
Méta-modélisation de sorties vectorielles et de la fonction objectif

Les propriétés d'intérêt considérées pour la méta-modélisation sont issues de la simulation d'écoulement. Ce sont donc des fonctions du temps, voire de l'espace, et leurs valeurs simulées peuvent être vues comme des vecteurs dont chaque composante correspond à un temps et/ou une maille donnée. La méta-modélisation de sorties vectorielles est ainsi un enjeu important en ingénierie de réservoir.

Ce chapitre est divisé en trois parties. Tout d'abord, nous résumons l'ensemble des propriétés et sorties d'intérêt de la simulation d'écoulement au sein de la partie 3.1. La méta-modélisation de ces sorties, généralement vectorielles, est un enjeu pour de nombreuses applications. Les méthodes proposées pour la méta-modélisation de sorties vectorielles dans des cadres simple ou multi-fidélité sont ensuite décrites en partie 3.2. Puis en partie 3.3, nous nous intéressons à la méta-modélisation de la fonction objectif quantifiant l'erreur entre données de production et réponses simulées. En particulier, nous proposons de l'approcher en utilisant les estimations des propriétés vectorielles à caler.

3.1 Propriétés d'intérêt de la simulation d'écoulement

En ingénierie de réservoir, on s'intéresse à l'évolution dynamique de la pression et des fluides dans le réservoir en réponse à un schéma de production. La grande majorité des sorties issues de la simulation d'écoulement sont donc des propriétés qui dépendent du temps. Ces propriétés sont calculées à chaque pas de temps considéré lors de la simulation d'écoulement. En pratique, on ne conserve leurs réponses qu'à certains temps, et notamment ceux pour lesquels on dispose de mesures de données dynamiques. Parmi les propriétés dynamiques d'intérêt, on peut distinguer les propriétés mesurées aux puits, les propriétés définies au niveau du réservoir, et les propriétés spatio-temporelles.

Les propriétés d'intérêt aux puits

Les propriétés d'intérêt aux puits regroupent notamment la pression en fond de puits (*bottomhole pressure*) et une mesure de la production des différents fluides, comme le débit d'huile. Pour l'eau et le gaz, on préfère souvent une expression de leur production sous la forme d'un rapport. Pour l'eau, on utilise la fraction d'eau (*watercut*). C'est une réponse prenant ses valeurs entre 0 et 1, qui désigne la part en eau produite par rapport à la part totale des fluides. Pour le gaz, on utilise fréquemment le ratio gaz/huile (*gaz/oil ratio*), qui désigne un rapport entre les volumes de gaz et d'huile produits en condition de pression et température de surface. Cette réponse permet de prendre en compte le gaz dissous au sein du pétrole brut produit. Selon le cas étudié et les options du simulateur d'écoulement,

de nombreuses autres sorties peuvent être simulées aux puits.

Les propriétés associées au réservoir

Ces propriétés représentent le comportement dynamique global au niveau du réservoir, et sont obtenues à partir des propriétés dynamiques aux différents puits. Il peut s'agir par exemple de débits ou de volumes cumulés pour un fluide donné, comme le volume cumulé de l'huile produite par le réservoir ou encore le débit total d'eau produite. Ces sorties fournissent des indicateurs globaux pouvant être utilisés dans des calculs économiques visant à estimer la valeur d'un schéma de production ou d'un emplacement pour un puits additionnel. Elles peuvent ainsi être utiles pour le choix d'une stratégie d'exploitation dans le cadre du développement du champ.

Les propriétés spatiales

Les propriétés spatiales sont définies dans chaque maille du réservoir. Il s'agit principalement de la distribution de la pression et des saturations en huile ou en eau dans le réservoir. La modélisation de ces propriétés est particulièrement intéressante lorsque l'on dispose de données sismiques répétées dans le temps, aussi appelées données de sismique 4D (variations d'impédance acoustique dans le temps par exemple). En effet, ces attributs sismiques peuvent être calculés via un modèle pétro-élastique à partir des cartes de pression et saturation fournies par la simulation d'écoulement. Les attributs sismiques obtenus peuvent ensuite être comparés aux données de sismique 4D au sein d'une fonction objectif adaptée et être pris en compte lors du processus de calage [Roggero et al., 2012].

La fonction objectif

La fonction objectif n'est pas une propriété de la simulation d'écoulement. C'est cependant une sortie d'intérêt majeure utilisée lors du calage d'historique afin de quantifier l'écart entre les données de production mesurées et celles simulées à l'aide du modèle réservoir. Nous en rappelons ci-dessous la définition introduite dans la section 1.2.2 :

$$FO(\mathbf{x}) = \frac{1}{2} \sum_{k=1}^K \omega_k \sum_{t=1}^{T(k)} \left(\frac{y_k(\mathbf{x}, t) - d_{kt}^{obs}}{\sigma_{kt}} \right)^2 \quad (3.1)$$

On rappelle que les données de production simulées sont notées $y(\mathbf{x}, t)$ et celles mesurées d^{obs} . Les indices $k \in [1, K]$ et $t \in [1, T(k)]$ font référence aux différentes sorties et aux différents temps pour lesquels ces sorties sont mesurées. Les coefficients ω et σ sont des poids définis par l'ingénieur. On note également que les données de production sont des réponses dynamiques. Leur dépendance au temps est indiquée par l'indice t . Les données de production sont généralement des propriétés aux puits, mais les propriétés associées au réservoir peuvent également être considérées dans le calcul de la fonction objectif.

Applications

La méta-modélisation de sorties vectorielles rend possible l'analyse de sensibilité simultanément sur les différents éléments du vecteur. Ainsi, dans [Douarce et al., 2014], une analyse de sensibilité réalisée au cours de la récupération de l'huile du réservoir permet de caractériser l'évolution de l'impact des différents surfactants et polymères testés dans le temps. De façon similaire, la méta-modélisation de

propriétés spatiales peut permettre de caractériser l'évolution de l'impact des différents paramètres dans l'espace [Marrel et al., 2015] [Da Veiga and Gervais, 2015]. La méta-modélisation de propriétés au niveau du réservoir est intéressante pour faire des prédictions probabilistes sur la production dans le temps [Feraille and Marrel, 2012]. Nous proposons également dans cette thèse d'utiliser des méta-modèles approchant les sorties de puits afin d'améliorer la modélisation de la fonction objectif, et ainsi accélérer le calage d'historique. Cette dernière approche est décrite en détail dans la partie 3.3.

3.2 Méta-modélisation de sorties vectorielles

La méta-modélisation d'une sortie vectorielle peut se faire à travers la construction de modèles de krigeage/co-krigeage multi-fidélité pour chacun des éléments de la sortie vectorielle. Cependant, la grande taille des sorties rend cette approche peu pratique, voire irréaliste pour les sorties spatiales dont la dimension est celle de la grille réservoir (jusqu'à plusieurs millions d'éléments). Une alternative consiste à réduire la dimension des sorties vectorielles afin de pouvoir les approcher avec un nombre limité de méta-modèles. Ceci peut se faire par exemple comme dans [Douarche et al., 2014] en réalisant une analyse en composantes principales (section 3.2.1) [Loeve, 1978], puis en appliquant la méta-modélisation par krigeage dans la base réduite obtenue (section 3.2.2). Cette approche est étendue ici au cadre multi-fidélité (section 3.2.3).

3.2.1 L'analyse en composantes principales (ACP)

L'approche consistant à construire un méta-modèle par élément du vecteur est coûteuse en temps de calcul. De plus, elle est peu efficace car elle ignore les corrélations existantes entre les différents éléments de la sortie vectorielle. Une approche plus rapide et efficace passe par la réduction de la dimension de la sortie vectorielle en appliquant par exemple une analyse en composantes principales, notée ACP. Selon le domaine d'application considéré, l'ACP est aussi connue sous le nom de transformation de Karhunen-Loève [Loeve, 1978] ou encore décomposition orthogonale aux valeurs propres.

