



HAL
open science

Apports des données gravimétriques GRACE pour l'assimilation de données altimétriques et in-situ dans un modèle de l'Océan Pacifique Tropical.

Frédéric Castruccio

► **To cite this version:**

Frédéric Castruccio. Apports des données gravimétriques GRACE pour l'assimilation de données altimétriques et in-situ dans un modèle de l'Océan Pacifique Tropical.. Océan, Atmosphère. Université Joseph-Fourier - Grenoble I, 2006. Français. NNT: . tel-00138506v1

HAL Id: tel-00138506

<https://theses.hal.science/tel-00138506v1>

Submitted on 26 Mar 2007 (v1), last revised 14 Jun 2007 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THESE

présentée par

Frédéric CASTRUCCIO

pour obtenir le titre de DOCTEUR
de l'UNIVERSITE JOSEPH FOURIER - GRENOBLE 1

(Arrêtés ministériels du 5 juillet 1984 et du 30 mars 1992)

Discipline : **Sciences de la Terre et de l'Univers**

Apports des données gravimétriques GRACE pour l'assimilation de données altimétriques et in-situ dans un modèle de l'Océan Pacifique Tropical.

Soutenue le 22 décembre 2006

Composition du jury :

M. Eric BLAYO	Président
M. Yves du PENHOAT	Rapporteur
M. Jacques BLUM	Rapporteur
M. Jacques VERRON	Directeur de thèse
M. Lionel GOURDEAU	Codirecteur de thèse
M. Pierre BRASSEUR	Examineur

Thèse préparée au sein du Laboratoire des Ecoulements Géophysiques et Industriels

Remerciements

Et voilà, trois ans de thèse s’achèvent ici. Je viens de mettre le point final à mon manuscrit et il ne me reste plus que la traditionnelle page des remerciements à écrire. Même si une thèse représente un travail personnel, elle est aussi et surtout la somme d’une multitude de contributions. J’ai un certain nombre de personnes à remercier dans cette ultime page, chose qui n’est pas si évidente que cela à rédiger...

En premier lieu, j’adresse un grand merci à Jacques Verron et Lionel Gourdeau, mes deux directeurs de thèse. Jacques, pour m’avoir fait mettre un premier pied dans le milieu de l’océanographie, en me proposant un stage de DEA, puis le deuxième en me permettant de poursuivre l’expérience dans le cadre de cette thèse au sein de l’équipe MEOM. Lionel, pour tes coups de pouce à distance, toujours bienvenus, pour ton aide afin de mieux appréhender la dynamique équatoriale et aussi pour m’avoir permis de mettre un pied... dans le Pacifique Tropical, cet océan que j’ai étudié pendant trois ans derrière mon ordinateur.

Je remercie également les membres du jury : Eric Blayo pour avoir accepté la présidence du jury, Jacques Blum et Yves du Penhoat pour m’avoir fait l’honneur d’être mes rapporteurs ainsi que Pierre Brasseur, pour l’intérêt que tu as porté à mon travail tout au long de ces trois ans et pour le recul que tu m’as apporté sur l’assimilation de données.

Comme je l’ai dit, une thèse repose essentiellement sur les interactions au sein, mais aussi en dehors de son équipe de recherche. Un grand merci donc à l’ensemble de l’équipe MEOM pour l’extraordinaire ambiance de travail, et pas seulement de travail d’ailleurs, qui y règne. Il ne fait aucun doute que chaque membre de l’équipe a joué un rôle dans le bon déroulement de ma thèse et je leur en suis très reconnaissant. Merci à toi Jean-Michel (qui forme avec Pierre la “Belgian connection” de l’équipe assimilation), pour ta patience et ta disponibilité sans faille pour résoudre l’ensemble des problèmes relatifs à l’assimilation de données, tant pratiques que théoriques, qui n’ont pas manqué d’émailler cette thèse. Merci à toi Jean-Marc, pour tes talents de débogueur hors pair et pour ton aide lors de la prise en main d’OPA. Merci à toi Bernard, pour les discussions passionnées et passionnantes sur l’océanographie et la modélisation. Je n’oublie pas non plus Josiane qui fait “entre autres” office de secrétaire et sans qui l’équipe ne serait pas ce qu’elle est.

Ces trois années n’auraient pas été aussi agréables sans l’indescriptible ambiance qui a régné dans le bureau des “twin theses”. Un grand merci pour ça à Grégoire, alias M. Broquet, auteur de la thèse jumelle (*Broquet, 2007*) et à Laurent, alias M. Potatoes Brodeau. Certaines personnes ont eu la “chance” de pouvoir profiter de cette vie débordante en occupant le bureau voisin. Je pense à Nadia, Gaëlle, Tong et David. Certaines ont même été aux premières loges, n’est-ce pas Clothilde... Merci à eux de ne pas avoir craqué. Les autres membres de l’équipe ont également pu suivre l’évolution alambiquée de nos “private jokes” par le biais du tableau du très convivial “espace convivialité”. Merci à multi-clopes, multi-poches, M. Brerron ou la fille de l’IIIIIMMMAG, pour ne citer qu’eux, les incontournables et bien involontaires piliers des dites “private jokes”. Merci à “l’université de tous les savoirs”¹ de nous avoir fait découvrir la foudre en boule... Un immense merci aux

¹<http://www.canal-u.education.fr/canalu/chainev2/utls/>

touillettes de EVE pour avoir servi d'exutoire à M. Broquet pendant près de deux ans (il doit en rester des vestiges dans les tiroirs de son bureau)... limitant ainsi grandement les débordements sonores. Merci aussi aux Fraggles Rock qui ont su prendre le relais des touillettes...

Merci également à tous les autres membres de l'équipe, et là je vais forcément en oublier donc je m'en excuse par avance... Merci donc à Yann, alias M. Ourmière, Léo et Gaëlle les dignes représentant du "MEOM d'en bas". Merci à Fabien, Jérôme, Florence, David et Sylvain, l'ancienne génération. Merci aussi à la nouvelle génération : Pierre, Clément, Angélique et Albanne. Merci à la génération des "sans génération", comprenez les permanents, plus communément appelés le "MEOM d'en haut" : Thierry, Emmanuel, Julien et Achim, à qui il faut ajouter Claire et Chafih. Merci à Anna pour la touche espagnole et à Sergey "c'est forrrmidable" Skachko et Denis pour la touche russe. Et voilà, j'ai failli oublier Marc, le roi du calembour, et Mélanie, miss colocalisation MEOM 2006.

Je remercie enfin mes collocs qui se sont succédé au 32 rue Mortillet, dans l'ordre : Jérémy mon frère, Pablo, Sonia et plus récemment Pol, encore un Belge (Vous avez remarqué comme les gens ont tendance à être Belge ?), sur le même refrain : "C'est quoi ta thèse déjà... ?". Je remercie aussi Mat, Cédric, Ben, Seb, etc... pour les sorties skis, VTT, les squashes, les apéros dinatoires et les "passages sous l'étiquette" toujours très utiles, pour ne pas dire nécessaires, afin de se changer les idées entre deux expériences d'assimilation... Pour finir, je tiens évidemment à remercier mes parents pour m'avoir toujours soutenu et m'avoir permis de poursuivre mes études si tardivement...

Table des matières

Introduction	1
I La circulation observée et les données d'observation	9
1 La dynamique du Pacifique Tropical	11
1.1 Introduction	12
1.2 L'état moyen	13
1.3 Le cycle saisonnier	16
1.4 La variabilité interannuelle	19
2 L'observation de l'océan : les données	25
2.1 Un petit historique de la mesure de l'océan	26
2.2 Les données satellites : le nouveau duo altimétrie/gravimétrie	28
2.2.1 La mesure altimétrique	30
2.2.2 La mesure gravimétrique	31
2.2.3 La topographie dynamique utilisée	34
2.3 Les données in-situ	36
2.3.1 Les données TAO	36
2.3.2 Les données XBT	36
II Les Outils	39
3 Le modèle	41
3.1 Introduction	42
3.2 Présentation du modèle OPA	42
3.3 La physique résolue par le modèle	43
3.3.1 Les équations primitives	43
3.3.2 La discrétisation spatiale et temporelle	45
3.3.3 Paramétrisation des échelles sous-mailles	45
3.3.4 Les conditions aux limites	48
3.3.5 Le calcul de la hauteur de la mer	50
3.4 La configuration utilisée	51
3.4.1 Le domaine d'étude	52
3.4.2 La résolution verticale	53
3.4.3 Les paramétrisations physiques	53
3.4.4 Les forçages atmosphériques	54

3.4.5	Récapitulatif	57
4	L'assimilation de données	59
4.1	Introduction	60
4.2	L'assimilation de données en général	61
4.2.1	Généralités	61
4.2.2	Les applications	62
4.2.3	Les différentes méthodes d'assimilation de données	64
4.3	La théorie du filtrage de Kalman	66
4.3.1	Positionnement du problème	66
4.3.2	Erreurs et densités de probabilité	67
4.3.3	Analyse optimale	69
4.3.4	La séquence d'assimilation	70
4.4	Le SEEK : de la théorie vers les applications réelles	71
4.4.1	Le principe de la réduction d'ordre	72
4.4.2	L'initialisation de l'espace réduit	73
4.4.3	L'évolution de la covariance des erreurs d'analyse et de prévision	74
4.4.4	L'algorithme SEEK	75
4.4.5	Les différentes variantes de l'algorithme SEEK	76
4.4.6	Le SEEK local	76
4.5	Implémentation du SEEK avec OPA surface libre	77
4.5.1	Le logiciel SESAM	77
4.5.2	La réinitialisation du modèle : le schéma IAU	78
4.5.3	Domaine d'assimilation	81
4.5.4	Le vecteur d'état	82
4.5.5	La spécification de l'espace d'erreur en rang réduit	83
4.6	Conclusion	84
III	Résultats	85
5	Présentation des expériences	87
5.1	Définition des conditions expérimentales	88
5.1.1	Une assimilation de topographie dynamique absolue. Pour quoi faire ?	88
5.1.2	Les différentes configurations du système d'observation	90
5.1.3	Les erreurs d'observation	92
5.1.4	L'état initial \mathbf{x}_0	93
5.1.5	Durée du cycle d'assimilation	93
5.2	Paramétrisation de l'espace d'erreur en rang réduit	94
5.2.1	Positionnement du problème	94
5.2.2	La paramétrisation retenue	95
5.3	Récapitulatif des expériences réalisées	96
6	Apports d'une topographie dynamique moyenne observée	99
6.1	Pertinence de la MDT GRACE dans le Pacifique Tropical	101
6.2	D'une surface moyenne synthétique à une surface moyenne observée	104
6.2.1	Apports de l'assimilation : écarts RMS	105
6.2.2	Apports de l'assimilation : comparaisons aux observations	106

6.3	Conclusion	112
7	Complémentarité des données altimétriques et in-situ	115
7.1	Vers l'assimilation conjointe d'altimétrie et de données in-situ	117
7.1.1	Apports de l'assimilation : écarts RMS	117
7.1.2	Apports de l'assimilation : comparaisons aux observations	119
7.2	Apports de la MDT observée GRACE pour l'assimilation conjointe	124
7.2.1	Synthèse des principaux résultats de la GRL	124
7.2.2	Letter : On the role of the GRACE mission in the joint assimilation of altimetric and TAO data in a tropical Pacific Ocean model	125
7.3	Conclusion	133
8	Une expérience interannuelle sur la période 1993-1998	135
8.1	Présentation de l'expérience	136
8.2	Apports de l'assimilation : écarts RMS	137
8.2.1	Apports pour la partie observée	138
8.2.2	Apports par rapport à des données indépendantes	138
8.3	Apports de l'assimilation : comparaisons aux observations	140
8.3.1	Niveau de la mer	140
8.3.2	Courants	144
8.3.3	Température	147
8.4	Discussion	153
8.5	Conclusion	155
9	Evolution des transports dans les couches de surface	159
9.1	Variabilité du flux zonal de masse	161
9.2	Discussion	172
	Conclusions et perspectives	175
	Annexes	183
A	Impact of ocean mean dynamic topography on satellite data assimilation	185
A.1	Introduction	186
A.2	Numerical model, assimilation method and data sets	187
A.2.1	Numerical model	187
A.2.2	Assimilation method	189
A.2.3	Data sets	189
A.3	Problem definition and experimental framework	193
A.3.1	Problem definition	193
A.3.2	Description of the experiments	194
A.4	Sensitivity tests : results	195
A.4.1	Surface currents	195
A.4.2	Barotropic streamfunction	196
A.4.3	Vertical temperature section at 55°W	198
A.5	Comparison with independent TS profiles	200

A.6	Specific case of the assimilation of TS profiles	202
A.7	Summary and conclusions	203
B	Une paramétrisation de l'espace d'erreur en rang réduit basée sur un ensemble de simulations perturbées	205
B.1	Positionnement du problème	205
B.2	La paramétrisation retenue	206
B.3	Discussion	208
C	A sequential data assimilation approach for estimating turbulent air-sea flux bulk parameters	213
C.1	Introduction	214
C.2	The ocean model	215
C.3	The assimilation scheme	216
C.4	Design of the assimilation experiments	219
C.4.1	Experimental setup	219
C.4.2	Parameterization of background error covariance	221
C.5	Comparison of the 3 variants of the assimilation scheme	222
C.5.1	Time evolution of temperature and salinity errors	222
C.5.2	Distribution of temperature and salinity errors	224
C.5.3	Quality of the parameter estimates	227
C.5.4	Correction of systematic errors	229
C.5.5	Quality of the flux estimates	231
C.5.6	Quality of the model forecasts	231
C.6	Conclusions	233
	Bibliographie	235
	Listes des figures	246
	Listes des tables	254

Introduction

Après la canicule de l'été 2003 en Europe ou le cyclone Katrina qui a dévasté la ville de la Nouvelle-Orléans en 2005, l'année 2006 n'a pas été épargnée par les catastrophes naturelles avec des inondations désastreuses en Ethiopie et en Corée du Nord ou encore le typhon le plus intense qu'ait connu la Chine depuis plus de 50 ans. L'apparente multiplication de ces événements exceptionnels a joué le rôle d'un véritable électrochoc sur l'opinion et la classe politique internationale. Ces phénomènes extrêmes sont souvent perçus comme des preuves qu'un dérèglement du climat nous frappe déjà. D'ailleurs un récent sondage CSA/Canal+ indique que 51 % des français craignent en premier lieu les conséquences du réchauffement de la planète, devant la menace terroriste (43 %). La pression se fait donc forte sur les scientifiques pour expliquer et surtout prévoir les aléas de notre climat à court, moyen et long terme.

L'océan représente le véritable chef d'orchestre du climat de notre planète, dont il couvre près de 70 % de la surface. En raison de sa très grande capacité thermique, environ 1200 fois supérieure à celle de l'atmosphère, et de son inertie, il est capable de générer, amplifier ou encore modérer des anomalies climatiques, comme par exemple le phénomène El Niño, le plus intense, mais aussi le plus emblématique signal de variabilité interannuelle du système climatique terrestre actuel. Les océans ont également un rôle prépondérant sur l'évolution du climat à plus long terme. On sait aujourd'hui que sans le piégeage de carbone par les océans, le taux de dioxyde de carbone (CO_2), l'un des principaux gaz à effet de serre, serait beaucoup plus élevé. A titre d'exemple, on estime que l'océan stocke actuellement 39000 milliards de tonnes de carbone contre seulement 600 milliards pour l'atmosphère et 610 milliards pour la biosphère. Chaque année l'homme injecte entre 6 et 7 milliards de tonnes de carbone dans l'atmosphère. Sur cette quantité, l'océan en capte au moins un tiers. L'évolution à moyen et long terme du climat est donc fortement conditionnée par la question : L'océan restera-t-il un piège à carbone dans les années à venir ?

Toutes les considérations évoquées ci-dessus font que la compréhension du système océan, ainsi que la prédiction de son évolution à diverses échelles temporelles est aujourd'hui une thématique scientifique porteuse et très dynamique dans la mesure où elle est en phase avec les attentes de la société.

Les scientifiques ont à leur disposition deux grands types d'outils afin de produire une description réaliste de l'océan : l'observation et la modélisation. La compréhension d'un phénomène physique complexe comme la dynamique océanique passe nécessairement par la mesure et l'acquisition de données d'observations de celui-ci. L'océan est cependant un milieu très difficile à observer. Si les satellites nous fournissent aujourd'hui une vision globale avec une couverture spatio-temporelle et une précision élevée de la surface des océans, les observations de l'océan profond restent en revanche rares, voire inexistantes en dessous de 2000 m.

Les océanographes ont également développé des modèles numériques pour étudier et décrire la "machine océan". Ces modèles sont basés sur les équations fondamentales régissant la dynamique des fluides géophysiques. L'explosion de la puissance de calcul disponible combinée à une meilleure compréhension des phénomènes régissant la dynamique océanique font que les modèles actuels reproduisent avec une grande fidélité la circulation observée de l'océan. Le caractère fortement non-linéaire, voir chaotique, de la dynamique de l'océan fait cependant qu'un modèle numérique est une source d'information insuffisante pour avoir une connaissance précise de l'évolution des différents phénomènes mis en jeu dans la physique océanique.

Afin de tirer le meilleur parti de ces deux sources d'informations indépendantes et ainsi améliorer notre compréhension du système océan mais aussi et surtout affiner la capacité prédictive des modèles numériques de circulation océanique, différentes méthodes d'assimilation de données ont été développées en océanographie. L'assimilation de données a pour objet de combiner modèle et observations de manière optimale. En pratique, on impose au modèle des contraintes supplémentaires pour forcer sa trajectoire à se maintenir le plus proche possible de la réalité. Récemment, des progrès significatifs ont été réalisés, tant sur plan du développement des méthodes d'assimilation de données appliquées à l'océanographie, que sur le plan du développement des modèles numériques de circulation océanique. Ces progrès, combinés à la mise en œuvre de systèmes d'observation, notamment par satellite, permettant la mesure permanente des paramètres physiques de l'océan, se sont concrétisés par l'émergence de l'océanographie dite opérationnelle. Cette approche intégrée comprenant mesures et outils fournit des analyses et des prévisions tridimensionnelles systématiques de l'océan. Citons par exemple le projet MERCATOR² représentant la composante française du programme international GODAE³ (Global Ocean Data Assimilation Experiment). L'expérience GODAE a pour but de démontrer la faisabilité d'une assimilation de données océaniques à l'échelle globale et en temps réel capable de fournir des descriptions régulières et complètes des champs océaniques, en support à l'océanographie opérationnelle, aux modèles de prévision saisonnière et à la recherche en océanographie. MERCATOR publie hebdomadairement des bulletins de prévision océanique à haute résolution de l'Atlantique Nord depuis le 17 janvier 2001. La mise en place du projet européen MERSEA⁴ (Marine Environmental and Security for the European Area) est un autre exemple de l'importance grandissante accordée à la prévision océanique au niveau international. MERSEA s'inscrit dans la branche océan du programme GMES⁵ (Global Monitoring for Environment and Security) et a pour vocation de fédérer les contributions européennes à GODAE. Les enjeux sociétaux sont nombreux et importants. La prévision systématique du "temps" océanique est utile pour de nombreuses applications au premier rang desquelles se trouve la prévision des risques environnementaux (suivi des pollutions par hydrocarbures, algues toxiques, ...) et climatiques (inondations, tempêtes et phénomènes climatiques extrêmes). La connaissance de l'état de l'océan à l'échelle de quelques semaines bénéficie également à la navigation commerciale, aux plates-formes offshore, à la pêche, etc... ou de manière plus anecdotique pour le routage des navires lors des courses au large. A plus long terme, les modèles de climat reposent sur des modélisations couplées entre l'atmosphère et l'océan. Les modèles d'océan sont donc indispensables aux modèles climatiques. Les sorties du prototype global basse résolution de MERCATOR sont actuellement utilisés comme forçage pour le modèle de prévision saisonnière de Météo-France.

Le caractère opérationnel de l'océanographie repose en grande partie sur la mise en œuvre d'un système d'observation fournissant des mesures permanentes de l'état de l'océan. Ces mesures sont de deux types : in-situ et satellites. Compte tenu de la difficulté d'observer l'océan, la distribution en temps et en espace des données in-situ reste irrégulière et insuffisante. Elles sont cependant indispensables car ce sont les seules qui nous renseignent sur l'océan profond (entre la surface et 2000 m de profondeur pour l'es-

²<http://www.mercator-ocean.fr>

³<http://www.bom.gov.au/bmrc/ocean/GODAE/>

⁴<http://www.mersea.eu.org>

⁵<http://www.gmes.info/>

sentiel). Les données satellites, en nous offrant une vision globale, continue et homogène de la surface des océans ont réellement permis d'envisager la mise en place de systèmes opérationnels en océanographie. Parmi ces données satellites, l'altimétrie fournit une information intégrée sur toute la colonne d'eau. Elle a joué un rôle de catalyseur dans les formidables développements scientifiques et applicatifs de l'océanographie ces dernières années, en particulier pour le développement de l'océanographie opérationnelle. L'élévation de surface libre des océans (SSH pour Sea Surface Height) est mesurée avec une précision centimétrique (estimée à 3 cm RMS) depuis octobre 1992 par le satellite TOPEX/Poseidon aujourd'hui relayé par son successeur Jason 1. Cependant, seule la partie résiduelle (l'anomalie de surface libre ou SLA pour Sea Level Anomaly) du signal altimétrique mesuré par les satellites peut être utilisée de manière fiable dans le cadre d'étude océanographique. Le niveau moyen de référence, la topographie dynamique moyenne (MDT pour Mean Dynamic Topography) est contaminée par des erreurs importantes sur le géoïde, en particulier pour les harmoniques élevées (harmoniques supérieures à 20 environ). L'absence d'une bonne topographie dynamique moyenne est un problème récurrent (*Blayo et al.*, 1994) pour l'assimilation de données altimétriques. Il est habituellement contourné en s'appuyant soit sur la surface moyenne du modèle (la topographie dynamique moyenne du modèle est supposée parfaite et sert à référencer les anomalies de surface libre) soit sur d'autres sources de données (en général in-situ) après un traitement plus ou moins élaboré (e.g. *Mercier*, 1986; *Le Grand*, 1998). Aucune de ces solutions n'est vraiment satisfaisante et la question des erreurs sur le niveau moyen de référence est très contemporaine, à la fois du point de vue de l'altimétrie et de la complémentarité multi-données des systèmes d'observations. Les récents travaux de *Birol et al.* (2004) sur la sensibilité de l'assimilation de données à la topographie dynamique moyenne utilisée pour référencer les anomalies de surface libre ont montré que les erreurs sur l'estimation de la topographie dynamique moyenne faisaient partie des facteurs ayant le plus fort impact sur le réalisme de la circulation océanique simulée avec assimilation de données altimétriques pour l'Atlantique Nord.

Récemment, avec le lancement des missions satellites CHAMP⁶ (Challenging Mini-satellite Payload) en 2000 et GRACE⁷ (Gravity Recovery and Climate Experiment) en 2002, dédiées à l'étude du champ de gravité terrestre, notre connaissance du géoïde s'est grandement améliorée. La mission européenne GOCE⁸ (Gravity Field and Steady-State Ocean Circulation Explorer), complémentaire de GRACE, devrait permettre d'accéder à un modèle de géoïde à haute résolution présentant une précision centimétrique pour des résolutions spatiales de 100 km à l'horizon 2007. La précision de 40 cm pour une résolution spatiale de 500 km correspondant à la précision avec laquelle le géoïde était connu avant le lancement des missions CHAMP et GRACE permet de se rendre compte de l'ampleur du gain consécutif aux missions satellites gravimétriques sur notre connaissance du géoïde. Ces importantes améliorations sur la résolution et la précision avec laquelle nous connaissons le géoïde, et donc la topographie dynamique moyenne des océans fait qu'il est aujourd'hui possible d'utiliser le signal altimétrique absolu observé par satellite dans un contexte d'expérience d'assimilation de données réalistes (*Gourdeau et al.*, 2003).

Les potentialités de l'utilisation d'une topographie dynamique moyenne observée plus réaliste que celle du modèle sont importantes pour l'assimilation de données altimétriques. En premier lieu, l'assimilation de la topographie dynamique absolue va permettre de

⁶<http://www.gfz-potsdam.de/pb1/op/champ/>

⁷<http://www.gfz-potsdam.de/pb1/op/grace/>

⁸<http://www.esa.int/esaLP/LPgoce.html>

contrôler la circulation moyenne simulée en plus de sa variabilité. L'utilisation d'une topographie dynamique moyenne plus réaliste devrait également permettre de régler, du moins en partie, le problème de la compatibilité entre données altimétriques et données in-situ. *Parent et al.* (2003) ont en effet montré que l'assimilation conjointe de données altimétriques et de profils de température in-situ dans la région du Pacifique Tropical pouvait conduire à une impasse en raison de la différence d'état moyen entre les deux types de données du fait des erreurs sur la surface moyenne du modèle servant à référencer les résidus altimétriques. En améliorant la compatibilité entre l'altimétrie et les données in-situ, l'utilisation d'une topographie dynamique moyenne observée devrait donc permettre de pleinement utiliser la complémentarité existant entre ces deux types de données au travers d'expérience d'assimilation conjointe efficace.

C'est dans ce contexte que s'inscrit ce travail de thèse, qui se propose d'explorer les potentialités et d'estimer les apports offerts par l'utilisation d'une topographie dynamique moyenne observée déduite des données gravimétriques GRACE. Pour cela, nous nous placerons dans le cadre d'expériences d'assimilation de données altimétriques afin d'aller plus loin que la simple vision de surface fournie par le satellite pour reconstruire la circulation océanique dans toutes ses dimensions spatio-temporelles.

La première difficulté inhérente à ce travail consistera à développer et mettre en œuvre une plate-forme d'assimilation de données basée sur le filtre SEEK (Singular Evolutive Extended Kalman filter), développé à Grenoble par *Pham et al.* (1998), et le modèle OPA (*Madec et al.*, 1998) dans sa configuration ORCA2 à surface libre afin de répondre aux problèmes spécifiques soulevés par l'assimilation de la topographie dynamique absolue observée par satellite. Le but est ici de contrôler l'écoulement moyen du modèle en plus de sa variabilité. La paramétrisation du filtre SEEK, qui repose habituellement sur la variabilité d'une simulation libre, devra en particulier être repensée afin d'efficacement utiliser l'information contenue dans les observations sur l'état moyen du système. Même si notre modèle est global, nous nous limiterons à l'étude du Pacifique Tropical. Plusieurs raisons justifient ce choix. D'une part, la dynamique équatoriale est à plus "grande échelle" que la dynamique océanique des moyennes latitudes; la relative basse résolution de la topographie dynamique moyenne observée est donc moins limitante dans cette région. D'autre part, la Pacifique Tropical, siège du phénomène El Niño, est une des régions les plus observées et étudiées de l'océan mondial. Le Pacifique Tropical est notamment échantillonné de manière permanente par un réseau de près de 70 mouillages (le réseau TAO/TRITON⁹) depuis l'année 1994. La présence de données in-situ concomitantes et présentant une couverture globale de la bande équatoriale (i.e $\pm 8^\circ$) de latitude a fortement influencé le choix de la région d'étude.

Dans un deuxième temps, nous nous attacherons à estimer l'impact de l'utilisation d'une topographie dynamique moyenne observée par satellite sur l'assimilation de données altimétriques. L'assimilation conjointe d'altimétrie et de données in-situ, de même que la complémentarité de ces observations seront discutées. Les apports de la topographie dynamique moyenne observée sur le problème de la compatibilité entre données altimétriques et données in-situ seront également estimés.

Ce manuscrit est organisé en 3 parties :

⁹<http://www.pmel.noaa.gov/tao/>

La première partie est dédiée à un rapide tour d’horizon de la dynamique du Pacifique Tropical et des observations de l’océan dont nous disposons pour caractériser la circulation de cette région. Nous nous intéresserons notamment aux données altimétriques et gravimétriques qui font l’originalité de ce travail.

La deuxième partie s’attache à détailler les outils, c’est-à-dire le modèle et la méthode d’assimilation utilisés pour ce travail d’un point de vue théorique. La mise en œuvre du SEEK avec le modèle OPA en surface libre sera également abordée, ainsi que l’étape cruciale de la paramétrisation de l’espace d’erreur en rang réduit du filtre SEEK.

La troisième et dernière partie est consacrée à l’étude des nouvelles capacités offertes par l’utilisation d’un signal de topographie dynamique absolu observé pour le contrôle par assimilation de données d’un modèle d’océan dans la région du Pacifique Tropical. Les apports d’une topographie dynamique moyenne observée pour l’assimilation de données seront d’abord discutés avec différentes configurations du système d’observation (altimétrie seule, in-situ seule, altimétrie et in-situ). La complémentarité des données, ainsi que le rôle de la topographie dynamique moyenne observée sur la compatibilité entre les données altimétriques et in-situ seront ensuite discutés. Pour finir, nous nous intéresserons aux apports de l’assimilation conjointe de données altimétriques et in-situ sur le réalisme de la dynamique du Pacifique Tropical simulée dans le cadre d’une expérience interannuelle sur la période 1993-1998.

Première partie

La circulation observée et les données d'observation

Chapitre 1

La dynamique du Pacifique Tropical

Sommaire

1.1	Introduction	12
1.2	L'état moyen	13
1.3	Le cycle saisonnier	16
1.4	La variabilité interannuelle	19

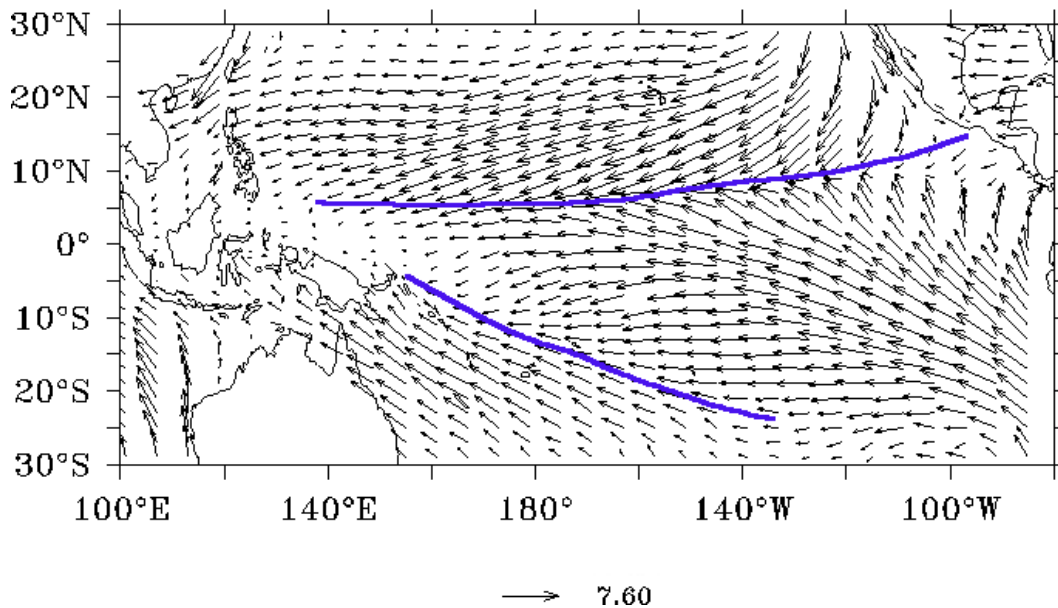


FIG. 1.2 – Climatologie COADS de la vitesse du vent et de la position moyenne de l'ITCZ et de la SPCZ.

1.2 L'état moyen

Si on se place du point de vue de l'atmosphère, et en simplifiant les choses, les régions tropicales reçoivent un excès d'énergie par rayonnement solaire alors que les régions polaires sont en déficit. On a donc un gradient de température négatif entre l'équateur et les pôles qui se traduit par un gradient de pression atmosphérique positif suivant le sens équateur-pôles. Cela entraîne une convergence des vents au niveau de l'équateur. On a donc une vaste cellule convective méridienne, appelée cellule de Hadley, qui exporte de la chaleur vers les pôles par sa branche supérieure et importe de l'air froid vers les régions tropicales par sa branche inférieure (Gill, 1982). Sous l'action de la force de Coriolis, l'air de la branche inférieure est dévié et donne naissance aux alizés, vents dominants qui soufflent d'Est en Ouest le long de l'équateur. Comme le montre la figure 1.2, il existe deux grandes régions de convergence des vents dans le Pacifique Tropical :

- la zone de convergence Inter Tropicale (ITCZ) qui traverse le Pacifique aux alentours de $5^{\circ}N$ et correspond à la zone de convergence des alizés de Sud Est et de Nord Est.
- la zone de convergence Sud Tropicale (SPCZ) qui s'étend de la Papouasie-Nouvelle Guinée à la Polynésie et correspond à la zone de convergence des alizés de Sud-Est et des vents du système de mousson associé à l'Australie.

Le système des alizés, couplé au fait que la force de Coriolis s'annule au niveau de l'équateur, crée une dynamique essentiellement zonale. La figure 1.4 montre clairement cette alternance de courants zonaux coulant tantôt vers l'Est et tantôt vers l'Ouest, qui caractérise la circulation du Pacifique Tropical. Le courant de surface vers l'Est, compris entre $3^{\circ}N$ et $10^{\circ}N$, est connu sous le nom de contre courant Nord équatorial (NECC pour North Equatorial Counter Current). Il est situé sous l'ITCZ (qui représente un minimum de tension zonale de vent) et s'écoule dans le sens opposé aux vents dominants avec des

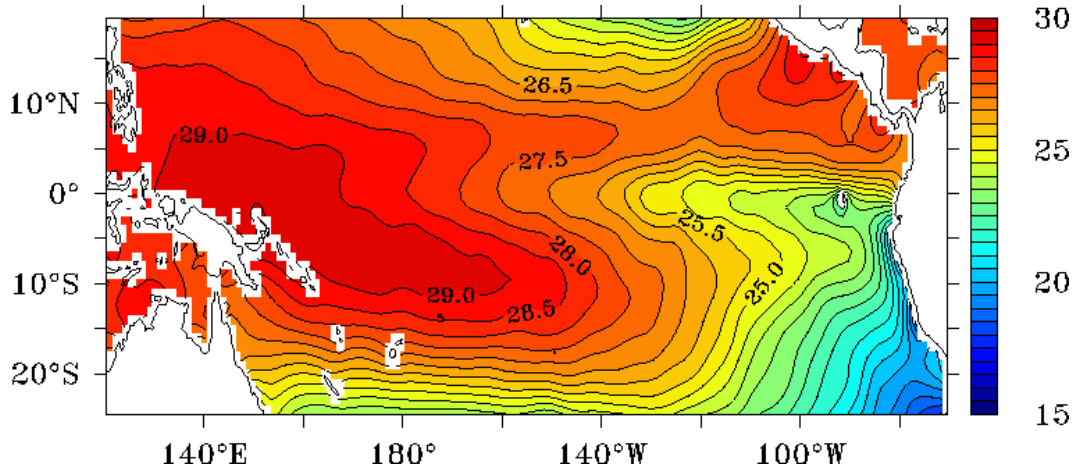


FIG. 1.3 – SST climatologique (Levitus, 1998).

vitesses en surface de l'ordre de 0.5 m/s. Cet étroit courant s'intercale entre deux larges courants s'écoulant vers l'Ouest, les courants Nord et Sud équatorial (NEC pour North Equatorial Current et SEC pour South Equatorial Current), au Nord de 10°N et au Sud de 3°N respectivement. Le NEC est relativement faible, avec des vitesses de surface de l'ordre de 0.2 m/s mais le SEC a un cœur très intense, situé entre 3°N et l'équateur, avec des vitesses qui peuvent dépasser 1 m/s. Un courant vers l'Est, le contre courant Sud équatorial (SECC pour South Equatorial Counter Current) est également présent aux alentours de 9°S et à l'Est de 180°.

Aux abords de l'équateur, le SEC accumule les eaux chauffées par l'intense rayonnement solaire des tropiques dans le Pacifique Ouest qui concentre les eaux parmi les plus chaudes de la planète avec des températures supérieures à 28° (cf figure 1.3). Cette vaste zone à l'Ouest de 180°, véritable réservoir d'eau chaude, est connue sous le nom de WarmPool. De par sa taille (de l'ordre de 10 millions de kilomètre-carrés), elle est la plus importante source de chaleur pour l'atmosphère et les hautes latitudes de l'océan Pacifique.

La thermocline qui sépare les eaux chaudes de surface des eaux froides abyssales s'approfondit donc en allant vers l'Ouest où elle atteint une profondeur de l'ordre de 150 m alors qu'elle fait surface dans l'Est du bassin (cf figure 1.5). Si l'on fait l'hypothèse que le gradient de pression est négligeable dans l'océan profond, cette pente de la thermocline implique, dans le cadre de l'équilibre hydrostatique, une force de pression dirigée vers l'Est. Les isopycnes, et donc la surface de l'océan, présentent une pente Est/Ouest. C'est une des caractéristiques de la dynamique équatoriale : son caractère barocline. Les variations du niveau de la mer sont le miroir de la profondeur de la thermocline. Le niveau de la mer, à l'équateur, est ainsi environ 60 cm plus haut à l'Ouest du bassin qu'à l'Est. Ce gradient de pression et cette pente zonale ne sont pas cantonnés à l'équateur et on les retrouve dans la quasi totalité du Pacifique Tropical. Loin de l'équateur, et sous la couche d'Ekman directement forcée par les vents, ce gradient de pression p_x induit un transport géostrophique v vers l'équateur dans les 2 hémisphères :

$$-fv + \frac{1}{\rho}p_x = 0 \quad (1.1)$$

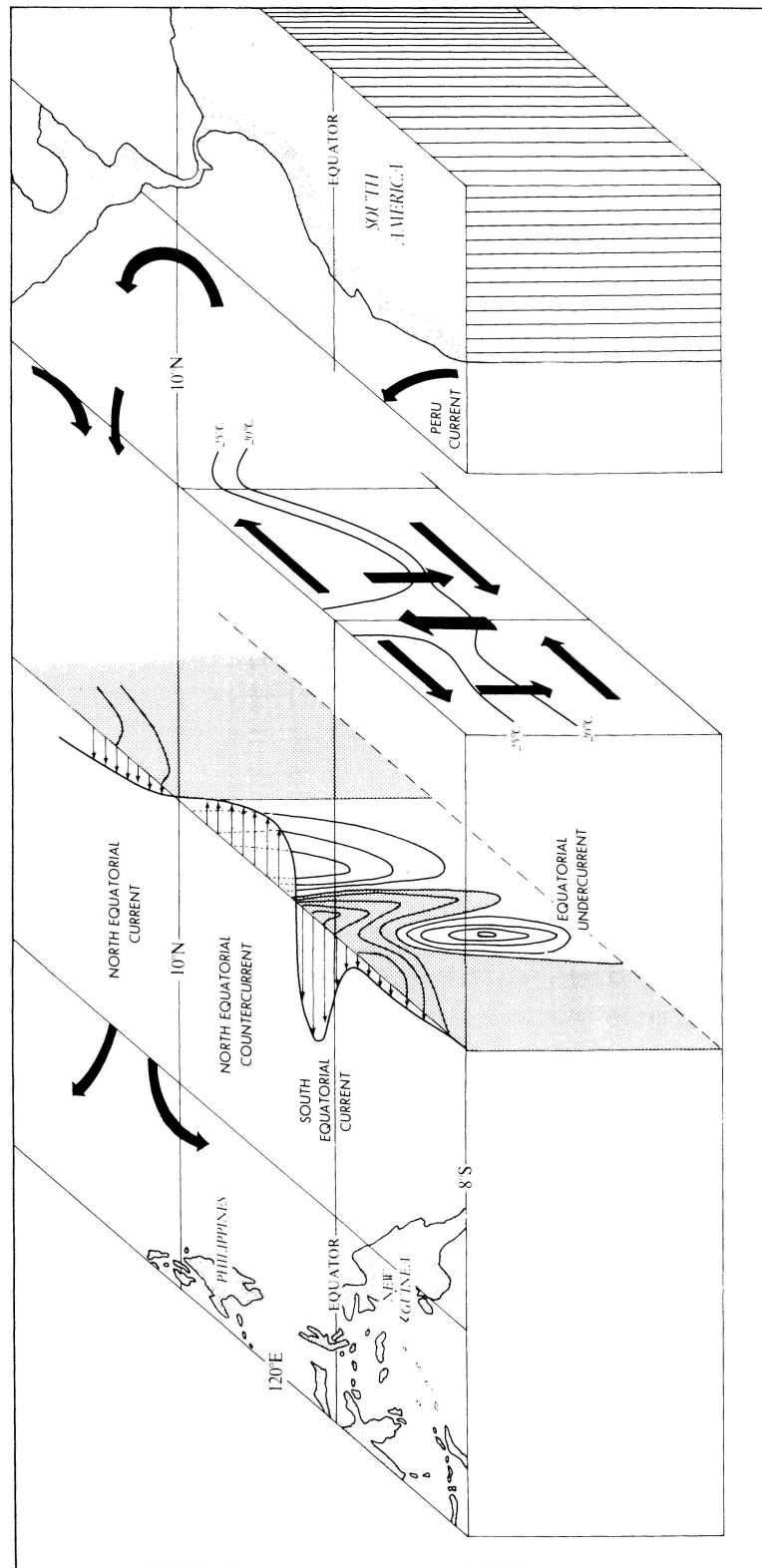


FIG. 1.4 – Circulation de surface et subsurface de l'océan Pacifique Tropical (source Philander (1990)).

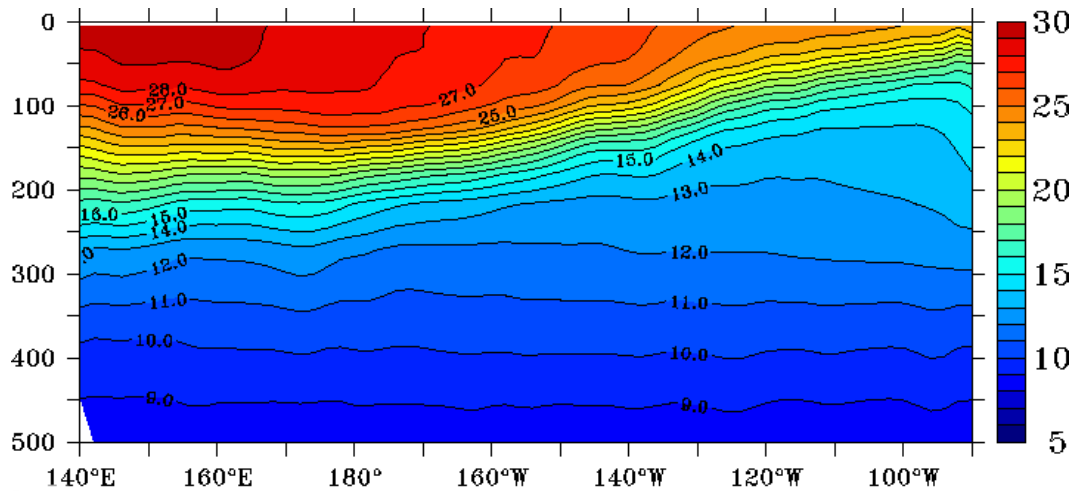


FIG. 1.5 – Section de température à l'équateur (climatologie Levitus (1998)).

où f est le paramètre de Coriolis qui change de signe quand on passe l'équateur. A l'équateur, la force de Coriolis disparaît ($f = 0$) et ne vient plus compenser la force de pression. Une partie des courants qui convergent vers l'équateur en profondeur va donc remonter vers la surface pour compenser les transports d'Ekman divergents de part et d'autre de l'équateur et ainsi former l'upwelling équatorial. L'autre partie va s'écouler vers l'Est, le long du gradient de pression, et former le sous-courant équatorial (EUC pour Equatorial Undercurrent) (cf figure 1.4). Ce jet, dont les vitesses peuvent dépasser 1.2 m/s, est centré sur la thermocline et a une extension verticale de l'ordre de 150 m. Bien que très étroit (environ 300 km), il est continu sur une distance longitudinale de plus de 10000 km.

La faible profondeur de la thermocline dans l'Est facilite les upwellings équatoriaux des eaux froides intérieures par les alizés si bien qu'une langue d'eau froide se développe de la côte Sud Américaine jusqu'à 180° (cf figure 1.3). On observe un fort gradient de température de surface entre l'Est, relativement froid, et l'Ouest du bassin. Ce gradient de température de surface vient renforcer les alizés. Les vents soufflent en effet à l'équateur le long des gradients de pression, des hautes vers les basses pressions. Or, de faibles pressions atmosphériques sont associées avec les fortes températures de surface à l'Ouest du Pacifique, tandis que des hautes pressions sont associées avec les basses températures à l'Est. De plus, en même temps qu'ils traversent le Pacifique d'Est en Ouest, les alizés prennent de la chaleur et de l'humidité à l'océan (par évaporation). La masse d'air, chaude et humide, devient donc très peu dense et s'élève au dessus de la WarmPool, où la forte convection (entretenu par la SST supérieure à 28°) engendre de nombreux nuages (cumulus à fort développement vertical) et de fortes pluies. Ces masses d'air ascendantes se dirigent ensuite vers l'Est, dans les couches supérieures de la troposphère, avant de redescendre au dessus des eaux froides de l'Est du Pacifique et ainsi fermer la cellule de Walker (Gill, 1982).

1.3 Le cycle saisonnier

Au cours de l'année, la position de l'ITCZ oscille entre 2°N et 10°N. Elle adopte sa position la plus boréale en été-automne (de l'hémisphère Nord), lorsque les alizés du Sud-

Est sont les plus intenses et franchissent l'équateur. Par contre, au printemps, lorsque les alizés faiblissent, l'ITCZ se trouve alors très proche de l'équateur ($1 - 2^\circ N$). Concernant la branche Ouest de la SPCZ, on observe le phénomène inverse. Pendant l'été (de l'hémisphère Nord), la SPCZ se rapproche de l'équateur alors qu'en hiver elle atteint $10^\circ S$ et les côtes australiennes.

Le cycle saisonnier de l'océan Pacifique Tropical est la réponse de l'océan à cette variabilité saisonnière des vents. La SST dans l'Est du Pacifique Tropical atteint ainsi un maximum pendant le printemps de l'hémisphère Nord, quand les alizés du Sud-Est faiblissent, et un minimum durant l'été-automne quand ils sont très intenses (cf figure 1.6). Ces changements de la température de surface ne sont qu'un aspect de la réponse de l'océan. Durant la période où les alizés de Sud-Est sont faibles, les upwellings équatoriaux sont minimums et les courants de surface du Pacifique Tropical sont faibles, excepté le NEC qui avec le renforcement des alizés de Nord Est est à son maximum saisonnier. On peut même observer une disparition du NECC dans les couches de surface du Pacifique central. En revanche, en subsurface, la vitesse et le transport de l'EUC atteignent leurs maximums saisonniers. A partir du mois de Mai, les alizés de Sud-Est se ré-intensifient alors que ceux de Nord-Est diminuent. La conséquence est une réactivation des upwellings équatoriaux et donc une baisse de la température de surface à l'équateur. La température de surface reste élevée, légèrement au Nord de l'équateur, ce qui crée un creux dans la thermocline aux alentours de $3 - 4^\circ N$. Plus au Nord, vers $8 - 10^\circ N$, la thermocline remonte car le transport d'Ekman est divergent à cet endroit. Cet accroissement de la pente méridienne de la thermocline se traduit par une intensification du NECC. Avec le renforcement des alizés de Sud-Est, on a également une intensification du SEC avec un maximum de vitesse juste au Nord de l'équateur. Le cisaillement longitudinal vers $4^\circ N$ devient tellement intense qu'il entraîne des instabilités. Des ondes, les TIWs (Tropical Instability Waves) apparaissent et se propagent vers l'Ouest le long du front de température avec des longueurs d'onde de l'ordre de 1000 km et des périodes d'environ 3 semaines.

Les alizés prévalent donc à l'Est de 180° . En revanche, les vents dans le Pacifique Ouest font partie du système de la mousson australienne et présentent donc une renverse saisonnière, avec un flux dominant qui s'oriente vers l'Est au printemps. Cela donne lieu à des courants équatoriaux présentant une structure verticale complexe. Le SECC, qui représente une bande de faible courant vers l'Est, imbriqué dans le SEC aux alentours de $9^\circ S$, est fortement contrôlé par cette renverse du flux dominant. Le SECC est dû à un minimum de tension de vents des alizés de Sud-Est et atteint son maximum saisonnier durant la mousson australienne (Décembre-Avril), avec des vitesses pouvant avoisiner les 0.3 ms^{-1} . Le SECC est difficilement identifiable durant le reste de l'année. A l'Est de 180° son intensité décroît rapidement et il est absent du Pacifique Est.

D'autre part, cette renverse saisonnière du flux vers l'Est coïncide, lorsque celui-ci est présent, avec le début d'El Niño, qui débute lui aussi par la propagation d'anomalies vers l'Est. Cette coïncidence temporelle rend plus difficile la détermination de la nature de l'anomalie de printemps (saisonnière ou interannuelle) et donc la prévision d'un possible El Niño.

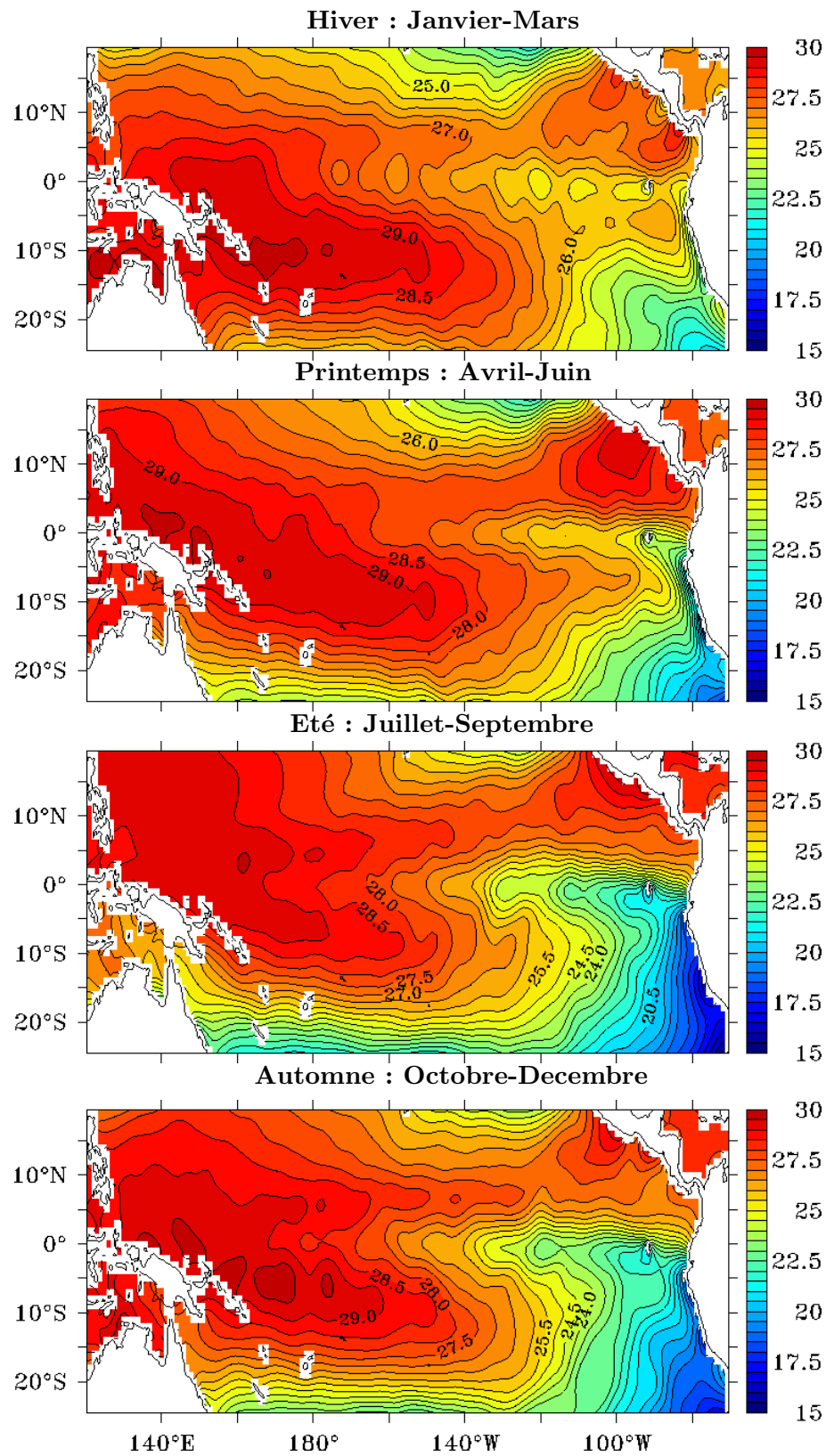


FIG. 1.6 – Climatologie saisonnière pour la SST (Levitus, 1998).

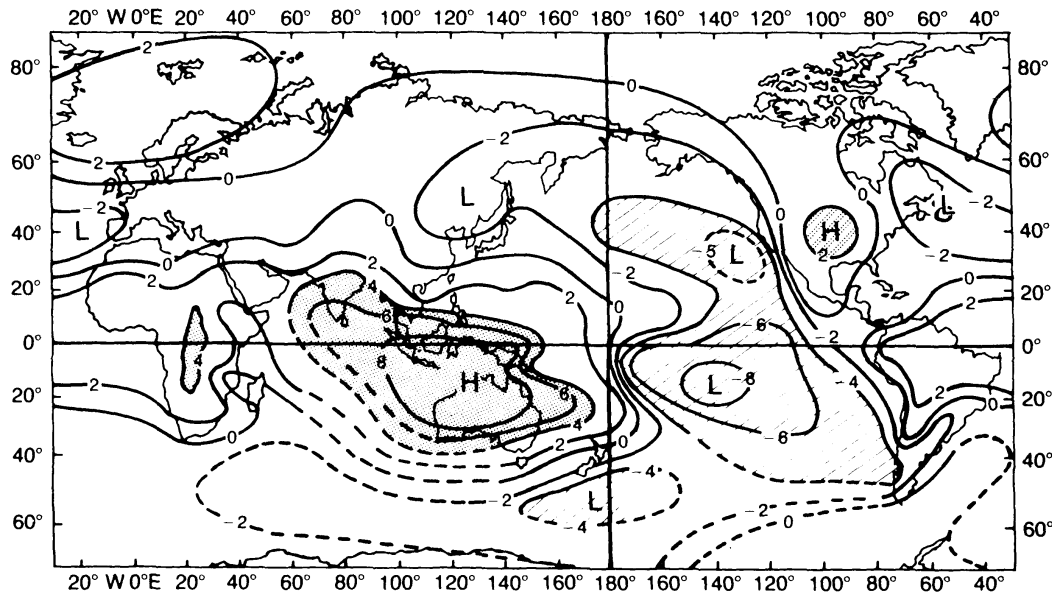


FIG. 1.7 – *Distribution spatiale des coefficients de corrélation ($\times 10$) entre les pressions atmosphériques de surface et la pression atmosphérique à Darwin, Australie. (source Trenberth et Shea (1987)).*

1.4 La variabilité interannuelle

El Niño et La Niña représentent les phases opposées du cycle ENSO, une oscillation du système couplé océan-atmosphère entre des conditions inhabituellement chaudes (El Niño) et froides (La Niña) dans le Pacifique centre équatorial. Les événements El Niño ont une périodicité de l'ordre de 3 à 7 ans avec une durée moyenne de 18 mois. El Niño signifie en espagnol le petit garçon ou l'enfant-Jésus. Il a été nommé ainsi par les pêcheurs des côtes Sud Américaines qui observaient épisodiquement un courant chaud responsable de la disparition des bancs d'anchois sur les côtes du Pérou et de l'Equateur, ceci au moment de Noël. Ce n'est que bien plus tard, dans le courant des années 1960 que les océanographes ont réalisé que les eaux anormalement chaudes n'étaient pas confinées le long des côtes du Pérou et étaient associées à des anomalies de grande échelle présentes dans l'ensemble du bassin tropical. L'oscillation australe (SO) est la composante atmosphérique du phénomène. Elle se caractérise par une oscillation des pressions de surface, donc par une redistribution des masses, entre la région Australie-Indonésie dans le Pacifique Ouest et la région du Pacifique Tropical Sud-Est. Quand la pression atmosphérique est relativement élevée dans la première région, elle est relativement basse dans la deuxième, et vice versa (cf figure 1.7).

Ce n'est qu'à la fin des années 60 que le couplage entre l'oscillation australe et El Niño fut mise en évidence par Bjercknes (1969). Ce couplage est flagrant sur la figure 1.8. L'indice SOI (pour Southern Oscillation index) nous renseigne sur l'intensité et la phase de l'oscillation australe. Il est basé sur la fluctuation mensuelle d'anomalie de pression entre Tahiti et Darwin. L'indice Niño-3.4 est quant à lui calculé à partir des anomalies mensuelles de SST dans la région $5^{\circ}N - 5^{\circ}S, 120^{\circ} - 170^{\circ}W$. Les anomalies supérieures à 0.5° correspondent aux événements El Niño, alors que celles inférieures à -0.5° sont à

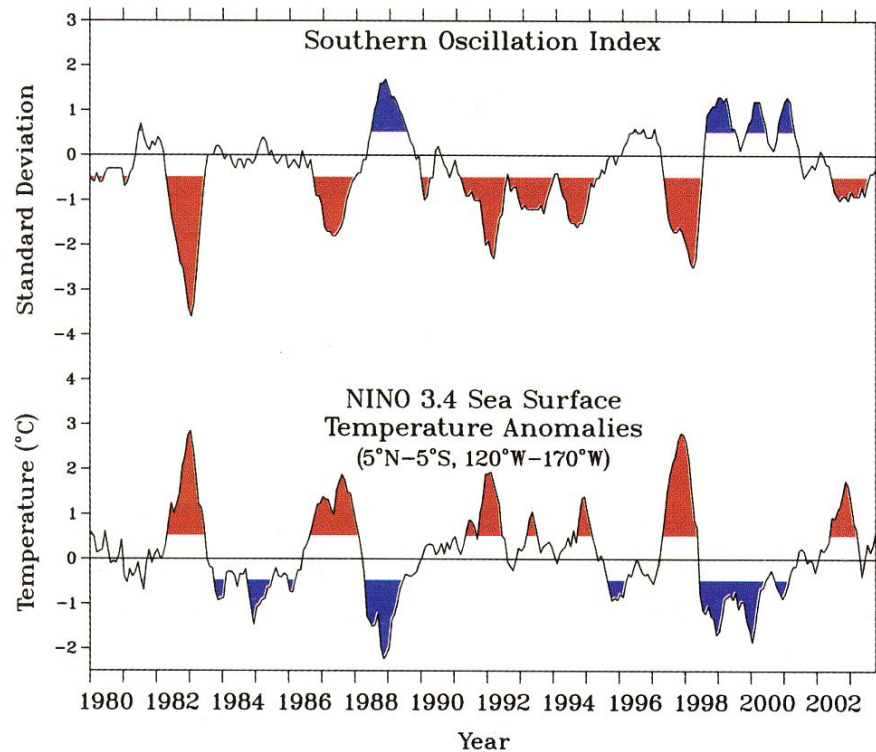


FIG. 1.8 – L'indice SOI et l'anomalie de SST pour la région Niño-3.4.

relier avec les conditions La Niña. Les périodes où l'indice SOI et/ou l'indice Niño-3.4 a une magnitude supérieure à 0.5 sont colorées de manière à souligner la corrélation entre les deux indices.

Les conditions “normales”, ont été détaillées dans la section 1.2 décrivant l'état moyen du Pacifique Tropical. Elles sont schématisées sur la figure 1.10.

Pendant El Niño, les alizés diminuent en intensité du fait de l'augmentation de la pression atmosphérique à l'Ouest et de sa diminution à l'Est. Du fait de ces vents plus faibles qui s'accompagnent généralement de coups de vent d'Ouest (WWB pour Westerly Wind Burst), il se crée un déséquilibre dans le gradient de pression zonal. La Warmpool migre alors vers le centre du bassin. Parallèlement, les WWB, en exerçant une pression sur la surface de la mer, génèrent deux types d'onde. Les ondes de Kelvin se propagent d'Ouest en Est à une vitesse proche de 3,0 m/s et les ondes de Rossby (nous nous limitons ici au premier mode méridien) se propagent d'Est en Ouest à une vitesse proche de 1,0 m/s. L'onde de Kelvin dite de “downwelling”, élève le niveau de la mer et enfonce la thermocline, alors que l'onde de Rossby dite d’“upwelling” produit les effets inverses. Ces ondes entraînent donc une plongée de la thermocline dans le bassin Est et une remontée dans le bassin Ouest. La pente Est-Ouest de la thermocline devient alors quasi nulle sur toute la largeur du Pacifique (cf figure 1.10). Le SEC change de sens pour former le courant Est équatorial (EEC pour East Equatorial Current) qui vient intensifier le NECC. En subsurface, l'EUC finit par disparaître. La capacité des upwellings équatoriaux à faire remonter les eaux froides de l'océan intérieur pour refroidir les eaux de surface est alors diminuée dans les régions où la thermocline s'est approfondie. On a donc un fort

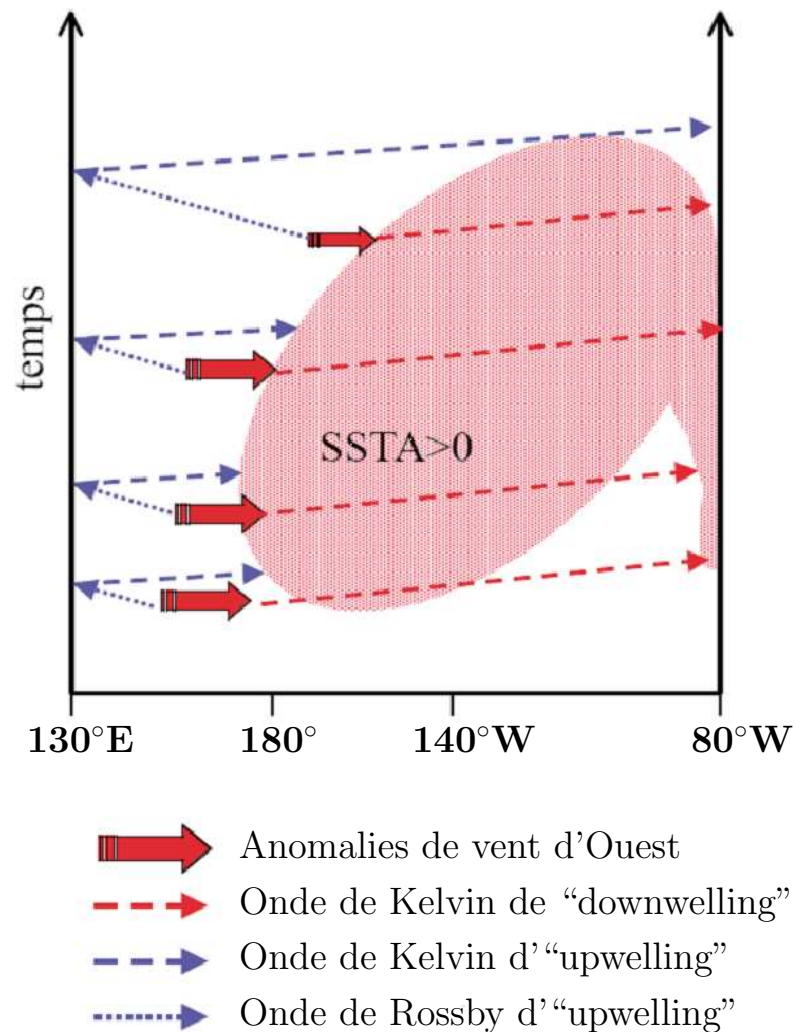


FIG. 1.9 – Représentation schématique de l'oscillateur retardé présentant la croissance et la terminaison d'une anomalie chaude.

réchauffement de la SST dans le centre et l'Est du Pacifique Tropical dû à la combinaison de deux phénomènes : (i) la réduction des apports d'eaux froides par le biais des upwellings équatoriaux et (ii) l'advection zonale de la Warmpool vers l'Est. Du côté atmosphérique, les cumulus et les fortes pluies associés à la branche ascendante de la cellule de Walker, qui sont habituellement situés dans l'Ouest du Pacifique, migrent eux aussi vers l'Est au delà de 180° suivant les eaux anormalement chaudes ($SST \geq 28^\circ$) dans l'Est du bassin. Ce déplacement vers l'Est de la zone de forte convection augmente la probabilité de sécheresse en Australie, en Indonésie et aux Philippines. Réciproquement, la probabilité de pluies torrentielles augmente dans les îles du Pacifique Central et le long des côtes Sud Américaines. Ce déplacement accentue les conditions El Niño (rétroaction positive). On est dans la phase croissante du phénomène. Par ailleurs, les pluies tropicales libèrent de l'énergie dans la moyenne et haute troposphère. Le déplacement de cette source d'énergie modifie la circulation atmosphérique globale et exporte l'influence d'El Niño sur la quasi totalité du globe (cf figure 1.1).

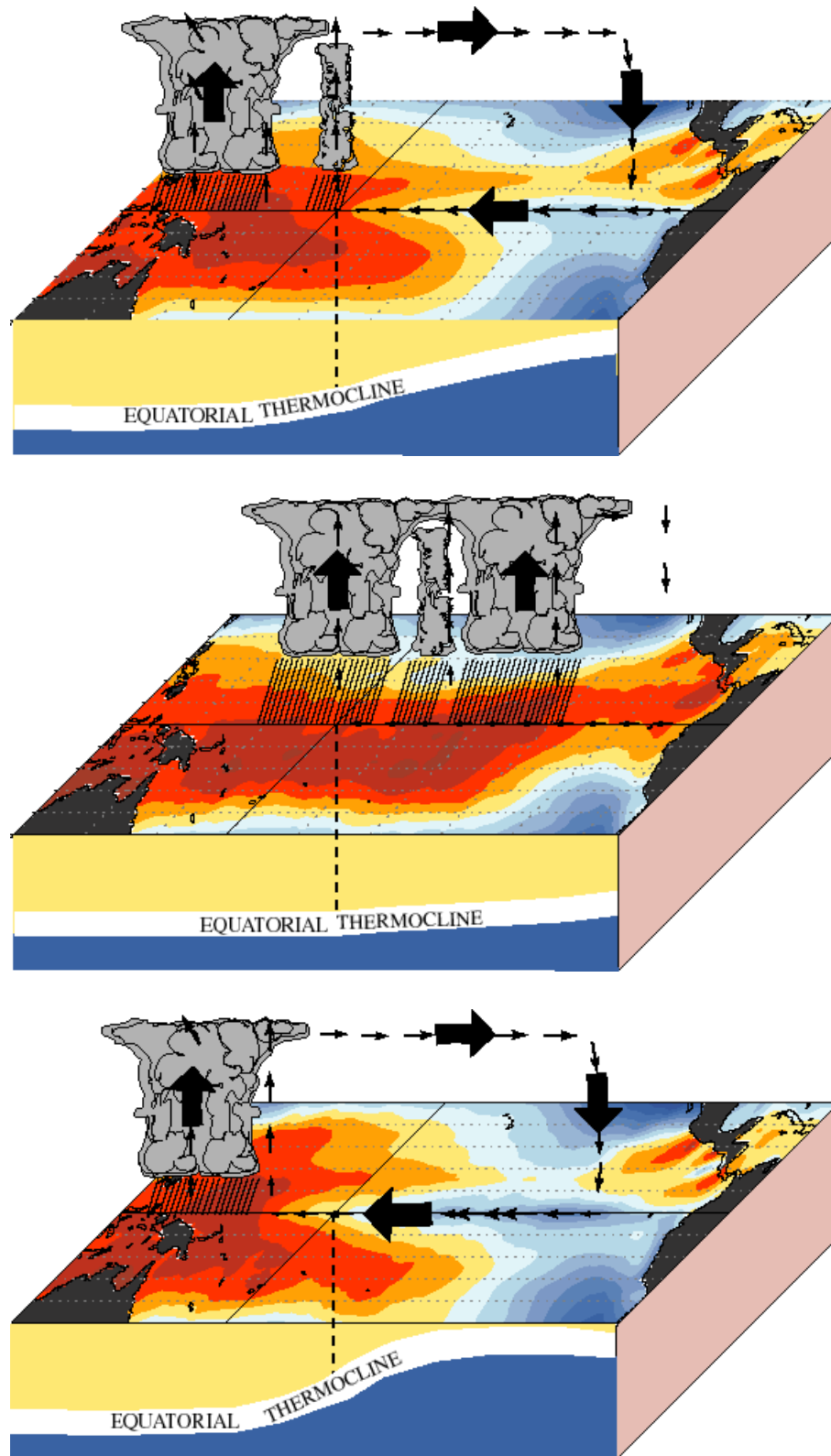


FIG. 1.10 – Représentation schématique du couplage océan-atmosphère en situation normale (en haut), en situation El Niño (au milieu) et en situation La Niña (en bas).

Un certain nombre de théories ont été développées au cours des années 80 et 90 pour expliquer la nature oscillante du phénomène ENSO. Elles sont toutes basées sur l'hypothèse, postulée par *Bjerknes* (1969) et présentée ci-dessus, qu'un feedback positif entre l'océan et l'atmosphère impliquant la cellule de Walker est responsable des anomalies chaudes de SST dans le Pacifique Central et le Pacifique Est (*Wang et Picaut*, 2003). Il n'y a en revanche pas de consensus pour expliquer la terminaison de l'évènement chaud. La théorie du "delayed oscillator", largement développée pendant les années 80 avec les travaux de *McCreary* (1983) et de *Suarez et Schopf* (1988) utilise les ondes de Rossby d'"upwelling" engendrées par les WWB comme feedback négatif (cf figure 1.9). En arrivant sur la côte Ouest du bassin, ces ondes se transforment, lors de leur réflexion, en ondes de Kelvin, toujours d'"upwelling". Ces dernières se propagent alors vers le Pacifique Central où elles affaiblissent le signal positif du niveau de la mer contribuant à la décroissance des anomalies chaudes en provoquant notamment une remontée de la thermocline et des eaux froides, plus profondes, vers la couche superficielle. Suite à cette décroissance, les anomalies de vent positives décroissent. Les ondes de Kelvin d'"upwelling", renforcées par des anomalies de vent d'Est dans le Pacifique Ouest, traversent alors le bassin et contribuent à la fin de l'évènement El Niño et au passage à La Niña. La Niña se caractérise par des alizés plus intenses que la normale, des eaux de surface également plus froides que la normale et un décalage des fortes pluies dans l'extrême Ouest du Pacifique (cf figure 1.10). En d'autres termes, les caractéristiques d'un épisode La Niña sont principalement celles d'un renforcement du cycle saisonnier selon le phasing de ENSO avec une intensification des alizés et une langue d'eau froide équatoriale pénétrant plus à l'Ouest de sa position normale. Les effets au niveau global de La Niña sont approximativement, mais pas exactement, les effets inverses d'El Niño. Il est à noter que cette période froide La Niña ne suit pas obligatoirement la période chaude El Niño.

Le théorie du "delayed oscillator" décrite ci-dessus et reposant sur un système de propagation des ondes de Kelvin et de Rossby reste critiquable en plusieurs points. Elle ignore les processus d'advection et considère le Pacifique Ouest comme une région inactive. Les années 90 ont vu apparaître de nouvelles théories pour expliquer la nature oscillante du phénomène ENSO. Elles reposent toutes sur le même feedback positif mais utilisent un feedback négatif différent pour expliquer le passage de la phase chaude à la phase froide. La théorie du "western Pacific oscillator", par exemple, insiste sur les anomalies de vents dans le Pacifique Ouest, qui selon *Weisberg et Wang* (1997) représentent le feedback négatif qui permet au système couplé d'osciller. *Jin* (1997a,b) avance quant à lui la théorie du "recharge-discharge oscillator" qui postule que les oscillations du système sont en fait dues à un stockage et un déstockage du contenu de chaleur des couches de surface de l'océan aux abords de l'équateur. Enfin, *Picaut et al.* (1997) ont proposé la théorie de l'"advective-reflective oscillator" qui présume que le mécanisme important pour ENSO n'est pas les réflexions au bord Ouest comme le suggère l'oscillateur retardé, mais les réflexions des ondes de Kelvin en ondes de Rossby au bord Est du Pacifique. Ces ondes en revenant vers le Pacifique Central seraient à même d'advecter les eaux chaudes vers le Pacifique Central et ainsi de contribuer à la terminaison de l'évènement chaud. Il est vraisemblable que tous les mécanismes intervenant dans les quatre théories ci-dessus opèrent dans la nature et sont donc à considérer pour expliquer la variabilité observée.

Chapitre 2

L'observation de l'océan : les données

Sommaire

2.1	Un petit historique de la mesure de l'océan	26
2.2	Les données satellites : le nouveau duo altimétrie/gravimétrie	28
2.2.1	La mesure altimétrique	30
2.2.2	La mesure gravimétrique	31
2.2.3	La topographie dynamique utilisée	34
2.3	Les données in-situ	36
2.3.1	Les données TAO	36
2.3.2	Les données XBT	36

2.1 Un petit historique de la mesure de l'océan

Les mesures et observations de l'océan ne datent pas d'aujourd'hui. Si les premières expéditions maritimes n'avaient pas pour but de décrire l'océan d'un point de vue scientifique mais plutôt de le cartographier et de trouver de nouvelles routes maritimes afin de développer le commerce entre les différentes parties du connues globe, elles ont quand même fourni de précieuses observations. En 1770, par exemple, une première carte du Gulf-Stream est réalisée par B. Franklin. C'est la pose du premier câble télégraphique sous-marin en 1851 entre la France et la Grande Bretagne qui a fait prendre conscience de la nécessité de mieux connaître les fonds marins (topographie, température, courants, etc...).

On considère généralement que l'expédition britannique du navire Challenger de 1872 à 1876, avec ses quelques 400 stations de mesures réparties dans les océans Atlantique, Pacifique et Indien, marque le début de la science "océanographique". D'autres expéditions ont suivi, notamment les 28 campagnes océanographiques du Prince Albert *I^{er}* de Monaco de 1885 à 1915. Néanmoins, les moyens dont on disposait à l'époque rendaient les campagnes de mesures et leurs exploitations très longues et surtout extrêmement coûteuses. Elles sont donc le plus souvent restées circonscrites à l'étude de certaines régions ou de certaines thématiques. Il a fallu attendre la deuxième partie du *XX^{ème}* siècle et ses formidables progrès technologiques, notamment l'apparition des mouillages autonomes hauturiers à partir des années 60, pour accéder à une réelle connaissance de la dynamique des océans.

Dans les années 1980-1990, le déploiement des mouillages autonomes ATLAS composant le réseau TAO/TRITON¹ dans le Pacifique Tropical a été motivé par l'évènement El Niño de 1982-1983, le plus intense de siècle à ce moment là, qui a causé, à l'échelle mondiale des pertes économiques à hauteur de plusieurs milliards de dollars et humaines à plus d'un millier de morts. Ce fort évènement a pourtant pris les scientifiques par surprise et il n'avait pas été prédit ni même détecté avant qu'il ait quasiment atteint son intensité maximum. Cette évènement souligna le besoin de données temps-réel du Pacifique Tropical. Près de dix ans auront été nécessaire à la mise en place le réseau qui a été terminé en décembre 1994. Depuis cette date, les mouillages TAO délivrent en temps réel des données météorologiques de surface (température de l'air, vent, précipitation, etc ...) et océanographiques de subsurface (température, salinité et courant pour certains mouillages). Dans le prolongement du succès scientifique du déploiement des 70 bouées constituant le réseau TAO/TRITON¹ dans le Pacifique Tropical, d'autres programmes ont vu le jour comme PIRATA² (Pilot Research Moored Array in the Tropical Atlantic), un programme d'océanographie opérationnelle mis en place en 1997 sous l'égide du programme international CLIVAR³ (Climate Variability and predictability) afin d'étudier les interactions océan-atmosphère dans l'Atlantique Tropical. Plus récemment, devant la nécessité de données d'observations synoptiques de l'océan profond, notamment pour les besoins des études climatiques et de l'océanographie opérationnelle, le projet ARGO⁴ a été initié. En effet, l'océan profond reste globalement sous observé. Cette lacune, résulte de l'absence de dispositifs automatiques peu coûteux permettant d'obtenir un profil sur toute la colonne d'eau, l'équivalent d'un ballon sonde qui explore les propriétés de l'atmosphère sur la ver-

¹<http://www.pmel.noaa.gov/tao/>

²<http://www.ifremer.fr/ird/pirata/>

³<http://www.clivar.org>

⁴<http://www.argo.net>

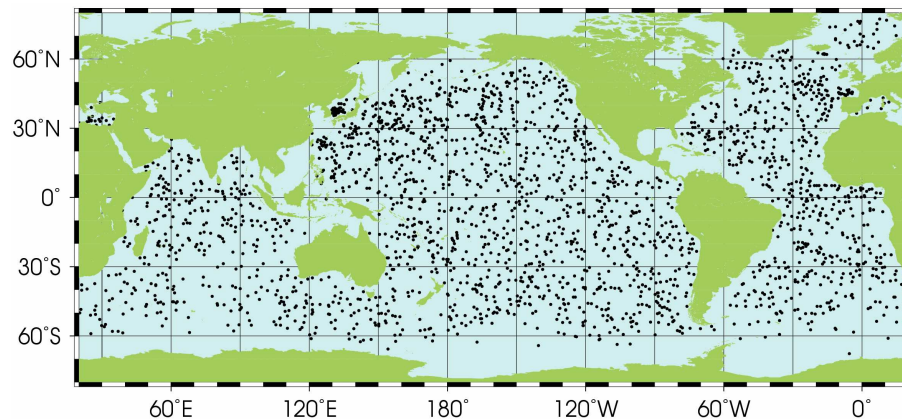


FIG. 2.1 – *Position des flotteurs ayant délivré des données dans les 30 derniers jours le 7 août 2006 (source site ARGO).*

ticale, mais pour l'océan. C'est devant ce constat que le déploiement du réseau de flotteurs profileurs ARGO⁴ a été initié en 2000 dans le cadre des programmes internationaux Global Climate Observing System⁵/Global Ocean Observing System⁶ (GCOS/GOOS). Le projet ARGO vise à déployer et à maintenir un réseau global d'environ 3000 flotteurs profilants, répartis sur un maillage de $3^\circ \times 3^\circ$, mesurant des profils de température et de salinité jusqu'à des profondeurs de 2000m, d'où leur nom de "profileur Argo", ainsi que la vitesse du courant à des profondeurs variables. Chaque flotteur remonte tous les dix jours à la surface, transmet ses données via des satellites, à un centre à terre, avant de replonger à 2000m. Dix-sept pays plus l'Union Européenne participent au réseau Argo. Ce dernier est toujours en cours de déploiement et compte aujourd'hui environ 2500 bouées profilantes réparties sur l'ensemble de l'océan mondial comme le montre la figure 2.1.

