

# Conclusions

---

L'objectif de cette thèse était double. Il s'agissait dans un premier temps, de caractériser le sous-espace d'erreur d'un modèle barotrope implémenté sur la zone du Golfe de Gascogne et de la Manche, en réponse aux incertitudes sur les forçages atmosphériques ; dans une seconde étape, nous avons pour objectif de proposer une méthodologie robuste permettant de contrôler le sous-espace d'erreur du modèle en utilisant les statistiques d'erreurs sous forme d'EOFs d'ensemble. La méthodologie, ainsi que la sensibilité de cette dernière à plusieurs paramètres majeurs de la configuration ont été testées dans le cadre d'expériences jumelles.

Dans une phase préliminaire, une partie du travail fut consacrée à la caractérisation des principaux processus barotropes induits par la réponse dynamique de l'océan au forçage météorologique. Pour ce faire, nous avons mis en place une configuration réaliste d'étude de cette dynamique, en utilisant le modèle barotrope MOG2D, implémenté sur la zone du Golfe de Gascogne et la Manche, et validé par comparaison aux observations marégraphiques disponibles sur les côtes espagnoles et françaises. La dynamique haute fréquence forcée par l'atmosphère a été ensuite détaillée durant la période d'étude, courant du 16 novembre au 1<sup>er</sup> décembre. Celle-ci est très variable à l'échelle du domaine : dans le Golfe de Gascogne, nous avons montré que la réponse océanique en élévation de surface était principalement pilotée par les variations de pression atmosphérique, alors qu'elle est essentiellement contrôlée par le vent dans la Manche et le long des côtes sud-ouest irlandaises. Des propagations d'ondes de Kelvin ont aussi été mises en évidence dans la Manche, générées par des variations brutales du champ de pression et de vent associées aux passages de perturbations atmosphériques dans le Nord du domaine.

La réponse océanique à la pression et au vent ainsi caractérisée, nous nous sommes intéressés à l'impact statistique des incertitudes des forçages atmosphériques sur cette dynamique barotrope. Notre premier objectif se résumait à la question suivante :

- **Comment caractériser les erreurs du modèle barotrope MOG2D (implémenté en**

### zone côtière) en réponse aux incertitudes sur les forçages atmosphériques ?

Pour y répondre, nous avons mis en place une méthode d'ensemble. Reprenant l'approche utilisée par Auclair *et al.* (2003), nous avons généré un ensemble de 300 perturbations atmosphériques, obtenues par combinaisons linéaires aléatoires des 10 principaux modes (EOFs) de variabilité spatio-temporelle du champs de forçage météorologique de référence. Nous avons construit par cette méthode un ensemble *a priori* de perturbations atmosphériques réalistes et dynamiquement cohérentes avec les champs de référence. Les erreurs du modèle ont ensuite été décrites par le biais de statistiques calculées à partir de l'ensemble *a posteriori* de simulations du modèle résultant. Nos questions plus spécifiques concernant les statistiques d'erreur étaient les suivantes :

- **Quelles sont les structures spatiales des erreurs étudiées ?**
- **Dans quelle mesure les statistiques d'erreurs obtenues vont-elles évoluer dans le temps ?**
- **Les hypothèses simplificatrices communément admises en zone hauturière sont-elles applicables dans notre configuration ?**

Tout d'abord, nous avons montré que l'**hypothèse d'homogénéité** des statistiques des erreurs étudiées n'est **pas applicable** dans notre étude en zone côtière et de plateau : la variance d'erreur (approchée par la variance d'ensemble) de l'élévation de surface atteint sa valeur maximale dans la Manche, et apparaît presque entièrement contrôlée par les erreurs du champ de vent ; elle est beaucoup plus faible dans le Golfe de Gascogne, et principalement générée par les erreurs de pression, elles-mêmes relativement faibles par construction. En ce qui concerne les courants barotropes, l'impact des erreurs atmosphériques sur la modélisation des vitesses s'est révélé différent dans la partie profonde et sur le plateau : les erreurs associées sont essentiellement localisées dans la frange côtière peu profonde où elles sont fortement pilotées par les erreurs dans la tension de vent, et quasiment nulles en zone plus profonde. D'autre part, les statistiques obtenues ont clairement montré une **non-stationnarité** des erreurs modélisées. Elles mettent en évidence des **régimes d'erreur** évoluant rapidement, et fortement corrélés aux développements des erreurs dans les forçages météorologiques. Les durées typiques de ces régimes d'erreur sont de l'ordre de 24-36 heures pour l'élévation de surface, et 12-36 heures pour les vitesses barotropes. Enfin, nous avons décrit ces structures d'erreurs sous forme de structures covariantes d'erreurs, par le biais d'EOFs d'ensemble multivariés (les variables océaniques et atmosphériques sont considérées). Ces modes de variabilité de l'erreur du modèle, calculés à partir d'états de l'ensemble modèle à différents instants (afin de prendre en compte le caractère instationnaire de l'erreur), ont permis de décrire l'intense régime d'erreur se développant en Manche, mais aussi de souligner la complexité des structures d'erreurs étudiées dans les zones proches des côtes et/ou du talus. De plus, ces structures d'erreurs covariantes se sont avérées **anisotropes**. Cette étude confirme ainsi que **les hypothèses habituellement faites sur les caractéristiques des erreurs dans les processus dynamiques hauturiers ne sont pas applicables en zone côtière et de plateau**.

