

N° d'ordre: 00000

THÈSE

présentée en vue de l'obtention du grade de
DOCTEUR DE L'UNIVERSITÉ de STRASBOURG

devant l'Université de Strasbourg

pour obtenir

le grade de : DOCTEUR DE L'UNIVERSITÉ DE STRASBOURG
Mention SCIENCE DE LA TERRE

par

Alexandre ABELLAN

Équipe d'accueil : IMFS
École Doctorale : Institut de Mécanique des Fluides de Strasbourg
Composante universitaire : STRASBOURG

Titre de la thèse :

Construction de fonctions objectifs
&
Modélisation a priori de l'apport de mesures en géosciences :
une approche statistique

À soutenir le 14 Décembre 2009 devant la commission d'examen

M. :	Ghislain	DE MARSILY	Président
MM. :	Frédéric	DELAY	Rapporteurs
	Michel	CARA	
MM. :	Benoît	NETINGER	Examineurs
M. :	Philippe	ACKERER	Directeur de thèse

Remerciements

Mes plus vifs remerciements vont aux personnes qui ont contribué au bon déroulement et à l'aboutissement de cette thèse. En premier lieu, je tiens à exprimer ma gratitude à Benoit Noëtinger, Professeur et directeur associé à l'IFP et à Philippe Ackerer, directeur de Recherche du LHyGeS qui m'ont donné la chance et l'opportunité de réaliser ces travaux et de les mener à bien dans des conditions exceptionnelles. Je retiens leurs expériences ainsi leurs conseils avisés et la disponibilité constante dont ils ont toujours fait preuve.

Mes remerciements et ma profonde gratitude vont également aux membres du Jury. J'exprime ainsi toute ma reconnaissance à Monsieur G. De Marsily, Professeur émérite et membre de l'Académie des Sciences de Paris pour avoir accepté d'examiner mon travail et le remercie pour sa lecture attentive et toutes ses critiques ainsi que de l'intérêt réel et sincère qu'il a manifesté sur la teneur de ce rapport.

Je suis très honoré que Monsieur F. Delay, Professeur de l'université de Poitier et Monsieur M. Cara, directeur de l'École et Observatoire des Sciences de la Terre, aient accepté d'être rapporteur de ce travail.

C'est avec sincérité que je salue les membres du département Ingénierie de Réservoir et Simulation des écoulements en milieux poreux et tout particulièrement F. Roggero pour son accueil et ses conseils avisés, Y. Gautier pour toutes nos longues discussions intéressantes, I. Zabalza pour son soutien et sa collaboration scientifique fructueuse, M. Feraille pour ses idées et remarques fort instructives.

Il me serait impossible de ne pas saluer Mme et M. Knossow et ma famille pour le soutien immense et les encouragements qu'ils n'ont cessé de me prodiguer tout au long de cette entreprise.

Enfin, je remercie ma femme, Myriam pour l'écoute, l'engagement et la patience qu'elle a du acquérir pendant ces trois années. Ce travail de thèse lui est dédié.

Résumé

La caractérisation de réservoir ainsi que la quantification des incertitudes sont des enjeux majeurs en ingénierie de réservoir. L'une des difficultés vient de l'impossibilité pratique de disposer d'une description exhaustive des roches souterraines, ce qui impose une analyse probabiliste du réservoir. Dans un premier temps, il est donc indispensable de déterminer un cadre théorique et un cadre pratique permettant d'estimer *au mieux* les paramètres du modèle tout en fournissant des outils de calcul efficaces. Dans un second temps, l'intégration de données observées dynamiquement ou *calage d'historique* se fait au moyen d'une *fonction objectif* qui mesure l'écart entre les données de production simulées par le modèle numérique de réservoir et les données réelles mesurées sur le terrain. Sa minimisation aboutit à l'estimation des paramètres du modèle sans pour autant être en accord avec la géologie présente. Pour éviter toute inconsistance, on intègre avant l'étude, toute l'information *a priori* disponible, dans la fonction objectif. Le cadre bayésien retenu permet de définir cette nouvelle fonction objectif de façon naturelle. Dans un troisième temps, l'introduction de la notion d'entropie en probabilité et de la divergence de Kullback-Leibler permettent de donner un sens précis à la notion d'information contenue dans les mesures. Ainsi, l'étude des variations par rapport au nombre d'observations permet de quantifier le gain d'information sur la caractérisation du réservoir. Des tests ont été effectués dans le cadre du krigeage, de l'interprétation des essais de puits et de l'écoulement diphasique montrant la pertinence de l'approche.

Mots clés : Théorie bayésienne, information de Fisher, divergence de Kullback-Leibler, entropie de Shannon, optimisation non-linéaire, calage d'historique, problème inverse.

Résumé Étendu

Construction de fonctions objectifs

&

Modélisation *a priori* de l'apport de mesures en Géosciences :
une approche statistique

À la différence de nombreux domaines industriels où les ingénieurs définissent les systèmes sur lesquels ils travaillent, en ingénierie de réservoir pétrolier ceux-ci leur sont imposés. Un gisement d'huile ou de gaz n'est pas un objet dont tous les *éléments* ont été assemblés. De plus, les dimensions du réservoir de l'ordre du kilomètre carré en surface et de l'ordre du mètre en profondeur en rendent la caractérisation très difficile. Actuellement, un des défis en ingénierie pétrolière est d'arriver à donner de meilleures descriptions de l'architecture interne des réservoirs afin d'en prévoir le comportement dynamique et ainsi, d'obtenir des prévisions de production fiables ce qui permet la mise en place de stratégies optimisant la récupération d'hydrocarbure à moindre risque et à moindre coût.

En pratique, l'ingénieur de réservoir recourt à des simulations numériques d'écoulement qui, pour différents ensembles de paramètres décrivant un état initial ou état *a priori* du réservoir, fournissent les profils de production correspondant. Les données *réelles* observées telles que la pression ou les débits au puits d'huile, de gaz, d'eau sont le résultat d'un processus physique complexe dont la description implique la modélisation des écoulements polyphasiques en milieu poreux. Ce processus est décrit à l'aide d'un système d'équations de transport non-linéaires fortement couplées. Prises en des lieux fixes et à des instants donnés de l'évolution du gisement, ces données sont disponibles car directement reliées à l'objet même de l'exploration pétrolière. Le comportement du réservoir apparaît généralement comme une boîte noire et ces données apportent une information importante qu'il faut judicieusement *analyser* et *intégrer* dans la caractérisation du réservoir. Elles permettront l'estimation des paramètres du modèle de simulation gouvernant le phénomène physique et par la suite de construire un ou plusieurs modèles prédictifs.

L'ingénieur de gisement est donc amené à modifier progressivement son modèle ini-

tial afin de le calibrer à l'ensemble des données disponibles. Cette opération est très coûteuse en temps, et repose bien souvent sur une approche empirique faisant largement appel à l'expérience de l'ingénieur. Des techniques permettant d'automatiser ce processus font l'objet de recherches très actives. On cherche ainsi à caractériser le réservoir en estimant les paramètres à partir des données observées et on parle alors de résolution de *problèmes inverses*. Celle-ci est difficile et délicate car l'analyse de ces données nécessite un *post-traitement* complexe avant leur utilisation. Il existe de plus un problème dit de support de mesures. En effet, en caractérisation de réservoir, les observations sont faites à plusieurs échelles. On peut noter par exemple que l'échelle d'observation de la pression aux puits est différente de celle d'une perméabilité observée en laboratoire. Le choix de cette échelle d'observation sera à prendre en compte comme un paramètre extérieur au phénomène physique. Enfin, la non-linéarité et la non-inversibilité des opérateurs mathématiques accentuent la difficulté de résolution du problème inverse.

C'est dans ce contexte d'incertitudes que nous cherchons à mettre en place un formalisme théorique permettant d'inférer les paramètres *explicatifs* du modèle et de quantifier l'apport des observations sur cette inférence.

Le formalisme *bayésien* est particulièrement bien adapté pour prendre en compte à la fois les notions de problème inverse et d'incertitudes. On suppose que le réservoir peut être décrit à l'aide de M paramètres regroupés dans le vecteur $\mathbf{m} \in \mathcal{M} \subseteq \mathbb{R}^M$.

Si l'on dispose d'une *information* dite *a priori* sur ces paramètres avant même d'effectuer une mesure et qu'il est possible de définir à l'aide de cette information une distribution de probabilité *a priori* notée $\pi(\mathbf{m})$, alors un choix peut être fait de représenter les paramètres du modèle à l'aide de cette distribution¹.

On suppose alors que l'on effectue des mesures physiques sur le système (observations de pressions aux puits, débits de production, diagraphies...). On dispose alors d'un vecteur $\mathbf{d} \in \mathcal{D} \subseteq \mathbb{R}^N$ de dimension N collectant ces observations.

Enfin, on suppose l'existence d'une fonction \mathcal{F} de $\mathcal{M} \rightarrow \mathcal{D}$ qui à tout vecteur $\mathbf{m} \in \mathcal{M}$ associe le vecteur $\mathbf{d} \in \mathcal{D}$. Soit encore de façon mathématique :

$$\begin{aligned} \mathcal{M} &\longmapsto \mathcal{D} \\ \mathbf{m} &\longmapsto \mathbf{d} = \mathcal{F}[\mathbf{m}] \end{aligned}$$

Cette fonction permet alors de définir un modèle paramétrique, c'est-à-dire une relation reliant les réponses \mathbf{d} du modèle aux paramètres \mathbf{m} . Il est alors possible d'explicitement la distribution de probabilité des observations conditionnellement aux paramètres. Cette distribution désignée par la fonction de vraisemblance $\rho(\mathbf{d}|\mathbf{m})$ contient l'information *théorique* disponible apportée par le modèle.

1. En effet, c'est un choix subjectif de représenter les paramètres du modèle par des variables \mathbf{m} , par une valeur moyenne $\langle \mathbf{m} \rangle$ ou par une distribution $\pi(\mathbf{m})$ sachant que ces variables sont observées par une réponse moyennée sur une fenêtre temporelle avec des conditions initiales peu précises.

La distribution *a posteriori* définie par la formule de Bayes s'interprète alors comme la combinaison entre l'information *a priori* et l'information *théorique*. Cette distribution est définie par la formule suivante :

$$\rho(\mathbf{m}|\mathbf{d}) = \frac{\rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m})}{\int_{\mathcal{M}} \rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m})d\mathbf{m}}$$

L'intérêt de notre approche est de déterminer si cette nouvelle distribution de probabilité permettant d'estimer les paramètres \mathbf{m} du modèle après avoir effectué les observations \mathbf{d} est différente de la distribution *a priori*. Le problème revient à quantifier l'apport des observations \mathbf{d} sur la distribution *a posteriori*. Cet apport d'information contenue dans l'observation sur les paramètres est défini en *mesurant* l'écart entre les distributions *a posteriori* et *a priori*. On calcule la quantité d'information transmise entre l'état *a priori* du système et son état *a posteriori* :

$$\pi(\mathbf{m}) \longrightarrow \rho(\mathbf{m}|\mathbf{d})$$

Selon cet *écart*, on pourra déterminer si l'observation \mathbf{d} est significative ou non dans l'étude. C'est-à-dire, si les paramètres à estimer seront modifiés par l'intégration de cette observation.

Le critère d'information permettant de calculer la *distance* entre les deux distributions est basé sur la notion d'entropie mutuelle qui mesure la dépendance statistique entre variables aléatoires. On peut quantifier l'apport d'information d'une nouvelle mesure en évaluant l'entropie de la distribution de probabilité *a posteriori* relative à la distribution de probabilité *a priori*. La divergence de Kullback-Leibler calcule cette entropie relative et est définie par :

$$K(\rho(\mathbf{m}|\mathbf{d})||\pi(\mathbf{m})) = \int_{\mathcal{M}} \rho(\mathbf{m}|\mathbf{d}) \log \frac{\rho(\mathbf{m}|\mathbf{d})}{\pi(\mathbf{m})} d\mathbf{m} \quad (1)$$

Puisque l'on s'intéresse à l'apport potentiel d'un processus de mesure et donc que les observations ne sont pas disponibles *au préalable*, il convient de définir un critère d'information moyenne. Pour ce faire, on prend l'espérance mathématique de la divergence de Kullback-Leibler sur l'espace des observations. Le critère d'information moyenne est alors défini par la fonctionnelle de Zellner :

$$\langle I(\pi(\mathbf{m})) \rangle = \int_{\mathcal{D}} \sigma(\mathbf{d}) \int_{\mathcal{M}} \rho(\mathbf{m}|\mathbf{d}) \log \frac{\rho(\mathbf{m}|\mathbf{d})}{\pi(\mathbf{m})} d\mathbf{m} d\mathbf{d} \quad (2)$$

où la distribution marginale $\sigma(\mathbf{d})$ est définie par :

$$\sigma(\mathbf{d}) = \int_{\mathcal{M}} \rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m}) d\mathbf{m} \quad (3)$$

Dans le cadre bayésien, on peut redéfinir le critère d'information moyenne à l'aide de l'entropie de Shannon \mathcal{E} telle que :

$$\langle I(\pi(\mathbf{m})) \rangle = \mathcal{E}\{\sigma(\mathbf{d})\} - \mathbb{E}^{\pi}[\mathcal{E}\{\rho(\mathbf{d}|\mathbf{m})\}] \quad (4)$$

On peut interpréter l'information moyenne comme la différence entre l'information fournie par la distribution des observations et l'information théorique moyenne espérée *a priori*.

Si il est possible de linéariser l'opérateur \mathcal{F} , alors on obtient une relation (linéaire) simple entre les observations et les paramètres du système telle que :

$$\mathbf{d} = \mathbf{F} \cdot \mathbf{m} + \boldsymbol{\nu} \quad (5)$$

où $\boldsymbol{\nu}$ symbolise l'erreur faite sur l'évaluation des observations. Généralement, on suppose que $\boldsymbol{\nu}$ suit une loi gaussienne de moyenne nulle et de covariance \mathbf{C}_d telle que $\boldsymbol{\nu} \sim \mathcal{N}(0, \mathbf{C}_d)$. La matrice \mathbf{F} est considérée indépendante de \mathbf{m} et de l'inférence. On suppose généralement que le paramètre \mathbf{m} représentant par exemple la perméabilité du milieu suit une loi gaussienne de moyenne \mathbf{m}_p et de covariance \mathbf{C}_p connues *a priori*. Celle-ci est dite *conjuguée* à la distribution de vraisemblance $\mathbf{d}|\mathbf{m}$. La distribution *a posteriori* est alors gaussienne de moyenne \mathbf{m}_q et de covariance \mathbf{C}_q telle que $\mathbf{m}|\mathbf{d} \sim \mathcal{N}(\mathbf{m}_q, \mathbf{C}_q)$. Ces considérations nous permettent de définir explicitement le critère d'information moyenne :

$$\langle I(\pi(\mathbf{m})) \rangle = \frac{1}{2} \log \frac{|\mathbf{C}_p|}{|\mathbf{C}_q|} = \frac{1}{2} \log \left| \mathbf{C}_p \cdot \mathbf{C}_q^{-1} \right| \quad (6)$$

Cette dernière expression représente un rapport entre les volumes de corrélations obtenus avant et après avoir effectué une mesure. Le calcul de cette quantité peut devenir important lors de l'inversion de la matrice de covariance *a posteriori*. Il existe une bijection entre l'espace des paramètres et l'espace des observations et on préfère exprimer ce critère dans l'espace des observations. Après calculs algébriques, on montre que le critère est égal à :

$$\langle I(\pi(\mathbf{m})) \rangle = \frac{1}{2} \log \left| \mathbb{I}_N + \mathbf{C}_d^{-1} \cdot \mathbf{F} \mathbf{C}_p \mathbf{F}^t \right| \quad (7)$$

où \mathbb{I}_N est la matrice identité d'ordre N , avec N représente le nombre total de mesures.

C'est cette formule que nous allons utiliser par la suite. Dans les différents exemples que nous avons traités, l'opérateur de modélisation permet d'estimer le champ de perméabilité.

Dans le premier exemple, nous étudions le cas de l'estimation du champ de perméabilité lorsque celui-ci est modélisé par krigeage. Ainsi, cherchons nous à définir l'échantillon spatial optimal des observations permettant d'estimer au mieux le champ de perméabilité. Dans notre deuxième exemple, l'estimation de la perméabilité est obtenue par la perméabilité apparente qui sera calculée au moyen d'un essai de puits (observation de la pression au puits). Notre but est alors de définir l'échantillon temporel optimal d'acquisition de pression. Enfin, dans notre dernier exemple, nous estimons la perméabilité à l'aide d'acquisition sismique. Pour simplifier le traitement des cubes de données sismiques, nous étudions l'évolution des lignes de courant. Dans le cadre diphasique, nous avons cherché à définir un échantillon spatio-temporel optimal. Nous allons présenter plus en détail ces trois exemples.

Exemple I

Le premier exemple considéré est celui du krigeage, qui est une méthode d'interpolation spatiale issue de la géostatistique. On procède à des mesures en différents emplacements de l'espace (en pratique des puits) et les techniques de krigeage permettent d'interpoler linéairement les données spatiales en des positions autres, fournissant ainsi une estimation de l'ensemble du champ de perméabilité.

Le nombre d'observations étant nettement inférieur au nombre de paramètres, on exprime dans ce cas l'information moyenne dans l'espace des observations. On suppose les N observations indépendantes. On suppose que la fonction de covariance d'erreurs sur les observations est définie par $\mathbf{C}_d = \sigma_d^2 \cdot \mathbb{I}_N$ où σ_d^2 est la variance d'erreur des observations supposée connue. On considère le variogramme gaussien pour la covariance *a priori* de la perméabilité, le problème étant alors de déterminer les N positions optimales $\mathbf{x}_1, \dots, \mathbf{x}_N$ des observations maximisant le critère d'information moyenne :

$$\langle I(\pi(\mathbf{m})) \rangle = \frac{1}{2} \log \left| \mathbb{I}_N + \frac{\sigma_0^2}{\sigma_d^2} \mathbf{F} \cdot \left(\exp - \left(\frac{\mathbf{x}_i - \mathbf{x}_j}{\lambda_c} \right)^2 \right)_{i,j} \cdot \mathbf{F}^t \right| \quad i, j \in [1, N] \quad (8)$$

où l'opérateur $\mathbf{F} \in \mathbb{R}^{N \times M}$ est défini tel que $\mathbf{F}_{ij} = \delta_{ij}$, λ_c la longueur de corrélation et σ_0 la variance du champ de perméabilité considéré. Des résultats numériques sont donnés pour différentes longueurs de corrélation.

On a aussi pu quantifier la notion de redondance des mesures. Dans le cas du krigeage, lorsque les mesures sont indépendantes, ce qui veut dire en pratique que les puits d'observation sont espacés d'au moins deux longueurs de corrélation, la quantité d'information varie linéairement avec le nombre de puits. En augmentant le nombre de mesures, on comprend intuitivement qu'il existe un nombre d'observations tel que l'ajout d'une nouvelle information a un apport moyen marginal. L'information moyenne $\langle I \rangle$ devient proportionnelle à $\ln(N)$, N étant le nombre total de mesures. Ainsi l'apport d'une mesure est inversement proportionnel au nombre total de mesures prises en compte.

Exemple II

Les résultats du krigeage étant cohérent avec les résultats connus, nous avons étendu notre approche à l'interprétation des essais de puits, sous l'hypothèse classique de fluide faiblement compressible. En se donnant un budget de N mesures de pression au puits sur une fenêtre temporelle fixée, on cherche l'échantillonnage maximisant le critère d'information moyenne. L'équation de diffusion au puits déduite de la loi de Darcy modélise l'écoulement monophasique d'un fluide faiblement compressible et fournit l'évolution de la pression au cours du temps sous l'effet d'un terme source Q . La pression P est ici la

seule inconnue du problème direct d'évolution. Le système est donné par :

$$\begin{cases} \phi \mu c_T \partial_t P(\mathbf{r}, t) = \nabla \cdot (k(\mathbf{r}) \nabla P(\mathbf{r}, t)) + Q \frac{\mu}{h} \delta(\mathbf{r} - \mathbf{r}_w) H(t) \\ \forall \mathbf{r}, P(\mathbf{r}, t = 0) = 0 \\ \forall t, P(r \rightarrow +\infty, t) = 0 \end{cases}$$

où μ est la viscosité du fluide, ϕ la porosité du milieu, c_T la compressibilité totale du fluide, $\delta(\mathbf{r} - \mathbf{r}_w)$ est la fonction de Dirac au puits en \mathbf{r}_w et $H(t)$ est la fonction de Heaviside en temps. Les spécialistes de l'interprétation des essais de puits déduisent des observations de la pression des propriétés du champ de perméabilité $k(\mathbf{r})$.

De façon à simplifier la formulation du problème, il est utile d'introduire la perméabilité apparente du milieu par la relation suivante :

$$k_{app}(t) = \frac{\mu Q}{4\pi h t \frac{dP(\mathbf{r}=\mathbf{r}_w, t)}{dt}} \quad (9)$$

où $P(\mathbf{r} = \mathbf{r}_w, t)$ est la pression au puits à l'instant t . Après linéarisation du problème, on définit explicitement le critère d'information moyenne à l'aide d'un noyau pondérant la perméabilité. La fonction de covariance des perméabilités entre deux temps d'observations t et τ est définie par :

$$\langle k_{app}(t), k_{app}(\tau) \rangle = \int_{\mathbb{R}^2} d\xi \frac{G(\xi, \tau)}{2\pi D_0 \tau} \int_{\mathbb{R}^2} d\mathbf{r} \frac{G(\mathbf{r}, t)}{2\pi D_0 t} \langle k(\mathbf{r}), k(\xi) \rangle \quad (10)$$

où G est le noyau pondérateur défini par Oliver et :

$$\begin{aligned} \langle k(\mathbf{r}), k(\xi) \rangle = k_G^2 \cdot \exp & \left(\frac{1}{2} \frac{\sigma_0^2}{\exp \frac{r^2}{\ell_c^2}} \mathcal{I}_0 \left(\frac{r^2}{\ell_c^2} \right) + \frac{1}{2} \frac{\sigma_0^2}{\exp \frac{\xi^2}{\ell_c^2}} \mathcal{I}_0 \left(\frac{\xi^2}{\ell_c^2} \right) \right) \cdot \\ & \left(\exp \left(\frac{\sigma_0^2}{\exp \frac{r^2 + \xi^2}{2\ell_c^2}} \mathcal{I}_0 \left(\frac{r\xi}{\ell_c^2} \right) \right) - 1 \right) \end{aligned}$$

avec \mathcal{I}_0 la fonction de Bessel modifiée d'ordre 0. On obtient des résultats numériques cohérents avec l'intuition. En effet, il apparaît que les temps d'observations optimaux se concentrent aux temps courts et sont de plus en plus éloignés aux temps longs ce qui reflète que l'essentiel de l'information fournie par un essai de puits est surtout concentrée au puits.

Exemple III

On a finalement cherché à appliquer cette approche à l'intégration de données de sismique 3D. Il est important de disposer de techniques permettant de planifier au départ l'organisation des campagnes d'acquisition, eu égard aux investissements réalisés par les compagnies dans le domaine.

Pour simplifier le problème, et ne pas avoir à traiter les aspects strictement sismiques, on suppose que l'on est capable d'acquérir N cartes de saturation à N temps différents. L'inconnue cherchée est encore le champ de perméabilité hétérogène du milieu. On a utilisé les approches lagrangiennes au premier ordre pour la simulation des fronts de saturation afin de déterminer, dans une approche 4D, la fréquence optimale d'acquisition. L'approche perturbative nous fournit l'approximation au premier ordre de l'évolution du front de saturation dans un milieu supposé hétérogène.

On suppose que le champ de vitesse \mathbf{u} est stationnaire et constant en espace. On suit l'hypothèse selon laquelle la perturbation de la vitesse locale est la somme de deux fluctuations qui dérivent d'une part du couplage visqueux et d'autre part de l'hétérogénéité propre du milieu. Ainsi, avons nous relié la perturbation du front de saturation à la perturbation de la vitesse, elle-même exprimée en fonction de la fluctuation de la perméabilité et de la fonction de Green associée à notre problème. On obtient dans l'espace de Fourier, l'évolution d'un mode α de perturbation de position du front x_α de saturation au temps t . Elle est donnée par :

$$\delta x_\alpha(t) = \frac{f'(S_f)}{\phi} e^{A|\alpha|t} \int_{-\infty}^t d\tau e^{-A|\alpha|\tau} \delta u_{bx\alpha}(c\tau) \quad (11)$$

où f est le flux fractionnaire total, S_f la saturation au front, c la vitesse du front de saturation dans un milieu homogène, A un coefficient dépendant de la mobilité², ϕ la porosité du milieu et $\delta u_{bx\alpha}$ un mode de la perturbation de la vitesse.

En considérant un champ de perméabilité isotrope et pour des temps suffisamment longs³, on déduit de ce qui précède la fonction de covariance entre deux fronts de saturation à deux temps différents dans l'espace réel :

$$C(h, \Delta t) = \frac{k_0 \ell_c^2}{4} \int_0^1 d\beta \beta^{d/2-2} e^{-\theta\beta} \left(\frac{1+A^2}{1-A^2} (f_{\ell_c^2/4}(\beta) - 1) - g_{h, \Delta t}(\beta) \right) \quad (12)$$

où h est la distance entre les deux fronts, Δt la différence entre les temps d'acquisition, f et g sont deux fonctions définies chapitre 8. La quantité $\frac{k_0 \ell_c^2}{4}$ est définie en fonction de la fonction de covariance du champ de perméabilité dans l'espace de Fourier. Le choix de la fonction de covariance est dépendant de l'approche de linéarisation du problème. On note que dans une approche stochastique permettant de calculer les moments longitudinaux du déplacement du front de saturation, nous avons obtenu une forme différente

2. Le coefficient A est négatif dans le cas stable.

3. L'analyse est faite aux temps longs car cela permet de négliger la zone de raréfaction due au caractère diphasique immiscible de l'écoulement.

et plus complexe de la fonction de covariance.

Conclusions et perspectives

La théorie de l'information nous est apparue comme un outil indispensable permettant de quantifier, puis d'optimiser la stratégie de mesure. Les premiers résultats relatifs au krigeage ont été concluants et nous ont permis d'aborder des domaines moins intuitifs comme notamment ceux de l'essai de puits et de l'acquisition sismique. En fonction d'un nombre fini d'observations spatiales ou temporelles, nous avons pu définir un échantillon optimal grâce à la maximisation de l'information moyenne.

Dans notre étude, nous avons linéarisé les opérateurs mathématiques en les perturbant et utilisé une formulation analytique. Une perspective immédiate serait d'étudier l'évolution de l'information (ou de l'information moyenne) avec les opérateurs non-linéarisés. On pourrait envisager de définir une fréquence d'acquisition pour surveiller l'évolution d'un système plus complexe (triphase) ou l'évolution de l'injection de CO₂ dans un réservoir.

Liste des communications :

- *Designing Optimal Data acquisition schemes using Kullback-Leibler Divergence within the Bayesian Framework*, ECMOR XI, 2008, Bergen, Norway
- *Quantification et optimisation de l'information apporté par de nouvelles données*, Séminaire scientifique à GDF (Gaz de France), July, 2007
- *History matching using Kullback-Leibler divergence*, Poster presentation, Chemical and Materials Engineering Graduate Research Symposium, Energy Day at University of Alberta, Canada
- *Optimizing History Matching in Oil Reservoir Using Kullback Leibler Distance in Bayesian Framework*, SIAM Conference on Mathematics & Computational Issues in the Geosciences, March 2007, Santa Fe, New Mexico

Liste des publications :

- A. Abellan et B. Nøttinger, *Optimizing subsurface field data acquisition using information theory*, Mathematical Geosciences, (à paraître 2009-2010)

Table des matières

Remerciements	i
Résumé	iii
Résumé Étendu	v
Table des matières	xiii
1 Introduction	1
2 Éléments d'ingénierie de réservoir	7
1 Ingénierie de réservoir : Objectifs et enjeux	7
2 Modélisation physique des écoulements	8
2.1 Modélisation statique du réservoir : modèle géologique	8
2.2 Modélisation dynamique	9
3 Caractérisation des réservoirs par problèmes inverses	10
3.1 Calage d'historique	10
3.2 Identification des paramètres et problème inverse	11
3 Problèmes inverses en ingénierie de réservoir	13
1 Formalisation du problème inverse déterministe	13
1.1 Fonction objectif et optimisation	14
1.2 Spécificités des problèmes inverses en ingénierie de réservoir	14
2 Formalisation du problème inverse probabiliste	15
2.1 Principe et intérêt de la Géostatistique	15
2.2 Distribution de probabilité	16
2.3 Réalisation de champs géostatistiques	17
2.4 Fonction objectif et optimisation	18
4 Problèmes inverses et mesure de l'information	21
1 États d'information du système	22
1.1 Problème direct	22
1.2 Résolution du problème inverse dans le cadre probabiliste	23
2 Cadre probabiliste	24
2.1 Conjonction d'informations	24
2.2 Gain d'information et apprentissage	24
3 Comportement asymptotique de l'information moyenne	27
3.1 Fonctionnelle d'information	27

3.2	Critères d'information	28
3.3	Cas d'un paramètre unidimensionnel	29
3.4	Cas d'un paramètre multidimensionnel	30
5	Modèle linéarisé Gaussien	33
1	Modèle de régression	34
2	Noyau de covariance	34
3	Résolution du problème inverse linéarisé	35
3.1	Définition de la distribution <i>a priori</i>	35
3.2	Linéarisation autour du maximum de vraisemblance	35
3.3	Linéarisation autour de la moyenne <i>a priori</i>	36
4	Quantité d'information	37
4.1	Apport d'information	37
4.2	Apport d'information moyenne	37
6	Optimisation de mesures statiques	41
1	Estimateurs du champ de perméabilité	42
1.1	Estimation par cokrigage simple	43
1.2	Estimation par maximum a posteriori	44
2	Optimisation de l'information	46
2.1	Cas asymptotique	46
2.2	Maximisation de l'information moyenne	46
2.2.1	Méthode séquentielle	46
2.2.2	Méthode globale	48
2.3	Concentration des observations	50
3	Conclusion	51
7	Optimisation des temps d'acquisition de mesures	53
1	Principe	53
2	Équation d'état	54
3	Solution analytique et linéarisation	54
3.1	Solution en milieu homogène	54
3.2	Approximation et solution numérique en milieu hétérogène	55
4	Essai de puits en milieu hétérogène	58
4.1	Propriétés stochastiques de la perméabilité apparente	59
4.2	Approche perturbative de l'équation de diffusivité	60
4.3	Fonction de covariance de la perméabilité apparente	63
5	Optimisation	65
6	Résultats numériques	67
6.1	Optimisation séquentielle	67
6.2	Optimisation globale	67
7	Conclusion	70

8	Optimisation des temps d'observation dans un cadre diphasique	73
1	État du problème	74
2	Modèle Analytique	74
3	Linéarisation du modèle et dynamique du front	77
4	Calcul de la fonction de covariance	83
5	Conclusion	84
	Conclusion	87
	A Théorie de l'Information et Formalisme Bayésien	91
	B Calcul de la distribution <i>a posteriori</i> dans le cas linéaire gaussien	99
	C Calcul de l'information apportée par une mesure gaussienne	101
	D Formule de Schur	103
	E Calcul du déterminant après agrégation de p points	105
	F Linéarisation autour du maximum de vraisemblance	109
	G Fonction de covariance de la perméabilité apparente	115
	H Simulateur d'essai de puits monophasique 3D	121
	I Fonction de covariance du front de saturation	123
	J Calcul de la mesure spectrale	129
	Bibliographie	131

Chapitre 1

Introduction

L'étude des réservoirs à hydrocarbures a pour objectif l'exploitation d'un réservoir de pétrole ou de gaz à moindre coût et à moindre risque. Les gisements les plus rentables économiquement sont les gisements les plus facilement accessibles et les plus simples à exploiter. Mais ces champs s'épuisent et sont de moins en moins nombreux. Dans les zones géopolitiquement sûres, ceux restant à exploiter sont souvent dans des zones d'accès difficile, comme par exemple, sous la mer à de grandes profondeurs. Leurs exploitations demandent des technologies très sophistiquées et coûteuses. De ce fait, la caractérisation est d'importance croissante en ingénierie puisqu'elle permet une meilleure gestion des risques dans les prises de décision.

Si dans de nombreux domaines industriels les ingénieurs définissent les systèmes sur lesquels ils travaillent, en ingénierie de réservoir, ces systèmes leur sont imposés. Un gisement d'huile ou de gaz n'est pas un objet dont tous les *éléments* ont été assemblés. De plus, les dimensions du réservoir de l'ordre du kilomètre en longueur et de l'ordre du mètre en profondeur rendent la caractérisation très difficile. Actuellement, un des défis en ingénierie pétrolière est d'arriver à donner de meilleures descriptions de l'architecture interne des réservoirs afin d'en prévoir le comportement dynamique et ainsi, d'obtenir des prévisions de production fiables. Ceci peut permettre la mise en place de stratégies optimisant la récupération à moindre risque.

En pratique, l'ingénieur de réservoir recourt à des simulations numériques d'écoulement qui, pour différents ensembles de paramètres décrivant un état initial ou état *a priori* du réservoir, fournissent les profils de production correspondant. Les données *réelles* observées telles que la pression sont le résultat d'un processus physique (à savoir des écoulements polyphasiques en milieux poreux) dont le mécanisme non-linéaire reste complexe. Prises en des lieux fixes et à des instants donnés de l'évolution du gisement, ces données sont disponibles car directement reliées à l'objet même de l'exploration pétrolière. Le comportement du réservoir apparaît généralement comme une boîte noire et ces données apportent une information importante qu'il faut judicieusement *analyser* et *intégrer* dans la caractérisation du réservoir. Elles permettront l'estimation des paramètres gouvernant le phénomène physique et par la suite de construire un ou plusieurs

modèles prédictifs.

On met ainsi en place un processus itératif qui permet de caractériser le réservoir en estimant les paramètres à partir des données observées. On parle alors de résolution de problèmes inverses. Ce *problème inverse* est difficile et délicat car ces données nécessitent souvent un *post-traitement* avant leur utilisation. De plus, il existe un problème dit de support de mesures où l'échelle d'observation d'une pression aux puits est différente de celle d'une perméabilité observée en laboratoire et de celle utilisée lors de la simulation. Le choix de l'échelle d'observation est un paramètre extérieur au phénomène physique qu'il faudra prendre en compte. Enfin, la non-linéarité et la non-inversibilité des opérateurs mathématiques accentuent la difficulté de résolution. C'est dans ce contexte d'incertitudes que nous cherchons à mettre en place un formalisme théorique permettant d'inférer les paramètres *explicatifs* du modèle.

Pour ce qui est de la résolution de problèmes inverses, le cadre *bayésien* s'est imposé comme un outil particulièrement pertinent dans nos travaux. Supposons que l'on dispose d'un vecteur $\mathbf{d} \in \mathcal{D}$ de dimension N symbolisant les observations et considérons que ces observations sont décrites grâce à M paramètres regroupés dans le vecteur $\mathbf{m} \in \mathcal{M}$. Ainsi, supposons nous l'existence d'une fonction \mathcal{F} de $\mathcal{M} \rightarrow \mathcal{D}$ qui à tout vecteur \mathbf{m} associe le vecteur \mathbf{d} .

$$\begin{aligned} \mathcal{M} &\mapsto \mathcal{D} \\ \mathbf{m} &\longrightarrow \mathbf{d} = \mathcal{F}[\mathbf{m}] \end{aligned}$$

Si l'on dispose d'une certaine *information* dite *a priori* sur ces paramètres avant même d'effectuer une mesure et qu'il est possible de définir à l'aide de cette information une distribution de probabilité *a priori* notée $\pi(\mathbf{m})$, alors on peut choisir de représenter les paramètres du modèle à l'aide de cette distribution¹.

De plus, disposant d'un modèle paramétrique, c'est-à-dire d'une relation reliant les réponses \mathbf{d} du modèle aux paramètres \mathbf{m} , il est possible d'explicitier la distribution de probabilité des observations conditionnellement aux paramètres. Cette distribution désignée par la fonction de vraisemblance $\rho(\mathbf{d}|\mathbf{m})$ contient l'information disponible apportée par le modèle.

En combinant l'information *a priori* et l'information *théorique*, nous pouvons construire la distribution *a posteriori* $\rho(\mathbf{m}|\mathbf{d})$ à l'aide de la formule de Bayes :

$$\rho(\mathbf{m}|\mathbf{d}) \propto \rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m})$$

où \propto désigne l'égalité au facteur de normalisation près.

1. En effet, c'est un choix subjectif de représenter les paramètres du modèle par des variables \mathbf{m} , par une valeur moyenne $\langle \mathbf{m} \rangle$ ou par une distribution $\pi(\mathbf{m})$ sachant que ces variables sont observées par une réponse moyennée sur une fenêtre temporelle avec des conditions initiales peu précises.

Il est possible de quantifier l'apport d'information contenue dans l'observation \mathbf{d} sur les paramètres \mathbf{m} en *mesurant* l'écart entre les distributions *a posteriori* et *a priori* et de savoir par conséquent si cet apport d'information est significatif. On calculera alors la quantité d'information transmise entre l'état *a priori* du système et son état *a posteriori* :

$$\pi(\mathbf{m}) \longrightarrow \rho(\mathbf{m}|\mathbf{d})$$

Selon l'*écart* entre les deux distributions, on pourra déterminer si une observation est significative ou non dans l'étude et si le jeu de paramètres se voit modifié par l'intégration de cette observation.

Le critère d'information permettant de calculer la *distance* entre les deux distributions est basé sur la notion d'entropie. Cette fonction mathématique correspond à la quantité d'information contenue par une source et quantifie l'incertitude liée à un événement aléatoire. On parle alors d'entropie d'une loi de probabilité. Initialement, la notion d'entropie fut définie dans le cadre discret. Le passage au cas continu n'est pas simple car il faut utiliser une définition de l'entropie qui prend en compte la notion d'état de *non-information* où l'entropie est nulle. La notion d'entropie, tout comme la vitesse, n'est plus absolue mais relative à un état *a priori*. En d'autres termes, on est informatif soit par rapport à un état de non-information, soit par rapport à un état de référence. L'apport d'information d'une nouvelle mesure revient à évaluer l'entropie de la distribution de probabilité *a posteriori* relative à la distribution de probabilité *a priori*. La divergence de Kullback-Leibler calcule l'entropie relative entre les deux distributions de probabilité telle que :

$$K(\rho(\mathbf{m}|\mathbf{d})||\pi(\mathbf{m})) = \int_{\mathcal{M}} \rho(\mathbf{m}|\mathbf{d}) \log \frac{\rho(\mathbf{m}|\mathbf{d})}{\pi(\mathbf{m})} d\mathbf{m} \quad (1.1)$$

Dans la mesure où les observations ne sont pas disponibles *a priori*, il convient de définir un critère d'information moyenne. On prend alors l'espérance mathématique de la divergence de Kullback-Leibler sur l'espace des observations. L'information moyenne est alors définie par la mesure de l'information de Lindley [1] :

$$\langle I(\pi(\mathbf{m})) \rangle = \int_{\mathcal{D}} \sigma(\mathbf{d}) \int_{\mathcal{M}} \rho(\mathbf{m}|\mathbf{d}) \log \frac{\rho(\mathbf{m}|\mathbf{d})}{\pi(\mathbf{m})} d\mathbf{m} d\mathbf{d} \quad (1.2)$$

où la distribution des observations est définie par la distribution marginale $\sigma(d)$ telle que :

$$\sigma(\mathbf{m}) = \int_{\mathcal{M}} \rho(\mathbf{d}|\mathbf{m}) \pi(\mathbf{m}) d\mathbf{m} \quad (1.3)$$

Dans le cadre bayésien, le critère d'information moyenne est exprimé à l'aide l'entropie de Shannon \mathcal{E} tel que :

$$\langle I(\pi(\mathbf{m})) \rangle = \mathcal{E}\{\sigma(\mathbf{d})\} - \mathbb{E}^{\pi}[\mathcal{E}\{\rho(\mathbf{d}|\mathbf{m})\}] \quad (1.4)$$

Le critère d'information moyenne peut être alors exprimé autant dans l'espace des paramètres que dans l'espace des observations. Selon le problème posé, ces espaces sont

de grande taille et le calcul du critère d'information devient conséquent. Il faudra alors choisir la formulation adéquate. Dans notre étude, nous avons considéré une approximation au premier ordre.

En linéarisant l'opérateur \mathcal{F} , on obtient une relation simple entre les observations et les paramètres du système :

$$\mathbf{d} = \mathbf{F} \cdot \mathbf{m} + \nu \quad (1.5)$$

où ν symbolise l'erreur faite sur l'évaluation des observations. Généralement, on suppose que ν suit une loi gaussienne de moyenne nulle et de covariance \mathbf{C}_d telle que $\nu \sim \mathcal{N}(0, \mathbf{C}_d)$.

Dans les différents exemples que nous avons traités, l'opérateur de modélisation \mathcal{F} permet d'estimer le champ de perméabilité. Dans le premier exemple, nous étudions le cas de l'estimation du champ de perméabilité lorsque celui-ci est modélisé par krigeage. Ainsi, cherchons nous à définir l'échantillon spatial optimal des observations permettant d'estimer au mieux le champ de perméabilité. Dans notre deuxième exemple, l'estimation de la perméabilité est obtenue par la perméabilité apparente qui sera calculée au moyen d'un essai de puits (observation au la pression au puits). Notre but est alors de définir l'échantillon temporel optimal d'acquisition de pression. Enfin, dans notre dernier exemple, nous estimons la perméabilité à l'aide d'acquisition sismique. Pour simplifier le traitement des cubes de données sismiques, nous étudions l'évolution des lignes de courant. Dans le cadre diphasique, nous avons cherché à définir un échantillon spatio-temporel optimal. Nous allons présenter plus en détail ces trois exemples.

Dans le chapitre 6, le premier exemple considéré est celui du krigeage, qui est une méthode d'interpolation spatiale issue de la géostatistique. En différentes positions de l'espace, typiquement aux puits, des mesures de perméabilité sont faites et permettent d'estimer linéairement les données spatialisées. On comprend intuitivement qu'il existe un nombre d'observations tel que l'ajout d'information n'apporte plus rien sur l'estimation. On suppose généralement que le paramètre \mathbf{m} représentant la perméabilité du milieu suit une loi gaussienne de moyenne \mathbf{m}_p et de covariance \mathbf{C}_p connues *a priori*. Le nombre d'observations étant nettement inférieur au nombre de paramètres, on exprime l'information moyenne dans l'espace des observations. On suppose que les observations sont indépendantes et on pose $\mathbf{C}_d = \sigma_d^2 \cdot \mathbb{I}_N$. En considérant un variogramme gaussien pour la covariance *a priori* de la perméabilité, le problème est alors de déterminer les N positions $\mathbf{x}_1, \dots, \mathbf{x}_N$ des observations maximisant le critère d'information moyenne :

$$\langle I(\pi(\mathbf{m})) \rangle = \frac{1}{2} \log \left| \mathbb{I}_N + \frac{\sigma_0^2}{\sigma_d^2} \mathbf{F} \cdot \left(\exp - \left(\frac{\mathbf{x}_i - \mathbf{x}_j}{\lambda_c} \right)^2 \right)_{i,j} \cdot \mathbf{F}^t \right| \quad i, j \in [1, N] \quad (1.6)$$

où λ_c est la longueur de corrélation et σ_0 la variance du champ de perméabilité considéré.

Au cours du chapitre 7, cette approche a été étendue à l'intégration de données d'essai de puits où l'on cherche à obtenir l'échantillonnage temporel maximisant le critère

d'information moyenne. L'équation de diffusion au puits déduite de la loi de Darcy modélise l'écoulement monophasique d'un fluide faiblement compressible, en l'occurrence l'huile lourde. Elle modélise l'évolution de la pression au cours du temps sous l'effet d'un terme source Q , lié au débit de production. La pression est la seule inconnue du problème d'évolution. Le système est donné par :

$$\begin{cases} \phi\mu c_T \partial_t P(\mathbf{r}, t) = \nabla \cdot (k(\mathbf{r})\nabla P(\mathbf{r}, t)) + Q \frac{\mu}{h} \delta(\mathbf{r})H(t) \\ \forall \mathbf{r}, P(\mathbf{r}, t = 0) = 0 \\ \forall t, P(r \rightarrow +\infty, t) = 0 \end{cases}$$

En prenant la définition de la perméabilité apparente du milieu et en développant en série le terme de pression, on définit explicitement le critère d'information moyenne à l'aide d'un noyau pondérant la perméabilité. La fonction de covariance des perméabilités entre deux temps d'observations t et τ est définie par

$$\langle k_{app}(t), k_{app}(\tau) \rangle = \int_0^{+\infty} \frac{G(\boldsymbol{\xi}, \tau)}{2\pi D_0 \tau} \int_0^{+\infty} \langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle \frac{G(\mathbf{r}, t)}{2\pi D_0 t} d\mathbf{r} d\boldsymbol{\xi} \quad (1.7)$$

où G est le noyau pondérateur défini par Oliver et la covariance des perméabilités est définie par :

$$\langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle = k_G^2 \cdot \exp \left(\frac{1}{2} \frac{\sigma_0^2}{\exp \frac{r^2}{\ell_c^2}} \mathcal{I}_0 \left(\frac{r^2}{\ell_c^2} \right) + \frac{1}{2} \frac{\sigma_0^2}{\exp \frac{\xi^2}{\ell_c^2}} \mathcal{I}_0 \left(\frac{\xi^2}{\ell_c^2} \right) \right) \cdot \left(\exp \left(\frac{\sigma_0^2}{\exp \frac{r^2 + \xi^2}{2\ell_c^2}} \mathcal{I}_0 \left(\frac{r\xi}{\ell_c^2} \right) \right) - 1 \right)$$

où \mathcal{I}_0 est la fonction de Bessel modifiée d'ordre 0. Ainsi après la résolution du problème inverse, il apparaît que les temps d'observations optimaux se concentrent aux temps courts et sont de plus en plus éloignés aux temps longs. Ceci reflète que l'information d'un essai de puits est concentrée au puits et permet ainsi de caractériser les propriétés du réservoir aux abords du puits.

De la même manière, dans le chapitre 8, on a cherché à utiliser les approches lagrangiennes au premier ordre pour la simulation des fronts de saturation afin de déterminer les temps d'acquisition optimaux pour des observations répétées de sismique 3D. On utilise pour cela une approche perturbative de l'évolution du front de saturation dans un milieu supposé hétérogène. On suppose que le champ de vitesse \mathbf{u} est stationnaire et constant en espace. On suit l'hypothèse selon laquelle la perturbation de la vitesse locale est la somme de deux fluctuations dépendant d'une part du couplage visqueux et d'autre part de l'hétérogénéité du milieu. On relie ainsi la perturbation du front de saturation à la perturbation de la vitesse, elle-même exprimée en fonction de la fluctuation de la perméabilité et de la fonction de Green du problème associé. Dans l'espace

de Fourier, l'évolution d'un mode de perturbation de position du front x_α au temps t est donnée par :

$$\delta x_\alpha(t) = \frac{f'(S_f)}{\phi} e^{A|\alpha|t} \int_{-\infty}^t e^{-A|\alpha|\tau} \delta u_{bx\alpha}(c\tau) d\tau \quad (1.8)$$

où f est le flux fractionnaire total en fluide injecté, c la vitesse constante du front de saturation dans un milieu homogène, $A < 0$ un coefficient de stabilité dépendant de la mobilité², ϕ la porosité du milieu et $\delta u_{bx\alpha}$ la fluctuation de la vitesse pour le mode de Fourier α .

De cette formule, on obtient en utilisant la propriété d'isotropie du champ de perméabilité, la fonction de covariance entre deux fronts de saturation à deux temps différents dans l'espace réel et aux temps longs³ :

$$C(h, \Delta t) = \frac{k_0 \ell_c^2}{4} \int_0^1 d\beta \beta^{d/2-2} e^{-\theta\beta} \left(\frac{1+A^2}{1-A^2} (f_{\ell_c^2/4}(\beta) - 1) - g_{h,\Delta t}(\beta) \right) \quad (1.9)$$

où h est la distance entre les deux fronts, Δt la différence entre les temps d'acquisition, f et g sont deux fonctions définies chapitre 8. La quantité $\frac{k_0 \ell_c^2}{4}$ est définie en fonction de la fonction de covariance du champ de perméabilité dans l'espace de Fourier. Le choix de la fonction de covariance est dépendant de l'approche de linéarisation du problème. On note que dans une approche stochastique permettant de calculer les moments longitudinaux du déplacement du front de saturation, nous avons obtenu une forme différente et plus complexe de la fonction de covariance.

Nous concluons par un chapitre récapitulant nos résultats ainsi que les perspectives associées. La théorie de l'information nous est apparue comme un outil indispensable à l'optimisation. Cela nous a permis dans un premier temps de redécouvrir des résultats admis intuitivement. Dans le cas de sismique 3D répétée, dans la mesure où l'acquisition sismique représente une part importante dans le budget d'étude, il nous est apparu indispensable d'établir un formalisme servant à définir l'échantillonnage optimal d'acquisition.

Nous aimerions que nos travaux soient étendus au cas d'un calage d'historique des débits de production, polyphasique et plus généralement au cadre de distributions de probabilité diverses.

Avant de définir le contexte de résolution des problèmes inverses en ingénierie de réservoir, ainsi que l'importance significative de l'apport d'information, nous consacrerons le chapitre suivant à une présentation concise du métier et des enjeux de l'ingénierie de réservoir.

2. Le coefficient A est négatif dans le cas stable

3. L'analyse est faite aux temps long car cela permet de négliger la zone de raréfaction

Chapitre 2

Éléments d'ingénierie de réservoir

Sommaire

1	Ingénierie de réservoir : Objectifs et enjeux	7
2	Modélisation physique des écoulements	8
2.1	Modélisation statique du réservoir : modèle géologique	8
2.2	Modélisation dynamique	9
3	Caractérisation des réservoirs par problèmes inverses	10
3.1	Calage d'historique	10
3.2	Identification des paramètres et problème inverse	11

Dans ce premier chapitre, nous donnons une vue d'ensemble des intérêts et des enjeux de l'ingénierie et de la caractérisation de réservoir.

L'étude des réservoirs à hydrocarbures a pour objectif l'exploitation d'un réservoir de pétrole ou de gaz à moindre coût et à moindre risque. Les gisements les plus rentables économiquement sont les gisements les plus facilement accessibles et les plus simples à exploiter. Mais ces champs s'épuisent et sont de moins en moins nombreux. Dans les zones géopolitiquement sûres, ceux restant à exploiter sont souvent dans des zones d'accès difficile, comme par exemple, sous la mer à de grandes profondeurs. Leurs exploitations demandent des technologies très sophistiquées et coûteuses. De ce fait, la caractérisation est d'importance croissante en ingénierie puisqu'elle permet une meilleure gestion des risques dans les prises de décision.

1 Ingénierie de réservoir : Objectifs et enjeux

Un gisement pétrolier est formé d'un ou plusieurs réservoirs souterrains, poreux et contenant des hydrocarbures sous forme liquide et/ou gazeuse. Un réservoir est une roche comportant un grand nombre de micro-cavités inter-connectées appelées pores ($\sim 10^{-8}$ m), dans laquelle la circulation des fluides est possible. On dit alors de la roche réservoir qu'elle est poreuse et perméable. La structure du gisement ou du réservoir est

limitée par des barrières imperméables qui piègent les hydrocarbures.

L'étude d'un gisement permet d'élaborer un schéma de développement. L'ingénieur cherche alors à définir dans une stratégie donnée un scénario de production optimisant la récupération des hydrocarbures. Le schéma de développement détermine par exemple le nombre et l'emplacement des puits. Nous entendons par scénario de production les conditions possibles de production parmi les différentes options telles que le type de récupération (injection d'eau ou de gaz, production primaire sous l'effet d'un aquifère actif), débits de production et d'injection, contraintes de fonctionnement de puits etc.

Pour optimiser la récupération, l'ingénieur doit améliorer sa connaissance du gisement et adapter son schéma de développement aux nouvelles acquisitions d'informations. La mise en place d'une telle étude va permettre de prendre des décisions optimales pour le développement.

Durant toute la durée de vie du champ, les différentes caractéristiques du gisement sont étudiées afin d'élaborer un modèle géologique et un modèle d'écoulement. Ces modèles sont construits en utilisant toute la connaissance disponible, à savoir les données géologiques et les interprétations des experts sur ces données. Mais celles-ci sont par nature incertaines, du fait des erreurs de mesure. De plus, elles sont généralement insuffisantes pour représenter complètement les hétérogénéités présentes dans le gisement.