Cependant, aussi précises soient ces observations, on se heurte toujours au problème de la couverture spatiale des données. Pour avoir une vision globale, nécessaire à la bonne compréhension de la dynamique océanique il faudrait échantillonner l'océan mondial avec une résolution spatio-temporelle élevée, et cela de manière permanente, ce qui est bien entendu pratiquement et financièrement irréalisable à ce jour.

L'apparition des techniques spatiales d'observation de la Terre et de son environnement a permis aux scientifiques de poser les problèmes dans un cadre global. De nombreux instruments ont été embarqués pour mesurer la température, la couleur ou la diffusion à la surface de l'eau. Mais ce sont les missions altimétriques qui ont révolutionné la vision des océans. Des progrès considérables ont été réalisés ces trente dernières années grâce à l'avènement des observations satellites des océans qui ont enfin offert cette vision globale, continue et homogène, nécessaire à la bonne compréhension des processus et de la dynamique de l'océan par la communauté océanographique. Ces données représentent un formidable complément des données in-situ, les seuls capables de fournir des mesures sur la verticale.

C'est au début des années 70 que les satellites ont commencé à transmettre régulièrement des informations sur la physique, la chimie et la dynamique des océans. C'est aussi l'époque des premières mesures altimétriques pour déterminer la topographie de la sur-

⁵<http://www.wmo.ch/web/gcos/gcoshome.html>

⁶<http://www.ioc-goos.org>

face océanique. Les Etats-Unis sont les premiers à avoir fait voler un altimètre dédié à bord de satellites comme Seasat en 1978, puis Geosat (*Douglas et Cheney*, 1990) en 1985. Avec les années 90, de nouvelles missions altimétriques sont lancées : ERS-1⁷ (en 1991), Topex/Poséidon⁸ (en 1992), ERS-2⁷ (en 1995). Depuis décembre 2001, le satellite Jason⁹ assure la continuité de la série de mesures effectuées depuis 1992 par le satellite Topex/Poséidon. Depuis le lancement des premiers satellites altimétriques, les données altimétriques ont été largement utilisées par la communauté océanographique pour mieux comprendre le système océan et son évolution à diverses échelles spatiales et temporelles (*Fu et Cazenave*, 2001). Le satellite Topex-Poséidon, de par la précision de ces mesures, a, plus que tout autres, marqué un tournant capital dans l'étude des mouvements océaniques, contribuant à rendre indispensable l'altimétrie satellitaire. A titre d'exemple, les observations altimétriques fournies par Topex-Poséidon au cours de son premier mois d'exploitation ont permis d'établir la carte de la topographie de l'océan mondial avec une précision supérieure à ce que l'on était parvenu à faire à partir de l'ensemble des cents et quelques années de données in-situ collectées depuis la campagne Challenger. En plus de la variabilité du niveau de la mer, les satellites nous donnent aujourd'hui accès à la température de surface, l'état de la mer, les vents, ou encore la couleur de l'eau. Depuis l'année 2000 et le lancement de la mission CHAMP¹⁰ (CHALLENGING Minisatellite Payload) relayée par la mission GRACE¹¹ (Gravity Recovery and Climate Experiment) en 2002 les satellites nous donnent également un accès direct à des observations du géoïde terrestre. Le géoïde représentant la surface équipotentielle du champ de pesanteur terrestre assimilable au niveau de référence des océans. Cette grandeur est un complément indispensable de la mesure altimétrique comme nous allons le voir dans ce chapitre. Les développements sont constants. La mission GOCE¹² (Gravity Field and Steady-State Ocean Circulation Explorer) prévue pour 2007 nous donnera accès à une géoïde haute résolution. La mission SMOS¹³ s'envolera à l'horizon 2007 afin d'effectuer la première cartographie à l'échelle globale de l'humidité des sols et de la salinité de surface (SSS) des océans. Elle sera suivie de peu par la mission américaine AQUARIUS¹⁴ dédiée à l'étude de la SSS en 2008.

Dans la suite de ce chapitre, nous détaillerons plus particulièrement les données utilisées dans le cadre des expériences d'assimilation réalisées au cours de ce travail de thèse. Ceci englobe les données assimilées mais aussi les données utilisées afin de valider nos différentes simulations.

2.2 Les données satellites : le nouveau duo altimétrie/gravimétrie

Depuis le lancement des premiers satellites altimétriques dans les années 70, les données altimétriques ont été largement utilisées par la communauté océanographique afin de mieux comprendre le système océan (*Fu et Cazenave*, 2001). La topographie de la sur-

⁷<http://earth.esa.int/ers/>

⁸http://www.cnes.fr/html/_112.810_.php

⁹http://www.cnes.fr/html/_112.788_.php

¹⁰<http://www.gfz-potsdam.de/pb1/op/champ/>

¹¹<http://www.gfz-potsdam.de/pb1/op/grace/>

¹²<http://www.esa.int/esaLP/LPgoce.html>

¹³http://www.cnes.fr/html/_112.821_.php

¹⁴<http://aquarius.nasa.gov>

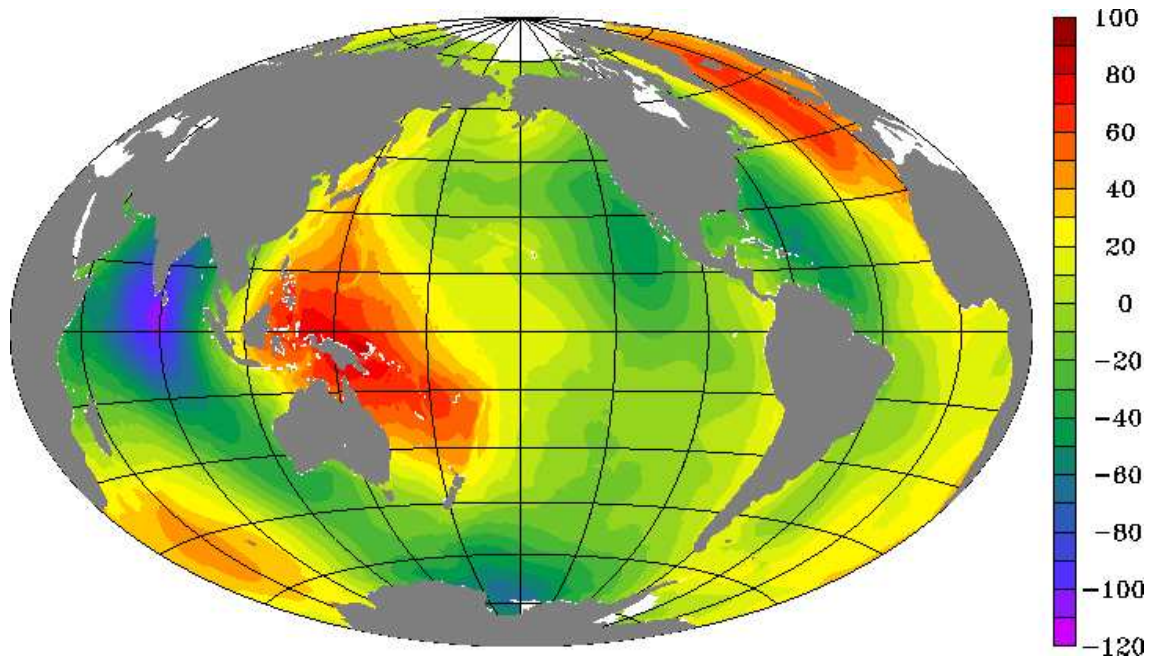


FIG. 2.2 – Surface moyenne (MSSH) en mètres mesurée par altimétrie satellite (Hernandez et al., 2001).

face océanique mesurée par les satellites altimétriques fournit une information intégrée sur toute la colonne d'eau. Elle correspond à la mesure du niveau des océans (SSH) à un instant donné par rapport à une ellipsoïde de référence (cf section 2.2.1). L'altimétrie a joué un rôle de catalyseur dans les formidables développements scientifiques et applicatifs de l'océanographie ces dernières années, en particulier pour le développement de l'océanographie opérationnelle. Cependant, seule la partie résiduelle (SLA) du signal altimétrique mesuré par les satellites peut être utilisée de manière fiable pour l'étude de la circulation océanique. En effet, les satellites altimétriques mesurent avec une précision centimétrique (une erreur de 3 cm RMS est couramment admise) et une résolution unique la surface océanique, ou SSH. Or, cette surface océanique représente plusieurs effets combinés (les effets dus à la circulation océanique appelés topographie dynamique et les effets dus aux variations d'attraction terrestre appelés géoïde). La topographie dynamique, qui est la variable pertinente en océanographie, est, en théorie, obtenue en retranchant la hauteur de géoïde à la SSH mesurée par l'altimétrie. Cependant, la topographie dynamique est contaminée par des erreurs importantes sur le géoïde, en particulier pour les harmoniques élevées (harmoniques supérieures à 20 environ). C'est donc la moyenne temporelle de la SSH, la surface moyenne océanique ou MSSH qui est retranchée à la SSH, donnant ainsi l'accès à la partie variable du signal océanique. La figure 2.2 montre la MSSH CLS01 proposée par Hernandez et al. (2001). Elle a été calculée à partir de 7 ans de données (de 1993 à 1999) et en utilisant quatre satellites : GEOSAT, ERS1&2 et TOPEX-Poseidon. Elle représente le produit de surface moyenne océanique le plus abouti à ce jour. C'est notamment cette MSSH qui sert de référence pour le calcul des anomalies (SLA) distribuées par AVISO¹⁵. On notera que la figure 2.2 fait apparaître des structures spatiales très marquées avec

¹⁵<http://www.aviso.oceanobs.com>

un minimum de l'ordre de -120 m au sud de l'Inde et un maximum avoisinant les 100 m au large de la Papouasie-Nouvelle-Guinée. Elle est très éloignée des surfaces moyennes utilisées par la communauté océanographique et simulée par les OGCMs. La MSSH est dominée, aux grandes échelles par les effets dus aux variations de la gravité et est donc très proche du géoïde. Elle peut d'ailleurs être utilisée pour estimer le géoïde (*Rapp*, 1995). Cette méthode ne présente cependant pas d'intérêt pour les océanographes puisque cette estimation du géoïde inclue la topographie dynamique.

2.2.1 La mesure altimétrique

Afin de mieux comprendre le problème, revenons au principe de la mesure altimétrique. La figure 2.3 présente schématiquement les grandeurs caractéristiques de la mesure altimétrique. Le principe est simple : l'altimètre émet une onde radar vers la surface océanique et reçoit en retour l'écho réfléchi par la surface de l'océan, qui forme pour les ondes électromagnétiques une surface réfléchissante avec un faible taux d'absorption. Après traitement du signal, le temps de parcours donne la distance R entre le satellite et la surface océanique. Le niveau des océans s'obtient alors par simple différence entre l'orbite du satellite H et la distance altimétrique :

$$SSH = H - R \quad (2.1)$$

La mesure altimétrique SSH , qui est donc la hauteur de l'océan référencée par rapport à une ellipsoïde de référence, est une valeur intégrale qui représente plusieurs effets combinés :

- La hauteur du géoïde qui reflète les variations du champ de gravité terrestre d'un point à l'autre de la Terre. Cette surface représente la topographie qu'aurait l'océan en absence de toutes perturbations (vent, marées, courants, etc...). Ces variations sont liées à la distribution in-homogène des champs de masse et de densité dans la croûte et le manteau terrestre. Elles représentent des amplitudes à grande échelle de plusieurs centaines de mètres. On notera que le champ de gravité terrestre présente une variabilité temporelle due à des redistributions globales de masse, mais cette dernière est généralement faible, lente et à grande échelle. Elle sera négligé dans ce qui suit.
- La topographie dynamique qui est la variable pertinente en océanographie. Elle représente, en effet, la signature de la circulation océanique qui comprend une partie permanente stationnaire (circulation permanente liée à la rotation de la terre, aux vents permanents, etc.) et une partie fortement variable (liée aux vents, à la variabilité saisonnière, etc.). La DT présente des amplitudes de l'ordre du mètre en moyenne.

Pour accéder à la topographie dynamique qui est donc la SSH référencée au géoïde, il suffit donc, en théorie, de retrancher la hauteur du géoïde G à la hauteur du niveau de la mer :

$$DT = SSH - G \quad (2.2)$$

En pratique, on commence à peine à avoir une connaissance suffisamment précise du géoïde pour déduire la topographie dynamique par la méthode "directe" (équation 2.2).

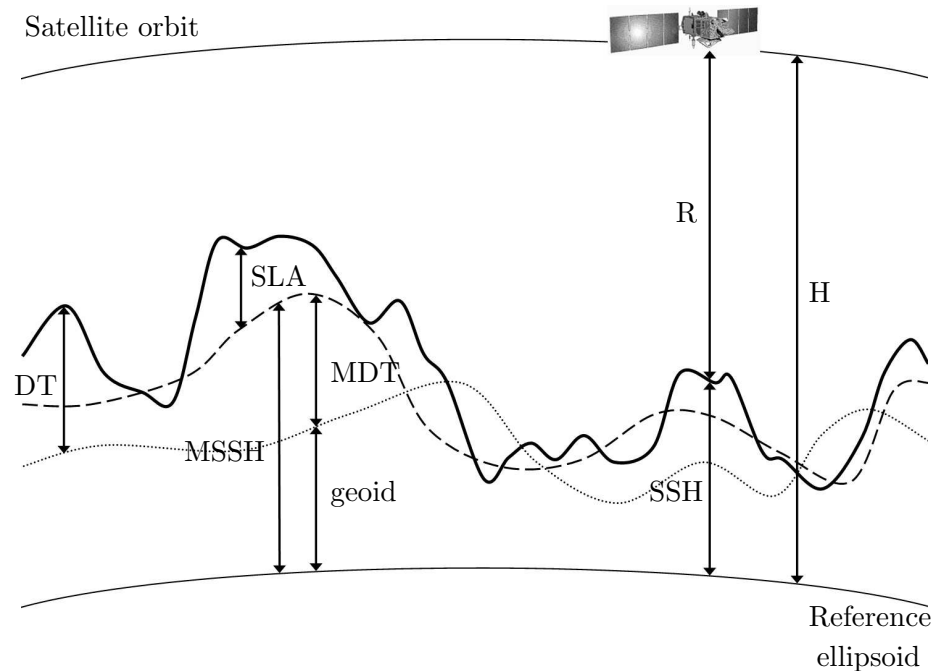


FIG. 2.3 – Principe de la mesure altimétrique.

Jusqu'à présent, c'est donc la surface moyenne océanique (géoïde + circulation permanente) présentée sur la figure 2.2 qui était retranché à la SSH, donnant ainsi accès à la partie variable du signal océanique (SLA) avec une précision élevée. On notera que la DT, respectivement la MDT, sont couramment et abusivement appelées SSH, respectivement MSSH, par les modélisateurs. La plupart des OGCMs faisant l'hypothèse que le champ de gravité présente une symétrie sphérique, ces deux quantités sont en effet identiques dans l'espace modèle.

2.2.2 La mesure gravimétrique

Depuis maintenant une trentaine d'années, le champ de gravité terrestre est étudié, notamment en suivant et étudiant les perturbations des trajectoires d'un certain nombre de satellites en orbite autour de la Terre, conduisant à des améliorations significatives sur l'estimation du géoïde (Nerem *et al.*, 1995). Si cette méthode a permis de développer des modèles de géoïde précis, surtout pour les grandes longueurs d'onde (Lemoine *et al.*, 1998; Biancale *et al.*, 2000), elle s'avère en revanche inadaptée à fournir les petites échelles du géoïde nécessaires à une large gamme d'applications, et notamment à la détermination précise d'une MDT. Les modèles de géoïde les plus récents et les plus précis basés sur les perturbations des orbites satellites sont les modèles EGM96S (Lemoine *et al.*, 1998) complet jusqu'à l'harmonique 70 et GRIM5-S1 (Biancale *et al.*, 2000) complet jusqu'à l'harmonique 99. Lemoine *et al.* (1998) proposent également le modèle EGM96 complet jusqu'à l'harmonique 360 et incorporant des données altimétriques et des données gravimétriques in-situ (et donc inadéquate pour référencer l'altimétrie). La précision pour les modèles purement satellites ne dépasse pas 40 cm avec une résolution spatiale de 500 km. Les principales limitations de la méthode sont dues à l'atténuation du champ de gravité avec l'altitude, la disparité des données et la difficulté de modélisation des forces non-

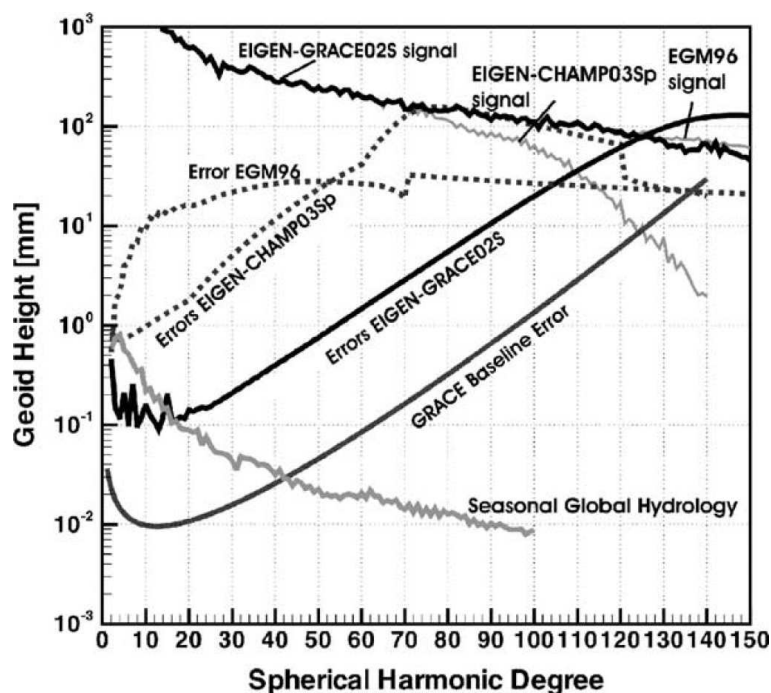


FIG. 2.4 – Amplitude du signal et de l'erreur par degré pour différents modèles de géoïde (source Reigber et al. (2005)).

gravitationnelles (Biancale et al., 2000). Ces erreurs restent donc incompatibles avec une utilisation de ce type de géoïde pour référencer les données mesurées par les satellites altimétriques (Le Provost et Brémond, 2003).

Avec le lancement en 2000 de la mission CHAMP¹⁰ (CHALLENGING Minisatellite Payload), pour la première fois le champ de gravité terrestre a pu être déterminé à partir des instruments d'un satellite unique. CHAMP embarque en effet un antenne GPS pour une estimation précise de la position et de la vitesse, un accéléromètre afin de mesurer toutes les forces non-gravitationnelles agissant sur le satellite et deux "star-cameras" pour une orientation précise du satellite dans le référentiel géocentrique. Grâce à la connaissance précise de l'orbite du satellite et à la mesure directe des forces non-gravitationnelles venant perturber cette orbite, les modèles du champ de gravité terrestre se sont fortement améliorés par rapport à la période pré-CHAMP. La mission GRACE¹¹ (Gravity Recovery and Climate Experiment) lancée le 17 mars 2002 a pour principal objectif de cartographier le champ de gravité terrestre et sa variabilité pour une période de 5 ans avec une précision sans précédent pour les moyennes et grandes longueurs d'onde (de quelques centaines de kilomètres à 40000 km). C'est un projet conjoint entre la NASA (National Aeronautics and Space Administration) et le DLR (Deutsches Zentrum für Luft- und Raumfahrt). La configuration de GRACE est basée sur l'héritage de la mission CHAMP (les mêmes instruments sont embarqués avec toutefois un accéléromètre plus précis) et repose sur deux satellites strictement identiques qui se suivent sur la même orbite (environ 220 kilomètres séparent les 2 satellites). Les orbites des 2 satellites, qui sont dépendantes des effets de la gravité locale, sont influencées par ces effets avec une très légère différence de phase et sont donc perturbées différenciellement. Ces différences de perturbation engendrent des modi-

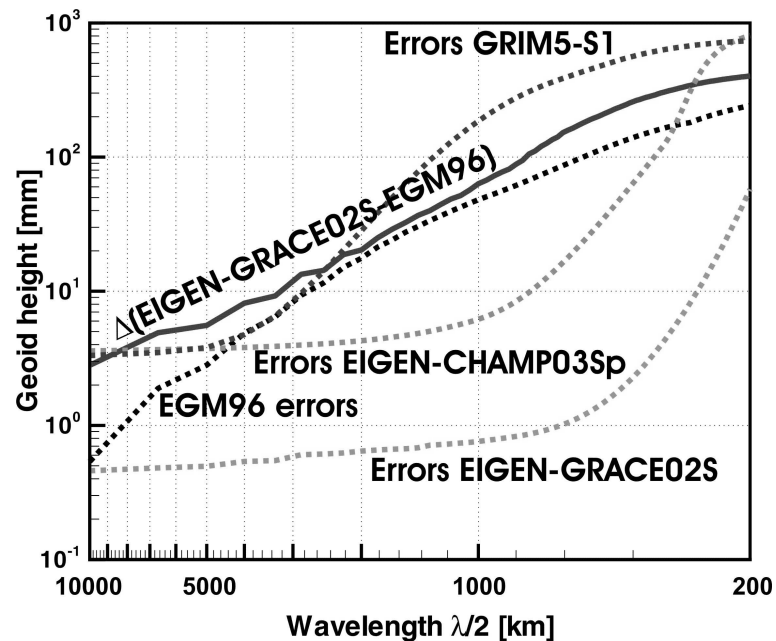


FIG. 2.5 – Erreur en fonction de la résolution spatiale pour différents modèles de géoïde (source Reigber et al. (2005)).

fications de la distance inter-satellite. Une mesure très précise de cette distance et de son taux de variation (avec une précision meilleure que, respectivement, $10 \mu\text{m}$ et $1 \mu\text{ms}^{-1}$) est utilisée pour enregistrer les variations hautes fréquences de la réponse au signal gravitationnel et ainsi améliorer la résolution du champ de gravité estimé. Il faudra cependant attendre la mission GOCE¹² (Gravity Field and Steady-State Ocean Circulation Explorer) prévue pour 2007 pour réellement connaître les variations à petites échelles du champ de gravité. Cette mission fait partie du programme “Living Planet” de l’agence spatiale européenne (ESA) et a pour principal objectif de déterminer le géoïde avec une précision de 1-2 cm, et ce, pour une résolution spatiale meilleure que 100 km.

Le modèle de géoïde utilisé au cours de ce travail de thèse est le modèle EIGEN-GRACE02S du GeoForschungsZentrum (GFZ) de Potsdam, en Allemagne (Reigber et al., 2005). Il a été calculé à partir de 110 jours de données GRACE et a été dérivé uniquement à partir des perturbations de l’orbite des deux satellites. Ce modèle est complet jusqu’à l’harmonique 150. La figure 2.4 montre l’amplitude du signal et des erreurs par degré pour EIGEN-GRACE02S, EIGEN-CHAMP03Sp (le dernier modèle CHAMP multi-annuel) et EGM96 (Lemoine et al., 1998), un modèle incorporant des données altimétriques et des données gravimétriques in-situ, donc non affecté par l’atténuation. L’allure de l’erreur est classique pour un modèle uniquement satellite et traduit une atténuation excessive des courtes longueurs d’onde du géoïde. Le signal n’est pas atténué jusqu’à l’harmonique 120 comme le montre la comparaison avec EGM96 qui n’est pas affecté par l’atténuation vu qu’il incorpore des données in-situ. Ceci représente un gain important par rapport à EIGEN-CHAMP03Sp qui commence à être atténué dès le degré 70. EIGEN-GRACE02S est à peu près un ordre de magnitude plus précis pour les grandes échelle que les modèles de géoïde basés sur les données CHAMP et plus de deux ordres de magnitude plus précis que les derniers modèles pré-CHAMP basés seulement sur les perturbations d’orbites satellites.

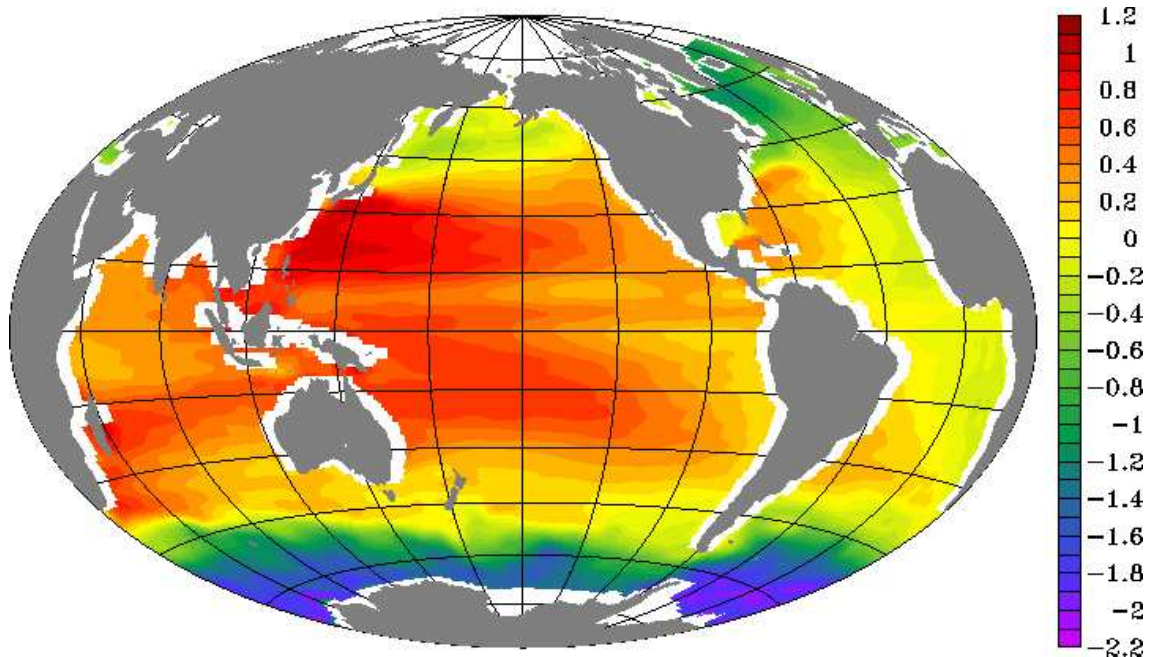


FIG. 2.6 – *Topographie dynamique moyenne (MDT) en mètres référencé au géoïde EIGEN-GRACE02S.*

La figure 2.5 montre l'amplitude de l'erreur en fonction de la résolution spatiale pour différents modèles de géoïde. Là aussi, EIGEN-GRACE02S représente une nette amélioration comparée aux modèles précédemment disponibles. La précision est supérieure au millimètre jusqu'à la résolution de 600 km et reste supérieure au centimètre jusqu'à la résolution de 275 km. Elle tombe ensuite à environ 10 cm à la résolution de 150 km. Les progrès sont donc substantiels à moyenne échelle comparés au précédent produit disponible. Ainsi, la précision centimétrique obtenue jusqu'à l'harmonique 75 représente un gain en résolution spatiale de 1000 à 275 km comparée au modèle de géoïde pré-CHAMP et de 400 à 275 km comparée au dernier modèle CHAMP multi-annuel et satellite uniquement.

2.2.3 La topographie dynamique utilisée

Une topographie dynamique moyenne, ci après dénommée MDT GRACE, est calculée par méthode directe en faisant la différence entre la MSSH CLS01 (*Hernandez et al.*, 2001) et le géoïde EIGEN-GRACE02S présenté ci dessus. Cette topographie dynamique moyenne est représentée sur la figure 2.6. La méthode directe qui s'apparente à une simple différence (cf équation 2.2) n'est en fait pas si évidente que cela à mettre en œuvre. En effet, le contenu spectral des deux surfaces (la MSSH et le géoïde) est très différent ce qui complique singulièrement la tâche. La MSSH contient des échelles spatiales jusqu'à, environ 10-20 km alors que le géoïde EIGEN-GRACE02S n'est défini que jusqu'au degré 150 (soit 133 km). Le calcul d'une simple différence donnera donc une MDT avec une résolution de 133 km, plus les échelles spatiales contenues dans la MSSH entre 10-20 km et 133 km, qui ne sont pas résolues par le géoïde. Ces signaux contiennent entre autres les plus petites échelles du géoïde, directement contraintes par la topographie. Pour résoudre le problème, on procède à une décomposition en harmoniques sphériques (HS) de cette

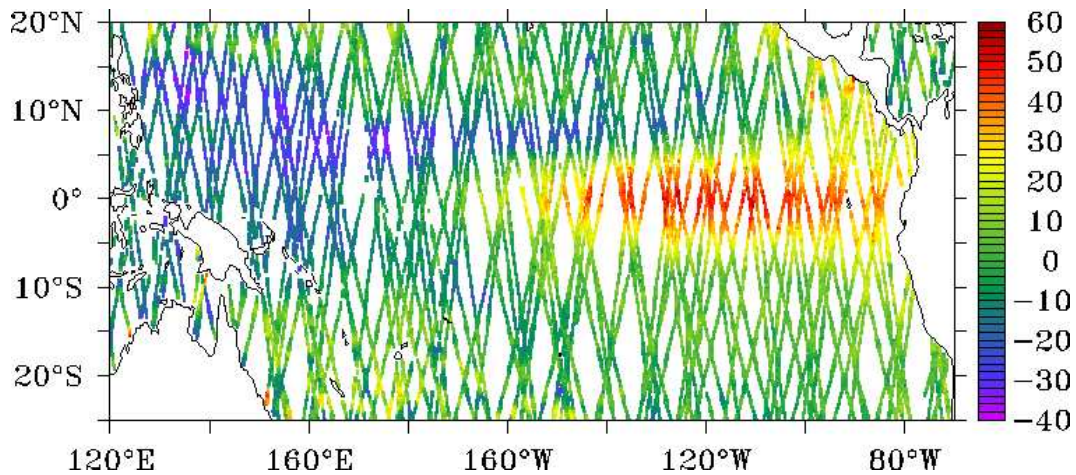


FIG. 2.7 – Anomalies de surface le long des traces pour l’analyse du 15 novembre 1997.

différence (la décomposition en HS d’un signal qui comporte des “trous”, ici les continents, est à elle seule une tâche très complexe). Ces HS sont ensuite tronquées pour tracer la MDT avec le degré de filtrage voulu. La topographie dynamique moyenne déduite du géoïde EIGEN-GRACE02S utilisée dans la cadre de cette thèse a été estimée lors de travaux dédiés menés au CNES¹⁶ (Centre National d’Etude Spatiale). Ces travaux sont dans la continuité de ceux déjà effectués à partir de la solution EIGEN2 intégrant les données du satellite CHAMP dans le cadre d’un projet soutenu au PNTS (*Gourdeau et al., 2003*).

Au final, la MDT GRACE utilisée au cours de ce travail de thèse est complète jusqu’à l’harmonique 60 (résolution spatiale de 333 km) et présente une erreur cumulée estimée à 4 cm à cette résolution.

Cette MDT est, dans un deuxième temps, combinée avec un produit de SLA afin de reconstruire un signal de topographie dynamique absolue.

$$\begin{aligned}
 DT &= SSH - geoid \\
 &= SSH - MSSH + MSSH - geoid \\
 DT &= SLA + MDT
 \end{aligned}
 \tag{2.3}$$

Ceci est identique à la méthode directe dans la mesure où le même produit de MSSH est utilisé pour évaluer la MDT et pour générer la SLA (cf équation 2.3). Les données de SLA “le long des traces” des satellites Topex/POSEIDON⁸, ERS1 et ERS2⁷ produites par Ssalto/Duacs et distribuées par AVISO¹⁵, avec le support du CNES¹⁷ ont été utilisées. Ces données sont référencées par rapport à la MSSH CLS01 (*Hernandez et al., 2001*), soit la même MSSH que celle ayant servi à évaluer le géoïde EIGEN-GRACE02S. La figure 2.7 illustre la carte des observations de SLA obtenues pour l’analyse de 15 novembre 1997. Ces données, additionnées à la MDT GRACE (cf équation 2.3) vont nous fournir la topographie dynamique absolue qui sera assimilée dans le modèle.

¹⁶<http://www.cnes.fr>

¹⁷<http://www.cnes.fr>

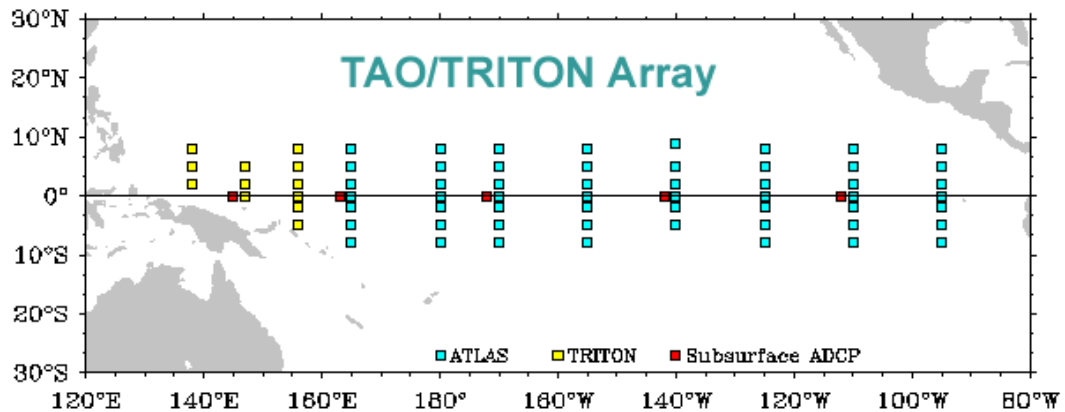


FIG. 2.8 – Position des mouillages TAO le long du rail équatorial.

2.3 Les données in-situ

Les données in-situ, les seules capables de fournir des mesures sur la verticale représentent un complément important, pour ne pas dire nécessaire, des données altimétriques. Le Pacifique Tropical est une des parties de l'océan mondial la mieux observée, grâce notamment au projet TOGA (Tropical Ocean Global Atmosphere), qui a mis en place l'ENSO Observing System (McPhaden *et al.*, 2001) afin d'étudier, de mieux comprendre et de prédire les évènements El Niño.

2.3.1 Les données TAO

Une des principales composantes de l'ENSO Observing System est le réseau de mouillages TAO/TRITON¹ (Tropical Atmosphere Ocean/Triangle Trans-Oceans Buoy Network). Il consiste en un réseau de 69 mouillages le long du rail équatorial (i.e. entre 8°N et 8°S). La figure 2.8 montre la position de chacun des mouillages.

Nous utiliserons ici les profils de température de subsurface sous forme de moyenne journalière. Les profils comportent 10 points de mesure entre la surface et 500 mètres de profondeur (20, 40, 60, 80, 100, 120, 140, 180, 300 et 500 mètres). Il est à noter que lors de la conception du réseau TAO, l'accent a été mis sur la couverture à grande échelle de la température de l'océan dans la couche supérieure, des courants à l'équateur et des vents. Le réseau a été conçu afin d'étudier et de prévoir les oscillations ENSO qui sont contrôlées, en partie, par la propagation d'ondes équatoriales à grandes échelles qui redistribuent le contenu de chaleur de la couche de mélange dans le bassin de l'océan Pacifique Tropical. Dans le cadre de cette thèse en revanche, les profils de températures TAO vont être assimilés dans un OGCM afin de contrôler le champ de masse du modèle. La couverture spatiale très disparate des données (surtout zonalement avec des mouillages tous les 10-15°) pourra dans ce cas être une limitation.

2.3.2 Les données XBT

Les profils verticaux de température XBT (eXpendable BathyThermograph), XCTD (Expendable Conductivity, Temperature and Depth Sensors) ou XBT provenant, res-

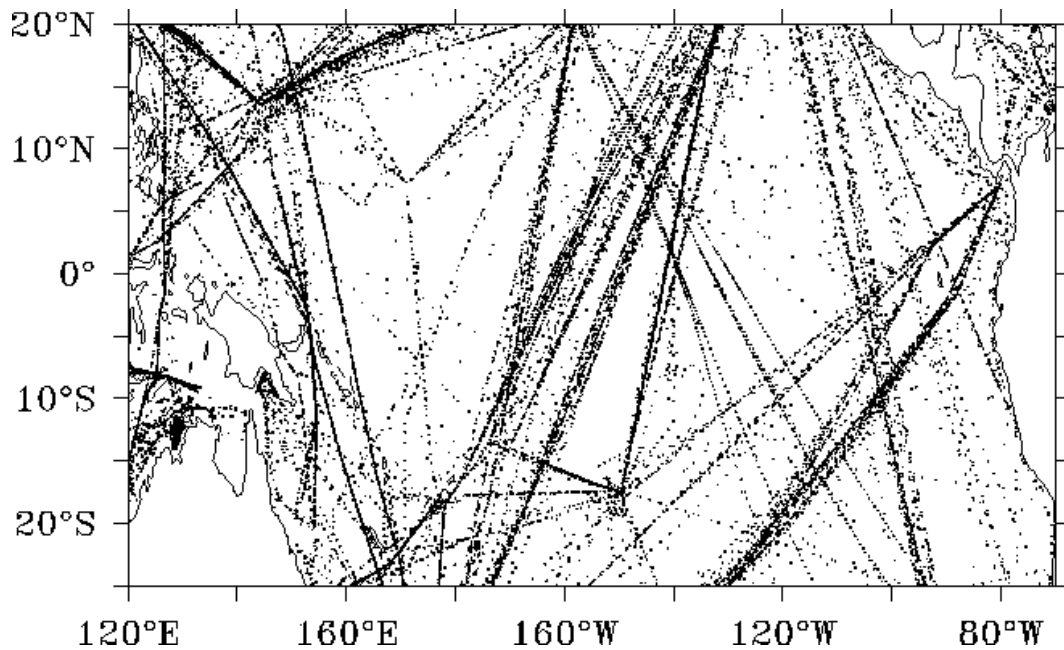


FIG. 2.9 – Position des profils XBT sur la période 1993-1998.

pectivement, de bateaux scientifiques et de bateaux VOS (Voluntary Observing Ship) représentent une deuxième source d'information importante. En plus des classiques campagnes océanographiques, des navires de commerce volontaires sont utilisés pour échantillonner la température entre la surface et environ 700 m de profondeur le long de leur ligne de navigation. La répartition des données est très inhomogène tant spatialement comme le montre la figure 2.9 donnant la position de l'ensemble des profils disponibles sur la période 1993-1998, que temporellement.

Les données ont été récupérées par le biais de site internet du projet CORIOLIS¹⁸. Nous disposons d'environ 60000 profils sur la période présentant une couverture spatio-temporelle très irrégulière. Certaines lignes de navigation comme Auckland-Tokyo, Auckland-San Fransisco ou encore Papeete-San Fransisco sont par contre relativement bien échantillonnées et pourront être utilisées pour la validation des différentes simulations.

¹⁸<http://www.coriolis.eu.org>

Deuxième partie

Les outils

Chapitre 3

Le modèle

Sommaire

3.1	Introduction	42
3.2	Présentation du modèle OPA	42
3.3	La physique résolue par le modèle	43
3.3.1	Les équations primitives	43
3.3.2	La discrétisation spatiale et temporelle	45
3.3.3	Paramétrisation des échelles sous-maïlles	45
	La paramétrisation sous-maïlle horizontale	46
	Le schéma de diffusion verticale TKE	47
3.3.4	Les conditions aux limites	48
3.3.5	Le calcul de la hauteur de la mer	50
3.4	La configuration utilisée	51
3.4.1	Le domaine d'étude	52
3.4.2	La résolution verticale	53
3.4.3	Les paramétrisations physiques	53
3.4.4	Les forçages atmosphériques	54
	Forçages "flux spécifié"	54
	Forçages "bulk"	55
3.4.5	Récapitulatif	57

3.1 Introduction

L'océanographie physique est une science relativement jeune, surtout quand on la compare à son pendant pour l'atmosphère, la météorologie dynamique. Elle a connu un développement très rapide depuis quelques années, notamment grâce aux observations par satellite de l'océan. La compréhension du système océan, ainsi que la prévision de son évolution à diverses échelles temporelles est aujourd'hui une thématique scientifique porteuse et dynamique dans la mesure où elle est en phase avec les préoccupations de la société. Le contexte actuel, avec la prise de conscience populaire du changement climatique et les interrogations sur l'évolution à court et moyen terme de notre climat est en grande partie à l'origine de l'effort consenti sur l'étude de la dynamique des océans. Avec une masse 300 fois supérieure à celle de l'atmosphère et une capacité de stockage de la chaleur 1200 fois plus grande que l'atmosphère, la très grande inertie des océans fait qu'ils jouent un rôle de régulateur du climat terrestre. Les océans stockent également 65 fois plus de carbone que l'atmosphère et sans le piégeage de carbone par les océans, le taux de dioxyde de carbone dans l'atmosphère serait beaucoup plus élevé. C'est pour cela que le couplage entre l'océan et l'atmosphère est fondamental pour la compréhension et l'étude du climat. Les mêmes équations d'évolution spatio-temporelle gouvernent la dynamique de ces deux fluides : les équations de Navier-Stokes. Elles sont basées sur le principe de la conservation de la quantité de mouvement et s'écrivent en repère tournant de la façon suivante :

$$\frac{D\vec{u}}{Dt} = -\frac{\vec{\nabla}P}{\rho} + \vec{g} - 2\vec{\Omega} \times \vec{u} + \vec{D} + \vec{F} \quad (3.1)$$

où t est le temps, \vec{u} le vecteur vitesse de la particule de fluide considérée, P sa pression, ρ sa densité, \vec{g} l'accélération de la pesanteur et $\vec{\Omega}$ le vecteur rotation de la Terre. D et F sont les termes de dissipation et de source de quantité de mouvement. Le terme $\frac{D}{Dt}$ représente la dérivée Lagrangienne (en suivant la particule). Elle est reliée à la dérivée Eulérienne (en un point fixe) $\frac{\partial}{\partial t}$ par la relation $\frac{D}{Dt} = \frac{\partial}{\partial t} + \vec{u} \cdot \vec{\nabla}$.

Ensuite, la différenciation entre les modèles d'atmosphère et les modèles d'océan repose sur le choix des variables composant le vecteur d'état et sur les approximations utilisées. Dans ce qui suit, nous allons nous focaliser sur les modèles océaniques. Le modèle utilisé au cours de ce travail de thèse est un modèle forcé et nous n'avons donc pas de modèle atmosphérique. En effet, l'atmosphère est considérée comme étant extérieure au système et son influence est prise en compte par l'intermédiaire de termes de forçage utilisés pour contraindre le modèle d'océan à l'interface air-mer.

3.2 Présentation du modèle OPA

Parmi les différents modèles océaniques possibles, le choix s'est porté sur le modèle aux équations primitives OPA8.2¹ (Océan PARallélisé; *Madec et al.*, 1998) qui a été développé au LOCEAN (Laboratoire d'Océanographie et du Climat - Expérimentation et Analyse Numérique) et dans la communauté océanographique européenne pour simuler la circulation océanique régionale et globale. Le code OPA est à la base du système NEMO²

¹<http://www.lodyc.jussieu.fr/opa/>

²<http://www.lodyc.jussieu.fr/NEMO/>

(Nucleus for European Modelling of the Ocean) développé dans le cadre du projet européen MERSEA³ (Marine Environmental and Security for the European Area). Il est et a été largement utilisé dans le cadre de projets français et européens (CLIPPER⁴, DRACKAR⁵, MERCATOR⁶ MERSEA³), et son code est en constante évolution. Cet OGCM a été mis en œuvre dans une grande variété d'études océanographiques. La liste ci-après est très loin d'être exhaustive mais présente quelques unes des études sur la dynamique du Pacifique Tropical réalisées avec le modèle OPA. *Grima et al.* (1999) ont étudié la sensibilité du modèle aux forçages de vents utilisés et a notamment montré que les vents satellites (ERS1&2) produisent une thermocline et des courants zonaux plus réalistes. *Vialard et al.* (2001) ont pour leur part investigué les mécanismes affectant la température de surface lors du très intense El Niño de 1997. L'impact des WWB dans le déclenchement de l'El Niño de 1997 a été étudié par *Lengaigne et al.* (2002). *Lengaigne et al.* (2003) ont également utilisé la configuration ORCA2 afin d'étudier la sensibilité de la dynamique du Pacifique Tropical à la paramétrisation du mélange horizontal. Plus récemment la validation d'une simulation décennale sur la période 1948-1999 de la dynamique du Pacifique Tropical a été réalisée par *Alory et al.* (2005). OPA a également été utilisé dans le Pacifique Tropical afin de réaliser des expériences d'assimilation de données. En particulier, par *Parent* (2000) qui a réalisé des expériences d'assimilation de l'anomalie de hauteur de la mer (SLA) mesurée par T/P+ERS sur la période 1994-1998 dans une configuration Pacifique Tropical du modèle OPA utilisant l'hypothèse de toit rigide (cf section 3.3.5). *Durand* (2003) a lui réalisé un travail plus prospectif sur l'assimilation de la salinité de surface (SSS) motivé par la perspective des missions européennes SMOS⁷ et américaine AQUARIUS⁸. Plus récemment, *Robert* (2004) a travaillé sur l'hybridation des systèmes d'assimilation SEEK et 4D-Var (cf chapitre 4) en utilisant la configuration Pacifique Tropical de OPA. Cette liste est là encore loin d'être exhaustive.

Dans le cadre de ce travail de thèse, nous avons besoin d'un modèle performant, fiable et relativement peu coûteux. J'ai opté pour la configuration globale ORCA2. En effet, cette configuration a été validée et présente des performances très satisfaisantes dans la région du Pacifique Tropical (*Lengaigne et al.*, 2003; *Alory et al.*, 2005). Cette configuration paraît en adéquation avec les objectifs de cette thèse. Dans la suite de ce chapitre, je donnerai un bref aperçu de la physique résolue par le modèle. Le lecteur désirant de plus amples informations sur les problématiques de la modélisation océanique et la paramétrisation des modèles d'océan pourra, par exemple, se référer au livre édité par *Chassignet et Verron* (1998)

3.3 La physique résolue par le modèle

3.3.1 Les équations primitives

Comme nous l'avons dit précédemment, le modèle OPA résout les équations fondamentales de la dynamique des fluides en milieu tournant. Ces équations, issues de la mécanique des fluides, sont :

³<http://www.mersea.eu.org>

⁴<http://www.ifremer.fr/lpo/clipper/>

⁵<http://www.ifremer.fr/lpo/drakkar/>

⁶<http://www.mercator-ocean.fr>

⁷<http://smos.cnes.fr/SMOS/Fr/index.htm>

⁸<http://aquarius.nasa.gov/>

- Les équations de mouvement ou équations de Navier-Stokes, issues du principe de conservation de la quantité de mouvement
- L'équation de continuité, issue du principe de conservation de la masse.
- L'équation d'évolution de la température, issue du principe de conservation de l'énergie.
- L'équation d'évolution de la salinité, issue de la loi de conservation des substances dissoutes.
- L'équation d'état de l'eau de mer qui lie température, salinité et densité.

Compte tenu des échelles géophysiques et du milieu étudié, un certain nombre d'approximations spécifiques peuvent être réalisées :

- L'approximation sphérique : les surfaces géopotentielle sont supposées être des sphères et le vecteur de l'accélération gravitationnelle \vec{g} est donc parallèle au rayon de la Terre.
- L'approximation d'eau peu profonde : la profondeur de l'océan est faible devant le rayon de la terre (3.8 km en moyenne devant 6400 km).
- L'approximation de Boussinesq : au vu des faibles variations de densité de l'eau de mer par rapport à sa valeur de référence $\rho_0 = 1020 \text{kg.m}^{-3}$, les variations de densité sont négligées, excepté dans les termes de flottabilité.
- L'hypothèse d'incompressibilité : la densité étant considérée comme quasi-constante, l'équation de conservation de la masse $\frac{DM}{Dt} = 0$ implique que la vitesse est non divergente $\nabla \cdot \vec{u} = 0$.
- L'approximation hydrostatique : dans la composante verticale de l'équation de la quantité de mouvement, l'accélération verticale et le terme de Coriolis sont négligés. Il en résulte un équilibre entre le gradient de pression vertical et les forces de flottabilité.
- L'hypothèse de fermeture turbulente : les flux turbulents correspondant aux processus d'échelle inférieure à celle de la discrétisation des variables (les effets "sous-mailles") peuvent être paramétrés en fonction des variables grandes échelles.

Dans le cadre des hypothèses énoncées ci-dessus, on peut reformuler les équations fondamentales de la dynamique des fluides comme suit. On notera que le modèle présente la particularité d'utiliser une condition limite de surface libre (*Roulet et Madec, 2000*), par opposition à l'approximation dite de "toit rigide" qui fut et est encore largement utilisée dans les OGCMs. Ce point fera l'objet d'une section dédiée (voir section 3.3.5). L'utilisation d'une surface libre nécessite, avec le modèle OPA, l'ajout du terme $gT_c \nabla_h \frac{\partial \eta}{\partial t}$ aux équations primitives. Ce terme peut être vu comme une force de capillarité et sert à résoudre un certain nombre de problèmes numériques liés spécifiquement aux mouvements de la surface libre (cf section 3.3.5). En pratique, l'effet de cette force additionnelle est de filtrer tous les mouvements ondulatoires de la surface (tels que les ondes de gravité) de période plus courte qu'un paramètre temporel T_c fixé a priori.

Au final, cet ensemble d'hypothèses et d'approximations nous conduit au système d'équations primitives discrétisables suivant :

$$\frac{\partial \vec{u}_h}{\partial t} = - \left\{ (\nabla \times \vec{u}) \times \vec{u} + \frac{1}{2} \nabla (\vec{u}^2) \right\}_h - f \vec{k} \times \vec{u}_h - \frac{1}{\rho_0} \nabla_h p + g T_c \nabla_h \frac{\partial \eta}{\partial t} + \vec{D}^u \quad (3.2)$$

$$\frac{\partial p}{\partial z} = -\rho g \quad (3.3)$$

$$\nabla \cdot \vec{u} = 0 \quad (3.4)$$

$$\frac{\partial T}{\partial t} = -\nabla \cdot (\vec{u} T) + D^T \quad (3.5)$$

$$\frac{\partial S}{\partial t} = -\nabla \cdot (\vec{u} S) + D^S \quad (3.6)$$

$$\rho = \rho(T, S, p) \quad (3.7)$$

avec $\vec{u} = \vec{u}_h + w\vec{k} = (u, v, w)$ le vecteur vitesse (en ms^{-1}), T la température potentielle (en $^\circ\text{C}$), S la salinité (en psu), ρ la masse volumique in-situ (en kgm^{-3}), p la pression (en bar), η le niveau de la surface libre (en m), z la coordonnée verticale ascendante, \vec{k} le vecteur unitaire ascendant, f le paramètre de Coriolis fonction de la latitude ϕ ($f = 2\Omega \sin \phi$). D^u , D^T , D^S , sont les paramètres des processus sous-maille sous la forme d'une diffusion pour la vitesse, la température et la salinité. Les termes D^u , D^T et D^S seront explicités plus loin.

3.3.2 La discrétisation spatiale et temporelle

L'ensemble de ces équations primitives est discrétisé spatialement selon une méthode de différences finies centrées du second ordre. La discrétisation verticale est en niveau z . Les variables sont discrétisées selon une grille de type Arakawa C (*Arakawa et Lamb, 1977*). Les mailles sont centrées sur les grandeurs scalaires (température, salinité, pression, ...) et les grandeurs vectorielles sont calculées au centre de chaque face de la maille (u , v , w) (cf figure 3.1).

Temporellement, un schéma "*Leapfrog*" (cf équation 3.8) est utilisé, stabilisé par l'utilisation d'un filtre d'Asselin (*Asselin, 1972*) permettant d'éviter la dissociation des modes pairs et impaires (cf équation 3.9).

$$\text{Leapfrog : } \quad u^{t+\Delta t} = u^{t-\Delta t} + 2\Delta t \text{ RHS}^t \quad (3.8)$$

$$u_f^t = u^t + \gamma \left(u_f^{t-\Delta t} - 2u^t + u^{t+\Delta t} \right) \quad (3.9)$$

(l'indice f dénote les valeurs filtrées)

3.3.3 Paramétrisation des échelles sous-mailles

La discrétisation spatio-temporelle utilisée afin de pouvoir résoudre les équations primitives du modèle représente une limite forte interdisant la représentation explicite de l'ensemble des processus de taille inférieure à la maille. Or, la circulation océanique résulte

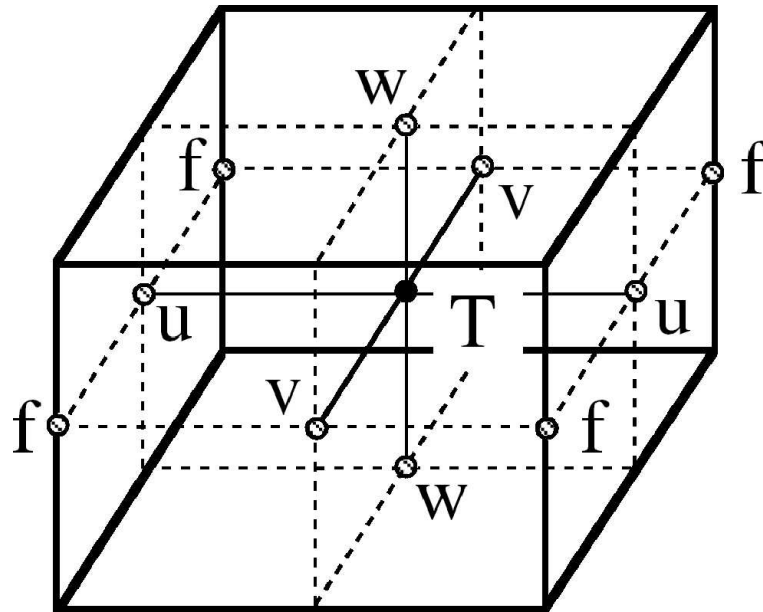


FIG. 3.1 – Disposition des variables sur la grille C . T indique le point où les grandeurs scalaires sont calculées. (u, v, w) les points où les grandeurs vectorielles sont calculées.

du couplage non-linéaire entre un ensemble de processus dont les échelles caractéristiques sont comprises entre des longueurs bien inférieures à la taille de la maille (turbulence de sub-méso-échelle, turbulence dans la couche limite de surface, viscosité, etc...) et des longueurs de l'ordre du millier de kilomètres (pour la circulation générale). L'impact des effets "sous-maille" (provenant des termes d'advection dans les équations de Navier-Stokes) sur les processus explicitement résolus est important. Il est donc nécessaire de les paramétrer. En suivant l'hypothèse de fermeture turbulente, il est possible d'exprimer ces flux turbulents en fonction des variables grandes échelles. Un formalisme de type "diffusion" est adopté (les termes D^x dans les équation primitives de la section 3.3.1). On suppose que les flux turbulents dépendent linéairement des gradients de grandes échelles. Compte tenu de la forte anisotropie des mouvements océaniques (due notamment au contrôle exercé par la gravité et la rotation), la paramétrisation des flux horizontaux et verticaux est traitée séparément.

$$D^x = D^{hx} + D^{vx}$$

La paramétrisation sous-maille horizontale

Dans notre configuration, les termes de diffusion sont exprimés différemment pour les traceurs et pour la vitesse.

Pour les traceurs, la paramétrisation proposée par *Gent et McWilliams* (1990) est utilisée. Cette paramétrisation a pour effet de réduire l'énergie potentielle moyenne de l'océan. Le terme de diffusion s'exprime alors comme la somme d'un terme diffusif (sous la forme d'un laplacien à travers les isopycnes) et d'un terme advectif.

$$D^{hT} = \nabla \cdot (A^{hT} \mathfrak{R} \nabla T) + \nabla (U^* T) \quad (3.10)$$

A^{hT} étant le coefficient de diffusion, $U^* = (u^*, v^*, w^*)$ la partie non divergente du transport tourbillonnaire et \mathfrak{R} un opérateur (3×3) prenant en compte les pentes entre les surfaces isopycnales et les surfaces géopotentiellelles.

Pour la vitesse, en revanche, le terme de diffusion agit au travers des surfaces géopotentiellelles et sa formulation se fait en fonction de χ , la divergence du champ de vitesse horizontale et ζ , la vorticité relative.

$$D^{hU} = \nabla_h(A^{hm}\chi) - \nabla_h(A^{hm}\zeta k) \quad (3.11)$$

Le schéma de diffusion verticale TKE

Les principales sources de turbulence verticale (déferlement d'ondes internes, mélange vertical, échange de quantité de mouvement et de chaleur à l'interface air/mer, etc...) ont des longueurs caractéristiques bien inférieures à la résolution du modèle. Ces flux turbulents sont particulièrement vigoureux dans la couche limite de surface. Ces processus non explicitement résolus sont capitaux pour la dynamique de la couche mélangée ainsi que pour la ventilation de l'océan profond et nécessitent une paramétrisation adaptée. Les flux paramétrés dépendent linéairement des gradients verticaux locaux des variables à grandes échelles, par analogie avec la diffusion moléculaire. La paramétrisation des flux turbulents s'exprime pour la vitesse, la température et la salinité par

$$\vec{D}^{vu} = \frac{\partial}{\partial z} \left(A^{vm} \frac{\partial \vec{u}_h}{\partial z} \right) \quad (3.12)$$

$$D^{vT} = \frac{\partial}{\partial z} \left(A^{vT} \frac{\partial T}{\partial z} \right) \quad (3.13)$$

$$D^{vS} = \frac{\partial}{\partial z} \left(A^{vT} \frac{\partial S}{\partial z} \right) \quad (3.14)$$

Le coefficient de viscosité turbulente A^{vm} et le coefficient de diffusion turbulente A^{vT} sont calculés en utilisant le modèle de fermeture turbulente TKE (*Blanke et Delecluse, 1993*) d'ordre 1.5, basé sur l'équation pronostique de l'énergie cinétique turbulente \bar{e} . L'évolution temporelle de \bar{e} (cf équation 3.16) est le résultat du bilan entre la production de \bar{e} par le cisaillement vertical, sa destruction par la stratification, sa diffusion verticale et sa dissipation.

$$\bar{e} = \frac{1}{2}(u'^2 + v'^2 + w'^2) \quad (3.15)$$

$$\frac{\partial \bar{e}}{\partial t} = \frac{\partial}{\partial z} \left(K_e \frac{\partial \bar{e}}{\partial z} \right) + K_m \left(\frac{\partial U}{\partial z} \right)^2 - K_\rho N^2 - C_\epsilon \frac{\bar{e}^{\frac{3}{2}}}{l_\epsilon} \quad (3.16)$$

Où u' , v' , w' sont les perturbations respectives par rapport aux moyennes temporelles de u , v et w , K_e , K_m et K_ρ les coefficients de diffusion respectivement de l'énergie, de la quantité de mouvement et des traceurs et N la fréquence de Brünt-Vaisälä. Les coefficients de viscosité turbulente A^{vm} et de diffusion turbulente A^{vT} sont estimés comme le produit d'une vitesse turbulente et d'une longueur caractéristique. Afin d'éliminer les instabilités numériques associées à une diffusion verticale trop faible, une valeur seuil est imposée pour A^{vm} et A^{vT} .

$$K_m = \max(\nu_b, C_\kappa l_\kappa \sqrt{\bar{e}}) = A^{vm} \quad (3.17)$$

$$K_\rho = \max(\kappa_b, \frac{K_m}{P_{rt}}) = A^{vT} \quad (3.18)$$

$$K_e = \alpha_e K_m \quad (3.19)$$

$$l_\kappa = \min(\frac{\sqrt{2\bar{\epsilon}}}{N}, \text{abs}(z)) \quad (3.20)$$

$$l_\epsilon = l_\kappa \quad (3.21)$$

P_{rt} étant le nombre de Prandtl.

Dans les écoulements océaniques, des instabilités hydrostatiques ($N < 0$) peuvent apparaître (convection hivernale par exemple). Ces instabilités vont engendrer une convection verticale qui va rétablir l'équilibre hydrostatique. Dans un modèle hydrostatique, ces instabilités ne peuvent se dissiper. Bien que le schéma TKE prenne en compte la stabilité de la stratification dans le calcul de la viscosité et de la diffusion turbulente par l'intermédiaire de la fréquence de Brünt-Vaisälä (quand N est négatif, le terme de destruction de $\bar{\epsilon}$ par la stratification devient un terme source dans l'équation 3.16), le schéma TKE se révèle inefficace à fournir des coefficients suffisamment élevés pour dissiper ce type d'instabilité et homogénéiser la colonne d'eau. Pour paramétrer la convection, les coefficients de viscosité turbulente A^{vm} et de diffusion turbulente A^{vT} sont spécifiés à une valeur très grande ($A^{vm} = A^{vT} = 1 \text{ m}^2\text{s}^{-1}$) lorsque la fréquence de Brünt-Vaisälä est négative, c'est à dire dans les régions où la stratification est instable.

3.3.4 Les conditions aux limites

La configuration utilisée, qui sera décrite dans la partie suivante, est une configuration globale. Elle permet, à ce titre de s'affranchir du problème des frontières ouvertes. Les conditions aux limites du modèle concernent donc uniquement les côtes et la topographie fond (bathymétrie), d'une part, et les échanges avec l'atmosphère à l'interface air/mer, d'autre part.

Echanges avec le continent :

- Le long des côtes et au fond, des conditions de type Neumann (pas de flux à travers les frontières) sont appliquées pour les flux de chaleur et de sel. La vitesse normale à la frontière est nulle : $w = -\vec{u}_h \cdot \nabla_h H = 0$.
- La condition de non-glissement ("no slip") est utilisée le long des côtes pour les vitesses horizontales.

$$u = v = 0 \quad (3.22)$$

- En ce qui concerne le frottement de fond, la paramétrisation utilisée fait l'hypothèse que le frottement est proportionnel à la vitesse horizontale dans la couche de fond.

$$A^{vm} \left(\frac{\partial \vec{u}_h}{\partial z} \right)_{z=H} = \vec{F}_h = r \vec{u}_h^b \quad (3.23)$$

\vec{u}_h^b est la vitesse horizontale dans la couche de fond
 r un coefficient de frottement en ms^{-1}

- A l'embouchure des principales rivières, les apports en eau douce sont pris en compte (ces apports sont en fait pris en compte par le biais du flux d'eau douce à l'interface air/océan).

Echanges avec l'atmosphère :

- Prise en compte du flux de chaleur : A la surface ($z=0$), le flux de chaleur à l'interface océan-atmosphère Q_{NET} est utilisé comme condition limite pour le flux de chaleur turbulent dans l'équation pronostique de la température (cf équation 3.5) :

$$\vec{D}_{z=0}^{vT} = \left(A^{vT} \frac{\partial \vec{T}}{\partial z} \right)_{z=0} = \frac{Q_{NET}}{\rho_0 C_{pw}} \quad (3.24)$$

C_{pw} est la capacité calorifique de l'eau de mer

Il existe deux types de forçage : les forçages "à flux spécifié" et les forçages de type "bulk". Dans les deux cas, le but est d'estimer le flux de chaleur à l'interface océan-atmosphère Q_{NET} . En flux spécifié (*Barnier, 1998*), le flux de chaleur (préalablement calculé à partir d'observations et de modèles d'atmosphère) est directement spécifié. Dans la réalité, le flux de chaleur dépend de la différence de température entre la surface de l'océan et l'atmosphère. C'est pour mieux représenter ce couplage entre l'océan et l'atmosphère que les forçages bulk ont été introduits (*Large et al., 1997*). Ces derniers utilisent les conditions atmosphériques (vent, humidité, température de l'air à la surface et couverture nuageuse) associées à la température de surface de l'océan simulée pour calculer de manière interactive le flux de chaleur à l'interface océan-atmosphère selon des formules semi-empiriques (Bulk formulae). Le modèle OPA offre la possibilité de choisir entre ces deux types de forçage. Ce point sera discuté plus en détail dans la partie suivante.

- Prise en compte du flux de quantité de mouvement : la condition limite de surface sur les vitesses horizontales est donnée par le forçage en vent. A la surface ($z=0$), le frottement exercé par le vent est utilisé comme condition limite dans l'équation de conservation de la quantité de mouvement (cf équation 3.2) :

$$\vec{D}_{z=0}^{vu} = \left(A^{vm} \frac{\partial \vec{u}_h}{\partial z} \right)_{z=0} = \frac{\vec{\tau}}{\rho_0} \quad (3.25)$$

$\vec{\tau}$ est le vecteur de tension de vent à la surface de l'océan

Le forçage par le vent nécessite donc la connaissance du vecteur tension de vent à la surface de l'océan. A titre d'exemple, ce genre d'information peut être accessible par des formulations empiriques du type :

$$\vec{\tau} = \rho_a C_D |U_{10} - U_{sfc}| \vec{u}_{10},$$

U_{10} est la vitesse du vent à une altitude de 10 mètres

U_{sfc} est la vitesse du courant de surface océanique

C_D est un coefficient d'échange turbulent

ρ_a est la densité de l'air

- Prise en compte du flux d'eau douce : l'échange d'eau douce en surface modifie la salinité. Le flux d'eau douce par évaporation (E), précipitation (P), apport des rivières et fonte de la glace (R) intervient donc dans l'équation de conservation du sel (cf équation 3.6) sous la forme d'un terme de concentration-dilution εS ($\varepsilon = E - P - R$) en lieu et place d'un mélange diffusif vertical.

$$\vec{D}_{z=0}^{vS} = \left(A^{vT} \frac{\partial \vec{S}}{\partial z} \right)_{z=0} = \varepsilon S_{z=0} \quad (3.26)$$

3.3.5 Le calcul de la hauteur de la mer

Le modèle présente la particularité d'utiliser une condition limite de surface libre introduite par *Roulet et Madec* (2000). Le dénivellement de la surface libre η se déduit de l'équation suivante :

$$\frac{\partial \eta}{\partial t} = \vec{u}_{z=\eta} \cdot \vec{n} + P + R - E \quad (3.27)$$

avec η la dénivellement de la surface libre relativement à la surface libre moyenne ; \vec{n} un vecteur unitaire perpendiculaire à la surface ($\vec{n} = (\frac{\partial \eta}{\partial x}, \frac{\partial \eta}{\partial y}, 1)$) ; \vec{u} la vitesse et E , P et R les trois composantes du flux d'eau douce.

La surface libre η est animée par des ondes de gravité externes (barotropes). La célérité de ces ondes ($c = \sqrt{gH}$) pose un problème numérique évident dans les OGCMs, ces derniers utilisant un grand pas de temps (typiquement $O(1\text{heure})$). Comme nous l'avons déjà indiqué, un processus ne peut exister dans un modèle que s'il est résolu spatialement. Il n'est de plus numériquement stable que s'il est résolu temporellement (critère CFL⁹). La résolution explicite des ondes longues de gravité animant la surface libre imposerait donc un pas de temps extrêmement faible, ce qui engendrerait des coûts de calcul prohibitifs.

Toute méthode d'introduction d'une surface libre dans un OGCM utilisant un grand pas de temps est basée sur un filtrage temporel (*Roulet et Madec*, 2000) afin d'atténuer toutes les ondes de gravité externes non résolues et ainsi préserver le critère CFL. Dans le modèle OPA, le choix d'introduire le filtrage temporel directement dans les équations de Naviers Stokes a été fait. Comme nous l'avons déjà indiqué dans l'équation 3.2, cela se traduit en pratique par l'introduction d'un terme supplémentaire $-gT_c \frac{\partial \eta}{\partial t}$. Ce terme peut être vu comme une force de capillarité qui agit comme un filtre passe-bas d'ordre 1, avec une période de coupure T_c . En suivant *Roulet et Madec* (2000), la constante de filtrage T_c déterminée par analyse de la stabilité est fixée à la valeur optimale $T_c = 2\Delta T$.

Le choix d'utiliser une surface libre a été motivé par diverses raisons. Contrairement à ce que l'on pourrait penser au premier abord, il n'y a en effet aucune raison fondamentale de ne pas utiliser l'approximation toit rigide pour ce travail de thèse. Cette hypothèse est finalement une très bonne hypothèse pour la circulation générale océanique de grande échelle et l'assimilation de données altimétriques s'opère très bien en toit rigide (*Testut et al.*, 2003; *Pinardi et al.*, 1995). Si l'on a préféré travailler en surface libre plutôt qu'en toit rigide, c'est tout d'abord pour rester en phase avec la communauté modélisatrice. Dans la nouvelle version du code OPA (la version 9.0), l'option toit rigide est toujours disponible,

⁹Le critère de Courant Friedrichs Levy impose une vitesse "numérique" supérieure à la célérité des processus que l'on veut représenter ($\frac{\Delta X}{\Delta T} > c$). Il fournit donc une limite supérieure sur le pas de temps étant donné le pas d'espace et la célérité : $\Delta T < \frac{\Delta X}{c}$

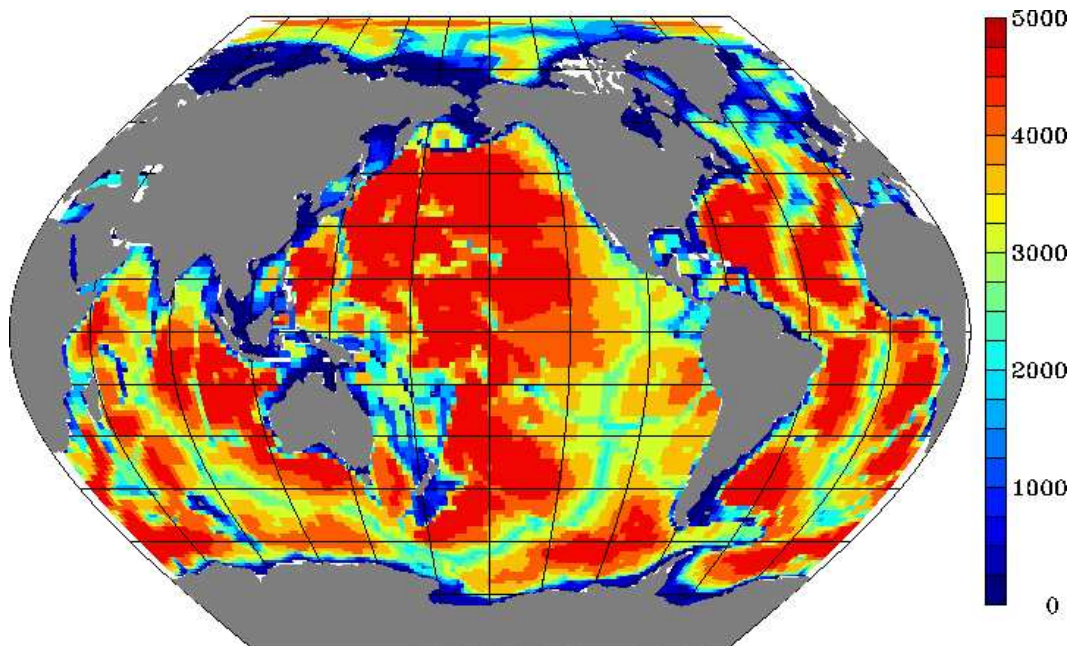


FIG. 3.2 – Le domaine du modèle et la topographie de fond de ORCA2.

mais n'est plus utilisée. L'ensemble des configurations développées dans le cadre du projet de modélisation DRAKKAR (*Barnier et al., 2006*) utilise une surface libre. D'autre part, l'utilisation d'une surface libre simplifie singulièrement le processus d'assimilation. En effet, l'approximation de toit rigide impose que le champ de vitesse barotrope du modèle soit non divergent. Or il n'y a aucune raison intrinsèque pour que le champ de vitesse analysé satisfasse cette contrainte. Cela impose donc de "corriger" le champ de vitesse analysé en remplaçant la partie barotrope divergente par une partie barotrope non divergente (*Testut, 2000*). Avec la condition limite de surface libre, le modèle peut très bien redémarrer avec un champ de courant divergent et l'assimilation se fait de manière beaucoup plus naturelle. C'est la force de tension superficielle et artificielle rajoutée dans l'équation de quantité de mouvement qui se charge de dissiper, en quelques pas de temps, les ondes de gravité externes générées lors du redémarrage du modèle par la partie divergente du champ de vitesse analysé.

3.4 La configuration utilisée

La configuration utilisée au cours de ce travail de thèse est la configuration globale basse résolution ORCA2 (Figure 3.2). Le choix d'utiliser une configuration globale alors que notre région d'intérêt se limite au seul Pacifique Tropical peut paraître paradoxal. En fait les considérations qui nous ont poussé à faire ce choix sont multiples. D'une part, sur un plan technique, cette configuration était la configuration de développement du modèle OPA. Elle est donc bien optimisée et est relativement peu coûteuse en temps de calcul. D'autre part, l'utilisation d'une configuration globale permet de s'affranchir du problème des frontières ouvertes qui n'est jamais vraiment trivial. De plus, l'utilisation de la confi-

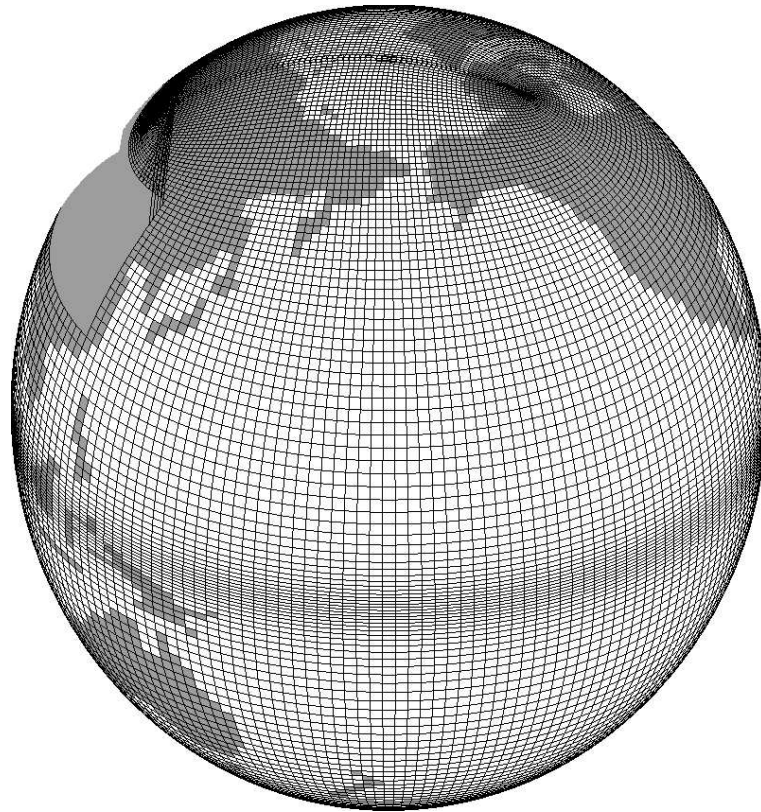


FIG. 3.3 – Maillage et masque des points T en surface dans le Pacifique.

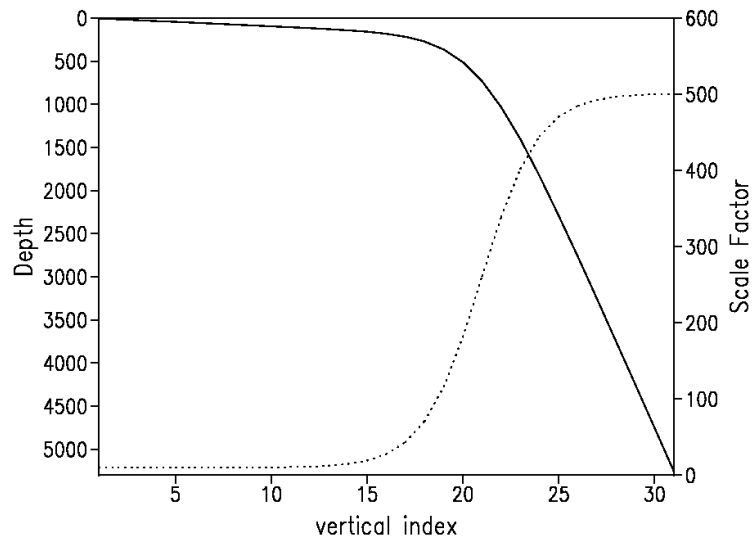
guration ORCA2 permet une synergie intéressante avec le projet européen MERSEA¹⁰ (Skachko et al., 2007), ainsi qu’avec le projet d’océanographie opérationnelle MERCATOR¹¹. La configuration ORCA2 (nommé PSY2G dans l’arborescence des configurations MERCATOR) a en effet été retenue par MERCATOR pour servir de prototype lors du développement d’un premier système opérationnel global d’analyse océanique. C’est aujourd’hui le cas. Depuis le mois d’Octobre 2005, le système PSY3V1 basé sur la configuration globale ORCA025 au $1/4^\circ$ effectue chaque semaine des analyses opérationnelles de l’océan mondial en assimilant l’ensemble des observations altimétriques disponibles. En parallèle, le système PSY2G est également maintenu opérationnellement et fournit des analyses hebdomadaires qui sont notamment utilisées pour la prévision saisonnière par Météo France.