Considérons que la réponse dynamique que l'on souhaite approcher $y(\mathbf{x}, t)$ avec $\mathbf{x} \in \mathbb{R}^d$, $t \in [1, T]$ soit la réalisation d'un processus Gaussien centré $Y_t(\mathbf{x})$, tel que $\mathbf{E}[Y_t] = 0$ pour tout $t \in [1, T]$. L'application au cas Gaussien du théorème de Karhunen-Loève [Revuz, 1997], Y_t admet la décomposition suivante :

$$Y_t(\mathbf{x}) = \sum_{l=1}^{\infty} A_l(\mathbf{x})\phi_l(t) \quad (3.2)$$

avec ϕ_l des fonctions formant une base orthonormale et A_l des variables Gaussiennes indépendantes définies par $A_l(\mathbf{x}) = \int_1^T Y_t(\mathbf{x})\phi_l(t)dt$. L'intérêt d'une telle décomposition est de pouvoir séparer les effets du temps (ou de l'espace) de ceux des paramètres incertains.

Généralement, seuls quelques termes de la base sont suffisants pour fournir une bonne approximation de Y_t . Une version tronquée à l'ordre L de l'Equation 3.2 est alors :

$$Y_t(\mathbf{x}) \approx \sum_{l=1}^L A_l(\mathbf{x})\phi_l(t) \quad (3.3)$$

Sous cette forme tronquée, les fonctions orthonormales ϕ_l sont choisies de telle sorte que l'approximation soit aussi bonne que possible au sens des moindres carrés, pour la valeur de L considérée. Elles

sont définies séquentiellement de manière à ce que les i premières fonctions fournissent la meilleure approximation possible avec i termes. Par ailleurs, on peut quantifier la part de la variance du champ reproductible par l'approximation tronquée à l'ordre L . En général, on choisit L tel qu'au moins 95% de la variance de la sortie puisse être reproduite par la décomposition.

On remarque qu'il est possible de considérer une décomposition similaire pour une sortie spatiale $y(\mathbf{x}, u, v, w)$ définie sur un espace (u, v, w) , comme par exemple une carte de la pression dans le réservoir à un temps donné. Il faut alors remplacer les fonctions orthonormales $\phi_l(t)$ par $\phi_l(u, v, w)$ dans les Equations 3.2 et 3.3.

En pratique, il est possible de calculer une version empirique de l'Équation 3.3 à partir d'un échantillon de réponses. On parle alors de transformation de Karhunen-Loève. Les différents termes de la décomposition sont calculés en appliquant une ACP sur l'échantillon de réponses.

3.2.2 ACP et krigeage

Nous présentons ici une méthode pour la construction d'un méta-modèle d'une sortie vectorielle reposant sur l'utilisation combinée de l'expression tronquée de la transformation de Karhunen-Loève (Équation 3.3) et du krigeage. Comme illustré par la Figure 3.1, cette méthode consiste à réaliser une décomposition en base réduite de la réponse d'intérêt par ACP, puis à construire des modèles de krigeage approchant les coefficients de projection sur cette base. On obtient ainsi un méta-modèle de la réponse vectorielle avec un nombre réduit de modèles de krigeage. On note qu'une telle approche a déjà été appliquée avec succès dans divers domaines tels que l'automobile [Xiao et al., 2010], l'environnement [Marrel et al., 2015] ou l'ingénierie de réservoir [Douarche et al., 2014].

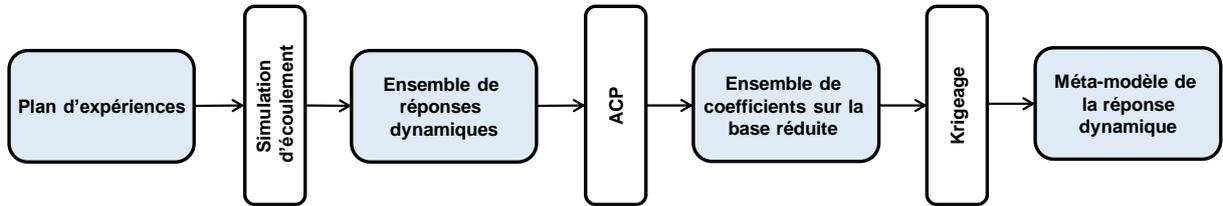


Figure 3.1 : Schéma de la méta-modélisation d'une sortie vectorielle en utilisant une analyse en composantes principales (ACP) dans un cadre simple fidélité.

Description de la méthode

On considère que l'on dispose de n simulations de la réponse dynamique $y(\mathbf{x}, t)$, $\mathbf{x} \in \mathbb{R}^d$ et $t \in [1, T]$ sur le plan d'expériences $\mathbf{D} = \{\mathbf{x}_1; \dots; \mathbf{x}_n\}$. On note $y(\mathbf{D}, t)$ cet ensemble de réponses simulées. Suivant l'Équation 3.3, on peut écrire l'approximation suivante de $y(\mathbf{x}, t)$ pour tout $\mathbf{x} \in \mathbf{D}$:

$$y(\mathbf{x}, t) \approx \bar{y}(t) + \sum_{l=1}^L \alpha_l(\mathbf{x}) \phi_l(t) \quad (3.4)$$

avec $\bar{y}(t)$ la réponse moyenne de $y(\mathbf{D}, t)$ au temps t , ϕ_l des fonctions formant une base orthonormale et α_l les coefficients de projection dans cette base. Le terme $\bar{y}(t)$ apparaît dans cette expression car il est nécessaire de centrer $y(\mathbf{D}, t)$ afin de pouvoir appliquer la décomposition de l'Équation (3.2). Une étape préliminaire consiste donc à centrer l'ensemble des simulations $y(\mathbf{D}, t)$ par rapport à la réponse moyenne \bar{y} . On note $(y(\mathbf{D}, t))_{cent}$ l'ensemble des simulations centrées. La base réduite ϕ_l est calculée en appliquant une décomposition en composantes principales à $(y(\mathbf{D}, t))_{cent}$. Les coefficients $\alpha_l(\mathbf{x})$ sont

ensuite obtenus par projection des simulations centrées $y(\mathbf{x}, t) - \bar{y}(t)$ pour chaque $\mathbf{x} \in \mathbf{D}$ sur la base $\phi_l : \alpha_l(\mathbf{x}) = \sum_{t=1}^T (y(\mathbf{x}, t) - \bar{y}(t)) \phi_l(t)$.

On remarque que, pour l donné, on peut associer à chaque point du plan d'expériences $\mathbf{x} \in \mathbf{D}$ une valeur du coefficient de projection $\alpha_l(\mathbf{x})$. On peut alors construire un modèle de krigeage approchant le coefficient α_l pour chaque $l \in [1, L]$. Nous notons $\hat{\alpha}_l$ les estimateurs de krigeage associés à ces modèles. Un estimateur de $y(\mathbf{x}, t)$ pour un point quelconque de l'espace des paramètres $\mathbf{x}^* \in \mathbb{R}^d$ est alors donné par :

$$\hat{y}(\mathbf{x}^*, t) = \bar{y}(t) + \sum_{l=1}^L \hat{\alpha}_l(\mathbf{x}^*) \phi_l(t) \quad (3.5)$$

Dans cette thèse, nous utilisons cet estimateur pour la construction d'un méta-modèle de sortie vectorielle dans un contexte de simple fidélité. On note que la méthode reste la même lorsque la sortie d'intérêt est spatiale (voir [Marrel et al., 2015], [Da Veiga and Gervais, 2015]).

En pratique, on observe que peu de termes de la décomposition en composantes principales suffisent à une bonne approximation, même lorsque la taille des sorties vectorielles est importante. On a alors $L \ll T$. Cette approche permet ainsi de réduire très fortement le nombre de modèles de krigeage nécessaires à la méta-modélisation d'une sortie vectorielle.

Algorithme

Nous récapitulons ci-dessous les étapes de l'algorithme de construction d'un méta-modèle d'une sortie vectorielle en utilisant l'analyse en composantes principales (ACP) et le krigeage, dans un contexte simple fidélité.

1. Centrer $y(\mathbf{D}, t)$ sur $\bar{\mathbf{y}}$. On note $(y(\mathbf{D}, t))_{cent}$ l'ensemble des réponses centrées.
2. Appliquer une ACP à $(y(\mathbf{D}, t))_{cent}$ pour calculer ϕ_l avec $l \in [1, L]$. On ne garde que les L premiers termes permettant d'expliquer 95% de la variance de $(y(\mathbf{D}, t))_{cent}$.
3. Pour $l = 1, \dots, L$ et $i = 1, \dots, n$: calculer $\alpha_l(\mathbf{x}_i) = \sum_{t=1}^T (y(\mathbf{x}_i, t) - \bar{y}(t)) \phi_l(t)$.
4. Pour $l = 1, \dots, L$: construire un modèle de krigeage $\hat{\alpha}_l$ de α_l à partir de $(\mathbf{x}_i, \alpha_l(\mathbf{x}_i))_{i=1, \dots, n}$.
5. Construire un méta-modèle de $y(\mathbf{x}, t)$ en utilisant les estimateurs $\hat{\alpha}_l$ selon l'Equation 3.5.

