

# Nouveaux outils pour l'analyse de sensibilité et de risques

---

L'exploration pétrolière se base sur des données très parcellaires et difficiles à interpréter à l'échelle d'un bassin sédimentaire (carottes, cuttings, affleurements...) ainsi que sur des informations indirectes (enregistrement sismique, suintements de fluides en surface...) qui rendent les analyses de systèmes pétroliers complexes et les résultats incertains. Elle fait également appel à des investissements financiers très importants (plusieurs milliards d'euros). Les taux de succès de forages pétroliers d'exploration varient entre 10 et 20% pour les zones vierges d'exploration et peuvent monter jusqu'à 30% pour des bassins sédimentaires mieux connus. Les coûts de forages varient, eux, de quelques millions à plusieurs dizaines de millions d'euros par puits. L'estimation des risques de succès dans la phase d'exploration est donc essentielle pour les prises de décision d'investissements dans un portfolio de prospects pétroliers.

Les années 80 ont vu la formalisation du concept de système pétrolier (ex : Dow, 1972 ; Perrodon, 1980 ; Demaison, 1984) qui réunit l'ensemble des éléments et des mécanismes aboutissant à des accumulations d'hydrocarbures dans les bassins sédimentaires. L'estimation des risques exploratoires requiert la prise en compte des incertitudes qui affectent la connaissance du bassin sédimentaire. Concrètement, les explorateurs pétroliers segmentent l'étude des risques liés à la découverte d'un nouveau prospect, supposés indépendants, en plusieurs études portant sur les éléments du système pétrolier (présence et efficacité). Il s'agit notamment d'étudier la probabilité de présence d'un réservoir, de présence d'un piège, de l'existence d'une charge en hydrocarbures et de l'accumulation effective des hydrocarbures dans le piège (ex : CCOP, 2000 ; SPE, 2001). Ces différentes analyses indépendantes sur le fonctionnement et la mise en place d'un système pétrolier sont très souvent présentées sous formes de cartes. Elles se basent sur l'analyse des environnements de dépôts sédimentaires (Gross Depositional Environment Maps, GDE) issue de l'interprétation des données de puits et de sismique. Ces cartes de risques individuels (Component Common Risk Segment Maps, CRS maps) sont ensuite combinées pour déterminer et cartographier, à l'échelle du bassin sédimentaire, le risque exploratoire global (ex : Le Guerroué *et al.*, 2016).

Ce chapitre présente de nouveaux outils permettant à la fois d'intégrer la modélisation de bassin au sens large dans ce type d'approche et d'approfondir l'analyse et la compréhension des risques liés à l'exploration pétrolière. On présente ainsi dans une première partie des travaux permettant de cartographier et ainsi de mieux comprendre l'influence des paramètres de modélisation sur la probabilité de présence d'un réservoir pétrolier. L'approche proposée est illustrée sur un modèle stratigraphique synthétique du Golfe du Mexique réalisé avec DionisosFlow, logiciel développé à IFPEN. Une seconde partie décrit une méthode permettant d'obtenir des cartes de type CRS pour deux types de risques : le risque de présence d'un réservoir dans le Golfe du Mexique et le risque sur la maturité d'une roche-mère dans le bassin du Levant.

## 2.1 Analyse de sensibilité à l'échelle régionale

### 2.1.1 Contexte

L'un des grands défis actuels de l'exploration pétrolière concerne le renouvellement des réserves en hydrocarbures par la découverte de nouveaux champs pétroliers. Les concepts de la stratigraphie sismique et séquentielle (e.g. *Vail et al.*, 1977, *Catuneanu et al.*, 2009), très utilisés dans les procédures d'exploration pétrolière, sont très utiles pour analyser la dynamique des bassins sédimentaires à l'origine de la distribution des réservoirs pétroliers. Cependant, ils sont souvent limités à des études en deux dimensions et à des approches qualitatives. Depuis les années 1970 (e.g. *Harbaugh et Bonham-Carter*, 1970 ; *Merriam*, 1972), des outils de modélisation stratigraphique ont été conçus afin de reconstruire de manière dynamique l'histoire des bassins sédimentaires, d'améliorer la compréhension des processus physiques en lien avec la sédimentation, le climat et la tectonique et de prévoir la distribution des propriétés lithologiques notamment dans les réservoirs, les couvertures et les roches mères (ex : *Granjeon et Joseph*, 1999 ; *Mann et Zweigel*, 2008 ; *Chauveau et al.*, 2014). Ces modèles sont de plus en plus utilisés par l'industrie pétrolière car ils participent à l'amélioration de l'interprétation des données géologiques, fournissent un moyen de tester de manière quantitative le rôle des différents mécanismes impliqués dans les processus de sédimentation, permettent de prévoir et d'extrapoler des résultats dans des zones mal contraintes, favorisent l'interdisciplinarité, mettent en évidence des résultats parfois peu intuitifs et aident à illustrer les concepts géologiques imaginés par les explorateurs. Ces outils jouent donc un rôle majeur dans la découverte de nouvelles opportunités dans des zones frontières telles que les formations sous salifères, les dépôts marins profonds, ou les piémonts.

Le recours à des modèles numériques permet généralement une prise en compte des incertitudes plus aisée et plus quantitative. Les analyses de sensibilité et d'incertitudes sont ainsi couramment utilisées dans les études de modélisations de bassin ou de modélisation stratigraphique. Elles restent toutefois manuelles et sont très souvent cantonnées au test de valeurs et de scénarios extrêmes (ex : *Burgess*, 2006 ; *Csato et Granjeon*, 2013). Au mieux sont-elles basées sur des méthodes de Monte Carlo qui requièrent plusieurs milliers de simulations, les rendant inadaptées aux besoins de l'industrie du fait de la durée des simulations (*Lothe*, 2005 ; *Falivene*, 2012 ; *Tommeras*, 2016). On voit toutefois croître l'intérêt pour des outils permettant de traiter rigoureusement ce sujet, d'une part afin de mieux comprendre les modèles et la physique, et d'autre part afin d'obtenir des représentations statistiques permettant une amélioration de la capacité prédictive des modèles.

La plupart des outils disponibles pour la prise en compte des incertitudes portent sur des résultats de modélisation scalaires (volume d'un réservoir, quantité d'hydrocarbures produite par une roche mère...). Ces approches ont ouvert la voie à la prise en compte des incertitudes en modélisation de bassin mais ne couvrent pas encore l'essentiel des besoins. En effet, de nombreuses méthodes d'analyse en exploration, notamment pour l'estimation des risques exploratoires, reposent sur l'interprétation de cartes (CRS maps) ou de courbes (ex. : évolution de la pression le long d'un futur puits de forage pour prévenir les risques d'explosion).

Nous proposons donc d'utiliser une méthode alternative permettant des interprétations robustes des incertitudes liées aux modèles de bassins sur la base d'un nombre réduit de simulations en comparaison avec des approches de type Monte Carlo. La méthode se base sur les travaux de *Feraille et Marrel* (2012) qui utilisent des méta-modèles afin d'approcher les valeurs d'intérêt du simulateur. La méthode consiste à utiliser des fonctions analytiques pour estimer les résultats du simulateur dans l'espace d'incertitude sur la base de quelques résultats de simulation. Lorsque ces méta-modèles représentent de manière convenable les résultats du simulateur, il devient possible de les utiliser pour la quantification des incertitudes (ex : *Bouquet*, 2013). Cette approche a, par la suite, été étendue à l'étude de résultats de simulation plus complexes tels que des cartes et des courbes (*Marrel et Perrot*, 2012; *Douarche et al.*,

2014; Marrel *et al.*, 2015). C'est cette dernière possibilité que nous allons illustrer sur des problématiques d'exploration pétrolière. Elle a également été récemment mise en œuvre par Bruneau (2016) pour l'analyse des mécanismes régissant la distribution de la matière organique dans le bassin de Paris.

Pour une meilleure compréhension des travaux réalisés, on présente d'abord sommairement les fondements mathématiques des méthodes utilisées pour le traitement des incertitudes sur des sorties fonctionnelles de simulation telles que des cartes. Le potentiel de la méthode est par la suite illustré par une application sur la détermination de l'influence spatiale des différents processus à l'œuvre dans la distribution des réservoirs pétroliers d'un intervalle stratigraphique du Golfe du Mexique.

## 2.1.2 Méthode de quantification des incertitudes sur des cartes

### Approche par modèles proxy

Le modèle analytique approchant la réponse du simulateur  $y$ , également appelé méta-modèle, modèle proxy ou surface de réponse, est construit à partir d'une base d'apprentissage de quelques simulations. Le nombre de ces simulations va dépendre du nombre de paramètres incertains intégrés dans l'étude et de la complexité de la réponse. Ce modèle analytique peut être construit sur une base polynomiale, par réseau de neurones ou par des processus gaussiens. Cette dernière approche est utilisée pour la construction de nos méta-modèles (se référer à Sacks *et al.* (1989) ou Forrester et Keane (2009) pour plus de détails).

Du fait de la topographie et des mécanismes en jeu, les paramètres incertains d'un modèle stratigraphique vont avoir des effets sensiblement différents dans les différentes zones d'un même bassin sédimentaire. Ces différents effets peuvent être déterminés par la construction d'un méta-modèle dans chacune des cellules du modèle. Cette manière de procéder ne permet toutefois pas de tenir compte de l'environnement immédiat de chaque cellule considérée et se révèle coûteuse en temps de calcul et en mémoire (une carte compte en général entre 2500 et 250000 cellules). La méthode utilisée ici, et reprise de Marrel et Perrot (2012) et de Douarche *et al.* (2014), consiste à réaliser une analyse en composantes principales des cartes résultant des quelques simulations réalisées afin de réduire le nombre de modèles proxy à construire. Dans cette analyse en composantes principales, les mailles jouent le rôle des individus et les simulations celui des variables. Les principes de la méthode mathématique utilisée sont brièvement rappelés avant qu'elle ne soit appliquée à une représentation synthétique du Golfe du Mexique.

### Analyse en composantes principales des cartes résultats

On considère ici une propriété d'intérêt  $y(u, \theta)$  dépendant de la position dans l'espace  $u$  et d'un ensemble de paramètres incertains  $\theta$ . La décomposition de Karhunen-Loève (Loève, 1978) permet de caractériser  $y(u, \theta)$  par une suite de fonctions du type:

$$\text{Équation 15: } y(u, \theta) = \sum_{k=1}^{\infty} \alpha_k(\theta) \phi_k(u)$$

où les fonctions  $\phi_k(u)$  forment une base orthonormale et où les  $\alpha_k(\theta)$  sont les coefficients de projection sur cette base. Les fonctions de base sont orthogonales et triées par ordre décroissant de variance expliquée de la réponse. Cette décomposition revient donc à définir de nouvelles variables décorrélatées  $\alpha_k$ . Ces composantes vont être approchées par des méta-modèles de type processus gaussiens car en pratique, on ne connaît pas la forme analytique de la réponse.

Un échantillon d'apprentissage de taille  $n$ ,  $(\theta_i)_{i=1 \dots n}$ , est généré en utilisant une méthode d'hypercube latin (McKay *et al.*, 1979). Avec cette approche les paramètres varient de manière simultanée et une bonne couverture de l'espace d'incertitude est assurée. On simule alors les réponses d'intérêt

correspondantes ( $y(u, \theta_i)_{i=1\dots n}$ ) dans chacune des  $N$  mailles du modèle. On utilise ensuite cet échantillon des réponses du simulateur pour réaliser la décomposition Karhunen-Loève de la sortie, ce qui revient à faire une analyse en composantes principales (ACP). Pour chaque élément de l'échantillon, on a donc :

$$\text{Équation 16: } y(u, \theta_i) = \sum_{k=1}^M \hat{\alpha}_k(\theta_i) \hat{\phi}_k(u)$$

où les  $\hat{\phi}_k$  sont les estimations des fonctions de base qui correspondent aux composantes principales de l'ACP et les  $\hat{\alpha}_k$  sont les estimations des coefficients de projection sur cette base. La décomposition linéaire est ici finie, avec  $M$  termes ( $M \leq N$ ). En pratique, cette décomposition peut être « compressée », dans le sens où on ne garde qu'un nombre  $m$  limité, de fonctions de base qui capturent l'essentiel de la variance de la réponse :

$$\text{Équation 17: } y(u, \theta_i) \approx \sum_{k=1}^m \hat{\alpha}_k(\theta_i) \hat{\phi}_k(u)$$

$m$  est choisi de telle sorte que la combinaison linéaire en résultant corresponde à un pourcentage donné de la variance totale de la réponse de l'échantillon. En pratique, on cherche en général à reproduire 90 à 95% de la variance. Si la variance totale de l'échantillon n'est pas conservée, le modèle approché de  $y$  ne reproduira pas exactement les valeurs simulées ( $y(u, \theta_i)_{i=1\dots n}$ ) aux points de l'échantillon ( $\theta_i)_{i=1\dots n}$  (équation 17).