A ceci s'ajoute un grand nombre de possibilités de développement ou de production d'un réservoir. Par conséquent, pour étudier le comportement d'un gisement pétrolier de manière précise, une démarche scientifique adaptée et rigoureuse est requise.

2 Modélisation physique des écoulements

2.1 Modélisation statique du réservoir : modèle géologique

La modélisation statique du réservoir est réalisée pendant la phase de délinéation et au fur et à mesure de l'acquisition des données lors du développement du champ. Le modèle décrit les propriétés du réservoir avant la mise en production, c'est-à-dire lorsque le réservoir est à l'équilibre hydrostatique et thermodynamique. Cette modélisation est imprécise du fait de l'importance des coûts relatifs à l'acquisition des données. La difficulté est renforcée par le fait que le spécialiste dispose uniquement d'informations quantitatives sur une partie du gisement. Ces informations proviennent généralement de mesures sur carottes ou de diagraphies. D'autres sources d'information existent comme la sismique qui, quant à elle, fournit une information sur l'ensemble des réservoirs. Le géologue et l'ingénieur réservoir doivent donc extrapoler ces données ponctuelles à la totalité du champ.

La construction de ce modèle géologique est réalisée en trois grandes étapes :

Construction d'un modèle structural : Cette première modélisation provient principalement de la sismique associée aux mesures aux puits. Elle permet d'obtenir une image des différentes couches sédimentaires de réservoir, d'estimer l'emplacement des réservoirs dans le gisement en localisant les fonds et couvertures des réservoirs, de caractériser les principaux événements de fracturation, à savoir les failles principales, et enfin les contacts entre les fluides, notamment le contact huile-eau.

Construction d'un modèle stratigraphique : Ce modèle est une discrétisation du modèle structural sur un maillage cartésien fin.

Construction d'un modèle en faciès : Ce modèle détermine la géométrie et les proportions de chaque faciès et leur attribue des propriétés pétrophysiques. Il est obtenu à partir de la sismique, des mesures aux puits et des propriétés des carottes. Rappelons qu'un faciès caractérise les conditions sédimentaires et environnementales du sous-sol, chaque faciès ayant des propriétés spécifiques (type de roche, perméabilité, porosité, etc.).

Remarque 2.1 *Il existe une différence de plusieurs ordres de grandeur entre l'échelle d'observation des données sismiques, typiquement de l'ordre de la dizaine de mètres, l'échelle d'observation des mesures sur carottes de l'ordre de quelques centimètres et enfin la taille des mailles du modèle de l'ordre de 200 à 300 mètres horizontalement. Par conséquent l'utilisation conjointe des données sismiques et des mesures sur carottes impose des choix susceptibles d'introduire des incertitudes sur le modèle.*

La construction d'un modèle de réservoir revient donc à déterminer, entre autres, la structure géologique du sous-sol (hétérogénéité), les limites du réservoir, le type de sédimentation (éolien, fluvial), le type de piège (faille), etc. La caractérisation du réservoir est dans un premier temps statique et ne prend pas en considération ces propriétés dynamiques. Le modèle ainsi construit devra être mis à jour lors de futures phases de développement ou lors d'acquisitions de nouvelles données, ce qui impose de savoir modéliser le réservoir de façon dynamique.

2.2 Modélisation dynamique

Une étude des propriétés des fluides et de leurs types de déplacement est nécessaire lors de la mise en production du gisement par simple déplétion. Dans un gisement, il peut y avoir un écoulement simultané d'une ou plusieurs phases fluides : eau, huile et gaz. En début de production, il n'y a parfois qu'un seul fluide produit – l'huile, et lorsque l'on injecte de l'eau et/ou du gaz, deux ou trois phases fluides peuvent se déplacer dans la roche. Déterminer le type d'écoulement dans un réservoir est donc une tâche complexe. Ainsi, devons-nous nous intéresser à l'écoulement polyphasique dans un milieu poreux. Pour une description physique complète des phénomènes, on pourra se reporter notamment aux ouvrages de K. Aziz et A. Settari [2] ou D.-W. Peaceman [3].

La simulation numérique nécessite de résoudre des équations aux dérivées partielles faisant intervenir des grandeurs pétrophysiques dépendant de la nature des roches. Hormis dans quelques cas très simples, les systèmes d'équations régissant les modèles mathématiques n'ont pas de solutions explicites. Il faut donc faire appel à des méthodes d'approximation numériques. L'accroissement des performances informatiques (vitesse de calcul) et l'amélioration des algorithmes de résolution des systèmes linéaires ont grandement contribué au développement de la simulation numérique pour l'étude des réservoirs pétroliers.

Un facteur important pour la modélisation d'écoulement est la taille du maillage. Il est toujours préférable de travailler sur une grille de simulation aussi fine que possible afin de prendre en compte les hétérogénéités du réservoir. Cependant, pour des raisons pratiques évidentes, il convient de faire un compromis entre la finesse de la modélisation et les temps de calculs raisonnables. Ainsi, les simulations d'écoulement sont généralement effectuées sur une grille dont la dimension des mailles est plus grande que celle du maillage géologique. On parle alors de maillage grossier.

Pour ce faire, on fait appel à des techniques de mise à l'échelle pour calculer les valeurs caractéristiques équivalentes du réservoir sur ces mailles. Il faut noter que ce changement d'échelle augmente l'incertitude sur la description du réservoir.

Il est ainsi possible de construire un modèle dynamique en étudiant les propriétés dynamiques des fluides présents et de fournir des prévisions de production basées sur la simulation du phénomène qu'il faut désormais comparer aux *vraies* données issues de la production.

3 Caractérisation des réservoirs par problèmes inverses

Le modèle dynamique présenté dans le paragraphe précédent tient uniquement compte des propriétés du réservoir et de l'écoulement des fluides. Cependant, il est nécessaire de prendre en considération les données de production.

3.1 Calage d'historique

Les décisions des ingénieurs ne peuvent être basées sur des modèles numériques n'ajustant pas les données de production déjà observées. L'intégration de ces données dans le modèle numérique d'écoulement est désignée sous le terme de *calage d'historique*. Ainsi cherche-t-on à calibrer le modèle numérique aux données observées. Le modèle calibré restitue donc l'historique de production du gisement ainsi que son comportement en estimant les paramètres du gisement.

L'estimation des paramètres de ce modèle est un problème d'optimisation. On cherchera à minimiser l'écart entre l'historique réel de production et les estimations obtenues par le simulateur.

Remarque 2.2 *Le calage d'historique aboutit généralement à une mise à jour du modèle géologique, mais ceci est loin d'être systématique. Dans certains cas, celui-ci peut être incompatible avec les données géologiques, ce qui impose alors une remise en question plus profonde de la modélisation.*

3.2 Identification des paramètres et problème inverse

La fiabilité de la réponse du simulateur est fortement conditionnée par les valeurs initiales des paramètres du modèle, les conditions limites, etc... Dans un cadre idéal, les valeurs de production mesurées sur le terrain doivent coïncider avec la solution obtenue par le simulateur. En pratique, ce résultat est loin d'être atteint car les systèmes d'équations aux dérivées partielles impliqués sont résolus à l'aide de schémas numériques et ne donnent qu'une approximation du résultat souhaité. De plus, les conditions aux limites sont difficiles à déterminer précisément.

Notre but est d'estimer ces paramètres à l'aide de l'information disponible. Les données dynamiques constituent une information importante renseignant sur le comportement du gisement. Pour ce faire, il faut commencer à exploiter le gisement. Dès lors que celui-ci est en phase de production, il convient de procéder à un calage de l'historique de production pour contraindre les paramètres du modèle numérique d'écoulement. Une fonction coût ou fonction objectif est alors associée aux contraintes du réservoir. Sa minimisation permet alors de mettre à jour, au cours de l'exploitation du gisement, les paramètres. On intègre les nouvelles données de production pour ainsi aboutir à une meilleure compréhension des caractéristiques du réservoir. On représente schématiquement le processus de calage d'historique en figure 2.1.

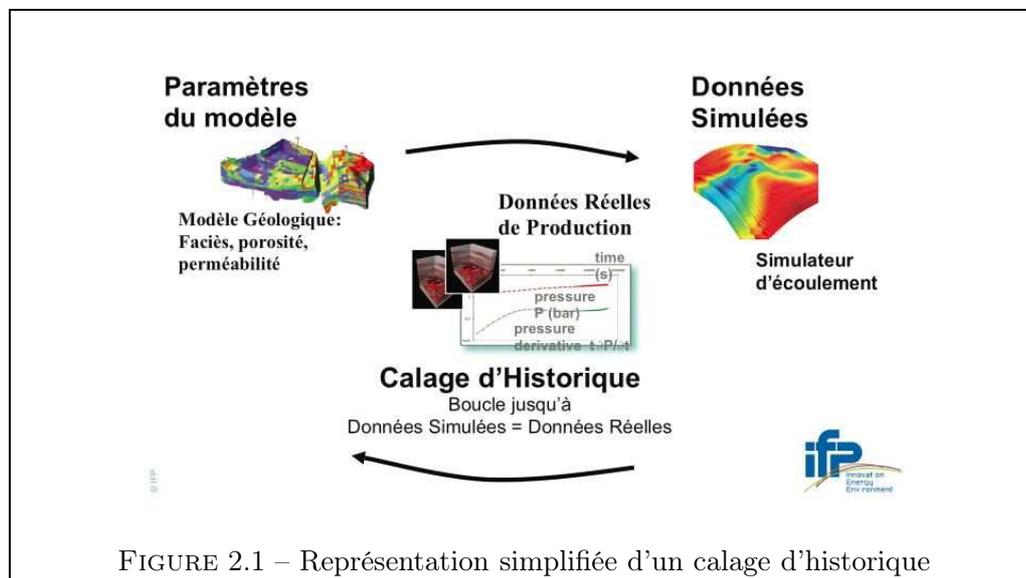


FIGURE 2.1 – Représentation simplifiée d'un calage d'historique

Dans un premier temps, le modèle géologique est construit à l'aide de l'information

disponible ou donnée par les géologues. Modulo des techniques de mises à l'échelle et le modèle initial construit, il est possible d'utiliser un simulateur de réservoir dans le but de modéliser la production du modèle initial. Dans un second temps, on confronte via la fonction objectif les données réelles de production et les productions simulées. Les données réelles de production dont on dispose permettent ainsi de juger si le modèle initial doit être mis à jour ou non. Le processus de calage d'historique converge alors vers un système de paramètres optimal (par rapport à la fonction objectif) reproduisant au mieux les données de production. Nous voyons par construction que le système de paramètres ainsi construit ne sera pas unique et donc, que ces différents modèles fourniront des prévisions et des scénarios de production différents.

Chapitre 3

Problèmes inverses en ingénierie de réservoir

Sommaire

1	Formalisation du problème inverse déterministe	13
1.1	Fonction objectif et optimisation	14
1.2	Spécificités des problèmes inverses en ingénierie de réservoir .	14
2	Formalisation du problème inverse probabiliste	15
2.1	Principe et intérêt de la Géostatistique	15
2.2	Distribution de probabilité	16
2.3	Réalisation de champs géostatistiques	17
2.4	Fonction objectif et optimisation	18

Nous présentons dans ce chapitre une vue d'ensemble de la caractérisation de réservoir par problème inverse en définissant dans sa version formelle la fonction objectif. Après avoir rappelé brièvement les difficultés rencontrées lors de la résolution de problèmes inverses, nous terminons par la présentation de son formalisme dans un contexte probabiliste.

1 Formalisation du problème inverse déterministe

La construction d'un modèle de réservoir prédictif est obtenue par la résolution d'un problème inverse. Ces problèmes demeurent difficiles à résoudre, essentiellement en raison de leur forte non-linéarité. Les méthodes de résolution les plus classiques reposent sur l'estimation par la méthode des moindres carrés, mais présentent l'inconvénient majeur de conduire à des problèmes de minimisation possédant de nombreux minima locaux.

1.1 Fonction objectif et optimisation

Le calage d'historique de production permet de déterminer la valeur de certains paramètres caractéristiques du gisement en tenant compte de l'historique des mesures : débits et pressions aux puits, données sismiques 2D, 3D et 4D, etc. L'objectif de cette approche est de restituer le comportement dynamique d'un gisement sur la durée totale de l'acquisition des données par ajustement des paramètres pétrophysiques.

La méthode des gradients développée par Anterion *et al.*[4] permet de calculer les sensibilités des grandeurs observées par rapport aux paramètres caractérisant le réservoir. L'ingénieur sait alors qu'une variation de la valeur d'un paramètre de très faible gradient ne saurait permettre une meilleure restitution des observations.

Le calage opère par la minimisation d'une fonctionnelle appelée *fonction objectif* ou *fonction coût*. Cette minimisation est généralement réalisée en utilisant des techniques d'optimisation et de contrôle optimal. On pourra consulter Chavent [5]. Cette optimisation nécessite la définition d'un état adjoint. On consultera Roggero et Guérillot [6]. La formulation générale du problème d'optimisation est :

$$\left\{ \begin{array}{l} \min_{\mathbf{m} \in \mathcal{M}} \mathcal{J}(\mathbf{m}) = \mathcal{R}(\mathbf{d}^{obs}, \mathcal{F}[\mathbf{m}]) \\ c(\mathbf{m}) \end{array} \right.$$

où l'on désigne par \mathbf{d}^{obs} le vecteur des *vraies* observations et $\mathcal{F}[\mathbf{m}]$ le vecteur des observations simulées. La fonction \mathcal{R} est définie comme une fonction risque qui devra être explicitée, et enfin, $c(\mathbf{m})$ représente les contraintes imposées sur le paramètre \mathbf{m} .

1.2 Spécificités des problèmes inverses en ingénierie de réservoir

Les problèmes inverses tels que le calage d'historique présentent de nombreuses difficultés liées à leurs non-linéarités, leurs grandes tailles, peu de données de départ (paramétrisation difficile), ce qui entraîne un coût important en temps de calcul. De plus, ces problèmes sont généralement mal posés¹ : la solution du problème inverse ne dépend pas de façon continue des données. En effet, une faible variation sur les données \mathbf{d} peut induire d'importantes variations sur l'inférence de \mathbf{m} .

Lors de la formulation et de la résolution d'un problème inverse, une attention particulière doit être portée aux points suivants :

Choix de la formulation : suivant l'expression du critère à minimiser (choix des paramètres à estimer, prise en compte d'*a priori*, choix du coefficient de régularisation) le problème de minimisation associé est plus ou moins bien posé. Une bonne compréhension du problème physique est alors nécessaire pour effectuer de bons choix.

1. Au sens d'Hadammard

Choix de la paramétrisation : le choix de la représentation discrète des paramètres à estimer est un problème non trivial. Ce choix a aussi une influence sur le conditionnement du problème de minimisation associé.

Identifiabilité : la question qui se pose toujours, que ce soit de façon théorique ou pratique, est de savoir si les mesures sont suffisantes pour estimer les paramètres d'intérêt et comment l'incertitude sur les mesures se répercute sur leurs estimations.

Les choix de la formulation et de la paramétrisation sont essentiels car ils impactent la résolution du problème inverse et de fait les prévisions de productions.

L'identifiabilité montre le besoin croissant de quantifier l'apport d'information d'une nouvelle mesure sur l'estimation des paramètres. Il convient de définir avant toute simulation, le nombre effectif de degrés de liberté du problème. Cette nouvelle mesure sera intrinsèquement liée à l'échelle d'observation du phénomène. En effet, si N observations permettent d'estimer $M \gg N$ paramètres, alors on sera en mesure de prendre une décision sur une nouvelle campagne d'acquisition de données, ou alors de contraindre la modélisation à M paramètres.

2 Formalisation du problème inverse probabiliste

L'information disponible ne permet pas de caractériser les propriétés du modèle de manière unique. Une approche stochastique de la caractérisation du réservoir permet de contrôler et de modéliser les incertitudes permanentes présentes dans le réservoir. Les méthodes stochastiques permettent alors de construire plusieurs modèles de réservoir en conservant les propriétés statistiques relatives à la structure spatiale des corps géologiques.

2.1 Principe et intérêt de la Géostatistique

La Géostatistique fut formalisée par Matheron [7] comme *l'application du formalisme des fonctions aléatoires à la reconnaissance et à l'estimation de phénomènes naturels*. La complexité des réservoirs ainsi que l'impossibilité de représenter toutes les hétérogénéités ont amené naturellement à une description stochastique des réservoirs pétroliers. Les paramètres des fonctions aléatoires tels que l'espérance, le type de variogramme ou encore la loi de distribution permettent de décrire les structures géologiques complexes ayant des distributions spatiales semblables pour la taille des corps (géologiques) ou la disposition dans l'espace et pour la propriété physique des différentes roches.

On suppose que le réservoir est décrit par une fonction aléatoire $y(\mathbf{r})$. Cette variable peut représenter une propriété de la roche qui peut être mesurée directement et localement sur le champ ou en laboratoire telle que la porosité ou la perméabilité. Supposons qu'en \mathbf{r} la valeur de y est $y(\mathbf{r})$.

Le problème est de caractériser et de modéliser la variabilité de $y(\mathbf{r})$ en espace. Celle-ci servira à estimer la valeur en une position différente \mathbf{r}' de $y(\mathbf{r}')$ où la mesure n'a pas encore été réalisée. Cette méthode géostatistique est le krigeage et fut développée empiriquement par Krige et formalisée par Matheron. Elle permet l'estimation de valeurs locales en considérant l'organisation spatiale des variables étudiées.

Ainsi, la Géostatistique consiste-t-elle en l'interprétation de chaque valeur $y(\mathbf{r}_i)$ comme une représentation particulière d'une variable aléatoire. De cet ensemble de variables aléatoires, on construit une fonction aléatoire notée $Y(\mathbf{r}) = \{y(\mathbf{r}_i)\}$. Cette fonction construite est généralement fortement erratique et possède de nombreuses discontinuités et anisotropies.

Pour caractériser la variabilité spatiale de $y(\mathbf{r})$, on définit la corrélation entre les différentes variables aléatoires $y(\mathbf{r}_i)$. Celle-ci dépend à la fois de la nature de la variable considérée et de la distance séparant les emplacements des mesures. La Géostatistique ne fait pas appel à la *Statistique descriptive* qui développe des techniques permettant une meilleure compréhension du type d'information disponible mais à la *Statistique inférentielle*. Celle-ci permet de tirer des conclusions fiables à partir de données d'échantillons statistiques.

2.2 Distribution de probabilité

La fonction de densité de probabilité ou simplement densité de probabilité caractérise la probabilité $dF(\mathbf{y})$ qu'une variable aléatoire prenne une valeur entre \mathbf{y} et $\mathbf{y} + d\mathbf{y}$. Ici, F est la fonction de répartition de \mathbf{y} . Soit :

$$\text{proba}(\mathbf{v} \in [\mathbf{y}, \mathbf{y} + d\mathbf{y}]) = dF(\mathbf{y}) \quad (3.1)$$

Si de plus la fonction de répartition est différentiable, la densité de probabilité p de y est alors définie par :

$$dF(\mathbf{y}) = p(\mathbf{y})d\mathbf{y} \quad (3.2)$$

Dans la cas où la variable est continue, la densité de probabilité décrit le comportement et l'évolution de la variable aléatoire. C'est une fonction positive, nécessairement inférieure à 1 et normalisée de sorte que :

$$\int p(\mathbf{y})d\mathbf{y} = 1 \quad (3.3)$$

Il existe une multitude de densités de probabilité. En géostatistique, la distribution normale multi-variée et la distribution log-normale multi-variée jouent des rôles importants. La densité de probabilité normale ou gaussienne multi-variée de \mathbf{y} est définie par :

$$\forall \mathbf{y} \in \mathbb{R}^N, \quad p(\mathbf{y}) = ((2\pi)^N |C|)^{-1/2} \exp\left(-\frac{1}{2}(\mathbf{y} - \langle \mathbf{Y} \rangle)^t \mathbf{C}^{-1}(\mathbf{y} - \langle \mathbf{Y} \rangle)\right) \quad (3.4)$$

où $\langle \mathbf{Y} \rangle$ et \mathbf{C} sont respectivement la moyenne et la covariance de \mathbf{y} .

Dans le cas continu, la matrice de covariance \mathbf{C} est définie par :

$$C_{ij} = \langle \mathbf{y}_i, \mathbf{y}_j \rangle \quad (3.5)$$

On dit que \mathbf{x} suit une distribution log-normale de paramètres $\langle \mathbf{y} \rangle$ et C si et seulement si $\mathbf{y} = \log \mathbf{x}$ suit une distribution normale de paramètres $\langle \mathbf{Y} \rangle$ et C . La densité de probabilité est définie par :

$$p(\mathbf{x}) = ((2\pi)|C|)^{-N/2} \prod_{i=0}^{i=N} \frac{1}{x_i} \exp \left(-\frac{1}{2} (\log \mathbf{x} - \langle \mathbf{y} \rangle)^t C^{-1} (\log \mathbf{x} - \langle \mathbf{y} \rangle) \right) \quad (3.6)$$

2.3 Réalisation de champs géostatistiques

Traditionnellement, les champs aléatoires Gaussiens sont générés à l'aide de la décomposition de Cholesky de la matrice de covariance du champ aléatoire. Rappelons les bases de la méthode de Cholesky. Considérons la matrice de covariance C . Alors, on peut décomposer cette dernière comme le produit d'une matrice triangulaire inférieure et de sa transposée.

$$C = LL^t \quad (3.7)$$

Si on considère un champ gaussien aléatoire Y de moyenne m et de matrice de covariance C , alors on peut écrire à l'aide d'un bruit blanc gaussien (vecteur aléatoire gaussien centré-réduit non corrélé) z :

$$Y = m + L \cdot z \quad (3.8)$$

Le coût de la décomposition dépend directement de la taille de l'échantillon à générer. Pour un champ aléatoire comportant N points, la matrice de covariance sera de dimension $N \times N$. La décomposition de Cholesky est envisageable en optimisation stochastique pour des champs d'un millier de points au maximum. Or, en ingénierie pétrolière, la taille des champs aléatoires est de l'ordre du million. Il faut donc une méthode rapide et efficace qui permette de contourner ce problème.

La méthode de la moyenne mobile a été développée initialement pour les champs gaussiens stationnaire 1D [8] et a été étendue aux dimensions supérieures 2D et 3D par Oliver [9]. Pour la génération de champ aléatoire, nous utilisons l'algorithme FFT-MA² développé à l'IFP [10]. Le champ $Y(\mathbf{r})$ est considéré comme un processus gaussien stationnaire et on considère sa représentation dans l'espace des fréquences à l'aide de la transformée de Fourier :

$$Y(\mathbf{q}) = \int e^{i\mathbf{q}\cdot\mathbf{r}} Y(\mathbf{r}) d\mathbf{r} \quad (3.9)$$

On a la relation inverse suivante :

$$Y(\mathbf{r}) = \frac{1}{(2\pi)^D} \int e^{i\mathbf{q}\cdot\mathbf{r}} Y(\mathbf{q}) d\mathbf{q} \quad (3.10)$$

Le processus gaussien $Y(\mathbf{r})$ étant stationnaire, on montre que le processus $Y(\mathbf{q})$ est aussi gaussien tel que :

$$\langle Y(\mathbf{r}) \rangle = 0 \implies \langle Y(\mathbf{q}) \rangle = 0 \quad (3.11)$$

$$\langle Y(\mathbf{r}), Y(\mathbf{r}') \rangle = C(\mathbf{r} - \mathbf{r}') \implies \langle Y(\mathbf{q}), Y(\mathbf{q}') \rangle = (2\pi)^D C(\mathbf{q}) \delta(\mathbf{q} - \mathbf{q}') \quad (3.12)$$

où C est la fonction de covariance associée au processus Y . On donne une représentation d'un champ de perméabilité log-normal figure (3.1) construit par la méthode FFTMA.

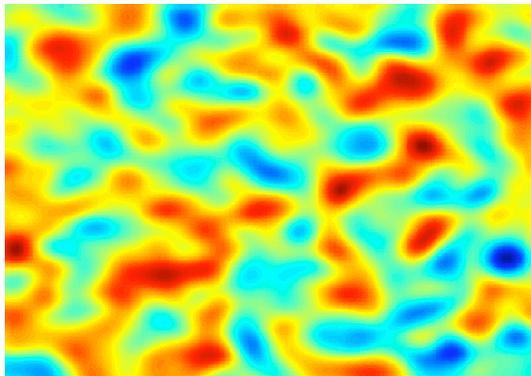


FIGURE 3.1 – Représentation d'un champ de perméabilité log-normal

2.4 Fonction objectif et optimisation

La résolution du problème inverse doit permettre l'estimation des paramètres pétrophysiques à l'aide des données d'observations. Du fait qu'il existe conjointement des erreurs de modélisation et des erreurs sur les observations, il devient naturel de proposer une approche du problème inverse sous un angle probabiliste.

On peut noter les deux principaux avantages à utiliser une méthode probabiliste pour résoudre les problèmes inverses. Premièrement, il est possible d'estimer à la fois l'incertitude et la valeur des paramètres inconnus. Les modèles étant construits, on pourra quantifier l'erreur sur les prévisions au sens des probabilités. Nous pouvons schématiser le processus d'inversion par la figure 3.2

Après calage des données de production, plusieurs modèles sont construits et fournissent différents scénarios de production. On peut alors quantifier les incertitudes des différents modèles en vue de les réduire. Dans notre étude, nous utilisons le formalisme bayésien afin de résoudre notre problème inverse. Ce formalisme est décrit en annexe A.

Nous venons de présenter dans ce chapitre une vue d'ensemble des problèmes inverses en ingénierie de réservoir. La fonction objectif est alors centrale dans l'estimation des paramètres pétrophysiques et géologiques. Or, ces différents paramètres sont différents par nature et parfois ne sont pas représentés à la même échelle (pressions, cubes sismiques et perméabilité par exemple). Un problème serait alors de devoir normaliser ces mesures avant d'optimiser la fonction objectif. Nous montrons dans le chapitre suivant qu'introduire la théorie de l'information permet d'optimiser nos paramètres en fonction du gain d'information que ceux-ci nous apportent.

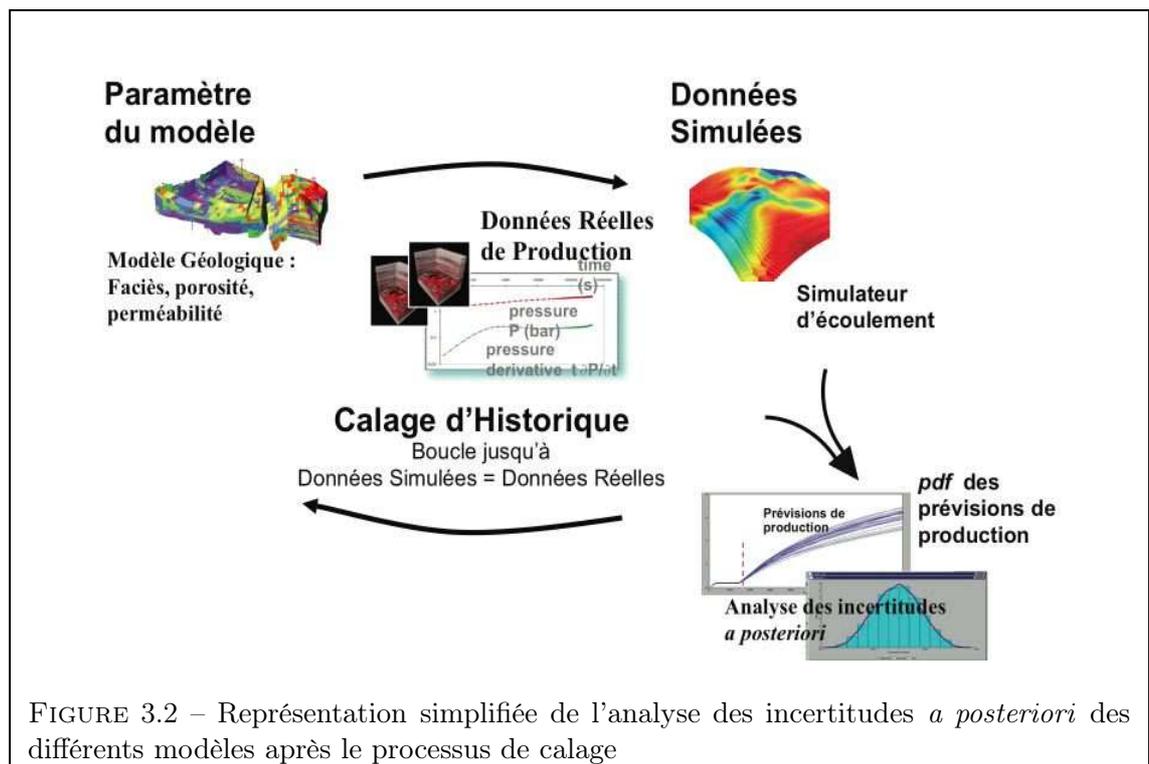


FIGURE 3.2 – Représentation simplifiée de l'analyse des incertitudes *a posteriori* des différents modèles après le processus de calage

Chapitre 4

Problèmes inverses et mesure de l'information

Sommaire

1	États d'information du système	22
1.1	Problème direct	22
1.2	Résolution du problème inverse dans le cadre probabiliste	23
2	Cadre probabiliste	24
2.1	Conjonction d'informations	24
2.2	Gain d'information et apprentissage	24
3	Comportement asymptotique de l'information moyenne	27
3.1	Fonctionnelle d'information	27
3.2	Critères d'information	28
3.3	Cas d'un paramètre unidimensionnel	29
3.4	Cas d'un paramètre multidimensionnel	30

Dans ce chapitre, nous définissons notre mesure de l'information et exposons son utilisation pour définir notre gain d'information. Ces deux critères seront en effet déterminants afin d'optimiser le processus de caractérisation de réservoir, notamment pour proposer un meilleur échantillonnage spatial ou temporel des mesures.

En ingénierie pétrolière, la caractérisation de réservoir correspond en particulier à l'estimation des propriétés pétrophysiques telles que la perméabilité et la porosité du milieu. Ces quantités ne peuvent généralement pas être mesurées directement sur l'ensemble du gisement. L'ingénieur dispose seulement de quelques mesures de perméabilité sur le gisement, ou alors de mesures indirectes de pression et de débit aux puits.

L'information disponible sur le milieu n'est donc pas connue continûment dans l'espace et dans le temps. De plus, les mesures sont généralement entachées d'un *bruit*, le plus souvent supposé gaussien et stationnaire. Le problème n'est donc pas seulement de prédire une évolution de la production ou encore un scénario exceptionnel, mais avant

tout de pouvoir estimer les paramètres explicatifs afin de disposer d'un modèle de simulation d'écoulement robuste.

Le problème inverse consiste donc à estimer des variables pétrophysiques telles que la perméabilité et la porosité au vu d'observations telles que les débits et les pressions aux puits. Ce problème se reformule en un problème d'optimisation.

Les variables d'observations et les variables pétrophysiques sont considérées comme des variables aléatoires sur un espace probabilisé et ne sont plus considérées comme des inconnues du problème. Notre problème inverse se redéfinit comme la caractérisation d'une distribution de probabilité pour le couple de variables d'observation et de paramètres.

L'introduction de la théorie de l'information dans les problèmes inverses en géosciences a été initialement faite par Tarantola [11]. Il postule que la solution du problème inverse peut être interprétée comme le résultat d'une *conjonction* d'états d'*information* comme nous le verrons plus en détails.

L'estimation de la perméabilité et de la porosité par inférence bayésienne fut introduite par Gavalas et Shah [12]. Ils considèrent un réservoir 1D et les observations proviennent des mesures des pressions au puits. Ils ajoutent à la fonction objectif l'information *a priori* disponible dans le réservoir. Ce processus de régularisation statistique contraint ainsi les paramètres du modèle à ne pas trop s'écarter des valeurs moyennes du modèle.

1 États d'information du système

Il est important en inversion de comprendre la relation entre le modèle dynamique et les observations résultantes. Si nous disposons d'un modèle direct nous permettant d'obtenir, en différents lieux de l'espace, des observations (simulées) à un temps donné et ceci depuis une idée *a priori* des valeurs des paramètres, alors on peut se demander *a posteriori* de quelles valeurs de paramètres proviennent les observations.

En production, les ingénieurs disposent de données réelles et la résolution du problème inverse doit donc permettre l'intégration de ces observations dans le modèle théorique. Généralement, l'intégration de ces observations se fait par la minimisation d'une *fonction coût* ou *fonction objectif* mesurant l'écart entre les observations réelles et les observations simulées.

1.1 Problème direct

Le problème continu direct peut être exprimé à l'aide de la fonctionnelle \mathcal{F} qui à tout champ $\mathbf{m}(\mathbf{x}) \in \mathcal{M}$ associe le vecteur d'observations $\mathbf{d}^* \in \mathcal{D}$ tel que :

$$\begin{aligned} \mathcal{M} &\longmapsto \mathcal{D} \\ \mathbf{m} &\longmapsto \mathbf{d}^* = \mathcal{F}[\mathbf{m}] \end{aligned} \quad (4.1)$$

Nous avons vu au chapitre précédent qu'il n'était pas possible de reproduire exactement les observations et de ce fait, on ajoute une erreur ou bruit $\boldsymbol{\eta}$ aux observations simulées. On suppose implicitement que les observations simulées ne sont en réalité qu'une moyenne de ce que l'on observerait réellement. On traduit généralement cette supposition en considérant $\boldsymbol{\eta}$ comme un bruit blanc gaussien stationnaire, indépendant de \mathbf{m} et de \mathcal{F} .

Remarque 4.1 *Définir l'erreur $\boldsymbol{\eta}$ comme une erreur gaussienne stationnaire est une hypothèse forte. C'est en effet une hypothèse d'école qui est le reflet d'une situation idéale asymptotique. Ces considérations permettent dans un premier temps de donner des résultats analytiques puis numériques.*

On définit alors la relation entre les observations réelles \mathbf{d}^{obs} et les paramètres \mathbf{m} par un modèle de régression général tel que

$$\mathbf{d}^{obs} = \mathcal{F}[\mathbf{m}] + \boldsymbol{\eta} \quad (4.2)$$

Pour obtenir \mathbf{m} , il suffirait dans un cadre idéal d'inverser l'opérateur \mathcal{F} . Mais dans la majorité des cas, l'inversion n'est pas possible directement ou explicitement et le problème revient à construire un estimateur *convenable* \mathcal{G} de \mathcal{F}^{-1} tel que

$$\begin{aligned} \mathcal{D} &\longmapsto \mathcal{M} \\ \mathbf{d}^{obs} &\longrightarrow \tilde{\mathbf{m}} = \mathcal{G}[\mathbf{d}^{obs}] = \mathcal{G}[\mathcal{F}[\mathbf{m}] + \boldsymbol{\eta}] \end{aligned} \quad (4.3)$$

De plus, considérant des distributions de probabilité sur les observations et les paramètres du système, il est possible à l'aide de l'entropie de définir plusieurs états d'information. Tarantola [11] a initialement introduit en Géosciences l'état d'*information* disponible dans le système et donne un schéma général pour la résolution du problème inverse dans le cadre probabiliste.

1.2 Résolution du problème inverse dans le cadre probabiliste

La résolution du problème inverse dans le cadre probabiliste est très largement traitée par Tarantola [11] et K. Mosegaard & Tarantola [13]. Cette résolution consiste en trois étapes, pour lesquelles une distribution de probabilité est définie, et une estimation de l'information est réalisée. Les 3 étapes de la résolution sont :

Paramétrisation du système : on cherche l'ensemble minimal des paramètres du modèle susceptible de permettre l'explication des observations du système tout en conservant la description géologique du milieu.

Problème direct : on décrit les lois physiques permettant de prédire les résultats de mesure de paramètres d'observations. Ici, dans le cas de l'ingénierie de réservoir, nous pouvons recourir à l'utilisation d'un modèle numérique de simulation d'écoulement polyphasique en milieu poreux.

Problème inverse : utilisation des quantités d'information précédentes, ainsi que les observations pour estimer les valeurs des paramètres.

2 Cadre probabiliste

Comme nous venons de le voir, la résolution du problème inverse peut se faire au moyen de distributions de probabilité. Considérons la variable $(\mathbf{d}, \mathbf{m}) \in \mathcal{X} = \mathcal{D} \times \mathcal{M}$ dans l'espace joint des observations et des paramètres et définissons les différentes distributions de probabilité qui permettront de résoudre le problème inverse comme solution d'une conjonction d'informations.

2.1 Conjonction d'informations

Le résultat de mesure \mathbf{d} peut être défini par une densité de probabilité $\rho(\mathbf{d})$. Nous pouvons dans un premier temps considérer que l'information *a priori* disponible est exogène des observations, c'est-à-dire qu'elle n'est pas définie en fonction de \mathbf{d} . Ainsi, la distribution dans l'espace \mathcal{X} décrite par la distribution jointe $\rho(\mathbf{d}, \mathbf{m})$ est exprimée par :

$$\rho(\mathbf{d}, \mathbf{m}) = \rho(\mathbf{d})\rho(\mathbf{m})$$

Nous pouvons de même définir la distribution $p(\mathbf{d}, \mathbf{m})$ qui représente l'information dite théorique sur la corrélation entre les observations et les paramètres. Elle contient l'information du modèle. En établissant la conjonction entre ces deux états d'informations, Tarantola [11] définit l'état d'information résultant par la distribution suivante :

$$\sigma(\mathbf{d}, \mathbf{m}) = k \frac{\rho(\mathbf{d}, \mathbf{m}) p(\mathbf{d}, \mathbf{m})}{\mu(\mathbf{d}, \mathbf{m})} \quad (4.4)$$

où la distribution $\mu(\mathbf{d}, \mathbf{m})$ est une distribution non-informative représentant l'état d'information homogène et k est une constante de normalisation.

2.2 Gain d'information et apprentissage

Dans une approche bayésienne, nous définissons l'ensemble $\mathcal{X} = \mathcal{D} \times \mathcal{M}$ comme le produit de l'ensemble des observations et l'ensemble des paramètres. Nous portons l'inférence sur le couple (\mathbf{d}, \mathbf{m}) . Il faut donc définir une distribution *a priori* sur \mathcal{X} .

La solution du problème inverse est donc la distribution marginale sur les paramètres de la distribution jointe définie en 4.4 telle que :

$$\sigma(\mathbf{m}) = \int_{\mathcal{M}} \sigma(\mathbf{d}, \mathbf{m}) \, d\mathbf{m} \quad (4.5)$$

L'égalité 4.4 est en réalité la version généralisée de la formule de Bayes [14].

Définition 4.1 Soit une observation $\mathbf{d} \sim \rho(\mathbf{d}|\mathbf{m})$ de \mathcal{D} . Si l'on dispose d'une distribution de probabilité *a priori* $\pi(\mathbf{m})$ sur le paramètre $\mathbf{m} \in \mathcal{M}$. Alors, le théorème de Bayes permet de construire la distribution *a posteriori* $\rho(\mathbf{m}|\mathbf{d})$:

$$\rho(\mathbf{m}|\mathbf{d}) = \frac{\rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m})}{\int_{\mathcal{M}} \rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m}) \, d\mathbf{m}} \quad (4.6)$$

En théorie de l'information, le gain d'information peut être calculé à l'aide de la divergence de Kullback-Leibler entre les distributions *a posteriori* et *a priori*. Il existe d'autres types de divergence, on pourra se référer à [15, 16] pour la définition des *f*-divergence introduites par Csiszár. Comme nous l'avons vu au Chapitre 2, cette quantité définit l'apport d'information transmise d'un état *pré*-mesure à un état *post*-mesure. C'est-à-dire :

$$D(\mathbf{d}) = K(\rho(\mathbf{m}|\mathbf{d})\|\pi(\mathbf{m})) = \int_{\mathcal{M}} \rho(\mathbf{m}|\mathbf{d}) \log \frac{\rho(\mathbf{m}|\mathbf{d})}{\pi(\mathbf{m})} d\mathbf{m} \quad (4.7)$$

La quantité d'information apportée par une mesure est l'espérance *a posteriori* de $\log \frac{\rho(\mathbf{m}|\mathbf{d})}{\pi(\mathbf{m})}$. Nous aimerions avoir un critère d'information avant d'effectuer toutes mesures. Ainsi, il convient de prendre la moyenne sur l'ensemble des observations possibles.

Définition 4.2 *On peut définir désormais la quantité d'information moyenne par :*

$$\langle I \rangle = I_M(\pi(\mathbf{m})) = \int_{\mathcal{D}} D(\mathbf{d}) \sigma(\mathbf{d}) d\mathbf{d} \quad (4.8)$$

En reprenant l'égalité $\sigma(\mathbf{d}, \mathbf{m}) = \rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m})$, l'information moyenne peut être aussi définie à l'aide de la divergence de Kullbak-Leibler. Ainsi,

$$\begin{aligned} \langle I \rangle &= \int_{\mathcal{D}} \int_{\mathcal{M}} \sigma(\mathbf{d}) \frac{\rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m})}{\sigma(\mathbf{d})} \log \frac{\rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m})}{\sigma(\mathbf{d})\pi(\mathbf{m})} d\mathbf{m} d\mathbf{d} \\ &= \int_{\mathcal{X}} \rho(\mathbf{d}, \mathbf{m}) \log \frac{\rho(\mathbf{d}, \mathbf{m})}{\sigma(\mathbf{d})\pi(\mathbf{m})} d(\mathbf{d}, \mathbf{m}) \\ \langle I \rangle &= K \left[\rho(\mathbf{d}, \mathbf{m}) \parallel \sigma(\mathbf{d})\pi(\mathbf{m}) \right] \end{aligned} \quad (4.9)$$

L'information moyenne est donc la divergence entre la distribution du modèle *joint* et la distribution du modèle *indépendant*. Ainsi les observations et les paramètres sont indépendants et le modèle joint est égal au modèle indépendant. L'information moyenne est nulle et une observation supplémentaire n'apporte donc aucune information.

L'entropie de Shannon permet de donner une définition de l'information moyenne en fonction de l'entropie de la distribution de vraisemblance et l'entropie de la distribution des observations.

Proposition 1 *Soit $\mathbf{d} \in \mathcal{D}$ un vecteur d'observations de distribution $\sigma(\mathbf{d})$. On suppose qu'il existe un vecteur de paramètres $\mathbf{m} \in \mathcal{M}$ pouvant expliquer le phénomène observé \mathbf{d} . On suppose que la distribution de vraisemblance $\rho(\mathbf{d}|\mathbf{m})$ connue et normalisée. On note $m(\mathbf{d}, \mathbf{m})$ la distribution non-informative sur \mathcal{X} et on suppose que $m(\mathbf{d}, \mathbf{m}) = \mu(\mathbf{d}^{obs})\varrho(\mathbf{m})$ où μ et ϱ sont respectivement les distributions non-informatives sur \mathcal{D} et \mathcal{M} .*

Alors, l'information moyenne apportée est exprimée par :

$$\langle I \rangle = \mathcal{E}_{\mathbf{d}}\{\sigma\} - \mathbb{E}^{\pi} [\mathcal{E}_{\mathbf{d}|\mathbf{m}}\{\rho\}] \quad (4.10)$$

Preuve On reprend la définition de l'information moyenne apportée par une observation \mathbf{d} de distribution de probabilité $\sigma(\mathbf{d})$.

$$\langle I \rangle = \int_{\mathcal{D}} \int_{\mathcal{M}} \rho(\mathbf{d}, \mathbf{m}) \log \frac{\rho(\mathbf{d}, \mathbf{m})}{\sigma(\mathbf{d})\pi(\mathbf{m})} \, \mathrm{d}\mathbf{m} \, \mathrm{d}\mathbf{d} \quad (4.11)$$

On introduit la distribution non-informative dans l'expression précédente,

$$\langle I \rangle = \int_{\mathcal{D}} \int_{\mathcal{M}} \rho(\mathbf{d}, \mathbf{m}) \left[\log \frac{\rho(\mathbf{d}, \mathbf{m})}{\mu(\mathbf{d})\varrho(\mathbf{m})} - \log \frac{\sigma(\mathbf{d})\pi(\mathbf{m})}{\mu(\mathbf{d})\varrho(\mathbf{m})} \right] \, \mathrm{d}\mathbf{m} \, \mathrm{d}\mathbf{d} \quad (4.12)$$

En développant la distribution jointe à l'aide de la distribution $\sigma(\mathbf{d})$ des observations, on a la relation suivante :

$$\rho(\mathbf{d}, \mathbf{m}) = \rho(\mathbf{m}|\mathbf{d})\sigma(\mathbf{d}) = \rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m}) \quad (4.13)$$

on obtient :

$$\langle I \rangle = \int_{\mathcal{D}} \int_{\mathcal{M}} \rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m}) \log \frac{\rho(\mathbf{d}|\mathbf{m})}{\mu(\mathbf{d})} \frac{\pi(\mathbf{m})}{\varrho(\mathbf{m})} \, \mathrm{d}\mathbf{m} \, \mathrm{d}\mathbf{d} \quad (4.14)$$

$$- \int_{\mathcal{D}} \int_{\mathcal{M}} \rho(\mathbf{d}, \mathbf{m}) \log \frac{\sigma(\mathbf{d})}{\mu(\mathbf{d})} \frac{\pi(\mathbf{m})}{\varrho(\mathbf{m})} \, \mathrm{d}\mathbf{m} \, \mathrm{d}\mathbf{d} \quad (4.15)$$

$$\begin{aligned} &= \int_{\mathcal{M}} \pi(\mathbf{m}) \int_{\mathcal{D}} \rho(\mathbf{d}|\mathbf{m}) \log \frac{\rho(\mathbf{d}|\mathbf{m})}{\mu(\mathbf{d})} \, \mathrm{d}\mathbf{d} \, \mathrm{d}\mathbf{m} \\ &+ \int_{\mathcal{M}} \pi(\mathbf{m}) \log \frac{\pi(\mathbf{m})}{\varrho(\mathbf{m})} \int_{\mathcal{D}} \rho(\mathbf{d}|\mathbf{m}) \, \mathrm{d}\mathbf{d} \, \mathrm{d}\mathbf{m} \\ &- \left[\int_{\mathcal{D}} \int_{\mathcal{M}} \rho(\mathbf{d}, \mathbf{m}) \log \frac{\pi(\mathbf{m})}{\varrho(\mathbf{m})} \, \mathrm{d}\mathbf{m} \, \mathrm{d}\mathbf{d} \right. \\ &\left. + \int_{\mathcal{D}} \int_{\mathcal{M}} \rho(\mathbf{d}, \mathbf{m}) \log \frac{\sigma(\mathbf{d})}{\mu(\mathbf{d})} \, \mathrm{d}\mathbf{m} \, \mathrm{d}\mathbf{d} \right] \end{aligned} \quad (4.16)$$

On reconnaît l'expression de l'entropie de Shannon. En simplifiant l'expression et en développant la distribution jointe, on obtient :

$$\langle I \rangle = - \int_{\mathcal{M}} \pi(\mathbf{m}) \mathcal{E}_{\mathbf{d}|\mathbf{m}}\{\rho\} - \int_{\mathcal{D}} \int_{\mathcal{M}} \rho(\mathbf{m}|\mathbf{d})\sigma(\mathbf{d}) \log \frac{\sigma(\mathbf{d})}{\mu(\mathbf{d})} \, \mathrm{d}\mathbf{m} \, \mathrm{d}\mathbf{d} \quad (4.17)$$

$$= - \int_{\mathcal{M}} \pi(\mathbf{m}) \mathcal{E}_{\mathbf{d}|\mathbf{m}}\{\rho\} - \int_{\mathcal{D}} \sigma(\mathbf{d}) \log \frac{\sigma(\mathbf{d})}{\mu(\mathbf{d})} \int_{\mathcal{M}} \rho(\mathbf{m}|\mathbf{d}) \, \mathrm{d}\mathbf{m} \, \mathrm{d}\mathbf{d} \quad (4.18)$$

Enfin, en utilisant le fait que la distribution $\rho(\mathbf{m}|\mathbf{d})$ est normalisée, on obtient le résultat souhaité :

$$\langle I \rangle = \mathcal{E}_{\mathbf{d}}\{\sigma\} - \mathbb{E}^{\pi} [\mathcal{E}_{\mathbf{d}|\mathbf{m}}\{\rho\}] \quad (4.19)$$

L'information moyenne est donc définie comme la différence entre l'incertitude sur la distribution des observations et ce que l'on attend *a priori* de l'incertitude sur la distribution de vraisemblance - information apportée par le modèle.

L'égalité est obtenue dans l'espace des observations, ce qui permet de réduire considérablement les dimensions des matrices de covariances mises en jeu.

Nous disposons d'un résultat théorique pour le gain d'information moyenne. On pourra alors maximiser ce critère par rapport à l'information *a priori* permettant de définir et de construire un échantillonnage optimal pour les observations.

Nous voyons que les premières difficultés seront de calculer l'entropie des différentes distributions de probabilité. Nous pouvons, dans le cas des distributions gaussiennes, calculer analytiquement ces entropies.

3 Comportement asymptotique de l'information moyenne

3.1 Fonctionnelle d'information

La fonctionnelle d'information que nous avons définie appartient à une famille plus large. Dans cette section, nous introduisons la famille de fonctionnelle d'information proposée par Venegas-Martinez [17]. Pour simplifier, on considère le cas où le paramètre \mathbf{m} est un réel, $M = 1$. L'extension au cas multidimensionnel dans le cas de mesures de l'information amène des difficultés conceptuelles discutées dans [18], [19], [20] et [21].

On suppose que le paramètre $\mathbf{m} \in \mathcal{M} \subseteq \mathbb{R}$ est de probabilité $\mathbb{P}_{\mathbf{m}}$ dont il est possible d'observer le vecteur \mathbf{d} et on suppose que cette probabilité a une distribution $\rho(\mathbf{d}|\mathbf{m})$ au sens de la dérivée de Radon-Nykodym par rapport à une mesure dominante σ -finie λ sur \mathbb{R} pour tout $\mathbf{m} \in \mathcal{M} \subseteq \mathbb{R}$:

$$\forall \mathbf{m} \in \mathcal{M} \subseteq \mathbb{R}, \quad \frac{d\mathbb{P}_{\mathbf{m}}}{d\lambda} = \rho(\mathbf{d}|\mathbf{m}) \quad (4.20)$$

Nous avons vu que l'approche bayésienne consiste en la définition d'une distribution *a priori* $\pi(\mathbf{m})$ décrivant l'information initiale disponible sur la vraisemblance de la valeur du paramètre \mathbf{m} . On pose alors que cette distribution est une distribution σ -finie par rapport à une mesure dominante μ sur \mathbb{R} . Une fois la distribution *a priori* définie, le théorème de Bayes nous permet de mettre à jour les paramètres du modèle par l'intégration des observations via la distribution *a posteriori*. Venegas-Martínez définit un système de fonctionnelles d'intérêt noté $\mathcal{V}_{\gamma,\alpha,\delta}$ et défini par :

$$\mathcal{V}_{\gamma,\alpha,\delta}(\pi) = \frac{1}{1-\gamma} \int_{\mathcal{M}} \pi(\mathbf{m}) G_{\gamma,\alpha,\delta}(\mathcal{I}(\mathbf{m}), \mathcal{E}\{\mathbf{d}|\mathbf{m}\}) d\mu(\mathbf{m}) \quad (4.21)$$

où $\mathcal{I}(\mathbf{m})$ est l'information de Fisher définie en (A.6) et $\mathcal{E}_{\mathbf{d}|\mathbf{m}}\{\rho\}$ est la neg-entropie¹ de

1. Entropie négative

Shannon de la distribution $\rho(\mathbf{d}|\mathbf{m})$ et :

$$G_{\gamma,\alpha,\delta}(\mathcal{I}(\mathbf{m}), \mathcal{E}_{\mathbf{d}|\mathbf{m}}\{\rho\}) = \log \left(\frac{\exp \left((\mathcal{E}_{\mathbf{d}|\mathbf{m}}\{\rho\}/\mathcal{I}(\mathbf{m}))^{1-\delta} (\mathcal{I}(\mathbf{m}))^{\frac{1-\gamma}{1+\alpha}} - \delta (\mathcal{I}(\mathbf{m}))^{1-\alpha} \right)}{\pi(\mathbf{m})^{1-\gamma}} \right) \quad (4.22)$$

3.2 Critères d'information

La fonctionnelle (4.21) permet de construire plusieurs critères d'information. On remarque que :

$$\mathcal{V}_{0,0,1}(\pi(\mathbf{m})) = - \int_{\mathcal{M}} \pi(\mathbf{m}) \log \pi(\mathbf{m}) d\mu(\mathbf{m}) \quad (4.23)$$

est la mesure d'information de Shannon ou la fonctionnelle de Jaynes.

$$\mathcal{V}_{1,1,1}(\pi(\mathbf{m})) = \lim_{\gamma \rightarrow 1} \mathcal{V}_{\gamma,1,1}(\pi(\mathbf{m})) = - \int_{\mathcal{M}} \pi(\mathbf{m}) \log \frac{\pi(\mathbf{m})}{(\mathcal{I}(\mathbf{m}))^{1/2}} d\mu(\mathbf{m}) \quad (4.24)$$

est la cross-entropie négative de Good dans son approche MEP (*Minimax Evidence Priors*) qui peut être réécrite à l'aide de l'information de Shannon et on obtient alors :

$$\mathcal{V}_{1,1,1} = \mathcal{V}_{0,0,1} + \int_{\mathcal{M}} \pi(\mathbf{m}) \log (\mathcal{I}(\mathbf{m}))^{1/2} d\mu(\mathbf{m}) \quad (4.25)$$

Le cas asymptotique où $(\gamma, \alpha, \delta) \rightarrow (0, 0, 0)$ définit la fonctionnelle suivante :

$$\mathcal{V}_{0,0,0}(\pi(\mathbf{m})) = \int_{\mathcal{D}} \sigma(\mathbf{d}) \int_{\mathcal{M}} \log \frac{\rho(\mathbf{d}|\mathbf{m})}{\pi(\mathbf{m})} d\mu(\mathbf{m}) d\lambda(\mathbf{d}) \quad (4.26)$$

Cette fonctionnelle est le critère d'information défini par Zellner [22] dans son approche MDIP (*Maximal Data Information Priors*). Cette dernière fonctionnelle peut se réécrire facilement avec la mesure d'information de Shannon. On obtient alors :

$$\mathcal{V}_{0,0,0}(\pi(\mathbf{m})) = \mathcal{V}_{0,0,1}(\pi(\mathbf{m})) + \mathbb{E}^{\pi}\{\rho(\mathbf{d}|\mathbf{m})\} \quad (4.27)$$

La méthode développée par Zellner a été proposée dans le but de choisir une distribution qui maximise l'information moyenne de la distribution de vraisemblance par rapport à l'information contenue dans la distribution *a priori*.

