

Chapitre 5

Application de la méthode stochastique alternative au cas d'étude synthétique

Avant de se lancer dans l'application proprement dite de la méthode stochastique alternative MSA développée dans le cadre de ce travail et présentée au chapitre 3, deux notions importantes utilisées par la suite sont développées afin de rendre leur compréhension plus aisée : la méthode de calcul de la distribution spatiale d'un ensemble de zones de protection et les techniques de seuillage.

5.1 Distribution spatiale d'un ensemble de zones de protection

Dans la MSA, un grand nombre de réalisations équiprobables du champ de conductivité hydraulique sont engendrées, toutes conditionnées par les données existantes. Chacun de ces champs est utilisé dans une délimitation numérique et déterministe de la zone de protection : d'abord l'équation d'écoulement (4.6) est résolue numériquement à l'aide du logiciel MODFLOW, puis le trajet de particules imaginaires (positionnées au centre de chaque cellule de la grille dans le champ d'écoulement calculé) est suivi par *particle tracking* dans la direction de l'écoulement (logiciel MODPATH). A partir du calcul des temps de transfert

respectifs de ces particules au puits et pour chaque réalisation k engendrée, une isochrone $\tau_k(\underline{x}, t)$ définie comme la limite de la zone de protection $CAP_k(\underline{x}, t)$ (*) est déterminée. Elle enveloppe toutes les positions initiales \underline{x} (représentant les coordonnées généralisées (x, y)) des particules dans l'espace, dont le temps de transfert au puits $T \leq t$. Toutes les zones de protection obtenues sont alors considérées comme les réalisations d'une fonction aléatoire. L'analyse de Monte Carlo de l'ensemble de ces zones de protection aboutit à une approximation numérique de la **distribution spatiale de l'ensemble des zones de protection** $[CAP(\underline{x}, t)]$, définie par :

$$[CAP(\underline{x}, t)] = \frac{1}{n} \sum_{k=1}^n (I(\underline{x}, t))_k \quad (5.1)$$

où n est le nombre de champs de conductivité hydraulique générés et $(I(\underline{x}, t))_k$ l'événement "particule capturée par le puits endéans un temps t " pour la $k^{\text{ième}}$ réalisation. Ce dernier vaut 1 si la particule positionnée en un point $\underline{x} = \underline{x}_0$ atteint le puits en un temps $T \leq t$ et zéro dans le cas contraire. La distribution spatiale définit donc la probabilité $p(CAP(\underline{x}_0, t))^{(**)}$ qu'une particule inerte (conservative) placée en un point $\underline{x} = \underline{x}_0$ atteigne le puits en un temps spécifié $T = t$.

Les **isocontours** $\Gamma^{(i)}(\underline{x}, t)$ connectent les points, qui ont la même probabilité $p(CAP(\underline{x}, t)) = i$ au sein de la distribution. La **zone d'incertitude ZI** au temps t est définie comme l'aire formée par les points vérifiant la condition $0 < p(CAP(\underline{x}, t)) < 1$, c'est-à-dire pour laquelle $\Gamma^{(0)} < \Gamma < \Gamma^{(1)}$. L'**isocontour médiane** $\Gamma^{(0.5)}(\underline{x}, t)$ de la distribution correspond à la ligne d'égale probabilité $p(CAP(\underline{x}, t)) = 0,5$ et délimite la zone de protection stochastique médiane, définie comme l'aire où $p(CAP(\underline{x}, t)) \geq 0,5$. La figure 5.1 illustre les notions de distribution spatiale et d'isocontour.

(*) Afin de garder les mêmes notations que dans la littérature : *CAP* pour *time-related capture zone* (zone de protection en anglais).

(**) Il est évident que l'intervalle de confiance associé à cette probabilité est fonction du nombre d'isochrones pris en compte dans son calcul. Il est d'autant plus faible que le nombre de réalisations est important. DUNCAN [1986] fournit une abaque permettant de définir l'intervalle de confiance à 95% pour une proportion.

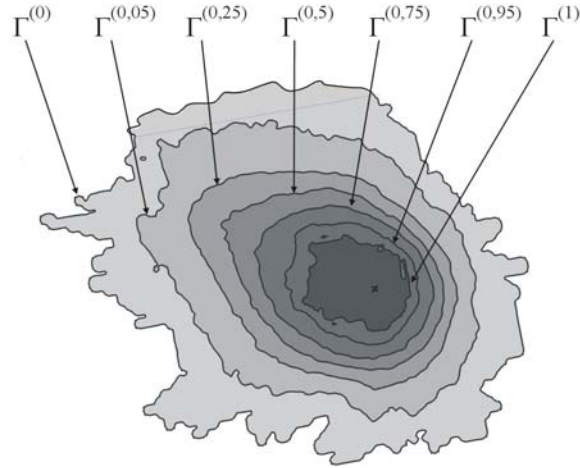


Figure 5.1 : Exemple d'une distribution spatiale d'un ensemble de zones de protection obtenues par l'analyse de Monte Carlo

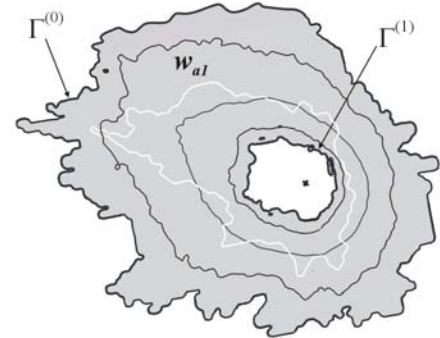
Dans le but de quantifier l'effet du conditionnement sur la distribution spatiale de la zone de protection, des mesures de performance ont été adoptées [LEEuwEN & al., 2000] :

- w_{a1} , illustré à la figure ci-contre, est défini comme la mesure de l'incertitude basée sur l'extension spatiale de la zone d'incertitude ZI au temps t :

$$w_{a1} = \int_{\underline{x}} f_{a1}(\underline{x}, t) d\underline{x} \quad (5.2a)$$

avec

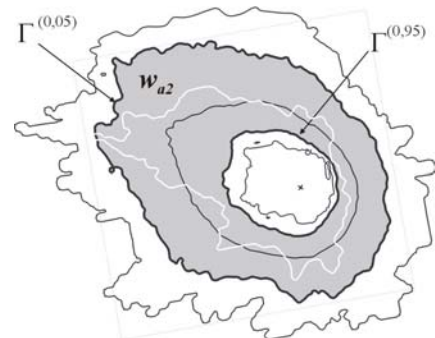
$$f_{a1}(\underline{x}, t) = \begin{cases} 1 & \text{si } 0 < p(CAP(\underline{x}, t)) < 1, \text{ c\`ad si } \Gamma^{(0)} < \Gamma < \Gamma^{(1)} \\ 0 & \text{sinon} \end{cases} \quad (5.2b)$$



Ce paramètre w_{a1} diminue avec la réduction de l'extension de la ZI.

- w_{a2} est défini comme la mesure de l'incertitude basée sur l'extension spatiale de la zone comprise entre $\Gamma^{(0,05)}$ et $\Gamma^{(0,95)}$ au temps t :

$$w_{a2} = \int_{\underline{x}} f_{a2}(\underline{x}, t) d\underline{x} \quad (5.3a)$$



avec

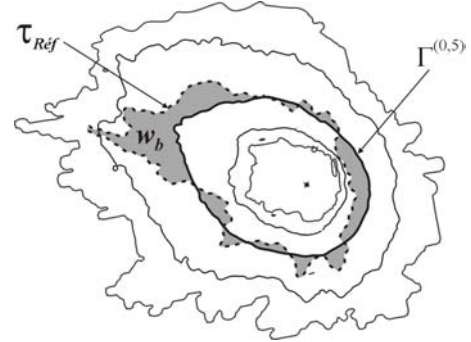
$$f_{a2}(\underline{x}, t) = \begin{cases} 1 & \text{si } 0,05 < p(CAP(\underline{x}, t)) < 0,95 \\ 0 & \text{sinon} \end{cases} \quad (5.3b)$$

- w_b correspond à la mesure de l'écart entre la position de l'isocontour médiane de la distribution $\Gamma^{(0,5)}(\underline{x}, t)$ et la position de l'isochrone de référence $\tau_{Réf}(\underline{x}, t)$:

$$w_b = \int_{\underline{x}} f_b(\underline{x}, t) d\underline{x} \quad (5.4a)$$

avec

$$f_b(\underline{x}, t) = \begin{cases} 1 & \text{si } (I(\underline{x}, t))_{Réf} = 0 \ \& \ p(CAP(\underline{x}, t)) > 0,5 \\ 1 & \text{si } (I(\underline{x}, t))_{Réf} = 1 \ \& \ p(CAP(\underline{x}, t)) < 0,5 \\ 0 & \text{sinon} \end{cases} \quad (5.4b)$$



où $(I(\underline{x}, t))_{Réf}$ est l'événement "particule capturée par le puits endéans un temps t " pour l'isochrone de référence $\tau_{Réf}(\underline{x}, t)$; il vaut 1 si la particule positionnée en un point $\underline{x} = \underline{x}_0$ atteint le puits en un temps $T \leq t$ et zéro dans le cas contraire. Lorsque la position de l'isocontour médiane se rapproche de celle de l'isochrone de référence, la valeur de w_b diminue. Elle vaut zéro quand ils coïncident.

5.2 Paramétrisation et techniques de seuillage

Le conditionnement des champs de conductivité hydraulique par les observations piézométriques est réalisé dans la MSA en calibrant les modèles d'écoulement correspondants sur base des observations. La calibration nécessite de paramétrer les champs de conductivité hydraulique générés (c'est-à-dire de réduire le nombre de paramètres ajustables du modèle). La technique de paramétrisation par zonation utilisée dans ce travail et illustrée à la figure 5.2 consiste, sur base de "seuils" prédéfinis (S_i), à découper l'intervalle de variation de la conductivité hydraulique en une série de classes (C_i) de valeur uniforme (K_{C_i}) représentant les paramètres ajustables du modèle et qui, par définition, respectent l'ordre relatif suivant $K_{C_i} < K_{C_{(i+1)}}$. La définition de chaque classe, et par conséquent leur répartition géométrique dans l'espace (zonation), sont déterminées par le choix de ces seuils.

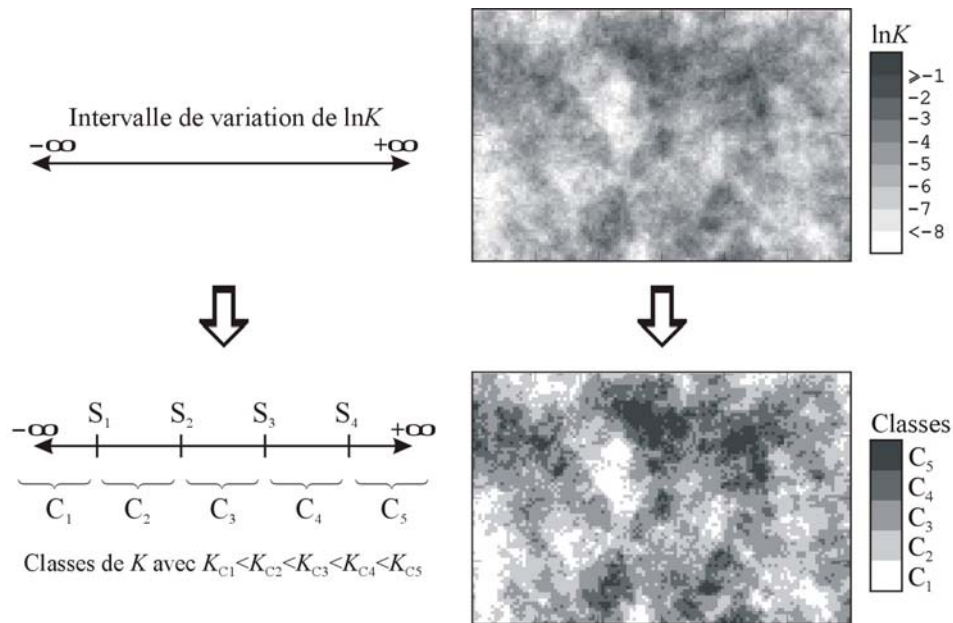


Figure 5.2 : Paramétrisation d'un champ de $\ln K$ (dans le cas où l'intervalle de variation de la conductivité hydraulique est divisé en cinq classes)

Cette paramétrisation nécessite la connaissance des seuils S_i . Trois techniques de seuillage (c'est-à-dire de détermination des seuils) ont été proposées dans le cadre de ce travail. Les deux premières dérivent de la méthode des gaussiennes seuillées (cf. chapitre 1), mais en diffèrent par le choix des seuils. Celui-ci n'est pas contrôlé par les proportions de chaque type de faciès (puisque aucun faciès n'est défini), mais par la proportion des données contenues dans chaque classe. La troisième technique teste différentes combinaisons de valeurs de seuils afin de déterminer celle qui minimise la variabilité des données regroupées par classe. Pour ces trois techniques de seuillage, la définition des seuils dépend du jeu de données de conductivité hydraulique utilisé.

D'une manière générale pour ces techniques de seuillage, la valeur de K_{C_i} attribuée à chaque classe correspond à la moyenne des données au sein de chacune de ces classes. Ces valeurs K_{C_i} sont des valeurs initiales qui seront optimisées lors de la calibration du modèle d'écoulement.

La paramétrisation permet donc de regrouper l'ensemble des données de conductivité hydraulique de même ordre de grandeur dans une même classe de valeur K_{C_i} uniforme. La conductivité hydraulique aux points de données ne sera donc plus égale à la valeur réellement mesurée mais à une valeur approchée de même ordre de grandeur. Ceci peut se justifier par le fait que les valeurs de conductivité hydraulique découlent de l'interprétation des essais de pompage et de l'estimation de l'épaisseur de l'aquifère. Une variation autour de ces valeurs peut dès lors être tolérée (dans des limites acceptables bien entendu).

Première technique de seuillage, notée TS_{unif} :

Cette technique de seuillage nécessite l'établissement de la fonction de distribution des données pour définir les seuils. Les seuils sont choisis de manière telle que le nombre de données dans chaque classe est identique. Ils correspondent donc aux percentiles de la fonction de distribution des données, définissant des classes de proportion identique (suivant une loi de distribution uniforme). La figure 5.3 illustre cette technique de seuillage dans le cas où l'intervalle de variation de la propriété est divisé en cinq classes : les quatre seuils correspondants sont alors associés aux valeurs des percentiles p20, p40, p60 et p80.

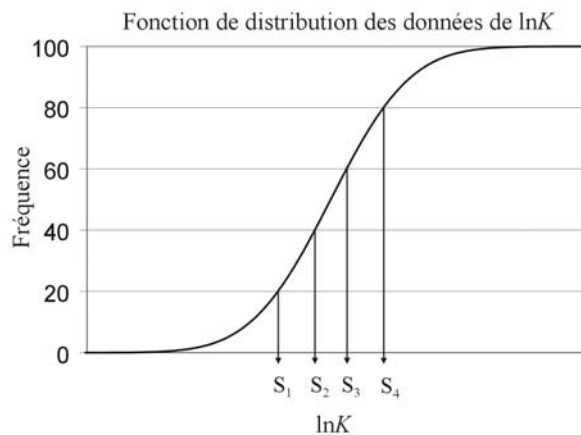


Figure 5.3 : Détermination des seuils S_i par la technique TS_{unif}

Deuxième technique de seuillage, notée TS_{gauss} :

Cette technique est similaire à la première, si ce n'est que la proportion des données dans chaque classe suit une loi de distribution gaussienne (comme $\ln K$) et non plus uniforme. La figure 5.4 illustre cette technique de seuillage pour cinq classes : les seuils sont dès lors définis par les valeurs des percentiles p8, p28, p72 et p92.

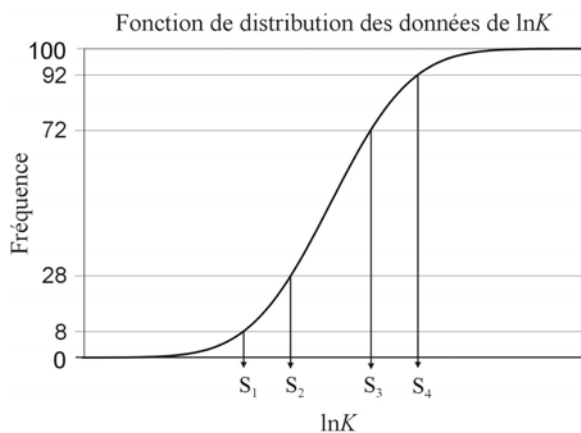


Figure 5.4 : Détermination des seuils S_i par la technique TS_{gauss}

Troisième technique de seuillage, notée TS_{fmin} :

Cette dernière technique suit une philosophie différente des deux autres. Elle part du principe que toutes les données de $\ln K$ rassemblées dans une même classe se verront, par paramétrisation, attribuer une même valeur (moyenne) de conductivité hydraulique. Afin de rester le plus cohérent possible, il paraît logique de regrouper dans une même classe les données dont les valeurs sont du même ordre de grandeur. Il est dès lors intéressant de trouver la combinaison des données pour laquelle la variabilité de l'ensemble des classes est minimisée. Cette technique nécessite de trier les données de $\ln K$ par ordre croissant. Pour toutes les combinaisons possibles de ces données réparties en un nombre N_c de classes, avec N_{d_i} le nombre de données de $\ln K$ dans chacune des classes (variant d'une combinaison à l'autre), une fonction f définie par l'équation (5.5a) est calculée.

$$f = \sum_{i=1}^{N_c} \sum_{j=1}^{N_{d_i}} (\ln K_{ij} - \overline{\ln K_i})^2 \quad (5.5a)$$

avec

$$\overline{\ln K_i} = \frac{1}{N_{d_i}} \sum_{j=1}^{N_{d_i}} \ln K_{ij} \quad , i = 1, N_c \quad (5.5b)$$

où $\sum_{i=1}^{N_c} N_{d_i} = N_D$ est le nombre total de données de $\ln K$.

Les seuils choisis sont ceux qui définissent les classes pour lesquelles la combinaison des données minimise cette fonction f . Cette technique revient à définir des classes à l'intérieur desquelles la variabilité des données est minimale. La figure 5.5 illustre cette technique de seuillage. L'algorithme permettant de déterminer la combinaison des données qui minimise la fonction f est illustré à l'annexe D. Cette technique TS_{fmin} ne dépend donc pas de la fréquence des données comme c'est le cas pour TS_{unif} et TS_{gauss} , mais de leur variabilité.

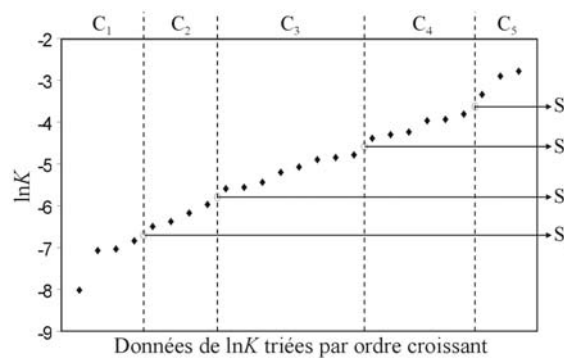


Figure 5.5 : Détermination des seuils S_i par la technique TS_{fmin}

5.3 Application de la MSA au cas d'étude synthétique

Dans les applications pratiques, une des grandes difficultés réside dans la construction du modèle conceptuel qui ne représente jamais parfaitement la réalité. Le modèle conceptuel utilisé pour l'application de la MSA au cas d'étude synthétique est le même que celui présenté au chapitre 4 et ce dans le but de ne pas prendre en compte l'influence sur les résultats des erreurs commises lors de la construction du modèle conceptuel.

Pour rappel, la grille de calcul, constituée de trois grilles réduites de taille identique, a une dimension de 1800 mètres de long sur 400 mètres de large. La partie centrale, plus raffinée, comprend des cellules carrées de 5 mètres de côté tandis que les deux extensions latérales sont composées de cellules rectangulaires de 30x5 mètres. Le modèle est constitué de deux couches surmontant le bedrock imperméable :

- la première correspond aux sables fins et limons de surface et est caractérisée par une conductivité hydraulique uniforme de 10^{-5} m/s ;
- la deuxième représente la couche graveleuse. Son champ de conductivité hydraulique, considéré comme la propriété incertaine, est divisé en trois selon la discrétisation de la grille de calcul (cf. figure 4.11) : une partie centrale constituant le site d'étude proprement dit, dont le champ hétérogène correspond aux simulations stochastiques générées (il varie donc d'une réalisation à l'autre) et deux extensions latérales situées de part et d'autre, dont la conductivité hydraulique est uniforme et de valeur constante égale à la moyenne géométrique des données de conductivité hydraulique.

Les conditions aux limites sont identiques sur les deux couches. Supposant un équilibre entre la nappe et le cours d'eau, un potentiel a été imposé sur cette frontière. Un potentiel a également été imposé sur les deux frontières latérales, déterminant ainsi le gradient de la nappe. Le long de la frontière représentant la limite de la plaine alluviale, on a considéré une alimentation en provenance du versant correspondant à un débit de $0,0064$ m³/s. Une recharge de 250 mm/an a été appliquée sur la couche supérieure du modèle. Enfin, une porosité efficace uniforme de 5% a été imposée sur l'entièreté de la couche graveleuse.

Les jeux de données conditionnantes créés au chapitre 4 à partir du champ de conductivité hydraulique de référence d'un cas d'étude synthétique consistent en des valeurs de conductivité hydraulique K , de hauteur piézométrique h et de résistivité électrique ρ . La MSA a été appliquée à ce cas d'étude synthétique et ce pour différents niveaux de connaissance, c'est-à-dire en combinant divers types de données conditionnantes dans le but d'évaluer l'effet de chacun d'entre eux individuellement. Cette méthode se base sur le fait que des données de hauteur piézométrique sont toujours disponibles, ce qui est généralement le cas sur les sites de captages, permettant alors de calibrer le modèle d'écoulement. Dans le cas

contraire, la méthode s'apparente à celle utilisée par VAN LEEUWEN [2000] (cf. chapitre 3) pour laquelle aucune calibration du modèle d'écoulement n'est réalisée.

Deux situations différentes ont été distinguées :

- **Conditionnement des champs de conductivité hydraulique par K et h ou K , ρ et h :** les données de conductivité hydraulique (variable principale), de hauteur piézométrique et de résistivité électrique conditionnent les réalisations engendrées ;
- **Conditionnement des champs de conductivité hydraulique par ρ et h :** cette situation correspond au cas particulier où le nombre de données de conductivité hydraulique est trop faible, voire nul, de sorte que ses propriétés statistiques ne peuvent être inférées de l'analyse géostatistique. Seules les observations piézométriques et les valeurs de résistivité électrique conditionnent les réalisations.