La construction du méta-modèle pour chaque fonction  $\hat{\alpha}_k(\theta)_{k=1\dots m}$  à partir des valeurs de l'échantillon  $(\theta_i, \hat{\alpha}_k(\theta_i))_{i=1\dots n}$  se fait ensuite par processus gaussiens comme expliqué précédemment. Le méta-modèle  $y^*$  de la sortie d'intérêt s'écrit alors, pour toute valeur des paramètres  $\theta$  :

$$\text{Équation 18: } y^*(u, \theta) = \sum_{k=1}^m \hat{\alpha}_k(\theta) \hat{\phi}_k(u)$$

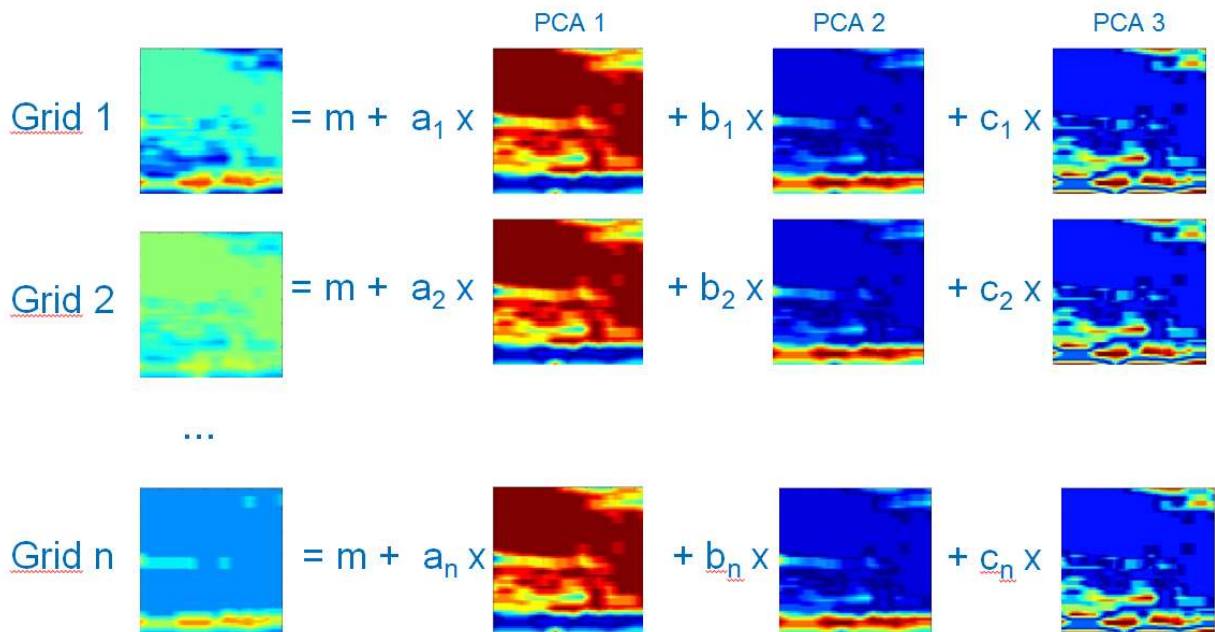


Figure 21 : Exemple de modèle approché où les réponses sont déterminées à partir de 3 composantes. Une surface de réponse (ou proxy) relie les valeurs des paramètres  $\hat{\alpha}_k$  aux valeurs des paramètres incertains  $\theta$  (Gervais et al., 2016).

*Détermination de la qualité du méta-modèle*

Le méta-modèle obtenu peut ensuite être contrôlé par différentes approches. Tout d'abord, on peut regarder l'erreur aux points de l'échantillon en lien avec la troncature de la base ( $Q2(troncature)$ ). On peut également regarder la qualité de la prévision obtenue dans chaque maille pour la propriété d'intérêt sur un ensemble de simulations indépendantes de celles utilisées pour construire les surfaces de réponse ( $Q2(confirmation)$ ). Ces deux critères de qualité sont utilisés par la suite pour estimer la qualité des méta-modèles générés :

$$\text{Équation 19: } Q2(troncature) = 1 - \frac{\sum_{i=1}^n (y(u, \theta_i) - y^*(u, \theta_i))^2}{\sum_{i=1}^n (y(u, \theta_i) - \bar{y}(u))^2}$$

$$\text{Équation 20: } Q2(confirmation) = 1 - \frac{\sum_{i=1}^{nt} (y(u, \theta_i) - y^*(u, \theta_i))^2}{\sum_{i=1}^{nt} (y(u, \theta_i) - \bar{y}(u))^2}$$

où  $\bar{y}$  représente la moyenne de  $y$  sur l'échantillon et où  $nt$  correspond à la taille d'un échantillon de simulations supplémentaires, non utilisé pour la détermination des méta-modèles. On peut remarquer que plus un critère de qualité  $Q2$  est proche de 1 et plus l'erreur est faible. Un modèle est généralement considéré comme prédictif lorsque les valeurs de  $Q2$  sont supérieures à 0,9. Une fois obtenu un modèle prédictif de la propriété d'intérêt, il est possible de l'utiliser pour réaliser par exemple :

- des propagations d'incertitudes : calcul de la réponse prédite pour un large échantillon des paramètres selon leur distribution *a priori*. Cette utilisation des méta-modèles est illustrée dans la section suivante pour l'obtention des cartes de probabilités.
- des analyses de sensibilité, notamment quantitatives avec l'estimation des indices de Sobol (GSA).

**Analyse de sensibilité par décomposition de la variance**

Sobol (1993) a proposé une approche de décomposition de la variance basée sur la somme des contributions des différents paramètres d'entrée du modèle. Elle a ensuite été formalisée et étendue par Saltelli et al. (1999; 2002; 2010). Elle est particulièrement intéressante car elle permet de mesurer la sensibilité sur l'ensemble de l'espace d'incertitude (méthode globale), elle s'applique à des réponses non-linéaires et permet de mesurer les effets d'interaction entre les paramètres.

Avec cette approche la réponse d'intérêt est vue comme une boîte noire qui peut se décomposer sous la forme suivante :

$$\text{Équation 21: } y(u, \theta) = f_0(u) + \sum_{i=1}^d f_i(u, \theta_i) + \sum_{i<j}^d f_{ij}(u, \theta_i, \theta_j) + \dots + f_{1,2,\dots,d}(u, \theta_1, \theta_2, \dots, \theta_d)$$

où  $f_0$  est une constante et  $f_i$  est une fonction de  $(u, \theta_i)$ , etc. L'ensemble des fonctions  $f_i$  forme une base orthogonale. Ceci conduit à exprimer les termes de la décomposition sous la forme suivante:

$$\text{Équation 22: } Var(y) = \sum_{i=1}^d Var_i + \sum_{i<j}^d Var_{ij} + \dots + Var_{1,2,\dots,d}$$

Où l'on voit apparaître l'effet de la variation de  $\theta_i$  seul (appelé effet principal de  $\theta_i$ ), et les effets couplés (effets d'interaction). L'Équation 23 illustre cette décomposition dans le cas d'une étude à deux paramètres.

Équation 23: 
$$1 = \frac{Var(f_1)}{Var(y)} + \frac{Var(f_2)}{Var(y)} + \frac{Var(f_{1,2})}{Var(y)} = S_1 + S_2 + S_{1,2}$$

En prenant un exemple avec  $d = 2$  et en normalisant l'expression précédente par  $Var(y)$ , on fait apparaître une mesure (indices de Sobol) de la sensibilité de la réponse aux variations des paramètres d'entrée. Les indices de premiers ordres,  $S_i$ , correspondent aux effets dits principaux des paramètres (effets de leur seule variation). L'effet total des paramètres est la deuxième mesure couramment utilisée. Elle caractérise l'effet de la variation de  $\theta_i$  en incluant l'ensemble de ses interactions avec les autres paramètres. On peut donc noter que la somme des effets totaux sera généralement supérieure à 1 car les effets d'interaction sont ajoutés à chacun des paramètres.

Dans la pratique les indices de sensibilités ne sont pas calculés analytiquement mais sont estimés par des méthodes de type Monte Carlo. Dans notre cas, l'approche suivie consiste à appliquer la méthode de Monte-Carlo sur le méta-modèle préalablement déterminé sur la base de processus gaussiens (Sudret 2008; Petropoulos et al., 2009). On obtient alors un résultat qui permet de cartographier l'influence des paramètres incertains du modèle et de la relier à la morphologie du bassin sédimentaire. C'est ce qui est illustré sur un cas synthétique du Golfe du Mexique dans la suite de ce travail.

### 2.1.3 Application à un intervalle sédimentaire du Golfe du Mexique

#### Modèle synthétique du Golfe du Mexique

L'outil de modélisation stratigraphique utilisé dans cette étude est DionisosFlow (Granjeon et Joseph, 1999), développé par IFPEN. La méthode d'estimation de la sensibilité du modèle est illustrée sur un cas synthétique en trois dimensions inspiré du Golfe du Mexique initialement présenté par Burgess *et al.* (2006).

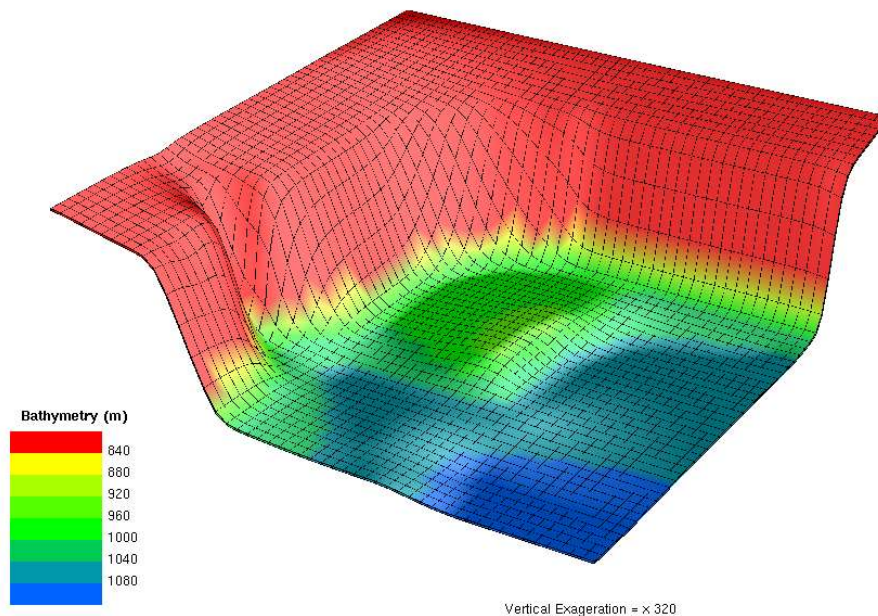


Figure 22 : Bathymétrie initiale du modèle synthétique du Golfe du Mexique. Le modèle a une résolution de 20kmx20km sur un domaine de 1000km x 1000km (Burgess et al., 2006). Le modèle décrit une zone de plateau continental suivie de fortes pentes jusqu'au bassin profond et un bassin présentant quelques reliefs reliés au plateau continental par des canyons sous-marins. Les apports sédimentaires proviennent de la marge nord-ouest, alimentée par le Mississippi.

Le modèle représente une marge passive sur un domaine de 1000km x 1000km. Le modèle décrit une zone de plateau continental suivie de fortes pentes jusqu'au bassin profond et un bassin présentant quelques reliefs reliés au plateau continental par des canyons sous-marins. La bathymétrie initiale du modèle est illustrée à la figure 22. Le domaine simulé est discrétisé en 50 x 50 cellules de 20km de résolution. Le remplissage sédimentaire du bassin est assuré par une source importante de sédiments de 50km de large (représentative du delta du Mississippi) sur la marge continentale nord-ouest du bassin. L'apport sédimentaire est un mélange de sable et d'argile considéré constant sur la durée de la simulation. La simulation rend compte de l'évolution du bassin sur une durée de 3 millions d'années par pas de temps successifs de 0,2 millions d'années. La distribution des sédiments dans le bassin est assurée par les mécanismes de transport diffusif. Les variations du niveau de la mer sont supposées sinusoïdales. La subsidence du bassin est supposée constante dans le temps. Nous ne présentons pas l'analyse stratigraphique des résultats du modèle de référence, décrits en détail par Burgess *et al.* (2006).

Afin de faciliter l'interprétation des résultats, les incertitudes sur le modèle n'ont pas été définies sur des paramètres mais sur des phénomènes physiques pouvant regrouper plusieurs paramètres. Il apparaît en effet plus intuitif de raisonner à l'échelle d'un processus physique (ex : transport plus ou moins efficace) qu'à l'échelle du paramètre contrôlant le processus physique (ex : coefficient de diffusion gravitaire du sable en domaine marin). Les incertitudes définies dans l'étude ont été choisies pour être représentatives des trois grands processus mis en jeu dans un modèle stratigraphique : l'apport de sédiments, leur transport et la dynamique de l'espace disponible (accommodation) pour leur dépôt (tableau 8).