Variance et validation croisée

On peut étendre certains résultats du krigeage à la méthode proposée pour la méta-modélisation de sorties vectorielles. Tout d'abord, on remarque que la construction de modèles de krigeage pour les coefficients α_l permet d'approcher les variables Gaussiennes A_l introduites dans l'Equation 3.2. En effet, l'espérance de la variable A_l est donnée par l'estimateur du méta-modèle de α_l , et sa variance par la variance de ce méta-modèle, notée \hat{s}_l^2 .

A partir des \hat{s}_l^2 , on peut calculer la variance \hat{u}^2 associée à l'estimateur \hat{y} de la sortie vectorielle. Les variables Gaussiennes A_l de l'Equation 3.2 étant indépendantes entre elles, la somme des variances est équivalente à la variance de la somme. Il vient alors :

$$\hat{u}^2(\mathbf{x}^*, t) = \sum_{l=1}^L \hat{s}_l^2(\mathbf{x}^*) \phi_l(t)^2 \quad (3.6)$$

On peut également obtenir une expression de la prédiction \hat{y}_{-i} et de la variance \hat{u}_{-i}^2 obtenues par validation croisée en ne considérant pas l'observation \mathbf{x}_i dans le calcul de l'estimateur \hat{y} . En effet, on peut utiliser les résultats de validation croisée des modèles de krigeage de α_l , notés $\hat{\alpha}_{-i,l}$ et $\hat{s}_{-i,l}^2$, pour remplacer $\hat{\alpha}_l$ et \hat{s}_l^2 dans les Equations 3.5 et 3.6, respectivement :

$$\hat{y}_{-i}(\mathbf{x}^*, t) = \bar{y}(t) + \sum_{l=1}^L \hat{\alpha}_{-i,l}(\mathbf{x}^*) \phi_l(t) \quad (3.7)$$

et

$$\hat{u}_{-i}^2(\mathbf{x}^*, t) = \sum_{l=1}^L \hat{s}_{-i,l}^2(\mathbf{x}^*) \phi_l(t)^2 \quad (3.8)$$

3.2.3 ACP et co-krigeage multi-fidélité

Pour la méta-modélisation d'une sortie vectorielle dans un contexte multi-fidélité, nous proposons une extension de la méthode décrite ci-dessus. Elle consiste à combiner l'expression tronquée de la transformation de Karhunen-Loève (Equation 3.3) au co-krigeage multi-fidélité. Cette approche est présentée ici pour deux niveaux de fidélité, mais se généralise à un nombre quelconque de niveaux.

Comme dans le cas simple fidélité, on peut réaliser une ACP sur un ensemble de réponses du niveau fin. On obtient alors une base réduite au niveau fin, ainsi que les coefficients de projection associés à cette base. On propose ensuite de projeter les réponses du niveau grossier sur la base réduite issue du niveau fin. On obtient ainsi des coefficients de projection sur les deux niveaux, dont on peut déduire des estimateurs par co-krigeage multi-fidélité. On se ramène ensuite à l'estimateur (Equation 3.5) dans lequel les estimateurs simple fidélité des coefficients de projection sont remplacés par leurs estimateurs multi-fidélité. Cette méthode peut être résumée par la Figure 3.2.

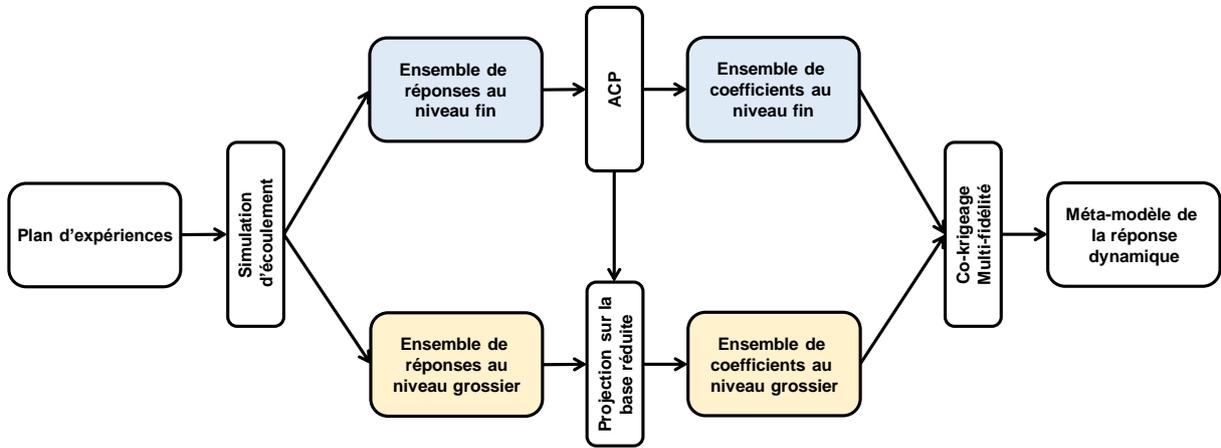


Figure 3.2 : Schéma de la méta-modélisation d'une sortie vectorielle en utilisant une analyse en composantes principales (ACP) dans un cadre multi-fidélité à deux niveaux.

Description de la méthode

On note $\mathbf{D}_f = \{\mathbf{x}_1; \dots; \mathbf{x}_{n_f}\}$ et $\mathbf{D}_g = \{\mathbf{x}_1; \dots; \mathbf{x}_{n_g}\}$ deux plans d'expériences imbriqués. On note $y_f(\mathbf{x}, t)$ pour $\mathbf{x} \in \mathbb{R}^d$ et $t \in [1, T]$ la réponse dynamique d'intérêt simulée sur le niveau fin, et $y_g(\mathbf{x}, t)$ la même propriété simulée sur le niveau grossier. On suppose que l'on connaît l'ensemble des réponses simulées sur les plans d'expériences fin et grossier, que l'on note respectivement $y_f(\mathbf{D}_f, t)$ et $y_g(\mathbf{D}_g, t)$.

On note également $\bar{y}_f(t)$ la réponse moyenne de $y_f(\mathbf{D}_f, t)$ au temps t .

En reprenant l'Équation 3.4, on trouve l'approximation suivante de $y_f(\mathbf{x}, t)$ pour tout $\mathbf{x} \in \mathbf{D}_f$:

$$y_f(\mathbf{x}, t) \approx \bar{y}_f(t) + \sum_{l=1}^L \alpha_l^f(\mathbf{x}) \phi_l^f(t) \quad (3.9)$$

Comme précédemment, la base ϕ_l^f est calculée en appliquant une décomposition en composantes principales sur l'ensemble des réponses simulées sur le niveau fin $y_f(\mathbf{D}_f, t)$ et centrées sur \bar{y}_f . Puis, on calcule les coefficients α_l^f en projetant les simulations centrées du niveau fin $y_f(\mathbf{x}, t) - \bar{y}_f(t)$ pour chaque $\mathbf{x} \in \mathbf{D}_f$ sur la base ϕ_l^f : $\alpha_l^f(\mathbf{x}) = \sum_{t=1}^T [y_f(\mathbf{x}, t) - \bar{y}_f(t)] \phi_l^f(t)$.