Lors de la construction du cas d'étude synthétique, deux jeux de données ont été définis pour chacun des types considérés (cf. paragraphe 4.3). Nous les noterons dans la suite de ce travail de la manière suivante :

- K_{15} , le jeu de 15 données de conductivité hydraulique ;
- K_{25} , le jeu de 25 données de conductivité hydraulique ;
- h_{15} , le jeu de 15 données de hauteur piézométrique ;
- h_{25} , le jeu de 25 données de hauteur piézométrique ;
- ρ_{150} , le jeu de 150 données de résistivité électrique ;
- ρ_{300} , le jeu de 300 données de résistivité électrique.

Notons que chacun des jeux de grande taille comprend toujours le jeu de petite taille et ce de manière à étudier l'influence du nombre de données sur les résultats en tentant d'éviter l'effet dû à la variation de données.

5.4 Conditionnement par K et h ou K , ρ et h

On considère trois jeux de données : un jeu de conductivités hydrauliques (récolté par exemple par essais de pompage aux puits), un jeu de résistivités électriques (récolté par tomographie électrique) et un jeu de hauteurs piézométriques (récolté par simple prise de mesure du niveau d'eau aux puits et piézomètres). Le schéma de la méthode stochastique alternative intégrant ces différents types de données est repris à la figure 5.6.

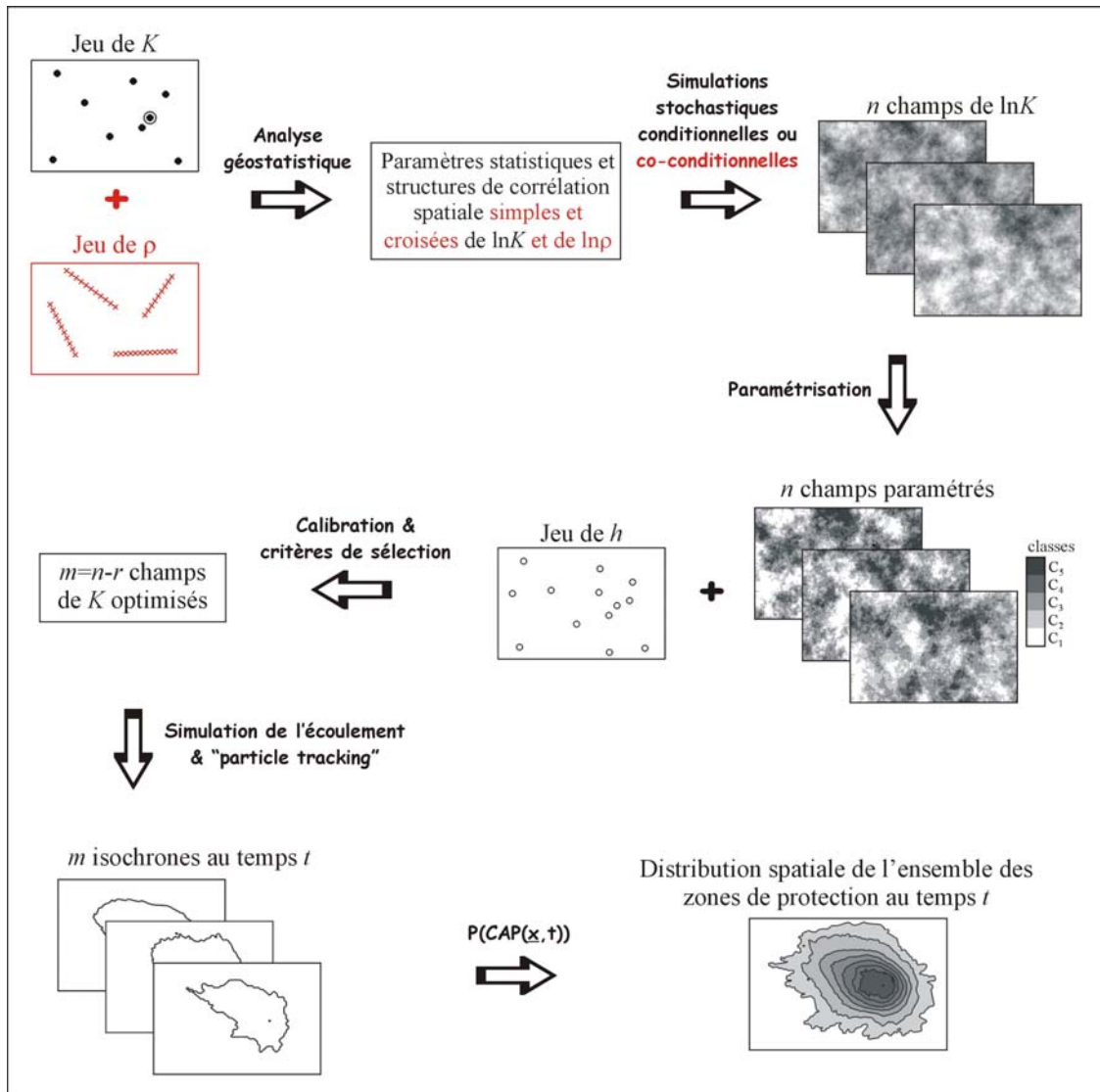


Figure 5.6 : MSA dans la situation où des données de K , de ρ et de h sont disponibles

5.4.1 Analyse géostatistique et génération des simulations stochastiques

Dans un premier temps, l'analyse exploratoire des données a été réalisée à l'aide du logiciel ISATIS. La figure 5.7 positionne les différents jeux de données sur la zone d'étude proprement dite. Contrairement aux jeux de K et de h qui sont isotopiques (valeurs prélevées aux mêmes points de mesure), les jeux de K et de ρ , et donc les jeux de h et de ρ , sont hétérotopiques (valeurs prélevées en des points de mesure différents).

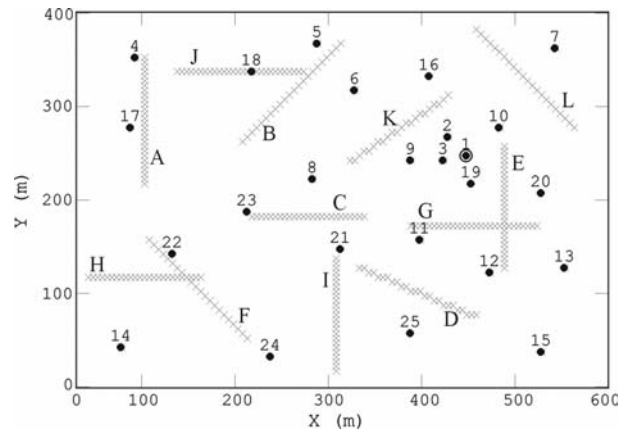


Figure 5.7 : Position des points de mesure pour les jeux de données K_{15} et h_{15} (constitué des points 1 à 15), K_{25} et h_{25} (points 1 à 25), ρ_{150} (profils A à F) et ρ_{300} (profils A à L)

Les histogrammes du logarithme népérien des données de K et de ρ , notés $\ln K_{15}$, $\ln K_{25}$, $\ln \rho_{150}$ et $\ln \rho_{300}$, présentés à la figure 5.8, mettent en évidence le caractère gaussien de ces distributions, ce qui est en accord pour la conductivité hydraulique avec les conclusions de nombreux auteurs tels que FREEZE [1975] et HOEKSEMA & KITANIDIS [1985]. En ce qui concerne la résistivité électrique, une corrélation existant entre $\ln K$ et $\ln \rho$ a été mise en évidence dans la plaine alluviale de la Meuse et utilisée pour construire les jeux de ρ du cas d'étude synthétique (cf. paragraphe 4.3.3). Nous travaillerons donc dans la suite avec $\ln K$ et $\ln \rho$ en faisant l'hypothèse que ces deux variables suivent une loi normale. Le tableau 5.1 présente les résultats statistiques obtenus sur les jeux de $\ln K$ et de $\ln \rho$.

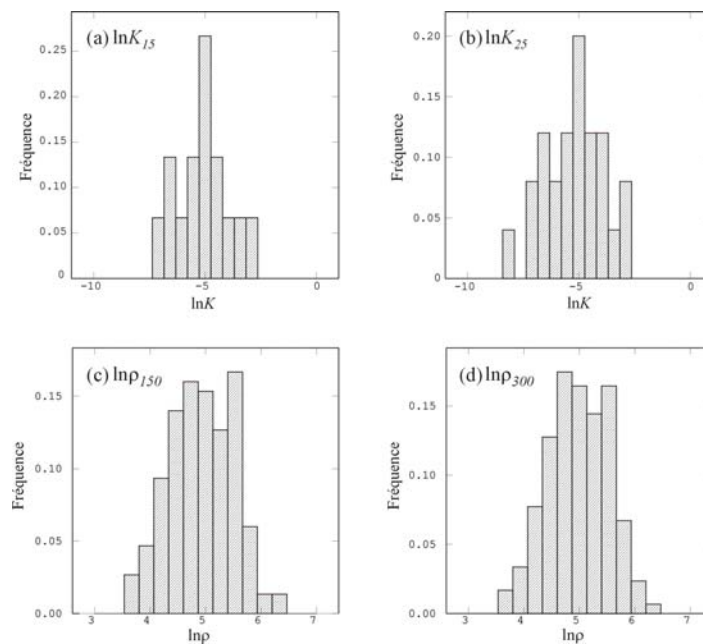


Figure 5.8: Histogrammes de (a) $\ln K_{15}$, (b) $\ln K_{25}$, (c) $\ln \rho_{150}$ et (d) $\ln \rho_{300}$

	<i>Minimum</i>	<i>Maximum</i>	<i>Moyenne</i> m_A	<i>Ecart-type</i> σ	<i>Variance</i> σ^2	<i>Coeff. Variat.</i> $ \sigma/m_A $
$\ln K_{15}$	-7,04	-2,90	-5,05	1,19	1,42	0,24
$\ln K_{25}$	-8,02	-2,78	-5,16	1,34	1,79	0,26
$\ln \rho_{150}$	3,69	6,42	4,93	0,57	0,33	0,12
$\ln \rho_{300}$	3,69	6,42	4,98	0,55	0,30	0,11

Tableau 5.1 : Statistiques sur les données de $\ln K$ et de $\ln \rho$ du cas d'étude synthétique

En raison du faible nombre de données de K (15 ou 25), les statistiques calculées sur les deux jeux diffèrent. Par contre, les variations entre les deux jeux de ρ sont extrêmement faibles.

Afin d'étudier l'effet de l'apport de données de prospection géophysique sur les résultats obtenus, deux cas seront distingués:

- cas A, conditionnement des simulations stochastiques par krigeage à partir d'un jeu de K (nécessitant l'inférence de la structure spatiale de $\ln K$).
- cas B, conditionnement des simulations stochastiques par cokrigeage à partir d'un jeu de K et d'un jeu de ρ (nécessitant l'inférence des structures spatiales simples et croisées de $\ln K$ et $\ln \rho$).

Cas A

Dans ce premier cas, les champs de $\ln K$ ont été générés uniquement sur base des données de K : des simulations stochastiques non-conditionnelles du champ de $\ln K$ ont d'abord été engendrées par la méthode de bandes tournantes, puis conditionnées à l'aide d'une étape supplémentaire de double krigeage (décrite au chapitre 1). La génération de ces simulations stochastiques a nécessité la modélisation des variogrammes expérimentaux de $\ln K_{15}$ et $\ln K_{25}$, présentés à la figure 5.9, afin de passer d'une fonction définie par des points à une fonction continue et possédant une expression mathématique. Plusieurs modèles théoriques ont été dans un premier temps ajustés de manière visuelle sur les variogrammes expérimentaux. Une validation croisée permettant de comparer l'impact des différents modèles sur les résultats de l'estimation a ensuite été réalisée afin de choisir, parmi ces modèles, celui présentant la meilleure adéquation avec les variogrammes expérimentaux. Un modèle de variogramme a été choisi pour chacun des deux jeux de données. Ces modèles possèdent tous deux une structure exponentielle caractérisée par une portée a de 150 mètres (correspondant à une échelle intégrale λ de 50 mètres) et un palier S égal à 1,6 pour K_{15} et 1,85 pour K_{25} , que l'on peut noter respectivement $\gamma_{\ln K_{15}}(h) = 1,6[1 - \exp(-h/50)]$ et $\gamma_{\ln K_{25}}(h) = 1,85[1 - \exp(-h/50)]$. Aucun effet de pépite n'a dû être introduit pour modéliser les variogrammes expérimentaux.

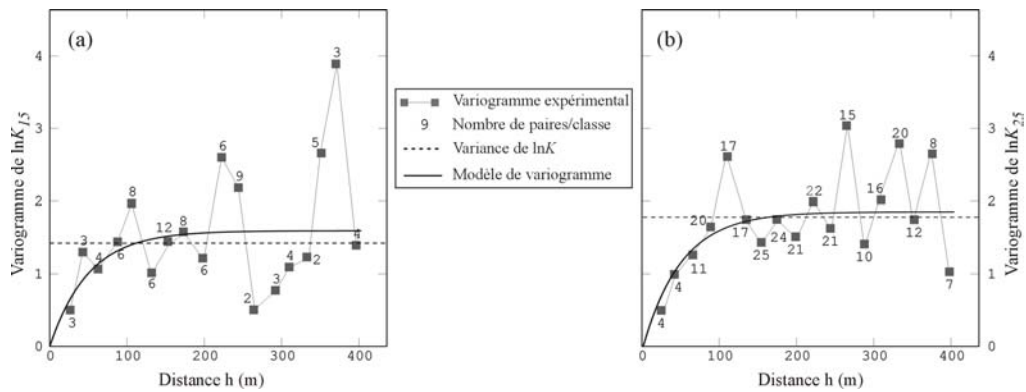


Figure 5.9 : Variogramme (a) de $\ln K_{15}$ et (b) de $\ln K_{25}$

Un modèle identique à celui utilisé pour générer le champ de conductivité hydraulique de référence aurait pu être choisi pour les deux jeux afin de pouvoir comparer les résultats. Cependant, il nous semblait plus judicieux de travailler dans des conditions réelles de terrain en analysant l'effet de l'apport des données de conductivité hydraulique non seulement par le conditionnement des simulations stochastiques mais également par la détermination des paramètres du modèle de variogramme.

A partir des paramètres de leur modèle de variogramme respectif, deux cents simulations stochastiques conditionnelles de $\ln K$ ont été générées pour chacun des deux jeux de données. Les simulations ainsi obtenues possèdent les mêmes propriétés statistiques spatiales que le jeu de données considéré et vérifient ces données aux points de mesure ; elles sont toutes équiprobables. Trois réalisations parmi les deux cents engendrées sont présentées à la figure 5.10 et ce pour les deux jeux de données conditionnantes. La comparaison des réalisations conditionnées par K_{15} et par K_{25} , permet de visualiser l'effet du conditionnement par les dix données supplémentaires. En effet, on remarque une modification du champ de conductivité hydraulique au voisinage de ces points de mesure supplémentaires.

Cas B

Le conditionnement des champs de $\ln K$ par les données de K et de ρ a été réalisé par génération de simulations stochastiques **co-conditionnelles** (sans qu'aucune corrélation existant entre ces deux variables ne soit établie). Pour ce faire, des simulations stochastiques non-conditionnelles ont d'abord été générées par la méthode des bandes tournantes, comme pour le cas A, puis un double **cokrigage** a été réalisée afin de les conditionner (cf. paragraphe 1.3.2.3). Par définition du cokrigage, cette étape supplémentaire nécessite l'inférence des structures spatiales simples et croisées de $\ln K$ et de $\ln \rho$. Puisque ces deux jeux de données sont complètement hétérotopiques, le variogramme croisé $\gamma_{\ln K, \ln \rho}(h)$ défini au

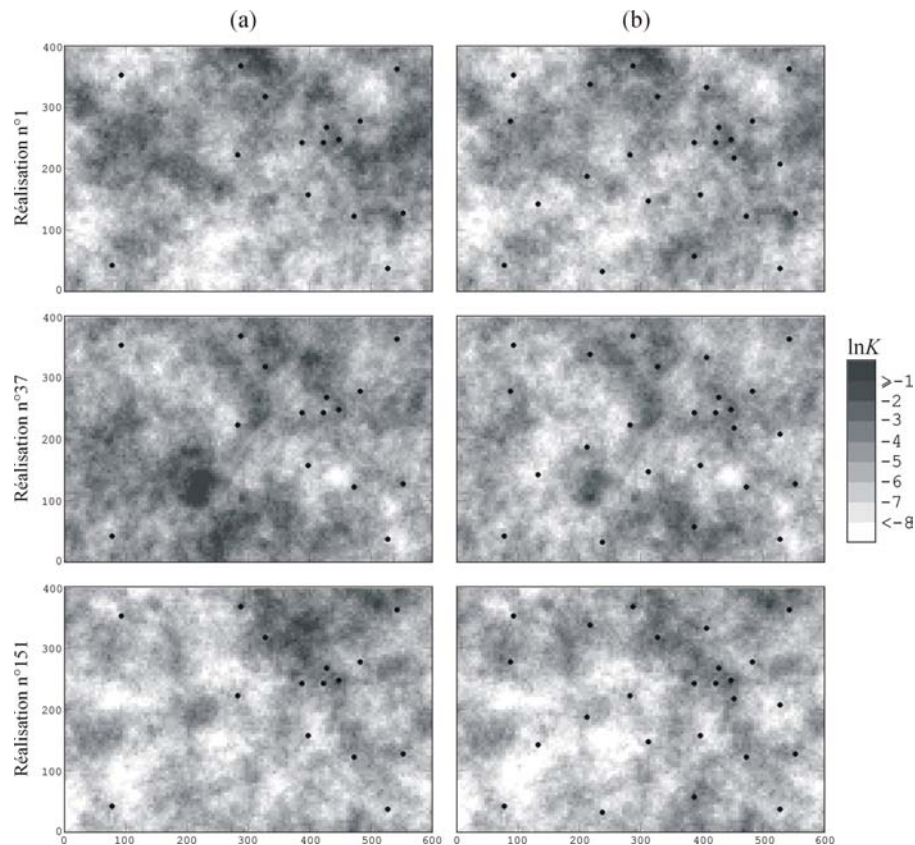


Figure 5.10 : Simulations stochastiques de $\ln K$ conditionnées (a) par le jeu K_{15} ou (b) par le jeu K_{25}

chapitre 1, équation (1.43), ne peut pas être déterminé. Par contre, la covariance croisée, dont l'équation (1.44) ne fait pas intervenir de données de type différent aux mêmes points de mesure, peut être calculée en considérant alors la stationnarité du deuxième ordre^(*).

De manière générale, l'ajustement des covariances (ou variogrammes) simples et croisées doit être mathématiquement cohérent de façon à garantir des variances positives. Le modèle le plus couramment employé permettant de passer d'une fonction définie par des points à une fonction continue est le **modèle linéaire de corégionalisation** [WACKERNAGEL, 1995 ; GOOVAERTS, 1997]. Dans un modèle linéaire de corégionalisation, les variables se décomposent linéairement en composantes élémentaires dont les structures (covariances ou variogrammes) simples et croisées sont des combinaisons linéaires de schémas élémentaires [DE MARSILY, 1986]. Ce modèle linéaire de corégionalisation est donc défini comme l'ensemble des $N_v \times N_v$ (N_v étant le nombre de variables) modèles de covariances (ou variogrammes) simples $C_{ii}(\underline{h})$ et croisées $C_{ij}(\underline{h})$, et s'écrit :

^(*) Pour rappel, une fonction aléatoire sera dite stationnaire d'ordre deux si en tout point elle admet une espérance indépendante du point d'appui, et si pour tout vecteur des distances \underline{h} , la covariance existe et ne dépend pas du point d'appui. La structure spatiale d'un processus stationnaire est donc conservée sur tout le domaine.

$$\underline{\underline{C}}(\underline{h}) = \sum_{l=0}^L \underline{B}_l g_l(\underline{h}) \quad (5.6)$$

ou

$$C_{ij}(\underline{h}) = \sum_{l=0}^L b_{ij}^l g_l(\underline{h}) \quad \forall i, j \quad (5.7)$$

où chaque fonction $g_l(\underline{h})$ est un modèle de covariance (ou de variogramme) admis et les matrices \underline{B}_l sont les matrices de corégionalisation dont les coefficients b_{ij}^l correspondent au palier ou à la pente du modèle $g_l(\underline{h})$. Dans le cas où $N_v=2$, le modèle linéaire de corégionalisation s'écrit :

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} = \sum_{l=0}^L \begin{bmatrix} b_{11}^l & b_{12}^l \\ b_{21}^l & b_{22}^l \end{bmatrix} g_l(\underline{h}) \quad (5.8)$$

Pour pouvoir accepter le modèle linéaire de corégionalisation, il suffit de vérifier que les matrices de corégionalisation sont toutes définies positives. Une matrice symétrique est définie positive si son déterminant et tous ses mineurs principaux sont positifs ou nuls. Chaque matrice de corégionalisation est donc définie positive si l'ensemble des inégalités suivantes sont satisfaites :

- tous les éléments diagonaux doivent être positifs ou nuls : $b_{11}^l \geq 0$ et $b_{22}^l \geq 0$
- le déterminant d'ordre 2 doit être positif ou nul: $\begin{vmatrix} b_{11}^l & b_{12}^l \\ b_{21}^l & b_{22}^l \end{vmatrix} = b_{11}^l b_{22}^l - b_{12}^l b_{21}^l \geq 0$

Une fois ajusté, un tel modèle permet le cokrigage des variables.

Un modèle linéaire de corégionalisation a été ajusté aux covariances expérimentales simples et croisées de $\ln K$ et $\ln \rho$ pour toutes les combinaisons des jeux de données. La figure 5.11 présente les covariances expérimentales simples et croisée pour la combinaison (K_{25}, ρ_{300}) ainsi que le modèle de corégionalisation ajusté.

Pour l'ensemble des combinaisons des jeux de données, le modèle linéaire de corégionalisation est isotrope et s'écrit :

$$\begin{bmatrix} C_{\ln K, \ln K}(h) & C_{\ln K, \ln \rho}(h) \\ C_{\ln \rho, \ln K}(h) & C_{\ln \rho, \ln \rho}(h) \end{bmatrix} = \underline{\underline{B}}_0 g_0(h) + \underline{\underline{B}}_1 \exp(-h/50) \quad (5.9)$$

où $g_0(h)$ est le modèle d'effet de pépite, n'existant que pour la covariance simple de $\ln \rho$, et $\exp(-h/50) = g_1(h)$, un modèle de structure exponentielle et de longueur de corrélation $\lambda = 50$ mètres (correspondant à une portée de 150 mètres).