Ce dernier critère peut être modifié en supposant que l'information de Fisher fournie par le modèle d'échantillonnage est significative. On définit le critère d'information suivant :

$$\mathcal{V}_{0,1,0}(\pi(\mathbf{m})) = \int_{\mathcal{D}} \sigma(\mathbf{d}) \int_{\mathcal{M}} \rho(\mathbf{m}|\mathbf{d}) \log \frac{\rho(\mathbf{d}|\mathbf{m})^{\mathcal{I}(\mathbf{m})^{1/2}}}{\pi(\mathbf{m})} d\mu(\mathbf{m}) d\lambda(\mathbf{d}) \quad (4.28)$$

Ces différentes fonctionnelles introduites par Venegas-Martínez permettent de donner un résultat important dans le cas unidimensionnel pour le critère d'information

moyenne défini en (4.7). Ce dernier a été introduit par Bernardo [23] pour contourner la difficulté dans la définition de l'*a priori* de Jeffrey dans le cas multidimensionnel et développé par la suite par Berger et Bernardo [24],[25],[26] et [27]. Ils définissent la notion d'*information manquante* sur le paramètre \mathbf{m} lors d'une expérience et construisent un *a priori de référence*. L'information manquante est quantifiée par (4.7) où l'espérance est prise par rapport à la distribution marginale sur les observations $\sigma(\mathbf{d}) = \int_{\mathcal{M}} \rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m})$.

3.3 Cas d'un paramètre unidimensionnel

Supposons désormais que l'on dispose d'un échantillon $\mathbf{d} = (d_1, \dots, d_N)$ de N observations indépendantes de distribution $\mathbb{P}_{\mathbf{m}}$, $\mathbf{m} \in \mathcal{M} \subseteq \mathbb{R}$. On a alors :

$$\frac{d\mathbb{P}_{\mathbf{m}}}{d\lambda} = \rho(\mathbf{d}|\mathbf{m}) = \prod_{k=1}^{k=N} \rho(d_k, \dots, d_N|\mathbf{m}) \quad (4.29)$$

pour tout $\mathbf{m} = (d_1, \dots, d_N)$ et tout $\mathbb{P}_{\mathbf{m}}$, $\mathbf{m} \in \mathcal{M} \subseteq \mathbb{R}$.

Venegas-Martínez donne un résultat asymptotique de (4.7) dans le cas unidimensionnel. Ce critère, appelé aussi mesure d'information de Lindley [1], est l'information moyenne sur le paramètre \mathbf{m} du modèle d'échantillonnage $\rho(\mathbf{d}|\mathbf{m})$ fourni par un échantillon \mathbf{d} de taille N . On suppose que la distribution *a priori* sur \mathbf{m} est $\pi(\mathbf{m})$. On donne explicitement le critère défini plus haut tel que :

$$\langle I \rangle^N(\pi) = \int_{\mathcal{D}} \sigma(\mathbf{d}) \int_{\mathcal{M}} \rho(\mathbf{m}|\mathbf{d}) \log \frac{\rho(\mathbf{m}|\mathbf{d})}{\pi(\mathbf{m})} d\mu(\mathbf{m}) d\lambda(\mathbf{d}) \quad (4.30)$$

où μ et λ sont des mesures de Lebesgue sur \mathbb{R} . On suppose par la suite que toutes les densités considérées sont Lebesgue-mesurables par rapport à \mathbf{m} et \mathbf{d} . Le théorème montre que l'*essentiel* du modèle statistique $\rho(\mathbf{d}|\mathbf{m})$ change lorsque la taille des échantillons augmente. Pour montrer cela, Venegas-Martínez réécrit (4.30) à l'aide des critères d'information définis plus haut. Ainsi :

$$\langle I \rangle^N(\pi) = \mathcal{V}_{\gamma,0,1}(\pi) + \log \sqrt{N} \quad (4.31)$$

$$- \int_{\mathcal{M}} \int_{\mathcal{D}} \log \left(\int T_N(w) W_N(w) d\mu(w) \right) \rho(\mathbf{d}|\mathbf{m}) \pi(\mathbf{m}) d\mu(\mathbf{m}) d\lambda(\mathbf{d}) \quad (4.32)$$

où

$$T_N(w) = \frac{\rho(\mathbf{d}|\mathbf{m} + \frac{w}{\sqrt{N}})}{\rho(\mathbf{d}|\mathbf{m})} \quad (4.33)$$

et

$$W_N(w) = \frac{\mathbf{m} + \frac{w}{\sqrt{N}}}{\mathbf{m}} \quad (4.34)$$

En supposant des conditions standards de régularité sur l'information de Fisher et des propriétés de convergence pour T_N et W_N , Venegas-Martínez montre que la convergence

en loi est telle que :

$$T_N(w) \xrightarrow{\mathcal{L}} \exp \left(w \sqrt{\mathcal{I}(\mathbf{m})} \left(Z - \frac{1}{2} w \sqrt{\mathcal{I}(\mathbf{m})} \right) \right) \quad (4.35)$$

où $Z \sim \mathcal{N}(0, 1)$. Ceci implique la convergence en loi :

$$\log \left(\int T_N(w) W_N(w) d\mu(w) \right) \xrightarrow{\mathcal{L}} \log \sqrt{2\pi/\mathcal{I}(\mathbf{m})} + \frac{1}{2} Z^2 \quad (4.36)$$

Ce qui permet d'obtenir au premier ordre le cas asymptotique :

$$\langle I \rangle^{(\infty)}(\pi) - \mathcal{V}_{1,1,1}(\pi) = -\mathcal{V}_{0,0,1}(\varphi) + \log \left(C \sqrt{N} \right) + o(1) \quad (4.37)$$

où $\varphi(z)$ est la distribution de $Z \sim \mathcal{N}(0, 1)$ et :

$$C = \left(\int \sqrt{\mathcal{I}(\mathbf{m})} d\mu(\mathbf{m}) \right)^{-1} \quad (4.38)$$

Le résultat montre la dépendance de l'information moyenne avec l'entropie croisée de Good définie par (4.24) qui est quant à elle indépendante du nombre de mesures N . Le cas asymptotique illustre le comportement logarithmique de l'information moyenne apportée par les mesures lorsque celles-ci augmentent.

3.4 Cas d'un paramètre multidimensionnel

Dans le cas multidimensionnel, le développement asymptotique de l'information moyenne se complique et les conditions de régularité sur les distributions ainsi que sur les limites sont plus fortes. Ghosal [28] donne une approximation au premier ordre pour le cas asymptotique $N \rightarrow \infty$ de $\langle I \rangle$ dans le cas multidimensionnel. Reprenant les travaux initiaux de Bernardo [23], Ghosal calcule, dans le cas d'un ouvert compact K sur l'ensemble des paramètres, cette approximation. Il sépare l'espace des paramètres en deux sous-espaces tels que $\mathcal{M} = \mathcal{M}_0 \times \mathcal{M}_M$ et tel que $\mathbf{m} = \{m_0\} \cup \{\mathbf{m} : \mathbf{m} = (\mathbf{m}_1, \dots, \mathbf{m}_M)\}$ et il suppose que la densité $\rho(\cdot|m_0, \mathbf{m})$ par rapport à la mesure de Lebesgue sur \mathbb{R} est régulière par rapport à \mathbf{m} et non-régulière par rapport à m_0 . Ainsi, la fonction $\mathbf{d} \rightarrow \rho(\mathbf{d}|m_0, \mathbf{m})$ est supposée continue pour tout \mathbf{m} à m_0 fixé et discontinue en des points dépendants seulement de m_0 . Les propriétés asymptotiques de ce genre de famille de distributions ont été étudiées par Ghosal et Samanta [29] utilisant la théorie développée par Ibragimov et Has'minskii [30] et Smith [31]. Ghosal suppose que la densité $\rho(\mathbf{d}|m_0, \mathbf{m})$ est à support sur l'intervalle $S(m_0) = [a_1(m_0), a_2(m_0)]$. On peut avoir la situation où $a_1 = -\infty$ ou $a_2 = +\infty$ mais pas de façon simultanée et Ghosal suppose que les deux cas limites suivants existent :

Pour tout $(m_0; \mathbf{m}) \in \mathcal{M}_0 \times \mathcal{M}_M$,

$$p(m_0, \mathbf{m}) = \lim_{m_0 \downarrow a_1(m_0)} \rho(\mathbf{d}|m_0, \mathbf{m}), \quad q(m_0, \mathbf{m}) = \lim_{m_0 \uparrow a_2(m_0)} \rho(\mathbf{d}|m_0, \mathbf{m}) \quad (4.39)$$

En choisissant :

$$c(m_0, \mathbf{m}) = a'_1(m_0)p(m_0, \mathbf{m}) - a'_2(m_0)q(m_0, \mathbf{m}) > 0 \quad (4.40)$$

le critère d'information moyenne défini comme l'espérance de la divergence de Kullback-Leibler entre la distribution *a posteriori* et la distribution *a priori*, admet un développement au premier ordre.

En supposant que pour toute distribution *a priori* $\pi(\cdot)$ positive, continue et dense sur un sous-ensemble compact K tel que $K \subset \mathcal{M}_0 \times \mathcal{M}_M$, $\langle I \rangle$ a alors lorsque $N \rightarrow \infty$, le développement asymptotique suivant :

$$\begin{aligned} \langle I \rangle = & \log \frac{N}{e} + \frac{d}{2} \log \frac{N}{2\pi e} \\ & + \int_K \pi(m_0, \mathbf{m}) \log \left(\frac{c(m_0, \mathbf{m}) |\mathcal{I}(m_0, \mathbf{m})|^{1/2}}{\pi(m_0, \mathbf{m})} \right) dm_0 d\mathbf{m} + o(1) \end{aligned} \quad (4.41)$$

où $|\mathcal{I}(m_0, \mathbf{m})|^{1/2}$ est le déterminant de la matrice de Fisher pour le paramètre (m_0, \mathbf{m}) . La démonstration est donnée par Ghosal dans [28].

Ghosal montre dans le cas multidimensionnel, que le terme dominant de l'ordre de $\log N$ est indépendant de la distribution *a priori*. Le résultat précédent montre bien le comportement logarithmique de l'information moyenne dans le cas multidimensionnel.

Dans notre cas, plus général, où la distribution sur le paramètre \mathbf{m} n'est pas supposée à support borné, les résultats de convergence ne peuvent pas être utilisés directement. Nous allons expliciter le critère d'information moyenne dans les chapitres suivants.

Chapitre 5

Modèle linéarisé Gaussien

Sommaire

1	Modèle de régression	34
2	Noyau de covariance	34
3	Résolution du problème inverse linéarisé	35
3.1	Définition de la distribution <i>a priori</i>	35
3.2	Linéarisation autour du maximum de vraisemblance	35
3.3	Linéarisation autour de la moyenne <i>a priori</i>	36
4	Quantité d'information	37
4.1	Apport d'information	37
4.2	Apport d'information moyenne	37

Dans ce chapitre nous donnons une définition de l'information moyenne lorsque le modèle est linéarisé. Nous supposons que les erreurs sur les observations sont gaussiennes.

En ingénierie pétrolière, les opérateurs de modélisation sont très sensibles aux conditions initiales et rendent la modélisation délicate. Dans le cadre bayésien, les distributions *a priori* jouent un rôle déterminant dans la construction de la distribution *a posteriori* ainsi que dans la mise à jour et l'estimation des paramètres du modèle. On cherchera donc à réduire l'influence de la non-linéarité des opérateurs mathématiques mis en jeu en les linéarisant autour de la moyenne *a priori* ou du maximum de vraisemblance. De plus, la distribution *a priori* sur le modèle de paramètres¹ est généralement considérée normale ou log-normale, ce qui permet d'explicitier la distribution *a posteriori* résultante. Ainsi, le modèle linéarisé Gaussien est utilisé fréquemment en première approximation car il permet d'obtenir directement des résultats analytiques pratiques ainsi que des résultats numériques en optimisation. Nous allons décrire brièvement celui-ci dans un premier temps, et dans un second temps, nous expliciterons ces critères d'optimisation.

1. On considère ici comme paramètres la perméabilité ou la porosité du milieu

1 Modèle de régression

Reprenons le modèle général tel que :

$$\mathbf{d}^{obs} = \mathcal{F}[\mathbf{m}] + \boldsymbol{\eta}$$

où $\mathbf{d}^{obs} \in \mathcal{D} \subseteq \mathbb{R}^N$ et $\mathbf{m} \in \mathcal{M} \subseteq \mathbb{R}^M$.

On considère le cas où les erreurs du modèle sont supposées gaussiennes centrées, de matrice de covariance \mathbf{C}_d :

$$\boldsymbol{\eta} \sim \mathcal{N}_N(0, \mathbf{C}_d) \quad (5.1)$$

De plus, on suppose que $\boldsymbol{\eta}$ est stationnaire. Ces hypothèses sont couramment utilisées mais il faut noter qu'elles sont fortes.

Le cas stationnaire impose que la covariance des erreurs entre deux points d'observations est fonction de leur distance, c'est-à-dire qu'elle est invariante par translation :

$$\forall \mathbf{x}, \mathbf{y} \in \mathcal{D}, \quad \mathbf{C}_d(\mathbf{x}, \mathbf{y}) = \psi(|\mathbf{x} - \mathbf{y}|)$$

Si de plus, on prend la distance euclidienne sur l'espace des observations, alors on se trouve dans le cas isotrope. Dans le cas contraire, on pourra toujours appliquer une transformation linéaire nous permettant de nous rapprocher du cas isotrope.

Ayant défini $\boldsymbol{\eta}$ gaussien, on déduit que la variable aléatoire $\mathbf{d}^{obs}|\mathbf{m}$ est gaussienne de loi :

$$\mathbf{d}^{obs}|\mathbf{m} \sim \mathcal{N}_N(\mathcal{F}[\mathbf{m}], \mathbf{C}_d) \quad (5.2)$$

La distribution de vraisemblance est donc :

$$\begin{aligned} \rho(\mathbf{d}^{obs}|\mathbf{m}) &= ((2\pi)^N |\mathbf{C}_d|)^{-1/2} \exp \left[\frac{-1}{2} (\mathbf{d}^{obs} - \mathcal{F}[\mathbf{m}])^t \mathbf{C}_d^{-1} (\mathbf{d}^{obs} - \mathcal{F}[\mathbf{m}]) \right] \\ &= ((2\pi)^N |\mathbf{C}_d|)^{-1/2} \exp \frac{-1}{2} \|\mathbf{d}^{obs} - \mathcal{F}[\mathbf{m}]\|_{\mathbf{C}_d^{-1}} \end{aligned} \quad (5.3)$$

où $\|\mathbf{x}\|_A^2 = \mathbf{x}^t A \mathbf{x}$.

2 Noyau de covariance

La matrice de covariance est définie positive et permet de construire une distance sur l'espace des observations. Lorsque l'on choisit le noyau de covariance, on impose implicitement la régularité du processus aléatoire. Ainsi si la covariance admet des dérivées d'ordre 2 alors le processus sera différentiable en moyenne quadratique. Si le phénomène physique admet des dérivées à tout ordre, il faudra choisir un noyau indéfiniment différentiable. Il faudra donc adapter le noyau de covariance à notre phénomène physique.

Une classe générale de noyaux admissibles est la classe de *Matérn* définie pour toute dimension de l'espace des observations par :

$$\mathbf{C}(h) = \frac{\sigma^2}{2^{\nu-1} \Gamma(\nu)} \left(\frac{2\nu^{1/2} h}{\rho} \right)^\nu \mathcal{K}_\nu \left(\frac{2\nu^{1/2} h}{\rho} \right) \quad (5.4)$$

où h est la distance Euclidienne, σ^2 est la variance du processus, alors que ρ et ν sont respectivement la largeur du noyau (ou longueur de corrélation) et sa régularité. $\mathcal{K}_\nu(\cdot)$ est la fonction de Bessel modifiée de seconde espèce².

3 Résolution du problème inverse linéarisé

La résolution du problème inverse dans le cadre bayésien passe par la définition d'une distribution *a priori* sur l'espace des paramètres \mathcal{M} . Les premiers résultats sont donnés dans un cadre informatif, c'est-à-dire le cas où une information *a priori* est disponible.

3.1 Définition de la distribution *a priori*

En pratique, il est fréquent de choisir la distribution *a priori* $\pi(\mathbf{m})$ gaussienne, de paramètres \mathbf{m}_p et \mathbf{C}_p respectivement moyenne et fonction de covariance. Soit \mathbf{m} de loi :

$$\mathbf{m} \sim \mathcal{N}_M(\mathbf{m}_p, \mathbf{C}_p) \quad (5.5)$$

La distribution est définie comme suit :

$$\begin{aligned} \pi(\mathbf{m}) &= ((2\pi)^M |\mathbf{C}_p|)^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{m} - \mathbf{m}_p)^t \mathbf{C}_p^{-1} (\mathbf{m} - \mathbf{m}_p)\right] \\ &= ((2\pi)^M |\mathbf{C}_p|)^{-1/2} \exp\left[-\frac{1}{2} \|\mathbf{m} - \mathbf{m}_p\|_{\mathbf{C}_p^{-1}}^2\right] \end{aligned} \quad (5.6)$$

Remarque 5.1 *La définition de la distribution a priori est subjective et donc très discutée en inversion bayésienne. Dans notre première approche, nous avons défini une distribution a priori de la même famille (gaussienne) que la distribution de vraisemblance et par conséquent la distribution a posteriori résultante sera également définie dans la même famille. On parle alors d'a priori conjugué [32].*

3.2 Linéarisation autour du maximum de vraisemblance

On peut linéariser l'opérateur \mathcal{F} autour du maximum de vraisemblance $\tilde{\mathbf{m}}$ à l'aide du développement de Taylor à l'ordre 1 :

$$\mathcal{F}[\mathbf{m}] = \mathcal{F}[\tilde{\mathbf{m}}] + \mathcal{H} \cdot (\mathbf{m} - \tilde{\mathbf{m}}) \quad (5.7)$$

où $\mathcal{H} = \left(\frac{\partial \mathcal{F}}{\partial \mathbf{m}}\right)_{\mathbf{m}=\tilde{\mathbf{m}}}$ est le gradient de l'opérateur \mathcal{F} en $\tilde{\mathbf{m}}$.

Le problème de régression est alors :

$$\mathbf{d}^{obs} - (\mathcal{F}[\tilde{\mathbf{m}}] - \mathcal{H} \cdot \tilde{\mathbf{m}}) = \mathcal{H} \cdot \mathbf{m} + \boldsymbol{\eta} \quad (5.8)$$

2. $\mathcal{K}_\nu(x) = \frac{\pi}{2} \frac{I_{-\nu}(x) - I_\nu(x)}{\sin(\pi\nu)}$ où $I_\nu = \left(\frac{x}{2}\right)^\nu \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(\nu + k + 1)} \left(\frac{x}{2}\right)^{2k}$

Soit en posant $\mathbf{d} = \mathbf{d}^{obs} - (\mathcal{F}[\tilde{\mathbf{m}}] - \mathcal{H} \cdot \tilde{\mathbf{m}})$, on obtient le problème linéaire suivant :

$$\mathbf{d} = \mathcal{H} \cdot \mathbf{m} + \boldsymbol{\eta} \quad (5.9)$$

Alors, $\tilde{\mathbf{m}}$ est solution du problème de minimisation :

$$\tilde{\mathbf{m}} = \arg \min_{\mathbf{m} \in \mathcal{M}} \|\mathbf{d} - \mathcal{H} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 = (\mathcal{H}^t \mathbf{C}_d^{-1} \mathcal{H})^{-1} \mathcal{H}^t \mathbf{C}_d^{-1} \mathbf{d} \quad (5.10)$$

On remarque que la distribution de la variable aléatoire $\mathbf{d}^{obs} | \mathbf{m}$ est égale à la distribution de la variable aléatoire $\mathbf{d} | \mathbf{m}$:

$$\rho(\mathbf{d}^{obs} | \mathbf{m}) = ((2\pi)^N |\mathbf{C}_d|)^{-1/2} \exp \frac{-1}{2} \|\mathbf{d}^{obs} - \mathcal{F}[\mathbf{m}]\|_{\mathbf{C}_d^{-1}}^2 \quad (5.11)$$

$$= ((2\pi)^N |\mathbf{C}_d|)^{-1/2} \exp \frac{-1}{2} \|\mathbf{d} - \mathcal{H} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 = \rho(\mathbf{d} | \mathbf{m}) \quad (5.12)$$

En posant π la distribution *a priori* sur le paramètre \mathbf{m} , on calcule la distribution *a posteriori* par :

$$\rho(\mathbf{m} | \mathbf{d}) \propto \pi(\mathbf{m}) \rho(\mathbf{d} | \mathbf{m}) = \pi(\mathbf{m}) ((2\pi)^N |\mathbf{C}_d|)^{-1/2} \exp \frac{-1}{2} \|\mathbf{d} - \mathcal{H} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 \quad (5.13)$$

On remarque que :

$$\|\mathbf{d} - \mathcal{H} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 = \|\mathbf{d} - \mathcal{H} \cdot \tilde{\mathbf{m}}\|_{\mathbf{C}_d^{-1}}^2 + \|\tilde{\mathbf{m}} - \mathbf{m}\|_{\mathcal{H}^t \mathbf{C}_d^{-1} \mathcal{H}}^2 \quad (5.14)$$

Par substitution de (5.14) dans l'expression précédente (5.13), on observe que la distribution *a posteriori* peut être exprimée comme le produit de deux distributions gaussiennes :

$$\rho(\mathbf{m} | \mathbf{d}) \propto \pi(\mathbf{m}) |\mathcal{H}^t \mathbf{C}_d^{-1} \mathcal{H}|^{1/2} \mathcal{N}_d(\mathcal{H} \cdot \tilde{\mathbf{m}}, \mathbf{C}_d^{-1}) \otimes \mathcal{N}_m(\tilde{\mathbf{m}}, \mathcal{H}^t \mathbf{C}_d^{-1} \mathcal{H}) \quad (5.15)$$

On peut enfin noter que $|\mathcal{H}^t \mathbf{C}_d^{-1} \mathcal{H}|$ est le déterminant de la matrice d'information de Fisher de la variable $\mathbf{d} | \mathbf{m}$ dans le cas linéaire. La linéarisation autour du maximum de vraisemblance ne donne pas directement une distribution *a posteriori* gaussienne. Pour plus de simplicité, nous décidons de linéariser l'opérateur autour de la moyenne *a priori* \mathbf{m}_p comme nous allons le voir.

3.3 Linéarisation autour de la moyenne *a priori*

On reprend la distribution de vraisemblance dans le cas linéaire :

$$\rho(\mathbf{d} | \mathbf{m}) = ((2\pi)^N |\mathbf{C}_d|)^{-1/2} \exp \frac{-1}{2} \|\mathbf{d} - \mathcal{H} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 \quad (5.16)$$

où dans le cas de la linéarisation autour de la moyenne *a priori* :

$$\mathbf{d} = \mathbf{d}^{obs} - (\mathcal{F}[\mathbf{m}_p] - \mathcal{H} \cdot \mathbf{m}_p) \quad (5.17)$$

$$\text{et } \mathcal{H} = \left(\frac{\partial \mathcal{F}}{\partial \mathbf{m}} \right)_{\mathbf{m}=\mathbf{m}_p}.$$

Si nous considérons la distribution de vraisemblance gaussienne comme précédemment et la distribution *a priori* gaussienne, alors il en résulte une distribution *a posteriori* gaussienne de moyenne et covariance données par [11] :

$$\begin{cases} \mathbf{m}_q = \mathbf{m}_p + \mathbf{C}_{\mathcal{H}}^t (\mathcal{H}\mathbf{C}_p\mathcal{H}^t + \mathbf{C}_d)^{-1}(\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_p) \\ \mathbf{C}_q = \mathbf{C}_p - \mathbf{C}_p\mathcal{H}^t (\mathcal{H}\mathbf{C}_p\mathcal{H}^t + \mathbf{C}_d)^{-1}\mathcal{H}\mathbf{C}_p \end{cases}$$

4 Quantité d'information

On mesure la quantité d'information en calculant la divergence de Kullback-Leibler entre la distribution *a posteriori* et la distribution *a priori*.

4.1 Apport d'information

On montre que l'information apportée par une mesure gaussienne d'*a priori* gaussien (cf. Annexe C) est telle que :

$$I(\pi(\mathbf{m})) = \frac{1}{2} \log |\mathbf{C}_q^{-1}\mathbf{C}_p| + \frac{1}{2} \|\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_p\|_{\mathbf{H}}^2 - \frac{1}{2} \text{tr} \left[\mathcal{H}^t (\mathcal{H}\mathbf{C}_p\mathcal{H}^t + \mathbf{C}_d)^{-1} \mathcal{H}\mathbf{C}_p \right] \quad (5.18)$$

$$\text{où } \mathbf{H} = (\mathcal{H}\mathbf{C}_p\mathcal{H}^t + \mathbf{C}_d)^{-1} \mathcal{H} \mathbf{C}_p \mathcal{H}^t (\mathcal{H}\mathbf{C}_p\mathcal{H}^t + \mathbf{C}_d)^{-1}$$

On remarque que le premier terme de l'expression est le rapport de deux déterminants, respectivement les déterminants de la matrice de covariance *a priori* et de la matrice de covariance *a posteriori*. Dans le cas particulier des distributions gaussiennes, cela correspond à la différence d'entropie entre ces deux distributions.

Le deuxième terme de l'expression est plus familier puisqu'il représente l'écart entre les observations et l'approximation faite au 1er ordre. Il correspond à la norme au carré de l'erreur du modèle numérique.

La dernière expression renvoie à la modélisation et aux considérations faites sur les matrices de covariance *a priori* et des covariances d'erreurs d'observations.

4.2 Apport d'information moyenne

On peut déduire de (5.18) la quantité moyenne d'information apportée par une mesure gaussienne d'*a priori* gaussien en calculant l'espérance par rapport à la mesure

$\sigma(\mathbf{d})$ des observations. C'est-à-dire :

$$\begin{aligned}
\langle I \rangle &= \mathbb{E}[I \mid \mathbf{d}] = \int_{\mathcal{D}} I(\pi(\mathbf{m})) \sigma(\mathbf{d}) d\mathbf{d} \\
&= \frac{1}{2} \mathbb{E} \left[\log | \mathbf{C}_q^{-1} \mathbf{C}_p | + \|\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_p\|_{\mathbf{H}}^2 - \text{tr} \left[\mathcal{H}^t (\mathcal{H} \mathbf{C}_p \mathcal{H}^t + \mathbf{C}_d)^{-1} \mathcal{H} \mathbf{C}_p \right] \mid \mathbf{d} \right] \\
&= \frac{1}{2} \log | \mathbf{C}_q^{-1} \mathbf{C}_p | + \frac{1}{2} \mathbb{E} \left[\|\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_p\|_{\mathbf{H}}^2 \mid \mathbf{d} \right] - \text{tr} \left[\mathcal{H}^t (\mathcal{H} \mathbf{C}_p \mathcal{H}^t + \mathbf{C}_d)^{-1} \mathcal{H} \mathbf{C}_p \right] \\
&= \frac{1}{2} \log | \mathbf{C}_q^{-1} \mathbf{C}_p | + \frac{1}{2} \text{tr} \left[\mathbf{H} \mathbb{E} \left[\|\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_p\|^2 \mid \mathbf{d} \right] \right] - \frac{1}{2} \text{tr} \left[\mathcal{H}^t (\mathcal{H} \mathbf{C}_p \mathcal{H}^t + \mathbf{C}_d)^{-1} \mathcal{H} \mathbf{C}_p \right]
\end{aligned}$$

En remarquant que :

$$\mathbb{E} \left[\|\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_p\|^2 \mid \mathbf{d} \right] = \mathbf{C}_d + \mathcal{H} \mathbf{C}_p \mathcal{H}^t \quad (5.19)$$

on obtient après simplification que la quantité moyenne d'information est :

$$\langle I \rangle = \frac{1}{2} \log | \mathbf{C}_q^{-1} \mathbf{C}_p | \quad (5.20)$$

On retrouve simplement la différence d'entropie entre les distributions *a priori* et *a posteriori* dans le cas gaussien. D'un point de vue géométrique, le calcul de l'information moyenne définit un rapport de volume de corrélation dans l'espace des paramètres. Sachant que la matrice de covariance *a posteriori* est exprimée par rapport à la matrice de covariance *a priori*, il est possible de définir un *design* maximisant le critère de l'information moyenne avant toute optimisation.

En ingénierie de réservoir, le nombre de paramètres à estimer peut être très important. L'inversion de la matrice *a posteriori* devient ainsi une contrainte. Pour contourner ce problème, il serait plus utile de donner une expression de l'information moyenne non plus dans l'espace des paramètres mais dans l'espace des observations.

En reprenant l'égalité (5.20), on explicite l'information moyenne $\langle I \rangle$ telle que :

$$2\langle I \rangle = \log | \mathbf{C}_q^{-1} \mathbf{C}_p | = -\log | \mathbf{C}_p^{-1} \mathbf{C}_q | \quad (5.21)$$

$$= -\log | \mathbb{I}_M - \mathcal{F}^t (\mathcal{F} \mathbf{C}_p \mathcal{F}^t + \mathbf{C}_d)^{-1} \mathcal{F} \mathbf{C}_p | \quad (5.22)$$

$$= -\log | \mathbb{I}_N - \mathcal{F} \mathbf{C}_p \mathcal{F}^t (\mathcal{F} \mathbf{C}_p \mathcal{F}^t + \mathbf{C}_d)^{-1} | \quad (5.23)$$

$$= -\log | \mathbf{C}_d (\mathbf{C}_d + \mathcal{F} \mathbf{C}_p \mathcal{F}^t)^{-1} | \quad (5.24)$$

$$= \log | \mathbf{C}_d + \mathcal{F} \mathbf{C}_p \mathcal{F}^t | - \log | \mathbf{C}_d | \quad (5.25)$$

Après calculs algébriques et en utilisant l'égalité de Schur (Annexe D), on obtient l'égalité obtenue en (4.10)

$$\begin{aligned}
\langle I \rangle &= \mathcal{E}_{\mathbf{d}} \{ \sigma \} - \mathbb{E}^{\pi} \left[\mathcal{E}_{\mathbf{d} \mid \mathbf{m}} \{ \rho \} \right] \\
\langle I \rangle &= \frac{1}{2} \log | \mathbb{I}_N + \mathbf{C}_d^{-1} \cdot \mathcal{F} \mathbf{C}_p \mathcal{F}^t | \quad (5.26)
\end{aligned}$$

La taille de la matrice est alors de l'ordre du nombre d'observations. La seule la matrice à inverser est la matrice $\mathbf{C}_d \in \mathbb{R}^{N \times N}$.

Ainsi peut on directement utiliser cette mesure de l'apport d'information pour proposer un plan d'acquisition de mesures. Par exemple, pour améliorer la caractérisation de la perméabilité aux abords d'un puits, à quel fréquence faudrait-il réaliser une nouvelle mesure de pression ? Dans les chapitres suivants, nous allons présenter trois cas d'applications dans lesquels l'estimation de la perméabilité est obtenue grâce à un interpolateur linéaire puis une pression au puits et enfin une acquisition de données spatio-temporelles de sismique 3D répétées.

Chapitre 6

Optimisation du nombre de mesures : cas de l'estimation d'un champ de perméabilité à partir de mesures statiques

Sommaire

1	Estimateurs du champ de perméabilité	42
1.1	Estimation par cokrigeage simple	43
1.2	Estimation par maximum a posteriori	44
2	Optimisation de l'information	46
2.1	Cas asymptotique	46
2.2	Maximisation de l'information moyenne	46
2.3	Concentration des observations	50
3	Conclusion	51

Dans ce chapitre, nous présentons notre approche dans le cas du krigage. En effet, le krigage est très souvent utilisé en ingénierie pétrolière pour estimer un champ de perméabilité en tout point du réservoir, par interpolation des données obtenues par carottes et diagraphies aux puits.

Nous allons appliquer les résultats précédents dans un cadre linéaire simple. Le but est de déterminer la répartition spatiale optimale du nombre de mesures et ainsi définir un seuil au delà duquel l'ajout d'une nouvelle mesure n'influe pas sur la quantité d'information déjà présente. On suppose alors que le champ de perméabilité peut être représenté par un processus aléatoire. La distribution sous-jacente doit représenter au mieux les hétérogénéités du milieu. On suppose que l'on accorde un budget permettant d'effectuer un nombre fini $N \in \mathbb{N}$ de mesures locales. On cherche alors à définir l'influence du nombre N^* de volumes de corrélation c'est-à-dire le nombre de volumes indépendants présents en fonction des propriétés statistiques du champ de perméabilité

ainsi construit. Les observations ainsi obtenues servent à réduire les incertitudes. Nous allons expliciter deux techniques communément utilisées qui permettent de réduire ces incertitudes. La première, le krigeage, est largement utilisée en ingénierie pétrolière. La seconde, l'analyse bayésienne, donne un cadre plus général lorsque l'on travaille avec des distributions de probabilité. Le but est alors de déterminer une stratégie optimale d'acquisition de mesures et il faudra définir un critère d'optimisation exogène aux observations.

Le *spatial design* est un champ de recherche très actif et mature dans la littérature. Cressie [33] définit une approche complète en construisant un modèle aléatoire lui servant pour la prédiction de la concentration d'un polluant à des lieux où les observations ne sont pas disponibles. Il définit un échantillonnage optimal en minimisant la variance de krigeage ainsi qu'une mesure moyenne et une mesure maximum de l'erreur de prédiction. Il montre ainsi que celle-ci diminue lorsque l'on ajoute un site d'observation. Une approche plus récente utilisant le formalisme bayésien et la théorie de l'information a été proposé par Zideck [34]. Il calcule la quantité d'information apportée par la mesure de jauges. La méthode prend en compte les hyperparamètres, soient les paramètres des distributions de probabilité des inconnues, dans le processus d'estimation. Il considère un vecteur gaussien dont la covariance admet une distribution de Wishart. L'entropie est fonction des hyperparamètres de la matrice de Wishart et d'une fonction générique.

Dans notre étude, le variogramme est connu et ses paramètres sont estimés selon des considérations géologiques. Marchant [35] utilise une approche par krigeage classique à l'aide d'un variogramme d'ordre 1 (variogramme dit sphérique) pour la définition d'une fonction objectif en prenant en compte l'inverse de la matrice d'information de Fisher dans le but de construire un schéma optimal d'échantillonnage. La fonction objectif est similaire à (4.30). Dans un cadre complètement bayésien, Fuentes [36] utilise l'entropie de Shannon pour définir un réseau optimal tel que la perte d'information statistique est minimale.

Dans notre cas, les approches recherchant le maximum d'entropie ne convergent pas nécessairement vers une formation géologique réaliste et nous préférons utiliser la divergence de Kullback-Leibler. Dans ce cas, l'essentiel de l'information statistique est contenu dans notre modèle *a priori* et respecte donc l'information géologique.

1 Estimateurs du champ de perméabilité

Le krigeage est une méthode d'interpolation spatiale à base de noyaux. En modélisant la réponse du simulateur comme la réalisation d'un processus aléatoire, le krigeage permet de construire une prédiction du comportement du système simulé en des points où nous ne disposons pas d'observations.

Dans ce cas, l'expression de \mathcal{F} peut être modélisée à l'aide de la fonction de Dirac. On attribue la valeur 1 aux points d'observations de coordonnées x_i $i = 1, \dots, N$ et 0 sinon. Les erreurs dues aux observations seront alors des erreurs de mesures.

La relation entre les observations et le processus de mesure peut être définie par la relation suivante :

$$\mathbf{d} = \mathcal{F} \cdot \mathbf{m} + \boldsymbol{\eta} \quad (6.1)$$

avec $\mathbf{d} \in \mathcal{D} \subseteq \mathbb{R}^N$ et $\mathbf{m} \in \mathcal{M} \subseteq \mathbb{R}^M$.

Dans le cas du krigeage, l'erreur vient des mesures et on prend généralement la distribution de $\boldsymbol{\eta}$ gaussienne centrée de covariance diagonale. Soit :

$$\boldsymbol{\eta} \sim \mathcal{N}_N(0, \sigma_d^2 \cdot \mathbb{I}_N) \quad (6.2)$$

où σ_d^2 est défini comme la variance d'erreur d'observation.

On considère alors les deux problématiques suivantes :

- (1) Étant donné un budget de N mesures, comment doit-on répartir ces N mesures afin qu'elles maximisent l'information moyenne disponible ?
- (2) Étant donné un champ contenant $N - 1$ puits, où doit-on placer le N -ème puits afin qu'il maximise le critère d'information moyenne ?

Nous présentons maintenant deux techniques d'estimation de la perméabilité à partir d'observations : le cokrigeage simple et le maximum *a posteriori*.

1.1 Estimation par cokrigeage simple

Le krigeage est défini, nous l'avons vu, comme une méthode d'interpolation spatiale. Ainsi résoud-on un problème classique d'estimation d'une fonction en un point particulier à partir de valeurs connues de cette fonction mais en d'autres points.

Dans le cas d'un processus stationnaire \mathbf{m} , l'estimation par cokrigeage simple suppose connue la valeur du premier moment tel que :

$$\mathbb{E}\mathbf{m} = \mathbf{m}_p \quad (6.3)$$

Le krigeage produit une estimation \mathbf{m}^* de \mathbf{m} sur une cible comme illustré figure 6.1.

Le krigeage s'intéresse alors à l'erreur :

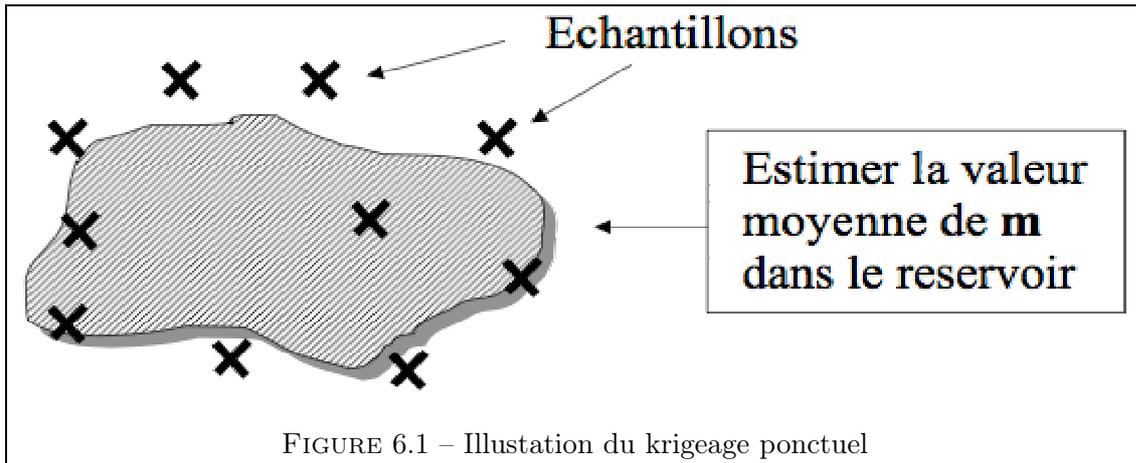
$$\boldsymbol{\eta} = \mathbf{m} - \mathbf{m}^* \quad (6.4)$$

En géostatistique linéaire, le krigeage obéit à des contraintes. Considéré comme un estimateur, le krigeage est tel que l'erreur d'estimation est :

- (1) une combinaison linéaire des données,
- (2) d'espérance nulle (nonbiais),
- (3) de variance minimale (optimalité).

Le krigeage postule que l'estimation de \mathbf{m} est obtenue par la combinaison linéaire suivante :

$$\mathbf{m}^* = \sum_{\alpha} \lambda_{\alpha} \mathbf{m}_{\alpha} + \mathbf{m}_p \quad (6.5)$$



où les pondérateurs λ_α sont les inconnues du problème. On cherche alors à minimiser la variance de l'erreur η .

On a alors les relations suivantes :

$$\mathbf{d} = \mathcal{F} \cdot \mathbf{m} + \boldsymbol{\eta} \quad \begin{cases} Cov(\mathbf{d}) = Cov(\mathcal{F} \cdot \mathbf{m} + \boldsymbol{\eta}) = \mathcal{F} \mathbf{C}_p \mathcal{F}^t + \mathbf{C}_d \\ Cov(\mathbf{d}\mathbf{m}) = \mathcal{F} \cdot \mathbf{C}_p \end{cases}$$

L'estimation \mathbf{m}^* de \mathbf{m} par cokrigeage simple donne alors :

$$\mathbf{m}^* = \mathbf{m}_p + \Lambda^t \cdot (\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_p) \quad (6.6)$$

avec $\Lambda = Cov(\mathbf{d})^{-1} \cdot Cov(\mathbf{d}\mathbf{m})$.

On pourra reprocher au modèle précédent le fait qu'il n'est pas général car il suppose un comportement raisonnablement linéaire de \mathcal{F} . Nous allons à présent utiliser le formalisme bayésien pour l'estimation du paramètre \mathbf{m} .

1.2 Estimation par maximum a posteriori

Dans le cas où la distribution de probabilité des erreurs d'observation est considérée comme une distribution gaussienne, le modèle précédent :

$$\mathbf{d} = \mathcal{F} \cdot \mathbf{m} + \boldsymbol{\eta} \quad (6.7)$$

est remplacé par la loi de probabilité gaussienne multidimensionnelle :

$$\mathbf{d} \sim \mathcal{N}_N(\mathcal{F} \cdot \mathbf{m}, \mathbf{C}_d) \quad (6.8)$$

où nous le rappelons, \mathcal{N}_N symbolise la loi multi-gaussienne de dimension N et \mathbf{C}_d est la fonction de covariance des erreurs d'observations.

Dans le cadre bayésien, une méthode d'estimation des paramètres du modèle consiste à déterminer l'ensemble des paramètres \mathbf{m} maximisant la distribution *a posteriori*. Dans le cas où la dimension de l'espace des paramètres est grande, cette méthode peut devenir coûteuse.

On cherche ainsi \mathbf{m}_{MAP} tel que :

$$\mathbf{m}_{MAP} = \arg \max_{\mathbf{m} \in \mathcal{M}} \frac{\rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m})}{\int_{\mathcal{M}} \rho(\mathbf{d}|\mathbf{m}')\pi(\mathbf{m}')d\mathbf{m}'} \quad (6.9)$$

Le dénominateur de la distribution *a posteriori* ne dépend pas de \mathbf{m} et ne joue pas ici un rôle dans la maximisation. On remarque que l'estimation par MAP coïncide avec l'estimation par maximum de vraisemblance lorsque la distribution *a priori* est uniforme.

Lorsque les distributions sont supposées gaussiennes, l'estimation par MAP revient au problème suivant :

$$\mathbf{m}_{MAP} = \arg \min_{\mathbf{m} \in \mathcal{M}} \left\{ \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 + \|\mathbf{m} - \mathbf{m}_p\|_{\mathbf{C}_p^{-1}}^2 \right\} \quad (6.10)$$

Ainsi dans le cas linéaire, l'estimateur MAP est donné par :

$$\mathbf{m}_{MAP} = (\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} + \mathbf{C}_p^{-1})^{-1} \cdot (\mathcal{F}^t \mathbf{C}_d^{-1} \cdot \mathbf{d} + \mathbf{C}_p^{-1} \cdot \mathbf{m}_p) \quad (6.11)$$

Dans le cas linéaire défini en (6.1), on pourra remarquer que l'estimation par cokrigage simple donne le même résultat que l'estimation par maximisation *a posteriori* :

$$\mathbf{m}_{MAP} = \mathbf{m}^* \quad (6.12)$$

Contrairement au krigeage, l'estimation par maximum *a posteriori* reste toujours possible lorsque \mathcal{F} est non-linéaire. On bénéficie de plus du formalisme probabiliste et d'une quantification des incertitudes dans le cas de l'estimation des modèles prédictifs. La méthode MAP apparait alors comme une méthode de régularisation dont la pénalité est $-\log(\pi(\mathbf{m}))$.

Le noyau de covariance *a priori* est généralement pris gaussien et ainsi l'expression de la fonction de covariance est alors :

$$\mathbf{C}_p(\mathbf{x}_i, \mathbf{x}_j) = \sigma_0^2 \cdot \exp - \left(\frac{\mathbf{x}_i - \mathbf{x}_j}{\lambda_c} \right)^2 \quad i, j \in [1, M] \quad (6.13)$$

où $\lambda_c \in]0, +\infty[$ est défini comme la longueur de corrélation entre \mathbf{x}_i et \mathbf{x}_j et σ_0^2 est la variance du champ de perméabilité. Avec ces considérations, l'information moyenne s'écrit explicitement :

$$\langle I \rangle = \frac{1}{2} \log \left| \mathbb{I}_N + \frac{\sigma_0^2}{\sigma_d^2} \mathcal{F} \cdot \left(\exp - \left(\frac{\mathbf{x}_i - \mathbf{x}_j}{\lambda_c} \right)^2 \right)_{i,j} \cdot \mathcal{F}^t \right| \quad i, j \in [1, M] \quad (6.14)$$

2 Optimisation de l'information

2.1 Cas asymptotique

Pour simplifier l'écriture, on pose $\mu^2 = \sigma_0^2/\sigma_d^2$ comme le rapport entre la variance *a priori* du paramètre \mathbf{m} et la variance d'erreurs sur les observations. Dans le cadre du krigeage simple, l'information moyenne est bornée par les deux cas asymptotiques suivants. Le cas où la longueur de corrélation est nulle et le cas où elle est infinie. On explicite l'information moyenne définie par :

$$\langle I \rangle_{\lambda_c \rightarrow 0} = \lim_{\lambda_c \rightarrow 0} \frac{1}{2} \log \begin{vmatrix} 1 + \mu^2 & \mu^2 \exp \frac{-\|\mathbf{x}_1 - \mathbf{x}_2\|^2}{\lambda_c^2} & \mu^2 \exp \frac{-\|\mathbf{x}_1 - \mathbf{x}_3\|^2}{\lambda_c^2} & \dots & \dots \\ \mu^2 \exp \frac{-\|\mathbf{x}_2 - \mathbf{x}_1\|^2}{\lambda_c^2} & 1 + \mu^2 & \mu^2 \exp \frac{-\|\mathbf{x}_2 - \mathbf{x}_3\|^2}{\lambda_c^2} & \dots & \dots \\ \mu^2 \exp \frac{-\|\mathbf{x}_3 - \mathbf{x}_1\|^2}{\lambda_c^2} & \mu^2 \exp \frac{-\|\mathbf{x}_3 - \mathbf{x}_2\|^2}{\lambda_c^2} & 1 + \mu^2 & \dots & \dots \\ \vdots & & & \ddots & \end{vmatrix} \quad (6.15)$$

L'information moyenne asymptotique est directement calculée grâce à cette dernière formule.

$$\lambda_c \rightarrow 0 \quad \langle I \rangle = \frac{1}{2} \log \left| \mathbb{I}_N + \mu^2 \cdot \mathbb{I}_N \right| = \frac{N}{2} \log(1 + \mu^2) \quad (6.16)$$

$$\lambda_c \rightarrow +\infty \quad \langle I \rangle = \frac{1}{2} \log \left| \mathbb{I}_N + \mu^2 \cdot \mathbf{1}_N \right| = \frac{1}{2} \log \left| 1 + N \cdot \mu^2 \right| \quad (6.17)$$

où l'on note $\mathbf{1}_N$ la matrice pleine dont tous les coefficients sont égaux à 1.

2.2 Maximisation de l'information moyenne

Dans le cas général, on cherche la répartition des N observations telle que la distribution résultante maximise l'information moyenne. Deux approches différentes d'optimisation peuvent être envisagées : séquentielle ou globale. On représente figure (6.2) un champ de perméabilité de variogramme gaussien.

2.2.1 Méthode séquentielle

La méthode séquentielle peut être utilisée pour deux problèmes différents. Dans un premier temps, on peut supposer qu'il existe au préalable des observations provenant de puits d'observations existants. Le problème revient alors à optimiser l'emplacement *a priori* d'une nouvelle observation à l'aide du critère d'information moyenne.