Dans une deuxième étape, les statistiques d'erreurs obtenues sous forme d'EOFs d'ensemble multivariés ont été utilisés dans le schéma d'assimilation de données séquentiel SEQUOIA, couplé au modèle océanique, dans le but de contrôler les erreurs de ce dernier. Les performances de ce système pour corriger et contraindre l'erreur du modèle induite par les incertitudes dans les forçages atmosphériques ont été étudiées dans un cadre d'expériences jumelles. Dans ce contexte, des observations sont extraites d'une simulation de *contrôle* et assimilées dans une simulation *libre* (ces deux simulations sont tirées de l'ensemble de trajectoires perturbées du modèle). Par conséquent, les structures d'erreur que l'on cherche à corriger (représentées synthétiquement par la différence *libre-contrôle*) sont plus 'riches' que celles décrites par la base d'EOFs.

Tout d'abord, nous nous sommes attachés à calibrer le système d'assimilation SEQUOIA/MOG2D au moyen de tests de sensibilité à différents paramètres de la configuration. Les questions spécifiques soulevées en introduction à ce manuscrit étaient les suivantes :

- **Quels types de diagnostics peut-on utiliser pour caractériser les performances du système ?**
- **L'utilisation des statistiques d'erreur calculées sous forme d'EOFs d'ensemble permet-elle de contraindre efficacement le modèle ?**
- **Dans quelle mesure des statistiques d'erreur indépendantes du temps peuvent-elles permettre de contrôler le développement de ces erreurs ?**
- **Est-il suffisant de corriger les seules variables océaniques pour assurer un contrôle efficace de ces erreurs ?**

Un **jeu de diagnostics** d'assimilation de données, basés sur des critères statistiques, a préalablement été choisi pour mener à bien ces tests. Certains utilisent des critères **adaptés au cadre des expériences jumelles**, tels que la mesure de distance *rms* (resp. la corrélation) entre les trajectoires *assimilée* et *contrôle*, permettant de qualifier l'efficacité de la correction par comparaison à la distance *rms* (resp. la corrélation) entre *libre* et *contrôle*. D'autres, comme la méthode dite du "*J<sub>min</sub>*" basée sur les travaux de Talagrand (1998) et Sadiki (2005), permettent de qualifier la **cohérence intrinsèque du schéma d'assimilation**. L'ensemble de ces diagnostics s'est avéré utile pour qualifier la performance du système.

Dans une expérience d'assimilation de référence, pour laquelle les données étaient simulées à partir d'un réseau de 10 marégraphes existants, nous avons pu valider le système. Celui-ci permet de réduire l'erreur de façon significative, pour toutes les variables du vecteur d'état. Nous avons constaté, au travers des différentes expériences de sensibilité menées, que ces résultats étaient sensibles à certains paramètres clés. Il est apparu **nécessaire de contraindre non seulement les variables océaniques mais aussi atmosphériques**, pour assurer un contrôle efficace de l'erreur. Ce résultat nous confirme ainsi la très forte sensibilité du modèle aux conditions atmosphériques et aux erreurs associées. D'autre part, nous avons noté qu'en dépit de ces corrections satisfaisantes, la croissance de l'erreur demeurait significative entre deux analyses ; dans un test de sensibilité à la fréquence d'assimilation et aux performances en prévision,