On propose ensuite de calculer les coefficients de projection sur le niveau grossier. Pour cela, on projette l'ensemble des réponses simulées au niveau grossier $y_g(\mathbf{D}_g, t)$ sur la base ϕ_l^f : pour tout $\mathbf{x} \in \mathbf{D}_g$, $\alpha_l^g(\mathbf{x}) = \sum_{t=1}^T y_g(\mathbf{x}, t) \phi_l^f(t)$. Ainsi, pour chacune des valeurs de l , on dispose de $\alpha_l^f(\mathbf{D}_f, t)$ et $\alpha_l^g(\mathbf{D}_g, t)$. On peut alors construire un méta-modèle approchant α_l^f par co-krigeage multi-fidélité pour tout $l \in [1, L]$. On note $\hat{\alpha}_l^f$ les estimateurs de co-krigeage multi-fidélité associés à ces modèles. On obtient ainsi un estimateur pour $y_f(\mathbf{x}, t)$ similaire à celui introduit dans le cas simple fidélité (Équation 3.5), pour $\mathbf{x}^* \in \mathbb{R}^d$:

$$\hat{y}_f(\mathbf{x}^*, t) = \bar{y}_f(t) + \sum_{l=1}^L \hat{\alpha}_l^f(\mathbf{x}^*) \phi_l^f(t) \quad (3.10)$$

Dans la méthode proposée, les coefficients α_l^g sont issus d'une projection des réponses du niveau grossier sur la base réduite calculée par l'ACP appliquée aux réponses du niveau fin. Toutefois, ces coefficients ne sont pas adaptés à l'approximation de y_g , contrairement aux coefficients α_l^f qui peuvent être utilisés pour approcher y_f (cf. Équation 3.9). On peut cependant définir \tilde{y}_g à partir des coefficients α_l^g comme :

$$\tilde{y}_g(\mathbf{x}^*, t) = \bar{y}_f(t) + \sum_{l=1}^L \hat{\alpha}_l^g(\mathbf{x}^*) \phi_l^f(t) \quad (3.11)$$

avec $\hat{\alpha}_l^g$ les estimateurs de krigeage approchant α_l^g pour tout $l \in [1, L]$. \tilde{y}_g peut être vu comme un estimateur de y_f sur le niveau grossier, mais ce n'est pas un estimateur de y_g .

Algorithme

Nous récapitulons ci-dessous les étapes de l'algorithme de construction d'un méta-modèle d'une sortie vectorielle en utilisant l'analyse en composantes principales (ACP) et le co-krigeage, dans un contexte multi-fidélité à deux niveaux.

1. Centrer $y_f(\mathbf{D}_f, t)$ sur \bar{y}_f . On note $(y_f(\mathbf{D}_f, t))_{cent}$ l'ensemble des réponses centrées.
2. Appliquer une ACP à $(y_f(\mathbf{D}_f, t))_{cent}$ pour calculer ϕ_l^f avec $l \in [1, L]$. On ne garde que les L premiers termes permettant d'expliquer 95% de la variance de $(y_f(\mathbf{D}_f, t))_{cent}$.
3. Pour $l = 1, \dots, L$ et $i = 1, \dots, n$: calculer $\alpha_l^f(\mathbf{x}_i) = \sum_{t=1}^T (y_f(\mathbf{x}_i, t) - \bar{y}_f(t)) \phi_l^f(t)$.
4. Pour $l = 1, \dots, L$ et $i = 1, \dots, n$: calculer $\alpha_l^g(\mathbf{x}_i) = \sum_{t=1}^T y_g(\mathbf{x}_i, t) \phi_l^f(t)$.

5. Pour $l = 1, \dots, L$: construire un modèle de co-krigeage multi-fidélité de α_l^f à partir de $(\mathbf{x}_i, \alpha_l^f(\mathbf{x}_i))_{\mathbf{x}_i \in \mathbf{D}_f}$ et $(\mathbf{x}_i, \alpha_l^g(\mathbf{x}_i))_{\mathbf{x}_i \in \mathbf{D}_g}$.
6. Construire un méta-modèle de $y_f(\mathbf{x}, t)$ en utilisant les estimateurs $\hat{\alpha}_l^f$ selon l'Equation 3.10.

Remarques

Lors du calcul des coefficients α_l^g , les réponses du niveau grossier n'ont pas besoin d'être préalablement centrées avant leur projection sur la base ϕ_l^f . En effet, soustraire la réponse moyenne du niveau grossier (ou bien celle du niveau fin) à $y_g(\mathbf{x}, t)$ pour $\mathbf{x} \in \mathbf{D}_g$ se traduit par l'ajout d'une constante sur les coefficients α_l^g . Plus précisément, on obtient $\alpha_l^g(x) + c_l$ pour tout $\mathbf{x} \in \mathbf{D}_g$ avec c_l une constante dépendant de l . Cette constante est donc identique pour un ensemble $\alpha_l^g(\mathbf{D}_g)$ donné. Cela change alors la réponse moyenne du niveau grossier lors du calcul du modèle de co-krigeage multi-fidélité, mais l'estimateur en résultant n'est pas modifié.

On note également que, pour appliquer la méthode proposée, il est nécessaire que les réponses $y_f(\mathbf{x}, t)$ et $y_g(\mathbf{x}, t)$ soient définies pour les mêmes temps ou positions. Prenons l'exemple d'un méta-modèle approchant des cartes de pression et plaçons-nous dans un cadre multi-fidélité défini par un changement de résolution de la grille réservoir. Il est alors nécessaire, pour pouvoir calculer les coefficients α_l^g , de raffiner les cartes de pression simulées avec le modèle de réservoir grossier afin que leur résolution corresponde à celle des cartes simulées avec le modèle de réservoir fin. On remarque que le raffinement concerne la résolution de la grille, il n'implique pas nécessairement une mise à l'échelle des propriétés (cf. section 4.2.3).

Pour finir, la généralisation de cette approche à un nombre quelconque de niveaux de fidélité est directe. Il suffit de projeter les réponses des différents niveaux grossiers sur la base réduite calculée au niveau fin pour obtenir les coefficients de projection sur les différents niveaux grossiers. Toutefois, les performances et limites de cette méthode pour plus de deux niveaux de fidélité n'ont pas été étudiées au cours de cette thèse.

Variance et validation croisée

Comme présenté précédemment dans le cadre simple fidélité, on peut étendre certains résultats du co-krigeage multi-fidélité à la méta-modélisation proposée pour les sorties vectorielles.

En notant respectivement $\hat{s}_{f,l}^2$ et $\hat{s}_{g,l}^2$ les variances sur les niveaux fin et grossier associées aux modèles de co-krigeage multi-fidélité des coefficients α_l^f , on a l'expression suivante pour les variances \hat{u}_f^2 et \hat{u}_g^2 associées aux estimateurs \hat{y}_f et \tilde{y}_g , respectivement :

$$\hat{u}_f^2(\mathbf{x}^*, t) = \sum_{l=1}^L \hat{s}_{f,l}^2(\mathbf{x}^*) \left[\phi_l^f(t) \right]^2 \quad (3.12)$$

et

$$\hat{u}_g^2(\mathbf{x}^*, t) = \sum_{l=1}^L \hat{s}_{g,l}^2(\mathbf{x}^*) \left[\phi_l^f(t) \right]^2 \quad (3.13)$$

On peut également obtenir des expressions des prédictions et variances de validation croisée pour les deux niveaux, à partir des résultats de validation croisée des différents modèles de co-krigeage multi-fidélité approchant α_l^f . On rappelle que les résultats de validation croisée sur le niveau grossier

s'obtiennent en considérant que l'observation $\mathbf{x}_i \in \mathbf{D}_g$ est inconnue. Pour le niveau fin, on considère que l'observation $\mathbf{x}_i \in \mathbf{D}_f$ est inconnue sur les deux niveaux de fidélité.

Sur le niveau fin, la prédiction de \hat{y}_f par validation croisée, notée \hat{y}_{-i}^f , s'obtient en remplaçant $\hat{\alpha}_i^f$ par la prédiction $\hat{\alpha}_{i,-i}^f$ dans l'Equation 3.10. La variance de validation croisée associée, notée $\hat{u}_{f,-i}^2$, s'obtient en remplaçant $\hat{s}_{f,l}^2$ par $\hat{s}_{f,l,-i}^2$ dans l'Equation 3.12.

Sur le niveau grossier, la prédiction de \tilde{y}_g par validation croisée, notée $\tilde{y}_{g,-i}$, s'obtient en remplaçant $\hat{\alpha}_i^g$ par la prédiction $\hat{\alpha}_{i,-i}^g$ dans l'Equation 3.11. La variance de validation croisée associée, notée $\tilde{u}_{g,-i}^2$, s'obtient en remplaçant $\hat{s}_{g,l}^2$ par $\hat{s}_{g,l,-i}^2$ dans l'Equation 3.13.

3.3 La méta-modélisation de la fonction objectif

Dans cette partie, nous nous intéressons à la méta-modélisation de la fonction objectif. Disposer d'un méta-modèle prédictif de la fonction objectif présente de nombreux avantages pour le calage d'historique. On peut s'en servir pour éliminer les paramètres non-influents en réalisant une analyse de sensibilité, pour adopter une approche probabiliste pour le calage, ou encore pour réduire le nombre de simulations nécessaires à la minimisation de la fonction objectif et ainsi accélérer le calage.