3.4.1 Le domaine d’étude

Cette configuration est basée sur une grille de type Mercator (i.e. les facteurs de forme méridien varient en cosinus de la latitude). Le maillage horizontal est déduit d’une méthode semi-analytique (Madec et Imbard, 1996) et présente la particularité, dans l’hémisphère Nord, d’utiliser 2 pôles (voir Figure 3.3) afin de contourner le problème de la singularité au pôle nord. L’utilisation d’un double pôle permet également de conserver, si on le souhaite,

¹⁰<http://www.mersea.eu.org>

¹¹<http://www.mercator-ocean.fr>

FIG. 3.4 – *Caractéristiques du maillage vertical.*

un facteur d'anisotropie proche de 1 sur la quasi totalité du domaine.

La résolution zonale est 2° . La résolution méridienne est maximale à l'équateur (0.5°) et augmente jusqu'à 2° à partir de $10^\circ N$ et $10^\circ S$ (cf figure 3.3). Ce raffinement de maillage méridional à l'équateur permet d'améliorer la dynamique équatoriale. Du fait du raffinement de maillage, la grille est anisotrope dans les régions tropicales. La distance zonale entre deux points de grille est de 222 km à l'équateur alors que la distance méridienne est de seulement 55.5 km. Compte tenu de la dynamique spécifique essentiellement zonale des régions équatoriales, cette anisotropie du maillage ne pose, cependant, pas de problème.

3.4.2 La résolution verticale

Sur la verticale, le domaine s'étend de la surface à la profondeur de 5000 m. La discrétisation verticale découpe la colonne d'eau en 31 couches. Le maillage vertical est défini à partir d'une fonction analytique de la profondeur z (*Madec et Imbard, 1996*). La résolution est de 10 m en surface et diminue jusqu'à 500 m au fond (cf figure 3.4). Dans les 600 premiers mètres (épaisseur de l'océan suivant laquelle la dynamique du Pacifique Tropical s'exerce principalement) on a 21 des 31 niveaux du modèle (cf figure 3.4).

3.4.3 Les paramétrisations physiques

Je donne ici un rapide tour d'horizon des paramètres physiques utilisés dans notre configuration. Ces paramètres sont les paramètres classiquement utilisés avec la configuration ORCA2 et ont donc à ce titre fait l'objet de validations préalables, notamment au LOCEAN.

- Le pas de temps : il est fixé à 1h36mn de manière à satisfaire la condition numérique de stabilité, ce qui équivaut à 15 itérations par jour.
- La diffusion horizontale pour les vitesses : le coefficient de viscosité turbulente dépend de la position géographique. Il vaut $40000 \text{ m}^2\text{s}^{-1}$ partout et est réduit à $2000 \text{ m}^2\text{s}^{-1}$ dans les régions équatoriales excepté près des bords ouest.

- La diffusion isopycnale pour les traceurs : la diffusion pour les traceurs se fait au travers des surfaces isopycnales avec un coefficient de diffusion de $2000 \text{ m}^2\text{s}^{-1}$.
- Paramétrisation de *Gent et McWilliams* (1990) : Un terme d'advection tourbillonnaire avec un coefficient fonction du taux de croissance des instabilités baroclines (variant en général entre $15 \text{ m}^2\text{s}^{-1}$ et $3000 \text{ m}^2\text{s}^{-1}$) est ajouté au terme de diffusion pour les traceurs.
- Conditions aux frontières latérales : pas de flux de chaleur et de sel et condition de non-glissement appliquée le long des côtes.
- Conditions au fond : pas de flux de chaleur et de sel au fond de l'océan. La paramétrisation de la couche de fond proposée par *Beckmann et Döscher* (1997) est appliquée le long des talus continentaux. Un coefficient de friction linéaire est utilisé pour les vitesses.
- Convection : la convection est paramétrée en fixant le coefficient de viscosité verticale à $1 \text{ m}^2\text{s}^{-1}$ en cas d'instabilité hydrostatique ($N < 0$).
- Forçages : l'océan reçoit des flux de chaleur, d'eau douce et de quantité de mouvement de l'atmosphère et/ou de la glace de mer. Le flux solaire pénètre les premiers mètres de l'océan de surface.

3.4.4 Les forçages atmosphériques

Au cours de cette thèse deux types de forçages ont été utilisés, un jeu de forçage à flux spécifié et un jeu de forçage bulk. Notre choix s'est porté sur ces deux jeux de forçages car ils sont représentatifs de l'état de l'art en la matière ; ils sont tous deux également utilisés dans des études de modélisation océanique (*Vialard et al.*, 2001; *de Boyer Montégut et al.*, 2005).

Forçages "flux spécifié"

Pour nos expériences en flux spécifié, le choix s'est porté sur la réanalyse du Centre Européen de Prévision Météorologique à Moyen Terme (CEPMMT¹²) couvrant la période 1957-2002 (ERA40). Les flux de chaleur Q_{NET} et d'eau douce ε sont issus de la réanalyse CEPMMT à une fréquence journalière.

Dans le cadre de l'approximation des flux spécifiés, le flux de chaleur Q_{tot} utilisé pour forcer le modèle est composé d'un flux de chaleur net Q_{NET} qui est spécifié par le champ de forçage, et d'un terme de correction de flux Q_{corr} qui lui est diagnostiqué de manière interactive.

$$Q_{tot} = Q_{NET} + Q_{corr}$$

Le terme de correction de flux Q_{corr} a une fonction essentielle pour un modèle à flux spécifié. Il permet de restituer un degré de liberté au système et assure la cohérence entre les forçages appliqués et la dynamique océanique simulée (*Barnier et al.*, 1995). En effet, sans ce terme, le transport méridien de chaleur du modèle serait contraint de s'ajuster rigoureusement au transport méridien de chaleur imposé par le biais des forçages. La dynamique propre, à grande échelle, du modèle d'océan aurait alors du mal à se développer

¹²<http://www.ecmwf.int/>

librement et serait très fortement contrainte par le transport de chaleur méridien imposé par les forçages. De plus, les flux utilisés pour forcer le modèle ont été calculés en utilisant une température de surface (SST) climatologique ou observée. Il est plausible que la température de surface prédite par le modèle OPA diffère par endroit de la SST ayant servi aux calculs des flux spécifiés ce qui débouche sur une incohérence entre le système dynamique et les forçages appliqués. Le fort couplage entre le flux de chaleur à la surface de l'océan et la SST est simulé par l'ajout du flux Q_{corr} qui s'écrit comme un terme de relaxation vers une SST observée. Ceci permet de résoudre, ou du moins de contrôler, les problèmes décrits ci-dessus. La SST climatologique mensuelle de Levitus (*Levitus, 1998*) est utilisée pour calculer Q_{corr} . Ce terme de correction de flux peut finalement être vu comme une rétroaction de l'océan sur le flux de chaleur. Il s'exprime sous la forme :

$$Q_{corr} = \left(\frac{\partial Q}{\partial T} \right)_{T_S^{clim}} \cdot (T_S^{clim} - T_S) \quad (3.28)$$

Dans notre étude, la valeur du coefficient $\left(\frac{\partial Q}{\partial T} \right)_{T_S^{clim}}$ est fixée à $-40W.m^{-2}$ sur l'ensemble du domaine.

En plus du rappel en SST, la salinité de surface (SSS) du modèle est relaxée vers la salinité climatologique mensuelle de Levitus. Contrairement au terme de rappel en SST, le rappel en SSS n'a aucune signification physique. Compte tenu des fortes incertitudes sur le flux d'eau douce ε , on utilise ce terme pour prévenir les éventuelles dérives de la SSS du modèle.

En ce qui concerne les forçages du vent, le vecteur de tension du vent à la surface de l'océan $\vec{\tau}$ est calculé à partir des observations des diffusomètres laser embarqués sur les satellites ERS1 et ERS2 complétées par les observations in-situ aux points de mouillage TAO (*Menkes et al., 1998*). Les vents satellites sont en effet de meilleure qualité en terme d'intensité que ceux du CEPMMT (qui sont pour la plupart sous estimés) et sont connus pour produire une thermocline et des courants zonaux plus réalistes, en particulier avec la modèle OPA (*Grima et al., 1999*).

Forçages "bulk"

Une mesure directe des échanges de chaleur à l'interface océan/atmosphère est impossible à ce jour. Les flux à travers la surface de l'océan doivent donc être estimés à partir des variables atmosphériques à la hauteur anémométrique (normalement 10 m au dessus de la surface de l'océan) en utilisant les formules aérodynamiques semi-empiriques bulk. Les variables atmosphériques (vent, humidité, température de l'air à la surface et couverture nuageuse) du National Center for Environmental Prediction (NCEP¹³) sont utilisées pour diagnostiquer les flux de chaleur Q_{NET} et d'eau douce ε à l'interface océan/atmosphère de manière interactive. Les précipitations moyennes mensuelles CMAP (CPC Merged Analysis of Precipitation) proposées par *Xie et Arkin (1996)* sont utilisées en lieu et place des précipitations du modèle NCEP qui présentent des erreurs importantes comparées aux observations (notamment dans la région de la SPCZ).

Les flux de chaleur Q_{NET} et d'eau douce ε à l'interface océan/atmosphère s'écrivent comme suit :

$$Q_{NET} = Q_{SEN} + Q_{LAT} + Q_{LW} + Q_{SW} \quad (3.29)$$

¹³<http://www.ncep.noaa.gov>

$$\varepsilon = E - P - R \quad (3.30)$$

Q_{SEN} étant le flux de chaleur sensible, Q_{LAT} le flux de chaleur latente, Q_{LW} le flux par rayonnement infra-rouge et Q_{SW} le flux solaire radiatif. E , P et R sont les 3 termes du bilan d'eau douce, respectivement l'évaporation, les précipitations et les apports des rivières.

Un bref aperçu de la manière d'estimer ces différents flux à la surface de l'océan est donné ci-après. Pour une description détaillée des formules bulk utilisé par OPA, on se référera à la description du modèle CLIO (Coupled Large-scale Ice Ocean) par *Goosse et al.* (1999).

Flux turbulent de chaleur sensible et latente

Les flux turbulents de chaleur sensible et latente sont calculés à partir des équations de transfert océan-atmosphère classiques (*Large et Pond*, 1982) :

$$Q_{SEN} = \rho_a c_p^a C_H W_{10} (T_a - T_w) \quad (3.31)$$

$$Q_{LAT} = \rho_a L_e C_E W_{10} (q_a - q_w) \quad (3.32)$$

où ρ_a est la densité de l'air, W_{10} , T_a et q_a sont la vitesse, la température et l'humidité spécifique de l'air, c_p^a et L_e sont la capacité calorifique et la chaleur latente de vaporisation de l'air et T_w et q_w sont la température de surface et l'humidité spécifique (qui est supposée à saturation) de l'océan.

Flux par rayonnement infra-rouge

Le flux par rayonnement infra-rouge est paramétré en suivant *Berliand et Berliand* (1952) :

$$Q_{LW} = \epsilon \sigma_{sb} T_a^4 (0.39 - 0.05 \sqrt{e_a}) (1 - \chi c_{ld}^2) + 4 \epsilon \sigma_{sb} T_a^3 (T_w - T_a) \quad (3.33)$$

où e_a (en mb) est la pression de vapeur déduite de q_a , ϵ l'émissivité de la surface, σ_{sb} la constante de Stephan-Boltzmann, c_{ld} la couverture nuageuse et $(1 - \chi c_{ld}^2)$ un facteur de correction pour prendre en compte les effets des nuages.

Rayonnement solaire incident

La formulation de rayonnement solaire incident se fait suivant la formule proposée par *Zillmann* (1972) :

$$Q_{SW} = (1 - \alpha)(1 - 0.62 c_{ld} + 0.0019 \beta) Q_{CLEAR} \quad (3.34)$$

où α est l'albédo de l'océan, β l'altitude du soleil à midi et Q_{CLEAR} le rayonnement solaire à la surface de l'océan par temps clair.

Evaporation

L'évaporation se déduit facilement du flux turbulent de chaleur latente :

$$E = \frac{Q_{LAT}}{L_e} \quad (3.35)$$

Concernant le flux de quantité de mouvement, le vecteur de tension du vent à la surface de l'océan $\vec{\tau}$ n'est pas calculé de manière interactive à partir de la vitesse du vent, des courants de surface et des formules aérodynamiques empiriques. La tension du vent $\vec{\tau}$ est donc directement spécifiée au modèle. Comme pour la configuration utilisant les flux spécifiés, les tensions de vent proposées par *Menkes et al.* (1998) sont utilisées.

Du fait que le couplage entre le flux de chaleur et la température de surface de l’océan se fait de manière interactive, il n’est plus nécessaire de rajouter un terme de rétro-action sur la SST comme précédemment. En revanche, on conserve le rappel sur la SSS climatologique.

3.4.5 Récapitulatif

Les informations concernant le modèle et la configuration utilisée sont synthétisées dans le tableau 3.1.

Le domaine		
Dimensions de la grille	181 × 149 × 31 (= 836039 mailles)	
Grille	2° long., 0.5° lat. à l’équateur → 2° lat.	
La physique		
ΔT	5760s.	
Diffusion horizontale (traceur)	paramétrisation de <i>Gent et McWilliams</i> (1990) $A^{hT} = 2000m^2s^{-1}$	
Diffusion horizontale (vitesse)	fonction de la position géographique : $A^{hm} = 2000m^2s^{-1}$ dans les tropiques et $40000m^2s^{-1}$ ailleurs	
Diffusion verticale (traceur et vitesse)	A^{vm} et A^{vT} estimés en utilisant le schéma TKE	
Les forçages		
Flux de chaleur	flux CEPMMT	NCEP bulk
Flux de QDM	vent ERS/TAO	vent ERS/TAO
Précipitations	/	<i>Xie et Arkin</i> (1996)
Rappel	SST et SSS	uniquement SSS

TAB. 3.1 – Principales caractéristiques de la configuration utilisée.

Chapitre 4

L'assimilation de données

Sommaire

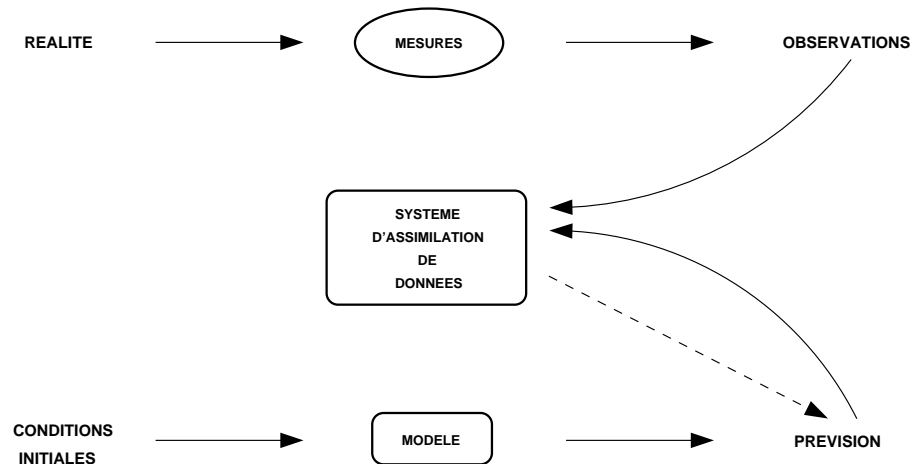
4.1	Introduction	60
4.2	L'assimilation de données en général	61
4.2.1	Généralités	61
4.2.2	Les applications	62
4.2.3	Les différentes méthodes d'assimilation de données	64
	Le nudging	64
	Les méthodes variationnelles	64
	Les méthodes statistiques	66
4.3	La théorie du filtrage de Kalman	66
4.3.1	Positionnement du problème	66
4.3.2	Erreurs et densités de probabilité	67
4.3.3	Analyse optimale	69
4.3.4	La séquence d'assimilation	70
4.4	Le SEEK : de la théorie vers les applications réelles	71
4.4.1	Le principe de la réduction d'ordre	72
4.4.2	L'initialisation de l'espace réduit	73
4.4.3	L'évolution de la covariance des erreurs d'analyse et de prévision	74
4.4.4	L'algorithme SEEK	75
4.4.5	Les différentes variantes de l'algorithme SEEK	76
4.4.6	Le SEEK local	76
4.5	Implémentation du SEEK avec OPA surface libre	77
4.5.1	Le logiciel SESAM	77
4.5.2	La réinitialisation du modèle : le schéma IAU	78
4.5.3	Domaine d'assimilation	81
4.5.4	Le vecteur d'état	82
4.5.5	La spécification de l'espace d'erreur en rang réduit	83
4.6	Conclusion	84

4.1 Introduction

Comme nous venons de le voir dans le chapitre précédent, la modélisation du système océan comporte des simplifications et des approximations. Les paramétrisations des processus non explicitement résolus et les conditions auxiliaires (conditions initiales et aux limites) restent mal connues et/ou entachées d'erreurs. Ainsi la dynamique simulée par les modèles présente des erreurs (tant sur l'état moyen que sur la variabilité) quand on la confronte à la réalité de l'océan. Les océanographes ont une deuxième source d'information à leur disposition pour décrire l'état de la machine océan : les observations. Il existe deux types d'observations, les observations in-situ et les observations à distance qui pour l'essentiel sont des observations satellitaires. Le nombre, la résolution et la précision des observations n'ont cessé de s'améliorer au cours des dernières années. Pour autant, les observations restent elles aussi entachées d'erreurs et sont toujours discontinues en temps et en espace, excepté pour la surface où les satellites nous fournissent une vision synoptique. Il est donc impossible de reconstruire un champ quadridimensionnel de l'état de l'océan à partir des seules observations.

Ni les observations, ni les modèles ne sont donc capables de fournir individuellement une description détaillée et cohérente de l'océan. C'est devant ce constat que les méthodes d'assimilation de données ont fait leur apparition en océanographie. L'objectif général de l'assimilation de données est de combiner diverses sources d'information afin de produire une estimation optimale de l'état d'un système dynamique et de son évolution. Les différentes techniques d'assimilation de données ne sont pas vraiment récentes. Les fondements théoriques remontent à peu près au milieu du $XX^{\text{ème}}$ siècle, alors que les méthodes les plus sophistiquées ont vu le jour dans les années 70. L'apparition de l'assimilation de données en océanographie est cependant relativement récente (courant des années 80). A cela diverses explications, mais la principale est sans nul doute le fait que l'océan est un milieu difficile à observer. Jusque dans les années 70, le seul moyen d'observation de l'océan était l'organisation de coûteuses campagnes océanographiques qui ne permettaient qu'une description limitée et ponctuelle (dans l'espace et le temps) de l'état de l'océan. A partir des années 70, les mouillages instrumentés dans l'océan hauturier ont fait leur apparition, fournissant des séries temporelles longues mais laissant entier le problème de la couverture spatiale limitée. Il aura fallu attendre la fin des années 80 et surtout le début des années 90 avec le lancement du satellite Topex/Poseidon en août 1992 pour que les océanographes aient à leur disposition une base d'observation du niveau de la mer précise, globale et continue dans le temps. Ce n'est que depuis le début des années 90 que les deux composantes principales d'un système d'assimilation, à savoir un modèle de prévision et un système d'observation, existent dans le cadre de l'océanographie. Ceci explique en partie le retard relatif de l'assimilation de données en océanographie par rapport à son pendant pour l'atmosphère, la météorologie, où l'assimilation de données est utilisée depuis longtemps. La nécessité de prévision du temps (pour l'atmosphère) représente également un besoin sociétal plus fort que la prévision de l'état de l'océan ce qui explique pourquoi les efforts de recherche se sont d'abord concentrés sur la météorologie.

Dans ce chapitre, nous commencerons par donner un aperçu du contexte scientifique, avant d'exposer rapidement les différentes méthodes d'assimilation existantes. Nous aborderons ensuite plus en détail le filtrage de Kalman et son approximation en rang réduit développé à Grenoble, le filtre SEEK (Singular Evolutive Extended Kalman filter). Nous finirons en passant en revue les développements pratiques implémentés avec le SEEK, dé-

FIG. 4.1 – *Système d'assimilation de données.*

veloppements nécessaires à la mise en œuvre d'une plate-forme d'assimilation basée sur OPA en surface libre et le SEEK.

4.2 L'assimilation de données en général

4.2.1 Généralités

Les différents problèmes d'assimilation de données appartiennent à la vaste catégorie des problèmes inverses. La thématique de base des problèmes inverses est la reconstruction d'une ou plusieurs grandeurs à partir d'un ensemble d'observations caractérisant la ou les grandeurs que l'on cherche à reconstruire. Les méthodes d'assimilation de données se distinguent des autres méthodes inverses par le fait qu'elles s'appliquent à des systèmes dynamiques. Les grandeurs que l'on cherche à reconstruire sont gouvernées par des équations d'évolution.

L'assimilation de données regroupe l'ensemble des méthodes mathématiques et numériques qui permettent de reconstituer, d'améliorer la connaissance de l'état du système étudié (pour nous l'océan) en combinant de manière optimale l'information mathématique contenue dans les équations du modèle et l'information physique provenant des données (observations). Si les modèles numériques sont capables de reconstruire l'état du système sur tout le domaine spatial et temporel ils n'en restent pas moins imparfaits. En effet, les modèles actuellement utilisés en océanographie comportent des simplifications et des approximations (cf chapitre 3). Les observations de l'océan, bien que de plus en plus précises, restent entachées d'erreurs, incomplètes et ne couvrent pas tout le domaine spatio-temporel. Il est alors important, afin de tirer le meilleur parti de ces deux sources d'informations indépendantes, de les combiner (cf figure 4.1) et ainsi d'améliorer notre compréhension du système océan mais aussi et d'affiner la capacité prédictive des modèles numériques de circulation océanique. De manière simpliste on peut dire que les observations vont guider le modèle sur une trajectoire réaliste, tandis que le modèle va fournir une interpolation spatio-temporelle dynamique des observations. Historiquement, les météorologues ont joué un rôle de précurseurs en matière d'assimilation de données en dynamique des fluides géophysiques. Depuis une trentaine d'années, la plupart des

développements méthodologiques ont été réalisés en relation avec la météorologie. L'existence de modèles atmosphériques relativement fiables et d'observations suffisamment bien échantillonnées ayant joué un rôle de catalyseur. Aujourd'hui, plusieurs méthodes ont été développées directement par les assimilateurs océanographes comme le filtre SEEK (*Pham et al.*, 1998) ou le Kalman d'ensemble (*Evensen*, 2003).

Il existe différentes méthodes d'assimilation de données. On les classe généralement en deux groupes : les méthodes stochastiques de filtrage ou méthodes séquentielles qui reposent sur une approche statistique comme l'algorithme du filtre de Kalman et les méthodes variationnelles qui reposent sur la théorie du contrôle optimal. Dans un cadre purement linéaire, les techniques séquentielles et variationnelles sont équivalentes.

4.2.2 Les applications

En pratique, les principaux objectifs et potentialités de l'assimilation de données océanographiques sont :

- Analyse des modèles de circulation océanique et amélioration de la compréhension du système océan

Les observations, bien qu'imparfaites, sont une source d'information indispensable (dans le sens où elles reproduisent fidèlement la réalité de l'océan) à la compréhension de la dynamique océanique. L'assimilation présente l'intérêt majeur de permettre de combiner de manière optimale cette information avec les autres informations disponibles sur le système, principalement : un modèle dynamique (qui peut être décrit par un ensemble d'équations aux dérivées partielles non linéaires), des statistiques relatives aux erreurs d'observation et aux erreurs du modèle. Cela permet ainsi d'aller au delà de la simple vision instantannée fournie par les observations. Dans l'optique d'un système d'assimilation performant, cela débouche sur une vision quadridimensionnelle optimale de l'état de l'océan et une augmentation du réalisme des simulations et des processus simulés. La conséquence directe étant une meilleure compréhension de la dynamique de l'océan et de sa variabilité.

- Estimation de paramètres d'un modèle physique

Comme nous l'avons expliqué dans le chapitre 3, les modèles numériques utilisent de nombreuses paramétrisations pour représenter les processus sous-maille, ou non explicitement résolus. La mauvaise connaissance de ces processus couplée aux approximations consenties par les modélisateurs pour paramétrer ces processus complexes font qu'il existe une grande incertitude sur les coefficients de ces paramétrisations. L'assimilation de données permet d'évaluer et d'optimiser certains de ces paramètres afin de les réintroduire dans le modèle libre afin de donner des simulations plus réalistes. L'estimation de paramètres par assimilation de données a un impact grandissant sur la modélisation de l'océan. Elle permet non seulement d'évaluer des paramètres du modèle lui-même, mais aussi des paramètres extérieurs au modèle pour peu que ces derniers soient suffisamment bien corrélés avec les variables observées. Dans les études interdisciplinaires par exemple, un des problèmes auxquels se confrontent les modélisateurs biologistes est l'impossibilité de mesurer les taux in-situ (par exemple le taux de broutage, ou le taux de mortalité). Avec l'estimation de paramètres par assimilation de

données, on peut utiliser des quantités que l'on sait mesurer, comme la concentration de plancton par exemple et un modèle biogéochimique pour estimer ces taux. La "régionalisation" des paramètres peut également tirer un grand parti de l'assimilation de données.

- Optimisation des réseaux d'observation

Un autre domaine dans lequel l'assimilation de données a un rôle à jouer est l'optimisation des réseaux d'observation. Comme nous l'avons présenté, l'observation de l'océan a fait un réel bond en avant avec les satellites et la télédétection. Il reste cependant encore beaucoup à faire. L'amélioration des réseaux d'observation de l'océan, in-situ et satellites, représente un coût humain et financier considérable. L'assimilation peut servir à les optimiser. Elle peut également servir de banc d'essai pour les systèmes d'observation à venir. Le but est alors de contribuer à la détermination des caractéristiques optimales des observations du futur système en vue de leur assimilation dans des modèles. Ces expériences de simulation de système d'observation sont de plus en plus utilisées en océanographie. Par exemple, *Verron et Holland* (1989) ont étudié l'impact des données altimétriques satellitaires sur la reconstruction de la circulation océanique, selon différents paramètres orbitaux ainsi que selon le niveau de bruit des observations. Le travail de thèse de *Durand* (2003) constitue un travail prospectif sur l'impact des futures données satellites de salinité de surface fournies par le satellite SMOS et AQUARIUS. Récemment dans l'équipe MEOM, *Debost* (2004) a travaillé sur l'évaluation des performances de différents types de constellations satellitaires pour la reconstruction de la circulation océanique à meso-échelle par assimilation de données. Cette liste est loin d'être exhaustive.

- Les prévisions

Cette application peut être vue comme une finalité pour l'assimilation de données. Elle n'en reste pas moins la plus délicate à mettre en place. Le fait de pouvoir réaliser des prévisions de l'état de l'océan à différentes échelles de temps ouvre un large spectre d'applications, tant au niveau de la gestion des ressources halieutiques, que de la prévention de dégradations environnementales, ou encore de la prévention des risques naturels ainsi que la prévision climatique. L'émergence, ces dernières années, de l'océanographie opérationnelle a donné une toute autre dimension à ce secteur d'application de l'assimilation de données. Le projet français MERCATOR¹ en est un exemple significatif. Il est aujourd'hui possible d'obtenir des prévisions de manière opérationnelle de l'état de l'océan mondial à 15 jours (c'est à dire que les données d'observations de l'océan sont collectées et assimilées de manière systématique dans un OGCM afin prévoir l'évolution de l'état de l'océan comme cela est fait pour l'atmosphère en météorologie). Les prévisions sont réalisées à haute (1/4°), voir très haute résolution (1/15°) pour l'Atlantique Nord. Ces bulletins du "temps" de l'océan sont en libre accès sur le site internet de MERCATOR. La création de grands projets fédérateurs internationaux tel que GODAE² (Global Ocean Data Assimilation Experiment), qui a débuté en 2003, ou encore MERSEA³ (Marine EnviRonment and Security for the European Area),

¹<http://www.mercator.com>

²<http://www.bom.gov.au/GODAE/>

³<http://www.mersea.eu.org>

qui a débuté en 2004, montre les enjeux que représente, aujourd'hui, l'assimilation de données en mode prévisionnel appliquée à l'océan. En ce qui concerne les prévisions à plus long terme (prévisions saisonnières et climatiques), les modèles de climat reposent sur des modélisations couplées entre l'atmosphère et l'océan. Les modèles d'océan sont donc indispensables aux modèles climatiques. Les sorties du prototype global basse résolution de MERCATOR sont actuellement utilisées comme forçage pour le modèle de prévision saisonnière de Météo-France.

Un nombre relativement conséquent de méthodes d'assimilation de données a été mis en œuvre pour répondre aux problèmes posés ci-dessus. Ces méthodes reposent sur deux théories principales qui définissent les deux grands courants de la communauté des assimilateurs en océanographie. La première se fonde sur une approche de type contrôle optimal et est à l'origine de ce que l'on appelle généralement les méthodes variationnelles. La deuxième est basée, quand à elle, sur une estimation statistique optimale et a débouché sur un ensemble de méthodes que l'on regroupe sous le terme de méthodes statistiques ou séquentielles. Ces deux grands types de méthodes d'assimilation de données sont détaillées dans la section suivante. A côté de ces deux grands types de méthodes, dites optimales, il y a les méthodes empiriques, dont nous donnerons un exemple ci-après avec le nudging.

4.2.3 Les différentes méthodes d'assimilation de données

Le nudging

L'algorithme du nudging consiste à introduire dans le système d'équations dynamiques un terme de relaxation vers les observations. Le modèle apparaît alors comme une contrainte faible et le terme de rappel force les variables du modèle à coller avec les observations. Cette méthode, d'abord utilisée en météorologie, a été examinée avec succès dans un contexte océanographique, notamment par (*Verron et Holland*, 1989; *Verron*, 1992; *Blayo et al.*, 1994). Les résultats sont généralement encourageants, mais cette technique simple ne tient pas compte des erreurs respectives des observations et du modèle.

Plus récemment, l'algorithme BFN du nudging direct et rétrograde (back and forth nudging), a été introduit par *Arnoux et Blum* (2005) afin d'identifier l'état initial d'un système à partir d'observations réparties en temps et/ou en espace. Cet algorithme consiste à résoudre dans un premier temps l'équation directe avec le terme de nudging, puis en repartant de l'état final ainsi obtenu, à résoudre les mêmes équations de façon rétrograde avec un terme de rappel opposé à celui du nudging direct. On obtient ainsi une estimation de l'état initial. Ce procédé est alors répété de façon itérative jusqu'à la convergence de l'état initial. Cet algorithme est donc à rapprocher des méthodes variationnelles, dans la mesure où il revient à estimer l'état initial du système par itérations successives.

Les méthodes variationnelles

Cette approche de l'assimilation de données est issue de la théorie du contrôle optimal (*Lions*, 1968). Dans le domaine météorologique, les méthodes variationnelles semblent devoir s'imposer pour le futur dans la plupart des centres opérationnels importants :

Météo-France⁴, ECMWF⁵, NCEP⁶. Il s'agit de prendre en compte globalement, dans la phase d'analyse, une série temporelle d'observations distribuées sur un intervalle $[0, T]$. Le problème se ramène à la minimisation d'une fonction coût $J(x) = J_o(x) + J_b(x)$, composée de la somme de l'écart quadratique entre la solution du modèle sur une période $[0, T]$ et une série temporelle d'observations distribuées sur la même période (le terme J_o), et de l'écart quadratique entre les variables de contrôles x et une ébauche de ces variables (le terme J_b). La variable de contrôle généralement privilégié pour les modèles de circulation océanique est la condition initiale, mais plusieurs travaux récents (*Vidard, 2001; Durbiano, 2001*) montrent qu'il est possible de contrôler l'erreur modèle. Si la condition initiale est retenue comme variable de contrôle, le terme J_b se résume alors à l'écart entre l'ébauche et la trajectoire corrigée à l'instant initial. Ce deuxième terme joue un rôle régularisant pour la solution corrigée obtenue. On considère en effet que l'ébauche est un état pertinent du système et qu'il ne faut pas trop s'en écarter.

L'objectif de la formulation variationnelle telle qu'on la rencontre classiquement en océanographie revient à estimer l'état initial x_0 qui minimise la fonction coût $J(x)$. Parmi les méthodes qui existent pour minimiser $J(x)$, citons la méthode adjointe proposée par *Le Dimet et Talagrand* (1986). Cette méthode est couramment utilisée en océanographie. Elle utilise le modèle adjoint, transposé du modèle direct linéarisé, pour calculer le gradient de la fonction coût. Si le gradient respecte un critère de minimisation déterminé, on considère que l'état initial est optimal. Dans le cas contraire, on calcule un nouvel état initial en fonction du gradient et de la fonction coût. On répète ce processus itératif jusqu'à ce que l'on obtienne une convergence satisfaisante.

La méthode adjointe pour la minimisation de la fonction coût $J(x)$ est performante pour les problèmes de grande taille, ce qui explique qu'elle soit fréquemment utilisée pour des applications océanographiques. Elle présente cependant plusieurs inconvénients. Elle devient très vite extrêmement coûteuse en temps de calcul comme en espace de stockage dans le cas d'un modèle d'océan réaliste tridimensionnel car elle requiert plusieurs intégrations du modèle direct, mais également du modèle adjoint. Le coût de calcul de la minimisation itérative de la fonction coût peut donc rapidement devenir prohibitif. Sur-tout, elle nécessite le développement du modèle adjoint qui est une tâche tout sauf triviale. Enfin, dans le cas de modèles à haute résolution, les échelles de décorrélation temporelle des champs peuvent devenir assez courtes. La dynamique non-linéaire de l'océan limite alors la durée pendant laquelle le modèle peut être approximé par son linéaire tangent et donc la durée de la période d'assimilation pendant laquelle la condition initiale est un paramètre de contrôle (*Luong, 1995*). On est alors contraint de procéder séquentiellement, en effectuant la minimisation sur des fenêtres temporelles plus courtes (*Luong et al., 1998*). On minimise alors séquentiellement la fonction coût sur chacune des périodes de sorte à reconstruire la trajectoire optimale globale. On perd alors la continuité de la trajectoire et le coût global de calcul augmente. Dans ce dernier contexte, il devient plus difficile de distinguer les méthodes variationnelles de la deuxième grande famille de méthode d'assimilation, les méthodes statistiques.

⁴<http://www.meteofrance.com>

⁵<http://www.ecmwf.int/>

⁶<http://www.ncep.noaa.gov/>

Les méthodes statistiques

L'approche statistique séquentielle repose quant à elle sur une estimation, à un instant donné, de l'état optimal du système en fonction des connaissances statistiques disponibles. Les méthodes statistiques permettent d'avoir, à moindre coût, un accès direct à la statistique des erreurs d'analyses. L'étape d'analyse, au cours de laquelle une correction de l'état courant du système est évaluée en fonction des erreurs respectives du modèle et des observations au sens des moindres carrés, est suivie d'une phase de prévision. Cette étape de prévision est assurée par le modèle qui se charge de faire évoluer les champs analysés ainsi que la covariance d'erreur suivant sa propre dynamique. On parle d'estimation optimale dans la mesure où l'état estimé minimise la variance de l'erreur. Dans ce contexte de la théorie de l'estimation statistique optimale, l'état analysé n'est donc pas forcément une solution du modèle, contrairement à l'état analysé obtenu dans le cadre de la théorie du contrôle optimal qui cherche la solution du modèle qui minimise l'écart aux observations.

Les différentes techniques d'assimilations statistiques optimales utilisées en océanographie dérivent du filtre de Kalman (*Kalman*, 1960). En effet, l'approche du filtrage de Kalman (KF) fournit une solution rigoureuse au problème du BLUE (Best Linear Unbiased Estimation), qui est à l'origine de la théorie de l'estimation optimale. Dans la mesure où l'algorithme du KF est à la base du filtre SEEK présenté dans la section 4.4 et utilisé pour les expériences d'assimilation réalisées au cours de cette thèse, nous nous attarderons un peu plus longuement sur les fondements théoriques et les hypothèses permettant d'en écrire les équations.

4.3 La théorie du filtrage de Kalman

Dans ce qui suit, nous utiliserons la notation proposée par *Ide et al.* (1997) qui est la plus communément admise pour décrire les problèmes d'assimilation. Le but de cette section n'est pas d'exposer de manière extensive la théorie du KF, mais plutôt d'introduire les notions essentielles nécessaires à la bonne compréhension des problèmes discutés ultérieurement. Le lecteur désirant de plus amples détails sur les fondements théoriques de l'estimation optimale et du filtrage de Kalman pourra se référer à la littérature dédiée (i.e. *Gelb*, 1974).

4.3.1 Positionnement du problème

Voyons comment on arrive aux équations du filtre de Kalman au travers d'un exemple simple. Supposons que l'on connaisse le vecteur d'état de l'océan \mathbf{x}_i^a à l'instant t_i . Le vecteur d'état est de dimension n . Nous avons également à notre disposition un modèle physique décrivant l'évolution du vecteur d'état de l'océan de l'instant t_i à l'instant t_{i+1} qui peut être mis sous la forme d'un opérateur matriciel linéaire $\mathbf{M}_{i,i+1}$.

$$\mathbf{x}_{i+1}^f = \mathbf{M}_{i,i+1} \mathbf{x}_i^a \quad (4.1)$$

où \mathbf{x}_{i+1}^f est le vecteur "prévu" (forecast en anglais) décrivant l'état du système à l'instant t_{i+1} . À l'instant t_{i+1} , nous disposons également d'une autre source d'information importante, les observations collectés dans le vecteur \mathbf{y}_{i+1} de dimension p . Le filtre de Kalman permet de calculer le meilleur estimé (au sens des moindres carrés) de l'état "vrai" (true

en anglais) de l'océan \mathbf{x}_{i+1}^t à l'instant t_{i+1} en fonction de ces deux sources d'information indépendantes, moyennant la connaissance des erreurs relatives sur \mathbf{x}_{i+1}^f et \mathbf{y}_{i+1} .

4.3.2 Erreurs et densités de probabilité

La précision de la prévision de l'état du système \mathbf{x}_{i+1}^f à l'instant t_{i+1} dépend de deux choses : l'erreur sur la condition initiale et l'erreur modèle. L'erreur entre la condition initiale \mathbf{x}_i^a et l'état vrai de l'océan à l'instant t_i peut s'écrire sous forme vectorielle $\epsilon_i^a = \mathbf{x}_i^a - \mathbf{x}_i^t$. Il est évident que l'on n'a pas accès à la valeur de ϵ_i^a . En revanche, on peut faire un certain nombre d'hypothèses sur ses propriétés statistiques : supposons que \mathbf{x}_i^a est non biaisé ($\langle \epsilon_i^a \rangle = 0$) et que son erreur ϵ_i^a suit une loi normale. Sa densité de probabilité s'écrit donc :

$$\epsilon_i^a \longrightarrow \mathcal{N}(0, \mathbf{P}_i^a) \sim \exp \left[-\frac{1}{2} \epsilon_i^{aT} \mathbf{P}_i^{a-1} \epsilon_i^a \right] \quad (4.2)$$

où $\mathbf{P}_i^a = \langle \epsilon_i^a \epsilon_i^{aT} \rangle$ est la matrice de covariance des erreurs associées à \mathbf{x}_i^a de dimension $n \times n$. Par construction, la matrice de covariance des erreurs est symétrique et définie positive. La deuxième source d'erreur sur l'état prédit est imputable à l'opérateur modèle $\mathbf{M}_{i,i+1}$ qui est imparfait. L'erreur modèle s'écrit :

$$\eta_i = \mathbf{M}_{i,i+1} \mathbf{x}_i^t - \mathbf{x}_{i+1}^t \quad (4.3)$$

Là encore, on ne connaît pas l'erreur modèle mais on peut émettre l'hypothèse qu'elle est distribuée normalement et centrée ($\langle \eta_i \rangle = 0$) :

$$\eta_i \longrightarrow \mathcal{N}(0, \mathbf{Q}_i) \sim \exp \left[-\frac{1}{2} \eta_i^T \mathbf{Q}_i^{-1} \eta_i \right] \quad (4.4)$$

où $\mathbf{Q}_i = \langle \eta_i \eta_i^T \rangle$ est la matrice $n \times n$ de covariance des erreurs modèle. En plus des hypothèses émises ci-dessus, on suppose également que l'erreur sur la condition initiale est décorrélée de l'erreur modèle, soit $\langle \epsilon_i^a \eta_i^T \rangle = 0$. Dans le cas d'un modèle réaliste de circulation générale de l'océan, il est clair que ces hypothèses sur la statistique des erreurs sont des hypothèses fortes et peu réaliste. En particulier, il est fréquent, pour ne pas dire de rigueur d'obtenir un biais modèle. Ces hypothèses sont cependant nécessaires afin d'écrire les équations de l'étape d'analyse du KF. La figure 4.2 présente une vision schématique du diagramme des erreurs dans l'espace d'état.

En partant des hypothèses énoncées ci-dessus, on en déduit aisément l'erreur sur le premier estimé, ou erreur de prévision, qui est la différence entre l'état prédit \mathbf{x}_{i+1}^f et l'état vrai \mathbf{x}_{i+1}^t de l'océan à l'instant t_{i+1} :

$$\begin{aligned} \epsilon_{i+1}^f &= \mathbf{x}_{i+1}^f - \mathbf{x}_{i+1}^t \\ &= \mathbf{M}_{i,i+1} \mathbf{x}_i^a - (\mathbf{M}_{i,i+1} \mathbf{x}_i^t - \eta_i) \\ &= \mathbf{M}_{i,i+1} \epsilon_i^a + \eta_i \end{aligned} \quad (4.5)$$

Dans le cas d'un opérateur modèle $\mathbf{M}_{i,i+1}$ linéaire, on peut facilement déterminer les propriétés statistiques de cette erreur de prévision. En effet, en combinant l'équation 4.5 avec les équations 4.2 et 4.4, on montre relativement aisément que l'erreur de prévision est non biaisée ($\langle \epsilon_{i+1}^f \rangle = \mathbf{M}_{i,i+1} \langle \epsilon_i^a \rangle + \langle \eta_i \rangle = 0$) et est distribuée normalement :

$$\epsilon_{i+1}^f \longrightarrow \mathcal{N}(0, \mathbf{P}_{i+1}^f) \sim \exp \left[-\frac{1}{2} \epsilon_{i+1}^{fT} \mathbf{P}_{i+1}^{f-1} \epsilon_{i+1}^f \right] \quad (4.6)$$

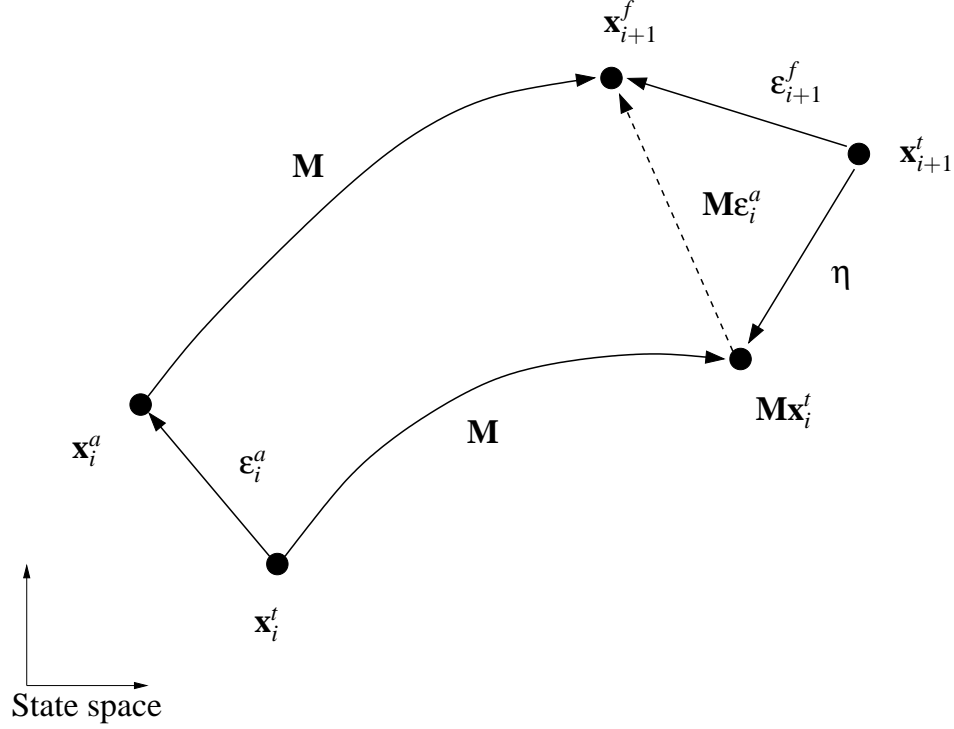


FIG. 4.2 – Représentation vectorielle des erreurs d'analyse et de prévision dans l'espace d'état (Source (Brasseur, 2006)).

où la matrice de covariance des erreurs du premier estimé \mathbf{P}_{i+1}^f est définie par :

$$\begin{aligned}
 \mathbf{P}_{i+1}^f &= \langle \boldsymbol{\varepsilon}_{i+1}^f \boldsymbol{\varepsilon}_{i+1}^{fT} \rangle \\
 &= \mathbf{M}_{i,i+1} \langle \boldsymbol{\varepsilon}_i^a \boldsymbol{\varepsilon}_i^{aT} \rangle \mathbf{M}_{i,i+1}^T + \langle \boldsymbol{\eta}_i \boldsymbol{\eta}_i^T \rangle \\
 &= \mathbf{M}_{i,i+1} \mathbf{P}_i^a \mathbf{M}_{i,i+1}^T + \mathbf{Q}_i
 \end{aligned} \tag{4.7}$$

L'équation 4.7 est la première équation fondamentale du filtre de Kalman. Si l'on regarde cette équation sous un angle physique, la covariance des erreurs de prévision \mathbf{P}_{i+1}^f est égale à la covariance des erreurs initiale \mathbf{P}_i^a propagée par la dynamique du modèle $\mathbf{M}_{i,i+1} \mathbf{P}_i^a \mathbf{M}_{i,i+1}^T$. Les imperfections du modèle sont prises en compte au travers de la covariance des erreurs modèle \mathbf{Q}_i qui vient amplifier la covariance des erreurs de prévision.

Afin de combiner les observations disponibles à l'instant t_{i+1} avec la prévision faite par l'intermédiaire de notre modèle \mathbf{x}_{i+1}^f , nous avons besoin de connaître, ou plutôt d'estimer l'erreur intrinsèque de notre système d'observation. Le vecteur d'observation \mathbf{y}_{i+1} à l'instant t_{i+1} s'écrit dans l'espace observation à l'aide de l'opérateur d'observation \mathbf{H}_{i+1} de la façon suivante :

$$\mathbf{y}_{i+1} = \mathbf{H}_{i+1} \mathbf{x}_{i+1}^t + \boldsymbol{\varepsilon}_{i+1}^o \tag{4.8}$$

L'opérateur \mathbf{H}_{i+1} est un simple opérateur de passage qui permet de calculer l'équivalent modèle aux point d'observation. Comme précédemment, on fait l'hypothèse que les erreurs d'observations sont non biaisées ($\langle \boldsymbol{\varepsilon}_{i+1}^o \rangle = 0$) et non corrélées avec les erreurs de prévision ($\langle (\mathbf{H}_{i+1} \boldsymbol{\varepsilon}_{i+1}^f) \boldsymbol{\varepsilon}_{i+1}^{oT} \rangle = 0$). La covariance des erreurs d'observation est $\mathbf{R}_{i+1} = \langle \boldsymbol{\varepsilon}_{i+1}^o \boldsymbol{\varepsilon}_{i+1}^{oT} \rangle$.

Attardons nous un peu sur cette matrice \mathbf{R}_{i+1} . Les erreurs d'observation quantifient l'écart entre les données d'observation et l'équivalent de la vérité modèle dans l'espace d'observation (i.e., $\mathbf{H}_{i+1}\mathbf{x}_{i+1}^t$). Elles sont donc la combinaison de deux types d'erreur bien distinctes, les erreurs dues à la précision des mesures \mathbf{E}_{i+1} , et les erreurs de représentativité \mathbf{F}_{i+1} qui proviennent de l'écart entre la vérité modèle et la réalité.

$$\mathbf{R} = \mathbf{E}_{i+1} + \mathbf{F}_{i+1} \quad (4.9)$$

Les erreurs de mesure \mathbf{E}_{i+1} sont pour l'essentiel imputables à la précision des appareils de mesure et sont donc totalement indépendantes du modèle utilisé. En revanche, les erreurs de représentativité \mathbf{F}_{i+1} sont entièrement dépendantes du modèle utilisé et englobent, par exemple, les phénomènes non représentables numériquement sur la grille du modèle, ainsi que l'ensemble des phénomènes physiques non résolus par les équations du modèle. Autant il est relativement aisé d'estimer la matrice \mathbf{E}_{i+1} , autant \mathbf{F}_{i+1} est souvent difficilement quantifiable. Là encore, afin de pouvoir écrire les équations du KF, on supposera que la distribution de probabilité des erreurs de mesure est gaussienne

$$\epsilon_{i+1}^o \longrightarrow \mathcal{N}(0, \mathbf{R}_{i+1}) \sim \exp \left[-\frac{1}{2} \epsilon_{i+1}^{oT} \mathbf{R}_{i+1}^{-1} \epsilon_{i+1}^o \right] \quad (4.10)$$

4.3.3 Analyse optimale

La densité de probabilité 4.6 nous donne la probabilité a priori $P(\mathbf{x}_{i+1}^t)$ alors que 4.10 nous renseigne sur la probabilité d'obtenir les observations \mathbf{y}_{i+1} sachant la vérité modèle $P(\mathbf{y}_{i+1}|\mathbf{x}_{i+1}^t)$. On en déduit alors facilement la probabilité de la vérité modèle connaissant les observations en utilisant la formules de Bayes :

$$P(\mathbf{x}_{i+1}^t|\mathbf{y}_{i+1}) = \frac{P(\mathbf{y}_{i+1}|\mathbf{x}_{i+1}^t)P(\mathbf{x}_{i+1}^t)}{P(\mathbf{y}_{i+1})} \quad (4.11)$$

La solution de notre problème inverse est l'état qui maximise cette probabilité conditionnelle. Une formulation détaillée des méthodes d'assimilation en utilisant l'approche probabiliste est donné par *van Leeuwen et Evensen* (1996). On cherche donc à maximiser la densité de probabilité de $P(\mathbf{x}_{i+1}^t|\mathbf{y}_{i+1})$, ce qui revient à maximiser $P(\mathbf{y}_{i+1}|\mathbf{x}_{i+1}^t)P(\mathbf{x}_{i+1}^t)$ d'après l'équation 4.11. En utilisant 4.6 et 4.10, il vient :

$$P(\mathbf{y}_{i+1}|\mathbf{x}_{i+1}^t)P(\mathbf{x}_{i+1}^t) \sim \exp \left[-\frac{1}{2} \left\{ \epsilon_{i+1}^{oT} \mathbf{R}_{i+1}^{-1} \epsilon_{i+1}^o + \epsilon_{i+1}^{fT} \mathbf{P}_{i+1}^{f-1} \epsilon_{i+1}^f \right\} \right] \quad (4.12)$$

L'estimation optimale de \mathbf{x}_{i+1}^t est donc le vecteur d'état maximisant 4.12, ou de manière équivalente, minimisant

$$J = \epsilon_{i+1}^{oT} \mathbf{R}_{i+1}^{-1} \epsilon_{i+1}^o + \epsilon_{i+1}^{fT} \mathbf{P}_{i+1}^{f-1} \epsilon_{i+1}^f \quad (4.13)$$

En incorporant la définition des erreurs dans l'équation 4.13, on en déduit que la combinaison optimale entre la prévision modèle et les observations correspond au minimum de la fonction coût

$$J(\mathbf{x}) = (\mathbf{y}_{i+1} - \mathbf{H}_{i+1}\mathbf{x})^T \mathbf{R}_{i+1}^{-1} (\mathbf{y}_{i+1} - \mathbf{H}_{i+1}\mathbf{x}) + (\mathbf{x}_{i+1}^f - \mathbf{x})^T \mathbf{P}_{i+1}^{f-1} (\mathbf{x}_{i+1}^f - \mathbf{x}) \quad (4.14)$$

Cette équation quadratique contient deux termes qui quantifient l'écart aux observations et l'écart à la prévision du modèle, pondérés par leurs covariances d'erreurs respectives. En

utilisant la méthode des variations, on obtient une équation implicite pour l'état optimal, ou état analysé \mathbf{x}_{i+1}^a

$$\delta J(\mathbf{x}) = 0 \longrightarrow \mathbf{x}_{i+1}^a = \mathbf{x}_{i+1}^f + \mathbf{P}_{i+1}^f \mathbf{H}_{i+1}^T \mathbf{R}_{i+1}^{-1} (\mathbf{y}_{i+1} - \mathbf{H}_{i+1} \mathbf{x}_{i+1}^a) \quad (4.15)$$

qui peut être résolue explicitement après quelques développements algébriques

$$\mathbf{x}_{i+1}^a = \mathbf{x}_{i+1}^f + \mathbf{P}_{i+1}^f \mathbf{H}_{i+1}^T (\mathbf{H}_{i+1} \mathbf{P}_{i+1}^f \mathbf{H}_{i+1}^T + \mathbf{R}_{i+1})^{-1} (\mathbf{y}_{i+1} - \mathbf{H}_{i+1} \mathbf{x}_{i+1}^f) \quad (4.16)$$

L'état analysé est donc obtenu en corrigeant la prévision du modèle par l'écart entre les observations et l'ébauche (le vecteur innovation $\mathbf{d}_{i+1} = \mathbf{y}_{i+1} - \mathbf{H}_{i+1} \mathbf{x}_{i+1}^f$), pondéré par la matrice de gain de Kalman de dimension $n \times p$ qui s'écrit :

$$\mathbf{K}_{i+1} = \mathbf{P}_{i+1}^f \mathbf{H}_{i+1}^T (\mathbf{H}_{i+1} \mathbf{P}_{i+1}^f \mathbf{H}_{i+1}^T + \mathbf{R}_{i+1})^{-1} \quad (4.17)$$

L'équation 4.17 est la deuxième équation fondamentale du filtre de Kalman. Elle met en balance la covariance de l'erreur de prévision et la covariance de l'erreur totale (i.e., la somme de la covariance de l'erreur de prévision et de l'erreur d'observation) projetée dans l'espace des observations. Si l'on regarde les cas limites que sont un système d'observation parfait ($\mathbf{R} \sim 0$ et $\mathbf{H} \sim \mathbf{I}$), la matrice de gain de Kalman converge vers la matrice identité et l'état analysé colle parfaitement aux observations. A l'opposé, si l'on considère un modèle parfait, soit $\mathbf{P}^f \sim 0$, le gain de Kalman ainsi que la correction sur la prévision tendent vers zéro.

On montre que \mathbf{K}_{i+1} correspond à la minimisation de la trace de la matrice de covariance des erreurs d'analyse (*Miller et Cane, 1989*) donnée par

$$\begin{aligned} \mathbf{P}_{i+1}^a &= \mathbf{P}_{i+1}^f - \mathbf{P}_{i+1}^f \mathbf{H}_{i+1}^T (\mathbf{H}_{i+1} \mathbf{P}_{i+1}^f \mathbf{H}_{i+1}^T + \mathbf{R}_{i+1})^{-1} \mathbf{H}_{i+1} \mathbf{P}_{i+1}^f \\ &= (\mathbf{I} - \mathbf{K}_{i+1} \mathbf{H}_{i+1}) \mathbf{P}_{i+1}^f \end{aligned} \quad (4.18)$$

La séquence d'assimilation est alors obtenue en répétant ce cycle prévision/analyse en séquence. La figure 4.3 illustre la séquence décrite ci-dessus.

4.3.4 La séquence d'assimilation

L'état analysé 4.16 à l'instant t_{i+1} peut servir de condition initiale pour une nouvelle prévision à l'instant t_{i+2} . A l'instant t_{i+2} , de nouvelles observations sont disponibles si bien que le processus peut être répété de manière récursive. Pour résumer, l'algorithme du KF contient deux étapes principales : l'étape de prévision au cours de laquelle on fait évoluer le vecteur d'état du modèle ainsi que les covariances d'erreur associées de l'instant t_i à l'instant t_{i+1} , et l'étape d'analyse qui corrige la prévision en utilisant les observations disponibles à l'instant t_{i+1} . On rappelle ici l'ensemble de la séquence du filtre de Kalman, étendu au modèle faiblement non-linéaire M et à l'opérateur H qui peut lui aussi être non-linéaire. On parle alors de filtre de Kalman Etendu (EKF) (*Jazwinski, 1970*).

On part de la condition initiale \mathbf{x}_i^a et \mathbf{P}_i^a , les équations de l'étape de prévision sont :

$$\mathbf{x}_{i+1}^f = M_{i,i+1} \mathbf{x}_i^a \quad (4.19)$$

et

$$\mathbf{P}_{i+1}^f = M_{i+1} \mathbf{P}_i^a M_{i+1}^T + \mathbf{Q}_i \quad (4.20)$$

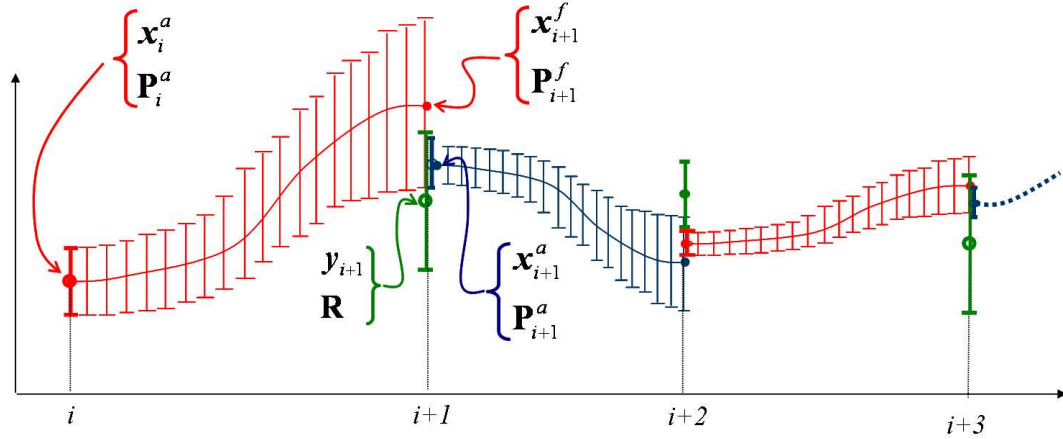


FIG. 4.3 – Représentation conceptuelle de la séquence d'assimilation du filtre de Kalman (Source Brasseur (2006)).

où \mathbf{M} est le modèle linéaire tangent dérivé de $M_{i,i+1}$. La covariance des erreurs est donc propagée par le modèle linéarisé.

L'étape de prévision est suivie d'une étape d'analyse pendant laquelle \mathbf{y}_{i+1} est utilisé pour corriger \mathbf{x}_{i+1}^f :

$$\mathbf{x}_{i+1}^a = \mathbf{x}_{i+1}^f + \mathbf{K}_{i+1}(\mathbf{y}_{i+1} - H_{i+1}\mathbf{x}_{i+1}^f) \quad (4.21)$$

où \mathbf{K}_{i+1} le gain de Kalman s'écrit

$$\mathbf{K}_{i+1} = \mathbf{P}_{i+1}^f \mathbf{H}_{i+1}^T (\mathbf{H}_{i+1} \mathbf{P}_{i+1}^f \mathbf{H}_{i+1}^T + \mathbf{R}_{i+1})^{-1} \quad (4.22)$$

où \mathbf{H}_{i+1} est l'opérateur d'observation linéarisé.

4.4 Le SEEK : de la théorie vers les applications réelles

Le filtre de Kalman et sa variante le filtre de Kalman Etendu qui s'applique aux modèles faiblement non-linéaires, que nous avons introduits dans la section précédente, posent de sérieux problèmes dès lors que l'on cherche à les appliquer à des configurations océanographiques réalistes. Ces problèmes sont de deux types : d'une part des problèmes liés à la méconnaissance des paramétrisations statistiques et d'autre part des problèmes de nature technologique. En effet, il n'existe pas de moyen simple d'estimer les covariances d'erreurs \mathbf{P} , \mathbf{R} et \mathbf{Q} . La grande imprécision dans la connaissance de ces statistiques d'erreurs a pour conséquence principale la perte de l'optimalité de la correction estimée lors de l'étape d'analyse. Les performances du EKF sont donc directement liées à la validité des hypothèses et approximations consenties pour spécifier les covariances d'erreurs \mathbf{P} , \mathbf{R} et \mathbf{Q} . Les problèmes de nature technologique évoqués ci-dessus sont encore plus limitants dans la mesure où ils interdisent purement et simplement l'applicabilité du EKF à une configuration océanographique réaliste. En effet, le EKF nécessite l'expression explicite des matrices \mathbf{P}^a , \mathbf{P}^f et \mathbf{Q} , qui sont toutes de dimension $n \times n$ ou n est la dimension du vecteur d'état du modèle (ici $n \sim 10^6$). Le stockage de telles matrices ainsi que la propagation de

la matrice \mathbf{P}^a par la dynamique linéarisée du modèle, dépasse, et de très loin, la capacité de stockage, mais aussi la puissance de calcul, des machines actuelles.

Le filtre SEEK développé à Grenoble par *Pham et al.* (1998) introduit une approximation supplémentaire afin de rendre l'EKF applicable à des configurations de grande ampleur comme c'est le cas en océanographie. Dans le cadre du filtre SEEK, l'espace d'erreurs en rang complet est approché par un sous-espace de rang réduit. Avec ce principe de réduction d'ordre (cf section suivante), le but est de diminuer le coût numérique et de stockage de l'algorithme du KF, tout en préservant les caractéristiques essentielles qui font la spécificité du KF. Un certain nombre d'arguments, autres que la seule réduction des coûts de calcul, vont dans le sens de l'utilisation d'un espace réduit. Premièrement, le système océan peut être considéré comme un système dynamique possédant un attracteur étrange (*Pham et al.*, 1998). L'existence d'un attracteur a d'ailleurs été démontrée pour les équations de Navier-Stokes (*Lions et al.*, 1997). La géostrophie, par exemple, est un des modes dominants de cet attracteur. Le fait de ne pas corriger le modèle suivant des directions dont on sait qu'il y a de fortes chances qu'elles soient rejetées du fait de la nature attractive du système paraît sensé. Deuxièmement, l'espace d'état est défini par le nombre de degrés de liberté qui est fixé par la discrétisation spatiale du modèle. Or, ce nombre est très largement supérieur au nombre de processus dynamiques qui nous intéressent dans le système. Troisièmement, on connaît vraisemblablement beaucoup mieux la dynamique que les erreurs, et le manque de connaissance des statistiques d'erreurs rendrait probablement l'utilisation d'un KF en rang complet superflu (*Cane et al.*, 1996), tout particulièrement pour le système océan qui est, de plus, sous-observé. Le lecteur pourra trouver une synthèse récente par *Brasseur et Verron* (2006) des développements du filtre SEEK appliqué à l'assimilation de données en océanographie depuis le papier original de *Pham et al.* (1998).

4.4.1 Le principe de la réduction d'ordre

Le principe de réduction d'ordre se fonde sur le fait que la matrice de covariance des erreurs \mathbf{P} est une matrice hermitienne (réelle et symétrique) d'ordre n . Elle est donc diagonalisable avec des valeurs propres réelles et des vecteurs propres orthogonaux. Cette matrice peut donc s'écrire :

$$\mathbf{P} = \mathbf{N}_n \mathbf{\Lambda}_n \mathbf{N}_n^T \quad (4.23)$$

où $\mathbf{\Lambda}_n$ est une matrice diagonale d'ordre n contenant les valeurs propres et \mathbf{N}_n est une matrice d'ordre $n \times n$ contenant les vecteurs propres associés.

La réduction d'ordre consiste à ne conserver que les r premiers vecteurs propres pour exprimer \mathbf{P} . L'espace complet des erreurs est alors réduit à un sous-espace dont la dimension r correspond désormais au nouveau rang de \mathbf{P} . Cette approximation est motivée par le fait que le spectre des valeurs propres décroît généralement très rapidement pour les raisons invoquées ci-dessus.

$$\mathbf{P} = \mathbf{N} \mathbf{\Lambda} \mathbf{N}^T \quad (4.24)$$

où $\mathbf{\Lambda}$ est une matrice diagonale d'ordre r contenant les valeurs propres et \mathbf{N} est une matrice diagonale d'ordre $r \times r$ contenant les vecteurs propres associés.

Par la suite, nous utiliserons la notation suivante. Nous définissons la matrice \mathbf{S} contenant les vecteurs propres multipliés par la racine carrée de leur valeur propre associée :

$$\begin{aligned}\mathbf{P} &= \mathbf{N}\mathbf{\Lambda}\mathbf{N}^T \\ \mathbf{P} &= \mathbf{N}\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}\mathbf{N}^T \\ \mathbf{P} &= \mathbf{S}\mathbf{S}^T\end{aligned}\tag{4.25}$$

où \mathbf{S} est donc une matrice de dimension $n \times r$.

4.4.2 L'initialisation de l'espace réduit

Comme tout filtre de Kalman, l'initialisation du filtre SEEK nécessite un premier estimé \mathbf{x}_0 à l'instant t_0 ainsi qu'une matrice de covariance des erreurs du premier estimé \mathbf{P}_0 . Une des caractéristiques des méthodes d'assimilation issues du KF, et donc du SEEK, est de présenter une grande dépendance face à l'initialisation de \mathbf{P}_0 (*Ballabrera-Poy et al.*, 2001). Une méthode couramment utilisée dans le cadre de la réduction d'ordre pour initialiser le SEEK dans la région du Pacifique Tropical est d'utiliser l'état moyen comme premier estimé, et de construire la matrice \mathbf{S}_0 décrivant le sous-espace d'erreurs initial à partir de la variabilité du système (*Cane et al.* (1996); *Verron et al.* (1999); *Gourdeau et al.* (2000)). Une décomposition en fonctions empiriques orthogonales (EOF) d'une série temporelle d'états du modèle permet de rendre compte de cette variabilité. Cette approche suppose implicitement que les hypothèses suivantes sont vérifiées :

- Le modèle est non biaisé :

$$\langle \mathbf{x} \rangle = \langle \mathbf{x}^t \rangle = \bar{\mathbf{x}}\tag{4.26}$$

- Les covariances du modèle \mathbf{x} et de la vérité modèle \mathbf{x}^t sont identiques

$$\langle (\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T \rangle = \langle (\mathbf{x}^t - \bar{\mathbf{x}})(\mathbf{x}^t - \bar{\mathbf{x}})^T \rangle\tag{4.27}$$

Il est alors possible, en prenant l'état moyen comme état initial, de calculer la covariance des erreurs du premier estimé

$$\begin{aligned}\mathbf{S}_0\mathbf{S}_0^T &= \langle (\mathbf{x}_0 - \mathbf{x}^t)(\mathbf{x}_0 - \mathbf{x}^t)^T \rangle \\ &= \langle (\bar{\mathbf{x}} - \mathbf{x}^t)(\bar{\mathbf{x}} - \mathbf{x}^t)^T \rangle \\ \mathbf{S}_0\mathbf{S}_0^T &= \langle (\mathbf{x} - \bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}})^T \rangle\end{aligned}\tag{4.28}$$

Il est cependant fréquent que les hypothèses 4.26 et 4.27 ne soient pas respectées lorsque l'on se place dans un contexte d'expérience réaliste. Si la variabilité est effectivement utilisée pour la construction de la matrice de covariance des erreurs du premier estimé, il est souvent préférable de partir d'un état modèle relatif à l'état du jour. Par ailleurs, dans le cadre de la réduction d'ordre, seul les r premiers modes de variabilité du modèle sont conservés. Le résultat de l'analyse EOFs étant influencé par les unités des variables (température, salinité, vitesses, ...) qui composent le vecteur d'état du modèle, il est nécessaire de les normaliser en utilisant une métrique appropriée. Une normalisation par les écarts types correspondants est généralement utilisée. Le choix de cette métrique décidera en fin de compte de l'ordre des modes dominants, elle est donc d'autant plus

sensible que la réduction d'ordre est importante. Une base de vecteurs EOFs n'est qu'une façon d'initialiser l'espace d'erreur réduit du filtre SEEK. Dans la littérature, on trouve des auteurs utilisant la méthode de Monte-Carlo (*Evensen, 1994*), ou encore les vecteurs singuliers, de Lyapunov ou issus du breeding (*Durbiano, 2001*).

4.4.3 L'évolution de la covariance des erreurs d'analyse et de prévision

Supposons que la matrice de covariance des erreurs d'analyse \mathbf{P}^a puisse s'exprimer dans le sous-espace d'erreurs \mathbf{S} issu de la décomposition en ordre réduit. On note :

$$\mathbf{P}^a = \mathbf{S}^a \mathbf{S}^{aT} \quad (4.29)$$

La matrice de covariance des erreurs de prévision (cf section précédente) s'écrit pour l'EKF

$$\mathbf{P}_{i+1}^f = \mathbf{M}_{i,i+1} \mathbf{P}_i^a \mathbf{M}_{i,i+1}^T + \mathbf{Q}_i \quad (4.30)$$

où \mathbf{Q}_i est l'erreur modèle. En injectant 4.29 dans l'équation ci-dessus, il vient :

$$\begin{aligned} \mathbf{P}_{i+1}^f &= \mathbf{M}_{i,i+1} \mathbf{S}_i^a \mathbf{S}_i^{aT} \mathbf{M}_{i,i+1}^T + \mathbf{Q}_i \\ \mathbf{P}_{i+1}^f &= (\mathbf{M}_{i,i+1} \mathbf{S}_i^a) (\mathbf{S}_i^{aT} \mathbf{M}_{i,i+1}^T)^T + \mathbf{Q}_i \end{aligned} \quad (4.31)$$

En posant :

$$\mathbf{M}_{i,i+1} \mathbf{S}_i^a = \tilde{\mathbf{S}}_{i+1}^f \quad (4.32)$$

il vient :

$$\mathbf{P}_{i+1}^f = \tilde{\mathbf{S}}_{i+1}^f \tilde{\mathbf{S}}_{i+1}^{fT} + \mathbf{Q}_i \quad (4.33)$$

Dans le cadre du SEEK, *Pham et al. (1998)* propose une paramétrisation simple de l'erreur modèle issue de l'automatique. Ils utilisent un coefficient ρ qu'ils appellent "facteur d'oubli". La covariance de l'erreur d'ébauche s'écrit alors :

$$\mathbf{P}_{i+1}^f = \mathbf{S}_{i+1}^f \mathbf{S}_{i+1}^{fT} = \frac{1}{\rho} \tilde{\mathbf{S}}_{i+1}^f \tilde{\mathbf{S}}_{i+1}^{fT} \quad (4.34)$$

avec le coefficient ρ tel que ($0 < \rho \leq 1$). Cette paramétrisation implique que l'erreur modèle \mathbf{Q}_i est de la forme :

$$\begin{aligned} \mathbf{Q}_i &= \mathbf{P}_{i+1}^f - \tilde{\mathbf{S}}_{i+1}^f \tilde{\mathbf{S}}_{i+1}^{fT} \\ &= \frac{1}{\rho} \tilde{\mathbf{S}}_{i+1}^f \tilde{\mathbf{S}}_{i+1}^{fT} - \tilde{\mathbf{S}}_{i+1}^f \tilde{\mathbf{S}}_{i+1}^{fT} \\ \mathbf{Q}_i &= \frac{1-\rho}{\rho} \tilde{\mathbf{S}}_{i+1}^f \tilde{\mathbf{S}}_{i+1}^{fT} \end{aligned} \quad (4.35)$$

Cette paramétrisation est certes relativement élémentaire puisque, au final, elle revient simplement à "dilater" l'erreur de prévision d'un facteur $\frac{1}{\rho}$ sans privilégier de directions d'erreur particulières, mais la connaissance précise de \mathbf{Q} est très difficilement accessible en pratique. D'autre part, elle présente l'avantage de ne pas engendrer de coût de calcul supplémentaire pendant l'étape d'analyse.

L'équation 4.32 fait apparaître l'évolution des modes d'erreur \mathbf{S}^a en ordre réduit autour de l'état central \mathbf{x}^a par le modèle linéaire tangent $\mathbf{M}_{i,i+1}$. Dans la pratique, notamment au sein de l'équipe MEOM, il est fréquent de se passer de l'écriture du code linéaire tangent. En effet, *Verron et al.* (1999) ont proposé une variante du calcul des statistiques d'erreurs de prévision par interpolation et non plus par évolution par la dynamique du modèle linéaire tangent. Dans ce cas, la covariance des erreurs de prévision en rang réduit se déduit de l'équation suivante :

$$\{\tilde{\mathbf{S}}_{i+1}^f\}_j = \frac{1}{\alpha} \left[M_{i,i+1}(\mathbf{x}_i^a + \alpha\{\mathbf{S}_i^a\}_j) - M_{i,i+1}\mathbf{x}_i^a \right] \quad j = 1, \dots, N_j \quad (4.36)$$

où $\{\}_j$ est la $j^{\text{ème}}$ colonne de la matrice et α un paramètre fixant l'amplitude de la perturbation selon chaque direction. Le nombre d'intégration modèle est ainsi égale au nombre de vecteur de la base réduite.

4.4.4 L'algorithme SEEK

Le principe de la réduction d'ordre appliqué à l'EKF permet, dans le cadre des hypothèses introduites ci-dessus, et après un certain nombre de calculs algébriques qui ne sont pas présentés ici (le lecteur désirant de plus amples informations sur la dérivation des équations du SEEK pourra, par exemple se référer à la thèse de *Testut* (2000)) d'introduire l'algorithme SEEK :

Initialisation (instant t_0)	
$\mathbf{x}_{i=0}$	$= \mathbf{x}_0 \quad (4.37)$
$\mathbf{P}_{i=0}$	$= \mathbf{S}^0 \mathbf{S}^{0T} \quad (4.38)$
Etape d'analyse	
\mathbf{P}_i^f	$= \mathbf{S}_i^f \mathbf{S}_i^{fT} \quad (4.39)$
\mathbf{K}_i	$= \mathbf{S}_i^f [\mathbf{I} + (\mathbf{H}_i \mathbf{S}_i^f)^T \mathbf{R}_i^{-1} (\mathbf{H}_i \mathbf{S}_i^f)]^{-1} (\mathbf{H}_i \mathbf{S}_i^f)^T \mathbf{R}_i^{-1} \quad (4.40)$
\mathbf{x}_i^a	$= \mathbf{x}_i^f + \mathbf{K}_i [\mathbf{y}_i - \mathbf{H}_i \mathbf{x}_i^f] \quad (4.41)$
\mathbf{P}_i^a	$= \mathbf{S}_i^f [\mathbf{I} + (\mathbf{H}_i \mathbf{S}_i^f)^T \mathbf{R}_i^{-1} (\mathbf{H}_i \mathbf{S}_i^f)]^{-1} \mathbf{S}_i^{fT} \quad (4.42)$
Etape de prévision	
\mathbf{x}_{i+1}^f	$= M_{i,i+1} \mathbf{x}_i^a \quad (4.43)$
$\{\tilde{\mathbf{S}}_{i+1}^f\}_j$	$= \frac{1}{\alpha} \left[M_{i,i+1}(\mathbf{x}_i^a + \alpha\{\mathbf{S}_i^a\}_j) - M_{i,i+1}\mathbf{x}_i^a \right] \quad j = 1, \dots, N_j \quad (4.44)$
\mathbf{P}_{i+1}^f	$= \mathbf{S}_{i+1}^f \mathbf{S}_{i+1}^{fT} = \frac{1}{\rho} \tilde{\mathbf{S}}_{i+1}^f \tilde{\mathbf{S}}_{i+1}^{fT} \quad (4.45)$

TAB. 4.1 – Les différentes étapes de l'algorithme SEEK.

4.4.5 Les différentes variantes de l'algorithme SEEK

A partir de l'algorithme SEEK que nous venons de décrire, on peut dériver 3 variantes qui se distinguent par la spécification de l'erreur d'ébauche à chaque étapes.

- **La variante évolutive**

C'est l'algorithme SEEK telle que nous venons de l'introduire. Le sous-espace d'erreur d'analyse évolue à la fois selon la statistique du filtre de Kalman et selon la dynamique du jour, conformément aux équations 4.44 et 4.45.

$$\mathbf{P}_{i+1}^f = \mathbf{M}_{i,i+1} \mathbf{P}_i^a \mathbf{M}_{i,i+1}^T + \mathbf{Q}_i = (\mathbf{M}_{i,i+1} \mathbf{S}_i^a)(\mathbf{M}_{i,i+1} \mathbf{S}_i^a)^T + \mathbf{Q}_i \quad (4.46)$$

La variante évolutive est très coûteuse. Elle nécessite un nombre d'intégration modèle égale au nombre de vecteur de la base réduite pour propager l'erreur d'analyse. Elle est relativement peu utilisée. De précédentes études (*Ballabrera-Poy et al.*, 2001) ont cependant montré qu'elle pouvait s'avérer utile, voir nécessaire, aux moyennes latitudes en raison de la dynamique fortement non-linéaire dans cette région.

- **La variante fixe**

Dans cette variante simplifiée, le sous espace d'erreur d'analyse n'évolue que suivant la statistique du filtre de Kalman et pas selon la dynamique du modèle. Cela revient à ne pas considérer l'équation 4.44. La formulation de l'erreur d'ébauche pour l'étape d'analyse suivante devient alors :

$$\mathbf{P}_{i+1}^f = \mathbf{P}_i^a + \mathbf{Q}_i = \mathbf{S}_i^a \mathbf{S}_i^{aT} + \mathbf{Q}_i \quad (4.47)$$

Par rapport à la variante évolutive, l'application de cette variante peut se révéler suffisante dans certain cas. C'est notamment le cas dans le Pacifique Tropical, comme l'ont montré *Verron et al.* (1999), où l'évolutivité des modes d'erreur d'analyse n'a pas d'impact significatif sur les résultats, et ce à cause de la dynamique quasi-linéaire de la région. Dans ce cas, les faibles gains de l'évolutivité sont à mettre en balance avec le surcoût considérable qu'impose la variante évolutive.