Paramètres ou processus incertains		Valeur minimale	Valeur maximale
<b>Accommodation</b>	Variation eustatique – Période (Ma)	0.5	2
	Variation eustatique – Amplitude (m)	5	50
	Taux de subsidence (m/Ma)	25	75
<b>Apports</b>	Apports sédimentaires (km <sup>3</sup> /Ma)	20000	80000
	Qualité de la source (% sable)	10	40
	Décharge fluviale (%)	50	200
<b>Transport des sédiments</b>	Coefficient de référence (km <sup>2</sup> /ka) - Diffusion du sable par flux d'eau en domaine marin	0.5	2
	Ratio d'efficacité de transport continental/marin (-)	50	100
	Ratio d'efficacité de transport argile/sable (-)	1.5	4.5

Tableau 8 : Ensemble des 9 paramètres incertains utilisés dans l'étude avec leurs intervalles de variation. Les incertitudes ont été choisies pour être représentatives des trois grands processus mis en jeu dans un modèle stratigraphique : l'apport de sédiments, leur transport et la dynamique de l'espace disponible pour leur dépôt.

L'incertitude sur la dynamique du bassin est prise en compte au travers de trois paramètres : l'amplitude et la périodes des variations eustatiques et le taux de subsidence. L'incertitude sur les apports sédimentaires est caractérisée par des variations des décharges fluviales dans le bassin, par la variation des apports sédimentaires ainsi que par la qualité de ces derniers (plus ou moins sableux). Enfin, les

incertitudes liées aux mécanismes de transport ont été largement remaniées par rapport à l'étude initiale de Burgess *et al.* (2006). Plutôt que de faire porter l'incertitude sur les coefficients de diffusion, trois paramètres de contrôle du transport ont été définis :

- 1- un coefficient de diffusion de référence (ici le coefficient de diffusion des sables par les flux d'eau en domaine marin),
- 2- un ratio permettant de contrôler la différence d'efficacité de transport entre les sédiments grossiers (sable) et les sédiments fins (argile),
- 3- un ratio permettant de contrôler la différence d'efficacité de transport entre les domaines marin et continental.

Cette approche permet d'une part d'assurer la cohérence entre les mécanismes de transport dans les différents domaines du bassin et pour les différents types de sédiments, et d'autre part de réduire le nombre de paramètres incertains (et donc le nombre de simulations à réaliser pour obtenir le méta-modèle). Elle permet également de faire porter l'incertitude sur des processus physiques plutôt que sur des paramètres indépendants, ce qui rend son interprétation plus aisée. Les intervalles d'incertitudes des neuf paramètres ou processus sélectionnés sont synthétisés au tableau 8.

La présence effective d'un réservoir fait partie des risques majeurs pris en compte dans les analyses de risques en exploration pétrolière. L'étude se focalise ainsi sur l'analyse de sensibilité du modèle sur la distribution des réservoirs dans le bassin. Ils doivent constituer un volume suffisamment grand pour être rentables et justifier les importants investissements financiers nécessaires à leur exploitation. Dans le modèle, les zones constituant des réservoirs seront donc définies à partir de l'épaisseur de sédiments et de la proportion de sable dans cette épaisseur. DionisoFlow ne fournit pas d'interprétation des résultats de simulation en termes d'éléments du système pétrolier, il se contente de fournir des résultats stratigraphiques bruts (distribution des classes de grains, profondeur de dépôt, flux d'eau dans la zone de dépôt...). Afin d'examiner l'effet des incertitudes du modèle sur la distribution des réservoirs, nous étudions leur sensibilité à la distribution de l'épaisseur de dépôt sédimentaire et à leur teneur en sable.

Comme décrit dans la présentation mathématique de la méthode, un ensemble de simulations d'apprentissage est nécessaire à la construction du méta-modèle utilisé pour les analyses de sensibilité. En plus de cet ensemble d'apprentissage, il est recommandé d'utiliser un ensemble indépendant de simulations pour déterminer la qualité du méta-modèle. Afin d'étudier le nombre de simulations d'apprentissage requis pour l'obtention d'un méta-modèle de bonne qualité, plusieurs lots de simulations d'effectif croissant sont utilisés : 30, 60, 90 et 120 simulations d'apprentissage. L'ensemble des simulations de contrôle, identique pour tous les lots testés comporte 50 simulations.



## 2.1.4 Résultats et discussion

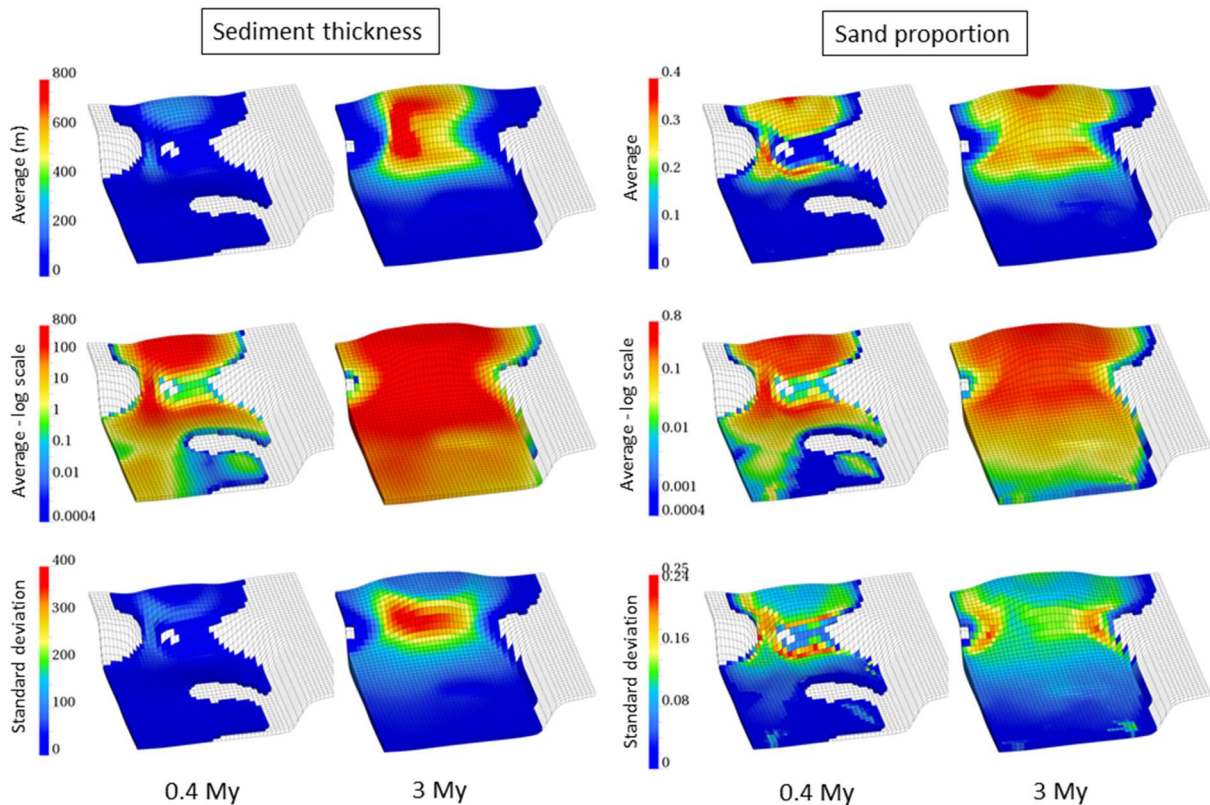


Figure 23 : Moyenne (haut), moyenne en échelle log (centre) et écart-type (bas), déterminés sur l'échantillon de 50 simulations de contrôle, pour l'épaisseur de sédiments (à gauche) et la proportion de sable (à droite) après 0,4 et 3Ma. Les cellules en gris correspondent à des zones sans dépôt sédimentaire.

La moyenne et la variance des épaisseurs sédimentaires et de la proportion de sable pour l'échantillon de simulation de contrôle sont illustrés à la figure 23. Les sédiments sont principalement déposés dans le delta. Le reste des sédiments est ensuite canalisé par le canyon sous-marin qui les redistribue dans le bassin profond principalement au pied de la marge continentale. Les résultats après 3Ma montrent que le canyon se remplit et que d'importants dépôts sédimentaires sont enregistrés le long de la marge nord-ouest du bassin, dans le prolongement du delta et de la source de sédiments. Les zones les plus sableuses sont le centre du delta et les pieds de pentes, qui correspondent aux zones de plus forte énergie où les argiles, plus fines que les sables, peuvent moins facilement se déposer. Les zones les moins riches en sables sont principalement localisées dans les parties distales et sur les bords du delta.

Pour chaque ensemble de simulations d'apprentissage, la base de vecteurs résultant de l'analyse en composantes principales a été tronquée de sorte d'expliquer au moins 98% de la variance des sorties d'intérêt (cf. Analyse en composantes principales des cartes résultats). La qualité des méta-modèles est ensuite calculée en utilisant l'ensemble des 50 simulations de contrôle. Les cartes de qualité qui en résultent sont présentées à la figure 24.

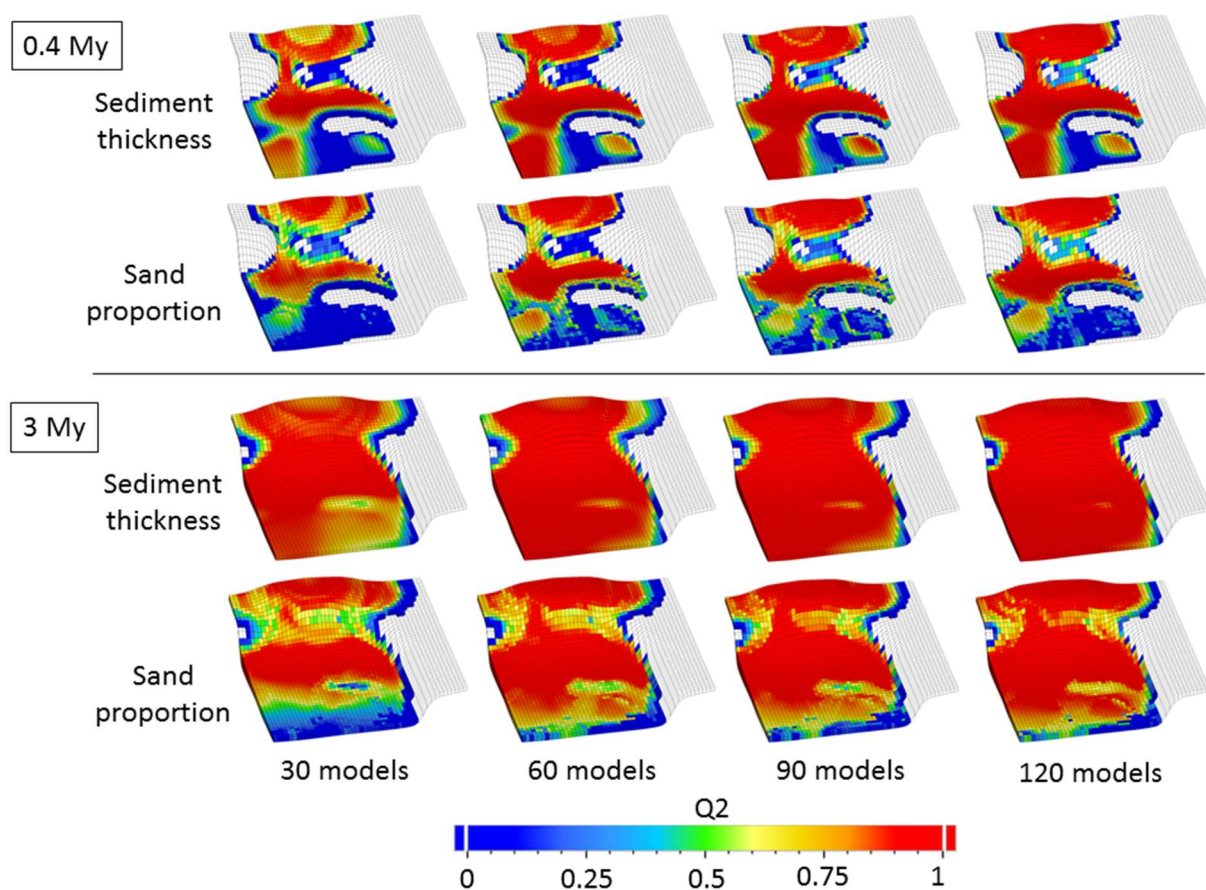


Figure 24 : Valeurs de  $Q2$  calculées pour les ensembles d'apprentissage de 30, 60, 90 et 120 simulations pour les épaisseurs de sédiments et la proportion de sable.