Dans ce cas, on cherche à optimiser l'emplacement \mathbf{x}_N de la N -ième observation sachant les $N - 1$ observations précédentes.

$$\mathbf{x}_N = \arg \max_{\mathbf{x} \in \mathcal{M}} \langle I(\mathbf{x} | \mathbf{x}_1, \dots, \mathbf{x}_{N-1}) \rangle$$

Dans un deuxième cas, on utilise la méthode séquentielle pour optimiser *a priori* les N observations en choisissant la première observation aléatoirement. On obtient alors

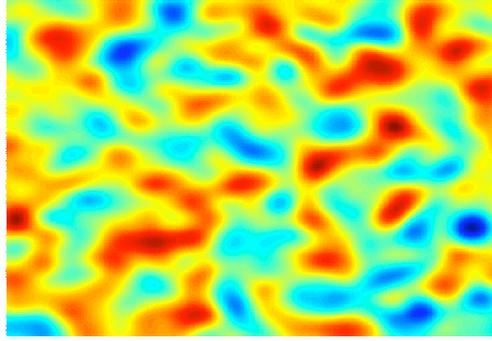


FIGURE 6.2 – Génération du champ de perméabilité de variogramme gaussien

le problème suivant. On cherche l'emplacement de N observations tel que :

$$\mathbf{x}_0 \text{ pris aléatoirement dans } \mathcal{M} \quad (6.18)$$

pour $i = 1, N$

$$\mathbf{x}_i = \arg \max_{\mathbf{x} \in \mathcal{M}} \langle I(\mathbf{x} | \mathbf{x}_1, \dots, \mathbf{x}_{i-1}) \rangle \quad (6.19)$$

fin

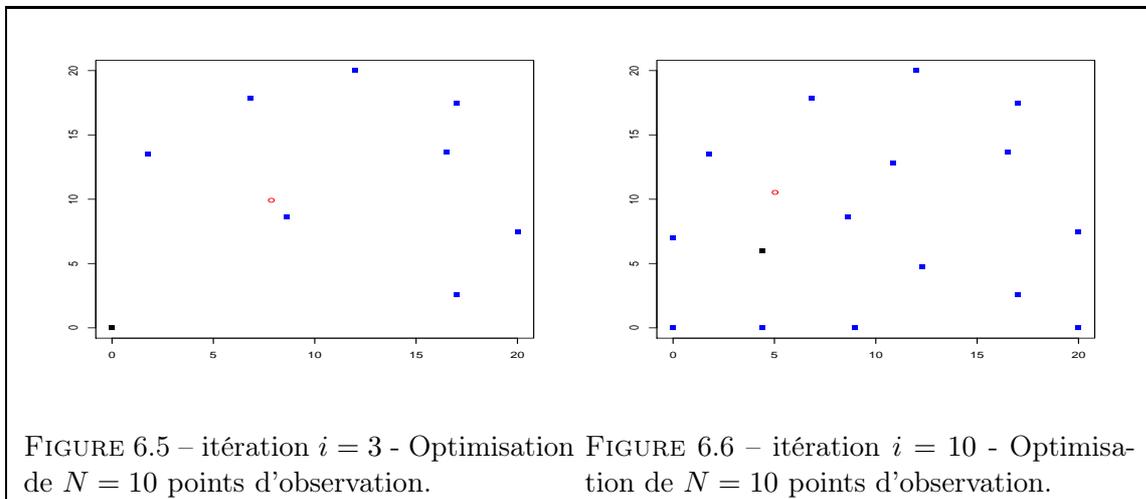
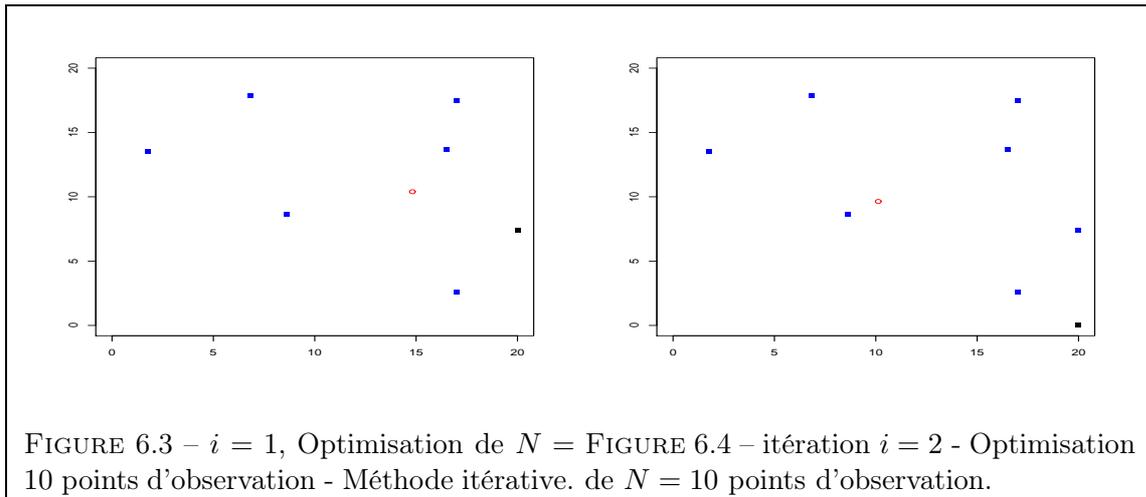
L'expression $\langle I(\mathbf{x} | \mathbf{x}_1, \dots, \mathbf{x}_{i-1}) \rangle$ est formulée par rapport à $i - 1$. On a alors :

$$\langle I(\mathbf{x} | \mathbf{x}_1, \dots, \mathbf{x}_{i-1}) \rangle = \frac{1}{2} \log \left| \mathbb{I}_N + \mathbf{C}_d^{-1} \mathcal{F} \cdot \begin{pmatrix} & & & \mathbf{x}_{1,i} \\ & & & \vdots \\ & \mathbf{C}_{i-1} & & \mathbf{x}_{i-1,i} \\ \mathbf{x}_{i,1} & \dots & \mathbf{x}_{i,i-1} & \mathbf{x}_{i,i} \end{pmatrix} \cdot \mathcal{F} \right|$$

où \mathbf{C}_{i-1} est la matrice de covariance *a priori* pour $i - 1$ mesures.

Le problème d'optimisation est alors un problème unidimensionnel toujours indépendant du maillage. On suppose initialement que l'on dispose de N_i observations et que l'on cherche à optimiser l'emplacement de N observations.

Dans l'exemple suivant, N_i est égal à 5 et N est égal à 10. Les carrés bleus sont les emplacements initiaux des observations. On cherche alors à optimiser (6.19). Le cercle rouge symbolise l'emplacement aléatoire du point courant \mathbf{x}_i . Quant au carré noir, il symbolise l'emplacement optimal. Le résultat pour $i = 10$ est représenté figure (6.6).

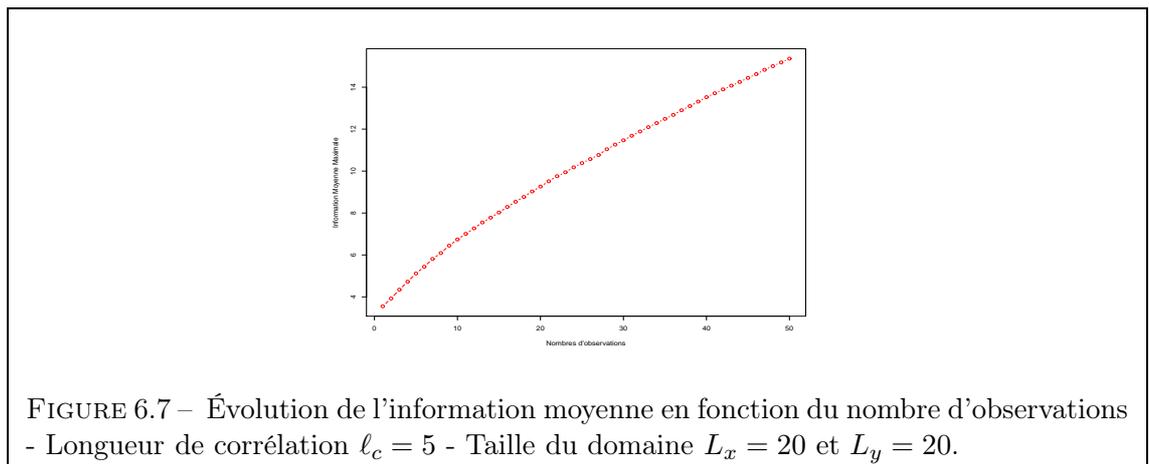


Que se passe-t-il lorsque nous augmentons le nombre d'observations? Existe-t-il un seuil où l'information moyenne n'évolue plus linéairement avec le nombre d'observations? On représente l'évolution de la valeur de l'information moyenne en fonction du nombre d'observations figure (6.7).

2.2.2 Méthode globale

La méthode séquentielle est une méthode de résolution élégante et fort adaptée au cas des gisements en cours de production pour lesquels des puits existent donc. Cependant après la phase d'exploration un premier plan de développement du champ comprenant le forage de plusieurs puits simultanément est mis en place.

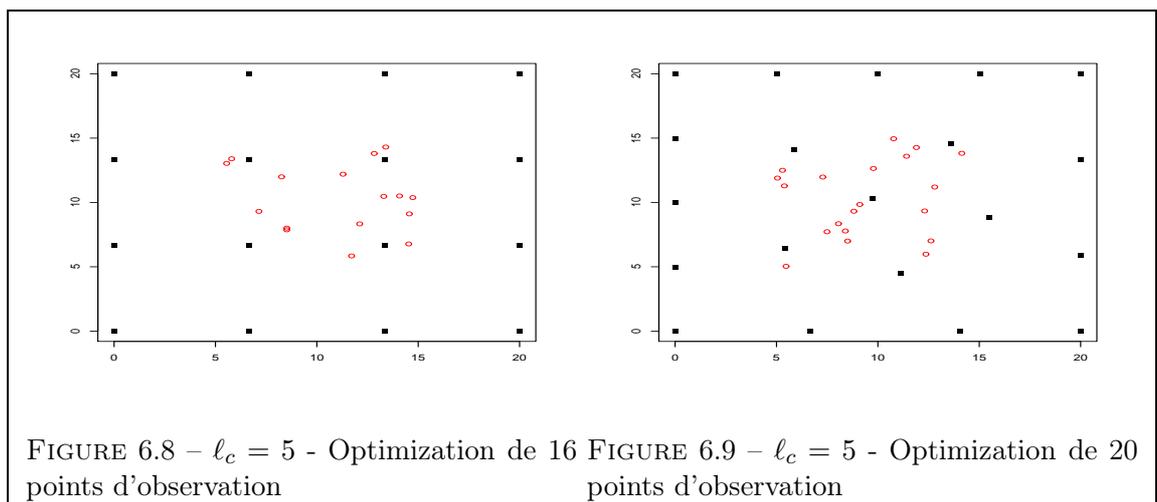
Cependant, lorsqu'un scénario de production est défini, le nombre d'observations N est choisi et il convient de définir le placement de tous les points au préalable. Pour N observations, $d \times N$ paramètres sont à optimiser, $d = 1, 2, 3$ étant la dimension de

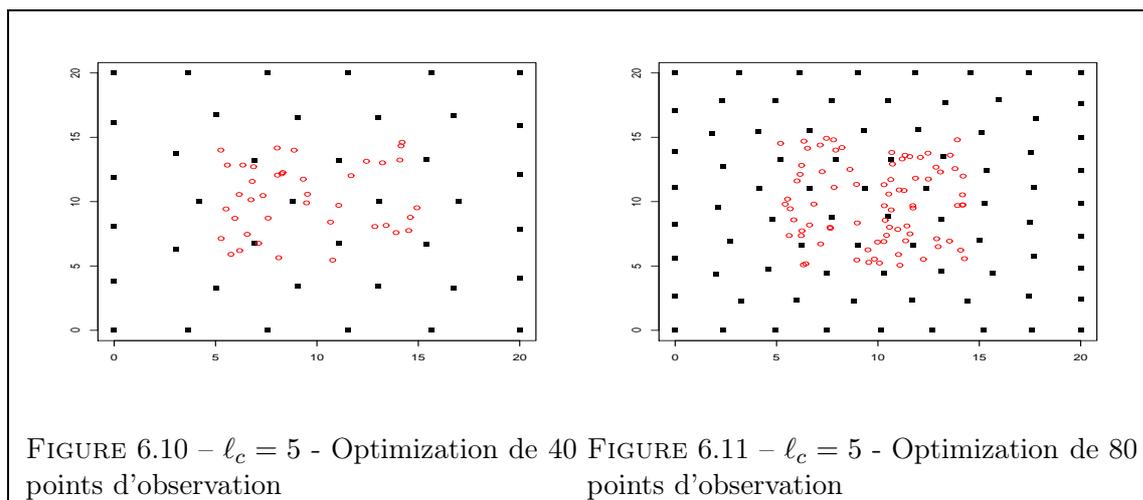


l'espace. Le problème d'optimisation est le suivant : on cherche l'emplacement des N observations tel que :

$$\begin{aligned} \mathbf{x} &= \mathbf{x}_1, \dots, \mathbf{x}_N \quad \text{pris dans } \mathcal{M} \\ \mathbf{x}^* &= \arg \max_{\mathbf{x} \in \mathcal{M}} \langle I(\mathbf{x}) \rangle \end{aligned} \quad (6.20)$$

L'optimisation globale est réalisée à l'aide de l'algorithme L-BFGS-B (*Limited Broyden-Fletcher-Goldfarb-Shanno Bounded*) de type quasi-Newton. Cette méthode permet de résoudre notre problème non-linéaire de grande taille en construisant par la méthode *BFGS* une approximation de l'inverse de la matrice Hessienne de la fonction $\langle I(\mathbf{x}) \rangle$ en analysant les différents gradients successifs. On pourra consulter [37],[38] pour les détails mathématiques de la méthode de descente ainsi que son implémentation. Pour différentes valeurs de N , on représente dans les figures suivantes les résultats.





Il apparaît clairement ici que la maximisation de l'information moyenne revient à une répartition très uniforme des points dans l'espace. Ce résultat est tout à fait en ligne avec l'intuition physique puisque nous considérons ici un champ isotrope.

Nous représentons figure (6.12) l'évolution de l'information moyenne $\langle I \rangle$ en fonction du nombre de mesures avec différentes longueurs de corrélation. On constate que l'augmentation de points permet un meilleur recouvrement. Selon la longueur de corrélation, il faudra plus d'observations pour faire décroître l'information moyenne.

La courbe de l'évolution de l'information moyenne figure (6.12) n'est plus asymptotique au maximum d'information à partir d'un nombre N^* de l'ordre du nombre de volumes de corrélation présents dans le domaine. Ceci permet de conclure sur la redondance de l'information. En effet, ce nombre N^* constitue donc un seuil à partir duquel l'information ajoutée par une nouvelle mesure décroît.

Il est clair qu'il n'est pas nécessaire d'effectuer plus de mesures sachant que ces mesures vont être corrélées. Le nombre N^* varie en fonction de la longueur de corrélation. On comprend intuitivement ce résultat sachant que si la longueur de corrélation est faible alors il faudra nécessairement plus de mesures pour expliciter le milieu. Dans le cas du krigeage, on conjecture que N^* est le nombre effectif de degrés de liberté.

2.3 Concentration des observations

Lorsque les mesures sont proches et que l'on considère que géologiquement le niveau d'hétérogénéité est raisonnable, on peut envisager de faire coïncider un certain nombre de points qui seront corrélés. Le problème est ici présenté en figure (6.13)

Du fait de l'agrégation, la répartition des observations est alors changée, ce qui se traduit directement dans le calcul du déterminant utilisé dans le critère de l'information moyenne. Le calcul est laissé en Annexe E.

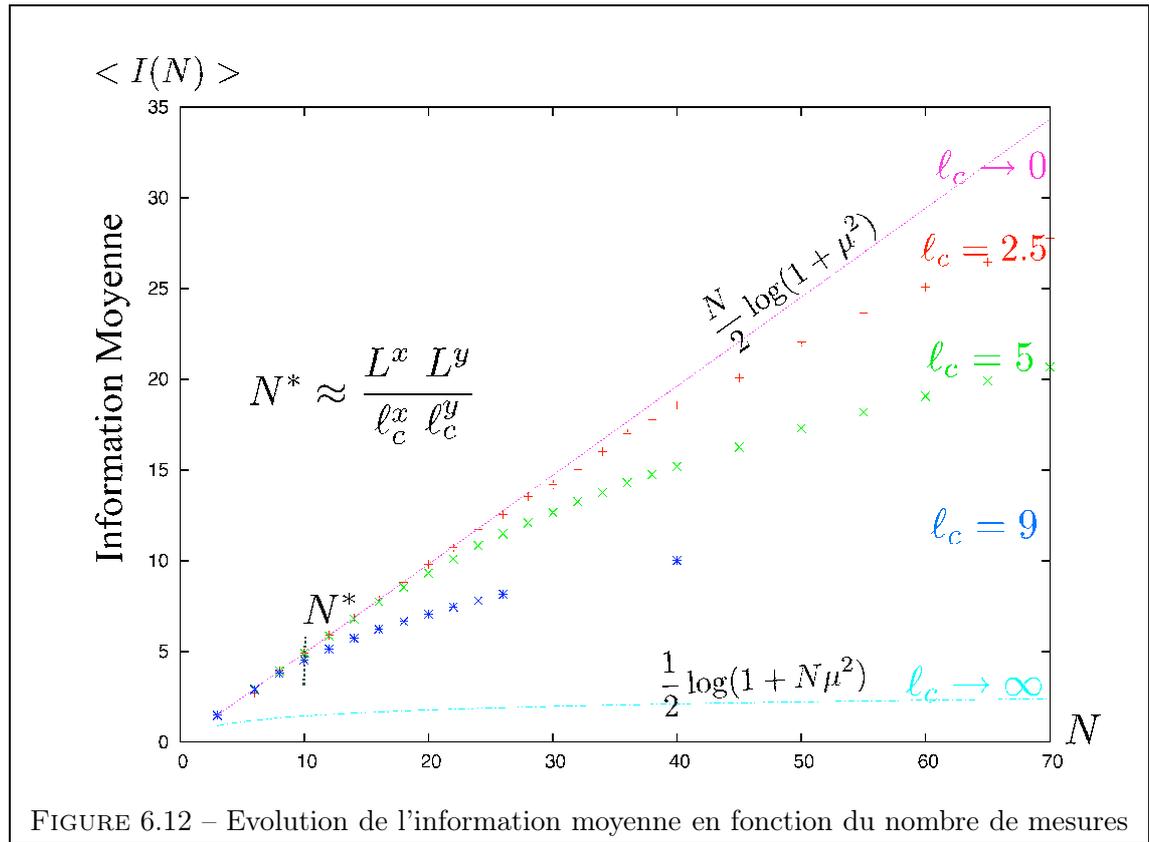


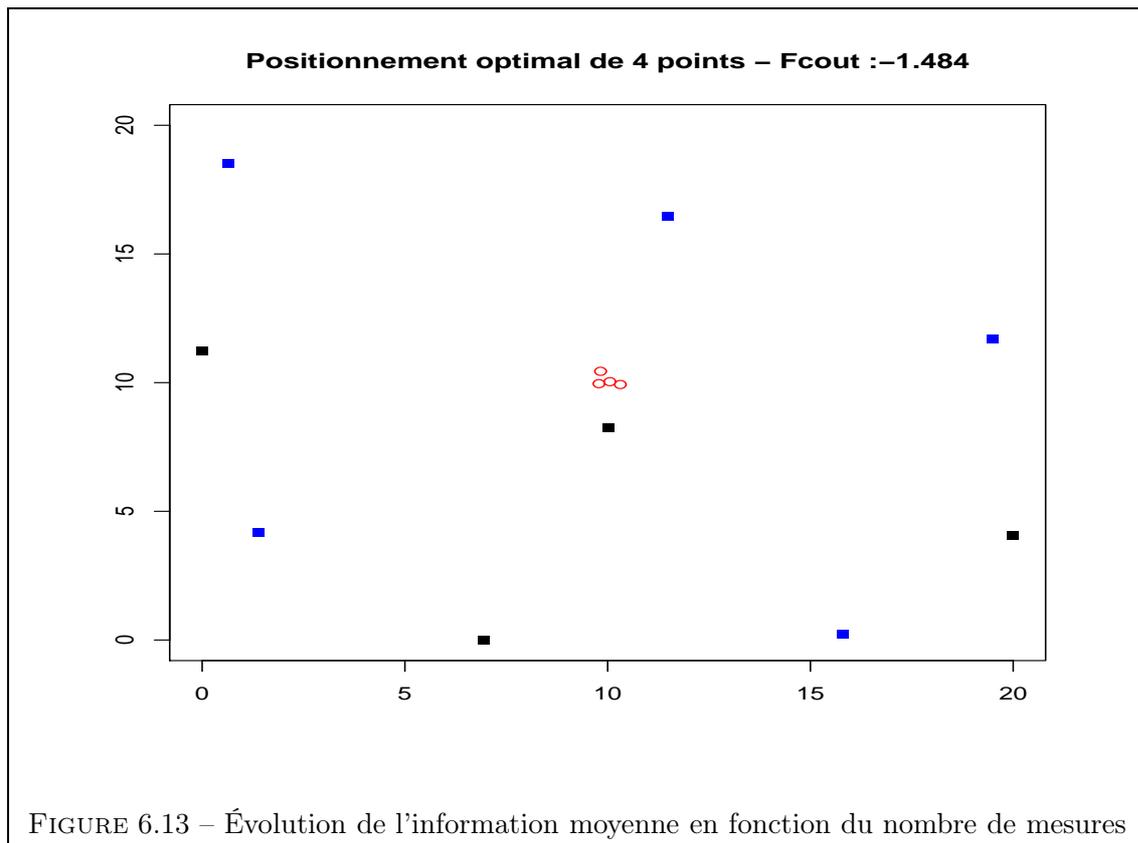
FIGURE 6.12 – Evolution de l’information moyenne en fonction du nombre de mesures

$$\begin{vmatrix}
 1 + \sigma_{pp} & \sigma_{pp} & \dots & \sigma_{pp} & \sigma_{pp+1} & \dots & \sigma_{pN} \\
 \sigma_{pp} & \ddots & & \vdots & \sigma_{pp+1} & \dots & \sigma_{pN} \\
 \vdots & & \ddots & \sigma_{pp} & \vdots & & \vdots \\
 \sigma_{pp} & \dots & \sigma_{pp} & 1 + \sigma_{pp} & \sigma_{pp+1} & \dots & \sigma_{pN} \\
 \sigma_{p+1p} & \dots & \sigma_{p+1p} & \sigma_{p+1p} & 1 + \sigma_{p+1p+1} & \dots & \sigma_{p+1N} \\
 \vdots & & & \vdots & \vdots & \ddots & \vdots \\
 \sigma_{Np} & \dots & \sigma_{Np} & \sigma_{Np} & \sigma_{Np+1} & \dots & \sigma_{NN}
 \end{vmatrix}
 =
 \begin{vmatrix}
 1 + p\sigma_{pp} & p\sigma_{pp+1} & \dots & p\sigma_{pN} \\
 \sigma_{p+1p} & 1 + \sigma_{p+1p+1} & \dots & \sigma_{p+1N} \\
 \vdots & \vdots & \ddots & \vdots \\
 \sigma_{Np} & \sigma_{Np+1} & \dots & \sigma_{NN}
 \end{vmatrix}$$

On remarque, qu’au point d’agrégation, le poids sur la covariance des mesures est alors multiplié par le nombre de mesures en question.

3 Conclusion

Dans ce chapitre nous nous sommes intéressés à une méthode d’estimation de la perméabilité basée sur l’intégration de données statiques : le krigeage. Sa formulation



dans le cadre bayésien, nous a permis de définir un critère d'information moyenne. En optimisant celui-ci, nous avons défini un *design* optimal *a priori* pour l'estimation du champs de perméabilité. Nous avons pu montrer qu'en fonction de la longueur de corrélation *a priori* du milieu, il existe une limite en dessous de laquelle, l'information apportée par une nouvelle mesure est redondante. Il est important de noter que cette limite est proportionnelle aux volumes de corrélation présents dans le domaine.

Ici, nous avons donc présenté une approche optimale pour l'intégration de données statiques. Qu'en est-il de l'intégration de données dynamiques ? Dans le chapitre suivant, nous utilisons les observations de pression aux puits et nous nous posons la question de savoir s'il existe, comme dans le cas spatial, une limite temporelle en dessous de laquelle toute acquisition en pression n'est pas informative.

Chapitre 7

Optimisation des temps d'acquisition de mesures : cas des essais de puits

Sommaire

1	Principe	53
2	Équation d'état	54
3	Solution analytique et linéarisation	54
	3.1 Solution en milieu homogène	54
	3.2 Approximation et solution numérique en milieu hétérogène	55
4	Essai de puits en milieu hétérogène	58
	4.1 Propriétés stochastiques de la perméabilité apparente	59
	4.2 Approche perturbative de l'équation de diffusivité	60
	4.3 Fonction de covariance de la perméabilité apparente	63
5	Optimisation	65
6	Résultats numériques	67
	6.1 Optimisation séquentielle	67
	6.2 Optimisation globale	67
7	Conclusion	70

Dans ce chapitre, nous présentons une application de notre approche de maximisation de l'information pour planifier la fréquence d'acquisition de mesures de pression au cours d'un essai de puits.

1 Principe

Le but d'un essai de puits est d'obtenir des renseignements sur un puits et sur le réservoir. Les essais de puits sont largement utilisés en ingénierie de réservoir pour obtenir des propriétés dynamiques des formations à une échelle plus grande que celle des

analyses de carottes effectuées en laboratoire. Pendant des années, l'interprétation a été limitée aux essais de puits des réservoirs de géométrie simple se comportant comme un réservoir homogène. Plus récemment, les conditions d'interprétation ont été élargies en traitant des réservoirs plus particuliers ayant par exemple des fractures naturelles.

Le principe de l'essai de puits consiste à mesurer dans un puits une valeur de pression. Des variations de débit sont effectuées. Elles ont pour but de provoquer une perturbation du régime des pressions existant. Les mesures de pression sont alors effectuées après perturbation. Ces mesures sont effectuées pendant la production ou lors d'une période d'arrêt du puits. En pratique, seule la période à débit nul est exploitable. L'évolution de la pression est ensuite interprétée pour dégager des informations sur l'écoulement et les propriétés du réservoir, comme classiquement la perméabilité. Dans le cas de l'essai de puits, notre fonction \mathcal{F} est alors le plus souvent représentée par des modèles analytiques simplifiés et des modèles numériques du problème d'écoulement monophasique en milieu poreux.

2 Équation d'état

L'équation de diffusivité au puits déduite de la loi de Darcy modélise l'écoulement monophasique d'un fluide faiblement compressible, en l'occurrence l'huile. Elle modélise l'évolution de la pression au cours du temps sous l'action d'un terme source Q . La pression P est la seule inconnue du problème. Le système est donné par :

$$\begin{cases} \phi\mu c_T \partial_t P(\mathbf{r}, t) = \nabla \cdot (k(\mathbf{r}) \nabla P(\mathbf{r}, t)) + Q \frac{\mu}{h} \delta(\mathbf{r}) H(t) \\ \forall \mathbf{r}, P(\mathbf{r}, t = 0) = 0 \\ \forall t, P(r \rightarrow +\infty, t) = 0 \end{cases}$$

L'équation d'état intègre des paramètres géologiques tels que ϕ la porosité du milieu supposée constante en espace et en temps, μ la viscosité du fluide, c_T la compressibilité totale du système roche-fluide et k le tenseur de perméabilité. Le puits de débit constant Q est considéré par un point et est caractérisé à l'aide de la fonction de Dirac $\delta(\mathbf{r})$ et de la fonction de Heavyside H . La condition aux limites du problème impose au cours du temps une condition de flux nul, c'est-à-dire un domaine supposé imperméable à l'écoulement. On retrouve cette condition dans la fonction de Dirac δ . Pour les grandes distances *i.e* $\mathbf{r} \rightarrow +\infty$, nous supposons que la variation de pression P est nulle.

3 Solution analytique et linéarisation

Nous rappelons ici la solution homogène de l'équation de diffusivité à l'aide de la transformée de Fourier-Laplace.

3.1 Solution en milieu homogène

Le problème homogène considère la diffusivité comme constante. On pose alors dans ce cas $D(\mathbf{r}) = D_0$. La solution du problème homogène est obtenue en calculant la

fonction de Green associée au problème. L'équation de la dérivée de la pression P_0 dans le cas homogène est gouvernée par l'équation différentielle suivante :

$$\partial_t P'_0(\mathbf{r}, t) = D_0 \nabla^2 P'_0(\mathbf{r}, t) + \delta(\mathbf{r})\delta(t) \quad (7.1)$$

On remarque que la dérivée de la pression est la fonction de Green du problème en milieu homogène, *i.e.* $G_0 = P'_0$. La transformée de Fourier-Laplace appliquée à l'équation précédente donne la solution analytique dans l'espace des fréquences :

$$P'_0(\mathbf{q}, s) = \frac{1}{s + D_0 q^2} \quad (7.2)$$

Il reste alors à appliquer la transformée de Fourier-Laplace inverse et à intégrer par rapport au temps. On obtient la solution analytique classique suivante de la pression en tout point du réservoir différent du puits :

$$\forall \mathbf{r} \neq 0, \quad P_0(\mathbf{r}, t) = \frac{1}{4\pi D_0} \int_{\frac{r^2}{4D_0 t}}^{+\infty} dv \frac{e^{-v}}{v} \quad (7.3)$$

C'est de cette relation que découle la formule de base de l'interprétation classique des essais de puits en milieu homogène :

$$P(\mathbf{r} \rightarrow 0, t) = \frac{1}{4\pi D_0} \left(-\gamma - \ln \left(\frac{r^2}{4D_0} \right) + \ln(t) \right) \quad (7.4)$$

où γ est la constante d'Euler. En dérivant cette relation par rapport au logarithme du temps, on obtient le hk du réservoir en milieu homogène tel que :

$$hk_0 = \frac{\mu Q}{4\pi \frac{dP(\mathbf{r}=0, t)}{d \ln(t)}} \quad (7.5)$$

D'autres études faites par Daviau [39], Bourdarot [40] relatives à l'évolution de la pression en milieu homogène permettent d'estimer certains paramètres relatifs à l'écoulement comme l'épaisseur du skin, les effets de capacité, les limites du réservoir, etc.

3.2 Approximation et solution numérique en milieu hétérogène

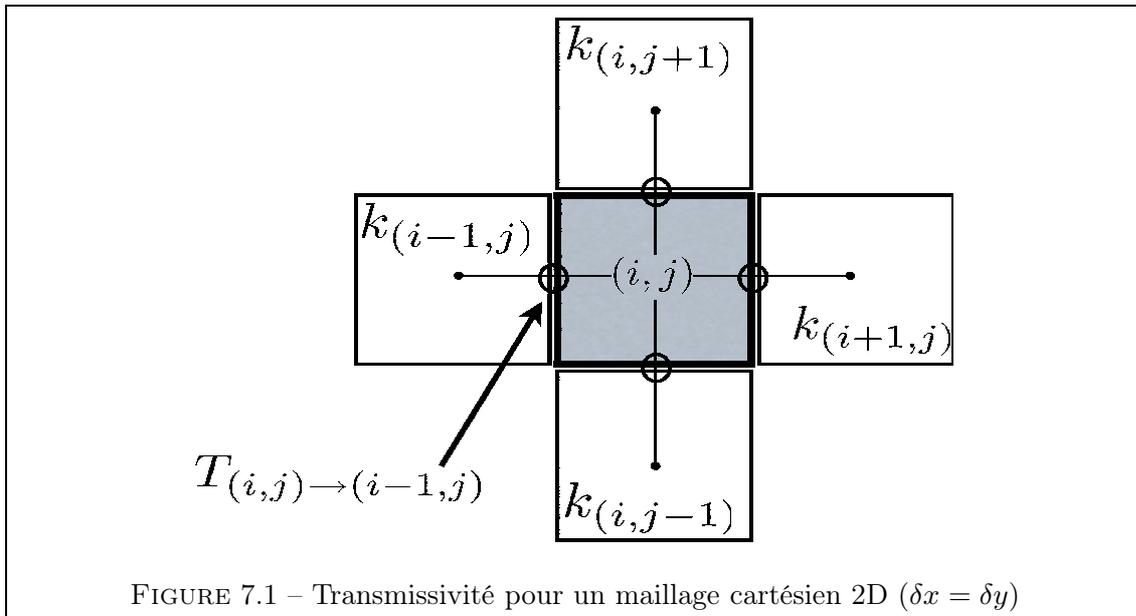
En milieu hétérogène, l'équation de diffusivité n'admet pas de solution analytique simple et implémentable facilement. Une approximation est nécessaire. Le réservoir est maillé de façon cartésienne et on suppose que les propriétés pétrophysiques sont constantes sur chaque maille. La discrétisation est faite par volumes finis en espace. La description mathématique utilisée est celle définie dans le simulateur d'écoulement monophasique 3D développé à l'IFP. On considère un réservoir comportant un puits situé au milieu de la maille m_w . La formule suivante synthétise les considérations précédentes. Pour une maille m , on obtient :

$$\phi V_m \mu c_T \frac{P_m^{n+1} - P_m^n}{t^{n+1} - t^n} = \sum_{v \in \mathcal{V}(m)} T_{m \rightarrow v} \cdot (P_v^{n+1} - P_m^{n+1}) + Q \frac{\mu}{h} \delta_{m, m_w} H(t_{n+1}) \quad (7.6)$$

où $V_m\phi$ est le volume poreux de la maille m et δ_{m,m_w} est le symbole de Kronecker qui vaut 1 si $m = m_w$ et 0 sinon. La résolution de l'équation de diffusivité aboutit à la résolution d'un système linéaire dont l'inconnue est la pression sur chaque maille du réservoir. On a pris comme approximation de la pression $P_m^n = P(\mathbf{r}_m, t_n) = P((i\delta x, j\delta y, k\delta z), n\delta t)$ où $\delta x, \delta y, \delta z$ et δt sont respectivement les pas de discrétisation en espace et en temps. Par simplicité, on prend $\delta x = \delta y = \delta z$. Le scalaire $T_{m \rightarrow v}$ représente la *transmissivité* entre la maille m et la maille $v \in \mathcal{V}(m)$. C'est un facteur de proportionnalité entre le flux sortant de la maille m et la différence de pression entre la pression calculée en m et la pression calculée en v . L'ensemble $\mathcal{V}(m)$ contient toutes les mailles voisines ayant une face commune. Dans le cas 1D, \mathcal{V} contient 2 éléments; dans le cas 2D, 4 éléments et dans le cas 3D, 6 éléments. Dans la mesure où la discrétisation est cartésienne, nous avons utilisé une pondération harmonique pour calculer la transmissivité. Soit encore, en posant $k(\mathbf{r}) = k_{i,j,k}$, on obtient particulièrement pour la maille voisine $(i+1, j, k)$:

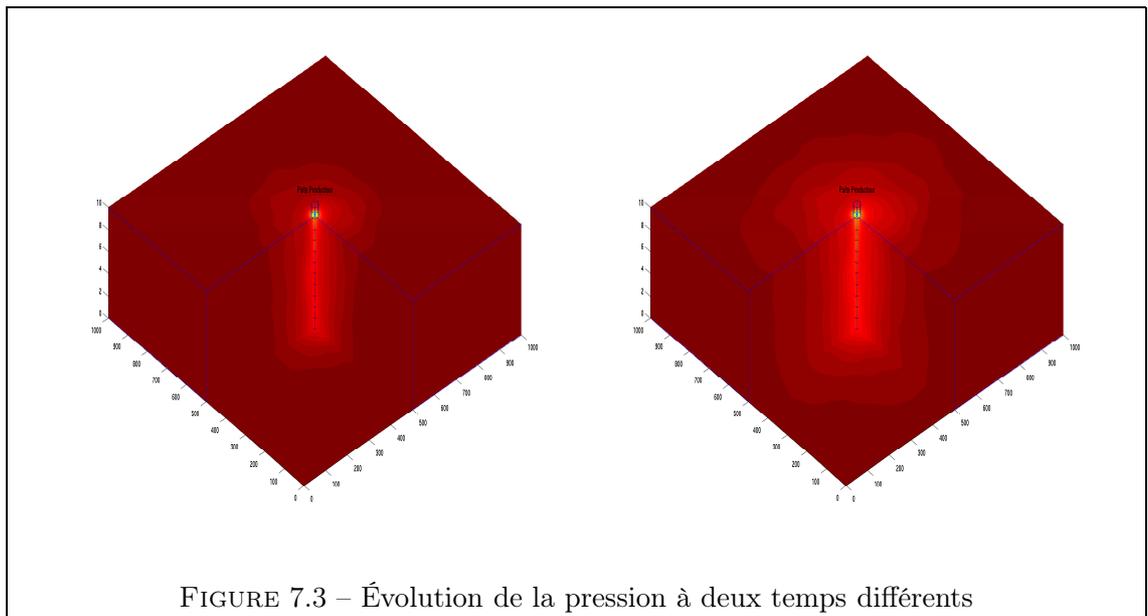
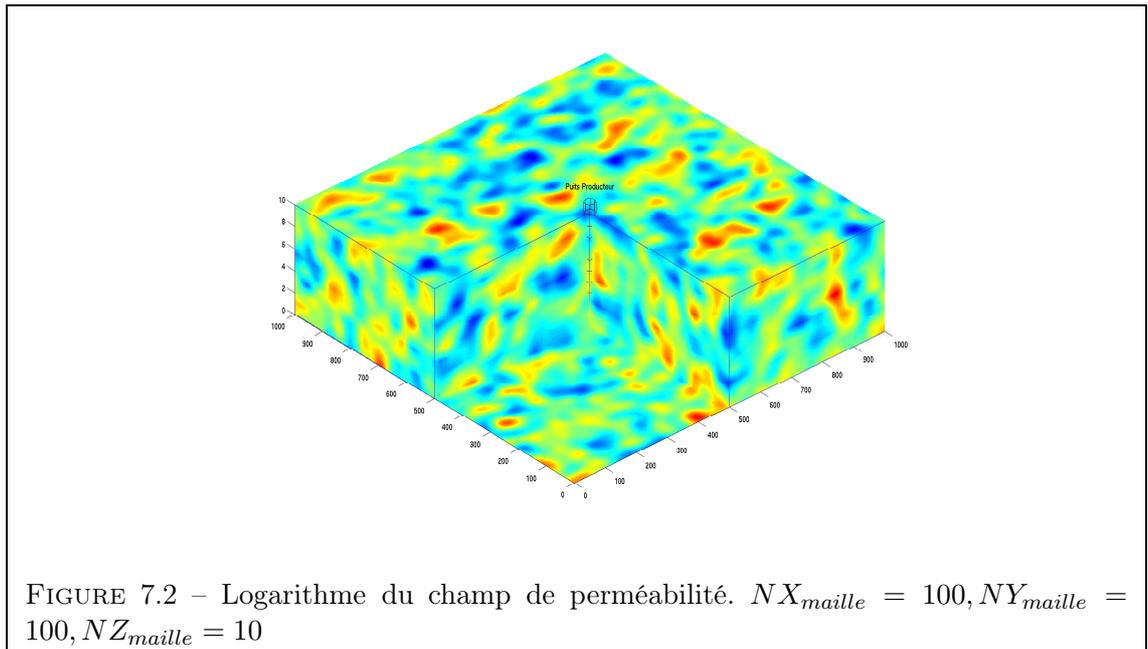
$$T_{m \rightarrow (i+1,j,k)} = 2(k_{i,j,k}^{-1} + k_{i+1,j,k}^{-1})^{-1} \quad (7.7)$$

Le principe du calcul de la transmissivité est illustré figure (7.1).

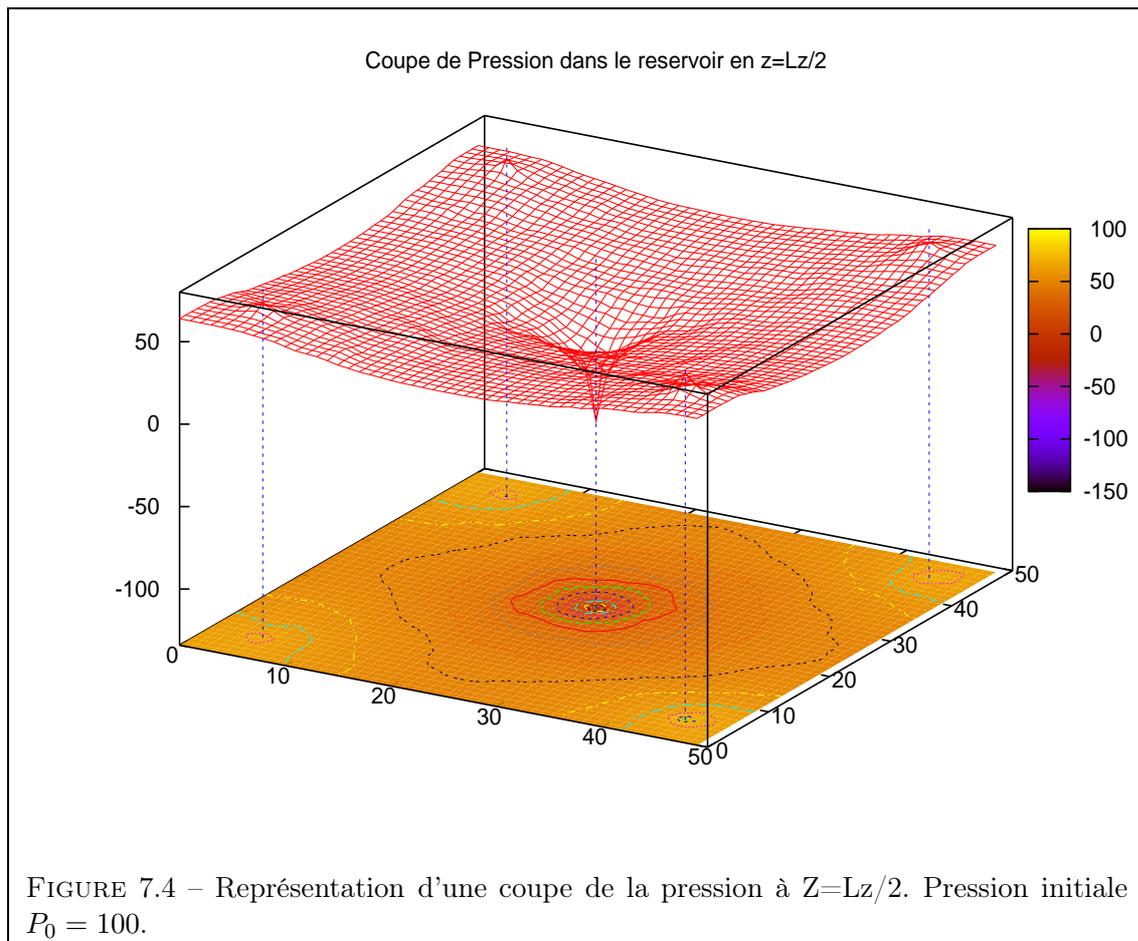


On peut alors simuler l'évolution de la pression dans le réservoir. On représente figure (7.2) le champ de perméabilité utilisé pour la simulation de l'évolution de la pression et en figure (7.3) on représente la pression à deux temps différents. Le puits producteur est situé au centre du réservoir. On choisit des temps relativement courts pour être dans les conditions de validité du modèle et ainsi considérer le réservoir comme infini.

On illustre l'évolution de la pression dans le cas d'un *five-spot*. Le puits producteur est situé au centre du réservoir, et les 4 puits injecteurs aux coins. On représente une



coupe de la pression au niveau $Z = Lz/2$ ainsi que ses isovaleurs en figure (7.4).



4 Essai de puits en milieu hétérogène

Dans le cas où la diffusivité n'est plus considérée comme constante, on redonne la définition de la notion de perméabilité apparente du milieu à partir de la pression au puits par la formule suivante :

$$k_{app}(t) = \frac{\mu Q}{4\pi h \frac{dP(t)}{d\ln(t)}} \quad (7.8)$$

Cette relation généralise la notion de perméabilité apparente en milieu homogène. La perméabilité apparente varie au cours du temps lorsque que le milieu est hétérogène. On définit le rayon d'investigation comme étant la zone influant sur la pression au puits. On introduit la définition du rayon d'investigation en utilisant la perméabilité apparente par la formule suivante :

$$R_{app}^2(t) = \frac{t k_{app}(t)}{\phi \mu c_T} \quad (7.9)$$

Pour des réservoirs de type radial, stratifié ou encore de type fracturé à double porosité, il existe des solutions analytiques pour l'évolution de la pression. On pourra consulter Streltsova [41] pour plus de détails. Plusieurs exemples montrent la dépendance de la pression au puits et la perméabilité du milieu. Dans le cas d'un réservoir fissuré, on pourra se référer à Daviau [39] et Bourdarot [40]. La perméabilité varie fortement en espace et ces considérations renforcent sa description stochastique. Dans ce contexte probabiliste, l'équation de diffusivité n'a pas de solution analytique simple exploitable et est résolue par des méthodes numériques types volumes finis ou différences finies.

4.1 Propriétés stochastiques de la perméabilité apparente

En utilisant un développement en perturbation à tous les ordres de l'équation de la diffusivité, on établit une relation avec la perméabilité du milieu. L'utilisation de la transformée de Fourier-Laplace, voir Indelman [42] ou encore, en utilisant les diagrammes de Feynman, voir Gautier [43], aboutit à une équation intégral-différentielle gouvernant l'espérance mathématique de la pression. De cette équation, on établit les propriétés asymptotiques aux temps courts et aux temps longs de l'espérance mathématique de la perméabilité apparente. Il apparaît qu'aux temps courts, l'espérance de la perméabilité apparente tend vers la moyenne harmonique de la perméabilité proche du puits. En notant k_w cette perméabilité, on obtient :

$$\lim_{t \rightarrow 0} \langle k_{app}(t) \rangle = \frac{4\pi h}{\mu Q} \left\langle \frac{1}{\frac{dP(t)}{d \ln(t)}} \right\rangle = \frac{1}{\left\langle \frac{1}{k_w} \right\rangle} \quad (7.10)$$

Remarque: *Si on considère que l'essai de puits est principalement sensible à la perméabilité au niveau du puits, alors la perméabilité apparente tend vers la perméabilité d'un milieu homogène. La moyenne de la perméabilité apparente ne tend plus comme précédemment vers la moyenne géométrique mais vers la moyenne arithmétique des perméabilités. Ce résultat diffère suivant que l'on choisit comme variable la perméabilité apparente ou la dérivée logarithmique de la pression au puits pour le calcul de la moyenne. On pourra consulter Gautier [43] pour plus de détails.*

Noetinger *et al.* [44] mettent en évidence la propriété d'auto-moyennage de la perméabilité apparente et montre qu'aux temps longs, celle-ci moyenne un grand nombre d'événements indépendants considérant ainsi l'essai de puits comme un processus d'homogénéisation. En effet, ils considèrent la perméabilité apparente comme une moyenne pondérée des perméabilités à petite-échelle dans le disque de rayon $R(t) \cong (2D_0 t)^{1/2}$ autour du puits. Après une période transitoire, le nombre $N(t) \approx \pi R^2(t)/\ell_c^2$ représente le nombre d'événements indépendants au temps t . Le théorème central-limite donne alors directement la variance de la perméabilité apparente qui est égale à $1/N(t) \approx 1/t$.

Mathématiquement, on montre que la perméabilité apparente converge, presque sûrement vers une valeur de perméabilité constante et intrinsèque au milieu que l'on appelle la perméabilité effective du milieu. On pourra consulter Goldenfeld [45], et Jikov [46] pour plus de détails sur l'homogénéisation des opérateurs différentiels. Soit encore, en

notant k_{eff} la perméabilité effective du milieu :

$$\lim_{t \rightarrow \infty} \langle k_{app} \rangle = k_{eff} \quad (7.11)$$

Gautier et Noetinger [47] montrent que la perméabilité apparente est en moyenne proportionnelle à $1/t$:

$$\left\langle k_{app}(t) - k_{eff} \right\rangle \propto \frac{1}{t} \quad (7.12)$$

Dans un milieu infini 1D, il existe une relation analytique entre les paramètres du variogramme de la perméabilité et les deux premiers moments de la perméabilité apparente. Romeu [48] et Yortos [49] obtiennent ce résultat par un développement en perturbation à l'ordre 2 par rapport à la perméabilité de l'équation de diffusivité.

4.2 Approche perturbative de l'équation de diffusivité

Dans le paragraphe suivant nous exposons le calcul d'une approximation de la perméabilité apparente, qui nous sera ensuite utile pour définir la fonction de covariance de la perméabilité apparente entre deux temps d'acquisition de pression.

On considère la fonction de Green $G_0(\mathbf{r}, \boldsymbol{\zeta}; t, \tau)$ du problème en milieu homogène solution de l'équation différentielle suivante :

$$\partial_t G_0 = D_0 \nabla^2 G_0 + \delta(t - \tau) \delta(\mathbf{r} - \boldsymbol{\zeta}) \quad (7.13)$$

et la dérivée de la pression en milieu homogène obéissant à l'équation différentielle suivante :

$$\partial_t P'_0(t, \mathbf{r}) = D_0 \nabla^2 P'_0 + \delta(t) \delta(\mathbf{r}) \quad (7.14)$$

où l'on définit $P'_0 = \partial_t P_0$.

On suppose qu'il est possible de développer en puissance du logarithme de la variance de la perméabilité, les termes de diffusivité et de pression :

$$D(\mathbf{r}) = D_0 + \sigma_{\ell n} D_1(\mathbf{r}) + \sigma_{\ell n}^2 D_2(\mathbf{r}) + \dots \quad (7.15)$$

$$P(\mathbf{r}, t) = P_0(\mathbf{r}, t) + \sigma_{\ell n} P_1(\mathbf{r}, t) + \sigma_{\ell n}^2 P_2(\mathbf{r}, t) + \dots \quad (7.16)$$

Reprenons l'équation de la dérivée de la pression en milieu hétérogène décrite par :

$$\partial_t P' = \nabla \cdot (D(\mathbf{r}) \nabla P') + \delta(\mathbf{r}) \delta(t) \quad (7.17)$$

En utilisant le développement en perturbation de la pression, on généralise cette dernière équation à l'ordre n telle que :

$$\partial_t P'_n = \nabla \cdot \left(\sum_{j=0}^n D_j(\mathbf{r}) \nabla P'_{n-j} \right) \quad (7.18)$$

On applique la transformée de Laplace à cette dernière équation et on la multiplie par la fonction de Green en milieu homogène. En intégrant sur le domaine, on obtient la nouvelle équation :

$$\int_{\mathbb{R}^2} sP'_n G_0 - D_0 \nabla^2 (P'_n) G_0 d\zeta = \int_{\mathbb{R}^2} \nabla \cdot \left(\sum_{j=1}^n D_j(\zeta) \nabla P'_{n-j} \right) \times G_0 d\zeta \quad (7.19)$$

En appliquant le théorème de Green, on obtient du fait des conditions de bord nulles :

$$\int_{\mathbb{R}^2} (sG_0 - D_0 \nabla^2 G_0) \times P'_n d\zeta = - \int_{\mathbb{R}^2} \left(\sum_{j=1}^n D_j(\zeta) \nabla P'_{n-j} \right) \cdot \nabla G_0 d\zeta \quad (7.20)$$

Soit encore, d'après l'équation (7.13) appliquée à la transformée de Laplace :

$$P'_n(\mathbf{r}, s) = - \int_{\mathbb{R}^2} \left(\sum_{j=1}^n D_j(\zeta) \nabla P'_{n-j} \right) \cdot \nabla G_0 d\zeta \quad (7.21)$$

On réécrit ce dernier résultat en explicitant le terme de diffusivité ainsi que la fonction de Green du problème en milieu homogène. Soit, au puits $\mathbf{r} = 0$:

$$P'_n(\mathbf{r} = 0, s) = - \int_{\mathbb{R}^2} \left(\sum_{j=1}^n \frac{k_j(\zeta)}{\phi \mu c_T} \nabla P'_{n-j}(\zeta, s) \right) \cdot \nabla P'_0(\zeta, s) d\zeta \quad (7.22)$$

On reprend l'expression de la dérivée de la pression en milieu homogène et on applique la transformée de Fourier inverse. On obtient :

$$P'_0(\zeta, s) = \int_{\mathbb{R}^2} d\mathbf{q} e^{-i\zeta \cdot \mathbf{q}} P'(\mathbf{q}, s) \quad (7.23)$$

$$= \int_{\mathbb{R}^2} d\mathbf{q} \frac{e^{-i\zeta \cdot \mathbf{q}}}{s + D_0 q^2} \quad (7.24)$$

En utilisant la représentation de Schwinger, on obtient :

$$P'_0(\zeta, s) = \int_0^{+\infty} dt \int_{\mathbb{R}^2} d\mathbf{q} e^{-i\zeta \cdot \mathbf{q}} e^{-t(s + D_0 q^2)} \quad (7.25)$$

$$= \int_0^{+\infty} dt e^{-st} \int_{\mathbb{R}^2} d\mathbf{q} e^{-i(\mathbf{q}\sqrt{D_0 t}) \cdot \zeta / 2\sqrt{D_0 t} - (q\sqrt{D_0 t})^2} \quad (7.26)$$

On reconnaît au facteur près une intégrale de Poisson en dimension 2. Et donc :

$$P'_0(\zeta, s) = \int_0^{+\infty} dt e^{-st} \frac{e^{-\zeta^2 / 4D_0 t}}{4\pi D_0 t} \quad (7.27)$$

On remarque par identification que cette dernière relation est la transformée de Laplace de $P'_0(\zeta, t)$. Elle a une expression analytique (on pourra consulter Abramowitz [50]) telle que :

$$P'_0(\zeta, s) = \frac{1}{2\pi D_0} \mathcal{K}_0 \left(\zeta \sqrt{\frac{s}{D_0}} \right) \quad (7.28)$$

De cette expression, on calcule l'expression de son gradient par rapport à ζ :

$$\nabla P'_0(\zeta, s) = \frac{-1}{2\pi D_0} \sqrt{\frac{s}{D_0}} \mathcal{K}_1 \left(\zeta \sqrt{\frac{s}{D_0}} \right) \quad (7.29)$$

On utilise ce dernier résultat pour expliciter la dérivée de la pression à l'ordre 1 que l'on injecte dans l'expression (7.22).