nous avons pu constater que le système assurait un contrôle de l'erreur efficace en prévision à court-terme, mais plus limité à moyen terme : **l'horizon de prévision a été mesuré à environ 30 heures**, dans la configuration actuelle du système. Cependant, cette limite semble être conditionnée par le temps de persistance de la correction des variables atmosphériques, fixé à 24 heures dans les expériences. Enfin, la question critique de la dépendance temporelle des statistiques d'erreurs utilisées dans le système d'assimilation a été soulevée : en regard de la forte non-stationnarité des variances d'erreurs du modèle, il n'était pas assuré qu'une base d'EOFs indépendante du temps (mais contruite à partir d'échantillons pris en plusieurs instants de l'ensemble) puisse représenter efficacement les structures d'erreurs se développant au cours de la simulation, et par la même les contrôler. Une **base d'EOFs d'erreurs 'améliorée'**, calculée préalablement aux expériences et spécifique de chacun des instants d'analyse, permet certes des **corrections plus efficaces**, mais **n'améliore que très légèrement le contrôle de l'erreur** et les performances en prévision. Cependant, il faut garder à l'esprit que nous avons réalisé dans cette expérience un compromis entre une configuration de type OI (peu coûteuse en utilisation avec des EOFs moyens) et un schéma de type filtre de Kalman d'ensemble (avec un coût numérique plus élevé), et que celle-ci donne des résultats somme toute satisfaisants.

Enfin, le système ainsi validé a été utilisé pour la comparaison et l'étude de performances de différents réseaux d'observation. Tous les réseaux considérés sont existants ou réalistes et permettent de combiner des données d'élévation de surface (provenant d'observations *in situ*, simulées en différentes stations marégraphiques, et altimétriques) et de vitesses barotropes (simulant aussi des mesures *in situ* effectuées par des radars HF et des bouées ancrées réelles). Nous cherchions principalement à répondre à la question suivante :

- **Quels types de données et de réseau d'observations sont efficaces pour contrôler les erreurs du modèle ?**

Il est clairement apparu que les réseaux les plus efficaces sont ceux proposant une **fréquence régulière d'observation et une couverture spatiale étendue**. Ainsi, les performances de correction des réseaux marégraphiques sont nettement plus élevées que celles des réseaux altimétriques : un réseau de 21 marégraphes assimilés toutes les 2 heures assure environ 30% de correction de l'erreur d'élévation de surface, contre 8% pour le réseau combinant les satellites *Jason*, *Topex*, *GFO* et *Envisat*. La couverture spatiale assurée par les traces altimétriques, essentiellement hauturière, et le sous-cycle d'observation associé (3 jours) ne permettent pas de contrôler efficacement la plupart des structures d'erreurs du modèle dont on a montré qu'elles se développaient essentiellement en zone côtière et selon des régimes temporels rapides (de l'ordre de 24 heures). Ces résultats rejoignent globalement ceux de Baptiste Mourre qui montrait en 2004 (voir aussi Mourre *et al.*, 2005) qu'un réseau de 20 marégraphes assimilés à 6 heures permettait de contraindre très efficacement les erreurs du même modèle induites par les incertitudes sur la bathymétrie. En outre, s'il mettait en évidence une complémentarité notable entre réseaux marégraphiques et altimétriques dans sa configuration d'étude (utilisant notamment une méthode d'assimilation plus complexe que celle considérée dans cette étude), nous

ne sommes pas arrivés à de telles conclusions dans notre cas. Cependant, la combinaison d'un réseau marégraphique, fournissant des données d'élévation de surface, et d'un réseau de radars HF et de bouées, fournissant des mesures de courant, révèle une complémentarité intéressante, notamment en terme de correction des vitesses : celle-ci passe ainsi d'environ 14% dans le cas du réseau de 10 marégraphes assimilés à 2 heures, à près de 25% dans le cas de la combinaison de données de SLA et de vitesses.

De nombreuses perspectives s'ouvrent finalement à ce travail. En termes de modélisation de la dynamique barotrope induite par les forçages atmosphériques dans la zone du Golfe de Gascogne et la Manche, il sera intéressant de visualiser l'impact de champs de forçages météorologiques à une résolution spatio-temporelle plus fine, comme celles des champs issus des modèles météorologiques ALADIN et AROME, développés à Météo-France. Par ailleurs, nous avons calibré le système par l'intermédiaire du diagnostic de cohérence interne (test du  $J_{min}$ ), d'où il ressort que les erreurs atmosphériques modélisées ont été spécifiées de manière réaliste. Néanmoins, des ensembles atmosphériques issus d'AROME (obtenus dans le cadre d'une collaboration avec Météo-France) permettraient peut-être une meilleure caractérisation des erreurs atmosphériques, que l'on pourrait alors utiliser pour perturber encore plus efficacement les champs météorologiques de référence. Il serait notamment intéressant de poursuivre l'étude de sensibilité à ce type de forçage haute résolution dans la période suivant celle que nous avons étudiée, *i.e.* celle de décembre 1999, pendant laquelle des tempêtes d'une rare intensité ont frappé les côtes ouest européennes et ont engendré des phénomènes de surcotes importantes.