Plusieurs enseignements peuvent être tirés de ces résultats. Il est d'abord évident que la qualité du méta-modèle augmente avec le nombre de simulations d'apprentissage. Toutefois la qualité du méta-modèle est déjà relativement bonne avec l'ensemble d'apprentissage le plus petit. Il est d'ailleurs intéressant de remarquer que la qualité du modèle est très bonne ( $>0,9$ ) dans les zones où les dépôts sédimentaires et la variance de la réponse d'intérêt sont les plus importants comme le delta et les pieds de pente (cf. figure 23). Même lorsque le nombre de simulations dans l'ensemble d'apprentissage est faible ( $\sim 3$  fois le nombre de paramètres incertains), la qualité du modèle est très bonne dans les zones plus intéressantes pour l'analyse de sensibilité (zones avec des dépôts sédimentaires et présentant une forte variation des résultats). L'augmentation de la taille de l'ensemble d'apprentissage permet donc principalement d'améliorer la qualité du modèle dans les zones de faible variance et moins concernées par les mécanismes de sédimentation. On peut enfin noter que l'épaisseur de sédiments semble plus facile à prévoir que la distribution de sable. Cette observation peut être reliée au fait que la morphologie du bassin contrôle certainement plus fortement la distribution des épaisseurs de sédiment que la proportion de sable.

La qualité de la prévision des méta-modèles obtenus est également illustrée au travers de la figure 25 qui permet de comparer la prévision obtenue sur la base du méta-modèle aux résultats du modèle DionisosFlow pour les valeurs des paramètres incertains correspondant à l'une des simulations de contrôle. Il apparaît clairement que le méta-modèle est capable de reproduire les grandes tendances des résultats de DionisosFlow même avec le plus réduit des ensembles d'apprentissage. L'augmentation de la qualité de la prévision est principalement visible sur les limites de dépôts sédimentaires qui correspondent aux zones de plus faibles sédimentations.

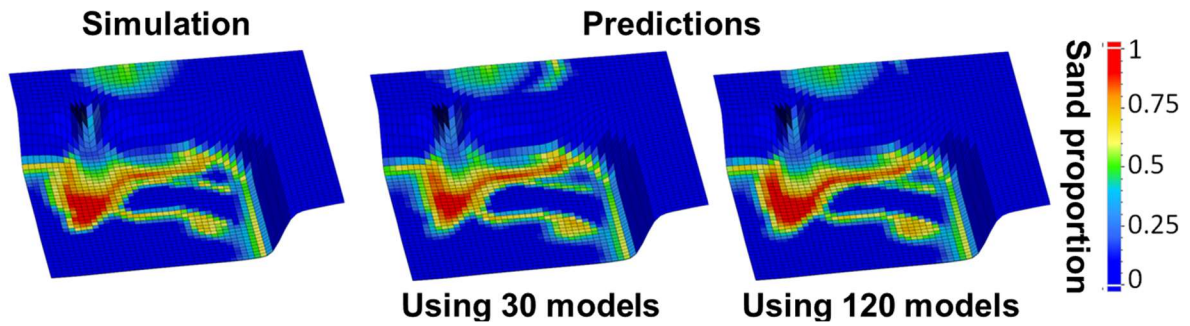


Figure 25 : Proportions de sable simulées par DionisosFlow (à gauche) et prévues par le méta-modèle après 0.4 millions d'années pour les ensembles d'apprentissage de 30 simulations (au milieu) et 120 simulations (à droite).

Dans la suite de l'étude, les résultats présentés sont ceux construits sur l'ensemble d'apprentissage de 120 simulations. La figure 26 présente, pour chacun des deux résultats d'intérêt, les effets totaux des paramètres incertains ayant la contribution la plus élevée à leur variance.

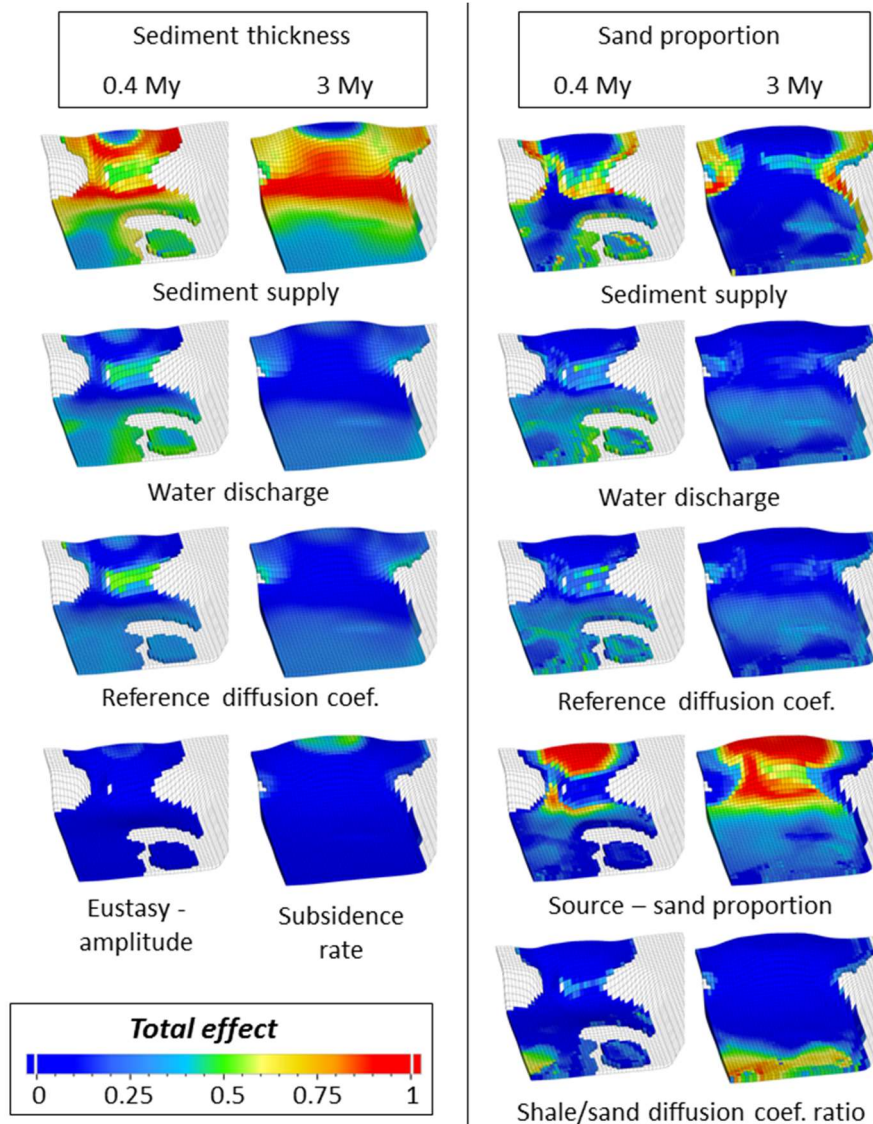


Figure 26 : Effets totaux des paramètres les plus influents sur la variance des résultats d'intérêt.

Les résultats suggèrent que l'apport sédimentaire, la décharge fluviale et le coefficient de diffusion de référence ont des effets significatifs sur les deux réponses. Sans surprise, l'apport sédimentaire est le paramètre le plus influent sur les épaisseurs de sédiments. Un autre paramètre peut toutefois être localement plus influent. C'est le cas du taux de subsidence dans certaines zones du delta. Bien que n'étant pas les principaux responsables de la variance de l'épaisseur, la décharge fluviale et le coefficient de diffusion de référence contrôlent en partie les épaisseurs sur les pentes de la marge nord-ouest et dans les zones les plus distales du bassin. La dynamique du bassin (amplitude des variations eustatiques et subsidence) joue, elle, un rôle plus important sur les dépôts enregistrés dans le delta. L'analyse des résultats sur différentes durées suggère que les variations eustatiques pourraient avoir principalement un effet à haute fréquence alors que le taux de subsidence aurait un effet de basse fréquence.

En termes de distribution de la proportion de sable, là encore sans surprise, le contributeur majeur à la variance des résultats est la composition de la source (proportion de sable dans les apports). Ses effets vont se retrouver principalement dans les zones à fort taux de sédimentation que sont le delta et les pieds de pente (cf. figure 26). Le volume d'apport sédimentaire joue un rôle majeur sur les limites de dépôts sédimentaires comme les bordures du delta ou les franges des reliefs dans le bassin profond. Dans la partie la plus distale du bassin profond, aux limites de dépôts sédimentaires, la décharge fluviale, le ratio de capacité de transport entre l'argile et le sable et le coefficient de diffusion de référence, qui contrôlent en partie la distance relative de transport du sable, jouent également un rôle significatif.

### 2.1.5 Synthèse

Ce travail illustre l'utilisation d'une nouvelle méthode d'analyse de sensibilité globale, réalisée à partir d'un nombre limité de simulations, à l'échelle d'un bassin sédimentaire et sur des résultats de simulation sous forme de cartes permettant ainsi une interprétation plus en lien avec la géologie du bassin sédimentaire. Cette méthode repose sur la construction d'un méta-modèle obtenu par interpolation par processus gaussiens des sorties du simulateur décomposés en une base de vecteur par analyse en composantes principales. Ce méta-modèle analytique constitue une estimation des résultats donnés par le simulateur en tout point de l'espace d'incertitudes. Lorsqu'il est de bonne qualité, il peut être utilisé en lieu et place du simulateur pour réaliser des analyses de sensibilité globale.

La méthode est illustrée sur un cas synthétique 3D représentatif de la marge continentale nord du golfe du Mexique. L'étude a porté sur la sensibilité de la distribution en sable et en épaisseur de sédiments qui sont deux variables importantes de la définition des zones constituant de potentiels réservoirs pétroliers. Un effort de paramétrisation des incertitudes a été réalisé afin de permettre des interprétations porteuses de sens pour des explorateurs de bassin. Les résultats obtenus grâce à l'analyse de sensibilité permettent de déterminer les principales sources d'incertitude sur la distribution des réservoirs pétroliers. Elles permettent également de mieux appréhender le rôle des différents paramètres dans la phase de calibration d'un modèle aux données de contrainte (épaisseurs issues de l'interprétation sismique ou proportion de sable le long d'un puits de forage par exemple).

Dans l'analyse d'un réservoir pétrolier, pour une même épaisseur et une même proportion de sable, la fréquence des alternances entre des couches majoritairement sableuses et des couches majoritairement argileuses affecte la qualité du réservoir. Il est en effet préférable d'avoir la plus grosse colonne continue de sable afin de favoriser la production du réservoir. Un critère tenant compte de la stratification pourrait donc également être pris en compte dans l'analyse de sensibilité (Le Guerroué *et al.*, 2016).

Le temps disponible est très souvent un facteur très limitant des études menées par les équipes d'exploration des compagnies pétrolières. Une simulation stratigraphique réalisée avec un outil de simulation du type de DionisosFlow sur un bassin réel et avec une résolution relativement fine (mailles de 2 à 5km de côté) peut prendre entre plusieurs dizaines de minutes et plusieurs heures. Réaliser ne serait-ce que quelques dizaines de simulations peut ainsi devenir rédhibitoire dans des études

opérationnelles. Des travaux sont envisageables pour réduire encore le nombre de simulations nécessaires à la réalisation d'analyses de sensibilité. Il est en effet possible de réaliser dans un premier temps un nombre très réduit de simulations, de construire le méta-modèle correspondant puis d'ajouter peu à peu des simulations supplémentaires afin d'améliorer la qualité du méta-modèle et ce jusqu'à atteindre la qualité désirée ( $Q2 > 0,9$  par exemple). Cette méthode permettrait de ne réaliser que le nombre de simulations indispensables à l'obtention d'un méta-modèle de qualité suffisante.

Afin de rendre la méthode réellement applicable en conditions opérationnelles dans les compagnies d'exploration pétrolière, il faudrait également pondérer l'analyse de sensibilité par la capacité des résultats obtenus à reproduire les données géologiques disponibles (épaisseurs issues de l'interprétation sismique ou proportion de sable le long d'un puits de forage par exemple). Pour cela des travaux permettant de pondérer les différentes prévisions effectuées avec le modèle proxy par une vraisemblance calculée sur la cohérence du modèle avec les données de contrainte pourraient être envisagés (méthode bayésienne par exemple). Cette approche suppose toutefois de définir une mesure de la distance entre les résultats du modèle et les données de contrainte. Cette définition n'est pas évidente et des travaux sont en cours à IFPEN pour transcrire mathématiquement la distance entre les observations géologiques et les résultats de simulation.

Enfin, si la méthode que nous proposons donne accès à des analyses de sensibilités sur des résultats présentés sous forme de cartes permettant une meilleure interprétation par des explorateurs pétroliers, il n'est pas encore possible de prendre en compte des incertitudes sur des paramètres complexes comme des cartes. Or, les résultats obtenus suggèrent que la morphologie du bassin peut-être un contributeur majeur à l'incertitude sur les distributions des réservoirs pétroliers. Par exemple, dans l'étude réalisée ici, la subsidence a été supposée homogène sur l'ensemble du bassin alors qu'il serait beaucoup plus correct d'introduire des variations spatiales dans les incertitudes sur ce processus. De même, les incertitudes sur la bathymétrie initiale, pourtant élevées, n'ont pas été prises en compte. Des améliorations de l'approche sont envisageables sur la base des travaux de Iooss et Ribatet (2009) pour la prise en compte d'entrées fonctionnelles complexes.