Pour $n = 1$, l'expression de la dérivée de la pression au puits est alors :

$$P'_1(\mathbf{r} = 0, s) = - \int_{\mathbb{R}^2} d\zeta \frac{k_1(\zeta)}{\phi \mu c_T} \left(\frac{1}{2\pi D_0} \right)^2 \frac{s}{D_0} \mathcal{K}_1^2 \left(\zeta \sqrt{\frac{s}{D_0}} \right) \quad (7.30)$$

En remarquant de plus que $P'(\mathbf{r}, s) = sP(\mathbf{r}, s)$, la relation précédente devient :

$$P_1(\mathbf{r} = 0, s) = - \int_{\mathbb{R}^2} d\zeta \frac{k_1(\mathbf{r})}{\phi \mu c_T} \left(\frac{1}{2\pi D_0} \right)^2 \frac{1}{D_0} \mathcal{K}_1^2 \left(\zeta \sqrt{\frac{s}{D_0}} \right) \quad (7.31)$$

Finalement, l'expression analytique $P_1(\mathbf{r} = 0, t)$ est donnée en calculant la transformée de Laplace inverse de $\mathcal{K}_1^2 \left(\zeta \sqrt{\frac{s}{D_0}} \right)$. On pourra consulter l'annexe de Oliver, [51] pour le détail du calcul. On obtient alors, à l'ordre 1 l'expression de la dérivée de la pression au puits :

$$P_1(\mathbf{r} = 0, t) = \int_{\mathbb{R}^2} d\zeta \frac{k_1(\zeta)}{\phi \mu c_T} \left(\frac{1}{2\pi D_0} \right)^2 \frac{-1}{2D_0 t} e^{\frac{-\zeta^2}{2D_0 t}} \mathcal{K}_1 \left(\frac{\zeta^2}{2D_0 t} \right) \quad (7.32)$$

On dérive maintenant par rapport au temps t l'expression précédente et on obtient ainsi :

$$P'_1(\mathbf{r} = 0, t) = \int_{\mathbb{R}^2} d\zeta \frac{k_1(\zeta)}{\phi \mu c_T} \left(\frac{1}{2\pi D_0} \right)^2 \frac{-1}{4D_0 t^2} W(\zeta, t) \quad (7.33)$$

où W est le noyau pondérateur des perméabilités de Oliver [51] défini par :

$$W(\zeta, t) = W(z) = z e^{-z} (2\mathcal{K}_1(z) (1 - \frac{1}{z}) + \mathcal{K}_0(z) + \mathcal{K}_2(z)) \quad (7.34)$$

où $z = \frac{\zeta^2}{2D_0 t}$. On notera l'analogie entre z et la variable réduite de Boltzman.

On reprend la définition de la perméabilité apparente en explicitant la pression à l'aide de son développement et on donne son expression au premier ordre. Soit :

$$k_{app}(t) = \frac{\mu Q}{4\pi h t (P'_0 + \sigma_{ln} P'_1 + \dots)} = \frac{1}{\frac{4\pi h t P'_0}{\mu Q} + \frac{4\pi h t}{\mu Q} \sigma_{ln} P'_1 + \dots} \quad (7.35)$$

On reconnaît l'expression de k_0 . L'expression de k_{app} devient au premier ordre :

$$k_{app}(t) = k_0 - k_0^2 \frac{4\pi h t}{\mu Q} \sigma_{ln} P'_1(0, t) \quad (7.36)$$

On prend la définition de k_0 telle que $k_0 = D_0 \cdot \phi \mu c_T$ et le fait que $\frac{Q}{\phi c_T h} = 1$. Après simplification dans l'expression de $P'_1(0, t)$, on obtient :

$$k_{app}(t) = k_0 + \frac{\sigma_{ln}}{2\pi D_0 t} \int_{\mathbb{R}^2} d\mathbf{r} k_1(\mathbf{r}) G(\mathbf{r}, t) \quad (7.37)$$

où $G(\mathbf{r}, t) = \frac{1}{2} W(\mathbf{r}, t)$.

Soit encore, en utilisant la propriété telle que :

$$\int_{\mathbb{R}^2} d\mathbf{r} G(\mathbf{r}, t) = 1 \quad (7.38)$$

on obtient l'approximation au premier ordre de la perméabilité apparente suivante :

$$k_{app}(t) = \int_{\mathbb{R}^2} d\mathbf{r} k(\mathbf{r}) G(\mathbf{r}, t) \quad (7.39)$$

avec $k(\mathbf{r}) = k_0 + \frac{\sigma_{ln}}{2\pi D_0 t} k_1(\mathbf{r})$.

Cette approximation va nous permettre de définir la fonction de covariance de la perméabilité apparente entre deux temps d'acquisition de pression.

4.3 Fonction de covariance de la perméabilité apparente

Disposant d'une approximation de la perméabilité apparente, nous pouvons définir au premier ordre la fonction de covariance de la perméabilité apparente entre deux temps d'acquisition.

On reprend l'expression de la perméabilité apparente en utilisant le fait qu'au premier ordre :

$$\sigma_{ln} k_1(\mathbf{r}) = k(\mathbf{r}) - k_0 \quad (7.40)$$

et on obtient :

$$k_{app}(t) - k_0 = \frac{1}{2\pi D_0 t} \int_{\mathbb{R}^2} d\mathbf{r} (k(\mathbf{r}) - k_0) G(\mathbf{r}, t) \quad (7.41)$$

On utilise cette dernière relation pour définir la fonction de covariance de la perméabilité apparente au premier ordre :

$$\langle k_{app}(t) - k_0, k_{app}(\tau) - k_0 \rangle = \left\langle \int_{\mathbb{R}^2} d\mathbf{r} (k(\mathbf{r}) - k_0) \frac{G(\mathbf{r}, t)}{2\pi D_0 t}, \int_{\mathbb{R}^2} d\boldsymbol{\xi} (k(\boldsymbol{\xi}) - k_0) \frac{G(\boldsymbol{\xi}, \tau)}{2\pi D_0 \tau} \right\rangle \quad (7.42)$$

En considérant que k_0 et que le noyau de pondération G sont constants par l'opérateur de la moyenne spatiale, nous obtenons l'expression suivante :

$$\langle k_{app}(t), k_{app}(\tau) \rangle = \int_{\mathbb{R}^2} d\mathbf{r} \frac{G(\mathbf{r}, t)}{2\pi D_0 t} \int_{\mathbb{R}^2} d\boldsymbol{\xi} \frac{G(\boldsymbol{\xi}, \tau)}{2\pi D_0 \tau} \langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle \quad (7.43)$$

L'expression de la fonction de covariance de la perméabilité apparente au premier ordre est reliée au calcul de la fonction de covariance des perméabilités $\langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle$. En pratique, la perméabilité est considérée comme une fonction constante par morceau et

$\langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle$ est alors la fonction de covariance des perméabilités entre deux couronnes respectivement de rayon r et ξ .

Remarque: Nous utilisons la moyenne géométrique pour représenter la perméabilité sur une couronne afin d'être cohérent avec le comportement asymptotique de la perméabilité apparente. En effet, l'expression (7.41) de la perméabilité apparente caractérise une moyenne harmonique pondérée sur des couronnes dont le rayon augmente et dont la valeur des perméabilités fluctue de moins en moins autour de la moyenne considérée.

On choisit de représenter le champ de perméabilité sur l'ensemble du domaine par une fonction aléatoire Gaussienne Y centrée (*i.e.* $\langle Y(\mathbf{u}) \rangle = 0$) de \mathbf{u} , caractérisée par sa fonction de covariance $C(\mathbf{u}, \mathbf{u}') = \langle Y(\mathbf{u}), Y(\mathbf{u}') \rangle$ telle que :

$$k(\mathbf{r}) = k_G \cdot \exp Y(\mathbf{r}) \quad (7.44)$$

où k_G est la moyenne géométrique des perméabilités telle que $k_G = \exp \langle \log k(\mathbf{r}) \rangle$. On se place dans un cadre cartésien et on définit la perméabilité sur une couronne $C_{\mathbf{r}}$ à l'aide de sa moyenne géométrique par :

$$k(\mathbf{r}) = \exp \left\{ \frac{1}{|C_{\mathbf{r}}|} \int_{C_{\mathbf{r}}} \log k_{\mathbf{s}} \, ds \right\} = k_G \cdot \exp \left\{ \frac{1}{|C_{\mathbf{r}}|} \int_{C_{\mathbf{r}}} Y(\mathbf{s}) \, ds \right\} \quad (7.45)$$

On peut voir la quantité $k(r)$ comme étant une moyenne spatiale orthoradiale de la log-perméabilité et ainsi :

$$k(r) = \exp \left\{ \frac{1}{2\pi} \int_0^{2\pi} d\theta \log(k(r \cdot \cos(\theta), r \cdot \sin(\theta))) \right\} \quad (7.46)$$

$$= k_G \cdot \exp \left\{ \frac{1}{2\pi} \int_0^{2\pi} d\theta Y_r(\theta) \right\} \quad (7.47)$$

où $Y_r(\theta) = Y(r \cdot \cos(\theta), r \cdot \sin(\theta))$. Nous choisissons de nouveau une représentation de la perméabilité à l'aide d'un variogramme gaussien de longueur de corrélation ℓ_c et de variance σ_0^2 connues. L'expression de la fonction de covariance de la perméabilité entre deux couronnes est alors explicitée par :

$$\langle k(r), k(\xi) \rangle = k_G^2 \cdot \exp \left(\frac{1}{2} \frac{\sigma_0^2}{\exp \frac{r^2}{\ell_c^2}} \mathcal{I}_0 \left(\frac{r^2}{\ell_c^2} \right) + \frac{1}{2} \frac{\sigma_0^2}{\exp \frac{\xi^2}{\ell_c^2}} \mathcal{I}_0 \left(\frac{\xi^2}{\ell_c^2} \right) \right) \times \left(\exp \left(\frac{\sigma_0^2}{\exp \frac{r^2 + \xi^2}{2\ell_c^2}} \mathcal{I}_0 \left(\frac{r\xi}{\ell_c^2} \right) \right) - 1 \right)$$

où \mathcal{I}_0 est la fonction de Bessel modifiée d'ordre 0. On donne les détails du calcul en Annexe G.

On retrouve la propriété d'une fonction de covariance telle que :

$$\langle k(r), k(\xi) \rangle = \langle k(\xi), k(r) \rangle \quad (7.48)$$

Blanc *et al.* [52] montrent que la perméabilité apparente est une moyenne des perméabilités locales pondérée par un noyau dont la taille du support est le rayon d'investigation de l'essai de puits $R(t) \cong (2D_0t)^{1/2}$. Pour le calcul de la fonction de covariance, on utilise l'approximation suivante du noyau d'Oliver :

$$\frac{G(\mathbf{r}, t)}{2\pi D_0 t} \sim \delta(\mathbf{r} - \sqrt{2D_0 t} \cdot \mathbf{e}_r) \quad (7.49)$$

où \mathbf{e}_r est un vecteur unitaire tel que $\mathbf{e}_r^2 = 1$.

On représente ainsi la perméabilité apparente par :

$$k_{app}(t) \approx \delta(\mathbf{r} - \sqrt{2D_0 t} \cdot \mathbf{e}_r) \cdot k(\mathbf{r}) \quad (7.50)$$

On reprend l'expression de la covariance (7.43) et l'expression (7.48), et on obtient l'approximation de la fonction de covariance de la perméabilité apparente :

$$\langle k_{app}(t), k_{app}(\tau) \rangle \approx k_G^2 \cdot \exp \left(\frac{1}{2} \frac{\sigma_0^2}{\exp t \frac{2D_0}{\ell_c^2}} \mathcal{I}_0 \left(t \frac{2D_0}{\ell_c^2} \right) + \frac{1}{2} \frac{\sigma_0^2}{\exp \tau \frac{2D_0}{\ell_c^2}} \mathcal{I}_0 \left(\tau \frac{2D_0}{\ell_c^2} \right) \right) \times \\ \left(\exp \left(\frac{\sigma_0^2}{\exp(t + \tau) \frac{D_0}{\ell_c^2}} \mathcal{I}_0 \left(\sqrt{t \cdot \tau} \frac{2D_0}{\ell_c^2} \right) \right) - 1 \right)$$

On donne une représentation graphique de la fonction $(t, \tau) \rightarrow \langle k_{app}(t), k_{app}(\tau) \rangle$ figure (7.5).

5 Optimisation

On cherche la fréquence optimale des temps d'acquisition dans le cas d'un essai de puits décrit par le système (2). On cherche alors à déterminer le vecteur $\mathbf{t} = (t_1, \dots, t_N) \in \mathcal{T}$ des temps qui maximise l'information moyenne apportée par l'observation de la perméabilité apparente. Soit encore, le problème d'inférence suivant :

$$k_{app}(t) = \mathcal{F}[t, k(\mathbf{r})] + \boldsymbol{\eta} \quad (7.51)$$

où $\boldsymbol{\eta}$ symbolise l'erreur faite sur les observations. Celles-ci sont supposées indépendantes deux à deux et caractérisées par un bruit blanc gaussien centré de variance σ_d connue. On utilise la linéarisation au premier ordre définie précédemment et on obtient :

$$k_{app}(t) = \int_{\mathbb{R}^2} d\mathbf{r} k(\mathbf{r}) G(\mathbf{r}, t) + \boldsymbol{\eta} \quad (7.52)$$

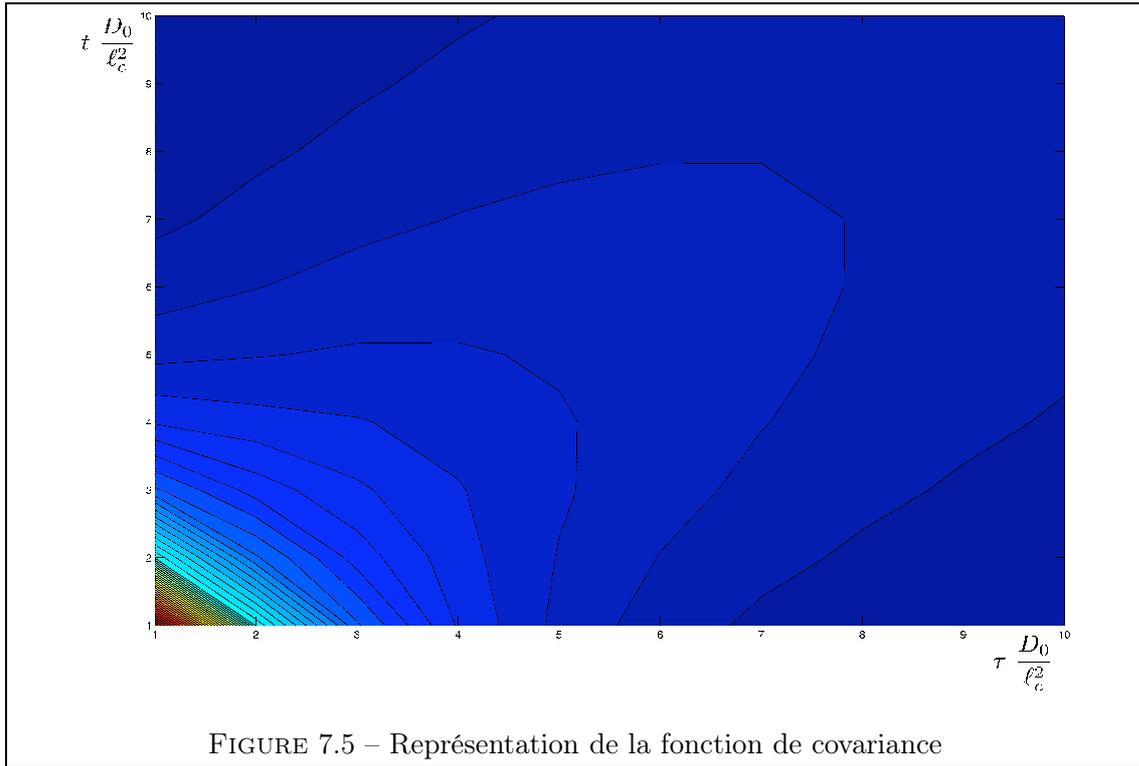


FIGURE 7.5 – Représentation de la fonction de covariance

On reprend l'approximation faite sur le noyau d'Oliver (7.49) et on obtient le problème linéaire suivant :

$$k_{app}(t) = k(\sqrt{2D_0t}) + \boldsymbol{\eta} \quad (7.53)$$

où k est exprimé par la moyenne spatiale orthoradiale (7.47). On pose alors le vecteur des observations \mathbf{d} de N composantes tel que :

$$\mathbf{d} = (k_{app}(t_1), \dots, k_{app}(t_N)) \quad (7.54)$$

On reprend la définition de l'information moyenne $\langle I \rangle$ et on définit le critère moyen suivant :

$$\langle I \rangle(t_1, \dots, t_N) = \frac{1}{2} \log \left| \mathbf{I}_N + \langle k_{app}(t_i), k_{app}(t_j) \rangle_{1 \leq i, j \leq N} \right| \quad (7.55)$$

Soit encore en prenant l'approximation (7.53) :

$$\langle I \rangle(t_1, \dots, t_N) = \frac{1}{2} \log \left| \mathbf{I}_N + \left\langle k(\sqrt{2D_0t_i}), k(\sqrt{2D_0t_j}) \right\rangle_{1 \leq i, j \leq N} \right| \quad (7.56)$$

où l'expression de la covariance est celle calculée en (7.51).

Le problème d'optimisation est alors le suivant :

Déterminer l'échantillon temporel de la pression $\mathbf{t}^* = (t_1, \dots, t_N)$ tel que :

$$\mathbf{t}^* = \underset{(t_1, \dots, t_N) \in \mathcal{T}^N}{\operatorname{argmax}} \left. \frac{1}{2} \log \left| \mathbf{I}_N + \left\langle k(\sqrt{2D_0 t_i}), k(\sqrt{2D_0 t_j}) \right\rangle_{1 \leq i, j \leq N} \right| \right|$$

où \mathcal{T} est la période d'étude de l'essai de puits.

6 Résultats numériques

Supposons que l'on dispose d'un budget permettant d'effectuer $N = 20$ mesures de pression au puits. On cherche alors l'échantillon temporel $\mathbf{t}^* = (t_1, \dots, t_{20})$ qui maximise l'expression (7.56) de l'information moyenne. De même que pour l'optimisation précédente, deux cas peuvent être envisagés : optimisation séquentielle ou globale.

6.1 Optimisation séquentielle

Dans cette approche, on cherche quel sera le temps t_{n+1} de la prochaine mesure de pression au puits qui sera la plus informative sachant que t_1, \dots, t_n mesures ont déjà été effectuées. Ce nouveau temps d'acquisition est estimé en optimisant le critère d'information moyenne. On définit alors t_{n+1} par :

$$t_{n+1} = \underset{t_{n+1} \in \mathcal{T}}{\operatorname{argmax}} \langle I \rangle(t_{n+1} | t_1, \dots, t_n)$$

On prend un intervalle d'étude suffisamment large pour que les N points puissent être représentés. En effet, la période d'étude de l'essai de puits devient une contrainte et un problème numérique pour l'optimisation. Lorsque celle-ci est courte et que la taille de l'échantillon est importante, les temps d'acquisition auront tendance à prendre les valeurs extrêmes.

On prend la valeur initiale $t_0 = 0$ et on optimise les $N - 1$ temps restants. On illustre le résultat de l'optimisation figure (7.6). En abscisse, on représente la répartition temporelle de l'échantillon et en ordonnée l'indice correspondant au temps de l'acquisition.

On constate l'évolution logarithmique de la répartition de l'échantillon. Nous observons clairement que l'information contenue dans l'essai de puits est plus importante aux temps courts, c'est-à-dire lors des premières acquisitions alors qu'aux temps longs le processus d'auto-moyennage donne une information moins significative. Cette répartition rend bien compte du phénomène physique et de la caractérisation de l'essai de puits.

6.2 Optimisation globale

Dans cette approche, nous traitons directement l'échantillon dans sa globalité. Contrairement à l'approche précédente, la problématique est posée lorsque l'on cherche à pla-

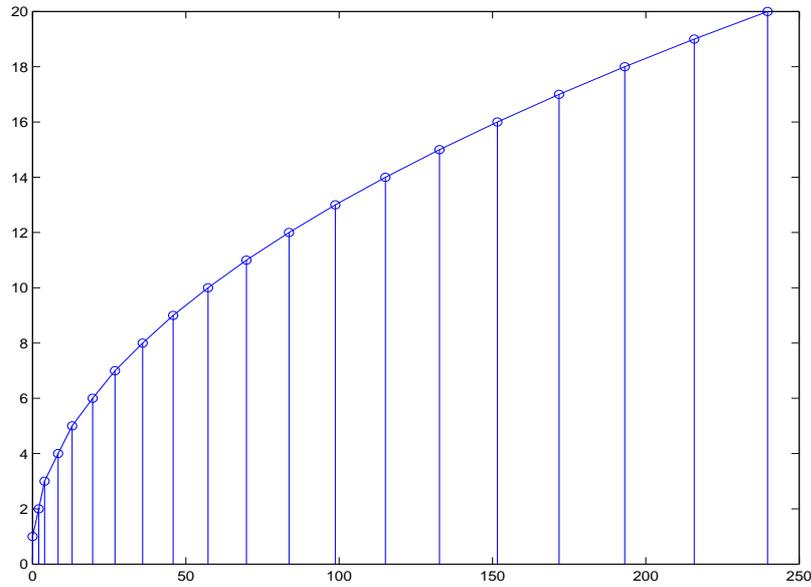


FIGURE 7.6 – Optimisation séquentielle *a priori* des N mesures temporelles de pression au puits.

nifier un échantillon temporel dans un budget. On se donne un intervalle d'acquisition et on répartit aléatoirement les N temps dans cet intervalle. L'optimisation est réalisée à l'aide de la méthode L-BFGS-B [38] que nous avons utilisée précédemment pour l'optimisation de la répartition spatiale de données (cf chapitre 6). Le problème d'optimisation est alors de définir le vecteur $\mathbf{t}^* = (t_0, \dots, t_{N-1})$ qui maximise le critère (7.56) de l'information moyenne. On retrouve le problème suivant :

$$\mathbf{t}^* = \underset{\mathbf{t} = (t_0, \dots, t_{N-1}) \in \mathcal{T}}{\operatorname{argmax}} \langle I \rangle(t_0, \dots, t_{N-1})$$

En figure (7.7), nous illustrons la répartition initiale de l'échantillon temporel et en figure (7.8) le résultat de l'optimisation.

Comme pour l'optimisation séquentielle, nous constatons que la répartition temporelle de mesure de pression est concentrée aux temps courts et plus espacée aux temps longs. Cela caractérise les propriétés stochastiques de l'essai de puits. La répartition des temps d'acquisition reflète l'évolution temporelle de l'information moyenne contenue dans un essai de puits. Le premier avantage que nous constatons est que le temps final ne limite pas le nombre d'acquisitions. En effet, l'optimisation globale permet de répartir l'ensemble des temps d'acquisitions sur l'intervalle entier et d'optimiser ainsi la densité de la répartition, ce qui n'était pas possible avec la méthode séquentielle.

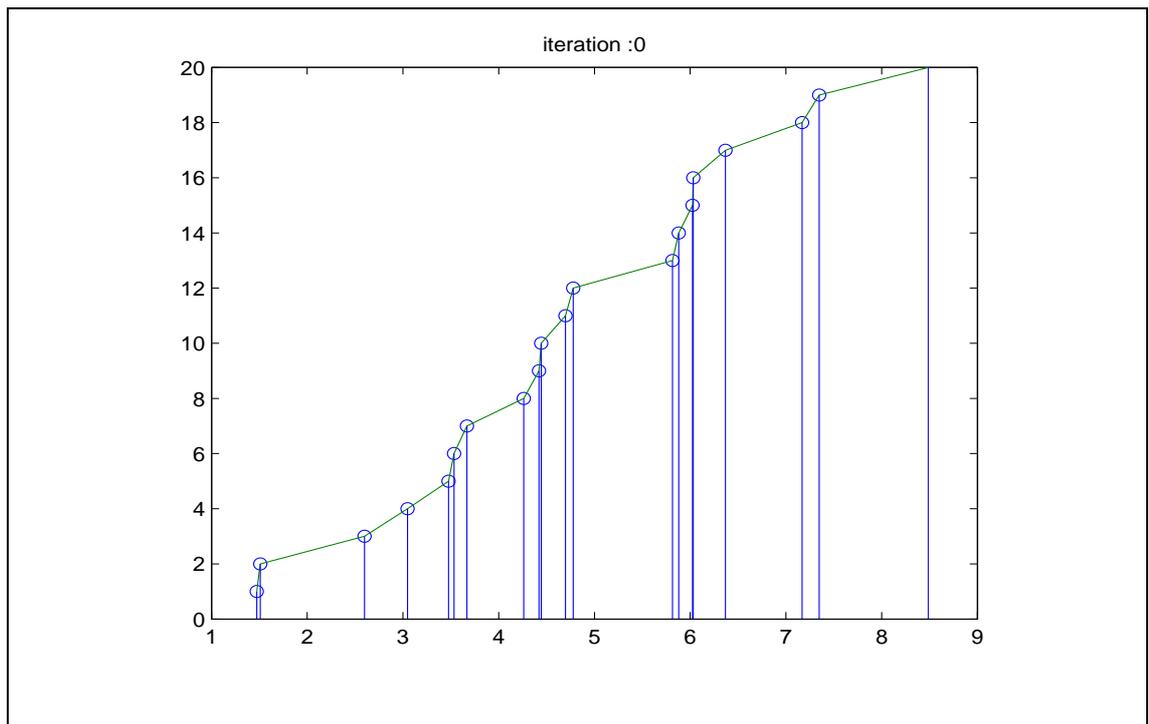


FIGURE 7.7 – Optimisation globale *a priori* des $N = 20$ mesures temporelles de pression au puits. Iteration initiale

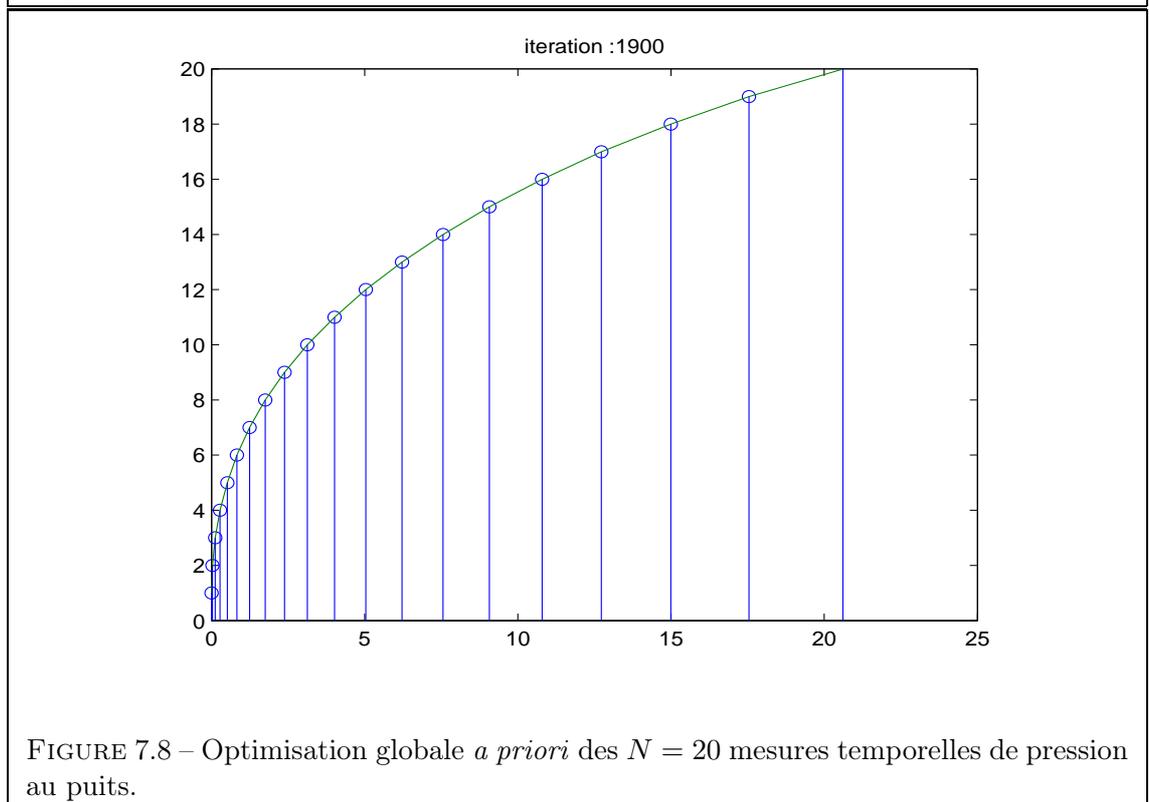


FIGURE 7.8 – Optimisation globale *a priori* des $N = 20$ mesures temporelles de pression au puits.

Comme dans le chapitre précédent, nous représentons en figure (7.9) l'évolution de l'information moyenne en fonction de la longueur de corrélation. Nous constatons que l'information moyenne évolue logarithmiquement en fonction du nombre d'observations. Comme dans le cas précédent, plus la longueur de corrélation est importante, plus l'information moyenne est petite. Inversement, plus la longueur de corrélation est petite, et plus l'information moyenne est grande.

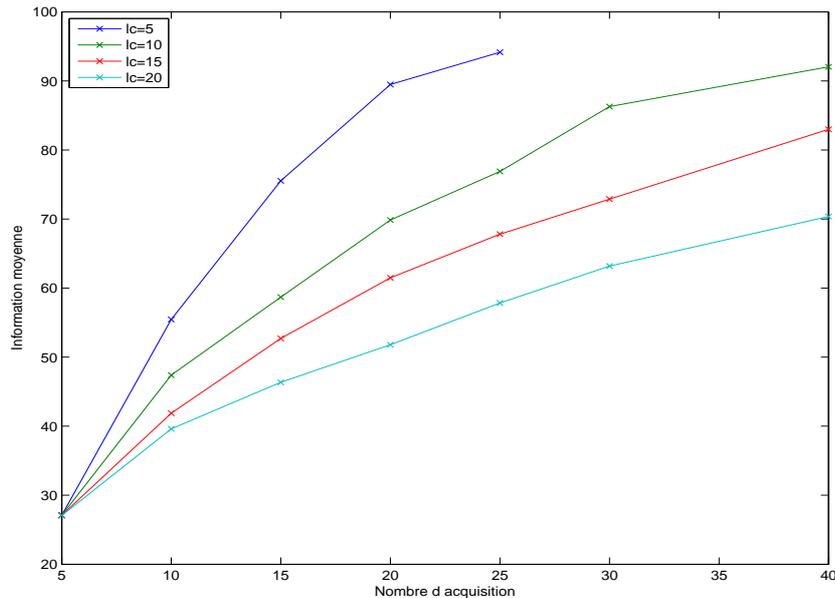


FIGURE 7.9 – Évolution de l'information moyenne en fonction du nombre de mesures en pression.

7 Conclusion

Dans ce chapitre, nous avons appliqué le critère de l'information moyenne dans le cas d'un essai de puits et nous avons cherché à optimiser la répartition temporelle de l'échantillon.

Dans notre étude, nous avons choisi le cas monophasique car la linéarisation des opérateurs mathématiques était possible. Nous avons pu ainsi obtenir une formulation analytique exploitable.

Nous avons utilisé l'expression de la perméabilité apparente pour estimer la perméabilité du milieu en fonction de la pression au puits. L'expression de la perméabilité

apparente nous a alors permis de quantifier l'apport d'information moyenne sur une nouvelle mesure en fonction de la fonction de covariance des perméabilités.

Une approximation de la fonction de covariance a été réalisée en remplaçant le noyau de pondération d'Oliver par une fonction de Dirac. L'optimisation prend ainsi en compte le paramètre de diffusion D_0 dans le milieu homogène. Nous avons pu montrer que l'information moyenne contenue dans un essai de puits est essentiellement située aux abords du puits. En effet, le caractère d'auto-moyennage de l'essai de puits ne permet pas d'obtenir une bonne estimation de la perméabilité apparente lorsque les temps d'acquisition sont longs. Il en est de même pour l'estimation des paramètres pétrophysiques. Il existe ainsi une zone d'investigation où l'information de l'essai de puits permet de bien restituer les paramètres du milieu. Il devient alors illusoire de vouloir estimer pleinement les paramètres pétrophysiques au delà de cette limite.

Nous avons simplifié la modélisation en ne prenant pas en compte des modifications de la perméabilité au niveau même du puits (*skin effect*). Une perspective directe serait donc de définir deux zones de perméabilité. La première serait proche du réservoir (modélisant un *skin*) et la seconde serait juxtaposée à la première (représentant la perméabilité du réservoir).

Dans le chapitre suivant, nous avons étendu notre approche à l'optimisation des temps d'acquisitions dans un cas diphasique. Comme dans ce chapitre, nous allons présenter le développement mathématique de la linéarisation ainsi que la recherche de la fonction de covariance qui permettra de définir notre critère d'information.

Chapitre 8

Optimisation des temps d'observation dans un cadre diphase

Sommaire

1	État du problème	74
2	Modèle Analytique	74
3	Linéarisation du modèle et dynamique du front	77
4	Calcul de la fonction de covariance	83
5	Conclusion	84

Dans ce chapitre, nous étudions l'intérêt de notre formulation de l'information moyenne pour l'optimisation des temps d'acquisition de sismiques 3D.

Les résultats précédents nous ont amenés à déterminer si dans le cadre de la sismique 3D répétée, l'information moyenne permet de définir un échantillonnage optimal des fréquences d'acquisition. L'acquisition sismique traite des volumes de données importants pouvant aller jusqu'au tera giga d'octets d'information. Malgré des moyens de plus en plus puissants dédiés aux post-traitements de cette information, les temps de calculs et les techniques impliqués pour l'acquisition jouent un rôle de plus en plus important dans les stratégies visant à la récupérations des hydrocarbures. On cherche alors à définir cette stratégie optimale de mesure.

Au lieu de travailler directement avec les cubes d'acquisition sismique, on simplifie le problème en travaillant avec l'évolution du front de saturation. Pour ce faire, on considère l'injection d'un fluide 1 dans un milieu poreux hétérogène, balayant le milieu initialement saturé en fluide 2. La direction de l'écoulement est selon la direction x . Enfin, on suppose le milieu hétérogène seulement en perméabilité absolue et on suppose que les courbes de perméabilités relatives et la distribution de la porosité sont connues ainsi que leurs caractéristiques statistiques.

Le milieu poreux réel est par nature hétérogène à différentes échelles variant du millimètre au kilomètre et influe considérablement sur le comportement du déplacement du fluide. Utilisant le résultat de King et Dunayevsky [53], Noëtinger [54] et Artus [55],[56] ont proposé un modèle de déplacement de front au premier ordre en utilisant une approche perturbante et retrouvent le critère de stabilité obtenu par Hagoort [57] et King et Dunayevsky [53]. Ils ont souligné le fait que dans le cas diphasique le tenseur des vitesses dépend de manière non linéaire de la distribution de la saturation via les fonctions de perméabilités relatives.

1 État du problème

Les travaux de Saffman et Taylor [58], et Chuoke *et al.* [59] sont une référence pour l'étude 2D du déplacement de deux fluides immiscibles, incompressibles, de viscosités différentes dont le déplacement est considéré sans résidu. La largeur de la zone de transition entre les deux fluides est considérée fine par rapport aux dimensions caractéristiques du domaine. Ils montrent que le rapport entre les deux viscosités des fluides permet de définir si le front de saturation est instable ou non à l'aide de l'équation de continuité de la pression et de la composante normale du champ de vitesse au front. Lorsque le front est instable, celle-ci augmente avec le temps en créant des structures géométriques complexes, souvent désignées sous le nom de "fingers". On pourra consulter Homsy [60] pour plus de détails.

Lorsque le milieu n'est plus considéré comme homogène, la perméabilité du milieu ainsi que la perméabilité relative des fluides ont un rôle dans l'instabilité du déplacement du front. En pratique, les courbes de perméabilités relatives sont données expérimentalement et n'ont pas de représentation non-linéaire. Lorsque nous cherchons à résoudre l'équation hyperbolique décrivant l'évolution de la saturation, la non-linéarité de la perméabilité relative fait apparaître des chocs dans la solution. Buckley et Leverett [61] ont souligné l'importance de ces chocs sur le comportement du front de saturation et de l'écoulement du fluide.

Hagoort [57] prend en compte les effets de la perméabilité relative dans les écoulements diphasiques 2D. Il montre que la stabilité du front est dépendante d'un rapport de mobilité (dépendante de la perméabilité relative) que nous définirons plus tard.

2 Modèle Analytique

On suppose que la gravité est négligée. L'écoulement de chaque fluide $i = o, w$ dans le milieu poreux est décrit par la loi de Darcy-Muskat généralisé aux écoulements diphasiques, dont la vitesse \mathbf{u} en \mathbf{r} est définie à l'aide du gradient de pression p_i par :

$$\mathbf{u}(\mathbf{r})_i = -kr_i(S_i, \mathbf{r}) \frac{K(\mathbf{r})}{\mu_i} \nabla p_i(\mathbf{r}, t) \quad (8.1)$$

où K est un tenseur de second ordre représentant la perméabilité du milieu en \mathbf{r} , ν_i est la viscosité de la phase i et p_i est la pression de la phase i en \mathbf{r} ; kr_i représente la

perméabilité relative de la phase i à la saturation S_i en \mathbf{r} . La saturation d'un fluide en milieu poreux est définie comme le rapport entre le volume de pores occupé par ce fluide et le volume de pores du volume élémentaire représentatif du milieu poreux. Pour des fluides incompressibles, la conservation du volume permet d'écrire pour les deux fluides :

$$\sum_{i=o,w} S_i = 1. \quad (8.2)$$

Par soucis de simplification des notations, on notera par la suite $S = S_w$ la saturation du fluide injecté (*i.e* alors $S_o = 1 - S_w$). La loi de Darcy est généralisée par la relation suivante :

$$\mathbf{u}(\mathbf{r}, t)_i = -\lambda_i(\mathbf{r}, S_i) \nabla p_i(\mathbf{r}, t) \quad (8.3)$$

où λ_i est la mobilité de la phase i en \mathbf{r} , définie par :

$$\lambda_i(\mathbf{r}, S_i) = kr_i(S_i, \mathbf{r}) \frac{K(\mathbf{r})}{\mu_i} \quad (8.4)$$

On suppose qu'à notre échelle d'observation, la dispersion due à l'hétérogénéité du milieu poreux est plus significative que celle induite par les effets capillaires, et donc que ceux-ci peuvent être négligés [62],[63]. On a alors :

$$p_o = p_w = p \quad (8.5)$$

En l'absence d'effets capillaire, la loi de Darcy pour le système peut alors se réécrire :

$$\mathbf{u}(\mathbf{r}, t) = \sum_i \mathbf{u}(\mathbf{r}, t)_i = -\lambda \nabla p \quad (8.6)$$

où λ est la mobilité totale du fluide, définie par :

$$\lambda = \sum_i \lambda_i(\mathbf{r}, S_i) \quad (8.7)$$

$$= K(\mathbf{r}) \left(\frac{kr_o(S_o, \mathbf{r})}{\mu_o} + \frac{kr_w(S_w, \mathbf{r})}{\mu_w} \right) \quad (8.8)$$

On note que la perméabilité absolue assimilée à un champ hétérogène est supposée isotrope. Ainsi, sa valeur est donc une fonction scalaire des coordonnées. Les courbes des fonctions de perméabilité relative sont généralement déduites du modèle de Corey [64] et elles sont fonctions de la saturation telles que :

$$kr_w(S_w) = S^2, \quad kr_o(S_o) = (1 - S)^2 \quad (8.9)$$

En l'absence de pression capillaire, l'équation de conservation de la masse pour la phase i est pour un volume de pore élémentaire traduite par :

$$\frac{\partial}{\partial t} (\phi \rho_i S_i) + \nabla \cdot (\rho_i \mathbf{u}_i) = 0 \quad (8.10)$$

où ρ_i est la densité du fluide i et où ϕ est la porosité du milieu poreux supposée constante. En supposant les fluides et la roche incompressibles, les équations de conservation de la masse sont exprimées comme suit :

$$\phi \frac{\partial S_w}{\partial t} + \nabla \cdot \mathbf{u}_w = 0 \quad (8.11)$$

$$\phi \frac{\partial S_o}{\partial t} + \nabla \cdot \mathbf{u}_o = 0 \quad (8.12)$$

La condition d'incompressibilité est alors traduite par la relation suivante :

$$\nabla \cdot \mathbf{u} = 0 \quad (8.13)$$

On reprend l'équation de Darcy pour le fluide injecté et on utilise la loi de Darcy pour le système diphasique. Avec la relation telle que $\nabla p_0 = \nabla p_w = \nabla p$, on obtient facilement l'expression de la vitesse du fluide injecté en fonction de la vitesse totale du système :

$$\mathbf{u}_w = \frac{\lambda_w(S)}{\lambda(S)} \mathbf{u} = \frac{\lambda_w(S)}{\lambda_w(S) + \lambda_o(S)} \mathbf{u} = f(S) \mathbf{u} \quad (8.14)$$

où la fonction f , dépendante des perméabilités relatives est définie comme le flux fractionnaire de Buckley-Leverett du fluide injecté [65]. On reprend maintenant l'équation de conservation de la masse pour le fluide injecté et on utilise la propriété d'incompressibilité. On obtient ainsi :

$$\phi \frac{\partial S}{\partial t} + \nabla \cdot \mathbf{u}_w = 0 \quad (8.15)$$

$$\rightarrow \phi \frac{\partial S}{\partial t} + \nabla \cdot (f(S) \mathbf{u}) = 0 \quad (8.16)$$

$$\rightarrow \phi \frac{\partial S}{\partial t} + f'(S) \mathbf{u} \cdot \nabla S = 0 \quad (8.17)$$

Cette équation connue sous le nom d'équation de Buckley-Leverett est une équation différentielle de type hyperbolique. Cette formulation montre la relation entre la saturation en fluide injecté et la vitesse totale, mettant un peu plus en évidence le couplage des équations. Ce couplage dit *visqueux* a un rôle significatif dans l'écoulement diphasique. En particulier, pour un écoulement selon la direction x , elle se réécrit pour le fluide injecté :

$$\frac{\partial S}{\partial t} + \mathbf{v} \cdot \frac{\partial S}{\partial x} = 0 \quad (8.18)$$

où la vitesse \mathbf{v} dépend de la saturation en fluide injecté et est définie par :

$$\mathbf{v}(S) = \frac{\mathbf{u} df}{\phi dS} \quad (8.19)$$

Lorsque le milieu est homogène, cette dernière équation peut être résolue avec la méthode des caractéristiques. La solution contient un choc (discontinuité). La vitesse \mathbf{v} caractérise la vitesse constante de propagation des fronts de saturation. L'écoulement est alors purement convectif et les fronts évoluent linéairement en temps. L'analyse mathématique de l'équation hyperbolique permet de définir 3 zones distinctes dans la distribution de la saturation :

- une zone de raréfaction où la saturation diminue avant d'atteindre le front,
- une zone de saut à l'interface entre la saturation en fluide injecté 1 et la saturation 2,
- une zone en amont du front où la saturation en fluide injecté 1 n'a pas atteint le fluide 2.

Finalement, on obtient une dernière équation en pression grâce à la définition de la mobilité totale :

$$\nabla (\lambda(\mathbf{r}, S) \nabla p(\mathbf{r}, t)) = 0 \quad (8.20)$$

3 Linéarisation du modèle et dynamique du front

La linéarisation consiste en l'introduction d'une petite perturbation dans les équations d'états décrites plus haut. Nous chercherons alors à étudier l'évolution de cette perturbation.

Dans un milieu supposé homogène, les fronts de saturation suivent simplement l'équation de Buckley-Leverett. De plus, le champ de vitesse \mathbf{u} est stationnaire et constant en espace. Il peut être défini par ses deux composantes telles que :

$$u_x = u_0 \quad (8.21)$$

$$u_y = 0 \quad (8.22)$$

Chaque front de saturation se déplace à la vitesse \mathbf{v} constante définie par :

$$\mathbf{v} = \frac{f'(S)}{\phi} \mathbf{u} \quad (8.23)$$

où la fonction f , définie avec l'équation (8.14), est le flux fractionnaire en fluide 1. Dans notre étude, nous supposons que f' conduit à la formation d'un choc au front x_f de saturation S_f et à une onde de raréfaction en amont. En milieu homogène, le front de saturation est défini au cours du temps par :

$$x_0(t) = \frac{f'(S)}{\phi} u_0 t = ct \quad (8.24)$$

où

$$c = \frac{f'(S)}{\phi} u_0 \quad (8.25)$$

On se place désormais dans un milieu hétérogène et on suppose que le front de saturation subit une perturbation δx autour de sa position théorique x_0 en milieu homogène. L'expression du front de saturation s'exprime au premier ordre par :

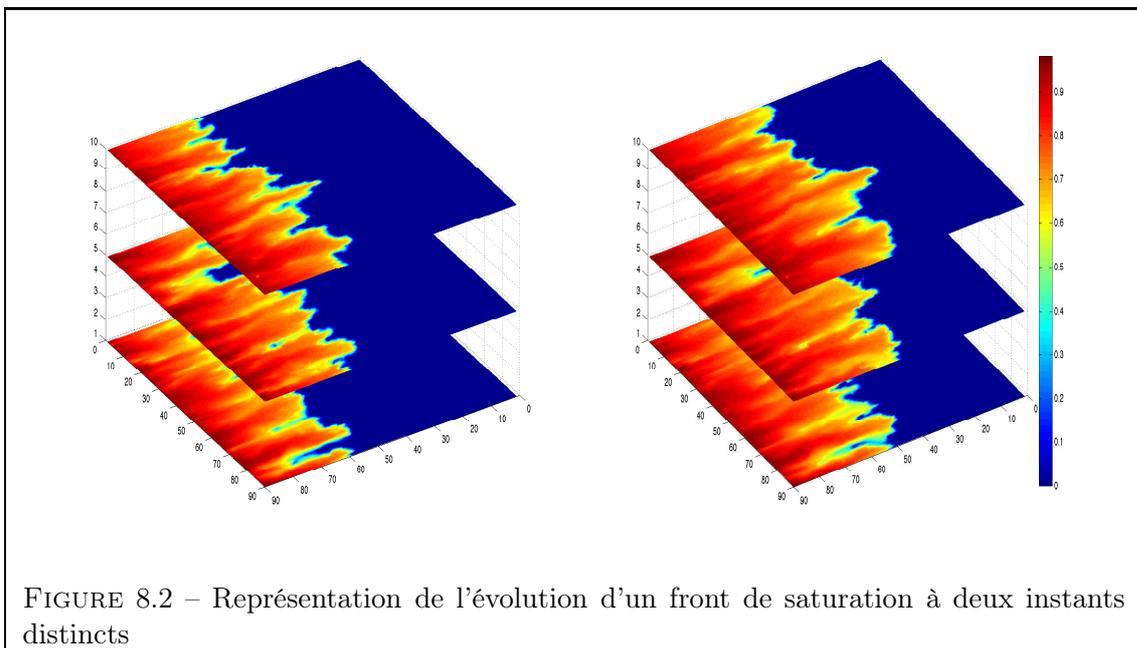
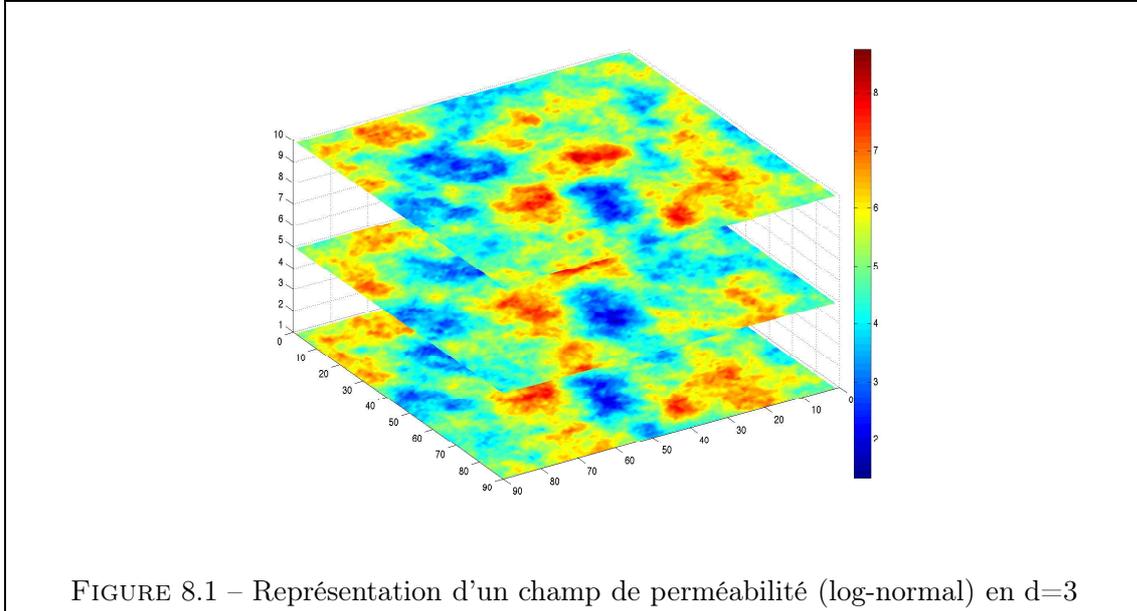
$$x(\mathbf{y}, t) = x_0(t) + \delta x(\mathbf{y}, t) \quad (8.26)$$

On réécrit cette perturbation dans le domaine fréquentiel le long de la composante \mathbf{y} , et on obtient :

$$x(\mathbf{y}, t) = x_0(t) + \frac{1}{(2\pi)^{d-1}} \int_{-\infty}^{+\infty} \delta x(\mathbf{q}, t) \cdot e^{i\mathbf{q}\cdot\mathbf{y}} d\mathbf{q} \quad (8.27)$$

où d est la dimension de l'espace. La perturbation exprimée dans le domaine fréquentiel permettra de raisonner en terme de filtrage et d'approximation hautes fréquences.

On représente la position du front de saturation à différents temps figure (8.2).



L'analyse de la stabilité du front joue un rôle dans la construction du modèle linéaire. Lorsque l'on cherche à optimiser et à assister la récupération des hydrocarbures,

il est nécessaire de comprendre les effets de stabilité du front. Lorsque celui-ci est instable, il apparaît naturellement une digitation dans l'avancée du front de saturation qui a pour conséquence de modifier radicalement les temps de percée, et *in fine*, de diminuer le volume d'huile à récupérer. Un modèle important dans l'étude des instabilités du front est celui de Saffman et Taylor [58]. Ils considèrent le déplacement de deux fluides immiscibles, dont le front de saturation sépare le domaine en deux régions : l'une complètement saturée en fluide injecté et l'autre complètement en fluide en place. Ils montrent dans ce cas particulier où les effets capillaires sont négligés que le rapport des mobilités pilote l'instabilité du front. Lorsque la mobilité du fluide injecté est plus importante que celle du fluide en place, alors il y a instabilité.

Dans le cas d'un milieu poreux, l'évolution du front est décrite par 3 phénomènes comme nous l'avons vu plus haut et il faut alors considérer le couplage des équations dans l'étude des perturbations. Nous donnons ici le résultat majeur de King et Dunayevsky [53]. En prenant en compte les effets de la zone de raréfaction et en négligeant les effets capillaires, ils donnent l'évolution $\delta x(\mathbf{q}, t)$ de la perturbation du front en fonction de la perturbation de la vitesse $\delta u(\mathbf{q}, t)$ liée au couplage pour des petites longueurs d'ondes \mathbf{q} par le système d'équations suivant :

$$\partial_t \delta x(\mathbf{q}, t) = \xi \delta u(\mathbf{q}, \xi), \quad \xi = x/t$$

où

$$\delta u(\mathbf{q}, \xi) = |q| \frac{M_f - 1}{M_f + 1} \sqrt{\frac{\lambda(\xi)}{\lambda(c)}} \exp(-p(c - \xi)) \delta x(\mathbf{q}, \xi) + \frac{\lambda(\xi)}{t} \frac{d\lambda^{-1}}{d\xi} \delta x(\mathbf{q}, \xi) \quad (8.28)$$

Ici $p = |q|t$ est le nombre d'onde adimensionnel. Aux temps longs ($t \rightarrow +\infty$), au niveau du front, on retrouve le résultat de Hagoor [57] tel que :

$$\partial_t \delta x(\mathbf{q}, t) = cA|q|\delta x(\mathbf{q}, t) \quad (8.29)$$

où

$$A = \frac{M_f - 1}{M_f + 1} \quad (8.30)$$

où M_f est le rapport des mobilités totales au passage du front tel que :

$$M_f = \frac{\sum_i \lambda^{(i)}(S_f)}{\sum_i \lambda^{(i)}(S_{min})} \quad (8.31)$$

avec S_{min} la saturation irréductible en fluide injecté.

