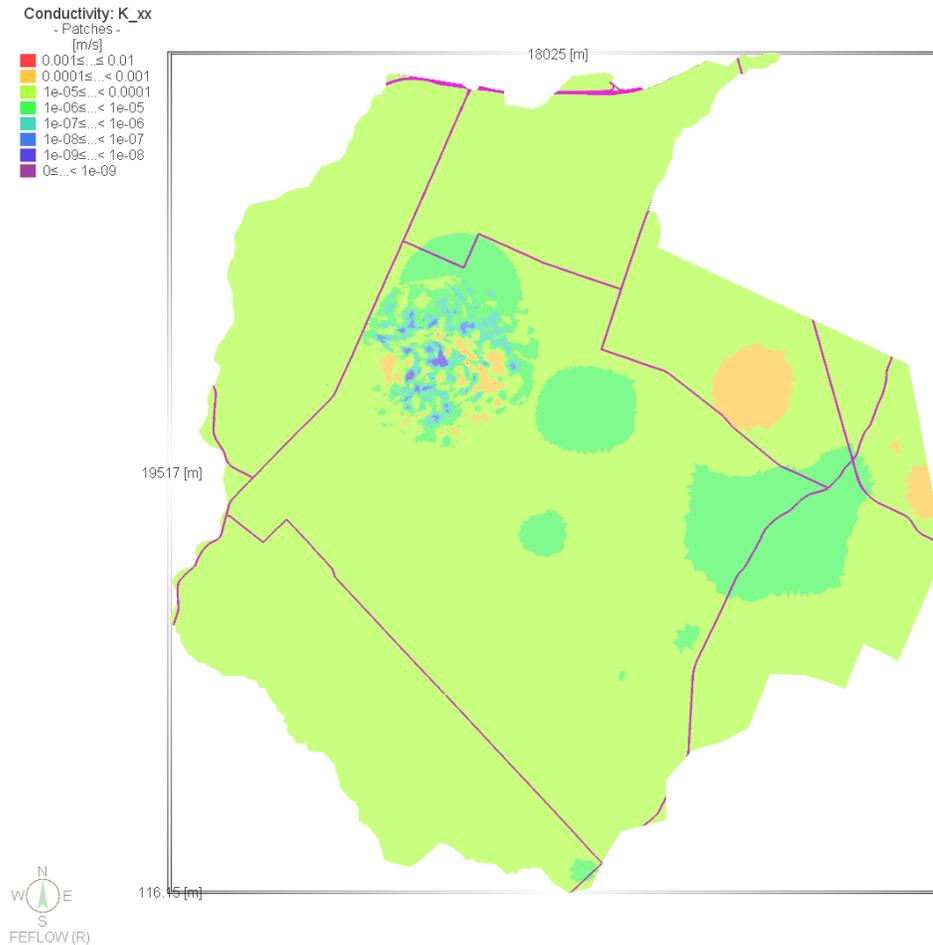


Figure 5.17 – À gauche, les puits de mesures de la conductivité hydraulique dans les trois unités modélisées. À droite, les puits d'observation et de pompage qui seront utilisés lors du processus d'assimilation. Les limites du site de l'UTES sont reprises des Figures 5.15 et 5.16.

Les simulations de conductivités hydrauliques ont été réalisées à l'aide du logiciel Isatis. Lors du processus d'assimilation, les recharges, conditions frontières et les ratios  $K_r = K_x/K_z$  ont été imposés aux valeurs issues de la calibration régionale. Les champs de perméabilités géostatistiques et

l'assimilation ont donc uniquement porté sur la perméabilité horizontale isotrope et les coefficients d'emmagasinement spécifique.

Au regard de la distribution spatiale de données de perméabilités issues de forage, et de la disponibilité des données d'essais transitoires, la zone locale a été restreinte au centre du modèle autour des anciennes lagunes. La position des anciennes lagunes était identifiée sur la Figure 5.15. La Figure 5.18 indique clairement la zone locale simulée avec des valeurs variables. Cette couche du roc démontre que la moyenne régionale de perméabilité dans le roc est plus élevée que ce qui était mesuré dans les données locales. Cette différence pourrait provenir de la perméabilité équivalente pour le modèle régional, intégrant du roc sain et du roc fracturé. Dans le cas des données d'essais locaux, les perméabilités identifiaient indépendamment les zones de roc sain et fracturé. Dans le cas du roc, la transition entre le modèle local et régional a été assurée en utilisant les résidus par rapport aux valeurs régionales. En assignant des résidus à 0 sur les bordures du domaine, la transition est douce, réduisant les instabilités numériques.



**Figure 5.18 – Modèle FEFLOW de perméabilité dans une couche du roc. L’insert géostatistique local présente des variabilités spatiales lissées dans le modèle régional. Le réseau routier est indiqué en magenta.**

Il a été décidé de bâtir 106 modèles initiaux. Ce nombre est proche des 100 modèles cités dans la littérature (Raanes, 2015a; Evensen, 2009a) et anticipe la non-convergence de quelques modèles lors des multiples simulations-assimilations. L’objectif étant de rester autour de 100 modèles à la fin du processus pour garder de la variabilité, et donc de l’incertitude dans l’ensemble.

### 5.5.1 Définition de l’ensemble initial pour le processus d’assimilation

3 régions individuelles ont été simulées, une pour le coeur de l’esker, une pour les bords de l’esker et une pour le roc. Ces points de supports pour les simulations ont été extraits du modèle régional montré sur la Figure 5.11. Chaque région a été simulée individuellement en utilisant unique-

ment les données de perméabilités issues des horizons géologiques correspondants. 200 réalisations individuelles de perméabilités ont été générées pour chaque région, en anticipation des modèles qui ne convergeraient pas lors des simulations d'écoulement. Chaque modélisation a été obtenue grâce aux simulations de bandes tournantes (TBS) d'Isatis Neo. 2400 bandes ont été utilisées pour chaque simulation.

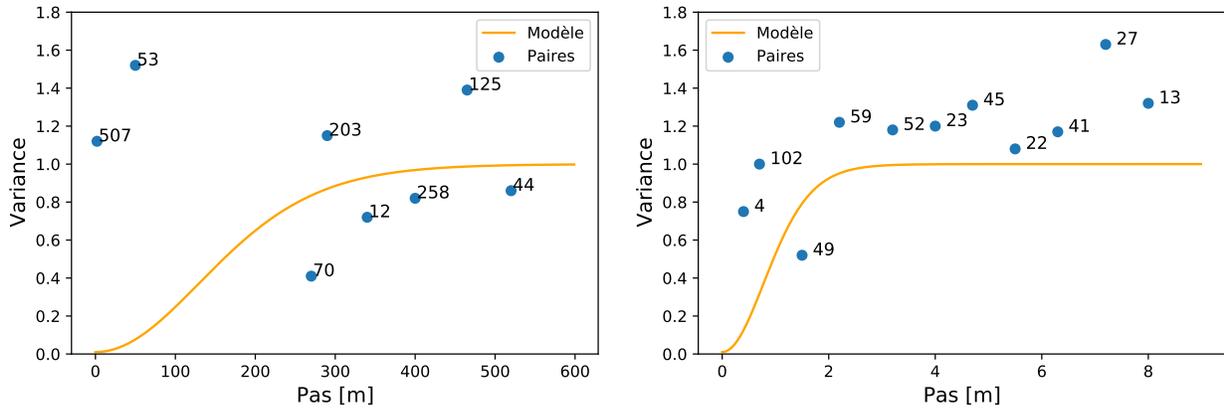
### Coeur de l'esker

Pour l'unité géologique définie comme étant le coeur de l'esker, 18 puits et 86 mesures de perméabilité étaient disponibles. L'esker est une unité quasiment autonome dans le modèle régional, la perméabilité a donc été simulée directement dans le modèle local. La jonction entre le modèle local et le modèle régional représente une toute petite portion du modèle global, ce qui limite les risques d'instabilités numériques lors des simulations d'écoulement.

L'analyse des mesures de perméabilités locales et de l'analyse exploratoire précédente pour le modèle régional a conduit à choisir le variogramme suivant après transformation gaussienne des  $\log_{10}$  de perméabilité :

- Variance totale de  $1 \text{ m}^2$ ;
- Effet pépite de  $0.01 \text{ m}^2$ ;
- Modèle Cauchy généralisé de 500 m en horizontal, 3 m en vertical et de contribution  $0.99 \text{ m}^2$ .

Le variogramme et les paires locales sont donnés sur la Figure 5.19. Le variogramme omnidirectionnel reprend les nombreuses paires verticales à courtes portées. Le modèle en rouge sur la Figure 5.19a représente le modèle horizontal. Le modèle Cauchy généralisé a été retenu pour assurer une transition spatiale douce entre les éléments de perméabilités et donc éviter des problèmes lors des simulations numériques d'écoulement. Dans le plan horizontal, un modèle omnidirectionnel a été retenu par manque d'information pour évaluer les variations de variances en fonction des azimuts.



(a) Variogramme omnidirectionnel pour l'unité géologique du coeur de l'esker. (b) Variogramme vertical pour l'unité géologique du coeur de l'esker.

Figure 5.19 – Variogrammes modélisés pour l'unité géologique du coeur de l'esker.

Le résultat d'une simulation de  $\log_{10}$  de perméabilité est donné sur la Figure 5.20 avec les points conditionnant aux puits supportant la simulation.

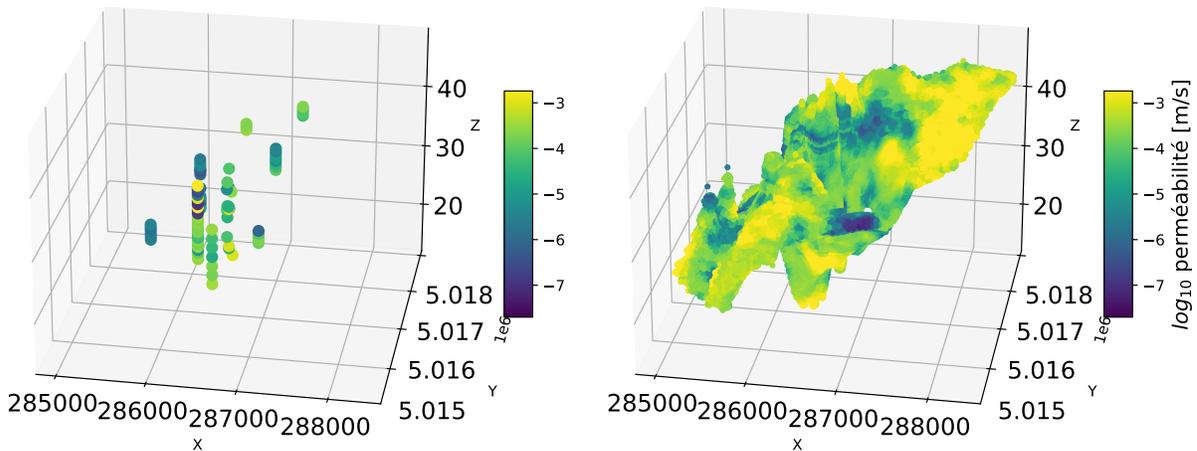


Figure 5.20 – À gauche, puits de mesures utilisés pour les simulations géostatistiques. À droite, une réalisation de perméabilité pour l'unité du coeur de l'esker.

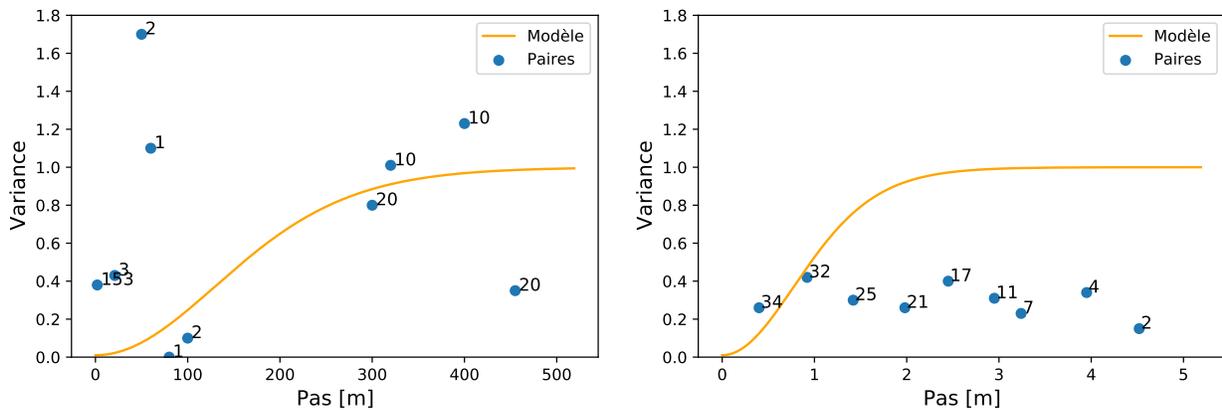
### Bordures de l'esker

Les bordures de l'esker sont une unité géologique globalement sous-échantillonnée. Spatialement elle est bien définie par l'analyse des données sismiques, mais uniquement 9 puits et 37 mesures

l'échantillonnage. Les variogrammes obtenus en appliquant la même logique que pour le coeur de l'esker sont difficilement exploitables. Le manque de données a conduit à utiliser le même modèle de variogramme que dans le cas du coeur de l'esker :

- Variance totale de  $1 \text{ m}^2$ ;
- Effet pépite de  $0.01 \text{ m}^2$ ;
- Modèle Cauchy généralisé de 500 m en horizontal, 3 m en vertical et de contribution 0.99  $\text{m}^2$ .

Les variogrammes et les paires locales sont donnés sur la Figure 5.21.



(a) Variogramme omnidirectionnel pour l'unité géologique des bordures de l'esker. (b) Variogramme vertical pour l'unité géologique des bordures de l'esker.

**Figure 5.21 – Variogrammes modélisés pour l'unité géologique des bordures de l'esker.**

Le résultat d'une simulation de  $\log_{10}$  de perméabilités est donnée sur la Figure 5.22 avec les points conditionnant aux puits supportant la simulation.

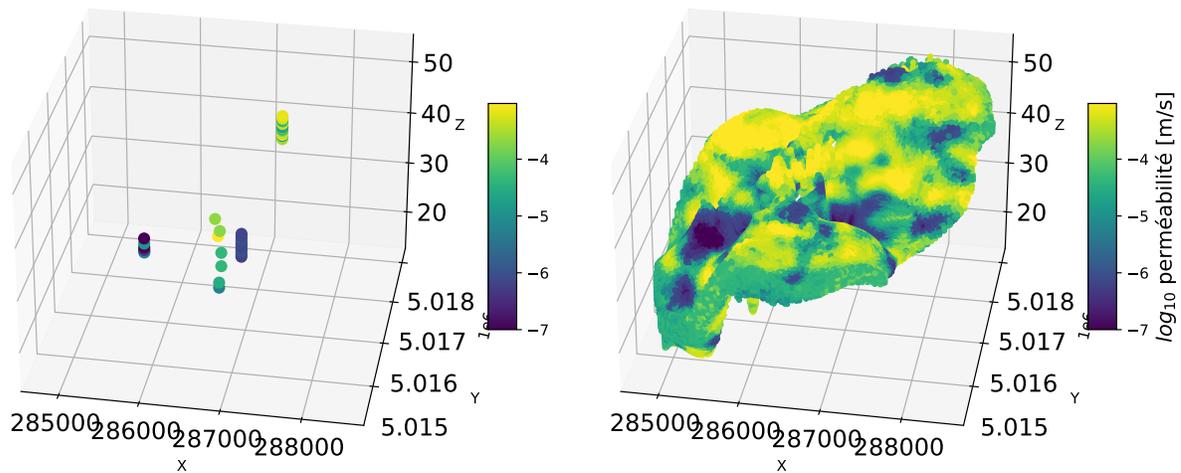


Figure 5.22 – À gauche, puits de mesures utilisés pour les simulations géostatistiques. À droite, une réalisation de perméabilité pour l'unité des bordures de l'esker.

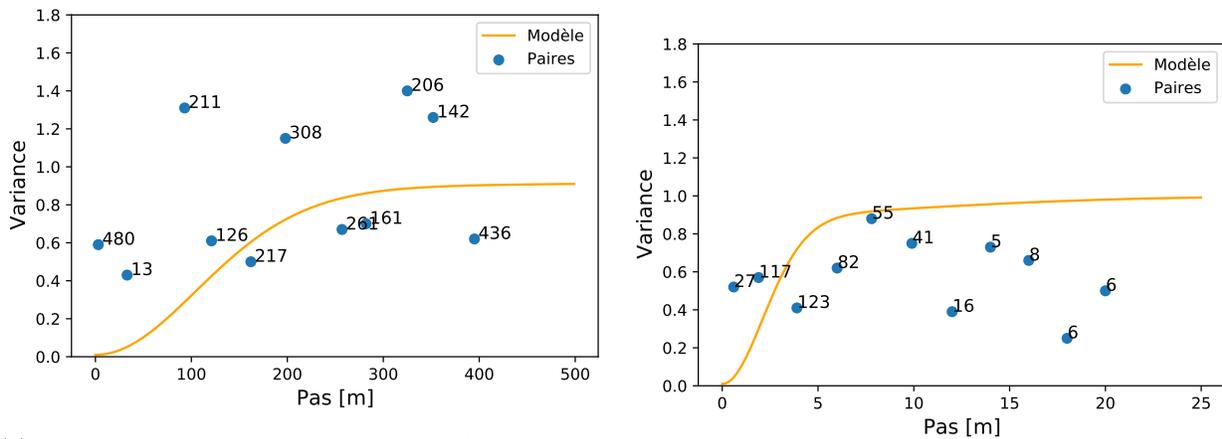
## Roc

Dans le cas du roc, 15 puits et 111 mesures de perméabilités étaient disponibles. Pour assurer la transition entre les modèles de perméabilité locaux et régionaux, les résidus ont été simulés entre les mesures de perméabilités dans les forages et la perméabilité du modèle régional. De cette manière, il a été possible d'imposer des valeurs de résidus à 0 à l'interface entre modèle local et régional, et assurer une continuité dans les champs de perméabilités. Cette continuité permet d'assurer une meilleure convergence dans les simulations d'écoulement.

L'analyse des mesures de perméabilités locales et de l'analyse exploratoire précédente pour le modèle régional ont conduit à choisir le variogramme suivant pour la transformation gaussienne des résidus :

- Variance totale de  $1 \text{ m}^2$ ;
- Effet pépite de  $0.01 \text{ m}^2$ ;
- Modèle Cauchy généralisé de 500 m en horizontal, 10 m en vertical et de contribution  $0.89 \text{ m}^2$ ;
- Modèle Cauchy généralisé de 5000 m en horizontal, 50 m en vertical et de contribution  $0.1 \text{ m}^2$ .

Le variogramme et les paires locales sont donnés sur la Figure 5.23. Le variogramme omnidirectionnel reprend les nombreuses paires verticales à courtes portées. Le modèle en rouge sur la Figure 5.23a représente le modèle horizontal. Le modèle Cauchy généralisé a été retenu pour assurer une transition spatiale douce entre les éléments de perméabilités et donc éviter des problèmes lors des simulations numériques d'écoulement. Dans le plan horizontal, le manque d'information a contraint d'utiliser un modèle omnidirectionnel pour évaluer les variations de variances en fonction des azimuts.



(a) Variogramme omnidirectionnel pour l'unité géologique du roc.

(b) Variogramme vertical pour l'unité géologique du roc.

**Figure 5.23 – Variogrammes modélisés pour l'unité géologique du roc.**

Le résultat d'une simulation de résidus est donné sur la Figure 5.24 avec les points conditionnant aux puits supportant la simulation.

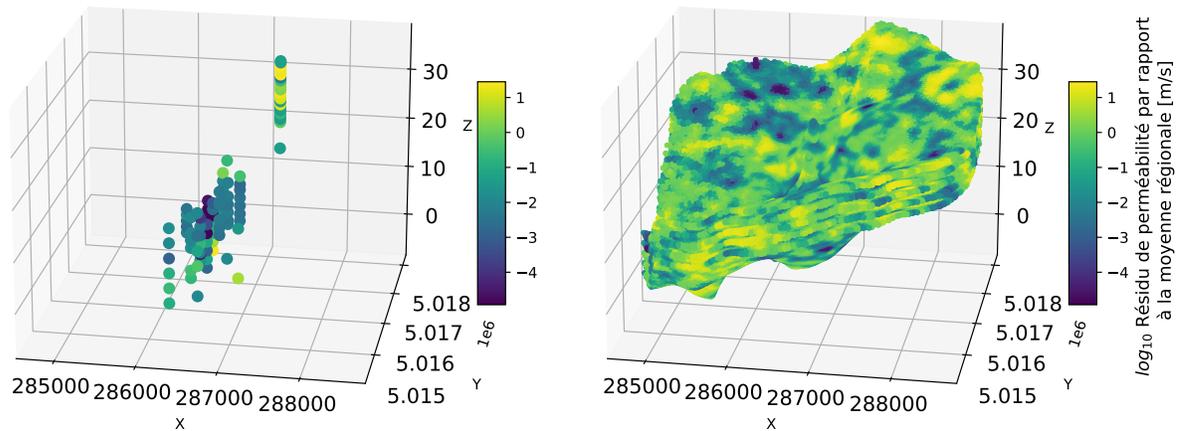


Figure 5.24 – À gauche, puits de mesures utilisés pour les simulations géostatistiques. À droite, une réalisation de résidus de perméabilité pour l'unité de roc.

Les textures obtenues sur les Figures 5.20, 5.22 et 5.24 traduisent les variabilités à courtes échelles recherchées. Pour les 2 unités de l'esker, cela correspond aux processus de déposition plus ou moins énergétiques. Pour l'unité géologique du roc, cela correspond aux zones de fracturations qui permettent une communication hydraulique entre l'esker et le roc.

À la fin de cette étape, 200 modèles de perméabilités locaux ont été obtenus pour créer l'ensemble initial. Depuis cet ensemble initial, des simulations d'écoulement ont été effectuées avec les deux essais de pompages. À la suite de ces pompages, 106 modèles qui convergeaient numériquement ont été retenus. Pour obtenir 106 modèles convergeant, il a été nécessaire de rouler 151 modèles. Cette première élimination de modèles a permis d'enlever de l'ensemble initial les modèles de perméabilités qui ne permettaient pas les pompages des débits observés lors des essais de pompage.

Le travail de modélisation géostatistique, traduisant les hypothèses du modèle conceptuel, est considéré satisfaisant en regardant les résultats de l'ensemble initial. Cet ensemble a été construit à partir de données régionales permanentes et de données locales de perméabilité issues d'essais de pompage. Les observations transitoires simulées dans les 106 modèles initiaux étaient bien distribuées autour des mesures dans les simulations. Cette validation croisée entre données permanentes et transitoire a permis d'être confiants pour utiliser cet ensemble initial pour le processus d'assimi-

lation. La RMSE moyenne de l'ensemble initial était d'environ 1.5 m comme le démontre les Figures 5.33 et 5.34, traduisant la qualité des modèles initiaux obtenus.

## 5.6 Essais de pompages

### 5.6.1 Essai de pompage numéro 1

La performance des ensembles va être suivie sur deux essais de pompages distincts. L'essai 1 est suivi par 12 puits, la position de ces puits autour des puits de pompage est donnée sur la Figure 5.25. Les puits F-102, F-106 et F-124 sont séparés en deux unités, A et C. La partie crépinée A suit les charges hydrauliques dans l'esker, les puits C suivent les valeurs dans le roc.

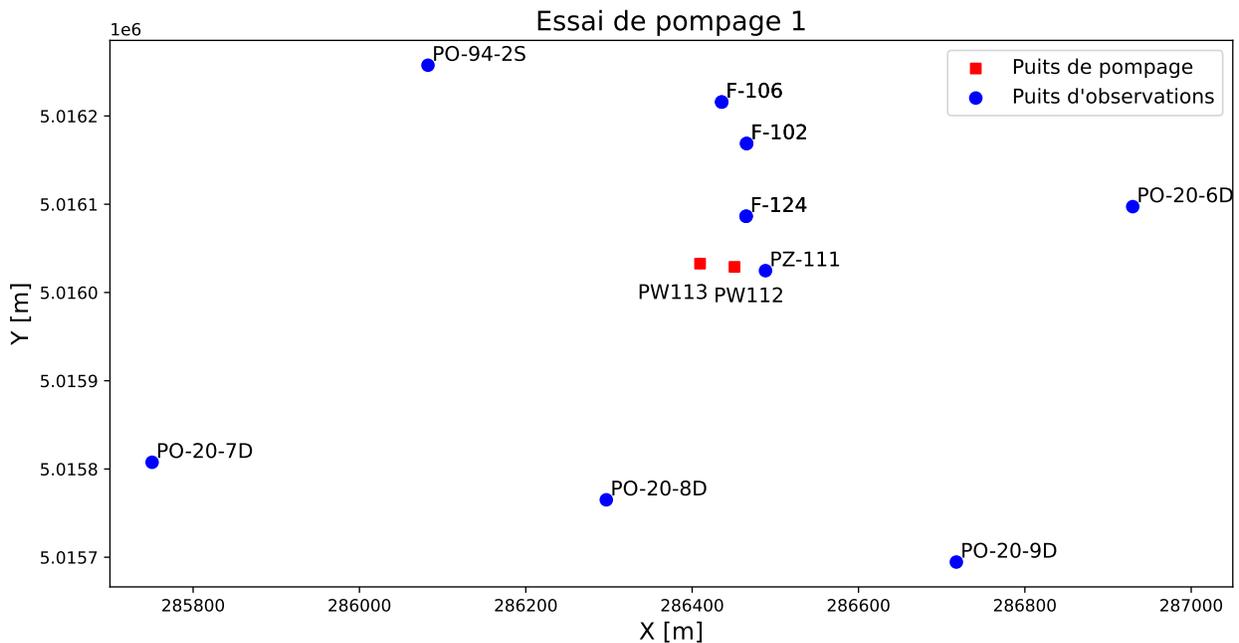


Figure 5.25 – Position des puits d'observations et de pompage pour l'essai de pompage 1.

Deux puits de pompages PW112 et PW113 ont vu leurs débits varier tout au long de l'essai comme indiqué sur la Figure 5.26. Pour des raisons de contraintes techniques, il n'a pas été possible d'interrompre l'intégralité du pompage. Le puits PW113 a vu son débit augmenter pour compenser l'arrêt du puits PW112.

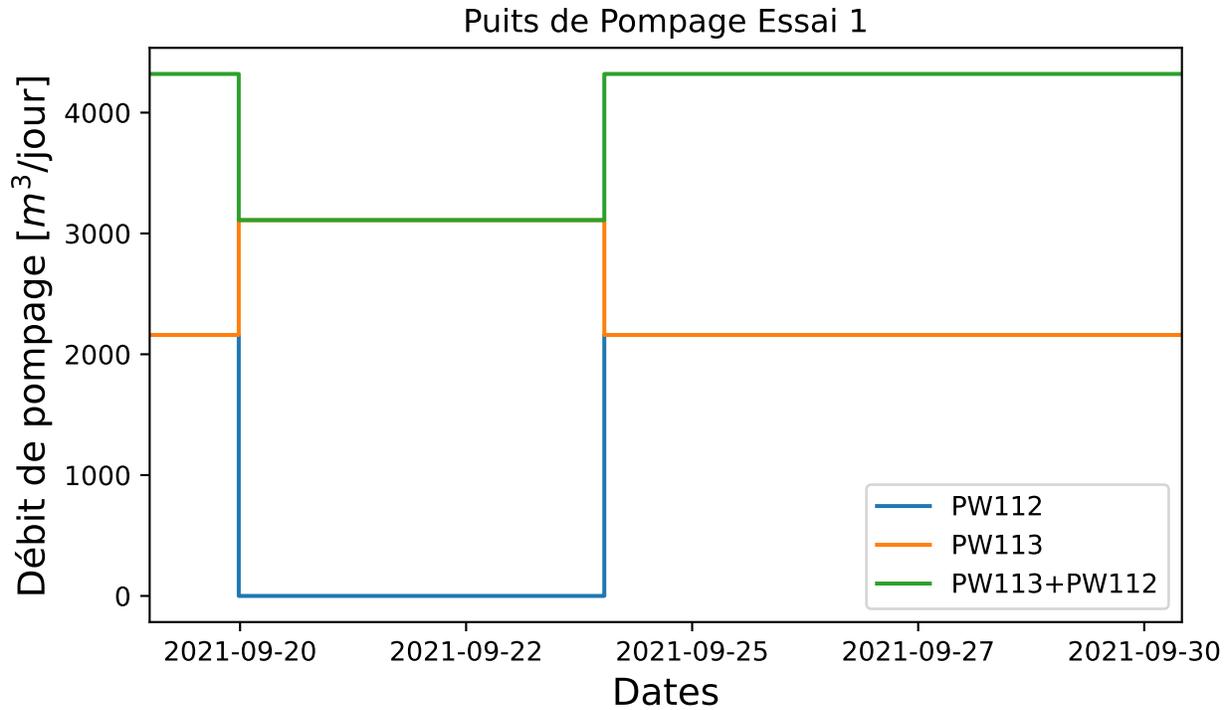


Figure 5.26 – Débits prélevés par les puits de pompage PW112 et PW113 lors de l’essai de pompage 1.

Un exemple des charges hydrauliques mesurées lors de l’essai de pompage 1 est donné sur la Figure 5.27. Les réponses sont dynamiquement très différentes en fonction de l’éloignement des puits d’observations à l’essai de pompage. L’amplitude varie également d’un puits d’observation à un autre. Ces charges traduisent un sous-sol complexe avec des unités hydrauliques potentiellement déconnectées.

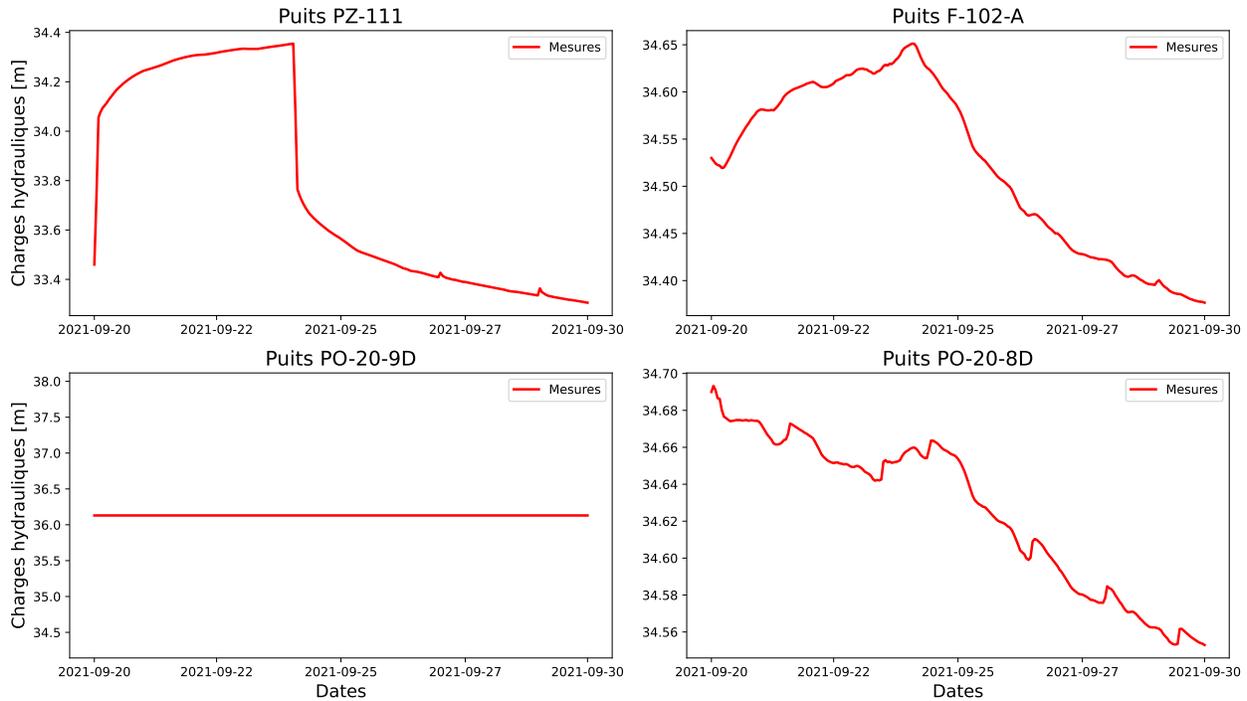


Figure 5.27 – Charges hydrauliques mesurées sur 4 puits de l'essai de pompage 1.

## 5.6.2 Essai de pompage numéro 2

L'essai 2 est suivi par 13 puits, la position de ces puits autour du puits de pompage est donnée sur la Figure 5.28. Le puits Z11 est séparé en deux unités, A et C. La partie crépinée A suit les charges hydrauliques dans l'esker, le puits C suit les valeurs dans le roc.

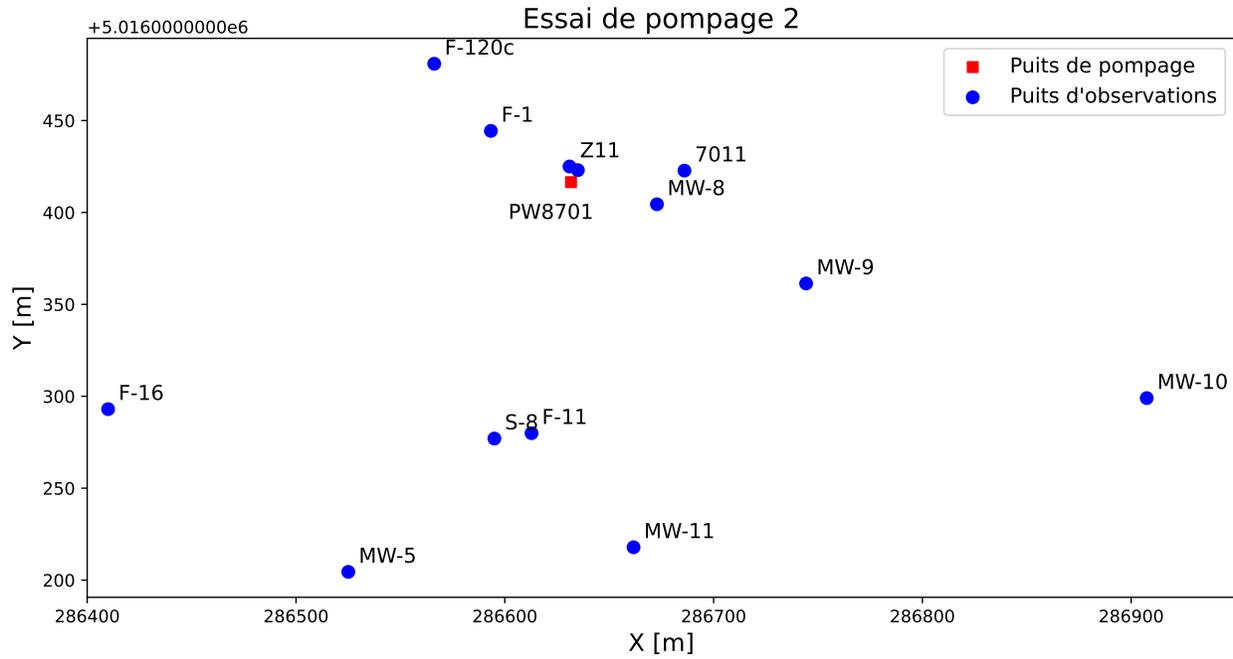


Figure 5.28 – Position des puits d’observations et de pompage pour l’essai de pompage 2.

Le puits de pompage PW8701 a été suivi lors de l’essai de pompage. Les débits prélevés sont indiqués sur la Figure 5.29. Lors de l’essai de pompage, de très fortes précipitations ont eu lieu. Ces dernières n’ont pas pu être correctement intégrées lors de la modélisation transitoire.

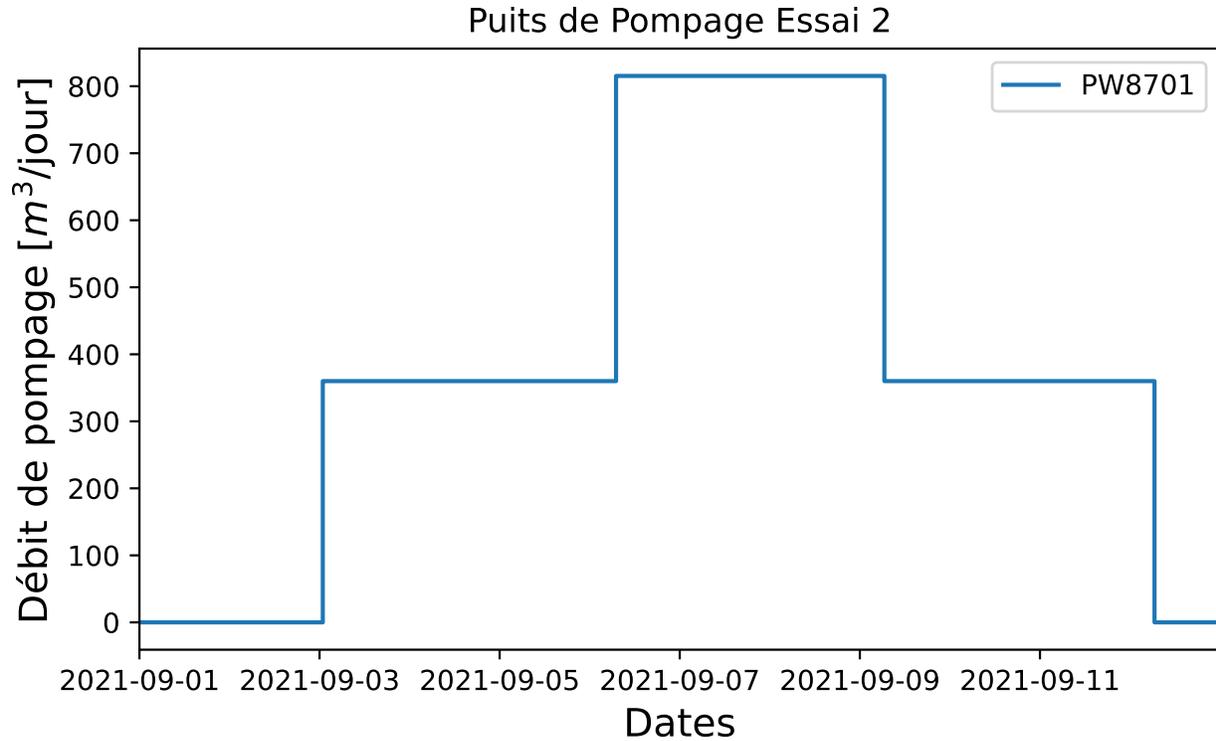


Figure 5.29 – Débits prélevés par le puits de pompage PW8701 lors de l'essai de pompage 2.

Une représentation des charges hydrauliques mesurées pendant l'essai de pompage 2 est donnée sur la Figure 5.30. Certaines réponses ont tendance à indiquer une augmentation de la charge hydraulique lors de l'essai (Puits F-120C et F-1) tandis que d'autres puits démontrent une diminution de la charge (MW-10 et Z11A). À nouveau, les variations d'amplitudes et de dynamiques traduisent un sous-sol complexe.

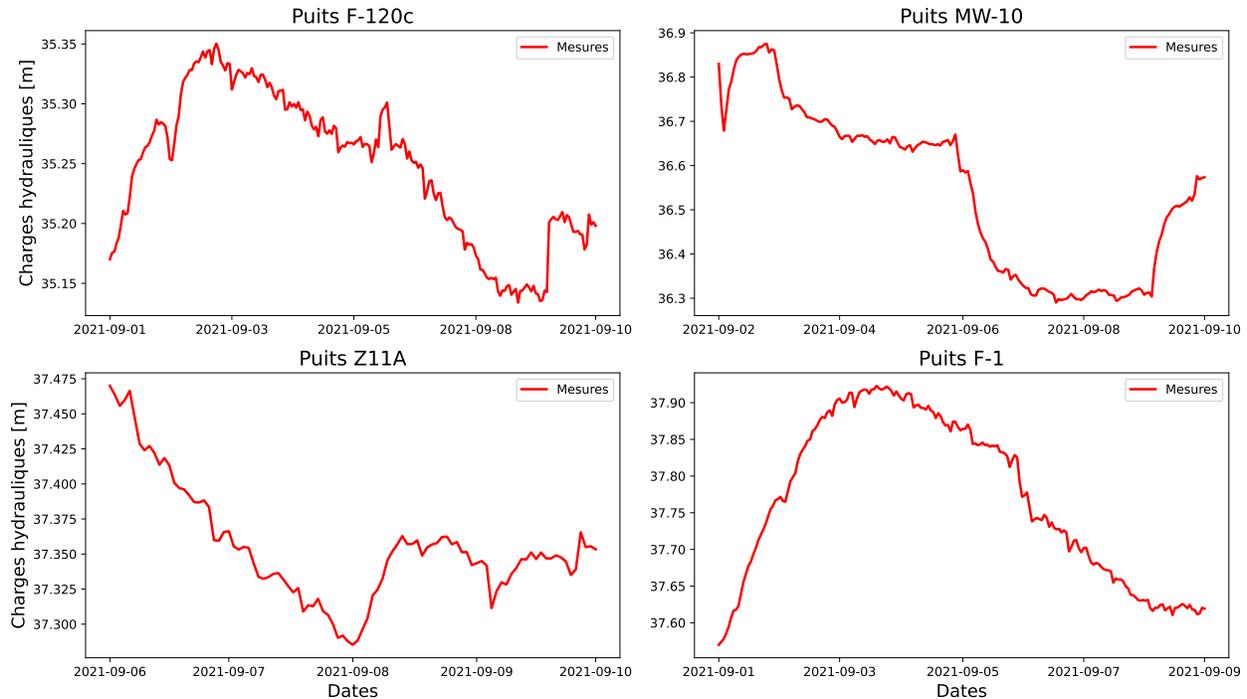


Figure 5.30 – Charges hydrauliques mesurées sur 4 puits de l’essai de pompage 12.

Ces deux essais de pompages procurent l’intégralité des données dynamiques nécessaires à l’assimilation de l’ensemble des modèles. Ces deux essais présentent des limites avec des sources externes ayant pu affecter les essais. L’erreur relative sera donc forcément plus élevée que lors du cas idéal de pompage lors du cas synthétique.