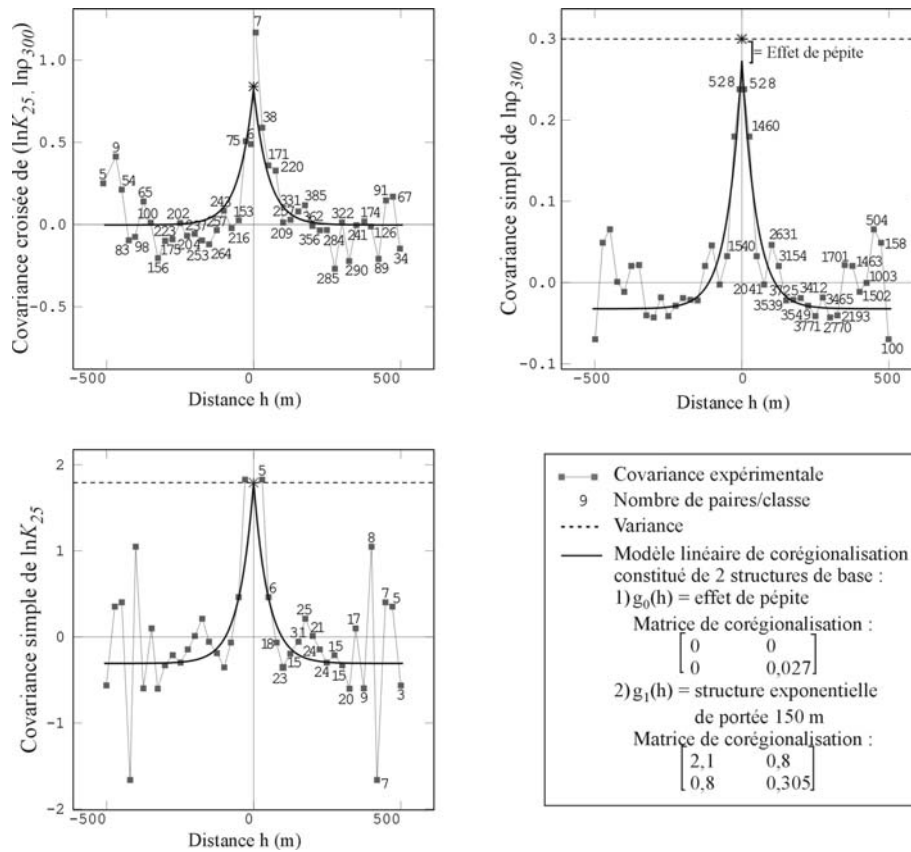


Figure 5.11 : Modèle linéaire de corégionalisation ajusté aux covariances expérimentales simples et croisées de $\ln K$ et $\ln \rho$ (pour K_{25} et ρ_{300})

Le tableau 5.2 reprend les éléments et les déterminants des matrices de corégionalisation \underline{B}_0 et \underline{B}_1 pour chacune des combinaisons de données.

	b_{11}^0	$b_{12}^0 = b_{21}^0$	b_{22}^0	$Dét(\underline{B}_0)$	b_{11}^1	$b_{12}^1 = b_{21}^1$	b_{22}^1	$Dét(\underline{B}_1)$
K_{15}, ρ_{150}	0	0	0,027	0	1,8	0,75	0,33	0,0315
K_{15}, ρ_{300}	0	0	0,027	0	1,8	0,7	0,31	0,068
K_{25}, ρ_{150}	0	0	0,027	0	2,1	0,8	0,33	0,053
K_{25}, ρ_{300}	0	0	0,027	0	2,1	0,8	0,31	0,011

Tableau 5.2 : Eléments et déterminants des matrices de corégionalisation \underline{B}_0 et \underline{B}_1 pour chacune des combinaisons de données (avec $1=\ln K$ et $2=\ln \rho$)

Ce tableau montre que toutes les matrices de corégionalisation \underline{B}_0 et \underline{B}_1 sont bien définies positives. En effet, leurs éléments diagonaux sont tous positifs ou nuls ainsi que leur déterminant respectif.

Deux cents simulations stochastiques co-conditionnelles du champ de $\ln K$ ont dès lors été générées à partir des paramètres de ce modèle linéaire de corégionalisation pour chacune des quatre combinaisons des jeux de données : $(\ln K_{15}, \ln \rho_{150})$, $(\ln K_{15}, \ln \rho_{300})$, $(\ln K_{25}, \ln \rho_{150})$ et

($\ln K_{25}$, $\ln \rho_{300}$). Trois réalisations de $\ln K$ parmi les deux cents engendrées sont représentées à la figure 5.12.

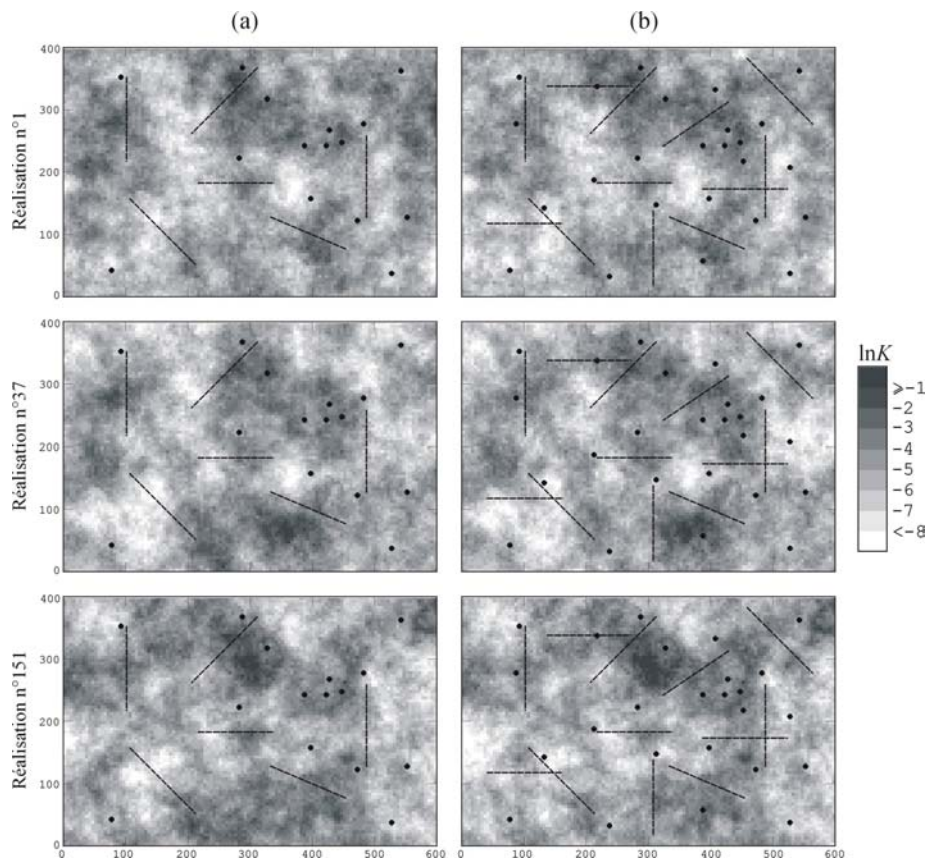


Figure 5.12 : Simulations stochastiques de $\ln K$ conditionnées
 (a) par les jeux K_{15} et ρ_{150} , (b) par les jeux K_{25} et ρ_{300}

5.4.2 Paramétrisation

L'étape suivante de la MSA a consisté à paramétrer tous les champs de $\ln K$ obtenus afin de pouvoir calibrer les modèles d'écoulement par les observations piézométriques. Pour rendre un problème inverse "mieux posé" (cf. chapitre 2) et obtenir une solution valide, le nombre de paramètres ajustables doit être inférieur au nombre d'observations piézométriques. De cette manière, chacune des zones contient suffisamment d'observations pour permettre l'optimisation des paramètres associés. Cinq paramètres ajustables ont dès lors été choisis pour le cas d'étude synthétique. Ils ont été définis en divisant l'intervalle de variation des données de K en cinq classes de valeur uniforme dont les valeurs initiales (avant optimisation) correspondent à la moyenne des données dans chaque classe. Les valeurs des quatre seuils correspondants ont été déterminées pour les jeux K_{15} et K_{25} par la technique de seuillage TS_{fmin} décrite précédemment (cf. paragraphe 5.2). Afin d'améliorer la compréhension, le choix de cette technique ne sera justifié qu'au paragraphe 5.4.4. Les données de $\ln K$ ont tout

d'abord été triées par ordre croissant, puis la combinaison des données minimisant la fonction f (cf. équations 5.5a et 5.5b) a été calculée à l'aide de l'algorithme illustré à l'annexe D. Les résultats sont présentés à la figure 5.13.

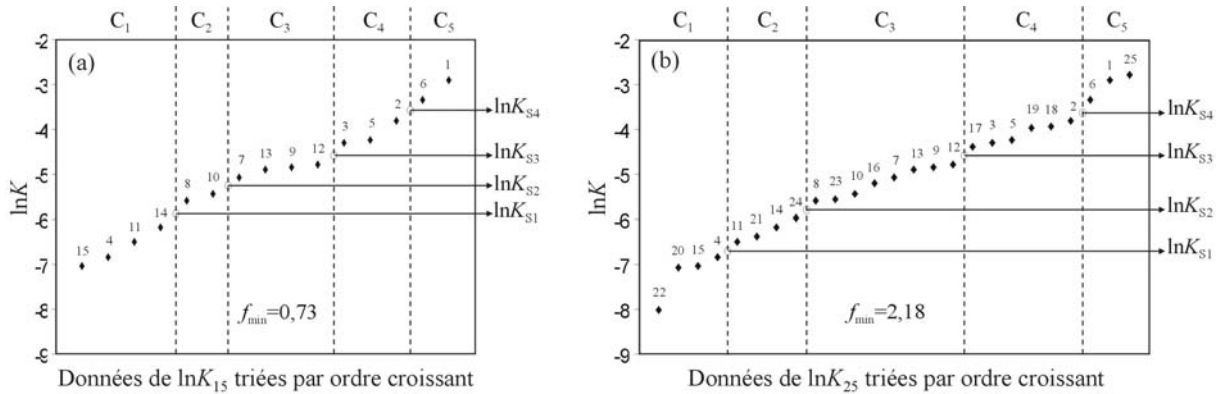


Figure 5.13 : Combinaison des données de $\ln K$ minimisant la fonction f (a) pour K_{15} et (b) pour K_{25}

Ces deux combinaisons fournissent pour chacun des jeux K_{15} et K_{25} des classes à l'intérieur desquelles la variabilité des données de $\ln K$ est minimale. Le tableau 5.3 reprend les quatre valeurs de seuils – $\ln K_{S1}$, $\ln K_{S2}$, $\ln K_{S3}$, $\ln K_{S4}$ – définissant ces classes.

	$\ln K_{S1}$	$\ln K_{S2}$	$\ln K_{S3}$	$\ln K_{S4}$
$\ln K_{15}$	-5,88	-5,26	-4,54	-3,58
$\ln K_{25}$	-6,67	-5,78	-4,59	-3,58

Tableau 5.3 : Valeur des seuils calculés par la technique de seuillage $TS_{f_{\min}}$

Comme le montre ce tableau, l'apport de dix valeurs de $\ln K$ supplémentaires modifie la valeur des seuils, qui dépendent donc du jeu de données.

Sur base de ces seuils, tous les champs de $\ln K$ ont alors été découpés en zones de valeur uniforme représentant les cinq paramètres ajustables du modèle d'écoulement, que l'on peut associer à différents "hydrofaciès" (zones de propriétés hydrodynamiques semblables). Sur les figures 5.14 et 5.15 montrant la paramétrisation des réalisations présentées aux figures 5.10 et 5.12 respectivement, on peut voir que, contrairement à l'étude de VASSOLO & al. [1998], la répartition spatiale de la conductivité hydraulique varie d'une réalisation à l'autre tout en intégrant une certaine variabilité spatiale.

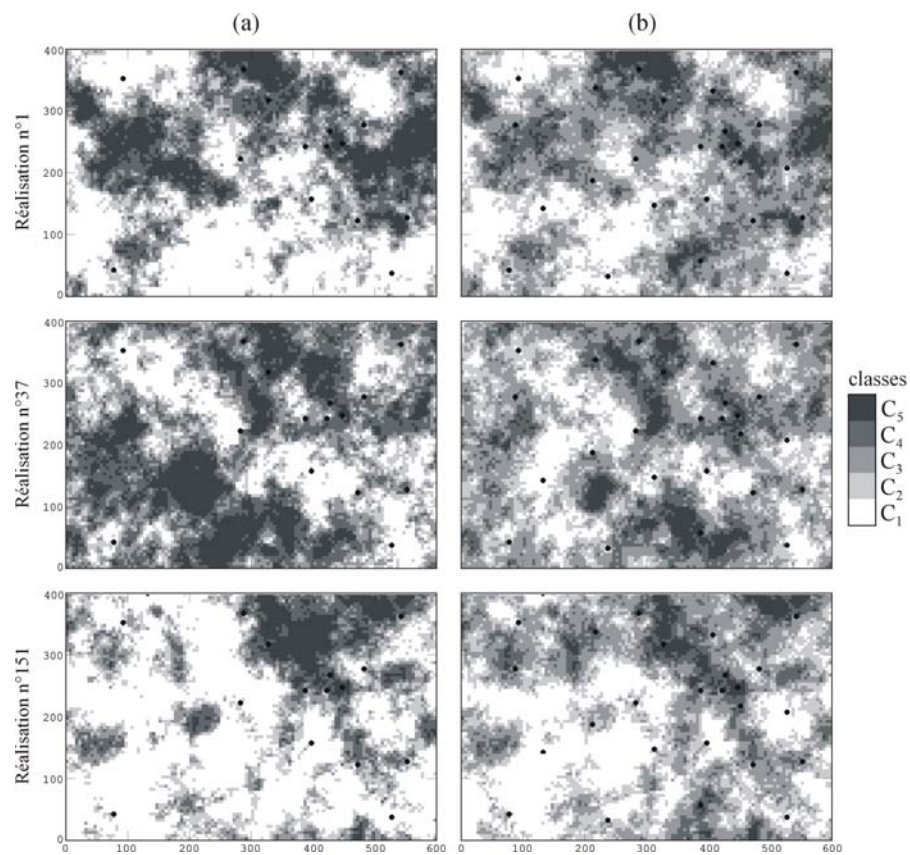


Figure 5.14 : Champs conditionnés et paramétrés (a) par le jeu K_{15} ou (b) par le jeu K_{25}

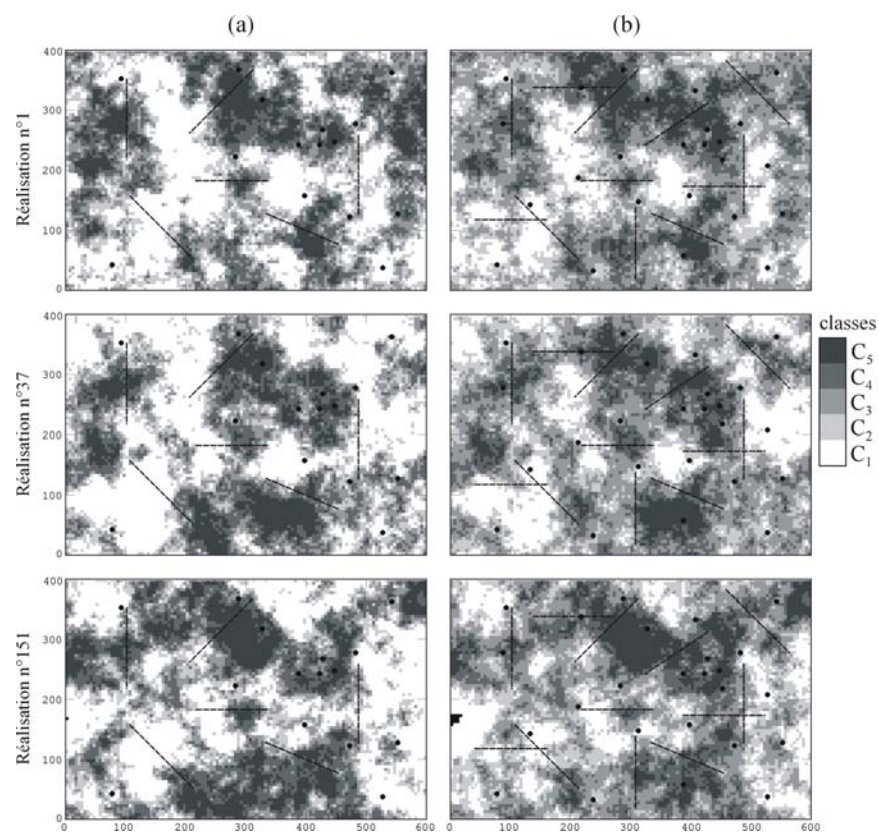


Figure 5.15 : Champs (a) conditionnés par K_{15} et ρ_{150} et paramétrés par K_{15} , (b) conditionnés par K_{25} et ρ_{300} et paramétrés par K_{25}

5.4.3 Calibration, critères de sélection, simulation du modèle d'écoulement et *particle tracking*

Comme mentionné précédemment, chacun des champs de conductivité hydraulique introduit dans le modèle d'écoulement est composé de trois parties :

- une partie centrale constituée du champ paramétré pour lequel une valeur de conductivité hydraulique initiale, qui sera optimisée par calibration du modèle, est attribuée à chaque paramètre ;
- deux extensions latérales de mêmes dimensions pour lesquelles la conductivité hydraulique est constante et uniforme. Rappelons que pour le champ de référence, une valeur de conductivité hydraulique de $4,2 \cdot 10^{-3}$ m/s correspondant à la moyenne géométrique de la conductivité hydraulique de la simulation stochastique non-conditionnelle a été introduite. Afin de se placer dans des conditions réelles de terrain, cette valeur sera égale à $6,4 \cdot 10^{-3}$ m/s pour les champs conditionnés par le jeu K_{15} et $5,7 \cdot 10^{-3}$ m/s pour K_{25} , correspondant à la valeur de la moyenne géométrique du jeu de conductivité hydraulique donné. Il est bon de noter que ces valeurs, différentes de celle du champ de référence, pourront influencer les résultats et qu'il faudra en tenir compte lors des comparaisons.

Pour chacun de ces champs de conductivité hydraulique, la calibration des modèles d'écoulement correspondants a été réalisée de façon automatique sur base des données de hauteur piézométrique, et ce à l'aide du code d'inversion PEST [DOHERTY & *al.*, 1994]. Pour rappel, la technique de calibration automatique du modèle d'écoulement utilisée tente d'optimiser la valeur des paramètres en minimisant la fonction objectif Φ , définie ici comme la somme pondérée du carré des écarts entre les observations piézométriques mesurées h'_i et calculées h_i par le modèle et qui s'écrit :

$$\Phi = \sum_{i=1}^{N_{obs}} [w_i (h'_i - h_i)]^2 \quad (5.10)$$

où N_{obs} est le nombre d'observations piézométriques conditionnantes et w_i est le poids attribué à la $i^{\text{ème}}$ observation.

Puisque les données de hauteur piézométrique proviennent du cas d'étude synthétique, elles ont toutes été "prélevées" de la même manière et sont exemptes d'erreur de mesure. Des poids w_i identiques ont dès lors été utilisés. La valeur de ces poids a été choisie de manière à normer la fonction objectif par le nombre d'observations conditionnantes. L'équation 5.10 s'écrit alors :

$$\Phi = \frac{1}{N_{obs}} \sum_{i=1}^{N_{obs}} [(h_i' - h_i)]^2 \quad (5.11)$$

Les graphiques de la figure 5.16 montrent l'évolution de la moyenne des fonctions objectifs minimisées (après calibration du modèle d'écoulement) en fonction du nombre de réalisations générées pour les différentes combinaisons des données testées – (K_{15}, h_{15}) , $(K_{15}, \rho_{150}, h_{15})$, $(K_{15}, \rho_{300}, h_{15})$, $(K_{15}, h_{25})^{(*)}$, (K_{25}, h_{25}) , $(K_{25}, \rho_{150}, h_{25})$ et $(K_{25}, \rho_{300}, h_{25})$.

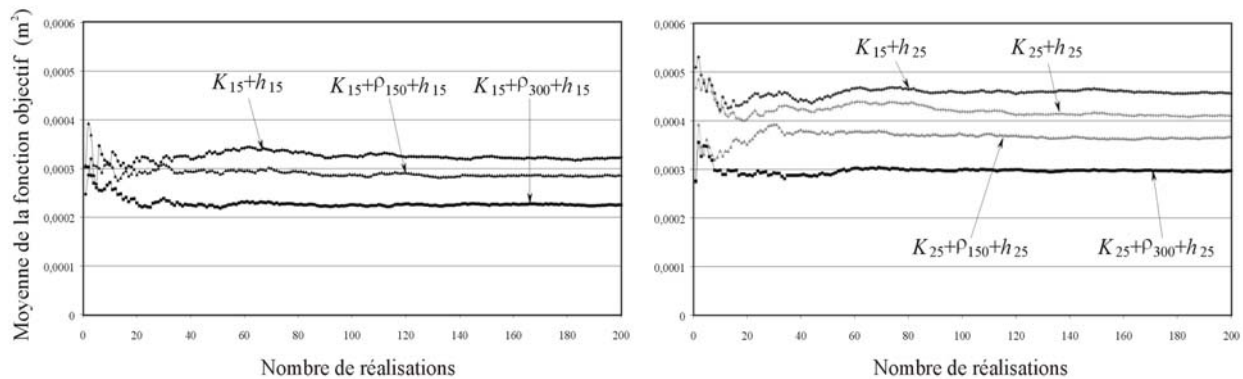


Figure 5.16 : Evolution de la moyenne des fonctions objectifs en fonction du nombre de réalisations générées (a) pour h_{15} et (b) pour h_{25}

Les résultats obtenus avec les deux jeux de hauteurs piézométriques de dimensions différentes (h_{15} et h_{25}) ne peuvent être comparés puisque la fonction objectif a été minimisée sur base d'un nombre différent d'observations. En effet, pour un même nombre de paramètres ajustables (cinq dans notre cas), plus le nombre d'observations piézométriques augmente, plus il est difficile de vérifier l'ensemble des valeurs mesurées, ce qui ne traduit pas pour autant une moins bonne calibration du modèle d'écoulement, mais un moins bon ajustement de l'ensemble des observations piézométriques. La figure 5.16 montre que pour un même jeu d'observations piézométriques, la moyenne des fonctions objectifs diminue lorsque des données conditionnantes additionnelles sont introduites, traduisant une meilleure calibration du modèle d'écoulement.