- **La variante stationnaire**

Encore plus simple que la variante fixe, dans cette version, la covariance d'erreur d'analyse \mathbf{P}_i^a est uniquement diagnostiquée, dans la mesure où elle ne sert pas à évaluer la covariance d'erreur d'ébauche \mathbf{P}_{i+1}^f de l'étape suivante. Cela revient à ne pas considérer les équations 4.42, 4.44 et 4.45. La matrice de covariance des erreurs de l'instant initial est alors systématiquement utilisée comme matrice de covariance d'erreur d'ébauche. Cette variante est très similaire à l'interpolation optimale

$$\mathbf{P}_i^f = \mathbf{P}_0 \quad (4.48)$$

4.4.6 Le SEEK local

La réduction d'ordre a permis la faisabilité numérique du SEEK pour des modèles de circulation générale de l'océan. Par contre, la troncature de la base lors de la réduction

d'ordre pose des problèmes de fiabilité des faibles covariances, et donc en particulier, des covariances à très grande distance. En effet, tel que nous l'avons introduit ci-dessus, le filtre SEEK utilise des modes d'erreurs que l'on qualifiera de "globaux", issus d'une décomposition EOFs classique d'une série temporelle d'états du modèle. On parle de modes d'erreurs globaux parce qu'ils sont définis sur l'ensemble du domaine d'étude. On voit alors que l'utilisation de tels modes multivariés globaux va se traduire par la propagation de l'information d'une seule observation à tout le domaine étudié. Ce comportement se déduit facilement des équations du SEEK. L'utilisation de modes d'erreurs globaux n'est pas à remettre en cause fondamentalement, dans la mesure où dans le cas d'un espace d'erreur en rang complet et parfaitement spécifié, la combinaison de tous les modes dans l'espace d'étape conduirait à annuler les covariances à grande distance quand ces dernières sont nulles dans la réalité. En revanche, la troncature du sous-espace d'erreurs en un nombre limité de modes empêche l'annulation de ces covariances et introduit donc des effets indésirables à grande distance.

Afin de corriger ce problème, *Testut* (2000) a développé une variante, dite locale, du SEEK. Cette approche consiste à annuler explicitement les corrélations à grande distance par le biais d'une paramétrisation locale de l'erreur d'ébauche. Cette paramétrisation consiste à introduire a priori l'échelle maximale de corrélation horizontale des erreurs lors de l'analyse. On définit donc des bulles d'influence au-delà desquelles toutes les corrélations sont imposées à zéro. Au final, on utilise donc toujours des modes multi-variés globaux mais avec une influence locale afin de s'affranchir des problèmes à grande distance résultant de la réduction d'ordre.

Dans notre région d'intérêt qu'est le Pacifique Tropical, la dynamique est essentiellement zonale et s'exerce à très grande échelle. La correction globale peut donc a priori paraître bien adaptée pour des expériences d'assimilation réalistes dans cette région. Le SEEK global a d'ailleurs été utilisé pour de précédentes études dans cette région par *Parent* (2000) et *Durand* (2003). Cependant, pour les raisons invoquées ci-dessus, j'ai préféré utiliser l'approche locale que je viens d'introduire. Les zones d'influences utilisées sont, en revanche, anisotrope et de grande taille, afin de respecter les spécificités de la dynamique locale.

4.5 Implémentation du SEEK avec OPA surface libre

4.5.1 Le logiciel SESAM

Malgré l'apparente simplicité de l'algorithme SEEK (cf tableau 4.1), dans les faits, les difficultés pratiques inhérentes à la mise en place d'un système d'assimilation séquentielle basé sur le SEEK sont importantes. Le programme SESAM⁷ (System of Sequential Assimilation Modules) a été développé au sein de l'équipe MEOM par J-M. Brankart, C-E. Testut et L. Parent (*Testut*, 2000) afin de fournir à l'assimilateur un code permettant la mise en œuvre du SEEK dans différentes configurations. Le fait d'avoir SESAM à ma disposition, a grandement facilité la phase implémentation du SEEK avec le modèle OPA dans sa version utilisant une surface libre.

Le code est constitué de différents modules, écrits en Fortran 90. A titre d'exemple, SESAM représente un code d'environ 80000 lignes, ce qui est équivalent au volume du code

⁷<http://www-meom.hmg.inpg.fr/Web/Assimilation/SESAM/>

de circulation générale de l'océan OPA. SESAM gère la phase d'analyse du SEEK, mais également un grand nombre d'autres tâches relatives à l'assimilation de données comme la gestion des données d'observation du système, le calcul d'écart RMS par rapport aux observations, ou encore l'estimation de paramètres statistiques et le calcul de bases EOFs. Un des gros avantages de SESAM est qu'il n'est pas lié à une configuration ou à un modèle particulier. Ainsi, préalablement à ce travail de thèse, il a été utilisé dans des contextes aussi variés que l'assimilation de SSH et de SST dans un modèle de l'Atlantique Nord à haute résolution (*Testut et al.*, 2003), l'assimilation de SSH dans un modèle du Pacifique Tropical (*Parent et al.*, 2003), l'assimilation de SSH et de SST dans un modèle à couche isopycnes de l'Atlantique Nord (*Brankart et al.*, 2003) ou hybride (*Birol et al.*, 2004), et plus récemment, pour l'assimilation de SSH et SST dans un modèle couplé physique-biogéochimie de l'Atlantique Nord (*Berline*, 2005). Cette liste est loin d'être exhaustive. Pour permettre une telle diversité d'applications, le module d'assimilation et le modèle sont complètement indépendants. Le modèle s'arrête à l'instant d'analyse t_i . Le vecteur d'état \mathbf{x}_i est alors transmis à la chaîne d'assimilation. Le module SESAM se charge de calculer l'état analysé à l'instant t_i , ce dernier est à son tour transmis au modèle sous forme d'une condition initiale que l'on fait évoluer par la dynamique du modèle jusqu'à l'instant d'analyse suivant t_{i+1} . Le vecteur d'état \mathbf{x}_i est transmis sous la forme d'un fichier Netcdf et l'ensemble de l'expérience d'assimilation est géré par un méta-script spécifique à chaque expérience d'assimilation et qui est donc en grande partie codé par l'utilisateur.

4.5.2 La réinitialisation du modèle : le schéma IAU

Le redémarrage du modèle est toujours une étape problématique lorsque l'on fait de l'assimilation séquentielle. Dans un cadre de modélisation, le modèle numérique d'océan redémarre en utilisant un fichier "restart". Ce fichier contient l'ensemble des informations nécessaires afin de permettre au modèle de redémarrer comme s'il ne s'était pas arrêté.

Le modèle OPA utilise un schéma Leap-frog (cf équation 3.8) pour l'intégration temporelle. Ceci nécessite donc la connaissance de l'état modèle à deux pas de temps successifs, les champs "before" (à l'instant $t - \Delta t$) et "now" (à l'instant t). Or, compte tenu des approximations faites lors de l'analyse statistique, il n'est pas raisonnable de penser estimer la tendance à l'instant t avec suffisamment de précision pour pouvoir l'utiliser dans le cadre du redémarrage du modèle. Ce que l'on fait classiquement pour contourner ce problème et de forcer le modèle à redémarrer selon un schéma temporel d'Euler (cf équation 4.49). Ce dernier ne nécessite que la connaissance de l'état "now" :

$$u^{t+\Delta t} = u^t + \Delta t \text{ RHS}^t \quad (4.49)$$

Au redémarrage du modèle, après une analyse, on remplace donc les champs "before" et "now" par le champs analysé \mathbf{x}_a et on impose un pas de temps en Euler pour l'intégration temporelle. Cette méthode bien qu'efficace, n'est pas sans poser un certain nombre de problèmes (on pourra consulter *Parent et al.* (2003) pour plus de détail), et, notamment, il n'est plus équivalent de faire 10 simulations de 5 jours et une simulation de 50 jours avec un tel mode de redémarrage.

Dans le cadre de cette thèse, j'ai donc préféré opter pour une méthode incrémentale, l'IAU (pour Incremental Analysis Updating) pour la réinitialisation du modèle après l'analyse SEEK. Cette méthode, comme la plupart des développements méthodologiques en assimilation de données provient de la météorologie. Elle a été introduite par *Bloom et al.*

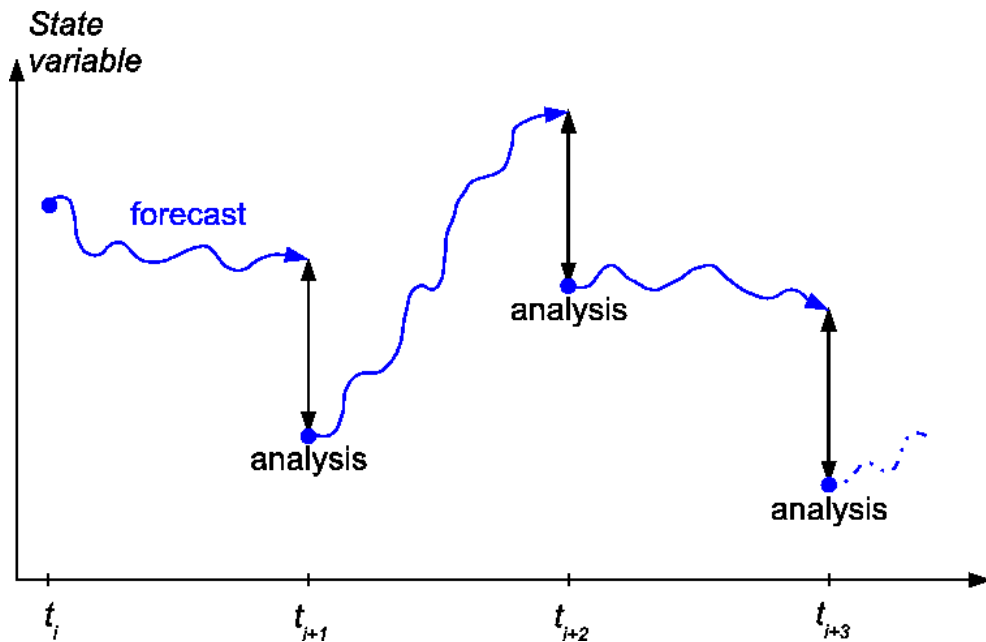


FIG. 4.4 – Schéma de l'assimilation séquentielle classique.

(1996). Le principe de base de l'IAU est d'incorporer directement l'incrément d'analyse $\delta = x^a - x^f$ dans les équations pronostiques du modèle sous la forme d'un terme de forçage. En plus de résoudre le problème du redémarrage du modèle après l'analyse, cette méthode agit comme une méthode d'assimilation séquentielle continue, ce qui conduit à obtenir une trajectoire assimilée continue. La discontinuité entre la prévision x^f et l'analyse x^a est un des inconvénients majeurs des méthodes séquentielles : à cause de la suboptimalité de l'analyse statistique (en raison des approximations consenties et pas toujours vérifiées), cette discontinuité peut engendrer des chocs dans le modèle au moment du redémarrage, entraînant des oscillations erronées à haute fréquence, et un phénomène de rejet de l'information assimilée. Ces ondes transitoires introduites par l'étape d'analyse peuvent être considérées comme le résultat d'une correction imparfaite du vecteur d'état de l'océan, due à des covariances d'erreur mal évaluées. L'IAU est reconnu pour efficacement limiter ces problèmes.

Dans cette section, nous présenterons brièvement la méthode IAU mise en place au cours de cette thèse. Pour plus d'information sur le sujet et pour une étude d'impact détaillée, le lecteur pourra se référer à *Ourmières et al.* (2006). Dans une séquence d'assimilation classique, telle que celle du SEEK (voir figure 4.4), l'incrément d'analyse est appliqué en une seule fois, au moment de l'analyse afin de corriger la prévision modèle en tenant compte de l'information fournie par les observations disponibles. Il en résulte une trajectoire discontinue avec des sauts à chaque étape d'analyse. Comme nous l'avons déjà dit, ces discontinuités dans la trajectoire assimilée peuvent introduire des oscillations à hautes fréquences qui peuvent conduire à un phénomène de rejet de l'information assimilée.

Dans le cadre de l'algorithme IAU, en revanche, l'incrément d'analyse $\delta = x^a - x^f$ est appliqué en continu directement dans les équations pronostiques du modèle sous la forme d'un terme de forçage. En plus des avantages précédemment évoqués, cette tech-

nique présente donc également l'intérêt d'être simple à implémenter. Pour une variable pronostique V donnée, l'équation d'évolution devient :

$$\frac{\partial V}{\partial t} = M + \gamma(t)(V^a - V^f) \quad (4.50)$$

avec M les membres de l'équation d'évolution de la variable V et γ la fonction de répartition temporelle de l'incrément telle que

$$\int_0^{\Delta t} \gamma(t) dt = 1 \quad (4.51)$$

Δt étant la durée de la fenêtre temporelle d'assimilation.

Le lecteur notera les différences entre l'équation 4.50 et le nudging (équation 4.52) qui sera évoqué dans la section traitant de l'initialisation du filtre SEEK.

$$\frac{\partial V}{\partial t} = M + \gamma(t)(V^{obs} - V) \quad (4.52)$$

Dans l'équation 4.52, le terme de correction varie dans le temps, en fonction de l'état instantané de la variable et ne tient absolument pas compte des erreurs respectives des observations et du modèle. Dans l'équation 4.50, en revanche, l'incrément $V^a - V^f$ reste constant (temporellement) sur toute la fenêtre d'assimilation. De plus, la méthode d'obtention de l'incrément est beaucoup plus sophistiquée dans la mesure où l'incrément est issu d'une analyse SEEK. Il prend donc en compte les erreurs relatives du modèle et des observations.

Les applications de l'IAU sont répandues pour l'assimilation dans les modèles de circulation atmosphérique (*Bloom et al.*, 1996; *DeWeaver et Nigam*, 1997; *Zhu et al.*, 2003). Elle sont par contre beaucoup plus restreintes, mais également plus récentes, pour les applications océaniques (*Carton et al.*, 2000; *Huang et al.*, 2002; *Alves et al.*, 2004; *Ourmières et al.*, 2006). L'ensemble de ces auteurs mettent en avant les propriétés de filtre passe-bas pour les ondes d'ajustement ainsi que le caractère continu de la solution analysée qu'offre l'utilisation de l'IAU.

On trouve de nombreuses variantes de la méthode IAU dans la littérature. Elles diffèrent essentiellement par l'instant auquel est calculé l'incrément d'analyse et la fenêtre temporelle d'application de cet incrément. La méthode originale utilisée ici est une variante de celle proposée par *Bloom et al.* (1996). La figure 4.5 illustre la séquence incrémentale utilisée. Dans un premier temps, on fait une prévision par intégration du modèle de l'instant t_i à l'instant t_{i+1} exactement comme dans le cadre de l'algorithme SEEK classique. Une analyse SEEK à l'instant t_{i+1} est alors réalisée. On en déduit un incrément $\delta = x^a - x^f$ pour les variables pronostiques du modèle, soit la température, la salinité et les vitesses zonale et méridienne. La fonction de répartition $\gamma(t)$ est choisie constante, soit $\gamma(t) = 1/\Delta t$. On refait alors une intégration modèle de l'instant t_i à l'instant t_{i+1} en utilisant ces incréments sous la forme de termes de forçage additionnels dans les équations pronostiques de la température, de la salinité et de la vitesse horizontale :

$$\frac{\partial T}{\partial t} = - \nabla \cdot (\vec{u} T) + D^T + \frac{1}{\Delta t}(T^a - T^f) \quad (4.53)$$

$$\frac{\partial S}{\partial t} = - \nabla \cdot (\vec{u} S) + D^S + \frac{1}{\Delta t}(S^a - S^f) \quad (4.54)$$

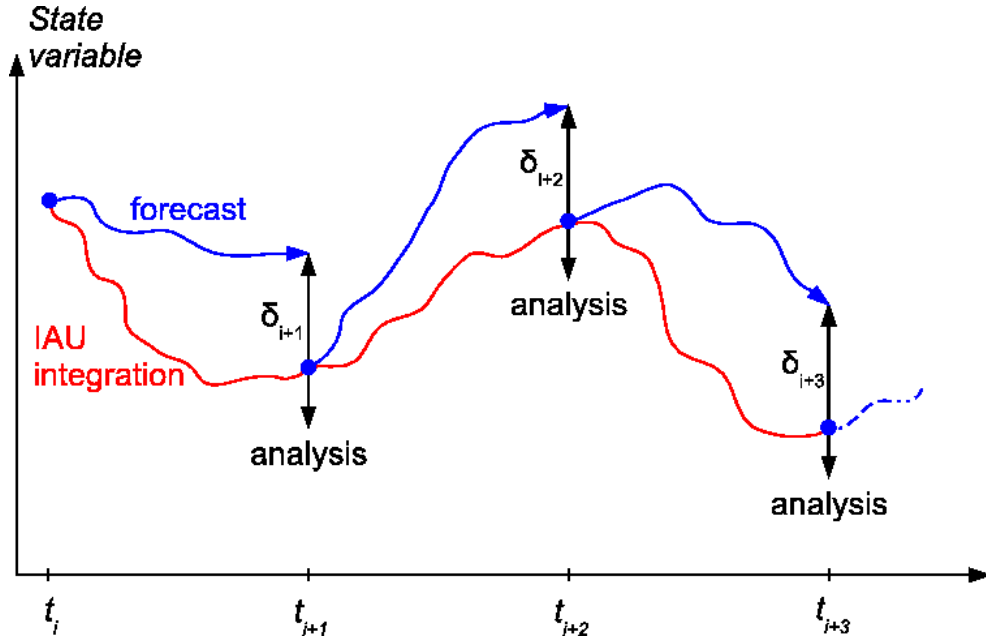


FIG. 4.5 – Schéma IAU.

$$\begin{aligned} \frac{\partial \vec{u}_h}{\partial t} = & - \left\{ (\nabla \times \vec{u}) \times \vec{u} + \frac{1}{2} \nabla (\vec{u}^2) \right\}_h \\ & - f \vec{k} \times \vec{u}_h - \frac{1}{\rho_0} \nabla_h p + g T_c \nabla_h \frac{\partial \eta}{\partial t} + \vec{D}^u + \frac{1}{\Delta t} (\vec{u}^a - \vec{u}^f) \end{aligned} \quad (4.55)$$

A la fin de cette intégration en mode “forcé IAU”, on obtient un nouvel état analysé. Un fichier restart classique du modèle qui contient les états “now” et “before” est généré à la fin de l’intégration avec correction IAU. Ce fichier est utilisé comme condition initiale pour faire une nouvelle prévision de l’instant t_{i+1} à l’instant t_{i+2} et le modèle peut redémarrer normalement en utilisant le schéma Leap-frog dès le premier pas de temps pour l’intégration temporelle.

4.5.3 Domaine d’assimilation

L’océan Pacifique Tropical constitue notre région d’intérêt. Ce travail s’inscrit dans la continuité des travaux utilisant le filtre SEEK dans cette région initiés par *Verron et al.* (1999) et *Gourdeau et al.* (2000), et plus récemment *Parent* (2000) et *Durand* (2003). Pour différentes raisons, nous avons cependant opté pour la configuration globale ORCA2. Ce choix permet, entre autre, de s’affranchir des problèmes, toujours délicats, des frontières ouvertes. Ce choix permet également une synergie avec MERCATOR et l’océanographie opérationnelle, puisque c’est la configuration retenue par MERCATOR pour le développement d’un premier système opérationnel français à l’échelle globale. C’est d’ailleurs une réalité depuis le mois d’Octobre 2005 avec la mise en ligne des bulletins globaux au $1/4^\circ$ basé sur la configuration ORCA025, appelé PSY3 dans l’arborescence des configuration MERCATOR. Par ailleurs, cette configuration a longtemps été la configuration de développement du code OPA, et à ce titre, elle est numériquement bien optimisée. Par exemple, un an de simulation en mode libre (i.e. sans assimilation de données) représente

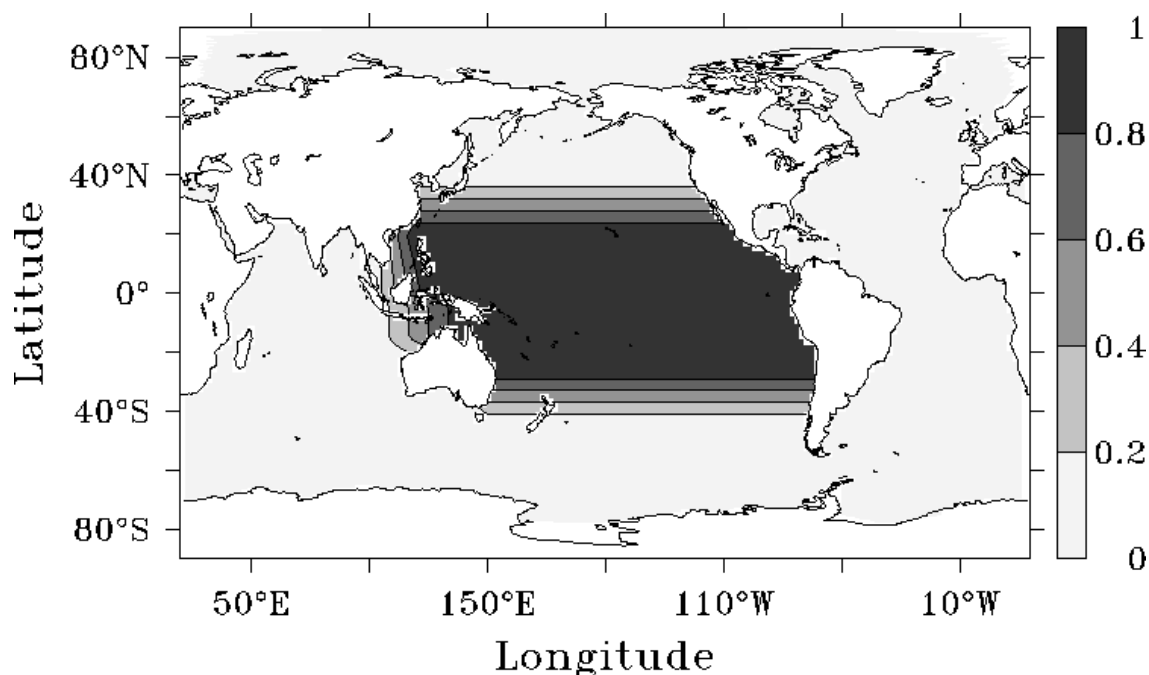


FIG. 4.6 – Coefficient appliqué à l'incrément d'analyse. Isocontours tous les 0.2.

un coût de calcul d'environ 1h30 sur uqbar, la machine vectorielle de l'IDRIS, le centre de calcul du CNRS.

L'existence d'un domaine d'assimilation restreint par rapport au domaine du modèle a cependant demandé quelques aménagements de la plate-forme d'assimilation. Le Pacifique Tropical, qui est la zone où l'assimilation est totalement effective, est défini entre $25^{\circ}S$ à $20^{\circ}N$ méridionalement et entre $120^{\circ}E$ à la côte américaine zonalement, en suivant *Durand et al.* (2002). Cela permet d'inclure toute les régions clés de la variabilité du Pacifique Tropical. Aux frontières de cette zone, des zones tampons sont définies de manière à progressivement relâcher la contrainte de l'assimilation de données. Cette zone de raccord à une extension de 20° . Dans la pratique, on multiplie l'incrément d'analyse par un simple coefficient. Ce coefficient vaut un dans la zone de travail, décroît linéairement de un à zéro dans les zones tampons et est fixé à zéro dans tout le reste du domaine. L'extension des zones tampons correspond à typiquement plusieurs fois l'échelle de décorrélation des champs dynamiques de grande échelle du Pacifique Tropical (*Meyers et al.*, 1991).

4.5.4 Le vecteur d'état

Le vecteur d'état représente le jeu de variables minimum contenant toute l'information sur le système. La connaissance du vecteur d'état permet de diagnostiquer toutes les propriétés du système à un instant donné. Toutes les variables du vecteur d'état sont par définition indépendantes.

Pour le modèle OPA, le vecteur d'état se compose de la vitesse zonale U , de la vitesse méridienne V , de la température T , de la salinité S et de la dénivellation de la surface SSH. Il s'écrit sous la forme suivante :

$$\mathbf{x} = \begin{pmatrix} U & (x, y, z) \\ V & (x, y, z) \\ T & (x, y, z) \\ S & (x, y, z) \\ SSH & (x, y) \end{pmatrix}$$

Dans la version globale ORCA2 du modèle OPA, un champ 2D représente environ 27000 points et un champ 3D 840 000 points. Le vecteur d'état contient donc environ 3 500 000 points. En ce limitant à la région du Pacifique Tropical, on se ramène à environ 500 000 points. C'est ce vecteur d'état qui va être estimé par le filtre SEEK à chaque étape d'analyse.

4.5.5 La spécification de l'espace d'erreur en rang réduit

Les différentes version du SEEK présentées dans la section 4.4.5 se distinguent par la façon dont la base réduite évolue au cours du temps. Pour le type d'expérience d'assimilation que j'ai réalisé au cours de cette thèse, à savoir le contrôle de la structure de l'océan Pacifique Tropical à l'échelle du bassin, *Verron et al.* (1999) ont montré que l'évolution des modes d'erreur d'analyse par la dynamique du modèle (cf équation 4.44) n'a pas d'impact significatif sur les résultats. Au vu du surcoût important (n fois le modèle pour une base de n vecteurs propres), j'ai donc opté pour la version dite "stationnaire" du SEEK. Le sous-espace d'erreur d'analyse n'évolue donc pas et la matrice de covariance des erreurs initiale est systématiquement utilisée comme matrice de covariance des erreurs d'ébauche.

Comme nous l'avons brièvement introduit dans la section 4.4.2, l'initialisation du filtre SEEK nécessite un première estimé \mathbf{x}_0 à l'instant t_0 ainsi qu'une matrice de covariance des erreurs du premier estimé \mathbf{P}_0 . Le filtre SEEK est très sensible à cette phase d'initialisation (*Ballabrera-Poy et al.*, 2001) et la méthode utilisée afin d'initialiser le SEEK est très fortement liée au type d'erreur que l'on veut corriger. Cette étape est donc très importante dans la perspective d'obtenir une assimilation efficace. En plus de l'état \mathbf{x}_0 et de la covariance des erreurs associée \mathbf{P}_0 , il faut également spécifier la covariance des erreurs modèle \mathbf{Q} , ainsi que la covariance des erreurs d'observation \mathbf{R} .

Dans le cadre d'une assimilation d'une topographie dynamique absolue, on cherche à contraindre la variabilité de modèle libre, mais également son état moyen. La topographie dynamique moyenne qui est assimilée est donc différente de la topographie dynamique moyenne du modèle. La covariance des erreurs initiales \mathbf{P}_0 ne peut plus être approximée par la variabilité d'une simulation libre comme c'est classiquement fait (cf section 4.4.2). En effet, utiliser la variabilité d'une simulation libre pour paramétrer le SEEK reviendrait à ignorer toute une partie des directions de l'erreur que l'on cherche à corriger avec l'assimilation de données, celles qui engendrent une structure moyenne différente entre l'océan simulé et l'océan vrai. Par exemple, considérons le cas simple d'un océan à deux couches, une couche chaude de température T_1 en surface et une couche abyssale froide de température T_2 (cette configuration est une bonne approximation pour les océans tropicaux), le SEEK va traduire une erreur en SSH par un incrément en température au niveau de la thermocline alors que cette erreur peut, en fait, être la signature d'une mauvaise position moyenne de la thermocline. Au final, l'assimilation n'a donc que peu de chance d'être efficace, la statistique d'erreur ne contenant qu'une partie de l'erreur que l'on cherche à corriger. Il a donc fallu développer un certain nombre de méthodes spécifiques afin de

mieux estimer l'erreur et ainsi correctement initialiser le filtre SEEK. Ces méthodes seront discutées ultérieurement.

4.6 Conclusion

Ce chapitre avait pour objectif, dans un premier temps, de présenter de façon générale l'assimilation de données. Nous avons notamment donné une liste, loin d'être exhaustive, des différentes applications de l'assimilation de données. Les deux grandes familles de méthodes d'assimilation ont été présentées : les méthodes variationnelles et les méthodes statistiques. Parmi les méthodes statistiques, les fondements théoriques du filtre de Kalman ont été abordés, le filtre de Kalman étant à l'origine du filtre SEEK utilisé dans les expériences d'assimilation réalisées au cours de cette thèse. Dans un deuxième temps, nous avons introduit les aménagements réalisés afin d'implémenter le SEEK avec le modèle OPA en surface libre. La méthode incrémentale IAU utilisée pour réinitialiser le modèle après l'étape d'analyse a été détaillée en particulier. En résumé, la méthode d'assimilation utilisées pour les expériences présentées dans la suite de ce document est le filtre SEEK dans sa variante stationnaire. La base d'erreur en rang réduit est estimée à l'aide d'EOFs 3D globale multi-variées (la façon d'obtenir ces EOFs sera explicitée dans le chapitre suivant) mais avec une influence locale. Une méthode incrémentale IAU utilisée pour réinitialiser le modèle après l'analyse SEEK.

Troisième partie

Résultats

Chapitre 5

Présentation des expériences

Sommaire

5.1	Définition des conditions expérimentales	88
5.1.1	Une assimilation de topographie dynamique absolue. Pour quoi faire?	88
5.1.2	Les différentes configurations du système d'observation	90
5.1.3	Les erreurs d'observation	92
5.1.4	L'état initial \mathbf{x}_0	93
5.1.5	Durée du cycle d'assimilation	93
5.2	Paramétrisation de l'espace d'erreur en rang réduit	94
5.2.1	Positionnement du problème	94
5.2.2	La paramétrisation retenue	95
5.3	Récapitulatif des expériences réalisées	96

Dans les chapitres 3 et 4, nous avons présenté les fondements théoriques et mathématiques du modèle OPA et du filtre SEEK. Les aspects techniques inhérents à leur association dans le cadre d'une plate-forme d'assimilation ont également été abordés, notamment le logiciel SESAM qui représente la colonne vertébrale de notre système d'assimilation. Le chapitre 2 s'est attaché à décrire les observations du Pacifique Tropical que nous avons à notre disposition, ainsi que leurs caractéristiques. Dans ce chapitre, nous détaillerons les caractéristiques techniques des expériences d'assimilation présentées dans la suite du document.

Il est évident que toutes les expériences d'assimilation réalisées au cours de cette thèse ne sont pas discutées dans ce manuscrit. Notamment, préalablement à la mise en place des expériences dites "en données réelles" que je présente ici, des expériences jumelles et cousines ont été réalisées afin de valider notre système d'assimilation dans un cadre simplifié et nous permettant de contrôler le niveau de difficulté. Ces expériences préliminaires nous ont permis de mieux comprendre la nature du problème inhérent à l'assimilation d'une topographie dynamique absolue. Elles nous ont également donné des pistes intéressantes pour l'étape, ô combien importante, de la paramétrisation de l'erreur d'ébauche qui fait l'objet de la section 5.2.

Nous ne présenterons ici que les caractéristiques des expériences discutées dans la suite de ce manuscrit de thèse. J'ai volontairement choisi de limiter au maximum le nombre de paramètres susceptibles d'être modifiés d'une simulation à l'autre. Au delà de la simplification de la présentation des résultats, c'est plus dans une optique d'inter-comparaison entre les différentes simulations avec assimilation de données que ce choix a été fait.

5.1 Définition des conditions expérimentales

5.1.1 Une assimilation de topographie dynamique absolue. Pour quoi faire ?

En premier lieu, il convient de rappeler que la topographie dynamique est la grandeur que simulent les OGCMs. Il est donc tout à fait naturel d'assimiler cette grandeur pour contraindre le modèle plutôt que la SLA référencée avec le niveau moyen du modèle comme cela est couramment fait. C'est uniquement à cause des trop fortes erreurs sur le géoïde (cf chapitre 2) qui viennent contaminer la topographie dynamique que les assimilateurs se sont limités à l'utilisation de la seule partie résiduelle du signal altimétrique mesuré par les satellites. L'absence d'une bonne topographie dynamique moyenne est un problème récurrent (*Blayo et al.*, 1994) pour l'assimilation de données altimétriques. Il est habituellement contourné de deux façons. La première approche consiste à s'appuyer sur le niveau moyen du modèle (la topographie dynamique moyenne du modèle est supposée parfaite et sert à référencer les anomalies de surface libre observées). Cette approche a notamment été utilisée par *Verron et al.* (1999), *Gourdeau et al.* (2000) et *Parent et al.* (2003) dans le Pacifique Tropical. Dans cette région, les OGCMs sont relativement performants. Le niveau moyen simulé est donc assez proche de la réalité et l'approximation peut donc se justifier. La deuxième approche s'appuie sur des données d'observation (en général in-situ) afin d'évaluées après un traitement plus ou moins élaboré une topographie dynamique moyenne synthétique (*Mercier*, 1986; *Le Grand*, 1998; *Niiler et al.*, 2003). L'utilisation d'une MDT synthétique en lieu et place de la MDT modèle permet de reconstruire un signal altimétrique absolu et ainsi de pouvoir contrôler l'écoulement moyen du

modèle en plus de sa variabilité. Dans les régions où la circulation simulée par les OGCMs présente des erreurs marquées, on peut difficilement continuer à faire l'hypothèse que la surface moyenne simulée est parfaite. Cette deuxième approche est, par exemple, habituellement utilisée dans l'atlantique Nord (*Testut et al.*, 2003; *Brankart et al.*, 2003; *Biol et al.*, 2004) afin de pouvoir corriger la position moyenne de la veine de courant du Gulf Stream qui reste généralement trop collée au talus continental dans les OGCMs actuels, au lieu de se détacher et continuer vers l'Est à partir du Cap Hatteras (*Testut et al.*, 2003; *Brankart et al.*, 2003; *Biol et al.*, 2004).

Aucune des solutions n'est vraiment satisfaisante et le problème de la définition d'une topographie dynamique moyenne pour l'assimilation de la SLA observée par les satellites altimétriques reste d'actualité. *Biol et al.* (2004), ont illustré la sensibilité de l'assimilation de données à la MDT utilisée en utilisant un filtre SEEK et le modèle HYCOM (HYbrid Coordinate Ocean Model) proposé par *Bleck* (2002) dans sa configuration Atlantique Nord 1/3°. Cet article, dont je suis co-auteur, a été publié dans *Marine Geodesy*. Il est fourni en annexe A de ce document. Bien que basé sur le SEEK, le système d'assimilation mis en place pour cette étude utilise la méthode proposée par *Cooper et Haines* (1996) basée sur la conservation de la vorticité potentielle sur les isopycnes en profondeur, pour passer l'incrément d'analyse en SSH au modèle, et non la méthode statistique classique du filtre SEEK. Les trois expériences, utilisant trois MDTs différentes pour référencer la SLA T/P+ERS, montrent d'importantes différences en terme de circulation océanique, de transport et de champs thermo-halin simulés. *Biol et al.* (2004) en concluent que les erreurs sur l'estimation de la MDT font parties des facteurs ayant le plus fort impact sur la qualité de la circulation océanique simulée avec assimilation de données altimétriques.

Les données gravimétriques, dont la précision et la résolution n'ont cessé de s'améliorer (cf chapitre 2) sous l'impulsion des missions satellites dédiées présentes (CHAMP et GRACE) et à venir (GOCE), nous donnent aujourd'hui accès à une MDT satellite observée plus fidèle à la réalité de l'océan mondial. Ce type de MDT observée, utilisée pour référencer les produits de SLA, produits déjà très largement utilisés avec les succès que l'on connaît dans la cadre d'expériences d'assimilation de données pour contrôler la variabilité des OGCMs, va maintenant nous donner accès à un signal de topographie dynamique absolue observé et précis. L'utilisation de ce signal absolu dans la cadre d'expériences d'assimilation de données va ainsi nous permettre de contrôler, en plus de la variabilité, la circulation moyenne des OGCMs. Ceci représente le premier objectif de l'assimilation de DT : l'amélioration de la structure moyenne de l'écoulement simulé.

Un autre intérêt d'un signal altimétrique absolu pour l'assimilateur océanographe est la perspective d'une assimilation conjointe multi-données (comme par exemple l'altimétrie et les profils de température TAO) plus efficace. En effet, compte tenu de la nature différente des données et de leurs erreurs respectives (notamment les erreurs sur le géoïde qui se traduisent par une erreur sur la moyenne du signal de DT), il n'est pas évident que ce type d'assimilation conjointe fonctionne correctement. Les difficultés inhérentes à ce type d'expérience sont notamment discutées par *Parent et al.* (2003). Ces derniers ont montré que la réalisation d'expériences assimilant de l'altimétrie et des profils de température in-situ TAO pouvait conduire à une impasse à cause de la différence d'état moyen entre les deux types de données. *Parent et al.* (2003) ont réalisé deux expériences assimilées, une n'assimilant que la SLA et l'autre n'assimilant que les profils de température TAO. L'information sur le contenu thermique moyen de l'océan présent dans les profils de température TAO s'est traduit pas une forte modification de la surface moyenne du modèle. Une er-

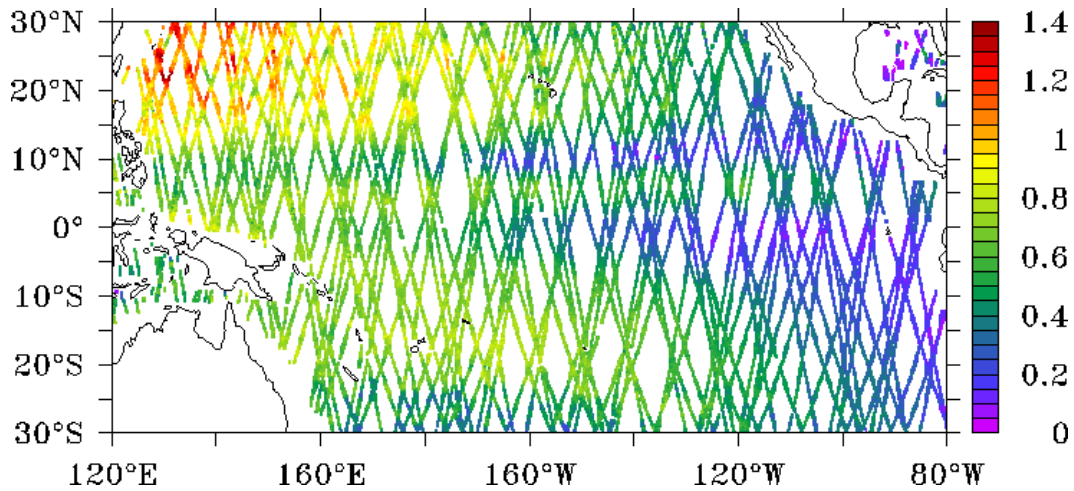


FIG. 5.1 – Topographie dynamique absolue le long des traces pour l'analyse du 22 septembre 1995. La fenêtre temporelle d'observation est de 5 jours.

reur de plus de 5 cm RMS est ainsi diagnostiquée entre les deux MDT simulées. Ceci a conduit *Parent et al. (2003)* à conclure que l'assimilation conjointe d'altimétrie et de profils de température TAO, bien que très prometteuse, n'était envisageable que moyennant la connaissance d'une MDT suffisamment précise afin d'assurer la compatibilité, entre eux, des deux jeux de données d'observation. Sur ce dernier point, les données gravimétriques, et les données GRACE en particulier, ont également un rôle à jouer en fournissant une MDT plus réaliste et donc une meilleure compatibilité avec les autres sources d'observation de l'océan, comme les mouillages TAO. Ceci représente le deuxième objectif de l'assimilation de DT : l'amélioration de la compatibilité entre l'altimétrie et les autres sources d'observation de l'océan.

5.1.2 Les différentes configurations du système d'observation

Afin d'estimer l'impact de l'utilisation d'une MDT absolue sur l'assimilation de données, nous avons retenu différentes configurations pour le système d'observation.

- **Assimilation de la topographie dynamique absolue uniquement.** Une première configuration du système d'observation consiste à ne considérer que la DT. Cette configuration permettra de valider notre système d'assimilation de données et de évaluer l'impact de l'assimilation de la DT par rapport à l'assimilation des seules anomalies de SLA. La DT comme le montre la figure 5.1 présente une très bonne couverture spatio-temporelle et est globale.

Cependant, l'information est par définition une information de surface et il est vraisemblable, compte tenu du degré de liberté du système que l'extrapolation vers le subsurface par la seule statistique du filtre SEEK ne permette pas de contrôler efficacement les couches intermédiaires.

- **Assimilation des profils de température in-situ TAO uniquement.** On considère ici que la seule source d'observation de l'océan est constituée de profils de température in-situ TAO. La nature des observations TAO est très différente de celle de la DT. Le réseau TAO a été conçu afin d'étudier le phénomène ENSO qui

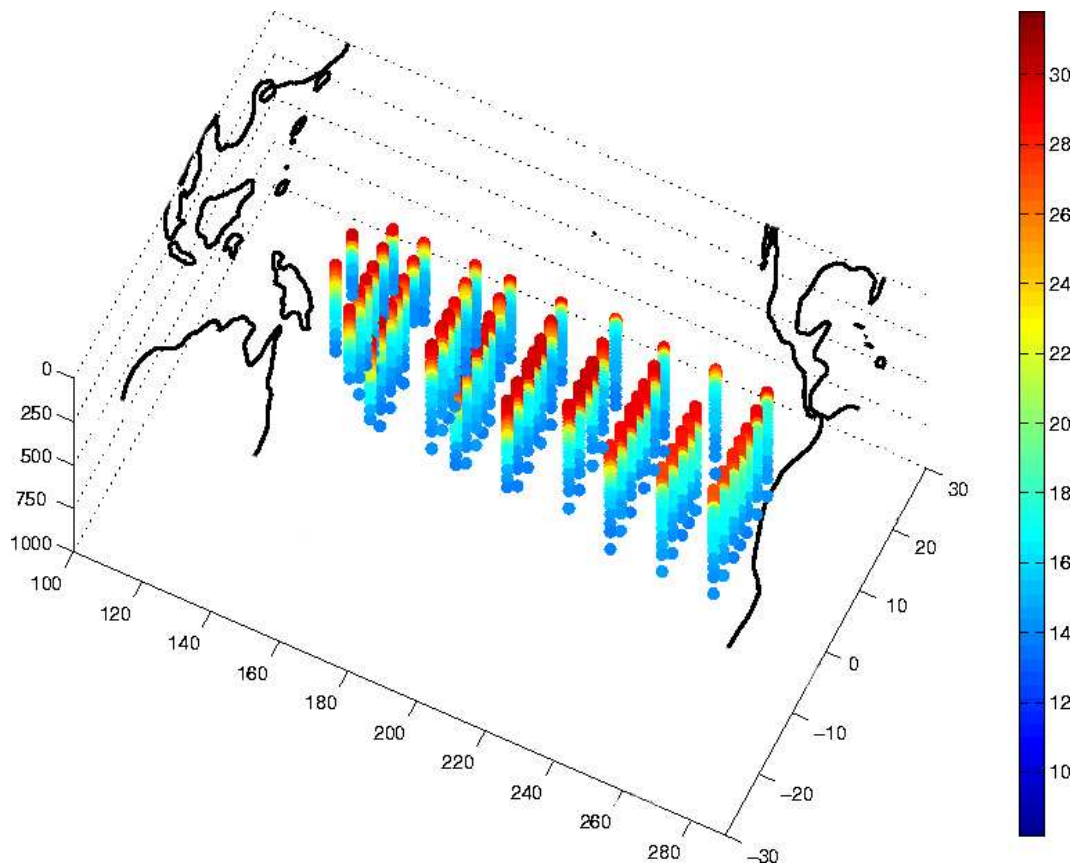


FIG. 5.2 – Profils *in-situ* de température TAO pour l'analyse du 22 septembre 1995. La fenêtre temporelle d'observation est de 5 jours.

est contrôlé, en partie, par la propagation d'ondes équatoriales à grande échelle qui redistribuent le contenu de chaleur de la couche de mélange dans le bassin de l'océan Pacifique Tropical. La configuration du réseau TAO a donc été optimisée pour la mesure de ce type d'onde, et non pour le contrôle d'un OGCM par assimilation de données.

Comme on le voit sur la figure 5.2, on a une observation qui est cette fois-ci 3D, mais qui présente une couverture spatiale à grande échelle de la température dans la couche supérieure de l'océan. De plus, les observations se cantonnent au rail équatorial (i.e. entre 8°S et 8°N) et ne couvrent donc qu'une partie de notre domaine d'étude.

- **Assimilation conjointe de la DT absolue et des profils de température *in-situ* TAO.** Dans cette dernière configuration, la DT et les profils de température *in-situ* TAO vont être assimilés conjointement dans le modèle. Cette configuration est a priori la plus intéressante. D'une part, la quantité d'information disponible est plus importante. D'autre part ces deux jeux de données sont complémentaires. Les données altimétriques présentent une couverture globale et répétitive mais ne fournissent pas une vision tridimensionnelle du milieu. Les profils de température donnent quant à eux accès à la structure de subsurface

mais elle reste largement sous-échantillonnée, même dans le Pacifique Tropical où l'on dispose des données du réseau TAO. Il est donc naturel de considérer ces deux jeux de données conjointement pour contraindre plus efficacement le modèle. Ce type d'assimilation multi-données complique cependant fortement la tâche de l'assimilateur et introduit notamment le problème de la compatibilité des données que je viens de discuter. Au final, il n'est pas évident que cette configuration soit apte à efficacement contraindre le modèle dans un cas réaliste. C'est notamment un des enjeux de l'utilisation de la MDT GRACE, qui en nous fournissant un MDT observée et plus réaliste devrait permettre de résoudre, du moins en partie, les problèmes d'incompatibilité des données précédemment constatés.

5.1.3 Les erreurs d'observation

L'erreur de d'observation, de covariance \mathbf{R} , est la somme de plusieurs contributions caractérisant l'écart entre les données d'observation et l'équivalent de la vérité modèle dans l'espace d'observation. La première contribution est l'erreur de mesure (de covariance \mathbf{E}). La deuxième est la partie du signal vrai qui n'est pas représentable par le modèle (à cause de la résolution de la grille, ou de phénomènes non résolus par le modèle). C'est l'erreur de représentativité (de covariance \mathbf{F}). Dans un contexte de réduction d'ordre, il existe une troisième contribution qui est liée à la fraction de l'erreur qui n'est pas incluse dans l'espace d'erreur en rang réduit. C'est l'erreur de troncature et elle doit être réduite autant que faire se peut en choisissant un espace d'erreur en rang réduit approprié.

Fournir une estimation réaliste de \mathbf{R} est loin d'être une chose évidente. L'erreur de représentativité est par exemple très difficile à diagnostiquer. Pour autant, tous les schémas d'assimilation avancés nécessitent, au même titre qu'une paramétrisation de l'erreur d'ébauche, une paramétrisation adéquate de l'erreur d'observation, afin d'estimer un état analysé combinant de manière optimale un modèle et des observations. Ces matrices définissent en particulier le poids respectif qui est donné au modèle et aux observations au moment de l'analyse et conditionnent donc fortement le résultat. Pour autant, alors qu'un certain nombre d'études ont été menées sur la paramétrisation de l'erreur d'ébauche, peu d'études ont été dédiées à une estimation fine de l'erreur d'observation. La matrice \mathbf{R} ainsi que la façon de la diagnostiquer dépendent : (i) du système d'observation (par l'intermédiaire de la matrice \mathbf{E}), (ii) de de la configuration modèle (par l'intermédiaire de la matrice \mathbf{F}), et (iii) de la méthode d'assimilation (par l'intermédiaire de l'erreur de troncature) et sont donc étroitement liées à la plate-forme d'assimilation utilisée. *Fukumori et al.* (1999) ont proposé une technique pour évaluer les caractéristiques statistiques des erreurs sur les SLA. Plus récemment, Fabien Durand a travaillé sur l'estimation de la covariance des erreurs d'observation \mathbf{R} pour les données in-situ du réseau TAO. Ce travail a fait l'objet d'un poster à la conférence internationale "Dynamic Planet 2005" à Cairns en Australie et un papier est en cours de préparation. Ces études représentent un point de départ intéressant pour mettre en place une paramétrisation de la matrice \mathbf{R} plus réaliste que la simple matrice diagonale généralement utilisée dans les expériences d'assimilation de données. Cependant, l'estimation et l'utilisation d'une telle matrice \mathbf{R} dans le cadre d'expériences d'assimilation de données réalistes reste un travail de grande ampleur et sort largement des objectifs et du cadre de cette thèse.

Au final, l'erreur d'observation est paramétrée de façon très simple pour ce travail. La matrice de covariance d'erreur d'observation \mathbf{R} utilisé pour les expériences avec assimila-

tion de données qui suivent est une matrice diagonale avec, respectivement 5 cm RMS et 0.4 °C RMS d'erreur pour l'altimétrie et les profils de température in-situ TAO.

5.1.4 L'état initial \mathbf{x}_0

Le vecteur d'état utilisé pour initialiser les simulations est commun à l'ensemble des expériences présentées dans la suite du document. Il est issu d'une simulation relaxée vers les champs climatologiques de température et de salinité.

En effet, on observe dans tous les OGCMs actuels une dérive de l'état moyen du modèle au cours du temps. Si l'on reprend les notations introduites dans le chapitre 4 et que l'on cherche à caractériser l'évaluation temporelle de l'erreur ϵ_i d'une simulation libre en fonction de l'erreur au début du cycle i et de l'erreur modèle, on montre sans peine, en faisant l'hypothèse que le modèle $M_{i,i+1}$ peut être approché par le modèle linéaire tangent $\mathbf{M}_{i,i+1}$ sur le cycle i , que l'erreur ϵ_i à la fin de chaque cycle est régie par l'équation :

$$\epsilon_i = \mathbf{M}_{i-1,i}\epsilon_{i-1} + \eta_i \quad (5.1)$$

Physiquement, l'erreur à la fin du cycle i correspond bien à la propagation par l'opérateur modèle de l'erreur existant à la fin du cycle précédent, à laquelle il faut ajouter l'erreur générée par le modèle durant le cycle. Si l'on exprime maintenant l'erreur au cycle i en fonction de l'erreur sur la condition initiale ($i = 0$) et des erreurs du modèle pour les cycles successifs, il vient :

$$\epsilon_i = \eta_i + \mathbf{M}_{i-2,i-i}\eta_{i-1} + \dots + \mathbf{M}_{1,2}\eta_1 + \mathbf{M}_{0,1}\eta_0 \quad (5.2)$$

Cette formule montre que l'erreur sur la simulation libre est le résultat de la propagation par le modèle de l'erreur initiale, ainsi que des erreurs du modèle qui se sont insérées dans le système au fil du temps. Rien n'empêche donc que la moyenne temporelle de l'erreur de la simulation libre soit non nulle (même si $\langle \eta_i \rangle = 0 \quad \forall i$) :

$$\bar{\epsilon} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \epsilon_i \neq 0 \quad (5.3)$$

Afin de faciliter la tâche de l'assimilation de données, plutôt que d'utiliser une condition initiale provenant d'une simulation libre qui a pu s'écarter fortement de la réalité, nous avons choisi d'utiliser une condition initiale issue d'une simulation utilisant une relaxation sur toute la colonne d'eau vers les champs climatologiques mensuels interpolés journalièrement de température et de salinité proposé par *Levitus* (1998).

5.1.5 Durée du cycle d'assimilation

Le choix de la durée du cycle d'assimilation est forcément le résultat d'un compromis. D'un côté, du fait de la mauvaise connaissance des paramétrisations statistiques, il faut une quantité d'observations suffisante pour que l'analyse SEEK soit efficace. Par conséquent, il faut un cycle d'assimilation avec une durée suffisamment longue de manière à collecter assez d'observations de l'état du système océan. D'un autre côté, en allongeant la durée du cycle d'assimilation, on s'expose à une dérive de l'erreur lors de la prévision. De plus, étant donné que l'on fait l'approximation que toutes les observations collectées au cours d'un cycle d'assimilation sont disponibles à l'instant de l'analyse, un cycle d'assimilation long

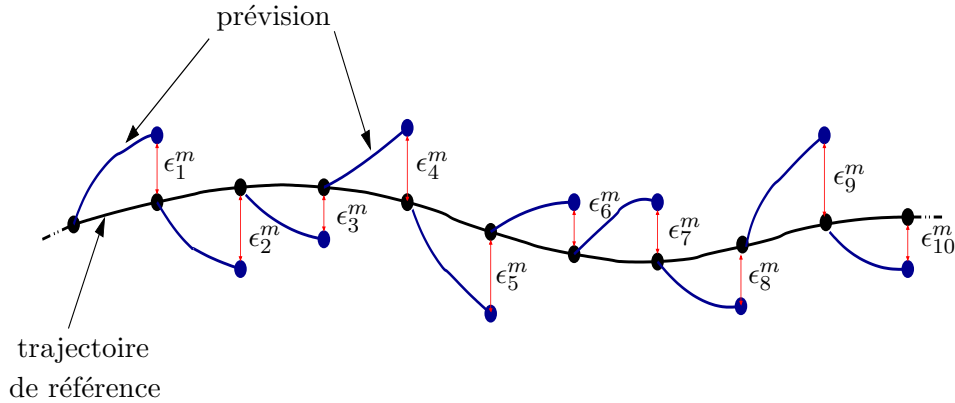


FIG. 5.3 – Estimation de l'erreur à l'échéance de la prévision.

impliquerait un décalage temporel des observations par rapport à leur instant d'acquisition potentiellement important. En pratique, nous avons opté pour une durée du cycle d'assimilation de 5 jours, durée qui donne des résultats satisfaisants et permet un bon échantillonnage de l'océan par notre système d'observation (cf figures 5.1 et 5.2).

5.2 Paramétrisation de l'espace d'erreur en rang réduit

5.2.1 Positionnement du problème

La méthode introduite dans la section 4.4.2 pour initialiser le filtre SEEK s'avère souvent inadaptée, particulièrement dans le cas d'expérience réaliste. La statistique d'erreur sur la condition initiale est en effet correctement estimée, dans le cadre des hypothèses présentées dans la section 4.4.2, mais l'erreur sur la condition initiale, du fait que l'on utilise l'état moyen comme condition initiale, est très importante. Il est souvent préférable d'utiliser un état modèle relatif à l'état du jour afin de faciliter la tâche de l'assimilation. Dans ce cas, les 4 approximations suivantes sont faites : (i) la covariance de la variabilité du système peut être utilisée comme un proxy de la covariance d'erreur initiale, (ii) la statistique de la variabilité du modèle est représentative de la statistique de la variabilité du système, (iii) un échantillonnage d'états instantanés du modèle permet d'estimer la variabilité du modèle, et (iv) la covariance de cet ensemble peut être approchée par les r premières EOFs d'une décomposition EOF de cet ensemble. En utilisant cette méthode, on obtient une matrice de covariance de rang r qui est utilisée pour paramétrer la matrice de covariance d'erreur initiale. L'hypothèse sous-jacente reste que la moyenne de la simulation libre est représentative de la moyenne de l'océan vrai. Comme je l'ai déjà dit à plusieurs reprises, cette hypothèse forte n'est généralement pas vérifiée. L'intérêt de l'assimilation d'une topographie dynamique absolue étant de contraindre l'état moyen du modèle, en plus de sa variabilité, nous avons été amenés à développer un protocole spécifique pour paramétrer la matrice de covariance d'erreur initiale \mathbf{P}_0 .

A ce stade, il est important de se poser la question de la nature de l'erreur que l'on cherche à contraindre par l'assimilation. C'est en effet en répondant à cette question que l'on pourra déterminer une paramétrisation du SEEK adaptée, et donc une paramétrisation qui permettra d'utiliser l'information contenue dans les observations de manière efficace. Mon problème consiste à contrôler la dérive du champ de masse du

modèle. Cette dérive est due à une accumulation d'erreur modèle (cf section 5.1.4) et se traduit par l'apparition d'une différence entre l'état moyen observé et l'état moyen simulé par le modèle. On cherche donc à évaluer l'erreur modèle à l'échéance de la prévision. Dans le cadre d'expériences cousines (i.e. expériences jumelles dans lesquelles la source d'erreur est une erreur modèle et non une erreur sur la condition initiale) non montrées ici, j'ai pu confirmer que la variabilité de la simulation de référence n'était pas adaptée pour paramétrer le filtre SEEK afin de corriger ce type d'erreur, l'analyse dégradant systématiquement les résultats en terme d'écart RMS, en comparaison de la prévision. En revanche, une décomposition EOFs des différences entre la simulation de référence et la simulation perturbée à l'échéance de la prévision est apparue être une manière adéquate de paramétrer le SEEK dans ce cadre expérimental.

5.2.2 La paramétrisation retenue

Au lieu d'utiliser une décomposition EOFs de la variabilité modèle pour évaluer \mathbf{P}_0 , comme cela est couramment fait avec le filtre SEEK, nous avons utilisé une procédure d'ensemble pour paramétrer l'espace d'erreur en rang réduit. Pour cela, nous avons supposé qu'une simulation libre (i.e. sans assimilation) avec une forte relaxation vers le climatologie était à même de fournir une bonne approximation de la circulation moyenne de l'océan réel. Dans cette simulation de référence, un terme de relaxation newtonienne est ajouté dans les équations pronostiques de la température et la salinité :

$$\frac{\partial T}{\partial t} = \dots + \gamma(T - T_0) \quad (5.4)$$

$$\frac{\partial S}{\partial t} = \dots + \gamma(S - S_0) \quad (5.5)$$

où γ est l'inverse d'une échelle de temps et T_0 et S_0 sont les champs de température et de salinité observés. Nous avons utilisé une valeur forte pour γ (équivalente à une échelle de temps de 1 jour) et les champs climatologiques mensuels proposés par *Levitus* (1998) interpolés journalièrement pour spécifier T_0 et S_0 . Le nudging est une des fonctionnalités proposées par le modèle OPA et n'a donc pas nécessité de développement supplémentaire. En plus de la relaxation sur les champs T-S, nous avons aussi relaxé les courants de surface vers les courants géostrophiques déduits de la MDT GRACE. Le principe est le même que pour la température et la salinité :

$$\frac{\partial u}{\partial t} = \dots + \gamma(u - u_0) \quad (5.6)$$

$$\frac{\partial v}{\partial t} = \dots + \gamma(v - v_0) \quad (5.7)$$

u et v étant les composantes zonales et méridiennes du vecteur vitesse horizontal. Il n'est, en revanche, pas prévu dans le code OPA de pouvoir relaxer le modèle vers un champ de vitesse observé et cela a donc nécessité un certain nombre d'aménagements dans le code.

Une simulation forcée à l'aide d'un forçage climatologique annuel calculé à partir des forçages inter-annuels sur la période 1993-1998 et relaxée comme expliqué ci-dessus est réalisée. La trajectoire de référence ainsi obtenue est considérée comme étant représentative du cycle saisonnier moyen du Pacifique Tropical. Cette trajectoire de référence est utilisée pour réinitialiser le modèle libre tous les 5 jours (i. e. la durée de notre fenêtre temporelle

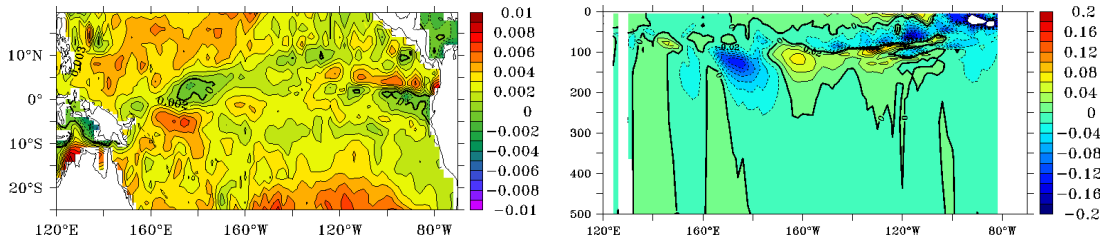


FIG. 5.4 – Premier mode 3D multivarié de la covariance d’erreur initiale \mathbf{P}_0 pour la SSH (à gauche en m), la température à l’équateur (à droite en °C).

d’assimilation) pendant 1 an (cf figure 5.3). On génère ainsi un ensemble de différences entre la simulation de référence et la prévision à 5 jours du modèle libre (i. e. prévision du modèle libre car sans assimilation ni nudging) $\epsilon_1^m, \epsilon_2^m, \epsilon_3^m, \dots, \epsilon_n^m$. Cette ensemble de 73 membres fournit une estimation de la statistique de l’erreur de prévision à 5 jours entre le modèle libre et la simulation de référence qui, dans le cadre des hypothèses mises en place, est une approximation du cycle saisonnier moyen de l’océan vrai. Au final on a donc une estimation de la statistique de l’erreur de prévision du modèle à l’échéance de l’étape de prévision de la séquence d’assimilation.

La condition initiale utilisée pour nos simulations avec assimilation de données étant un état modèle issu de la simulation de référence ci-dessus (cf section 5.1.4), la covariance de cette ensemble paraît appropriée pour estimer la matrice de covariance des erreurs initiales \mathbf{P}_0 , nécessaire à la paramétrisation du filtre SEEK. Le rang de cette matrice est réduit en utilisant un nombre limité (30 dans ce cas) d’EOFs 3D multivariées décrivant les modes dominants de la covariance de l’ensemble. Le nombre de 30 EOFs peut paraître élevé comparé au nombre de modes classiquement retenus pour paramétrer le SEEK (typiquement une dizaine de modes). Cependant, la structure de l’erreur ainsi estimée paraît plus complexe que la structure de la variabilité de la simulation libre qui est dominée par les quelques premiers modes. Avec 10 modes, on ne capture ici que 53% de la variance totale du signal. L’utilisation des 30 premiers modes permet en revanche de représenter plus de 85% de la variance du signal total. La figure 5.4 illustre les structures présentes dans le premier mode d’erreur 3D multivarié obtenu avec cette paramétrisation.

5.3 Récapitulatif des expériences réalisées

Afin de répondre aux objectifs de la problématique de cette thèse, un certain nombre d’expériences réalistes d’assimilation de données ont été réalisées. Elles utilisent différentes configuration de système d’observation (données altimétriques seule, données in-situ seule, données altimétriques et in-situ conjointement). Elle seront discutées dans les chapitres suivants. Dans un premier temps, nous évaluerons les apports d’une topographie dynamique moyenne observée à l’assimilation de données altimétriques T/P+ERS (chapitre 6). Nous nous intéresserons ensuite au rôle important de la topographie dynamique moyenne GRACE pour la complémentarité des données altimétriques (T/P+ERS) et des données in-situ (TAO) (chapitre 7). Nous terminerons en analysant l’impact de l’assimilation de données sur de la dynamique du Pacifique Tropical simulée sur la période 1993-1998 couvrant l’El Niño le plus intense du $XX^{\text{ème}}$ siècle (chapitre 8 et 9).

Dans une perspective de clarté, nous avons regroupé les différentes simulations réalisées

dans le tableau 5.1 suivant :

TAB. 5.1 – *Les différentes expériences réalisées.*

EXP1	SLA + MDT modèle
EXP2	SLA + MDT GRACE
EXP3	TOA
EXP4	SLA + MDT GRACE + TAO
EXP1-GRL ^a	SLA + MDT modèle + TAO
EXP2-GRL ^a	SLA + MDT GRACE + TAO

^a modèle forcé en flux spécifiés

Chapitre 6

Apports d'une topographie dynamique moyenne observée à l'assimilation de données altimétriques T/P+ERS

Sommaire

6.1	Pertinence de la MDT GRACE dans le Pacifique Tropical . . .	101
6.2	D'une surface moyenne synthétique à une surface moyenne observée	104
6.2.1	Apports de l'assimilation : écarts RMS	105
	Topographie dynamique	105
	Température	105
6.2.2	Apports de l'assimilation : comparaisons aux observations	106
	Niveau de la mer	108
	Courants de surface	109
	Structure thermique verticale	111
6.3	Conclusion	112

Maintenant que les objectifs techniques (c'est à dire la mise en œuvre de notre système d'assimilation) sont atteints, nous allons aborder les objectifs scientifiques de ce travail. L'absence d'une bonne connaissance de la MDT est un problème récurrent (*Blayo et al.*, 1994) qui est contourné en s'appuyant soit sur un modèle numérique (la MDT du modèle est supposée parfaite et sert de référence aux résidus altimétriques), soit sur d'autres sources de données (en général in-situ) après un traitement plus ou moins élaboré (e.g. *Mercier* (1986); *Le Grand* (1998)). Aucune de ces solutions n'est vraiment satisfaisante. Les récentes missions gravimétriques CHAMP (2000) et surtout GRACE (2002) ont permis d'améliorer notre connaissance du géoïde de manière importante (cf section 2 et ainsi d'accéder à une MDT observée présentant une résolution et une précision rendant son utilisation envisageable dans la cadre d'expérience d'assimilation de données. L'utilisation de la MDT observée pour référencer les SLA T/P+ERS devrait enfin permettre d'apporter une réponse satisfaisante aux problèmes de la définition du niveau de référence pour l'assimilation de données altimétriques dans un OGCM. L'objet de ce chapitre est précisément d'estimer l'impact des données gravimétriques et de la MDT observée GRACE sur l'assimilation de données altimétriques dans la région du Pacifique Tropical.

Le comportement du modèle libre dans la région du Pacifique Tropical a été préalablement validé pour des configurations très similaires à celle mise en place pour ce travail (*Vialard et al.* (2001), *Lengaigne et al.* (2003), *Alory et al.* (2005)). Le but de ce travail n'étant pas une validation exhaustive de la simulation libre, je n'y consacrerai pas une section dédiée. D'une manière générale, la simulation libre reproduit bien les grandes structures caractéristiques de la dynamique du Pacifique Tropical en terme de courants de surface, de température et de salinité et les structures simulées sont en adéquation avec les études précédemment citées.

Nous utiliserons dans ce chapitre les configurations expérimentales présentées dans le chapitre précédent (chapitre 5) dédié à la présentation des caractéristiques techniques des expériences d'assimilation, et plus particulièrement, à la présentation de la paramétrisation de l'erreur d'ébauche. Ce chapitre sera d'ailleurs l'occasion de valider les approximations consenties lors de cette étape particulièrement sensible.

Avant de s'intéresser à la validation des simulations avec assimilation de données à proprement dite, je fournirai, dans un premier temps, une rapide validation océanographique de la MDT GRACE. Comme je l'ai discuté dans le chapitre précédent (chapitre 5), le fait que l'on dispose aujourd'hui d'une MDT satellite observée avec une résolution et une précision suffisante pour être utilisée dans le cadre d'expériences d'assimilation est un des points clé de ce travail. Encore faut-il que cette MDT observée reproduise fidèlement, aux erreurs de mesure près, la réalité de l'océan Pacifique Tropical. C'est en effet par le biais de l'amélioration du réalisme de la MDT utilisée pour référencer le signal altimétrique que l'on peut espérer améliorer la circulation moyenne de la simulation avec assimilation de données. L'amélioration de la compatibilité entre l'altimétrie et les observations in-situ passe également par une MDT plus réaliste.

Nous passerons ensuite à la validation des différentes expériences avec assimilation de données. L'évaluation de la qualité d'une simulation assimilée est une tâche complexe qui demande beaucoup de temps. Le but n'est pas ici de fournir une description exhaustive de chaque simulation avec assimilation de données, mais plutôt de donner une idée des performances relatives des configurations utilisées. Pour se faire, nous commencerons par nous assurer de la convergence de la partie observée du vecteur d'état vers les observations. Ce n'est qu'après que nous évaluerons l'impact de l'assimilation sur le reste du vecteur

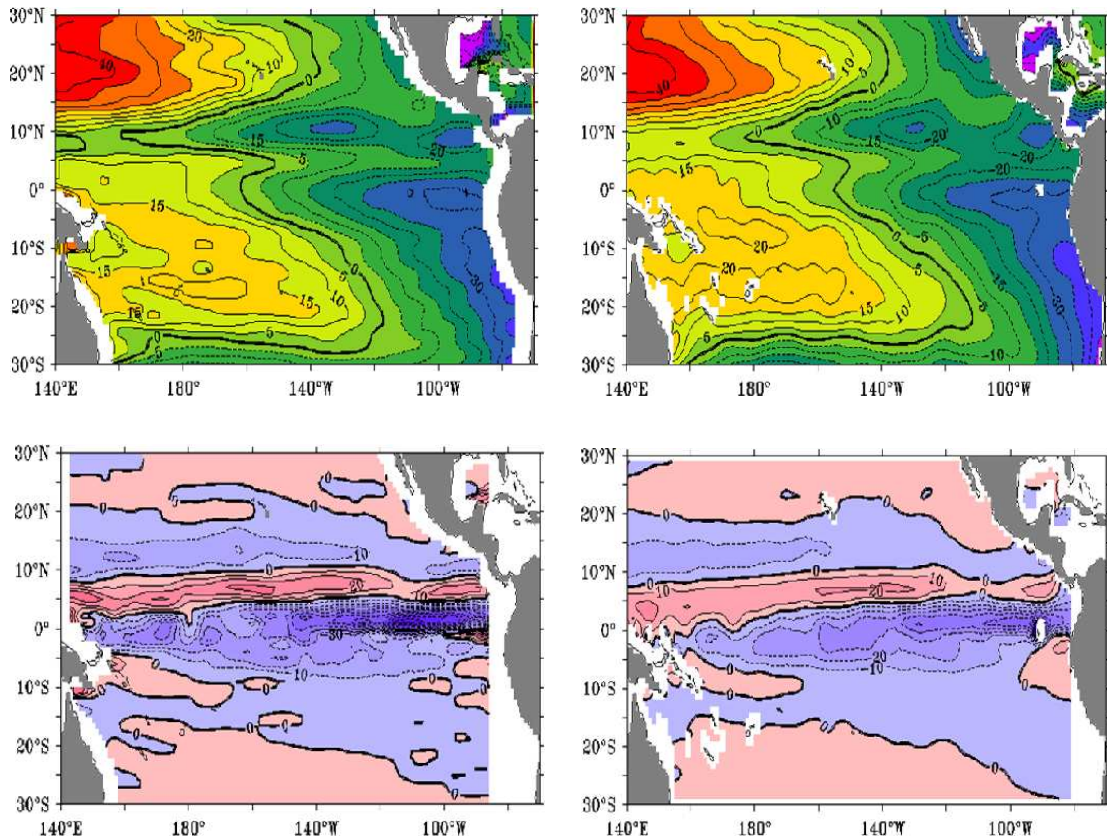


FIG. 6.1 – MDT déduite des données satellite GRACE (à gauche) et de la climatologie Levitus (à droite). Isocontours tous les 5 cm. En bas : vitesse zonale géostrophique moyenne à la surface déduite respectivement de la MDT GRACE (à gauche) et de la MDT Levitus (à droite). Isocontours tous les 10 cm^{-1} .

d'état, et que nous confronterons nos résultats à des données indépendantes.

6.1 Pertinence de la MDT GRACE dans le Pacifique Tropical

En améliorant notre connaissance du géoïde (cf chapitre 2), les données gravimétriques GRACE nous donnent accès à une MDT observée avec une précision et une résolution sans précédent. La MDT utilisée ici est complète jusqu'à l'harmonique 60 (résolution spatiale de 333 km) et présente une erreur cumulée estimée à 4 cm à cette résolution.

Un premier test pour valider, sur un plan océanographique, la solution est de la comparer avec les autres sources d'observations disponibles. Le problème est que l'on ne dispose que de peu d'observations fiables de la topographie dynamique. Une solution pour contourner ce problème est de confronter la MDT GRACE à la climatologie Levitus en terme de topographie dynamique moyenne et de courant géostrophique (Figure 6.1). Il faut cependant garder à l'esprit que la climatologie est très lisse. En particulier, les échelles utilisés pour établir la climatologie de *Levitus* (1998) ne sont probablement pas adaptées au signal méridien à l'équateur. Les deux solutions sont proches pour les grandes échelles. Afin

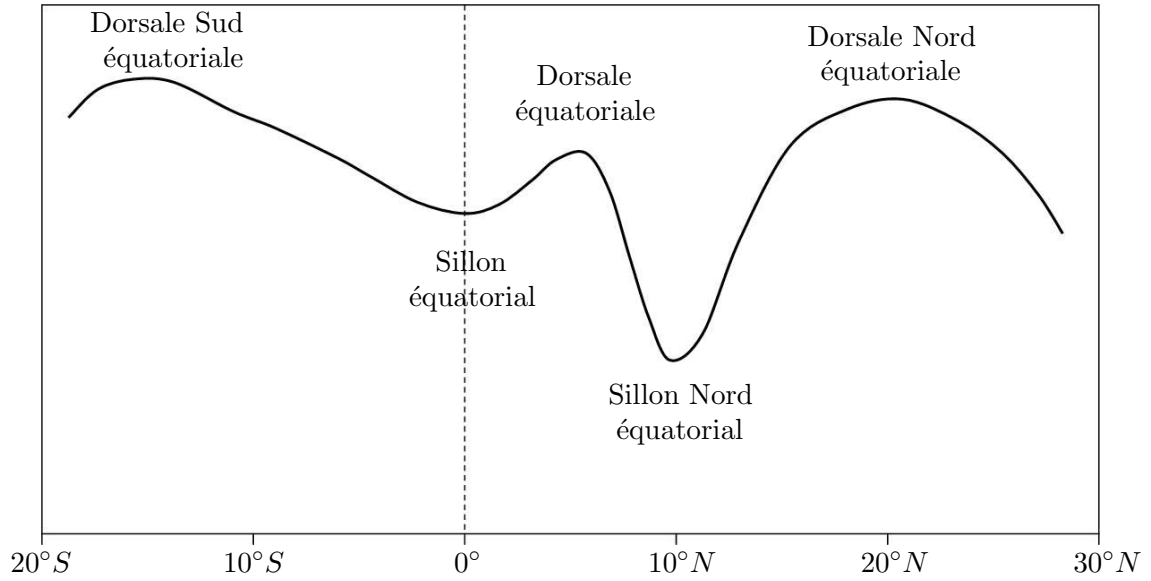


FIG. 6.2 – Représentation schématique d'une coupe méridienne de topographie dynamique.

de faciliter la discussion, nous introduisons les conventions de nom suivantes : la dorsale équatoriale correspond au maximum de topographie dynamique au Nord de l'équateur vers 5°N et le sillon Nord équatorial, au minimum de topographie dynamique vers 10°N (cf figure 6.2). Les gradients méridiens de la MDT GRACE, en particulier dans la région de la dorsale équatoriale et du sillon Nord équatorial sont plus marqués. On note par ailleurs que la MDT GRACE fournit des courants zonaux beaucoup plus énergétiques, ce qui est en adéquation avec les gradients méridiens de MDT plus forts. Le calcul des courants géostrophiques est très sensible au gradient méridional de MDT, particulièrement proche de l'équateur où la force de Coriolis disparaît. L'estimation des courants dans cette région est donc un test rigoureux et révélateur pour valider la MDT GRACE. La carte des courants géostrophiques montre bien l'ensemble des courants caractérisant la dynamique de cette région (cf chapitre 1) avec le NEC, au nord de 10°N, le NECC entre 3°N et 10°N, et le SEC au sud de 3°N. Le SECC est également présent autour de 9°S dans le Pacifique Ouest. L'intensité de l'ensemble de ces courants est en accord avec les valeurs que l'on retrouve dans la littérature et données dans le chapitre 1 ce qui confirme que la MDT déduite de la climatologie est trop lisse alors que la MDT GRACE semble plus réaliste.

Une autre façon de valider la MDT GRACE est de la confronter avec les données in-situ. Dans le Pacifique Tropical, le réseau TAO nous fournit, entre autres, une observation présentant une couverture globale à grande échelle de la hauteur dynamique. La hauteur dynamique est l'intégrale sur la verticale de l'anomalie de volume spécifique :

$$D(p_1, p_2) = \int_{p_1}^{p_2} \delta(T, S, p) dp \quad (6.1)$$

où p_1 et p_2 sont deux pressions de référence, δ l'anomalie de volume spécifique (densité de référence pour une eau à $T = 0^\circ$ et $S = 35$), T la température, S la salinité et P la pression. La hauteur dynamique en 0 et 500 dbar est évaluée en utilisant les profils verticaux de température in-situ TAO et une estimation de la salinité calculée à l'aide

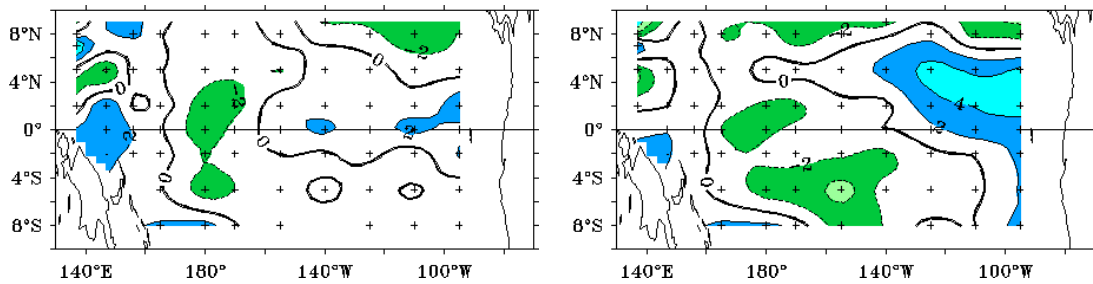


FIG. 6.3 – Différence pour la période 1993-2001 entre la hauteur dynamique moyenne TAO entre 0-500 dbar et, respectivement, la MDT GRACE (à gauche) et MDT déduite de la climatologie Levitus (à droite). Isocontours tous les 2 cm^{-1} .

d'une relation empirique locale T-S déduite à partir de la climatologie. Le calcul de la hauteur dynamique est donc limitée par (i) la référence à 500 dbar et (ii) l'utilisation d'une estimation statistique de la salinité. Au final, l'erreur totale sur l'estimation de la hauteur dynamique est estimée à environ 2-3 cm pour l'ensemble de réseau TAO (*Delcroix et al.*, 1994) et peut atteindre localement 6 cm dans la région spécifique de la Warm Pool (dans cette région le signal de hauteur dynamique est plus significativement affecté par la salinité (*Maes*, 1998)).

La topographie dynamique est calculée en faisant la somme entre les anomalies de hauteur de la mer TOPEX/Poseidon et la MDT GRACE d'une part, et entre les anomalies de hauteur de la mer TOPEX/Poseidon et la MDT Levitus d'autre part. Les DT ainsi obtenues ont été cartographiées sur une grille de $1^\circ \times 1^\circ$ et échantillonnées tous les 5 jours sur la période 1993-2001. La figure 6.3 montre la différence moyenne entre la hauteur dynamique issue des TAO entre 0-500 dbar et respectivement la DT calculée à partir de la MDT GRACE et la DT calculée à partir de la MDT Levitus. La différence moyenne est de 2.4 cm lorsque l'on considère la MDT Levitus et décroît à 1.95 cm lorsque la MDT GRACE est utilisée.