## 5.7 Résultats

Les résultats obtenus sur le cas réel sont le résultat du workflow présenté sur la Figure 3.1. Les modèles conceptuel et hydrogéologique ont été bâtis par Claprood *et al.* (2022). Le modèle géostatistique local a été construit pour construire l’ensemble de modèles nécessaires à l’assimilation. Les performances vont être comparées avec différents schémas d’analyses locales.

### 5.7.1 Évaluation des performances des schémas de localisation envisagés

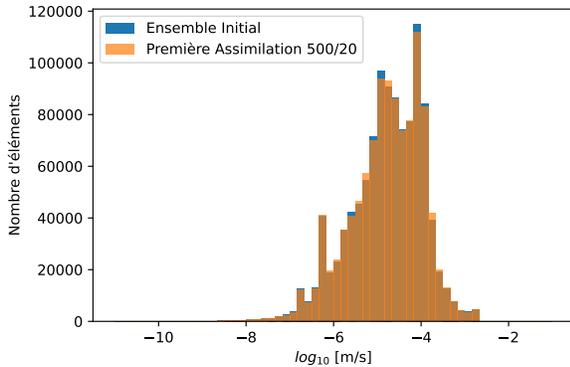
Initialement, 4 schémas d'assimilations ont été envisagés pour tester les performances de la méthode sur un cas réel. Les performances attendues sont inférieures à celles observées sur les cas synthétiques. Le nombre de paramètres et de puits à assimiler étant beaucoup plus important. De plus, uniquement les charges hydrauliques seront à disposition, ce qui apportera moins d'informations que le couple charges hydrauliques / concentrations du cas synthétique.

Dans un premier temps, l'assimilation brute de toutes les données en une fois sans analyse locale sera testée. Ceci sera la performance de base. Ensuite l'assimilation de toutes les données en une seule fois, mais avec uniquement des paramètres dans un voisinage de 500 m horizontal et 20 m vertical autour des points d'observations sera évaluée. Le voisinage calculé est issu de la méthode détaillée précédemment, dans ce schéma-ci, tous les poids d'observations sont de 1 à l'intérieur du voisinage. Ici l'étude portera uniquement sur l'impact d'un nombre de paramètres réduits sans atténuation en fonction de la distance. Enfin, deux schémas sont testés pour évaluer l'impact de la nouvelle méthode de localisation, avec des distances de 500/20 m et 250/15 m. Dans ces deux méthodes, l'assimilation se fait puits par puits avec un poids d'observation qui diminue en s'éloignant des points d'observation pour atteindre 0 aux distances imposées.

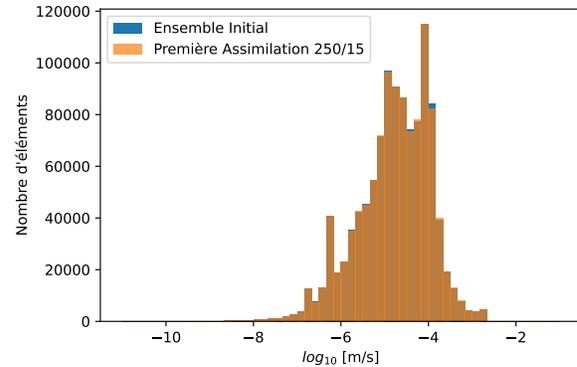
L'assimilation complète multiple prend 6 jours pour réaliser les 4 itérations d'un schéma. Pour étudier l'impact initial des assimilations avec et sans analyse locale, les 4 schémas ont été testés au départ après une itération. Si les modèles permettent une amélioration de l'ensemble, la RMSE des charges hydrauliques devrait déjà s'améliorer au bout d'une itération.

Les distributions de perméabilités moyennes et individuelles affichées sur les Figures 5.31 et 5.32 indiquent que dans le cadre de ce modèle hautement complexe, avec de nombreuses observations, il n'est pas possible de conserver une distribution de perméabilité proche de celle initiale sans utiliser une assimilation des données, puits par puits avec analyse locale. Dans toutes les assimilations présentées ici, la moyenne de l'ensemble est conservée, mais la distribution et la variance sont perdues pour les assimilations brutes de tous les puits en une fois. Pour honorer les hypothèses des modèles conceptuels et les perméabilités échantillonnées, il serait souhaitable de conserver une distribution proche de l'ensemble initial plutôt qu'une distribution trop lissée. En assimilant toutes les données en une étape, le processus n'arrive pas à détecter des corrélations viables. La mise à jour ressemble donc à un bruit gaussien proportionnel à l'anomalie de perméabilité dans l'ensemble appliquée sur

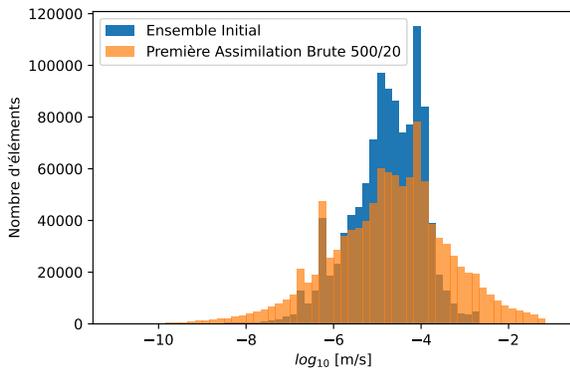
chaque élément du modèle de perméabilité. Dans les Figures 5.31a et 5.31b, l'amplitude de la mise à jour sur les perméabilités est faible, permettant de conserver une distribution statistique proche de celle de l'ensemble initial. Le schéma 250/15 impose des mises à jour moyennes plus faibles que dans le cas du schéma 500/20. Ceci correspond aux résultats observés dans le cas synthétique. Lorsque le voisinage de l'analyse locale est réduit, le nombre de paramètres d'états mis à jour est logiquement plus faible.



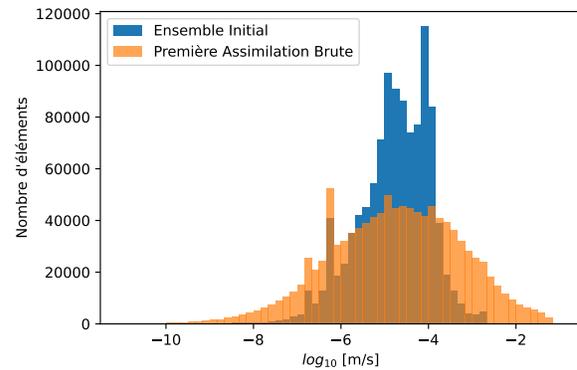
(a) En orange, distribution moyenne de l'ensemble après une assimilation avec un schéma puits par puits et analyse locale 500/20.



(b) En orange, distribution moyenne de l'ensemble après une assimilation avec un schéma puits par puits et analyse locale 250/15.



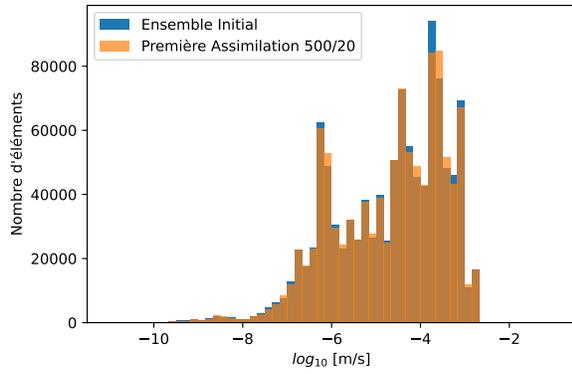
(c) En orange, distribution moyenne de l'ensemble après une assimilation de toutes les données en une étape et analyse locale 500/20.



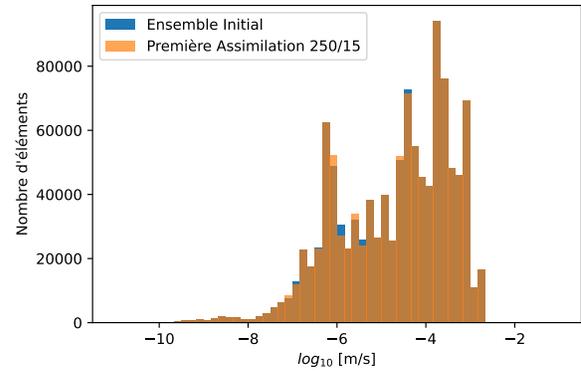
(d) En orange, distribution moyenne de l'ensemble après une assimilation de toutes les données sans analyse locale.

**Figure 5.31 – Évolution des distributions moyennes de perméabilités entre l'ensemble initial et 4 schémas d'assimilations après une itération. L'histogramme bleu sur les 4 sous-figures correspond à la distribution moyenne de l'ensemble initial.**

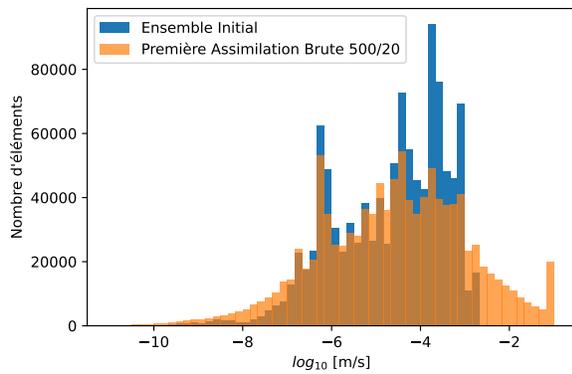
Les distributions de perméabilités mises à jour les Figures 5.32c et 5.32d indiquent une distribution de fortes valeurs aux alentours de 0.1 m/s. Cet artefact provient d'une limite imposée, modifiant toute valeur au-dessus de 0.1 m/s à 0.1 m/s après assimilation. Cette limite haute était imposée pour assurer une convergence des modèles numériques d'écoulements.



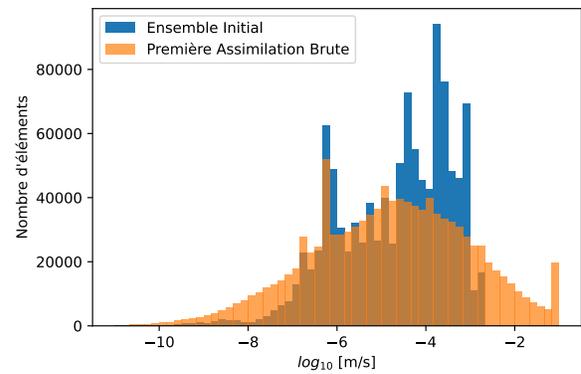
(a) En orange, distribution d'un modèle après une assimilation avec un schéma puits par puits et analyse locale 500/20.



(b) En orange, distribution d'un modèle après une assimilation avec un schéma puits par puits et analyse locale 250/15.



(c) En orange, distribution d'un modèle après une assimilation de toutes les données en une étape et analyse locale 500/20.



(d) En orange, distribution d'un modèle après une assimilation de toutes les données sans analyse locale.

**Figure 5.32 – Évolution d'une distribution individuelle de perméabilités entre l'ensemble initial et 4 schémas d'assimilations après une itération. L'histogramme bleu sur les 4 sous-figures correspond à la distribution de l'ensemble initial.**

Pour évaluer l'impact de ces différentes distributions de paramètres sur les simulations de charges hydrauliques au bout d'une itération d'assimilation, les RMSE sur les charges hydrauliques ont été calculées sur l'ensemble de puits séparés en 2 essais de pompage. L'essai de pompage 1 regroupe 12 puits d'observations qui sont donnés sur la Figure 5.25. L'essai de pompage numéro 2 regroupe quant à lui 13 puits donnés sur la Figure 5.28.

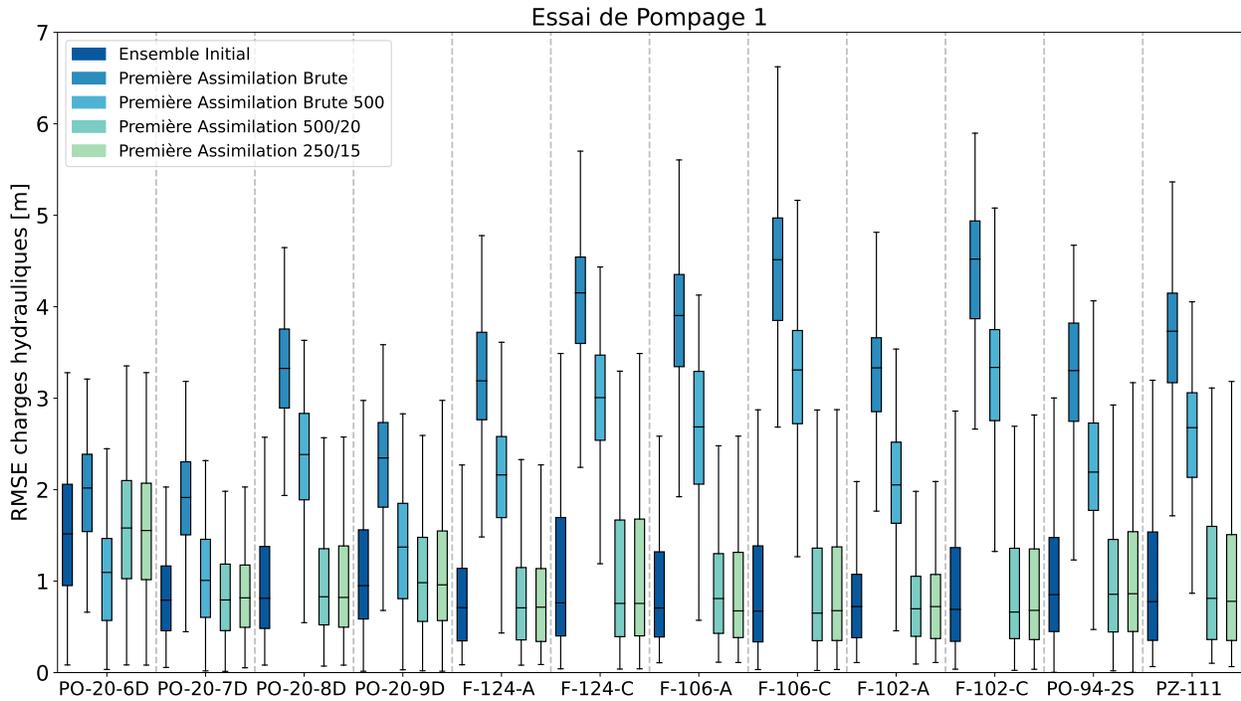


Figure 5.33 – Erreur moyenne au carré (RMSE) des charges hydrauliques au droit des 12 puits de l’essai de pompage 1 après une itération d’assimilation. Chaque couleur correspond à un schéma d’assimilation différent.

Pour l’essai de pompage 1 décrit sur la Figure 5.33, la RMSE de l’ensemble initial sur les charges hydrauliques était inférieure à 1 m en moyenne sur tous les puits à l’exception du puits PO-20-6D. Dans le cas de l’assimilation brute de toutes les données en une fois, les RMSE atteignent maintenant 4 m sur de nombreux puits de cet essai. L’assimilation en une seule fois en mettant uniquement à jour les paramètres dans un voisinage de 500 m des observations atteint des RMSE de 3 m. Ces deux schémas dégradent donc fortement l’ensemble. Dans le cas de l’assimilation puits par puits, avec différentes portées d’analyse locale, les RMSE restent quasiment identiques à celles de l’ensemble initial après une assimilation.

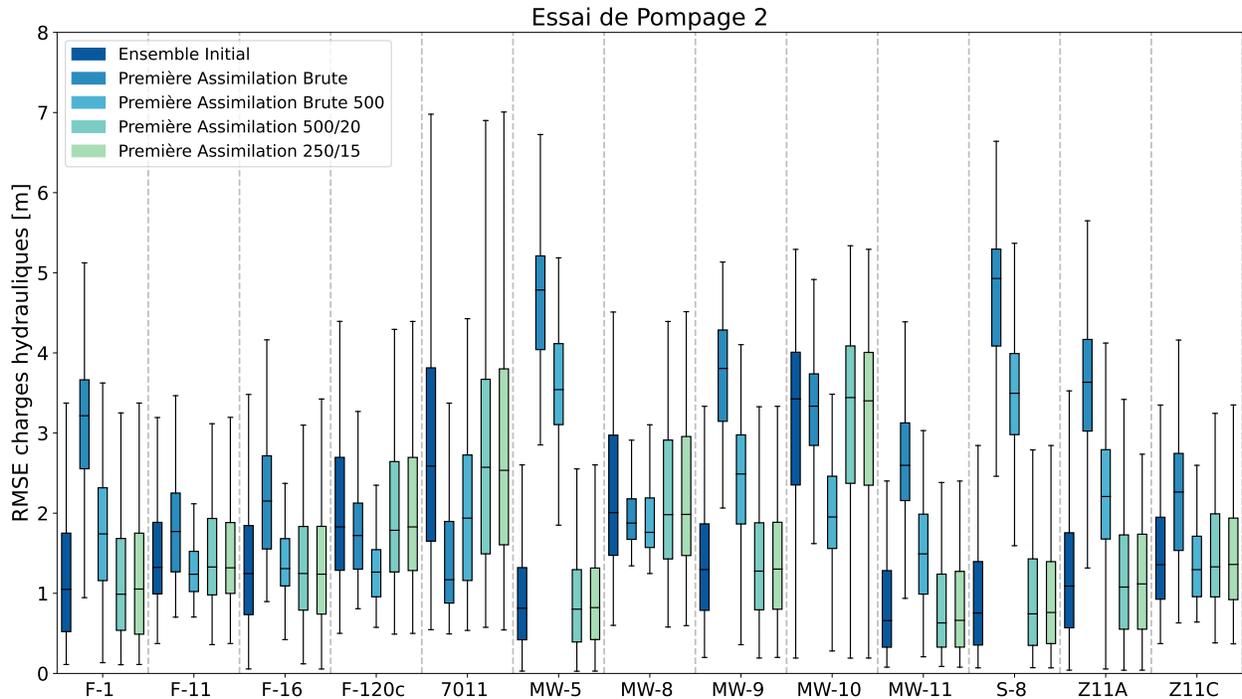


Figure 5.34 – Erreur moyenne au carré (RMSE) des charges hydrauliques au droit des 13 puits de l’essai de pompage 2 après une itération d’assimilation. Chaque couleur correspond à un schéma d’assimilation différent.

Pour l’essai de pompage 2 détaillé sur la Figure 5.34, les RMSE étaient majoritairement comprises entre 1 et 2 m, avec un maximum de 3 m sur le puits MW-10. Les assimilations de toutes les données en une seule fois, avec ou sans analyse locale, produisent à nouveau des RMSE supérieures à celle de l’ensemble initial. Cependant, la méthode brute avec une analyse locale de 500/20 permet d’améliorer la RMSE sur les puits F-120c, 7011 et MW-10. Cette amélioration se fait au détriment de la RMSE sur les autres puits, spécialement MW-5, MW-9 et S-8. Les assimilations puits par puits avec différentes portées d’analyse locale n’impactent quasiment pas les RMSE.

Les modifications drastiques au niveau de la RMSE des charges hydrauliques sur les nombreux puits indiquent que dans le cas de ce modèle complexe, une assimilation brute de tous les puits n’est pas capable d’améliorer l’ensemble. Uniquement les assimilations puits par puits permettent de ne pas diverger au bout d’une itération d’assimilation. Cependant, à cette étape, il n’est pas possible de conclure sur l’efficacité des méthodes d’assimilations puits par puits avant de conduire les 4 itérations d’assimilation.

À l'issue de cette première assimilation, il a été décidé d'éliminer les méthodes d'assimilation sans assimilation individuelles puits par puits. L'assimilation complète avec 4 itérations prenant environ 6 jours, il ne semblait pas logique de continuer l'assimilation avec les méthodes brutes, l'ensemble ayant déjà divergé au bout d'une itération. L'apport de la nouvelle méthode avec assimilation puits par puits et analyse locale permet de ne pas diverger au bout de la première itération d'assimilation. De plus, cette nouvelle méthode permet de mettre à jour les distributions de perméabilités en les conservant très proches de la distribution initiale.

De plus, l'assimilation de tous les paramètres et toutes les observations en une étape induit une matrice de gain de Kalman de 1094771 par plus de 4000 observations. Cette matrice seule occupe plus de 40Gb de mémoire. Cette même matrice est fortement réduite lors de l'assimilation puits par puits, sans localisation elle occupera au maximum 1094771 paramètres par 239 observations. Cette dimension peut encore être réduite en utilisant une analyse locale, réduisant le nombre de paramètres mis à jour autour du point d'observation.

### 5.7.2 Évaluation des performances des schémas d'analyse locale retenus

Pour l'évaluation finale, uniquement des schémas d'assimilation avec des assimilations puits par puits et analyse locale ont été retenus. Différentes portées d'analyse locale ont été étudiées.

Dans le même temps, un facteur multiplicatif de l'amplitude de la mise à jour a été défini. Les mises à jour reposent sur une moyenne de la mise à jour apportée par chaque puits assimilé. Ce qui peut résulter en une mise à jour très faible sur les paramètres d'états après une mise à jour de l'ensemble avec beaucoup de puits.

Les cas synthétiques étaient contraints par plus de types de données et par uniquement 3 puits assimilés par rapport au modèle de Mercier. Ce modèle plus complexe repose uniquement sur les charges hydrauliques et sur l'assimilation de 12 puits pour l'essai 1 et 13 pour l'essai 2. Dans ces conditions la mise à jour moyenne reposant sur l'apport de multiple puits propose une amplitude trop faible pour avoir un impact significatif sur l'ensemble en 4 itérations.

Lorsqu'un puits d'observation mets à jour la conductivité hydraulique dans son voisinage immédiat, celle-ci est atténué par l'apport moyen très faible provenant des puits d'observations plus éloignés. Le but est toujours de garder l'information provenant de tous les puits proches. La moyenne

des apports de chaque puits d'observations donnera toujours la direction de la mise à jour des paramètres d'états, mais un facteur multiplicatif permettra d'amplifier l'amplitude de cette mise à jour.

Le facteur d'amplification multiplicatif retenu est de 4. Ce nombre a été choisi au regard du cas de Mercier. La Figure 5.35 indique que dans le cas d'une analyse locale de 500/20, 61% des éléments sont mis à jour par 4 puits ou plus pour l'essai 1 et 50% pour l'essai 2.

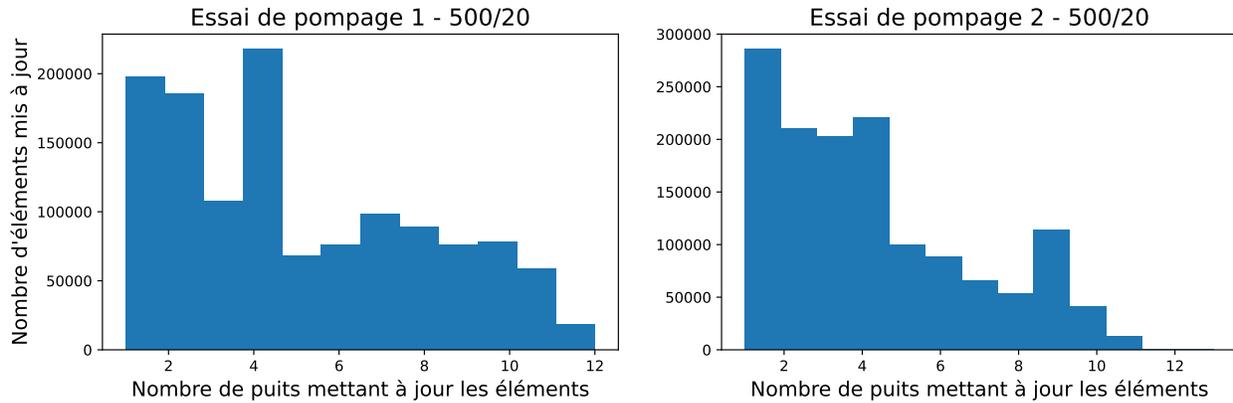


Figure 5.35

La Figure 5.36 indique que dans le cas d'une analyse locale de 1000/30, 93% des éléments sont mis à jour par 4 puits ou plus pour l'essai 1 et 92% pour l'essai 2.

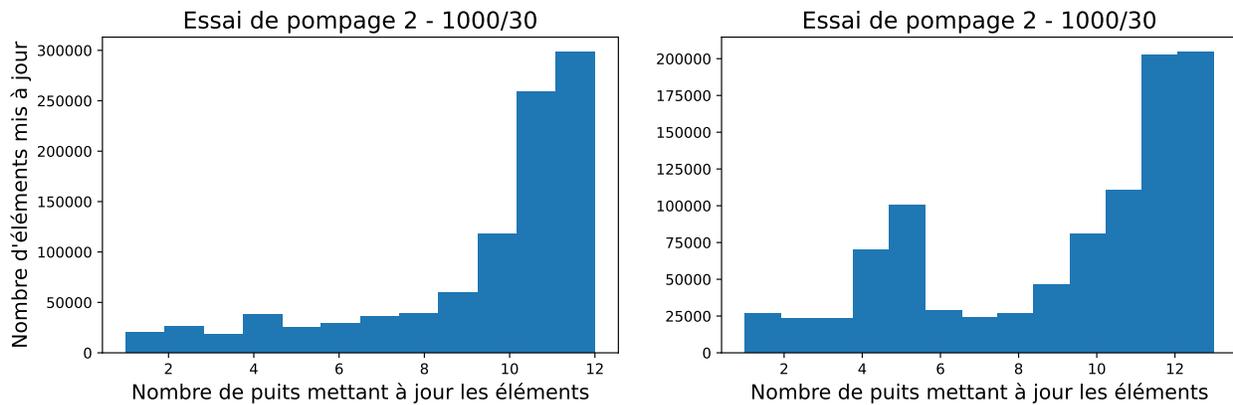


Figure 5.36

Le facteur de 4 a donc été retenu au regard du nombre de puits d'observations multiples mettant à jour les paramètres d'états de perméabilité. De plus, ce facteur de 4 n'a pas montré d'impact sur la stabilité numérique des simulations.

Les schémas d'analyse locale d'assimilations retenus pour effectuer les 4 itérations complètes ont été les suivants :

- Assimilation puits par puits, analyse locale 250/15 m;
- Assimilation puits par puits, analyse locale 250/15 m, facteur multiplicatif de 4;
- Assimilation puits par puits, analyse locale 500/20 m;
- Assimilation puits par puits, analyse locale 500/20 m, facteur multiplicatif de 4;
- Assimilation puits par puits, analyse locale 1000/30 m, facteur multiplicatif de 4.

Les schémas d'analyses locales de 250/15 m et 500/20 m présentaient des impacts négligeables sur les RMSE des deux essais sur les Figures 5.33 et 5.34 au bout d'une itération.

Chacune de ces assimilations sera présentée en détail dans les paragraphes suivants en détaillant la RMSE à chacune des itérations du processus d'assimilation.

L'assimilation avec le schéma 250/15 m ne présente aucune modification significative de la RMSE sur aucun puits. L'assimilation 500/20 présente de petites améliorations sur la moyenne de tous les puits à l'exception du PZ-111. Le schéma 250/15 amplifié indique des variations faibles par rapport aux valeurs initiales, à l'exception des puits PO-20-6D et PO-20-8D qui voient leurs RMSE moyennes augmenter de 0.5 m et 0.2 m. Le schéma 500/20 amplifié montre une augmentation de la RMSE de plus de 10cm sur quasiment l'ensemble des puits. Les puits PO-20-8D, PO-20-9D et F-106-A présentent une dégradation de la RMSE de plus de 50 cm. Enfin, le schéma 1000/30 amplifié indique lui aussi une augmentation moyenne de la RMSE de plus de 10 cm sur quasiment tous les puits. Les puits PO-20-6D, PO-20-7D, PO-20-8D et PO-20-9D indiquent une augmentation supérieure de la RMSE de 0.5 m par rapport à l'ensemble initial.

La variance est conservée pour tous les puits et schémas d'assimilation, à l'exception du puits PZ-111 avec la méthode 1000/30 amplifiée. Pour ce puits la variance est divisée par deux par rapport à l'ensemble initial.

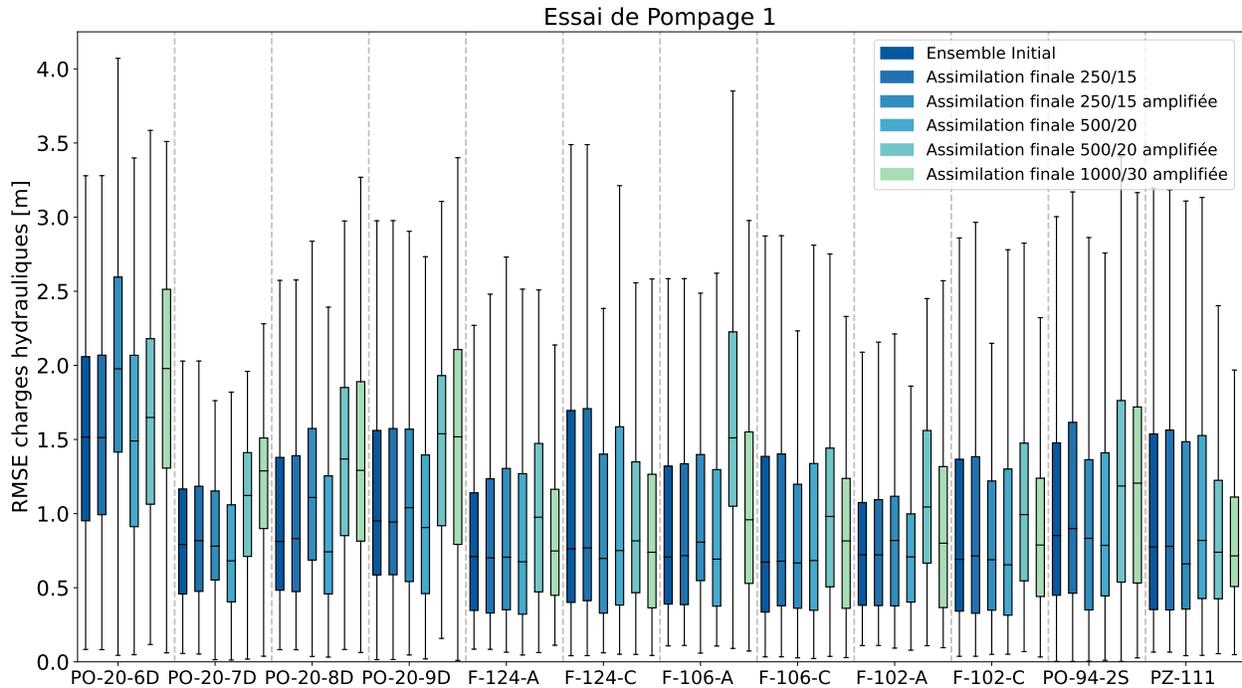


Figure 5.37 – RMSE des charges hydrauliques à la fin des 4 schémas d’assimilation pour les puits de l’essai de pompage 1. La RMSE de l’ensemble initial est rappelée pour évaluer la performance.

Les RMSE des puits de l’essai de pompage 2 sont données sur la Figure 5.38. À la fin du processus d’assimilation, le schéma 250/15 amplifié présente des mises à jour faibles par rapport à l’ensemble initial à l’exception des puits F-120C, 7011, MW-10 et Z11C avec des diminutions respectives de 10 cm, 30 cm, 15 cm et 10 cm. Une légère réduction de la variance est visible sur les puits S-8 et MW-10.

L’assimilation avec le schéma 250/15m ne présente aucune modification significative de la RMSE sur aucun puits.

Le schéma 500/20 indique une mise à jour de la RMSE plus faible que pour le schéma 250/15 amplifié. La majorité des puits est soumise à une amélioration inférieure à 5cm. Les puits F-120-C et 7011 ont vu leur RMSE diminuer de 10 cm. Le puits MW-10 a une RMSE plus élevée de 10 cm que l’ensemble initial. Les puits F-120-C et MW-11 ont subi une réduction de variance.

Le schéma 500/20 amplifié démontre une amélioration de la RMSE par rapport à l’ensemble initial supérieur à 50cm sur les puits F-120-C, 7011, MW-8 et MW-10. La RMSE a augmenté de plus de 10cm sur les puits MW-5, Z11A et Z11C. Le reste des puits a vu sa RMSE être modifiée

entre plus ou moins 10 cm par rapport à l'ensemble initial. La variance a été réduite légèrement sur les puits F-120-C, MW-8 et MW-10. Elle a été divisée par 3 sur le puits 7011. La variance a augmenté sur le puits Z11C.

Enfin, le schéma 1000/30 amplifié indique une augmentation de la RMSE supérieure à 10 cm sur les puits F-1, MW-5, MW-9, MW-10, MW-11, et Z11A. Une réduction supérieure à 10 cm est notée sur les puits F-120C, 7011 et MW-8. Le reste des puits a vu sa RMSE varier d'une quantité inférieure à 10cm. La variance a été légèrement réduite au droit des puits F-11, F-120-C, MW-5, MW-8 et Z11-C.

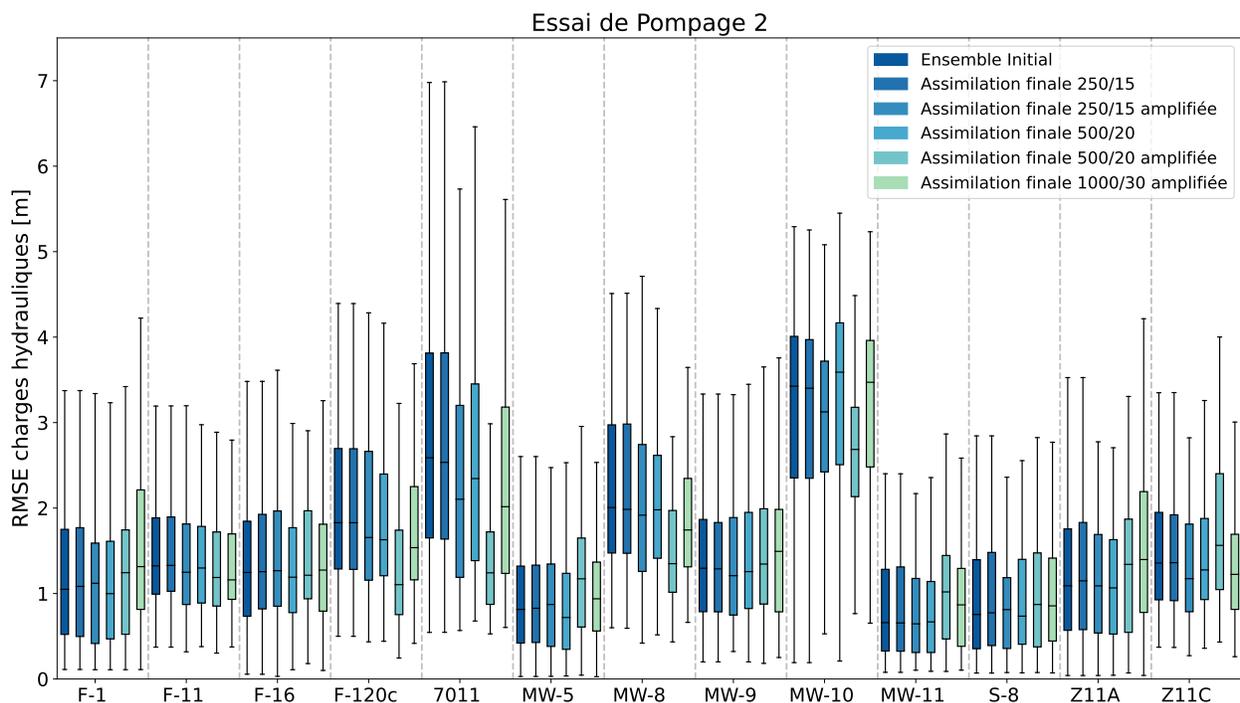


Figure 5.38 – RMSE des charges hydrauliques à la fin des 4 schémas d'assimilation pour les puits de l'essai de pompage 2. La RMSE de l'ensemble initial est rappelée pour évaluer la performance.

### Assimilation puits par puits, analyse locale 250/15m

Les Figures 5.39 et 5.40 montrent l'évolution de l'ensemble pour le schéma 250/15 amplifié à chaque itération du processus d'assimilation.

Aucun des puits ne présente d'évolution de la RMSE moyenne des charges hydrauliques, ni sur les puits de l'essai 1, ni sur les puits de l'essai 2.

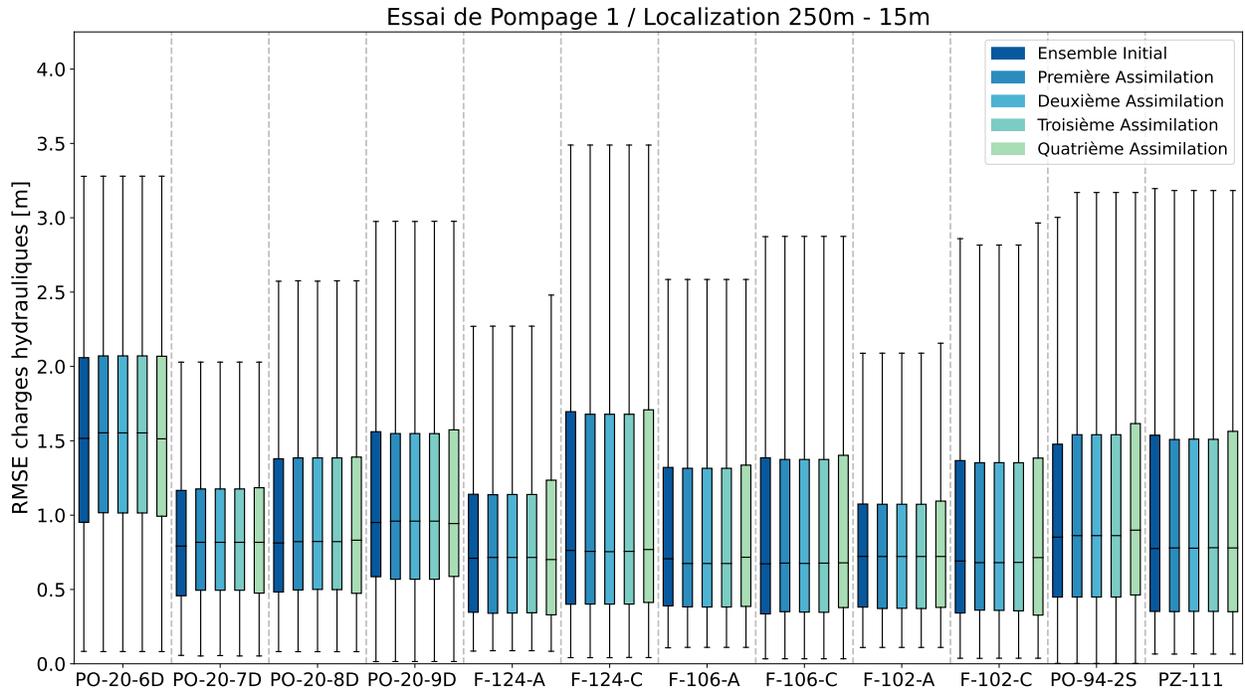


Figure 5.39 – RMSE des charges hydrauliques à chaque itération du processus d’assimilation pour les puits de l’essai de pompage 1 et le schéma 250/15.

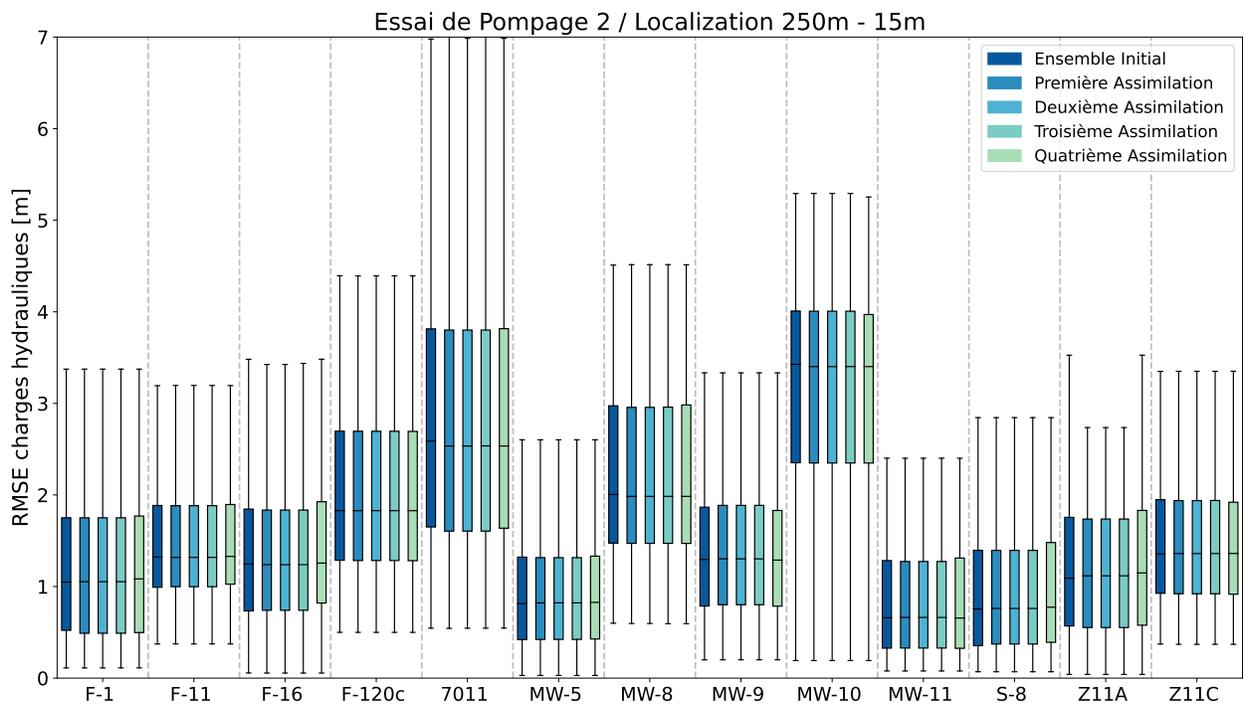


Figure 5.40 – RMSE des charges hydrauliques à chaque itération du processus d’assimilation pour les puits de l’essai de pompage 2 et le schéma 250/15.

La variance est restée identique à celle de l'ensemble initial pour tous les puits.