Les paramètres ajustables de chacun des champs de conductivité hydraulique ont ainsi été optimisés par résolution du problème inverse. Cependant, pour certaines réalisations, les valeurs optimisées de conductivité hydraulique ne respectent plus l'ordre relatif des classes défini par seuillage, soit $K_{C1} < K_{C2} < K_{C3} < K_{C4} < K_{C5}$. Ces permutations dans l'ordre des classes créent des champs de conductivité hydraulique qui peuvent être considérés comme inacceptables d'un point de vue géologique. En effet, malgré une fonction objectif minimisée

(*) La combinaison (K_{25}, h_{15}) n'a pas été testée car, de manière générale, le nombre d'observations piézométriques est supérieur ou égal au nombre de données de conductivité hydraulique puisqu'un niveau d'eau peut être mesuré là où un essai de pompage a été réalisé.

et donc des hauteurs piézométriques mieux ajustées aux points de mesure, le champ des vitesses d'écoulement peut s'avérer erroné ; les résultats de transport qui en découlent sont alors médiocres. Afin de ne pas prendre ce genre de résultat en compte dans le calcul de la distribution spatiale des zones de protection, une sélection des réalisations a été réalisée sur base d'un critère d'ordre (qui peut être assimilé à l'intégration indirecte de données *soft* provenant de l'expérience du géologue). Les réalisations ne satisfaisant pas ce critère sont alors rejetées du processus, réduisant leur nombre de n à $m=n-r$, r étant le nombre de réalisations rejetées. Deux critères d'ordre ont été testés :

- *critère d'ordre strict* (Cr1) pour lequel l'ordre relatif des valeurs de conductivité hydraulique doit être rigoureusement respecté ;
- *critère d'ordre souple* (Cr2) pour lequel des permutations entre classes voisines sont tolérées. Ceci peut être justifié par le fait que les essais de pompage fournissent un ordre de grandeur de la conductivité hydraulique et non une valeur précise, ce qui rend possible, dans des limites acceptables, une certaine fluctuation des valeurs.

Le tableau 5.4 reprend, pour toutes les combinaisons des données considérées, le nombre m de réalisations acceptées par chacun des deux critères de sélection Cr1 et Cr2 au départ d'un nombre total de deux cents réalisations.

	K_{15}, h_{15}	K_{15}, h_{25}	K_{25}, h_{25}	$K_{15}, \rho_{150}, h_{15}$	$K_{15}, \rho_{300}, h_{15}$	$K_{25}, \rho_{150}, h_{25}$	$K_{25}, \rho_{300}, h_{25}$
Cr1	1	7	7	12	16	15	34
Cr2	42	53	58	62	81	109	103

Tableau 5.4 : Nombre de réalisations acceptées par chacun des critères de sélection Cr1 et Cr2

D'une manière générale, plus le nombre de données conditionnantes augmente, plus le nombre de réalisations acceptables augmente. Le critère Cr1 est un critère très, voire peut être trop restrictif, tandis que le critère d'ordre Cr2 permet une plus grande souplesse.

On obtient donc m champs de conductivité hydraulique optimisés par calibration du modèle d'écoulement et non rejetés par le critère de sélection choisi. Pour chacun de ces champs, le modèle d'écoulement a été simulé (résolution du problème direct), puis les isochrones associées à des temps de transfert de 1, 5, 10 et 20 jours ont été délimitées par *particle tracking*. Enfin, les distributions spatiales $[CAP(x,t)]$ ont été calculées comme décrit au paragraphe 5.1.

Par définition, les simulations stochastiques conditionnelles vérifient les valeurs de K aux points de mesure. Cependant, les étapes de paramétrisation des champs et de calibration du modèle d'écoulement, utilisées dans la MSA, modifient ces valeurs, et ce de manière différente pour chaque réalisation. On est dès lors amenés à se demander si on ne perd pas trop d'information sur les valeurs de K en introduisant ces étapes. La figure 5.17 donne, en chacun des points de mesure, la valeur "mesurée" comparée à la distribution des valeurs de l'ensemble des champs acceptés par le critère de sélection Cr2 (critère le moins sélectif) pour les combinaisons suivantes : (K_{15}, h_{15}) , (K_{15}, h_{25}) , $(K_{15}, \rho_{300}, h_{15})$ et $(K_{25}, \rho_{300}, h_{25})$.

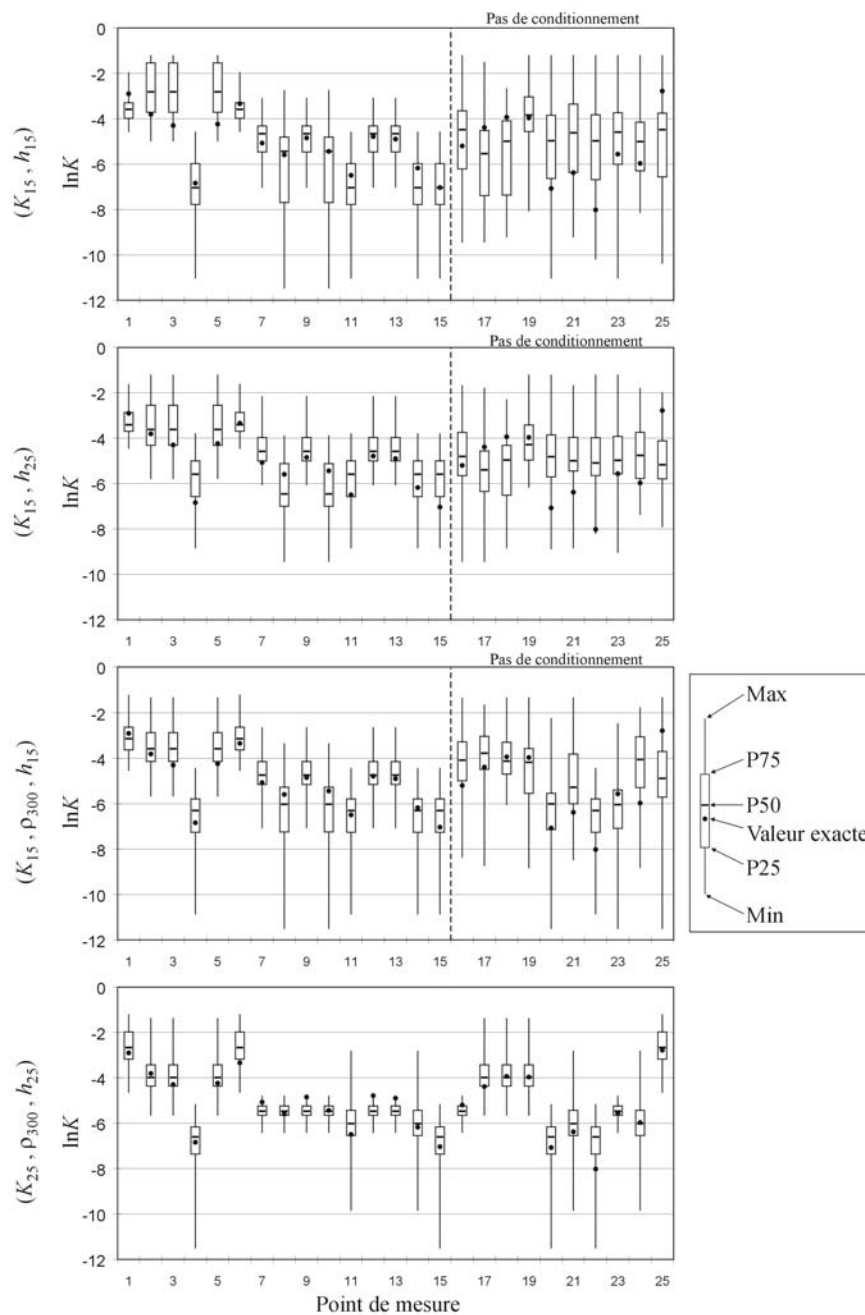


Figure 5.17 : Comparaison, en chacun des points de mesure, entre la valeur "mesurée" (point noir) et la distribution des valeurs de l'ensemble des champs acceptés par le critère de sélection Cr2 (représenté par un "box plot") pour différentes combinaisons des jeux de données

Cette figure montre que plus le conditionnement est important, plus l'incertitude est faible et plus la précision de l'estimation est bonne. Il faut tout de même noter que :

- lorsque le jeu K_{15} est utilisé, seules les 15 premiers points de mesure conditionnent les réalisations. Il est donc logique que l'estimation des 10 derniers points soit moins bonne ;
- les points de mesure possédant la même distribution appartiennent à une même classe ; chacune des distributions est donc associée à une des cinq classes définies par seuillage, différent suivant le jeu de K donné.

En comparant les résultats obtenus pour les combinaisons (K_{15}, h_{25}) et $(K_{15}, \rho_{300}, h_{15})$ par rapport à ceux de (K_{15}, h_{15}) , on remarque que l'introduction de 10 données supplémentaires de h apporte la même amélioration sur les valeurs de K pour les 15 premiers points de mesure que l'intégration de 300 valeurs de ρ . Cependant, ce jeu de ρ améliore également la définition de K au droit des 10 autres points de mesure (et dès lors probablement pour l'ensemble des valeurs du champ). L'introduction des valeurs de ρ permet d'améliorer la précision de l'estimation ("recentrage" de la distribution), sans pour autant réduire l'incertitude sur la valeur.

5.4.4 Analyse des résultats et discussion

Dans un but comparatif, on a considéré une situation non-conditionnelle pour laquelle aucune donnée de K , de ρ , ou de h n'a été prise en compte. En l'absence de donnée conditionnante, c'est la méthode de délimitation des zones de protection développée par VAN LEEUWEN [2000] qui a été utilisée (cf. chapitre 3). Deux cents simulations stochastiques non-conditionnelles du champ de conductivité hydraulique ont été générées sur base des mêmes paramètres S (palier) et λ (longueur de corrélation) que ceux fixés pour générer le champ de conductivité hydraulique de référence, c'est-à-dire $S = 2$ et $\lambda = 50$ mètres. Les champs de conductivité hydraulique introduits dans le modèle d'écoulement sont constitués d'une partie centrale correspondant à ces simulations stochastiques non-conditionnelles et de deux extensions latérales sur lesquelles une conductivité hydraulique uniforme de $4,2 \cdot 10^{-3}$ m/s, identique à celle du champ de référence, a été appliquée. Le modèle d'écoulement a été directement simulé pour chacun des champs ainsi obtenus. Aucune calibration n'a été réalisée et aucun critère de sélection n'a été appliqué. Cependant, pour certains champs, un dénoyage du puits dû à des valeurs de conductivité hydraulique simulées trop faibles au voisinage de celui-ci a été observé. Ces champs de conductivité hydraulique ont été éliminés de l'analyse de Monte Carlo puisqu'ils ne vérifient pas les conditions "réelles" d'écoulement. Pour les 184 champs restants, les isochrones à 1, 5, 10, et 20 jours ont été déterminées par *particle tracking*, puis les distributions spatiales correspondantes ont été calculées.

Dans ce qui suit, afin d'éviter toute confusion possible, nous indiquerons par $[CAP(x,t)]$ la distribution spatiale obtenue pour le cas **non-conditionnel**. Pour désigner un conditionnement, nous utiliserons une barre verticale suivie du ou des jeux de données conditionnantes. Par exemple, la distribution spatiale obtenue par application du critère de sélection Cr2 à partir de champs de conductivité hydraulique conditionnés par les jeux K_{15} et h_{15} sera notée $[CAP(x,t)|K_{15},h_{15},Cr2]$.

Pour rappel, deux cas ont été distingués : un premier cas (cas A) dans lequel les simulations stochastiques ont été conditionnées uniquement à partir d'un jeu de K et un second (cas B) dans lequel les simulations stochastiques ont été conditionnées à la fois à partir d'un jeu de K et d'un jeu de ρ . Le cas B nous permettra d'étudier l'effet de l'introduction des données de résistivité électriques sur les résultats obtenus. Dans un premier temps, aussi bien pour le cas A que pour le cas B, seuls les résultats obtenus par l'application du critère de sélection **Cr2** seront décrits.

Cas A : conditionnement par K et h

Les figures 5.17 et 5.18 présentent respectivement l'évolution des distributions $[CAP(x,t)]$ et $[CAP(x,t)|K_{15},h_{15},Cr2]$ en fonction du temps de transfert. Ces distributions donnent une idée de l'incertitude que l'on a sur la position de la zone de protection ; le contour noir indique la position de l'isocontour médiane $\Gamma^{(0,5)}$ et le contour blanc, celle de l'isochrone de référence obtenue pour le champ de conductivité hydraulique de référence.

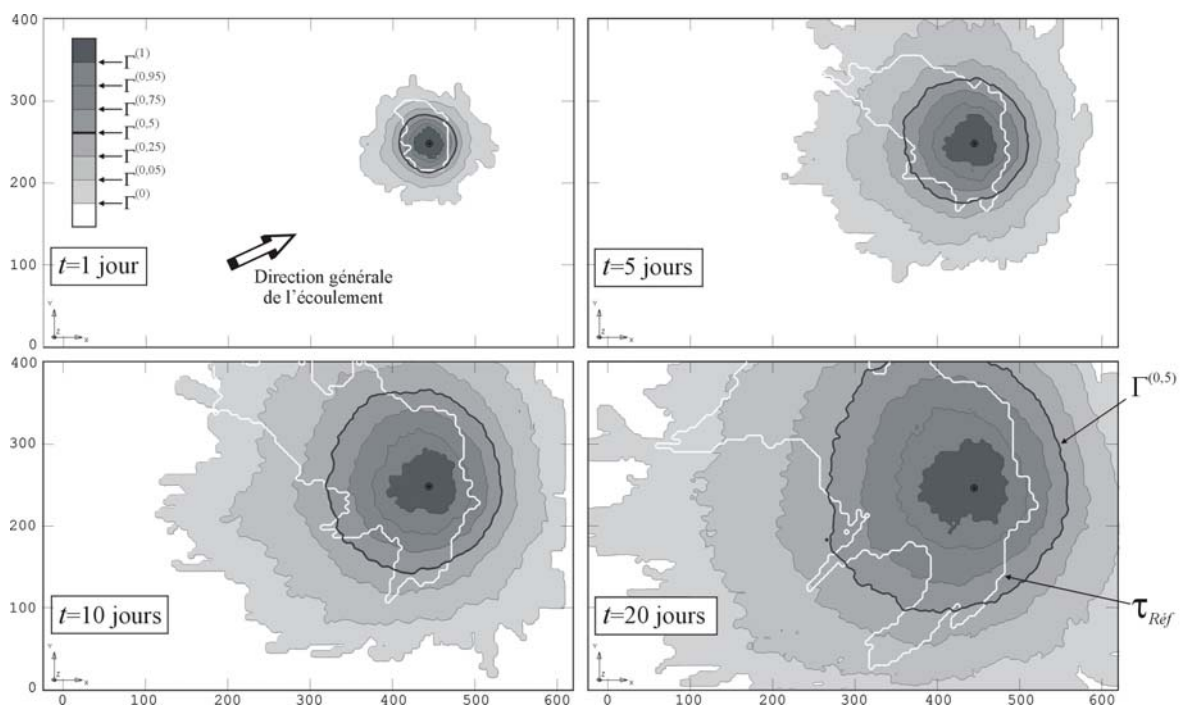


Figure 5.17 : Evolution de $[CAP(x,t)]$ en fonction du temps de transfert

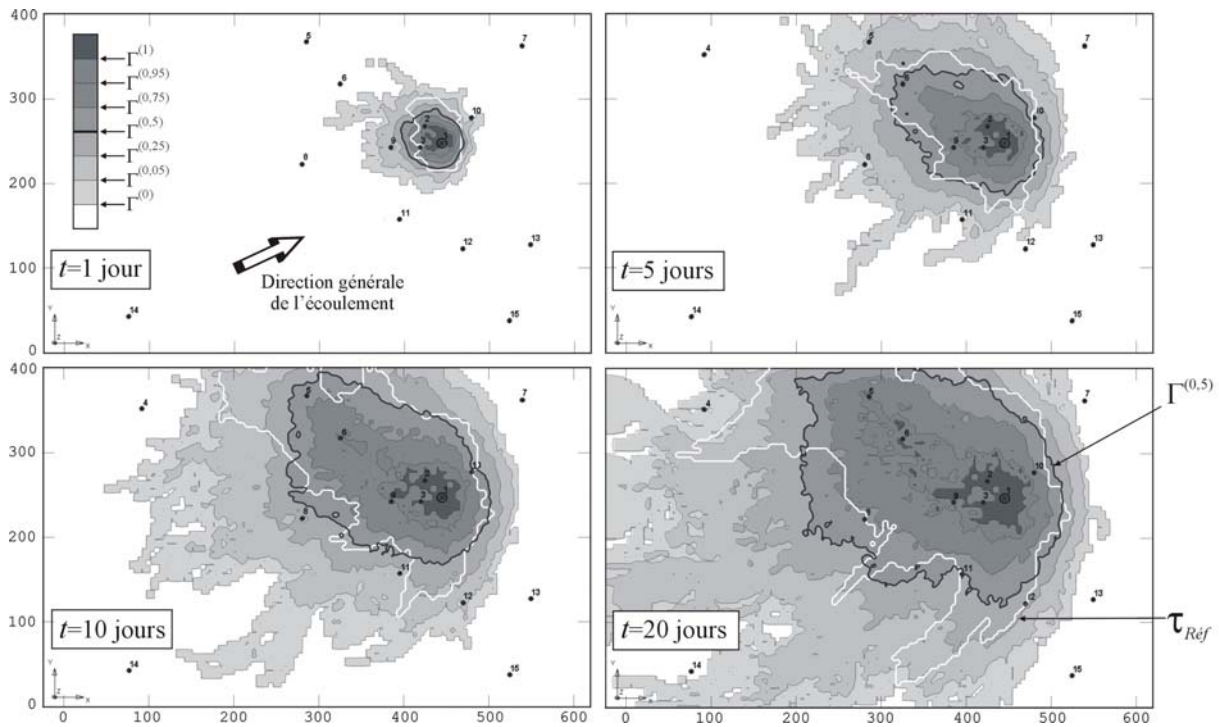


Figure 5.18 : Evolution de $[CAP(\underline{x},t)|K_{15},h_{15},Cr2]$ en fonction du temps de transfert

Comme on pouvait s'y attendre de manière intuitive, l'incertitude sur la localisation de la zone de protection croît avec le temps de transfert pour un même nombre de points de mesure. En effet, lorsque le temps de transfert augmente, la particule parcourt une distance plus importante et intègre donc plus de variabilité de la conductivité hydraulique, ce qui augmente l'incertitude. Ce constat est à mettre en relation avec la portée (ou la longueur de corrélation) de la conductivité hydraulique [VARLIJEN & SHAFER, 1991] : lorsque les distances parcourues sont plus petites que la portée, ce qui est le cas pour de faibles temps de transfert, toutes les valeurs de K rencontrées par les particules sont corrélées et donc plus homogènes que celles rencontrées pour des temps de transfert plus longs (variabilité moins importante).

En comparant les figures 5.17 et 5.18, on remarque que le conditionnement diminue considérablement l'écart entre $\Gamma^{(0,5)}$ et $\tau_{Réf}$: les isocontours des distributions $[CAP(\underline{x},t)]$ et en particulier $\Gamma^{(0,5)}$ sont presque circulaires autour du puits d'extraction, tandis que ceux des distributions $[CAP(\underline{x},t)|K_{25},h_{25},Cr2]$ ont une forme beaucoup plus asymétrique, se rapprochant beaucoup plus de celle de l'isochrone de référence $\tau_{Réf}$. L'écart le plus important entre l'isocontour $\Gamma^{(0,5)}$ et $\tau_{Réf}$ est observé en amont du puits. On remarque également que l'incertitude sur la localisation des zones de protection décroît sensiblement à l'aval du puits quand les données de K et h sont introduites dans l'analyse. De manière générale, cette incertitude est affectée par le gradient général d'écoulement. Elle est quantitativement plus

grande en amont du puits de pompage en raison de la dissymétrie du cône de rabattement qui est plus étalé vers l'amont (figure 5.19).

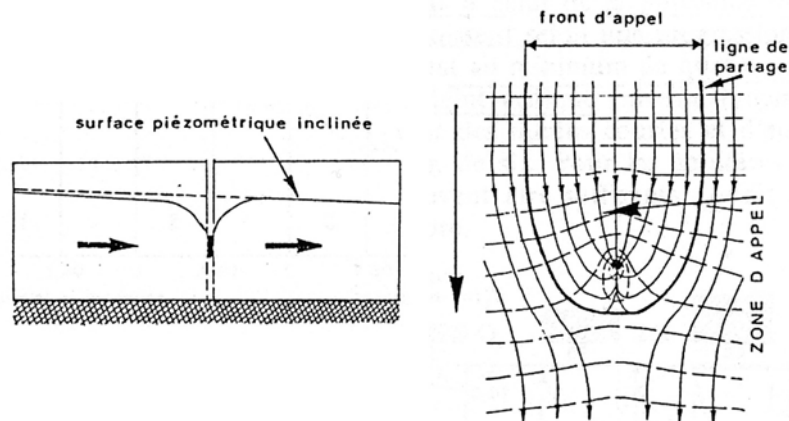


Figure 5.19 : Géométrie du cône de rabattement produit par pompage au puits [tiré de CASTANY, 1982]

La différence quantitative entre l'étendue de la distribution spatiale à l'aval et à l'amont du puits, d'autant plus forte que le temps de transfert augmente, est beaucoup plus marquée sur les distributions $[CAP(x,t)|K_{15}, h_{15}, Cr2]$ en raison de piézométries mieux ajustées, non seulement par calibration des modèles d'écoulement (à l'aide du jeu de h), mais également par conditionnement des simulations stochastiques (à l'aide du jeu de K). En effet, on observe que l'évolution de l'incertitude de $[CAP(x,t)|K_{15}, h_{15}, Cr2]$ en aval du puits a tendance à être freinée lorsque le temps de transfert augmente, contrairement à ce qui se passe en amont et à ce qui se passe pour les distributions $[CAP(x,t)]$. Ceci résulte probablement du fait que la partie avale de la ligne de partage (délimitant la zone d'appel du puits) se trouve pour l'ensemble des réalisations à une distance correspondant à un temps de transfert inférieur au temps de transfert désiré, n'augmentant donc plus l'incertitude dans cette direction. En amont, au contraire, la zone d'appel est très étalée. Les distances de transfert des particules vers le puits dans cette partie du champ d'écoulement sont donc plus grandes engendrant une incertitude plus prononcée.