Le plus souvent, la méta-modélisation est directe, c'est-à-dire que le méta-modèle est construit en utilisant un ensemble de valeurs de la fonction objectif calculées à partir de résultats de simulations. Nous suggérons ici d'approcher la fonction objectif à partir des méta-modèles construits pour les sorties dynamiques intervenant dans son calcul. Ainsi, plutôt que de considérer un unique modèle de krigeage/co-krigeage multi-fidélité pour approcher cette fonction complexe, on considère un nombre important de méta-modèles approchant des propriétés moins complexes. On espère ainsi améliorer la méta-modélisation de la fonction objectif et réduire le nombre d'évaluations nécessaire à l'obtention d'un méta-modèle prédictif.

Toutes les sections de cette partie traitent successivement d'une application dans un cadre simple fidélité, puis multi-fidélité. On commence par présenter en section 3.3.1 l'approche classique pour la méta-modélisation de la fonction objectif. Ensuite, en section 3.3.2, nous décrivons en détail l'approche alternative que nous proposons pour la méta-modélisation de la fonction objectif. L'approche proposée admet également une expression de la variance sur l'estimateur de la fonction objectif, ainsi que des prédictions et variances issues de la validation croisée. Ces expressions sont importantes dans le cadre d'une planification séquentielle d'expériences (*cf.* Chapitre 5). Elles sont détaillées dans la section 3.3.3.

3.3.1 Méta-modélisation directe de la fonction objectif

Méta-modélisation directe de la fonction objectif en simple fidélité

La fonction objectif est une fonction scalaire dépendant des paramètres incertains $\mathbf{x} \in \mathbb{R}^d$. Considérons que l'on dispose de n simulations d'écoulement du modèle réservoir aux n points du plan d'expériences \mathbf{D} . On peut alors calculer les n valeurs correspondantes $FO(\mathbf{D})$ à partir de l'ensemble des réponses dynamiques simulées $y_k(\mathbf{D}, t)$ pour chaque propriété $k \in [1, K]$. Il est ensuite possible d'obtenir un méta-modèle de la fonction objectif par krigeage en utilisant $(\mathbf{x}, FO(\mathbf{x}))_{\mathbf{x} \in \mathbf{D}}$. On note FO^{dir} un tel méta-modèle. La notation *dir* indique que le méta-modèle est direct au sens où on le construit directement pour la fonction objectif.

Méta-modélisation directe de la fonction objectif en multi-fidélité

On se place maintenant dans le cadre multi-fidélité à deux niveaux, et on se donne deux plans d'expériences imbriqués $\mathbf{D}_f \subset \mathbf{D}_g$ pour lesquels les simulations d'écoulement sont réalisées sur les niveaux fin et grossier, respectivement. Il est alors possible de calculer la valeur de la fonction objectif pour les simulations au niveau fin (notée FO^f), la valeur de la fonction objectif pour les simulations au niveau grossier (notée FO^g), et d'en déduire un méta-modèle de co-krigeage multi-fidélité $FO^{dir.f}$ de la fonction objectif au niveau fin. On note que la construction de ce méta-modèle dépend de $FO^{dir.g}$, le modèle de krigeage construit à partir des valeurs de FO^g sur \mathbf{D}_g . On remarque aussi que la généralisation de cette approche pour un nombre quelconque de niveaux de fidélité est immédiate.

3.3.2 Méta-modélisation vectorielle de la fonction objectif - méthode

Nous proposons dans cette section une méthode alternative pour la construction d'un méta-modèle de la fonction objectif. Elle consiste à substituer les données de production simulées $y_t(\mathbf{x})$ dans l'Équation 3.1 par un méta-modèle les approchant.

Description de la méthode dans un cadre simple fidélité

On peut définir pour chaque propriété $k \in [1, K]$ la fonction objectif FO_k associée :

$$FO_k(\mathbf{x}) = \frac{1}{2} \sum_{t=1}^{T(k)} \left(\frac{y_k(\mathbf{x}, t) - d_{kt}^{obs}}{\sigma_{kt}} \right)^2 \quad (3.14)$$

Il est alors possible de construire un méta-modèle de FO_k en remplaçant la propriété simulée $y_k(\mathbf{x}, t)$ par le méta-modèle simple fidélité (Équation 3.5) combinant ACP et krigeage. L'estimateur \widehat{FO}_k^{vec} associé s'écrit alors :

$$\widehat{FO}_k^{vec}(\mathbf{x}) = \frac{1}{2} \sum_{t=1}^{T(k)} \left(\frac{\bar{y}_k(t) + \sum_{l=1}^{L(k)} \hat{\alpha}_{kl}(\mathbf{x}) \phi_{kl}(t) - d_{kt}^{obs}}{\sigma_{kt}} \right)^2 \quad (3.15)$$

En répétant cette procédure pour l'ensemble des propriétés, on obtient un méta-modèle FO^{vec} approchant la fonction objectif et ayant pour estimateur :

$$\begin{aligned} \widehat{FO}^{vec}(\mathbf{x}) &= \sum_{k=1}^K \omega_k \widehat{FO}_k^{vec}(\mathbf{x}) \\ &= \frac{1}{2} \sum_{k=1}^K \omega_k \sum_{t=1}^{T(k)} \left(\frac{\bar{y}_k(t) + \sum_{l=1}^{L(k)} \hat{\alpha}_{kl}(\mathbf{x}) \phi_{kl}(t) - d_{kt}^{obs}}{\sigma_{kt}} \right)^2 \end{aligned} \quad (3.16)$$

Extension au cadre multi-fidélité

De la même manière, il est possible d'estimer les fonctions objectifs FO_k au niveau fin à partir des méta-modèles multi-fidélité des propriétés correspondantes. On obtient alors les estimateurs suivants :

$$\widehat{FO}_k^{vec,f}(\mathbf{x}) = \frac{1}{2} \sum_{t=1}^{T(k)} \left(\frac{\bar{y}_k^f(t) + \sum_{l=1}^{L(k)} \hat{\alpha}_{kl}^f(\mathbf{x}) \phi_{kl}^f(t) - d_{kt}^{obs}}{\sigma_{kt}} \right)^2 \quad (3.17)$$

On en déduit un méta-modèle multi-fidélité $FO^{vec,f}$ de la fonction objectif au niveau fin dont l'estimateur est donné par :

$$\widehat{FO}^{vec,f}(\mathbf{x}) = \frac{1}{2} \sum_{k=1}^K \omega_k \sum_{t=1}^{T(k)} \left(\frac{\bar{y}_k^f(t) + \sum_{l=1}^{L(k)} \hat{\alpha}_{kl}^f(\mathbf{x}) \phi_{kl}^f(t) - d_{kt}^{obs}}{\sigma_{kt}} \right)^2 \quad (3.18)$$

On peut également définir un méta-modèle $FO^{vec,g}$ de la fonction objectif au niveau grossier en utilisant l'estimateur \tilde{y}_g (Equation 3.11) pour approcher chaque sortie y_k^f . L'estimateur associé à ce méta-modèle s'écrit :

$$\widehat{FO}^{vec,g}(\mathbf{x}) = \frac{1}{2} \sum_{k=1}^K \omega_k \sum_{t=1}^{T(k)} \left(\frac{\bar{y}_k^f(t) + \sum_{l=1}^{L(k)} \hat{\alpha}_{kl}^g(\mathbf{x}) \phi_{kl}^f(t) - d_{kt}^{obs}}{\sigma_{kt}} \right)^2 \quad (3.19)$$

On remarque que $\widehat{FO}^{vec,f}$ et $\widehat{FO}^{dir,f}$ sont tous les deux des estimateurs de FO^f , la fonction objectif calculée avec les réponses du niveau fin. En revanche, contrairement à $\widehat{FO}^{dir,g}$, $\widehat{FO}^{vec,g}$ n'est pas un estimateur de FO^g . C'est un estimateur de FO^f sur le niveau grossier défini par les coefficients α_l^g .