**Assimilation puits par puits, analyse locale 250/15m, facteur multiplicatif de 4**

Les Figures 5.41 et 5.42 montrent l'évolution de l'ensemble pour le schéma 250/15 amplifié à chaque itération du processus d'assimilation.

Les puits PO-20-6D, PO-20-8D, PO-20-9D et F-102-A indiquent une détérioration de la RMSE à chaque itération. Les puits PO-20-7D, F-124-A, F-106-C, F-102-C et PO-94-2S ne montrent pas d'évolution significative de la RMSE tout au long du processus d'assimilation. Les puits F-124-C et PZ-111 sont les deux puits qui montrent une légère réduction de la RMSE durant le processus d'assimilation.

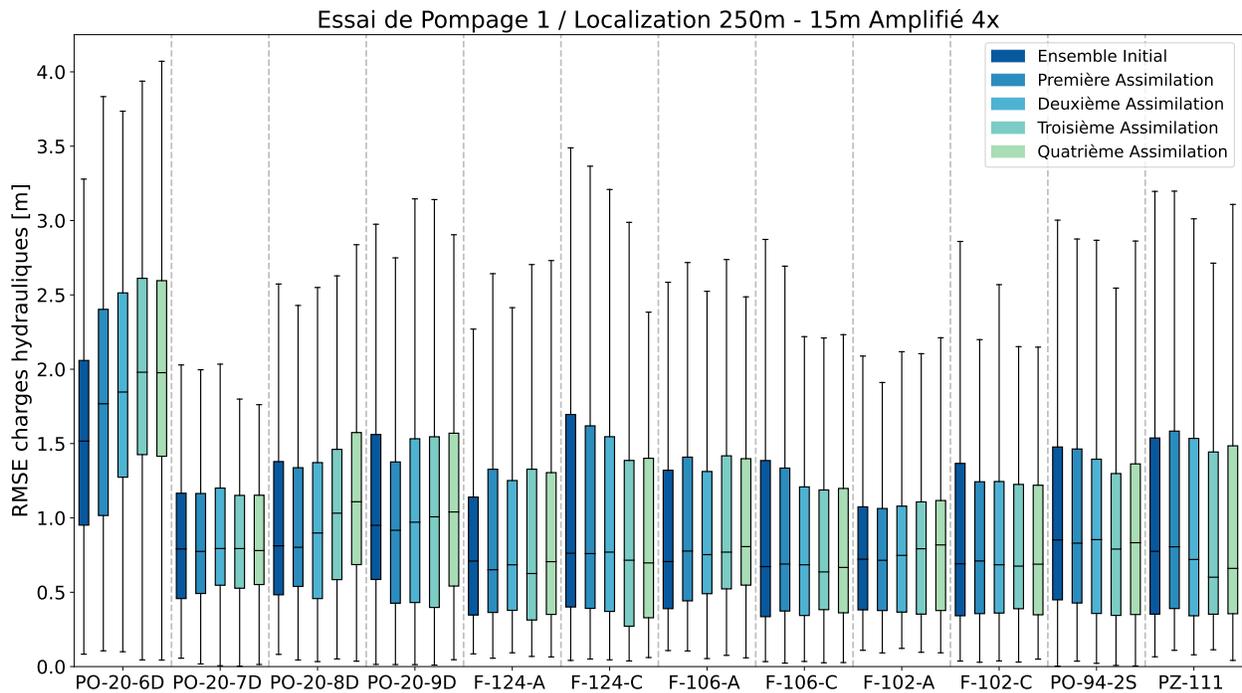


Figure 5.41 – RMSE des charges hydrauliques à chaque itération du processus d'assimilation pour les puits de l'essai de pompage 1 et le schéma 250/15 amplifié.

Les puits de l'essai de pompage 2 sont représentés sur la Figure 5.42. Une réduction de la RMSE est visible sur les puits F-120-C, 7011, MW-10 et Z11C, avec respectivement 15cm, 50cm, 20cm et 15cm. Les autres puits de l'ensemble ne présentent pas de variation significative de la RMSE entre l'ensemble initial et final.

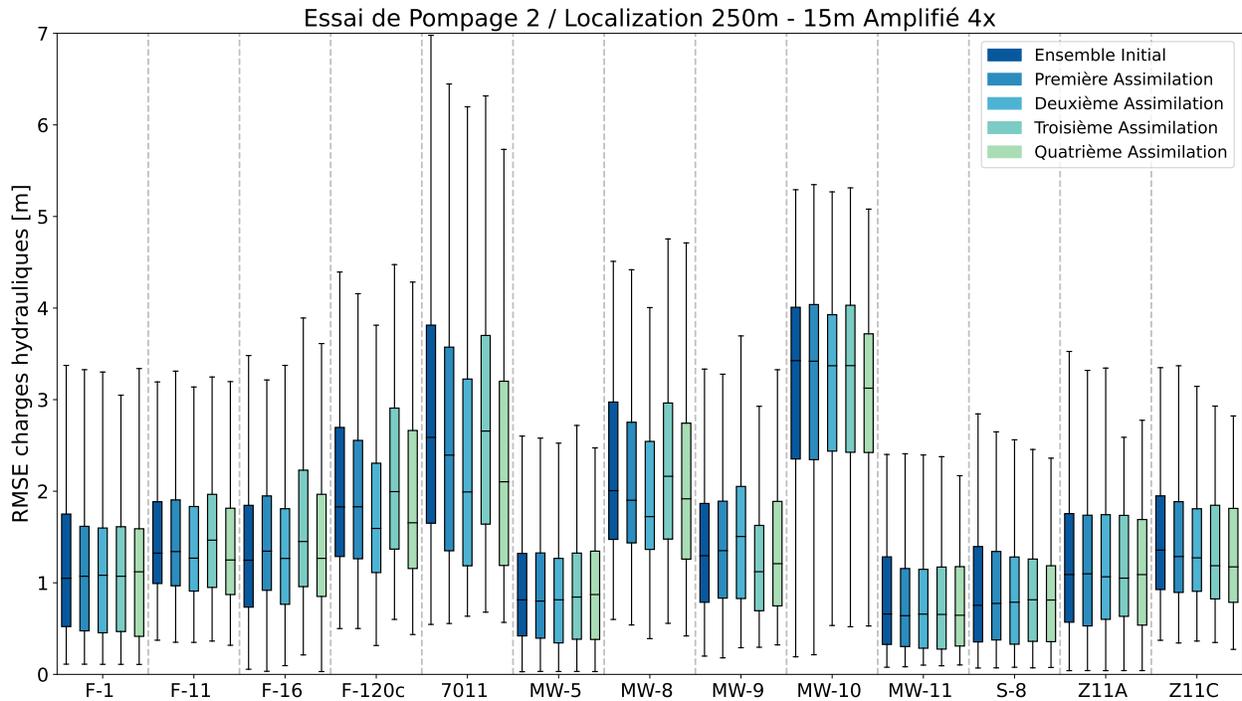


Figure 5.42 – RMSE des charges hydrauliques à chaque itération du processus d’assimilation pour les puits de l’essai de pompage 2 et le schéma 250/15 amplifié.

L’évolution de la RMSE est moins linéaire en chaque itération d’assimilation sur les puits de l’essai 2 que ceux de l’essai 1. Les directions de mise à jour des paramètres d’états sont moins définies à chaque itération pour l’essai 2. La variance est légèrement réduite sur les puits F-106-C et S-8, montrant que la variance de l’ensemble ne s’effondre pas lors du processus complet d’assimilation.

### Assimilation puits par puits, analyse locale 500/20m

La Figure 5.43 indique l’évolution de la RMSE à chaque itération d’assimilation du processus multiple pour les puits de l’essai 1. La RMSE moyenne a été réduite de quelques centimètres sur les puits PO-20-7D, PO-20-8D, PO-20-9D et PO-94-2S. Les autres puits ne présentent pas de variation significative dans la RMSE. Une légère réduction de la variance est visible sur les puits PO-20-8D, F-102-A, F-102-C et PO-94-2S. La variance a augmenté sur le puits F-124-A.

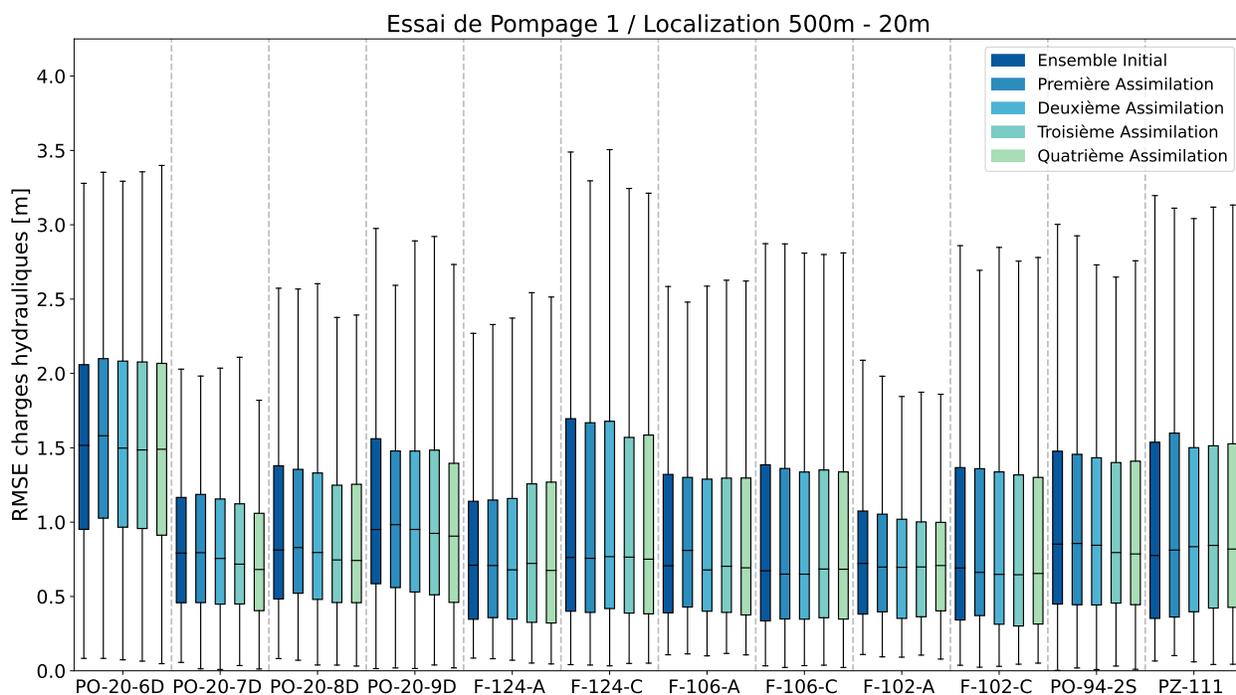


Figure 5.43 – RMSE des charges hydrauliques à chaque itération du processus d’assimilation pour les puits de l’essai de pompage 1 et le schéma 500/20.

La Figure 5.44 indique l’évolution de la RMSE à chaque itération d’assimilation du processus multiple pour les puits de l’essai 2. La RMSE moyenne sur les puits F-120C et 7011 a été réduite de plus de 15cm entre l’ensemble initial et final. La moyenne a été réduite de quelques centimètres sur les puits F-1, F-11, F-16, MW-5, MW-8, S-8, Z11A et Z11C. Cependant, ces réductions ne présentent pas d’amplitudes significatives. La RMSE du puits MW-10 a vu sa RMSE moyenne augmenter à chaque itération, pour atteindre une dizaine de centimètres de plus que la RMSE du puits dans l’ensemble initial.

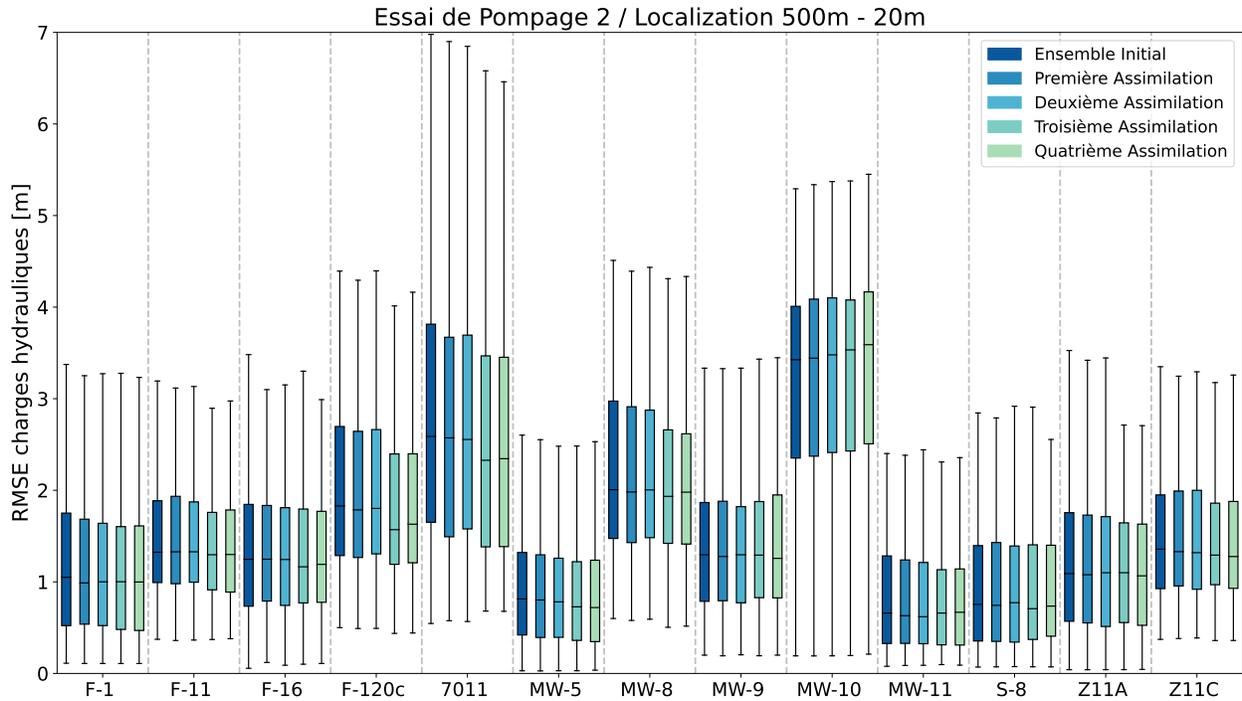


Figure 5.44 – RMSE des charges hydrauliques à chaque itération du processus d'assimilation pour les puits de l'essai de pompage 2 et le schéma 500/20.

Une légère réduction de la variance est notée sur les puits MW-8, MW-11 et Z-11C. Globalement, le schéma 500/20 conserve la variance de l'ensemble initial, indiquant une très bonne résilience à l'effondrement de la variance de l'ensemble.

#### Assimilation puits par puits, analyse locale 500/20m, facteur multiplicatif de 4

La Figure 5.45 indique l'évolution de la RMSE à chaque itération d'assimilation du processus multiple pour les puits de l'essai 1. La RMSE moyenne a augmenté sur l'ensemble des puits, à chaque itération à l'exception du puits PZ-111. La moyenne de la RMSE a augmenté de 16.3cm pour les puits de l'essai 1. L'augmentation la plus importante est d'environ 75cm sur le puits F-106-A. La variance est restée globalement identique tout au long du processus. Le puits F-106-A a vu sa variance de RMSE augmenter tandis que le puits PZ-111 a subi une réduction de variance.

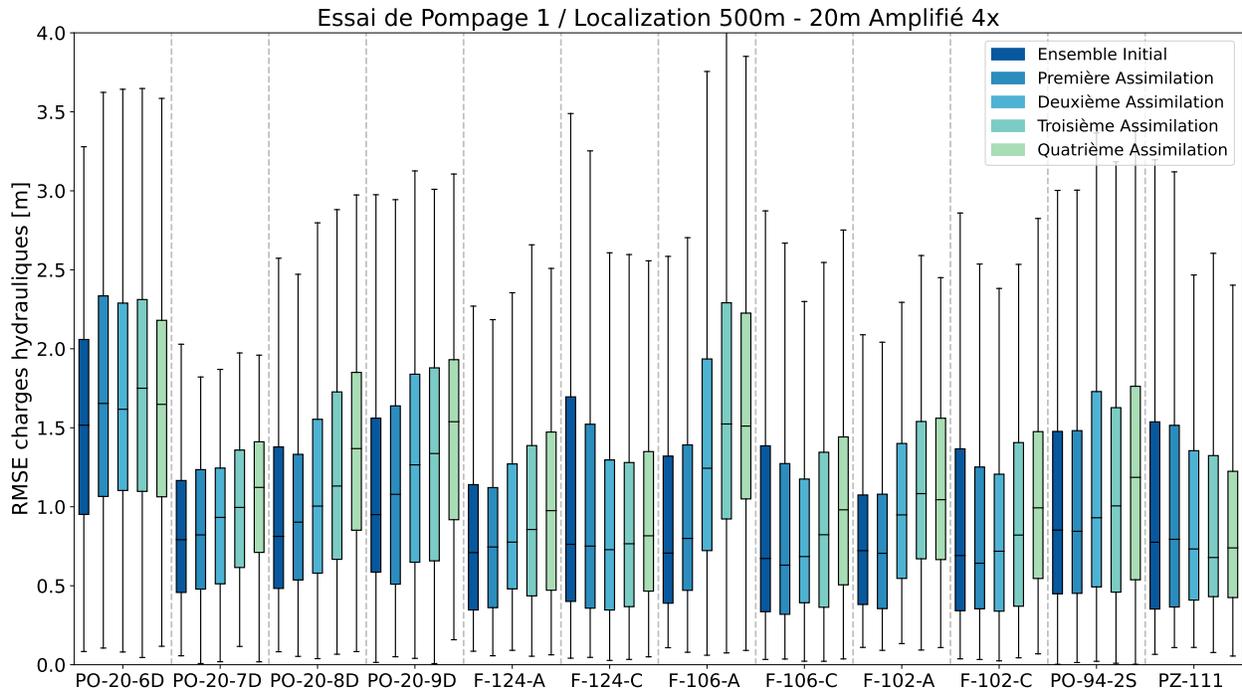


Figure 5.45 – RMSE des charges hydrauliques à chaque itération du processus d’assimilation pour les puits de l’essai de pompage 1 et le schéma 500/20 amplifié.

La Figure 5.46 indique l’évolution de la RMSE à chaque itération d’assimilation du processus multiple pour les puits de l’essai 2. Les puits F-1, MW-5, MW-11, S-8, Z11A et Z11C montrent une augmentation de la RMSE moyenne de plus de 10 centimètres entre l’ensemble initial et final. Les puits F-11 ne F-16 présentent pas d’évolution significative de la RMSE. La réduction de RMSE entre les ensembles initiaux et finaux atteint 75cm sur le puits F-120C, 125cm sur le puits 7011, 50 cm sur le puits MW-8 et 75cm sur le puits MW-10.

La variance des RSME de charges hydrauliques pour l’essai 2 à diminuée sur les puits F-120C, 7011, MW-8 et MW-10. Les autres puits ont conservé la variance de l’ensemble initial.

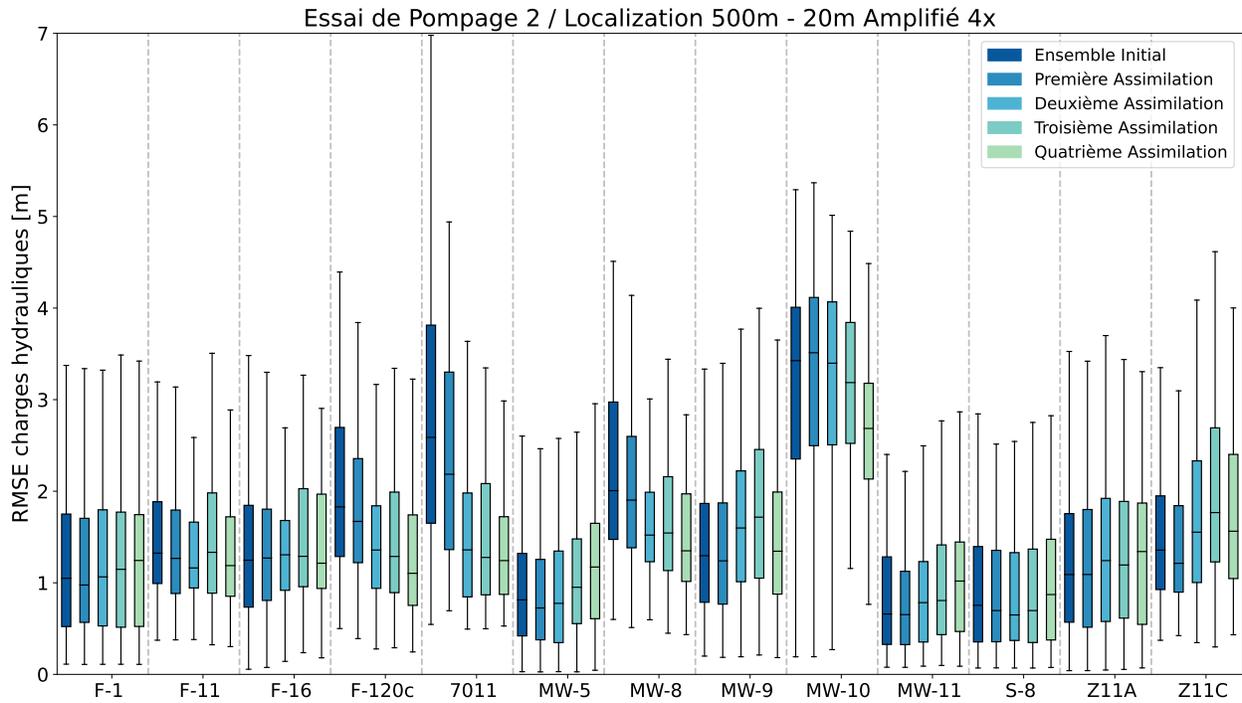


Figure 5.46 – RMSE des charges hydrauliques à chaque itération du processus d’assimilation pour les puits de l’essai de pompage 2 et le schéma 500/20 amplifié.

#### Assimilation puits par puits, analyse locale 1000/30m, facteur multiplicatif de 4

La Figure 5.47 montre les charges hydrauliques à chaque itération de l’assimilation pour le schéma 1000/30 amplifié pour tous les puits de l’essai 1. La RMSE augmentait à chaque itération pour tous les puits à l’exception des puits F-124-A, F-124-C et PZ-111. L’augmentation moyenne de RMSE sur tous les puits est de 10.3cm. Elle est supérieure à 50 cm sur les puits PO-20-8D et PO-94-2S. La variance est restée identique à celle de l’ensemble initial pour tous les puits, sauf le puits PZ-111 qui montre une réduction de variance.

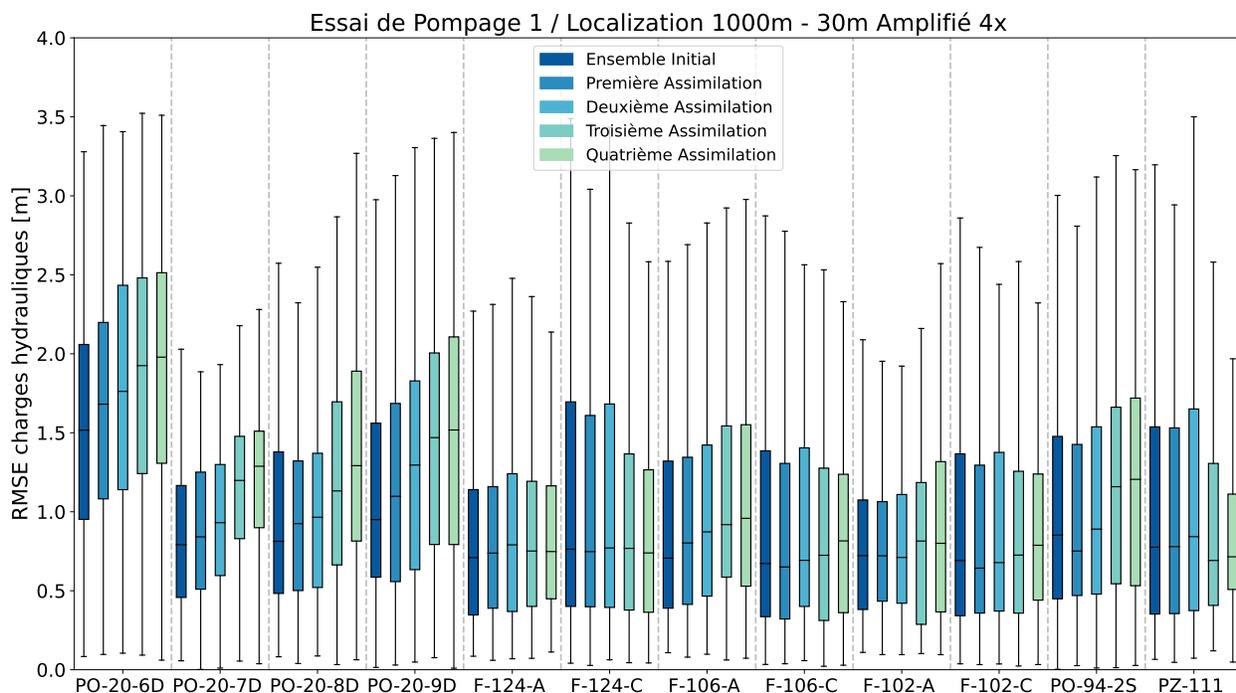


Figure 5.47 – RMSE des charges hydrauliques à chaque itération du processus d'assimilation pour les puits de l'essai de pompage 1 et le schéma 1000/30 amplifié.

La Figure 5.47 montre les charges hydrauliques à chaque itération de l'assimilation pour le schéma 1000/30 amplifié pour tous les puits de l'essai 2. La RMSE moyenne a augmenté de plus de 10 cm entre les ensembles initiaux et finaux sur les puits F-1, MW-5, MW-9, MW-11 et Z11A. La RMSE moyenne a diminué de plus de 10 cm entre les ensembles initiaux et finaux sur les puits F-120C, 7011, MW-8 et Z11C. La réduction atteinte 50cm sur le puits 7011. La variance sur tous les puits est restée comparable à la variance de l'ensemble initial.

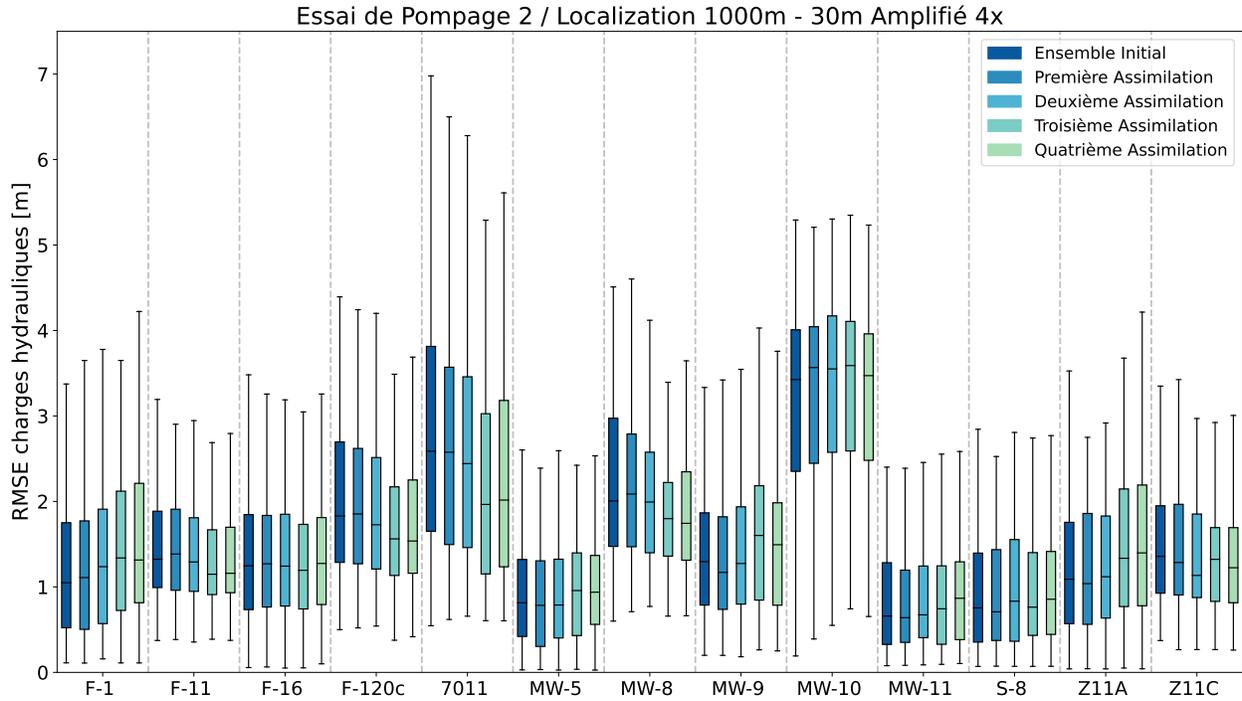


Figure 5.48 – RMSE des charges hydrauliques à chaque itération du processus d’assimilation pour les puits de l’essai de pompage 2 et le schéma 1000/30 amplifié.

Un résumé des performances finales de ces 5 schémas est donné ci-dessous dans le Tableau 5.1. La RMSE moyenne de l’essai 1 pour l’ensemble initial était de 1.05 m, celle pour l’essai 2 était de 1.66 m. Toutes les assimilations présentent des difficultés pour améliorer l’essai 1, avec uniquement le schéma sans amplification 500/20 m arrivant à réduire la RMSE. Tous les schémas ont réussi à améliorer l’essai 2 à l’exception du schéma d’analyse 250/15 m. L’essai 1 présentant une RMSE plus faible dans l’ensemble initial pourrait expliquer ce comportement. Les schémas 250/15 m 4x et 500/20 sont ceux qui présentent le meilleur compromis sur les deux essais. L’amélioration maximale sur l’essai 2 est obtenue avec le schéma 500/20 m 4 avec une RMSE réduite de 22.6 cm. En contrepartie la RMSE moyenne sur l’essai 1 a augmenté de 16.3 cm.

**Tableau 5.1 – Synthèse de l'évolution des RMSE après la dernière assimilation par rapport à l'ensemble initial pour chacun des schémas étudiés.**

Schémas	Évolution	Évolution	Évolution	Évolution	RMSE	RMSE
	RMSE	RMSE	RMSE	RMSE	moyenne	moyenne
	essai 1	essai 2	essai 1	essai 2	essai 1	essai 2
250/15m	+1.94 %	+0.99 %	+1.9 cm	+0.8 cm	1.07 m	1.67 m
250/15m 4	+1.07 %	-5.93 %	+1.8 cm	-11.3 cm	1.07 m	1.55 m
500/20m	-3.79 %	-3.08 %	-3.8 cm	-5.1 cm	1.01 m	1.61 m
500/20m 4	+16.97 %	-6.80 %	+16.3 cm	-22.6 cm	1.21 m	1.43 m
1000/30m 4	+10.14 %	-1.06 %	+10.3 cm	-6.2 cm	1.15 m	1.6 m

Les résultats de ces assimilations sur la RMSE des charges hydrauliques après 4 assimilations sont donnés sur la Figure 5.33 pour l'essai 1. La RMSE moyenne de l'essai 1 pour l'ensemble initial était de 1.05 m, celle pour l'essai 2 était de 1.66 m.

## 5.8 Comparaison avec la méthode PESTPP-IES existante

Une comparaison a été réalisée avec PESTPP-IES à partir des mêmes modèles FEFLOW , et des mêmes observations (White *et al.*, 2020). PESTPP-IES est un algorithme d'Ensemble Smoother localisé développé par l'United States Geological Survey (USGS). Il permet de travailler directement sur des modèles FEFLOW.

Cette comparaison permet de valider les apports de la méthode développée dans ce doctorat. Les RMSE par rapport à l'ensemble initial en relatif sont données dans le Tableau 5.2.

**Tableau 5.2 – Synthèse de l'évolution des RMSE relatives en pour cent par rapport à l'ensemble initial. Comparaison entre la méthode développée dans ce doctorat et la méthode PESTPP-IES.**

	Évolution en % de la RMSE Ensemble final 500/20 Amplifié	Évolution en % de la RMSE PESTPP-IES
PO-20-6D	-29.04	-64.60
PO-20-7D	-22.11	568.75
PO-20-8D	-11.45	12.05
PO-20-9D	-35.84	-81.29
F-124-A	-16.27	-45.45
F-124-C	-11.46	126.89
F-106-A	7.40	-62.29
F-106-C	-9.82	15.08
F-102-A	-14.48	-51.46
F-102-C	-14.21	74.16
PO-94-2S	-7.38	41.27
PZ-111	-8.06	-34.50
F-1	-8.52	0.24
F-11	-29.62	-24.55
F-16	-1.97	-78.20
F-120c	-61.49	-34.85
7011	-44.85	-34.90
MW-5	-8.23	-2.05
MW-8	-42.78	-30.91
MW-9	-3.23	0.12
MW-10	-7.20	2.61
MW-11	-11.82	-6.27
S-8	-8.44	-3.22
Z11A	-8.37	-1.74
Z11C	-32.36	-31.23
Médiane	-11.46	-6.27
Moyenne	-17.66	10.15

Cette comparaison a été menée sur les charges relatives aux puits lors des essais de pompage. L'étude des charges relatives uniquement permettait de contrôler la réponse dynamique des systèmes, et non plus la somme des réponses dynamiques et statiques avec la valeur absolue imposée par la moyenne. Si les paramètres d'état sont proches des paramètres réels, la réponse dynamique devrait être proche des mesures corrigées des valeurs à l'instant initial de l'essai.

La méthode nouvellement développée permet une amélioration des charges hydrauliques relatives lors des deux essais de pompage. Ces améliorations sont présentes sur l'ensemble des puits, à l'exception du puits F106-A qui présente une détérioration de 7.4% par rapport à l'ensemble initial. L'amélioration moyenne est de 17.66% et la réduction médiane de la RMSE est de -11.46%

Les résultats ont été comparés avec la méthode itérative d'ensemble implémentée dans le logiciel PESTPP-IES. La médiane est légèrement moins bonne avec cette méthode, avec une amélioration de 6.27%. Cependant, la moyenne sur la totalité des puits est une détérioration de 10.15%. Cette détérioration est due à deux puits, PO-20 - 7D avec une détérioration de 568.75% et F-124-C avec 126.89%.

La différence pourrait provenir du manque d'analyse locale dans PESTPP-IES, ou dans la résolution différente des mises à jour de paramètres d'états. L'analyse locale permet de contraindre une mise à jour dans la bonne direction pour la quasi-intégralité des puits avec la nouvelle méthode développée d'analyse puits par puits et type de données par type de données.

Les réalisations individuelles de quelques puits sont données sur les Figures suivantes.

Les puits des Figures 5.49 et 5.50 indiquent une dynamique différente lors de l'assimilation entre la méthode développée et celle de PESTPP-IES. Il faut cependant souligner que les valeurs d'erreurs étant de l'ordre de quelques centimètres, ces dernières tombent dans les incertitudes numériques de simulation.

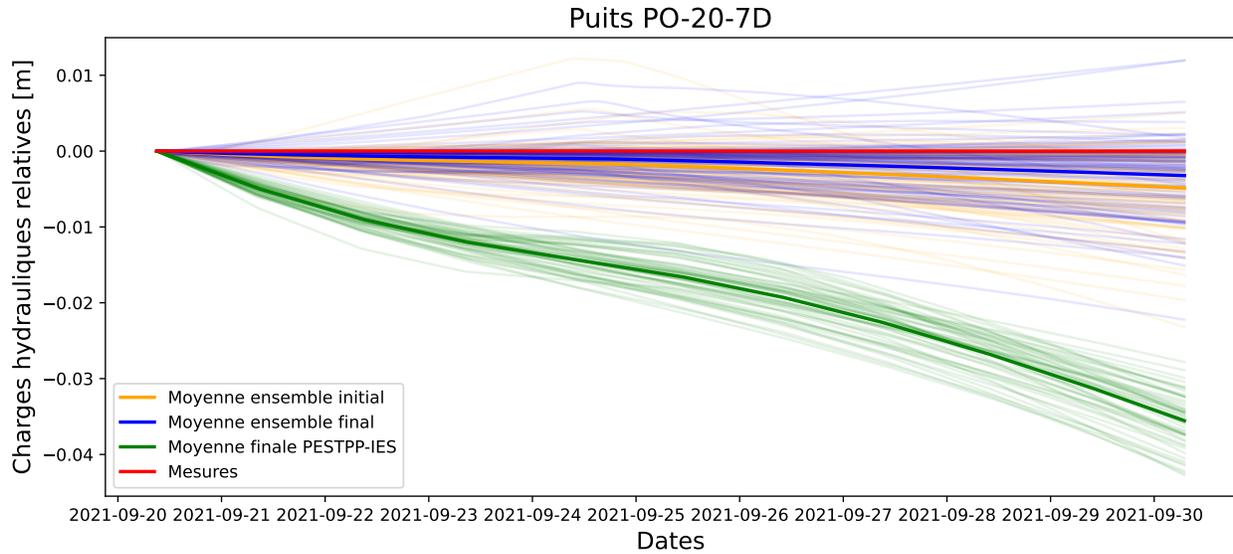


Figure 5.49 – Charges hydrauliques relatives par rapport à la valeur initiale au début de l'essai. Mesure en rouge, ensemble initial en orange, ensemble final de la méthode 500/20 4x en bleu, et ensemble final PESTPP-IES en vert.

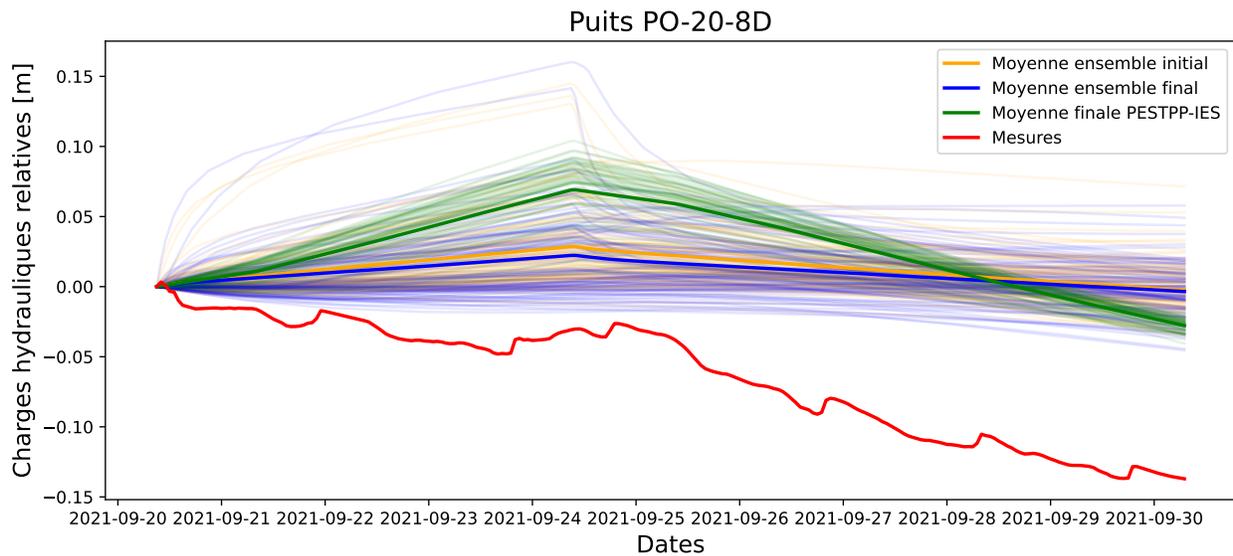


Figure 5.50 – Charges hydrauliques relatives par rapport à la valeur initiale au début de l'essai. Mesure en rouge, ensemble initial en orange, ensemble final de la méthode 500/20 4x en bleu, et ensemble final PESTPP-IES en vert.

Le puits de la Figure 5.51 montre une meilleure assimilation avec la méthode PESTPP-IES. Cependant, cette assimilation dégrade très fortement les charges hydrauliques sur la Figure 5.52. Les deux points d'observations F-124-A et F-124-C sont situés sur un même puits, avec le C

dans le roc et le A dans l'esker. Sans analyse locale, l'assimilation de PESTPP-IES a privilégié une fenêtre d'observations par rapport à l'autre. La méthode développée a permis d'améliorer les deux fenêtres d'observations.

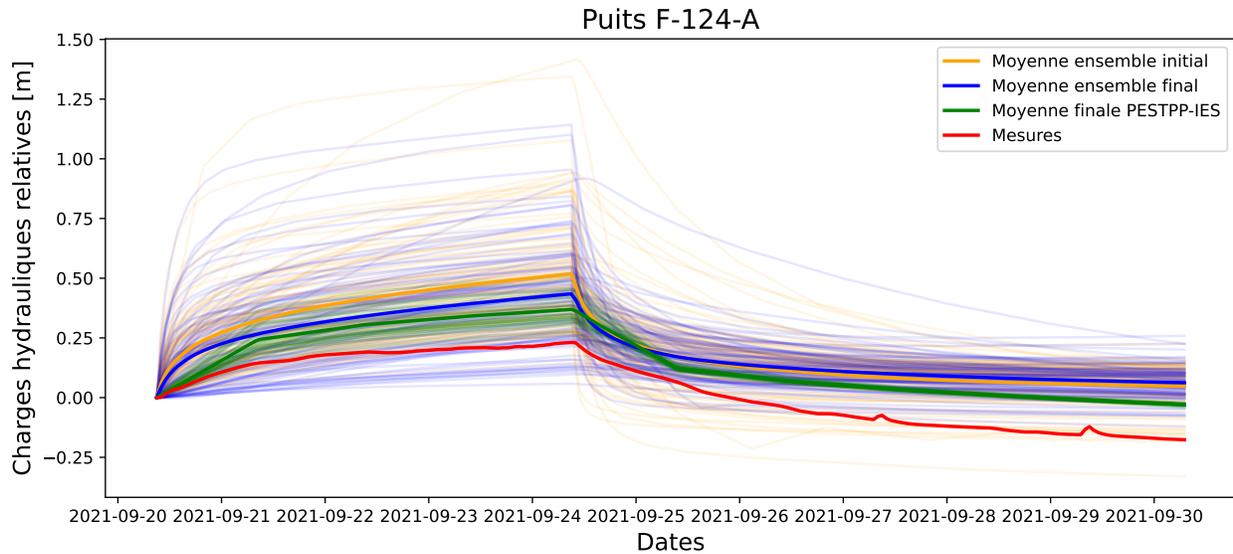


Figure 5.51 – Charges hydrauliques relatives par rapport à la valeur initiale au début de l'essai. Mesure en rouge, ensemble initial en orange, ensemble final de la méthode 500/20 4x en bleu, et ensemble final PESTPP-IES en vert.