La figure 5.20 compare les distributions spatiales $[CAP(x,t)|K, h, Cr2]$ à 5 et 20 jours pour différentes combinaisons des données de K et de h .

Sur cette figure, on constate une décroissance de l'incertitude ainsi qu'une meilleure concordance entre $\Gamma^{(0,5)}$ et $\tau_{Réf}$ quand le nombre de données conditionnantes augmente. L'amélioration observée est d'autant plus prononcée que le temps de transfert est important puisque le calcul de ces distributions intègre alors une plus grande quantité d'informations.

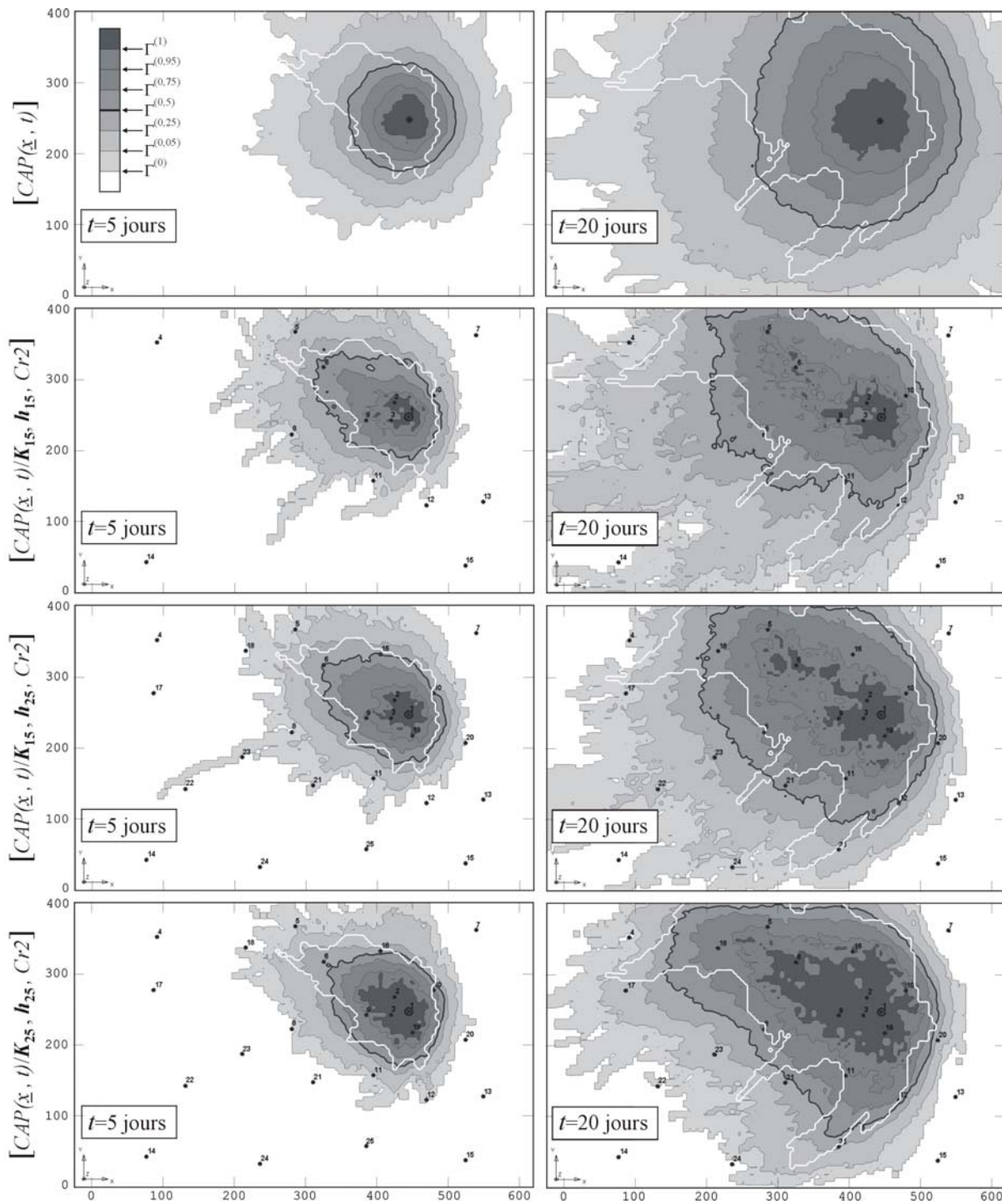


Figure 5.20 : $[CAP(x,t)|K,h,Cr2]$ à 5 et 20 jours pour différentes combinaisons des données de K et h

Les mesures de performance, w_{a1} relatif à l'extension de la ZI (région comprise entre $\Gamma^{(0)}$ et $\Gamma^{(1)}$), w_{a2} relatif à l'extension de la zone comprise entre $\Gamma^{(0,05)}$ et $\Gamma^{(0,95)}$, et w_b relatif à la déviation de $\Gamma^{(0,5)}$ par rapport à $\tau_{Réf}$, permettent de quantifier les résultats. L'amélioration due au conditionnement pour ces trois mesures de performance peut être estimée par les valeurs de Ω_{a1} , Ω_{a2} et Ω_b , définies par VAN LEEUWEN & al. [2000] comme suit :

$$\Omega_i = \frac{w_i^{nc} - w_i^c}{w_i^{nc}}, \quad i = a1, a2, b \quad (5.12)$$

où w_i^c est la mesure de la quantité w_i dans un cas conditionnel et w_i^{nc} est la mesure de cette même quantité dans le cas non-conditionnel.

Afin de pouvoir juger, en terme de **sécurité**, de la qualité des distributions spatiales obtenues, une quantité ws_i est calculée. Elle permet de vérifier que la distribution décrit de manière adéquate la variation de l'isochrone de référence $\tau_{Réf}$ (et donc de la zone de protection "réelle"), c'est-à-dire que $\tau_{Réf}$ est bien englobée dans la zone formée par $\Gamma^{(0)}$ (ws_1) ou par $\Gamma^{(0,05)}$ (ws_2). La mesure ws_i est définie de la manière suivante :

$$ws_i = \int_{\underline{x}} fs_i(\underline{x}, t) d\underline{x} \quad , i = 1,2 \quad (5.13a)$$

avec

$$fs_i(\underline{x}, t) = \begin{cases} 1 & \text{si } (I(\underline{x}, t))_{Réf} = 1 \text{ \& } p(CAP(\underline{x}, t)) \leq bs_i \\ 0 & \text{sinon} \end{cases} \quad , i = 1,2 \quad (5.13b)$$

où $bs_1 = 0$ quand on considère l'isocontour $\Gamma^{(0)}$ et $bs_2 = 0,05$ quand on considère l'isocontour $\Gamma^{(0,05)}$.

Cette quantité ws_i calcule donc la surface de l'isocontour $\tau_{Réf}$ située en dehors de l'isocontour $\Gamma^{(0)}$ ou $\Gamma^{(0,05)}$ suivant le cas. Elle vaut zéro lorsque $\tau_{Réf}$ est complètement enveloppée par l'isocontour considéré. C'est une mesure "optimiste" de l'erreur puisqu'on ne prend en compte que l'isocontour extérieur pour définir la zone de protection (ou zone de "confiance").

Les graphiques 5.21a à 5.21c reportent respectivement les valeurs de w_{a1} , w_{a2} et $w_b^{(*)}$ des distributions $[CAP(\underline{x}, t)|K, h, Cr2]$ en fonction des jeux de données conditionnantes pour les différents temps de transfert spécifiés, ainsi que les valeurs de ws_1 (figure 5.21a) ou ws_2 (figure 5.21b) relatives à ces distributions. Aux figures 5.21d à 5.21f, on retrouve les quantités Ω_{a1} , Ω_{a2} et Ω_b associées respectivement w_{a1} , w_{a2} et w_b .

(*) Les unités de mesures sont exprimées en nombre de cellules de 5m x 5m de côté, soit 25 m².

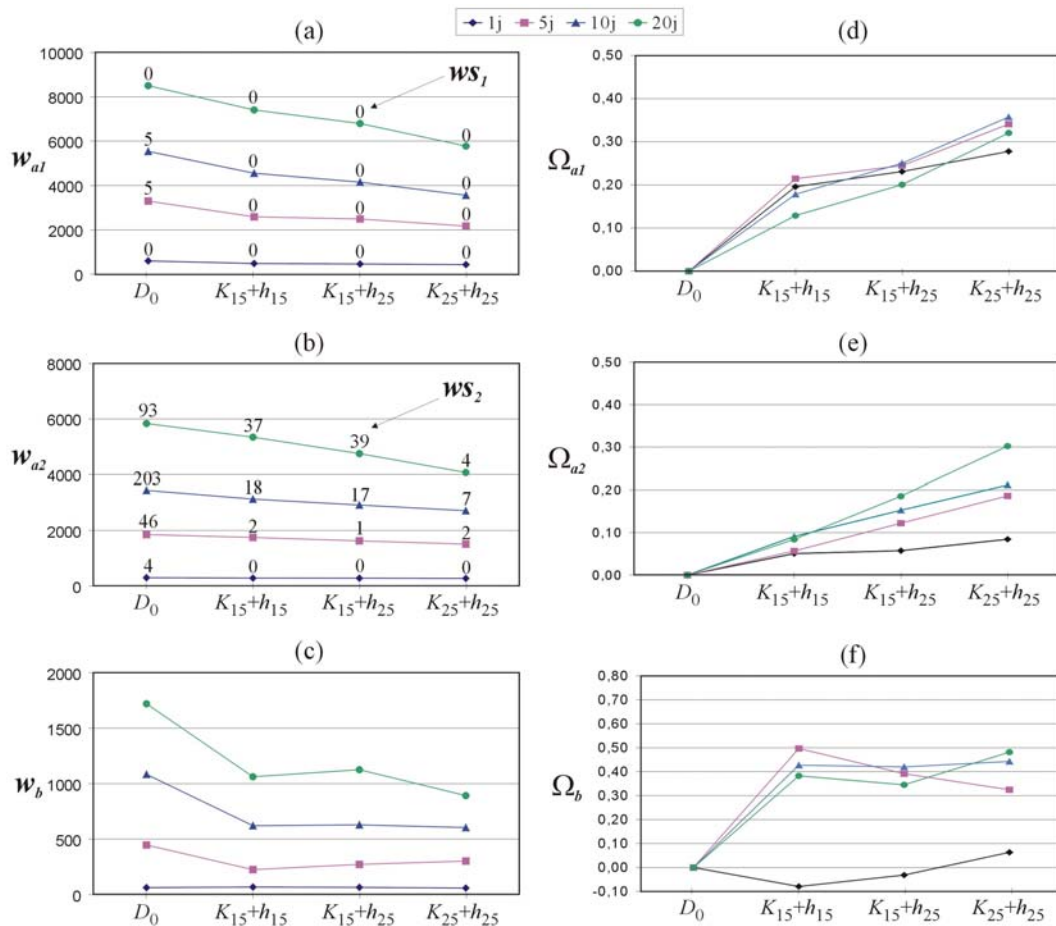


Figure 5.21 : Evolution de (a) w_{a1} , (b) w_{a2} , (c) w_b et (d) Ω_{a1} , (e) Ω_{a2} , (f) Ω_b en fonction des jeux de données conditionnantes K et h considérés (D_0 faisant référence au cas non-conditionnel), pour différents temps de transfert

L'observation des graphiques 5.21a et 5.21b montre clairement une réduction de w_{a1} et w_{a2} lorsqu'un plus grand nombre de données conditionnantes est pris en considération dans l'analyse. Cette réduction, d'autant plus importante que le temps de transfert est grand, provient d'une meilleure estimation des champs de conductivité hydraulique lorsque des données additionnelles sont utilisées pour les conditionner, tant par diminution de la variance de krigeage lors de la génération des simulations stochastiques conditionnelles (conditionnement par K) que par optimisation des zones de conductivité hydraulique lors de la calibration du modèle d'écoulement (conditionnement par h). La forte réduction de w_{a1} pour (K_{25}, h_{25}) est également dû, comme expliqué au paragraphe 5.4.3, à la meilleure estimation de la conductivité hydraulique appliquée sur les extensions latérales du champ, plus proche de celle de référence.

Les valeurs de ws_1 (figure 5.21a) indiquent que l'isochrone τ_{ref} se trouve bien à l'intérieur de la zone formée par l'isocontour $\Gamma^{(0)}$ pour tous les cas conditionnels considérés et tous les temps de transfert, ce qui est rassurant ! L'isocontour $\Gamma^{(0)}$ décrit donc de manière adéquate la variation de l'isochrone de référence. Par contre, d'après les valeurs de ws_2 (figure 5.21b),

l'isocontour $\Gamma^{(0,05)}$ n'englobe pas, pour les cas considérés, l'entièreté de $\tau_{Réf}$, si ce n'est pour de faibles temps de transfert. Néanmoins, le conditionnement apporte une amélioration traduite par une diminution de la surface de $\tau_{Réf}$ à l'extérieur de $\Gamma^{(0,05)}$, malgré une réduction de l'extension de l'incertitude. En règle générale, on remarque que plus le temps de transfert augmente, plus cette surface est importante.

L'amélioration de w_{a1} et w_{a2} par rapport au cas non-conditionnel est mise en évidence à l'aide des graphiques 5.21d et 5.21e représentant respectivement les valeurs de Ω_{a1} et Ω_{a2} . La réduction de la ZI par l'introduction de données conditionnantes est décrite par la croissance de Ω_{a1} : cette croissance, un peu plus importante par l'apport de K que de h , est pratiquement la même pour tous les temps de transfert. La superficie de la ZI est réduite de plus de 30% par rapport au cas non-conditionnel quand on prend en compte 25 données de K et de h . Au contraire, la croissance de Ω_{a2} est identique quel que soit le type de données introduites, mais dépend fortement du temps de transfert : pour 25 données de K et de h , la réduction varie entre 10% (pour $t = 1j$) et 30% (pour $t = 20j$) par rapport au cas non-conditionnel. L'amélioration Ω_{a2} augmente donc avec le temps de transfert.

Les figures 5.21c et 5.21f indiquent, pour les temps de transferts importants, une nette réduction de w_b par rapport au cas non-conditionnel (amélioration de l'ordre de 40 à 50%), lorsqu'un conditionnement est réalisé. Cependant, l'apport d'informations supplémentaires n'engendre pas d'amélioration sensible. A l'inverse, pour un temps de transfert de 1 jour, le conditionnement par un faible nombre de données de K détériore l'écart entre $\Gamma^{(0,5)}$ et $\tau_{Réf}$ par rapport au cas non-conditionnel. Ce phénomène est attribué à la faible étendue de la distribution par rapport à la taille des cellules du modèle ainsi qu'au nombre peu important de données conditionnantes dans la zone concernée.

Il faut tout de même noter que l'amélioration de Ω_{a1} , Ω_{a2} et Ω_b observée dans tous les cas est sous-estimée. Ceci est dû au fait que les champs de K du cas non-conditionnel sont constitués de simulations stochastiques non-conditionnelles générées, en l'absence de données, sur base du même modèle de variogramme que celui utilisé pour engendrer le champ de conductivité hydraulique de référence (soit les paramètres S et λ "exacts") et d'une valeur de K uniforme appliquée aux extensions latérales du modèle, identique à celle du champ de référence. Les champs de K des cas conditionnels, quant à eux, ont été construits sur base des paramètres du modèle de variogramme S et λ et d'une valeur de K uniforme, tous trois déduits des données. Ils sont donc propres à chaque jeu considéré et non pas "exacts" comme dans le cas non-conditionnel.

Nous sommes maintenant en mesure de justifier le choix de la technique de seuillage utilisée. Les différentes techniques proposées plus haut – TS_{unif} , TS_{gauss} et TS_{fmin} – ont été testées pour les combinaisons (K_{15}, h_{15}) et (K_{25}, h_{25}) puis leurs résultats ont été comparés. A partir de chacune de ces techniques, quatre valeurs de seuil ont été déterminées, subdivisant ainsi l'intervalle de variation de la conductivité hydraulique en cinq classes de valeur uniforme afin de se placer dans les mêmes conditions que précédemment. L'identification de ces seuils par la technique TS_{fmin} a été décrite au paragraphe 5.4.2. Pour les deux autres techniques, la détermination des seuils a été réalisée sur base des fonctions de distribution des données de K présentées à la figure 5.22.

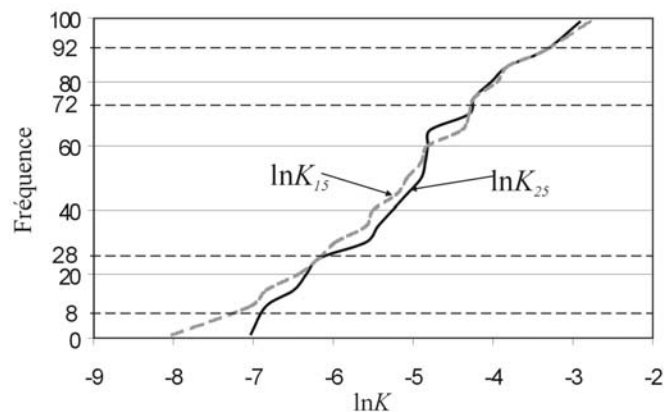


Figure 5.22 : Fonction de distribution des données de K_{15} et de K_{25}

Les quatre valeurs de seuils déterminées par TS_{unif} correspondent aux percentiles p20, p40, p60 et p80 de la fonction de distribution tandis que pour TS_{gauss} , elles correspondent aux percentiles p8, p28, p72 et p92. Ces percentiles ont été calculés à l'aide du logiciel d'analyse statistique STATLETS (développé par StatPoint, LLC), téléchargeable à partir du site <http://www.statlets.com>. Le tableau 5.5 reprend les valeurs des seuils calculées par chacune des techniques.

	$\ln K_{15}$				$\ln K_{25}$			
	$\ln K_{S1}$	$\ln K_{S2}$	$\ln K_{S3}$	$\ln K_{S4}$	$\ln K_{S1}$	$\ln K_{S2}$	$\ln K_{S3}$	$\ln K_{S4}$
TS_{unif}	-6,34	-5,26	-4,82	-4,02	-6,44	-5,50	-4,82	-3,95
TS_{gauss}	-6,84	-5,59	-4,30	-3,34	-7,05	-6,07	-4,27	-3,12
TS_{fmin}	-5,88	-5,26	-4,54	-3,58	-6,67	-5,78	-4,59	-3,58

Tableau 5.5 : Valeurs des seuils calculées par les différentes techniques de seuillage

La MSA a ensuite été poursuivie, comme schématisé à la figure 5.6, jusqu'à l'obtention des distributions spatiales. La figure 5.23 compare les mesures de performance obtenues dans chaque cas ainsi que leur amélioration par rapport au cas non-conditionnel.

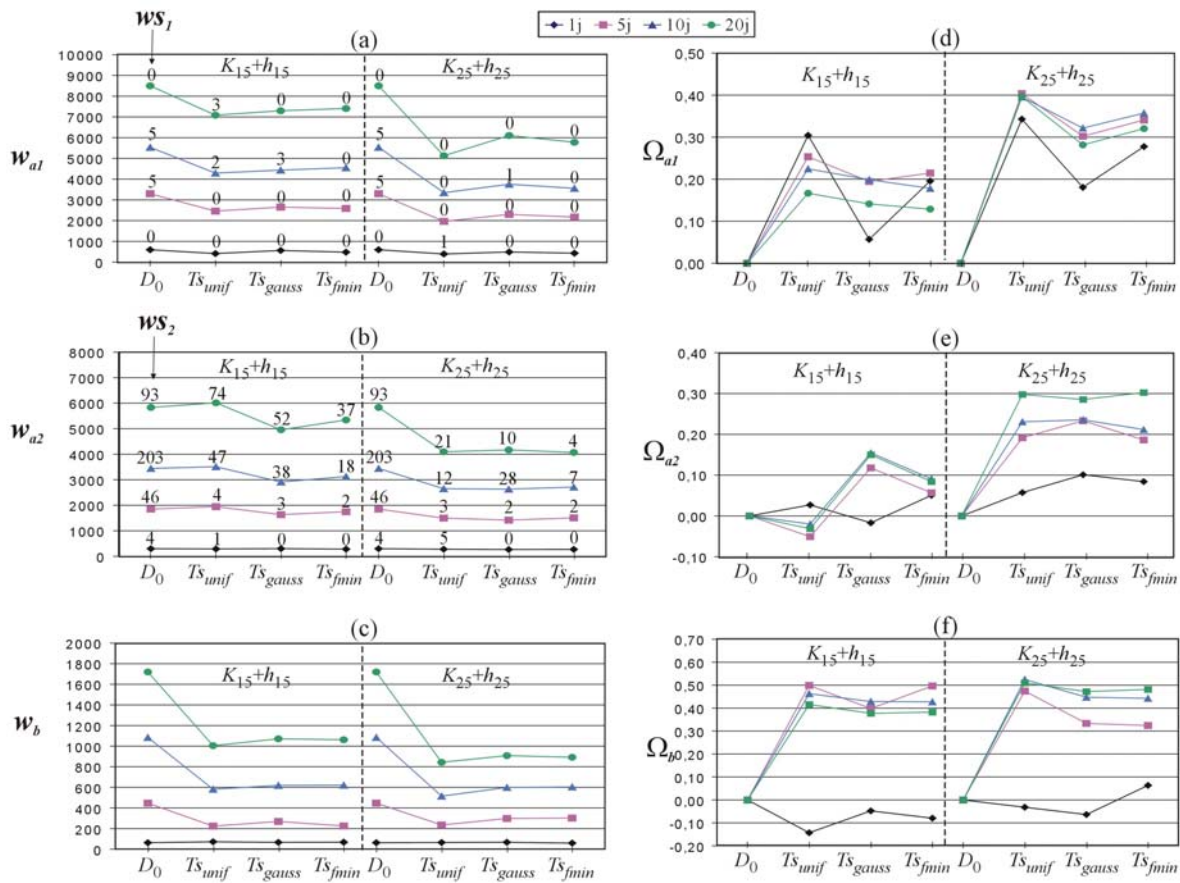


Figure 5.23 : Evolution de (a) w_{a1} , (b) w_{a2} , (c) w_b et (d) Ω_{a1} , (e) Ω_{a2} , (f) Ω_b en fonction de la technique de seuillage utilisée pour les jeux de données conditionnantes K et h (D_0 faisant référence au cas non-conditionnel), et pour différents temps de transferts

En comparant ces résultats, aucune technique de seuillage n'est mise en avant : l'amélioration de Ω_{a1} est plus grande pour TS_{unif} , celle de Ω_{a2} l'est pour TS_{gauss} et celle de Ω_b est pratiquement indépendante de la technique de seuillage choisie. Cependant, les valeurs de w_{s1} (figure 5.23a) montrent que la distribution englobe totalement l'isochrone $\tau_{Réf}$ lorsque TS_{fmin} est utilisée, contrairement aux autres techniques, et ce aussi bien pour (K_{15}, h_{15}) que pour (K_{25}, h_{25}) . De plus, les valeurs de w_{s2} sont plus faibles pour les distributions obtenues par TS_{fmin} , traduisant une meilleure description de $\tau_{Réf}$ par la zone formée de l'isocontour $\Gamma^{(0,05)}$. Ces deux dernières remarques nous paraissent essentielles lors de la détermination des zones de protection, c'est la technique de seuillage TS_{fmin} qui a été choisie. En effet, une zone d'incertitude ne comprenant pas l'isochrone de référence pourrait conduire à une sous-protection du puits.