Dans le champ hétérogène, le couplage des équations en pression et en saturation a la conséquence suivante : la perturbation du champ de perméabilité implique de même une perturbation du champ de pression et de saturation. Ainsi, peut-on supposer que le front de saturation est dépendant des perturbations liées au champ de perméabilité et au couplage visqueux. Nous prenons ici l'hypothèse de Noetinger *et al.* [66] qui considèrent que la perturbation de vitesse locale est la somme de deux fluctuations distinctes. En

2D, ils utilisent le résultat de King et Dunayevsky [53] dans leur approximation au premier ordre de l'analyse de stabilité des fronts de saturation. La première fluctuation, noté δu_a , est liée au couplage visqueux tandis que la seconde perturbation, noté δu_b , est directement liée à l'hétérogénéité du milieu poreux. On pose alors que :

$$\delta \mathbf{u}(x, \mathbf{y}, t) = \delta \mathbf{u}_a(x, \mathbf{y}, t) + \delta \mathbf{u}_b(x, \mathbf{y}, t) \quad (8.32)$$

En intégrant le résultat de Hagoort aux considérations précédentes, Artus montre que l'évolution d'un mode particulier de perturbation des plans de saturation résout l'équation suivante :

$$\begin{aligned} \partial_t \delta \hat{x}(t) &= c|q|A\delta \hat{x}(t) + \frac{f'(S_f)}{\phi} \delta \hat{\mathbf{u}}_b(x_f, t) \\ \partial_t \delta x(\mathbf{q}, t) &= c|q|A\delta x(\mathbf{q}, t) + \frac{c}{u_0} \delta \mathbf{u}_b(x_f, \mathbf{q}, t) \end{aligned} \quad (8.33)$$

Dans un cadre purement stochastique, la perturbation δx_α est un exemple de processus Gaussien à variance liée. Le processus stochastique est alors considéré comme un processus d'Ornstein-Uhlenbeck [67] unique solution de l'équation de Langevin. L'expression de $\mathbf{u}_b(x, \mathbf{q}, t)$ représente l'influence des perturbations de perméabilité sur le champ de vitesse, lorsque les profils des cartes de saturation sont solutions des équations de Buckley-Leverett, c'est-à-dire que les profils sont non-perturbés en milieu homogène. On linéarise au premier ordre la perméabilité K , la pression P et l'hétérogénéité u_b du champ de perméabilité sur le champ de vitesse et on exprime ces perturbations en mode de Fourier le long de la direction y . On obtient alors :

$$K(x, \mathbf{y}) = K_0 + \frac{1}{(2\pi)^d} \int_{-\infty}^{+\infty} d\mathbf{q} d q_1 \delta K(q_1, \mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{y}} e^{iq_1 \cdot y} \quad (8.34)$$

$$p(x, \mathbf{y}, t) = p_0(x, \mathbf{y}, t) + \frac{1}{(2\pi)^{d-1}} \int_{-\infty}^{+\infty} d\mathbf{q} \delta p(x, \mathbf{q}, t) e^{i\mathbf{q}\cdot\mathbf{y}} \quad (8.35)$$

$$\mathbf{u}_b(x, \mathbf{y}, t) = \mathbf{u}_0 + \frac{1}{(2\pi)^{d-1}} \int_{-\infty}^{+\infty} d\mathbf{q} \delta u(x, \mathbf{q}, t) e^{i\mathbf{q}\cdot\mathbf{y}} d\mathbf{q} \quad (8.36)$$

Par simplicité d'écriture, on notera lorsqu'il n'y a pas d'ambiguïté $\delta \hat{p}(x, t) = \delta p(x, \mathbf{q}, t)$ la transformée de Fourier selon \mathbf{y} de $p(x, \mathbf{y}, t)$.

On injecte le développement précédent à l'approximation de la loi de Darcy au premier ordre et nous exprimons les composantes du vecteur de la perturbation δu de la vitesse u en fonction de la perturbation de la perméabilité et de la perturbation de la pression.

$$\mathbf{u} = -\lambda \nabla p \quad (8.37)$$

$$\mathbf{u}_0 + \delta \mathbf{u} = -\kappa(S) (K_0 + \delta K) \nabla p = -\lambda \nabla (p_0 + \delta p) \quad (8.38)$$

$$\mathbf{u}_0 + \delta \mathbf{u} = -\kappa(S) K_0 \nabla p_0 - \kappa(S) \delta K \nabla p_0 - \lambda \nabla \delta p \quad (8.39)$$

On remarque que la loi de Darcy pour $\mathbf{u}_0 = -\kappa(S) K_0 \nabla p_0$ émerge de la linéarisation. On obtient alors une expression pour la perturbation de la vitesse :

$$\delta \mathbf{u} = -\kappa(S) \delta K \nabla p_0 - \lambda \nabla \delta p \quad (8.40)$$

$$\delta \mathbf{u} = \mathbf{u}_0 K_0^{-1} \delta K - \lambda \nabla \delta p \quad (8.41)$$

On reprend l'expression de la perturbation δp de la pression dans le domaine de Fourier telle que :

$$\begin{aligned}\nabla \delta p &= \nabla \cdot \left(\int_{\mathbb{R}^{d-1}} \frac{d\mathbf{q}}{(2\pi)^{d-1}} \delta \hat{p}(x, t) e^{i\mathbf{q}\cdot\mathbf{y}} \right) \\ &= \int_{\mathbb{R}^{d-1}} \frac{d\mathbf{q}}{(2\pi)^{d-1}} (i\mathbf{q}\delta \hat{p} + \partial_x \delta \hat{p}) e^{i\mathbf{q}\cdot\mathbf{y}} \\ \lambda \nabla \delta p &= \int_{\mathbb{R}^{d-1}} \frac{d\mathbf{q}}{(2\pi)^{d-1}} (i\mathbf{q}\lambda \delta \hat{p} + \lambda \partial_x \delta \hat{p}) e^{i\mathbf{q}\cdot\mathbf{y}}\end{aligned}\quad (8.42)$$

Soit encore, plus schématiquement, avec l'opérateur de la transformée de Fourier en \mathbf{y} :

$$\mathcal{F}[\lambda \nabla \delta p] = i\mathbf{q}\lambda \delta \hat{p} + \lambda \partial_x \delta \hat{p} \quad (8.43)$$

On obtient ainsi dans le domaine fréquentiel, la partie réelle et la partie imaginaire de la perturbation $\delta \hat{\mathbf{u}}$ de la vitesse dans le domaine de Fourier :

$$\mathcal{R}[\delta \hat{\mathbf{u}}] = \frac{u_0}{K_0} \delta \hat{K} - \lambda \partial_x \delta \hat{p} \quad (8.44)$$

$$\mathcal{I}[\delta \hat{\mathbf{u}}] = \lambda \delta \hat{p} \quad (8.45)$$

On applique alors la condition d'incompressibilité. On obtient alors :

$$\nabla \cdot \delta \mathbf{u} = \mathbf{u}_0 K_0^{-1} \nabla \delta K - \nabla \lambda \nabla \delta p = 0 \quad (8.46)$$

On applique alors la transformée de Fourier \mathcal{F} à cette dernière relation :

$$\mathbf{u}_0 K_0^{-1} \left(i\mathbf{q} \delta \hat{K}(x) + \partial_x \cdot \delta \hat{K}(x) \right) - (i\mathbf{q} \mathcal{F}[\lambda \nabla \delta p] + \partial_x \cdot \mathcal{F}[\lambda \nabla \delta p]) = 0 \quad (8.47)$$

On injecte ce résultat dans l'équation (8.47). On obtient alors, l'expression des parties réelle et imaginaire :

$$\mathcal{I}[\mathcal{F}[\nabla \cdot \delta \mathbf{u}]] = \mathbf{u}_0 K_0^{-1} \delta \hat{K}(x) + \lambda \partial_x \delta \hat{p} + \partial_x (\lambda \delta \hat{p}) \quad (8.48)$$

$$\mathcal{R}[\mathcal{F}[\nabla \cdot \delta \mathbf{u}]] = \mathbf{u}_0 K_0^{-1} \partial_x \hat{K}(x) - \partial_x \lambda \partial_x \delta \hat{p} + q^2 \lambda \delta \hat{p} \quad (8.49)$$

On obtient dans l'espace de Fourier une équation différentielle partielle dont l'inconnue est la perturbation de la pression et dont le terme source est une perturbation de la perméabilité. En prenant la partie réelle de cette dernière équation, nous avons :

$$\partial_x \lambda \partial_x \delta \hat{p} - q^2 \lambda \delta \hat{p} = \frac{u_0}{K_0} \partial_x \delta \hat{K} \quad (8.50)$$

Cette équation différentielle est linéaire, non-homogène à coefficients variables. On reconnaît l'équation de Helmholtz avec terme source. Pour obtenir une expression de $\delta \mathbf{u} = \delta \mathbf{u}_b$, il faut résoudre cette dernière équation différentielle. On utilise pour cela une fonction de Green G définie par :

$$G(x, x', \mathbf{q}) = \frac{1}{|\mathcal{W}|} \left(p^1(x) p^2(x') H(x' - x) + p^1(x') p^2(x) H(x - x') \right) \quad (8.51)$$

où H est la fonction de Heaviside, p^1, p^2 sont les solutions de l'équation homogène et

$$\mathcal{W} = \begin{pmatrix} p^1 & p^2 \\ -\lambda \frac{\partial p^1}{\partial x} & -\lambda \frac{\partial p^2}{\partial x} \end{pmatrix}$$

Pour les détails des calculs des fonctions p^1, p^2 et de la fonction de Green, on pourra consulter M.J. King et V.A. Dunayevsky ([53]). On obtient alors formellement :

$$\delta \hat{p}(x, t) = \left\{ \int_{-\infty}^{x_f} + \int_{x_f}^{+\infty} \right\} dx' G(x, x', \mathbf{q}) \frac{u_0}{K_0} \frac{\partial \delta \hat{K}}{\partial x'} \quad (8.52)$$

où x_f est la position du front. L'évolution d'un mode de perturbation de position du front dans l'espace de Fourier est solution au premier ordre de l'équation différentielle partielle suivante :

$$\partial_t \delta x(\mathbf{q}, t) = c|q|A \delta x(\mathbf{q}, t) + \frac{c}{u_0} \delta \hat{\mathbf{u}}_b(x_f) \quad (8.53)$$

où

$$\delta \hat{\mathbf{u}}(x_f) = \frac{u_0}{K_0} \int_{-\infty}^{+\infty} \Gamma(x_f, x, \mathbf{q}) \delta \hat{K}(x) dx \quad (8.54)$$

$$\Gamma(x_f, x, \mathbf{q}) = \delta(x_f - x) + \lambda(S_f) \frac{\partial G(x_f, x, \mathbf{q})}{\partial(x_f) \partial x} \quad (8.55)$$

Il est possible de représenter la fluctuation $\delta \mathbf{u}_b(x_f, \mathbf{y})$ de la vitesse au front x_f dans le domaine de Fourier complet pour les variables (x, \mathbf{y}) directement en fonction de la perméabilité $K(x, \mathbf{y})$. On utilise ici le résultat de Pavel [68],[69], telle que l'expression de la perturbation de la vitesse dans le domaine de Fourier est exprimée par la relation suivante :

$$\frac{c}{u_0} \delta \mathbf{u}_b(q_1, \mathbf{q}) = q^2 \int_{-\infty}^{+\infty} \frac{dq_1}{2\pi} \frac{e^{iq_1 \cdot x_f}}{q_1^2 + q^2} \left(1 + \frac{iAq_1}{q} \right) K(q_1, \mathbf{q}) \quad (8.56)$$

La position du front x_f peut être exprimée à l'aide de la vitesse du front c . On a alors la relation simple $x_f = ct$. On reprend l'équation de la dynamique du front 8.33 et on applique la transformée de Fourier en temps. On obtient ainsi :

$$\begin{aligned} -i\omega \delta x(\mathbf{q}, w) &= cAq \delta x(\mathbf{q}, w) + q^2 \int_{-\infty}^{+\infty} dt e^{i\omega t} \int_{-\infty}^{+\infty} \frac{dq_1}{2\pi} \frac{e^{iq_1 \cdot (ct)}}{q_1^2 + q^2} \left(1 + \frac{iAq_1}{q} \right) K(q_1, \mathbf{q}) \\ &= cAq \delta x(\mathbf{q}, w) + \int_{-\infty}^{+\infty} \frac{dq_1}{2\pi} \frac{q^2}{q_1^2 + q^2} \left(1 + \frac{iAq_1}{q} \right) K(q_1, \mathbf{q}) 2\pi \delta(w + cq_1) \end{aligned} \quad (8.57)$$

Soit encore,

$$\begin{aligned} -i\frac{w}{c} \delta x(\mathbf{q}, w) &= A|q| \delta x(\mathbf{q}, w) + \frac{q^2}{(w/c)^2 + q^2} \left(1 - \frac{iAw}{qc} \right) \frac{K(\frac{-w}{c}, \mathbf{q})}{c} \\ \delta x(\mathbf{q}, w) &= \frac{-q^2}{(iw/c + Aq) ((w/c)^2 + q^2)} \left(1 - \frac{iAw}{qc} \right) \frac{K(\frac{-w}{c}, \mathbf{q})}{c} \end{aligned} \quad (8.58)$$

Grâce à ce résultat, il est alors possible de calculer la fonction de covariance entre deux fronts de saturation à deux temps différents. Cette dernière sera alors utilisée pour le calcul de l'information moyenne.

4 Calcul de la fonction de covariance

Reprenons l'expression de la perturbation du front dans le mode de Fourier-Laplace. En reprenant l'expression de $q_1 = -w/c$, on obtient l'expression suivante, pour la perturbation du front de saturation :

$$\delta x(\mathbf{q}, w) = \mathcal{M}(q_1, \mathbf{q}) \frac{K(q_1, \mathbf{q})}{c} \quad (8.59)$$

où

$$\mathcal{M}(q_1, \mathbf{q}) = \frac{-q^2}{(Aq - iq_1)(q^2 + q_1^2)} \left(1 + i \frac{Aq_1}{q} \right) \quad (8.60)$$

Il est alors possible de définir une fonction de covariance entre deux fronts de saturation notée C et définie par :

$$C(\mathbf{y}, t; \boldsymbol{\xi}, \tau) = \langle x(\mathbf{y}, t), x(\boldsymbol{\xi}, \tau) \rangle \quad (8.61)$$

Soit encore en explicitant l'expression du front :

$$C(\mathbf{y}, t; \boldsymbol{\xi}, \tau) = x_0(t)x_0(\tau) + \langle \delta x(\mathbf{y}, t), \delta x(\boldsymbol{\xi}, \tau) \rangle \quad (8.62)$$

En utilisant le fait que la perturbation soit en moyenne nulle, $\langle \delta x \rangle = 0$, on obtient la relation suivante :

$$C(\mathbf{y}, t; \boldsymbol{\xi}, \tau) = x_0(t)x_0(\tau) + \langle \delta x(\mathbf{y}, t) \cdot \delta x(\boldsymbol{\xi}, \tau) \rangle \quad (8.63)$$

On applique les transformées inverses afin de calculer l'expression de la moyenne dans le domaine de Fourier-Laplace. On obtient alors :

$$\langle \delta x(\mathbf{y}, t) \cdot \delta x(\boldsymbol{\xi}, \tau) \rangle = \int \frac{dw}{2\pi} \frac{d\mathbf{q}}{(2\pi)^{d-1}} \frac{dw'}{2\pi} \frac{d\mathbf{q}'}{(2\pi)^{d-1}} e^{i\mathbf{y} \cdot \mathbf{q} + i\boldsymbol{\xi} \cdot \mathbf{q}'} \langle \delta x(\mathbf{q}, w) \cdot \delta x(\mathbf{q}', w') \rangle e^{iwt + iw'\tau} \quad (8.64)$$

On reprend alors l'expression (8.58) et on exploite les propriétés isotropiques statistiques du champ de perméabilité telles que :

$$\langle K(q_1, \mathbf{q}) K(q'_1, \mathbf{q}') \rangle = D(q_1, \mathbf{q}) (2\pi)^d \delta(q_1 + q'_1) \delta(\mathbf{q} + \mathbf{q}') \quad (8.65)$$

où D est la fonction logarithme de la perméabilité. On obtient alors :

$$\langle \delta x(\mathbf{q}, w) \cdot \delta x(\mathbf{q}', w') \rangle = \mathcal{M}(q_1, \mathbf{q}) \mathcal{M}(-q_1, -\mathbf{q}) D(q_1, \mathbf{q}) \quad (8.66)$$

La fonction de covariance est alors exprimée par :

$$\langle \delta x(\mathbf{y}, t) \cdot \delta x(\boldsymbol{\xi}, \tau) \rangle = \int \frac{dq_1}{2\pi} \frac{d\mathbf{q}}{(2\pi)^{d-1}} \mathcal{M}(q_1, \mathbf{q}) \mathcal{M}(-q_1, -\mathbf{q}) D(q_1, \mathbf{q}) e^{i\mathbf{q} \cdot (\vec{y} - \boldsymbol{\xi})} e^{iq_1 c(t - \tau)} \quad (8.67)$$

Dans notre étude, nous avons choisi une fonction gaussienne pour l'expression de D . Le calcul analytique de la fonction de covariance est donné en Annexe I. On aboutit à l'expression suivante :

$$\langle \delta x(\mathbf{y}, t) \cdot \delta x(\boldsymbol{\xi}, \tau) \rangle = \frac{k_0 \ell_c^2}{4} \int_0^1 d\beta \beta^{d/2-2} e^{-\theta\beta} \left(\frac{1+A^2}{1-A^2} (f_{\ell_c^2/4}(\beta) - 1) - g_{\bar{\mathbf{u}}, \Delta t}(\beta) \right) \quad (8.68)$$

où on définit $\mathbf{u} = \mathbf{y} - \boldsymbol{\xi}$, $\Delta t = t - \tau$ et $\theta = \frac{u^2 + (c\Delta t)^2}{\ell_c^2}$. Avec :

$$f_{m, \Delta t}(\beta) = \frac{1}{A} \frac{e^{-\frac{(c\Delta t)^2 \beta}{4m}} (A^2 / (1 - (1 - A^2)\beta) - 1)}{\sqrt{1 - (1 - A^2)\beta}}$$

$$g_{\bar{\mathbf{u}}, \Delta t}(\beta) = \frac{u^2}{u^2 + (c\Delta t)^2} + (1 - \beta) \left(\frac{d}{2} \left(1 - \frac{u^2}{u^2 + (c\Delta t)^2} \right) - \frac{1}{2} \right)$$

La fonction de covariance caractérise les propriétés d'un champ stationnaire et isotrope. Soit encore,

$$\langle \delta x(\mathbf{y}, t), \delta x(\boldsymbol{\xi}, \tau) \rangle = C(\mathbf{y} - \boldsymbol{\xi}; t - \tau)$$

On peut ainsi illustrer cette dernière fonction $\frac{\langle \delta x(\mathbf{y}, t) \cdot \delta x(\boldsymbol{\xi}, \tau) \rangle}{k_0 \ell_c^2 / 4}$ sur les axes adimensionnels (u/ℓ_c) et ($c\Delta t/\ell_c$). Nous définissons aussi son variogramme :

$$\gamma(\mathbf{u}, \Delta t) = \frac{k_0 \ell_c^2}{2} \int_0^1 d\beta \beta^{d/2-2} e^{-c\Delta t\beta/\ell_c^2} (1 - e^{-u^2\beta/\ell_c^2}) \left(\frac{1+A^2}{1-A^2} (f_{\ell_c^2/4}(\beta) - 1) - g_{\bar{\mathbf{u}}, \Delta t}(\beta) \right) \quad (8.69)$$

Nous représentons le variogramme figure (8.4) Une fonction d'information moyenne est alors exprimée à l'aide de la fonction de covariance. En reprenant les considérations des chapitres précédents, on définit le critère de l'information moyenne par :

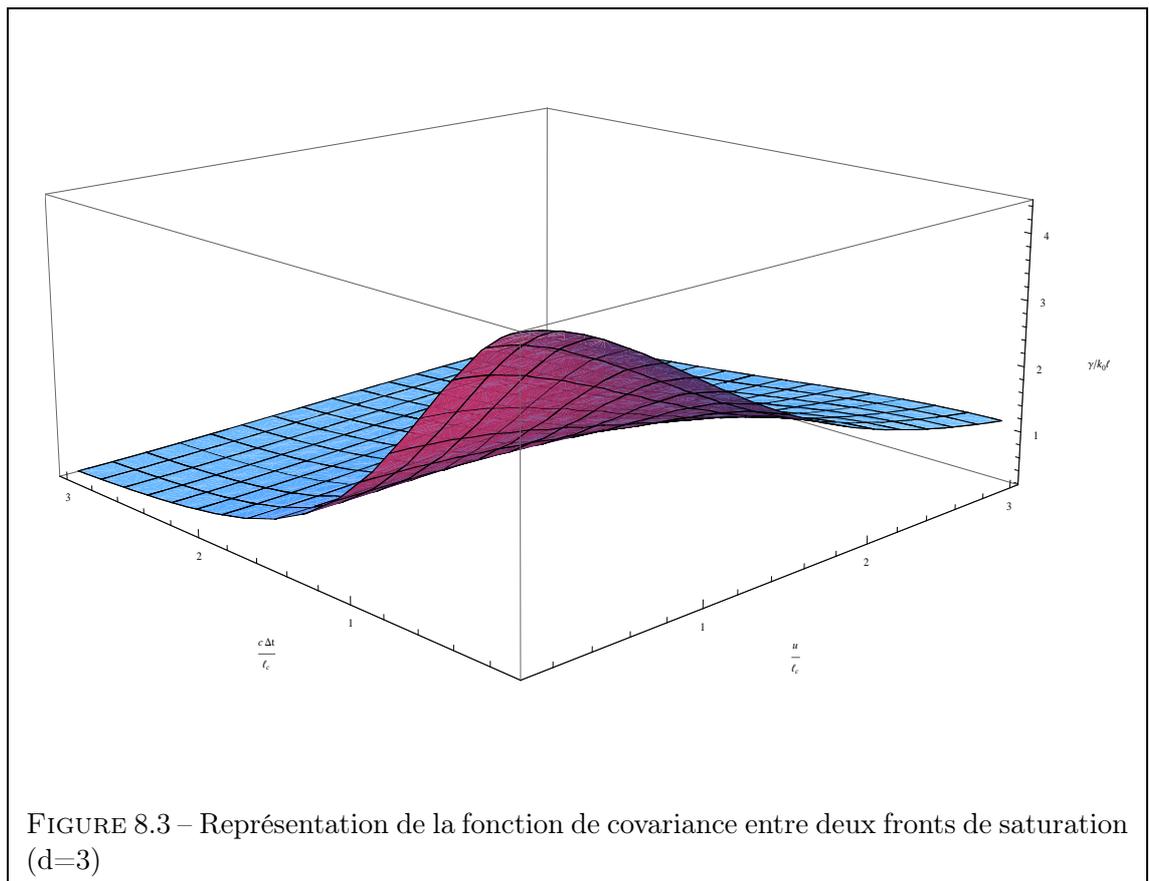
$$\langle I \rangle(t_1, \dots, t_n) = \frac{1}{2} \log |\mathcal{M}_{ij}| \quad (8.70)$$

où

$$\begin{aligned} \mathcal{M}_{ii} &= 1 + x_0^2(t_i) + \langle \delta x(\mathbf{y}, t_i) \cdot \delta x(\boldsymbol{\xi}, t_i) \rangle \\ \forall i > j, \mathcal{M}_{ij} &= x_0(t_i)x_0(t_j) + \langle \delta x(\mathbf{y}, t_i) \cdot \delta x(\boldsymbol{\xi}, t_j) \rangle \end{aligned}$$

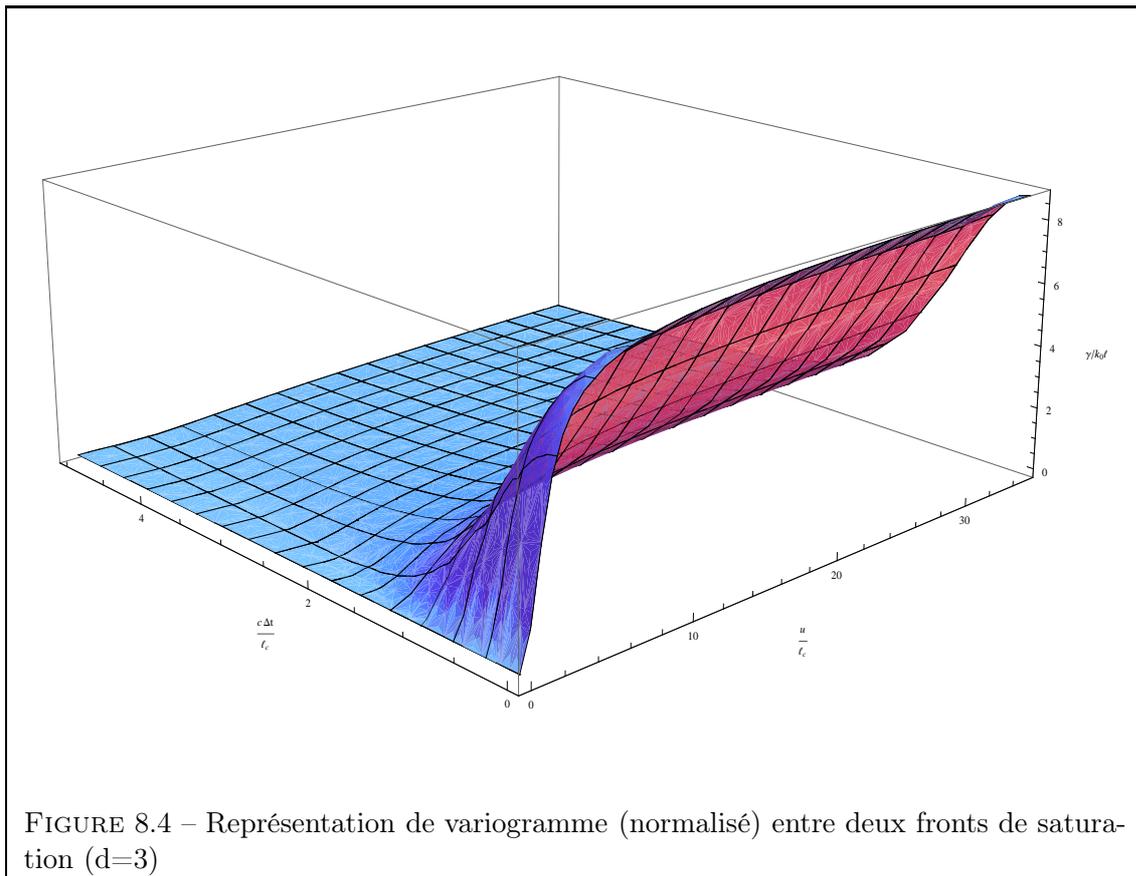
5 Conclusion

Dans ce chapitre, nous avons utilisé les lignes de courant pour suivre l'évolution du front de saturation. Nous avons défini notre critère d'information moyenne pour ainsi définir une fréquence temporelle optimale.



Une perspective directe serait de définir une stratégie cherchant à observer l'évolution d'un système. La formule que nous obtenons pour l'information moyenne est plus difficile à manier pour l'optimisation et ne permet pas de donner des résultats numériques directs comme dans les chapitres précédents. En définissant une méthode de calcul rapide et efficace de l'information moyenne, il est possible de définir aisément une stratégie optimale. Nous remarquons aussi que l'optimisation sera possible si et seulement si, les fronts de saturation sont stables ($A < 0$). On note aussi que le paramètre qui pilote la covariance est le paramètre $\theta = \frac{u^2 + (c\Delta t)^2}{\ell_c^2}$. On a ainsi un rapport de deux distances : l'une représentant la distance entre les deux fronts de saturation et l'autre représentant la longueur de corrélation. Lorsque les deux fronts de saturation se distancent d'au moins une longueur de corrélation, alors ceux-ci seront indépendants.

Une perspective à long terme serait de pouvoir estimer et ainsi de définir une stratégie permettant de suivre l'évolution du front en fonction du temps de percée d'un fluide.



Conclusion

Initialement introduites par Tarantola en géosciences, les probabilités en problème inverse permettent de mieux prendre en compte les incertitudes et sont désormais incontournables dans le processus d'intégration de données statiques et dynamiques ainsi que pour la mise à jour du modèle de réservoir. Cette dernière se fait naturellement par la formule de Bayes.

Ainsi contruit-on une distribution de probabilité dite *a posteriori* à partir d'une distribution *a priori* sur les paramètres du modèle et de la distribution de vraisemblance qui, grâce au modèle théorique de simulation, fournit une distribution de probabilité idéalisée des observations. La distribution *a posteriori* sur les paramètres permet donc la conjonction de toutes les informations disponibles du système étudié.

Dans notre étude, nous avons cherché à définir des stratégies optimales en maximisant un critère d'information. La théorie de l'information suggère de nombreux outils pour caractériser et quantifier l'information à partir d'une distribution de probabilité. Pour comparer deux distributions entre elles, nous avons utilisé la divergence de Kullback-Leibler. Pseudo-distance très utilisée en statistiques, elle nous a permis de définir un critère d'information dont la maximisation conduit à la construction d'un *design a priori* des mesures.

Dans notre étude, nous avons utilisé des distributions gaussiennes pour représenter idéalement les hétérogénéités. Pour obtenir un critère d'information exploitable, nous avons linéarisé nos opérateurs mathématiques. Le modèle ainsi linéarisé, nous avons explicité la divergence de Kullback-Leibler entre la distribution *a posteriori* et la distribution *a priori*. Afin d'effectuer la moyenne mathématique sur les observations, nous avons calculé la distribution de probabilité des observations à l'aide de la distribution de vraisemblance (le modèle mathématique) et la distribution *a priori*. Dans le cas linéaire gaussien, cette distribution est une distribution gaussienne dont la moyenne et la matrice de covariance peuvent être explicitées. Le critère d'information est ensuite moyenné à l'aide de la distribution des observations. Nous définissons alors un critère d'information moyenne sur toutes les observations possibles avant toute optimisation. C'est ce critère qui nous permet d'optimiser *a priori* l'apport de nouvelles mesures.

Dans le cas des distributions gaussiennes, le critère d'information moyenne est défini par le déterminant du rapport des matrices de covariance *a priori* et *a posteriori*

sur les paramètres. Dans nos exemples, le paramètre considéré est la perméabilité et la taille de la matrice de covariance est alors de l'ordre du million. L'inversion est alors fortement coûteuse en temps de calculs. Pour contourner ce problème, nous utilisons la bijection qui existe entre l'espace des paramètres et l'espace des observations. Grâce à la formule de Bayes, nous avons défini un critère d'optimisation moyenne dans l'espace des observations. Ce nouveau critère s'exprime avec les matrices de covariance sur les observations et la taille des matrices est alors réduite. Cette bijection est possible car les distributions sur les observations sont clairement explicitées et ne découlent pas d'une optimisation ou d'une approximation par un algorithme du type de type MCMC (Monte Carlo Markov Chain), SVM (*Support Vector Machine*) ou encore filtre bayésien particulière.

Les trois exemples que nous traitons sont associés à la détermination du champ de perméabilité à partir de trois types de mesures : des mesures statiques, des données d'essai de puits et des données de sismique 3D. Dans le premier cas, l'estimation est statique en espace et modélisée grâce au krigeage, dans le deuxième cas elle est dynamique en temps et dans le dernier cas, elle est spatio-temporelle.

Pour le krigeage, l'opérateur mathématique est une simple fonction de Dirac et le critère d'optimisation est immédiat. Son expression permet de comprendre intuitivement les résultats. Nous avons mis en évidence une limite statistique en dessous de laquelle une nouvelle mesure ne fait pas croître *a priori* significativement l'information moyenne. Cette limite s'exprime en fonction du nombre de volumes de corrélation présents dans le réservoir.

On montre que la répartition optimale des observations est uniforme dans le réservoir. La nouvelle observation se placera dans une zone dont le volume est de l'ordre du volume de corrélation. Lorsque le nombre de volumes de corrélation est égal au nombre d'observations, l'information moyenne décroît et de ce fait, l'information apportée par de nouvelles mesures est redondante. On montre que si l'on choisit d'agréger un nombre p d'observations corrélées, alors la variance de cette observation sera divisée par p . Les résultats concluants obtenus pour le krigeage ont justifié l'hypothèse d'une limite en dessous de laquelle l'information apportée par une nouvelle mesure ne fait pas croître l'information moyenne.

Nous nous sommes alors intéressés à l'intégration de données dynamiques pour l'estimation de la perméabilité. Les observations temporelles de la pression au puits permettent à l'aide de la définition de la perméabilité apparente d'estimer en moyenne la perméabilité. Le modèle d'écoulement est alors celui d'un système monophasique de porosité constante et la perméabilité est définie comme un processus aléatoire isotrope et stationnaire. Nous avons linéarisé nos opérateurs mathématiques en utilisant l'opérateur de Fourier-Laplace (resp. Fourier-Laplace inverse). La perméabilité apparente est alors explicitée à l'aide d'un produit de convolution entre le noyau pondérateur de Oliver et la perméabilité sur une couronne, considéré comme une moyenne ortho-radiale de la log-perméabilité. Nous avons alors donné une définition de la covariance de la

perméabilité apparente. Nous avons introduit ce résultat dans le critère d'information moyenne. Il nous a alors été possible d'optimiser un échantillon temporel optimal en maximisant ce critère. Nous avons constaté que l'optimisation préconise de concentrer les acquisitions aux temps courts et de les espacer aux temps longs. Ceci met en évidence le caractère d'auto-moyennage de l'essai de puits. En effet, plus les acquisitions en pression sont prises aux temps longs, plus l'information fournie par celles-ci sera moyennée. Il est alors normal de constater que l'échantillon *a priori* qui maximise l'information a des temps concentrés aux temps courts. Comme dans le cas du krigeage, la longueur de corrélation a une influence sur les courbes d'information moyenne. L'information moyenne décroît lentement en fonction de la longueur de corrélation. Il faut beaucoup plus de temps d'acquisition en pression pour estimer *a priori* le champ de perméabilité lorsque la longueur de corrélation est petite. L'intégration de données dynamiques dont le résultat sur l'optimisation est moins intuitif que pour le cas du krigeage est obtenue par l'approximation de la fonction de covariance. Une perspective directe serait alors d'utiliser la fonction complète de la covariance à proprement parler, plutôt qu'une approximation, ce qui aboutirait à estimer une intégrale double sur tout le réservoir.

D'autres applications peuvent être envisagées. On peut représenter les deux profils de l'évolution de l'information moyenne, calculée sur toutes les observations disponibles, et l'évolution de l'information. Si la courbe de l'évolution de l'information révèle une grande déviation par rapport à celle de l'information moyenne, cela signifierait que une anomalie lors de l'intégration des données dynamiques. L'étude de l'évolution de ces deux courbes pourrait alors nous amener à repenser le modèle de réservoir choisi *a priori*.

Nous avons étudié l'intégration de données sismiques 3D répétées. Celles-ci prennent une part de plus en plus importante dans le budget d'une étude. Définir une stratégie *a priori* d'intégration de ces données en fonction d'un critère d'information devient de ce fait crucial. Nous avons utilisé un modèle d'écoulement diphasique et nous avons considéré le champ de perméabilité comme un processus stochastique gaussien stationnaire et isotrope. La linéarisation du modèle a nécessité des hypothèses sur le couplage des équations en pression et en saturation. De plus, nous supposons que les deux fluides sont séparés directement par un front. On a négligé la zone de raréfaction présente derrière ce front. De plus, on suppose que le déplacement du front est stable et qu'il ne présente pas de digitation trop importante. Nous avons alors défini une expression de la covariance entre deux fronts de saturation à deux temps différents. Cette expression, initialement de quatre variables, s'exprime simplement par une fonction de deux variables : la différence entre les deux temps d'acquisition et la distance entre les deux fronts. Nous avons alors défini l'expression de l'information moyenne. Difficile à optimiser, nous donnons son expression théorique. Nous notons qu'à la distance entre les deux fronts de saturation s'ajoute le terme de la vitesse de propagation multiplié par le temps entre les deux fronts. Cette vitesse est celle de la vitesse de propagation du front dans un milieu homogène. La somme de ces deux termes est divisée par la longueur de corrélation du champ de perméabilité. Ceci donne une idée de la répartition des temps d'acquisition pour le cas de la sismique : si le champ de perméabilité était supposé homogène alors chaque profil de saturation apporterait la même information.

Dans notre cas, pour que l'information apportée par une nouvelle acquisition ne soit pas redondante, il faut que l'espace entre les fronts de saturation soit supérieur à la longueur de corrélation. L'optimisation du critère d'information que nous avons défini permet alors de surveiller l'avancement d'un front. Lorsque l'on injecte du CO_2 dans un réservoir, il faut avant tout assurer l'étanchéité du site de stockage sur des dizaines ou des centaines d'années. L'optimisation de l'information moyenne permettra de définir un échantillon temporel pour la surveillance de l'évolution du panache de CO_2 .

Annexe A

Théorie de l'Information et Formalisme Bayésien

La théorie de l'information a été introduite par Shannon [70] sous l'influence des théoriciens de l'informatique tels Turing [71], von Neumann ou encore Wiener. Le problème est alors celui de la communication entre une source et un récepteur : la source émet un message que le récepteur lit. On voudrait quantifier l'« *information* » que contient chaque message émis. Par exemple, dans le cas de la communication, il est clair que si l'émetteur dit toujours la même chose, la quantité d'information apportée par une répétition supplémentaire est nulle.

La mesure du contenu d'information peut s'effectuer à l'aide de grandeurs définies sous différents angles théoriques : entropie de Von Neumann en théorie quantique, complexités algorithmiques ou complexités de Kolmogorov ou encore quantité statistique d'information. C'est cette dernière que nous allons brièvement présenter.

Dans l'approche bayésienne, nous utilisons l'*information* contenue dans le système étudié afin de construire différentes distributions de probabilités.

Entropie

Le terme entropie fut utilisé par Shannon dans sa théorie de l'information, par analogie à la notion d'entropie existante en thermodynamique. L'entropie d'une source possède ainsi des propriétés similaires, comme notamment l'extensivité.

Entropie de Shannon

L'*entropie de Shannon* est une fonction mathématique qui correspond à la quantité d'information contenue ou délivrée par une source d'information. La définition de l'entropie d'une source selon Shannon est telle que plus la source est redondante, moins elle contient d'information.

Intuitivement, elle peut être vue comme mesurant la quantité d'*incertitude* liée à un événement aléatoire. Une autre manière de voir est de définir la quantité d'information

portée par le signal : l'information fournie par chaque évènement est fonction de son incertitude.

Il existe plusieurs définitions de l'entropie suivant que la variable aléatoire se trouve dans un espace discret ou continu.

Définition A.1 Dans le cas discret, l'entropie d'une variable aléatoire discrète X , avec n états possibles de probabilité P dans un espace \mathcal{X} est définie par

$$\mathcal{E}_P(X) = - \sum_{x \in \mathcal{X}} P(X = x) \ln P(X = x) \quad (\text{A.1})$$

Utilisant la définition de l'espérance sur un espace discret, on a comme valeur de l'entropie

$$\mathcal{E}_P(X) = \mathbb{E}_P[-\ln P(X = x)] \quad (\text{A.2})$$

Il s'agit de l'information moyenne qu'on obtient si on tire un élément x de \mathcal{X} . En théorie de l'information $I_P = -\ln P(X = x)$ est définie comme l'information propre de X et par conséquent l'entropie est vue comme la valeur moyenne de l'information propre.

Il faut noter que la quantité d'information dépend de la distribution P plus que d'un élément x particulier et on parle alors d'entropie d'une loi de probabilité.

À la différence du cas discret, l'entropie ne reste pas invariante par changement de variables dans le cas continu. Dans le cas discret, le passage à la limite ne permet pas directement d'obtenir l'expression de l'entropie continue et on a besoin d'une mesure de référence. On définit alors l'entropie comme suit.

Définition A.2 Considérons la variable aléatoire continue X à valeur dans \mathcal{X} de densité de probabilité p . L'entropie de p est exprimée à l'aide de la divergence de Küllback-Leibler par rapport à la loi uniforme μ sur \mathcal{X} par

$$\mathcal{E}_p = - \int_{\mathcal{X}} p(x) \ln \frac{p(x)}{\mu(x)} dx = \mathbb{E}_p \left[- \ln \frac{p}{\mu} \right] \quad (\text{A.3})$$

On interprète cette mesure en théorie de l'information comme la quantité qui doit être fournie pour passer d'une distribution μ à la distribution p .

Remarque A.1 La divergence de Küllback-Leibler aura un rôle particulier dans notre processus d'inversion. En effet, en calculant la divergence de la distribution a posteriori par rapport à la distribution a priori, on verra qu'il est possible de définir un critère permettant de quantifier l'information apportée par une mesure.

Exemple 1 Si on prend le cas d'une variable gaussienne de moyenne m et variance σ^2 , alors l'entropie est donnée par :

$$\begin{aligned} \mathcal{E}_P(X) &= -\mathbb{E}_P \left[\ln \left\{ (2\pi\sigma^2)^{-1/2} \exp \frac{-1}{2} \frac{(X - m)^2}{\sigma^2} \right\} \right] \\ &= \frac{1}{2} \ln(2\pi\sigma^2) + \frac{1}{2} \end{aligned}$$

Exemple 2 Dans le cas d'un vecteur \mathbf{X} gaussien N -dimensionnel tel que $\mathbf{X} \sim \mathcal{N}_N(\mathbf{m}, \mathbf{C})$,

$$\begin{aligned} \mathcal{E}_P(\mathbf{X}) &= -\mathbb{E}_P \left[\ln \left\{ ((2\pi)^N |\mathbf{C}|)^{-1/2} \exp \frac{-1}{2} (\mathbf{X} - \mathbf{m})^\top \mathbf{C}^{-1} (\mathbf{X} - \mathbf{m}) \right\} \right] \\ &= \frac{1}{2} \ln(2\pi)^N |\mathbf{C}| + \frac{N}{2} \end{aligned}$$

L'entropie d'une loi de probabilité est interprétée comme une mesure de dispersion ou une mesure du volume de corrélation. Elle est maximale quand toutes les possibilités sont équiprobables.

Remarque A.2 Le principe général de maximum d'entropie, énoncé par Jaynes [72], consiste à dire que la loi de probabilité à choisir parmi un ensemble de loi lors d'une modélisation doit être celle qui maximise l'entropie. Les contraintes proviennent alors des informations dont on dispose. Par exemple, sous contrainte de support, on choisira une loi uniforme, sous contrainte de moyenne, on choisira une loi exponentielle, sous contrainte de variance, on considérera une loi gaussienne, etc.

Les distributions de probabilité nous permettent de définir les différentes quantités d'information du système étudié. Nous donnons dans le paragraphe suivant leurs définitions.

Théorie de l'information

Selon la théorie de l'information, des données contiennent de l'information lorsque celles-ci ne sont que peu compressibles et qu'elles sont complexes.

Kullback est l'un des premiers à établir les nombreux liens qui existent entre la théorie de l'information et la statistique mathématique [73],[74]. Dans un cadre statistique, parmi un ensemble d'événements, une information désigne un ou plusieurs événements possibles. En théorie, l'information diminue l'incertitude. En pratique, la quantité d'information peut saturer et empêcher la prise de décision¹.

Information de Shannon

Soit X et Y deux variables aléatoires respectivement de loi p et q . L'information conjointe désigne l'entropie de la loi jointe. L'information de Shannon représente la quantité d'information présente en double dans X et dans Y . On définit l'information de Shannon par :

$$IM(X, Y) = \mathcal{E}_p(X) + \mathcal{E}_q(Y) - \mathcal{E}(X, Y) \quad (\text{A.4})$$

On peut définir de la même manière l'entropie relative des mesures. On cherche alors à quantifier l'information : l'élément n'est plus tiré selon la loi q mais selon la loi p .

1. En théorie de la décision, on considère qu'il faut appeler *information* seulement ce qui est susceptible d'avoir un effet sur nos décisions.

L'entropie relative appelée aussi divergence ou discrimination est alors :

$$\mathcal{E}(p|q) = \mathbb{E} \left[\ln \frac{p}{q} \right] \quad (\text{A.5})$$

Une autre quantité d'information très utilisée en statistique est l'information de Fisher.

Information de Fisher

L'information de Fisher est une notion statistique introduite par R.A. Fisher qui quantifie l'information relative à un paramètre contenu dans une distribution. Considérons la variable d'observation \mathbf{d} permettant d'estimer le paramètre inconnu \mathbf{m} . On se place dans un modèle paramétrique $(E, \mathbb{P}_{\mathbf{m}}, \mathbf{m} \in \mathcal{M})$.

- Il existe une mesure λ dite dominante σ -finie sur le modèle. On peut ainsi choisir la densité des lois $\mathbb{P}_{\mathbf{m}}$ par rapport à cette mesure dominante, pour tout \mathbf{m}

$$\frac{d \mathbb{P}_{\mathbf{m}}}{d\lambda}(\mathbf{d}) = \rho(\mathbf{d}|\mathbf{m})$$

et on suppose que cette densité est strictement positive.

- \mathcal{M} est un ouvert de \mathbb{R}^N
- Pour λ -presque pour tout $\mathbf{d} \in E$, l'application $\mathbf{d} \rightarrow \rho(\mathbf{d}|\mathbf{m})$ est différentiable sur \mathcal{M} .
- On peut dériver l'application $\mathbf{m} \rightarrow \int \rho(\mathbf{d}|\mathbf{m}) d\lambda(\mathbf{d})$ et sa dérivée est obtenue par interversion des signes intégrale et dérivée :

$$\partial_{\mathbf{m}_j} \int \rho(\mathbf{d}|\mathbf{m}) d\lambda(\mathbf{d}) = \int \partial_{\mathbf{m}_j} \rho(\mathbf{d}|\mathbf{m}) d\lambda(\mathbf{d})$$

Sous ces hypothèses, on appelle matrice d'information de Fisher, la matrice

$$\mathcal{I}(\mathbf{m}) = \mathbb{E} \left[\nabla_{\mathbf{m}^\top} \log \rho(\mathbf{d}|\mathbf{m}) \cdot \nabla_{\mathbf{m}} \log \rho(\mathbf{d}|\mathbf{m}) \middle| \mathbf{m} \right]$$

Lorsque l'application $d \rightarrow \rho(\mathbf{d}|\mathbf{m})$ est deux fois dérivables, alors on peut définir la matrice d'information de Fisher à l'aide de la matrice Hessienne. On obtient ainsi

$$\mathcal{I}(\mathbf{m}) = -\mathbb{E} \left(\frac{\partial^2}{\partial \mathbf{m}_i \partial \mathbf{m}_j} \log \rho(\mathbf{d}|\mathbf{m}) \middle| \mathbf{m} \right)_{1 \leq i, j \leq M} \quad (\text{A.6})$$

La notion d'information de Fisher dans la statistique est liée à la définition de la notion de score du modèle.

Score

Le vecteur score est le gradient du logarithme de la fonction de vraisemblance du modèle. On le définit comme suit :

$$\mathbf{S}_{\mathbf{m}} = \nabla_{\mathbf{m}} \ln \rho(\mathbf{d}|\mathbf{m}) = \left(\frac{1}{\rho(\mathbf{d}|\mathbf{m})} \frac{\partial}{\partial \mathbf{m}_i} \rho(\mathbf{d}|\mathbf{m}) \right)_{1 \leq i \leq M}$$

C'est un vecteur centré, c'est-à-dire que

$$\mathbb{E} \left(\nabla_{\mathbf{m}} \ln \rho(\mathbf{d}|\mathbf{m}) \middle| \mathbf{m} \right) = 0$$

On remarque ainsi que la matrice d'information de Fisher est la matrice de variance-covariance du vecteur score.

La fonction **score** mesure la qualité de discrimination entre les différents paramètres du modèle au vu des observations et quantifie la qualité du modèle.

Remarque A.3 *L'information de Fisher ainsi que la fonction de score ont un intérêt lorsque l'estimation des paramètres par la formule de Bayes fait intervenir des hyperparamètres qui pilotent les distributions utilisées. Lorsque les hyperparamètres sont connus, la matrice d'information de Fisher est constante et ne joue donc pas un rôle significatif dans le processus d'inférence.*

Le formalisme bayésien

Il existe deux écoles de Statistique. La première regroupe des théoriciens de l'échantillonnage, ou plus généralement des probabilistes fréquentistes, qui cherchent à avoir des méthodes fonctionnant à chaque fois, en ayant le moins d'hypothèses possibles. La deuxième école, celle des Bayésiens, font de l'inférence qui prend en compte toute l'information disponible. Ils essaient de répondre à une question particulière sachant un ensemble particulier de données.

La théorie de l'échantillonnage est la méthode la plus largement utilisée en Statistique, et il est fréquent de voir chez beaucoup d'auteurs les quantités telles que les intervalles de confinement, les p -valeurs, etc. Bien que dans beaucoup de cas, l'approche bayésienne des probabilités et l'approche fréquentiste sont assez similaires, on peut aussi trouver des cas, où il existe des différences significatives, où l'approche fréquentiste échoue devant l'approche bayésienne.

Dans ce paragraphe, nous donnons le formalisme bayésien qui permettra de définir et de construire la distribution de probabilité *a posteriori*.

Définition A.3 *On appelle modèle statistique bayésien tout modèle statistique $(\mathcal{D}, (\mathbb{P}_{\mathbf{m}})_{\mathbf{m} \in \mathcal{M}})$ tel que l'espace des paramètres \mathcal{M} soit muni d'une loi de probabilité π , dite loi a priori (une tribu \mathcal{T} étant fixée sur \mathcal{M}).*

Distribution de vraisemblance

La distribution de vraisemblance donne la répartition des observations sachant un modèle. Son entropie donne l'incertitude du modèle courant et ainsi l'information contenue dans ce modèle.

Définition A.4 *On dit qu'un modèle statistique est dominé s'il existe une mesure σ -finie μ sur \mathcal{D} telle que, pour tout $\mathbf{m} \in \mathcal{M}$, la loi soit absolument continue par rapport à μ . Dans un tel cas, on dit que μ est une dominante du modèle et on appelle vraisemblance du modèle toute application ρ de $\mathcal{D} \times \mathcal{M}$ dans \mathbb{R}^+ telle que, pour $\mathbf{m} \in \mathcal{M}$, l'application partielle $\rho(\cdot, \mathbf{m})$ de \mathcal{D} dans \mathbb{R}^+ soit une densité de $\mathbb{P}_{\mathbf{m}}$ par rapport à μ .*

Dans notre étude, nous abordons seulement le cas paramétrique. Nous supposons que les observations d_1, d_2, \dots, d_N proviennent de lois de probabilité paramétrique, c'est-à-dire que chaque d_i a une distribution de densité $\rho_i(d_i | m_i, d_1, \dots, d_{i-1})$. Le modèle est défini synthétiquement par la notation suivante $\mathbf{d} \sim \rho(\mathbf{d} | \mathbf{m})$.

Distribution *a priori*

La distribution *a priori* représente l'information disponible sur le paramètre \mathbf{m} , c'est-à-dire préalable à l'observation \mathbf{d} . Il faut pouvoir déterminer la distribution *a priori* π de sorte qu'elle maximise le gain d'information.

Remarque A.4 *La définition de la distribution a priori dans l'analyse statistique bayésienne est importante car elle est la meilleure façon de résumer l'information (et ainsi le manque d'information) sur le paramètre \mathbf{m} ainsi que l'incertitude résiduelle. Elle permet implicitement l'incorporation de cette information inexacte dans le processus d'inversion ou de décision.*

Distribution *a posteriori*

La distribution de vraisemblance et la distribution *a priori* définies, il est alors possible de construire complètement le modèle bayésien. Celui-ci construit une distribution dite *a posteriori* de la variable aléatoire $\mathbf{m} | \mathbf{d}$ définie par le théorème de Bayes telle que

$$\sigma(\mathbf{m} | \mathbf{d}) = \frac{\rho(\mathbf{d} | \mathbf{m}) \pi(\mathbf{m})}{\int_{\mathcal{M}} \rho(\mathbf{d} | \mathbf{m}) \pi(\mathbf{m}) \, \mathbf{d}\mathbf{m}} \quad (\text{A.7})$$

En terme statistique, la distribution *a posteriori* actualise l'information contenue dans la distribution *a priori* sur le paramètre \mathbf{m} en extrayant l'information contenue dans la distribution de *vraisemblance* de l'observation \mathbf{d} . On peut réécrire l'égalité précédente synthétiquement

$$\text{posterior} = \frac{\text{vraisemblance} \times \text{prior}}{\text{évidence}} \quad (\text{A.8})$$

Distributions résultantes

Il est possible de définir différentes distributions autres que la distribution *a posteriori*.