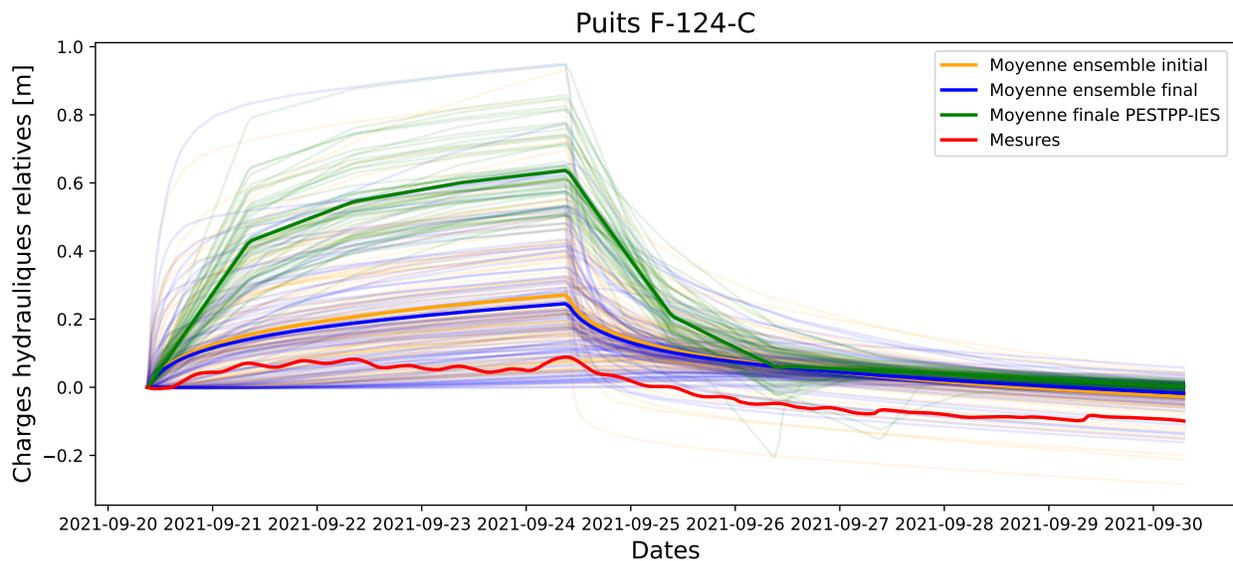
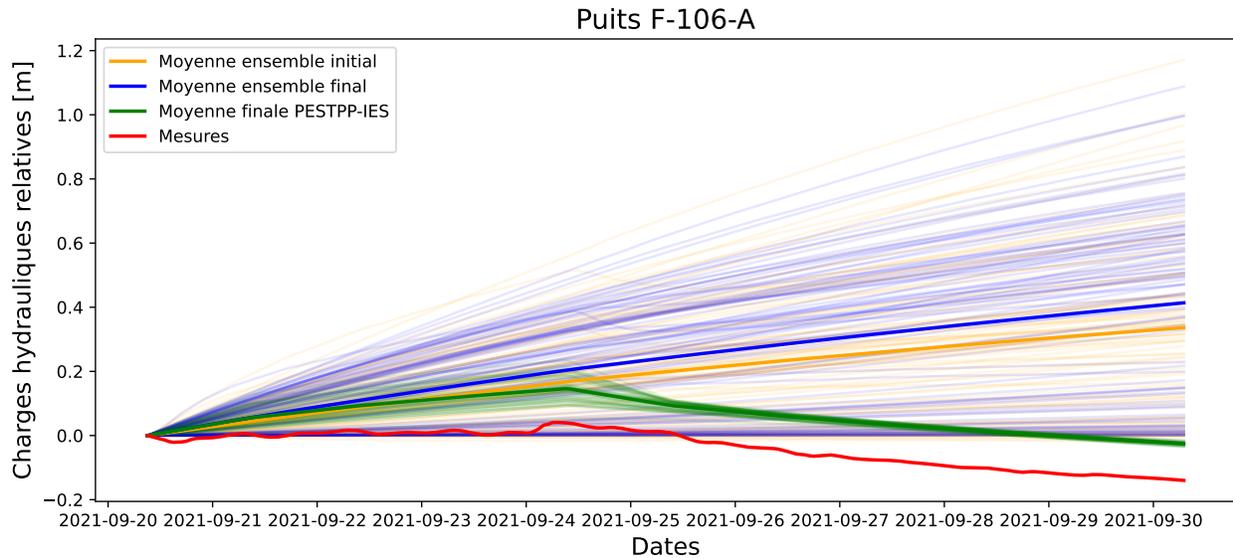


Figure 5.52 – Charges hydrauliques relatives par rapport à la valeur initiale au début de l'essai. Mesure en rouge, ensemble initial en orange, ensemble final de la méthode 500/20 4x en bleu, et ensemble final PESTPP-IES en vert.

La Figure 5.53 montre l'assimilation de meilleure qualité grâce à la méthode PESTPP-IES. La nouvelle méthode n'a pas permis d'améliorer ce puits.



**Figure 5.53 – Charges hydrauliques relatives par rapport à la valeur initiale au début de l'essai. Mesure en rouge, ensemble initial en orange, ensemble final de la méthode 500/20 4x en bleu, et ensemble final PESTPP-IES en vert.**

La Figure 5.54 décrit le puits PZ-111. Ce dernier présente une RMSE plus faible d'après le Tableau 5.2 pour la méthode PESTP-IES. Cependant, il est possible de voir que la méthode PESTPP-IES a perdu la réponse dynamique au pompage. La méthode développée a permis de conserver cette dynamique.

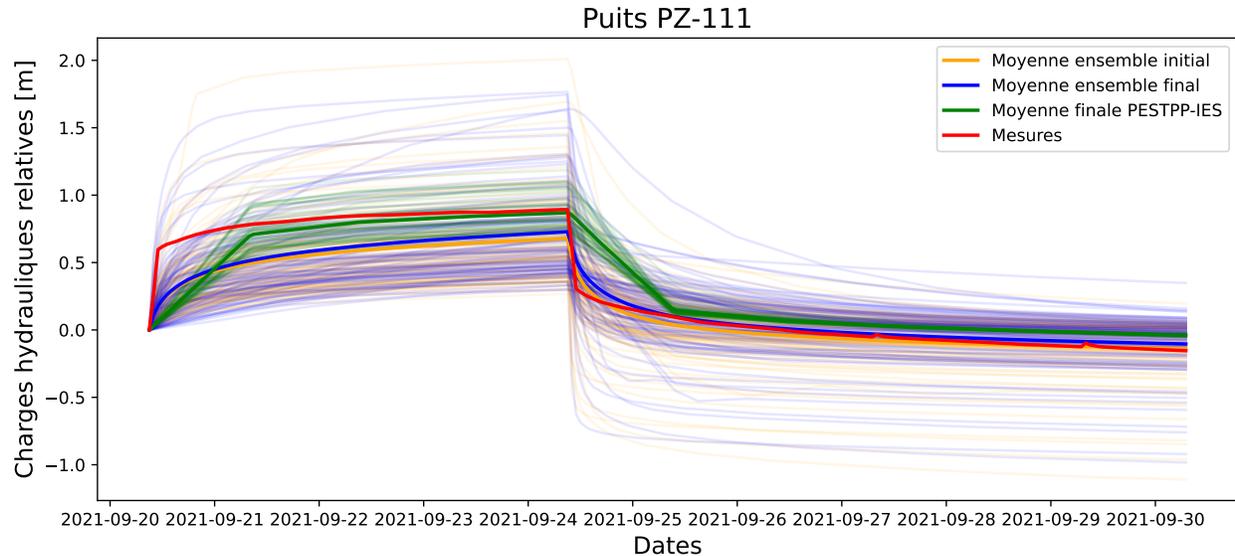


Figure 5.54 – Charges hydrauliques relatives par rapport à la valeur initiale au début de l’essai. Mesure en rouge, ensemble initial en orange, ensemble final de la méthode 500/20 4x en bleu, et ensemble final PESTPP-IES en vert.

L’ensemble des Figures indique que la méthode PESTPP-IES a tendance à provoquer un effondrement de la variance pour les ensembles. Les résultats affichés sont obtenus après une itération seulement, une deuxième itération provoque l’effondrement complet. La méthode développée permet donc une meilleure réduction de la RMSE moyenne et médiane, tout en conservant une variance dans l’ensemble pour évaluer les incertitudes.

## 5.9 Discussion

Tous les schémas testés avec la méthode développée présentent de très bonnes résiliences contre l’effondrement de la variance. Ce problème d’effondrement de variance est observé lors de l’utilisation de méthodes d’ensemble dans la littérature. Les schémas évalués parviennent à conserver la variance de l’ensemble initial.

Tous les schémas présentent une assimilation de meilleure qualité sur les puits de l’essai 2 que ceux de l’essai 1. Le Tableau 5.1 reprend l’évolution de la RMSE entre l’ensemble initial et final pour les 4 schémas. Cette différence pourrait provenir d’une covariance plus faible entre les paramètres d’états et les observations du puits de l’essai 2. Elle pourrait aussi provenir du fait que l’essai 1

avait une RMSE moyenne initiale de 1.05 m, contre 1.66 m sur l'essai 2. Il était donc plus facile d'améliorer les observations sur les puits de l'essai 2 que ceux de l'essai 1.

Le facteur d'amplification permet d'améliorer les observations les plus éloignées des mesures. En contrepartie, cette amélioration impose une modification à l'ensemble des puits alentour, altérant la RMSE par rapport à une assimilation sans amplification. Il reviendra donc à l'utilisateur de choisir s'il préfère une amélioration légère moyenne de tout le puits, ou une amélioration centrée sur les puits les plus éloignés des mesures.

Le schéma 250/15 conserve 103 modèles sur les 106 de l'ensemble initial. Pour le schéma 500/20, il reste 99 modèles dans l'ensemble à la fin des 4 assimilations. Les schémas 500/20 amplifié et 1000/30 amplifié conservent 97 modèles dans les ensembles. Le schéma d'analyse locale 250/15 amplifié quant à lui conserve 94 modèles à la fin du processus d'assimilation. Ces réductions dans les dimensions d'ensembles sont respectivement de 2.8%, 6.6%, 8.5% et 11.3%. Ces réductions, bien qu'existantes, restent limitées et permettent de conserver environ 100 modèles comme souhaité lors de la construction des ensembles. Ces ensembles conservent assez de modèles pour explorer l'incertitude à la fin du processus d'assimilation.

Les modèles éliminés dans les ensembles sont ceux qui ne convergent plus numériquement après une légère modification des paramètres d'états. Ce sont donc des modèles qui étaient équiprobables au niveau des paramètres d'états, mais ceux-ci ne permettaient pas d'honorer de manière stable les observations dynamiques induites par des essais de pompage. Il est important de détecter ces modèles, car ceux-ci ne permettent pas d'honorer les mesures de terrain.

La nouvelle méthode développée à partir d'une assimilation puits par puits avec analyse locale permet d'améliorer les charges hydrauliques. Cette amélioration passe par une détection des modèles numériquement instable et une modification des paramètres d'état. Les ensembles finaux sont donc numériquement plus stables et les valeurs de charges simulées sont plus proches que celles mesurées sur le terrain. La portée de l'analyse locale est un critère crucial pour la qualité de l'assimilation. Dans le cas d'un voisinage trop faible et sans facteur d'amplification, la mise à jour des éléments est trop faible pour avoir un impact sur les charges hydrauliques. Lorsque la portée d'analyse locale est trop importante, même avec amplification, la mise à jour détériore la RMSE moyenne de l'ensemble si les deux essais sont considérés. Les schémas 250/15 m amplifiés, 500/20 m amplifié et 500/20 m parviennent à améliorer les ensembles lors du processus d'assimilation. Sans amplification

l'amélioration est faible, mais présente sur les puits des deux essais. Avec amplification, la RMSE des puits de l'essai 2 est fortement réduite au détriment de la RMSE des puits de l'essai 1.

La création de l'ensemble initial de modèles à partir des modèles conceptuels et des données précédentes est extrêmement importante. Il faut que l'ensemble se distribue le mieux possible autour des mesures, ceci assurera une assimilation de données plus efficace.

Le résultat final dynamique en partant d'un modèle régional calibré en statique atteint une RMSE d'environ 1.3 m. La RMSE a été améliorée de manière plus visible sur les puits de l'essai 2 que sur ceux de l'essai 1. Cependant la RMSE des puits de l'essai 1 était majoritairement inférieure à 1 m dans l'ensemble initial, traduisant un travail important de modélisation en amont. La RMSE de l'ensemble initial pour l'essai 2 était plus importante, avec une valeur dépassant 2 m au puits 7011 et une dépassant 3 m au puits MW-10. L'essai 2 laissait donc plus de place à l'amélioration par rapport à l'essai 1.

Les résultats sur les charges hydrauliques à la fin du processus d'assimilation sont satisfaisants au vu de la complexité du modèle et du nombre de puits à calibrer.

Dans le futur, l'objectif serait d'étudier la possibilité de créer un ensemble initial à partir de multiples variogrammes, traduisant l'incertitude sur les propriétés spatiales des unités géologiques, et non plus uniquement à partir de simulations géostatistiques basées sur un variogramme par unité géologique.

Il serait également intéressant de tester un voisinage adaptatif, qui varierait à chaque itération de l'assimilation multiple. Il serait possible d'envisager une diminution de la portée d'analyse locale à chaque itération. Chaque itération de l'assimilation multiple permettrait de capturer une information à différente résolution spatiale.



# Conclusion

L'objectif de ce doctorat était de développer une méthode d'assimilation d'ensemble applicable à des problématiques hydrogéologiques. La littérature avait montré l'efficacité des méthodes d'ensembles dans le domaine pétrolier pour mettre à jour les modèles géologiques de réservoirs. Cependant, dans le domaine hydrogéologique, les filtres de Kalman n'avaient pas démontré d'efficacité réelle. Cette différence de résultats pouvait provenir de la quantité d'information disponible pour mettre à jour les modèles.

La méthode développée dans ce doctorat repose sur deux principes. Dans un premier temps, une assimilation puits par puits afin de conserver l'amplitude relative de chaque type d'anomalie observée. Dans un second temps, un schéma d'analyse locale afin de contraindre un sous-ensemble de paramètres d'états à assimiler. L'idée derrière cette combinaison était de réduire la nécessité d'utiliser une localisation purement mathématique appliquée sur la matrice de covariance. La combinaison de ces deux apports a été démontrée sur le cas synthétique par rapport aux méthodes d'assimilation d'ensemble classique.

Les résultats étaient prometteurs sur le cas synthétique, la nouvelle méthode développée permettait une amélioration des performances par rapport au schéma classique. Les RMSE de concentration étaient également améliorées en assimilant uniquement les observations de charges hydrauliques. L'analyse locale développée, combinée à l'assimilation puits par puits, fonctionnait très bien sur un cas synthétique de 3 puits d'observations. Cependant, rien ne garantissait que cette nouvelle méthode marcherait sur un modèle réel plus complexe.

Le modèle des lagunes de Mercier comportait 24 puits d'observations à calibrer, avec une complexité de modèle d'écoulement bien supérieur. Le modèle de Mercier comportait 1094771 valeurs

de perméabilité ajustable contre 68490 pour le cas synthétique. Les performances ne pouvaient évidemment pas être similaires.

Malgré tout, la méthode développée a permis d'améliorer les ensembles. Sans analyse locale et assimilation puits par puits, les ensembles divergeaient au bout d'une itération de l'assimilation multiple. La méthode développée dans ce doctorat a donc permis d'améliorer la méthode initialement développée par Raanes (2015a) pour les cas de réservoir pétrolier sur une problématique hydrogéologique. Les conclusions des cas synthétiques et réels permettent de définir deux recommandations pour la définition du schéma d'analyse locale. La première est de choisir une portée d'analyse locale de l'ordre de grandeur de la portée des variogrammes au minimum. La seconde est de définir une portée suffisante pour mettre à jour chaque paramètre physique du modèle au moins une fois, l'idéal étant d'avoir une couverture multiple provenant de plusieurs assimilations.

La qualité de l'assimilation dépend de la qualité de l'ensemble initial. L'ensemble initial doit proposer des modèles qui entourent les mesures de terrain. Si des mesures ne sont pas comprises dans l'ensemble, ce dernier n'arrivera que très difficilement, ou même pas du tout, à corriger les modèles. Les étapes de construction des modèles conceptuels, géostatistique et hydrogéologiques sont donc toujours importantes. L'apport de la méthode d'ensemble développée sera de pouvoir quantifier l'incertitude sur les prédictions hydrauliques.

# Références

- Altaf MU, Butler T, Mayo T, Luo X, Dawson C, Heemink AW & Hoteit I (2014). A Comparison of Ensemble Kalman Filters for Storm Surge Assimilation. *Monthly Weather Review*, 142(8):2899–2914. DOI:10.1175/MWR-D-13-00266.1.
- Anderson JL & Anderson SL (1999). A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts. *Monthly Weather Review*. DOI:10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2.
- BAPE (1994). Rapport d'enquête et d'audience publique: Restauration du lieu contaminé de Mercier. Bureau d'Audiences Publiques sur l'Environnement. 241 pages.
- Bauer P, Thorpe A & Brunet G (2015). The quiet revolution of numerical weather prediction. *Nature*, 525(7567):47–55. DOI:10.1038/nature14956.
- Brun J & Brisebois D (1994). La plate-forme du Saint-Laurent et les Appalaches. *Géologie du Québec. Les Publications du Québec*.
- Calcagno P, Chilès JP, Courrioux G & Guillen A (2008). Geological modelling from field data and geological knowledge: Part I. Modelling method coupling 3D potential-field interpolation and geological rules. *Physics of the Earth and Planetary Interiors*, 171(1):147–157. DOI:10.1016/j.pepi.2008.06.013.
- Carrassi A, Bocquet M, Bertino L & Evensen G (2018). Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *WIREs Climate Change*, 9(5):e535. DOI:10.1002/wcc.535.
- Caumon G, Gray G, Antoine C & Titeux T (2013). Three-dimensional implicit stratigraphic model building from remote sensing data on tetrahedral meshes : theory and application to a regional model of La Popa Basin, NE Mexico - CRPG. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 51(3):1613.
- Chen Y & Oliver D (2017). Localization and regularization for iterative ensemble smoothers. *Computational Geosciences*, 21:1–18. DOI:10.1007/s10596-016-9599-7.
- Chen Y & Oliver DS (2013). Levenberg–Marquardt forms of the iterative ensemble smoother for efficient history matching and uncertainty quantification. *Computational Geosciences*, 17(4):689–703. DOI:10.1007/s10596-013-9351-5.
- Chilès JP & Delfiner P (1999). *Geostatistics: Modeling Spatial Uncertainty, 2nd Edition* / Wiley. <https://www.wiley.com/en-us/Geostatistics+2022-02-01>.

- Claproud M, Gloaguen E, Béraud T, Blouin M, Dupuis C, Ferron P, Ouellet M, Chaussé M, Martel R, Paradis D & Ballard JM (2022). A Case Study Using Seismic Reflection and Well Logs to Reduce and Quantify Uncertainty During a Hydrogeological Assessment. *Frontiers in Water*, 3.
- Crestani E, Camporese M, Bau D & Salandin P (2013). Ensemble Kalman Filter Versus Ensemble Smoother for Assessing Hydraulic Conductivity via Tracer Test Data Assimilation. *Hydrology and Earth System Sciences*, 17:1517–1531. DOI:10.5194/hessd-9-13083-2012.
- Croteau A, Nastev M & Lefebvre R (2010). Groundwater Recharge Assessment in the Chateauguay River Watershed. *Canadian Water Resources Journal*, 35:451–468. DOI:10.4296/cwrj3504451.
- De la Varga M, Schaaf A & Wellmann F (2019). GemPy 1.0: Open-source stochastic geological modeling and inversion. *Geoscientific Model Development*, 12:1–32. DOI:10.5194/gmd-12-1-2019.
- Defforge CL, Carissimo B, Bocquet M, Bresson R & Armand P (2019). Improving CFD atmospheric simulations at local scale for wind resource assessment using the iterative ensemble Kalman smoother. *Journal of Wind Engineering and Industrial Aerodynamics*, 189:243–257. DOI:10.1016/j.jweia.2019.03.030.
- Denis C (1991). Caractérisation hydrogéologique du substratum rocheux fracture du site pollue de Ville-Mercier. Mémoire de maîtrise, Université du Québec à Chicoutimi, Chicoutimi.
- Dong Y, Gu Y & Oliver DS (2006). Sequential assimilation of 4D seismic data for reservoir description using the ensemble Kalman filter. *Journal of Petroleum Science and Engineering*, 53(1):83–99. DOI:10.1016/j.petrol.2006.03.028.
- Emerick A & Reynolds A (2012). History Matching Time-Lapse Seismic Data using the Ensemble Kalman Filter with Multiple Data Assimilations. *Computational Geosciences*, 16. DOI:10.1007/s10596-012-9275-5.
- Emerick AA & Reynolds AC (2013). Ensemble smoother with multiple data assimilation. *Computers & Geosciences*, 55:3–15. DOI:10.1016/j.cageo.2012.03.011.
- Evensen G (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research : Oceanics*, 99(C5):10143–10162. DOI:10.1029/94JC00572.
- Evensen G (1997). Advanced Data Assimilation for Strongly Nonlinear Dynamics. *Monthly Weather Review*, 125(6):1342–1354. DOI:10.1175/1520-0493(1997)125<1342:ADAFSN>2.0.CO;2.
- Evensen G (2003). The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4):343–367. DOI:10.1007/s10236-003-0036-9.
- Evensen G (2004). Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynamics*, 54:539–560. DOI:10.1007/s10236-004-0099-2.
- Evensen G (2009a). *Data Assimilation: The Ensemble Kalman Filter*. Springer Science & Business Media.
- Evensen G (2009b). The ensemble Kalman filter for combined state and parameter estimation. *IEEE Control Systems Magazine*, 29(3):83–104. DOI:10.1109/MCS.2009.932223. Conference Name: IEEE Control Systems Magazine.

- Galli A & Meunier G (1987). Study of a Gas Reservoir Using the External Drift Method. *Geostatistical Case Studies*. Matheron G & Armstrong M (éditeurs), Springer Netherlands, Quantitative Geology and Geostatistics, pages 105–119. Dordrecht.
- Globensky Y (1986). Géologie de la région de St-Chrysostome et de Lachine. Ministère de l'énergie et des ressources du Québec. Québec.
- Goovaerts P (1997). *Geostatistics for Natural Resources Evaluation*. Oxford University Press.
- Houtekamer PL, Mitchell HL, Pellerin G, Buehner M, Charron M, Spacek L & Hansen B (2005). Atmospheric Data Assimilation with an Ensemble Kalman Filter: Results with Real Observations. *Monthly Weather Review*, 133(3):604–620. DOI:10.1175/MWR-2864.1.
- Hunt BR, Kostelich EJ & Szunyogh I (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D: Nonlinear Phenomena*, 230(1):112–126. DOI:10.1016/j.physd.2006.11.008.
- Janjić T, Nberger L, Albertella A, Schröter J & Skachko S (2011). On Domain Localization in Ensemble-Based Kalman Filter Algorithms. *Monthly Weather Review*, 139(7):2046–2060. DOI:10.1175/2011MWR3552.1.
- Ji J, Tong Q, (Leon) Wang L, Lin CC, Zhang C, Gao Z & Fang J (2018). Application of the EnKF method for real-time forecasting of smoke movement during tunnel fires. *Advances in Engineering Software*, 115:398–412. DOI:10.1016/j.advengsoft.2017.10.007.
- Johns CJ & Mandel J (2008). A two-stage ensemble Kalman filter for smooth data assimilation. *Environmental and Ecological Statistics*, 15(1):101–110. DOI:10.1007/s10651-007-0033-0.
- Journel AG & Huijbregts CJ (1978). *Mining Geostatistics*. The Blackburn Press, Caldwell, N.J.
- Kalman RE (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35–45. DOI:10.1115/1.3662552.
- Lefebvre R (2007). Contexte hydrogéologique et potentiel de réhabilitation du site des anciennes lagunes de Mercier (Québec), Canada. INRS – Eau, Terre et Environnement. Québec.
- Li L, Stetler L, Cao Z & Davis A (2018). An iterative normal-score ensemble smoother for dealing with non-Gaussianity in data assimilation. *Journal of Hydrology*, 567:759–766. DOI:10.1016/j.jhydrol.2018.01.038.
- Mallet JL (1992). GOCAD A Computer Aided Design Program for Geological Applications. *Three-Dimensional Modeling with Geoscientific Information Systems*. Turner AK, éditeur, Springer Netherlands, NATO ASI Series, pages 123–141. Dordrecht.
- Mitchell HL, Houtekamer PL & Pellerin G (2002). Ensemble Size, Balance, and Model-Error Representation in an Ensemble Kalman Filter. *Monthly Weather Review*, 130(11):2791–2808. DOI:10.1175/1520-0493(2002)130<2791:ESBAME>2.0.CO;2.
- Molson JW & Frind EO (2022). SALTFLOW: Density dependent flow and mass transport model in three dimensions. *User Guide, Université Laval, University of Waterloo, Ontario*.
- Nastev M & Lamontagne C (2010). Survol Hydrogéologique de l'Aquifère Transfrontalier du Bassin Versant de la Rivière Châteauguay, Canada - États Unis. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques*, 35(4):359–376. DOI:10.4296/cwrj3504359.

- Pontlevoy O (2004). Modélisation hydrogéologique pour supporter la gestion du système aquifère de la région de Ville-Mercier. Mémoire de maîtrise, Université du Québec, Institut national de la recherche scientifique, Québec.
- Pontlevoy O, Lefebvre R, Therrien R, Martel R, Ouellet M, Lamontagne C & C R (2004). Numerical modeling of groundwater flow in interconnected granular and rock aquifers at the Ville Mercier DNAPL-contaminated site, Quebec, Canada.
- Poulin M (1977). *Groundwater contamination near a liquid waste lagoon, Ville Mercier, Québec*. Mémoire de maîtrise, University of Waterloo, Ontario.
- Raanes P (2015a). On the ensemble Rauch-Tung-Striebel smoother and its equivalence to the ensemble Kalman smoother. *Quarterly Journal of the Royal Meteorological Society*, 142:n/a–n/a. DOI:10.1002/qj.2728.
- Raanes PN (2015b). *Improvements to ensemble methods for data assimilation in the geosciences*. <http://purl.org/dc/dcmitype/Text>, University of Oxford.
- Raanes PN, Stordal AS & Evensen G (2019). Revising the stochastic iterative ensemble smoother. *Nonlinear Processes in Geophysics*, 26(3):325–338. DOI:10.5194/npg-26-325-2019.
- Robin MJL, Gutjahr AL, Sudicky EA & Wilson JL (1993). Cross-correlated random field generation with the direct Fourier Transform Method. *Water Resources Research*, 29(7):2385–2397. DOI:10.1029/93WR00386.
- Sakov P & Bertino L (2011). Relation between two common localisation methods for the EnKF. *Computational Geosciences*, 15(2):225–237. DOI:10.1007/s10596-010-9202-6.
- Sakov P, Evensen G & Bertino L (2010). Asynchronous data assimilation with the EnKF. *Tellus A: Dynamic Meteorology and Oceanography*, 62(1):24–29. DOI:10.1111/j.1600-0870.2009.00417.x.
- Skjervheim JaA, Evensen G, Hove J & Vabø JG (2011). An Ensemble Smoother for assisted History Matching. OnePetro.
- Tremblay T (2008). *Hydrostratigraphie et géologie du quaternaire dans le bassin-versant de la rivière Châteauguay, Québec*. <https://archipel.uqam.ca/1814/>. 2019-06-11.
- Tremblay T, Hunter J, Lamontagne C & Nastev M (2010). High Resolution Seismic Survey in a Contaminated Esker Area, Chateauguay River Watershed, Quebec. *Canadian Water Resources Journal / Revue canadienne des ressources hydriques*, 35(4):417–432. DOI:10.4296/cwrj3504417.
- Wang Y, Hu K, Ren L & Lin G (2019). Optimal observations-based retrieval of topography in 2D shallow water equations using PC-EnKF. *Journal of Computational Physics*, 382:43–60. DOI:10.1016/j.jcp.2019.01.004.
- White JT, Hunt RJ, Fienen MN & Doherty JE (2020). Approaches to highly parameterized inversion: PEST++ Version 5, A software suite for parameter estimation, uncertainty analysis, management optimization and sensitivity analysis. U.S. Geological Survey. Reston, VA, 64 pages.
- Wu J, Remy N & Boucher A (2009). *Applied Geostatistics with SGeMS: A user's guide*.
- Zhou H, Gómez-Hernández JJ, Hendricks Franssen HJ & Li L (2011). An approach to handling non-Gaussianity of parameters and state variables in ensemble Kalman filtering. *Advances in Water Resources*, 34(7):844–864. DOI:10.1016/j.advwatres.2011.04.014.

Zovi F, Camporese M, Hendricks Franssen HJ, Huisman JA & Salandin P (2017). Identification of high-permeability subsurface structures with multiple point geostatistics and normal score ensemble Kalman filter. *Journal of Hydrology*, 548:208–224. DOI:10.1016/j.jhydrol.2017.02.056.



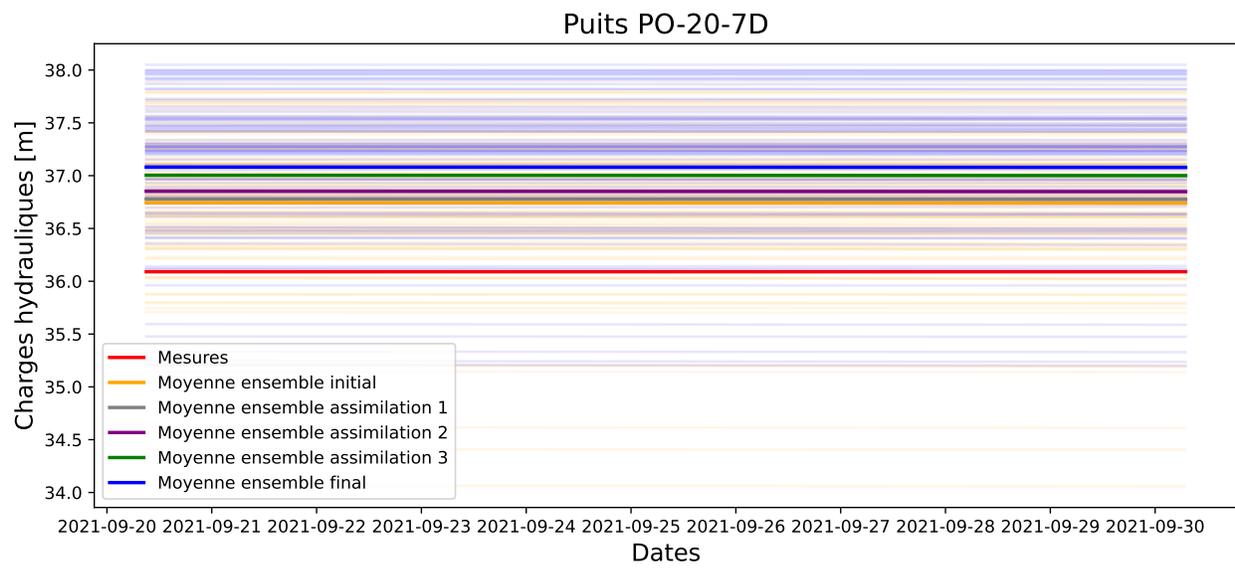
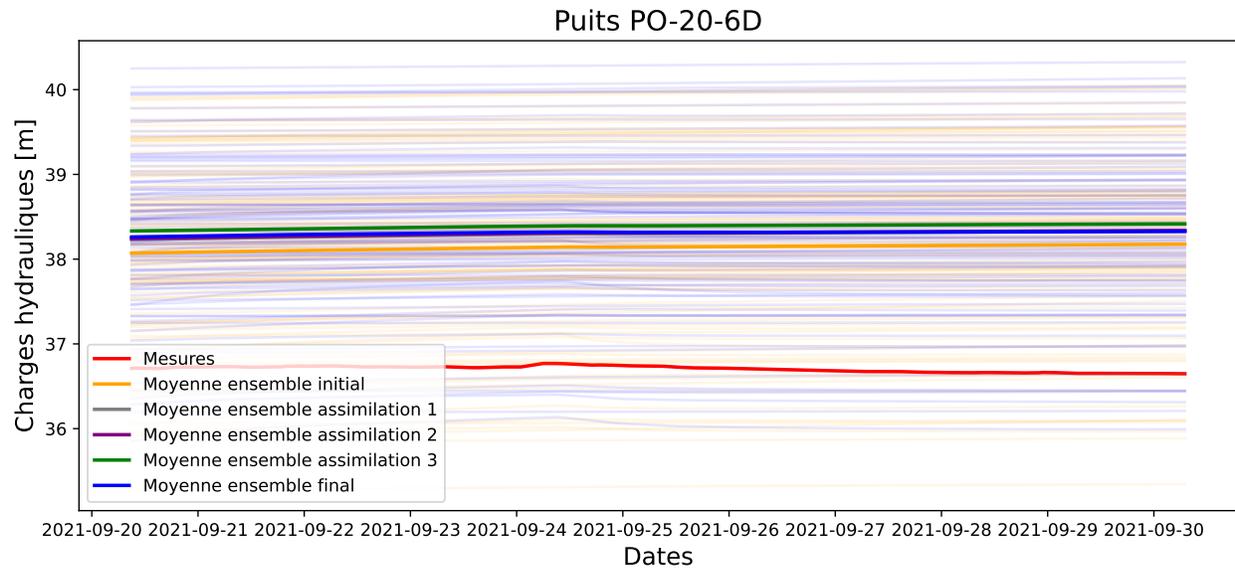
## Annexe A

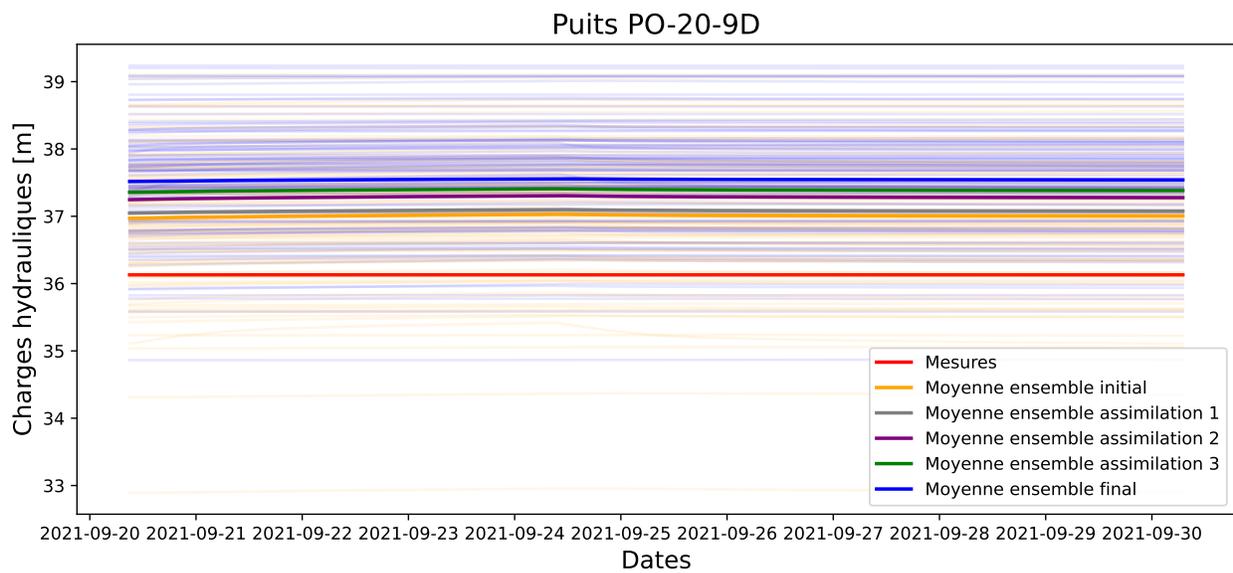
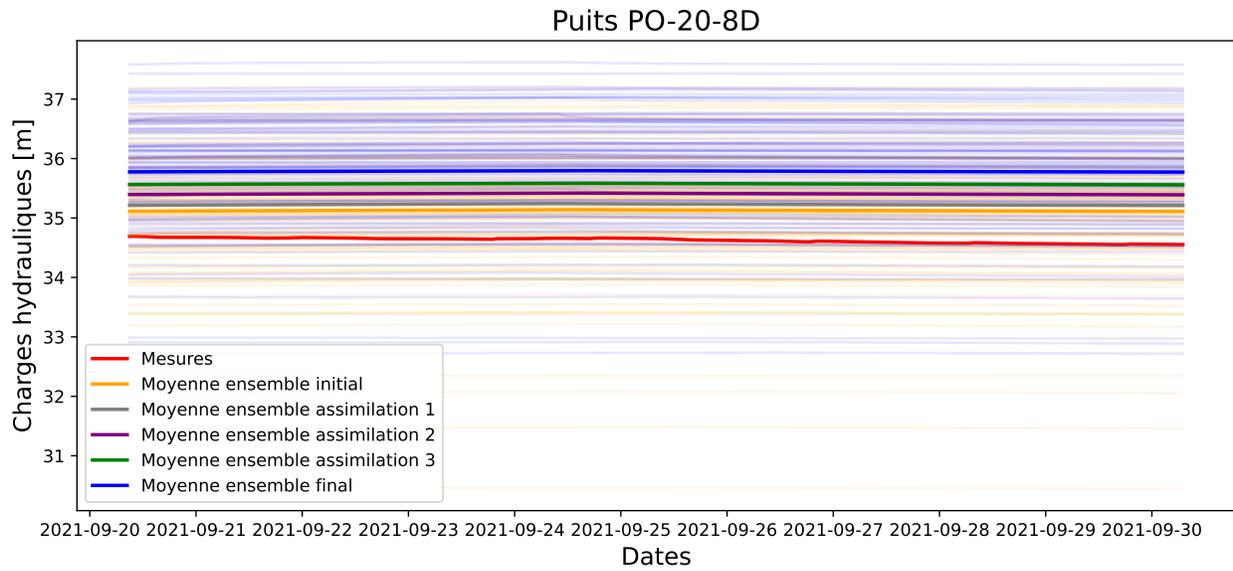
# Courbes individuelles de charges hydrauliques pour le schéma 500/20 amplifié

Pour rappel, le schéma 500/20 amplifié correspond à une assimilation puits par puits avec une localisation définie autour de chacun des puits à chaque assimilation. La résultante de toutes les mise à jour est calculée en additionnant les mise à jour calculées pour chaque paramètres du modèle au droit de chaque puits. La mise à jour est amplifiée d'un facteur 4, ce qui revient à augmenter l'amplitude de la mise à jour finale, sans modifier sa direction.

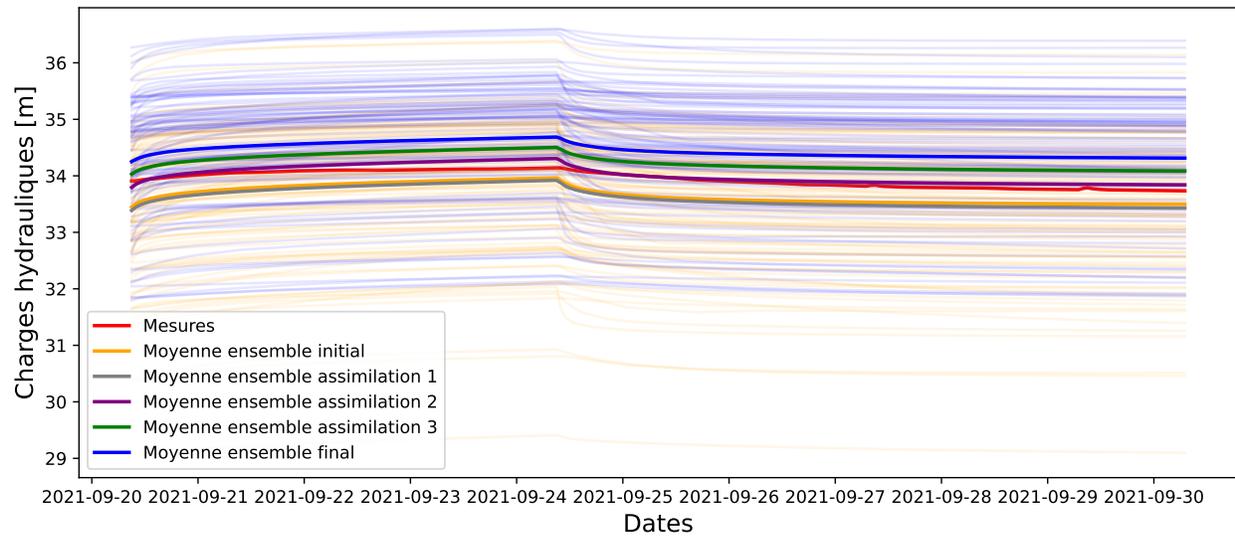
Sur l'ensemble des figures suivantes, les mesures sont affichées avec un trait solide rouge. Les moyennes des ensembles sont données en orange, gris, violet, vert et bleu. Les réalisations individuelles de chaque modèle de l'ensemble sont données uniquement pour l'ensemble initial (orange transparent) et l'ensemble final (bleu transparent). Dans le cas de ce schéma précis, le nombre de modèle dans l'ensemble à diminué à chaque assimilation. Il est passé de 106 à 106, puis à 102, puis à 99 pour terminer avec 97 modèles dans l'ensemble final. Cette réduction à permis d'éliminer les modèles numériquement les plus instables tout au long du processus.

Il est possible de constater que la moyenne de l'ensemble final n'est pas automatiquement améliorer par rapport à l'ensemble final. Cependant les modèles très éloignés des mesures n'ont pas réussis à être mis à jour, il ne convergeaient plus numériquement et ont donc été éliminés durant le processus d'assimilation. C'est l'élimination de ces modèles éloignés qui a aidée à réduire la RMSE.