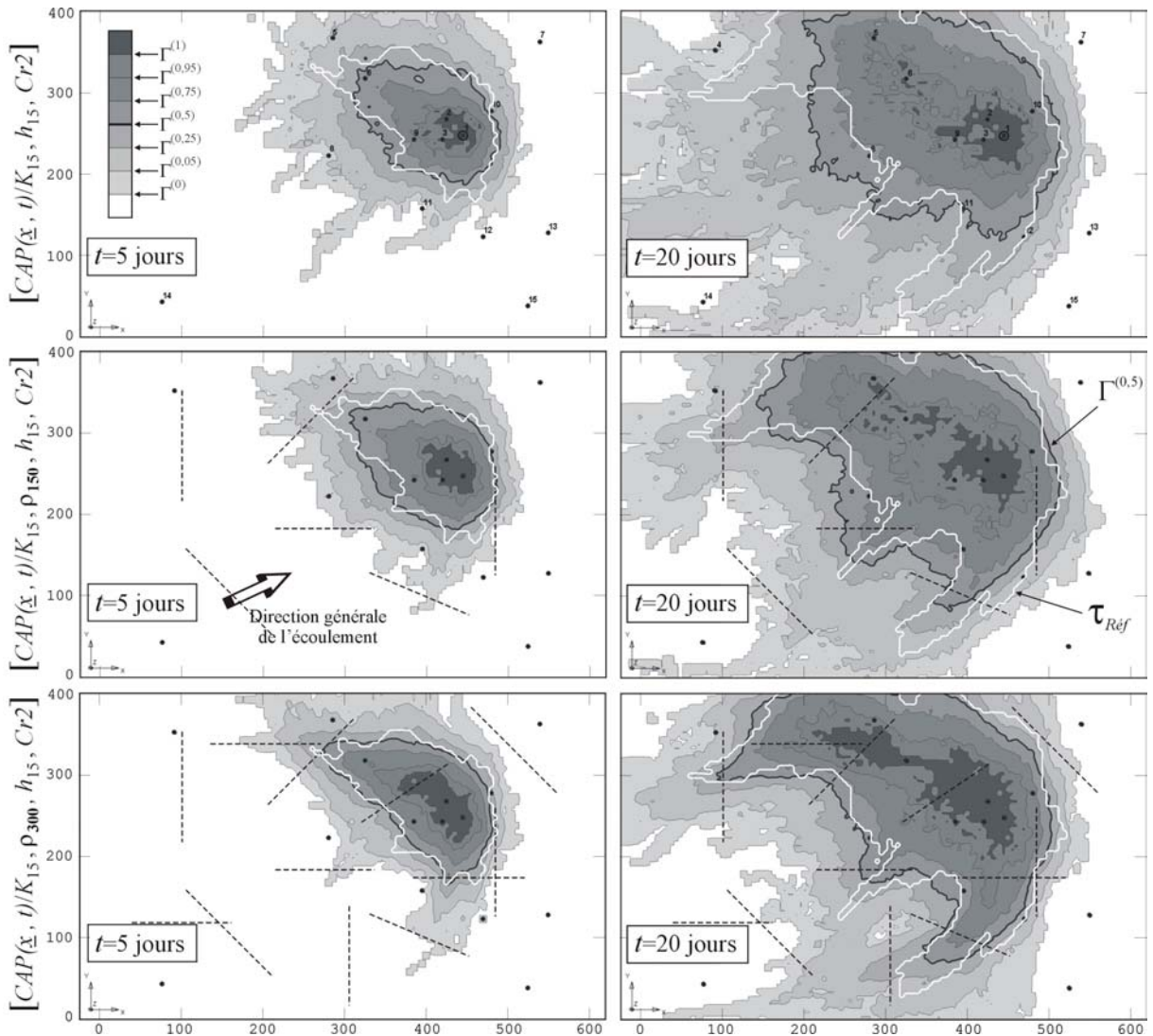


Figure 5.24 : $[CAP(x, t)|K_{15}, \rho, h_{15}, Cr2]$ à 5 et 20 jours pour les différentes combinaisons des données de K , ρ et h considérées

Cas B : conditionnement par K , ρ et h

Les figures 5.24 et 5.25 comparent les distributions spatiales avec et sans apport de données de résistivité électrique ρ dans la MSA, respectivement pour les différentes distributions $[CAP(x, t)|K_{15}, \rho, h_{15}, Cr2]$ et $[CAP(x, t)|K_{25}, \rho, h_{25}, Cr2]$ à 5 et 20 jours.

Lorsque le nombre de données de ρ augmente, on constate une amélioration incontestable de la correspondance entre $\Gamma^{(0,5)}$ et $\tau_{Réf}$ (et de l'allure générale de la distribution), mais également une décroissance de l'incertitude sur la localisation de la zone de protection. Ces améliorations sont plus prononcées sur la figure 5.24 pour laquelle le nombre de données conditionnantes de K et de h est plus faible et donc où les données de ρ apportent plus d'information. En effet, pour une densité plus grande de mesures de K , la grille de conditionnement, par apport de ρ , devient si dense que les points de conditionnement voisins

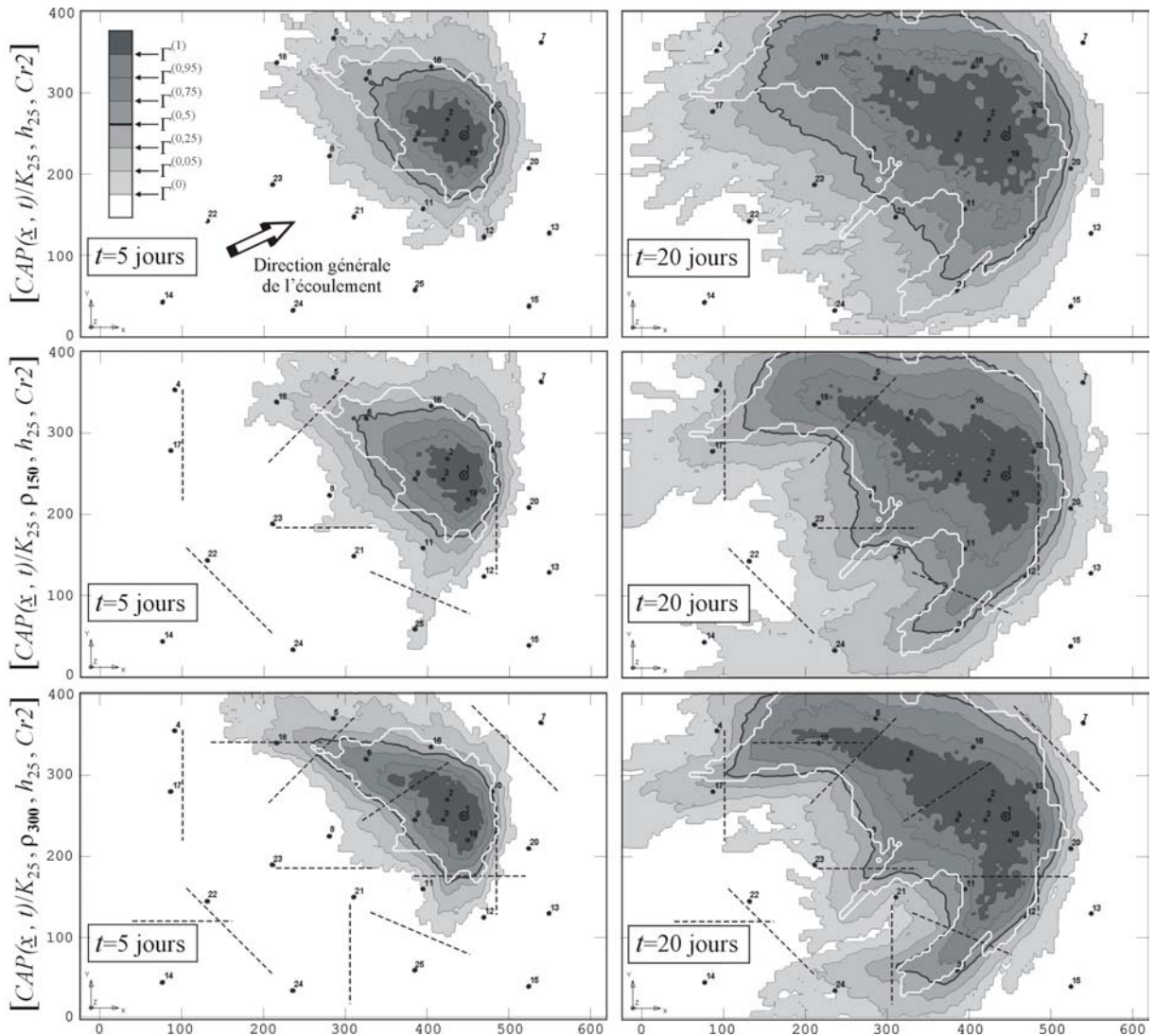


Figure 5.25 : $[CAP(x, t) | K_{25}, \rho, h_{25}, Cr2]$ à 5 et 20 jours pour les différentes combinaisons des données de K , ρ et h considérées

sont hautement corrélés et réduisent donc très peu l'incertitude. C'est un effet d'écran dû à la redondance des données qui a également été observé par d'autres auteurs, mais dans le cas d'un accroissement de données de K uniquement [RUBIN & DAGAN, 1992 ; VAN LEEUWEN & al., 2000].

En amont du puits de pompage, l'étendue de la distribution montre une variabilité considérable (beaucoup plus importante qu'en aval), dépendant principalement de la localisation des points de mesure – l'incertitude décroît là où de nouvelles données sont introduites –, mais dépendant également des valeurs de conductivité hydraulique : en effet, dans les régions de conductivité hydraulique élevée, là où les distances parcourues par les particules sont plus grandes et donc où plus de variabilité naturelle de K est rencontrée, l'incertitude est plus importante.

Les mesures de performance w_i ainsi que leur amélioration Ω_i par rapport au cas non-conditionnel sont toutes reprises graphiquement à la figure 5.26. Elles mettent en évidence le rôle joué par l'introduction des résistivités électriques dans la MSA.

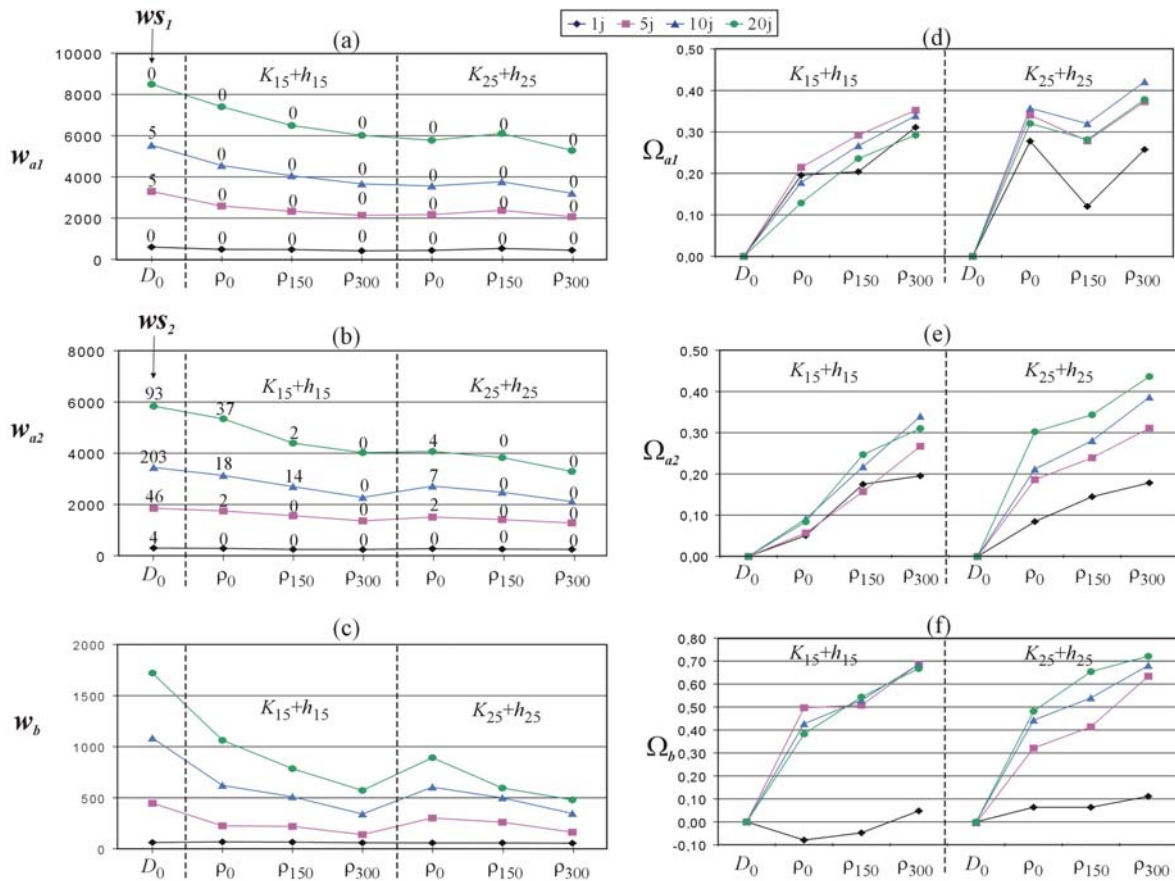


Figure 5.26 : Evolution de (a) w_{a1} , (b) w_{a2} , (c) w_b et (d) Ω_{a1} , (e) Ω_{a2} , (f) Ω_b en fonction des données conditionnantes K , ρ et h considérées, pour différents temps de transfert (D_0 faisant référence au cas non-conditionnel, ρ_0 au cas conditionnel sans apport de ρ , ρ_{150} et ρ_{300} aux cas conditionnels avec apport de ρ)

L'observation des graphiques de w_{a1} (fig. 5.26a) et w_{a2} (fig. 5.26b) montre clairement une réduction de l'extension de l'incertitude lorsqu'un plus grand nombre de données conditionnantes est pris en considération dans l'analyse, d'autant plus importante que le temps de transfert est grand. Cette réduction provient d'une meilleure estimation des champs de conductivité hydraulique lorsque des données additionnelles sont utilisées pour les conditionner. On remarque, en comparant les valeurs de w_{a1} et w_{a2} pour les combinaisons (K_{15} , ρ_{300} , h_{15}) et (K_{25} , h_{25}), que l'introduction d'un grand nombre de résistivités électriques (dans notre cas 300) donne des résultats identiques (si pas meilleurs) à ceux obtenus par l'apport de dix mesures supplémentaires de K et de h .

La figure 5.26a affiche des valeurs de w_{s1} nulles pour tous les cas considérés, indiquant que ces distributions décrivent bien l'isochrone $\tau_{Réf}$. A la figure 5.26b, les valeurs de w_{s2} diminuent lorsque le nombre de données de ρ augmente et ce malgré une réduction de l'extension de la zone considérée, traduisant une meilleure définition de la distribution. Pour certaines combinaisons, $(K_{15}, \rho_{300}, h_{15})$, $(K_{25}, \rho_{150}, h_{25})$ et $(K_{25}, \rho_{300}, h_{25})$, les données fournissent même suffisamment d'information pour que $\tau_{Réf}$ soit totalement enveloppée par l'isocontour $\Gamma^{(0,05)}$.

Les graphiques 5.26d et 5.26e, respectivement Ω_{a1} et Ω_{a2} , mettent en évidence l'amélioration des mesures de performance w_{a1} et w_{a2} par rapport au cas non-conditionnel. La croissance de Ω_{a1} est peu dépendante du temps de transfert. Pour les jeux conditionnants K_{15} et h_{15} , cette croissance est la même par augmentation du nombre de ρ et peut atteindre plus de 30% par rapport au cas non-conditionnel pour le jeu ρ_{300} , équivalent à ce qui a été observé par augmentation du nombre de K et de h (de 15 à 25). On remarque, pour les jeux K_{25} et h_{25} et les temps de transfert élevés, une faible dégradation de Ω_{a1} lorsque l'on passe de 0 à 150 données de ρ . En cause, la grande amélioration (plus de 30%) déjà apportée lors du conditionnement effectué uniquement par les jeux K_{25} et h_{25} (sans intégration de ρ). De plus, d'après les distributions de la figure 5.25, l'apport de données de ρ , dans ce cas précis, met en évidence des zones de plus grande conductivité hydraulique, ce qui, comme vu précédemment, augmente fortement l'incertitude dans ces régions. La diminution de l'incertitude en d'autres endroits ne compensant pas suffisamment cette augmentation, la valeur de Ω_i décroît. Cette dégradation par rapport au cas non-conditionnel est beaucoup plus marquée sur la figure 5.26d pour un faible temps de transfert ($t = 1j$) et apparaît toujours mais de manière moins prononcée quand ρ augmente. Néanmoins, pour les temps de transfert élevés, la combinaison $(K_{25}, \rho_{300}, h_{25})$ permet d'atteindre jusqu'à 40% de réduction de la ZI.

La croissance de Ω_{a2} dépend moins du temps de transfert dans le cas B lorsqu'on augmente le nombre de ρ , que lorsque le nombre de K et de h augmente comme au cas A (cf. figure 5.21e). Pour la combinaison (K_{25}, h_{25}) , elle est même pratiquement indépendante de t . Contrairement à Ω_{a1} , il n'y a pas de dégradation par rapport au cas non-conditionnel car les isocontours $\Gamma^{(0,05)}$ et $\Gamma^{(0,95)}$ sont moins affectés que $\Gamma^{(0)}$ et $\Gamma^{(1)}$ par les valeurs extrêmes d'une réalisation unique.

Sur la figure 5.26c, on observe une réduction importante de w_b lorsque le nombre de données de ρ augmente, traduisant une moins grande déviation de $\Gamma^{(0,5)}$ par rapport à $\tau_{Réf}$, contrairement à ce qui est observé au cas A par augmentation du nombre de K ou de h (cf. figure 5.21c). La figure 5.26f permet de se rendre compte de cette amélioration qui atteint, au maximum, près de 70% (contre 40 à 50% sans apport de ρ). L'introduction de données de ρ

dans la MSA joue donc un rôle essentiel dans l'amélioration de l'allure de $\Gamma^{(0,5)}$ par rapport à $\tau_{Réf}$.

La figure 5.27 permet d'apprécier les résultats obtenus pour le conditionnement maximal par les jeux K_{25} , ρ_{300} et h_{25} . L'amélioration par rapport aux distributions de la figure 5.18, pour lesquelles un conditionnement par K_{15} et h_{15} a été réalisé, est considérable pour chacun des temps de transfert envisagés, ce qui met en évidence l'intérêt de l'intégration des données de prospection géophysique et plus particulièrement (dans notre cas) de résistivités électriques dans la génération des champs de conductivité hydraulique pour la délimitation des zones de protection.

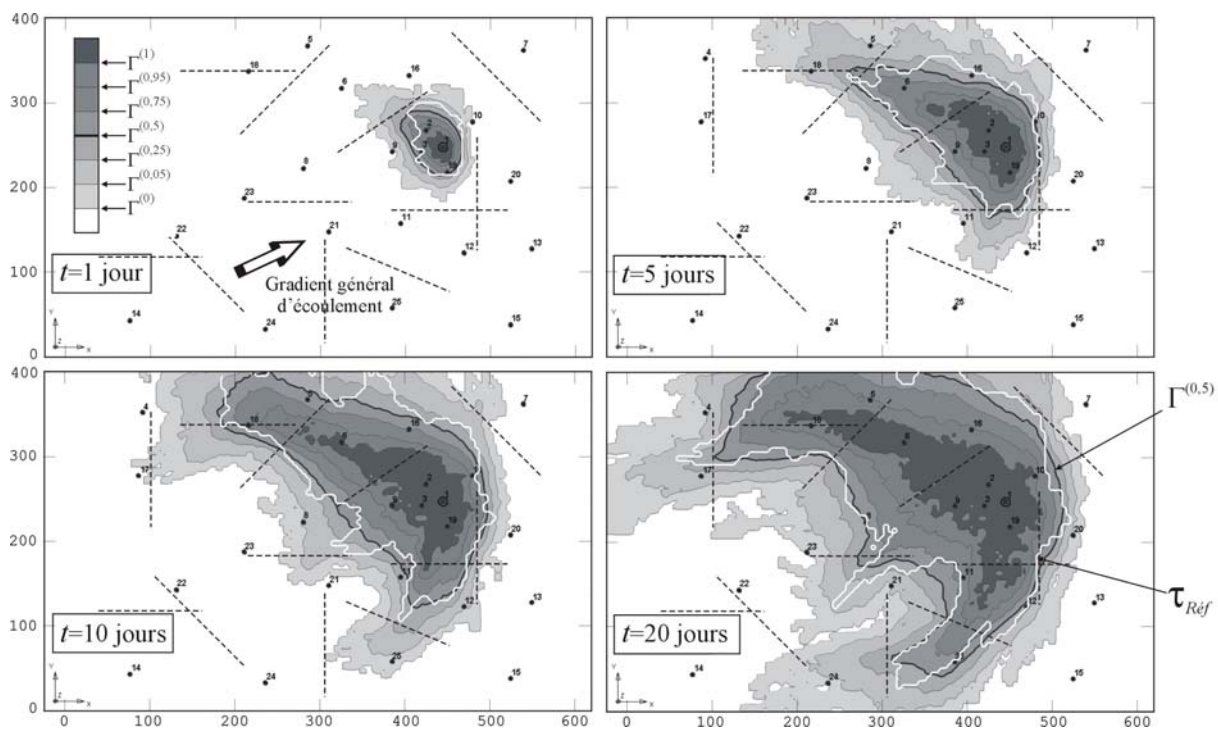


Figure 5.27 : Evolution de $[CAP(x,t)|K_{25}, \rho_{300}, h_{25}, Cr2]$ en fonction du temps de transfert

Application du critère d'ordre strict Cr1

Il est bon de rappeler que le critère d'ordre strict Cr1 élimine un nombre important de réalisations (cf. tableau 5.4) et notamment dans le cas de la combinaison (K_{15}, h_{15}) pour laquelle il ne reste plus qu'une seule réalisation sur les deux cents générées. Vu le faible nombre de réalisations restantes, il est dès lors impossible, pour la plupart des combinaisons, d'évaluer l'isocontour $\Gamma^{(0,05)}$ (donc w_{a2} et Ω_{a2}) et en particulier pour le cas (K_{15}, h_{15}) , d'évaluer également l'isocontour $\Gamma^{(0)}$ (donc w_{a1} et Ω_{a1}). Pour bien faire, il aurait fallu augmenter le nombre de simulations stochastiques conditionnelles engendrées pour que le

nombre de champs acceptés par le critère Cr1 soit représentatif. Néanmoins, d'après la figure 5.28, présentant les distributions $[CAP(x,t)|K, \rho, h, Cr1]$ à 20 jours obtenues par application du critère d'ordre strict Cr1, aussi bien pour le cas A que pour le cas B, on peut remarquer que l'isochrone unique obtenue par le conditionnement (K_{15}, h_{15}) est une assez bonne approximation de l'isochrone de référence τ_{ref} , mais malheureusement pas suffisante pour la décrire totalement.