La figure 6.4 montre quant à elle l'écart type de cette différence. On obtient un écart type moyen de 4.6 cm si l'on utilise la MDT Levitus comme référence, et un écart type moyen de 4.3 cm pour la DT calculée à partir de la MDT GRACE. Les plus fortes erreurs sont concentrées au Nord, dans la région de fort cisaillement entre de NECC et le SEC. Dans cette région, les gradients méridionaux de DT sont fort, et il est vraisemblable que la résolution de la MDT GRACE (333 km) est encore trop faible pour les représenter correctement.

En conclusion, la topographie dynamique moyenne satellite et indépendante de toute observation in-situ calculée à partir du géoïde EIGEN-GRACE02S et de la MSSH CLS01 (*Hernandez et al.*, 2001) semble en accord avec la réalité de la dynamique du Pacifique Tropical. Les résultats présentent une nette amélioration par rapport à la climatologie, mais aussi et surtout par rapport aux résultats obtenus avec les données gravimétriques CHAMP (voir *Gourdeau et al.* (2003)). La résolution et la précision de la MDT GRACE et des courants géostrophiques associés sont maintenant compatibles avec son utilisation pour des études océanographiques dans le Pacifique Tropical. En particulier, les erreurs entre les hauteurs dynamiques déduites des mouillages TAO et la topographie dynamique sont réduites lorsque l'on utilise la MDT GRACE pour référencer le signal altimétrique. Cette meilleure compatibilité entre les données altimétriques et les données in-situ (dans

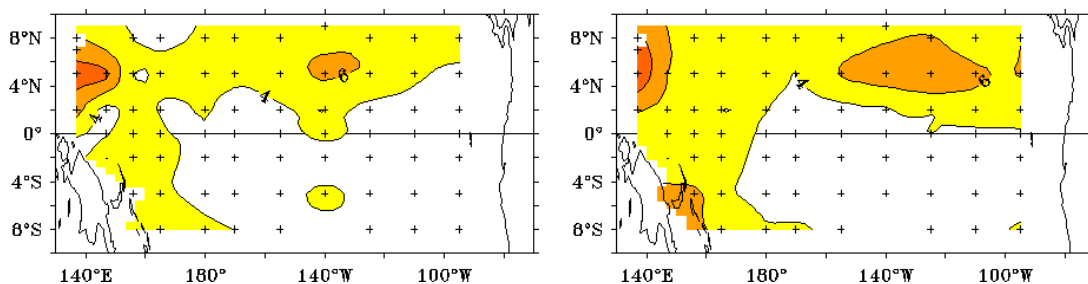


FIG. 6.4 – Différence RMS pour la période 1993-2001 entre la hauteur dynamique TAO entre 0-500 dbar et le signal altimétrique de TOPEX/Poseidon référencé, respectivement, à la MDT GRACE (à gauche) et à la MDT Levitus (à droite). Isocontours tous les 2 cm^{-1} .

ce cas les données TAO) devrait permettre une assimilation plus efficace, en particulier dans la perspective où l'on veut assimiler ces deux jeux de données conjointement. On peut donc espérer tirer un meilleur parti de la complémentarité des données altimétriques et des données in-situ en utilisant la MDT GRACE. L'amélioration du réalisme de la topographie dynamique moyenne devrait également permettre d'efficacement contraindre la circulation moyenne simulée avec assimilation de donnée altimétrique seule.

6.2 D'une surface moyenne synthétique à une surface moyenne observée

Dans cette section, nous allons nous intéresser à l'impact de l'utilisation de la topographie dynamique moyenne observée GRACE en lieu et place de la topographie dynamique moyenne du modèle pour l'assimilation de données altimétriques. L'assimilation de la SLA T/P et/ou T/P+ERS a déjà fait l'objet de nombreuses études (Verron *et al.* (1999); Fukumori *et al.* (1999); Gourdeau *et al.* (2000); Parent *et al.* (2003)). Toutes s'accordent pour montrer une franche amélioration de la variabilité reproduite avec assimilation de données par rapport à la variabilité simulée par le modèle libre. L'assimilation de la topographie dynamique observée (SLA T/P+ERS référencée avec la MDT GRACE) devrait permettre d'améliorer le réalisme de la circulation moyenne simulée avec assimilation de données, en plus de sa variabilité.

Afin d'estimer l'impact de l'utilisation d'une topographie dynamique moyenne observée et réaliste (cf section 6.1) sur l'assimilation de données altimétriques, deux expériences ont été réalisées. Ces deux expériences sont strictement identiques, si ce n'est en ce qui concerne la MDT utilisée. La première expérience (EXP1) simule une situation pré-GRACE et utilise la MDT modèle (comme cela a couramment été fait jusqu'à maintenant) pour référencer l'altimétrie alors que la deuxième utilise la MDT satellite observée. Les deux expériences sont initialisées au 1^{er} janvier 1993 avec la même condition initiale (issue d'une simulation relaxée vers la climatologie afin de limiter l'erreur sur la CI) et s'étendent sur l'année 1993. Le modèle est forcé par le biais de forçage de type bulk (cf chapitre 3) provenant du centre NCEP. Les vents ERS+TAO (Menkes *et al.*, 1998) sont utilisés pour spécifier le flux de quantité de mouvement à l'interface océan/atmosphère. Le SEEK utilise la paramétrisation de l'espace d'erreur en rang réduit basée sur des écarts avec la simulation

de référence relaxée vers la climatologie (cf section 5.2).

6.2.1 Apports de l'assimilation : écarts RMS

Topographie dynamique

Dans un premier temps, nous allons nous assurer de la convergence des variables observées vers les observations assimilées. La figure 6.5 montre l'évolution temporelle de l'écart RMS (RMSD) entre les observations de topographie dynamique, d'une part, et la DT de la simulation libre (en vert) et les deux simulations avec assimilation de données EXP1 et EXP2 (respectivement en cyan et bleu) d'autre part. L'observation de topographie dynamique est définie comme étant la somme entre le MDT GRACE et l'anomalie de SLA T/P+ERS le long des traces. L'observation de DT est donc une observation assimilée dans le cas de l'expérience EXP2 assimilant le signal altimétrique absolu alors qu'elle est en partie indépendante (la partie moyenne du signal) pour l'expérience EXP1 n'assimilant que la SLA. Il faut garder cette idée en tête lorsque l'on compare les courbes de RMSD de la figure 6.5. Je suis conscient du fait que cela pose un problème au niveau de la rigueur de la comparaison des résultats, mais la MDT GRACE étant la topographie dynamique moyenne la plus proche de la réalité du Pacifique Tropical (cf section 6.1) que nous ayons à notre disposition, il me semble difficile de ne pas l'utiliser pour référencer la SLA T/P+ERS afin d'évaluer la topographie dynamique simulée par le modèle. La RMSD moyenne sur la période passe de 7.4 cm pour la simulation libre, à respectivement 7.3 cm pour la simulation EXP1 assimilant la SLA T/P+ERS référencée avec la surface moyenne du modèle et 5.5 cm pour la simulation EXP2 assimilant la SLA T/P+ERS référencée avec la topographie dynamique moyenne GRACE. L'évolution temporelle des deux courbes de RMSD pour l'expérience EXP1 et l'expérience EXP2 présentent des similarités assez nettes. On a l'impression que l'on a décalé la courbe de l'expérience EXP2 de 1.8 cm RMS vers les observations. Ceci indique deux choses : (i) l'erreur sur la variabilité, et donc la variabilité simulée dans les deux expériences avec assimilation de données est proche et (ii) la réduction de la RMSD pour l'expérience EXP2 par rapport à l'expérience EXP1 est la conséquence directe de structure plus réaliste de la topographie dynamique moyenne GRACE par rapport à la MDT modèle.

Température

Un test beaucoup plus pertinent pour estimer la valeur de l'assimilation de données est de comparer les résultats avec des données non assimilées. Par rapport aux deux simulations EXP1 et EXP2, les profils de températures collectés par les mouillages TAO entre la surface et 500 m de profondeur sont des données indépendantes. La figure 6.6 montre l'évolution des écarts RMS par rapport à ces observations indépendantes. La RMSD moyenne est égale à 1.32 °C pour la simulation libre. Elle reste quasi inchangée pour la simulation assimilant la SLA T/P+ERS référencée avec la MDT modèle avec une RMSD de 1.35 °C pour l'expérience EXP1. L'utilisation de la MDT GRACE pour référencer la SLA T/P+ERS permet de réduire cette erreur à 1.15°C dans l'expérience EXP2. On peut faire la même remarque que pour la RMSD par rapport à la DT. L'évolution temporelle de la RMSD par rapport aux données TAO pour les deux expériences assimilées est très similaire. Là encore cela semble indiquer que la variabilité simulée est proche et que c'est la structure moyenne du champ de température qui a été améliorée dans l'expérience EXP2.

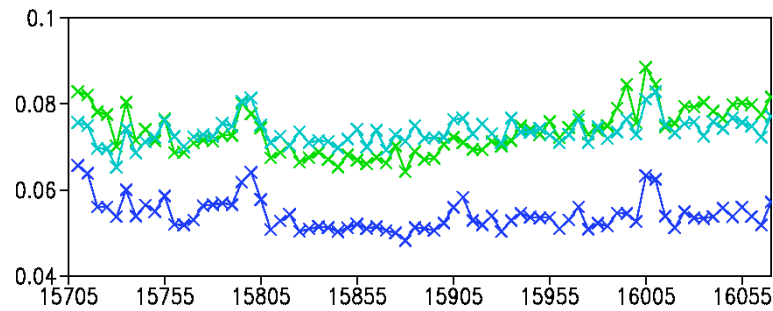


FIG. 6.5 – *Ecart RMS (en cm) par rapport à la topographie dynamique (SLA + MDT) pour la simulation libre (en vert) et les expériences avec assimilation de données EXP1 (en cyan) et EXP2 (en bleu).*

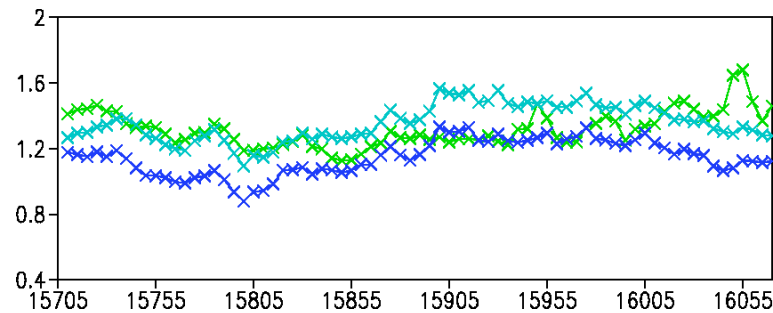


FIG. 6.6 – *Ecart RMS (en °C) par rapport aux profils de température TAO indépendant pour la simulation libre (en vert) et les expériences avec assimilation de données EXP1 (en cyan) et EXP2 (en bleu).*

En conclusion de cette première série d'expériences, on constate que le comportement de l'assimilation en terme d'écart RMS est positif. L'assimilation du signal altimétrique absolu a permis de réduire la RMSD par rapport à la topographie dynamique observée de près de 2 cm RMS. Il semble que la variabilité simulée dans les deux simulations avec assimilation de données soit comparable ce qui est un point positif. Les améliorations entre l'expérience EXP1 et l'expérience EXP2 sont vraisemblablement des améliorations sur la circulation moyenne simulée. On retiendra surtout, le fait que l'on parvient à réduire la RMSD par rapport aux profils de température TAO qui sont des données indépendantes pour ces deux expériences d'assimilations, en utilisant la topographie dynamique moyenne GRACE pour référencer la SLA T/P+ERS. Ce point est un argument fort à mettre au crédit de l'utilisation de la topographie dynamique moyenne GRACE en lieu et place de la MDT modèle.

6.2.2 Apports de l'assimilation : comparaisons aux observations

La section précédente nous a montré que l'assimilation de la SLA T/P+ERS référencée avec la topographie dynamique moyenne observée GRACE permettait de réduire les écarts RMS par rapport aux données assimilées, mais aussi par rapport aux données TAO indépendantes. Les écarts RMS sont importants, dans la mesure où ils permettent de quantifier la réduction de l'erreur par rapport aux observations et de s'assurer de la cohérence

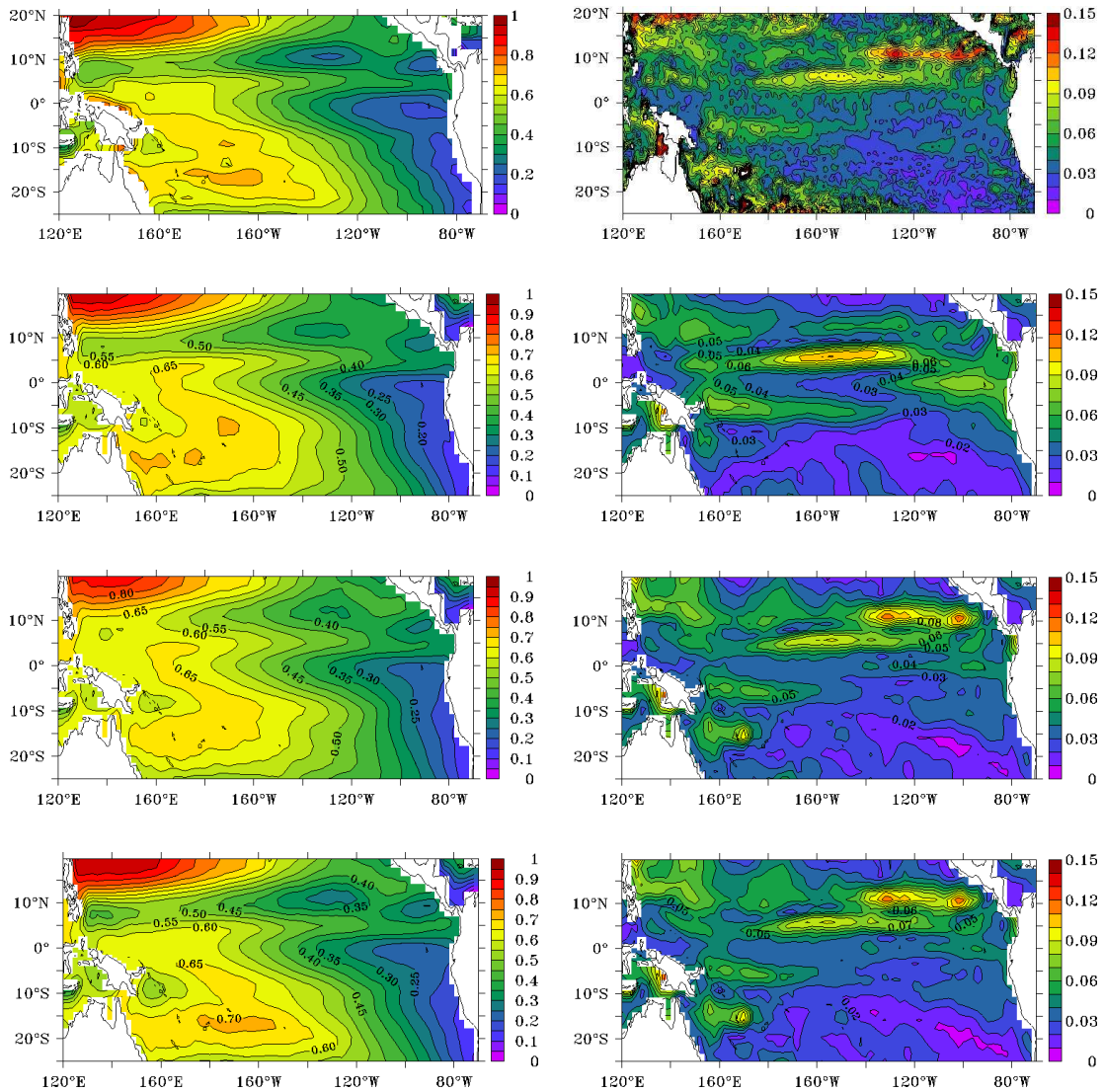


FIG. 6.7 – MDT (à gauche en m) et écart type de la DT (à droite en m) observé, de la simulation en mode libre, de la simulation avec assimilation de SLA EXP1 et de la simulation avec assimilation de DT EXP2, respectivement de haut en bas.

statistique de notre système d'assimilation. Une réduction des écarts RMS par rapport aux observations (surtout par rapport aux observations assimilées) ne garantit par contre en rien que le réalisme de la circulation simulée a été améliorée. Nous allons maintenant regarder l'impact de l'assimilation sur la dynamique simulée. Pour cela, nous confronterons, en terme de structures moyenne et de variabilité, la simulation libre ainsi que les deux expériences avec assimilation de données avec différentes sources d'observations.

Niveau de la mer

En premier lieu, il convient de regarder l'influence de l'assimilation de données sur le niveau moyen et la variabilité du niveau de la mer simulé. Pour simplifier le discours, dans ce qui suit je parlerai d'assimilation de SLA T/P+ERS pour désigner l'assimilation de la SLA T/P+ERS référencée en utilisant la MDT modèle et d'assimilation DT T/P+ERS pour désigner l'assimilation de la SLA T/P+ERS référencée en utilisant la MDT observée GRACE. La figure 6.7 montre la MDT observée et l'écart type du signal altimétrique (en haut). Les MDT et la variabilité de la DT simulées avec assimilation de SLA T/P+ERS et de DT T/P+ERS sont également montrées, respectivement de haut en bas. Le modèle sans assimilation reproduit déjà de manière satisfaisante le niveau moyen et la variabilité du niveau de la mer dans la bande $\pm 15^\circ$ (Vialard *et al.*, 2001). On note l'importante différence entre les structures de MDT des deux simulations avec assimilation de données. L'expérience EXP2 assimilant la DT T/P+ERS est plus proche de la MDT observée. Le sillon Nord équatorial est mieux représenté dans la simulation EXP2, en particulier les gradients méridionaux sont plus en accord avec les observations. Le maximum de MDT dans le Pacifique Ouest au niveau de l'équateur est également plus réaliste. Ce maximum est trop important et trop à l'ouest dans la simulation libre et la simulation EXP1. Concernant la variabilité, en revanche, il y a très peu de différence entre les deux simulations avec assimilation de données. Ceci est un point positif dans la mesure où l'assimilation de SLA permettait déjà d'améliorer de manière significative la variabilité simulée (Parent *et al.*, 2003). Les très faibles différences sur les cartes de variabilité de la topographie dynamique (cf figure 6.7) des deux expériences avec assimilation de données confirment ce que l'on avait déjà constaté en regardant les courbes de RMSD de la figure 6.5. Le fait d'assimiler la SLA T/P+ERS référencée avec le MDT GRACE au lieu de la MDT modèle comme cela est classiquement fait, ne dégrade pas la représentation de la partie variable du signal altimétrique. La variabilité de la DT simulée avec assimilation de donnée est plus proche de la variabilité observée par les satellites altimétriques. Les deux zones de fortes variabilité situées de part et d'autre de l'équateur aux environs de 5° de latitude et associées à la propagation d'ondes de Rossby (Dewitte *et al.*, 2003) sont réduites, en particulier leur extension zonale. Ces ondes modifient la profondeur de la thermocline et ont une forte signature en DT. La simulation libre présente un autre maximum de variabilité centré sur l'équateur dans le bassin Est. Ce maximum est associé avec la propagation d'ondes de Kelvin. Le modèle libre surestime également cette variabilité. L'assimilation permet de corriger cette erreur. La zone de forte variabilité située à 10°N dans les observations est en revanche quasi inexistante dans la simulation libre. Dans cette région, le pompage d'Ekman est assez fort et se traduit par un signal sur la colonne d'eau clairement visible avec l'altimétrie (Kessler, 1990). La variabilité associée au pompage d'Ekman est mieux représentée dans les simulations assimilées, mais elle reste encore sous-estimée par rapport à la variabilité observée.

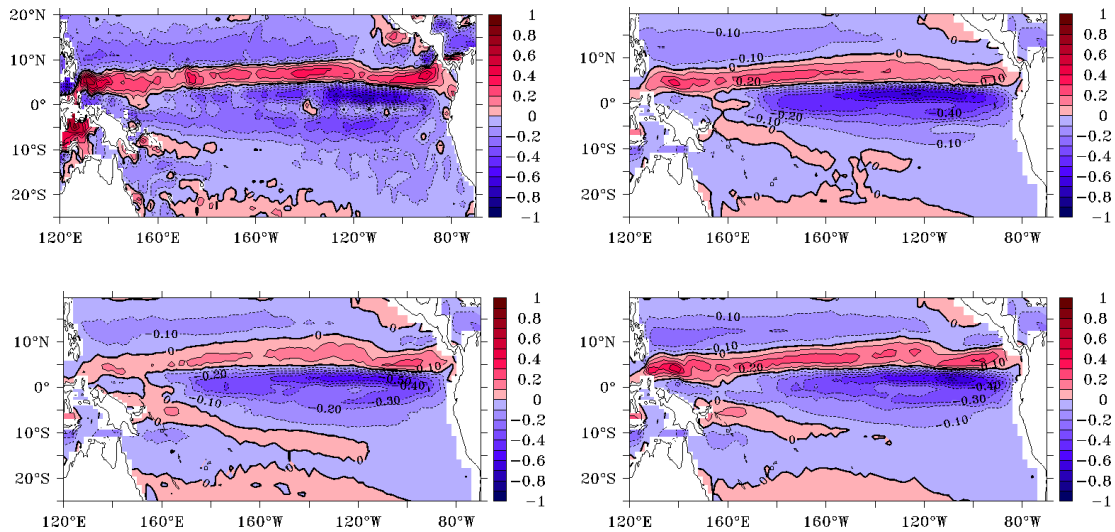


FIG. 6.8 – Vitesse zonale moyenne à 15 m de profondeur (en ms^{-1}). Respectivement de gauche à droite et de haut en bas, la climatologie proposée par Niiler (2001), la moyenne pour l'année 1993 pour la simulation libre, l'EXP1 et l'EXP2.

En résumé, la variabilité du niveau de la mer simulée par les deux simulations avec assimilation de données est très similaire. La variabilité simulée avec assimilation de données est plus en accord avec les observations altimétrique que la variabilité de la simulation libre qui est déjà de bonne qualité. En terme de niveau moyen, l'utilisation de la topographie dynamique moyenne observée permet de sensiblement améliorer le réalisme de la simulation.

Courants de surface

L'étude des courants de surface est particulièrement révélatrice dans les régions tropicales du réalisme des gradients zonaux de topographie dynamique. La figure 6.8 montre la moyenne sur l'année 1993 des courants zonaux de surface à 15 m de profondeur. La carte des courants de surface climatologique proposée par Niiler (2001) est également fournie sur la figure 6.8. Cette climatologie a été obtenue en suivant par satellite la trajectoire des bouées dérivantes SVP (Surface Velocity Program, Niiler (2001)) déployée dans le cadre du programme international WOCE¹ (World Ocean Climat Experiment). Le relevé de la position des bouées nous fournit une observation de la vitesse lagrangienne et donne ainsi accès aux courants de surface (ou plutôt aux courants aux abords de la surface, les bouées dérivantes suivant les courants intégrés sur la profondeur de dérive qui est de l'ordre de 30 m). J'insiste sur le fait que les vitesses pour les simulations libre et assimilées sont également données à 15 m de profondeur pour permettre une comparaison rigoureuse. La première constatation est que la circulation de surface simulée par le modèle libre est de très bonne qualité. Tous les principaux courants caractérisant la circulation du Pacifique Tropicale sont présents. Les courants de surface de l'EXP1 assimilant la SLA T/P+ERS montrent une diminution de l'intensité de NECC dans l'Ouest du bassin ce qui est un point négatif, ce courant étant déjà trop faible dans la simulation libre par rapport à la

¹<http://www.woce.org>

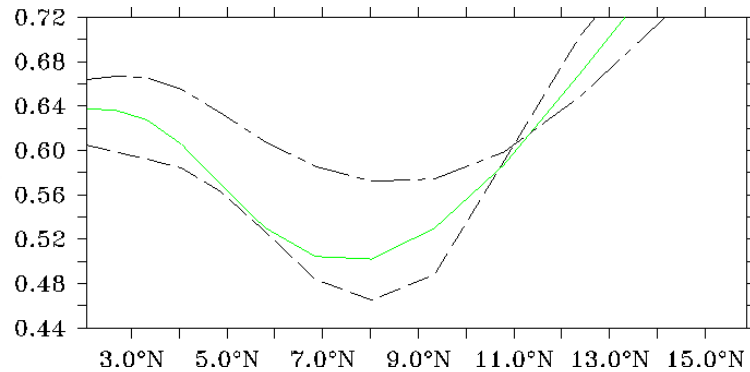


FIG. 6.9 – Section de MDT à $140^{\circ}E$ pour la MDT GRACE (en tiret), la MDT modèle (en tiret mixte) et la MDT de la simulation libre sur l'année 1993 (en vert).

climatologie de Niiler (2001). Cette diminution de l'intensité du NECC est imputable au fait que la dorsale et le sillon Nord équatorial ne sont pas assez marqués dans la simulation EXP1 (cf figure 6.7). Le relativement mauvais comportement de l'EXP1, ou plutôt le relativement très bon comportement de la simulation libre sur ce point s'explique par le fait que toutes les simulations utilisent la même condition initiale au 1^{er} janvier 1993. Or cette condition initiale est issue d'une simulation relaxée vers la climatologie (cf section 5.1.4). Le champ de masse du modèle libre n'ayant pas dérivé de manière significative au cours de la première année de simulation, la DT simulée et notamment les gradients méridionaux de la dénivellation de la surface libre sont proches de ceux observés. Pour une simulation plus longue, les erreurs modèles vont s'accumuler au cours du temps et la structure moyenne de l'océan simulé va s'écarter de la réalité de l'océan Pacifique Tropical. Cette dérive d'état moyen est visible sur le niveau moyen de la mer. Ce point est illustré par la figure 6.9 qui montre une section méridienne de la topographie dynamique moyenne à $140^{\circ}E$. On voit clairement sur la figure que la MDT simulée sur l'année 1993 (en vert) est proche de la MDT observée (en tiret) avec un sillon Nord équatorial bien marqué, alors que la MDT simulée sur la période 1992-1998 (en tiret mixte) qui est la MDT servant de référence pour la SLA T/P+ERS dans la simulation EXP1, fait apparaître un sillon Nord beaucoup moins marqué. Un écart supérieur à 10 cm existe entre la MDT GRACE et la MDT modèle au niveau du sillon Nord équatorial. Les gradients méridionaux moyens de DT moins forts, présents dans la simulation EXP1 dans cette région, se traduisent par des courants zonaux et donc par un NECC moins intense. Ce point illustre bien l'aspect crucial du réalisme de la topographie dynamique moyenne utilisée pour référencer l'altimétrie.

Dans l'EXP2 assimilant la DT T/P+ERS, le NECC est renforcé par rapport à la simulation libre se qui est en accord avec la climatologie de courant observée. En particulier, le démarrage du NECC dans l'extrême Ouest du bassin est mieux simulé. Ce point est à mettre au crédit de la meilleure représentation des structures de la MDT dans cette région grâce à l'utilisation de la MDT GRACE. Dans l'Est du bassin, on observe également un renforcement du NECC ce qui est un autre point positif de l'assimilation. Ce renforcement est visible dans les expériences EXP1 et EXP2. Il est la conséquence de la meilleure représentation de la variabilité du niveau de la mer associée au pompage d'Ekman dans cette région (Kessler, 1990). Cette variabilité est très mal simulée par le modèle libre (cf figure 6.7).

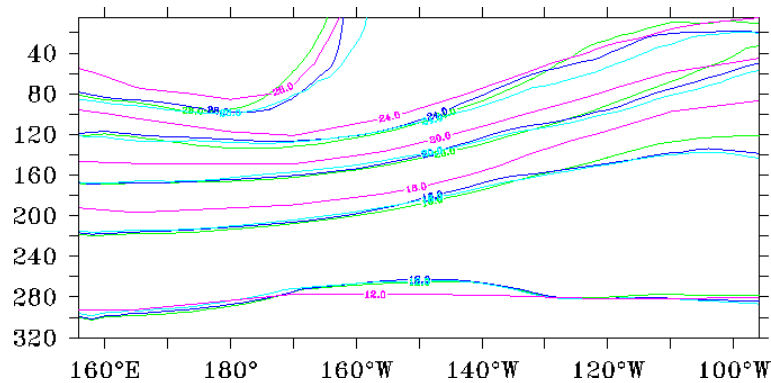


FIG. 6.10 – Profondeur moyenne pour l'année 1993 des isothermes 12, 16, 20, 24 et 28°C le long de l'équateur pour la simulation libre (en vert), les expériences avec assimilation de données EXP1 (en cyan) et EXP2 (en bleu) et les données indépendantes TAO (en magenta).

Un dernier point positif imputable à la meilleure représentation des structures à grande échelle dans l'expérience EXP2, assimilant la SLA T/P+ERS référencée par rapport à la MDT GRACE, est le renforcement du NEC. Les vitesses du NEC sont supérieures à 10 cm^{-1} sur toute la largeur du bassin et supérieures à 20 cm^{-1} dans l'Ouest conformément à *Niiler* (2001). La MDT GRACE est connue pour représenter tous les principaux courants géostrophiques (*Tapley et al.*, 2003).

En résumé, les apports de l'utilisation d'une topographie dynamique moyen sur la structure des courants moyens simulés avec assimilation de données sont importants et bénéfiques. L'intensification du NECC grâce à la meilleure représentation des gradients méridionaux de topographie dynamique, ou encore le renforcement de NEC en sont les principaux exemples.

Structure thermique verticale

Nous allons maintenant regarder comment l'information de surface fournie par l'altimétrie est propagée en profondeur par la statistique du SEEK. La figure 6.10 montre la position moyenne à l'équateur pour l'année 1993 d'un certain nombre d'isothermes pour les trois simulations, libre, EXP1 et EXP2. La position des isothermes déduites des données TAO est également tracée sur la figure 6.10. Je rappelle que les données TAO sont des données indépendantes vis à vis de ces trois simulations. Comme on le voit clairement sur la figure 6.10, il existe une erreur assez marquée entre la position moyenne des isothermes de la simulation libre et celles des observations TAO. Les isothermes sont systématiquement trop profondes dans le Pacifique Ouest avec un écart supérieur à 20 m à 160°E pour les trois isothermes caractérisant la position de la thermocline (les isothermes 16, 20 et 24°C étant habituellement considérées comme étant représentatives de la position du haut, du cœur et du bas de la thermocline respectivement). La pente de la thermocline est également trop forte dans la simulation libre.

L'assimilation de données de la SLA T/P+ERS référencée par rapport à la MDT modèle ou par rapport à la MDT GRACE, n'a pas permis de repositionner les isothermes à la profondeur observée par les TAO. La paramétrisation de l'espace d'erreur utilisée n'est visiblement pas assez fine pour permettre ce type de correction à partir de la seule

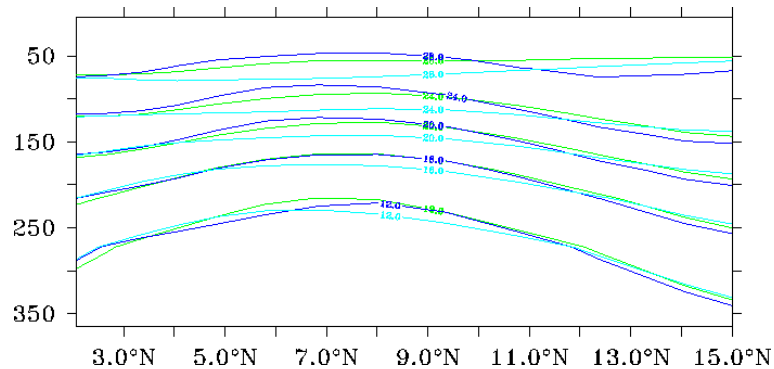


FIG. 6.11 – Section de température moyenne à $140^{\circ}E$ pour la simulation libre (en vert), et les expériences avec assimilation de données EXP1 (en cyan) et EXP2 (en bleu).

information altimétrique. Compte tenu des approximations faites, il semble que nous ayons besoin de plus d'informations sur l'état du système pour arriver à corriger cette erreur.

En revanche, l'assimilation a permis de modifier les pentes des isothermes en accord avec les dénivellations moyennes de la MDT. A l'équateur, la différence du gradient zonal entre la MDT modèle et la MDT satellite est assez faible (ce qui montre encore une fois que le modèle et les vents ERS+TAO sont très bons), et la différence en profondeur sur la pente de la thermocline est limitée et peu visible sur la figure 6.10. La figure 6.11, montrant une section de température méridienne à $140^{\circ}E$, dans une région où nous avons constaté des différences marquées en terme de MDT entre les différentes simulations dans la section précédente, permet d'illustrer ce point de manière plus flagrante.

On note également que la base de la thermocline dans l'Est du bassin (isotherme 16° sur la figure 6.10) est particulièrement mal représentée dans la simulation libre, mais également dans les deux simulations avec assimilation de données. Les upwellings côtiers, le long des côtes américaines, sont mal représentés du fait de la basse résolution zonale du modèle (2°) ce qui explique pourquoi les eaux en subsurface sont trop chaudes dans nos simulations.

En résumé, les apports de l'assimilation et d'une topographie dynamique moyenne observée sont plus mitigés en ce qui concerne la structure thermique simulée sur la verticale. Le point positif est que l'on parvient à corriger les pentes des isothermes en accord avec les dénivellations moyennes de la MDT utilisées pour référencer l'altimétrie. Une MDT plus réaliste se traduit donc par des gradients de températures en profondeur plus réalistes comme on a pu s'en rendre compte sur la figure 6.11. Cette meilleure représentation de la pente des isothermes, grâce à l'utilisation de la topographie dynamique moyenne GRACE, explique la réduction de près de $0.2^{\circ}C$ sur la RMSD par rapport aux observations de température TAO constatée dans la section 6.2.1. Le point négatif est que l'assimilation de données ne permet pas de repositionner les isothermes à la profondeur moyenne observée par les TAO.

6.3 Conclusion

A ce stade, nous avons mené à bien une expérience d'assimilation de données assimilant la topographie dynamique absolue (SLA T/P+ERS référencée avec la topographie dyna-

mique moyenne GRACE) observée de l’océan Pacifique Tropical avec le filtre SEEK. La correction, lors de l’analyse, se fait uniquement par le biais de la statistique du SEEK. Aucune hypothèse supplémentaire n’est rajoutée. Le gain en terme d’écarts RMS est marqué. Il se fait progressivement dans un premier temps (environ 4 mois) et la RMSD par rapport aux observations reste ensuite assez stable avec une valeur proche de 5.5 cm RMS. La simulation libre est déjà de très bonne qualité grâce notamment à l’utilisation des vents ERS+TAO. L’assimilation a permis d’améliorer la topographie dynamique moyenne, avec une meilleure représentation des structures et des gradients à grande échelle, en particulier le long de 10°N où la simulation libre présente des déficiences marquées. Cette amélioration se traduit notamment par des courants plus intenses (en particulier le NEC et NECC). La paramétrisation de l’espace d’erreur en rang réduit, basée sur les anomalies entre le modèle libre et une simulation de référence, semble donc bien adaptée pour l’assimilation de la topographie dynamique absolue observée. Elle reste cependant fortement dépendante de la validité des hypothèses faites sur la simulation de référence. L’extrapolation de l’information altimétrique vers la subsurface reste par contre plus problématique. L’assimilation permet d’améliorer la représentation de la variabilité du champ de température en profondeur (la corrélation entre une anomalie du niveau de la mer et une anomalie dans la position de la thermocline étant très forte, cela n’est pas très surprenant). L’assimilation de la SLA T/P+ERS référencée avec la topographie dynamique moyenne GRACE permet également d’améliorer le réalisme de la pente des isothermes en profondeur. Ceci se traduit par une réduction de RMSD par rapport aux observations de températures indépendantes TAO de près de 0.2°C, la RMSD moyenne passant de 1.32°C pour la simulation libre à 1.15°C pour l’expérience avec assimilation de données. L’assimilation d’altimétrie seule s’avère en revanche incapable de corriger l’erreur sur le positionnement moyen de la thermocline. A cela deux explications : (i) le champ thermohalin de la climatologie Levitus, et par là même le champ thermohalin de la simulation de référence, reste entaché d’erreurs (le thermocline est trop diffuse et trop profonde) même s’il est en meilleur accord avec les observations de la réalité de l’océan (et en particulier avec les TAO) que le champ thermohalin de la simulation libre, et (ii) compte tenu des approximations faites sur la paramétrisation du SEEK, l’utilisation d’une observation de surface unique ne suffit pas pour extrapoler de façon fiable l’information contenue dans les observations vers la subsurface en utilisant la seule statistique du filtre SEEK.

Si l’on replace cette première série de résultats dans le cadre plus général des objectifs de cette thèse, une première série de conclusions peut être tirée. Nous avons montré la faisabilité de l’assimilation d’une topographie dynamique absolue. L’utilisation d’une topographie dynamique moyenne observée, et surtout plus réaliste que la surface moyenne du modèle classiquement utilisée pour référencer la SLA T/P+ERS permet une amélioration du réalisme de la surface moyenne simulée. L’amélioration topographie dynamique moyenne simulée avec assimilation de données se traduit par une amélioration générale de la circulation moyenne simulée. Cependant, des problèmes persistent lors de l’extrapolation en profondeur de la correction par la statistique du SEEK.

Chapitre 7

Complémentarité des données altimétriques T/P+ERS et des données in-situ TAO. Rôle de la topographie dynamique moyenne GRACE.

Sommaire

7.1	Vers l'assimilation conjointe d'altimétrie et de données in-situ	117
7.1.1	Apports de l'assimilation : écarts RMS	117
	Topographie dynamique	117
	Température	118
7.1.2	Apports de l'assimilation : comparaisons aux observations	119
	Niveau de la mer	119
	Courants de surface	122
	Structure thermique verticale	123
7.2	Apports de la MDT observée GRACE pour l'assimilation conjointe	124
7.2.1	Synthèse des principaux résultats de la GRL	124
7.2.2	Letter : On the role of the GRACE mission in the joint assimilation of altimetric and TAO data in a tropical Pacific Ocean model	125
7.3	Conclusion	133

Le chapitre précédent a montré que l'assimilation de la seule topographie dynamique, même si elle permet des améliorations substantielles en surface, a un impact plus limité en subsurface. Notre système semble corriger les gradients en profondeur, en accord avec les gradients de la topographie dynamique assimilée. Il est par contre incapable de repositionner la thermocline et les autres isothermes à la profondeur observée par le réseau indépendant de mouillages TAO.

On voit ici les limites d'un système d'observation mono-données et surtout qui est limité à la seule surface de l'océan. L'altimétrie nous fournit une observation synoptique d'une grande précision (cf figure 5.1, page 90). Chacun des trois satellites T/P, ERS1 et ERS2 fournit environ 12000 observations du niveau de la mer le long de leurs traces respectives pour chaque analyse (seul un point sur trois est conservé dans les fichiers d'observations soit un point tous les 21 km). Durant la période 1993-1998, on dispose d'au moins deux des trois satellites, excepté pour une période d'un peu plus d'un an entre la fin de l'année 1993 et le début de l'année 1995, où seul T/P fournit des observations. On a donc un flux de données d'observations du niveau de la mer d'environ 24000 observations pour chaque analyse. Compte tenu de la résolution d'ORCA2, cela correspond à environ 3500 observations sur la grille modèle, soit la quasi totalité du premier niveau. Le problème ne vient donc pas de la quantité d'observations mais bien du fait qu'elles se limitent à la seule surface. Les nombreuses approximations consenties lors de la mise en œuvre et de la paramétrisation de la statistique d'erreur du filtre SEEK (approximations qui ont pour principal effet la perte du caractère optimal de l'analyse) font que l'on ne parvient pas à extrapoler la correction en profondeur de façon fiable.

Afin de faciliter la tâche de l'assimilation et de limiter l'influence des paramétrisations du SEEK, et donc des approximations consenties pour estimer ces paramétrisations, nous avons assimilé conjointement au signal altimétrique les données des profils de température in-situ TAO. Les données TAO représentent une source d'observation synoptique et facilement accessible. L'assimilation de données de température est de nature différente par rapport à celle des données altimétriques. D'une part, la température est une des variables pronostiques du modèle directement contrainte par les incréments IAU sous la forme de termes de forçage additionnels (cf section 4.5.2), contrairement à la topographie dynamique, qui n'était pas directement corrigée par l'assimilation (la correction constatée sur la topographie dynamique peut en fait être vue comme la réponse du modèle à la correction effectuée par l'assimilation sur le champ thermohalin du modèle). D'autre part, la couverture spatiale des observations TAO ne recouvre que la bande $\pm 8^\circ$ et les sections TAO sont espacées de $10/15^\circ$ en longitude (cf section 2.3.1). En revanche, les observations TAO donnent accès à une vision tridimensionnelle de l'océan.

Un système d'observation, combinant l'altimétrie et les mouillages TAO, est a priori plus intéressant. D'une part, la quantité d'information disponible est plus importante. On dispose en effet d'environ 10000 observations TAO pour chaque analyse. Sachant que les observations TAO sont des moyennes journalières et compte tenu de la résolution du modèle, cela correspond à environ 1000 observations sur la grille modèle, mais cette fois en 3D. D'autre part ces deux jeux de données sont complémentaires. En les assimilant conjointement, on va cependant introduire un nouveau problème, celui de compatibilité entre les données. Les données in-situ comme les données de température TAO sont des données absolues. Par opposition, seule la composante résiduelle de l'altimétrie est connue avec une grande précision, la topographie dynamique moyenne de référence étant contaminée par des erreurs sur le géoïde. Ces erreurs ont pour conséquence de possibles désaccords entre

l'information fournie par l'altimétrie et l'information fournie par les TAO. Ces désaccords peuvent se montrer problématiques comme cela a été illustré par *Parent et al.* (2003). L'amélioration du réalisme de la topographie dynamique moyenne grâce aux données GRACE devrait permettre de limiter les problèmes de compatibilité entre les données altimétriques et in-situ.

7.1 Vers l'assimilation conjointe d'altimétrie et de données in-situ

Afin de d'estimer l'impact du passage d'une assimilation mono-données à une assimilation multi-données, nous avons réalisé deux nouvelles expériences : l'expérience EXP3, n'assimilant que les données TAO, et l'expérience EXP4, assimilant conjointement la SLA T/P+ERS référencée avec la topographie dynamique moyenne GRACE et les profils de température in-situ TAO. Ces deux nouvelles expériences présentent les mêmes caractéristiques que les expériences EXP1 et EXP2. Elles utilisent les même forçages et s'étendent sur l'année 1993. En ce qui concerne le SEEK, la paramétrisation de l'espace d'erreur en rang réduit, basée sur les écarts avec la simulation de référence relaxée vers la climatologie (cf section 5.2), est utilisée.

7.1.1 Apports de l'assimilation : écarts RMS

Topographie dynamique

Comme je l'ai dit dans le chapitre précédent, la première étape pour valider une expérience avec assimilation de données est de s'assurer de la convergence vers les variables assimilées. L'évolution temporelle de la RMSD par rapport à la topographie dynamique observée est représentée sur la figure 7.1 pour la simulation libre et les différentes simulations avec assimilations de données. On note une légère dégradation en terme de RMSD par rapport à la topographie dynamique observée lorsque l'on passe d'une assimilation mono-données (l'expérience EXP2 n'assimilant que la DT T/P+ERS, en bleu sur la figure) à une assimilation multi-données (l'expérience EXP4 assimilant conjointement la DT T/P+ERS et les données de températures TAO, en noir sur la figure). La RMSD moyenne sur l'année 1993 passe de 5.5 cm à 6 cm. Comme nous l'avons vu dans la section 6.1, des

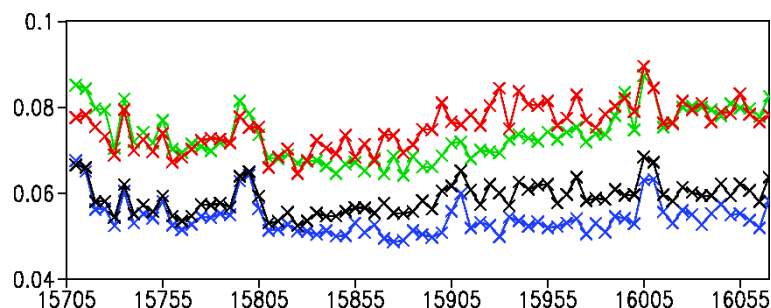


FIG. 7.1 – Écarts RMS (en cm) par rapport à la topographie dynamique (SLA + MDT) pour la simulation libre (en vert) et les expériences avec assimilation de données EXP2 (en bleu) et EXP3 (en rouge) et EXP4 (en noir).

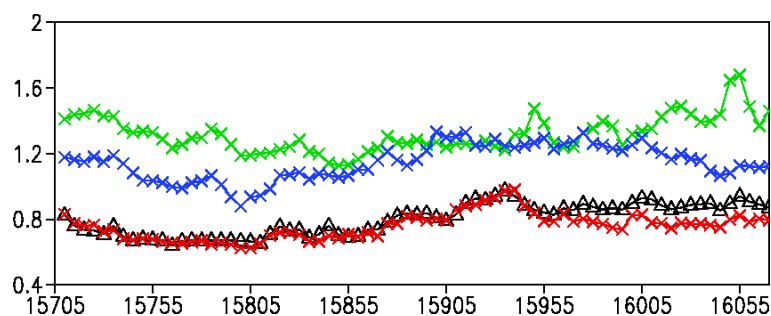


FIG. 7.2 – *Ecart RMS (en °C) par rapport aux profils de température TAO pour la simulation libre (en vert) et les expériences avec assimilation de données EXP2 (en bleu), EXP3 (en rouge) et EXP4 (en noir).*

erreurs persistent, en particulier dans l'hémisphère nord, entre les hauteurs dynamiques déduites des mouillages TAO et la topographie dynamique satellite observée obtenue en additionnant la MDT GRACE et la SLA T/P+ERS. Il n'est donc pas surprenant que l'assimilation conjointe de ces deux jeux de données se traduise par une légère dégradation des résultats en terme de RMSD, par rapport à la topographie dynamique observée, comparé à l'expérience EXP2 n'assimilant que la DT T/P+ERS. Le gain est en revanche important par rapport à la simulation libre (en vert) et la simulation n'assimilant que les TAO (en rouge) avec une RMSD moyenne par rapport à la topographie dynamique observée de l'ordre de 7.5 cm pour ces deux simulations.

Température

Voyons maintenant les apports de l'assimilation de données par rapport aux observations de température des mouillages TAO. Les TAO ne sont plus des données indépendantes vis à vis des expériences EXP3 et EXP4. La figure 7.2 montre l'évolution de la RMSD au cours de l'année 1993. Le gain en terme d'écart RMS par rapport aux observations est important. La RMSD moyenne sur l'année 1993 passe de 1.32°C pour la simulation libre, à 1.15°C pour l'EXP2 (les données TAO sont des données indépendantes pour cette expérience). Le gain est encore bien plus important pour les expériences assimilant les profils de température TAO. La RMSD moyenne diminue respectivement jusqu'à 0.76°C et 0.8°C pour les expériences EXP3 et EXP4. On note très peu de différence (0.04°C RMS) entre l'expérience assimilant les TAO seuls (expérience EXP3) et l'expérience assimilant conjointement la DT T/P+ERS et les TAO.

Les premiers résultats concernant cette deuxième série d'expériences sont très positifs. L'assimilation conjointe semble efficace en termes d'écarts RMS par rapport aux observations. La différence en terme de RMSD par rapport à la DT observée entre l'expérience EXP4 assimilant conjointement l'altimétrie et les TAO, et l'expérience EXP2 assimilant uniquement l'altimétrie, est faible (inférieur à 0.5cm RMS). De même, la différence en terme de RMSD vis à vis des données TAO assimilées entre les expériences EXP3 et EXP4 est minimale (de l'ordre de 0.04°C RMS). Il semble que la compatibilité entre les données assimilées soit effective et que notre système parvienne à les assimiler conjointement en réduisant simultanément l'erreur par rapport aux deux types d'observation. Ces résultats représentent une nette amélioration par rapport aux travaux antérieurs réalisés

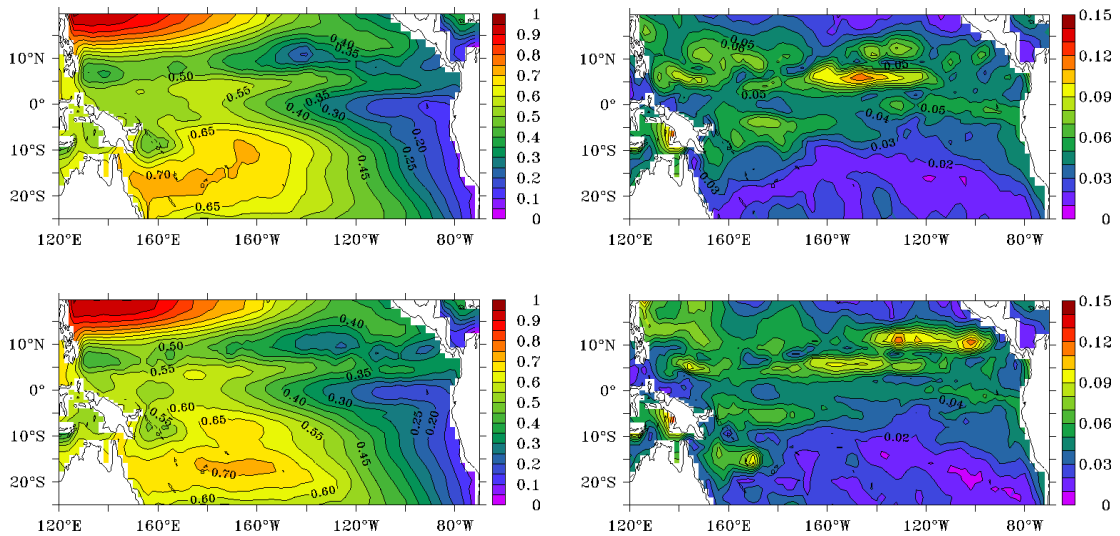


FIG. 7.3 – MDT (à gauche en m) et écart type de la DT (à droite en m) pour les simulations EXP3 (en haut) avec assimilation des profils de température TAO et EXP4 (en bas) avec assimilation conjointe de la DT et des profils de température TAO.

dans l'équipe MEOM (Parent, 2000).

7.1.2 Apports de l'assimilation : comparaisons aux observations

Voyons maintenant l'impact de l'assimilation conjointe de données altimétriques et in-situ sur la dynamique simulée. Comme nous l'avons fait dans le chapitre précédent, nous allons comparer, en termes de structures moyenne et de variabilité, les différentes simulations EXP2, EXP3 et EXP4 assimilant respectivement, altimétrie seule, TAO seuls et altimétrie et TAO conjointement avec les observations disponibles.

Niveau de la mer

En premier lieu, il convient de regarder l'influence de l'assimilation des TAO seuls, et de la DT T/P+ERS conjointement avec les TAO sur le niveau moyen et la variabilité du niveau de la mer simulé. Dans le chapitre précédent, nous avons vu que l'impact de l'assimilation de la DT T/P+ERS seule était très positif, avec une amélioration des structures moyennes et de la variabilité permettant de réduire la RMSD par rapport à la topographie dynamique observée à 5.5 cm. La figure 7.3 montre la MDT et la variabilité de la DT simulées pour les expérience avec assimilation de données EXP3 et EXP4. Les cartes de la MDT et de la variabilité de la DT de la simulation libre et des observations sont visibles sur la figure 6.7, page 107. Une première remarque concernant la topographie dynamique moyenne simulée avec assimilation de TAO uniquement est que la MDT, hors du rail 10°N-10°S, où l'on dispose des observations TAO, est très proche de MDT de la simulation libre. Cela n'est pas très surprenant compte tenu du fait que l'on utilise la variante dite locale du filtre SEEK (cf section 4.4.6). Les modes d'erreurs utilisés sont des modes multi-variés globaux mais avec une influence locale. C'est à dire que les corrélations à grande distance (dont on sait qu'elles sont peu fiables) sont explicitement imposées à

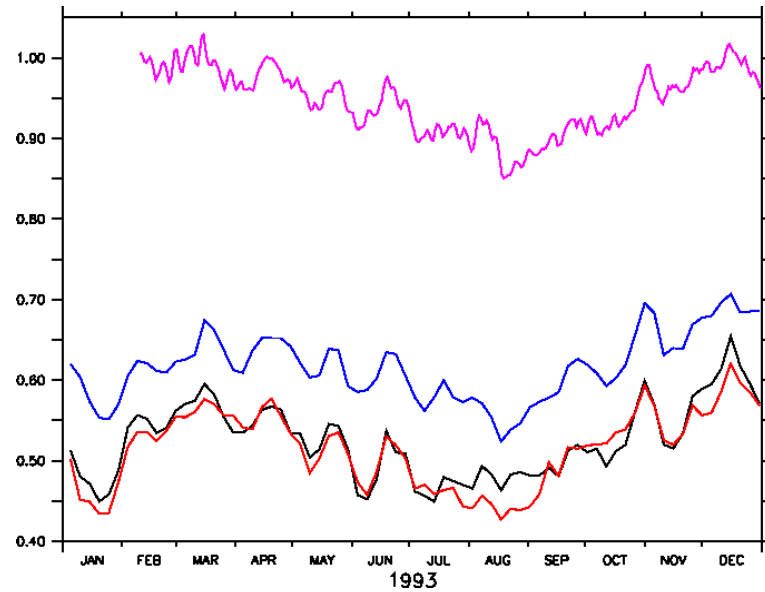


FIG. 7.4 – Evolution temporelle de la DT à l'équateur à 156° E de longitude de la simulation assimilant la topographie dynamique (en bleu), de la simulation assimilant les TAO (en rouge) et de la simulation assimilant conjointement la topographie dynamique et les TAO (en noir). L'évolution temporelle de la hauteur dynamique déduite des TAO au même point est tracée (en magenta). Pour la hauteur dynamique, le niveau de référence est arbitraire.

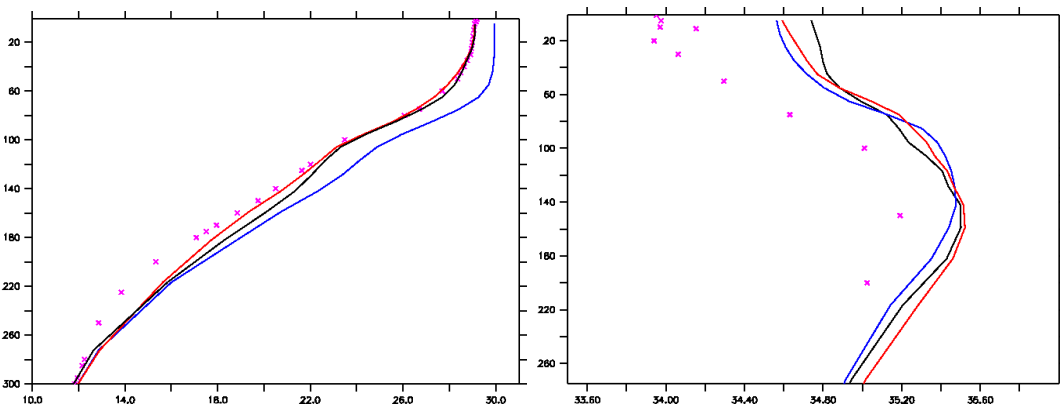


FIG. 7.5 – Profil de température (à gauche) et de salinité (à droite) moyen pour l'année 1993 au point 0° N- 156° E. Les profils des observations TAO (en magenta), de la simulation assimilant la topographie dynamique (en bleu), de la simulation assimilant les TAO (en rouge) et de la simulation assimilant conjointement la topographie dynamique et les TAO (en noir), sont tracés.

zéro. Il n'y a donc plus de correction hors des zones d'influence dont l'extension méridienne est fixée à 4 points de grille de part et d'autre du point où est calculé l'innovation. Compte tenu de l'anisotropie de la grille ORCA2 dans les régions tropicales, il n'y a pas de correction au delà de $\pm 14^\circ$ de latitude. Concentrons nous donc sur la bande échantillonnée par les TAO. On note un sillon Nord équatorial bien marqué, ce qui est en accord avec les observations satellites. Dans l'extrême Ouest du bassin, aux abords de l'équateur, l'erreur sur la MDT simulée avec assimilation de données TAO est assez forte, avec un niveau moyen trop bas comparé à celui de la MDT GRACE. Ce constat est aussi valable, mais dans une moindre mesure pour l'EXP4. Regardons en un point l'évolution temporelle de la DT pour essayer de mieux comprendre le problème. La figure 7.4 montre l'évolution temporelle pour les expériences EXP1, EXP2 et EXP4 de la DT simulée à l'équateur au point 156°E de longitude. La figure montre également l'évolution temporelle de la hauteur dynamique déduite du mouillage TAO à ce point. Le niveau de référence pour la hauteur dynamique TAO étant fixé de manière arbitraire, on ne peut pas comparer le niveau absolu des trois simulations avec celui de la hauteur dynamique. Il est par contre intéressant de noter la bonne corrélation entre le signal de hauteur dynamique TAO observé et le signal de DT simulé par les trois expériences avec assimilation de données, y compris avec l'expérience EXP2 n'assimilant pas les données TAO. Cette bonne corrélation illustre encore une fois la compatibilité entre l'altimétrie et les données in-situ TAO. L'assimilation des TAO modifie sensiblement le niveau absolu de la topographie dynamique simulée, avec un décalage quasiment constant sur toute l'année 1993 de près de 10 cm entre les EXP2 et EXP4, assimilant les profils TAO, et l'EXP1 assimilant l'altimétrie seule (cf figure 7.4). Dans le reste de la bande équatoriale, on ne retrouve pas, en revanche, de telles différences en terme de niveau moyen simulé. La région où se concentre ces fortes erreurs correspond à peu près à la région de la Warm Pool. On sait aujourd'hui que la salinité, et ses variations, a un impact non négligeable sur la topographie dynamique (*Maes*, 1998). La Warm Pool est une des régions où cet impact est particulièrement marqué (*Maes et al.*, 2002). Une des hypothèses que l'on peut avancer pour expliquer la différence significative constatée entre la DT simulée avec et sans assimilation de profils de température TAO dans la région de la Warm Pool est un mauvais contrôle de la salinité par l'assimilation de données. Cette hypothèse semble être confirmée par la figure 7.5 montrant les profils moyens de température et de salinité sur l'année 1993 pour les observations TAO et les expériences en mode assimilé au point $0^\circ\text{N}-156^\circ\text{E}$. On voit que l'assimilation des données de température TAO (EXP2 et EXP4) modifie assez significativement le profil de température moyen en ce point, et permet de réduire de manière importante l'erreur en température par rapport à l'EXP1 n'assimilant que la DT. Le profil de salinité simulé est en revanche peu modifié par l'assimilation de données. L'erreur par rapport à la salinité mesurée par le mouillage TAO reste importante pour cette période et en ce point. Le profil de température proche des observations alors que le profil de salinité présente des erreurs relativement importantes, combiné au fait que la hauteur dynamique dépend de manière significative de la salinité dans la région de la Warm Pool semble donc à même d'expliquer les différences constatées en terme de topographie dynamique sur les figures 7.3 et 7.4.

En ce qui concerne la variabilité de la DT, le passage à l'assimilation conjointe n'a pas modifié la variabilité simulée par rapport à la simulation assimilant uniquement l'altimétrie. Ceci est un point positif dans la mesure où la variabilité simulée était en accord avec la variabilité observée. Dans l'expérience EXP2 n'assimilant que les TAO, la variabilité simulée le long de 5°N semble surestimée par rapport à la variabilité observée par

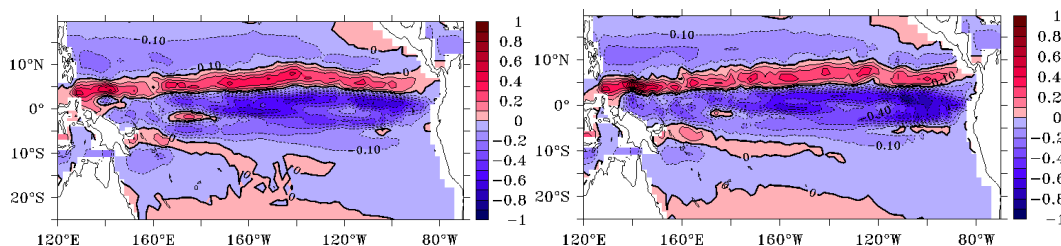


FIG. 7.6 – Vitesse zonale moyenne à 15 m de profondeur (en ms^{-1}) pour les simulation EXP3 (en haut) avec assimilation des profils de température TAO et de la simulation EXP4 (en bas) avec assimilation conjointe de la DT et des profils de température TAO.

l'altimétrie. L'assimilation des TAO a par contre permis de réduire la variabilité le long de l'équateur dans l'Est. La variabilité dans cette région était trop importante dans la simulation libre (cf figure 6.7). Hors de la bande $\pm 10^\circ$ de latitude, les modifications sont faibles pour les raisons évoquées ci-dessus.

En conclusion, l'assimilation des profils in-situ TAO seuls n'a un impact que dans la région contrôlée par les TAO. Ceci est la conséquence de l'utilisation de la variante locale du SEEK. Dans le guide d'onde équatorial, l'assimilation des profils in-situ TAO modifie assez fortement la topographie dynamique moyenne simulée. On observe une amélioration au niveau du sillon Nord équatorial, avec des gradients méridionaux plus forts. L'assimilation des profils in-situ TAO fait apparaître une région correspondant approximativement à la Warm Pool, où les erreurs par rapport à la topographie dynamique moyenne observée sont plus importantes. L'explication de la présence de cette erreur très localisée nécessite d'autres diagnostics. Une explication pourrait être les variations de la salinité insuffisamment contrôlées par l'assimilation de données et qui peuvent affecter de manière significative le signal de topographie dynamique dans cette région (Maes, 1998). L'analyse du niveau de la mer simulé avec assimilation conjointe d'altimétrie et de profils in-situ TAO confirme les résultats en terme de RMSD de la section précédente. La présence des TAO ne dégrade que très faiblement le niveau de la mer simulé avec assimilation. Le niveau moyen (excepté dans la région de la Warm Pool pour les raisons évoquées ci-dessus), ainsi que la variabilité sont très proches de l'expérience EXP2 assimilant l'altimétrie seule et des observations.

Courants de surface

Comme je l'ai déjà expliqué, dans le chapitre précédent, l'étude des courants zonaux est particulièrement révélatrice des modifications sur les gradients méridionaux de topographie dynamique. Voyons comment ils ont été modifiés par l'assimilation des TAO. Les courants zonaux à 15 m sont présentés sur la figure 7.6. On peut trouver une carte de la climatologie des courants zonaux déduite des bouées dérivantes SVP (Niiler, 2001) ainsi que les courants simulés en mode libre sur la figure 6.8, page 109. On ne note pas de différence marquée entre les deux cartes de courant. L'assimilation des TAO seuls a pour effet de renforcer le NECC par rapport aux courants de la simulation en mode libre et donc de se rapprocher des courants observés. Ceci est en accord avec l'amélioration de la MDT dans la région de la dorsale et du sillon Nord équatorial constatée sur la figure 7.3. Un point négatif de l'assimilation des profils TAO est la moins bonne représentation de

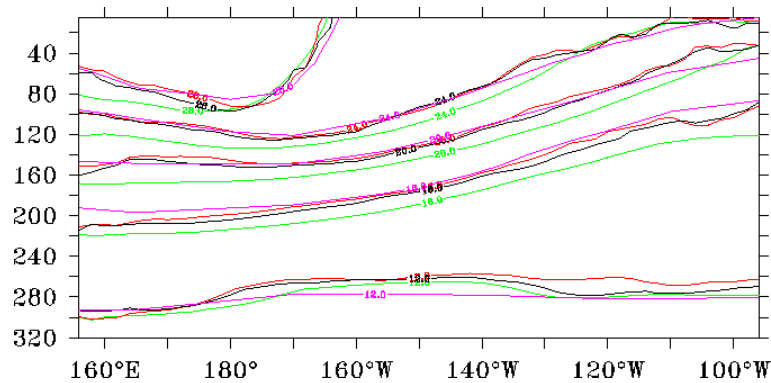


FIG. 7.7 – Profondeur moyenne pour l'année 1993 des isothermes 12, 16, 20, 24 et 28°C le long de l'équateur pour la simulation libre (en vert), les expériences avec assimilation de données EXP3 (en rouge) et EXP4 (en noir) et les données indépendantes TAO (en magenta).

la séparation en deux branches, la branche Nord et la branches Sud du SEC dans l'Est du bassin avec des vitesses vers l'Ouest trop importantes au niveau de l'équateur.

Structure thermique verticale

Les profils in-situ TAO nous fournissent maintenant une information sur la verticale. Voyons comment cette information est utilisée par le filtre SEEK pour corriger le champ thermique du modèle. La figure 7.7 est identique à la figure 6.10 mais montre la position des isothermes pour les observations TAO, l'expérience libre et les expériences en mode assimilé EXP3 et EXP4. Les observations TAO ne sont plus des observations indépendantes pour ces deux expériences assimilées. L'assimilation des données TAO a permis de réduire fortement les erreurs qui existaient entre la position moyenne des isothermes de la simulation libre, mais aussi de la simulation avec assimilation d'altimétrie seule et la position moyenne des isothermes observées par les mouillages TAO. Dans l'Est du bassin, aux alentours de 100°W, on a du mal à se rapprocher des températures observées. Dans la réalité, les îles Galápagos se situent juste à l'ouest de 90°W au niveau de l'équateur. Ces îles ne sont pas présentes dans la bathymétrie de la configuration modèle utilisée, ce qui peut être une source d'explication pour les différences constatées dans cette région.

En résumé, l'impact de l'assimilation des profils in-situ TAO sur le champ thermique du modèle est important. Contrairement à l'assimilation d'altimétrie seule, qui permettait de corriger les pentes des isothermes en profondeur sans pour autant parvenir à corriger l'erreur sur la position moyenne des isothermes simulées, l'assimilation des profils in-situ TAO, en nous fournissant une information sur la vertical permet de corriger cette erreur. Un autre point qu'il faut souligner concerne la faible différence qui existe entre la position des isothermes simulées avec assimilation de TAO seuls et avec assimilation conjointe d'altimétrie et de TAO. Cette faible différence traduit la bonne compatibilité entre l'altimétrie et les données de température TAO apportée par l'utilisation de la topographie dynamique moyenne observée GRACE.

7.2 Assimilation conjointe de l'altimétrie et des profils de température TAO : Apports de la MDT observée GRACE

Comme nous venons de le voir dans la section précédente, la seule assimilation d'altimétrie ou de profils de température TAO, même si elle permet de réduire un certain nombre d'erreurs présentes dans la simulation libre, ne permet pas de contrôler le modèle de manière satisfaisante. L'assimilation d'altimétrie seule permet par exemple d'améliorer le réalisme de la topographie dynamique moyenne ou des courants simulés, mais est incapable de corriger l'erreur sur la position moyenne en profondeur des isothermes en accord avec les observations indépendantes de température. L'assimilation des profils de température TAO seule permet de corriger l'erreur sur le champ thermohalin simulé mais ne permet pas, par exemple, de contrôler la variabilité ondulatoire du modèle. Nous avons également constaté que l'assimilation conjointe de la topographie dynamique référencée au géoïde GRACE et des profils TAO était efficace et permettait de se rapprocher simultanément de ces deux jeux de données d'observation. Ceci représente une nette amélioration par rapport aux travaux de *Parent* (2000) qui avaient constaté que l'assimilation conjointe des anomalies de surface libre T/P+ERS et des données profils de température TAO dans une configuration assez similaire à la mienne ne permettait pas d'amélioration, notamment en terme d'écart RMS. *Parent* (2000) avaient montré qu'ils parvenaient à se rapprocher de l'un ou l'autre des jeux d'observations en jouant sur le paramétrisation de l'erreur d'observation, mais pas des deux simultanément ce qui traduisait la présence d'incompatibilités entre les observations assimilées. Ces incompatibilités étaient la conséquence des erreurs sur la topographie dynamique moyenne du modèle utilisée par *Parent* (2000) pour référencer l'altimétrie. Il est vraisemblable que les améliorations constatées ici soient en grande partie imputables à la meilleure compatibilité entre les deux jeux de données offerte par l'utilisation de la topographie dynamique moyenne GRACE pour référencer les anomalies de niveau de la mer (cf section 6.1). Afin d'estimer les apports de l'utilisation d'un signal altimétrique absolu sur ce type d'assimilation, j'ai réalisé deux expériences. Une première assimilant les anomalies de niveau de la mer T/P+ERS référencées avec la surface moyenne du modèle, et une deuxième assimilant les anomalies de niveau de la mer T/P+ERS référencées avec la topographie dynamique moyenne GRACE. Conjointement à l'altimétrie, les profils de température TAO sont assimilés afin d'améliorer le contrôle de la subsurface dans les deux simulations. Le travail présenté dans cette section a fait l'objet d'une publication dans le journal *Geophysical Research Letters*. Je ferai une synthèse rapide des principaux résultats dans la section 7.2.1 ; la lettre est fournie dans la section 7.2.2.

7.2.1 Synthèse des principaux résultats de la GRL

Les récents développements de notre connaissance du géoïde terrestre, grâce notamment aux données gravimétriques satellites de la mission GRACE, ont permis l'exploitation de la topographie dynamique absolue observée dans le cadre d'expériences d'assimilation de données dans un modèle réaliste du Pacifique Tropical. Deux simulations avec assimilation de données, basées sur le système d'assimilation décrit dans les précédents chapitres, sont réalisées. Elles s'étendent sur l'année 1993 et utilisent la paramétrisation de l'espace d'erreur en rang réduit introduit dans la section 5.2. Les deux simulations diffèrent unique-

ment par le choix de la MDT : l'expérience EXP1 utilise le signal altimétrique référencé à la MDT modèle, alors que l'expérience EXP2 utilise la MDT GRACE pour reconstruire le signal de topographie dynamique. Les profils de température in-situ TAO sont assimilés conjointement au signal altimétrique.

La meilleure compatibilité entre les données assimilées présente dans l'expérience EXP2 se traduit par une assimilation plus efficace. Dans les deux expériences, on parvient ainsi à fortement réduire la RMSD avec les observations de température assimilées issues des mouillages TAO (plus de 60% en variance). En revanche, alors que dans l'expérience EXP2 utilisant la topographie dynamique moyenne satellite observée, on réduit en parallèle la variance de la RMSD sur la topographie dynamique absolue observée de plus de 40%, on parvient à une réduction de la RMSD d'à peine 10% dans l'expérience EXP1. Encore plus significatif, lorsque l'on évalue les deux simulations par rapport aux données indépendantes comme les profils de température XBTs collectés par les bateaux VOS au cours de l'année 1993, la variance de la RMSD est réduite d'environ 30% dans l'expérience EXP2, alors qu'elle reste inchangée pour l'expérience EXP1.

En conclusion, les apports d'une topographie dynamique moyenne observée et réaliste sont assez nets, en particulier en ce qui concerne la compatibilité entre les données satellites et les données in-situ. Cette meilleure compatibilité se traduit notamment par une assimilation conjointe plus efficace.

7.2.2 Letter : On the role of the GRACE mission in the joint assimilation of altimetric and TAO data in a tropical Pacific Ocean model

F. Castruccio⁽¹⁾, J. Verron⁽¹⁾, L. Gourdeau⁽²⁾, J. M. Brankart⁽¹⁾ and P. Brasseur⁽¹⁾

⁽¹⁾Laboratoire des Ecoulements Géophysiques et Industriels, Grenoble, France

⁽²⁾Institut de Recherche pour le Développement, Nouméa, New Caledonia

Received 23 January 2006 ; revised 31 May 2006 ; accepted 6 June 2006 ; published 26 July 2006.

Abstract

Recent advances in our knowledge of the earth geoid have made it possible to exploit absolute sea surface height measurements in realistic numerical modelling studies of the ocean. This letter provides evidence of the benefit of the GRACE referenced mean dynamic topography (MDT) for the simulation of the tropical Pacific ocean through the joint assimilation of altimetric data and of vertical temperature profiles from the TAO/TRITON array. Results are considered in relation to those obtained using a classical model MDT, based on the model itself, and are validated against independent XBT data. The use of the GRACE MDT leads to significantly improved results with respect to these independent data. If the compatibility between the altimetric sea surface height and the TAO/TRITON temperature improves the model simulation, some limitations exist, especially in the Warm Pool where a part of the sea surface height signature is associated to a salinity signal.

1. Introduction

Altimetry is today a major field of oceanographic developments, in particular for ocean monitoring and prediction through data assimilation into ocean models. However, only the residual component (the Sea Level Anomaly or SLA) of the altimetric measured signal (the Sea Surface Height or SSH) can reliably be used. The mean altimetric reference, the

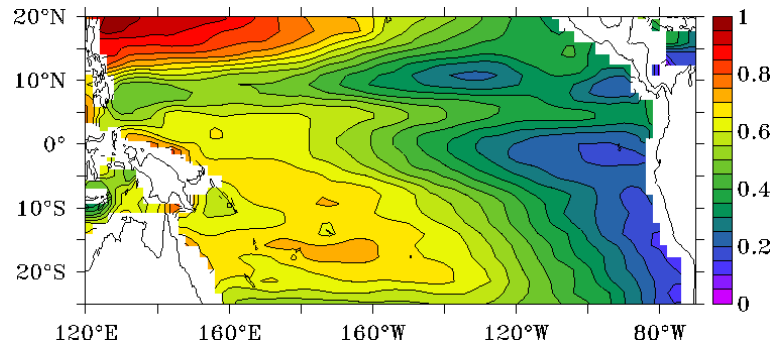


FIG. 7.8 – *GRACE MDT (in meter)*.

so called Mean Dynamic Topography (MDT) (e.g. *Birol et al. (2005)*), is contaminated by large geoid errors (especially with high order harmonics). The dynamic topography which is the relevant oceanic quantity, and results from the addition of the SLA and the MDT, was only accessible through the use of a synthetic MDT evaluated by indirect and often approximate approaches. The gravimetric satellite missions such as CHAMP (2000), GRACE (2002) and GOCE (2007) are changing this perspective by providing a more accurate geoid reference and the possibility of using the absolute altimetric signal (*Gourdeau et al. (2003)*; *Birol et al. (2005)*).

The objective of this work is to investigate the assimilation of this absolute altimetric signal and to show how this can impact the assimilation of multiple source data. The investigation is performed with a model of the tropical Pacific ocean where an easily accessible and rather synoptical set of in-situ data is available, thanks to the TAO/TRITON moorings, to complement the altimetric observations. The Singular Evolutive Extended Kalman (SEEK) filter is used to jointly assimilate the TOPEX/Poseidon and ERS SSH referenced to the GRACE geoid, and the TAO/TRITON temperature profiles. A set of two experiments is conducted : the first assumes a pre-GRACE situation using an artificial MDT from the model itself (as has been often done until now), while the second uses the new capability provided by a fully measured MDT computed with the GRACE geoid (hereafter denoted shortly as the “GRACE MDT”).

The following section, Section 2, describes the model and the data sets. Section 3 focuses on the assimilation system while Section 4 discusses the results of the data assimilation experiments and evaluates them in particular against independent XBTs. Section 5 draws some conclusions.

2. Model and data sets

2.1. The numerical model

The OPA OGCM (www.lodyc.jussieu.fr/opa) is based on the standard ocean primitive equations and uses a free surface formulation. The configuration is the so-called ORCA2 global configuration : the extratropical $2^\circ \times 2^\circ$ mesh has a meridional grid spacing refinement down to 0.5° in the tropics to improve the equatorial dynamics. Along the vertical, there are 31 z-levels. This model has been used extensively for tropical dynamic studies and validated accordingly (e.g. *Lengaigne et al. (2003)*).

In the tropics, ocean dynamics are critically determined by wind stresses. For this reason, the ERS scatterometer wind stresses complemented by TAO derived stresses, which

together tend to produce more realistic thermocline and zonal currents, are used as proposed by *Menkes et al.* (1998). Regarding heat and fresh water flux, the 1957-2002 ERA40 ECMWF reanalysis is used. A local feedback term is added to account for the coupling between the surface ocean heat flux and the sea surface temperature (SST) with a feedback coefficient of $-40 \text{ W m}^{-2} \text{ K}^{-1}$. A restoring term is also applied on sea surface salinity (SSS) to avoid unrealistic drift. The model is initialized with Levitus climatology. It is spun up over a period of 4 years and then the interannual experiment starts in January 1992. The year 1993 is chosen for the present analysis.

2.2. The observation data

2.2.1. TOPEX/Poseidon & ERS

The altimetric observations used for assimilation consist of along-track SLA from the TOPEX/Poseidon and ERS satellite altimeters. The data are processed by the AVISO Altimetry group (www.jason.oceanobs.com).

2.2.2. GRACE

The so called GRACE MDT (Figure 7.8) used in this study (EXP2) is the difference between a Mean Sea Surface Height (MSSH) and a geoid. The MSSH corresponds to a 7-year mean (1993-1999) based on the TOPEX/Poseidon, ERS1-2 and GEOSAT follow-on altimetric data (the same as the one used for the SLA computation). The geoid is deduced from the EIGEN-GRACE02S Earth gravity field from GFZ-Potsdam, developed to degree 60 and based on 110 days of GRACE data. This GRACE MDT solution reasonably fits the real tropical Pacific MDT. This is illustrated, for example, by the maps in Figure 7.9 which shows the difference between the 0/500 dbar TAO dynamic height and the GRACE MDT for the 1993-2001 period (Figure 7.9b). The TAO dynamic height is evaluated using salinity estimated from a local T-S relation computed from climatological data. The order of magnitude of the error due to the 500-dbar reference level and the inaccurate estimation of salinity is about 2-3 cm, growing up to 6 cm in the Warm Pool. The free model MDT is also shown on Figure 7.9a and compare poorly with TAO. The overall RMS difference (RMSD) is 4.0 cm for the free model MDT and only 1.9 cm for the GRACE MDT. Most of the GRACE remaining differences are concentrated in the western part of the Equatorial Pacific ocean.