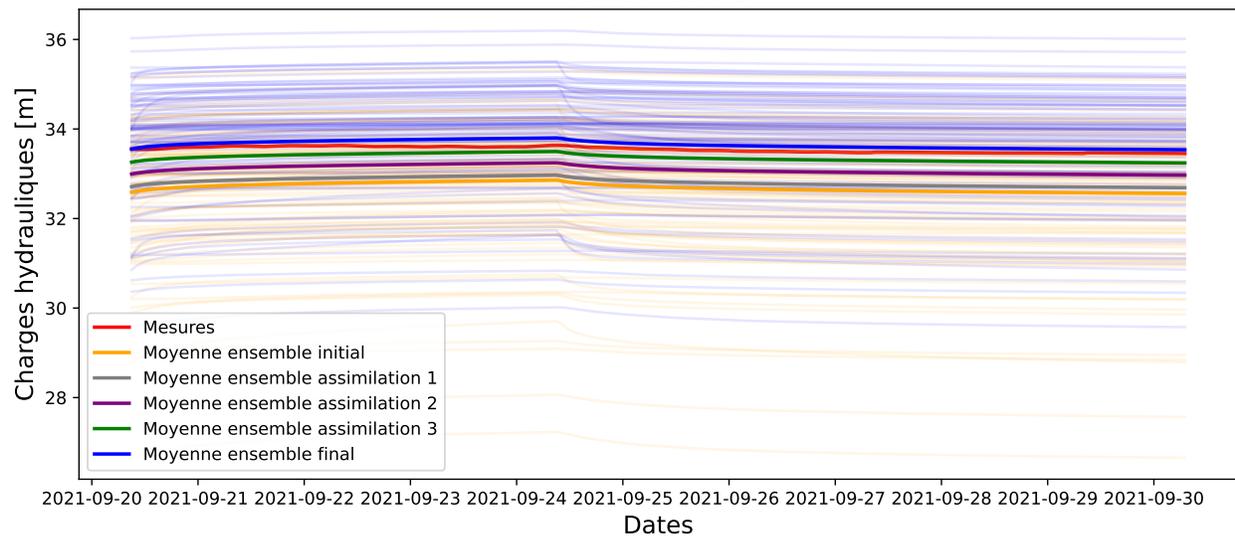


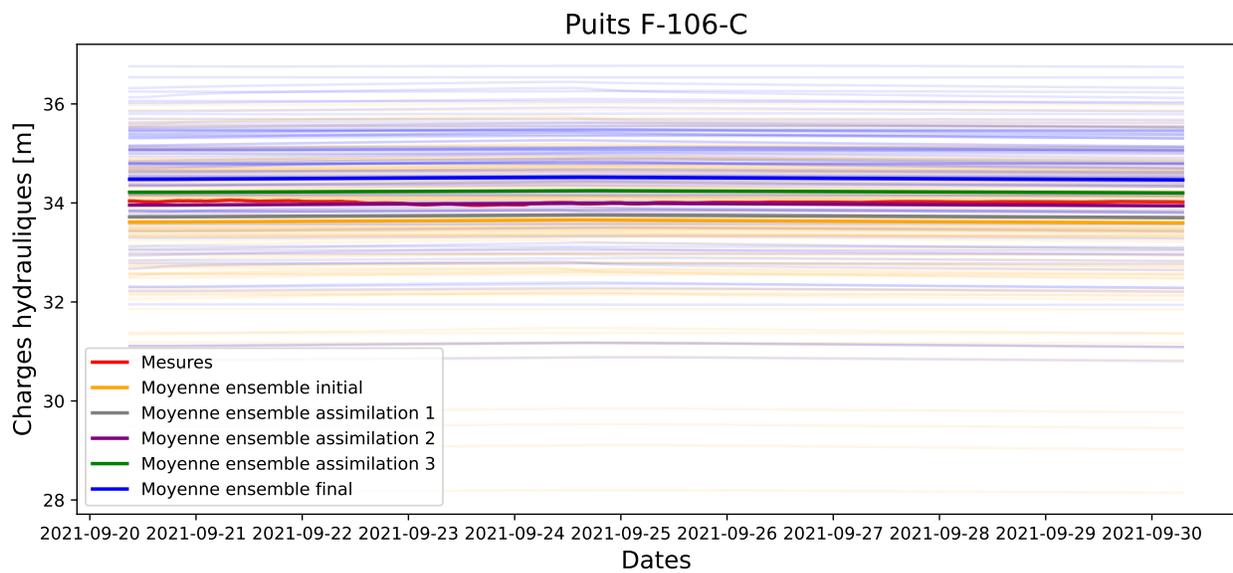
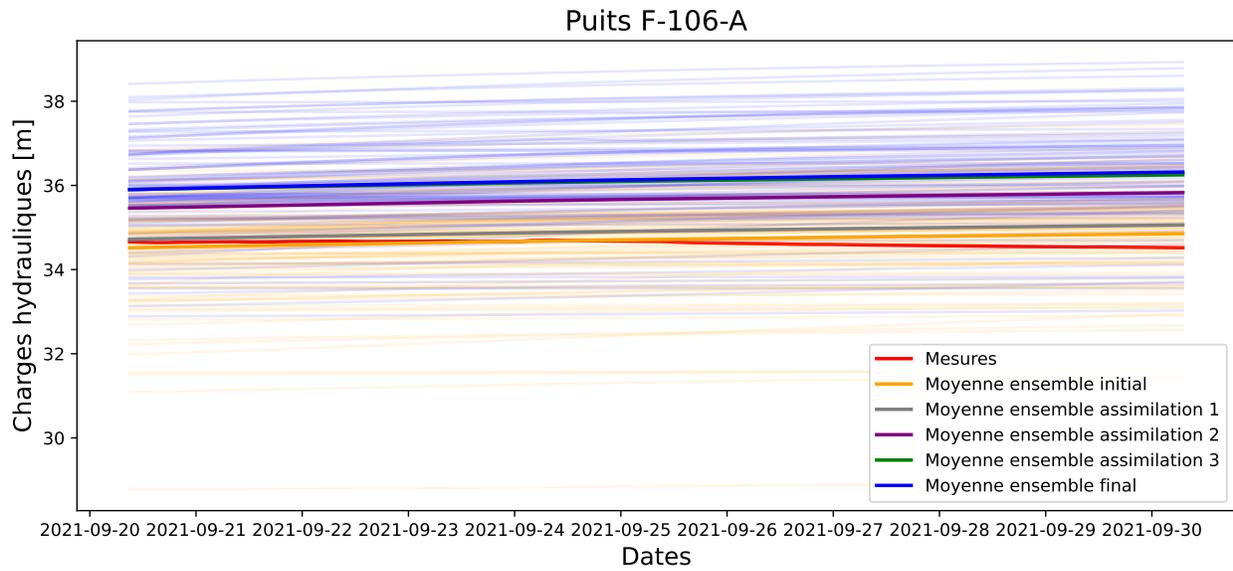


Puits F-124-A

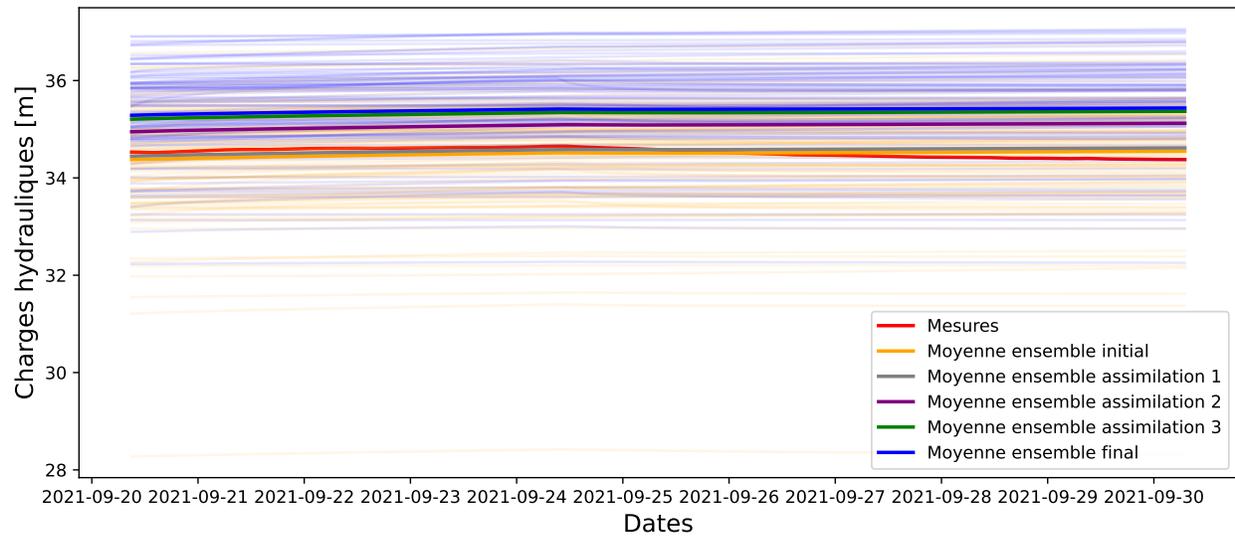


Puits F-124-C

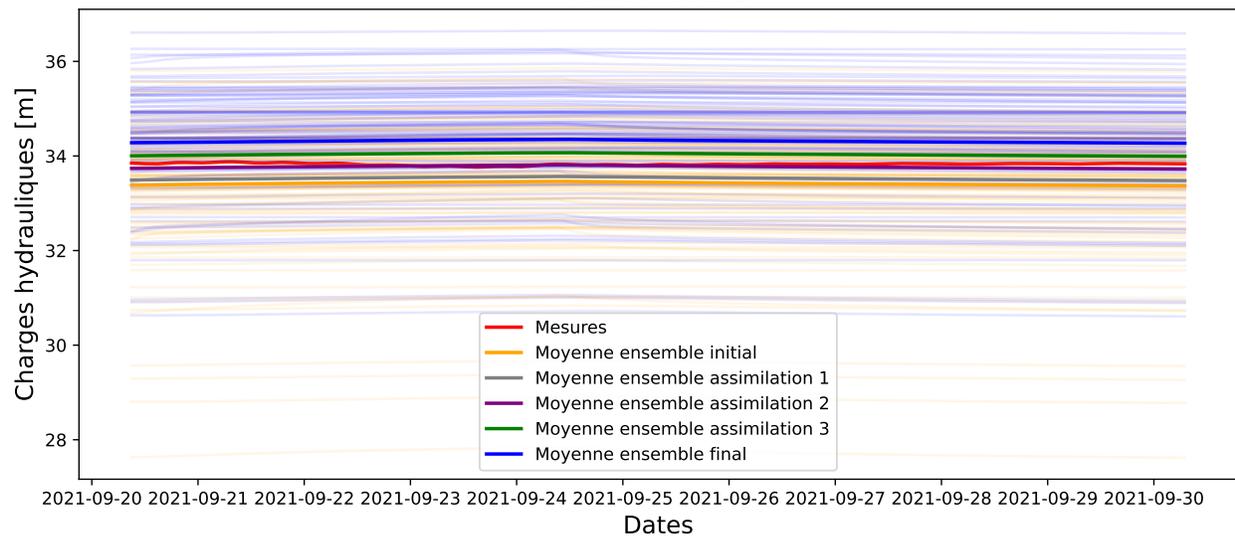


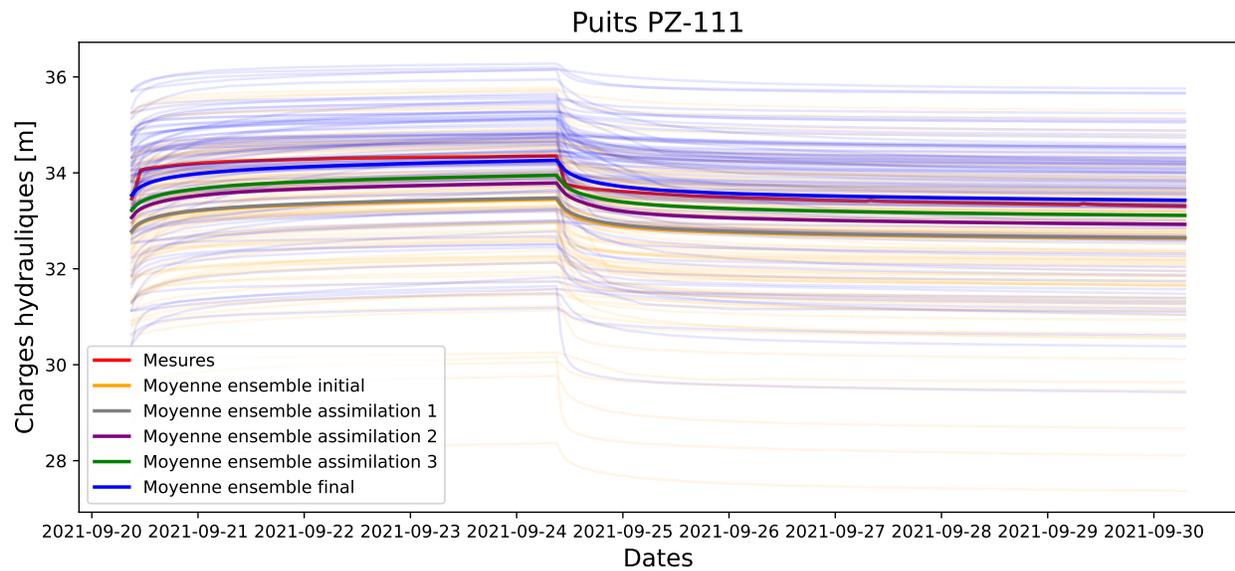
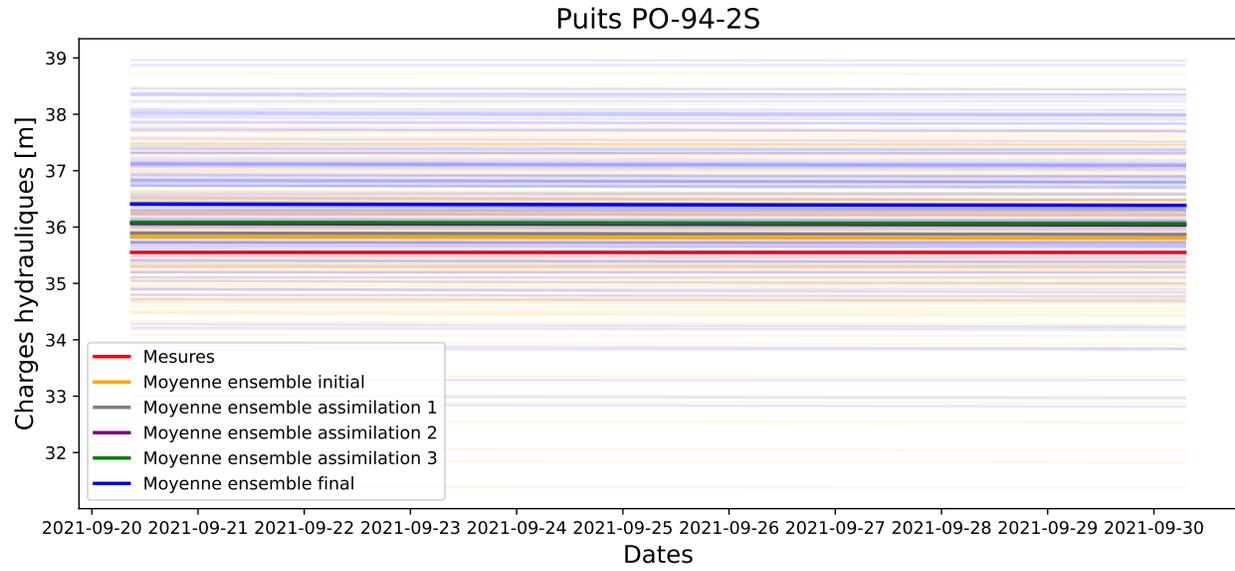


Puits F-102-A

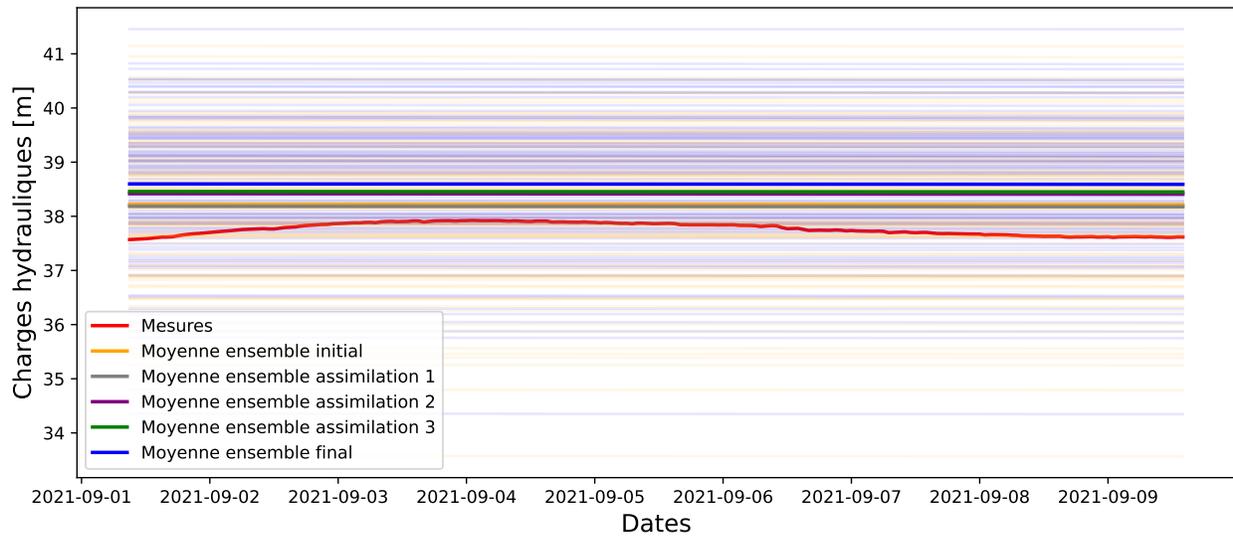


Puits F-102-C

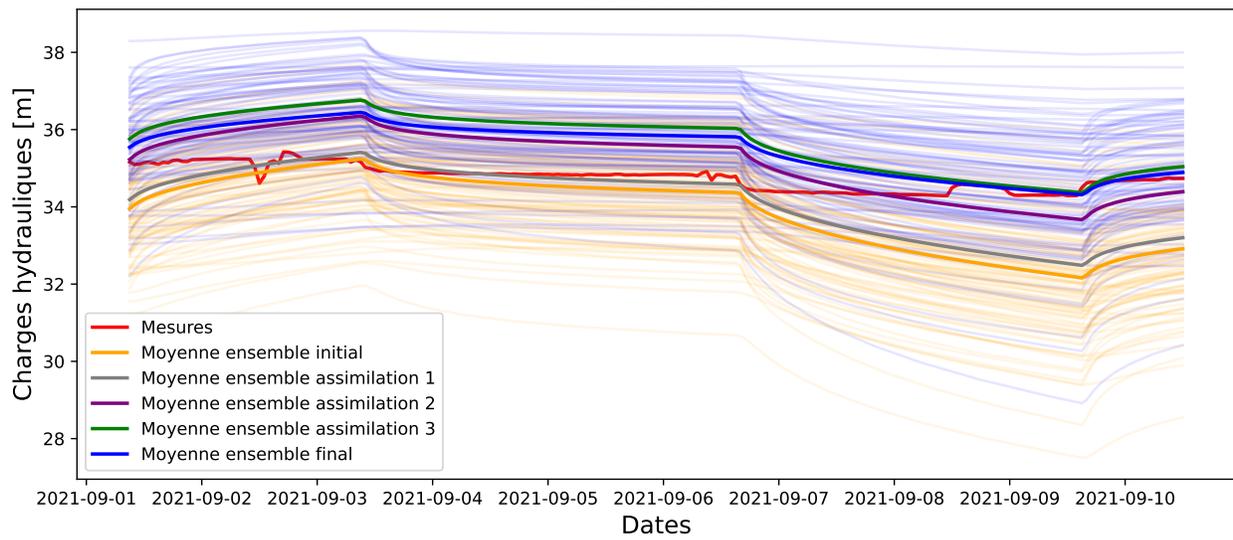


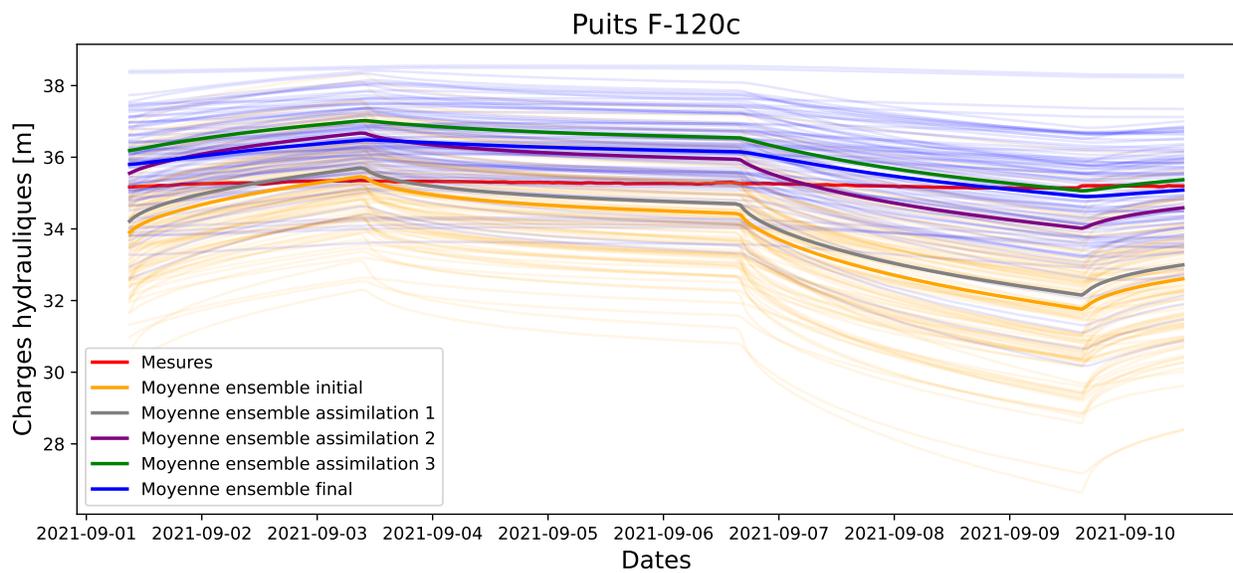
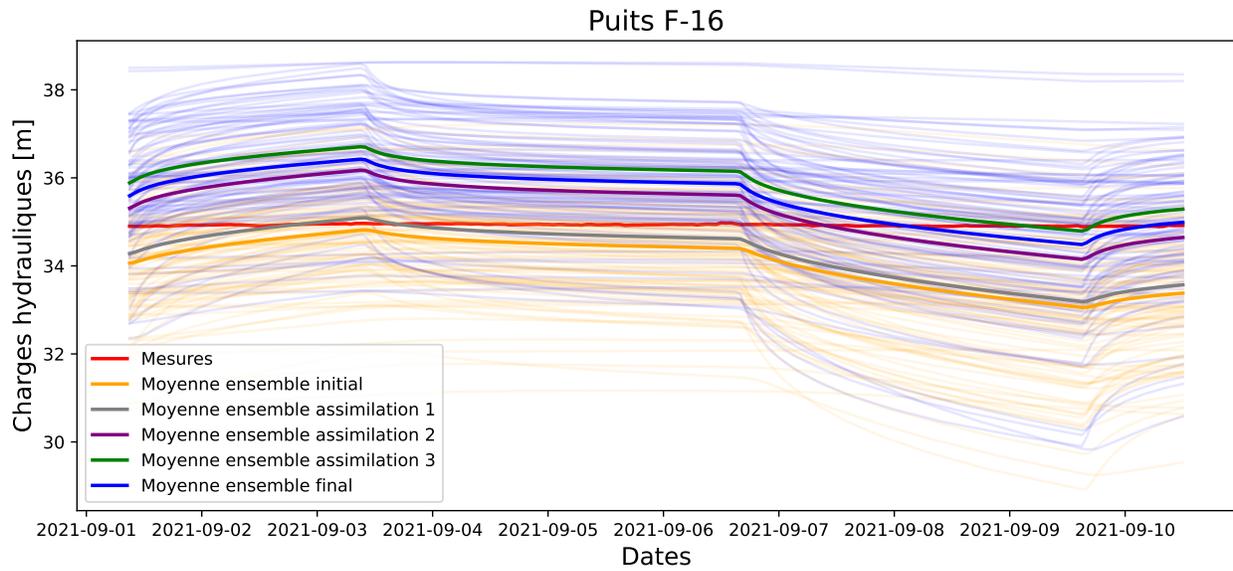


Puits F-1

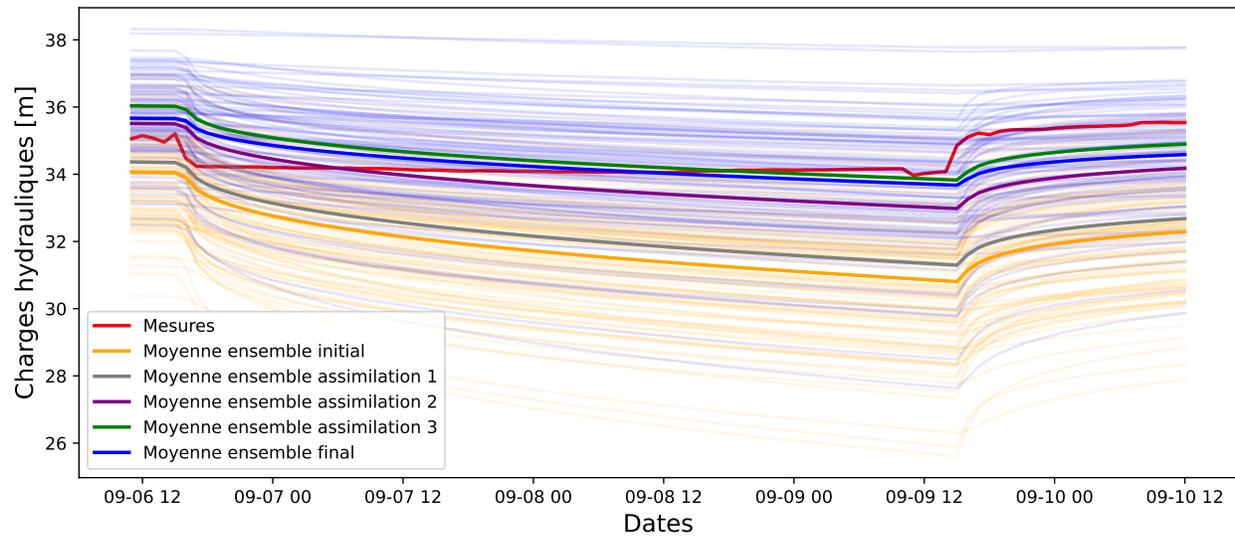


Puits F-11

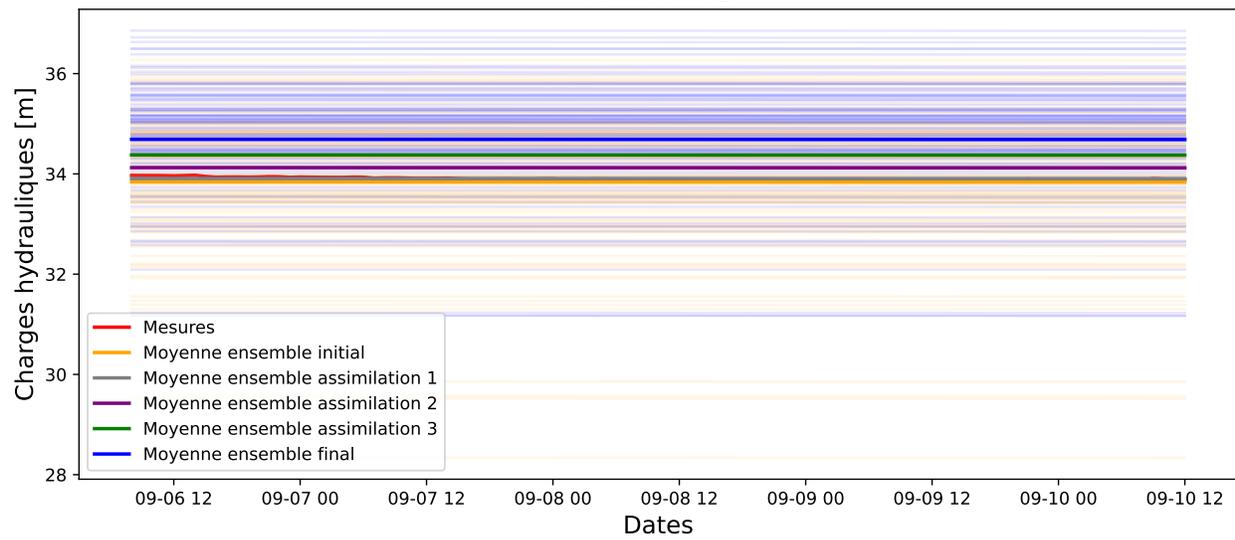


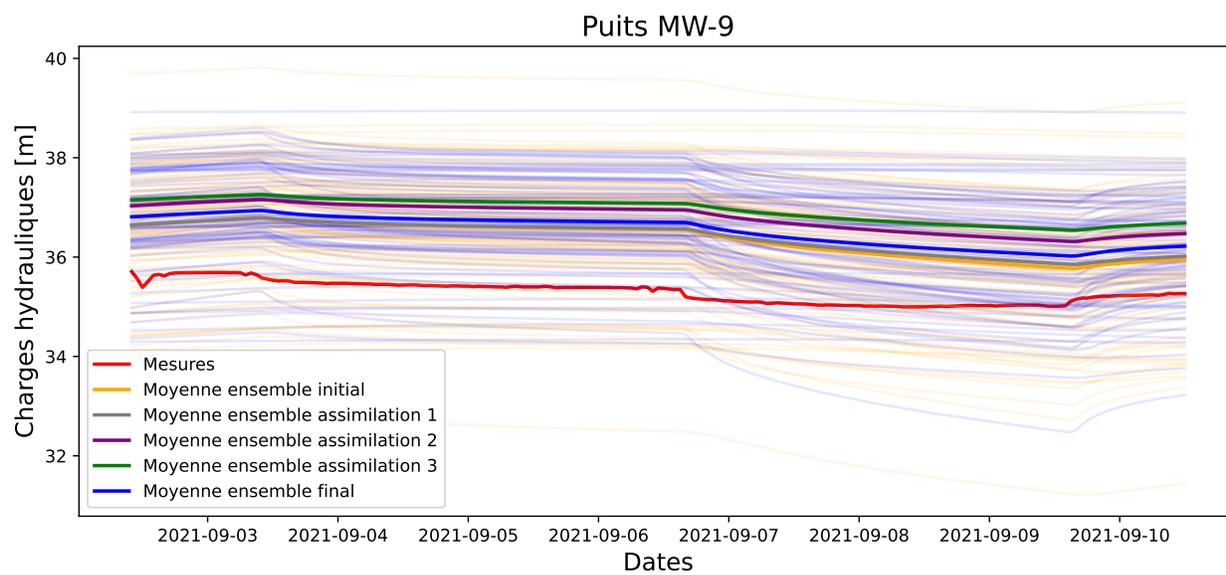
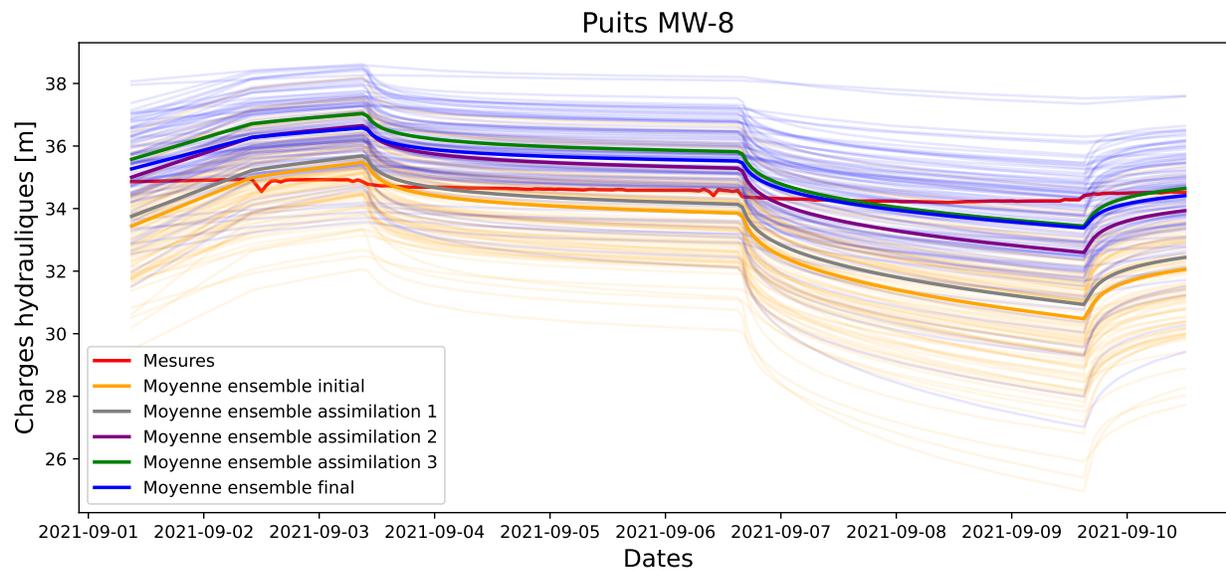


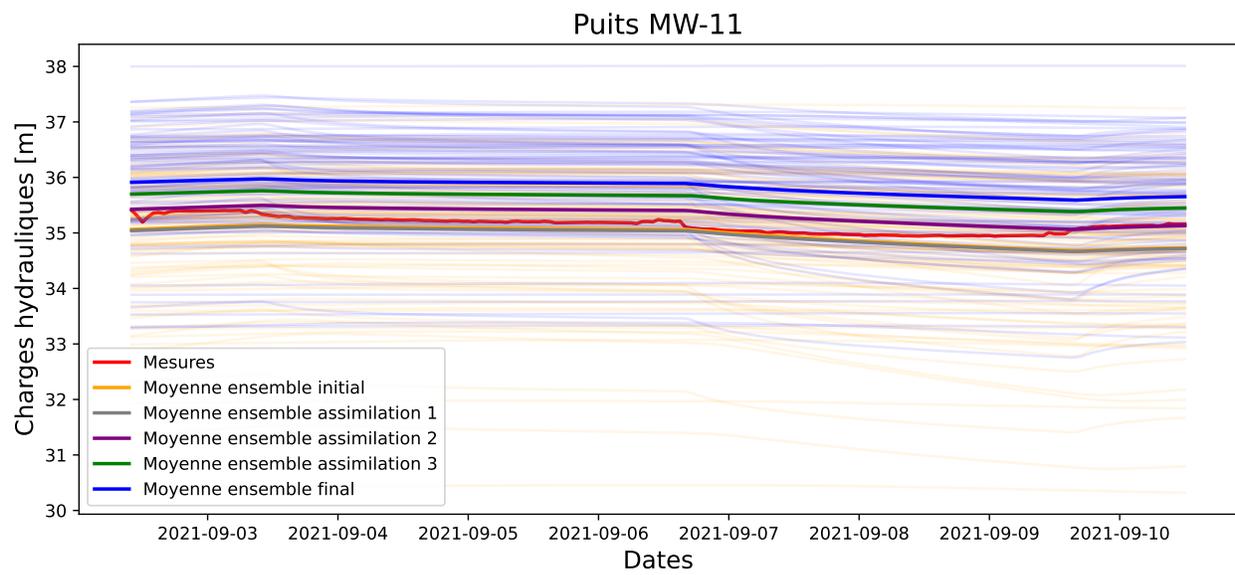
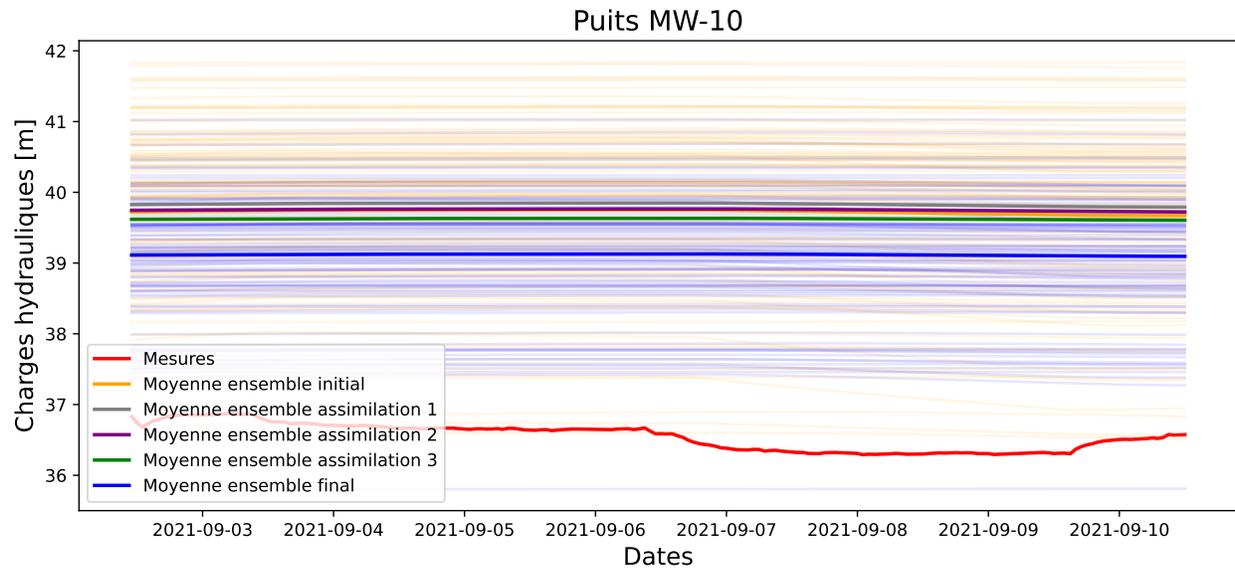
Puits 7011

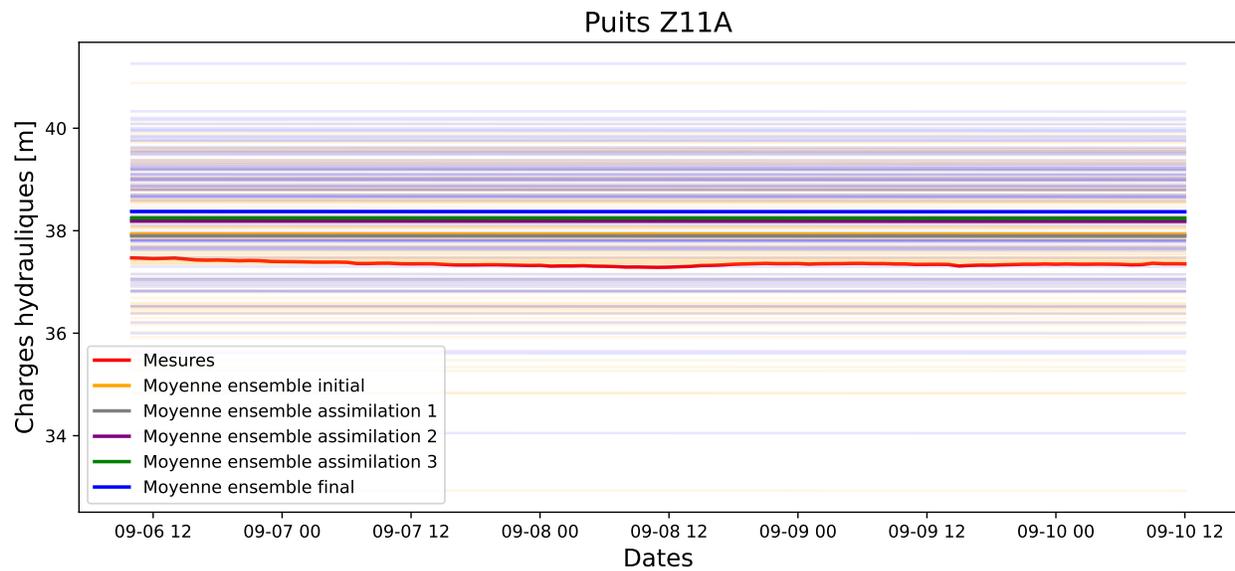
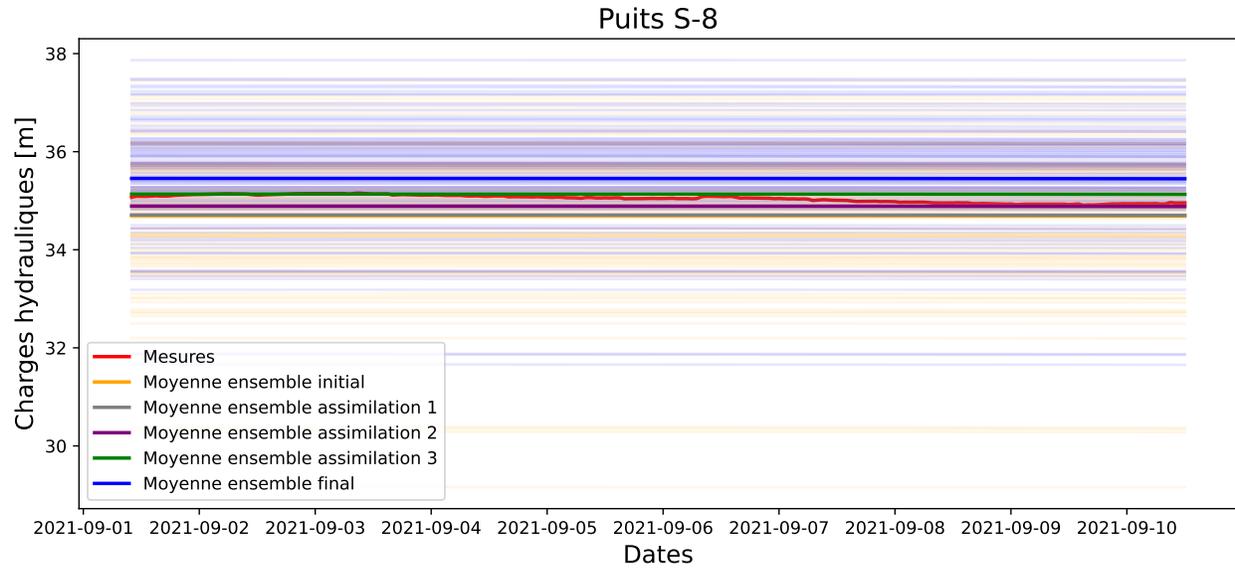