3.3.3 Méta-modélisation vectorielle de la fonction objectif - variance et validation croisée

Nous présentons dans cette section les expressions de la variance et les résultats de validation croisée pour la méta-modélisation vectorielle de la fonction objectif. Ces différents résultats sont importants dans le cadre d'une utilisation de cette approche au sein des méthodes de planification séquentielle d'expériences que nous proposons plus loin dans le Chapitre 5. Les notations employées sont semblables à celles utilisées en section 3.2.2 pour la version simple fidélité ou en section 3.2.3 pour la version multi-fidélité.

La méta-modélisation proposée pour la fonction objectif en simple fidélité repose sur la construction de modèles de krigeage approchant les coefficients α_{kl} pour tout $k \in [1, K]$ et $l \in [1, L]$ (Équation 3.16). Ces modèles de krigeage donnent l'expression d'un estimateur de krigeage $\hat{\alpha}_{kl}$ ainsi que de la variance \hat{s}_{kl}^2 associée à cet estimateur. Par ailleurs, on peut également calculer les prédictions issues de la validation croisée lorsque l'observation $\mathbf{x}_i \in \mathbf{D}$ est supposée inconnue. On note ainsi $\hat{\alpha}_{kl,-i}$ l'estimateur de validation croisée et $\hat{s}_{kl,-i}$ la variance associée. A partir de l'ensemble des estimateurs $\hat{\alpha}_{kl}$, \hat{s}_{kl}^2 , $\hat{\alpha}_{kl,-i}$ et $\hat{s}_{kl,-i}$, on peut exprimer la variance associée à l'estimateur \widehat{FO}^{vec} , ainsi que des prédictions de la fonction objectif par validation croisée.

Le raisonnement est identique en multi-fidélité. En effet, la méta-modélisation de la fonction objectif repose alors sur la construction de modèles de co-krigeage multi-fidélité approchant les coefficients α_{kl} . On note $\hat{\alpha}_{kl}^f$ et $\hat{\alpha}_{kl}^g$ les estimateurs de ces modèles ainsi que $\hat{s}_{f,kl}^2$ et $\hat{s}_{g,kl}^2$ les variances associées pour les niveaux fin et grossier, respectivement. On note également $\hat{\alpha}_{kl,-i}^f$ et $\hat{\alpha}_{kl,-i}^g$ les estimateurs de validation croisée, et $\hat{s}_{f,kl,-i}^2$, $\hat{s}_{g,kl,-i}^2$ les variances associées. Par ailleurs, on rappelle que les résultats de validation croisée sur le niveau grossier s'obtiennent en considérant que l'observation $\mathbf{x}_i \in \mathbf{D}_g$ est inconnue. Pour le niveau fin, on considère que l'observation $\mathbf{x}_i \in \mathbf{D}_f$ est inconnue sur les deux niveaux de fidélité.

Expression de la variance en simple fidélité

Pour pouvoir calculer une expression de la variance associée à l'estimateur \widehat{FO}^{vec} (Équation 3.16), on fait l'hypothèse que les méta-modèles FO_k^{vec} de la fonction objectif pour une propriété donnée $k \in [1, K]$ sont indépendants. On suppose donc que la fonction objectif pour une propriété donnée n'est pas corrélée avec la fonction objectif pour une autre propriété. La variance de la somme est alors la somme des variances et on peut écrire :

$$\text{Var}(FO^{vec}(\mathbf{x})) = \text{Var}\left(\frac{1}{2} \sum_{k=1}^K \omega_k FO_k^{vec}(\mathbf{x})\right) = \frac{1}{4} \sum_{k=1}^K \omega_k^2 \text{Var}(FO_k^{vec}(\mathbf{x})) \quad (3.20)$$

En pratique, une telle hypothèse est fautive car il est vraisemblable que l'on observe des corrélations entre certaines fonctions FO_k . Toutefois, ces corrélations sont inconnues, dépendantes du cas d'étude, et souvent négligeables. On peut donc considérer que la somme des variances de FO_k^{vec} pour toutes les propriétés $k \in [1, K]$ est une approximation raisonnable de la variance de FO^{vec} . On obtient finalement :

$$\text{Var}(FO_k^{vec}(\mathbf{x})) = 2 \sum_{i=1}^{T(k)} \sum_{j=1}^{T(k)} \frac{\sum_{l=1}^{L(k)} \phi_{kli} \phi_{klj} \hat{s}_{kl}^2(\mathbf{x})}{\sigma_{ki} \sigma_{kj}} \left(\sum_{l=1}^{L(k)} \phi_{kli} \phi_{klj} \hat{s}_{kl}^2(\mathbf{x}) + 2\mu_{ki} \mu_{kj} \right) \quad (3.21)$$

avec

$$\mu_{ki} = \left(\bar{\mathbf{y}}_{ki} + \sum_{l=1}^{L(k)} \phi_{kli} \hat{\alpha}_{kl} - d_{kt}^{obs} \right) \quad (3.22)$$

Expression de la variance en multi-fidélité

Comme précédemment, l'expression de la variance de $FO^{vec,f}$ est prise égale à la somme pondérée des variances de $FO_k^{vec,f}$ pour toutes les propriétés $k \in [1, K]$. On peut alors montrer que :

$$\text{Var}(FO_k^{vec,f}(\mathbf{x})) = 2 \sum_{i=1}^{T(k)} \sum_{j=1}^{T(k)} \frac{\sum_{l=1}^{L(k)} \phi_{kli}^f \phi_{klj}^f \hat{s}_{f,kl}^2(\mathbf{x})}{\sigma_{ki} \sigma_{kj}} \left(\sum_{l=1}^{L(k)} \phi_{kli}^f \phi_{klj}^f \hat{s}_{f,kl}^2(\mathbf{x}) + 2\mu_{ki}^f \mu_{kj}^f \right) \quad (3.23)$$

avec

$$\mu_{ki}^f = \left(\bar{\mathbf{y}}_{ki}^f + \sum_{l=1}^{L(k)} \phi_{kli}^f \hat{\alpha}_{kl}^f - d_{kt}^{obs} \right) \quad (3.24)$$

De façon similaire, la variance de $FO^{vec,g}$ peut être calculée par sommation des variances de $FO_k^{vec,g}$:

$$\text{Var}(FO_k^{vec,g}(\mathbf{x})) = 2 \sum_{i=1}^{T(k)} \sum_{j=1}^{T(k)} \frac{\sum_{l=1}^{L(k)} \phi_{kli}^g \phi_{klj}^g \hat{s}_{g,kl}^2(\mathbf{x})}{\sigma_{ki} \sigma_{kj}} \left(\sum_{l=1}^{L(k)} \phi_{kli}^g \phi_{klj}^g \hat{s}_{g,kl}^2(\mathbf{x}) + 2\mu_{ki}^g \mu_{kj}^g \right) \quad (3.25)$$

avec

$$\mu_{ki}^g = \left(\bar{y}_{ki} + \sum_{l=1}^{L(k)} \phi_{kli}^f \hat{\alpha}_{kl}^g - d_{kt}^{obs} \right) \quad (3.26)$$

Validation croisée en simple fidélité et multi-fidélité

La prédiction par validation croisée associée à FO^{vec} est notée \widehat{FO}_{-i}^{vec} . Son expression est obtenue en remplaçant $\hat{\alpha}_{kl}$ par $\hat{\alpha}_{kl,-i}$ dans l'Equation 3.16. La variance correspondant à cette prédiction est notée $\text{Var}(FO_{-i}^{vec})$. Elle est obtenue en remplaçant \hat{s}_{kl}^2 par $\hat{s}_{kl,-i}^2$ et $\hat{\alpha}_{kl}$ par $\hat{\alpha}_{kl,-i}$ dans l'Équation 3.21.

En multi-fidélité, la prédiction par validation croisée associée à $FO^{vec,f}$ est notée $\widehat{FO}_{-i}^{f,vec}$. Son expression est obtenue en remplaçant $\hat{\alpha}_{kl}^f$ par $\hat{\alpha}_{kl,-i}^f$ dans l'équation 3.18. La variance correspondant à cette prédiction est notée $\text{Var}(FO_{-i}^{f,vec})$. Elle est obtenue en remplaçant $\hat{s}_{f,kl}^2$ par $\hat{s}_{f,kl,-i}^2$ et $\hat{\alpha}_{kl}^f$ par $\hat{\alpha}_{kl,-i}^f$ dans l'Équation 3.23.