## 2.2 Méthode de détermination de cartes de risque en exploration pétrolière

### 2.2.1 Contexte

L'analyse des risques en exploration pétrolière se base sur la notion de système pétrolier. Les outils de modélisation de systèmes pétroliers permettent théoriquement de calculer les volumes d'hydrocarbures probabilisés dans les différents réservoirs d'un bassin sédimentaire en définissant des incertitudes sur les paramètres d'entrée du modèle. Cette utilisation des modèles est toutefois encore peu courante du fait de la difficulté à inclure toute l'incertitude dans un unique outil de modélisation et de la multiplicité des sources d'incertitudes. Les méthodes d'analyse des risques reposent ainsi plus généralement sur la décomposition des risques associés aux différents éléments du système pétrolier et sur leur analyse indépendante : source, réservoir, piège et charge en hydrocarbures.

Une méthode particulièrement diffusée actuellement dans l'industrie pétrolière consiste à cartographier les risques de chacun des éléments du système pétrolier à l'échelle des « Plays », définis comme un ensemble de champs et de prospects pétroliers appartenant à une même unité stratigraphique et pour lequel il existe une possibilité que l'ensemble des éléments du système pétrolier ait pu fonctionner (ex : Le Guerroué *et al.*, 2016). Cette méthode, appelée CRS mapping (Common Risk Segment Mapping), consiste en une analyse régionale des risques supposés indépendants. Une fois les cartes de risques obtenues pour chacun des éléments du système pétrolier, elles peuvent être combinées pour donner une image globale des zones les plus favorables à la présence effective d'une accumulation d'hydrocarbures (figure 27).

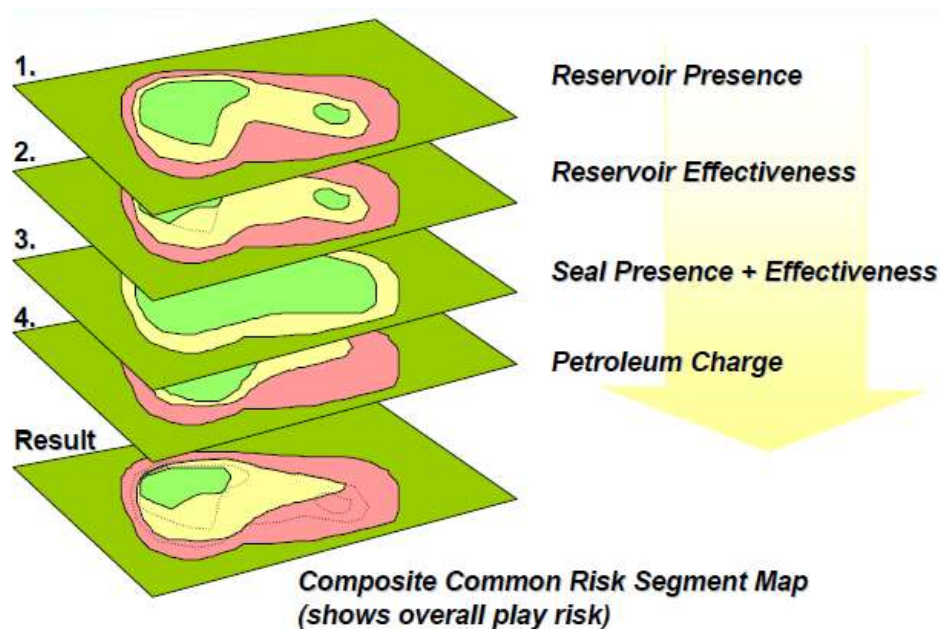


Figure 27 : Illustration de la méthode d'analyse des risques exploration dans un « play » pétrolier par CRS mapping. Une analyse régionale des risques est effectuée sur chacun des éléments du système pétrolier. Elles sont combinées dans une carte globale des risques (Composite Common Risk Segment Map). Le niveau de risque est représenté par un code couleur, du rouge (risque élevé) au vert (risque faible).

Cette analyse des risques peut être réalisée à partir d'un modèle géologique conceptuel construit par l'explorateur sur la base de son interprétation de l'ensemble des données à sa disposition par opposition à un modèle géologique obtenu par simulation numérique.

Nous proposons de montrer comment les outils de modélisation numériques de bassin (stratigraphique ou de systèmes pétroliers) peuvent être utilisés dans ce type d'approche. La méthode se base sur les travaux de Feraille et Marrel (2012) qui utilisent des méta-modèles et qui a été présentée à la section précédente. Le potentiel de la méthode est illustré par une estimation des probabilités de présence d'un réservoir sableux dans un play du Golfe du Mexique. Un deuxième exemple montre de quelle manière l'incertitude sur la détermination de la cinétique de transformation de la matière organique en hydrocarbures de la roche-mère du campanien du bassin du Levant peut être évaluée puis utilisée pour l'estimation des risques sur la maturité de cette dernière.

## 2.2.2 Méthode de cartographie des risques

La section précédente décrit une méthode d'analyse de sensibilité globale à l'échelle d'un bassin sédimentaire et sur des résultats de simulation sous forme de cartes. Le méta-modèle obtenu constitue une estimation des résultats donnés par le simulateur en tout point de l'espace d'incertitudes. Lorsqu'il est de bonne qualité il peut ainsi être utilisé en lieu et place du simulateur pour réaliser une analyse de risque de type Monte-Carlo.

Il s'agit de définir des distributions de probabilités des paramètres incertains sur l'intervalle d'incertitude utilisé puis de propager ces incertitudes sur la réponse d'intérêt, une carte dans notre cas, en utilisant le méta-modèle analytique. On obtient alors des distributions de probabilités pour la réponse d'intérêt dans chacune des cellules de la carte considérée. Lorsque l'analyse porte sur l'un des différents éléments du système pétrolier (présence d'un réservoir, maturité d'une roche-mère...) les résultats peuvent ensuite être intégrés dans une méthode d'analyse des risques de type CRS mapping.

## 2.2.3 Risque de présence d'un réservoir sableux

Pour certains risques liés au système pétrolier, comme la présence d'un réservoir, des critères doivent être utilisés pour filtrer les résultats bruts d'un modèle. En effet, les résultats d'un modèle stratigraphique ne sont pas directement des informations sur la nature des sédiments (réservoirs, couverture...) mais sur les conditions de dépôt (bathymétrie, énergie de l'environnement de dépôt...), leur épaisseur et leur composition sédimentaire (composition en classes de sédiments). Un réservoir peut ainsi être défini sur la base de critères comme l'épaisseur de sédiment et leur proportions en sable (Burgess *et al.*, 2006). Elle peut également prendre en compte la fréquence des alternances entre faciès majoritairement sableux et faciès majoritairement argileux (Le Guerroué *et al.*, 2016).

Pour illustrer la manière de réaliser la cartographie du risque de présence d'un réservoir avec cette approche, on reprend l'exemple du golfe du Mexique où l'on étudie un « play » correspondant à un intervalle de temps de 3 millions d'années. La distribution de probabilité de l'incertitude a été supposée uniforme pour l'ensemble des paramètres incertains (cf. Tableau 8). L'estimation des distributions de probabilité *a posteriori* des variables d'intérêt a été réalisée à partir d'un échantillon de 20000 valeurs des paramètres incertains, assurant une définition correcte des probabilités dans chacune des cellules des cartes de variables d'intérêt. La probabilité de présence d'un réservoir dans une cellule est ensuite directement donnée par la proportion de cas où les valeurs d'épaisseur et de proportion de sable respectent les critères préalablement définis parmi les 20000 échantillonnages des méta-modèles. La figure 28 présente les cartes de probabilités obtenues dans les cas où un réservoir est défini comme une zone du bassin pour laquelle l'épaisseur sédimentaire est supérieure à 300m et dont la teneur en sable est supérieure à 25% ou à 40%.



Sans grande surprise les résultats montrent que le delta, le canyon et les pieds de pente de la marge alimentée par des apports sédimentaires sont les zones les plus favorables au respect du critère sur les épaisseurs. Ils permettent toutefois de précisément cartographier la probabilité de respecter les critères de définition d'un réservoir. Le respect du critère sur la teneur en sable est particulièrement fort au cœur du système deltaïque. Les autres zones propices à l'accumulation de sable sont le reste de l'ensemble du système deltaïque, le canyon et les dépocentres en sortie du canyon et au pied de la marge. La carte finale de probabilité de présence d'un réservoir avec les critères retenus ici pour sa définition montre que la proportion de sable apparaît le principal facteur limitant. Des critères supplémentaires prenant en compte la manière dont les bancs sableux s'intercalent dans les dépôts argileux pourraient également être ajoutés afin de tenir compte d'autres critères de qualité du réservoir.

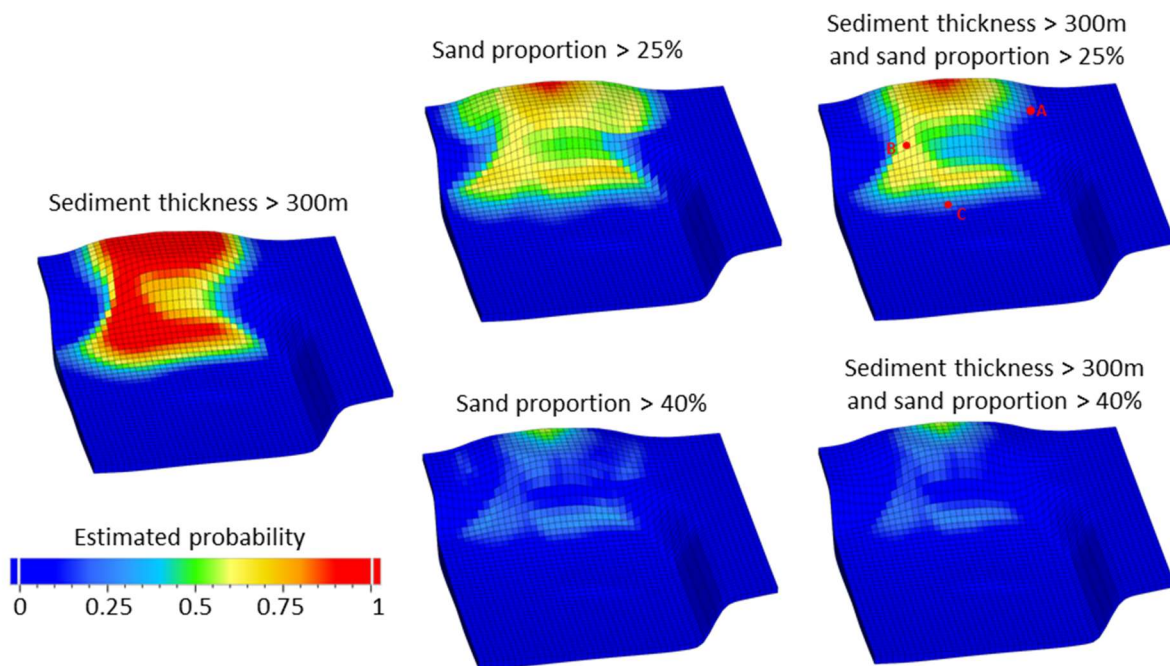


Figure 28 : Carte des probabilités de présence d'un réservoir estimée à partir d'un échantillon de 20000 valeurs des méta-modèles obtenus dans la section précédente : À gauche, probabilités que le play ait une épaisseur > 300m ; au milieu, probabilités que le play ait une teneur en sable > 25% (haut) ou 40% (bas) ; à droite, probabilités combinées que le play ait à la fois une épaisseur > 300m et une teneur en sable > 25% (haut) ou 40% (bas). Trois puits fictifs A, B et C ont été positionnés (image en haut à droite) afin de discuter plus en détails ces différents cas de figure.

L'analyse de risque et l'étude de la dynamique du bassin peuvent être poussées un peu plus loin en observant les distributions *a posteriori* des paramètres incertains correspondant à la présence de réservoirs dans les différentes zones du bassin. Trois puits fictifs A, B et C (figure 28) ont été positionnés respectivement à la bordure est du delta, dans le canyon et en limite des principaux dépôts au pied de la marge dans l'axe de la source des apports sédimentaires. L'interprétation de ces distributions de probabilités donne des résultats complémentaires à ceux précédemment obtenus avec l'analyse de sensibilité présentée à la section précédente (figure 29, figure 30 et figure 31). Ils montrent notamment que 4 des 9 paramètres incertains suivent des distributions de probabilités pratiquement uniformes, quel que soit le puits considéré (l'amplitude et la période des variations eustatiques, le taux de subsidence et le rapport entre la capacité de transport en domaine marin et en domaine continental). Ces paramètres ne jouent donc qu'un faible rôle dans la probabilité de présence d'un réservoir à ces positions. À l'inverse, la proportion de sable dans les apports sédimentaires joue un rôle majeur et régional sur la probabilité de présence d'un réservoir. Il est également intéressant de noter que les incertitudes sur le

flux d'eau et sur la capacité de transport de référence ne semblent pas influencer sur la probabilité de présence d'un réservoir dans le canyon (figure 30), mais elles jouent un rôle important dans le cas d'un puit positionné sur la marge du delta ou en bordure du dépo-centre (figure 29 et figure 31). Là encore des différences apparaissent car si le flux d'eau semble favoriser la présence de réservoir dans le bassin profond (les réservoirs ne sont observés que si le flux d'eau est élevé), il est contre-productif dans la partie deltaïque du système. Il en est de même mais dans une moindre mesure pour le ratio de capacité de transport entre les sables et les argiles.