Plusieurs questions peuvent être identifiées en perspective à l'étude d'assimilation.

### ◦ **Comment améliorer l'horizon de prévision du système d'assimilation ?**

Le taux de correction de l'erreur obtenu dans cette étude est de l'ordre de 25-30% (pour l'élévation de surface), associé à une limite de prévision de l'ordre de 30 heures. Ces valeurs sont certes satisfaisantes, mais demeurent un peu faibles. Il serait intéressant de s'appuyer sur les résultats d'une étude consacrée à l'impact de forçages atmosphériques haute résolution, comme proposée ci-dessus, et d'examiner la sensibilité au nombre de données assimilées. Un effort conséquent est cependant requis pour améliorer la modélisation des statistiques d'erreurs et notamment leur évolution temporelle. Pour ce faire, plusieurs solutions peuvent être avancées : toujours dans l'utilisation du schéma d'assimilation SEQUOIA, les EOFs d'ensemble pourraient être calculés à partir d'un plus grand nombre d'instantanés, afin d'échantillonner un maximum de structures d'erreurs du modèle. La prise en compte des *tendances* des signaux océaniques dans le vecteur d'état pourrait de même améliorer le contrôle de l'évolution temporelle de l'erreur lors des expériences d'assimilation. La mise en place de schémas d'assimilation plus avancés devrait aussi nous permettre d'affiner la modélisation de ces erreurs (au prix cependant d'un coût de calcul nettement plus élevé) : notamment, moyennant quelques modifications du code source de SEQUOIA, l'implémentation d'un schéma de type Filtre de Kalman d'Ensemble à Ordre Réduit (ROEnKF), pour lequel les statistiques d'ensemble sont calculées au moment de

l'analyse, permettrait de prendre en compte l'évolution temporelle de l'erreur du modèle, et d'effectuer une comparaison rigoureuse des résultats avec ceux obtenus dans la présente étude. Enfin, il serait probablement enrichissant d'examiner en parallèle les résultats fournis par un Filtre de Kalman d'Ensemble complet, et de les comparer à ceux présentés dans ce manuscrit, ceux de Baptiste Mourre, et ceux obtenus avec un ROEnKF.

◦ **Quel changements serait-il nécessaire d'apporter à la configuration actuelle afin de configurer le présent système d'assimilation en mode opérationnel ?**

Nous reprenons ici les exigences scientifiques et technologiques requises pour l'implémentation d'un système opérationnel, présentées en *section 4.2.5* et nous les commentons à la lumière des résultats obtenus dans l'étude :

- ▷ **utiliser une technologie multivariée** : nous avons montré que par l'intermédiaire d'un vecteur d'état et de statistiques d'erreurs multivariés, la correction apportée au champ d'élévation de surface est efficacement propagée au champs de vitesses barotropes. Par ailleurs, nous avons testé avec succès l'assimilation simultanée de plusieurs types d'observations de hauteur de mer et de vitesses.
- ▷ **être adapté à des vecteurs d'état de grande dimension à un coût raisonnable** : dans le noyau MANTA, le coût de l'analyse est réduit car l'inversion se fait dans l'espace des EOFs (espace réduit). Le temps supplémentaire moyen de trente analyses (pour une configuration de 120 données marégraphiques assimilées toutes les 12h selon des EOFs moyens) est de 1.7 % de plus que la durée d'une simulation *libre*, et ce pour un vecteur d'état dont la dimension est de l'ordre de  $10^6$ . Il serait donc possible d'augmenter significativement la taille du vecteur d'état (dans le cas de l'utilisation d'un modèle tridimensionnel ou d'un raffinement du maillage), du nombre de données, ... dans la mesure où la marge de progression du temps requis par l'assimilation de données est confortable (on estime qu'un schéma d'assimilation est "économique" jusqu'à un surcoût de 100% du temps d'intégration).
- ▷ **pouvoir fournir des prévisions à court terme** : dans le cadre d'expériences jumelles, nous avons montré que la prévision océanique était améliorée dans la limite d'une échéance de 30 heures ; il apparaît nécessaire de mesurer plus précisément l'impact de la persistance de la correction, à la fois océanique et atmosphérique, afin d'optimiser la configuration. Le problème de la qualité des champs atmosphériques utilisables devra être aussi examiné : il faudra probablement quantifier dans quelle mesure des produits atmosphériques prévus, associés à de brusques changements de qualité (à l'interface analyse/prévision) vont dégrader les solutions du modèle, et recalibrer le système en conséquence via la méthode du  $J_{min}$  (par exemple par le biais de modifications des variances d'erreurs de prévision). Nous n'avons pu tester le système développé dans une configuration temps réel, mais nous prévoyons de mettre en place un *prototype de bulletin océanique* en temps réel simulé, dans lequel les problématiques mentionnées ci-dessus pourront être analysées.
- ▷ **pouvoir fournir des estimations des erreurs d'analyse et de prévision** : le système