Finalement, la prédiction par validation croisée associée à $FO^{vec,g}$ est notée $\widehat{FO}_{-i}^{g,vec}$. Son expression est obtenue en remplaçant $\hat{\alpha}_{kl}^g$ par $\hat{\alpha}_{kl,-i}^g$ dans l'équation 3.19. La variance correspondant à cette prédiction est notée $\text{Var}(FO_{-i}^{g,vec})$. Elle est obtenue en remplaçant $\hat{s}_{g,kl}^2$ par $\hat{s}_{g,kl,-i}^2$ et $\hat{\alpha}_{kl}^g$ par $\hat{\alpha}_{kl,-i}^g$ dans l'Équation 3.25.

CHAPITRE 4

Etude de cas : PUNQ

Ce chapitre a fait l'objet d'un article publié dans Computational Geosciences [Thenon et al., 2016].

Afin de tester les méthodologies proposées pour la méta-modélisation dans des cadres simple et multi-fidélité, un cas multi-fidélité à deux niveaux inspiré du « benchmark » PUNQ-S3 [Floris et al., 2001] a été mis en place. Plus précisément, il est constitué de deux modèles réservoir qui diffèrent par la résolution de leur maillage. Le but de l'étude menée sur ce cas est la comparaison des approches simple et multi-fidélité pour la construction de méta-modèles. Les différentes méthodes pour la méta-modélisation (Chapitre 2), notamment pour l'approximation de sorties vectorielles (Chapitre 3), ont été appliquées à différentes sorties.

La partie 4.1 de ce chapitre décrit le cas d'étude, puis la procédure mise en place pour comparer les méta-modèles construits en simple et multi-fidélité. On s'intéresse en particulier à la prédictivité des méta-modèles par rapport au temps de calcul requis pour leur construction. Les résultats pour la méta-modélisation de sorties vectorielles (propriétés du réservoir, aux puits et spatiales) sont regroupés et discutés dans la partie 4.2. Les résultats concernant la méta-modélisation de la fonction objectif sont ensuite présentés et discutés dans la partie 4.3. Une application au calage d'historique est finalement présentée dans la partie 4.4.

4.1 Description du cas PUNQ multi-fidélité

Nous appelons PUNQ ou PUNQ multi-fidélité ce cas d'étude. Ce cas s'inspire de PUNQ-S3, un cas référence en ingénierie de réservoir présenté dans la section 4.1.1. Le cas PUNQ est constitué de deux modèles réservoir : PUNQ fin et PUNQ grossier. La construction de ces deux modèles est détaillée dans la section 4.1.2. La section 4.1.3 est ensuite dédiée au contrôle de ce cas d'étude à travers une comparaison des réponses dynamiques de PUNQ fin et grossier. Les différentes expériences menées sur le cas PUNQ sont décrites dans la section 4.1.4, ainsi que les paramètres considérés comme incertains.

4.1.1 PUNQ-S3 : un cas référence

Le cas d'étude PUNQ s'inspire du cas PUNQ-S3, un modèle synthétique construit à partir d'un vrai champ exploité par Elf Aquitaine. Ce cas est souvent utilisé dans la littérature comme cas d'application pour tester des méthodologies pour le calage d'historique. Il a été défini dans le cadre d'un projet européen visant à quantifier les incertitudes sur les prévisions de production [Floris et al., 2001]. L'ensemble des données nécessaires à la construction du cas PUNQ-S3 est disponible à l'adresse suivante: <http://www.imperial.ac.uk/earth-science/research/research-groups/perm/standard-models/>.

Structure du modèle

Le modèle PUNQ-S3 est défini sur une grille constituée de $19 \times 28 \times 5$ mailles, dont 1761 sont actives. Ces mailles mesurent 180 mètres selon les axes X et Y et sont d'épaisseur variable (~ 5 mètres). Le champ est délimité à l'est et au sud par une faille imperméable et il est connecté au nord et à l'ouest à un aquifère. La structure de son toit est visible sur la Figure 4.1. Il s'agit d'un dôme dont la partie supérieure est initialement remplie de gaz.

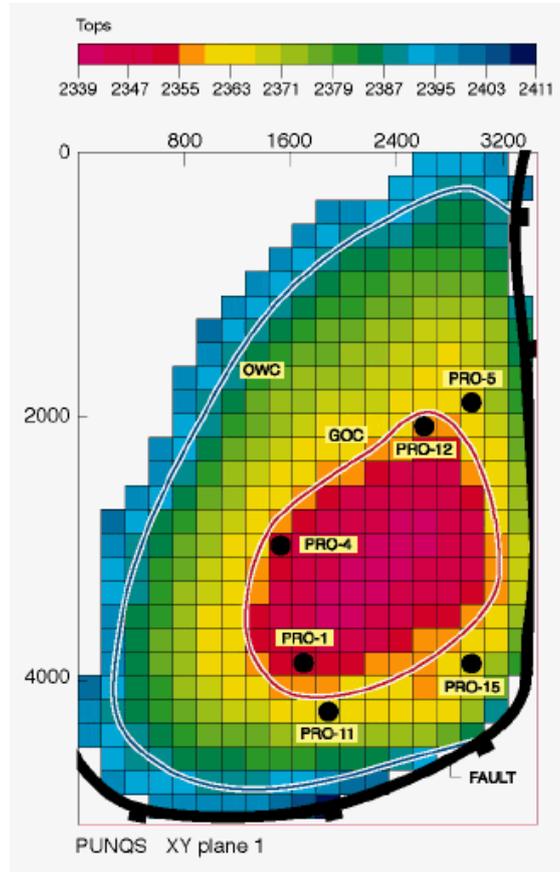


Figure 4.1 : Carte du toit et position des puits (figure adaptée de [Florin et al., 2001]).

Géologie

Le réservoir est composé de 5 couches correspondant à des sédiments issus d'un environnement de dépôt deltaïque en plaine côtière. Les couches 1, 3 et 5 sont constituées de chenaux fluviaux au milieu de mudstones venant du lit majeur. La couche 2 correspond à de l'argile d'origine marine ou lagunaire contenant quelques dépôts distaux de barres d'embouchure. La couche 4 est faite de sédiments lagunaires du delta ou de barres d'embouchure enrobées d'argile lagunaire.

Les couches 1, 3 et 5 contiennent des bancs de sable très poreux ($\phi > 20\%$), orientés entre 110° et 170° Sud-Est. Ces bancs d'environ 800 mètres de long sont entourés d'une matrice argileuse très peu poreuse ($\phi < 5\%$). Leur longueur et espacement varient selon les couches comme précisé dans le Tableau 4.1.

Dans la couche 2 se trouve de l'argile marine ou lagunaire contenant quelques dépôts de barres d'embouchure distales. Cela se traduit par quelques touches irrégulières de sédiments peu poreux ($\phi > 5\%$) parmi des sédiments argileux très peu poreux ($\phi < 5\%$).

La couche 4 contient des sédiments de barres d'embouchure ou du delta lagunaire dans une matrice d'argile lagunaire. On peut s'attendre à trouver des zones d'écoulement de formes lobées, avec une porosité intermédiaire ($\phi \sim 15\%$) au sein d'une matrice très peu poreuse ($\phi < 5\%$). Ces formes lobées peuvent être modélisées par des ellipses dont l'axe le plus long est perpendiculaire au paléocourant (qui est orienté 110° et 170° Sud-Est).

Couche	ϕ moyen (%)	Faciès	Longueur (m)	Espacement (km)
1	14	Chenal fluviatile	800	2-5
2	8	Argile lagunaire	-	-
3	14	Chenal fluviatile	1000	2-5
4	10	Barre d'embouchure	500-5000	10
5	14	Chenal fluviatile	2000	4-10

Tableau 4.1 : La porosité moyenne, le faciès sédimentaire et les estimations de la longueur et de l'espacement des principales zones réservoir par couche de PUNQ-S3.

Exploitation

Le réservoir est produit par 6 puits producteurs localisés autour du contact gaz-huile : PRO-1, PRO-4, PRO-5, PRO-11, PRO-12 et PRO-15 (*cf.* Figure 4.1). Le modèle ne comporte pas de puits injecteur car l'aquifère qui le borde exerce une forte pression sur le réservoir. Le planning de production original [Floris et al., 2001] s'étend sur 8 années. Il est constitué d'une première année de mise en production, suivie de 3 ans d'arrêt, puis de plusieurs années de production interrompues annuellement pendant 2 semaines pour des tests. Des objectifs de production d'huile communs aux 6 puits sont définis. Ils sont de 100, 200, 100 puis $50 \text{ m}^3/\text{jour}$ lors de la première année, puis de $150 \text{ m}^3/\text{jour}$ le reste du temps. Cependant, si la pression au fond du puits atteint la pression limite de 120 bar, l'objectif de production est annulé pour être remplacé par une production à la pression constante de 120 bar.