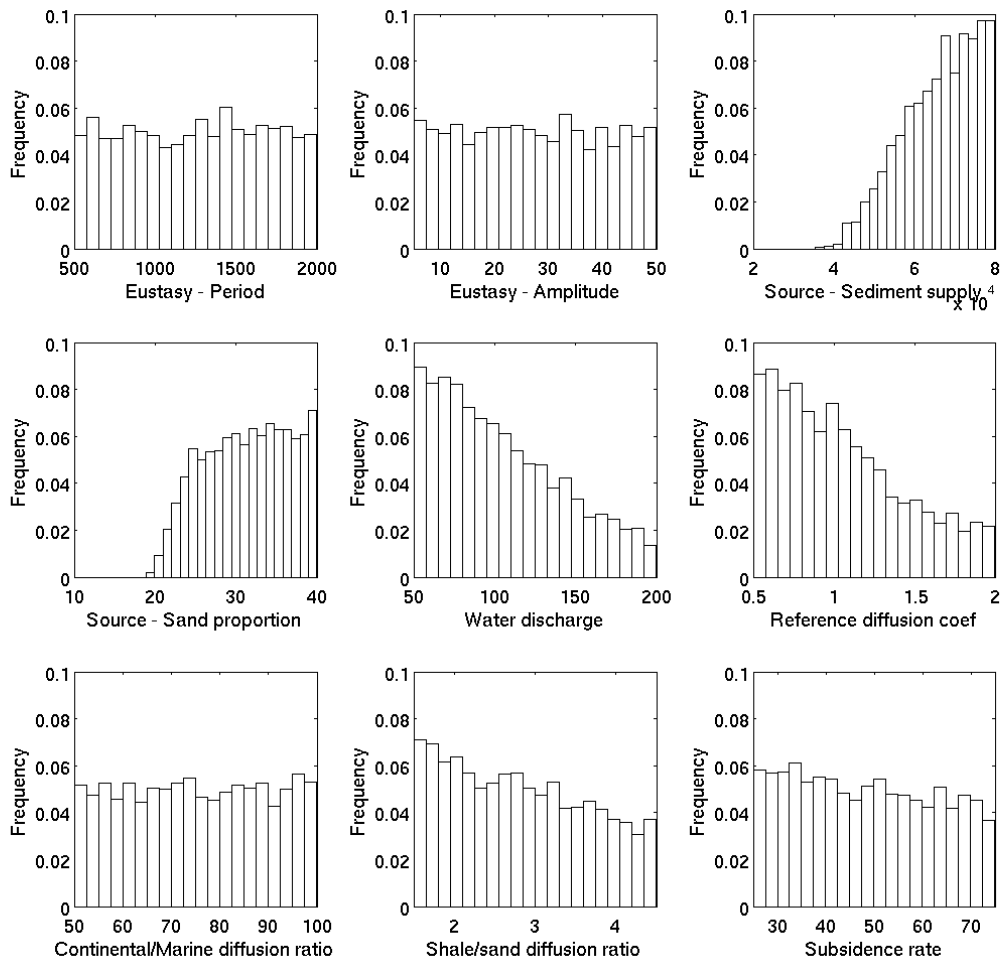


Figure 29 : Distribution des valeurs des paramètres incertains permettant de vérifier les critères de définition d'un réservoir (épaisseur > 300m et proportion de sable > 25%) pour le puits fictif A. Les distributions uniformes des variations eustatiques, des rapports de coefficients de diffusion entre domaines marin et continental et du taux de subsidence montrent leur faible effet sur la présence d'un réservoir. À cet endroit de la marge du delta, la présence de réservoirs est principalement associée à de forts apports sédimentaires et une forte proportion de sable. Elle est également préférentiellement reliée à de faibles flux d'eau, un transport faible des sédiments et notamment des argiles.

Enfin, un dernier comportement est observé avec l'incertitude sur l'apport sédimentaire global, qui ne semble pas avoir d'effet dans le canyon mais joue un rôle similaire important à la fois à la marge du delta et dans le bassin profond. Cette observation est probablement liée au fait que plus les apports sédimentaires sont élevés, plus ils vont se répandre sur une large zone permettant de trouver des réservoirs plus éloignés de la source. Le canyon n'est, lui, pas affecté car il constitue une dépression proche de la source d'apports sédimentaires qui se remplit quels que soient les apports de sédiments considérés.

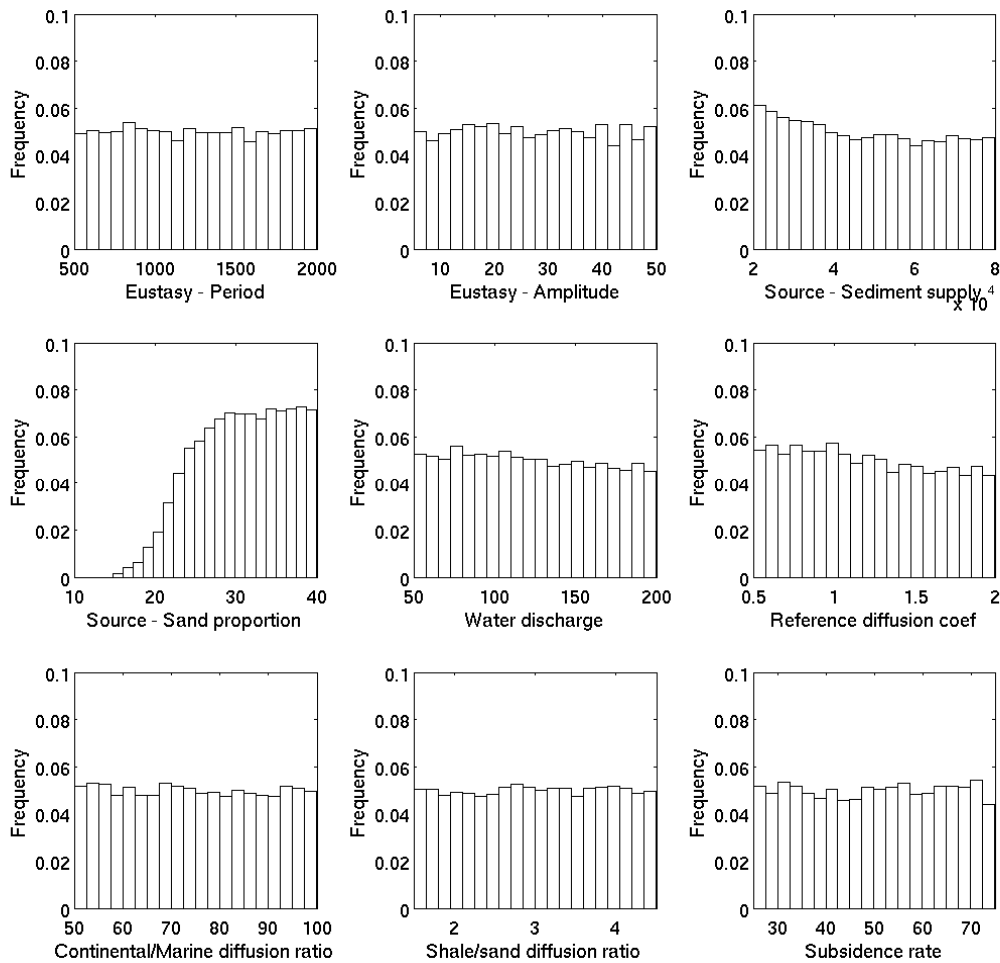


Figure 30 : Distribution des valeurs des paramètres incertains permettant de vérifier les critères de définition d'un réservoir (épaisseur > 300m et proportion de sable > 25%) pour le puits fictif B. Les distributions uniformes des variations eustatiques, des flux d'eau, du taux de subsidence et du transport des sédiments de manière globale montrent leur faible effet sur la présence d'un réservoir. Dans le canyon, la présence de réservoirs semble uniquement contrôlée par la proportion de sable dans les apports sédimentaires.

Si on souhaite approfondir l'analyse, on peut également s'intéresser à la forme des distributions de probabilité. On peut par exemple noter que la proportion de sable dans la source semble même agir comme un déclencheur de la possibilité d'obtention d'un réservoir pour la bordure du delta : en dessous de 20% de sable dans les apports, il n'est pas possible d'observer de réservoirs alors qu'après 25% il n'y a plus de différence (figure 29). On retrouve ce même seuil de 25% dans le canyon (figure 30) mais les effets y sont plus lissés : des réservoirs peuvent être observés même pour de faibles apports en sable. Enfin, dans le bassin (figure 31), on trouve une progression constante des apports en sable : plus la proportion de sable augmente plus on a de chance d'avoir des réservoirs. Pour le bassin, les apports sédimentaires vont jouer cet effet de seuil : en dessous de 26000km<sup>3</sup>/Ma il n'y a pas de formation possible de réservoir alors qu'il y a plus d'effet de l'apport sédimentaire au-dessus de 45000km<sup>3</sup>/Ma. Ce même paramètre a un effet sensiblement différent sur le bord du delta où aucun réservoir n'est observé pour des apports inférieurs à 28000km<sup>3</sup>/Ma mais où la présence des réservoirs augmente de façon proportionnelle avec l'apport sédimentaire.

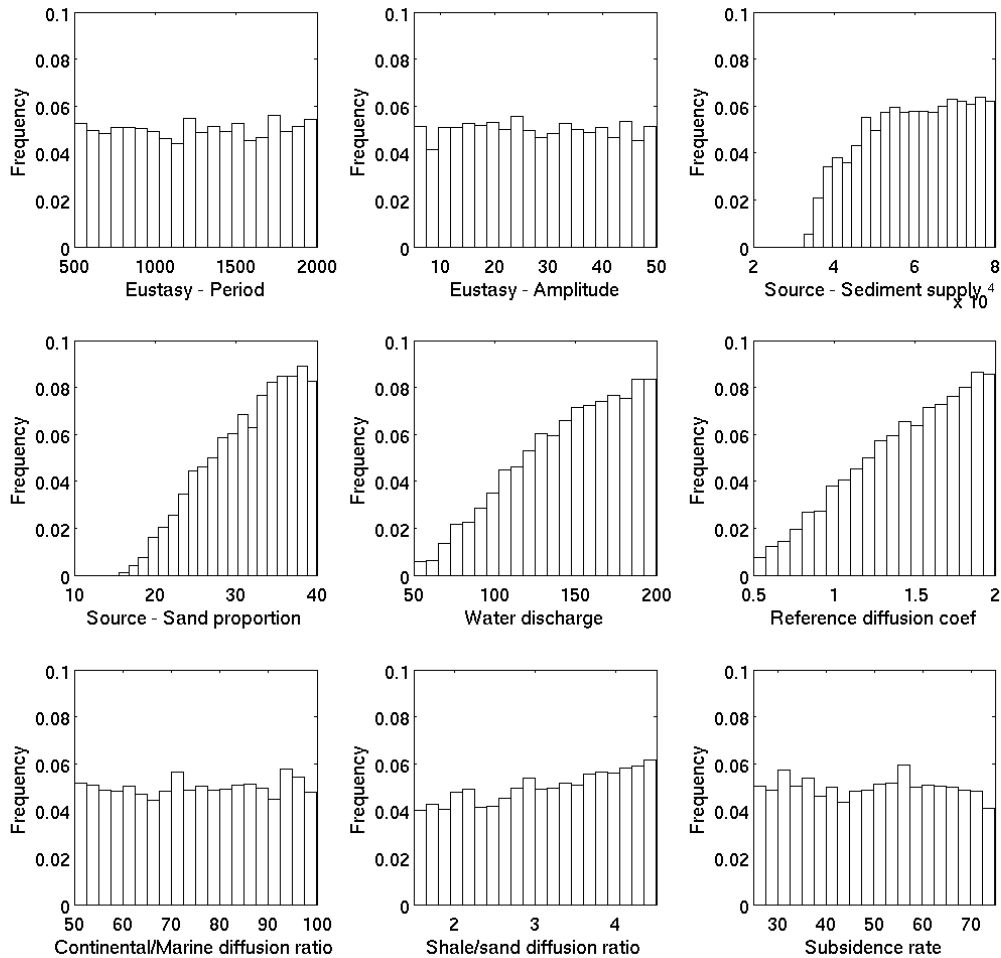


Figure 31 : Distribution des valeurs des paramètres incertains permettant de vérifier les critères de définition d'un réservoir (épaisseur > 300m et proportion de sable > 25%) à l'endroit du puits fictif C. Comme pour les abords du delta (figure 29), les distributions uniformes obtenues pour les variations eustatiques, les rapports de coefficients de diffusion entre domaines marin et continental et entre sable et argile ainsi que le taux de subsidence montrent leur faible effet sur la présence d'un réservoir. Dans cette zone du bassin profond, la présence de réservoirs est préférentiellement associée à de forts apports sédimentaires et une forte proportion de sable. Contrairement au delta, elle est également préférentiellement reliée à de forts flux d'eau et un coefficient de transport de référence élevé.