dans sa configuration actuelle ne permet pas de fournir de telles estimations. Ces points doivent donc être développés.

- ▷ **être "cognitif"** : l'Interpolation Optimale classique ne peut satisfaire à cette exigence dans la mesure où, par définition du schéma, le sous-espace d'erreur n'évolue pas. Cependant, nous avons construit dans cette étude une base d'EOFs d'erreurs prenant en compte la variabilité temporelle des erreurs du modèle; le système s'est avéré en mesure de corriger en partie l'erreur du modèle, et ses performances devraient être probablement améliorées dans le cas où il évoluerait vers une configuration plus avancée, de type ROEnKF, SEEK ou EnKF.
- ▷ **être robuste** : le système est stable et efficacement contrôlé, dans la mesure où nous n'avons pas observé au cours de l'étude, de développement d'instabilités, ni de transitoire violent, ni encore de génération d'oscillations à l'initialisation; la modélisation des covariances d'erreurs de prévision sous forme d'EOFs moyens tirés des simulations assure donc des corrections réalistes. Notons simplement que l'utilisation d'"EOFs du jour" produit une trajectoire à la fois légèrement meilleure et parfois plus instable, ce qui compliquerait la fourniture de barres d'erreur fiables. Nous n'avons pas réalisé de test de sensibilité du système à une dégradation des informations d'entrée (nombre d'observations, qualité des forçages atmosphériques . . . ), mais ceci pourrait être aisément réalisé dans le cadre du bulletin océanique mentionné précédemment.
- ▷ **être évolutif** : le code SEQUOIA a été construit avec un souci de modularité et d'évolutivité. Il a déjà été couplé avec succès au modèle tridimensionnel SYMPHONIE (développé par Patrick Marsaleix du Laboratoire d'Aérodynamique de Toulouse) dans le cadre de l'étude de Gabriel Jordà-Sanchez (2005) et celle de Claire Dufau-Julliand (en cours) réalisées au Pôle d'Océanographie Côtière. Comme mentionné précédemment, le code actuel peut aussi utiliser d'autres noyaux d'analyse, être configuré en Filtre de Kalman d'Ensemble, et moyennant quelques modifications, fonctionner en ROEnKF.
- ▷ **être accepté et publié par la communauté scientifique, et prêt à être testé dans un futur proche** : l'utilisation de techniques d'assimilation à ordre réduit appliquées à l'océanographie a été validée et implémentée dans plusieurs systèmes opérationnels tels que MERCATOR, MFS et ADRICOSM (via le code SOFA, précurseur de SEQUOIA). Par ailleurs, le système SEQUOIA/SYMPHONIE a été validé par Gabriel Jordà-Sanchez (thèse de doctorat, 2005). Enfin, en début d'année 2006, un article que j'ai rédigé avec d'autres auteurs, a été soumis en vue d'une publication au *Journal of Geophysical Research - Oceans* (Lamouroux *et al.*, 2006).

Les résultats obtenus dans le cadre de l'étude du système SEQUOIA/MOG2D nous ont permis de répondre de manière satisfaisante à un certain nombre des exigences requises par les systèmes opérationnels. D'autres points restent aussi à améliorer pour finaliser la validation de la mise en place opérationnelle du système d'assimilation étudié : il apparaît en effet nécessaire d'affiner l'étude de l'impact des observations en fonction de leur distribution géographique, et notam-

ment de comparer l'impact de données sur la plateau et sur la zone profonde, afin de conclure quant à l'intérêt et l'impact des différents types de mesures disponibles (marégraphiques, altimétriques et de bouées et radars HF) dans l'étude et le contrôle des erreurs considérées. Enfin, la *localisation* des covariances d'erreur semble un point de passage obligé dans la perspective de l'application à des domaines plus étendus.