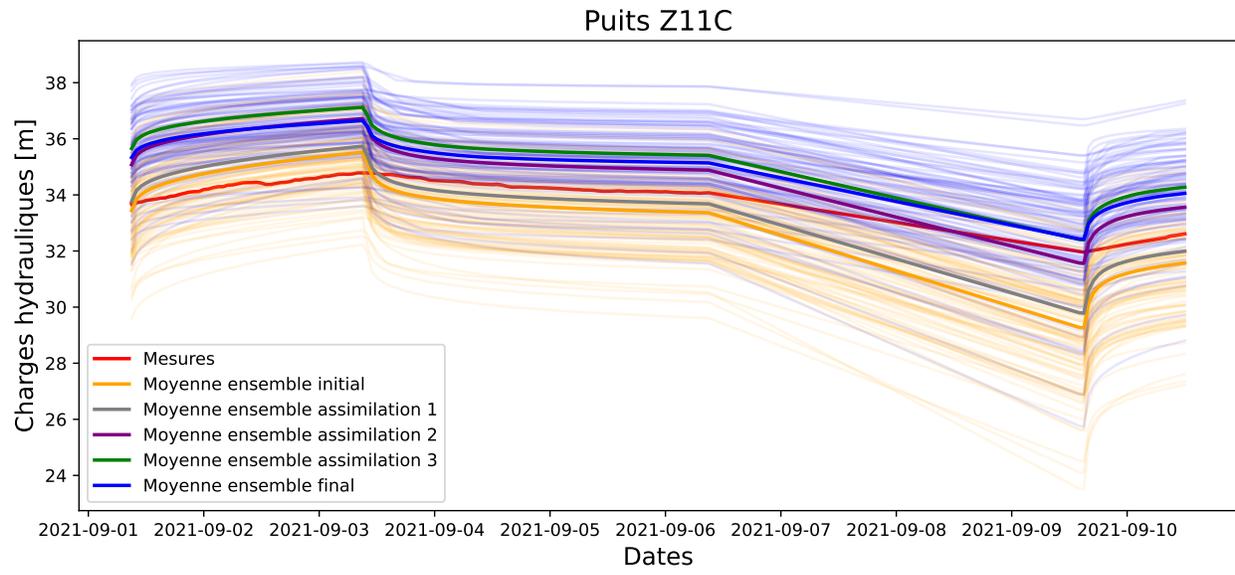
Puits MW-5











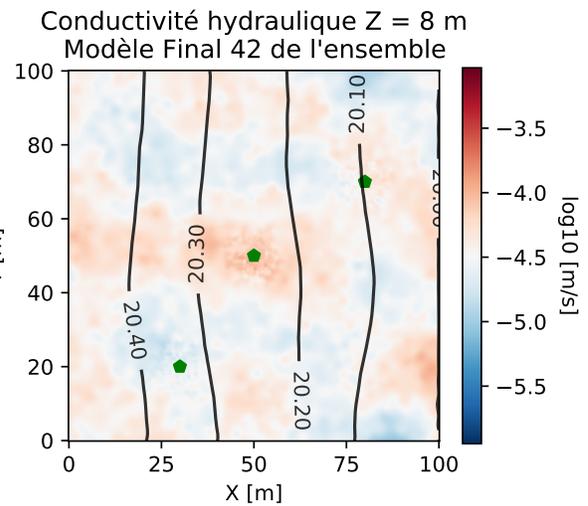
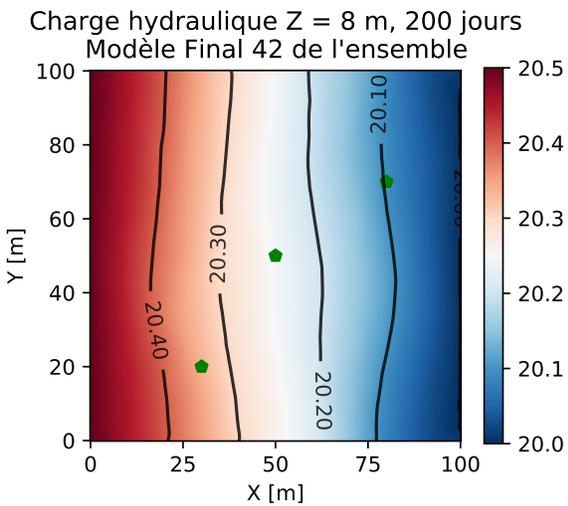
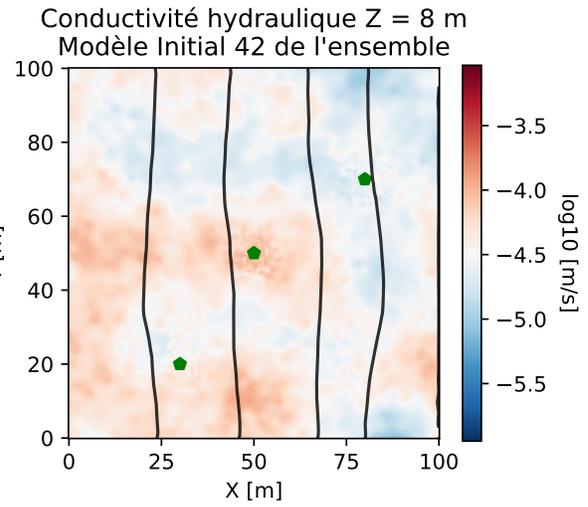
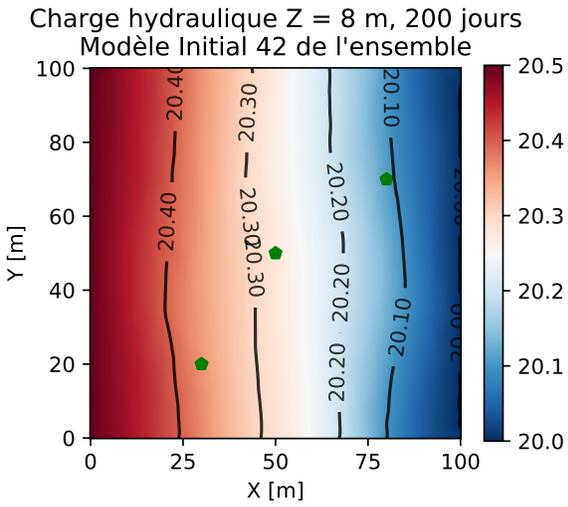
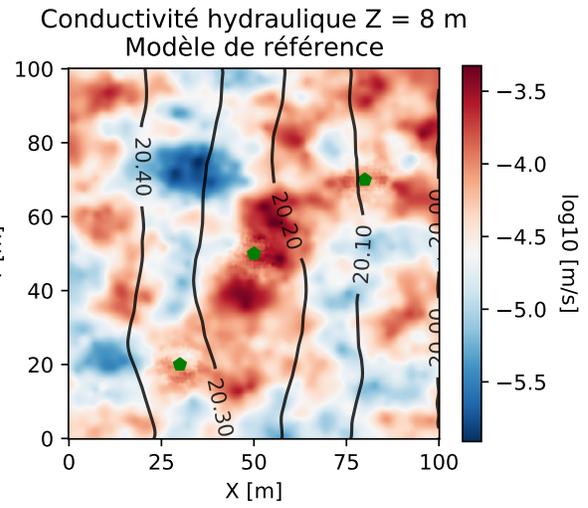
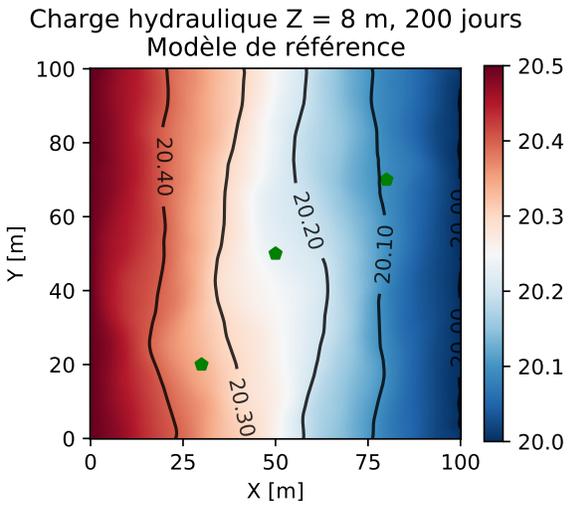
## Annexe B

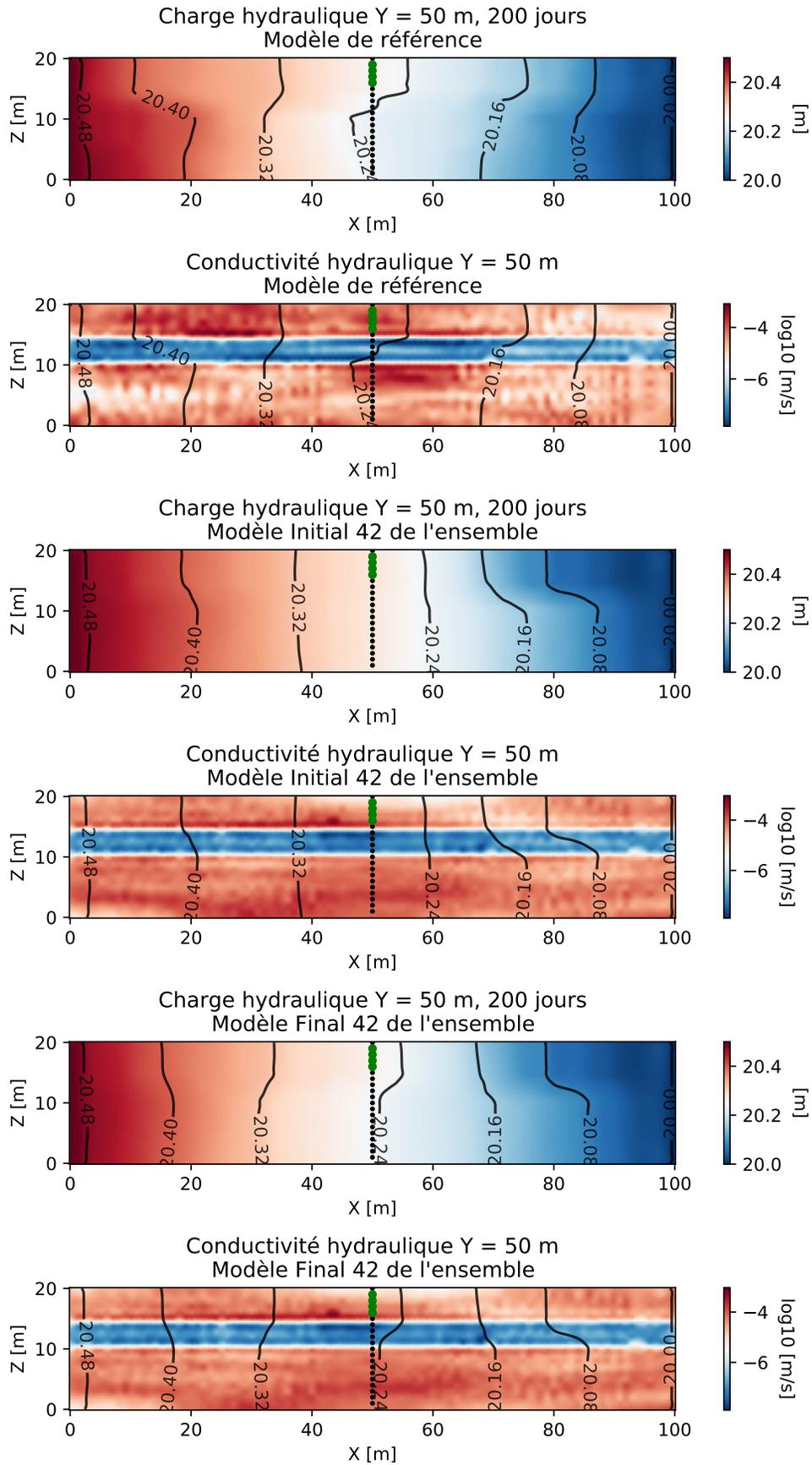
# Simulations numériques du cas synthétique

Les charges hydrauliques et concentrations simulées sont données pour le modèle de référence, un modèle de l'ensemble initial et le même modèle à la fin du schéma d'assimilation 100/20.

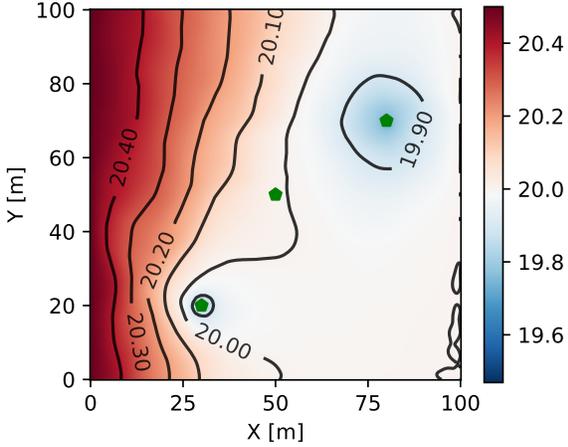
Les puits de pompage sont en vert sur les sections horizontales. Les crépines sont en vert sur les sections verticales en  $Y = 50\text{m}$ .

Le pompage était actif entre 200 et 1300 jours. Les sections à 200 jours indiquent donc des simulations avant modification de l'équilibre.

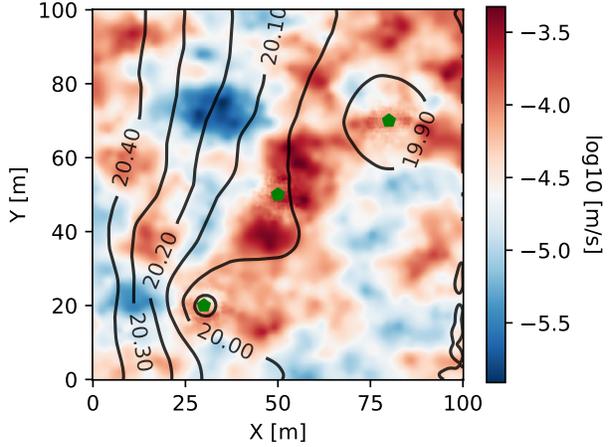




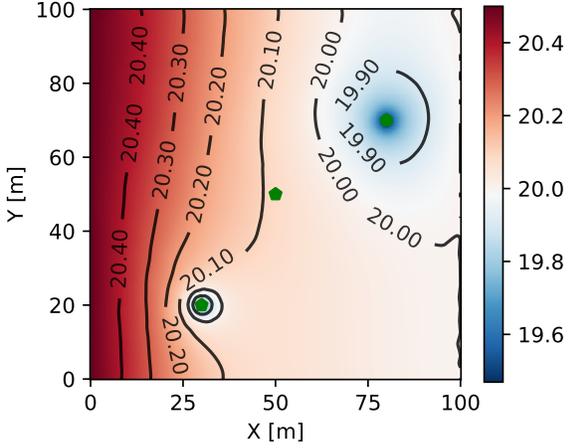
Charge hydraulique Z = 8 m, 800 jours  
Modèle de référence



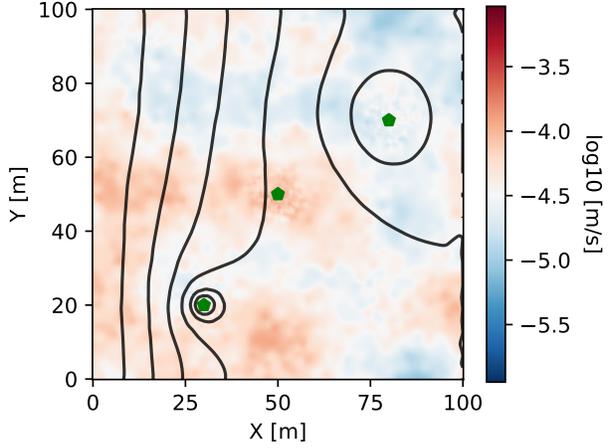
Conductivité hydraulique Z = 8 m  
Modèle de référence



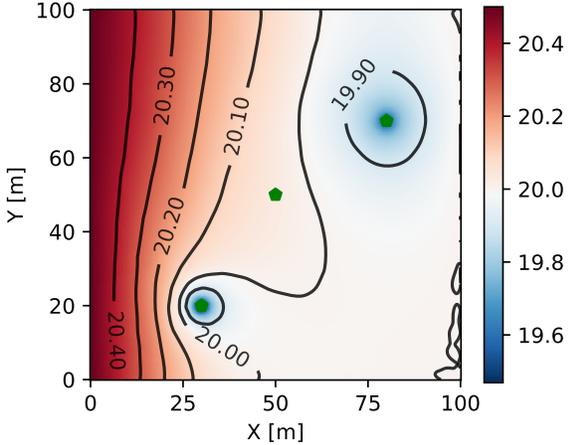
Charge hydraulique Z = 8 m, 800 jours  
Modèle Initial 42 de l'ensemble



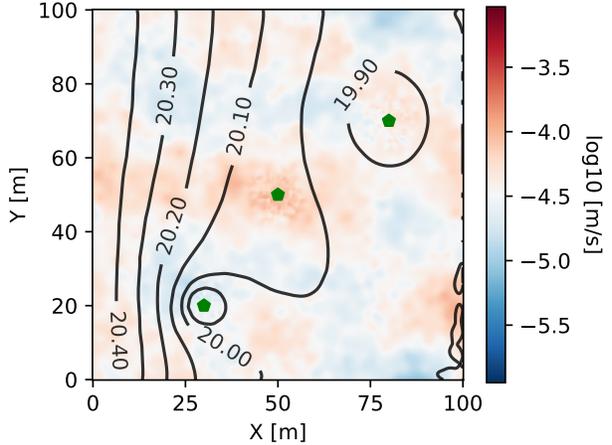
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble

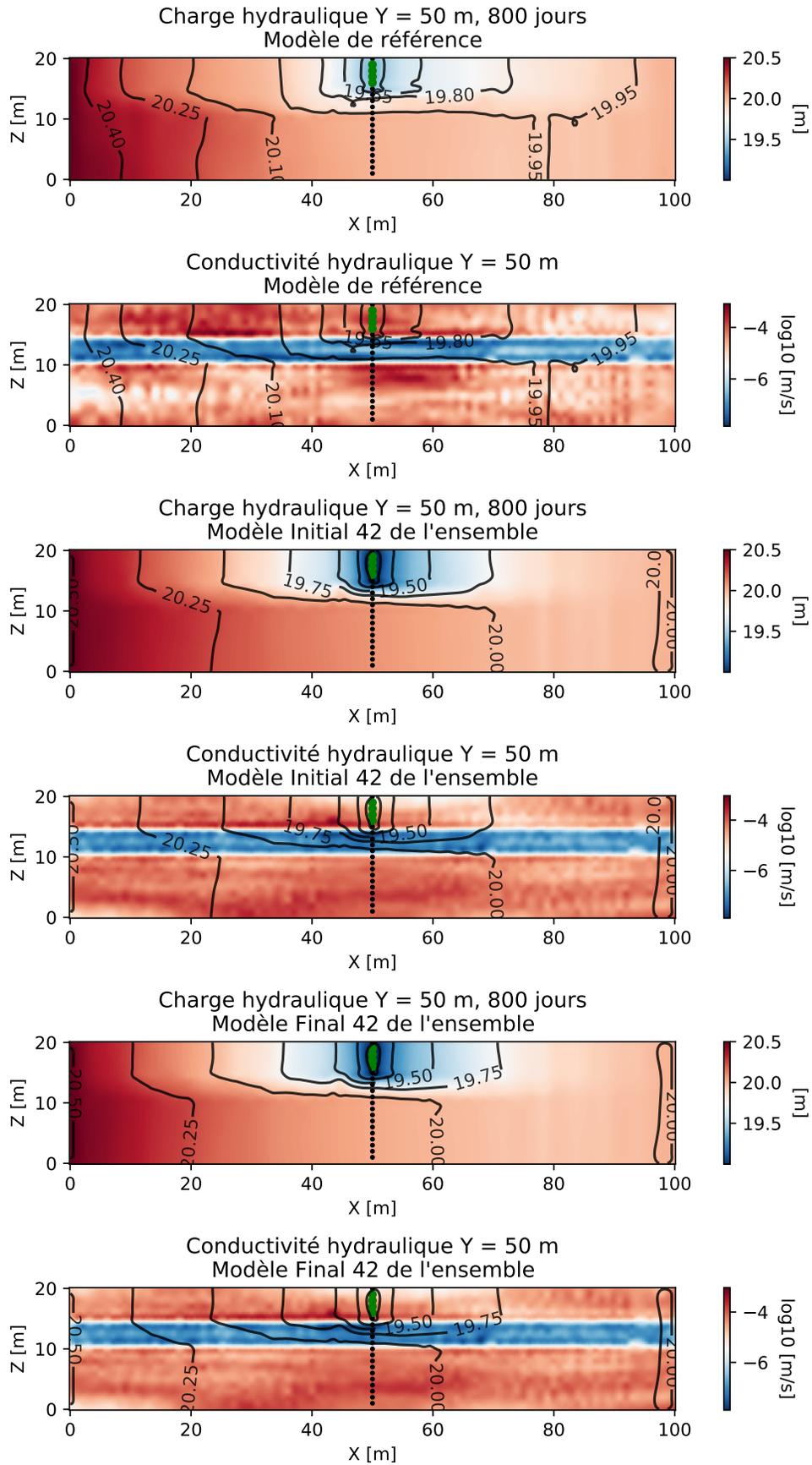


Charge hydraulique Z = 8 m, 800 jours  
Modèle Final 42 de l'ensemble

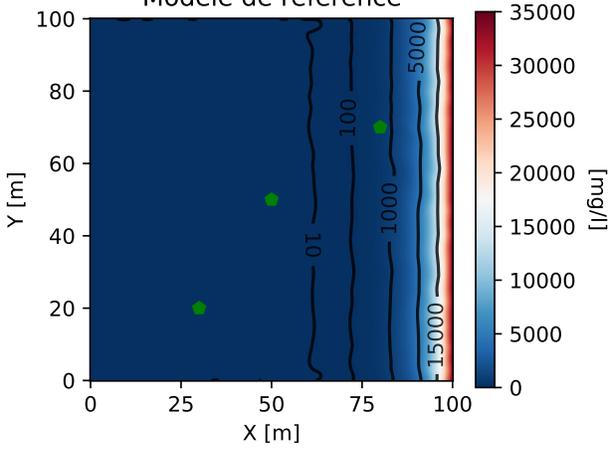


Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble

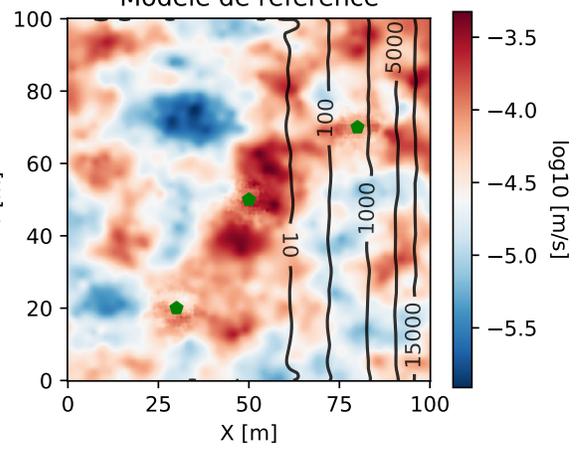




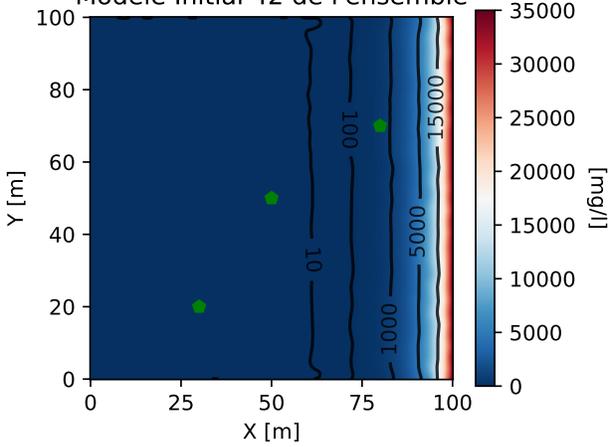
Concentration en sel Z = 8 m, 200 jours  
Modèle de référence



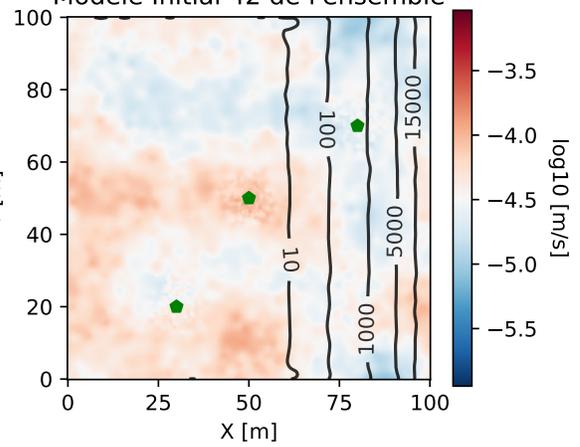
Conductivité hydraulique Z = 8 m  
Modèle de référence



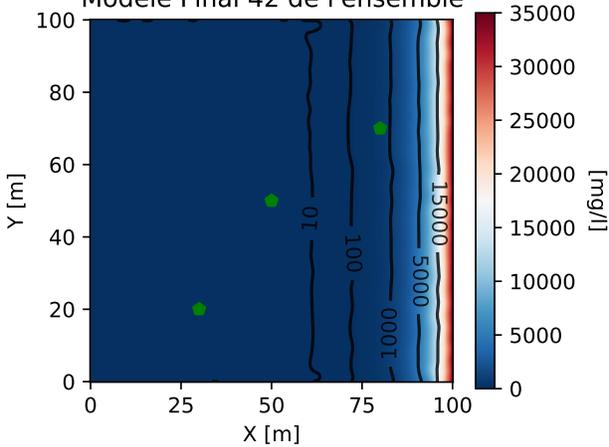
Concentration en sel Z = 8 m, 200 jours  
Modèle Initial 42 de l'ensemble



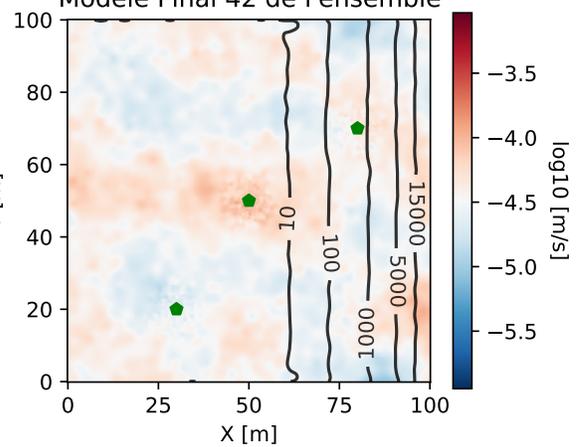
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



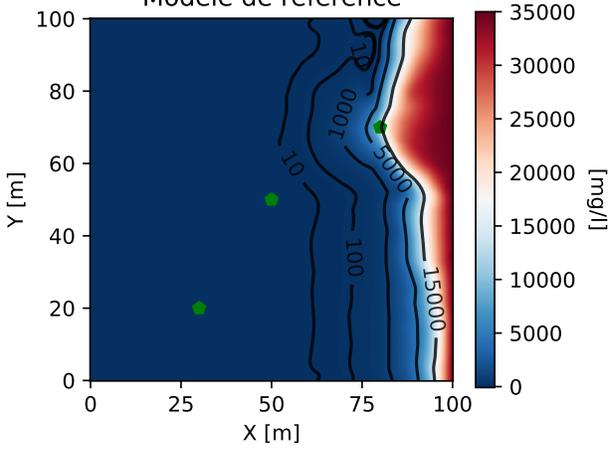
Concentration en sel Z = 8 m, 200 jours  
Modèle Final 42 de l'ensemble



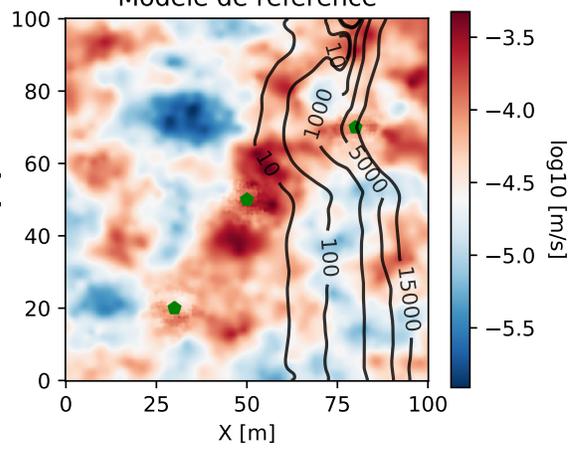
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



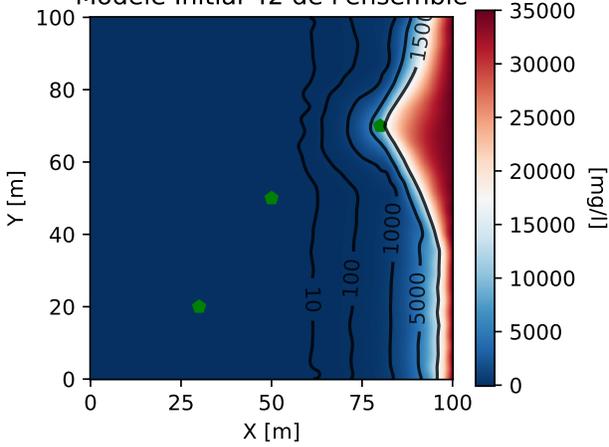
Concentration en sel Z = 8 m, 300 jours  
Modèle de référence



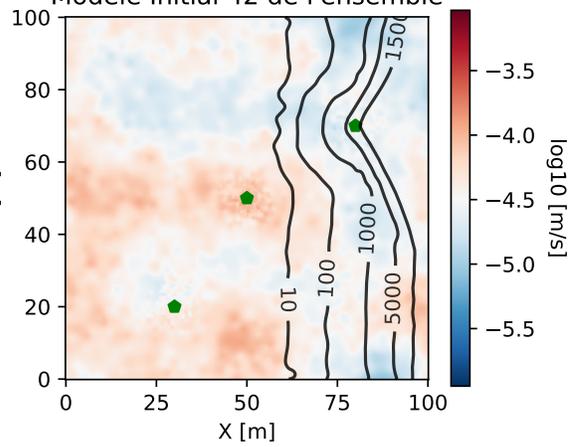
Conductivité hydraulique Z = 8 m  
Modèle de référence



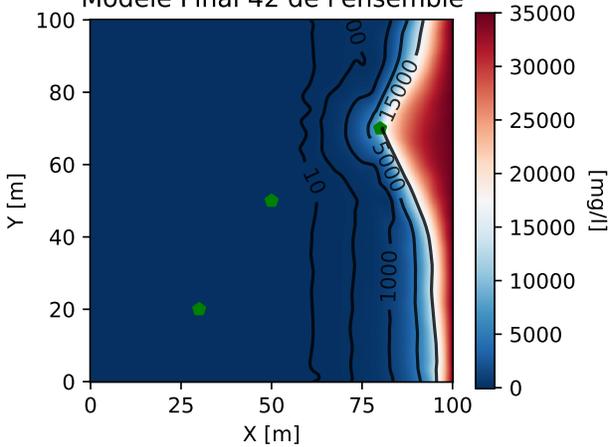
Concentration en sel Z = 8 m, 300 jours  
Modèle Initial 42 de l'ensemble



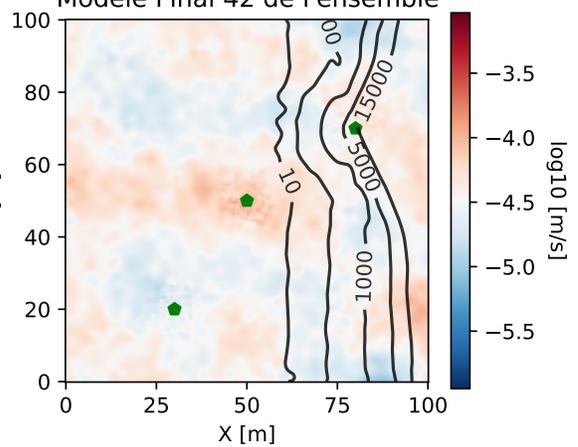
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



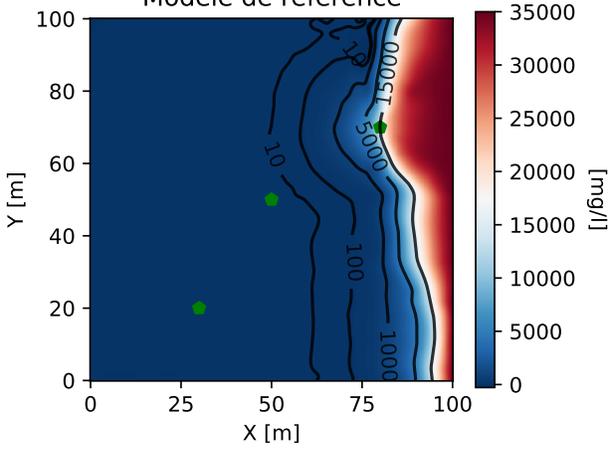
Concentration en sel Z = 8 m, 300 jours  
Modèle Final 42 de l'ensemble



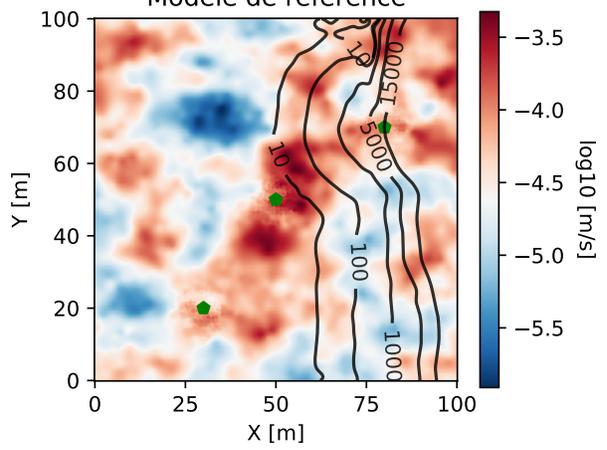
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



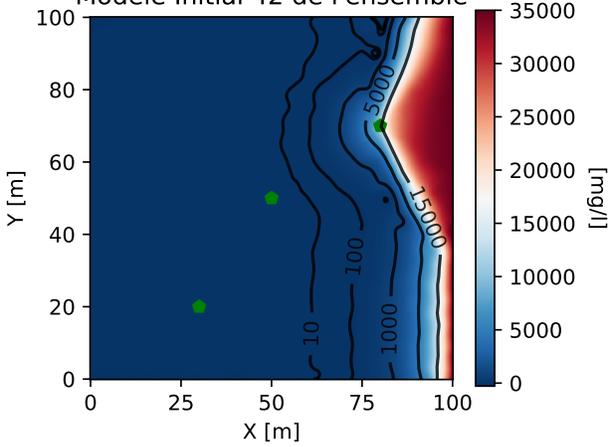
Concentration en sel Z = 8 m, 400 jours  
Modèle de référence



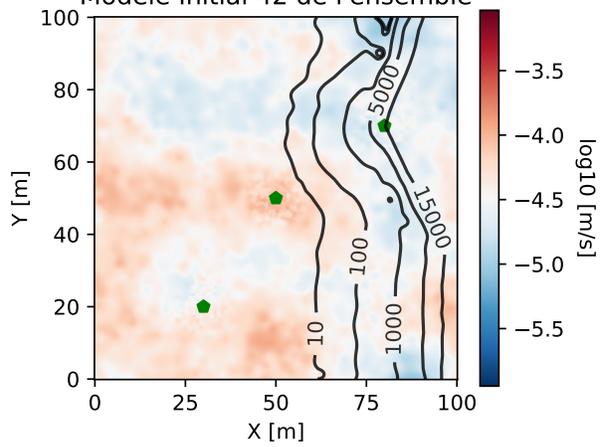
Conductivité hydraulique Z = 8 m  
Modèle de référence



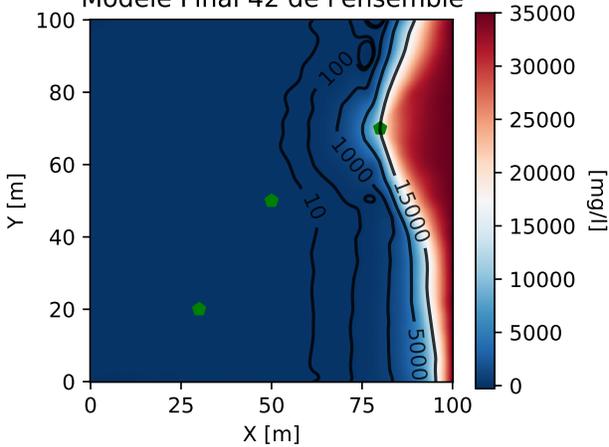
Concentration en sel Z = 8 m, 400 jours  
Modèle Initial 42 de l'ensemble



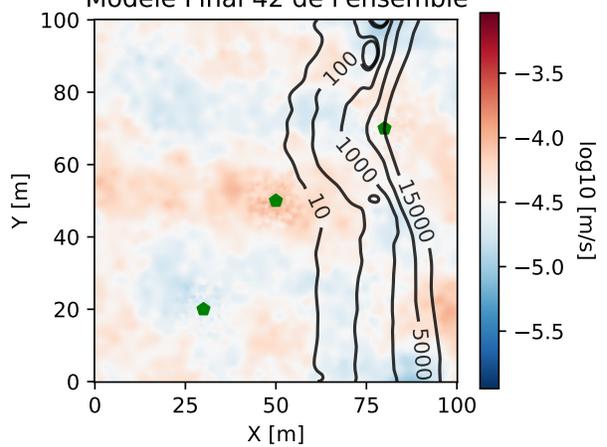
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



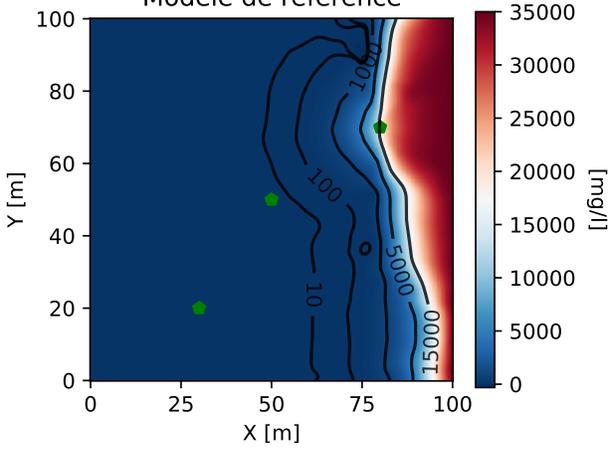
Concentration en sel Z = 8 m, 400 jours  
Modèle Final 42 de l'ensemble



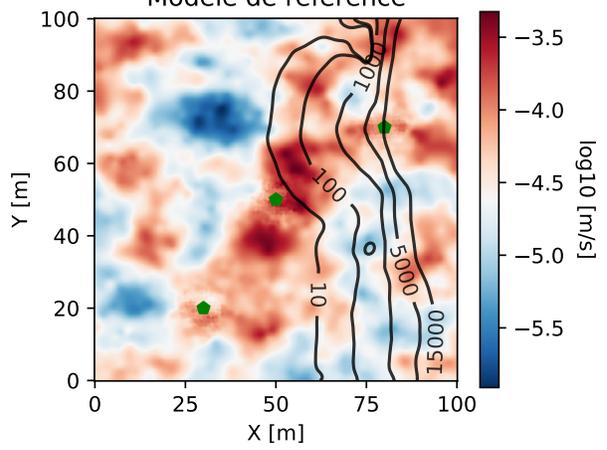
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



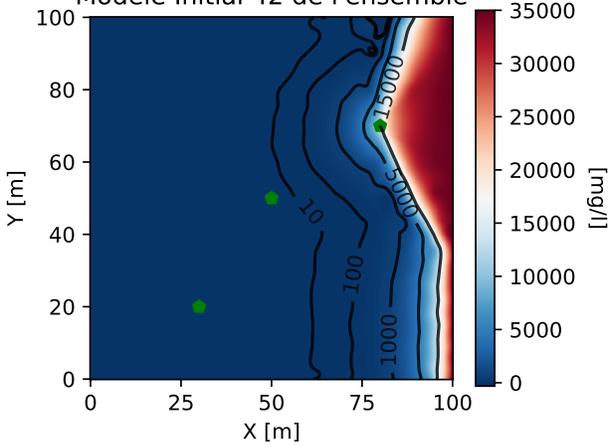
Concentration en sel Z = 8 m, 500 jours  
Modèle de référence



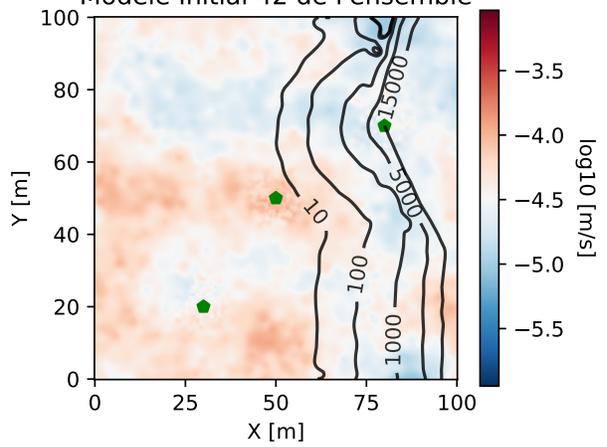
Conductivité hydraulique Z = 8 m  
Modèle de référence