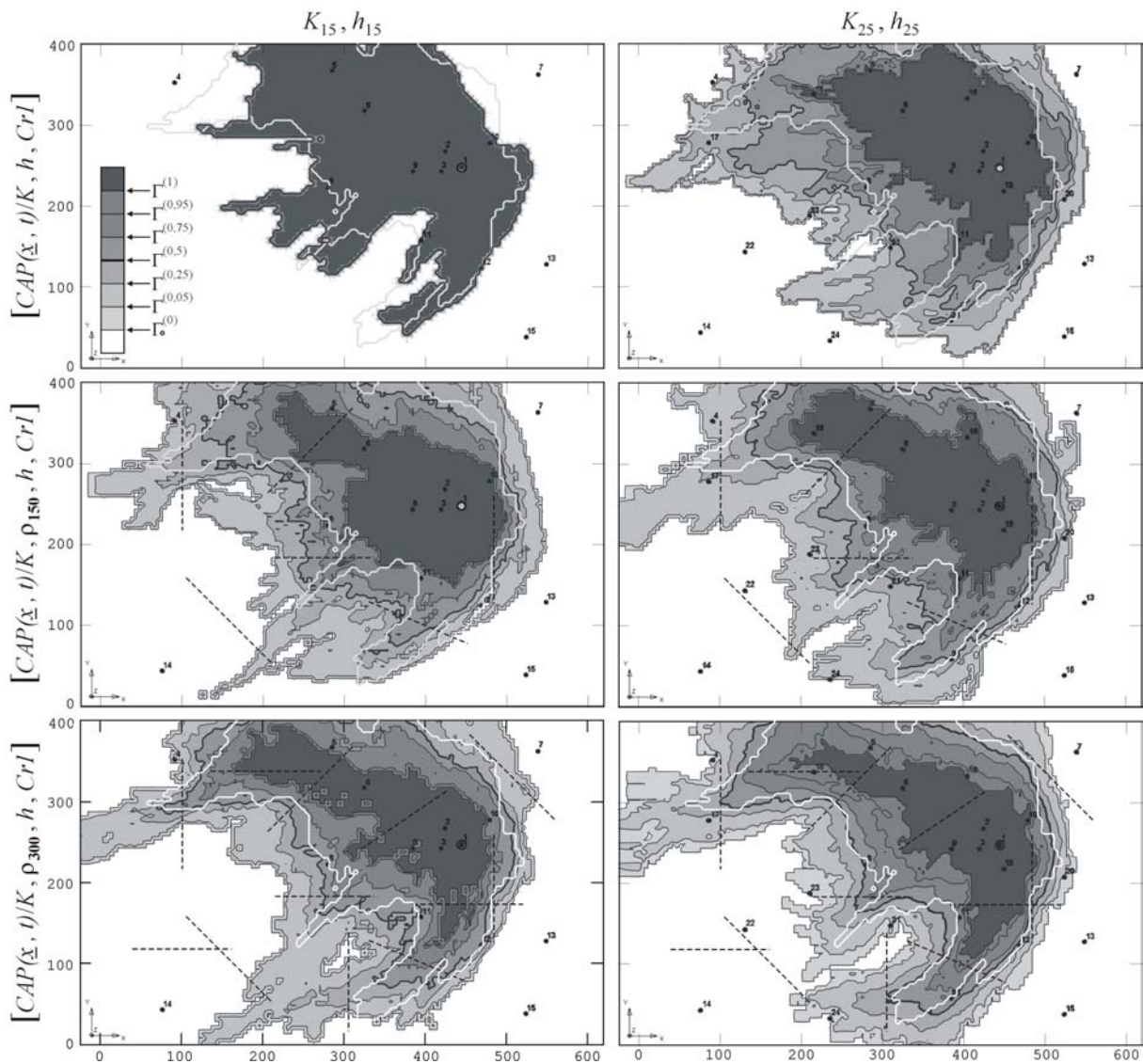


Figure 5.28 : $[CAP(x,t)|K, \rho, h, Cr1]$ à 20 jours pour les différentes combinaisons des données de K , ρ et h

A la figure 5.29, on peut comparer les mesures de performance w_{a1} et w_b associées à chacune des distributions ainsi que leur amélioration Ω_{a1} et Ω_b par rapport au cas non-conditionnel.

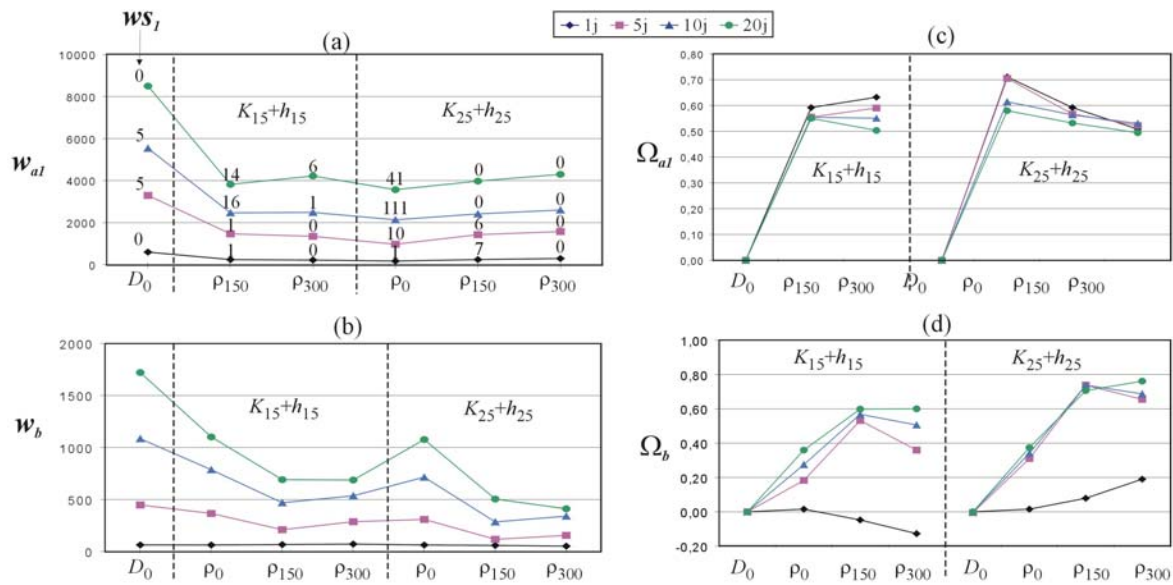


Figure 5.29 : Critère Cr1. Evolution de (a) w_{a1} et (b) w_b et (d) Ω_{a1} , (e) Ω_b en fonction des données conditionnantes K , ρ et h considérées, pour différents temps de transferts (D_0 faisant référence au cas non-conditionnel, ρ_0 au cas conditionnel sans apport de ρ et ρ_{150} , ρ_{300} au cas conditionnel avec apport de ρ)

La figure 5.29c montre une très grande amélioration Ω_{a1} , de l'ordre de 50% à 70%, dans tous les cas, ce qui est beaucoup plus important que ce qui a été observé par l'application du critère Cr2. En réalité, cette amélioration est due au fait que le nombre de réalisations prises en compte dans l'analyse est faible. En effet, lorsque ce nombre est trop faible, il ne permet pas la convergence^(*) de la distribution. On remarque également une décroissance de Ω_{a1} quand ρ augmente, provoqué non pas par une détérioration des résultats, mais par l'augmentation du nombre d'isochrones dans le calcul de la distribution spatiale. Les valeurs de w_{a1} reprises à la figure 5.29a diminuent avec l'introduction de données conditionnantes supplémentaires atteignant des valeurs nulles pour le conditionnement maximal (K_{25} , ρ_{300} , h_{25}). Malgré le faible nombre de réalisations considéré, la déviation w_b diminue lorsque ρ augmente, apportant une amélioration de près de 80% pour ce conditionnement maximal, semblable à ce qui a été observé pour Cr2.

Une comparaison réellement objective entre les deux critères Cr1 et Cr2 n'a pas été réalisée puisque la convergence n'a pas été atteinte pour Cr1. Cependant, ce critère Cr1, fort strict et demandant donc des efforts de calcul beaucoup plus importants pour converger, ne nous paraît pas indispensable, le critère Cr2 fournissant de très bons résultats.

^(*) La convergence de la distribution est atteinte lorsque l'ajout d'une réalisation supplémentaire ne modifie plus la distribution.

5.5 Conditionnement par ρ et h

L'objectif de ce travail étant de développer, en fonction des données disponibles, une méthode applicable à des aquifères réels, il est intéressant d'envisager le cas où le nombre de données de conductivité hydraulique est trop faible pour que les propriétés statistiques spatiales de ce paramètre puissent être inférées de l'analyse géostatistique, ce qui est souvent le cas sur les sites de captage.

Deux jeux de données ont été considérés : un jeu de résistivités électriques et un jeu de hauteurs piézométriques. Le schéma de la MSA est repris à la figure 5.30 dans cette situation où seules des données de ρ et de h sont disponibles.

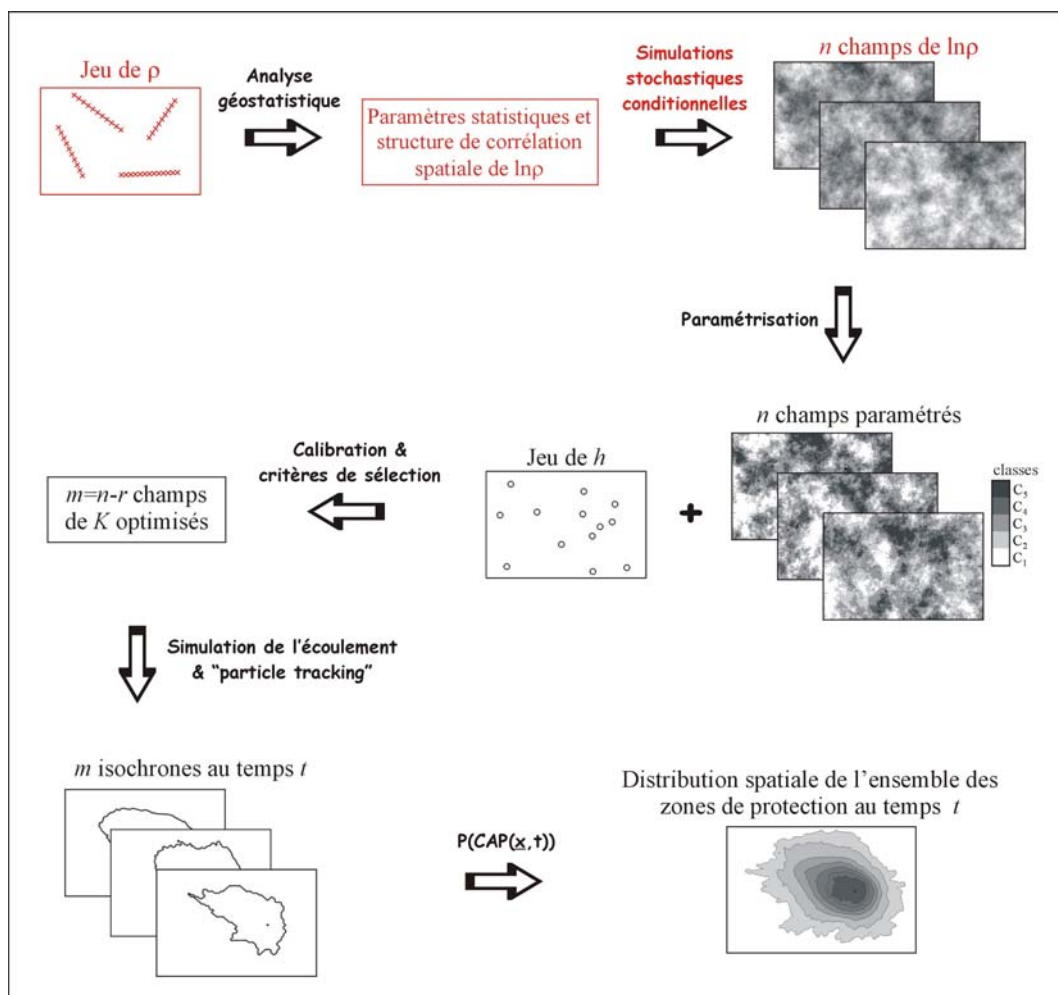


Figure 5.30 : MSA dans la situation où seules des données de ρ et de h sont disponibles

5.5.1 Analyse géostatistique et génération des simulations stochastiques

Les jeux de h et de ρ sont identiques à ceux de la situation précédente. Leur localisation sur la zone d'étude proprement dite est donc reprise à la figure 5.8. Comme mentionné précédemment, les histogrammes du logarithme népérien des données de ρ , notés $\ln\rho_{150}$ et $\ln\rho_{300}$, présentés à la figure 5.9 mettent en évidence le caractère gaussien de ces distributions. Les résultats statistiques obtenus sur ces deux jeux $\ln\rho_{150}$ et $\ln\rho_{300}$ sont identiques à ceux réalisés précédemment (cf. tableau 5.1).

Dans cette situation particulière, les champs générés ne consistent plus en des champs de $\ln K$ mais de $\ln\rho$, conditionnés uniquement sur les données de ρ . Pour ce faire, deux cents simulations stochastiques non-conditionnelles de $\ln\rho$ ont d'abord été engendrées par la méthode des bandes tournantes, puis conditionnées à l'aide d'une étape supplémentaire de double krigeage. Les simulations obtenues possèdent alors les mêmes propriétés statistiques spatiales que le jeu de données considéré et vérifient ces données aux points de mesure ; elles sont toutes équiprobables. Les paramètres – palier S et portée a (ou échelle intégrale λ) – relatifs à chacun des jeux $\ln\rho_{150}$ et $\ln\rho_{300}$, nécessaires pour créer ces simulations stochastiques, ont été déterminés par modélisation des variogrammes expérimentaux. Une validation croisée a été réalisée pour différents modèles théoriques ajustés de manière visuelle sur ces variogrammes expérimentaux afin de choisir pour chacun des deux jeux de ρ le modèle le plus adéquat, c'est-à-dire celui qui fournit la meilleure estimation. Les deux modèles choisis (figure 5.31) montrent un effet de pépité qui, dans notre cas synthétique, provient probablement de variations locales des valeurs de ρ . Ces modèles sont tous deux composés d'un effet de pépité égal à 0,027 et d'une structure exponentielle caractérisée par une portée a de 150 mètres (correspondant à une échelle intégrale λ de 50 mètres) et un palier S égal à 0,33 pour ρ_{150} et 0,31 pour ρ_{300} , que l'on peut noter respectivement :

$$\gamma_{\ln\rho_{150}}(h) = 0,027 + 0,33[1 - \exp(-h/50)] \quad \text{et} \quad \gamma_{\ln\rho_{300}}(h) = 0,027 + 0,31[1 - \exp(-h/50)]$$

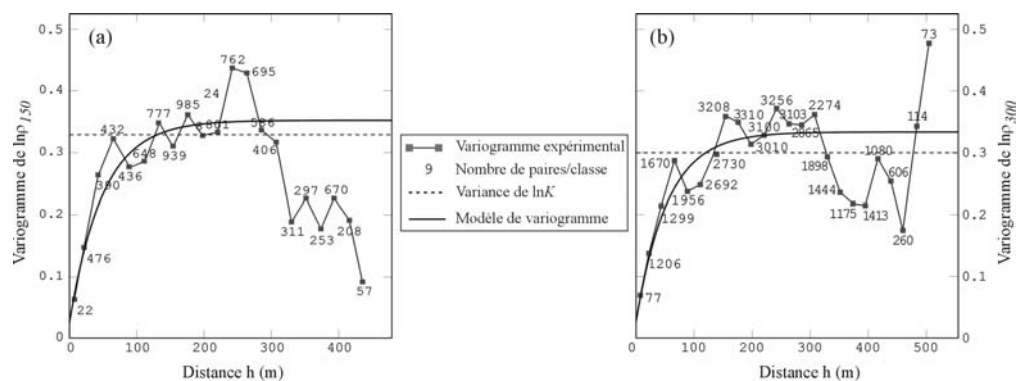


Figure 5.31 : Variogramme de (a) $\ln\rho_{150}$ et (b) $\ln\rho_{300}$

5.5.2 Paramétrisation

En acceptant l'hypothèse de l'existence d'une corrélation entre K et ρ , démontrée pour les aquifères de la plaine alluviale de la Meuse au paragraphe 4.3.3.1, on peut accepter que la répartition spatiale de K soit la même que celle de ρ . Sur base de cette hypothèse, tous les champs de $\ln\rho$ obtenus ont été paramétrés, avec comme intérêt principal de passer d'un champ de $\ln\rho$ à des zones de K de valeur uniforme et ainsi pouvoir calibrer les modèles d'écoulement par les observations piézométriques. Afin de se placer dans les mêmes conditions qu'à la situation précédente, l'intervalle de variation des données de ρ a été subdivisé en cinq classes de valeur uniforme, représentant les paramètres ajustables. Les valeurs des quatre seuils correspondants ont été déterminées pour chacun des jeux ρ_{150} et ρ_{300} par la technique de seuillage TS_{fmin} , assurant une variabilité minimale des données de $\ln\rho$ au sein de chaque classe. Les résultats obtenus à l'aide de l'algorithme illustré à l'annexe D sont présentés à la figure 5.32.

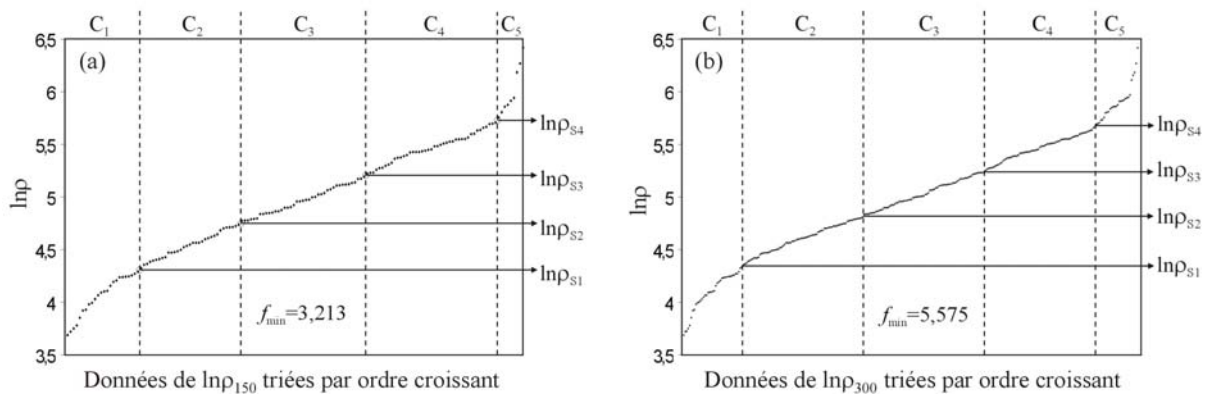


Figure 5.32 : Combinaison des données de $\ln\rho$ minimisant la fonction f (a) pour ρ_{150} et (b) pour ρ_{300}

Le tableau 5.6 reprend les quatre valeurs de seuils – $\ln\rho_{S1}$, $\ln\rho_{S2}$, $\ln\rho_{S3}$, $\ln\rho_{S4}$. Comme on peut le voir sur ce tableau, l'apport des 150 données supplémentaires ne modifie que très légèrement leurs valeurs.

	$\ln\rho_{S1}$	$\ln\rho_{S2}$	$\ln\rho_{S3}$	$\ln\rho_{S4}$
$\ln\rho_{150}$	4,31	4,76	5,21	5,73
$\ln\rho_{300}$	4,35	4,82	5,25	5,68

Tableau 5.6 : Valeurs des seuils calculées par la technique de seuillage TS_{fmin}

Sur base de ces seuils, les valeurs de $\ln\rho$ de chacun des champs ont été regroupés en cinq zones. A partir de l'hypothèse de l'existence d'une corrélation entre K et ρ , chaque zone de ρ peut être associée à une zone de K de valeur uniforme, représentant les cinq paramètres ajustables du modèle d'écoulement, que l'on peut associer à différents "hydrofaciès".

5.5.3 Calibration, critères de sélection, simulation du modèle d'écoulement et *particle tracking*

Chacun des champs paramétrés constitue, comme mentionné précédemment, la partie centrale du modèle, à laquelle on ajoute deux extensions latérales de mêmes dimensions et de conductivité hydraulique constante et uniforme. Cette valeur uniforme, fixée à 4.10^{-3} m/s pour les champs conditionnés par le jeu ρ_{150} et $4,6.10^{-3}$ m/s pour ρ_{300} , correspond à la moyenne géométrique du jeu de résistivité électrique envisagé, transformée en conductivité hydraulique à l'aide de la relation linéaire $\ln\rho-\ln K$ définie au chapitre 4. Lors de la comparaison des résultats, il faudra tenir compte de la possible influence de ces valeurs, différentes de celle du champ de référence ($4,2.10^{-3}$ m/s).