Ces distributions de probabilités *a posteriori* montrent également comment les données disponibles peuvent être utilisées pour mettre à jour les probabilités de présence d'un réservoir. Les données contraignant les modèles (épaisseurs mesurées par exemple) peuvent en effet modifier les distributions de probabilité de présence d'un réservoir en filtrant les réalisations non représentatives des mesures ou en pondérant les différentes réalisations par leur vraisemblance (approche bayésienne).

## 2.2.4 Détermination des risques de maturité de la roche-mère

La maturité de la roche-mère constitue un autre élément essentiel d'un système pétrolier. Lors de l'analyse des risques d'exploration pétrolière associés à un bassin sédimentaire, les explorateurs vont d'une part chercher à déterminer la zone de présence de la roche-mère mais également la maturité de cette dernière. En effet, ce paramètre, en combinaison avec le volume de la roche-mère et son potentiel pétrolier (sa capacité à générer des hydrocarbures) est une information directe sur la quantité d'hydrocarbures dans le système pétrolier.

La matière organique présente dans la roche-mère se transforme en hydrocarbures sous l'effet du temps et de la température (ex : McNab *et al.*, 1952; Pitt, 1961; Philippi, 1965 ; Louis et Tissot, 1967 ; Tissot, 1969). Cette transformation de la matière organique, dite de craquage, est contrôlée par des paramètres cinétiques. La plupart du temps, une formulation de type Arrhenius est employée pour modéliser cette transformation fonction de la température et du temps (ex : Ungerer *et al.*, 1986).

### Détermination de l'incertitude sur la cinétique de craquage de la matière organique

Il est possible d'essayer de déterminer les paramètres de la cinétique de craquage de la matière organique à partir d'analyses réalisées avec le Rock-Eval sur des échantillons de roche-mère immature (avant toute réaction de transformation de la matière organique). Une cinétique de transformation d'ordre 1 de la matière organique s'exprime sous la forme générique suivante :

$$\text{Équation 24: } \frac{dx}{dt} = -kx$$

où  $x$  représente la fraction restante de matière transformable,  $t$  représente le temps et  $k$  la constante de réaction.  $k$  est une fonction de la température qui s'exprime sous la forme de l'équation d'Arrhenius :

$$\text{Équation 25: } k = A \cdot \exp(-E/RT)$$

où  $E$  représente l'énergie d'activation,  $A$  est le facteur de fréquence de la réaction et  $T$  est la température. Ces équations peuvent être plus ou moins raffinées (cinétiques discrètes multi-énergétique, distribution gaussienne des énergies...) selon les approches considérées (se référer à Ducros, 2014, pour plus de détail).

La détermination des paramètres de la cinétique (c.à.d. les valeurs de  $A$  et  $E$ ) se fait généralement par calage des données obtenues lors des analyses Rock-Eval sur des échantillons immatures. Pour ce faire, on cherche à résoudre un problème inverse. Il apparaît généralement une dépendance entre  $A$  et  $E$  ne permettant pas d'obtenir une unique solution du problème d'optimisation. Cette dépendance est principalement due aux faibles amplitudes des rampes de pyrolyse utilisées dans les expériences de laboratoire (rampes de chauffe de 1 à 50°C/s en laboratoire contre des taux de chauffe allant jusqu'à 50°C/Ma dans les bassins sédimentaires) et à la différence entre la gamme de températures utilisée et celle du milieu géologique (températures de 250°C à 700°C en laboratoire contre 60°C à 200°C dans les bassins sédimentaires). Cette contrainte peut toutefois difficilement être levée du fait de la lenteur des réactions en jeu. Dans les bassins sédimentaires, les effets sont sensibles du fait de la durée « géologique » de l'ordre de plusieurs dizaines de millions d'années. Dans la pratique et en l'absence de contraintes additionnelles, les explorateurs de bassin ont opté pour une solution consistant à fixer, sur la base de leur expérience, la valeur de l'un des deux paramètres, généralement  $A$ , pour optimiser l'autre (Ungerer, 1989). Les conséquences de ces indéterminations sur les paramètres de la cinétique de transformation sont que les résultats peuvent être extrêmement différents pour les conditions de températures observées dans les conditions géologiques. Dans les années 1990, plusieurs auteurs ont étudié en détail la source de cette indétermination et les conséquences sur la quantité d'hydrocarbures générée par les roches-mères (ex : Nielsen et Dahl, 1991 ; Nielsen, 1993).

Dans ce travail nous avons utilisé les données obtenues par Bou Daher *et al.* (2014) qui a réalisé des analyses Rock-Eval sur la roche mère du Campanien du bassin du Levant dans l'est de la Méditerranée. Au lieu de fixer, comme cela est fait classiquement, l'un des deux paramètres de la cinétique, nous utilisons un outil (CougarFlow, commercialisé par Beicip-Franlab) permettant d'obtenir l'ensemble des couples ( $A$ ,  $E$ ) solutions du problème d'optimisation. Par la suite, nous analysons l'effet de l'incertitude sur la détermination de ces paramètres pour estimer le risque sur la maturité de la roche mère. La méthode choisie consiste à calculer avec Cougar un méta-modèle par processus gaussiens (cf. figure 32)

approchant la fonction d'erreur (équation 26) intégrant l'ensemble des analyses Rock-Eval (5 rampes de chauffe) :

$$\text{Équation 26: } FO = \sum_{\text{chauffés}} \sum_{\text{mesures}} (TR_{\text{mesuré}} - TR_{\text{simulé}})^2$$

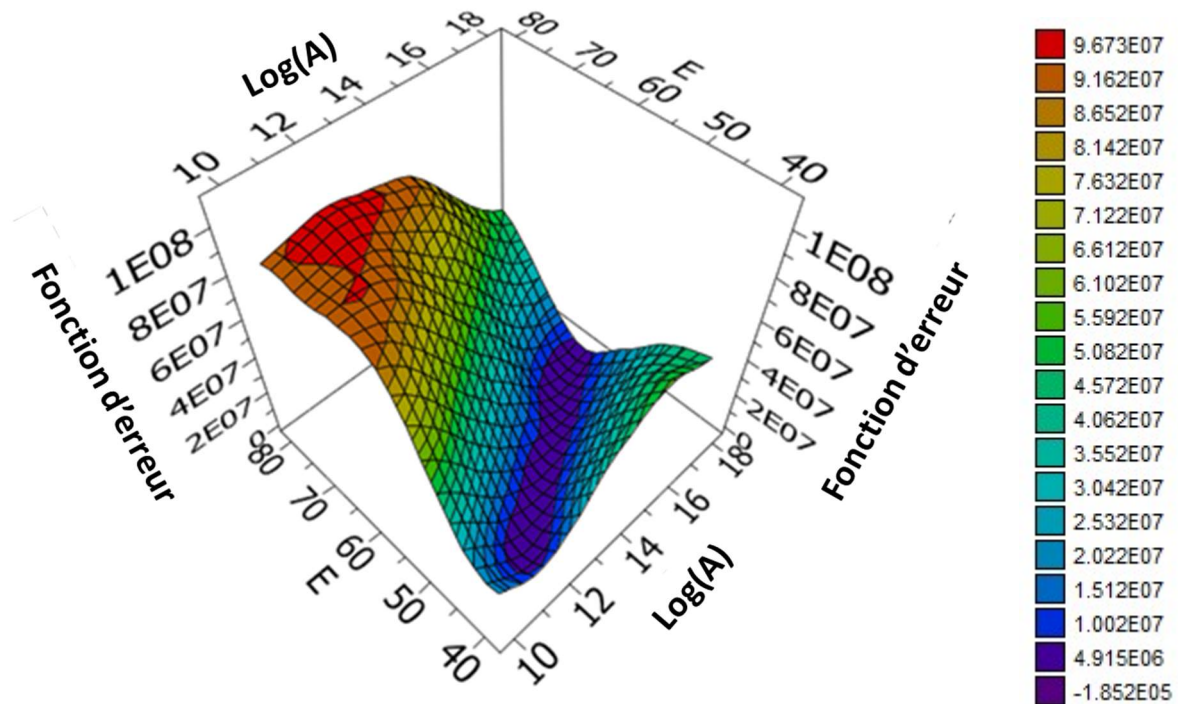


Figure 32 : Surface de réponse de la fonction coût pour la cinétique de transformation de la roche mère représentée en fonction des valeurs de A et de E. Plus la valeur est élevée et plus il y a de différence entre les simulations de transformation de la matière organique et les valeurs mesurées.

L'estimation de la fonction coût par la surface de réponse sur l'ensemble du domaine incertain est de bonne qualité (cf. tableau 9) selon les indicateurs fournis. 51 simulations ont été utilisées pour le calcul de la fonction coût et 49 simulations ont été utilisées pour la confirmation.

Indicateur	Valeur
Q2	0.984
Q2 (confirmation)	0.988

Tableau 9 : Indicateurs de la qualité de la surface de réponse

Le méta-modèle peut alors être utilisé pour déterminer les couples solutions du problème d'optimisation. Son analyse montre bien qu'il existe une « vallée » de solutions (zone en mauve sur la figure 32). On pourra en première approximation (et pour des raisons de simplification) considérer que la relation entre  $\log(A)$  et  $E$  est linéaire pour les couples solutions du problème d'optimisation et définie par :

$$\text{Équation 27 : } \log(A) = 0,3284 \times E - 2.8181$$

### Détermination des cartes de risques sur la maturité de la roche-mère

Une fois connue l'incertitude sur les paramètres de la cinétique, il est possible de la répercuter sur le calcul de la maturité de la roche-mère dans le bassin sédimentaire et de déterminer les risques associés. A partir d'un modèle numérique du bassin du Levant correspondant à une zone de 315km x 315km (figure 33, Bou Daher *et al.*, 2016), l'incertitude sur les paramètres de la cinétique est propagée sur la détermination de la maturité de la roche mère du campanien. Dans ce cas, la cinétique de la roche-mère du Campanien est l'unique paramètre incertain inclus dans l'analyse. Une méthode d'échantillonnage par hypercube latin est utilisée pour déterminer les 15 simulations réalisées avec TemisFlow, outil de modélisation de système pétrolier développé par IFPEN. Elles sont utilisées pour la construction du méta-modèle. Un second échantillonnage indépendant de 18 simulations est réalisé pour estimer la qualité du modèle obtenu.

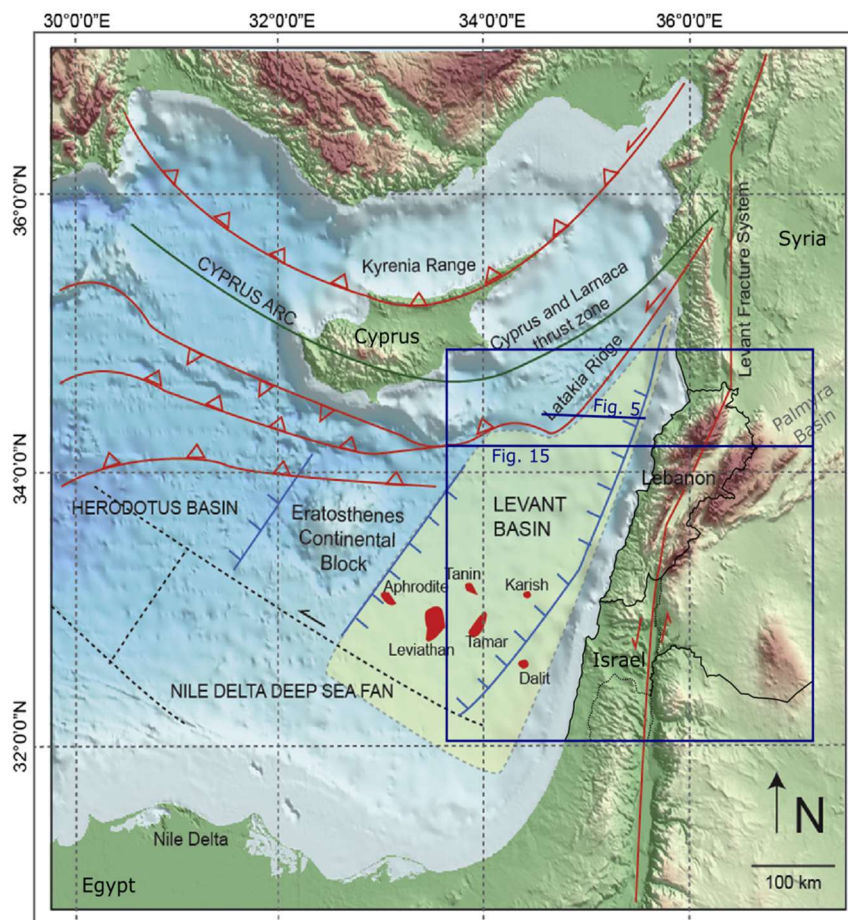


Figure 33 : Localisation du modèle du bassin du Levant pour l'estimation de la maturité de la roche-mère du Campanien (modifié de Hawie *et al.*, 2013)

Nombre de composantes	Part expliquée de la variance
2	93,6
3	97,7
4	99,2
5	99,7
6	99,9

Tableau 10 : Part de variance expliquée en fonction du nombre de composantes retenues.