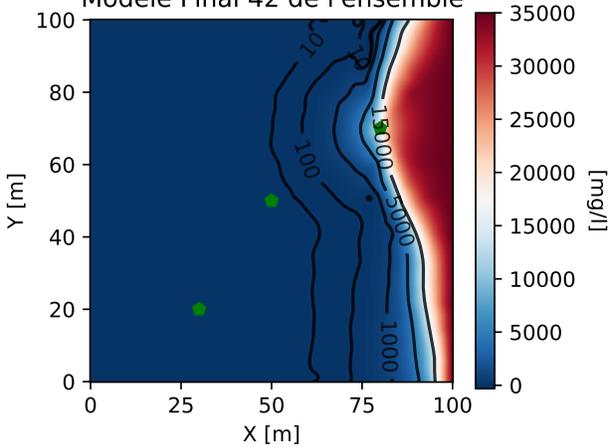
Concentration en sel Z = 8 m, 500 jours  
Modèle Initial 42 de l'ensemble



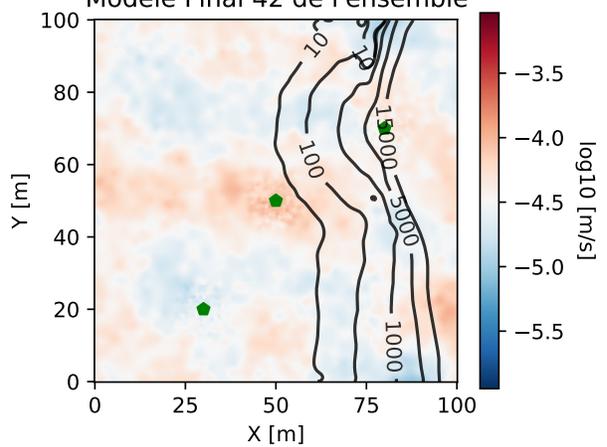
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



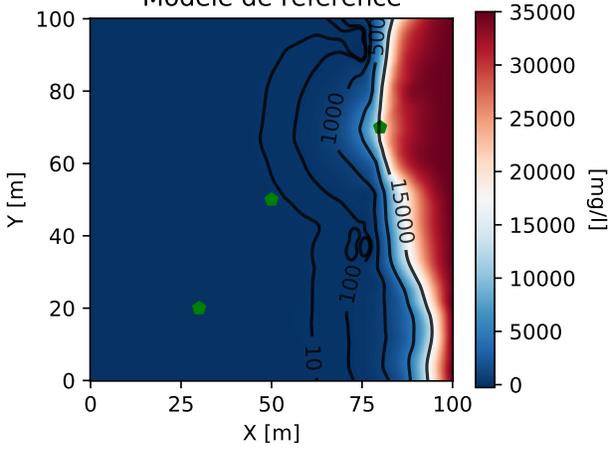
Concentration en sel Z = 8 m, 500 jours  
Modèle Final 42 de l'ensemble



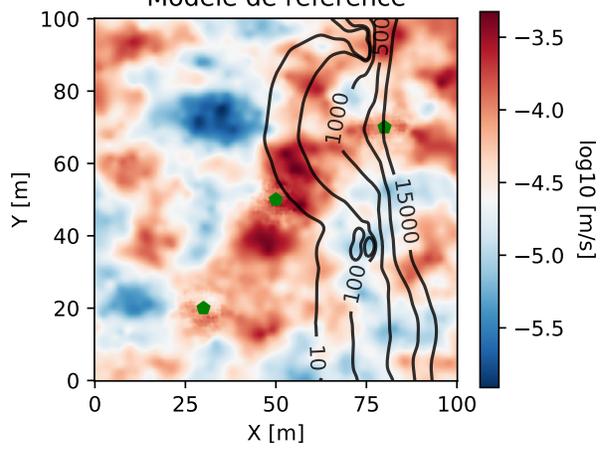
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



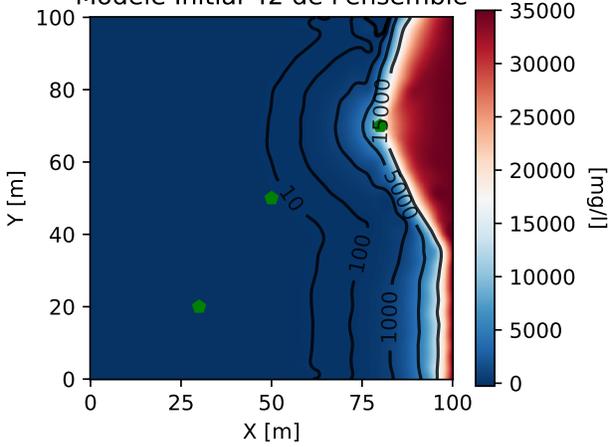
Concentration en sel Z = 8 m, 600 jours  
Modèle de référence



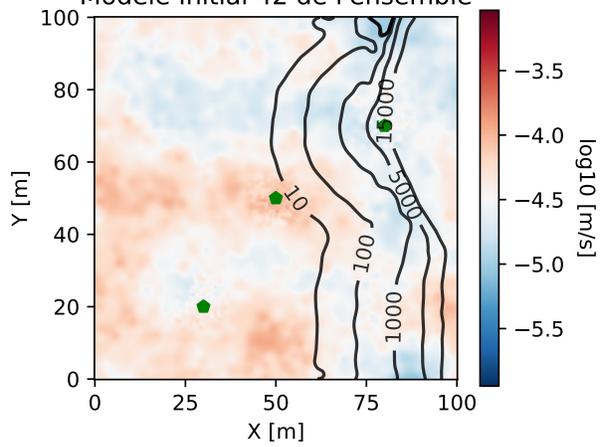
Conductivité hydraulique Z = 8 m  
Modèle de référence



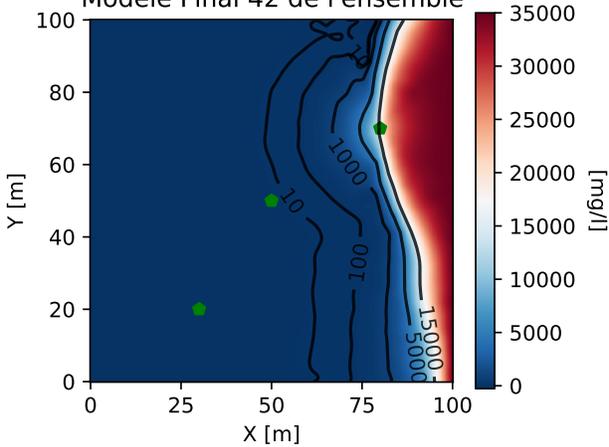
Concentration en sel Z = 8 m, 600 jours  
Modèle Initial 42 de l'ensemble



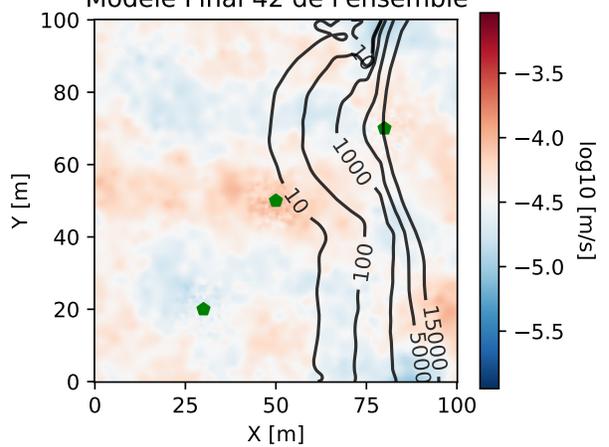
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



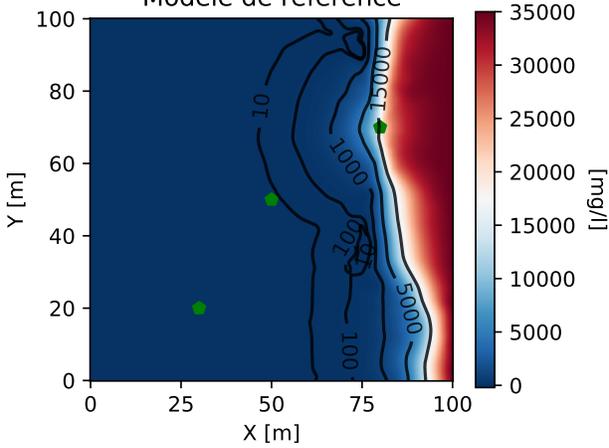
Concentration en sel Z = 8 m, 600 jours  
Modèle Final 42 de l'ensemble



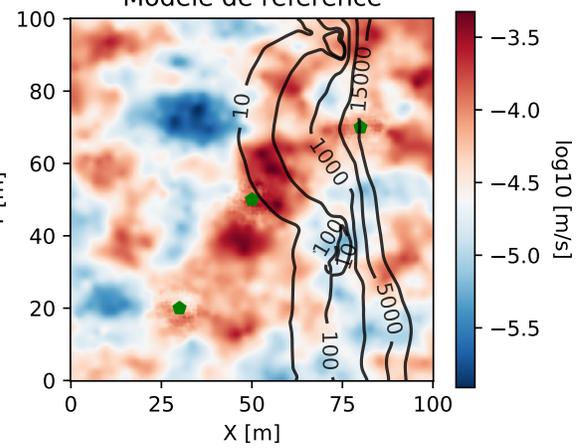
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



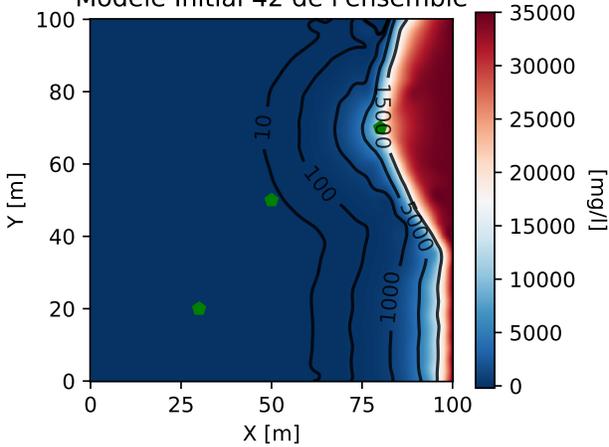
Concentration en sel Z = 8 m, 700 jours  
Modèle de référence



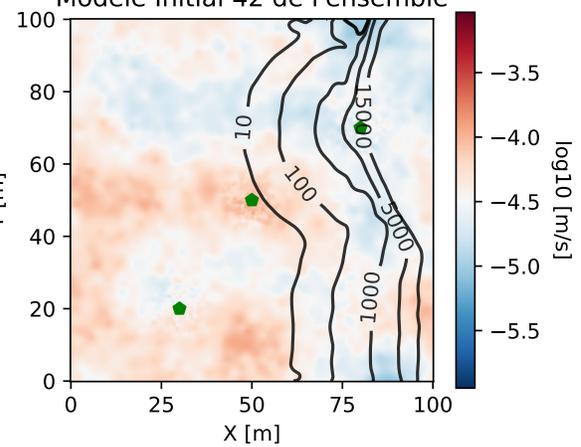
Conductivité hydraulique Z = 8 m  
Modèle de référence



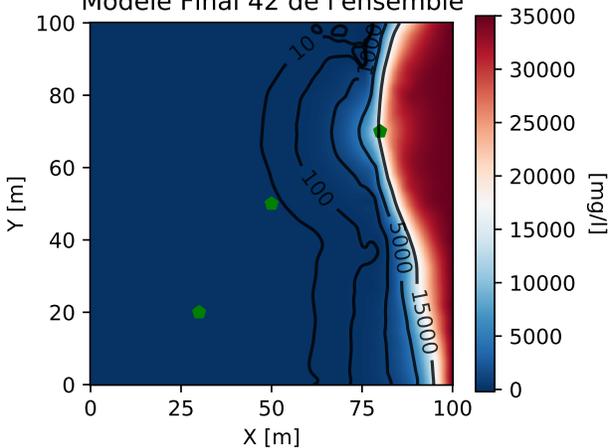
Concentration en sel Z = 8 m, 700 jours  
Modèle Initial 42 de l'ensemble



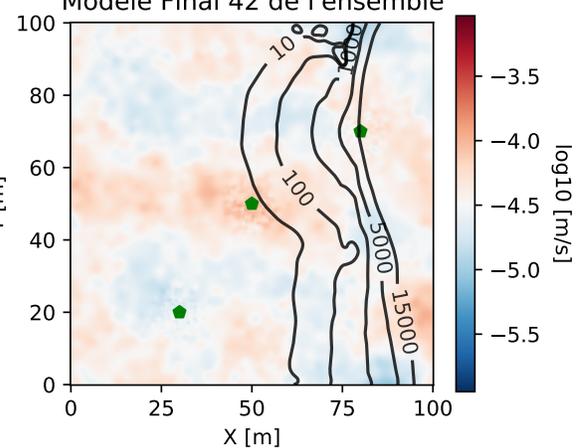
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



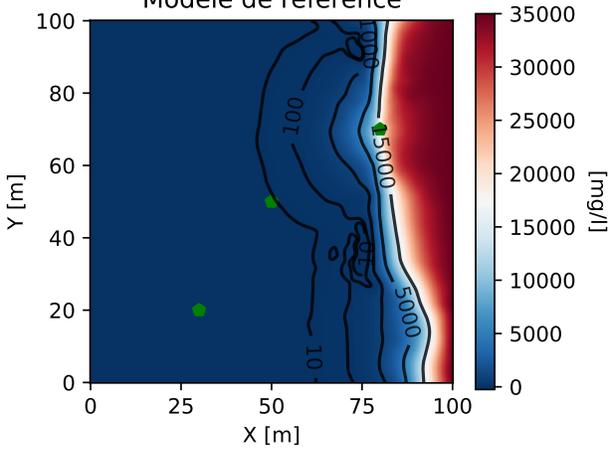
Concentration en sel Z = 8 m, 700 jours  
Modèle Final 42 de l'ensemble



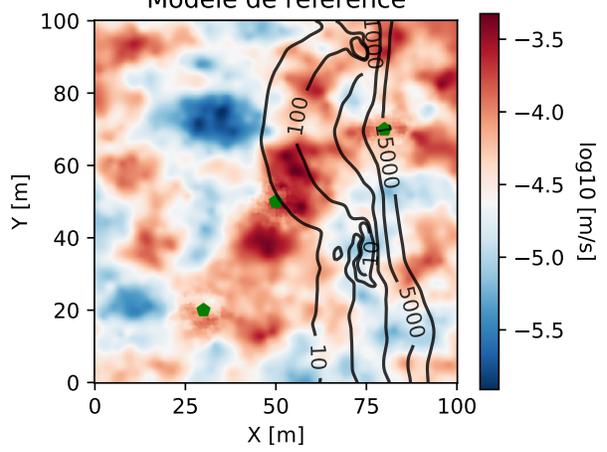
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



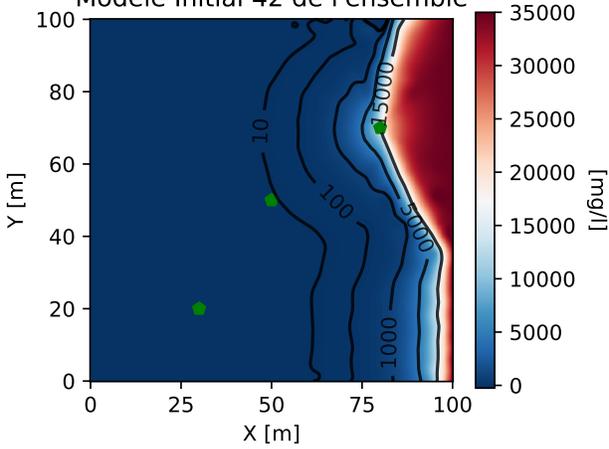
Concentration en sel Z = 8 m, 800 jours  
Modèle de référence



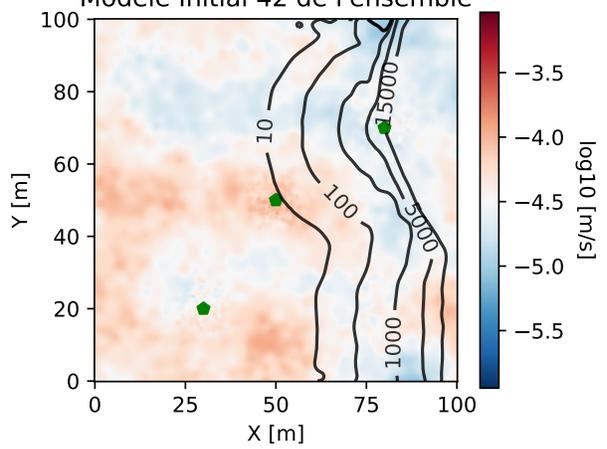
Conductivité hydraulique Z = 8 m  
Modèle de référence



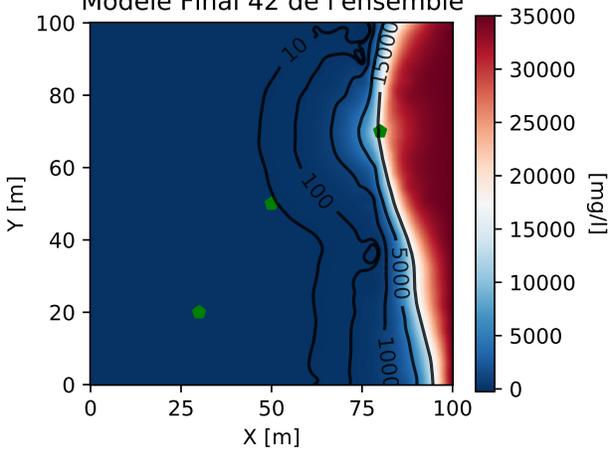
Concentration en sel Z = 8 m, 800 jours  
Modèle Initial 42 de l'ensemble



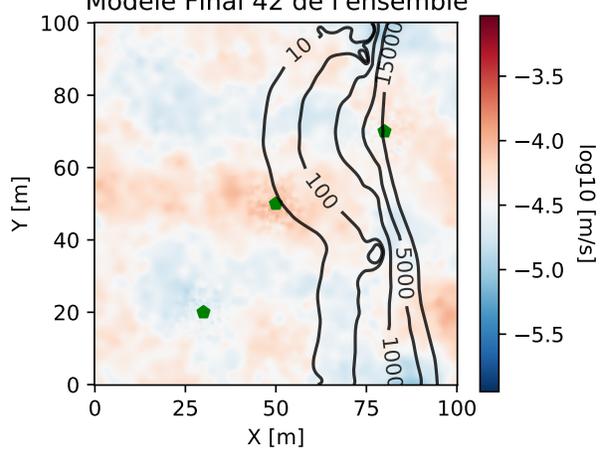
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



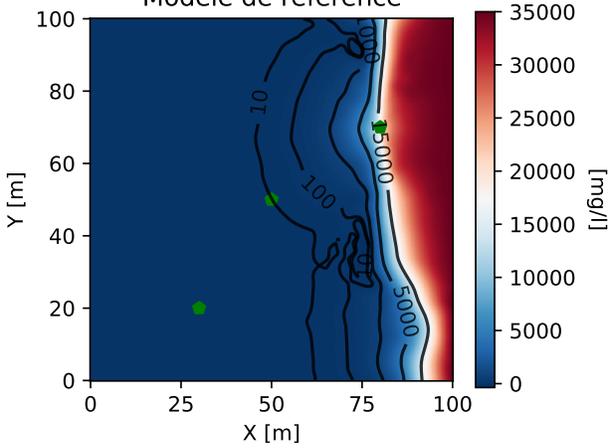
Concentration en sel Z = 8 m, 800 jours  
Modèle Final 42 de l'ensemble



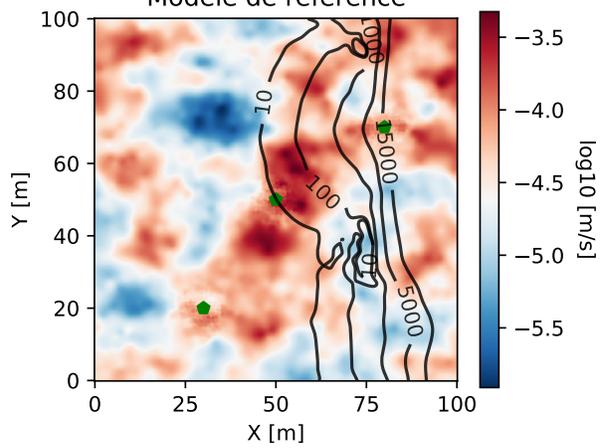
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



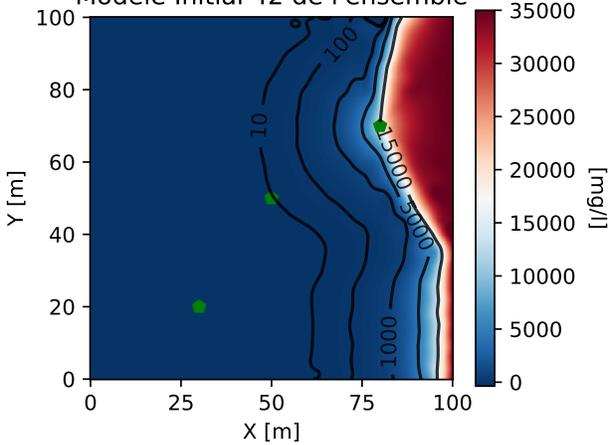
Concentration en sel Z = 8 m, 900 jours  
Modèle de référence



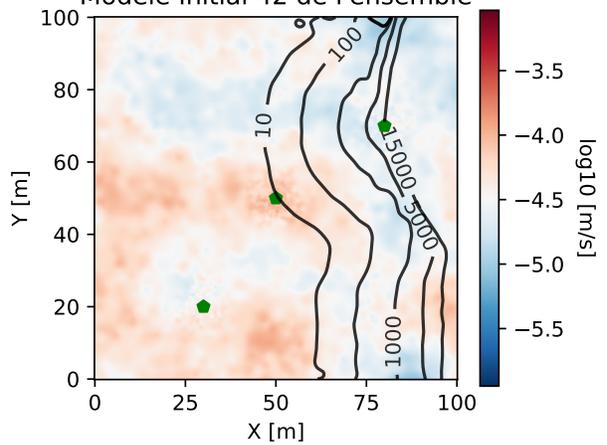
Conductivité hydraulique Z = 8 m  
Modèle de référence



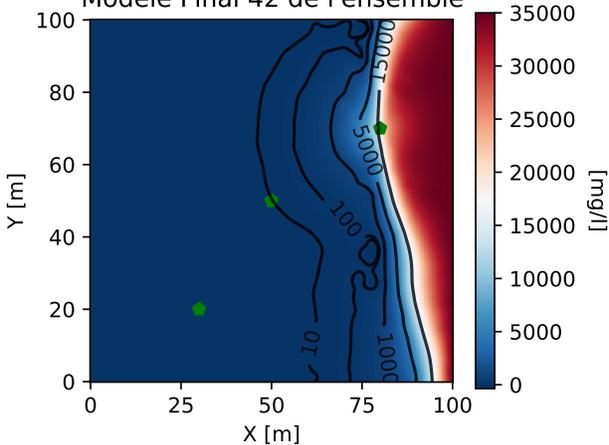
Concentration en sel Z = 8 m, 900 jours  
Modèle Initial 42 de l'ensemble



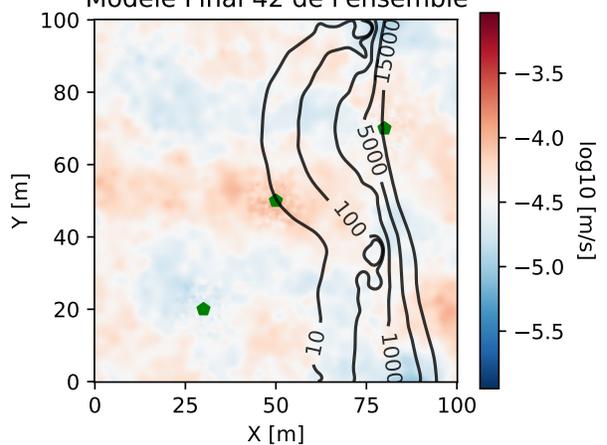
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



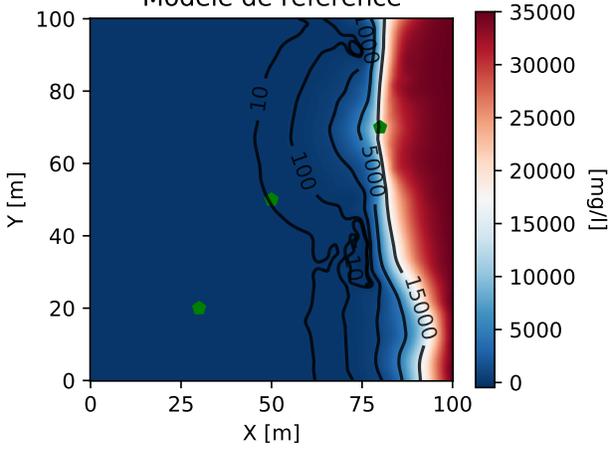
Concentration en sel Z = 8 m, 900 jours  
Modèle Final 42 de l'ensemble



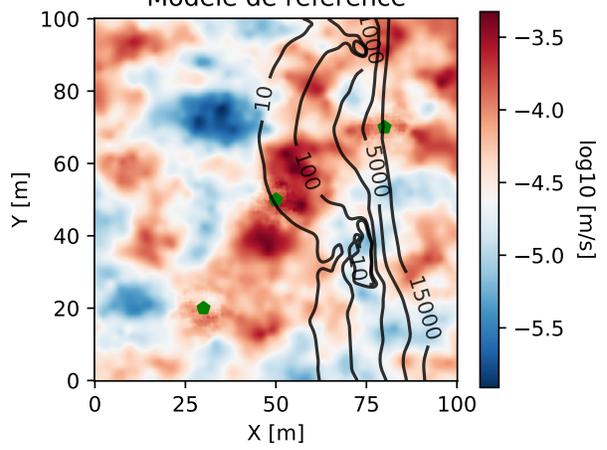
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



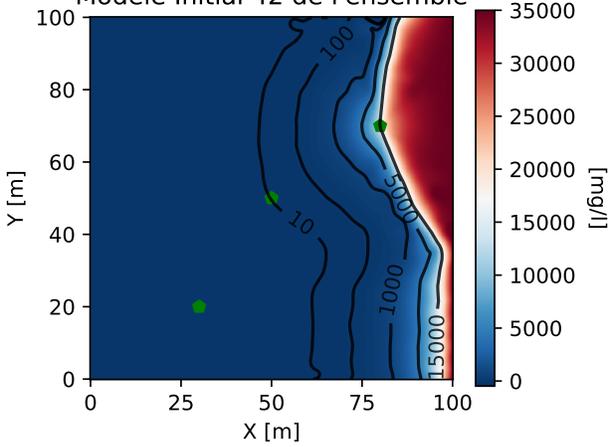
Concentration en sel Z = 8 m, 1000 jours  
Modèle de référence



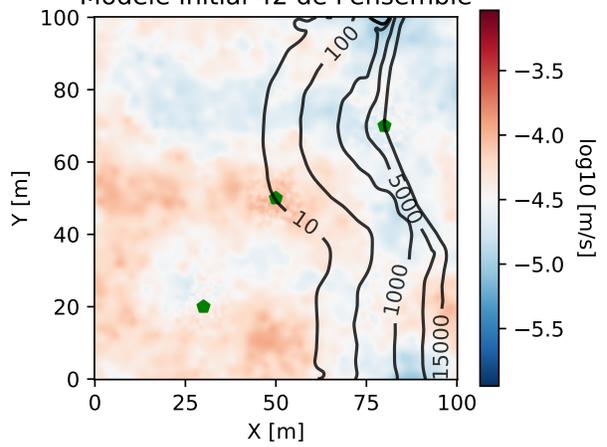
Conductivité hydraulique Z = 8 m  
Modèle de référence



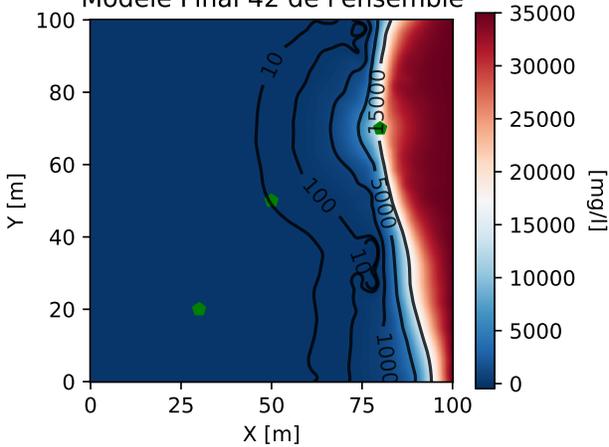
Concentration en sel Z = 8 m, 1000 jours  
Modèle Initial 42 de l'ensemble



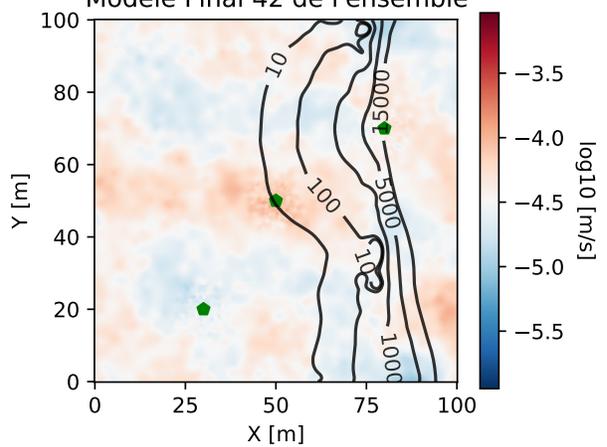
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



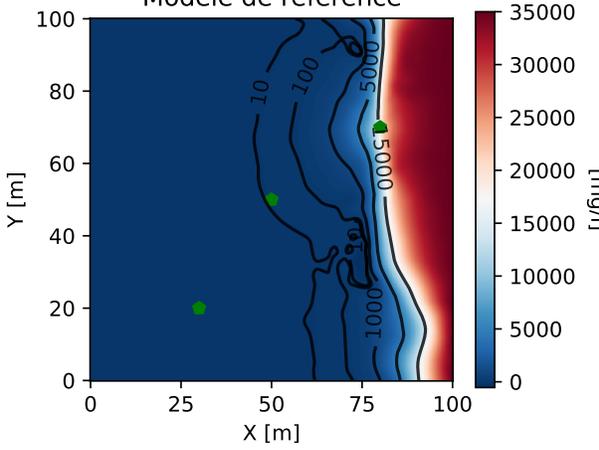
Concentration en sel Z = 8 m, 1000 jours  
Modèle Final 42 de l'ensemble



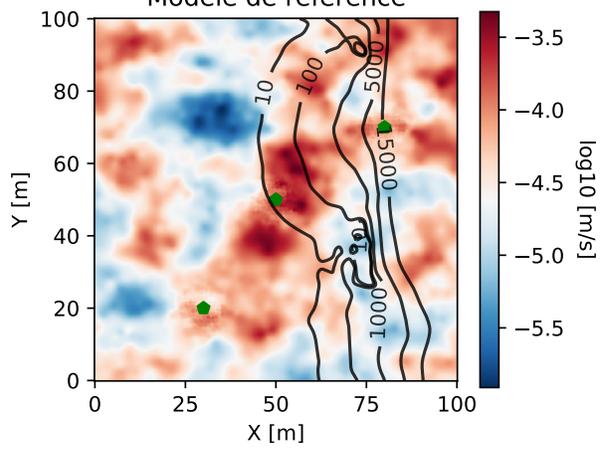
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



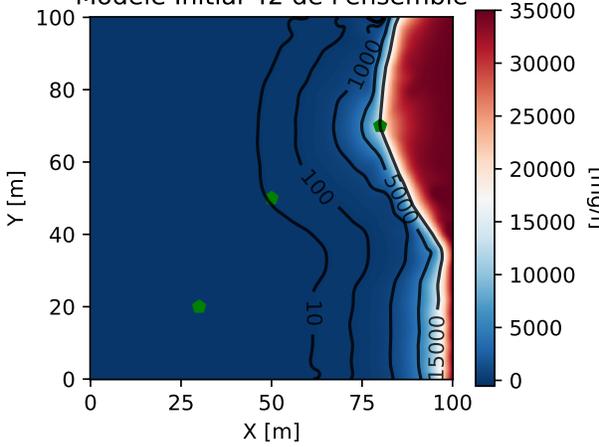
Concentration en sel Z = 8 m, 1100 jours  
Modèle de référence



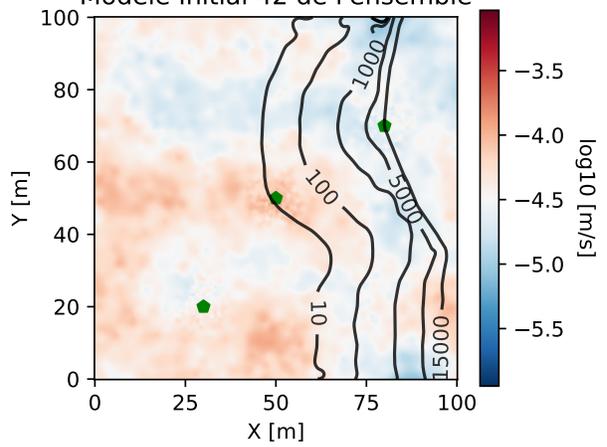
Conductivité hydraulique Z = 8 m  
Modèle de référence



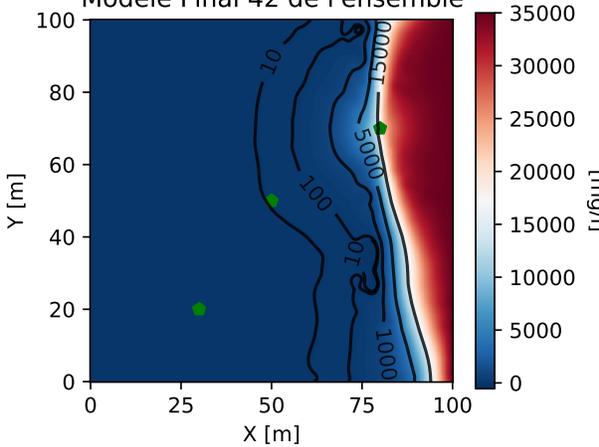
Concentration en sel Z = 8 m, 1100 jours  
Modèle Initial 42 de l'ensemble



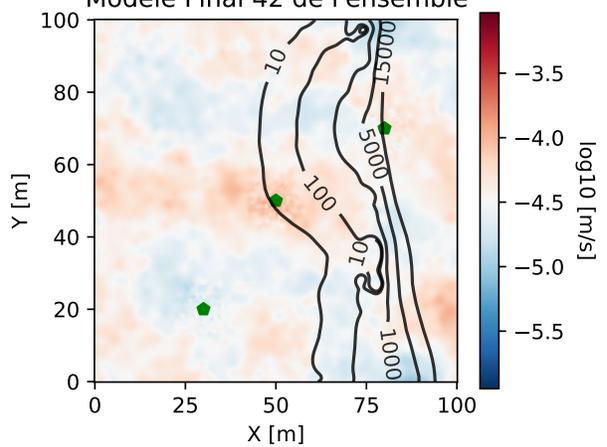
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



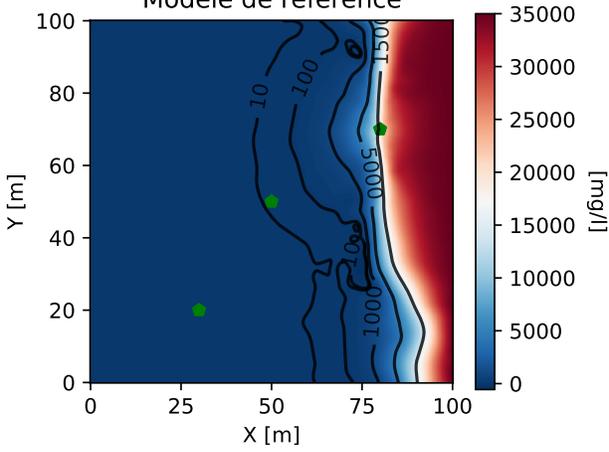
Concentration en sel Z = 8 m, 1100 jours  
Modèle Final 42 de l'ensemble



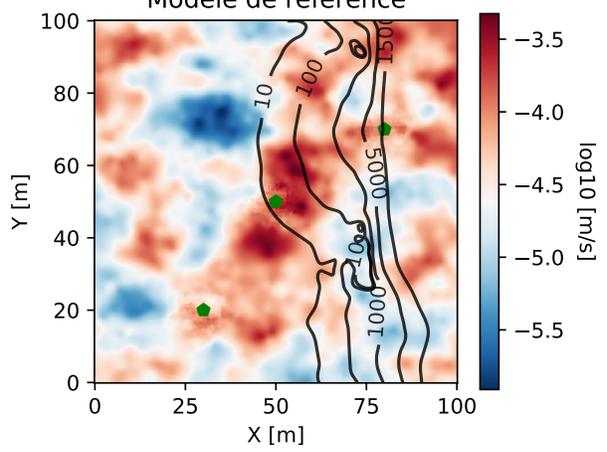
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



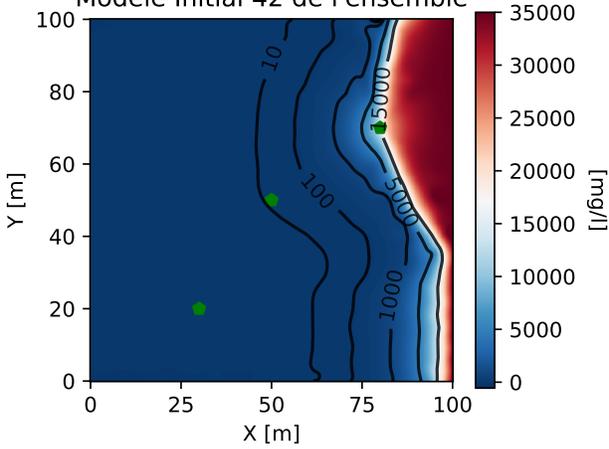
Concentration en sel Z = 8 m, 1200 jours  
Modèle de référence



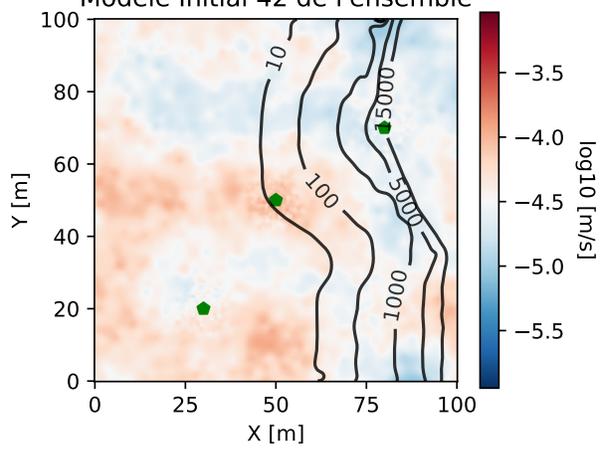
Conductivité hydraulique Z = 8 m  
Modèle de référence



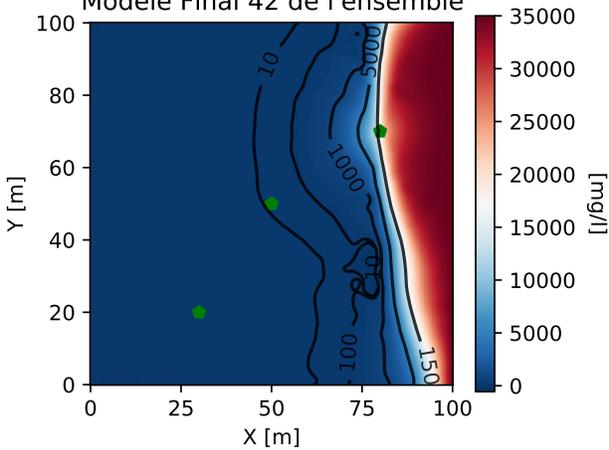
Concentration en sel Z = 8 m, 1200 jours  
Modèle Initial 42 de l'ensemble



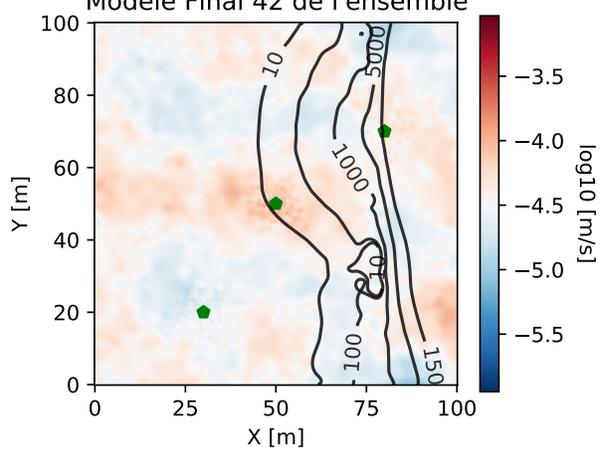
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



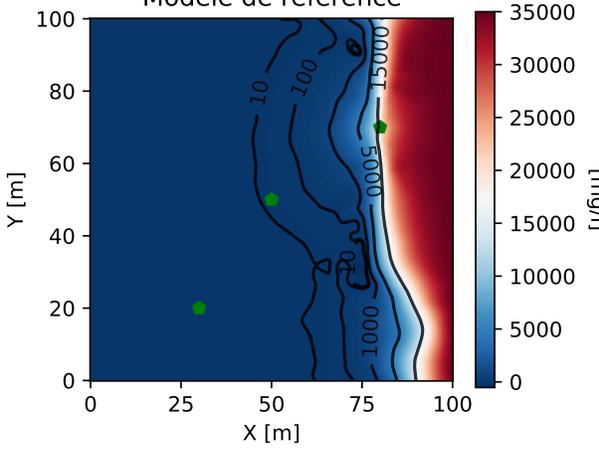
Concentration en sel Z = 8 m, 1200 jours  
Modèle Final 42 de l'ensemble