Les paramètres ajustables de chacun des champs de conductivité hydraulique ainsi créés ont été optimisés par calibration automatique du modèle d'écoulement sur base des données de h . Le graphique de la figure 5.33 montre l'évolution de la moyenne des fonctions objectifs minimisées (après calibration du modèle d'écoulement) en fonction du nombre de réalisations générées pour les différentes combinaisons des données testées – (ρ_{150}, h_{15}) , (ρ_{150}, h_{25}) , (ρ_{300}, h_{15}) et (ρ_{300}, h_{25}) .

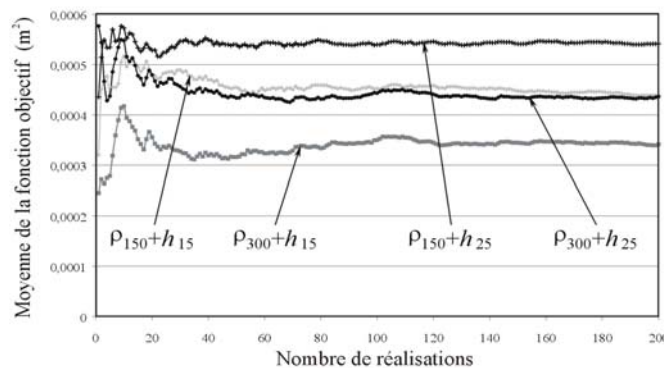


Figure 5.33 : Evolution de la moyenne des fonctions objectifs en fonction du nombre de réalisations générées (a) pour h_{15} et (b) pour h_{25}

Les résultats obtenus avec les jeux h_{15} et h_{25} ne peuvent être comparés puisque la fonction objectif a été minimisée sur base d'un nombre différent d'observations. Ce graphique indique cependant que pour un même jeu d'observations piézométriques, la moyenne des fonctions objectifs diminue lorsque le nombre de données de ρ conditionnantes augmente, ce qui traduit une meilleure calibration du modèle d'écoulement. Les réalisations n'étant pas conditionnées par K , les valeurs de la moyenne des fonctions objectifs sont plus élevées que celles obtenues à la figure 5.16 lors d'un conditionnement par K et h ou K , ρ et h .

Une sélection des réalisations a ensuite été appliquée sur base d'un des deux critères d'ordre Cr1 et Cr2 définis précédemment. Le tableau 5.7 reprend, pour toutes les combinaisons des données considérées, le nombre m de réalisations acceptées par chacun de ces critères, au départ d'un nombre total de deux cents réalisations.

	ρ_{150}, h_{15}	ρ_{150}, h_{25}	ρ_{300}, h_{15}	ρ_{300}, h_{25}
Cr1	1	1	4	7
Cr2	17	22	41	60

Tableau 5.7 : Nombre de réalisations acceptées par chacun des critères de sélection Cr1 et Cr2

D'une manière générale, plus le nombre de données conditionnantes augmente, plus le nombre de réalisations acceptables augmente. Le critère d'ordre Cr1 étant beaucoup trop strict (cf. fin du paragraphe 5.4.4), seul le critère Cr2 a été appliqué par la suite. De plus, dans ce cas particulier, le critère Cr2 nous paraît beaucoup mieux approprié. En effet, vu que les critères d'ordre sont appliqués à $\ln\rho$ (et non $\ln K$) qui est une information indirecte sur $\ln K$ et qu'il n'existe pas de relation linéaire parfaite entre ces deux variables, une variation autour de leur valeur mesurée est tout à fait plausible.

On obtient donc m champs de conductivité hydraulique optimisés par calibration du modèle d'écoulement et non rejetés par le critère de sélection Cr2. Pour chacun de ces champs, le modèle d'écoulement a été simulé (résolution du problème direct), puis les isochrones associées à des temps de transfert de 1, 5, 10 et 20 jours ont été délimitées par *particle tracking*. Enfin, les distributions spatiales $[CAP(\underline{x}, t)]$ ont été calculées comme décrit au paragraphe 5.1.

5.5.4 Analyse des résultats et discussion

Les distributions $[CAP(\underline{x}, t)|\rho, h, Cr2]$ obtenues ont été comparées à celles du cas non-conditionnel construit au paragraphe 5.4.4 et sont présentées à la figure 5.34. Sur cette figure, on constate une relativement bonne correspondance entre $\Gamma^{(0,5)}$ et $\tau_{Réf}$, d'autant meilleure que le nombre de données conditionnantes augmente. Cette correspondance est bien meilleure que dans le cas où les champs ont été conditionnés par K et h (cf. figure 5.20), et ce malgré la moins bonne calibration du modèle d'écoulement évoquée plus haut.

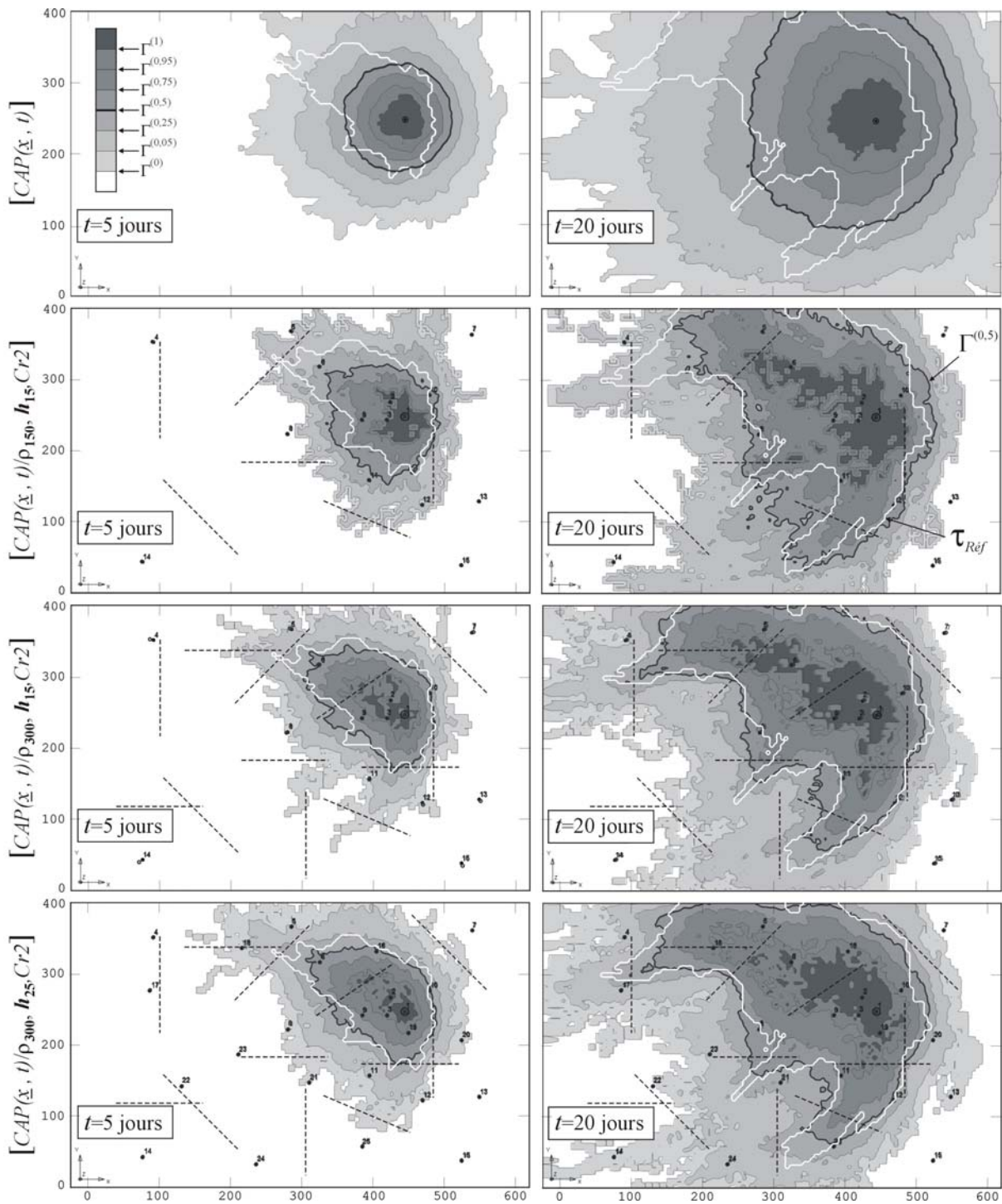


Figure 5.34 : $[CAP(\underline{x}, t) | \rho, h, Cr2]$ à 5 et 20 jours pour différentes combinaisons des données de ρ et h

Ces résultats sont quantifiés à la figure 5.35 : les graphiques 5.35a à 5.35b reportent respectivement les valeurs de w_{a1} et $w_b^{(*)}$ des distributions $[CAP(\underline{x}, t) | \rho, h, Cr2]$ en fonction des jeux de données conditionnantes pour les différents temps de transfert spécifiés, ainsi que les

(*) Vu le faible nombre de réalisations acceptées par le critère de sélection Cr2 et utilisées dans le calcul des distributions pour (ρ_{150}, h_{15}) , (ρ_{150}, h_{25}) , les quantités w_{a2} , Ω_{a2} et ws_2 correspondantes n'ont pu être évaluées.

valeurs de w_{s1} (figure 5.35a) relatives à ces distributions. Aux figures 5.35c à 5.35d, on retrouve les quantités Ω_{a1} et Ω_b associées respectivement à w_{a1} et w_b .

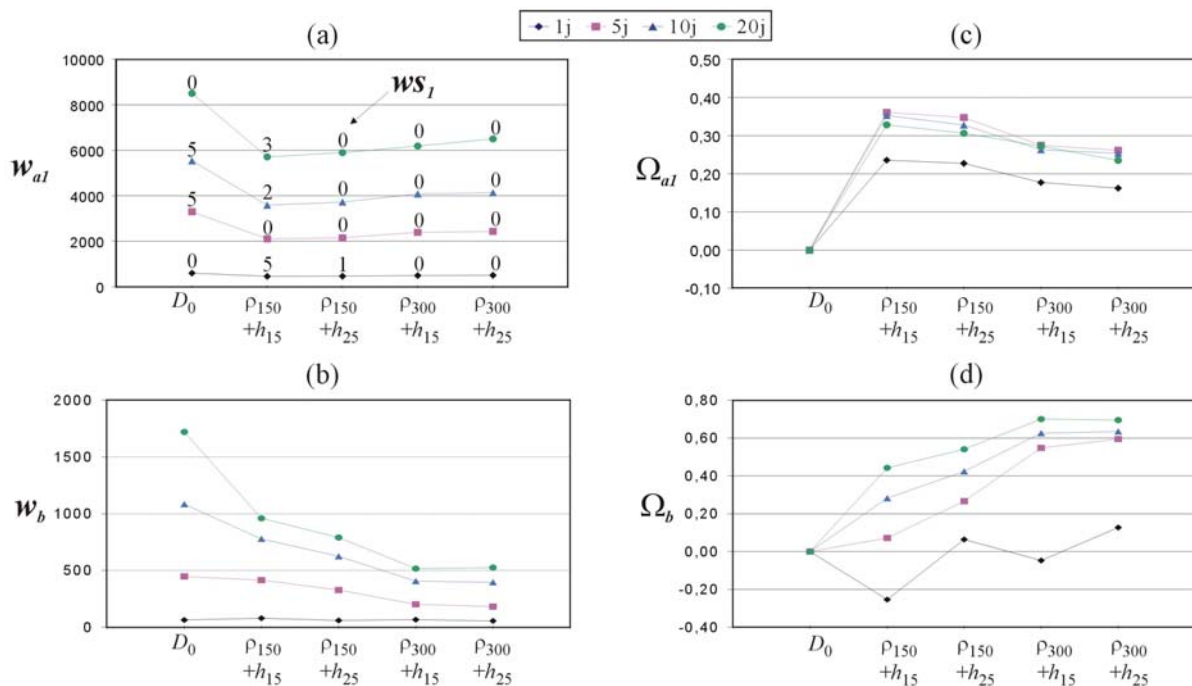


Figure 5.35 : Evolution de (a) w_{a1} , (b) w_b et (c) Ω_{a1} , (d) Ω_b en fonction des jeux de données conditionnantes ρ et h considérés pour différents temps de transfert (D_0 faisant référence au cas non-conditionnel)

L'observation des graphiques 5.35a et 5.35c montre une forte réduction de w_{a1} lorsque l'on conditionne les champs, se manifestant par une croissance de Ω_{a1} assez importante (de plus de 30% pour les temps de transfert élevés). Contrairement à ce que l'on pourrait attendre, w_{a1} augmente (et Ω_{a1} décroît) avec l'augmentation du nombre de données de ρ conditionnantes. Cette dégradation est due non seulement au fait que les données additionnelles de ρ mettent en évidence des zones de plus grande conductivité hydraulique, augmentant fortement l'incertitude dans ces régions, mais également au fait que l'estimation de la valeur de conductivité hydraulique appliquée aux extensions latérales du champ est meilleure (par rapport à celui du champ de référence) pour ρ_{150} que pour ρ_{300} (cf. paragraphe 5.5.3). L'augmentation du nombre de h ne provoque quant à elle pratiquement aucune variation des valeurs de w_{a1} (et donc de Ω_{a1}), mais améliore la description de $\tau_{Réf}$ par la distribution, puisqu'elle réduit les valeurs de w_{s1} . Ces dernières sont nulles pour toutes les distributions conditionnées par ρ_{300} .

Vu le faible nombre de réalisations restantes (après l'application du critère de sélection Cr2), utilisées dans le calcul des distributions pour les contours (ρ_{150}, h_{15}) , (ρ_{150}, h_{25}) , les quantités w_{a2} , Ω_{a2} et w_{s2} correspondantes n'ont pas été évaluées. Pour les combinaisons (ρ_{300}, h_{15}) , (ρ_{300}, h_{25}) par contre, les valeurs de w_{s2} ont été estimées : elles sont nulles pour

chaque distribution ce qui traduit une description parfaite de l'isochrone $\tau_{Réf}$ par l'isocontour $\Gamma^{(0,05)}$. Ce résultat n'a pas été observé lorsque le conditionnement a été réalisé uniquement par K et h (cf. figure 5.21).

D'après le graphique 5.35b, l'introduction des données supplémentaires de h joue un rôle plus important dans la réduction de w_b lorsque le jeu de ρ est petit. Ces observations supplémentaires n'apportent quasiment aucune réduction de w_b lorsque le jeu ρ_{300} est utilisé pour conditionner les simulations stochastiques (en raison de l'effet d'écran produit par la redondance des données) excepté pour de faibles temps de transfert. La figure 5.35d montre la grande amélioration apportée à l'écart entre $\Gamma^{(0,5)}$ et $\tau_{Réf}$ par rapport au cas non-conditionnel – près de 70% pour des temps de transfert élevés – par l'introduction de 300 données de ρ dans l'analyse, en l'absence de mesure de K . Ce chiffre est comparable à celui obtenu à partir d'un conditionnement maximum ($K_{25}, \rho_{300}, h_{25}$), c'est-à-dire lorsque le jeu K_{25} est introduit, ce qui démontre l'importance de l'apport des données de ρ dans la délimitation des zones de protection.

A nouveau, il faut insister sur le fait que les améliorations observées ici par rapport au cas non-conditionnel sont, sous-estimées. En effet, contrairement aux champs de K des cas conditionnels, ceux du cas non-conditionnel ont été construits sur base des mêmes paramètres S et λ , et de la même valeur de conductivité hydraulique uniforme appliquée aux extensions latérales du modèle que le champ de référence (valeurs "exactes").

5.6 Synthèse générale des résultats

Des jeux de données conditionnantes de conductivité hydraulique K , de hauteur piézométrique h et de résistivité électrique ρ ont été créés au chapitre 4 à partir du champ de conductivité hydraulique de référence d'un cas d'étude synthétique assimilé à un aquifère alluvial type et représentant la "réalité". La MSA a été appliquée à ce cas d'étude synthétique pour différents niveaux de connaissance, c'est-à-dire en combinant divers types de données conditionnantes. Deux situations différentes ont été distinguées selon que les champs de conductivité hydraulique aient été conditionnés ou non par les données de la variable principale K . Pour chacune de ces deux situations, plusieurs combinaisons des jeux de données ont été testées et comparées. Dans chacun des cas de figures, une quantification des résultats a été réalisée en terme (a) de l'extension spatiale de l'incertitude sur la localisation de la zone de protection, (b) de la déviation de l'isocontour $\Gamma^{(0,5)}$ par rapport à l'isochrone de référence $\tau_{Réf}$ et (c) de la description de $\tau_{Réf}$ par les isocontours $\Gamma^{(0)}$ et $\Gamma^{(0,05)}$.

Au sein de la MSA, différents choix ont été réalisés :

- Parmi les techniques de seuillage proposées, c'est la technique TS_{fmin} qui a été choisie. En effet, cette technique ne dépend pas de la fréquence des données comme c'est le cas pour TS_{unif} et TS_{gauss} , mais de leur variabilité. De plus, les distributions obtenues par cette technique décrivent parfaitement l'isochrone $\tau_{Réf}$, ce qui n'est pas toujours le cas pour les autres techniques, pouvant dès lors mener à une sous-protection du puits.
- Deux critères d'ordre ont été proposés sur base de l'ordre relatif des valeurs de conductivité hydraulique des classes défini par seuillage. Le critère Cr1, pour lequel cet ordre relatif doit être rigoureusement respecté, est fort strict et demande dès lors des efforts de calcul très important pour converger. Le deuxième critère d'ordre Cr2, plus souple, pour lequel des permutations entre classes voisines sont tolérées, fournit de bons résultats avec un nombre de réalisations acceptées suffisamment élevé pour converger. Les distributions calculées sur base du critère Cr2 enveloppent totalement l'isochrone $\tau_{Réf}$ et l'amélioration de la correspondance entre $\Gamma^{(0,5)}$ et $\tau_{Réf}$ est semblable à celui obtenu pour le critère Cr1, ce qui nous pousse à choisir le critère Cr2. Celui-ci se justifie par le fait que les essais de pompage fournissent un ordre de grandeur de la conductivité hydraulique et non une valeur précise, ce qui rend possible, dans des limites acceptables, une certaine fluctuation des valeurs.

Les résultats de la MSA ont permis de tirer plusieurs constatations générales. L'amplitude de l'incertitude et de l'écart entre $\Gamma^{(0,5)}$ et $\tau_{Réf}$ est affectée par le gradient général d'écoulement. Elle est beaucoup plus importante à l'amont du puits de pompage qu'à l'aval en raison de la dissymétrie du cône de rabattement qui crée une ligne de partage des eaux en aval, empêchant la croissance de l'incertitude dans cette direction. En amont, au contraire, la zone d'appel est très étalée ; les grandes distances de transfert engendrent donc une incertitude plus prononcée. On a également remarqué qu'en amont, l'extension spatiale de la distribution montre une variabilité considérable (beaucoup plus importante qu'en aval), dépendant principalement de la localisation des points de mesure – l'incertitude décroît là où de nouvelles données sont introduites –, mais dépendant également des valeurs de conductivité hydraulique. En effet, l'incertitude est plus importante dans les régions de conductivité hydraulique élevée, là où les distances parcourues sont plus grandes et donc où plus de variabilité naturelle de K est rencontrée.

Dans la première situation pour laquelle les champs de conductivité hydraulique ont été conditionnés par les données de la variable principale K , l'augmentation du nombre de données de K , de h , ou l'intégration des données de ρ , améliorent de manière générale l'extension ainsi que l'allure de la distribution spatiale. Cependant, l'introduction d'un grand nombre de données de ρ permet d'atteindre le même niveau de réduction de l'extension

spatiale de l'incertitude (représenté par une diminution des mesures de performance w_{a1} et w_{a2}) que celle obtenue par l'ajout de quelques données de K et h supplémentaires, tout en améliorant la description de l'isochrone $\tau_{Réf}$ (réduction de w_{S1} et w_{S2}). Pour tous les cas conditionnels considérés, l'isocontour $\Gamma^{(0)}$ enveloppe toujours bien l'isochrone $\tau_{Réf}$, ce qui n'est pas le cas de l'isocontour $\Gamma^{(0,05)}$ qui décrit beaucoup moins bien cet isochrone excepté lorsque les données de ρ sont introduites dans l'analyse. Ces nombreuses données de ρ fournissent bien souvent suffisamment d'information pour que l'isocontour $\Gamma^{(0,05)}$ décrive totalement l'isochrone $\tau_{Réf}$ et ce, malgré une réduction de l'extension spatiale de l'incertitude. Lorsqu'on réalise un conditionnement, l'écart entre $\Gamma^{(0,5)}$ et $\tau_{Réf}$ est fortement réduit par rapport au cas non-conditionnel. Toutefois, il n'est pas très sensible à l'apport de K ou de h supplémentaires (excepté pour un temps de transfert élevé). Par contre, l'introduction de ρ apporte une amélioration considérable. La réduction de l'écart entre $\Gamma^{(0,5)}$ et $\tau_{Réf}$ (représenté par une diminution de w_b) est alors bien plus importante que celle obtenue en ne prenant pas en compte les données de ρ . En raison d'un effet d'écran provoqué par la redondance des données, l'information apportée par l'introduction des données de ρ est plus importante lorsque le nombre de K et de h est faible.

L'objectif de ce travail étant de développer une méthode applicable à des aquifères réels, en fonction des données disponibles, il nous a paru intéressant d'envisager la situation où le nombre de données de conductivité hydraulique est trop faible pour que les propriétés statistiques spatiales de ce paramètre puissent être inférées de l'analyse géostatistique, ce qui est souvent le cas sur les sites réels. Les champs de conductivité hydraulique ont dès lors été conditionnés uniquement par les données de ρ et de h . Les résultats sont très encourageants : les distributions spatiales sont meilleures que celles obtenues dans la première situation en ne conditionnant ces champs que par K et h , non seulement en terme d'extension de l'incertitude, mais surtout en terme de correspondance entre $\Gamma^{(0,5)}$ et $\tau_{Réf}$.

De manière générale, dans la MSA, l'amélioration des résultats par l'apport de données conditionnantes provient d'une meilleure estimation des champs de conductivité hydraulique. En effet, ces données additionnelles permettent :

- de diminuer la variance de krigeage lors de la génération des simulations stochastiques conditionnelles (conditionnement par K ou par ρ), ou co-conditionnelles (conditionnement par K et ρ) ;
- d'optimiser les valeurs des zones de conductivité hydraulique par la calibration du modèle d'écoulement (conditionnement par h) ;
- d'éliminer les champs de conductivité hydraulique considérés comme irréalistes d'un point de vue géologique (application du critère de sélection Cr2) ;