(i) *Distribution jointe de (d, m) :*

$$\sigma(\mathbf{d}, \mathbf{m}) = \rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m}) \quad (\text{A.9})$$

(ii) *Distribution marginale de d :*

$$\sigma(\mathbf{d}) = \int_{\mathcal{M}} \rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m}) d\mathbf{m} \quad (\text{A.10})$$

(iii) *Distribution prédictive de \mathbf{d}^f , où $\mathbf{d}^f \sim g(\mathbf{d}^f|\mathbf{d}, \mathbf{m})$:*

$$\sigma(\mathbf{d}^f|\mathbf{d}) = \int_{\mathcal{M}} g(\mathbf{d}^f|\mathbf{d}, \mathbf{m})\sigma(\mathbf{m}|\mathbf{d}) d\mathbf{m} \quad (\text{A.11})$$

La théorie de l'information et le formalisme bayésien apparaissent comme des outils incontournables pour la quantification des incertitudes. La définition de l'entropie d'une variable aléatoire permet de mesurer l'incertitude associée à cette variable via sa distribution de probabilité. La formule de Bayes met à jour le modèle *a priori* via la distribution de vraisemblance et permet ainsi d'estimer un paramètre sachant les observations.

Annexe B

Calcul de la distribution *a posteriori* dans le cas linéaire gaussien

Considérons le modèle linéaire suivant :

$$\mathbf{d} = \mathcal{F} \cdot \mathbf{m} + \boldsymbol{\eta} \quad (\text{B.1})$$

où \mathcal{F} est déterministe et $\boldsymbol{\eta}$ est indépendant de \mathbf{m} . On considère le vecteur $\mathbf{d} \in \mathbb{R}^N$ de N observations, et le vecteur $\mathbf{m} \in \mathbb{R}^M$ des paramètres inconnus. On suppose que $\boldsymbol{\eta}$ est un processus gaussien centré de covariance \mathbf{C}_d . En supposant que notre distribution *a priori* π sur nos paramètres \mathbf{m} est gaussienne de moyenne \mathbf{m}_0 et covariance \mathbf{C}_0 , alors, on peut expliciter la distribution *a posteriori* :

$$\rho(\mathbf{m}|\mathbf{d}) \propto f(\mathbf{d}|\mathbf{m})\pi(\mathbf{m}) \quad (\text{B.2})$$

où f est notre distribution de vraisemblance telle que :

$$f(\mathbf{d}|\mathbf{m}) = ((2\pi)^N |\mathbf{C}_d|)^{-1/2} \exp \frac{-1}{2} \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 \quad (\text{B.3})$$

où l'on a défini la norme $\|\mathbf{x}\|_{\mathbf{A}}^2 = \mathbf{x}^t \mathbf{A} \mathbf{x}$. Alors,

$$\begin{aligned} \rho(\mathbf{m}|\mathbf{d}) &\propto f(\mathbf{d}|\mathbf{m})\pi(\mathbf{m}) \\ &\propto \exp \frac{-1}{2} \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 \cdot \exp \frac{-1}{2} \|\mathbf{m} - \mathbf{m}_0\|_{\mathbf{C}_0^{-1}}^2 \\ &\propto \exp \frac{-1}{2} \left\{ \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 + \|\mathbf{m} - \mathbf{m}_0\|_{\mathbf{C}_0^{-1}}^2 \right\} \\ &\propto \exp \frac{-1}{2} \left\{ \|\mathbf{d}\|_{\mathbf{C}_d^{-1}}^2 + \|\mathbf{m}_0\|_{\mathbf{C}_0^{-1}}^2 + \|\mathcal{F} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 + \|\mathbf{m}\|_{\mathbf{C}_0^{-1}}^2 - 2\mathbf{m}^t (\mathcal{F}^t \mathbf{C}_d^{-1} \mathbf{d} + \mathbf{C}_0^{-1} \mathbf{m}_0) \right\} \\ &\propto \exp \frac{-1}{2} \left\{ \|\mathbf{m}\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} + \mathbf{C}_0^{-1}}^2 - 2\mathbf{m}^t (\mathcal{F}^t \mathbf{C}_d^{-1} \mathbf{d} + \mathbf{C}_0^{-1} \mathbf{m}_0) \right\} \end{aligned}$$

On pose $\tilde{\mathbf{C}} = \mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} + \mathbf{C}_0^{-1}$. Et donc,

$$\begin{aligned} \rho(\mathbf{m}|\mathbf{d}) &\propto \exp \frac{-1}{2} \left\{ \|\mathbf{m}\|_{\tilde{\mathbf{C}}}^2 - 2\mathbf{m}^t (\mathcal{F}^t \mathbf{C}_d^{-1} \mathbf{d} + \mathbf{C}_0^{-1} \mathbf{m}_0) \right\} \\ &\propto \exp \frac{-1}{2} \left\{ \|\mathbf{m} - \tilde{\mathbf{C}}^{-1} (\mathcal{F}^t \mathbf{C}_d^{-1} \mathbf{d} + \mathbf{C}_0^{-1} \mathbf{m}_0)\|_{\tilde{\mathbf{C}}}^2 - \|\tilde{\mathbf{C}}^{-1} (\mathcal{F}^t \mathbf{C}_d^{-1} \mathbf{d} + \mathbf{C}_0^{-1} \mathbf{m}_0)\|_{\tilde{\mathbf{C}}^{-1}}^2 \right\} \\ &\propto \exp \frac{-1}{2} \left\{ \|\mathbf{m} - \tilde{\mathbf{C}}^{-1} (\mathcal{F}^t \mathbf{C}_d^{-1} \mathbf{d} + \mathbf{C}_0^{-1} \mathbf{m}_0)\|_{\tilde{\mathbf{C}}}^2 \right\} \end{aligned}$$

On reconnaît une distribution gaussienne de moyenne *a posteriori* :

$$\begin{aligned} \mathbf{m} &= (\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} + \mathbf{C}_0^{-1})^{-1} (\mathcal{F}^t \mathbf{C}_d^{-1} \mathbf{d} + \mathbf{C}_0^{-1} \mathbf{m}_0) \\ &= (\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} + \mathbf{C}_0^{-1})^{-1} (\mathcal{F}^t \mathbf{C}_d^{-1} \mathbf{d} + \mathbf{C}_0^{-1} \mathbf{m}_0 + \mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} (\mathbf{m}_0 - \mathbf{m}_0)) \\ &= \mathbf{m}_0 + (\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} + \mathbf{C}_0^{-1})^{-1} \mathcal{F}^t \mathbf{C}_d^{-1} (\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0) \end{aligned}$$

et la covariance *a posteriori*

$$\mathbf{C} = (\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} + \mathbf{C}_0^{-1})^{-1} \quad (\text{B.4})$$

Et prenant l'égalité matricielle de Woodburry :

$$(\mathbf{BCD} + \mathbf{A})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{C}^{-1} + \mathbf{D} \mathbf{A}^{-1} \mathbf{B})^{-1} \mathbf{D} \mathbf{A}^{-1}$$

et en remarquant l'égalité :

$$\begin{aligned} \mathcal{F}^t \mathbf{C}_d^{-1} (\mathcal{F} \mathbf{C}_0 \mathcal{F}^t + \mathbf{C}_d) &= \mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} \mathbf{C}_0 \mathcal{F}^t + \mathcal{F}^t \\ &= (\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} + \mathbf{C}_0^{-1}) \mathbf{C}_0 \mathcal{F}^t \end{aligned}$$

On obtient,

$$\mathbf{m} = \mathbf{m}_0 + \mathbf{C}_0 \mathcal{F}^t (\mathcal{F} \mathbf{C}_0 \mathcal{F}^t + \mathbf{C}_d)^{-1} (\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0) \quad (\text{B.5})$$

$$\mathbf{C} = \mathbf{C}_0 - \mathbf{C}_0 \mathcal{F}^t (\mathcal{F} \mathbf{C}_0 \mathcal{F}^t + \mathbf{C}_d)^{-1} \mathcal{F} \mathbf{C}_0 \quad (\text{B.6})$$

Annexe C

Calcul de l'information apportée par une mesure gaussienne

Soit la distribution conditionnelle des observations sachant les paramètres \mathbf{m} :

$$\rho(\mathbf{d}|\mathbf{m}) = ((2\pi)^N |\mathbf{C}_d|)^{-1/2} \exp \frac{-1}{2} \|\mathbf{d} - \mathcal{H} \cdot \mathbf{m}\|_{\mathbf{C}_d}^2$$

et la distribution *a priori* sur \mathbf{m} :

$$\rho(\mathbf{m}) = ((2\pi)^M |\mathbf{C}_0|)^{-1/2} \exp \frac{-1}{2} \|\mathbf{m} - \mathbf{m}_0\|_{\mathbf{C}_0}^2$$

La distribution *a posteriori* est aussi gaussienne définie par :

$$\rho(\mathbf{m}|\mathbf{d}) = (2\pi)^M |\mathbf{C}_q| \exp \frac{-1}{2} \|\mathbf{m} - \mathbf{m}_q\|_{\mathbf{C}_q}^2$$

où \mathbf{m}_q et \mathbf{C}_q sont respectivement la moyenne et la covariance *a posteriori* définies par :

$$\begin{aligned} \mathbf{m}_q &= \mathbf{m}_0 + \mathbf{C}_0 \mathcal{H}^t (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1} (\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_0) \\ \mathbf{C}_q &= \mathbf{C}_0 - \mathbf{C}_0 \mathcal{H}^t (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1} \mathcal{H} \mathbf{C}_0 \end{aligned}$$

Par définition, l'information I apportée par une mesure gaussienne d'*a priori* gaussien est donnée par la divergence de Kullback-Leiber entre les distributions *a posteriori* et *a priori* :

$$\begin{aligned} I &= I(\rho(\mathbf{m})) = \int_{\mathcal{M}} \rho(\mathbf{m}|\mathbf{d}) \log \frac{\rho(\mathbf{m}|\mathbf{d})}{\rho(\mathbf{m})} d\mathbf{m} \\ &= \frac{1}{2} \log \frac{|\mathbf{C}_0|}{|\mathbf{C}_q|} + \frac{1}{2} \text{tr} [\mathbf{C}_0^{-1} \mathbf{C}_q] + \frac{1}{2} \|\mathbf{m}_q - \mathbf{m}_0\|_{\mathbf{C}_0^{-1}}^2 - \frac{M}{2} \end{aligned}$$

On peut expliciter et simplifier l'écriture de I en reprenant les définitions de \mathbf{m}_q et de \mathbf{C}_q . Ainsi, on obtient :

$$\begin{aligned} \frac{1}{2} \text{tr}[\mathbf{C}_0^{-1} \mathbf{C}_q] &= \frac{1}{2} \text{tr} \left[\mathbb{I}_M - \mathcal{H}^t (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1} \mathcal{H} \mathbf{C}_0 \right] \\ &= \frac{M}{2} - \frac{1}{2} \text{tr} \left[\mathcal{H}^t (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1} \mathcal{H} \mathbf{C}_0 \right] \end{aligned}$$

$$\begin{aligned} \|\mathbf{m}_q - \mathbf{m}_0\|_{\mathbf{C}_0^{-1}}^2 &= (\mathbf{m}_q - \mathbf{m}_0)^t \mathbf{C}_0^{-1} (\mathbf{m}_q - \mathbf{m}_0) \\ &= (\mathbf{m}_q - \mathbf{m}_0)^t \mathcal{H}^t (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1} (\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_0) \\ &= (\mathbf{C}_0 \mathcal{H}^t (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1} (\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_0))^t \mathcal{H}^t (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1} (\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_0) \\ &= (\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_0)^t (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1} \mathcal{H} \mathbf{C}_0 \mathcal{H}^t (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1} (\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_0) \\ &= \|\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_0\|_{\mathbf{H}}^2 \end{aligned}$$

$$\text{où } \mathbf{H} = (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1} \mathcal{H} \mathbf{C}_0 \mathcal{H}^t (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1}$$

De même,

$$\begin{aligned} \frac{1}{2} \log \frac{|\mathbf{C}_0|}{|\mathbf{C}_q|} &= \frac{-1}{2} \log \frac{|\mathbf{C}_q|}{|\mathbf{C}_0|} = \frac{-1}{2} \log |\mathbf{C}_0^{-1} \mathbf{C}_q| \\ &= \frac{-1}{2} \log | \mathbb{I}_M - \mathcal{H}^t (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1} \mathcal{H} \mathbf{C}_0 | \\ &= \frac{-1}{2} \log | \mathbb{I}_N - \mathcal{H} \mathbf{C}_0 \mathcal{H}^t (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1} | \end{aligned}$$

en utilisant l'égalité de Schur.

Et donc, on obtient :

$$\begin{aligned} \frac{1}{2} \log \frac{|\mathbf{C}_0|}{|\mathbf{C}_q|} &= \frac{-1}{2} \log | \mathbf{C}_d (\mathbf{C}_d + \mathcal{H} \mathbf{C}_0 \mathcal{H}^t)^{-1} | \\ &= \frac{1}{2} \log | \mathbf{C}_d + \mathcal{H} \mathbf{C}_0 \mathcal{H}^t | - \frac{1}{2} \log | \mathbf{C}_d | \end{aligned}$$

On exprime donc l'information I :

$$2I = \log | \mathbf{C}_d + \mathcal{H} \mathbf{C}_0 \mathcal{H}^t | - \log | \mathbf{C}_d | + \|\mathbf{d} - \mathcal{H} \cdot \mathbf{m}_0\|_{\mathbf{H}}^2 - \text{tr} \left[\mathcal{H}^t (\mathcal{H} \mathbf{C}_0 \mathcal{H}^t + \mathbf{C}_d)^{-1} \mathcal{H} \mathbf{C}_0 \right]$$

Annexe D

Formule de Schur

La formule de Schur permet de décomposer le déterminant d'une matrice partitionnée en un produit à l'aide de ses éléments. Supposons que la matrice \mathbf{A}_1 est de taille $(m \times m)$ et \mathbf{A}_4 de taille $(n \times n)$, alors, on établit la relation, en supposant que la matrice \mathbf{A}_1 est non singulière.

$$\begin{aligned} \begin{vmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{vmatrix} &= \begin{vmatrix} \mathbf{I}_m & 0 \\ -\mathbf{A}_3\mathbf{A}_1^{-1} & \mathbf{I}_n \end{vmatrix} \begin{vmatrix} \mathbf{A}_1 & \mathbf{A}_1 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{vmatrix} = \begin{vmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ 0 & \mathbf{A}_4 - \mathbf{A}_3\mathbf{A}_1^{-1}\mathbf{A}_2 \end{vmatrix} \\ &= |\mathbf{A}_1| |\mathbf{A}_4 - \mathbf{A}_3\mathbf{A}_1^{-1}\mathbf{A}_2| \end{aligned}$$

On établit de même, en supposant que \mathbf{A}_4 est non singulière,

$$\begin{aligned} \begin{vmatrix} \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{vmatrix} &= \begin{vmatrix} \mathbf{I}_m & -\mathbf{A}_2\mathbf{A}_4^{-1} \\ 0 & \mathbf{I}_n \end{vmatrix} \begin{vmatrix} \mathbf{A}_1 & \mathbf{A}_1 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{vmatrix} = \begin{vmatrix} \mathbf{A}_1 - \mathbf{A}_3\mathbf{A}_4^{-1}\mathbf{A}_2 & 0 \\ \mathbf{A}_3 & \mathbf{A}_4 \end{vmatrix} \\ &= |\mathbf{A}_4| |\mathbf{A}_1 - \mathbf{A}_3\mathbf{A}_4^{-1}\mathbf{A}_2| \end{aligned}$$

En choisissant, astucieusement $\mathbf{A}_1 = \mathbf{I}_m$ et $\mathbf{A}_4 = \mathbf{I}_n$, on a l'égalité

$$\begin{aligned} \begin{vmatrix} \mathbf{I}_m & \mathbf{A}_2 \\ \mathbf{A}_3 & \mathbf{I}_n \end{vmatrix} &= \begin{vmatrix} \mathbf{I}_m & -\mathbf{A}_2\mathbf{I}_n^{-1} \\ 0 & \mathbf{I}_n \end{vmatrix} \begin{vmatrix} \mathbf{I}_m & \mathbf{I}_m \\ \mathbf{A}_3 & \mathbf{I}_n \end{vmatrix} \\ &= |\mathbf{I}_m| |\mathbf{I}_n - \mathbf{A}_3\mathbf{I}_m^{-1}\mathbf{A}_2| = |\mathbf{I}_n| |\mathbf{I}_m - \mathbf{A}_2\mathbf{I}_n^{-1}\mathbf{A}_3| \end{aligned}$$

Soit encore,

$$|\mathbf{I}_n - \mathbf{A}_3\mathbf{A}_2| = |\mathbf{I}_m - \mathbf{A}_2\mathbf{A}_3|$$

Annexe E

Calcul du déterminant après agrégation de p points

On considère le déterminant de la matrice suivante :

$$M = |\mathbb{I}_N + \mathbf{C}_d^{-1} \cdot \mathcal{F} \mathcal{C}_0 \mathcal{F}^t| = \begin{vmatrix} 1 + \sigma_{11} & \sigma_{12} & \dots & \sigma_{1p} & \sigma_{1p+1} & \dots & \sigma_{1N} \\ \sigma_{21} & \ddots & & \vdots & \sigma_{2p+1} & \dots & \sigma_{2N} \\ \vdots & & \ddots & \sigma_{p-1p} & \vdots & & \vdots \\ \sigma_{p1} & \dots & \sigma_{pp-1} & 1 + \sigma_{pp} & \sigma_{pp+1} & \dots & \sigma_{pN} \\ \sigma_{p+11} & \dots & \sigma_{p+1p-1} & \sigma_{p+1p} & 1 + \sigma_{p+1p+1} & & \sigma_{p+1N} \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{N1} & \dots & \sigma_{Np-1} & \sigma_{Np} & \sigma_{Np+1} & & \sigma_{NN} \end{vmatrix}$$

On suppose que l'on regroupe les p premiers points de la matrice et donc que les indices $\{1, 2, \dots, p-1\}$ sont égaux à p . On obtient alors la matrice suivante :

$$M = \begin{vmatrix} 1 + \sigma_{pp} & \sigma_{pp} & \dots & \sigma_{pp} & \sigma_{pp+1} & \dots & \sigma_{pN} \\ \sigma_{pp} & \ddots & & \vdots & \sigma_{pp+1} & \dots & \sigma_{pN} \\ \vdots & & \ddots & \sigma_{pp} & \vdots & & \vdots \\ \sigma_{pp} & \dots & \sigma_{pp} & 1 + \sigma_{pp} & \sigma_{pp+1} & \dots & \sigma_{pN} \\ \sigma_{p+1p} & \dots & \sigma_{p+1p} & \sigma_{p+1p} & 1 + \sigma_{p+1p+1} & & \sigma_{p+1N} \\ \vdots & & & & \vdots & \ddots & \vdots \\ \sigma_{Np} & \dots & \sigma_{Np} & \sigma_{Np} & \sigma_{Np+1} & & \sigma_{NN} \end{vmatrix}$$

Calculons M et évaluons le poids au point p .

$$C_i \leftarrow C_i - C_p, \quad i < p$$

$$M = \begin{vmatrix} 1 & 0 & \dots & \sigma_{pp} & \sigma_{pp+1} & \dots & \sigma_{pN} \\ 0 & \ddots & & \vdots & \sigma_{pp+1} & \dots & \sigma_{pN} \\ \vdots & & \ddots & \sigma_{pp} & \vdots & & \vdots \\ -1 & \dots & -1 & 1 + \sigma_{pp} & \sigma_{pp+1} & \dots & \sigma_{pN} \\ 0 & \dots & 0 & \sigma_{p+1p} & 1 + \sigma_{p+1p+1} & \dots & \sigma_{p+1N} \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \sigma_{Np} & \sigma_{Np+1} & \dots & \sigma_{NN} \end{vmatrix}$$

On additionne les $p - 1$ premières lignes à la ligne L_p :

$$L_p \leftarrow L_p + L_i, \quad i < p$$

$$M = \begin{vmatrix} 1 & 0 & \dots & \sigma_{pp} & \sigma_{pp+1} & \dots & \sigma_{pN} \\ 0 & \ddots & & \vdots & \sigma_{pp+1} & \dots & \sigma_{pN} \\ \vdots & & \ddots & \sigma_{pp} & \vdots & & \vdots \\ 0 & \dots & 0 & 1 + p\sigma_{pp} & p\sigma_{pp+1} & \dots & p\sigma_{pN} \\ 0 & \dots & 0 & \sigma_{p+1p} & 1 + \sigma_{p+1p+1} & \dots & \sigma_{p+1N} \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \sigma_{Np} & \sigma_{Np+1} & \dots & \sigma_{NN} \end{vmatrix}$$

On multiplie par p les $p - 1$ premières lignes, et on leur retranche la ligne L_p :

$$C_i \leftarrow pC_i, \quad i < p$$

$$M = \left(\frac{1}{p}\right)^{p-1} \begin{vmatrix} p & 0 & \dots & p\sigma_{pp} & p\sigma_{pp+1} & \dots & p\sigma_{pN} \\ 0 & \ddots & & \vdots & p\sigma_{pp+1} & \dots & p\sigma_{pN} \\ \vdots & & \ddots & p\sigma_{pp} & \vdots & & \vdots \\ 0 & \dots & 0 & 1 + p\sigma_{pp} & p\sigma_{pp+1} & \dots & p\sigma_{pN} \\ 0 & \dots & 0 & \sigma_{p+1p} & 1 + \sigma_{p+1p+1} & \dots & \sigma_{p+1N} \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \sigma_{Np} & \sigma_{Np+1} & \dots & \sigma_{NN} \end{vmatrix}$$

$$L_i \leftarrow L_i - L_p, \quad i < p$$

$$M = \left(\frac{1}{p}\right)^{p-1} \begin{vmatrix} p & 0 & \dots & -1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots & 0 & \dots & 0 \\ \vdots & & \ddots & -1 & \vdots & & \vdots \\ 0 & \dots & 0 & 1 + p\sigma_{pp} & p\sigma_{pp+1} & \dots & p\sigma_{pN} \\ 0 & \dots & 0 & \sigma_{p+1p} & 1 + \sigma_{p+1p+1} & \dots & \sigma_{p+1N} \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \sigma_{Np} & \sigma_{Np+1} & \dots & \sigma_{NN} \end{vmatrix}$$

Soit encore, en factorisant :

$$M = \begin{vmatrix} 1 & 0 & \dots & -1 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots & 0 & \dots & 0 \\ \vdots & & \ddots & -1 & \vdots & & \vdots \\ 0 & \dots & 0 & 1 + p\sigma_{pp} & p\sigma_{pp+1} & \dots & p\sigma_{pN} \\ 0 & \dots & 0 & \sigma_{p+1p} & 1 + \sigma_{p+1p+1} & \dots & \sigma_{p+1N} \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \sigma_{Np} & \sigma_{Np+1} & \dots & \sigma_{NN} \end{vmatrix}$$

On termine en additionnant les $p - 1$ premières colonnes à la colonne p ,

$$M = \begin{vmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \ddots & & \vdots & 0 & \dots & 0 \\ \vdots & & \ddots & 0 & \vdots & & \vdots \\ 0 & \dots & 0 & 1 + p\sigma_{pp} & p\sigma_{pp+1} & \dots & p\sigma_{pN} \\ 0 & \dots & 0 & \sigma_{p+1p} & 1 + \sigma_{p+1p+1} & \dots & \sigma_{p+1N} \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \sigma_{Np} & \sigma_{Np+1} & \dots & \sigma_{NN} \end{vmatrix}$$

Et ainsi :

$$M = \begin{vmatrix} 1 + p\sigma_{pp} & p\sigma_{pp+1} & \dots & p\sigma_{pN} \\ \sigma_{p+1p} & 1 + \sigma_{p+1p+1} & \dots & \sigma_{p+1N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{Np} & \sigma_{Np+1} & \dots & \sigma_{NN} \end{vmatrix}$$

Annexe F

Linéarisation autour du maximum de vraisemblance

Plaçons nous dans un cadre bayésien où toutes les distributions sont gaussiennes. Considérons le problème de régression général suivant.

Soit $\mathbf{d}|\mathbf{m} \sim \mathcal{N}_N(\mathcal{F} \cdot \mathbf{m}, \mathbf{C}_d)$ une observation sur l'espace $\mathcal{D} \subseteq \mathbb{R}^N$. Considérons le paramètre inconnu \mathbf{m} de $\mathcal{M} \subseteq \mathbb{R}^M$ de distribution *a priori* $\pi(\mathbf{m})$ et posons la relation suivante :

$$\mathbf{d} = \mathcal{F} \cdot \mathbf{m} + \boldsymbol{\eta} \quad (\text{F.1})$$

où $\boldsymbol{\eta} \sim \mathcal{N}_N(0, \mathbf{C}_d)$ est un bruit gaussien.

En posant la distribution *a priori* gaussienne :

$$\pi(\mathbf{m}) = ((2\pi)^M |\mathbf{C}_0|)^{-1/2} \exp \frac{-1}{2} \|\mathbf{m} - \mathbf{m}_0\|_{\mathbf{C}_0^{-1}}^2 \quad (\text{F.2})$$

on peut déduire la distribution *a posteriori* :

$$\rho(\mathbf{m}|\mathbf{d}) = \frac{\rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m})}{\int_{\mathcal{M}} \rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m}) d\mathbf{m}} = \frac{\rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m})}{\sigma(\mathbf{d})} \quad (\text{F.3})$$

Dans un premier temps, nous explicitons la distribution $\sigma(\mathbf{d})$.

$$\sigma(\mathbf{d}) = \int_{\mathcal{M}} \rho(\mathbf{d}|\mathbf{m})\pi(\mathbf{m}) d\mathbf{m} \quad (\text{F.4})$$

$$\sigma(\mathbf{d}) = ((2\pi)^N |\mathbf{C}_d|)^{-1/2} ((2\pi)^M |\mathbf{C}_0|)^{-1/2} \int_{\mathcal{M}} \exp \frac{-1}{2} \left[\|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 + \|\mathbf{m} - \mathbf{m}_0\|_{\mathbf{C}_0^{-1}}^2 \right] d\mathbf{m} \quad (\text{F.5})$$

En développant, on obtient :

$$\|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 + \|\mathbf{m} - \mathbf{m}_0\|_{\mathbf{C}_0^{-1}}^2 = \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0 - \mathcal{F} \cdot (\mathbf{m} - \mathbf{m}_0)\|_{\mathbf{C}_d^{-1}}^2 + \|\mathbf{m} - \mathbf{m}_0\|_{\mathbf{C}_0^{-1}}^2 \quad (\text{F.6})$$

$$= \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0\|_{\mathbf{C}_d^{-1}}^2 - 2(\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0)^t \mathbf{C}_d^{-1} \mathcal{F} \cdot (\mathbf{m} - \mathbf{m}_0) \quad (\text{F.7})$$

$$+ \|\mathbf{m} - \mathbf{m}_0\|_{\mathbf{C}_0^{-1} + \mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}$$

Soit, en remplaçant dans F.4, on obtient :

$$\begin{aligned} \sigma(\mathbf{d}) &= ((2\pi)^N |\mathbf{C}_d|)^{-1/2} ((2\pi)^M |\mathbf{C}_0|)^{-1/2} \exp \frac{-1}{2} \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0\|_{\mathbf{C}_d^{-1}}^2 \times \\ &\quad \int_{\mathcal{M}} \exp \frac{-1}{2} \left[\|\mathbf{m} - \mathbf{m}_0\|_{\mathbf{C}_0^{-1} + \mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 + (\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0)^t \mathbf{C}_d^{-1} \mathcal{F} \cdot (\mathbf{m} - \mathbf{m}_0) \right] d\mathbf{m} \end{aligned} \quad (\text{F.8})$$

$$\begin{aligned} &= ((2\pi)^N |\mathbf{C}_d|)^{-1/2} ((2\pi)^M |\mathbf{C}_0|)^{-1/2} \exp \frac{-1}{2} \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0\|_{\mathbf{C}_d^{-1}}^2 \times \\ &\quad \int_{\mathcal{M}} \exp \frac{-1}{2} \|\mathbf{m} - \mathbf{m}_0\|_{\mathbf{H}}^2 + \mathbf{b}^t \cdot (\mathbf{m} - \mathbf{m}_0) d\mathbf{m} \end{aligned} \quad (\text{F.9})$$

où l'on a posé $\mathbf{H} = \mathbf{C}_0^{-1} + \mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}$ et $\mathbf{b}^t = (\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0)^t \mathbf{C}_d^{-1} \mathcal{F}$.

Alors, puisque \mathbf{H} est symétrique

$$\begin{aligned} \sigma(\mathbf{d}) &= ((2\pi)^N |\mathbf{C}_d|)^{-1/2} ((2\pi)^M |\mathbf{C}_0|)^{-1/2} ((2\pi)^M |\mathbf{H}|^{-1})^{+1/2} \times \\ &\quad \exp \frac{-1}{2} \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0\|_{\mathbf{C}_d^{-1}}^2 \exp \frac{+1}{2} \|\mathbf{b}^t \cdot \mathbf{H}^{-1} \cdot \mathbf{b}\| \end{aligned} \quad (\text{F.10})$$

$$\begin{aligned} &= (2\pi)^{-N/2} (|\mathbf{C}_d| \cdot |\mathbf{C}_0| \cdot |\mathbf{H}|)^{-1/2} \times \\ &\quad \exp \frac{-1}{2} \left[\|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0\|_{\mathbf{C}_d^{-1}}^2 - \|\mathbf{b}^t \cdot \mathbf{H}^{-1} \cdot \mathbf{b}\| \right] \end{aligned} \quad (\text{F.11})$$

Or,

$$\mathbf{b}^t \cdot \mathbf{H}^{-1} \cdot \mathbf{b} = (\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0)^t \mathbf{C}_d^{-1} \mathcal{F} \cdot (\mathbf{C}_0^{-1} + \mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F})^{-1} \cdot \mathcal{F}^t \mathbf{C}_d^{-1} (\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0) \quad (\text{F.12})$$

$$= \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0\|_{\mathbf{C}_d^{-1} \mathcal{F} \cdot (\mathbf{C}_0^{-1} + \mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F})^{-1} \cdot \mathcal{F}^t \mathbf{C}_d^{-1}}^2 \quad (\text{F.13})$$

Donc,

$$\begin{aligned} \sigma(\mathbf{d}) &= (2\pi)^{-N/2} (|\mathbf{C}_d| \cdot |\mathbf{C}_0| \cdot |\mathbf{H}|)^{-1/2} \times \\ &\quad \exp \frac{-1}{2} \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0\|_{\mathbf{C}_d^{-1} - \mathbf{C}_d^{-1} \mathcal{F} \cdot (\mathbf{C}_0^{-1} + \mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F})^{-1} \cdot \mathcal{F}^t \mathbf{C}_d^{-1}}^2 \end{aligned} \quad (\text{F.14})$$

Avec l'identité de Woodburry, on obtient directement pour la matrice de covariance :

$$\mathbf{C}_d^{-1} - \mathbf{C}_d^{-1} \mathcal{F} \cdot (\mathbf{C}_0^{-1} + \mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F})^{-1} \cdot \mathcal{F}^t \mathbf{C}_d^{-1} = (\mathbf{C}_d + \mathcal{F} \mathbf{C}_0 \mathcal{F}^t)^{-1} \quad (\text{F.15})$$

On utilise cette même identité pour le calcul de \mathbf{H}^{-1} . Soit :

$$\mathbf{H}^{-1} = (\mathbf{C}_0^{-1} + \mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F})^{-1} = \mathbf{C}_0 - \mathbf{C}_0 \mathcal{F}^t (\mathbf{C}_d + \mathcal{F} \mathbf{C}_0 \mathcal{F}^t)^{-1} \mathcal{F} \mathbf{C}_0 \quad (\text{F.16})$$

$$= \mathbf{C}_0 (\mathbb{I}_M - \mathcal{F}^t (\mathbf{C}_d + \mathcal{F} \mathbf{C}_0 \mathcal{F}^t)^{-1} \mathcal{F} \mathbf{C}_0) \quad (\text{F.17})$$

On obtient, en utilisant l'égalité de Schur :

$$|\mathbf{C}_0|^{-1} |\mathbf{H}|^{-1} = |\mathbb{I}_M - \mathcal{F}^t (\mathbf{C}_d + \mathcal{F} \mathbf{C}_0 \mathcal{F}^t)^{-1} \mathcal{F} \mathbf{C}_0| \quad (\text{F.18})$$

$$= |\mathbb{I}_N - \mathcal{F} \mathbf{C}_0 \mathcal{F}^t (\mathbf{C}_d + \mathcal{F} \mathbf{C}_0 \mathcal{F}^t)^{-1}| \quad (\text{F.19})$$

$$= |\mathbb{I}_N - \mathcal{F}^t (\mathbf{C}_0^{-1} + \mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F})^{-1} \mathcal{F} \mathbf{C}_d^{-1}| \quad (\text{F.20})$$

Finalement,

$$|\mathbf{C}_d|^{-1} |\mathbf{C}_0|^{-1} |\mathbf{H}|^{-1} = |\mathbf{C}_d^{-1} - \mathbf{C}_d^{-1} \mathcal{F}^t (\mathbf{C}_0^{-1} + \mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F})^{-1} \mathcal{F} \mathbf{C}_d^{-1}| \quad (\text{F.21})$$

$$= |\mathbf{C}_d + \mathcal{F} \mathbf{C}_0 \mathcal{F}^t|^{-1} \quad (\text{F.22})$$

La distribution σ est donc gaussienne, indépendante de \mathbf{m} ,

$$\sigma(\mathbf{d}) = ((2\pi)^N |\mathbf{C}_d + \mathcal{F} \mathbf{C}_0 \mathcal{F}^t|)^{-1/2} \exp \frac{-1}{2} \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0\|_{(\mathbf{C}_d + \mathcal{F} \mathbf{C}_0 \mathcal{F}^t)^{-1}}^2 \quad (\text{F.23})$$

Nous reprenons la distribution *a posteriori*. On remarque que :

$$\frac{-1}{2} \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 = \frac{-1}{2} \|\mathbf{d} - \mathcal{F} \cdot \tilde{\mathbf{m}}\|_{\mathbf{C}_d^{-1}}^2 + \frac{-1}{2} \|\mathbf{m} - \tilde{\mathbf{m}}\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 \quad (\text{F.24})$$

où $\tilde{\mathbf{m}}$ est obtenu au maximum de vraisemblance, c'est-à-dire,

$$\tilde{\mathbf{m}} = \arg \min_{\mathbf{m} \in \mathcal{M}} \|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}\|_{\mathbf{C}_d^{-1}}^2 = (\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F})^{-1} \mathcal{F}^t \mathbf{C}_d^{-1} \mathbf{d} \quad (\text{F.25})$$

Si on intègre ce résultat dans la distribution *a posteriori*, on obtient :

$$\frac{\rho(\mathbf{m}|\mathbf{d})}{\pi(\mathbf{m})} = ((2\pi)^N |\mathbf{C}_d|)^{-1/2} \exp \frac{-1}{2} \|\mathbf{d} - \mathcal{F} \cdot \tilde{\mathbf{m}}\|_{\mathbf{C}_d^{-1}}^2 \exp \frac{-1}{2} \|\mathbf{m} - \tilde{\mathbf{m}}\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 \times \frac{1}{\sigma(\mathbf{d})} \quad (\text{F.26})$$

Calculons à présent l'information contenue dans une mesure.

Nous avons :

$$\log \frac{\rho(\mathbf{m}|\mathbf{d})}{\pi(\mathbf{m})} = \log((2\pi)^N |\mathbf{C}_d|)^{-1/2} - \frac{1}{2} \|\mathbf{d} - \mathcal{F} \cdot \tilde{\mathbf{m}}\|_{\mathbf{C}_d^{-1}}^2 - \frac{1}{2} \|\mathbf{m} - \tilde{\mathbf{m}}\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 - \log \sigma(\mathbf{d}) \quad (\text{F.27})$$

Nous obtenons l'information contenue dans une mesure en calculant l'espérance sachant \mathbf{m} de l'expression précédente. Soit,

$$I(\pi(\mathbf{m})) = \mathbb{E} \left[\log \frac{\rho(\mathbf{m}|\mathbf{d})}{\pi(\mathbf{m})} \middle| \mathbf{m} \right] \quad (\text{F.28})$$

$$= \log((2\pi)^N |\mathbf{C}_d|)^{-1/2} - \frac{1}{2} \|\mathbf{d} - \mathcal{F} \cdot \tilde{\mathbf{m}}\|_{\mathbf{C}_d^{-1}}^2 - \log \sigma(\mathbf{d}) - \frac{1}{2} \mathbb{E} \left[\|\mathbf{m} - \tilde{\mathbf{m}}\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 \middle| \mathbf{m} \right] \quad (\text{F.29})$$

Le calcul de $\mathbb{E} \left[\|\mathbf{m} - \tilde{\mathbf{m}}\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 \mid \mathbf{m} \right]$ fait intervenir la moyenne et la covariance *a posteriori* que nous noterons respectivement \mathbf{m}_q et \mathbf{C}_q . Explicitons le calcul. Soit,

$$\mathbb{E} \left[\|\mathbf{m} - \tilde{\mathbf{m}}\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 \mid \mathbf{m}, \mathbf{d} \right] = \mathbb{E} \left[\text{tr} \left(\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} \cdot \|\mathbf{m} - \tilde{\mathbf{m}}\|^2 \right) \mid \mathbf{m}, \mathbf{d} \right] \quad (\text{F.30})$$

$$= \text{tr} \left(\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} \cdot \mathbb{E} \left[\|\mathbf{m} - \tilde{\mathbf{m}}\|^2 \mid \mathbf{m}, \mathbf{d} \right] \right) \quad (\text{F.31})$$

$$= \text{tr} \left(\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} \cdot \mathbb{E} \left[\|\mathbf{m} - \mathbf{m}_q - (\tilde{\mathbf{m}} - \mathbf{m}_q)\|^2 \mid \mathbf{m}, \mathbf{d} \right] \right) \quad (\text{F.32})$$

$$= \text{tr} \left(\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} \cdot \left(\mathbb{E} \left[\|\mathbf{m} - \mathbf{m}_q\|^2 \mid \mathbf{m}, \mathbf{d} \right] + \|\tilde{\mathbf{m}} - \mathbf{m}_q\|^2 \right) \right) \quad (\text{F.33})$$

Or, $\mathbb{E} \left[\|\mathbf{m} - \mathbf{m}_q\|^2 \mid \mathbf{m}, \mathbf{d} \right] \equiv \mathbf{C}_q$ par définition. Donc,

$$\mathbb{E} \left[\|\mathbf{m} - \tilde{\mathbf{m}}\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 \mid \mathbf{m}, \mathbf{d} \right] = \text{tr} \left(\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} \cdot \left(\mathbf{C}_q + \|\tilde{\mathbf{m}} - \mathbf{m}_q\|^2 \right) \right) \quad (\text{F.34})$$

$$= \text{tr} \left[\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} \cdot \mathbf{C}_q \right] + \|\tilde{\mathbf{m}} - \mathbf{m}_q\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 \quad (\text{F.35})$$

On obtient la valeur de l'information contenue dans une mesure gaussienne,

$$\begin{aligned} I(\pi(\mathbf{m})) &= \log((2\pi)^N |\mathbf{C}_d|)^{-1/2} - \frac{1}{2} \|\mathbf{d} - \mathcal{F} \cdot \tilde{\mathbf{m}}\|_{\mathbf{C}_d^{-1}}^2 - \log \sigma(\mathbf{d}) \\ &\quad - \frac{1}{2} \text{tr} \left[\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} \cdot \mathbf{C}_q \right] - \frac{1}{2} \|\tilde{\mathbf{m}} - \mathbf{m}_q\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 \end{aligned} \quad (\text{F.36})$$

De même, à l'aide du calcul explicite de la distribution $\sigma(\mathbf{d})$, on peut déduire la valeur moyenne du contenu d'information en calculant l'espérance sachant \mathbf{d} de $I(\pi)$.

Ainsi, l'information moyenne est :

$$\langle I \rangle = \mathbb{E} \left[I(\pi) \mid \mathbf{d} \right] \quad (\text{F.37})$$

$$\begin{aligned} &= \log((2\pi)^N |\mathbf{C}_d|)^{-1/2} - \frac{1}{2} \text{tr} \left[\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} \cdot \mathbf{C}_q \right] - \frac{1}{2} \|\tilde{\mathbf{m}} - \mathbf{m}_q\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 \\ &\quad - \frac{1}{2} \mathbb{E} \left[\|\mathbf{d} - \mathcal{F} \cdot \tilde{\mathbf{m}}\|_{\mathbf{C}_d^{-1}}^2 \mid \mathbf{d} \right] + \mathbb{E} \left[-\log \sigma(\mathbf{d}) \mid \mathbf{d} \right] \end{aligned} \quad (\text{F.38})$$

On reconnaît la définition de l'entropie de Shannon telle que :

$$\mathbb{E} \left[-\log \sigma(\mathbf{d}) \mid \mathbf{d} \right] \equiv \mathcal{E}^{\mathbf{d}}\{\sigma\} = \frac{1}{2} \log(2\pi)^N |\mathbf{C}_d + \mathcal{F} \mathbf{C}_0 \mathcal{F}^t| + \frac{N}{2} \quad (\text{F.39})$$

On explicite l'autre terme d'espérance en introduisant sa moyenne $\mathcal{F} \cdot \mathbf{m}_0$ calculé plus haut.

$$\frac{1}{2} \mathbb{E} \left[\|\mathbf{d} - \mathcal{F} \cdot \tilde{\mathbf{m}}\|_{\mathbf{C}_d^{-1}}^2 \mid \mathbf{d} \right] = \frac{1}{2} \text{tr} \left(\mathbf{C}_d^{-1} \cdot \mathbb{E} \left[\|\mathbf{d} - \mathcal{F} \cdot \tilde{\mathbf{m}}\|^2 \mid \mathbf{d} \right] \right) \quad (\text{F.40})$$

$$= \frac{1}{2} \text{tr} \left(\mathbf{C}_d^{-1} \cdot \mathbb{E} \left[\|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0 - \mathcal{F} \cdot (\tilde{\mathbf{m}} - \mathbf{m}_0)\|^2 \mid \mathbf{d} \right] \right) \quad (\text{F.41})$$

$$= \frac{1}{2} \text{tr} \left(\mathbf{C}_d^{-1} \cdot \left(\mathbb{E} \left[\|\mathbf{d} - \mathcal{F} \cdot \mathbf{m}_0\|^2 \mid \mathbf{d} \right] + \|\mathcal{F} \cdot (\tilde{\mathbf{m}} - \mathbf{m}_0)\|^2 \right) \right) \quad (\text{F.42})$$

$$= \frac{1}{2} \text{tr} \left(\mathbf{C}_d^{-1} (\mathbf{C}_d + \mathcal{F} \mathbf{C}_0 \mathcal{F}^t) \right) + \text{tr} \left(\mathbf{C}_d^{-1} \cdot \|\mathcal{F} \cdot (\tilde{\mathbf{m}} - \mathbf{m}_0)\|^2 \right) \quad (\text{F.43})$$

$$= \frac{1}{2} \text{tr} \left(\mathbb{I}_N + \mathbf{C}_d^{-1} \mathcal{F} \mathbf{C}_0 \mathcal{F}^t \right) + \frac{1}{2} \|\tilde{\mathbf{m}} - \mathbf{m}_0\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 \quad (\text{F.44})$$

On conclut que l'information moyenne apportée par une mesure \mathbf{d} est

$$\begin{aligned} \langle I \rangle &= \frac{1}{2} \text{tr} \left(\mathbb{I}_N + \mathbf{C}_d^{-1} \mathcal{F} \mathbf{C}_0 \mathcal{F}^t \right) + \frac{1}{2} \|\tilde{\mathbf{m}} - \mathbf{m}_0\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 \\ &\quad + \log((2\pi)^N |\mathbf{C}_d|)^{-1/2} - \frac{1}{2} \text{tr} \left[\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} \cdot \mathbf{C}_q \right] - \frac{1}{2} \|\tilde{\mathbf{m}} - \mathbf{m}_q\|_{\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F}}^2 \\ &\quad + \frac{1}{2} \log(2\pi)^N |\mathbf{C}_d + \mathcal{F} \mathbf{C}_0 \mathcal{F}^t| + \frac{N}{2} \end{aligned} \quad (\text{F.45})$$

$$\begin{aligned} &= \frac{1}{2} \log |\mathbb{I}_N + \mathbf{C}_d^{-1} \mathcal{F} \mathbf{C}_0 \mathcal{F}^t| \\ &\quad + \frac{1}{2} \text{tr} \left(\mathbb{I}_N + \mathbf{C}_d^{-1} \mathcal{F} \mathbf{C}_0 \mathcal{F}^t \right) - \frac{1}{2} \text{tr} \left[\mathcal{F}^t \mathbf{C}_d^{-1} \mathcal{F} \cdot \mathbf{C}_q \right] \end{aligned} \quad (\text{F.46})$$

Annexe G

Fonction de covariance de la perméabilité apparente

A l'ordre $n = 1$, la perméabilité apparente est définie par :

$$k_{app}(t) = \frac{1}{2\pi D_0 t} \int_{\mathbb{R}^d} k(\mathbf{r}) G(\mathbf{r}, t) d\mathbf{r} \quad (\text{G.1})$$

On suppose que G est à l'ordre 1 constant par l'opérateur de la moyenne spatiale. Par conséquent, la moyenne spatiale de la perméabilité apparente s'exprime directement en fonction de la moyenne spatiale de la perméabilité pondérée par le noyau d'Oliver G , soit la relation suivante :

$$\langle k * G(\mathbf{r}, t) \rangle = \langle k \rangle * G(\mathbf{r}, t) \quad (\text{G.2})$$

L'expression de la fonction de covariance de la perméabilité entre deux temps d'observation t et τ est alors exprimée simplement par la fonction suivante :

$$\begin{aligned} \langle k_{app}(t), k_{app}(\tau) \rangle &= G(\mathbf{r}, t) \langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle G(\boldsymbol{\xi}, \tau) \\ &= \left(\frac{1}{2\pi D_0 \sqrt{t \cdot \tau}} \right)^2 \iint_{\mathbb{R}^2} \langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle G(\mathbf{r}, t) G(\boldsymbol{\xi}, \tau) d\mathbf{r} d\boldsymbol{\xi} \end{aligned} \quad (\text{G.3})$$

Dans la mesure où nous choisissons une réalisation lognormale de la perméabilité, un choix particulier de sa définition peut-être choisi. En effet, dans le cas d'une réalisation lognormale de la perméabilité, Sagar a montré numériquement qu'en effectuant une moyenne géométrique des perméabilités sur une couronne, son estimation était . Matheron a montré dans le cas cartésien que la perméabilité est exprimée en moyennant par :

$$k(\mathbf{r}) = k_G \cdot \exp \left\{ \frac{1}{|C_{\mathbf{r}}|} \int_{C_{\mathbf{r}}} \log k(\mathbf{s}) d\mathbf{s} \right\} \quad (\text{G.4})$$

Par simplicité d'écriture, on pose $k_G = 1$ et on définit la variable aléatoire $Z_{\mathbf{r}}$ par

$$Z_{\mathbf{r}} = \frac{1}{|C_{\mathbf{r}}|} \int_{C_{\mathbf{r}}} \log k(\mathbf{s}) d\mathbf{s} \quad (\text{G.5})$$

Par construction, la variable aléatoire $Z_{\mathbf{r}}$ est normale supposée centrée. Par conséquent, $\exp Z_{\mathbf{r}}$ est une variable aléatoire log-normale et on obtient alors :

$$\langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle = \langle \exp Z_{\mathbf{r}}, \exp Z_{\boldsymbol{\xi}} \rangle \quad (\text{G.6})$$

$$= \langle \exp Z_{\mathbf{r}} \cdot \exp Z_{\boldsymbol{\xi}} \rangle - \langle \exp Z_{\mathbf{r}} \rangle \langle \exp Z_{\boldsymbol{\xi}} \rangle \quad (\text{G.7})$$

$$= \langle \exp(Z_{\mathbf{r}} + Z_{\boldsymbol{\xi}}) \rangle - \langle \exp Z_{\mathbf{r}} \rangle \langle \exp Z_{\boldsymbol{\xi}} \rangle \quad (\text{G.8})$$

$$= \exp \frac{1}{2} \langle (Z_{\mathbf{r}} + Z_{\boldsymbol{\xi}})^2 \rangle - \exp \frac{1}{2} \langle Z_{\mathbf{r}}^2 \rangle \cdot \exp \langle Z_{\boldsymbol{\xi}}^2 \rangle \quad (\text{G.9})$$

Soit encore :

$$\langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle = \exp \frac{1}{2} (\langle Z_{\mathbf{r}}^2 \rangle + \langle Z_{\boldsymbol{\xi}}^2 \rangle) \left(\exp \langle Z_{\mathbf{r}} Z_{\boldsymbol{\xi}} \rangle - 1 \right) \quad (\text{G.10})$$

On définit la variable aléatoire Y par :

$$Y(\mathbf{s}) = \log \frac{k(\mathbf{s})}{k_G} \quad (\text{G.11})$$

On a d'après la définition de $Z_{\mathbf{r}}$:

$$Z_{\mathbf{r}} = \frac{1}{|C_{\mathbf{r}}|} \int_{C_{\mathbf{r}}} \log k(\mathbf{s}) \, d\mathbf{s} = \frac{1}{|C_{\mathbf{r}}|} \int_{C_{\mathbf{r}}} Y(\mathbf{s}) \, d\mathbf{s} \quad (\text{G.12})$$

Ainsi,

$$\langle Z_{\mathbf{r}}^2 \rangle = \frac{1}{|C_{\mathbf{r}}|^2} \iint_{C_{\mathbf{r}}} \langle Y_{\mathbf{r}}(\mathbf{s}) Y_{\mathbf{r}}(\mathbf{s}') \rangle \, d\mathbf{s} \, d\mathbf{s}' \quad (\text{G.13})$$

Par stationnarité, cette dernière égalité s'exprime seulement par une intégrale simple à l'aide de la fonction de covariance C telle que :

$$\langle Z_{\mathbf{r}}^2 \rangle = \frac{1}{|C_{\mathbf{r}}|^2} \iint_{C_{\mathbf{r}}} C(|\mathbf{r}_{\mathbf{s}} - \mathbf{r}_{\mathbf{s}'}|) \, d\mathbf{s} \, d\mathbf{s}' \quad (\text{G.14})$$

$$= \frac{1}{|C_{\mathbf{r}}|} \int_{C_{\mathbf{r}}} C(|\mathbf{r}_{\mathbf{s}} - \mathbf{r}_0|) \, d\mathbf{s} \quad (\text{G.15})$$

En choisissant la fonction de covariance gaussienne, on peut expliciter la dernière égalité comme suit,

$$\langle Z_{\mathbf{r}}^2 \rangle = \frac{1}{|C_{\mathbf{r}}|} \int_{C_{\mathbf{r}}} \sigma_0^2 \cdot \exp \frac{-1}{2} \left(\frac{|\mathbf{r}_{\mathbf{s}} - \mathbf{r}_0|}{\ell_c} \right)^2 \, d\mathbf{s} \quad (\text{G.16})$$

En paramétrant la courbe $C_{\mathbf{r}}$ par $(r_0 \cos(\theta), r_0 \sin(\theta))$, avec $\theta \in [0, 2\pi]$, on obtient par changement de variable :

$$\begin{aligned} \langle Z_{\mathbf{r}}^2 \rangle &= \frac{1}{2\pi r_0} \int_0^{2\pi} \sigma_0^2 \cdot \exp \frac{-1}{2} \left(\frac{((r_0 \cos \theta - r_0)^2 + r_0^2 \sin^2 \theta)}{\ell_c^2} \right) r_0 \, d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} \sigma_0^2 \cdot \exp \frac{-1}{2} \left(\frac{2r_0^2(1 - \cos \theta)}{\ell_c^2} \right) \, d\theta \\ &= \sigma_0^2 \cdot \exp \left(\frac{-r_0^2}{\ell_c^2} \right) \frac{1}{2\pi} \int_0^{2\pi} \exp \left(\frac{r_0^2}{\ell_c^2} \cdot \cos \theta \right) \, d\theta \\ &= \sigma_0^2 \exp \left(\frac{-r_0^2}{\ell_c^2} \right) \mathcal{I}_0 \left(\frac{r_0^2}{\ell_c^2} \right) \end{aligned}$$