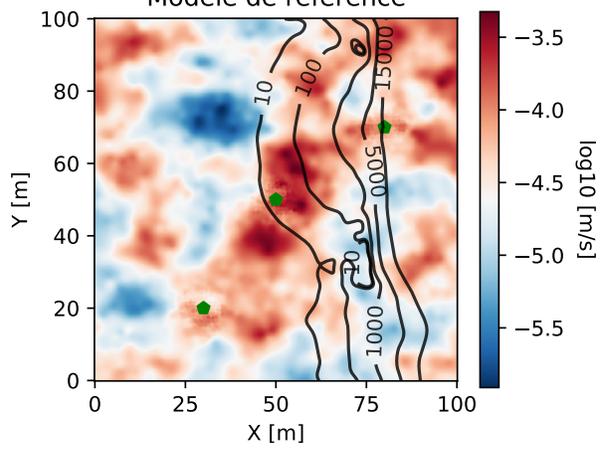
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



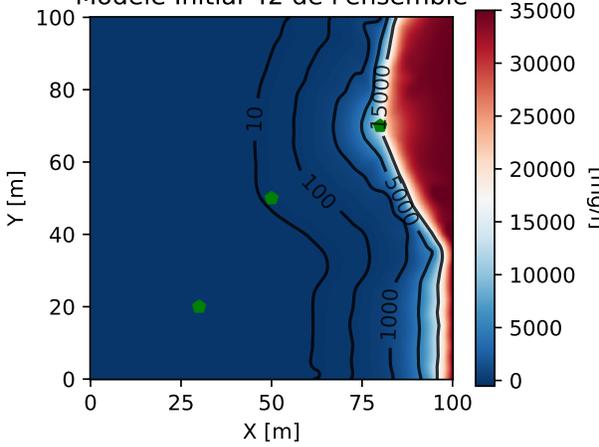
Concentration en sel Z = 8 m, 1300 jours  
Modèle de référence



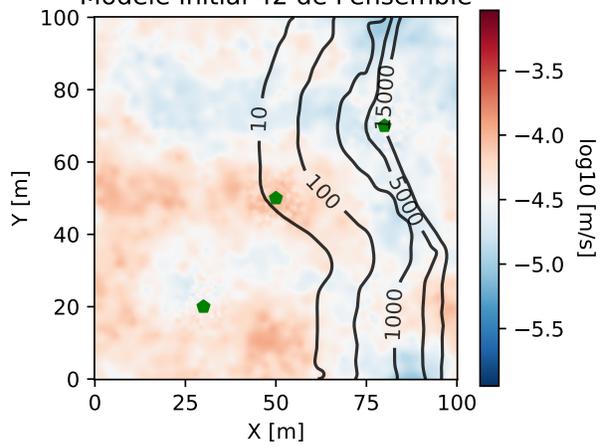
Conductivité hydraulique Z = 8 m  
Modèle de référence



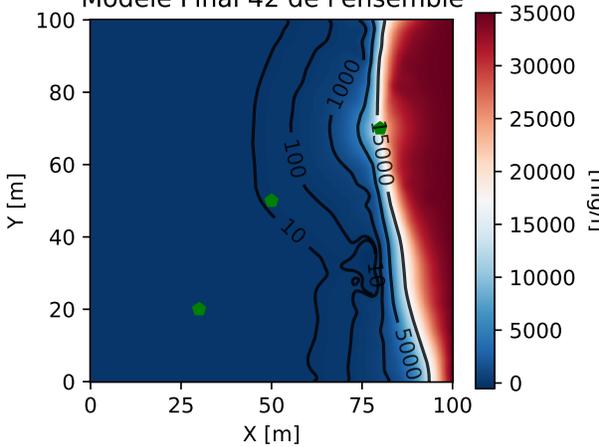
Concentration en sel Z = 8 m, 1300 jours  
Modèle Initial 42 de l'ensemble



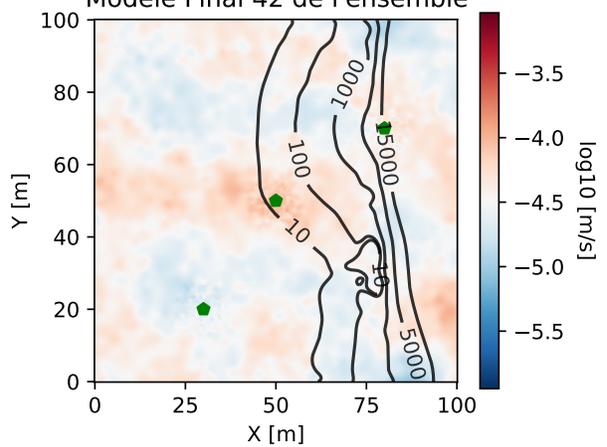
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



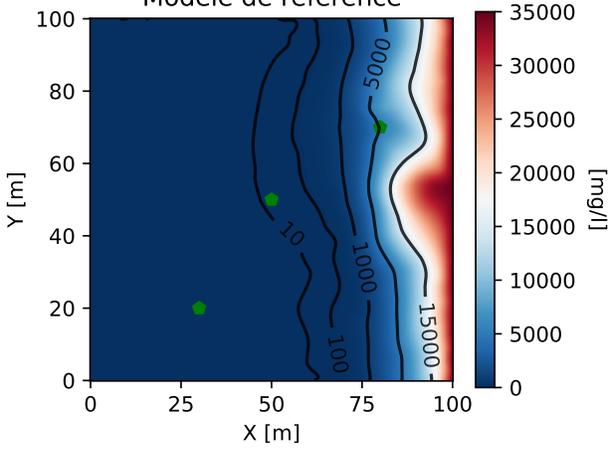
Concentration en sel Z = 8 m, 1300 jours  
Modèle Final 42 de l'ensemble



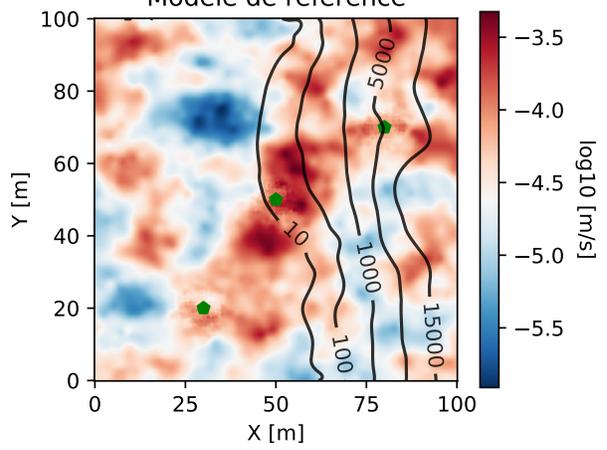
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



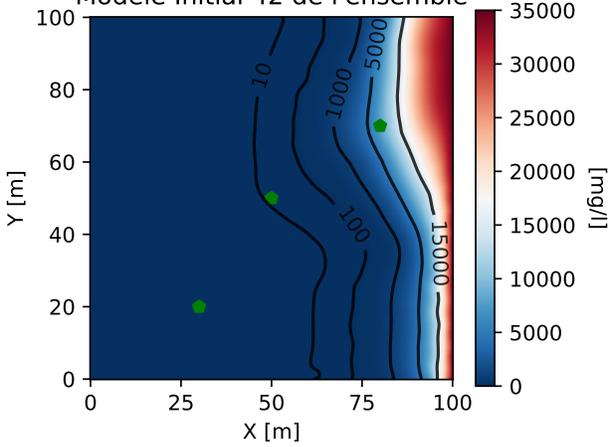
Concentration en sel Z = 8 m, 1400 jours  
Modèle de référence



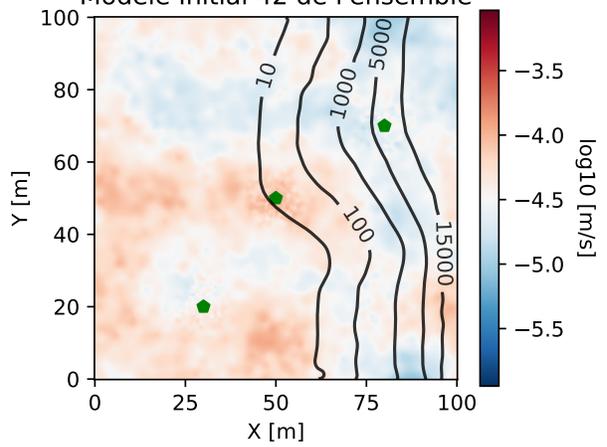
Conductivité hydraulique Z = 8 m  
Modèle de référence



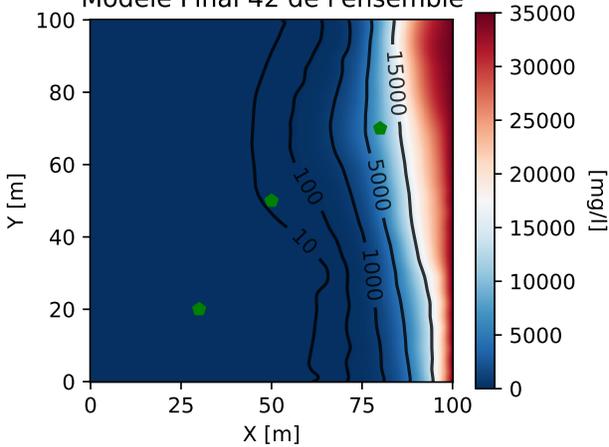
Concentration en sel Z = 8 m, 1400 jours  
Modèle Initial 42 de l'ensemble



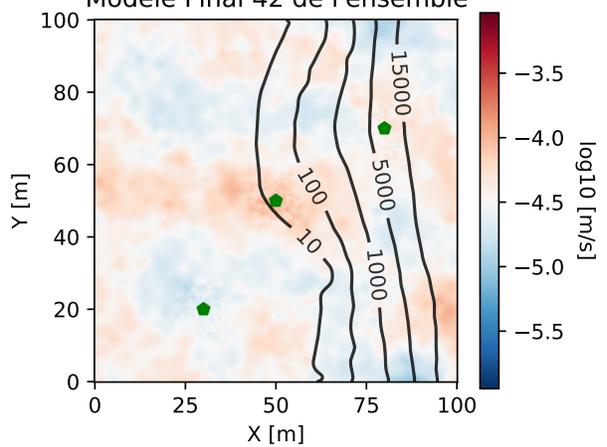
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



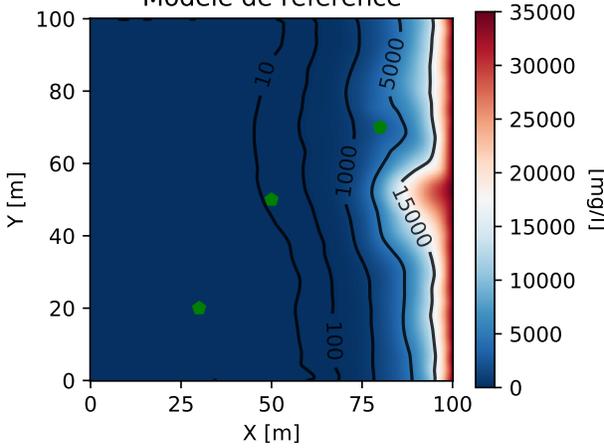
Concentration en sel Z = 8 m, 1400 jours  
Modèle Final 42 de l'ensemble



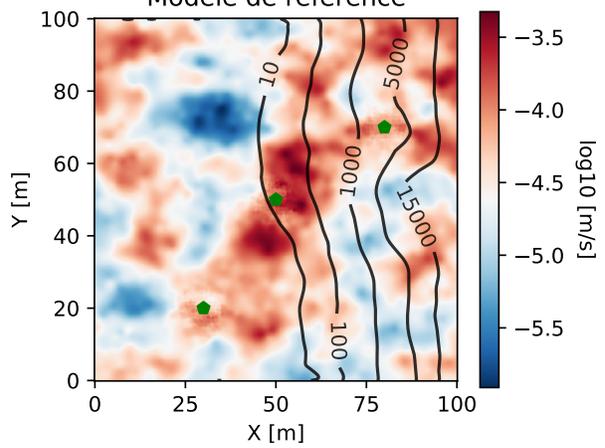
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



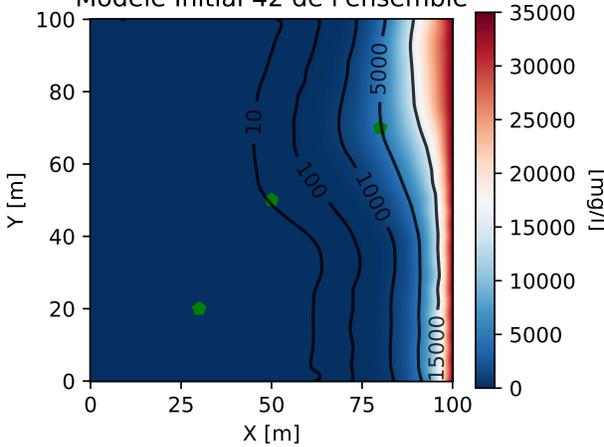
Concentration en sel Z = 8 m, 1500 jours  
Modèle de référence



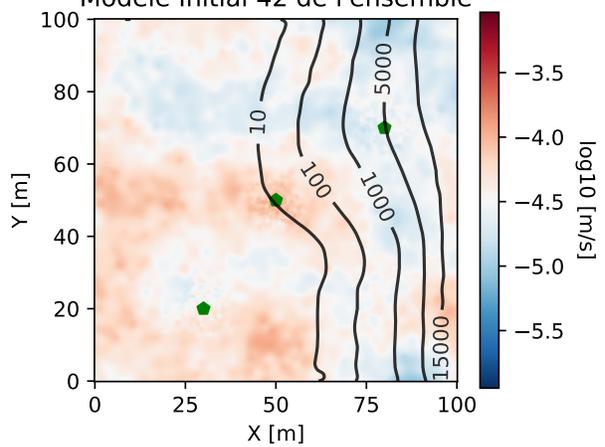
Conductivité hydraulique Z = 8 m  
Modèle de référence



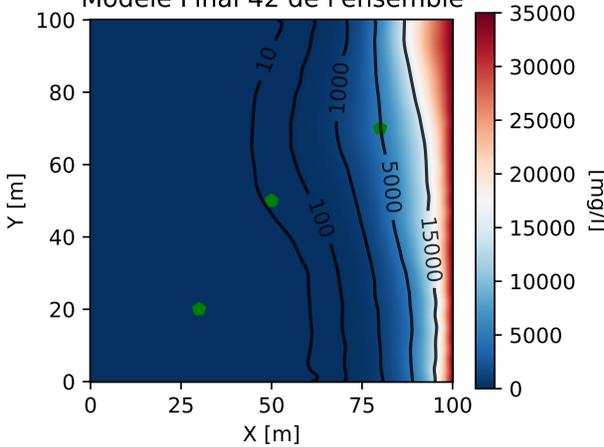
Concentration en sel Z = 8 m, 1500 jours  
Modèle Initial 42 de l'ensemble



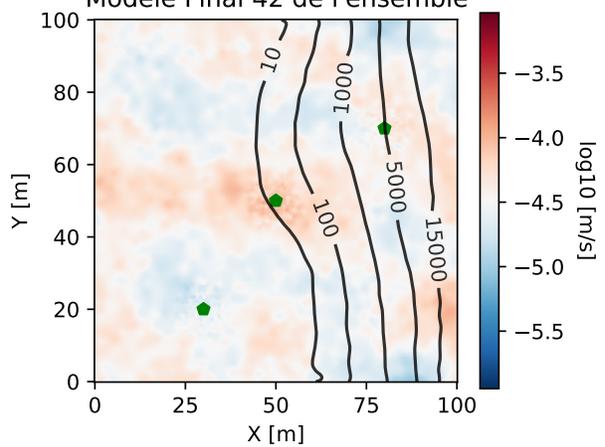
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



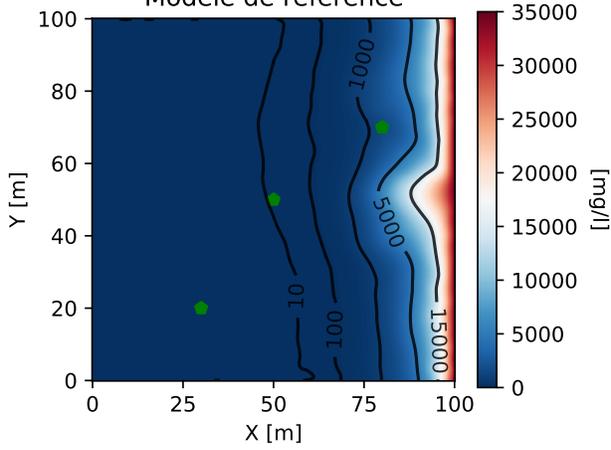
Concentration en sel Z = 8 m, 1500 jours  
Modèle Final 42 de l'ensemble



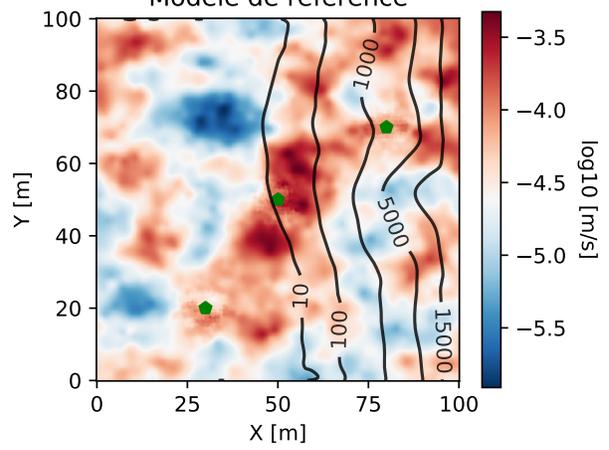
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



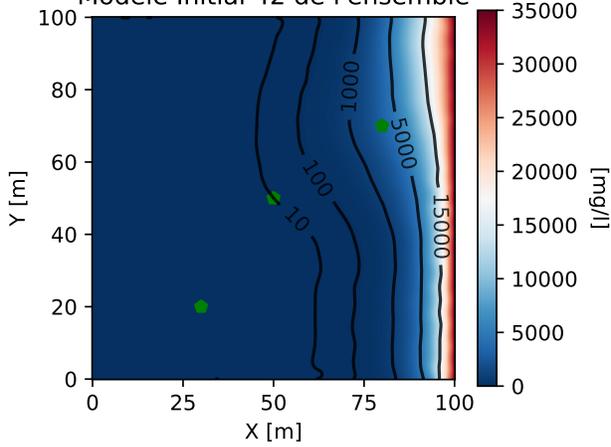
Concentration en sel Z = 8 m, 1600 jours  
Modèle de référence



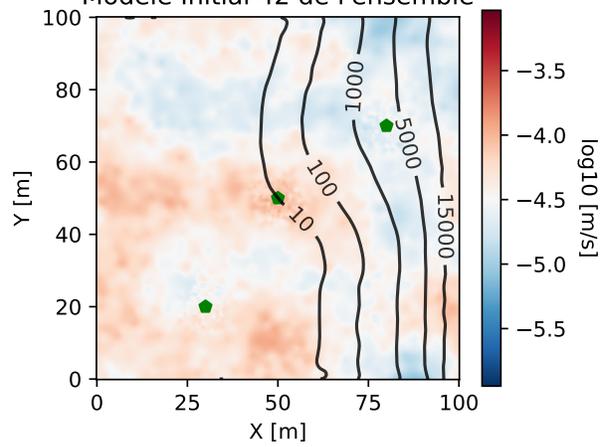
Conductivité hydraulique Z = 8 m  
Modèle de référence



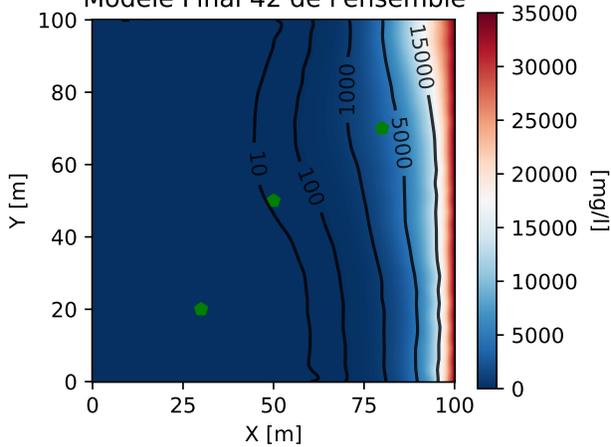
Concentration en sel Z = 8 m, 1600 jours  
Modèle Initial 42 de l'ensemble



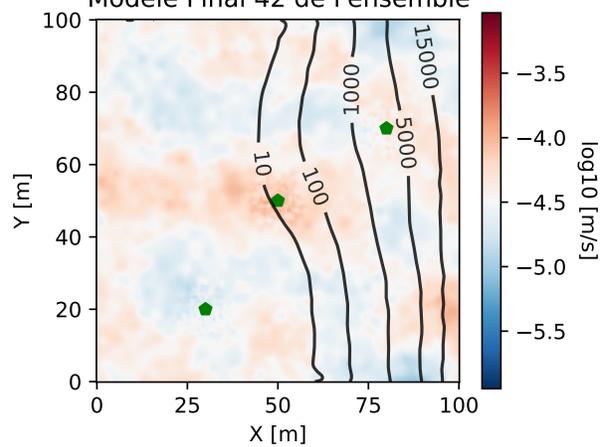
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



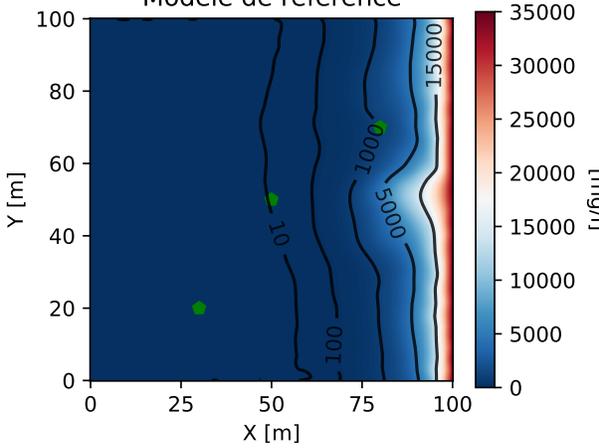
Concentration en sel Z = 8 m, 1600 jours  
Modèle Final 42 de l'ensemble



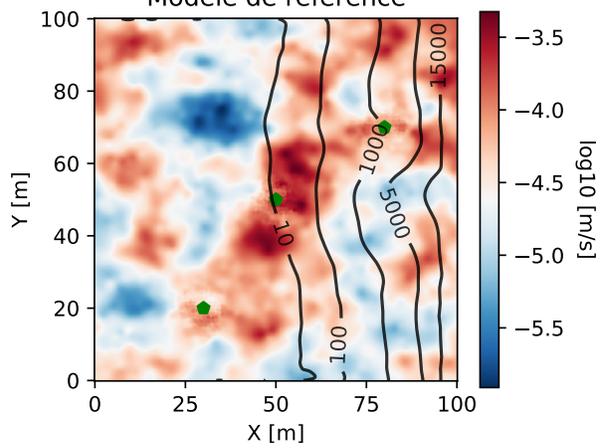
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



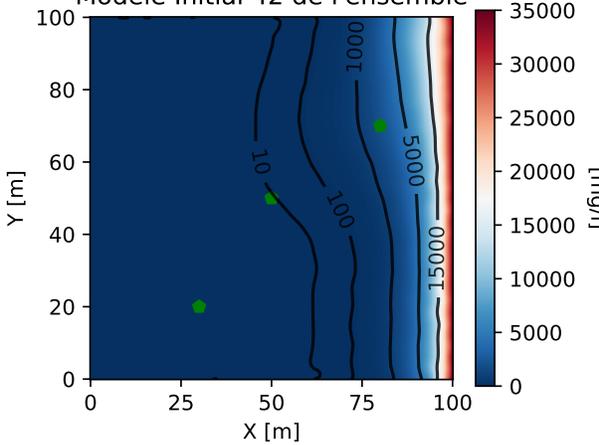
Concentration en sel Z = 8 m, 1700 jours  
Modèle de référence



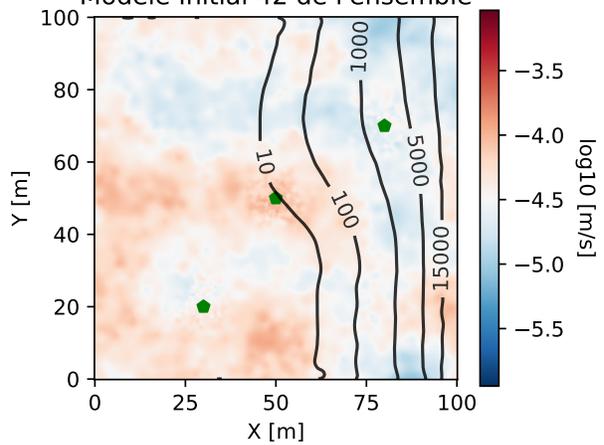
Conductivité hydraulique Z = 8 m  
Modèle de référence



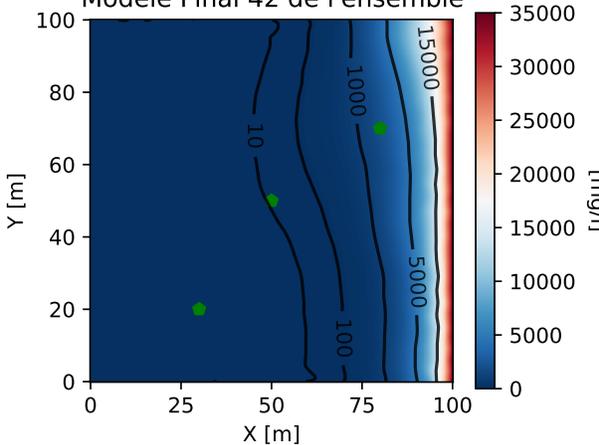
Concentration en sel Z = 8 m, 1700 jours  
Modèle Initial 42 de l'ensemble



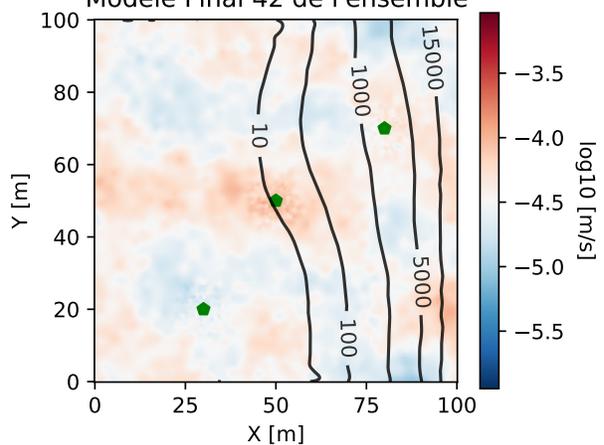
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



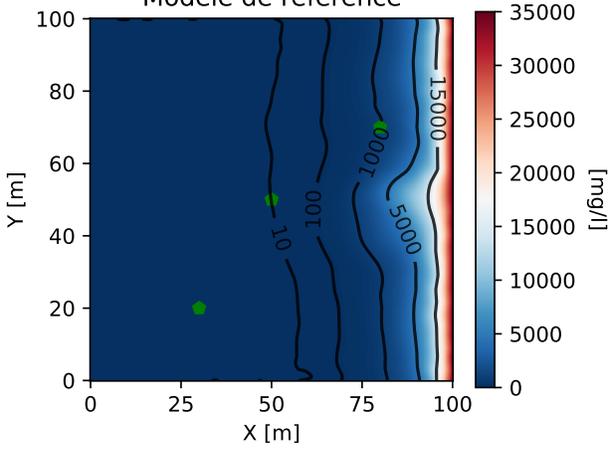
Concentration en sel Z = 8 m, 1700 jours  
Modèle Final 42 de l'ensemble



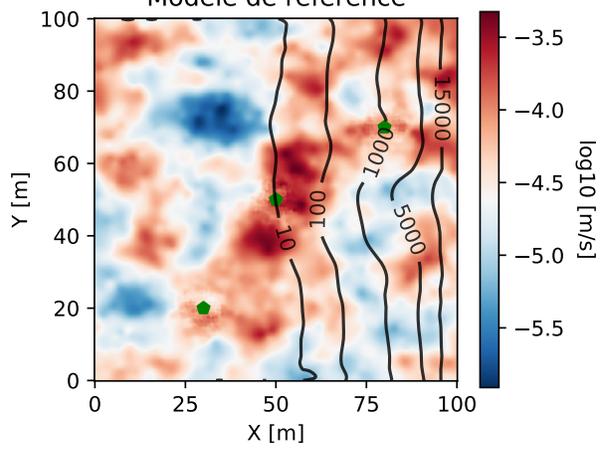
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



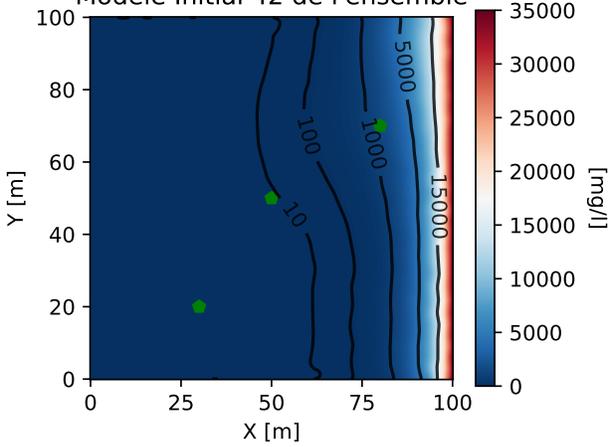
Concentration en sel Z = 8 m, 1800 jours  
Modèle de référence



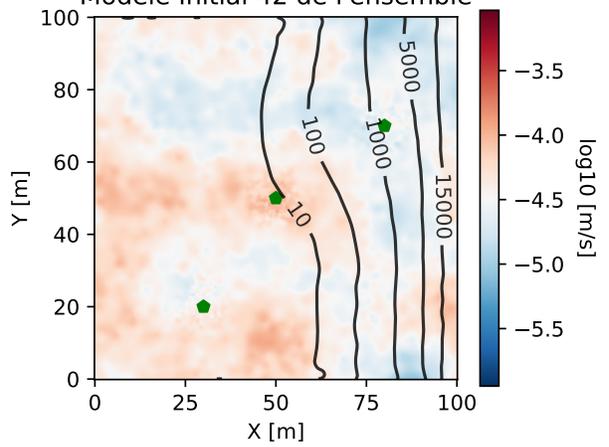
Conductivité hydraulique Z = 8 m  
Modèle de référence



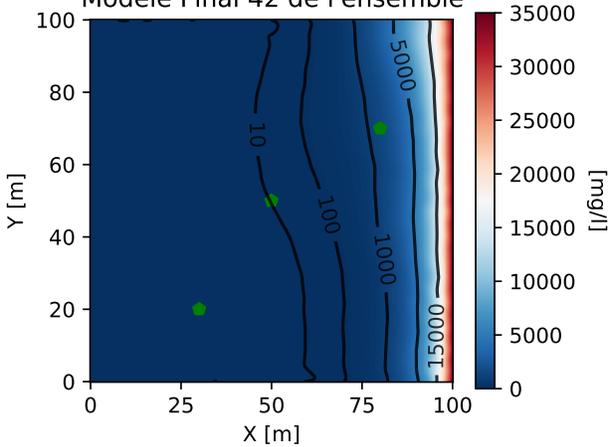
Concentration en sel Z = 8 m, 1800 jours  
Modèle Initial 42 de l'ensemble



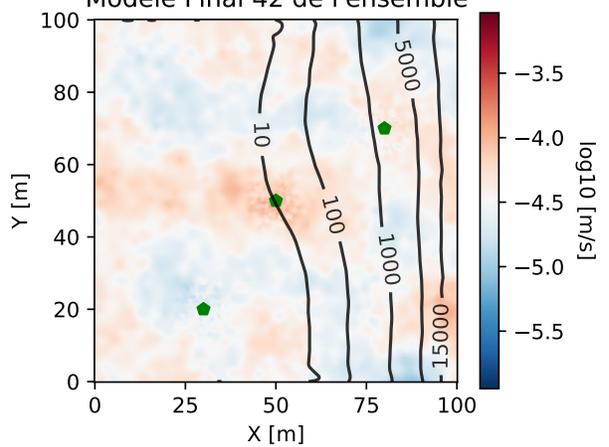
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



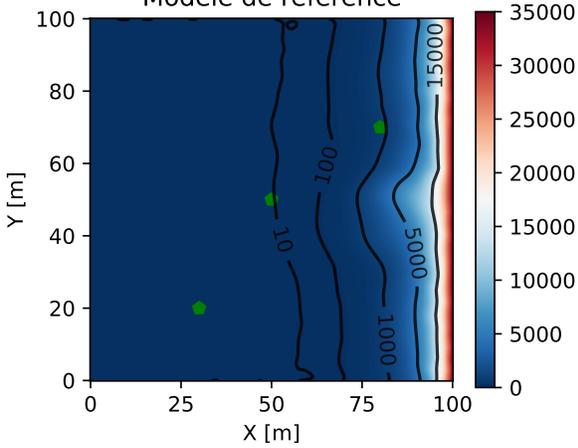
Concentration en sel Z = 8 m, 1800 jours  
Modèle Final 42 de l'ensemble



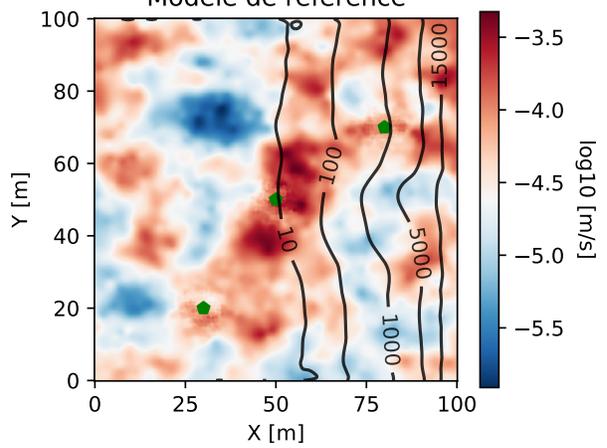
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



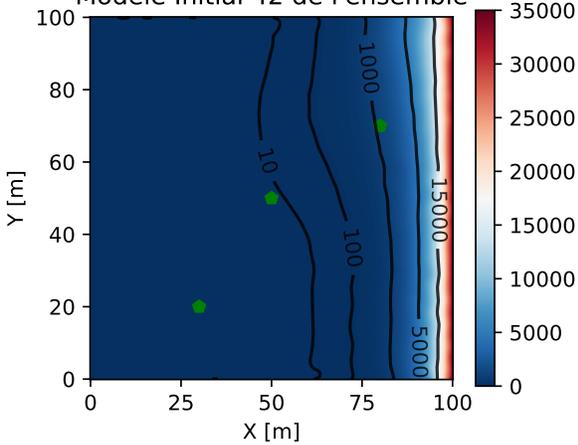
Concentration en sel Z = 8 m, 1900 jours  
Modèle de référence



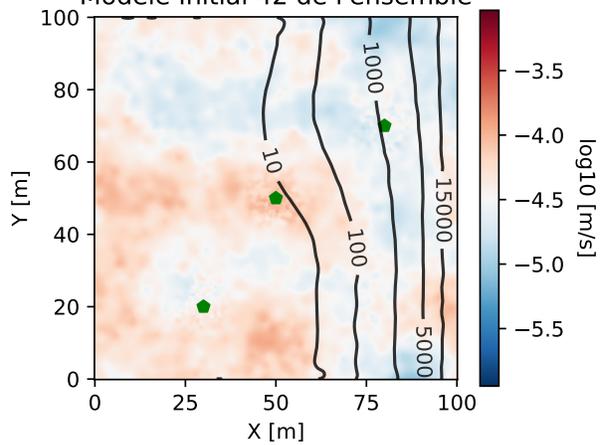
Conductivité hydraulique Z = 8 m  
Modèle de référence



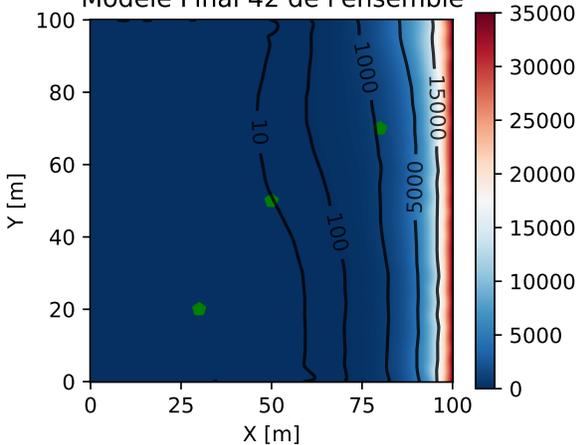
Concentration en sel Z = 8 m, 1900 jours  
Modèle Initial 42 de l'ensemble



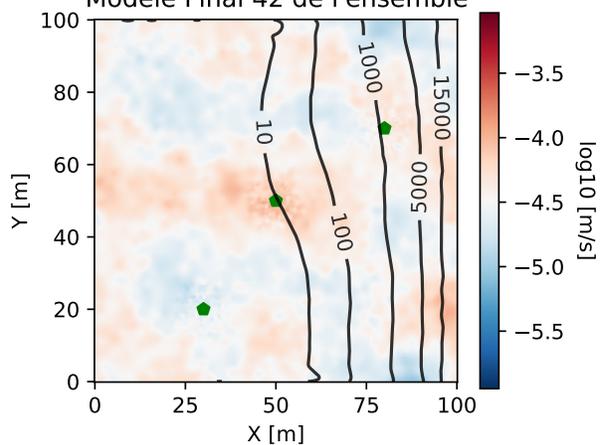
Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



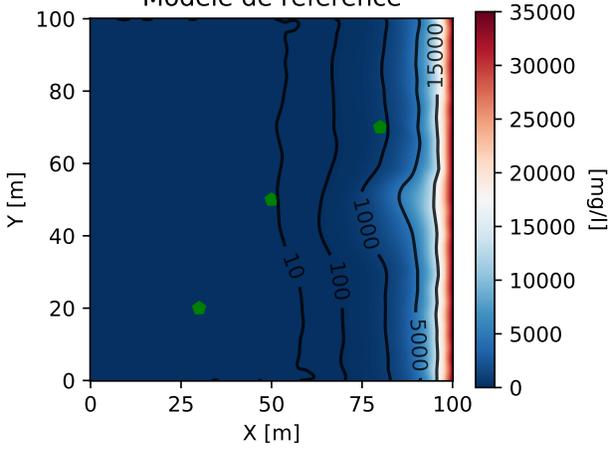
Concentration en sel Z = 8 m, 1900 jours  
Modèle Final 42 de l'ensemble



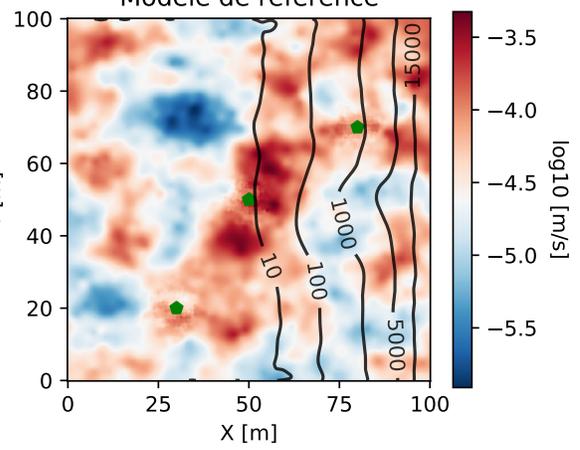
Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble



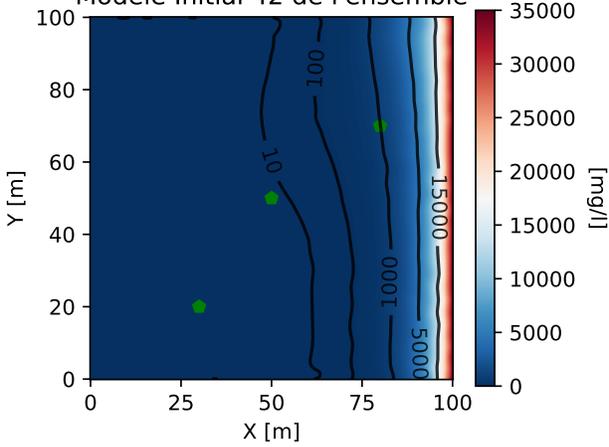
Concentration en sel Z = 8 m, 2000 jours  
Modèle de référence



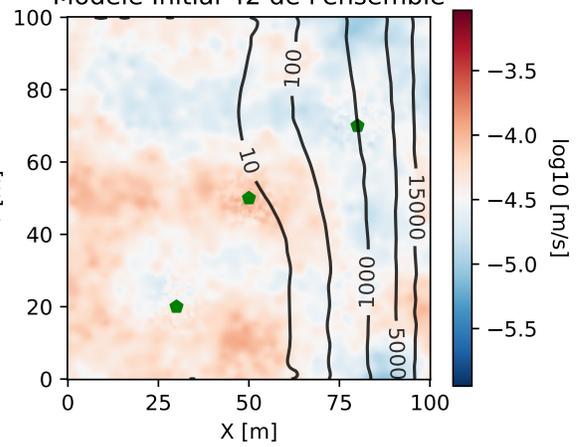
Conductivité hydraulique Z = 8 m  
Modèle de référence



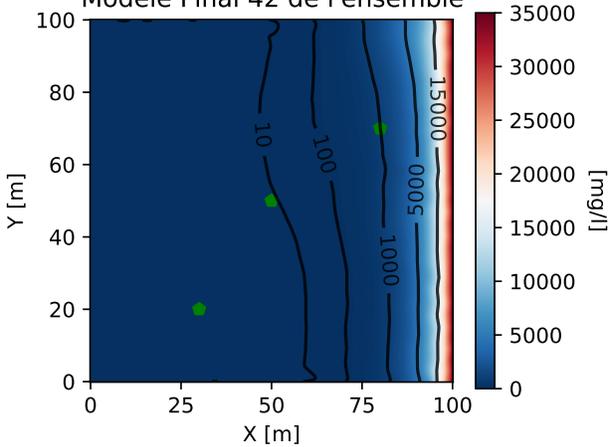
Concentration en sel Z = 8 m, 2000 jours  
Modèle Initial 42 de l'ensemble



Conductivité hydraulique Z = 8 m  
Modèle Initial 42 de l'ensemble



Concentration en sel Z = 8 m, 2000 jours  
Modèle Final 42 de l'ensemble



Conductivité hydraulique Z = 8 m  
Modèle Final 42 de l'ensemble