2.2.3. TAO

Altimetric data and in-situ temperature profiles are jointly assimilated into the model. Profiles from the TAO/TRITON array are used in this study (www.pmel.noaa.gov/tao). The TAO/TRITON array consists of about 70 moorings between 8°N - 8°S , 137°E - 95°W , providing high-quality temperature measurements in the upper 500 m.

2.2.4. XBT

In order to validate the thermal structure of our simulations, independent XBT profiles collected in the Tropical Pacific during the year 1993 are used. The profiles come from the CORIOLIS data base (www.coriolis.eu.org). We used around 17500 temperature profiles from the surface down to 700-meter depth over the year 1993.

3. Data assimilation scheme

3.1. Assimilation scheme

The assimilation scheme is based on the Singular Evolutive Extended Kalman (SEEK) filter introduced by *Pham et al.* (1998). This sequential method has already been used and assessed in the tropical Pacific (e.g. *Verron et al.* (1999)). In this study the forecast error

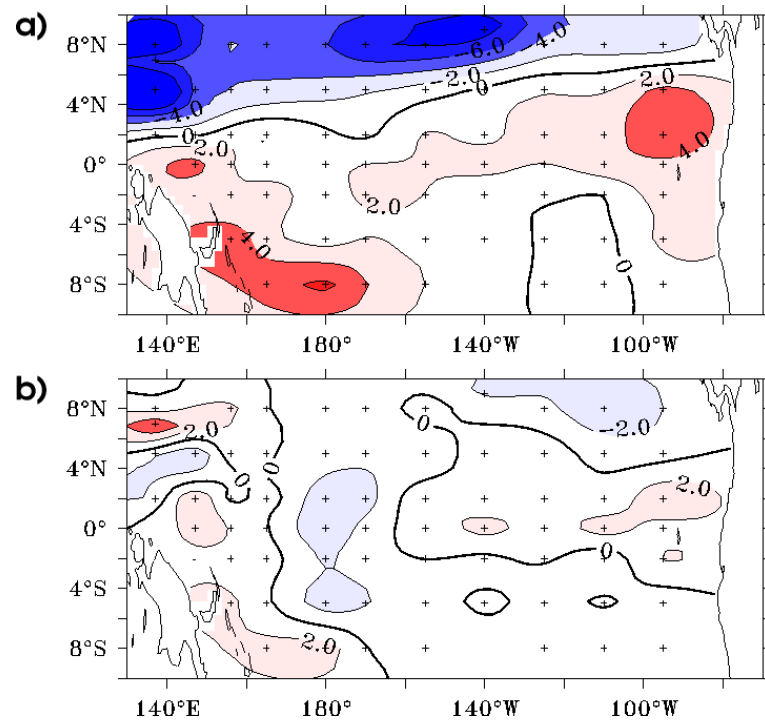


FIG. 7.9 – Difference for the 1993-2001 period between the mean 0/500 dbar TAO dynamic height and the model MDT (a) and the GRACE MDT (b). Contour interval is 2 cm. TAO moorings are located at crosses.

covariance matrix is kept constant in time as in *Gourdeau et al.* (2000), and the correction is applied on the full 3D model state vector (temperature, salinity, meridional and zonal velocity). Even though our model is global, data assimilation is only applied in the tropical Pacific domain defined -following *Durand et al.* (2003)- between 25°S and 20°N in latitude and from 120°E to the American coast in longitude. Buffer zones are used to smoothly connect the assimilated domain to extra tropical zones. In addition to the regular SEEK filter procedure an Incremental Analysis Update (IAU) algorithm was implemented to smooth out the time discontinuity of the solution resulting from intermittent corrections of the model state (see *Ourmières et al.* (2006)).

3.2. Parametrization of the error covariances

The parametrization of the error covariances is a primary factor in the efficiency of the assimilation process. An ensemble procedure is used to identify the appropriate reduced space for \mathbf{P}^f . It is first assumed that a model simulation with a strong relaxation towards climatological fields provides a good approximation of the mean true ocean. Then by using this reference run to reinitialize the free run every 5 days during one year, an ensemble of differences between this reference simulation and the free model 5-day forecasts without nudging is generated. This ensemble of 73 members represents the 5-day forecast error between the model and the reference trajectory. The covariance of this ensemble turns out to be an appropriate estimate of the model error covariance, needed to parameterize the SEEK filter. The rank of this matrix is reduced using a limited number (30) of three-dimensional, multivariate empirical orthogonal functions (EOFs) describing the dominant modes of the ensemble covariance.

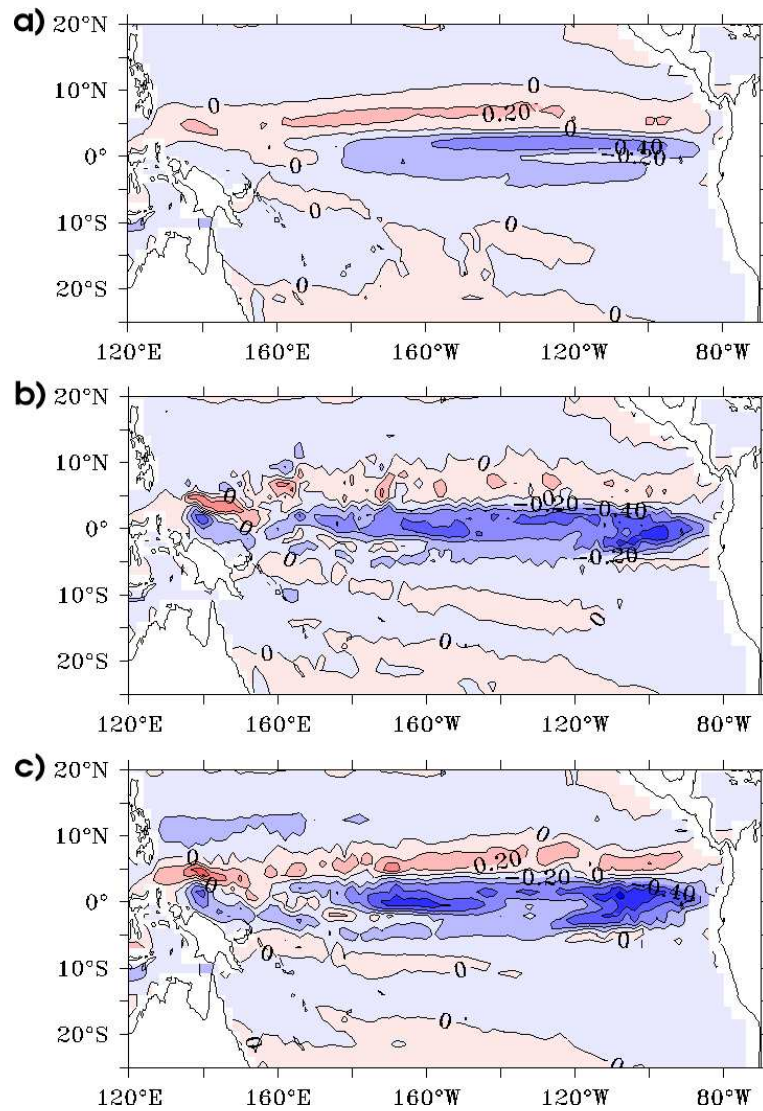


FIG. 7.10 – Mean zonal velocity for year 1993 at 25 m depth for the free run (a), for EXP1 (b) and for EXP2 (c). Contour interval is 0.2 ms^{-1} .

Regarding the parameterization of the observation errors covariance, a diagonal matrix is used with respectively 3 cm and 0.1°C standard deviation for altimetry and temperature data.

4. Data assimilation experiment

The data assimilation system described above is used to perform two assimilation experiments over the year 1993. As stated earlier, the two simulations only differ in the choice of the MDT : EXP1 uses altimetric data referenced to the model MDT, -which has a certain bias with regard to a TAO derived MDT (Figure 7.9a)-, while EXP2 uses altimetric data referenced to the GRACE MDT, -which has a less pronounced bias with regard to TAO (Figure 7.9b)-. Parent *et al.* (2003), in particular, have shown that performing a joint assimilation of SLA data with in-situ data may reach a deadlock due to the different

reference mean states of the two types of data. It is therefore expected here that the better data mean state compatibility present in EXP2 will contribute to provide a more efficient assimilation, making a better use of data complementarity.

A preliminary examination of the ability of the system to simulate the tropical Pacific circulation is illustrated in Figure 7.10 showing the mean zonal velocity fields at 25 m depth. The alternate bands of eastward- and westward-flowing currents characterizing the Tropical Pacific ocean circulation are present in the 3 runs. The free run currents are too weak. In Figure 7.10b the EXP1 mean zonal currents are uncharacteristically noisy for an annual mean, especially along 8°N, where the model produces the largest errors with regard to the TAO derived MDT, and where the incompatibility of the data sets is the most striking, as it is illustrated by the noisy small patterns. The EXP2 mean

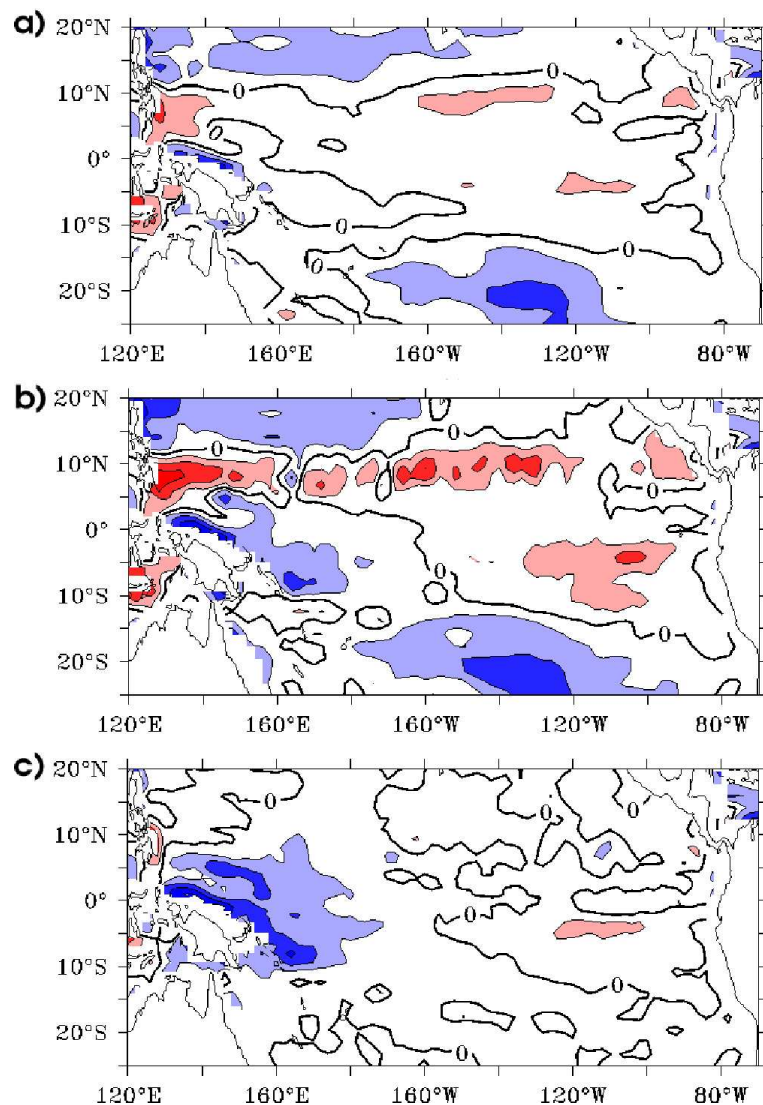


FIG. 7.11 – Difference maps between the simulated MDT for year 1993 and the satellite observed MDT for the free run (a), for EXP1 (b) and for EXP2 (c). Contour interval is 5 cm.

TAB. 7.1 – Standard deviation and percentage error variance reduction of the RMS misfit with respect to the SSH referenced to GRACE MDT and the TAO and XBT temperature profiles. EXP2* excludes the Western Equatorial region (see text).

	SSH	TAO	XBT
Free	7.62 cm	1.31 °C	1.52 °C
EXP1	7.18 cm	0.72 °C	1.51 °C
	-11%	-70%	-0.02%
EXP2	5.84 cm	0.82 °C	1.31 °C
	-41%	-61%	-26%
EXP2*	5.62 cm	0.79 °C	1.21 °C
	-45%	-63%	-38%

zonal currents are the closest to the drifter derived climatology proposed by P. Niiler (see *Lengaigne et al. (2003)*), thanks to the GRACE MDT which is known to feature all major geostrophic currents (*Tapley et al., 2003*).

Among the many diagnostics that could be shown, attention is drawn here to the predicted MDTs (with reference to GRACE) and to the vertical temperature profiles (compared to independent XBTs).

4.1. MDTs

The maps in Figure 7.11 show the difference between the simulated MDTs for the year 1993 and the observed MDT deduced from GRACE. The free-run RMSD (Figure 7.11a) is 4.4 cm over the whole domain. In EXP1 (Figure 7.11b), the RMSD increases to 5.7 cm. The solution is a long way from the observations, even further than the free run, especially along 8°N. Note that, assimilating the TAO temperature profiles only (experiment not shown) leads to a reduction in the RMSD, which is another evidence of the compatibility between the SSH referenced to GRACE geoid and in-situ TAO data. The EXP1 increase in the RMSD shows that the use of different types of observations can indeed be conflictual as noted by *Parent et al. (2003)*.

The assimilation system manages to drive the solution close to the observations in EXP2 (Figure 7.11c). The RMSD with respect to the observed MDT is 3.6 cm. The differences map reveals a narrow region of remaining large errors in the 10°S and 10°N strip west of the date line. The RMSD in this region increases up to 6.9 cm. Interestingly, the large errors area squares with the region of significant remaining errors with respect to the TAO dynamic height (see Figure 7.9b). Several hypotheses can be advanced to explain the concentration of errors within the Warm Pool. In particular, salinity variations insufficiently controlled by data assimilation, may significantly affect the SSH signal in this region (*Maes, 1998*). This can also account for the differences observed in Figure 7.9b in the same region. This suggests that both salinity and temperature data should be considered for joint assimilation with altimetric data.

The error patterns in Figure 7.11c led us to re-evaluate the EXP2 RMSD, excluding the 10°S and 10°N strip west of 180°E. The RMSD is then as low as 2.9 cm, thus confirming that a specific issue does exist in this sub-region.

4.2. Validation using independent XBT profiles

To further assess the results, we evaluated the experiments using independent XBT data. RMSD between the XBTs profiles over the whole domain and their model equivalent interpolated in space and time at profile location were computed and the results are shown

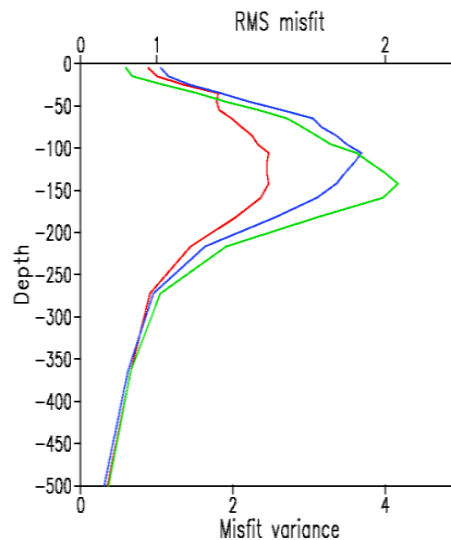


FIG. 7.12 – *Temperature RMSD with respect to independent XBT data for the 25°S and 20°N strip for year 1993 for the free run (green), for EXP1 (blue) and for EXP2 (red).*

in Figure 7.12. EXP1 is barely better than the free run, whereas EXP2 shows significant improvements for depths ranging from 40 m down to 250 m.

Table 7.1 summarizes the results. It presents RMSD with respect to the assimilated observations (SSH, TAO) and with respect to independent observation (XBT), the later being undoubtedly most meaningful to evaluate the performance of our assimilation system. When considering the SSH a significant RMSD decrease is seen from 7.6 cm RMS for the free run to 5.8 cm for EXP2, that is more than 40% error variance decrease. With respect to TAO data, the decrease in the error is even larger (more than 60% in variance). As far as the assimilation of TAO data is concerned, the presence of the GRACE MDT slightly degrades the performances from EXP1 to EXP2, which is a sign that data incompatibilities remains. Finally, and most importantly, with respect to XBT data, RMSD is 1.52°C for the free run, 1.51°C for EXP1 and 1.31°C for EXP2. RMSD is as low as 1.21°C when the West Equatorial Pacific ocean is excluded. This is a strong argument supporting the importance of GRACE data to jointly assimilate TAO and SSH observations.

5. Conclusion

The objective of this study was to assess how the use of a GRACE geoid could improve the modeling of tropical Pacific dynamics through data assimilation of altimetric and in-situ data. The study was motivated by the fact that defining the MDT reference is a recurrent problem in altimetric data assimilation studies and that ongoing and future satellite missions dedicated to the improvement of the earth geoid are likely to change this situation. The benefits brought by the use of a GRACE MDT in the assimilation system are clear. The first results are most positive, especially in terms of improved compatibility between in-situ and satellite data observations. This emphasizes that the observation data sets have to be coherent in order to be efficiently assimilated into an OGCM. Nevertheless, we identified a region (corresponding approximately to the Warm Pool) where both the observed MDT and the assimilated solution have large errors, thus suggesting that salinity

effects should be considered more carefully and that the use of in-situ salinity data may be required. The use of an absolute dynamic topography opens the way for the development of efficient assimilation systems based on multiple data sets (using altimetry and various in-situ data). Further improvements in the knowledge of the MDT itself might result from converging studies on assimilated data resolution and accuracy, both from space and in situ, on modeling and on assimilation techniques, within the context of operational oceanography systems (*Dobricic, 2005*).

Acknowledgments. This work was supported by the Centre National d'Etude Spaciales (CNES), the Centre National de Recherche Scientifique (CNRS) and the Groupe Mission Mercator Coriolis (GMMC). The research took place as part of the EU funded MERSEA project. The calculations were conducted with the support of IDRIS/CNRS.

7.3 Conclusion

Ce chapitre a été l'occasion d'illustrer par une simulation les apports d'une assimilation multi-données par rapport à une assimilation mono-données. Les écarts RMS par rapport aux observations sont réduits (cf section 7.1.1) avec assimilation conjointe de température in-situ et d'altimétrie. L'utilisation d'une topographie dynamique observée, comme nous l'avons illustré dans la section 7.1 et dans la lettre publiée dans GRL et fourni dans la section 7.2.2, permet une meilleure utilisation de la complémentarité des données altimétriques et de température in-situ TAO. La meilleure compatibilité présente entre l'altimétrie référencée avec la topographie dynamique observée et les données in-situ des mouillages TAO (cf section 6.1) est en grande partie à l'origine des améliorations constatées entre l'expérience assimilant conjointement l'altimétrie référencée par la surface moyenne du modèle et les données de température in-situ et l'expérience assimilant conjointement l'altimétrie référencée par la topographie dynamique observée et les données de température in-situ. L'apport de la topographie dynamique observée GRACE dans la complémentarité des données altimétriques et in-situ est un résultat fort à mettre au crédit de l'utilisation de la MDT GRACE.

Nous avons également confirmé les résultats obtenus par *Parent (2000)* en assimilant conjointement l'altimétrie référencée par la surface moyenne du modèle et les données de température in-situ TAO. Avec ce type de configuration, on arrive à réduire en terme d'écarts RMS par rapport aux observations d'un des jeux de données assimilées mais pas des deux simultanément. Le fait que l'on ne puisse pas converger simultanément vers les deux jeux d'observation illustre le fait qu'il existe des incompatibilités entre les données in-situ et l'altimétrie lorsque la surface moyenne du modèle est utilisée comme référence. Ces incompatibilités sont la conséquence des erreurs sur le champ thermohalin moyen simulé par le modèle, erreurs qui se traduisent notamment par des erreurs sur la surface moyenne simulée. Dans l'expérience présenté dans la lettre GRL, on réduit la RMSD par rapport aux données TAO. Cela est uniquement du à la nature de l'assimilation de température qui est une des variables pronostiques du modèle et est directement contrainte par les incréments IAU calculés lors de l'analyse SEEK. Je n'ai pas fait le test, mais il est vraisemblable qu'en jouant sur la matrice de covariance des erreurs d'observation \mathbf{R} , on arrive à se rapprocher de l'une ou l'autre des sources d'observation en fonction de la précision relative donnée à chaque jeu d'observations. Sur le plan de dynamique, la

présence de petites structures bruitées sur la moyenne annuelle des courants de surface de l'expérience assimilant conjointement l'altimétrie référencée avec la surface moyenne du modèle et les profils in-situ de température TAO (cf figure 7.10) est un autre signe d'incompatibilité entre les données assimilées.

En conclusion, l'utilisation d'une topographie dynamique moyenne observée permet une meilleure utilisation de la complémentarité existant entre des données altimétriques qui nous fournissent une vision globale avec une résolution spatiale et une précision élevée de la surface de l'océan, et les données in-situ, qui sont plus rares, mais viennent compléter l'altimétrie et sont les seules capables de fournir des mesures sur la verticale. Cette meilleure complémentarité entre les données altimétriques référencées avec la topographie dynamique observée et les données in-situ est la conséquence directe de la meilleure compatibilité présente entre la topographie dynamique observée et les données in-situ, notamment les profils de température TAO.

Suite à ce chapitre, nous allons regarder un peu plus en détail la physique simulée avec et sans assimilation de données en utilisant deux simulations longues s'étalant sur la période 1993-1998. Cette période comprend notamment l'évènement El Niño de 1997-1998, qualifié d'El Niño du siècle, car il fut par beaucoup d'aspect le plus intense du *XX^{ème}*. La simulation avec assimilation de données retenue assimile conjointement l'altimétrie référencée avec la MDT GRACE et les profils de température in-situ issus de mouillages TAO.

Chapitre 8

Evaluation des apports de l'assimilation de données dans une expérience interannuelle sur la période 1993-1998

Sommaire

8.1	Présentation de l'expérience	136
8.2	Apports de l'assimilation : écarts RMS	137
8.2.1	Apports pour la partie observée	138
8.2.2	Apports par rapport à des données indépendantes	138
8.3	Apports de l'assimilation : comparaisons aux observations . . .	140
8.3.1	Niveau de la mer	140
8.3.2	Courants	144
	Courants de surface	144
	Courants de subsurface	145
8.3.3	Température	147
	Comparaison avec les données TAO assimilées	147
	Comparaison avec les données XBT indépendantes	147
	Evolution de la profondeur de l'isotherme 20°C	152
8.4	Discussion	153
8.5	Conclusion	155

Dans le chapitre 6, nous avons évalué différentes configurations de simulation avec assimilation de données. Nous avons montré la faisabilité de l'assimilation d'une topographie dynamique absolue définie comme étant la somme des anomalies de surface libre le long des traces et la topographie dynamique moyenne observée (déduite des données gravimétriques GRACE). Nous avons également montré dans le chapitre 7 que l'utilisation d'une telle topographie dynamique ouvrait les portes d'une assimilation conjointe de données altimétriques et in-situ efficace. La meilleure compatibilité entre les différents types d'observations apportée par l'utilisation d'une topographie dynamique moyenne observée, et donc par l'utilisation d'un signal de topographie dynamique absolue observé, permet de faire une meilleure utilisation de la complémentarité entre les observations satellites du niveau de la mer et les données in-situ de température.

Dans le chapitre précédent, nous avons vu que notre système assimilant conjointement les données altimétriques et les données in-situ TAO permettait de réduire assez fortement les erreurs en terme d'écart RMS relativement aux deux jeux d'observations pour l'année 1993. La structure moyenne et la variabilité de la circulation simulée sur cette période est également apparue plus en accord avec la circulation observée du Pacifique Tropical. Nous allons maintenant nous intéresser un peu plus finement à la physique simulée avec et sans assimilation de données afin d'illustrer les impacts de l'assimilation. Pour cela, nous avons réalisé une réanalyse de l'évolution de la dynamique de l'océan Pacifique Tropical sur la période 1993-1998. Cette période englobe en particulier l'épisode El Niño/La Niña de 1997-1998, qui fût probablement le plus intense du *XX^{ème}* siècle. Nous ne pourrions évidemment pas regarder tous les aspects de la dynamique du Pacifique Tropical car cela demanderait de trop nombreuses analyses qui ne peuvent pas être menées par une personne seule dans la durée impartie pour une thèse. Nous tenterons donc de donner une vision générale de la dynamique simulée avant de regarder un peu plus finement quelques aspects spécifiques de la dynamique de l'océan Pacifique Tropical simulée avec assimilation de données.

8.1 Présentation de l'expérience

Dans ce chapitre, nous allons examiner plus en détail une simulation avec assimilation conjointe de données couvrant la période 1993-1998. Une simulation libre est également réalisée sur cette même période afin de pouvoir évaluer les apports et les déficiences du système d'assimilation mis en place pour cette expérience. Les chapitres 6 et 7 nous ont permis de se faire une première idée des performances relatives des différentes configurations mises en place, notamment en ce qui concerne le système d'observation et la paramétrisation de l'espace d'erreur en rang réduit du filtre SEEK. Pour cette simulation sur la période 1993-1998, nous avons choisi d'utiliser une expérience reprenant les caractéristiques de l'expérience EXP4 présentée dans le chapitre précédent. Cette expérience nous paraît en effet la plus aboutie. Pour mémoire, je rappelle ici succinctement ses caractéristiques.

La simulation démarre au 1^{er} janvier 1993 à partir d'une condition initiale issue d'une simulation relaxée vers la climatologie de Lévitius afin de limiter l'erreur sur le champ de masse simulé et ainsi faciliter la tâche de l'assimilation. Un forçage de type bulk (cf chapitre 3) est utilisé pour forcer le modèle à l'interface océan-atmosphère. Les données atmosphériques (température à 10 m, humidité relative, etc...) utilisées pour diagnostiquer les flux à l'interface océan-atmosphère proviennent du centre NCEP. Le flux de quantité de

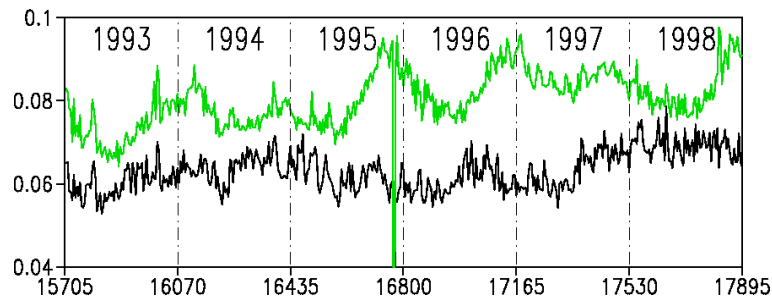


FIG. 8.1 – *Écarts RMS (en cm) par rapport à la topographie dynamique (SLA T/P+ERS + MDT) pour la simulation libre (en vert) et pour la simulation assimilée (en noir).*

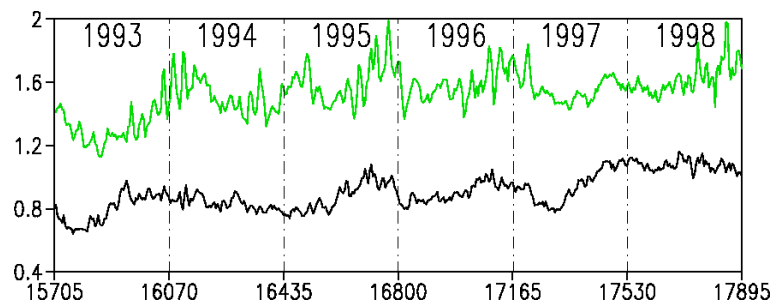


FIG. 8.2 – *Écarts RMS (en °C) par rapport aux profils de température TAO pour la simulation libre (en vert) et pour la simulation assimilée (en noir).*

mouvement est spécifié en utilisant les tensions de vents ERS+TAO proposées par *Menkes et al.* (1998). Ces tensions de vents sont connues pour produire une thermocline et des courants zonaux plus réalistes. Il n'y a pas de relaxation vers une SST observée. En revanche, le terme de relaxation vers la salinité de surface est conservé afin de prévenir d'éventuelles dérives de la SSS du modèle causées par les fortes incertitudes sur le flux d'eau douce. Pour ce qui est de la partie assimilation, le filtre SEEK est utilisé pour assimiler conjointement la topographie dynamique absolue mesurée par satellite et les données de température in-situ issues des mouillages TAO. La topographie dynamique absolue est définie comme étant la somme de la topographie dynamique moyenne GRACE (cf chapitre 2) et des anomalies de surface libre le long des traces mesurées par les satellites TOPEX/Poseidon et ERS1&2. La topographie dynamique observée est donc purement satellite et indépendante de toute observation in-situ. Le filtre implémenté utilise la variante dite stationnaire (cf chapitre 4) du SEEK. La paramétrisation de l'espace d'erreur en rang réduit retenue est la paramétrisation basée sur les écarts avec une simulation de référence relaxée vers la climatologie (cf chapitre 5).

8.2 Apports de l'assimilation : écarts RMS

Les résultats en terme de RMSD présentés dans le chapitre 7 nous ont déjà largement rassuré sur les performances du système d'assimilation conjointe DT+TAO mise en place pour cette expérience. Le but de cette section est de vérifier qu'il n'y a pas de dérive dans le temps de l'erreur par rapport aux observations. Il sera également intéressant de regarder

comment évolue la RMSD au cours du fort évènement interannuel El Niño/La Niña en fin de simulation.

8.2.1 Apports pour la partie observée

Les figures 8.1 et 8.2 montrent l'évolution temporelle des RMSD respectivement par rapport à la topographie dynamique assimilée et par rapport aux données de température assimilées. Le premier point intéressant est que la RMSD reste stable sur les 6 ans de la simulation. L'erreur de la simulation libre par rapport à la topographie dynamique est de 8.1 cm en moyenne sur les 6 années. L'assimilation permet de réduire cette erreur à 6.3 cm en moyenne sur la période 1993-1998. On note une légère augmentation de l'erreur pour le modèle libre au cours des 6 ans de simulation sur laquelle plusieurs maximas viennent se superposer. Par contre, l'erreur est stable pour la simulation avec assimilation de données, si ce n'est une légère augmentation pendant la période El Niño/La Niña de 1997-1998. Par rapport aux données de température TAO, la RMSD moyenne diminue de 1.54°C pour la simulation libre à 0.9°C pour la simulation avec assimilation de données. Par rapport à cette observation, la dérive de la RMSD sur la période 1993-1998 pour la simulation libre est moins évidente que par rapport à la topographie dynamique observée. Les maximas sont également moins marqués mais toujours présents. Il sont de plus bien corrélés avec les maximas de la RMSD par rapport à la topographie dynamique observée. Pour la simulation avec assimilation on retrouve une légère augmentation de l'erreur pendant la période El Niño/La Niña en fin de simulation.

8.2.2 Apports par rapport à des données indépendantes

Regardons maintenant comment a évolué l'erreur par rapport à des données indépendantes. La figure 8.3 montre les profils verticaux de RMSD par rapport aux données TAO et surtout par rapport aux données indépendantes de température XBT collectées sur tout le Pacifique Tropical par les navires VOS pendant la période 1993-1998. La position des profils est fournie sur la figure 2.9, page 37. Par rapport aux données TAO, nous avons fortement réduit l'erreur dans la région de la thermocline. Cela est dû au fait que l'assimilation a permis de correctement repositionner la thermocline qui est trop profonde dans la simulation libre, comme nous l'avons vu dans les deux chapitres précédents. Si l'on regarde le profil de RMSD par rapport aux données XBTs indépendantes, la réduction est moins forte. On passe d'une erreur moyenne de 1.64°C pour la simulation libre à 1.42°C pour la simulation assimilée. Il faut cependant se rappeler que les données XBTs sont réparties sur tout le domaine d'assimilation (cf figure 2.9) alors que les données TAO assimilées se limitent au seul guide d'onde équatorial (entre $\pm 8^\circ$ de latitude). Il y a donc une grande partie du domaine où l'on dispose de données XBTs pour évaluer nos simulations, mais où l'on ne dispose pas de données TAO afin de contrôler la température de subsurface. La figure 8.4 montre que dans la boîte Niño 3.4 (5°S - 5°N , 170°W - 120°W , donc dans une partie de notre domaine d'assimilation échantillonnée par le réseau TAO), la réduction de l'erreur est forte. Dans cette région, le modèle libre présente une erreur importante supérieure à 3°C au niveau du cœur de la thermocline. Cette erreur est réduite de plus de 1°C par l'assimilation de données.

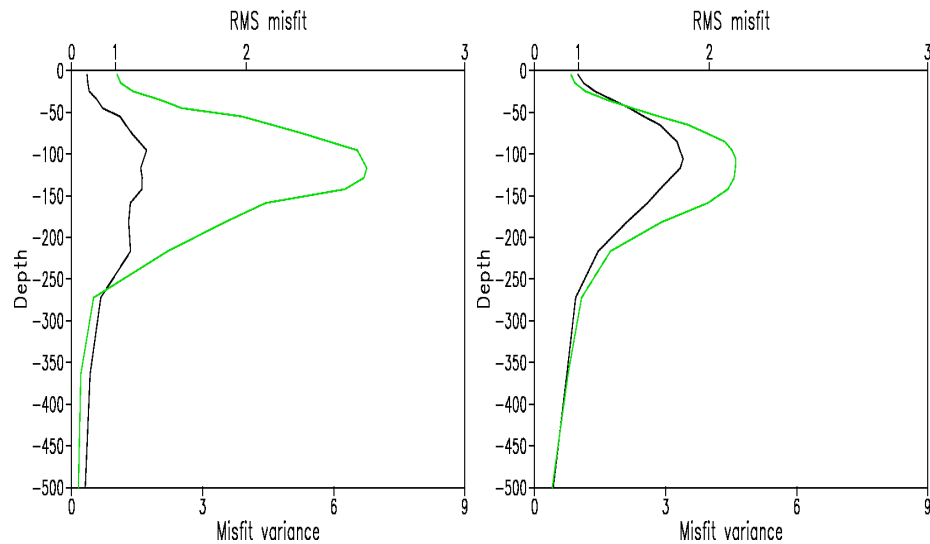


FIG. 8.3 – Profils d'écarts RMS (en °C) sur la verticale par rapport aux profils de température TAO (à gauche) et par rapport aux profils de température XBTs indépendantes collectés par les navire VOS (à droite) pour tout le Pacifique Tropical. Les RMSD pour la simulation libre (en vert) et pour la simulation assimilée (en noir) sont tracées.

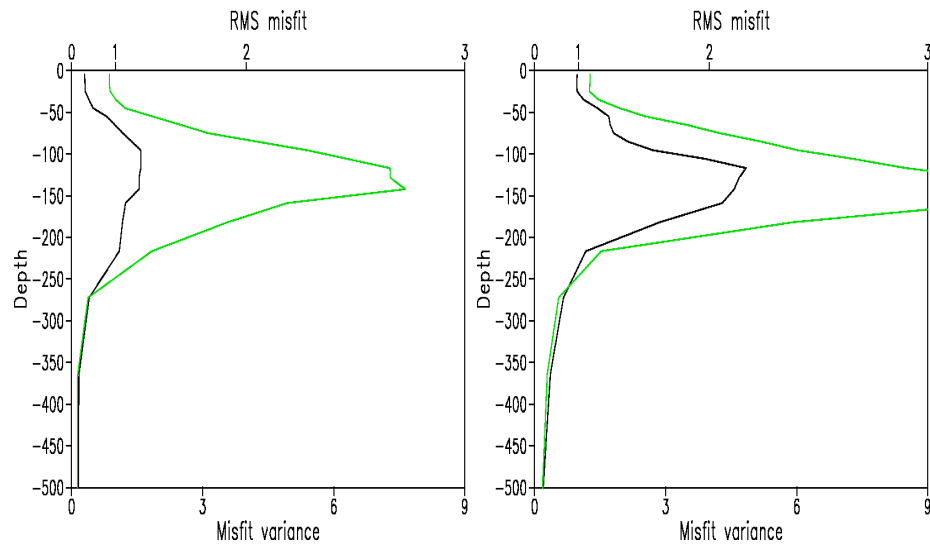


FIG. 8.4 – Identique à la figure 8.3 mais pour la boîte Niño 3.4 ($5^{\circ}S-5^{\circ}N$, $170^{\circ}W-120^{\circ}W$).

8.3 Apports de l'assimilation : comparaisons aux observations

Dans la section précédente, nous avons pu nous rendre compte que les bonnes performances en termes d'écart RMS que nous avons mis en évidence dans les chapitres 6 et 7 pour l'année 1993 se maintiennent sur la période pluri-annuelle 1993-1998. Notamment, pendant le fort El-Niño de 1997-1998, on ne constate qu'une très légère augmentation de l'erreur RMS par rapport à la topographie dynamique et aux températures TAO assimilées. Maintenant que la stabilité de notre système est avérée, nous allons nous intéresser aux grandeurs simulées avec et sans assimilation de données sur un plan plus physique.

8.3.1 Niveau de la mer

Le premier point à regarder, compte tenu du fait que l'on assimile la topographie dynamique absolue, est le niveau moyen et la variabilité du niveau de la mer simulée. Comme on pouvait s'y attendre au vu des résultats sur l'année 1993 uniquement, la topographie dynamique moyenne simulée a été fortement modifiée par l'assimilation de données (cf figure 8.5). Les structures de MDT simulées avec assimilation de données sont beaucoup plus en accord avec les structures de MDT observées, notamment dans la région du Pacifique Central Sud, où la topographie dynamique moyenne est plus réaliste et diffère assez sensiblement de la topographie dynamique de la simulation libre. La région de la dorsale et du sillon Nord équatorial est également significativement modifiée par l'assimilation. La dorsale est plus marquée et le sillon Nord plus creusé ce qui est en accord avec les observations. Dans le chapitre 6, nous avons constaté que le sillon du NECC sur la topographie dynamique moyenne de la simulation libre pour l'année 1993 était bien creusé. Ce sillon Nord est beaucoup moins marqué sur la MDT de la simulation libre pour la période pluri-annuelle 1993-1998. Ceci confirme l'hypothèse avancée dans le chapitre 6 pour expliquer les relativement très bonnes performances du modèle libre. Hypothèse selon laquelle le champ thermohalin du modèle libre n'avait pas significativement dérivé par rapport au champ thermohalin de la condition initiale (et donc de la climatologie) sur une seule année de simulation.

Toujours dans la région de la dorsale et du sillon Nord équatorial, la moyenne de la topographie dynamique simulée avec assimilation de données présente des petites structures qui semblent peu réalistes pour une moyenne sur 6 années de simulation. On ne retrouvait pas ce type de structure pour la moyenne sur la première année de simulation uniquement (cf section 7.1). Nous avons vu dans la section 6.1 que les erreurs les plus marquées entre la hauteur dynamique issues des mouillages TAO et la topographie dynamique observée calculée en sommant la MDT GRACE et la SLA T/P+ERS se trouvaient précisément dans cette région. Ceci n'est d'ailleurs pas très surprenant compte tenu du fait que les gradients méridionaux sont particulièrement forts dans la région du sillon Nord et que la résolution de la topographie dynamique moyenne GRACE (333km) reste encore un peu faible pour les représenter correctement. Les petites structures que l'on voit sur la MDT de la simulation avec assimilation de données (figure 8.5) dans la région de la dorsale et du sillon Nord équatorial sont probablement la signature des incompatibilités qui persistent entre les deux jeux de données d'observations assimilés dans cette région. D'autant que c'est également dans cette région que le modèle libre produit les plus fortes erreurs en terme de topographie dynamique simulée. Ce point sera de nouveau abordé un peu plus

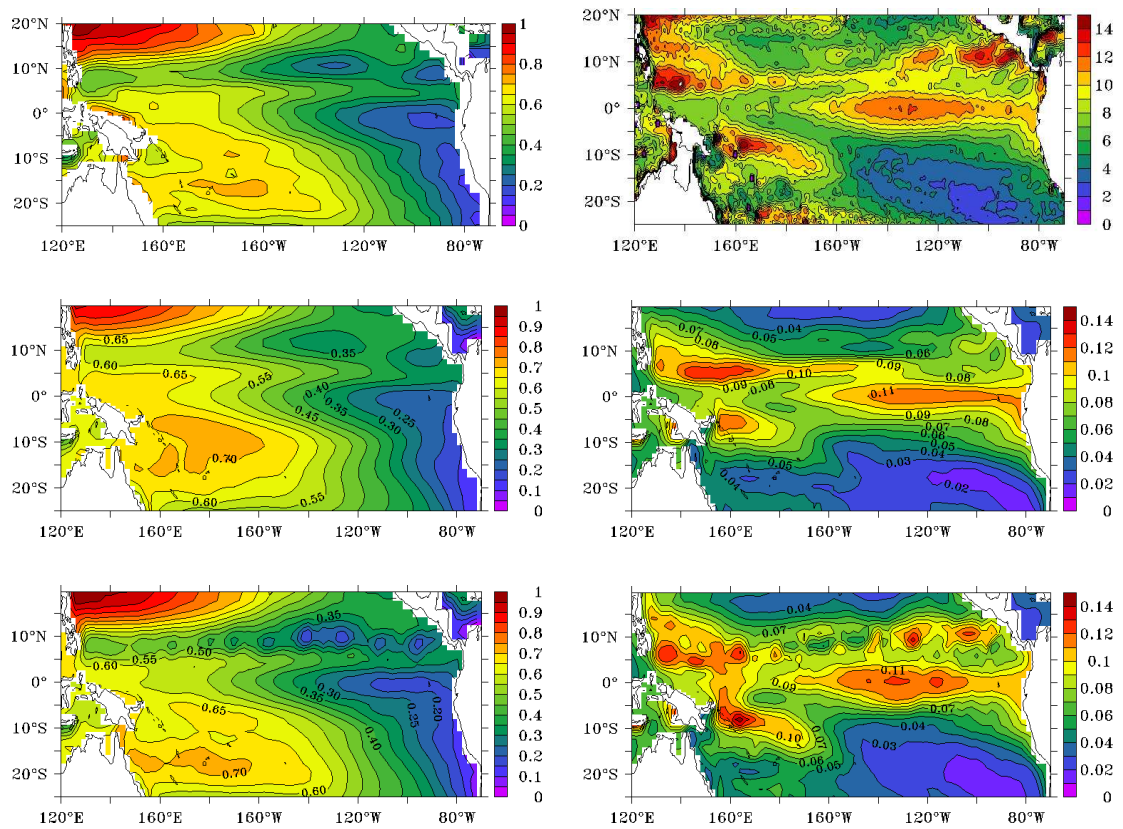


FIG. 8.5 – *MDT (à gauche en m) et écart type de la DT (à droite en m) observés, de la simulation en mode libre, de la simulation avec assimilation de données (respectivement de haut en bas).*

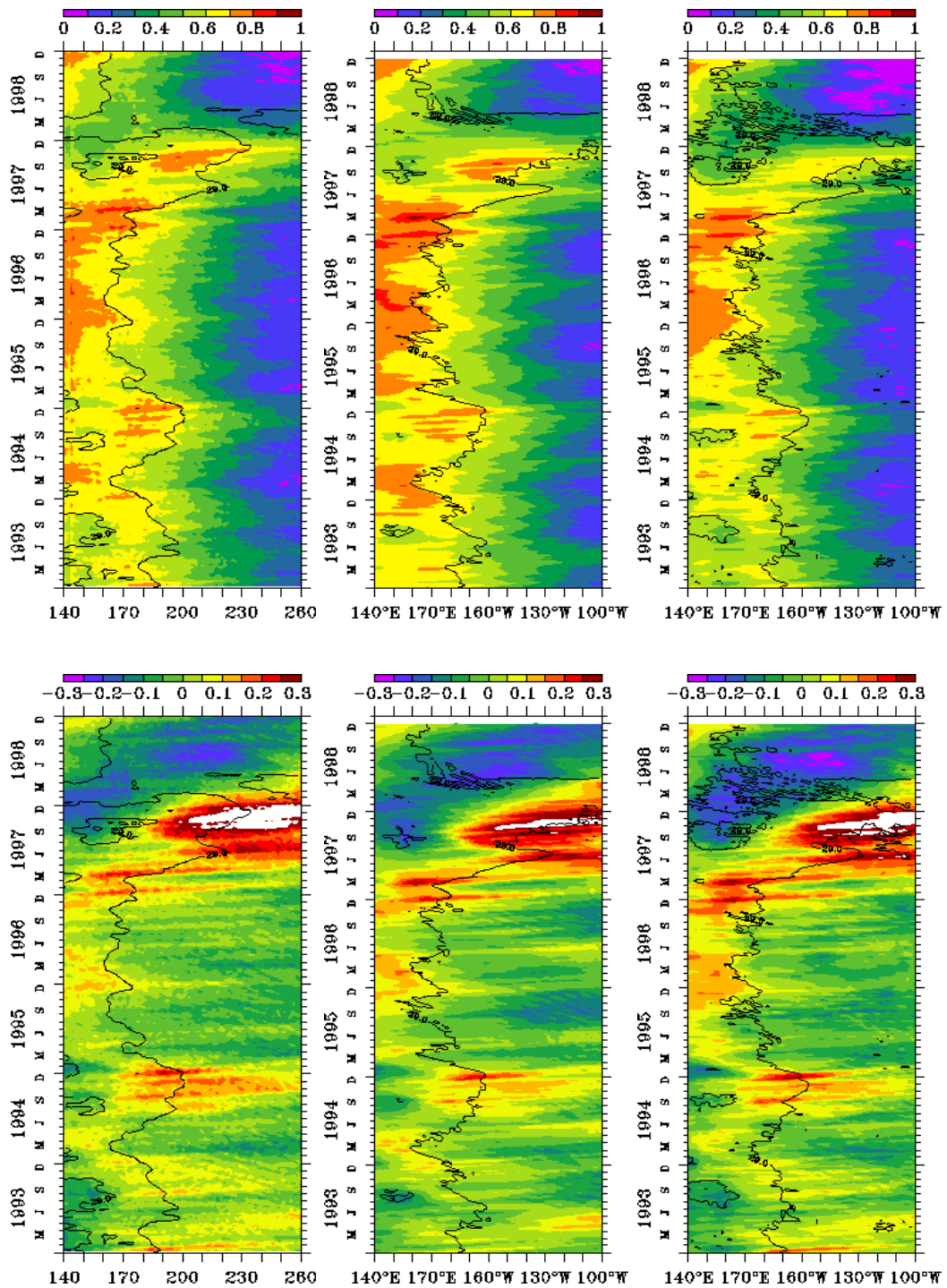


FIG. 8.6 – Diagramme longitude-temps à l'équateur de la topographie dynamique (en haut, en m) et de la SLA (en bas, en m). L'évolution de la position de l'isotherme 29°C est superposée sur chaque diagramme (pour les observations, la SST NCEP Reynolds hebdomadaire proposée par Reynolds et Smith (1994) est utilisée). Observation, à gauche. Simulation en mode libre, au milieu. Simulation avec assimilation de données, à droite.

tard dans la section 8.4.

En ce qui concerne la variabilité observée, elle est dominée par le fort El-Niño de 1997, avec un fort maximum de variabilité centré sur l'équateur dans l'est du bassin. Ce maximum correspond à la signature des ondes de Kelvin qui enfoncent la thermocline et se propagent d'Ouest en Est en contribuant à advecter les eaux chaudes du Pacifique Ouest vers l'Est du bassin lors de la phase croissante du phénomène. La simulation libre représente la variabilité du niveau de la mer de manière très satisfaisante. Le maximum de variabilité à l'équateur est bien présent mais il est cependant un peu sous-estimé. L'assimilation de données permet de renforcer ce maximum de variabilité et de le repositionner plus au centre du bassin ce qui est en accord avec les observations. Dans la simulation libre, le maximum de variabilité au niveau de l'équateur s'étend jusqu'à la côte américaine traduisant une sur-estimation de la variabilité saisonnière dans la région de la "cold tongue". Cette erreur est bien corrigée par l'assimilation. Un autre point positif de l'assimilation est le renforcement de la variabilité dans l'Est, à 10°N associé avec le fort pompage d'Ekman dans cette région. Ce point est d'autant plus positif que ce maximum de variabilité est quasi inexistant dans la simulation libre. Dans l'Ouest, on a diminué l'extension zonale du maximum de variabilité associé avec la propagation d'ondes de Rossby le long de 5°N, ce qui est en accord avec les observations. La variabilité simulée avec assimilation de données dans le Pacifique Sud tropical est globalement plus réaliste, avec une région de très faible variabilité dans le Sud-Est et une variabilité plus forte dans le Sud-Ouest. En revanche, un point négatif de l'assimilation de données est la trop forte variabilité au niveau de l'équateur dans l'Ouest du bassin, au niveau de 160°E.

L'évolution de la topographie dynamique et la propagation des anomalies de SLA le long de l'équateur est un très bon indicateur de la variabilité tant annuelle qu'interannuelle de la dynamique équatoriale, l'axe zonal étant privilégié. La figure 8.6 montre les diagrammes de Hovmueller (diagramme longitude-temps) de la DT et de la SLA pour les observations et pour les simulations libre et avec assimilation de données. Les petites structures présentes dans les diagrammes de la DT et la SLA observées sont liées à la résolution de la SLA (1/3°). L'El Niño de fin 1994, et surtout celui de 1997 sont très clairement visibles ainsi que La Niña 1998. La simulation libre est déjà de très bonne qualité. Cependant, on peut noter que le niveau de la mer simulé dans l'Ouest du bassin est généralement trop important. On retrouve également le fait que la variabilité annuelle est surestimée par le modèle libre dans l'Est du bassin avec des anomalies négatives trop importantes en Sept./Oct./Nov. pour les années 1993, 1995 et 1996. Cette période correspond au moment où les alizés de Sud-Est sont les plus intenses et donc l'upwelling équatorial le plus vigoureux. L'assimilation de données permet de bien corriger ces erreurs.

L'El Niño de 1997 dans la simulation libre est très bien représenté en terme de phase et l'amplitude n'est que très légèrement sous-estimée. Les vents ERS-TAO utilisés pour forcer le modèle y sont sûrement pour beaucoup. L'assimilation de données améliore le réalisme de l'anomalie positive de SLA lors de l'El Niño de 1997. Elle semble par contre avoir des effets plus mitigés lors de la transition de la phase chaude à la phase froide La Niña.

En résumé, l'impact de l'assimilation de données sur la topographie dynamique simulée sur la période 1993-1998 est globalement positif. Le réalisme du niveau de la mer simulé en terme de moyenne et de variabilité est amélioré. Notamment, la MDT simulée avec assimilation de données est plus en accord avec les observations, en particulier dans le Pacifique Central Sud et dans la région de la dorsale et du sillon Nord équatorial. La

variabilité ondulatoire est plus proche de celle observée par satellite. La variabilité annuelle est également améliorée, en particulier dans l'Est où elle est assez nettement sur-estimée par le modèle libre. L'apparition d'une zone de forte variabilité dans la simulation avec assimilation de données vers 10°N dans l'Est est également un point positif à mettre au crédit de l'assimilation. En revanche, la simulation avec assimilation de données sur-estime la variabilité dans le Pacifique Ouest à 160°E au niveau de l'équateur. L'assimilation fait également apparaître des petites structures irréalistes sur la topographie dynamique moyenne simulée dans la région de la dorsale et du sillon Nord équatorial. Ces petites structures sont vraisemblablement à relier aux incompatibilités qui persistent entre les données altimétriques et les données in-situ des mouillages TAO dans cette région. Ce point sera de nouveau abordé dans la section 8.4.

8.3.2 Courants

Courants de surface

Comme je l'ai dit dans le chapitre précédent, l'examen des courants de surface est un test rigoureux et révélateur dans les régions tropicales, les courants étant particulièrement sensibles aux gradients méridionaux de topographie dynamique avec la disparition de la force de Coriolis à l'équateur. La figure 8.7 montre les courants zonaux moyens à 15 m de profondeur pour la période 1993-1998. Je rappelle que le lecteur peut trouver une climatologie des courants zonaux à la même profondeur sur la figure 6.8, page 109. Les courants simulés avec assimilation de données sont beaucoup plus réalistes que les courants de la simulation libre. Le NECC a été significativement intensifié ce qui est en accord avec la climatologie de Niiler (2001). Le SEC est également plus intense. On observe bien la séparation du SEC en deux branches, la branche Nord et la branche Sud, dans le centre du Pacifique Tropical. En revanche, cette séparation est mal représentée dans l'extrême Est du bassin se qui se traduit par des vitesses en surface à l'équateur trop importantes. Ceci est un point négatif de l'assimilation de données. La mauvaise représentation de cette séparation du SEC en deux branches est un des défauts connus de la simulation ORCA2 (Lengaigne et al., 2003) et l'assimilation conjointe d'altimétrie et de données in-situ ne parvient visiblement pas à le corriger. Au Nord, le NEC atteint des vitesses supérieures à 20 cm s^{-1} ce qui est encore un peu faible par rapport à la climatologie déduite des bouées dérivantes SVP mais représente une nette amélioration par rapport aux vitesses moyennes du NEC dans la simulation libre. Dans l'hémisphère sud, l'assimilation de données se traduit par une amélioration assez nette du réalisme des courants de surface simulés avec un SECC qui reste bien confiné dans l'Ouest du bassin, au delà de 180° . Les courants vers

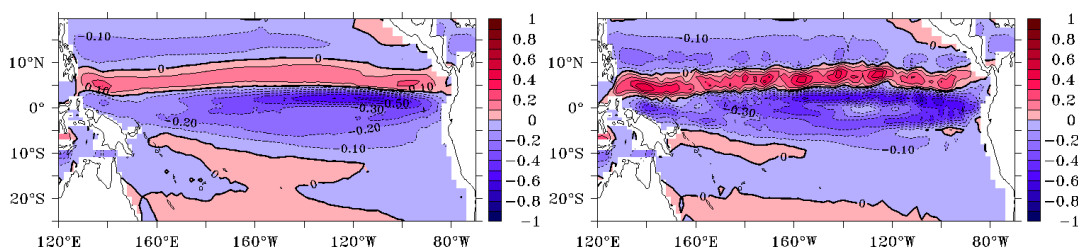


FIG. 8.7 – Vitesse zonale moyenne pour la période 1993-1998 pour la simulation libre (à gauche) et pour la simulation assimilée (à droite).

l'Est associés à la branche de retour du gyre subtropical sud sont présents uniquement au sud de 20°S comme dans la climatologie.

Courants de subsurface

L'EUC est une pièce importante de la dynamique du Pacifique Tropical (cf chapitre 1). Ce courant est très nettement visible sur la figure 8.8 montrant une section de vitesse zonale moyenne à l'équateur. L'EUC est plutôt bien simulé par le modèle libre en terme d'intensité. Cela n'est pas très surprenant, la pente zonale de la thermocline au niveau de l'équateur, et donc le gradient zonal de pression au niveau de la thermocline à l'équateur, étant bien simulée par le modèle libre (cf chapitre 6). Ce gradient étant le principal moteur de l'EUC (cf chapitre 1) il est normal que l'intensité de l'EUC soit correcte. On remarque des différences dans l'Ouest. Dans cette région, le cœur du courant simulé avec assimilation de données est plus intense et aussi moins profond. Ce dernier point est en accord avec la remontée des isothermes de la simulation avec assimilation de données (cf chapitre 7 et section suivante).

Afin de quantifier l'impact de l'assimilation sur les courants zonaux en profondeur, nous avons évalué nos simulations à l'aide des courants mesurés par les ADCP équipant certains mouillages TAO le long de l'équateur. Les profils moyens de courants zonaux sont calculés pour la période 1993-1998 et des comparaisons sont faites pour les longitudes 156°E, 165°E, 170°W, 140°W et 110°W qui sont les seules équipées de capteur ADCP. La figure 8.9 confirme ce que nous avons vu sur la figure 8.8. La structure verticale des courants à l'équateur est plutôt bien simulée, indépendamment du fait que l'on utilise l'assimilation de données ou non pour corriger le modèle excepté pour le profil à 156°E. Cependant, en ce point, nous n'avons des données que pour la période allant de février 1993 à novembre 1994. Les sorties modèles étant moyennées sur la période 1993-1998, on ne peut pas les comparer directement. Pour cela, il aurait fallu calculer les profils moyens pour la simulation libre et la simulation avec assimilation de données en utilisant uniquement les sorties lorsque des observations concomitantes sont disponibles. Cependant, les observations de courant zonal pour les 4 autres points de mesures étant continues dans le temps sur toute la période de simulation nous avons simplement calculé la moyenne temporelle sans se soucier de la disponibilité des observations. Cette approximation est justifiée pour les profils 165°E, 170°W, 140°W et 110°W, mais pas pour le profil situé à 156°E ce qui explique très certainement les différences assez marquées entre les courants observés et simulés en ce point.

Le maximum de vitesse de l'EUC simulé en mode libre est systématiquement légèrement trop faible. L'assimilation de données permet de limiter cette erreur. En particulier, le démarrage de l'EUC dans l'Ouest du bassin est amélioré et les vitesses simulées avec assimilation de données sont comparables aux observations in-situ de courant. Dans l'Est en revanche, pour le profil situé à 110°W, l'assimilation a un impact négatif sur la structure de courant simulée. L'assimilation a légèrement diminué l'intensité du cœur de l'EUC qui était déjà trop faible dans la simulation libre, mais a surtout dégradé la représentation des courants au dessus du cœur de l'EUC, avec des vitesses vers l'Ouest en surface beaucoup trop intenses par rapport aux observations TAO. Ces vitesses trop élevées sont la conséquence de la mauvaise représentation dans l'Est du bassin de la séparation en deux branches du SEC que l'on a constatée sur la carte des courants à 15 m de profondeur (cf figure 8.7).

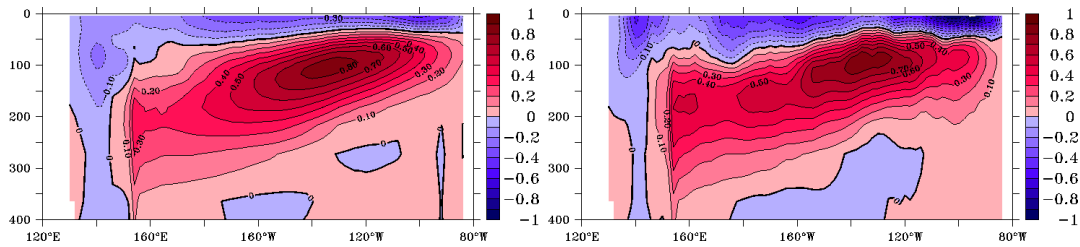


FIG. 8.8 – Section moyenne de vitesse zonale à l'équateur pour la simulation libre (à gauche) et la simulation avec assimilation de données (à droite).

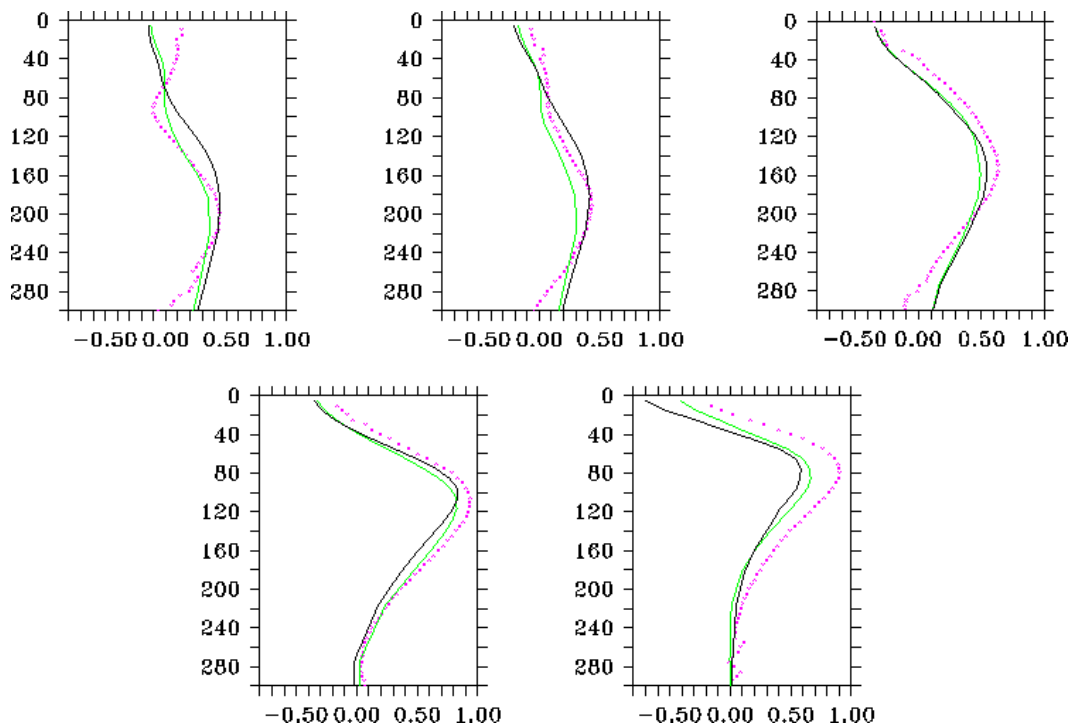


FIG. 8.9 – Profils moyens de vitesse zonale le long de l'équateur pour les longitudes $156^{\circ}E$, $165^{\circ}E$, $170^{\circ}W$, $140^{\circ}W$ et $110^{\circ}W$, respectivement de droite à gauche et de haut en bas. Profils pour la simulation libre (en vert), la simulation avec assimilation de données (en noir) et les observations ADCP TAO (en magenta).

8.3.3 Température

Comparaison avec les données TAO assimilées

Nous avons vu dans le chapitre précédent que l'assimilation conjointe des données de température TAO et de topographie dynamique se traduisait par une forte modification de la structure verticale de la simulation libre pour l'année 1993. La figure 8.10 confirme ce fort impact de l'assimilation de données. La figure 8.10 montre les écarts types de température pour les observations, la simulation libre et la simulation avec assimilation de données, sur la période 1993-1998 pour une section zonale à l'équateur et une section méridienne à 155°W. Les isocontours des champs de température moyens sont superposés. La variabilité de la simulation avec assimilation de données est plus en accord avec la variabilité mesurée par les mouillages TAO. Elle a été renforcée par rapport à celle de la simulation libre, mais reste toujours inférieure à la variabilité observée, en particulier dans l'Ouest du bassin. Cette modification de la variabilité du champ de température simulé s'accompagne d'une modification importante de la position des isothermes du champ moyen. Comme on l'avait constaté sur l'année 1993, les isothermes de la simulation libre sont globalement trop profondes d'une vingtaine de mètres. Cette erreur du modèle libre est corrigée par l'assimilation de données. En particulier, dans l'est du bassin sur la coupe zonale à l'équateur, la structure de la thermocline a été sensiblement modifiée, avec une thermocline beaucoup plus pincée que dans la simulation libre. La section méridienne à 155°W montre également une remontée des isothermes dans la simulation avec assimilation de données ainsi qu'une augmentation de la variabilité par rapport à la simulation libre, augmentation qui est en accord avec les observations. On note également une forte modification de la structure thermique simulée vers 8°N avec une thermocline beaucoup plus pincée et beaucoup moins profonde.

En résumé, les améliorations significatives sur le champ de température simulé avec assimilation de données conjointe altimétrique et in-situ entrevue sur l'année 1993 dans le chapitre 7, se confirment sur une période plus longue de 6 ans. La structure moyenne de la variabilité du champ de température simulé avec assimilation de données est en accord avec les observations de température assimilées. Dans la section suivante, nous allons nous assurer que cela est également le cas pour des données non assimilées.

Comparaison avec les données XBT indépendantes

Les données XBT ont une couverture spatio-temporelle très irrégulière. Il faut donc être prudent lorsque l'on travaille avec ce type de données, en particulier lorsque l'on veut les utiliser afin d'estimer la variabilité ou l'état moyen de l'océan. Les données XBT sont bien trop disparates pour calculer un état moyen global qui soit représentatif de la réalité moyenne du Pacifique Tropical. Il existe en revanche des lignes maritimes beaucoup mieux échantillonnées par les navires VOS. Nous avons sélectionné trois de ces lignes (cf figure 8.11), respectivement dans le Pacifique Ouest (Tokyo-Auckland), le Pacifique Central (San Francisco-Auckland) et le Pacifique Est (Panama-Auckland). Ces lignes commerciales maritimes particulièrement fréquentées vont être utilisées afin de calculer des sections verticales moyennes de température sur la période 1993-1998 représentatives du Pacifique Ouest, Central et Est respectivement. La figure 8.12 montre les diagrammes latitude-temps de la position des profils collectés le long des trois rails sélectionnés. Cette figure montre que les trois rails retenus sont bien échantillonnés sur toute la période de

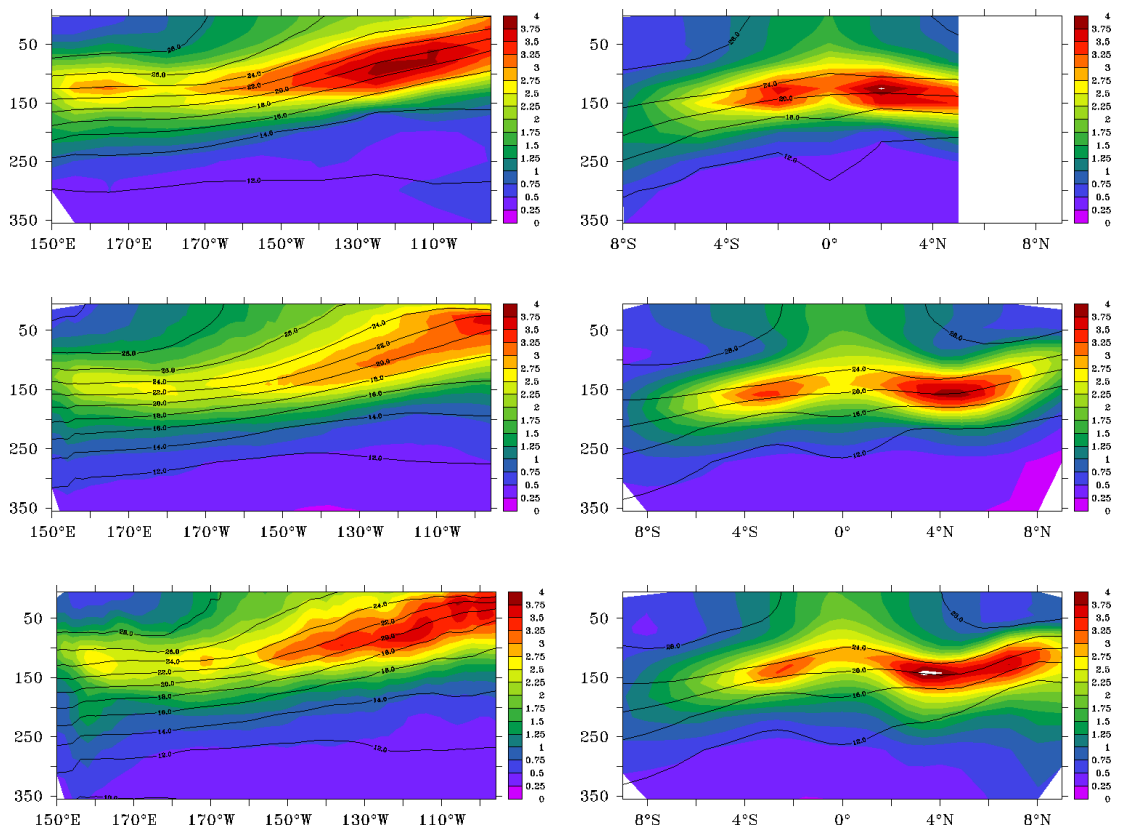


FIG. 8.10 – *Ecart type de température le long de l'équateur et pour une section méridienne à 155° W pour les observations TAO (en haut), la simulation libre (au milieu) et la simulation libre avec assimilation de données (en bas). Les isocontours du champ moyen de température sont superposés.*

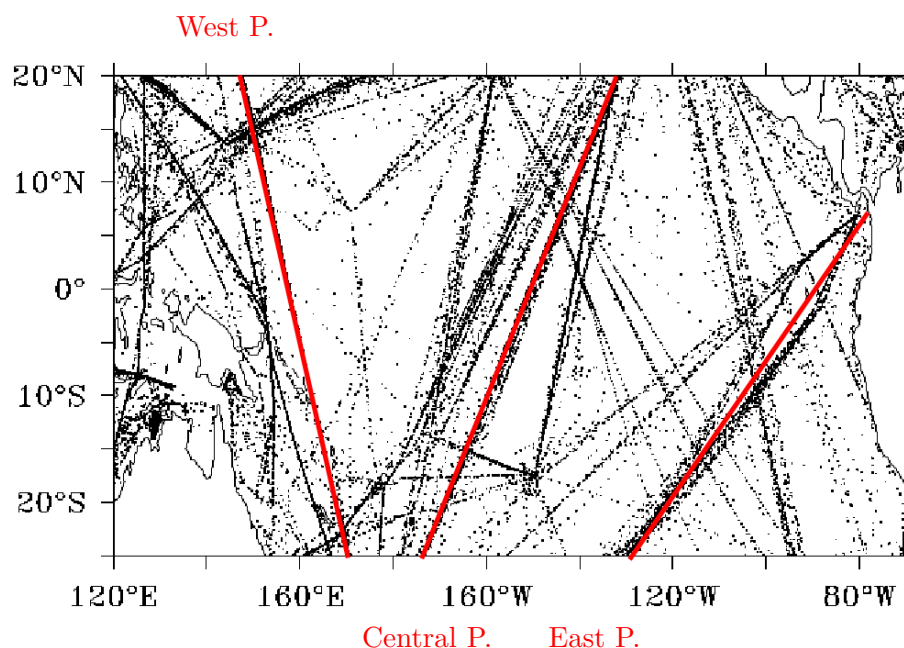


FIG. 8.11 – Position des rails West Pacific, Central Pacific et East Pacific échantillonné de manière régulière par les navire VOS. Ces rails correspondent respectivement aux lignes maritimes Tokyo-Auckland, San Francisco-Auckland et Panama-Auckland.

simulation. Ce bon échantillonnage spatio-temporel nous permet de s'assurer que les sections moyennes de température calculées à partir de ces données sont représentatives de la réalité du Pacifique Tropical sur cette période.

Les équivalents modèle des trois sections définies sur la figure 8.11 ont été extraits des sorties de la simulation libre et de la simulation avec assimilation de données. La figure 8.13 montre la différence de température moyenne le long de chacun des rails sur la période 1993-1998. Les isothermes du champ de température moyen observé par les XBTs sont superposées.

Les erreurs par rapport aux trois sections de température moyenne sont réduites de manière sensible, en particulier dans la région de la thermocline dans le guide d'onde équatorial où les erreurs de la simulation libre atteignaient localement plus de 5°C. Cela confirme que l'assimilation a permis de corriger, en partie au moins, l'erreur sur la position moyenne de la thermocline. Une autre constatation intéressante est que les erreurs par rapport à ces observations indépendantes ont été réduites même hors de la région contrôlée par les TAO. Ceci confirme que l'assimilation de topographie dynamique seule, même si elle ne parvient pas à corriger toutes les erreurs en profondeur, ne dégrade pas la représentation du champ de température en subsurface et bien au contraire permet même de corriger certaines erreurs, en particulier les erreurs sur les gradients à grande échelle comme nous l'avons vu avec l'expérience EXP2 dans le chapitre 6.

Le long du rail Central Pacific, l'assimilation de données fait apparaître un patch d'erreur assez forte en profondeur, sous l'isotherme 12°C. Cette erreur est très localisée et se situe aux environs de 10°N. Si l'on reprend la figure 8.12, on peut voir que ce point correspond à la longitude 140°W, soit juste au niveau d'un des mouillages du réseau TAO (9°N-140°W). Il semble donc que l'assimilation conjointe des données altimétriques et des

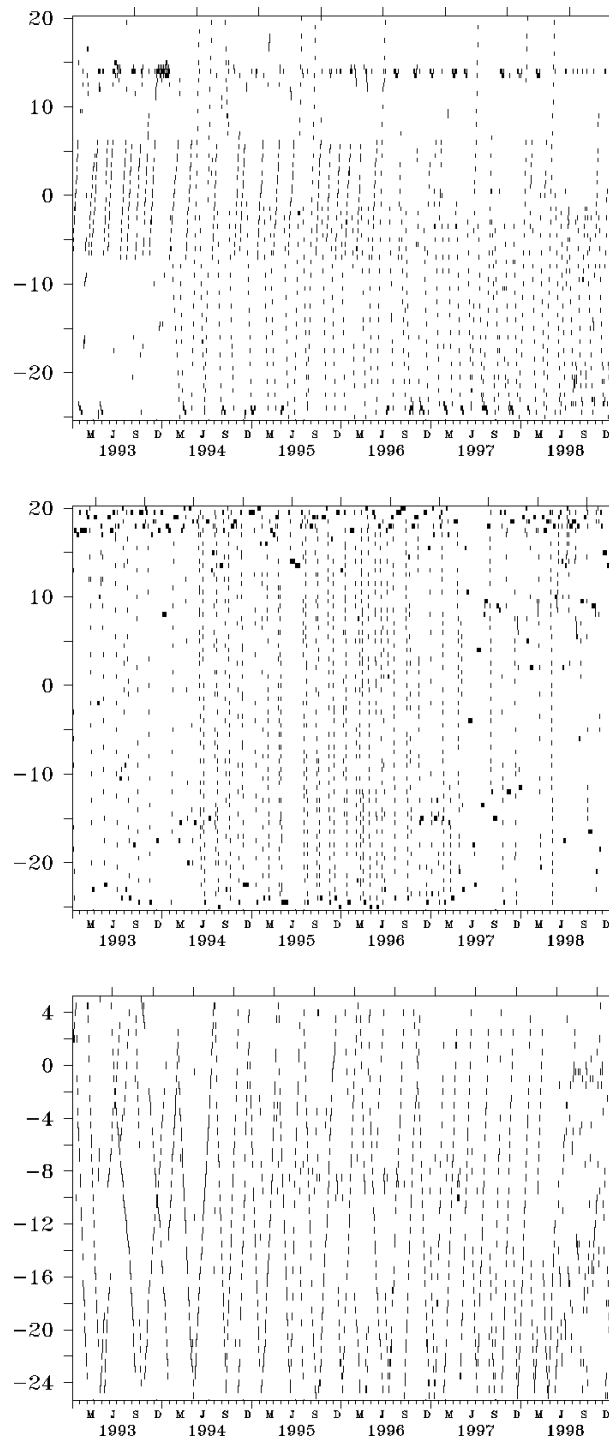


FIG. 8.12 – *Position des profils collectés le long des rails West P., Central P. et East P. respectivement de haut en bas en fonction du temps.*

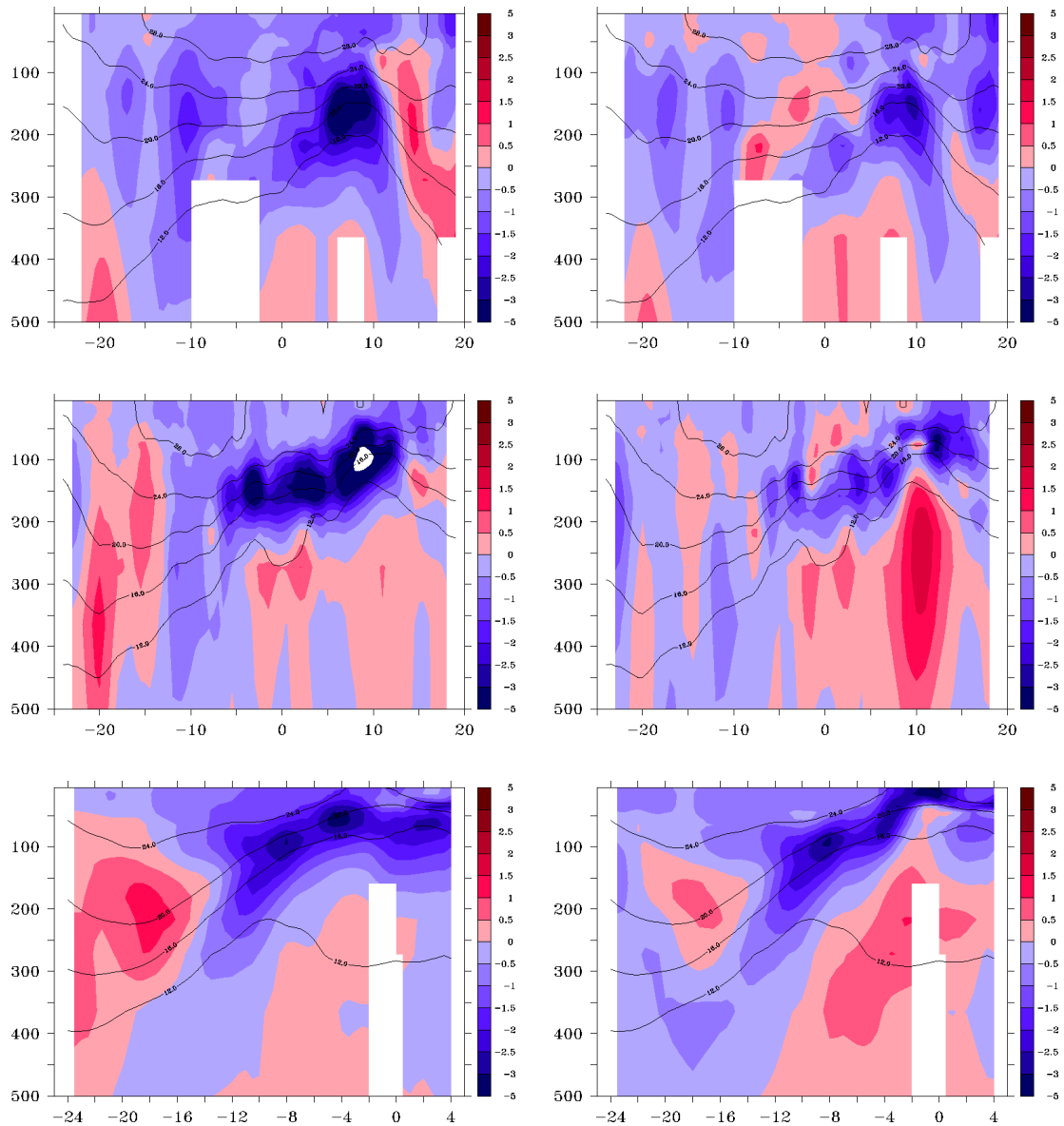


FIG. 8.13 – Différences de température moyenne sur la période 1993-1998 entre les données XBTs et la simulation libre (à gauche) et entre les données XBTs et la simulation avec assimilation de données (à droite) le long des rails West Pacific, Central P. et East P. définis sur la figure 8.11. Les isocontours du champ moyen de température observé par les XBTs sont superposés.

données de température in-situ TAO soit en cause. La figure 8.5, montrant des petites structures irréalistes sur la topographie dynamique moyenne simulée avec assimilation de données pour la période 1993-1998, nous avait déjà laissé penser qu'un problème persistait dans cette région. La présence de ce patch d'erreur localisé en profondeur au niveau d'un mouillage TAO semble également aller dans le sens d'un problème de compatibilité dans cette région. Dans la section suivante, nous nous intéresserons plus particulièrement à cette région afin d'essayer de mieux comprendre ce qui se passe le long de 8°N.