Nous utilisons par la suite un planning de production similaire étendu à 16,5 années. L'unique différence avec le planning original est que l'objectif de production par puits, en dehors de la phase de test, est ramené à $100 \text{ m}^3/\text{jour}$. Ce planning de production est résumé sur la Figure 4.2.

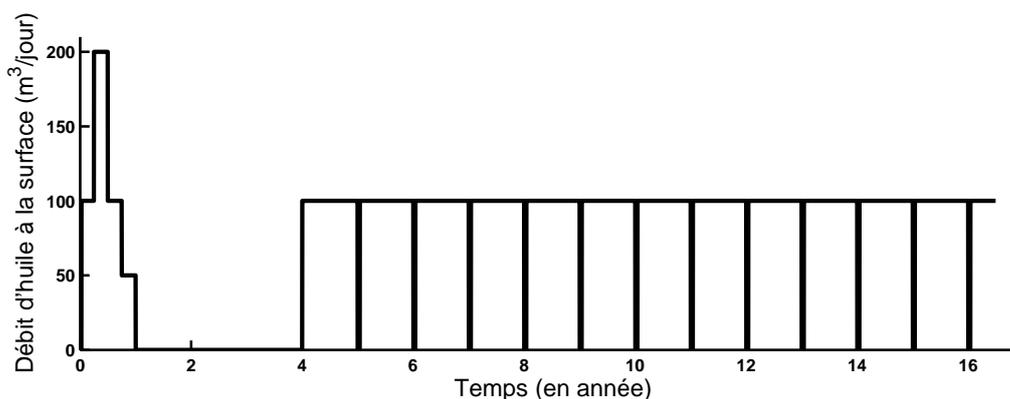


Figure 4.2 : Planning de production par puits utilisé pour le cas d'étude PUNQ.

4.1.2 Construction d'un cas PUNQ multi-fidélité

Si le cas PUNQ-S3 est constitué d'un nombre de mailles relativement faible comparé à un cas réel, ce n'est toutefois pas un cas d'étude simple. En effet, c'est un cas triphasique (eau, huile et gaz), défini en 3 dimensions, pour lequel le comportement des réponses dynamiques est relativement complexe à cause des contraintes de production définies aux puits. Néanmoins, il est rapide à simuler grâce à son nombre réduit de mailles. Il est donc pratique pour tester des méthodes de méta-modélisation. Afin de pouvoir étendre les tests à un cadre multi-fidélité, nous avons construit un cas multi-fidélité à partir du cas PUNQ-S3.

Définition des niveaux de fidélité

Nous considérons ici deux niveaux de fidélité correspondant à deux niveaux de résolution pour la grille du modèle.

1. Le « modèle grossier » possède une grille identique à celle du cas PUNQ-S3 d'origine. Elle contient $19 \times 28 \times 5$ mailles dont 1761 sont actives. On note ce modèle PUNQ grossier.
2. Le « modèle fin » possède une grille qui est obtenue via un raffinement par 3×3 selon les axes X et Y de la grille de PUNQ grossier. Elle contient donc $57 \times 84 \times 5$ mailles dont 15849 sont actives. On note ce modèle PUNQ fin.

Une simulation d'écoulement sur le modèle grossier va donner des réponses sur le niveau de fidélité grossier qu'on appellera aussi « réponses grossières ». De même, une simulation d'écoulement lancée sur le modèle fin va donner des réponses sur le niveau de fidélité fin, aussi appelées « réponses fines ». On note qu'une simulation d'écoulement utilisant le planning de production de la Figure 4.2 dure en moyenne 3 minutes pour PUNQ fin et 10 secondes pour PUNQ grossier.

Construction de PUNQ fin

Le modèle grossier ne diffère du modèle d'origine PUNQ-S3 que par la distribution des propriétés pétrophysiques (porosité et perméabilité) dans la grille. Le fichier de simulation utilisé pour PUNQ grossier est donc le même que pour PUNQ-S3 (après traduction du simulateur d'écoulement Eclipse [Eclipse, 2016] vers PumaFlow [PumaFlow, 2016], le simulateur d'écoulement développé par IFPEN).

Le modèle PUNQ fin est, quant à lui, construit à partir des fichiers de simulation de PUNQ grossier. Diverses adaptations ont été nécessaires pour obtenir un modèle cohérent:

- Changement du nombre de mailles selon X et Y ($19 \times 28 \rightarrow 57 \times 84$).
- Changement des dimensions des mailles selon X et Y (180 m \rightarrow 60 m).
- Mise à jour des coordonnées des puits.
- Raffinement des mailles inactives (1 maille inactive \rightarrow 9 mailles inactives).
- Raffinement des coefficients définissant les flux des aquifères (le coefficient d'une maille grossière est copié sur les 9 mailles fines associées).
- Raffinement de la profondeur du toit du réservoir (interpolation linéaire).
- Raffinement de l'épaisseur des mailles (interpolation linéaire).

On remarque que les puits dans le modèle PUNQ-S3 sont verticaux et passent par le centre des mailles. Le raffinement considéré pour la grille de PUNQ fin (3×3 selon les axes X et Y) implique que la mise à jour des coordonnées des puits est triviale. En effet, sur la grille fine le puits se positionne sur la maille fine centrale, obtenue suite au raffinement de la maille grossière où se situe ce puits.

Calcul des propriétés pétrophysiques au niveau fin

Pour obtenir des modèles fin et grossier cohérents, les propriétés pétrophysiques ont d'abord été simulées sur le modèle fin, puis mises à l'échelle sur le modèle grossier. Les propriétés pétrophysiques ont été simulées sur l'échelle fine en s'inspirant du modèle d'origine PUNQ-S3. La porosité a été générée indépendamment pour chacune des couches à l'aide de la méthode FFT-MA (*FFT Moving Average*) introduite par [Le Ravalec et al., 2000]. Les réalisations sont ensuite conditionnées en réalisant un double krigeage suivant la pratique usuelle décrite dans [Chilès and Delfiner, 1999]. Les valeurs utilisées pour le conditionnement aux puits sont données dans le Tableau 4.2.

Couche	PRO-1	PRO-4	PRO-5	PRO-11	PRO-12	PRO-15
1	0,08	0,22	0,23	0,08	0,08	0,28
2	0,06	0,06	0,07	0,09	0,11	0,10
3	0,10	0,11	0,21	0,24	0,10	0,19
4	0,15	0,16	0,15	0,13	0,18	0,20
5	0,24	0,21	0,09	0,15	0,24	0,28

Tableau 4.2 : Valeurs de porosité aux puits.

La distribution pour la perméabilité horizontale suit une loi log-normale. Les réalisations de perméabilité sont générées par cosimulation en considérant la porosité comme donnée secondaire. On utilise pour cela la méthode SGSim (*Sequential Gaussian Simulation*) [Journel and Deutsch, 1998]. Le coefficient de corrélation entre la porosité et le log de la perméabilité horizontale K_h est de 0,8. La perméabilité verticale K_v est prise égale à : $K_v = 0,1K_h$.

Les valeurs choisies pour les différents paramètres, ainsi que les caractéristiques des variogrammes utilisés durant cette procédure, sont spécifiées par couche dans le Tableau 4.3. Les réalisations en porosité et perméabilité sont finalement tronquées selon les valeurs limites données dans le Tableau 4.4. On remarque que les valeurs des Tableaux 4.2, 4.3 et 4.4 sont toutes semblables ou identiques aux valeurs fournies pour le cas de référence PUNQ-S3.

Couche	1	2	3	4	5
ϕ moyen	0,14	0,08	0,14	0,10	0,14
ϕ écart-type	0,11	0,04	0,11	0,06	0,11
$\log_{10}(K_h)$ moyen	2	1,4	2	1,6	2
$\log_{10}(K_h)$ écart-type	1	0,6	1	0,8	1
Portée (m)	1000	750	1500	750	1250
Ratio d'anisotropie	3,5	1	4	2	3
Azimut (degrés)	-60	-60	-45	60	-30

Tableau 4.3 : Caractéristiques des propriétés pétrophysiques par couche.