Le nombre de composantes dans la base réduite est sélectionné en fonction de la part de la variance expliquée. Le tableau 10 en donne l'évolution en fonction du nombre de composantes retenues. On choisit d'utiliser trois composantes, ce qui correspond à une explication de la variance de près de 98%.

Le méta-modèle est ensuite construit sur la base de ces trois vecteurs. La qualité du méta-modèle obtenu est représentée à la figure 34 sous la forme des critères de qualité  $Q2(troncature)$ , lié à l'erreur de troncature, et  $Q2(confirmation)$ , estimant la qualité de la prévision du modèle. La carte présentée à la figure 34(a), avec ses grandes zones de variance faible voire nulle, montre qu'une grande partie du bassin n'est pas affectée par l'incertitude sur la cinétique de transformation de la matière organique. Les principales zones pour lesquelles la maturité dépend fortement de cette incertitude sont les pentes de la marge continentale à l'est du bassin (centre du modèle), la zone de la mer morte au sud-est et le Latakieh Ridge au nord-ouest en direction de Chypre. Ces régions correspondent à des zones où la roche-mère du Campanien se trouve à un enfouissement intermédiaire : suffisamment enfouie pour atteindre des températures permettant un début de transformation et suffisamment peu enfouie pour que la transformation ne soit pas toujours complète. Les étroites zones bleues à l'est du bassin visibles sur la carte de la figure 34(c) correspondent aux reliefs du Mont Liban : la roche mère étudiée y a été érodée et n'est plus présente. Les larges zones bleues à l'ouest correspondent au bassin profond où la valeur de transformation de la matière organique est constante et égale à 100% quelle que soit la valeur du le paramètre incertain (figure 36).

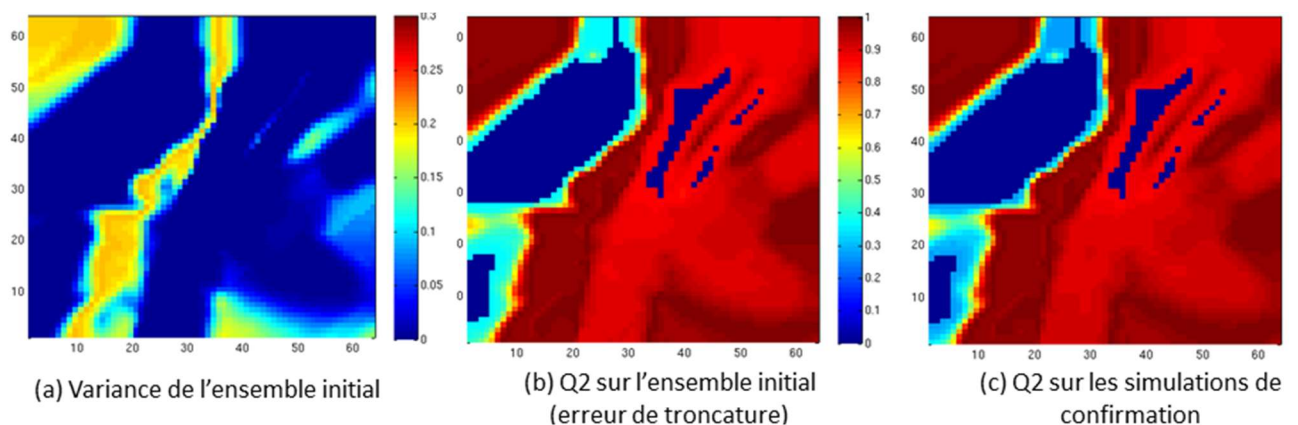


Figure 34 : (a) Carte de la variance de la réponse taux de transformation sur l'ensemble des 15 simulations, (b) Critère de qualité  $Q2(troncature)$  sur l'échantillon ayant servi à construire les modèles, (c) Critère de qualité  $Q2(confirmation)$  sur les simulations de confirmation. On voit nettement apparaître des zones de variance nulle correspondant à des zones toujours matures (100% de transformation) dans le bassin profond à l'ouest ou à des zones toujours immatures (0% de transformation) peu enfouies dans la partie continentale à l'est.

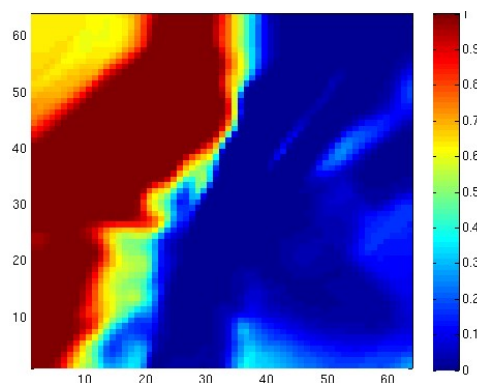


Figure 35 : Carte de la moyenne de la réponse sur l'ensemble des 15 simulations.



La figure 34(a) montre que le méta-modèle est globalement de bonne qualité avec un  $Q2(\text{confirmation})$  presque toujours supérieur à 0,9 dans les zones de variance non négligeable. Une fois le modèle approché validé, il est possible de calculer la distribution des réponses dans chacune des cellules de la carte. Il est ainsi possible de déterminer des cartes de percentiles (cf. figure 36) du taux de transformation de la matière organique.

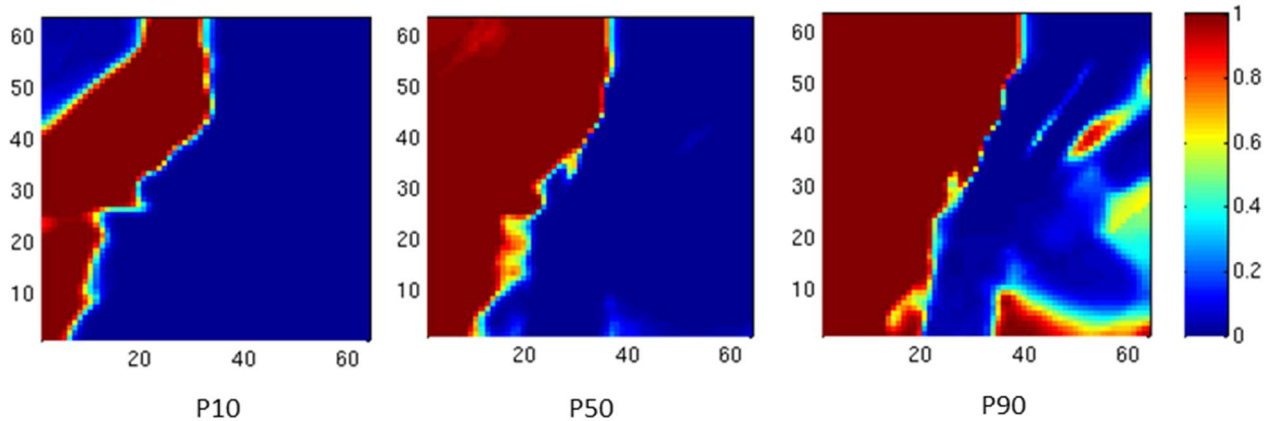


Figure 36 : Carte des percentiles P10 (gauche), P50 (centre) et P90 (droite) du taux de transformation estimée à partir du méta-modèle déterminé avec le lot de 15 simulations. La zone étudiée fait 315km x 315km. Le fait que même dans le cas le plus favorable la transformation de la partie côtière de la zone continentale à l'est du bassin soit nulle indique que l'incertitude sur la cinétique de transformation de la matière organique n'a aucun effet sur l'évaluation du niveau de maturité de la roche-mère dans cette zone. Il en est de même dans le bassin profond où même dans le cas le plus défavorable la transformation est complète. En revanche la cinétique va avoir un effet sur l'âge auquel la roche-mère génère les hydrocarbures, ce qui peut modifier l'analyse du système pétrolier

La figure 36 montre clairement que le bassin profond a généré l'ensemble de son potentiel pétrolier quelle que soit l'incertitude sur la cinétique (gauche). A l'opposé, la majeure partie de la plateforme continentale à l'est est immature. Enfin, il existe des zones, telles que les mini-bassins sur les contreforts du Mont Liban, la région de la mer Morte et la zone du Latakieh Ridge, où l'incertitude sur la cinétique contrôle le degré de maturité de la roche-mère. Par conséquent, dans ces zones, l'estimation des hydrocarbures générés par la roche-mère est plus risquée. Dans le bassin profond, même si l'incertitude sur la cinétique de transformation de la matière organique n'a aucun effet sur l'évaluation de son niveau de maturité, elle va en avoir un sur l'âge auquel la roche-mère génère les hydrocarbures, ce qui peut modifier profondément l'analyse du système pétrolier.

## 2.2.5 Synthèse

Les deux études présentées dans ce travail montrent l'intérêt de prendre en compte les incertitudes dans les études d'exploration afin de mieux délimiter les zones pouvant présenter un intérêt pétrolier. Ces études utilisent la puissance des outils numériques de modélisation géologique afin de pousser plus loin les analyses sur les risques liés à l'exploration pétrolière, dans un cadre rigoureux et sous des formes adaptées aux études classiques de risques. Elles illustrent en effet l'intérêt des réponses fonctionnelles par rapport aux méthodes utilisées jusqu'à présent, qui ne permettent de traiter que des valeurs scalaires, pour leur intégration dans les démarches courantes d'analyse des risques en exploration pétrolière (type CRS mapping). Les résultats obtenus montrent que l'analyse des incertitudes à partir de ces modèles est applicable sur les cas traités (bonne qualité des modèles approchés) et avec un faible nombre de simulations requis pour réaliser les études. Ils ouvrent également la voie à de nouvelles façons

d'interpréter les résultats de simulation qui peuvent être plus proches des méthodes traditionnelles des explorateurs pétroliers.

La première étude a permis de cartographier les zones d'un bassin sédimentaire les plus favorables à la présence d'un réservoir pétrolier sur les bases des incertitudes géologiques sur le remplissage sédimentaire du bassin. L'analyse des distributions de probabilité *a posteriori* de présence d'un réservoir a donné également accès à des informations géologiques importantes sur les environnements de dépôt de ces réservoirs. La deuxième étude a permis de déterminer précisément les zones du bassin concernées par un système pétrolier effectif et celles où la roche-mère n'a jamais pu générer d'hydrocarbures. Elle a également mis en évidence des zones plus risquées où l'incertitude sur la détermination de la cinétique de transformation de la roche-mère ne permet pas de connaître précisément les volumes d'hydrocarbures qui ont pu être générés. Elle permet ainsi de faire apparaître le risque sur le degré de maturité de la roche-mère en lien avec la détermination du modèle cinétique. Cette étude correspond toutefois à un instantané du risque sur la maturité. La cinétique a également des conséquences sur la période à laquelle les hydrocarbures sont générés. En allant un peu plus loin dans l'analyse, il serait possible de déterminer le risque sur la quantité d'hydrocarbures générée après la mise en place d'un réservoir et du piège associé.

Plusieurs travaux peuvent être envisagés pour approfondir l'évaluation des possibilités ouvertes par cette méthode de traitement des incertitudes dans les études en exploration pétrolière. La première étude a par exemple traité le cas d'un réservoir pétrolier clastique. Or, les outils de modélisation stratigraphique sont souvent utilisés pour déterminer les distributions lithologiques dans un bassin sédimentaire. Elles sont ensuite transférées dans un outil de modélisation de système pétrolier. Cette interprétation en lithologies se fait, comme pour la définition d'une « lithologie réservoir » que nous avons illustrée, à partir de combinaison de critères sur les résultats bruts du modèle. La méthode pourrait être généralisée à l'interprétation de l'ensemble des lithologies faite par les utilisateurs d'outils de modélisation stratigraphique.

Dans les deux études présentées, aucune donnée géologique de calibration n'a été utilisée pour pondérer les résultats de l'analyse de risque. L'utilisation d'approches bayésiennes, et les calculs de vraisemblance associés, pourrait être envisagée afin de déterminer les cartes de risque sous la contrainte des données disponibles (par exemple épaisseur effective des réservoir ou niveau de maturité de la roche-mère mesurée dans les puits de forage respectivement pour la première et la deuxième étude). Enfin, la méthode a été appliquée à des résultats se présentant sous la forme de cartes mais elle pourrait également être appliquée pour la prévision de la pression avant forage d'un puits (application de la méthode à une courbe de pression plutôt qu'à une carte). Ce type d'information pourrait être transmise aux foreurs. Ceux-ci bénéficieraient alors d'une meilleure estimation des incertitudes sur la pression au cours du forage, ce qui réduirait potentiellement les risques associés aux forages et qui sont sources de coûts importants.