En résumé, nous avons confirmé, par rapport à des données indépendantes, les améliorations significatives sur le champ de température simulé avec assimilation conjointe de données altimétriques. Nous avons également confirmé l'impact positif de l'assimilation sur le champ de température simulé hors de la région contrôlée par les TAO. Par ailleurs, la figure 8.13 fait apparaître un patch d'erreur très localisé, juste sous un des mouillages TAO. Ceci semble confirmer le fait qu'il persiste un problème de compatibilité le long de 8°N comme l'examen de la figure 8.5 l'avait déjà laissé présager.

Evolution de la profondeur de l'isotherme 20°C

Nous avons pu constater de fortes modifications (de l'ordre d'une vingtaine de mètres pour l'isotherme 20°C) de la position moyenne des isothermes à l'équateur entre la simulation libre et la simulation assimilant conjointement la DT T/P+ERS et les données TAO. Cette modification est en accord avec le champ moyen de température observée par les données TAO assimilées, mais également avec les données indépendantes XBT collectées par les navires VOS. Nous allons maintenant voir si l'assimilation a un impact sur l'évolution de la profondeur de la thermocline. Dans le Pacifique Tropical, la profondeur de l'isotherme 20°C est un bon proxy de la profondeur de la thermocline. La figure 8.14 montre l'évolution de la profondeur de l'isotherme 20°C moyennée entre 2°S et 2°N pour les observations TAO (à gauche) et simulée avec (à droite) et sans assimilation de données (au centre). La première chose qu'il convient de remarquer est la qualité de la simulation libre. La plupart des structures sont présentes. On note la signature sur la profondeur de l'isotherme 20°C de l'El Nino de 1994 avec un approfondissement de la thermocline dans l'Est du bassin et celle, beaucoup plus marqué de l'El Nino de 1997 pendant lequel la thermocline devient plate sur toute la largeur du bassin. On note aussi la signature de deux ondes de Kelvin particulièrement intenses au début de l'El Nino de 1997 (en janvier 1997 et en février/mars de la même année). Ces deux ondes ont probablement joué un rôle important dans le démarrage de l'El Niño de 1997 (*Lengaigne et al., 2002*), en particulier l'onde de Kelvin générée suite aux anomalies de vents d'Ouest de février/mars qui a enfoncé la thermocline sur toute la largeur du bassin (cf figure 8.14). Là encore la qualité des vents ERS+TAO y est probablement pour beaucoup. L'assimilation de données permet d'améliorer encore un peu plus le réalisme de la simulation. La profondeur de l'isotherme 20°C est systématiquement trop importante dans la simulation libre comme nous avons déjà eu l'occasion de le constater à plusieurs reprises. Cette erreur est corrigée par l'assimilation des données TAO. On voit aussi que la variabilité saisonnière de la profondeur de la thermocline est surestimée par le modèle libre dans l'Est du bassin alors que dans la simulation avec assimilation de données, cette variabilité paraît plus réaliste. Au printemps 1998, pendant le début de la phase La Nina, la profondeur de la thermocline simulée avec assimilation de données est assez nettement moins profonde que dans la simulation libre, surtout dans le Pacifique Central ce qui est en accord avec les observations.

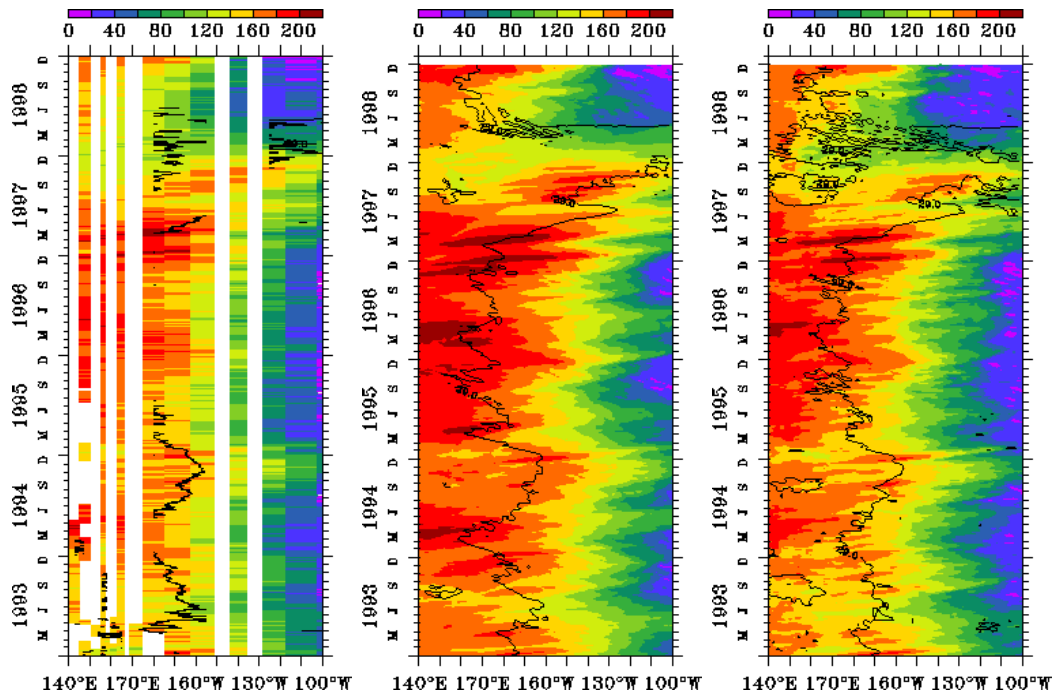


FIG. 8.14 – Diagramme longitude-temps à l'équateur de la profondeur de l'isotherme 20°C . L'évolution de la position de l'isotherme 29°C est superposée.

En résumé, l'assimilation de données a permis de corriger l'erreur sur la profondeur moyenne de la thermocline comme nous avons déjà pu le constater avec les sections verticales de température. L'évolution de la profondeur de la thermocline est en revanche très bien simulée par le modèle libre. L'assimilation permet tout de même quelques améliorations, en particulier sur le cycle saisonnier dans l'Est qui est surestimé avec une thermocline trop profonde en été/automne, lorsque les alizés du Sud-Est sont les plus intenses et par là même les upwellings équatoriaux les plus vigoureux. Ceci est en accord avec la surestimation de la variabilité de la topographie dynamique constatée sur la figure 8.6. Lors de l'année 1998 pendant laquelle les conditions de type La Niña prévalent, l'assimilation de données permet de corriger la profondeur de la thermocline qui est trop profonde, en particulier dans le Pacifique Central.

8.4 Discussion

Comme nous l'avons vu avec les figures 8.5 et 8.13, il semble que l'assimilation conjointe de données altimétriques et de données de température in-situ TAO pose quelques problèmes le long de 8°N . Lors de la rapide validation de la topographie dynamique moyenne observée dans le chapitre 6, c'est précisément dans cette région que nous avons relevé les plus fortes erreurs entre la topographie dynamique absolue (obtenue en sommant la SLA T/P+ERS et la MDT GRACE) et les hauteurs dynamiques issues des mouillages TAO. Ceci nous a amené à suspecter un problème de compatibilité dans cette région entre l'altimétrie et les données in-situ.

Afin de mieux se rendre compte de l'allure et de la répartition des erreurs sur la

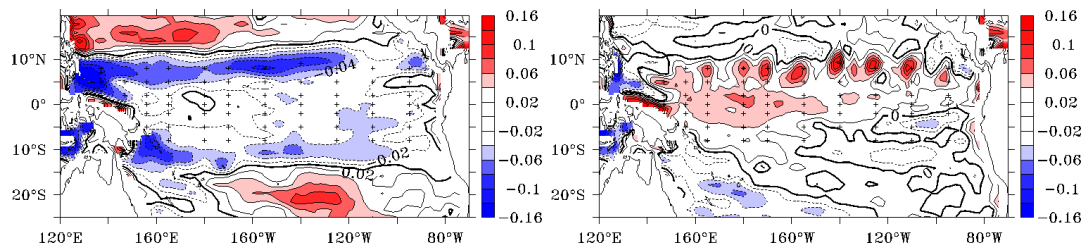


FIG. 8.15 – Carte des différences entre la MDT simulée pour la période 1993-1998 et la MDT observée par satellite pour la simulation libre (à gauche) et pour la simulation avec assimilation de données (à droite).

topographie dynamique moyenne simulée, nous avons calculé la différence entre la MDT satellite d'une part et les MDT de la simulation libre et de la simulation avec assimilation de données d'autre part. La figure 8.15 montre ces cartes de différence. Les croix sur la figure 8.15 représentent la position des différents mouillages TAO. On constate que l'assimilation de données a permis de fortement réduire les erreurs globales par rapport à la topographie dynamique moyenne observée. La RMSD sur tout le domaine passe de 4.5 cm pour la simulation libre à seulement 3.1 cm pour la simulation avec assimilation de données ce qui est un point très positif. La carte de différence entre la MDT observée et la MDT simulée avec assimilation de données montre cependant des bulles d'erreur autour de certains mouillages le long de la longitude 8°N. Ces bulles expliquent l'allure quelque peu bruitée de la topographie dynamique moyenne simulée avec assimilation de données (cf figure 8.5). Le fait qu'elles soient systématiquement localisées à la verticale d'un mouillage TAO est la signature d'une différence d'état moyen entre les données altimétriques et les profils in-situ TAO assimilés.

Pour illustrer cette différence d'état moyen entre les deux jeux de données assimilés, nous avons comparé la topographie dynamique moyenne simulée avec assimilation de données avec la topographie dynamique moyenne observée utilisée pour référencer les SLA T/P+ERS. La figure 8.16 montre deux sections de la topographie dynamique moyenne simulée avec et sans assimilation de données ainsi que la MDT observée par satellite. Nous avons utilisé la section 140°W, c'est à dire une section échantillonnée par le réseau TAO et la section 145°W qui elle se trouve entre deux sections TAO. Pour la section 140°W, nous avons également ajouté les hauteurs dynamique calculées à partir des observations in-situ de température des mouillage TAO. Le niveau de référence pour les hauteurs dynamiques est arbitraire. Il ne faut donc pas regarder la différence de niveau absolu. En revanche les hauteurs dynamiques nous donnent une précieuse information sur les gradients méridionaux du niveau de la mer mesurés par des données in-situ. La figure 8.16 illustre clairement le fait que les gradients méridionaux estimés par les hauteurs dynamiques TAO sont en accord avec les gradients méridionaux de la topographie dynamique moyenne GRACE, excepté pour le mouillage à 9°N, soit au niveau du sillon Nord équatorial. Dans cette région, les gradients méridionaux sont particulièrement forts comme on le voit sur la figure 8.16 et la résolution de la MDT satellite (333 km) est vraisemblablement encore trop faible pour les représenter de manière réaliste. La MDT GRACE présente cependant déjà une nette amélioration par rapport à la MDT de la simulation libre qui est encore beaucoup plus lisse (cf figure 8.16).

Si l'on regarde la topographie dynamique moyenne simulée avec assimilation, on re-

marque tout de suite qu'elle est en accord avec la MDT observée utilisée pour référencer la SLA T/P+ERS assimilée pour la section à 145°W. Pour la section TAO à 140°W, en revanche, la topographie dynamique moyenne simulée est également en accord avec la MDT GRACE, excepté pour la région du sillon Nord équatorial, aux alentours de 10°N. Le sillon Nord équatorial simulé à 140°W avec assimilation conjointe de données altimétriques et de température in-situ TAO est plus creusé et en accord avec les hauteurs dynamiques TAO, qu'il ne l'est dans la topographie dynamique moyenne observée par satellite et utilisée pour référencer la SLA T/P+ERS assimilée.

En résumé, nous avons une topographie dynamique moyenne simulée avec assimilation de données qui est proche des hauteurs dynamiques TAO lorsque les observations TAO sont disponibles mais qui est proche des observations satellites là où l'on ne dispose pas d'observations in-situ pour contraindre le modèle. Les deux jeux d'observations présentant des états moyens différents dans cette région particulière (cf chapitre 6), cela se traduit par les bulles d'erreurs que l'on voit sur la figure 8.15.

Ce comportement de l'assimilation de données s'explique par la nature et la résolution spatiale très différentes des données altimétriques et des données in-situ. Localement, au niveau du mouillage, l'information dominante est l'information de température TAO. Les données TAO sont journalières et continues dans le temps. Elles fournissent de plus une grande quantité d'informations sur la verticale (jusqu'à 40 observations pour un mouillage). La quantité d'informations au moment de l'analyse sur la température de subsurface en provenance d'un mouillage TAO est donc localement beaucoup plus importante que la quantité d'informations provenant de l'altimétrie. Il est donc normal que l'état analysé estimé par le filtre SEEK soit proche de l'état observé par les TAO au niveau des mouillages. Entre les mouillages en revanche, on ne dispose plus directement d'informations provenant du réseau TAO. L'influence des observations de température sur les corrections effectuées par le filtre SEEK sur l'état prévu se fait uniquement par l'intermédiaire des covariances d'erreurs utilisées pour paramétrer le filtre SEEK. Par contre, on dispose toujours de l'information altimétrique qui est une observation synoptique. Entre les mouillages, l'information dominante est donc l'information provenant des observations altimétriques. L'influence des observations de température des TAO diminue assez rapidement lorsque l'on s'éloigne de la position des mouillages du fait que l'on dispose d'une autre information entre les mouillages pour estimer une correction au moment de l'analyse. Ceci explique les bulles d'erreur que l'on trouve sur la figure 8.15 et l'influence beaucoup plus "locale" des observations TAO par rapport à la paramétrisation de la "localité" du filtre SEEK utilisée (les zones d'influences utilisées ont une extension zonale de 30° et une extension meridienne de l'ordre de 10° à 8°N compte tenu du raffinement de maillage à l'équateur).

8.5 Conclusion

Ce chapitre a été l'occasion de valider notre système d'assimilation de données sur une expérience longue couvrant la période 1993-1998 englobant l'El Niño du siècle en 1997-1998. L'assimilation conjointe de la topographie dynamique absolue et des profils in-situ de températures TAO s'est avérée efficace en terme d'écart RMS sur toute la période, y compris pendant l'événement El Niño de 1997-1998 durant lequel on note qu'une très légère augmentation de la RMSD par rapport à la topographie dynamique et aux températures TAO assimilées. L'assimilation a également permis de réduire la

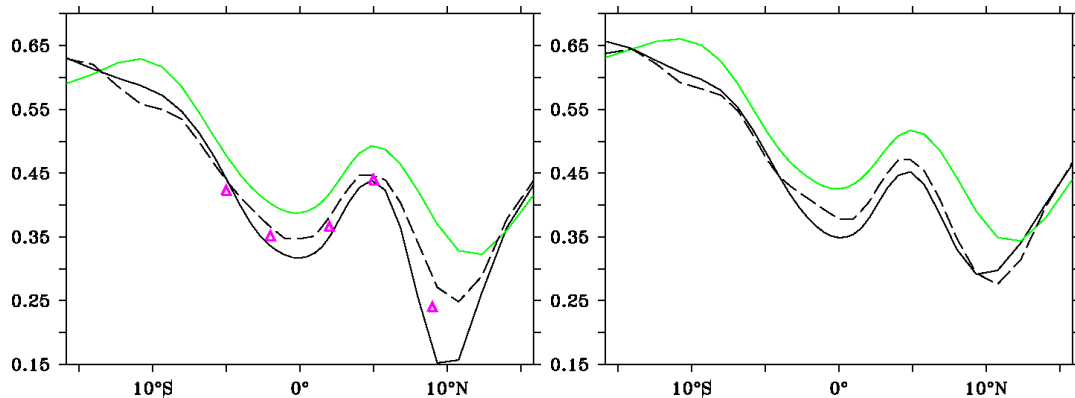


FIG. 8.16 – MDT le long des longitude $140^{\circ} W$ (à gauche) et $145^{\circ} W$ (à droite) pour la simulation libre (en vert), la simulation avec assimilation de données (en noir), et la MDT satellite observée (en tiret). Sur la section à $140^{\circ} W$, les hauteurs dynamiques issues des mouillages TAO sont également tracées (triangle magenta).

RMSD sur la période par rapport aux données indépendantes des XBTs collectées par les navires VOS, ce qui représente un test bien plus révélateur que la simple réduction de la RMSD par rapport aux données assimilées. Ces résultats sont globalement positifs et encourageants à plusieurs titres, notamment en ce qui concerne la compatibilité entre les données altimétriques référencées avec la topographie dynamique moyenne satellite observée déduite des données GRACE et les données in-situ de température provenant du réseau de mouillages TAO. Cette compatibilité permet d'utiliser de manière plus efficace la complémentarité qui existe entre les données altimétriques qui nous fournissent une vision synoptique avec une résolution spatiale et une précision élevée de la surface de l'océan et les données in-situ bien plus disparates mais qui nous renseignent sur la structure 3D de l'océan.

Nous avons également effectué un tour d'horizon des apports de l'assimilation de données sur la dynamique simulée du Pacifique Tropical. Les résultats sont satisfaisants, même si nous avons identifié des régions pour lesquelles des problèmes persistent. L'assimilation de données a permis d'améliorer de manière significative le réalisme de la topographie dynamique moyenne simulée. La variabilité de la topographie dynamique simulée avec assimilation de données est également plus en accord avec la variabilité observée par les satellites altimétriques. Cette représentation plus fidèle de la topographie dynamique moyenne de l'océan se traduit notamment par des courants de surfaces plus réalistes, en particulier le NEC et le NECC qui ont été assez fortement intensifiés par l'assimilation de données. Cette intensification est en accord avec la climatologie de courants de surface déduite des bouées dérivantes SVP proposée par Niiler (2001). L'assimilation n'a en revanche pas permis d'améliorer la représentation de la séparation en deux branches du SEC dans l'extrême Est du bassin. Sur la verticale, les courants zonaux simulés avec assimilation de données sont globalement plus proches des courants mesurés par les mouillages TAO équipés de courantomètres ADCP le long de l'équateur. En particulier le démarrage plus intense de l'EUC dans l'Ouest du bassin semble plus réaliste. Pour le mouillage à $110^{\circ} W$ l'assimilation dégrade les courants zonaux simulés au dessus du cœur de l'EUC avec des vitesses vers l'Ouest trop intenses. Cette surestimation de vitesses zonales est liée à la mauvaise représentation de la séparation du SEC dans cette région. En ce qui

concerne la température, l'assimilation conjointe a permis de correctement repositionner la thermocline qui est trop profonde dans la simulation libre. Cette meilleure représentation de la profondeur de la thermocline est confirmée lorsque l'on compare la simulation avec assimilations de données à des sections de températures moyennes calculées à partir des données indépendantes XBTs.

L'analyse de la simulation avec assimilation de données sur la période 1993-1998 nous a également permis de mettre en évidence une région pour laquelle l'assimilation conjointe de données altimétriques et in-situ est toujours problématique du fait de la différence d'état moyen entre la topographie dynamique absolue mesurée par satellite et les données in-situ. Cette région se situe le long du sillon Nord équatorial (vers 10°N) et correspond à la région où les gradients méridionaux de topographie dynamique sont les plus forts. Le manque de résolution spatiale de la MDT GRACE, empêche de représenter fidèlement ces forts gradients et se traduit par des erreurs sur la topographie dynamique observée par satellite. Ces erreurs engendrent à leur tour des incompatibilités entre les données altimétriques et les données in-situ, incompatibilités qui se révèlent problématiques lors de l'assimilation conjointe de ces deux types de données. On peut raisonnablement espérer que les dernières MDT calculées à partir du modèle de géoïde EIGEN4 basé sur un peu plus de deux années de données GRACE permettront de résoudre ce problème.

Chapitre 9

Evolution des transports dans les couches de surface du Pacifique Tropical sur la période 1993-1998

Sommaire

9.1	Variabilité du flux zonal de masse	161
9.2	Discussion	172

Dans le chapitre précédent, nous avons constaté que l'assimilation conjointe de données altimétriques et de données de température in-situ modifiait de manière significative les courants, en surface mais aussi en profondeur. Les courants simulés avec assimilation de données semblent plus en accord avec la climatologie de courant de surface déduite des bouées dérivantes SVP d'une part, et avec les profils de courants mesurés par les ADCP équipant certains des mouillages TAO le long de l'équateur d'autre part. Cette meilleure représentation des courants, et en particulier du NECC dont la variabilité aux échelles interannuelle impacte significativement l'évolution des événements ENSO, nous a conduit à nous intéresser aux transports simulés avec et sans assimilation de données. Les courants zonaux au voisinage de la surface sont impliqués dans la redistribution aux échelles interannuelles des eaux chaudes des couches supérieures du Pacifique Tropical. Durant El Niño, le NECC transportant des eaux chaudes vers l'Est s'intensifie. En parallèle, le SEC, courant relativement froid vers l'Ouest diminue. Pendant la période froide La Niña, la situation est inversée. Le bilan est donc une advection d'eaux chaudes vers l'Est durant El Niño mais une advection d'eaux froides vers l'Ouest durant La Niña. La compréhension et la quantification des variations des courants et de leurs transports sont donc importantes dans l'optique d'une compréhension plus fine de l'évolution des événements ENSO. Cette étude est préliminaire et demande à être développée. Il me semblait cependant intéressant de l'insérer dans ce manuscrit de thèse dans la mesure où elle illustre l'impact positif de l'assimilation de données sur la dynamique simulée du Pacifique Tropical, et les potentialités du système d'assimilation mis en place pour une analyse de la dynamique du Pacifique Tropical en vue d'élucider certains mécanismes encore mal compris.

Dans ce chapitre, nous allons nous intéresser plus spécifiquement à l'étude des transports zonaux simulés avec assimilation de données ainsi que de leur variabilité. L'étude des transports est importante pour mieux comprendre l'évolution de la dynamique du Pacifique Tropical, notamment aux échelles interannuelles. On dispose cependant de relativement peu d'observations de courants continus dans le temps. Seulement 5 mouillages TAO sont équipés de courantomètre et/ou capteurs ADCP et mesurent les profils de vitesse de la couche de surface (entre la surface et au mieux 450 m). Ces mouillages sont localisés à l'équateur aux longitudes 156°E, 165°E, 170°W, 140°W et 110°W. Ces données ont notamment été utilisées pour étudier les courants zonaux équatoriaux et leurs variabilités durant les précédents événements El Niño (*Halpern, 1987; McPhaden et Hayes, 1990; Kessler et McPhaden, 1995*). Ces études sont de plus restées régionales (Pacifique Est, Pacifique Ouest et Pacifique Central respectivement). *Izumo (2005)* a, quant à lui, utilisé les courants mesurés par les TAO à l'équateur aux longitudes 170°W, 140°W et 110°W pour construire une série temporelle entre 1980 et 2002 afin d'étudier plus spécifiquement la variabilité de l'EUC. Une autre source d'observation des courants dans les couches superficielles du Pacifique Tropical est donnée par les ADCP de coque des navires océanographiques. Ce type de données a été utilisé pour étudier les courants de surface et leurs variabilités durant les précédents événements El Niño (*Delcroix et al., 1992; Johnson et al., 2000, 2002*). Pour donner une idée de la quantité d'information disponible, *Johnson et al. (2002)* utilise 172 sections entre 138°E et 86°W, la plupart collectées durant les campagnes de maintenance du réseau TAO, de Juin 1985 à Décembre 2000 pour évaluer les courants et leurs variabilités.

L'utilisation des résultats de la simulation libre et de la simulation avec assimilation de données permet en revanche de disposer d'un champ de vitesse synoptique sur tout le Pacifique Tropical avec une fréquence temporelle de 5 jours pour toute la période 1993-

1998. La variabilité et les transports estimés sur la période seront donc bien représentatifs de la dynamique simulée par le modèle, et on peut le penser, de la dynamique du Pacifique Tropical. Les transports de la simulation libre présentent une variabilité assez similaire à celle de la simulation assimilée. L'assimilation a en revanche globalement renforcé l'intensité de courants simulés, en accord avec les observations (en particulier en surface). Cela se traduit par des transports zonaux plus importants. Dans la suite de ce chapitre, nous ne discuterons que de l'évolution des transports dans les couches de surfaces de la simulation avec assimilation de données pour les raisons évoquées ci-dessus.

9.1 Variabilité du flux zonal de masse

Le Pacifique Tropical est sujet à de larges modifications du transport de masse dans les couches superficielles de l'océan durant les événements El Niño-La Niña. Certains courants peuvent ainsi disparaître, voir carrément s'inverser durant les différentes phases du phénomène de variabilité interannuelle. Nous allons étudier ces importantes modifications des transports de masse dans les couches superficielles du Pacifique Tropical sur la période 1993-1998 en estimant les transports zonaux et leurs variabilités au travers de différentes sections méridiennes. Pour se faire, nous utiliserons les sorties de la simulation avec assimilation de données. Les transports zonaux sont estimés pour deux classes de densité. La couche de surface s'étend de la surface jusqu'à l'isopycne $\sigma_\theta = 24 \text{ kgm}^{-3}$ qui correspond à peu près au cœur de la pycnocline dans cette région. Cette couche de surface capture la plus grande partie des principaux courants de surface du Pacifique Tropical, i.e. le SEC, le NECC et l'EEC pour East Equatorial Current qui correspond au courant de surface vers l'Est advecté par les anomalies de vent d'Ouest le long de l'équateur (*Delcroix et al.*, 1992). Une petite fraction de l'EUC est également parfois prise en compte dans le débit intégré. La couche de subsurface contient des eaux présentant une densité comprise entre $24 \text{ kgm}^{-3} < \sigma_\theta < 26.25 \text{ kgm}^{-3}$. L'isopycne $\sigma_\theta = 26.25 \text{ kgm}^{-3}$ est proche de la base de la pycnocline à l'équateur qui est associée à des eaux proches de 13°C . Les transports de masse dans la couche de subsurface capturent l'essentiel de l'EUC ainsi que les extensions en profondeur des courants de surface. La figure 9.1 montre une section zonale de la vitesse moyenne sur la période 1993-1998 à l'équateur. Les isocontours des isopycnes $\sigma_\theta = 24 \text{ kgm}^{-3}$ et $\sigma_\theta = 26.25 \text{ kgm}^{-3}$ sont superposés afin d'illustrer la part des courants présents dans chacune des deux classes de densité définies. Les transports pour chaque classe de densité sont évalués pour des intervalles de 6 mois (printemps boréal du 1^{er} janvier au 30 juin et automne boréal du 1^{er} juillet au 31 décembre) en suivant *Johnson et al.* (2000). Spatialement, les transports dans les deux couches ont été calculés pour chaque intervalles de 1° de latitude. Les limites méridionales des différents courants sont définies par les reverses de courants intégrés sur la couche considérée.

Les figures 9.2 et 9.4 montrent respectivement les transports intégrés dans la couche de surface et la couche de subsurface par degrés de latitude et moyennés par période de 6 mois comme expliqué ci-dessus. Le tableau 9.1 synthétise les résultats présentés sur les figures 9.2 et 9.4. Sur la période 1993-1998 et en utilisant la méthode introduite ci-dessus pour définir l'expansion méridienne des différents courants, on obtient un transport moyen dans la couche de surface (i.e. $\sigma_\theta < 24 \text{ kg}^{-3}$) de $31 \pm 7 \text{ Sv}$ pour le SEC et de $15 \pm 5 \text{ Sv}$ pour le NECC et l'EEC quand celui-ci est présent. Dans la couche de subsurface (i.e. $24 \text{ kg}^{-3} < \sigma_\theta < 26.25 \text{ kg}^{-3}$), on obtient un transport moyen de $18 \pm 4 \text{ Sv}$ pour l'EUC.

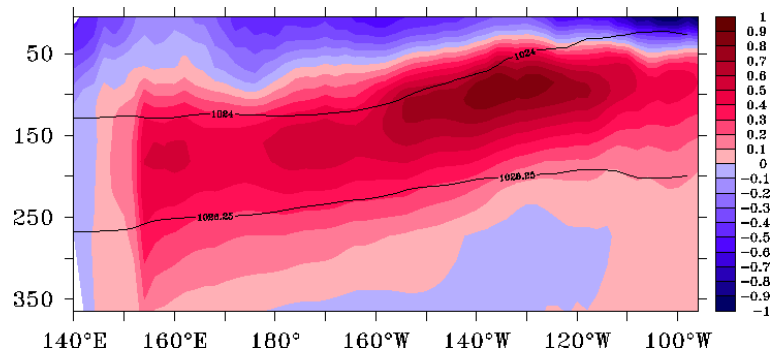
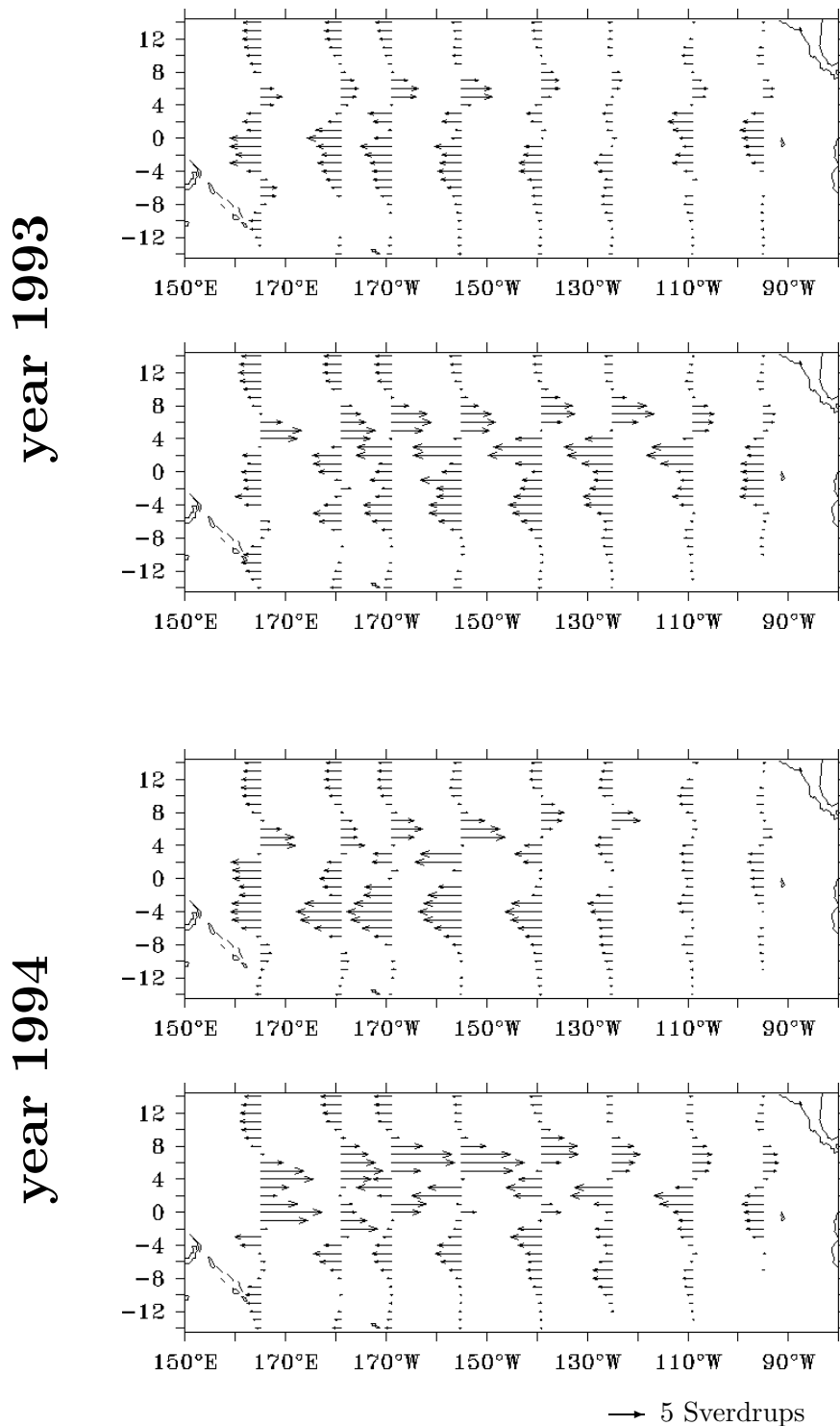


FIG. 9.1 – Section moyenne de vitesse zonale sur la période 1993-1998 à l'équateur. Les isopycnes $\sigma_\theta = 24 \text{ kgm}^{-3}$ et $\sigma_\theta = 26.25 \text{ kgm}^{-3}$ sont superposées.

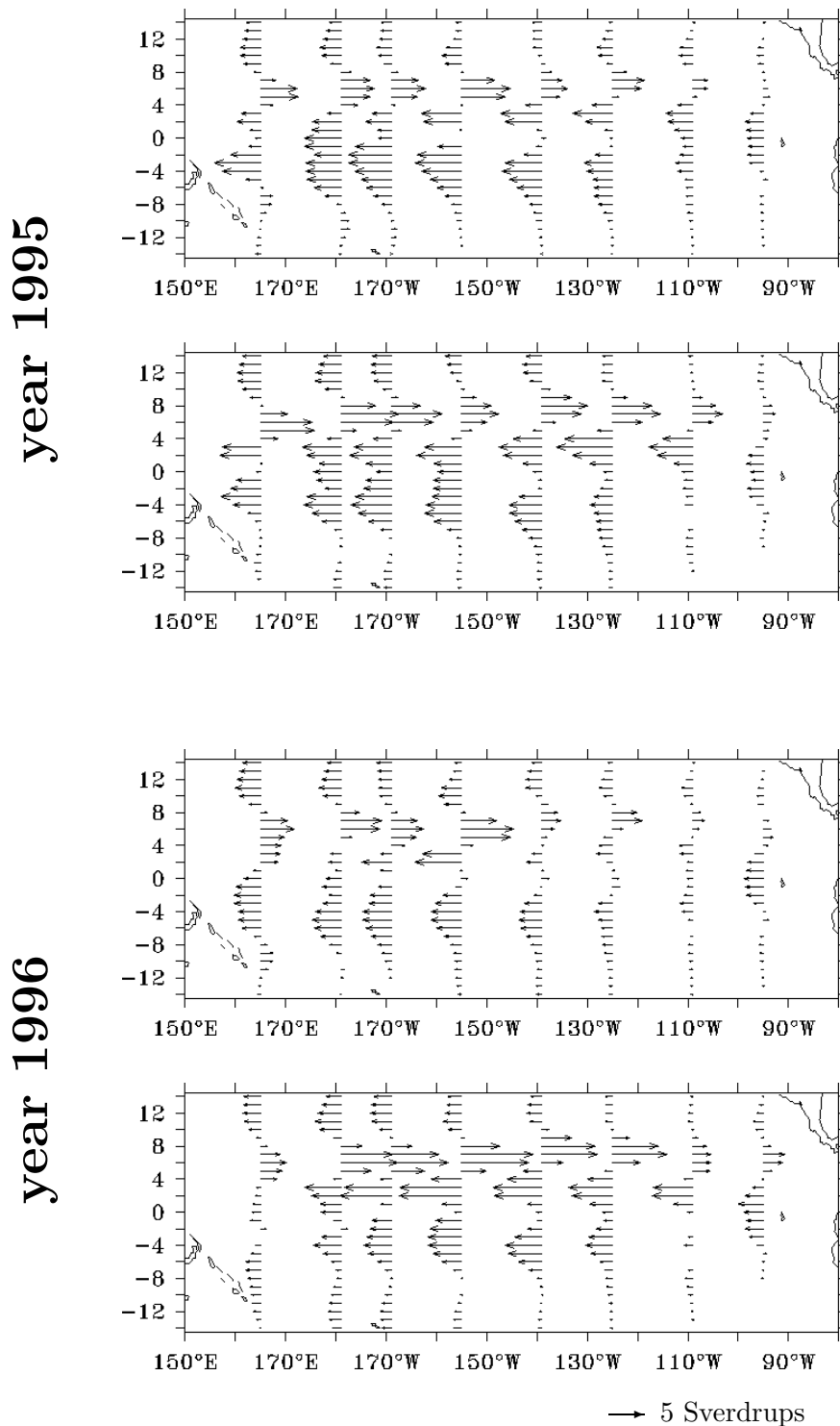
L'année 1993 fait apparaître un cycle saisonnier "normal" avec des courants de surface faibles au printemps (cf figure 9.2(a)). Le EUC, qui atteint son maximum saisonnier au printemps, lorsque les alizés de Sud-Est sont minimums (cf chapitre 1), est bien identifié sur la figure 9.4(a). Son transport n'est cependant que de 16 Sv à comparer aux 18 Sv obtenus en moyenne sur la période 1993-1998. En automne, avec le renforcement des alizés de Sud-Est, le SEC se renforce de manière significative (+ 11 Sv) de même que le NECC qui double quasiment par rapport au printemps 1993. De faibles conditions El Niño prévalent pendant l'automne 1993 comme le montre la figure 1.8 de la page 20 montrant l'évolution de l'indice SOI et de l'anomalie de SST pour la région Niño-3.4. Ces faibles conditions El Niño, dont une des caractéristiques est précisément un renforcement du NECC (*Taft et Kessler, 1991*), peuvent expliquer l'intensification marquée du NECC sur cette période par rapport à la saison précédente.

Le printemps 1994 fait apparaître un NECC faible en particulier dans l'Est du bassin où le NECC a quasiment disparu dans la couche de surface pour les sections à 110°W et 95°W (cf figure 9.2(a)). Le SEC est par contre relativement intense pour la saison, en particulier dans l'Ouest du bassin, avec des transports supérieurs à 30 Sv pour les sections allant de 140°W à 165°E . L'EUC est également très marqué dans la couche intermédiaire avec un transport moyen de l'ordre de 23 Sv en moyenne sur toute la largeur du bassin. Le transport augmente de 165°E à 125°W pour atteindre près de 30 Sv à cette longitude. En automne 1994, on retrouve des conditions de type El Niño (cf figure 1.8) mais elles sont cette fois plus marquées. Des anomalies de vent d'Ouest assez intenses se développent dans l'Ouest du Pacifique Tropical. Ces anomalies sont clairement visibles sur la figure 9.3 montrant les diagrammes longitude-temps à l'équateur des anomalies de la tension zonale des vents ETS+TAO utilisée pour forcer le modèle d'une part et de l'anomalie de température de surface simulée avec assimilation de données d'autre part. Ces anomalies de vent font apparaître l'EEC dans l'Ouest de bassin. Le SEC est très faible, en particulier dans l'Ouest où il a quasiment disparu de la couche de surface pour la section à 165°E , et est repoussé au Sud. Le NECC est particulièrement vigoureux, et se combine avec l'EEC pour transporter plus de 21 Sv vers l'Est du Pacific Tropical. Les masses d'eau advectées vers l'Est par l'EEC dans la bande équatoriale participent à la création des anomalies positives de SST visibles sur la figure 9.3. Le transport de l'EUC dans la couche intermédiaire a sensiblement diminué par rapport au printemps boréal et vaut 15 Sv en moyenne sur la largeur du bassin. Le transport réel de l'EUC est en fait inférieur à cette



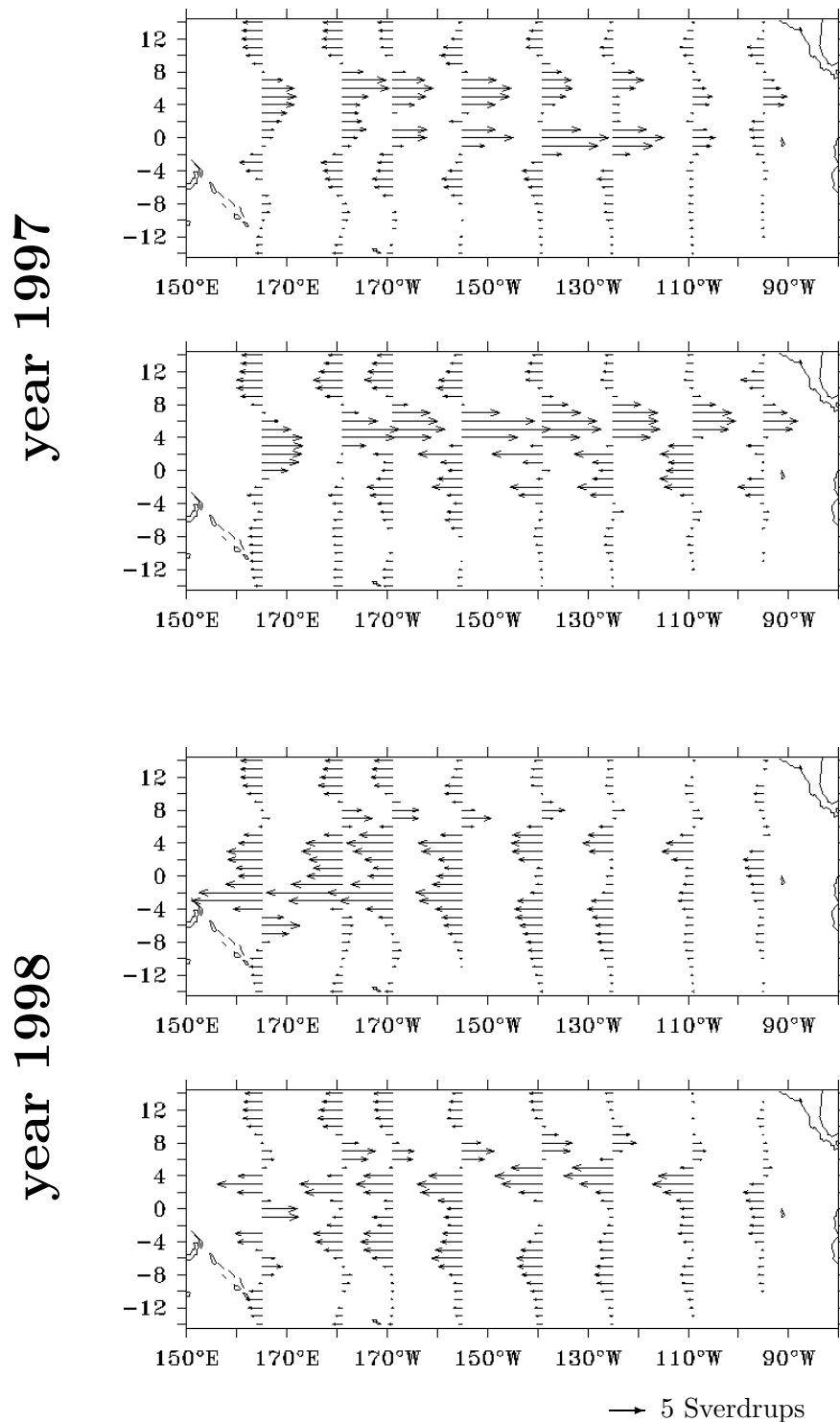
(a) First Part

FIG. 9.2 – Transport zonal de masse ($10^6 \text{m}^3 \text{s}^{-1}$) pour la couche de surface avec $\sigma_\theta < 24 \text{kgm}^{-3}$. Les vecteurs sont proportionnels au transport moyen dans la classe de densité pour chaque intervalle de 1° de latitude. Pour chaque année, le printemps boréal (en haut) et l'automne boréal (en bas).



(b) Second Part

FIG. 9.2 – (suite) *Transport zonal de masse ($10^6 m^3 s^{-1}$) pour la couche de surface avec $\sigma_\theta < 24 kg m^{-3}$. Les vecteurs sont proportionnels au transport moyen dans la classe de densité pour chaque intervalle de 1° de latitude. Pour chaque année, le printemps boréal (en haut) et l'automne boréal (en bas).*



(c) Third Part

FIG. 9.2 – (suite) *Transport zonal de masse ($10^6 m^3 s^{-1}$) pour la couche de surface avec $\sigma_\theta < 24 kg m^{-3}$. Les vecteurs sont proportionnels au transport moyen dans la classe de densité pour chaque intervalle de 1° de latitude. Pour chaque année, le printemps boréal (en haut) et l'automne boréal (en bas).*

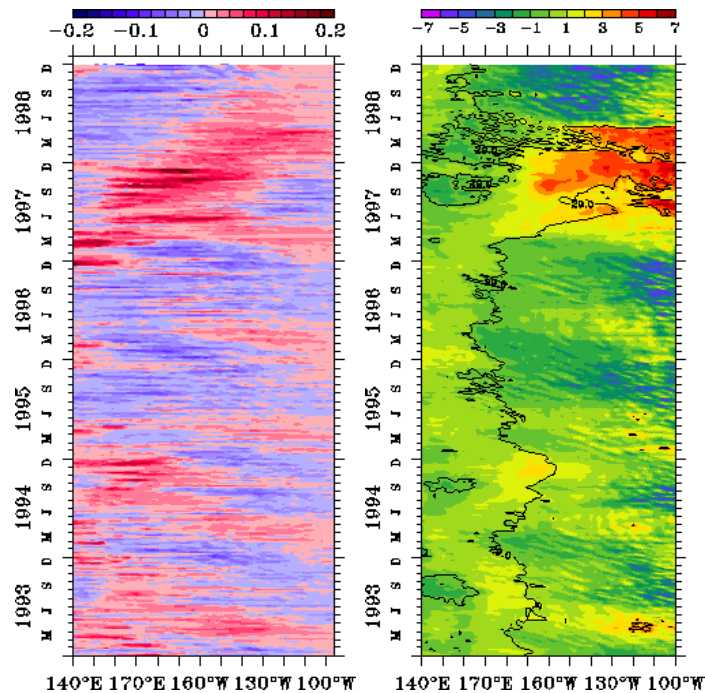


FIG. 9.3 – Diagramme longitude-temps à l'équateur de l'anomalie de la tension zonale de vent (à gauche) et de la SST (à droite) pour la simulation avec assimilation de données.

valeur de 15 Sv, mais une partie des extensions en profondeur du NECC sont prises en compte dans le calcul du transport total de l'EUC (cf figure 9.4(a)). Cela est dû à la façon dont sont définies les limites méridiennes des courants. Le NECC, qui est particulièrement intense sur cette période comme nous venons de le voir s'étend plus en profondeur qu'à l'accoutumée. Il n'y a plus de renverse de courant dans le transport intégré sur la couche intermédiaire pour délimiter le NECC et l'EUC qui transportent tous les deux les eaux vers l'Est du Pacifique Tropical. Le transport estimé de l'EUC prend donc en compte les eaux déplacées en profondeur par le NECC, en particulier pour les sections 165°E et 180°.

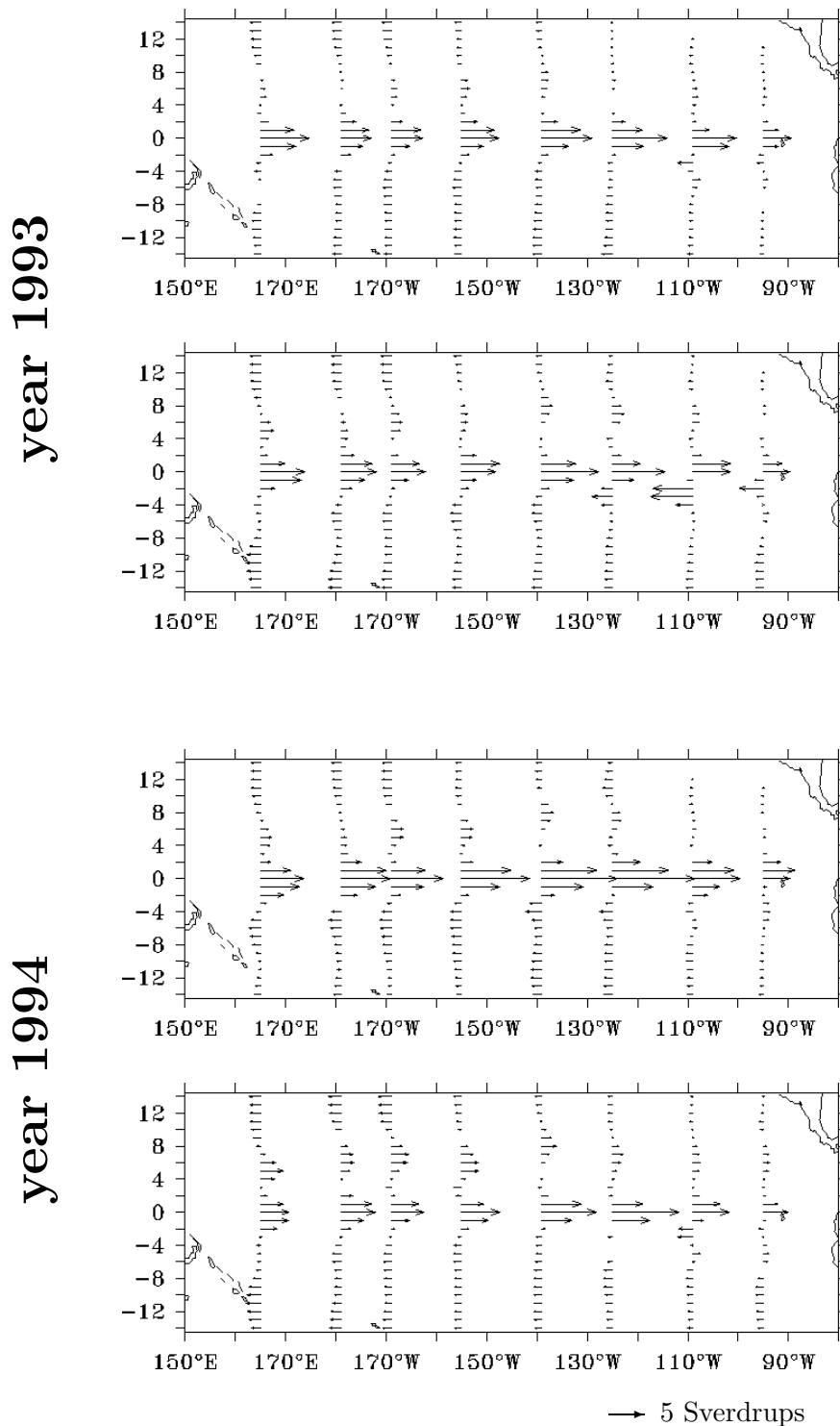
Au printemps 1995, les courants simulés avec assimilation de données dans la couche de surface sont assez intenses pour la saison. En particulier le SEC avec un transport de l'ordre de 36 Sv en moyenne sur toute la largeur du bassin présente un débit élevé par rapport au printemps boréal des deux années précédentes, avec respectivement 25 et 31 Sv pour les années 1993 et 1994. En profondeur, l'EUC est par contre assez vigoureux pour la saison avec un transport de 18 Sv. Avec le passage à l'automne, le SEC se réintensifie légèrement pour atteindre 39 Sv de même que le NECC qui transporte 16 Sv vers l'Est dans la couche de surface. Les deux branches du SEC sont clairement identifiées, excepté dans l'extrême Est du bassin (cf figure 9.2(b)). En subsurface, l'EUC qui devrait normalement être à son minimum saisonnier, reste étonnamment intense avec un transport de 19 Sv en moyenne sur la largeur du bassin, soit 1 Sv de plus que pour le printemps précédent.

Les courants de surface sont relativement faibles dans la couche de surface pour le printemps boréal de l'année 1996, le SEC transportant 25 Sv vers l'Ouest alors que le NECC ramène environ 12 Sv vers l'Est du bassin. L'EUC, avec un transport de 22 Sv est par contre plutôt vigoureux. Cette configuration de courant est en accord avec le cycle saisonnier du Pacifique Tropical. De faibles conditions La Niña prévalent durant l'automne

boréal de l'année 1996. Les deux branches du SEC sont intenses (cf figure 9.2(b)) et transportent 34 Sv d'Est en Ouest. Le NECC est également fort pour les sections allant de 165°E à 125°W avec un transport moyen sur toute la largeur du bassin de 18 Sv. Dans la couche de subsurface, l'EUC est clairement défini avec un transport moyen de 15 Sv. On note la présence d'extensions en profondeur du SEC et du NEC plus marquées que les années précédentes.

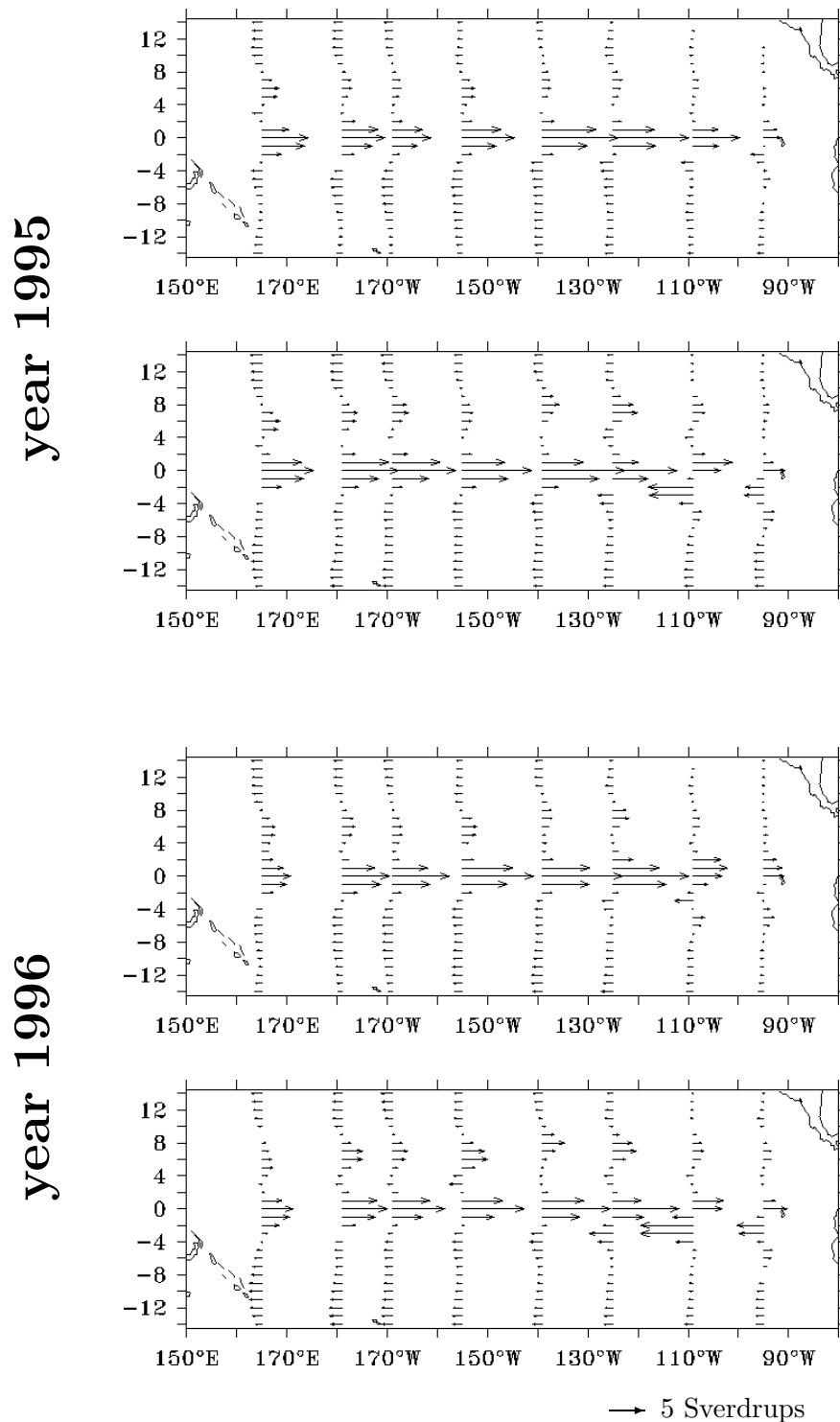
Le lecteur pourra trouver dans *Johnson et al.* (2000) le même type de figures que les figures 9.2 et 9.4 présentant les transports de masse zonaux intégrés par degrés de latitude et moyennés par période de 6 mois mais calculés à partir de 35 sections de courants zonaux collectées de septembre 1996 à novembre 1998 à l'aide des ADCPs de coque des navires chargés de la maintenance des mouillages ATLAS du réseau TAO. Il est difficile de comparer quantitativement les transports simulés avec les transports déduits des observations présentés par *Johnson et al.* (2000). D'une part, les courants ne sont pas totalement échantillonnés par les observations collectées par les navires de maintenance des bouées TAO. D'autre part, bien que les courants présentent généralement des structures cohérentes sur l'ensemble du bassin et pour une saison donnée, il faut garder à l'esprit qu'un ensemble de sections synoptiques réalisées pendant la même saison de 6 mois peut être aliasé selon une grande variété d'échelle temporelle. Une comparaison qualitative des courants simulés et observés est en revanche possible.

Le déclenchement de l'El Niño de 1997 durant le printemps boréal se traduit par un profond changement des transports simulés (cf figures 9.2(c) et 9.4(c)). Le SEC est extrêmement faible. La branche Nord a disparu dans l'Ouest du bassin et la branche Sud transporte très peu d'eau. Le transport moyen total du SEC tombe ainsi à seulement 12 Sv en moyenne sur toute la largeur du bassin soit moins de la moitié de son transport moyen sur la période 1993-1998 (cf tableau 9.2). De manière inattendue pour un printemps boréal, mais caractéristique lors d'un événement El Niño (*Taft et Kessler*, 1991), le NECC est très intense. Les fortes anomalies de vent d'Ouest (cf figure 9.3) font également apparaître un fort EEC comme le montre clairement la figure 9.2(c). Ces deux courants se combinent pour donner un transport vers l'Est qui atteint près de 30 Sv en moyenne sur toute la largeur du bassin et culmine à 41 Sv à 140°W. En profondeur, l'EUC est très faible dans l'Ouest du bassin mais se maintient dans l'Est (cf figure 9.4(c)). Comme nous l'avons vu sur la figure 8.14, page 153, la pente zonale de la thermocline diminue fortement pendant cette période pour devenir quasiment nulle sur toute la largeur du bassin au plus fort de l'El Niño. Cela se traduit par une réduction, voire une disparition du gradient de pression zonal à l'équateur au niveau de la thermocline. La forte diminution de l'EUC dans l'Ouest du Pacifique Tropical est consistante avec cette atténuation du gradient de pression zonal qui est le moteur de l'EUC comme je l'ai rapidement expliqué dans le chapitre 1 traitant de la dynamique du Pacifique Tropical. Dans l'Est, les anomalies positives de SST atteignent les côtes d'Amérique du Sud à partir du mois de Mars (cf figure 9.3). Le fort coup de vent d'Ouest en Mars 1997 génère une forte onde de Kelvin très clairement visible sur les figures 8.6 et 8.14 montrant respectivement les diagrammes longitude-temps de la SLA et de la profondeur de l'isotherme 20° utilisée comme proxy de la profondeur de la thermocline. Contrairement à l'onde de Kelvin générée par le coup de vent d'Ouest en Janvier 1997 (cf figure 9.3, 8.6 et 8.14), l'onde de Kelvin générée par le coup de vent d'Ouest de Mars 1997 qui est plus intense, d'une durée plus longue et présente également un fetsch plus long, parvient à enfoncer la thermocline sur toute la largeur du bassin. Dans l'Est, la thermocline ne s'approfondit de manière significative qu'à partir du mois d'Avril 1997. La pente de la



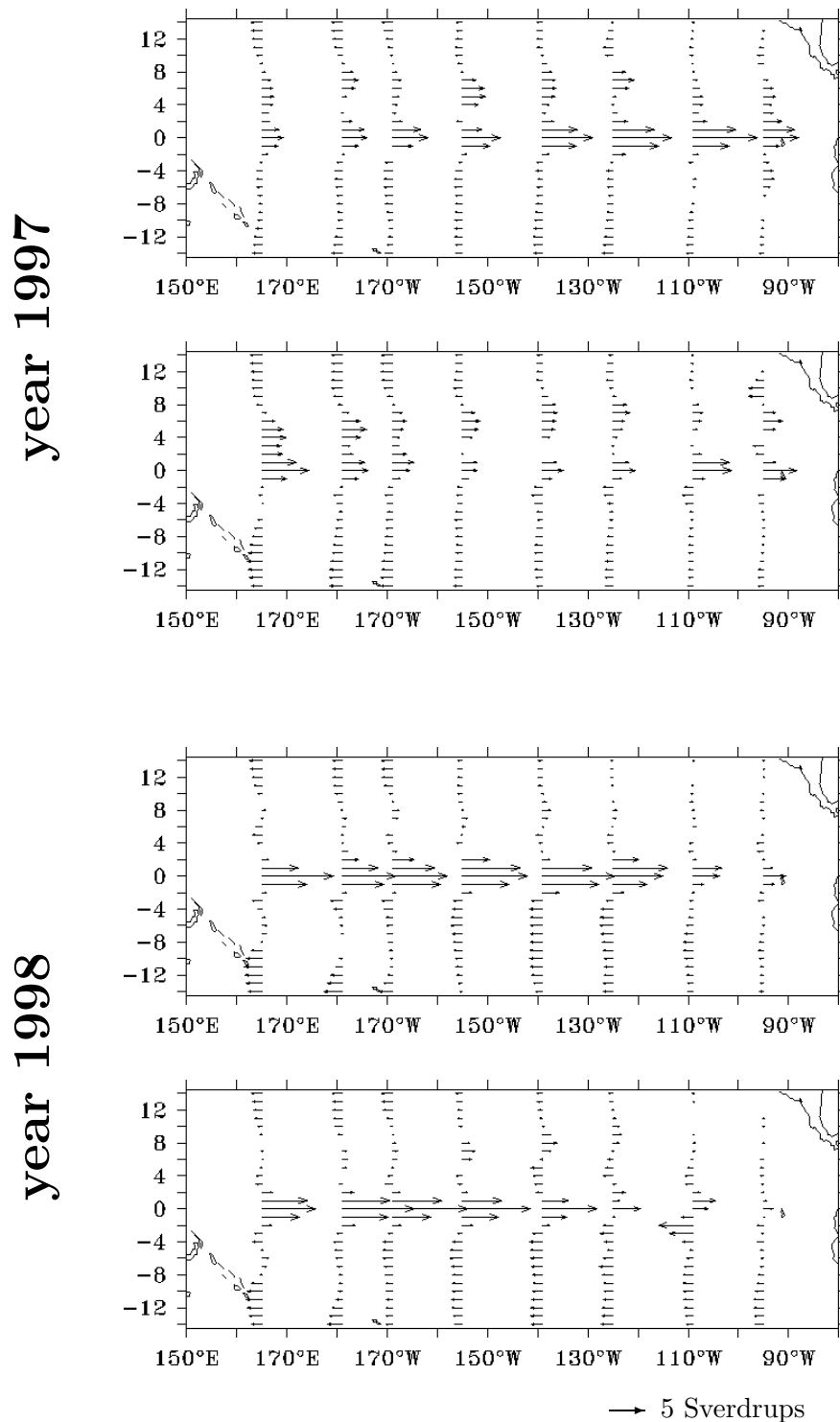
(a) First Part

FIG. 9.4 – Transport zonal de masse ($10^6 m^3 s^{-1}$) pour la couche intermédiaire avec $24 kg m^{-3} < \sigma_\theta < 26.25 kg m^{-3}$. Les vecteurs sont proportionnels au transport moyen dans la classe de densité pour chaque intervalle de 1° de latitude. Pour chaque année, le printemps boréal (en haut) et l'automne boréal (en bas).



(b) Second Part

FIG. 9.4 – Transport zonal de masse ($10^6 m^3 s^{-1}$) pour la couche intermédiaire avec $24 \text{ kg m}^{-3} < \sigma_\theta < 26.25 \text{ kg m}^{-3}$. Les vecteurs sont proportionnels au transport moyen dans la classe de densité pour chaque intervalle de 1° de latitude. Pour chaque année, le printemps boréal (en haut) et l'automne boréal (en bas).



(c) Third Part

FIG. 9.4 – Transport zonal de masse ($10^6 \text{m}^3 \text{s}^{-1}$) pour la couche intermédiaire avec $24 \text{ kg m}^{-3} < \sigma_\theta < 26.25 \text{ kg m}^{-3}$. Les vecteurs sont proportionnels au transport moyen dans la classe de densité pour chaque intervalle de 1° de latitude. Pour chaque année, le printemps boréal (en haut) et l'automne boréal (en bas).

thermocline, et donc le gradient de pression zonal au niveau de la thermocline, se maintient donc pendant une bonne partie du printemps boréal ce qui explique pourquoi le transport moyen intégré dans la couche de subsurface de l'EUC reste assez marqué dans l'Est du bassin sur cette période. Comme lors de l'El Niño de 1994, le transport diagnostiqué de 20 Sv pour l'EUC (cf tableau 9.1) apparaît surestimé par rapport au transport réel de l'EUC à cause des extensions en profondeur du NECC qui sont prises en compte dans le transport total estimé de l'EUC pour les raisons indiquées ci-dessus. Le transport de l'EUC seul est de l'ordre de 10 Sv dans l'Ouest et croît jusqu'à 24 Sv à 125°W. La persistance d'El Niño durant l'automne 1997 se reflète clairement dans les transports diagnostiqués dans la couche de surface et la couche de subsurface (cf figure 9.2(c) et 9.4(c) et 9.1). Comme pour le printemps précédent, le SEC est très faible, même si il s'est légèrement renforcé, en particulier dans l'Est du bassin. Ceci est dû à des alizés quasiment normaux sur une grande partie du Pacifique Tropical durant le mois de Juillet 1997 (cf figure 9.3). Le résultat de cet épisode de vent d'Est est une nette remontée de la thermocline (cf figure 8.14) pendant le mois d'Août et un renforcement assez marqué du SEC sur la même période. Cette courte réintensification du SEC se fait ressentir sur la moyenne saisonnière, en particulier pour les sections 125°W et 110°W. Malgré cela, le transport moyen sur la période est à peine supérieur à 20 Sv. Comme attendu pour un automne boréal, le NECC s'est renforcé, mais l'EEC est lui moins intense. Le transport combiné vers l'Est de ces deux courants, avec 27 Sv (cf tableau 9.1), reste ainsi quasiment le même que pour la saison précédente. Dans la couche de subsurface, l'EUC est extrêmement faible sur toute la largeur du bassin. Le transport estimé est de 14 Sv (cf tableau 9.1), mais prend là encore en compte les extensions en profondeur du NECC qui viennent participer au transport intégré vers l'Est dans la couche de subsurface pour les sections allant de 165°E à 155°W. Si l'on ne considère que les sections à l'Est de 155°W, on obtient un débit pour l'EUC de seulement 9 Sv. Cet EUC extrêmement faible est sans aucun doute la conséquence de la disparition de la pente zonale de la thermocline sur toute la largeur du bassin, comme on peut le voir sur la figure 8.14, qui se traduit par la disparition du gradient zonal de pression au niveau de la thermocline, gradient de pression à l'origine de l'EUC en temps normal.

Le passage en condition La Niña est très net sur les transports par couche de densité (cf figure 9.2(c) et 9.4(c) et tableau 9.1). Le SEC est très marqué avec un transport moyen de 45 Sv à comparer aux 30 Sv que transporte ce courant en moyenne sur la période 1993-1998. Le transport atteint presque 60 Sv dans l'Ouest du Pacifique Tropical. Les deux branches, excepté à 95°W, sont clairement définies. La branche Nord s'étend bien au delà de l'équateur. Le NECC est en revanche quasiment inexistant et ramène seulement 5 Sv d'Ouest en Est. Aucun transport significatif vers l'Est n'est trouvé dans la couche de surface. Même pour un printemps boréal, la quasi disparition du NECC dans la couche de surface (cf figure 9.2(c)), et ce pour l'ensemble des sections diagnostiquées, est remarquable. Dans la couche de subsurface, l'EUC est réapparu. Il est de plus assez intense excepté pour les sections à 110°W et 95°W. Le transport moyen obtenu est de 20 Sv (plus de 25 Sv si l'on exclut les deux sections les plus à l'Est). Les conditions La Niña persistent pendant l'automne 1998, avec un SEC toujours très fort, mais qui transporte un peu moins d'eau vers l'Ouest avec 37 Sv que pour la saison précédente. Ceci est particulièrement vrai pour l'Ouest du bassin (sections 165°E, 180° et 170°W), où le transport a diminué de plus de 15 Sv par rapport au printemps 1998. Dans l'Est en revanche, le SEC s'est plutôt renforcé. Le NECC s'est également légèrement réintensifié comme on pouvait s'y attendre

TAB. 9.1 – Moyenne saisonnière et écart type du transport zonal pour le SEC et le NECC dans la couche de surface ($\sigma_\theta < 26.25 \text{ kg m}^{-3}$) et pour l'EUC dans la couche intermédiaire ($24 \text{ kg m}^{-3} < \sigma_\theta < 26.25 \text{ kg m}^{-3}$).

Saison Boréale	Transport zonal dans la couche de surface pour le SEC et le NECC & EEC et dans la couche de subsurface pour l'EUC en Sverdrup ($10^6 \text{ m}^3 \text{ s}^{-1}$)		
	SEC	NECC & EEC	EUC
Automne 1998	-37 ± 10	7 ± 3	16 ± 9
Printemps 1998	-45 ± 13	5 ± 3	20 ± 8
Automne 1997	-23 ± 6	27 ± 6	14 ± 7 ^a
Printemps 1997	-12 ± 3	29 ± 10	20 ± 6 ^a
Automne 1996	-34 ± 13	18 ± 7	15 ± 5
Printemps 1996	-25 ± 7	12 ± 6	22 ± 7
Automne 1995	-39 ± 12	16 ± 5	19 ± 7
Printemps 1995	-36 ± 9	11 ± 5	18 ± 6
Automne 1994	-24 ± 8	21 ± 11	15 ± 5 ^c
Printemps 1994	-31 ± 9	9 ± 5	23 ± 5
Automne 1993	-36 ± 9	13 ± 4	15 ± 5
Printemps 1993	-25 ± 4	7 ± 3	16 ± 3
Moyenne 93-98	-31 ± 7	14 ± 5	18 ± 4

^a 165°E-155°W, combine les transports du NECC et de l'EEC

^b 165°E-95°W, combine les transports du NECC et de l'EEC

^c 165°E-180°, combine les transports du NECC et de l'EEC

pour un automne boréal, mais son transport reste très faible avec 7 Sv en moyenne à comparer au débit de 15 Sv obtenu en moyenne sur toute la période 1993-1998. Dans la couche de subsurface, l'EUC transporte 16 Sv en moyenne sur toute la largeur du bassin. De manière surprenante, il est particulièrement faible dans l'Est du bassin au niveau des trois sections (sections 125°W, 110°W et 95°W) les plus orientales. En excluant ces sections où le transport de l'EUC est le plus faible, on obtient un débit de 22 Sv, ce qui paraît assez élevé sachant que le débit moyen simulé sur la période est de 18 Sv et que l'EUC est à son minimum saisonnier au cours de l'automne boréal.

9.2 Discussion

La simulation avec assimilation de données altimétriques conjointement avec les données de température in-situ issues des mouillages du réseau TAO fait apparaître de profonds changements des courants et de leurs transports dans les couches de surface de l'océan Pacifique Tropical aux échelles interannuelles, mais aussi annuelles. Par exemple, le transport cumulé aux basses latitudes (i.e. entre $\pm 14^\circ$ de latitude) et dans les deux couches (i.e.

le transport intégré entre la surface et l'isopycne 26.25 kgm^{-3}) à 140°W passe de 20 Sv vers l'Ouest pour l'automne 1996 à 30 Sv vers l'Est pour la saison suivante, soit une renverse impressionnante du transport net en l'espace de 6 mois de l'ordre de 50 Sv (pour comparaison le Gulf Stream transporte environ 60 Sv au niveau du Cap Hatteras (*Berline et al.*, 2003)). Ce chiffre est du même ordre de grandeur que les 64 Sv estimé par *Johnson et al.* (2000) à partir de données in-situ. Les observations utilisées n'échantillonnent pas l'ensemble des courants, cela peut sans doute expliquer la différence entre les deux estimations.

Les transports moyennés par saison de 6 mois et calculés comme expliqué ci-dessus de la simulation libre (non montrés) présentent une variabilité très similaire à ceux de la simulation avec assimilation de données. L'évènement El Niño/La Niña de 1997/1998 est bien simulé en mode libre. On observe la renverse du transport net aux basses latitudes durant le printemps et l'automne 1997, ainsi que l'affaiblissement puis l'arrêt de l'EUC dans la couche de subsurface sur toute la largeur du Pacifique Tropical. En surface, l'apparition de l'EEC forcé par les WWB est également bien simulée. Les vents ERS+TAO de très bonne qualité y sont vraisemblablement pour beaucoup. L'assimilation a en revanche permis de renforcer de manière significative les courants ainsi que leurs transports. Ce constat est particulièrement vrai dans la couche de surface (i.e $\sigma_\theta < 24 \text{ kg}^{-3}$). Les transports moyens sur la période 1993-1998 diagnostiqués en mode libre sont respectivement de 25 ± 7 Sv pour le SEC et 11 ± 4 Sv pour le NECC. Pour mémoire, le transport du SEC et du NECC dans la simulation avec assimilation de données sont de 31 ± 7 Sv et 15 ± 5 Sv, ce qui représente un gain de 20% pour le SEC et de plus de 35% pour le NECC. Ces modifications sur les courants moyens simulés sont directement imputables à la modification par l'assimilation de données de l'état moyen simulé, notamment la modification de la topographie dynamique moyenne. Les courants, et leurs transports associés, simulés avec assimilation de données semblent plus en accord avec les observations du Pacifique Tropical (cf comparaison avec la climatologie de courants de surface dérivée des bouées dérivantes SVP proposée par *Niiler* (2001) et avec les profils TAO de courant zonal à l'équateur). Le renforcement du NECC, en particulier, qui apparaît être vraiment trop faible dans la simulation libre est un point positif à mettre au crédit de la meilleure représentation de la MDT dans la simulation avec assimilation de données.

Dans la couche de subsurface, le transport de l'EUC passe de 17 ± 3 Sv en mode libre à 18 ± 4 Sv avec assimilation de données. Le transport moyen de l'EUC a donc été assez faiblement modifié par l'assimilation de données. Cela n'est pas très surprenant compte tenu du fait que la pente zonale de la thermocline est déjà bien simulée par le modèle libre comme nous l'avons vu dans le chapitre précédent. Le démarrage de l'EUC a été renforcé dans la simulation avec assimilation de données, les transports pour les sections à l'Ouest de 140°W étant systématiquement plus importants que ceux de la simulation libre. Ceci est en adéquation avec l'intensification, par l'assimilation de données, du cœur de l'EUC dans l'Ouest du Pacifique Tropical que l'on avait constaté en confrontant les courants simulés, aux courants ADCP mesurés par certains mouillages TAO le long de l'équateur.

Cette étude reste très préliminaire. Elle apporte cependant une vision claire et robuste de l'évolution et de la variabilité des transports dans les couches de surface du Pacifique Tropical. Elle confirme notamment l'impressionnante renverse du transport net aux basses latitudes lors du début de l'El Niño de 1997. Il faudrait maintenant affiner les diagnostics afin d'expliquer les mécanismes mis en jeu, et en particulier mieux illustrer la façon dont l'assimilation de données modifie les transports par rapport à la simulation libre. Cette

étude illustre également les potentialités du système d'assimilation conjointe altimétrie/in-situ mis en place dans le cadre de cette thèse pour l'étude de la dynamique du Pacifique Tropical. Au vue des importantes modifications du champ de température moyen simulé avec assimilation de données, il me semblerait instructif de calculer les variations du volume d'eau chaude ($T > 20^\circ$) du Pacifique Tropical, une autre grandeur qui présente une forte variabilité interannuelle, et d'en comprendre les causes. *Jin* (1997a,b) avance d'ailleurs la théorie du "recharge-discharge oscillator" qui explique les oscillations ENSO par un phénomène de stockage et de déstockage du contenu de chaleur des couches de surface de l'océan Pacifique Equatorial.

Conclusions et perspectives

L'objectif de cette thèse était d'explorer la pertinence de l'utilisation d'une surface moyenne gravimétrique de type CHAMP ou GRACE afin de permettre de reconstruire au mieux les circulations océaniques dans le Pacifique Tropical par assimilation de données altimétriques, isolément ou en association avec des données in-situ.

La réalisation de cette objectif passait par la **mise en œuvre des outils de l'assimilation**. L'exploitation du signal altimétrique absolu dans les modèles numériques d'océan via les méthodes d'assimilation de données permet d'aller au delà de la "simple" vision de surface fournie par le satellite et de reconstruire les circulations océaniques dans toutes leurs dimensions spatio-temporelles.

Au delà de l'objectif premier de faisabilité, nous recherchions à mettre en évidence les **apports** bénéfiques potentiels de cette prise en compte **d'une surface moyenne gravimétrique observée**, tant sur de la dynamique simulée que sur le plan de la **complémentarité** multi-données. Il s'agissait donc d'analyser la qualité océanographique des résultats par référence à ce que nous connaissions des circulations du Pacifique Tropical, notamment au travers des observations non assimilées.