Et donc,

$$\langle Z_{\mathbf{r}}^2 \rangle = \sigma_0^2 \exp\left(\frac{-r_0^2}{\ell_c^2}\right) \mathcal{I}_0\left(\frac{r_0^2}{\ell_c^2}\right) \quad (\text{G.17})$$

De la même manière, en paramétrant la courbe C_{ξ} par $(\xi_0 \cos \theta, \xi_0 \sin \theta)$, avec $\theta \in [0, 2\pi]$, on obtient :

$$\langle Z_{\xi}^2 \rangle = \sigma_0^2 \exp\left(\frac{-\xi_0^2}{\ell_c^2}\right) \mathcal{I}_0\left(\frac{\xi_0^2}{\ell_c^2}\right) \quad (\text{G.18})$$

Le calcul de $\langle Z_{\mathbf{r}} Z_{\xi} \rangle$ est plus délicat. En reprenant (G.12), on obtient sur deux couronnes $C_{\mathbf{r}}$ et C_{ξ} , la covariance croisée :

$$\langle Z_{\mathbf{r}} Z_{\xi} \rangle = \frac{1}{|C_{\mathbf{r}}| |C_{\xi}|} \int_{C_{\mathbf{r}}} \int_{C_{\xi}} \mathbf{C}(|\mathbf{r}_s - \xi_u|) ds du \quad (\text{G.19})$$

En reprenant le paramétrage des deux courbes comme précédemment, on obtient l'égalité suivante :

$$\begin{aligned} \langle Z_{\mathbf{r}} Z_{\xi} \rangle &= \frac{1}{2\pi r_0} \frac{1}{2\pi \xi_0} \iint_0^{2\pi} \exp\left(\frac{-1}{2} \frac{|\mathbf{r}_s - \xi_u|^2}{\ell_c^2}\right) d\theta d\varphi \\ &= \frac{1}{(2\pi)^2} \frac{1}{r_0 \xi_0} \iint_0^{2\pi} \sigma_0^2 \exp\left(\frac{-1}{2} \frac{(r \cos \theta - \xi \cos \varphi)^2 + (r \sin \theta - \xi \sin \varphi)^2}{\ell_c^2}\right) r_0 d\theta \xi_0 d\varphi \\ &= \frac{\sigma_0^2}{(2\pi)^2} \iint_0^{2\pi} \exp\left(\frac{-1}{2} \frac{r^2 + \xi^2 - 2r\xi \cos(\theta - \varphi)}{\ell_c^2}\right) d\theta d\varphi \end{aligned}$$

En posant $t = \theta - \varphi$, avec $t \in [-\varphi, 2\pi - \varphi]$, on obtient :

$$\langle Z_{\mathbf{r}} Z_{\xi} \rangle = \frac{\sigma_0^2}{(2\pi)^2} \int_0^{2\pi} \exp\left(\frac{-1}{2} \frac{r^2 + \xi^2}{\ell_c^2}\right) \int_{-\varphi}^{2\pi - \varphi} \exp\left(\frac{r\xi}{\ell_c^2} \cos t\right) dt d\varphi$$

Du fait de la périodicité de $t \rightarrow \exp\left(\frac{r\xi}{\ell_c^2} \cos t\right)$, on obtient

$$\begin{aligned} \langle Z_{\mathbf{r}} Z_{\xi} \rangle &= \frac{\sigma_0^2}{(2\pi)^2} \int_0^{2\pi} \exp\left(\frac{-1}{2} \frac{r^2 + \xi^2}{\ell_c^2}\right) \int_0^{2\pi} \exp\left(\frac{r\xi}{\ell_c^2} \cos t\right) dt d\varphi \\ &= \frac{\sigma_0^2}{(2\pi)^2} \exp\left(\frac{-1}{2} \frac{r^2 + \xi^2}{\ell_c^2}\right) \int_0^{2\pi} d\varphi \int_0^{2\pi} \exp\left(\frac{r\xi}{\ell_c^2} \cos t\right) dt \\ &= \frac{\sigma_0^2}{2\pi} \exp\left(\frac{-1}{2} \frac{r^2 + \xi^2}{\ell_c^2}\right) \int_0^{2\pi} \exp\left(\frac{r\xi}{\ell_c^2} \cos \theta\right) d\theta \\ &= \sigma_0^2 \exp\left(\frac{-1}{2} \frac{r^2 + \xi^2}{\ell_c^2}\right) \mathcal{I}_0\left(\frac{r\xi}{\ell_c^2}\right) \quad (\text{G.20}) \end{aligned}$$

Finalement,

$$\begin{aligned} \langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle &= k_G^2 \cdot \exp \left(\frac{1}{2} \frac{\sigma_0^2}{\exp \frac{r^2}{\ell_c^2}} \mathcal{I}_0 \left(\frac{r^2}{\ell_c^2} \right) + \frac{1}{2} \frac{\sigma_0^2}{\exp \frac{\xi^2}{\ell_c^2}} \mathcal{I}_0 \left(\frac{\xi^2}{\ell_c^2} \right) \right) \cdot \\ &\quad \left(\exp \left(\frac{\sigma_0^2}{\exp \frac{r^2 + \xi^2}{2\ell_c^2}} \mathcal{I}_0 \left(\frac{r\xi}{\ell_c^2} \right) \right) - 1 \right) \end{aligned}$$

On retrouve la propriété de la covariance telle que :

$$\langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle = \langle k(\boldsymbol{\xi}), k(\mathbf{r}) \rangle \quad (\text{G.21})$$

On peut calculer les cas limites suivants. Sans difficultés, on a :

$$\begin{aligned} \langle k(\mathbf{r}), k(\mathbf{r}) \rangle &= \text{Var}(k(\mathbf{r})) = \lim_{\boldsymbol{\xi} \rightarrow \mathbf{r}} \langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle \\ &= k_G^2 \cdot \exp \left(\frac{\sigma_0^2}{\exp \frac{r^2}{\ell_c^2}} \mathcal{I}_0 \left(\frac{r^2}{\ell_c^2} \right) \right) \left(\exp \left(\frac{\sigma_0^2}{\exp \frac{r^2}{\ell_c^2}} \mathcal{I}_0 \left(\frac{r^2}{\ell_c^2} \right) \right) - 1 \right) \end{aligned} \quad (\text{G.22})$$

De même, on a :

$$\begin{aligned} \langle k(r_0), k(\infty) \rangle &= \lim_{\mathbf{r} \rightarrow r_0} \lim_{\boldsymbol{\xi} \rightarrow \infty} \langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle \\ &= \lim_{\boldsymbol{\xi} \rightarrow \infty} \lim_{\mathbf{r} \rightarrow r_0} \langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle \end{aligned} \quad (\text{G.23})$$

$$= \lim_{\boldsymbol{\xi} \rightarrow \infty} k_G^2 \cdot \exp \left(\frac{1}{2} \frac{\sigma_0^2}{\exp \frac{r_0^2}{\ell_c^2}} \mathcal{I}_0 \left(\frac{r_0^2}{\ell_c^2} \right) + \frac{1}{2} \frac{\sigma_0^2}{\exp \frac{\xi^2}{\ell_c^2}} \mathcal{I}_0 \left(\frac{\xi^2}{\ell_c^2} \right) \right) \quad (\text{G.24})$$

$$\begin{aligned} &\quad \left(\exp \left(\frac{\sigma_0^2}{\exp \frac{r_0^2 + \xi^2}{2\ell_c^2}} \mathcal{I}_0 \left(\frac{r_0}{\ell_c^2} \xi \right) \right) - 1 \right) \\ &= 0 \end{aligned} \quad (\text{G.25})$$

Reprenons le calcul de la corrélation entre les perméabilités aux temps t et τ .

$$\langle k_{app}(t), k_{app}(\tau) \rangle = \frac{1}{(2\pi D_0 \sqrt{t\tau})^2} \iint_{\mathbb{R}^2} \langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle G(\mathbf{r}, t) G(\boldsymbol{\xi}, \tau) d\mathbf{r} d\boldsymbol{\xi} \quad (\text{G.26})$$

où $G(\mathbf{r}, t) = G(z) = ze^{-z} (2\mathcal{K}_1(z)(1 - 1/z) + \mathcal{K}_0(z) + \mathcal{K}_2(z))$ avec $z = \frac{\mathbf{r}^2}{2D_0 t}$.

On réécrit cette dernière intégrale en scindant les deux termes. Et on obtient,

$$\langle k_{app}(t), k_{app}(\tau) \rangle = \int_0^{+\infty} \frac{G(\boldsymbol{\xi}, \tau)}{2\pi D_0 \tau} \int_0^{+\infty} \langle k(\mathbf{r}), k(\boldsymbol{\xi}) \rangle \frac{G(\mathbf{r}, t)}{2\pi D_0 t} d\mathbf{r} d\boldsymbol{\xi} \quad (\text{G.27})$$

Dans un premier temps, on peut approximer le noyau de pondération d'Oliver G par un dirac en $\sqrt{D_0 t}$ de telle manière que

$$G(\mathbf{r}, t) \approx \delta(r^2 - \sqrt{D_0 t}) \quad (\text{G.28})$$

Ainsi, une approximation de la corrélation entre les perméabilités apparentes est :

$$\begin{aligned} \langle k_{app}(t), k_{app}(\tau) \rangle &\approx \langle k(\sqrt{D_0 t}), k(\sqrt{D_0 \tau}) \rangle & (\text{G.29}) \\ &= \exp \left(\frac{1}{2} \frac{\sigma_0^2}{\exp \left(t \frac{D_0}{\ell_c^2} \right)} \mathcal{I}_0 \left(t \frac{D_0}{\ell_c^2} \right) + \frac{1}{2} \frac{\sigma_0^2}{\exp \left(\tau \frac{D_0}{\ell_c^2} \right)} \mathcal{I}_0 \left(\tau \frac{D_0}{\ell_c^2} \right) \right) \times \\ &\quad \left(\exp \left(\frac{\sigma_0^2}{\exp \left((t + \tau) \frac{D_0}{\ell_c^2} \right)} \mathcal{I}_0 \left((t + \tau) \frac{D_0}{\ell_c^2} \right) \right) - 1 \right) & (\text{G.30}) \end{aligned}$$

Annexe H

Simulateur d'essai de puits monophasique 3D

On fait l'hypothèse que la variation de pression suit l'équation de diffusivité classique [Daviau 1986, Romeu 1994] :

$$\phi\mu c_T \partial_t P(\mathbf{r}, t) = \nabla(k(\mathbf{r})\nabla P(\mathbf{r}, t)) + \frac{Q}{h}\chi_{\mathbf{r}-\mathbf{r}_w}$$

Le réservoir est discrétisé sur un maillage cartésien régulier. Les propriétés pétro-physiques sont supposées constantes sur chaque maille. La résolution de l'équation de diffusivité se fait par volumes finis.

On obtient pour chaque maille (ijk) :

$$\phi c_T \mathcal{V}_{ijk} \frac{P_{ijk}^n - P_{ijk}^{n-1}}{\Delta t} = \frac{1}{\mu} \sum_{\vartheta(ijk)} T_{(ijk) \rightarrow \vartheta(ijk)} \left(P_{\vartheta(ijk)}^n - P_{ijk}^n \right) + \frac{Q}{h\mu} \chi_{(ijk), i_w} \quad (\text{H.1})$$

où \mathcal{V}_{ijk} est le volume de la maille (ijk) , c_T la compressibilité totale du fluide et $P_{ijk}^n = P(\mathbf{x}_{ijk}, t_n) = P(i\delta x, j\delta y, k\delta z, n\Delta t)$. $\vartheta(ijk)$ représente les indices du voisinage de la maille (ijk) . On définit les transmissivités inter-blocs par $T_{(ijk) \rightarrow \vartheta(i)}$, où

$$T_{(ijk) \rightarrow (i+1, jk)} = 2(k_{ijk}^{-1} + k_{i+1, jk}^{-1})^{-1}$$

avec k_{ijk} la perméabilité en la maille (ijk) .

Explicitons le système linéaire permettant la résolution de (H.1).

$$\begin{aligned}
& \phi c_T \mathcal{V}_{ijk} P_{ijk}^n - \frac{\Delta t}{\mu} \sum_{\vartheta(ijk)} T_{(ijk) \rightarrow \vartheta(ijk)} \left(P_{\vartheta(ijk)}^n - P_{ijk}^n \right) = c_T \mathcal{V}_{ijk} \phi P_{ijk}^{n-1} + \frac{\Delta t}{h\mu} Q \chi_{(ijk), i_w} \\
& \phi c_T \mathcal{V}_{ijk} P_{ijk}^n + \frac{\Delta t}{\mu} \sum_{\vartheta(ijk)} T_{(ijk) \rightarrow \vartheta(ijk)} P_{ijk}^n - \frac{\Delta t}{\mu} \sum_{\vartheta(ijk)} T_{(ijk) \rightarrow \vartheta(i)} P_{\vartheta(ijk)}^n = c_T \mathcal{V}_{ijk} \phi P_{ijk}^{n-1} + \frac{\Delta t}{h\mu} Q \chi_{(ijk), i_w} \\
& \left(\phi c_T \mathcal{V}_{ijk} + \frac{\Delta t}{\mu} \sum_{\vartheta(ijk)} T_{(ijk) \rightarrow \vartheta(ijk)} \right) P_{ijk}^n - \frac{\Delta t}{\mu} \left(T_{i,i-1} P_{i-1,jk}^n + T_{i,i+1} P_{i+1,jk}^n + \right. \\
& \left. + T_{j,j-1} P_{i,j-1,k}^n + T_{j,j+1} P_{i,j+1,k}^n + T_{k,k-1} P_{ij,k-1}^n + T_{k,k+1} P_{ij,k+1}^n \right) = c_T \mathcal{V}_{ijk} \phi P_{ijk}^{n-1} + \frac{\Delta t}{h\mu} Q \chi_{(ijk), i_w}
\end{aligned}$$

où pour simplifier l'écriture, $T_{i,i+1} = T_{(i\bullet\bullet) \rightarrow i+1, \bullet\bullet}$.

On se place dans le repère cartésien, où $q = i + jN + kNM$ représente la position en (ijk) .

$$\begin{aligned}
& \left(\phi c_T \mathcal{V}_{ijk} + \frac{\Delta t}{\mu} \sum_{\vartheta(ijk)} T_{(ijk) \rightarrow \vartheta(ijk)} \right) P_{ijk}^n - \frac{\Delta t}{\mu} \left(T_{i,i-1} P_{i-1,jk}^n + T_{i,i+1} P_{i+1,jk}^n + \right. \\
& \left. + T_{j,j-1} P_{i,j-1,k}^n + T_{j,j+1} P_{i,j+1,k}^n + T_{k,k-1} P_{ij,k-1}^n + T_{k,k+1} P_{ij,k+1}^n \right) = c_T \mathcal{V}_{ijk} \phi P_{ijk}^{n-1} + \frac{\Delta t}{h} Q \chi_{(ijk), i_w} \\
& \left(\phi c_T \mathcal{V}_q + \frac{\Delta t}{\mu} \sum_{\vartheta(q)} T_{q \rightarrow \vartheta(q)} \right) P_q^n - \frac{\Delta t}{\mu} \langle T_{loc}, P_{loc}^n \rangle = c_T \mathcal{V}_q \phi P_q^{n-1} + \frac{\Delta t}{h} Q \chi_{q, q_w}
\end{aligned}$$

où $T_{loc} = [T_{q,q\pm 1} \ T_{q,q\pm N} \ T_{q,q\pm N^*M}]$ et $P_{loc}^n = [P_{q\pm 1}^n \ P_{q\pm N}^n \ P_{q\pm N^*M}^n]$

On suppose connue la matrice \mathbf{k} des perméabilités en tout point de l'espace Ω . On définit le produit matrice-vecteur. $\forall q \in \Omega$:

$$y_{[q]} + = \left(\phi c_T \mathcal{V}_q + \frac{\Delta t}{\mu} \alpha \right) x_{[q]} - \frac{\Delta t}{\mu} \langle T_{loc}, x_{loc} \rangle$$

où $\alpha = \sum T_{loc}$ et $x_{loc} = [x_{[q\pm 1]} \ x_{[q\pm N]} \ x_{[q\pm N^*M]}]$

le second membre est défini par :

$$b_{[q]} = \phi c_T \mathcal{V}_{[q]} P_{[q]}^0 + \frac{\Delta t}{h} Q \chi_{[q-q_w]}$$

où P_q^0 est la pression initiale dans le réservoir.

Annexe I

Fonction de covariance du front de saturation

On rappelle ici l'expression de la perturbation δx du front dans l'espace de Fourier-Laplace telle que :

$$\delta x(\mathbf{q}, w) = \frac{-q^2}{(Aq - q_1)((q_1)^2 + q^2)} \left(1 + i \frac{Aq_1}{q}\right) \frac{K(q_1, \mathbf{q})}{c} \quad (\text{I.1})$$

On calcule explicitement la fonction de corrélation entre deux fronts de saturation à deux temps différents. On pose :

$$C(\vec{\mathbf{y}}, t; \vec{\mathbf{x}}, \tau) = \int \frac{dq_1}{2\pi} \frac{d\vec{\mathbf{q}}}{(2\pi)^{d-1}} |\mathcal{M}(q_1, \vec{\mathbf{q}})|^2 K(q_1, \vec{\mathbf{q}}) e^{i\vec{\mathbf{q}} \cdot (\vec{\mathbf{y}} - \vec{\mathbf{x}})} e^{iq_1 c_0(t - \tau)} \quad (\text{I.2})$$

où $K(q_1, \vec{\mathbf{q}})$ est exprimée par sa mesure spectrale telle que :

$$K(q_1, \vec{\mathbf{q}}) = \int_0^{+\infty} dm \hat{\rho}(m) e^{-m(q_1^2 + \vec{\mathbf{q}} \cdot \vec{\mathbf{q}})} = \int_0^{+\infty} dm \hat{\rho}(m) e^{-m(q_1^2 + q^2)} \quad (\text{I.3})$$

et

$$|\mathcal{M}(q_1, \vec{\mathbf{q}})|^2 = \frac{1 + A^2}{1 - A^2} \left(\frac{1}{q_1^2 + A^2 q^2} - \frac{1}{q_1^2 + q^2} \right) - \frac{q^2}{(q_1^2 + q^2)^2} \quad (\text{I.4})$$

La fonction de covariance dépend de la variation entre les fronts de saturation $\vec{\mathbf{y}}$ et $\vec{\mathbf{x}}$, respectivement situés aux temps t et τ . On pose $\vec{\mathbf{u}} = \vec{\mathbf{y}} - \vec{\mathbf{x}}$ et $\Delta t = t - \tau$ et on obtient

$$\begin{aligned} C(\vec{\mathbf{u}}; \Delta t) &= \int_0^{+\infty} dm \hat{\rho}(m) \int \frac{dq_1}{2\pi} \frac{d\vec{\mathbf{q}}}{(2\pi)^{d-1}} \frac{1 + A^2}{1 - A^2} \left(\frac{1}{q_1^2 + A^2 q^2} - \frac{1}{q_1^2 + q^2} \right) e^{i\vec{\mathbf{q}} \cdot \vec{\mathbf{u}}} e^{iq_1 c \Delta t} e^{-m(q_1^2 + q^2)} \\ &\quad - \int_0^{+\infty} dm \hat{\rho}(m) \int \frac{dq_1}{2\pi} \frac{d\vec{\mathbf{q}}}{(2\pi)^{d-1}} \frac{q^2}{(q_1^2 + q^2)^2} e^{i\vec{\mathbf{q}} \cdot \vec{\mathbf{u}}} e^{iq_1 c \Delta t} e^{-m(q_1^2 + q^2)} \end{aligned} \quad (\text{I.5})$$

On exprime cette dernière intégrale par la différence entre l'intégrale :

$$I_1 = \int_0^{+\infty} dm \hat{\rho}(m) \int \frac{dq_1}{2\pi} \frac{d\vec{\mathbf{q}}}{(2\pi)^{d-1}} \frac{1 + A^2}{1 - A^2} \left(\frac{1}{q_1^2 + A^2 q^2} - \frac{1}{q_1^2 + q^2} \right) e^{i\vec{\mathbf{q}} \cdot \vec{\mathbf{u}} - mq^2} e^{iq_1 c \Delta t - mq_1^2}$$

et l'intégrale :

$$I_2 = \int_0^{+\infty} dm \hat{\rho}(m) \int \frac{dq_1}{2\pi} \frac{d\vec{q}}{(2\pi)^{d-1}} \frac{q^2}{(q_1^2 + q^2)^2} e^{i\vec{q}\cdot\vec{u} - mq^2} e^{iq_1 c\Delta t - mq_1^2} \quad (\text{I.6})$$

Calculons successivement ces deux intégrales. On utilise la propriété de la distribution Gamma telle que :

$$\frac{1}{q_1^2 + A^2 q^2} = \int_0^{+\infty} d\alpha e^{-\alpha(q_1^2 + A^2 q^2)}, \quad \frac{1}{q_1^2 + q^2} = \int_0^{+\infty} d\alpha e^{-\alpha(q_1^2 + q^2)} \quad (\text{I.7})$$

pour obtenir une intégrale d'exponentielles. L'intégrale I_1 peut se réécrire alors :

$$I_1 = \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} d\alpha \frac{1 + A^2}{1 - A^2} \int \frac{d\vec{q}}{(2\pi)^{d-1}} e^{i\vec{q}\cdot\vec{u} - (\alpha A^2 + m)q^2} \int \frac{dq_1}{2\pi} e^{iq_1 c\Delta t - (\alpha + m)q_1^2} \\ - \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} d\alpha \frac{1 + A^2}{1 - A^2} \int \frac{d\vec{q}}{(2\pi)^{d-1}} e^{i\vec{q}\cdot\vec{u} - (\alpha + m)q^2} \int \frac{dq_1}{2\pi} e^{iq_1 \Delta t - (\alpha + m)q_1^2} \quad (\text{I.8})$$

On reconnaît aux facteurs près les intégrales de Poisson,

$$I_1 = \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} d\alpha \frac{1 + A^2}{1 - A^2} \frac{1}{(4\pi)^{(d-1)/2}} \frac{e^{-u^2/4(\alpha A^2 + m)}}{(\alpha A^2 + m)^{(d-1)/2}} \\ \times \frac{1}{(4\pi)^{1/2}} \frac{e^{-(c\Delta t)^2/4(\alpha + m)}}{(\alpha + m)^{1/2}} \quad (\text{I.9})$$

$$- \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} d\alpha \frac{1 + A^2}{1 - A^2} \frac{1}{(4\pi)^{(d-1)/2}} \frac{e^{-u^2/4(\alpha + m)}}{(\alpha + m)^{(d-1)/2}} \\ \times \frac{1}{(4\pi)^{1/2}} \frac{e^{-(c\Delta t)^2/4(\alpha + m)}}{(\alpha + m)^{1/2}} \quad (\text{I.10})$$

Dans la première intégrale, on effectue le changement de variable $\alpha \rightarrow \alpha/A^2$. La deuxième intégrale restant inchangée, on obtient :

$$I_1 = \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} \frac{d\alpha/A^2}{(4\pi)^{d/2}} \frac{1 + A^2}{1 - A^2} \frac{e^{-u^2/4(\alpha + m)}}{(\alpha + m)^{(d-1)/2}} \frac{e^{-(cA\Delta t)^2/4(\alpha + A^2 m)}}{\frac{1}{A}(\alpha + A^2 m)^{1/2}} \\ - \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} \frac{d\alpha}{(4\pi)^{d/2}} \frac{1 + A^2}{1 - A^2} \frac{e^{-u^2/4(\alpha + m)}}{(\alpha + m)^{(d-1)/2}} \frac{e^{-(c\Delta t)^2/4(\alpha + m)}}{(\alpha + m)^{1/2}} \quad (\text{I.11})$$

$$= \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} \frac{d\alpha}{(4\pi)^{d/2}} \frac{1 + A^2}{1 - A^2} \frac{1}{A} \frac{e^{-u^2/4(\alpha + m)}}{(\alpha + m)^{(d-1)/2}} \frac{e^{-(cA\Delta t)^2/4(\alpha + A^2 m)}}{(\alpha + A^2 m)^{1/2}} \\ - \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} \frac{d\alpha}{(4\pi)^{d/2}} \frac{1 + A^2}{1 - A^2} \frac{e^{-u^2/4(\alpha + m)}}{(\alpha + m)^{(d-1)/2}} \frac{e^{-(c\Delta t)^2/4(\alpha + m)}}{(\alpha + m)^{1/2}} \quad (\text{I.12})$$

En regroupant les deux intégrales, on obtient après factorisation :

$$I_1 = \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} \frac{d\alpha}{(4\pi)^{d/2}} \frac{1+A^2}{1-A^2} \frac{e^{-u^2/4(\alpha+m)}}{(\alpha+m)^{d/2}} \times \left(\frac{1}{A} \left(\frac{\alpha+m}{\alpha+A^2m} \right)^{1/2} e^{-\frac{(cA\Delta t)^2}{4(\alpha+A^2m)}} - e^{-\frac{(c\Delta t)^2}{4(\alpha+m)}} \right) \quad (\text{I.13})$$

On effectue le changement de variable $\alpha \rightarrow m\alpha$ qui nous donne :

$$I_1 = \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} \frac{m d\alpha}{(4\pi)^{d/2}} \frac{1+A^2}{1-A^2} \frac{e^{-u^2/4m(\alpha+1)}}{m^{d/2}(\alpha+1)^{d/2}} \times \left(\frac{1}{A} \left(\frac{\alpha+1}{\alpha+A^2} \right)^{1/2} e^{-\frac{(cA\Delta t)^2}{4m(\alpha+A^2)}} - e^{-\frac{(c\Delta t)^2}{4m(\alpha+1)}} \right) \quad (\text{I.14})$$

$$= \int_0^{+\infty} dm \frac{m \hat{\rho}(m)}{(4m\pi)^{d/2}} \int_0^{+\infty} d\alpha \frac{1+A^2}{1-A^2} \frac{e^{-u^2/4m(\alpha+1)}}{(\alpha+1)^{d/2}} \times \left(\frac{1}{A} \left(\frac{\alpha+1}{\alpha+A^2} \right)^{1/2} e^{-\frac{(cA\Delta t)^2}{4m(\alpha+A^2)}} - e^{-\frac{(c\Delta t)^2}{4m(\alpha+1)}} \right) \quad (\text{I.15})$$

On simplifie cette dernière intégrale en effectuant le changement de variable $\beta \rightarrow 1/(1+\alpha)$. L'intégration se fait sur l'intervalle $[0, 1]$.

La mesure d'intégration est alors $d\alpha = -\beta^{-2}d\beta$.

L'expression de I_1 devient donc :

$$I_1 = \int_0^{+\infty} dm \frac{m \hat{\rho}(m)}{(4m\pi)^{d/2}} \int_0^1 d\beta \beta^{-2} \frac{1+A^2}{1-A^2} \beta^{d/2} e^{-u^2\beta/4m} \times \left(\frac{1}{A} \frac{e^{-\frac{(cA\Delta t)^2\beta}{4m(1-(1-A^2)\beta)}}}{\sqrt{1-(1-A^2)\beta}} - e^{-\frac{(c\Delta t)^2\beta}{4m}} \right) \quad (\text{I.16})$$

$$= \int_0^{+\infty} dm \frac{m \hat{\rho}(m)}{(4m\pi)^{d/2}} \int_0^1 d\beta \beta^{d/2-2} \frac{1+A^2}{1-A^2} e^{-(u^2+(c\Delta t)^2)\beta/4m} \times \left(\frac{1}{A} \frac{e^{-\frac{(c\Delta t)^2\beta}{4m}(A^2/(1-(1-A^2)\beta)-1)}}}{\sqrt{1-(1-A^2)\beta}} - 1 \right) \quad (\text{I.17})$$

$$= \int_0^{+\infty} dm \frac{m \hat{\rho}(m)}{(4m\pi)^{d/2}} \int_0^1 d\beta \beta^{d/2-2} \frac{1+A^2}{1-A^2} e^{-(u^2+(c\Delta t)^2)\beta/4m} (f_{m,\Delta t}(\beta) - 1) \quad (\text{I.18})$$

$$\text{où } f_{m,\Delta t}(\beta) = \frac{1}{A} \frac{e^{-\frac{(c\Delta t)^2\beta}{4m}(A^2/(1-(1-A^2)\beta)-1)}}}{\sqrt{1-(1-A^2)\beta}}$$

Calculons à présent l'intégrale I_2 en effectuant successivement les différents changements de variables effectués plus haut. Reprenons I_2 telle que :

$$I_2 = \int_0^{+\infty} dm \hat{\rho}(m) \int \frac{d\vec{q}}{(2\pi)^{d-1}} q^2 e^{i\vec{q}\cdot\vec{u}-mq^2} \int \frac{dq_1}{2\pi} \frac{e^{iq_1 c\Delta t - mq_1^2}}{(q_1^2 + q^2)^2} \quad (\text{I.19})$$

Avec la définition de la distribution Gamma telle que :

$$\frac{1}{(q_1^2 + q^2)^2} = \int_0^{+\infty} \frac{d\alpha}{\Gamma(2)} \alpha e^{-\alpha(q_1 + q^2)} = \int_0^{+\infty} d\alpha \alpha e^{-\alpha(q_1 + q^2)} \quad (\text{I.20})$$

l'intégrale I_2 peut se réécrire à l'aide d'une intégrale de Poisson. Soit :

$$\begin{aligned} I_2 &= \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} d\alpha \alpha \int \frac{d\vec{q}}{(2\pi)^{d-1}} q^2 e^{i\vec{q}\cdot\vec{u} - (\alpha+m)q^2} \int \frac{dq_1}{2\pi} e^{iq_1 c\Delta t - (\alpha+m)q_1^2} \\ &= \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} \frac{d\alpha}{(4\pi)^{1/2}} \alpha \\ &\quad \times \int \frac{d\vec{q}}{(2\pi)^{d-1}} q^2 e^{i\vec{q}\cdot\vec{u} - (\alpha+m)q^2} \frac{e^{-(c\Delta t)^2/4(\alpha+m)}}{(\alpha+m)^{1/2}} \end{aligned} \quad (\text{I.22})$$

$$\begin{aligned} &= \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} \frac{d\alpha}{(4\pi)^{1/2}} \frac{\alpha}{(\alpha+m)^{1/2}} e^{-(c\Delta t)^2/4(\alpha+m)} \\ &\quad \times \int \frac{d\vec{q}}{(2\pi)^{d-1}} q^2 e^{i\vec{q}\cdot\vec{u} - (\alpha+m)q^2} \end{aligned} \quad (\text{I.23})$$

Pour le calcul de l'intégrale vectorielle, on peut comme précédemment se ramener à celui d'une intégrale de Poisson en introduisant la dérivée suivant m de $m \rightarrow e^{i\vec{q}\cdot\vec{u} - (\alpha+m)q^2}$. Et ainsi,

$$q^2 e^{i\vec{q}\cdot\vec{u} - (m+\alpha)q^2} = -\partial_m e^{i\vec{q}\cdot\vec{u} - (\alpha+m)q^2} \quad (\text{I.24})$$

Les signes de la dérivée et de l'intégrale peuvent s'invertir et on obtient alors la

nouvelle expression telle que :

$$I_2 = \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} \frac{d\alpha}{(4\pi)^{d/2}} \frac{\alpha}{(\alpha+m)^{1/2}} e^{-(c\Delta t)^2/4(\alpha+m)} \times \partial_m \left(\int \frac{-d\vec{q}}{(2\pi)^{d-1}} e^{i\vec{q}\cdot\vec{u} - (\alpha+m)q^2} \right) \quad (\text{I.25})$$

$$= \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} \frac{d\alpha}{(4\pi)^{1/2}} \frac{\alpha}{(\alpha+m)^{1/2}} e^{-(c\Delta t)^2/4(\alpha+m)} \times \partial_m \left(\frac{-1}{(4\pi)^{(d-1)/2}} \frac{e^{-u^2/4(\alpha+m)}}{(\alpha+m)^{(d-1)/2}} \right) \quad (\text{I.26})$$

$$= \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} \frac{d\alpha}{(4\pi)^{d/2}} \frac{\alpha}{(\alpha+m)^{1/2}} e^{-(c\Delta t)^2/4(\alpha+m)} \times \partial_m \left(\frac{-e^{-u^2/4(\alpha+m)}}{(\alpha+m)^{(d-1)/2}} \right) \quad (\text{I.27})$$

$$= \int_0^{+\infty} dm \hat{\rho}(m) \int_0^{+\infty} \frac{d\alpha}{(4\pi)^{d/2}} \frac{\alpha}{(\alpha+m)^{1/2}} e^{-(c\Delta t)^2/4(\alpha+m)} \times e^{-u^2/4(\alpha+m)} \left(\frac{d-1}{2} (\alpha+m)^{-1} - \frac{u^2}{4} (\alpha+m)^{-2} \right) (\alpha+m)^{-\frac{d-1}{2}} \quad (\text{I.28})$$

Comme précédemment, on effectue le changement de variable $\alpha \rightarrow m\alpha$ pour avoir l'intégrale simplifiée suivante :

$$I_2 = \int_0^{+\infty} dm \frac{m\hat{\rho}(m)}{(4m\pi)^{d/2}} \int_0^{+\infty} d\alpha \frac{\alpha}{(\alpha+1)^{d/2}} e^{-(c\Delta t)^2/4m(\alpha+1)} \times e^{-u^2/4m(\alpha+1)} \left(\frac{d-1}{2} (\alpha+1)^{-1} - \frac{u^2}{4m} (\alpha+1)^{-2} \right) \quad (\text{I.29})$$

Il nous reste à effectuer le dernier changement de variable $\beta = 1/(\alpha+1)$ pour obtenir une formule équivalente à l'intégrale I_1 . Ainsi,

$$I_2 = \int_0^{+\infty} dm \frac{m\hat{\rho}(m)}{(4m\pi)^{d/2}} \int_0^1 d\beta \beta^{d/2-2} e^{-(c\Delta t)^2\beta/4m} \times e^{-u^2\beta/4m} (1-\beta) \left(\frac{d-1}{2} - \frac{u^2}{4m}\beta \right) \quad (\text{I.30})$$

$$= \int_0^{+\infty} dm \frac{m\hat{\rho}(m)}{(4m\pi)^{d/2}} \int_0^1 d\beta \beta^{d/2-2} e^{-(u^2+(c\Delta t)^2)\beta/4m} \times \left(\frac{d-1}{2} - \left(\frac{d-1}{2} + \frac{u^2}{4m} \right) \beta + \frac{u^2}{4m} \beta^2 \right) \quad (\text{I.31})$$

On remarque que le polynôme $\beta \rightarrow \beta^{d/2-2}(1-\beta)\left(\frac{d-1}{2} - \frac{u^2}{4m}\beta\right)$ admet des zéros en 0

et 1 respectivement de multiplicité $d/2 - 1$ et 1. On effectue plusieurs intégrations par partie successivement afin de conserver seulement le monome $\beta^{d/2-1}$. Soit :

$$\begin{aligned}
I_2 &= \int_0^{+\infty} dm \frac{m\hat{\rho}(m)}{(4m\pi)^{d/2}} \frac{d-1}{2} \int_0^1 d\beta \beta^{d/2-2} (1-\beta) e^{-(u^2+(c\Delta t)^2)\beta/4m} \\
&\quad - \int_0^{+\infty} dm \frac{m\hat{\rho}(m)}{(4m\pi)^{d/2}} \frac{u^2}{\int_0^1} d\beta \beta^{d/2-1} 4m(1-\beta) e^{-(u^2+(c\Delta t)^2)\beta/4m} \\
&= \int_0^{+\infty} dm \frac{m\hat{\rho}(m)}{(4m\pi)^{d/2}} \int_0^1 d\beta \beta^{d/2-2} e^{-(u^2+(c\Delta t)^2)\beta/4m} g_{\vec{u},\Delta t}(\beta) \quad (I.32)
\end{aligned}$$

où

$$g_{\vec{u},\Delta t}(\beta) = \frac{u^2}{u^2 + (c\Delta t)^2} + (1-\beta) \left(\frac{d}{2} \left(1 - \frac{u^2}{u^2 + (c\Delta t)^2} \right) - \frac{1}{2} \right) \quad (I.33)$$

Nous avons donc une expression analytique de la covariance entre deux fronts de saturation à deux temps donnés. On peut spécifier la covariance dans le cas où le variogramme est Gaussien, de variance globale k_0 et de longueur de corrélation ℓ_c . La mesure spectrale $\hat{\rho}(m)$ est alors exprimée par un Dirac (voir annexe J) :

$$\frac{\hat{\rho}(m)}{(4m\pi)^{d/2}} = k_0 \delta_{m-\ell_c^2/4} \quad (I.34)$$

Dans ce cas, les intégrales I_1 et I_2 se simplifient et on obtient les nouvelles expressions suivantes avec $\theta = \frac{u^2+(c\Delta t)^2}{\ell_c^2}$

$$\begin{aligned}
I_1 &= \frac{k_0\ell_c^2}{4} \int_0^1 d\beta \beta^{d/2-2} e^{-\theta\beta} \frac{1+A^2}{1-A^2} (f_{\ell_c^2/4}(\beta) - 1) \\
I_2 &= \frac{k_0\ell_c^2}{4} \int_0^1 d\beta \beta^{d/2-2} e^{-\theta\beta} g_{\vec{u},\Delta t}(\beta) \quad (I.35)
\end{aligned}$$

Finalement, on exprime analytiquement la fonction de covariance C :

$$\begin{aligned}
C(\vec{u}; \Delta t) &= I_1 - I_2 \\
&= \frac{k_0\ell_c^2}{4} \int_0^1 d\beta \beta^{d/2-2} e^{-\theta\beta} \left(\frac{1+A^2}{1-A^2} (f_{\ell_c^2/4}(\beta) - 1) - g_{\vec{u},\Delta t}(\beta) \right) \quad (I.36)
\end{aligned}$$

Annexe J

Calcul de la mesure spectrale

Lorsque le champ de perméabilité est considéré loi log-normal, sa fonction de corrélation est alors :

$$K(\mathbf{u}) = K_0 \exp\left(\frac{-u^2}{\ell_c^2}\right) \quad (\text{J.1})$$

où $u^2 = \mathbf{u} \cdot \mathbf{u}$ et ℓ_c est la longueur de corrélation du champ. La transformée de Fourier de K fait appel à l'intégrale de Poisson. La transformée de Fourier de K est alors,

$$\hat{K}(\mathbf{q}) = \int_{\mathbb{R}^d} d\mathbf{u} \exp\left(\frac{-u^2}{\ell_c^2}\right) \exp(-i\mathbf{q} \cdot \mathbf{u}) \quad (\text{J.2})$$

On obtient l'expression dans l'exponentielle à l'aide d'un carré remarquable,

$$\frac{-u^2}{\ell_c^2} - i\mathbf{q} \cdot \mathbf{u} = -(\mathbf{u}/\ell_c - i\mathbf{q}\ell_c/2)^2 - q^2\ell_c^2/4 \quad (\text{J.3})$$

On obtient :

$$\hat{K}(\mathbf{q}) = K_0 \exp(-q^2\ell_c^2/4) \int_{\mathbb{R}^d} d\mathbf{u} \exp -(\mathbf{u}/\ell_c - i\mathbf{q}\ell_c/2)^2 \quad (\text{J.4})$$

On effectue alors le changement de variable $\mathbf{v} = \mathbf{u}/\ell_c - i\mathbf{q}\ell_c/2$ avec $d\mathbf{v} = d\mathbf{u}/(\ell_c)^d$. Soit encore :

$$\hat{K}(\mathbf{q}) = K_0(\ell_c)^d(\pi)^{d/2} \exp(-q^2\ell_c^2/4) \quad (\text{J.5})$$

Ainsi, si l'on cherche $\hat{\rho}(m)$ tel que :

$$\hat{K}(\mathbf{q}) = \int_0^{+\infty} dm \hat{\rho}(m) \exp(-mq^2) \quad (\text{J.6})$$

alors, on obtient :

$$\hat{\rho}(m) = K_0 \cdot (4\pi m)^{d/2} \delta(m - q^2\ell_c^2/4) \quad (\text{J.7})$$

Bibliographie

- [1] D.V. Lindley. On the measure of information provided by an experiment. *Annals of Statistics*, 27 :986–1005, 1956.
- [2] Khalid Aziz and Antonin Settari. *Petroleum Reservoir Simulation*. Applied Science Publisher England, 1979.
- [3] New York D.-W. Peaceman. *Fundamentals of Numerical Reservoir Simulation*. Elsevier Science Publisher, 1977.
- [4] Anterion, Eymard, and Karcher. Use of parameter gradient for reservoir history matching. February 1989. SPE 18433.
- [5] G. Chavent, M. Dupuy, and P. Lemonnier. History matching by use of optimal control. *SPEJ*, 1(15) :74–86, 1975.
- [6] F. Roggero and D. Guerillot. Gradient method and bayesian formalism application to petrophysical parameter characterization. In Z.E. Heinemann and M. Kriebennegg, editors, *European Conference on the Mathematics of Oil Recovery*, 5, pages 271–282. ECMOR, September 1996.
- [7] G. Matheron, H. Beucher, C. de Fouquet, A. Galli, D. Guerillot, and C. Ravenne. Conditional simulation of the geometry of fluvio-deltaic reservoirs. pages 591–599. SPE Annual Technical Conference and Exhibition, Dallas, Texas, 1987. SPE 16753.
- [8] A.G. Journel. Geostatistics for conditional simulation of ore bodies. *Econ. geology*, 3(69) :673–687, 1974.
- [9] D.S. Oliver. Moving average for gaussian simulation in two and three dimensions. *Math. Geology*, 27(8) :939–960, 1995.
- [10] M. Le Ravalec, L.Y. Hu, and B. Noetinger. The fft moving average (fft-ma) method : an efficient tool for generating and conditioning gaussian simulations. *Math. Geol.*, 6(32), 2000.
- [11] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, 2005.
- [12] G.R Gavalas, P.C Shah, and J.H. Seinfeld. Reservoir history matching by bayesian estimation. *SPEJ*, pages 337–350, December 1976.
- [13] K. Mosegaard and A. Tarantola. *International Handbook of Earthquake & Engineering Seismology (Part A)*, pages 237–265. Academic Press, 2002.

- [14] M.À. Carreira-Perpiñan. *Continuous latent variable models for dimensionality reduction and sequential data reconstruction*. PhD thesis, Departement of Computer Science University of Sheffield, UK, February 2001.
- [15] I. Csiszar. Information type measures of differences of probability distribution and indirect observations. *Studia Math. Hungarica*, 2 :299–318, 1967.
- [16] I.J. Taneja. Refinement inequalities among symmetric divergence measures. *Mathematics Subject Classification*, 2000. URL <http://arXiv.org/abs/math/0505204v1>.
- [17] F. Venegas-Martínez. On information measures and prior distributions : a synthesis. *Morfismo*, 8(2) :27–50, 2004.
- [18] H. Jeffreys. *Theory of Probability*, 3rd. edition. Oxford University Press, 1961.
- [19] A. Zellner. *An introduction to Bayesian inference in Econometrics*. New York, Wiley, 1971.
- [20] G.E.P. Box and G.C. Tiao. *Bayesian Inference and Statistical Analysis*. Massachusetts, 1973.
- [21] J.O. Berger and J.M. Bernardo. On the development of reference priors. *Bayesian Statistics*, 4 :35–60, 1992.
- [22] A. Zellner. Past and recent results on maximal data information. *Journal of Statistical Research*, 32(1) :1–22, 1998.
- [23] J.M. Bernardo. Reference posterior distributions for bayes inference. *J. Roy. Statist., Soc. Ser B* 41, pages 113–147, 1979.
- [24] J.O Berger and J.M. Bernardo. Estimating a product of means : Bayesian analysis with reference priors. *J. Amer. Statist. Assoc.*, 84 :200–207, 1989.
- [25] J.O Berger and J.M. Bernardo. Reference priors in a variance components problem. *In Bayesian Analysis in Statistics and Econometrics*, 1992.
- [26] J.O Berger and J.M. Bernardo. Ordered group reference priors with application to a multinomial problem. *Biometrika*, 79 :25–37, 1992.
- [27] J.O Berger and J.M. Bernardo. On the development of the reference prior method. *In Bayesian Analysis*, 4, 1992.
- [28] S. Ghosal. Reference priors in multiparameter nonregular cases. *Test*, 6(1) :159–186, 1997.
- [29] S. Ghosal and T. Samanta. Asymptotic behaviour of bayes estimates and posterior distributions in multiparameter nonregular cases. *Math. Methods Statist.*, 4 :361–388, 1995.
- [30] I.A. Ibragimov and R.Z. Has'minski. *Statistical Estimation : Asymptotic Theory*. Springer-Verlag, New York, 1981.
- [31] R.L. Smith. Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72 :69–90, 1985.
- [32] C.P. Robert. *Le choix bayésien. Principes et pratique*. Springer-Verlag, Springer, 2006.

- [33] N. Cressie, C.A. Gotway, and M.O. Grondona. Spatial prediction from networks. *Chemometrics and Intelligent Laboratory Systems*, 7 :251–271, 1990.
- [34] J.V. Zidek, W. Sun, and D.N. Le. Designing and integrating composite networks for monitoring multivariate gaussian pollution fields. *Applied Statistics*, 49 :63–79, 2000.
- [35] B.P. Marchant and R.M. Lark. Optimized sample scheme for geostatistical surveys. *Mathematical Geology*, 39 :113–134, 2007.
- [36] M. Fuentes, A. Chaudhuri, and D.M. Holland. Bayesian entropy for spatial sampling design of environmental data, environmental and ecological statistics. *Environmental and Ecological Statistics*, 14 :323–340, 2007.
- [37] R. H. Byrd, P. Lu, and J. Nocedal. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific and Statistical Computing*, 15(5) : 1190–1208, 1995.
- [38] C. Zhu, RH Bryd, and J. Nocedal. L-bfgs-b : Algorithm 778 : L-bfgs-b, fortran routines for large scale bound constrained optimization. *ACM Transactions on Mathematical Software*, 23(4) :550–560, 1997.
- [39] F. Daviau. *Interprétation des essais de puits : Les méthodes nouvelles*. Technip, 1986.
- [40] G. Bourdarot. *Interprétation des essais de puits*. Technip, 1990.
- [41] T.D. Streltsova. *Well Testing in Heterogeneous Formations*. Wiley, New York, 1988. p. 413.
- [42] P. Indelman. Averaging of unsteady flow in heterogeneous media of stationary conductivity. *J. Fluid. Mech.*, 12(310) :39–60, 1996.
- [43] B. Noetinger and Y. Gautier. Use of the fourier-laplace transform and of diagrammatical methods to interpret pumping tests in heterogeneous reservoirs. *Advances in water Ressources*, 21(7) :581–590, 15 juin 1997.
- [44] B. Noetinger, V. Artus, and G. Zargar. The future of stochastic and upscaling method in hydrogeology. *J. Hydrogeol*, 8(13) :184–201, February 2005.
- [45] N. Goldenfeld. *Lectures on Phase Transitions and the Renormalization Group*. Addison Wesley. Frontiers in Physics, 1992.
- [46] VV. Jikov, SM. Kozlov, and OA. Oleinik. *Homogenization of Differential Operators*. Springer, Berlin Heidelberg New York, 1994.
- [47] Y. Gautier and B. Noetinger. Determination of geostatistical parameters using well test data. Number SPE 49278. SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, 27-30 September 1998.
- [48] R.K. Romeu. *Écoulements en milieux poreux hétérogènes : prise de moyenne de perméabilité en régime permanent et transitoire*. PhD thesis, Université Paris VI, 1994.
- [49] T. Yortos. Estimation of permeability variogram from stochastic analysis of pressure transients of multiple wells in heterogeneous reservoirs. SPE 71st Ann. Tech. Conf and Exh., Denver, Colorado, october 1996.

- [50] M. Abramowitz and I. Stegun. *Handbook of mathematical functions*. Dover Cooks Publisher, New York, 1972.
- [51] D. Oliver. The averaging process in permeability estimation from well test data. présenté au SPE 64th Ann. Tech. Conf. and Exh., San Antonio, 1989.
- [52] G. Blanc, D. Guérillot, D. Rahon, and F. Roggero. Building geostatistical models constrained by dynamic data : A posteriori constraints. Number SPE 35478, pages 19–33. European 3D Reservoir Modeling Conference, Norway, Stavanger, 1996.
- [53] M. J. King and V. A. Dunayevsky. Why waterflood works : a linearized stability analysis. *SPE 19648*, pages 187–200, 1989.
- [54] V. Artus and B. Nøttinger. Macrodispersion approach for upscaling of two-phase, immiscible flow in heterogeneous porous media. in 8th European Conference on the Mathematics of Oil Recovery, Freiberg, Germany, 2002.
- [55] V. Artus, B. Nøttinger, and L Ricard. Dynamics of the water oil front for two-phase, immiscible flow in heterogeneous porous media. 1 stratified media. *Transport in Porous Media*, 56 :283–303, 2004.
- [56] V. Artus, F. Furtado, B. Nøttinger, and F. Pereira. Stochastic analysis of two-phase immiscible flow in stratified porous media. *Computational and Applied Mathematics*, 23 :153–172, 2002.
- [57] J. Hagoort. Displacement stability of water drives in water-wet connate-water-bearing reservoirs. *SPE 4268*, pages 63–74, 1974.
- [58] P. G. Saffman and G. Taylor. The penetration of a fluid into a porous medium or hele-shaw cell containing a more viscous liquid. *Proc. R. Soc. London*, A245 : 312–329, 1958.
- [59] R. L. Chuoke, P. Van Meurs, and C. Van der Poel. The instability of slow, immiscible, viscous liquid-liquid displacements in permeable media. *Trans. AIME*, 216 : 188–194, 1959.
- [60] G. M. Homsy. Viscous fingering in porous media. *Ann. Rev. Fluid Mech.*, 19 : 271–311, 1987.
- [61] S. E. Buckley and M. C. Leverett. Mechanism of fluid displacement in sands. *Trans. AIME*, 146 :107–116, 1942.
- [62] P. Langlo and M. S. Espedal. Macrodispersion for two-phase, immiscible flow in porous media. *Advances in Water Resources*, 17 :297–316, 1995.
- [63] G. Dagan and V. Cvetkovic. Reactive transport and immiscible flow in geological media. i-general theory. *Proc. R. Soc. London*, 452 :285–301, 1996.
- [64] A. T. Corey. The inter-relation between gas and oil relative permeabilities. *Producers Monthly*, 19 :38–41, 1954.
- [65] G. I. Barenblatt, V. M. Entov, and V. M. Ryzhik. *Theory of Fluid Flows Through Natural Rocks*. Springer, 1990.
- [66] B. Nøttinger, V. Artus, and L. Ricard. Dynamics of the water-oil front for two-phase, immiscible flows in heterogeneous porous media. *Transport in Porous Media*, 56 :305–328, 2004.

-
- [67] Martin Jacobsen. Laplace and the origin of the Ornstein-Uhlenbeck process. *Bernoulli*, 2 :271–286, 1996.
- [68] P.E. Spesivtsev. The statistical characteristics of the displacement front in a randomly heterogeneous medium. *Journal of Mathematics and Mechanics*, 71 :424–431, 2007.
- [69] B. Nøttinger, P.E. Spesivtsev, and V. Teodorovich. Stochastic analysis of a displacement front in a randomly heterogeneous medium. *Fluid Dynamics*, 41 :830–842, 2006.
- [70] C.E. Shannon. A mathematical theory of communication. *AT&T Bell Laboratories Technical Journal*, 27 :379–423 and 623–656, 1948.
- [71] A.M. Turing. On computable numbers with an application to the Entscheidungsproblem. *Proc. London Math. Soc.*, 2(42) :230–265, 1936.
- [72] E.T. Jaynes. *Probability Theory*. MIT Press, 1950.
- [73] S. Kullback and R.A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22 :79–86, 1951.
- [74] S. Kullback. *Information Theory and Statistics*. Dover Publications, 1968.