De manière ultime, puisqu'une simulation avec assimilation de données fournit une vision haute résolution réellement quadridimensionnelle de l'océan, on pouvait espérer réaliser des diagnostics élaborés de diverses quantités intégrées permettant d'aider à **analyser** et élucider certains processus **physiques** encore obscurs de ces circulations.

L'ensemble de ces objectifs a été atteint, à l'exception du dernier qui n'a été abordé que de manière préliminaire pour les transports de masse zonaux.

Mise en œuvre des outils de l'assimilation

En vue de mesurer l'impact de l'utilisation des données gravimétriques et d'un signal altimétrique absolu sur la dynamique simulée de l'océan Pacifique Tropical, le premier défi à relever fut de nature technologique avec la mise en œuvre d'une plate-forme modèle-assimilation prenant en compte les spécificités liées à l'assimilation d'un signal altimétrique absolu. Le système mis en place est basé sur le modèle OPA (*Madec et al.*, 1998) dans sa configuration globale ORCA2 à surface libre et sur le filtre SEEK (*Pham et al.*, 1998). Le programme SESAM¹ développé au sein de l'équipe MEOM a grandement limité les difficultés pratiques inhérentes à l'implémentation du filtre SEEK avec le modèle OPA en surface libre.

Afin de résoudre l'étape problématique de la réinitialisation du modèle après l'étape d'analyse, problème commun à l'ensemble des méthodes séquentielles, nous avons implémenté une méthode incrémentale de type IAU (*Ourmières et al.*, 2006). Dans le cadre de cette méthode, l'incrément d'analyse évalué par le filtre SEEK est incorporé directement dans les équations pronostiques du modèle sous la forme d'un terme de forçage. En plus de résoudre le problème du redémarrage du modèle après l'analyse, cette méthode agit comme une méthode d'assimilation séquentielle continue, ce qui conduit à obtenir une trajectoire assimilée continue. Dans une perspective de réanalyse, la continuité de la trajectoire assimilée apportée par l'utilisation de l'IAU est un atout de notre système.

Un autre problème auquel il a fallu répondre est la paramétrisation de l'espace d'erreur en rang réduit. En effet, la méthode introduite par *Pham et al.* (1998) pour initialiser le filtre SEEK, reposant sur l'hypothèse sous-jacente que la moyenne de la simulation libre est représentative de la moyenne de l'océan vrai, est apparue inadaptée pour l'assimilation

¹<http://www-meom.hmg.inpg.fr/Web/Assimilation/SESAM/>

d'un signal altimétrique absolu. Nous avons donc été amenés à développer un protocole spécifique pour paramétrer la matrice de covariance d'erreur d'ébauche du filtre SEEK. Dans le cadre d'une assimilation d'une topographie dynamique absolue, on cherche à contraindre la variabilité de modèle libre, mais également son état moyen. La topographie dynamique moyenne assimilée est donc différente de la topographie dynamique moyenne d'une simulation libre. La covariance des erreurs initiales ne peut plus être approximée par la variabilité d'une simulation libre comme le préconise la théorie du SEEK. En effet, utiliser la variabilité d'une simulation libre pour paramétrer le SEEK reviendrait à ignorer toute une partie des directions de l'erreur que l'on cherche à corriger avec l'assimilation de données, celles qui engendrent une structure moyenne différente entre l'océan simulé et l'océan vrai. Pour répondre à ce problème, nous avons proposé d'utiliser un ensemble de différences entre une trajectoire de référence (obtenue grâce à une simulation relaxée vers la climatologie) et le modèle libre (sans assimilation ni nudging). La covariance de cette ensemble est apparue appropriée pour estimer la matrice de covariance des erreurs d'ébauche, nécessaire à la paramétrisation du filtre SEEK.

Le système d'assimilation mis en place s'est avéré efficace. Il nous a permis de mener à bien le premier objectif de cette thèse, à savoir la mise en œuvre d'expériences d'assimilation assimilant un signal altimétrique absolu observé par satellite. Après cette première partie consacrée au développement et à l'implémentation d'un système d'assimilation permettant de prendre en compte efficacement l'information contenue dans la topographie dynamique observée, l'apport principal de ce travail a consisté à estimer l'impact et évaluer les potentialités offertes par l'utilisation d'un signal altimétrique absolu pour reconstruire la circulation du Pacifique Tropical avec assimilation de données. Nous avons pu tirer les enseignements suivants des résultats obtenus.

Apport d'une topographie dynamique moyenne observée

Le premier enseignement que nous avons pu tirer est que l'assimilation des anomalies du niveau de la mer référencées avec la topographie dynamique moyenne observée déduite des données GRACE a permis de significativement améliorer le réalisme du niveau moyen de la mer simulé par le modèle. Comme on pouvait s'y attendre, cette meilleure représentation des structures et des gradients de topographie dynamique à grande échelle s'est traduit par des courants de surface plus réalistes (en particulier le NEC et le NECC).

En subsurface, l'assimilation a également permis de corriger un certain nombre de déficiences du modèle libre. La représentation de la variabilité du champ de température en profondeur a été améliorée ainsi que les gradients de température en subsurface avec notamment une amélioration du réalisme de la pente des isothermes en profondeur, en accord avec les gradients de la topographie dynamique assimilée. En revanche, l'assimilation d'altimétrie seule s'avère incapable de corriger l'erreur sur le positionnement moyen de la thermocline.

On voit probablement les limites d'un système d'observation mono-données et restreint à la seule surface de l'océan. En effet, la quantité d'observations altimétriques, relativement à la résolution du modèle est très importante. Nous avons estimé que l'on disposait d'observations du niveau de la mer pour la quasi-totalité des mailles du premier niveau de la grille modèle. Le problème ne vient donc pas de la quantité d'observations mais bien du fait qu'elles se limitent à la seule surface. Les nombreuses approximations consenties lors de la mise en œuvre et de la paramétrisation de la statistique d'erreur du filtre

SEEK (approximations qui ont pour principal conséquence la perte du caractère optimal de l'analyse) font que l'on ne parvient pas à extrapoler la correction en profondeur de façon suffisamment fiable pour corriger l'erreur sur le positionnement moyen des isothermes.

Devant ce constat, nous avons décidé d'assimiler conjointement au signal altimétrique les données de température in-situ TAO. Le réseau TAO nous fournit, en effet, des observations concomitantes du champ de température en surface, mais aussi et surtout en subsurface (entre la surface et 500 m). Ces données présentent une couverture globale à grande échelle de la bande équatoriale ($\pm 8^\circ$).

Si les données altimétriques et les profils de température in-situ sont des données éminemment complémentaires, leur assimilation conjointe a cependant introduit un problème de compatibilité entre les données. Compte tenu de la précision avec laquelle la topographie dynamique moyenne de référence, qui reste contaminée par des erreurs sur le géoïde, est connue, des "désaccords" subsistent entre l'information fournie par l'observation altimétrique et l'information fournie par l'observation des mouillages TAO. Ces désaccords peuvent se montrer problématiques comme cela a été montré par *Parent et al.* (2003).

Complémentarité des données altimétriques et in-situ

Le deuxième enseignement que nous avons pu tirer est que l'utilisation d'une topographie dynamique observée permet de grandement limiter les "désaccords" qui peuvent exister entre données altimétriques et données in-situ et apportent ainsi une réponse aux problèmes soulevés par *Parent et al.* (2003). Cela permet une meilleure utilisation de la complémentarité existant entre des données altimétriques qui nous fournissent une vision synoptique avec une résolution spatiale et une précision élevée de la surface de l'océan, et les données in-situ, qui sont plus rares, mais viennent efficacement compléter l'altimétrie et sont les seules capables de fournir des mesures sur la verticale. Cette meilleure complémentarité entre les données altimétriques référencées avec la topographie dynamique observée et les données in-situ est la conséquence directe de la meilleure compatibilité présente entre la topographie dynamique observée et les données in-situ, notamment les profils de température TAO. En effet, en améliorant le réalisme de la topographie dynamique moyenne, les données gravimétriques GRACE ont permis de limiter les désaccords entre l'information fournie pour les deux types d'observation et ainsi réaliser une assimilation conjointe efficace. Ceci s'est notamment traduit par une réduction simultanée et forte des écarts RMS par rapports aux deux jeux de données assimilées. Nous avons ainsi pu illustrer ces apports d'une assimilation multi-données par rapport à une assimilation mono-données. Nous avons confirmé ce résultat (et ceux de *Parent, 2000*) en réalisant une expérience assimilant conjointement l'altimétrie référencée par la surface moyenne du modèle et les données de température in-situ TAO. Avec ce type de configuration, on arrive à réduire les erreurs en terme d'écarts RMS par rapport aux observations de l'un des jeux de données assimilées mais pas des deux simultanément. Le fait que l'on ne puisse pas converger simultanément vers les deux jeux d'observation illustre le fait qu'il existe des incompatibilités entre les données in-situ et l'altimétrie lorsque la surface moyenne du modèle est utilisée comme référence. Ces incompatibilités sont la conséquence des erreurs sur le champ thermohalin moyen simulé par le modèle, erreurs qui se traduisent notamment par des erreurs sur la surface moyenne simulée.

L'amélioration de la compatibilité entre l'altimétrie et les données in-situ grâce aux données gravimétriques GRACE et la meilleure utilisation de la complémentarité entre les

deux types de données qui en découle est un résultat fort de cette thèse et a fait l'objet d'une publication dans la revue *Geophysical Research Letter* (*Castruccio et al.*, 2006).

Analyse physique

Nous avons, dans le chapitre 9, esquissé un regard sur la dynamique de l'océan Pacifique Tropical. Cette étude apporte une vision, certes préliminaire, mais claire et robuste de l'évolution et de la variabilité des transports dans les couches de surface du Pacifique Tropical. Elle confirme notamment l'impressionnante renverse du transport net aux basses latitudes lors du début de l'El Niño de 1997. Cette étude est préliminaire et demande à être affinée, mais elle illustre bien les apports de l'assimilation de multi-données sur la dynamique simulée du Pacifique Tropical. Elle démontre également les potentialités du système d'assimilation mis en place pour une analyse de la dynamique du Pacifique Tropical en vue d'élucider certains mécanismes encore mal compris.

Des difficultés résiduelles... mais un possible retour des modélisateurs vers les observateurs

Malgré les apports incontestables offerts par l'utilisation de la topographie dynamique moyenne observée déduite des données gravimétriques GRACE en lieu et place de la surface moyenne du modèle afin de référencer les anomalies du niveau de la mer mesurées par l'altimétrie, des difficultés subsistent encore.

Sur le plan de la méthode, avec la paramétrisation de l'espace d'erreur en rang réduit proposée, les directions d'erreurs évaluées sont contraintes par la climatologie dans la mesure où elles sont évaluées à partir d'erreurs entre le modèle libre et une simulation de référence relaxée vers *Levitus* (1998). Cette méthode permet en quelque sorte de "prédéfinir" des directions d'erreurs privilégiées et ainsi réduire le spectre des directions d'erreurs possibles lors de l'évaluation de la correction sur l'état prédit au moment de l'analyse SEEK. On facilite ainsi en quelque sorte la tâche de l'assimilation de données. Dans la mesure où ces directions prédéfinies sont correctes, cela ne pose aucun problème. En revanche, étant donné que l'on écarte a priori un certain nombre de modes d'erreurs, ce type de paramétrisation dépend très fortement de la validité des hypothèses faites sur la qualité de la simulation de référence.

Sur le plan de la compatibilité entre les données altimétriques et in-situ, nous avons montré que des problèmes persistent le long de 8°N, dans la région du sillon Nord équatorial. Dans cette région, les gradients méridionaux de topographie dynamique moyenne sont très forts. Le manque de résolution de la topographie dynamique moyenne observée (333 km) interdisant de représenter fidèlement ces gradients, nous avons mis en évidence des erreurs entre les signaux moyens mesurés par l'altimétrie d'une part, et le réseau de mouillages TAO, d'autre part. Nous avons également montré que cette différence d'états moyens était problématique dans le cadre d'une assimilation conjointe de ces deux jeux d'observations. Ce résultat est particulièrement intéressant et montre comment, par l'intermédiaire de l'utilisation de données d'observations dans le cadre d'expérience d'assimilation, on peut de manière rétroactive aider à identifier une insuffisance et ainsi améliorer la caractérisation requise des observations.

Les perspectives

Parmi les nombreuses pistes de travail sur lesquelles on peut s'engager à la suite de ce travail, j'en mentionnerais trois :

- Concernant les données. Les développements actuels sur les modèles de géoïde avec les données des satellites CHAMP et surtout GRACE sont extrêmement rapides. Sans compter qu'à l'horizon 2007 la mission GOCE nous fera encore franchir un palier en terme de précision et surtout de résolution avec pour objectif une précision centimétrique pour une résolution meilleure que 100 km. Depuis quelques mois, le dernier modèle purement satellite déduit des données GRACE, EIGEN-GRACE4S est disponible. Il est complet jusqu'à l'harmonique 150 (133 km de résolution). On peut raisonnablement espérer que l'augmentation de résolution apportée par EIGEN-GRACE4S permette de régler les problèmes résiduels de compatibilité entre les données altimétriques et les données in-situ mis en évidence dans la région du sillon Nord équatorial vers 10°N. Une autre perspective des plus intéressantes est de prolonger l'expérience d'assimilation jusqu'à nos jours afin de pouvoir utiliser les profils de température collectés par les bouées profilantes du réseau ARGO. Cela permettrait d'obtenir une observation globale du champ de température des couches supérieures du Pacifique Tropical (entre la surface et 2000 m de profondeur) et non plus des observations limitées au seul guide d'onde équatorial ($\pm 8^\circ$), soit une petite partie seulement de notre domaine d'étude.

- Concernant la méthode. Les flux atmosphériques représentent une source d'erreur importante, si ce n'est la plus importante, dans les OGCMs actuels. C'est la raison pour laquelle il serait nécessaire pour les systèmes d'assimilation de les contrôler, comme cela a d'ailleurs été souligné par Bill Large (*Large, 2006*) lors de l'école d'été GODAE de La Londe Les Maures. Bien que ceci dépasse largement le cadre de cette thèse, ce constat nous a conduit à envisager de nouvelles méthodes pour la paramétrisation de l'espace d'erreur en rang réduit. La tension du vent est une source d'erreurs importante dans la région du Pacifique tropical (*Grima et al., 1999; Wittenberg, 2003*). En partant de l'idée que l'erreur modèle est dominée par l'erreur sur le forçage en vent, nous avons proposé une méthode basée sur un ensemble de simulations obtenues en perturbant la tension du vent $\vec{\tau}$ à l'interface océan/atmosphère. Les premiers résultats sont encourageants (cf annexe B) mais des difficultés persistent. Je reste confiant sur le fait que cette approche est très intéressante et nécessite de plus amples investigations. Je n'ai malheureusement pas eu le temps d'aller plus en avant dans le cadre de ce travail de thèse.

Sur le plan de l'assimilation, nous avons vu que le SEEK dans sa variante stationnaire (qui s'apparente à l'interpolation optimale) permettait déjà d'obtenir de bons résultats. On peut légitimement se demander quelles seraient les apports liés à l'utilisation de variantes plus sophistiquées, mais aussi plus coûteuses comme la variante évolutive (*Ballabrera-Poy et al., 2001*), ou adaptative (*Brasseur et al., 1999; Durand et al., 2003*). Une autre piste intéressante, en particulier dans une optique de réanalyse, est l'hybridation du SEEK et du 4D-Var qui en théorie permet de tirer profit des avantages des deux méthodes (*Robert et al., 2006*).

- Concernant l'analyse physique. Dans le chapitre 9, nous avons montré qu'une simulation avec assimilation pouvait servir de base à une analyse de la dynamique du Pacifique Tropical. Il faudrait maintenant affiner et diversifier les diagnostics afin d'approfondir notre analyse. En particulier, il serait intéressant de mieux identifier et expliquer les processus

par l'intermédiaire desquels l'assimilation influence la dynamique simulée du Pacifique Tropical, et notamment les transports de masse zonaux, ou d'autres quantités comme par exemple la chaleur.

L'utilisation des sorties numériques de la simulation avec assimilation de données présente un réel intérêt pour une analyse et la compréhension de la dynamique du Pacifique Tropical. La thèse d'Angélique Melet qui vient de débiter et s'inscrit dans la continuité du travail présenté dans ce manuscrit devrait permettre d'aller dans ce sens en poussant bien plus en avant l'analyse physique de la dynamique simulée avec assimilation de données pour les objectifs du programme SPICE² (Southwest Pacific Circulation and Climate Experiment).

²<http://www.ird.nc/UR65/SPICE/>

Annexes

Annexe A

Impact of ocean mean dynamic topography on satellite data assimilation

F. Birol⁽¹⁾, J. M. Brankart⁽¹⁾, F. Castruccio⁽¹⁾, P. Brasseur⁽¹⁾, and J. Verron⁽¹⁾

⁽¹⁾Laboratoire des Ecoulements Géophysiques et Industriels, Grenoble, France

Received 1 December 2003; accepted 19 March 2004; published 1 November 2004.

Abstract

The response of an eddy-permitting ocean model to changes imposed by the use of different mean dynamic topographies (MDT) is analyzed in a multivariate assimilation context, allowing the evaluation of this impact, not only on the surface circulation, but also on the interior ocean representation. The assimilation scheme is a reduced-order sequential Kalman filter (SEEK). In a first set of experiments, high resolution sea surface temperature, along-track sea surface height and sea surface salinity from climatology are assimilated into a $1/3^\circ$ resolution North and Tropical Atlantic version of the HYCOM model. In a second experiment, in situ profile data are assimilated in addition to the surface measurements.

The first set of experiments illustrates that important differences in the representation of the horizontal model circulation pattern are related to differences in the MDT used. The objective of assimilation is to improve the representation of the 3D ocean state. However, the imperfect representation of the mean dynamic topography appears to be an important limiting factor with regard to the degree of realism obtained in the simulated flow.

Vertical temperature and salinity profiles are key observations to drive a general circulation ocean model toward a more realistic state. The second set of experiments shows that assimilating them in addition to sea surface measurements is a far from trivial exercise. A specific difficulty is due to inconsistencies between the dynamic topography diagnosed from in situ observations and that diagnosed from sea surface height. These two fields obtained from different data sources do not contain exactly the same information. In order to overcome this difficulty, a strategy is proposed and validated.

A.1 Introduction

Data assimilation for oceanographic applications has entered a new era since the advent of satellite missions that allow accurate measurements with high resolution and global coverage. In particular, radar altimetry provides quasi-global observations of the sea surface height (SSH) measured with respect to the reference ellipsoid. For this reason, SSH measurements have become one of the most important contributions to ocean prediction systems. However, the physical quantity that reflects the dynamic state of the ocean is the absolute ocean dynamic topography, corresponding to altimetric SSH minus the geoid. The geoid is an equipotential gravity surface that is different from the reference ellipsoid. Unfortunately, until now geoid uncertainties have been too large, at the scale of major ocean circulation features, to allow using them for oceanographic purposes, and in particular for the derivation of the absolute dynamic topography. As a result, for most oceanographic applications in general and for data assimilation in particular, the use of altimetric signals has been limited to sea level anomalies (SLA), variations of the sea level about the mean sea level (computed over a certain time span). The CHALLENGING Minisatellite Payload (CHAMP) mission (<http://op.gfz-potsdam.de/champ/>), launched in July 2000, was the first of a series of Geodetic missions that should lead to a precise estimate of the ocean geoid height and, consequently, to improved knowledge of absolute dynamic topography. In particular, the European Space Agency Gravity field and steady-state Ocean Circulation Explorer program (www.esa.int/esaLP/goce.html), scheduled for 2006, and the Gravity Recovery and Climate Experiment (GRACE) ([http://op.gfz-potsdam.de/grace/index GRACE.html](http://op.gfz-potsdam.de/grace/index%20GRACE.html)), launched in 2002, are designed to provide the data set required to obtain an accurate geoid (around 10 cm error) at a high spatial resolution (around 80 km). However, while substantial progress should be made in this field following these missions, the delivery of a product with sufficient precision will still take several years.

Since the quantity that can be assimilated in ocean circulation models is the absolute dynamic topography, we need to add information about the mean ocean state to the altimetric SLA. This distance between the geoid and the mean level of the ocean is called the Mean Dynamic Topography (MDT). (Note that the MDT is usually called mean sea surface height by ocean modelers, since most ocean models use a spherical geoid so that both quantities are identical). Different studies have provided different MDT products, with different characteristics and accuracy. A number of approaches exist. An inverse model of the ocean circulation constrained by observations can provide an estimate of the MDT (*Mercier, 1986*). A synthetic geoid approach consisting of combining satellite altimetry and in situ data can also be considered (*Mitchell et al., 1990*). Another approach is based on ocean model sea surface representation, with or without assimilation (*Killworth et al., 2001*). These different methods provide descriptions of the mean ocean circulation that are not necessarily the same.

Because of geostrophic balance, ocean surface currents are proportional to the gradient of sea surface height. The change in the slope of the sea level elevation imposed by the assimilation is expected to modify the simulated surface circulation. Surface currents as well as absolute dynamic topography contain the integrated signature of ocean processes within the whole water column. Therefore assimilating this quantity also implies making a correction to the ocean thermohaline structure. The use of imprecise MDT products to compute the absolute value of the ocean dynamic topography will impact the simulated

oceanic features. In addition to the sensitivity of surface currents, it is important to determine to which extent the choice of a particular MDT has an impact on the interior ocean properties estimated by the assimilation system. Both ocean model and data assimilation schemes have now reached a degree of accuracy where the realism of the imposed MDT can constitute a major limitation in forcing ocean models towards reality on regional scales. This is why it appears important to study the impact of state-of-the-art MDT products on assimilation hindcasts. A better knowledge of the sensitivity of ocean simulations to MDT uncertainties can also lead to significantly improved estimates of MDT.

With extensive hydrographic programs like ARGO, we have now entered a phase where the challenge is to combine optimally surface and subsurface information into a single estimation process. Through these observations, it is expected to better control the representation of subsurface fields/water masses. This also leads to a number of relevant and challenging questions. Some recent studies have been conducted to explore the impact of the assimilation of sparse in situ data compared to satellite data, in terms of spatial and temporal resolution (*Li et al.*, 2003). A specific question that needs to be addressed here relates to whether the dynamic topography from in situ observations is consistent with that obtained by the addition of SLA and the imposed MDT in a multivariate assimilation system. Indeed, there is no reason why these two fields, which are different in nature, obtained from different data sources and associated with different kind of errors, should contain completely compatible information.

The present study aims to demonstrate the sensitivity of a system assimilating altimetric data to different available MDT products. The detailed validation of the MDTs, which can be addressed independently of any assimilative system, is not the objective here. Such a validation or quantitative comparison would be dependent on the model and assimilation system used. Both the case where only sea surface data are assimilated and that where vertical temperature and salinity profiles are added are addressed. The next section describes the numerical model, assimilation method, and data sets. Following this, a section concentrates on problem definition and the experimental framework and another contains the results of the sensitivity study. Then a section deals with the specific question of the assimilation of temperature and salinity (TS) profiles. Finally, a summary and conclusions are presented.

A.2 Numerical model, assimilation method and data sets

A.2.1 Numerical model

The circulation model used is HYCOM, the HYbrid Coordinate Ocean Model (*Bleck, 2002; Chassignet et al., 2003; Halliwell, 2004*) developed at Rosenstiel School of Marine and Atmospheric Science (RSMAS). The computational domain is the North and Tropical Atlantic Ocean basin from 28°S to 70°N, including all the Mediterranean Sea. The horizontal grid resolution is given by $1/3^\circ \times 1/3^\circ \cos(\phi)$, where ϕ is the latitude. The Northern and Southern boundaries are closed and relaxation to climatology is applied within a 30 grid-point zone adjacent to the wall. The bathymetry is interpolated from the ETOPO5 database onto the model grid. Subgrid-scale mixing is parameterized using a biharmonic dissipation operator. Vertical mixing is governed by the K-Profile Parameterization (KPP) mixing scheme (*Large et al., 1994*).

The hybrid vertical coordinate is isopycnal in the open stratified ocean, z-level in the

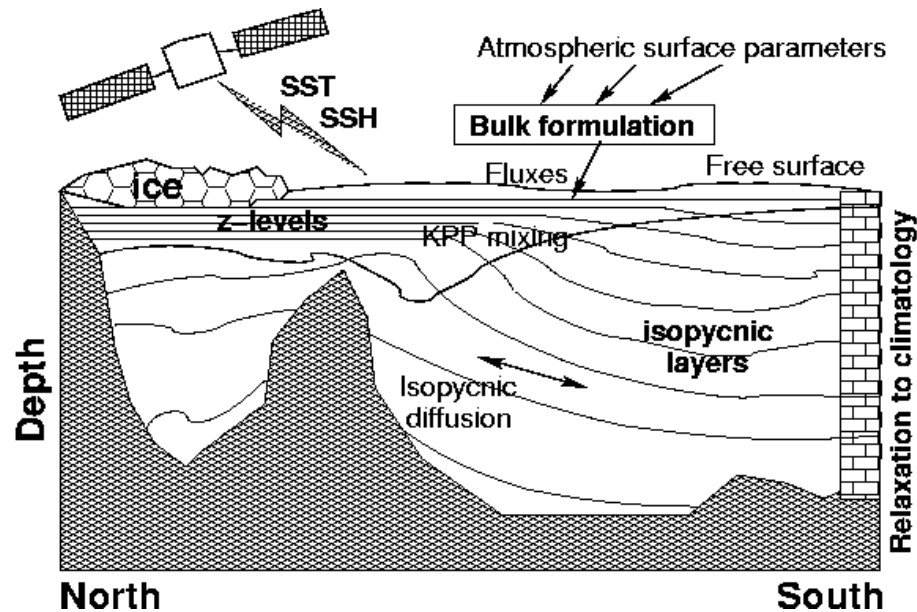


FIG. A.1 – Sketch of the HYCOM model configuration.

mixed layer and unstratified areas, and smoothly reverts to a terrain-following coordinate in shallow coastal regions (Figure A.1). In this version, the vertical density structure is represented by 26 hybrid layers¹. If the density of a given layer deviates from its isopycnic reference, the interfaces bounding this layer are adjusted, so that density is restored to its target value (adding denser or lighter water if the water is too light or too dense). If the layer interface approaches a prescribed minimum thickness, isopycnic conditions are not entirely restored, and a cushion function is used to recalculate a vertical or sigma coordinate location (depending on bottom depth). The lower layers are isopycnic and exist only if the corresponding model target density of each layer exists; otherwise they collapse to zero thickness. Hydrographic and momentum equations have been modified to handle a generalized vertical coordinate. In each layer, both density and salinity are advected. The theoretical basis for implementing the hybrid coordinate can be found in *Bleck et Boudra* (1981).

The model is initialized using the Levitus climatology and is then spun up from rest during a 9-year time period, before running a 1985-1996 interannual experiment (without assimilation). Bulk formulation is used and the model is driven by wind stress, wind speed, air temperature and humidity, precipitation, and longwave and shortwave surface radiation. During the spinup of the model, the forcing fields are derived from the Comprehensive Ocean-Atmosphere Data Set (COADS) monthly climatology. The sea surface salinity and sea surface temperature are relaxed towards the Levitus climatological fields with a 30-day timescale. During the 1985-1996 free model simulation and during the assimilation experiments, high frequency (every six hours) European Center for Medium-Range Weather Forecasts (ECMWF) atmospheric data sets are used.

¹The prescribed reference values are : 19.5, 20.25, 21.0, 21.75, 22.5, 23.25, 24.0, 24.77, 25.28, 25.77, 26.18, 26.57, 26.8, 27.03, 27.22, 27.38, 27.52, 27.64, 27.74, 27.82, 27.88, 27.94, 28.0, 28.06, 28.09 and 28.12.

A.2.2 Assimilation method

The assimilation method is derived from the SEEK filter, a reduced-order Kalman filter introduced by *Pham et al.* (1998). This sequential method has already been described and used in a number of studies (*Brasseur et al.*, 1999; *Penduff et al.*, 2002; *Durand et al.*, 2003). In the present implementation, we use the procedure described in detail by *Brankart et al.* (2003) and updated to take into account the hybrid nature of the vertical coordinate of HYCOM. The estimation vector is composed of the interface pressure, temperature and salinity for each of the 26 layers, in addition to the sea surface elevation. At every assimilation step, the misfit between observations and their model counterparts is evaluated and projected onto the dominant error modes. The background error is initialized using the system variability diagnosed from a free model run. In practice, we use the 10 dominant Empirical Orthogonal Function (EOFs) of a 3-year sample of model snapshots. A suboptimality is introduced in the scheme by keeping the background error covariance matrix unchanged from one assimilation cycle to the next. However, the use of spatially-dependant, three-dimensional, multivariate statistics derived from a free model integration relates the present method to a simplified, reduced-order Kalman filter. The gain is local, meaning that the control of a specified water column will depend only on the observations within a specified influence bubble. In practice, the size of each influence bubble is parameterized inside a box of 14×14 grid points. Observation errors are assumed to be independent; the corresponding prescribed values are given in section on data set which follows.

In order to reconcile statistical optimality with the physical constraints on the model state vector, an adjustment operator is introduced before model restart to restore the model constraints. A detailed justification of this operation can be found in *Brankart et al.* (2003), where a similar approach was successfully applied to the MICOM model. The adjustment step consists of three main operations. First, we consider that the correlation between surface observations and the state variables at depth is not sufficiently reliable. In practice, only the state variables located inside the mixed layer are updated if TS profiles are not assimilated. Second, we restore the hydrostatic stability of the water column. The objective is to generate a corrected analyzed state where layer density increases with depth and layer thicknesses are greater than their minimum value (HYCOM imposes a minimum thickness on each layer). In particular, if a layer has been reduced by too much during the statistical analysis, the excess reduction is transferred to the closest layer. Third, we apply a change along the Cooper and Haines mode (i.e., a vertical shift of the water column) to adjust the vertical structure of the water column to the analyzed dynamic topography (*Cooper et Haines*, 1996).

A.2.3 Data sets

Sea Surface Topography

The altimetric observations consist of along-track sea surface topography, obtained as the sum of TOPEX/Poseidon or ERS altimeter SLA, and the mean dynamic topography (see next section). The assimilation window is three days. Each analysis is computed using all the available data gathered within a 3-day interval (1.5 days before and after the analysis time). The accuracy of the SLA can be estimated at around 3 cm, except in the coastal zones, where tidal corrections are less accurate. We therefore decided to mask SLA

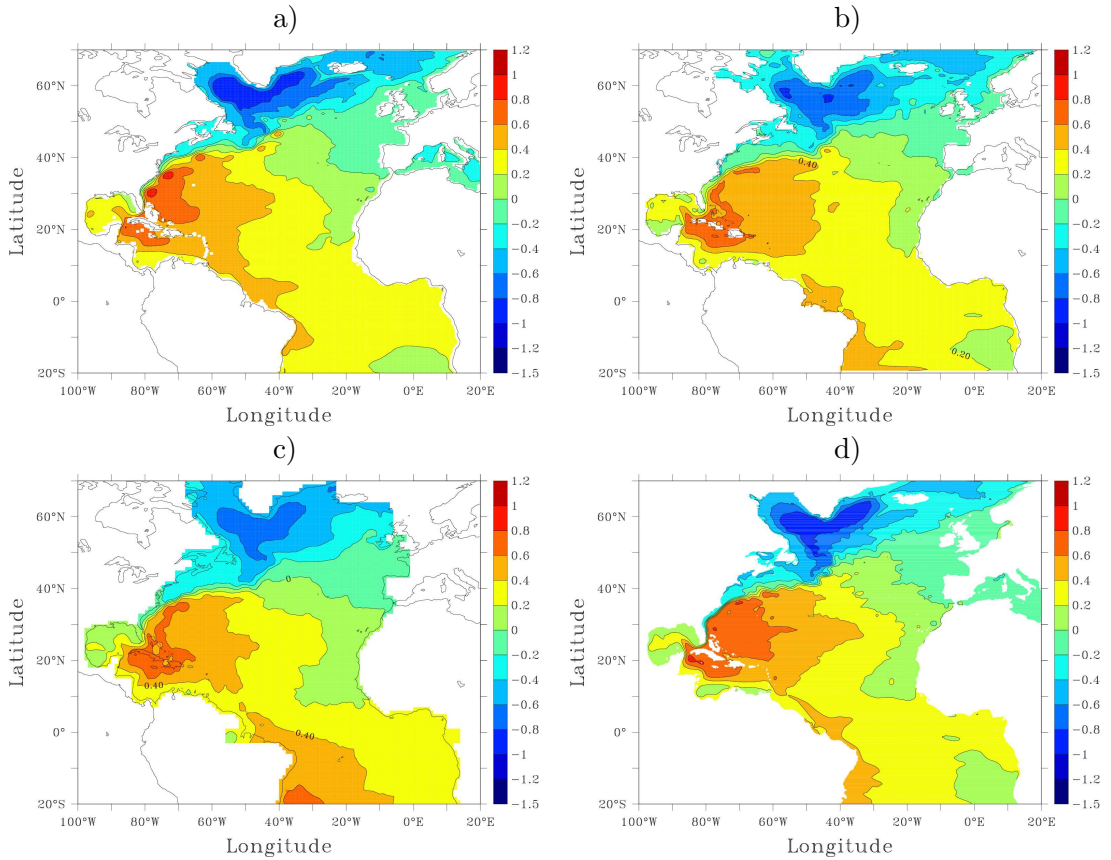


FIG. A.2 – *Mean Dynamic Topography (in meters)*. (a) *MDT0*, (b) *MDT1*, (c) *MDT2* and (d) *MDT3*.

data where the bottom depth is less than 150 m. In order to eliminate spurious effects due to the transition between zones where the model state is corrected by the assimilation of SLA and zones where it is not (corresponding to significant portions of the American and European continental shelves), we replace the sea surface topography by the mean dynamic topography in such coastal areas. A bulk error of 3 cm RMS was prescribed on the total sea surface topography data in the assimilation system.

For the present study, three different MDTs for the North and Tropical Ocean, computed with three different methods, were collected and used to perform three different assimilation experiments. The first, referred to as MDT1, is the MDT from *Rio et Hernandez* (2002), the second, referred to as MDT2, is from *Le Grand* (1998), and the third, referred to as MDT3, is derived from a $1/12^\circ$ Atlantic MICOM experiment (*Chassignet et Garraffo*, 2001).

MDT1 is a $1^\circ \times 1^\circ$ product, estimated from a multivariate statistical analysis using hydrographic data, surface drifter velocities and altimetry over the 1993-1999 period. It is based on the SMO CLS01 mean sea surface (*Hernandez et al.*, 2001) and the EIGEN2 geoid computed from CHAMP data. This MDT is produced to match the altimetric average removal period (1993-1999); the associated drawback is that in situ and drifter measurements available during this period provide insufficient sampling in some areas.

MDT2 is a $1^\circ \times 1^\circ$ product obtained by merging in situ measurements and altimetric

observations using a nonlinear inverse model (*Mercier, 1986*). Hydrographic data come from a compilation of 70 years and altimetric observations cover the 1993-1996 period. This approach is based on dynamical considerations but still strongly limited by resolution and accuracy of available data.

MDT3 is a 2-year average of a sea surface elevation time series from a $1/12^\circ$ Atlantic MICOM free run experiment performed with ECMWF forcing. This product has a high spatial resolution, but the associated mean field includes systematic model errors (mainly due to unresolved physics, numerical errors, errors in boundary conditions, or forcings).

Figure A.2 shows these three different MDT fields, which roughly correspond to three available types of MDT estimates. The 4-year average MDT computed from a $1/3^\circ$ Atlantic HYCOM free run experiment over the period 1993-1996 (named MDT0) is also shown for comparison. With regard to large-scale features, the four MDTs are fairly similar in terms of structure and amplitude, showing the classical picture of the mean ocean circulation in the North and Tropical Atlantic Oceans. The signature of the main currents is present.

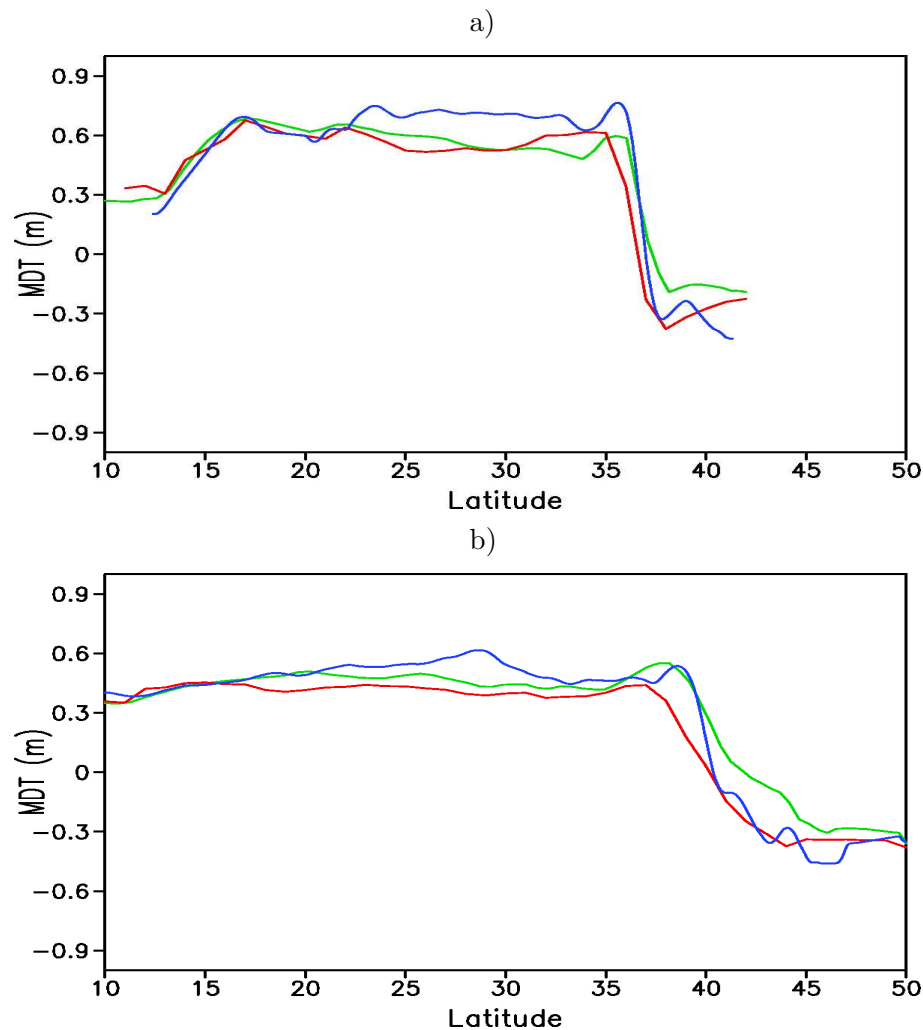


FIG. A.3 – Section of the different mean dynamic topographies (in meters) along (a) 72° W and (b) along 55° W. MDT1 is in green, MDT2 is in red and MDT3 is in blue.

All MDTs represent a realistic Gulf Stream separation at Cape Hatteras, except MDT0 that is affected by a significant model bias in the Gulf Stream region. The Gulf Stream (GS) overshoots in HYCOM $1/3^\circ$; this is a classical problem of ocean numerical models at eddy-permitting resolutions. The three MSSH fields differ mainly at the smaller scales, and also in their representation of several regional circulation features. In broad terms, MDT2 is smoother than the other two. This is due to the longer period of observations used to compute this field. In MDT3, the Florida Current, the Gulf Stream and North Atlantic Current (NAC) front are sharper than in the other MDTs. This is due to a better representation of the strength of ocean currents in a $1/12$ degree grid mean than in a lower resolution mean. In MDT1, MDT2, and MDT3, the NAC front turns north around 45°W , and then toward the northeast around 40°W , forming the North Atlantic Drift Current along the eastern flank of the Reykjanes Ridge. Many small-scale discrepancies between the different MDTs can be observed all along the resulting fronts. Another quantitative difference can be found in the subpolar gyre, with fronts associated with the Labrador Current and the East/West Greenland Current system that are significantly weaker in MDT2, illustrating a poorly defined subpolar gyre circulation.

An illustration of the differences of the Gulf Stream front representation by the different MDTs is given by Figure A.3, which represents cross-sections of the three MDT products along 72°W (Figure A.3a) and along 55°W (Figure A.3b). Both sections cross the Gulf Stream front, represented by a sharp MDT gradient. The difference across the Gulf Stream front is 10 cm larger in MDT3 than in MDT1 and MDT2. MDT1 has a weak front amplitude just off Cape Hatteras (20 cm lower than MDT2). MDT3 and, to a lesser extent MDT1, have secondary fronts south and north, associated to the inertial Gulf Stream recirculation, which are not present in MDT2. The Gulf Stream mean path is the same in MDT2 and MDT3, but is different in MDT1, with a too Northern position (Figure A.3b).

Sea Surface Temperature and Salinity

The Sea Surface Temperature (SST) observations assimilated in HYCOM consist of high-resolution composite AVHRR data (10 km resolution at the equator). The available SST data are gathered within the 3-day assimilation time window. An observation error of 0.4°C is prescribed. The SST data coverage can be strongly limited by the presence of clouds, particularly at high latitudes and during the winter season.

For multivariate data assimilation, inaccuracies in the specification of forecast error covariances can lead to inappropriate corrections of unobserved fields like SSS. In order to avoid such problems, we assimilated Levitus 98 monthly climatological Sea Surface Salinity (SSS) in addition to SST and SSH. Unlike SST or SLA, SSS measurements are sparse both in time and space. The European Space Agency SMOS mission that will be launched in the near future will provide SSS data every 10 days at 200 km resolution, but at the present time there is no high resolution SSS data set available. Because climatological fields are very smooth, we introduce a smoothing operator into the observation operator to compute the model equivalent to the data. In this way, only large-scale features in the model can be influenced by SSS data assimilation. For this reason, the error associated to the large-scale averaged SSS is set to 0.05 psu.

In-situ profiles

We could have chosen WOCE temperature and salinity profiles as data sources for our in situ data assimilation experiment. However, their spatial and temporal distribution remains sparse, with many sites having only temperature profiles and most of them having a poor vertical extension compared to local ocean depth. These features would have induced specific problems.

In the multivariate data assimilation experiment, we first want to illustrate important aspects of the complementarity between surface and subsurface data. Thus, we have chosen to extract the assimilated profiles from the Levitus 98 monthly climatology. In order to enable assimilation into HYCOM, they are transformed into hybrid coordinates, so that observations of temperature, salinity, and interface pressure are available for each layer and are effectively assimilated. Note that correlations between interface pressure errors and layer TS errors have been dropped in the background error covariance. Observation errors have been set at 0.2° , 0.1 psu, and 20 m, respectively, for temperature, salinity, and pressure measurements. This information is used with an adequate observation operator (same treatment as for climatological SSS), so that it only influences the large-scale features of the 3-D TS structure.

A.3 Problem definition and experimental framework

A.3.1 Problem definition

Since information about the geoid is not sufficiently precise, the mean (MSSH) of the altimetric signal SSH has to be removed, so that only the altimetric residuals can be accessed. To obtain the quantity assimilated into the model, that is, the absolute ocean dynamic topography (ADT), a mean dynamic topography (MDT) must be added to the SLA. Ideally, this MDT should correspond exactly to the average state of the ocean over the time period corresponding to the altimeter measurements. One simple solution would be to use the model mean ocean state over this time period. Unfortunately, this variable is affected by systematic errors, due to various model deficiencies. For this reason, it is necessary to use a more realistic mean ocean state. Our assimilation experiments will be performed by adding one of the MDT described in the section on data sets to the SLA.

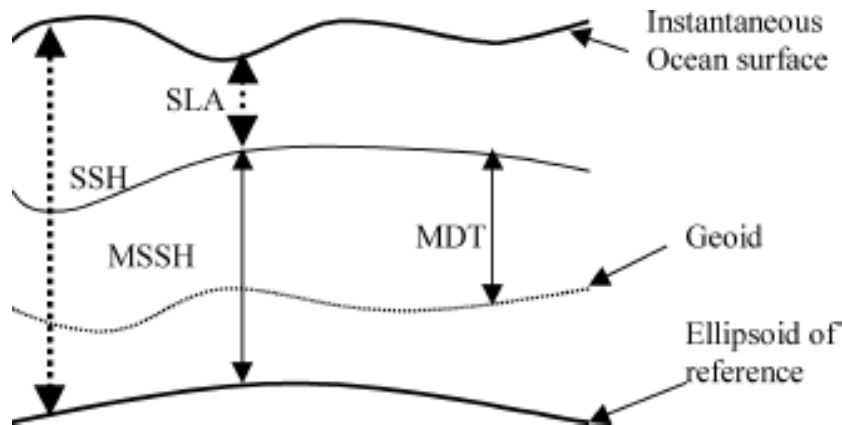
Figure A.4 illustrates the definition of the height measurements used. In summary :

$$ADT = SSH - MSSH + MDT \quad (\text{A.1})$$

MDT is a synthetic product and can then be altered by (1) errors associated with the data and the method used to compute this field, (2) bias related to nonuniform time and space coverage of the observation data sets used, and (3) bias introduced by the fact that the average state of the ocean is calculated over a time period which does not correspond to the altimeter measurements. These differences can lead to spurious effects in the assimilation experiments. Thus, let us define :

$$MDT(x, y) = S0 + S1(x, y) \quad (\text{A.2})$$

where $\int \int_D = S1(x, y) dx dy = 0$, D being the model ocean domain.

FIG. A.4 – *Definition of height measurements.*

S_1 is due to permanent ocean circulation patterns that vary geographically and determines a specific shape of the MDT. S_0 is the horizontal mean difference between the geoid and the MDT. S_0 is then purely conventional and defines a reference surface parallel to the geoid. In a system covering the global ocean domain, S_0 is conventionally equal to 0.

Problems arise because (1) the conventional reference may not be the same for the model and the data (for instance, our model is regional and was initialized with S_0 equal to 0 over the North and tropical Atlantic), and because (2) during the spinup, the model may have drifted from the initial climatological state. For these reasons, a significant horizontal mean difference exists between the MDT and the corresponding current model state. In the absence of vertical information (i.e., vertical profiles acting as an additional constraint in the assimilation system) that could track these misfits back to identifiable causes, no doubt we will face difficulties in interpreting this important part of the dynamic topography. This misfit can be of the same order of magnitude as the part due to sea-level variations. To account for this problem, an averaged misfit value S_0 is introduced, characteristic of the current model state. The S_0 component is removed from the corresponding MDT field, meaning that we do not correct the error corresponding to the model drift in the initial model state. Otherwise, layer interfaces could be shifted along the Cooper and Haines mode, resulting in unrealistic vertical displacement.

A.3.2 Description of the experiments

In a first set of experiments, four simulations are performed between October 1992 and December 1993. The first (called EXPT0) is a free-run simulation, without assimilation. It is defined as the control run. The other three differ from the first in terms of the assimilation of sea surface temperature, sea surface salinity, and sea surface height observations. The same initial condition and the same forcing fields were used for these experiments. They differ only in the MDT used to compute the absolute dynamic topography assimilated into the system : MDT1 for EXPT1, MDT2 for EXPT2, and MDT3 for EXPT3.

In a second stage, a fifth simulation, called EXPT4, is performed over the same period (October 1992-December 1993). In EXPT4, TS profiles are assimilated in addition to the surface data and MDT2 is used. The methodology is described in the section on the assimilation of TS profiles.

In this study, we present different diagnostics of these simulations. Focusing on the North Atlantic region, these diagnostics are all calculated from the model forecasts, before statistical correction, and over the year 1993 (from January to December).

A.4 Sensitivity tests : results

The system response is analyzed for the first set of simulations. The ability of the model, with assimilation, to reproduce the essential elements of the North Atlantic circulation is discussed. The sensitivity of the solution to the different MDTs is analyzed.

For this study, we assume that the assimilation results have already been validated. The diagnostics presented here are (1) the surface currents which are directly related to the changes in MDT through geostrophy and are crucial for climate studies and many operational applications ; and (2) the mean barotropic streamfunction as an expression of the vertically integrated circulation pattern. A temperature section along 55°W is also examined to consider the sensitivity of the solution in terms of thermohaline characteristics.

A.4.1 Surface currents

Figure A.5 shows the difference between the simulations in terms of the mean currents for 1993 at 50 m depths. The classical surface currents of the North Atlantic are clearly visible in all solutions but, as expected, there are also important regional differences.

The mean flow of the Florida Current is particularly strong in the control run and in EXPT3, where it exceeds 90 cm/s, while it is about 55 cm/s in EXPT1 and 65 cm/s in EXPT2. This result is coherent with the local gradients of the different MDTs, which are more realistic in EXPT0 and EXPT3 (*Fratantoni, 2001*). Although the corresponding current velocities are greater in EXPT3, the GS extension and the NAC pathway are fairly similar in all assimilation solutions. The NAC turns north around 45°W , and then northeast around 40°W . While it is around 30 cm/s in EXPT1 and exceeds 40 cm/s in EXPT3 and the control run, it is significantly weaker in EXPT2, with a mean speed of around 20 cm/s. The latter value is too low (*Fratantoni, 2001*). In all assimilation experiments, the Azores Current is shifted to the south and its flow slightly increased. From less than 10 cm/s in the control run, its speed reaches around 12 cm/s, with a maximum of 15 cm/s, in EXPT1, EXPT2, and EXPT3 near 25°W .

North of 50°N , the representation of the main surface current system is more variable from one experiment to another. In both the control run and EXPT3, west of 30°W , one branch of the NAC turns northward to join the Irminger basin, before crossing the Mid-Atlantic Ridge. This branch is much weaker in EXPT1 and EXPT2. The North Atlantic Drift Current is particularly strong in the control run, with the mean surface current exceeding 20 cm/s in some areas, which appears unrealistic. Observations highlight various branches of weak currents flowing to the northeast, rather than an intense and continuous flow (*Reverdin et al., 2003*). By contrast, in EXPT2, the surface currents are weak in the whole northeast region.

The East-West Greenland/Labrador Current system appears fairly different from one solution to another. The corresponding mean currents are reduced in all assimilation experiments, and particularly in EXPT2. With a mean velocity of about 35 cm/s in the control run, the Labrador current reaches 25 cm/s in EXPT1 and around 15 cm/s in EXPT2 and EXPT3. By comparison to *Reynaud et al. (1995)*, the latter value is too weak. With a mean

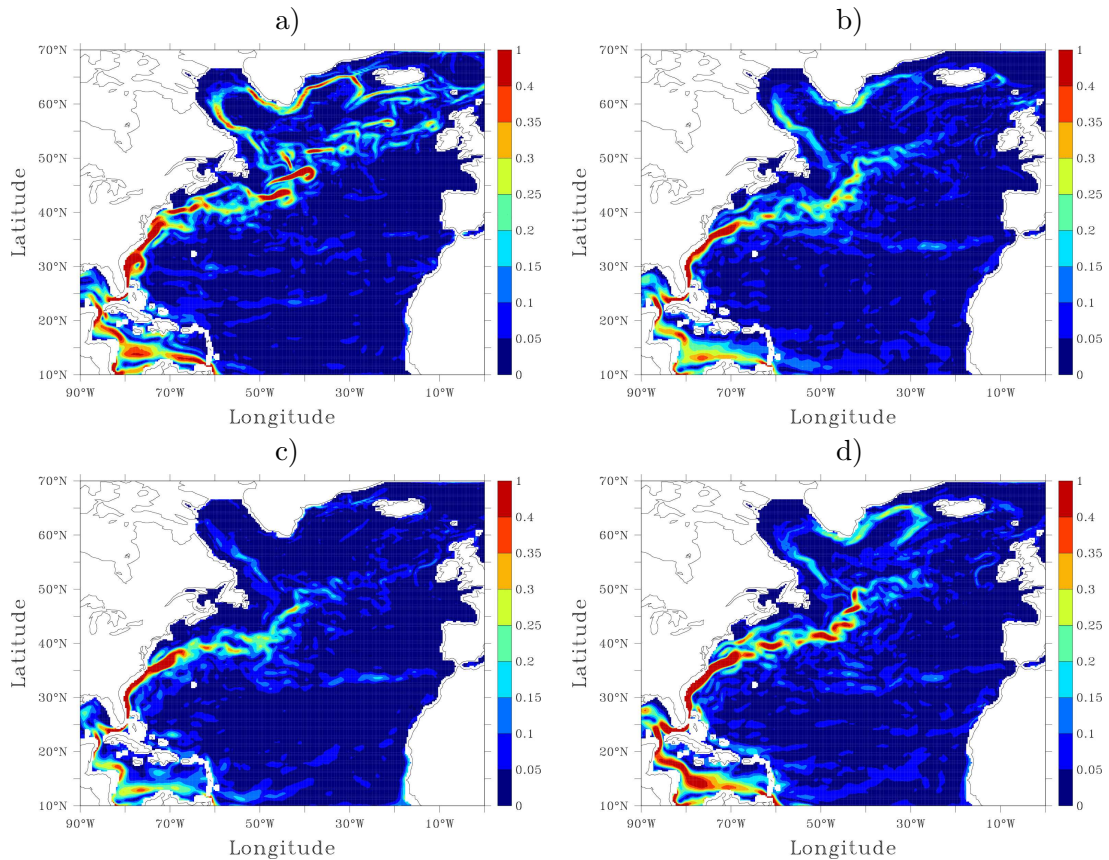


FIG. A.5 – Mean surface current velocity for (a) the control run *EXPT0*, and the (b) *EXPT1*, (c) *EXPT2* and (d) *EXPT3* assimilation experiment.

speed of about 35 cm/s in the control run, the West Greenland Current flows at around 5 cm/s in *EXPT2*, around 15 cm/s in *EXPT1*, and 20 cm/s in *EXPT3*. In *EXPT2*, all the West Greenland Current turns west near 61°N whereas it turns west near 63°N in the control run; in *EXPT1*, and *EXPT3*, different branches exist between 60°N and 64°N, which is in better agreement with observations (*Cuny et al.*, 2002).

For each of the experiments, the regional differences in the simulated surface currents can be largely attributed to differences in the MDTs, moderated by the effect of the assimilation. In *EXPT2*, the currents are generally weaker. In contrast, *EXPT3* exhibits the highest current values. This can be related to the MDT resolution; whereas, MDT2 is the smoothest MDT solution and MDT3, which was deduced from a 1/12° free-run simulation, has the highest resolution.

A.4.2 Barotropic streamfunction

Figure A.6 shows the 1993 mean barotropic streamfunction for the four simulations. Note first that changes introduced by assimilation are significant. Barotropic transport is increased by assimilation, and the increase is much larger in *EXPT1* and *EXPT3* than in *EXPT2*. At the global scale, all simulations show well-developed gyres of similar structure, though they all differ in amplitude. At the regional scale, however, differences in structure

may also be observed.

The Northwestern subtropical gyre limit has drifted southward with the assimilation, thus correcting the GS pathway. Both the transport associated with the Florida Current/Gulf Stream system and the associated recirculation are more pronounced in EXPT1 and EXPT3 than in the other solution. Numerous small cells are found along the GS and NAC paths, which have been shifted by the assimilation along the line of maximum gradient of mean dynamic topography. These small-scale structures are very different from one solution to the next but are much larger and stronger in EXPT1 and EXPT3. This feature can be related to the presence of small-scale structures along the subtropical front in MDT1 and MDT3 that are almost absent in MDT2.

The amplitude of the subpolar gyre also differs significantly. With a magnitude of approximately 30 Sv in the control run, it maintains approximately the same magnitude in EXPT1, increases to 50 Sv in EXPT3, and decreases to a value of 20 Sv in EXPT2. Significant differences are also observed in its structure. The location of maximum transport varies. Both subpolar gyre structure and maximum transport location observed in the solution are clearly correlated with the subpolar gyre structure and minimum mean sea level location of the different MDTs.

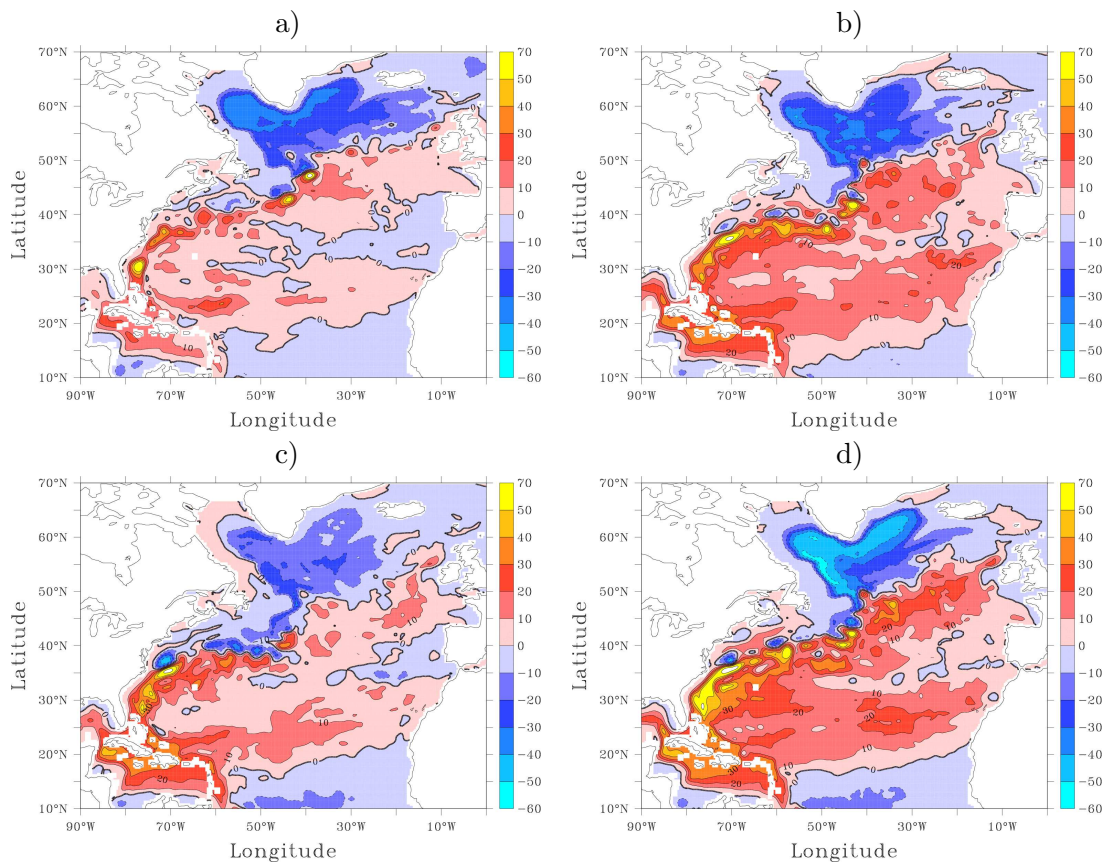


FIG. A.6 – Mean barotropic streamfunction for (a) the control run EXPT0, and the (b) EXPT1, (c) EXPT2 and (d) EXPT3 assimilation experiment.

A.4.3 Vertical temperature section at 55°W

As already observed in the previous section, the assimilation of surface data modifies the representation of not only the sea surface fields, but also the entire ocean state, through a vertical extrapolation of sea surface information. In this section, we analyze the sensitivity of the interior thermohaline properties to the use of different MDTs in the assimilation system.

Figure A.7a represents the thermal structure along the 55°W section from Levitus 98 climatology, while Figures A.7b, A.7c, A.7d, and A.7e show the differences along this section between the mean thermal structure of, respectively, the control run, EXPT1, EXPT2, EXPT3, and the climatological thermal structure of Figure A.7a. Assimilation strongly modifies the subsurface temperature structure and significant differences are observed from one experiment to another.

In the Labrador Sea, temperatures in the control run are more than 2°C too warm over the Northern and Southern limits of the Labrador Sea, where the West Greenland/Labrador Current system flows. In the assimilation runs, these anomalies have been reduced and even replaced by a cold anomaly in the Northern part of around 0.5°C in EXPT2. In the interior basin, the temperature is colder from the surface to the bottom in all assimilation runs, but the amplitude of cooling varies from one experiment to another.

The most marked anomaly with respect to the climatology is observed in the control experiment north of 40°N, just south of Newfoundland. Subsurface temperatures are as much as 5°C too warm in the upper 500 meters, associated with the large-scale “misplacement” of warm water from the Gulf Stream extension (located around 41°N) and the associated absence of colder Labrador Current water near the continental shelf. This anomaly has been considerably reduced in all assimilation experiments, especially in EXPT2 and EXPT3. The differences observed in the amplitude of this anomaly reduction are due to the presence of a small northern branch of the GS present near 45°N (its signature is also observed on Figure A.5), preventing colder water from the Labrador Sea turning southwest around Newfoundland. This current becomes weaker in EXPT2 than in EXPT1 and EXPT3, which explains the differences in the thermal structure observed in this area in the different assimilation simulations. Parallel to the observed decrease in the warm anomaly, a negative feature of 2°C amplitude appears north of 40°N, between depths of 500 m and 1000 m in EXPT2. This negative feature can be associated with the change in Labrador Sea water mass characteristics noted in the previous paragraph.

In the midlatitudes, between 20°N and 40°N, the control run is too warm between approximately 100 m and 600 m compared with the climatology. This feature is significantly amplified in both location and amplitude by the assimilation in EXPT3, an anomaly that corresponds to the presence of 18°C Mode Water formed in the Sargasso Sea. In the climatology, this water mass has a temperature of 17°C-18°C (Figure A.7a) and extends between depths of 200 m and 450 m. In all model runs, the water mass is too warm with a temperature of around 19°C. The misrepresentation of water mass structure results in a stronger positive anomaly in EXPT3, where it is too deep.

The results of this section illustrate the impact of the choice of the MDT product on different ocean estimates. Although the impact of assimilation is generally positive, the different experiments do not perform equally well everywhere. The surface currents and the deep properties are very sensitive to the change in mean dynamic topography, in particular where gradient differences are observed. In addition, the MDT, through different

mechanisms, also has an impact on a wide range of model oceanic processes, such as volume transport and thermohaline circulation which are shown to be very sensitive quantities.

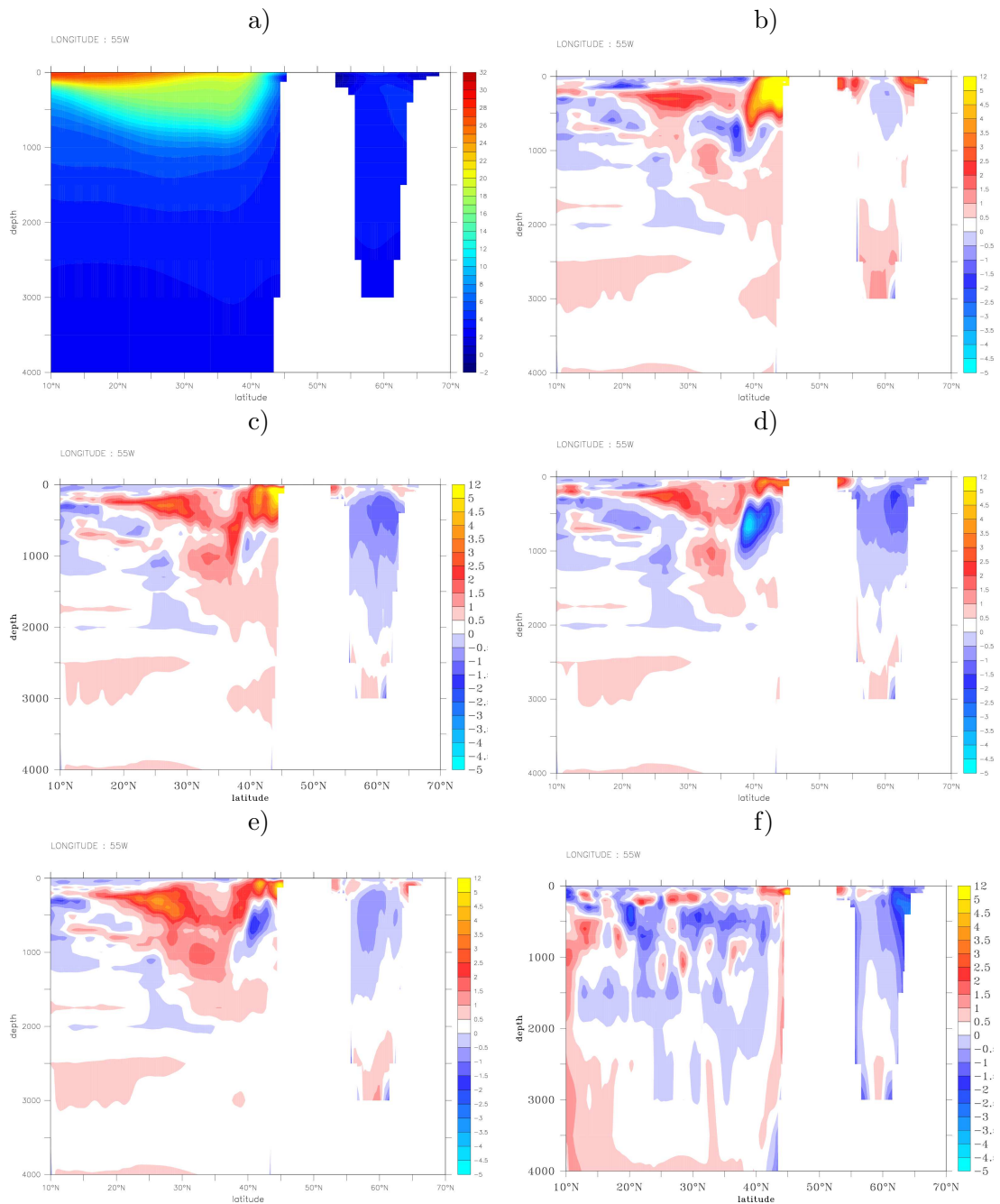


FIG. A.7 – (a) Temperature section along 55° W from Levitus climatology. Anomaly between the mean temperature from (b) the control run, (c) EXPT1, (d) EXPT2, (e) EXPT3, (f) EXPT4 and the climatological temperature field along 55° W.

A.5 Comparison with independent TS profiles

Another way to compare results is to validate each experiment using in situ data. For this purpose, the model thermal and salinity fields were compared to (a) the climatology and (b) an ensemble of XBT profiles collected in the North Atlantic. Figure A.8 shows the distribution of available XBTs in 1993.

RMS misfits between the XBTs or climatology and the equivalent model state were computed for both the Gulf Stream region and the Labrador basin and are shown in Figure A.9. For each assimilation experiment, the continuous-line curve represents the forecast and the dashed-line curve the analyses (after dynamical adjustment). There is little difference between analyses and forecasts, indicating a stable assimilation system. Here, the discussion is limited to the results of the experiments where no TS profiles are assimilated.

In the Gulf Stream region, where the results are particularly sensitive, EXPT1 and EXPT3 are closer than the free run to both climatology and XBTs down to a depth of at least 400 m. However, EXPT2 shows better results than EXPT1 and EXPT3 down to a depth of 700 m, with misfits remaining better than the free run. Note that the vertical structure of the temperature misfits exhibits a local maximum around 100 m depth, which is related

to the difficulty in correctly simulating the mixed layer depth. This maximum is greatly reduced by the assimilation, particularly in EXPT2. Note also that the error in salinity is also reduced, due to the assimilation of surface salinity observations.

In the Labrador region, the results are different. Here again, the impact of assimilation is positive, but the differences between the temperature misfits of EXPT1, EXPT2, and

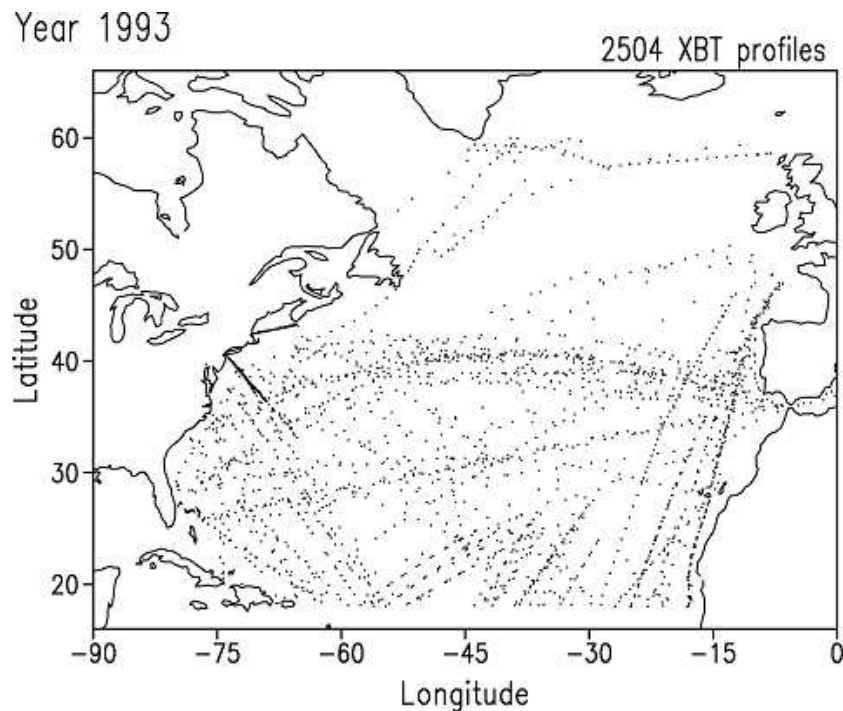


FIG. A.8 – Horizontal distribution of XBT profiles available during 1993 (SISMER).

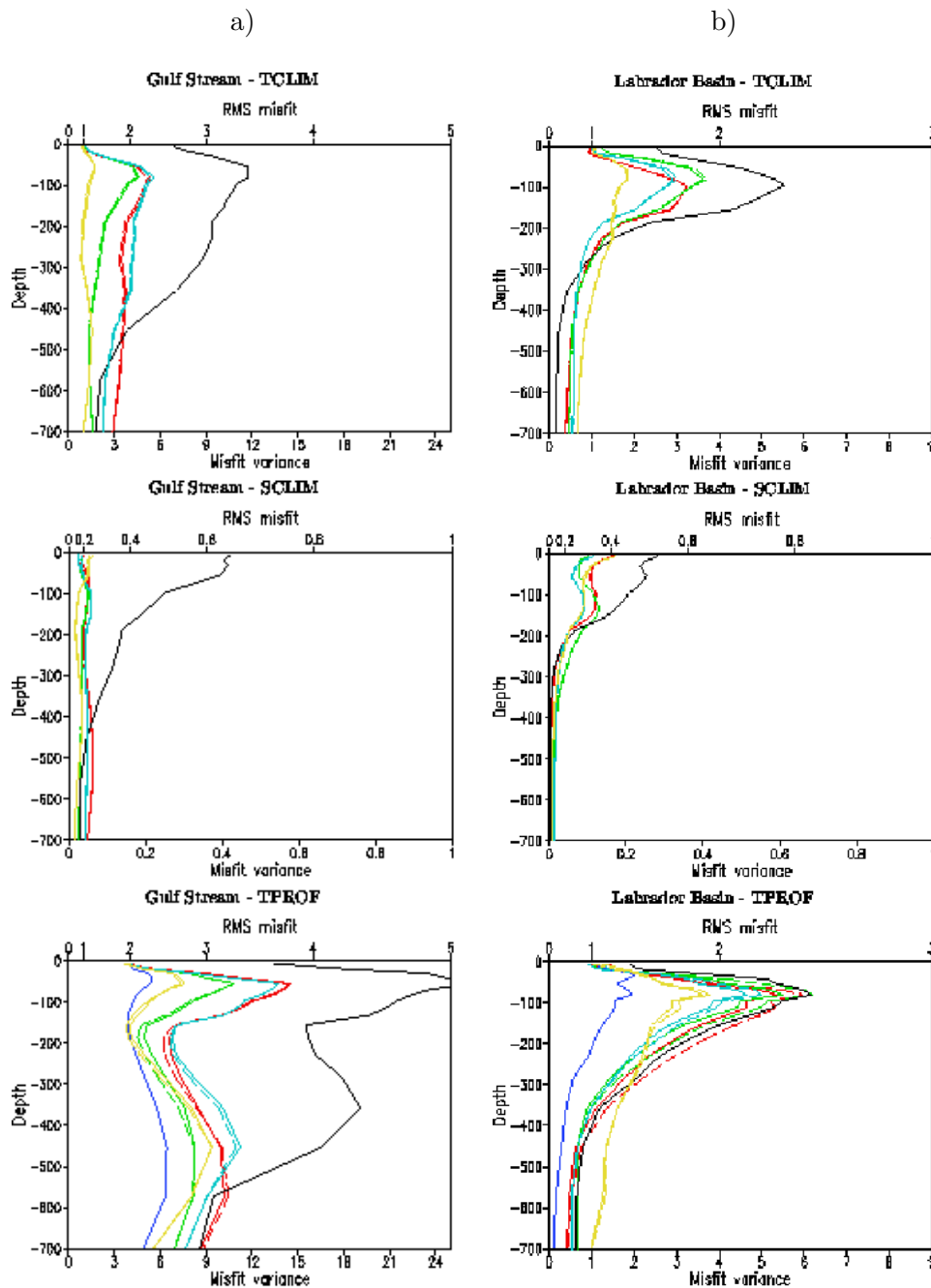


FIG. A.9 – Temperature ($^{\circ}\text{C}$) and salinity (PSU) RMS misfit with respect to TS climatology and to XBT profiles in (a) the Gulf Stream region and in (b) the Northeast region, down to 700 m and averaged for 1993. The figure shows plots of the free run (black curve), the EXPT1 (turquoise), EXPT2 (green), EXPT3 (red) and EXPT4 (yellow) solutions. For each assimilation experiment, the 3-day forecast is shown by the continuous-line curve and the analyses by the dashed-line curves.

EXPT3 are less marked, due to a much weaker vertical temperature gradient. In this region, EXPT1 shows slightly better results.

A.6 Specific case of the assimilation of TS profiles

In a multivariate case, the Kalman gain generates corrections based on specified cross-field error covariances. Errors in surface fields are related to errors in subsurface. Consequently, when the system is constrained only by surface observations, inaccuracy in cross-field covariances will lead to inaccurate corrections in water mass properties at depth. Given our limited knowledge of model errors and the associated correlations, this is a significant limitation on the quality of ocean predictions. The assimilation of new observations, such as TS fields, should moderate this problem.

However, this also means the introduction of new information types in the system, that is, of heterogeneous variables. The question we wish to address here is the problem of consistency between the sea level elevation corrections provided, on one hand, by the MDT information and, on the other, by the TS profiles. Indeed, there is no reason why these different data sources should contain exactly the same information about the $S0$ and $S1$ components defined previously in the section on problem definition. Note that this difficulty would not arise (1) if the model had not drifted from its initial climatological state, (2) if both fields were not associated with large errors, and (3) if the conventional sea 490 level reference was the same for the model and the data.

In order to overcome potential difficulties, we adopted the following strategy. A new $S0$ horizontal mean level was determined using in situ profile information, whereas the $S1$ component was prescribed using the MDT. This approach amounts to an estimation of the sea level topography in two steps. First, we assume no possible connection between errors in sea level elevation and errors in the subsurface fields. In practice, it means that we have dropped the cross-field correlation to zero in the forecast error covariance. Thus, the sea level elevation correction will be constrained only by sea surface observations. Second, modifications are introduced into the adjustment operator in order to take advantage of the TS data for the estimate of corrected sea level elevation. The density field is deduced from the analyzed vertical fields (after restoration of the hydrostatic assumption). The corresponding sea surface topography is calculated and compared to the analyzed sea surface topography prescribed by the filter. The horizontal mean difference between these two fields is then computed and added to the analyzed sea surface topography. In this way, we also expect to correct progressively the bias introduced in the initial state by model drift. This approach was used to perform EXPT4.

In order to test this method, we show in Figure A.10 the temporal evolution of mean sea surface heights averaged over the whole model domain corresponding to EXPT0, EXPT2, and EXPT4. The free run simulation drifts slowly away from its initial state with time because the model thermal content is increasing during the numerical experiment. Using surface observations only, the SEEK filter eliminates this model drift by controlling the mean sea level trajectory from the initial assimilation date (EXPT2 curve). The behavior of the EXPT4 curve is different. After a first stage in which mean sea surface height decreases, it stabilizes and starts reproducing the same seasonal variability as the EXPT2 curve : a new $S0$ value based on the TS information has been determined. It means that, with the proposed method, we are able to use climatological subsurface observations

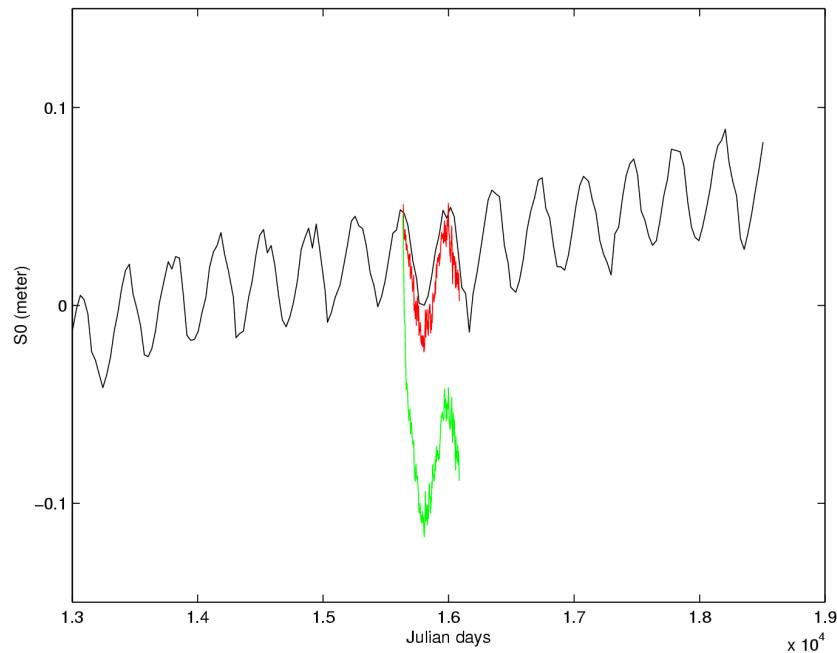


FIG. A.10 – *Time evolution of the sea surface elevation averaged over the model domain for EXPT0 (black), EXPT2 (red) and EXPT4 (green).*

correctly to reduce the error in the initial mean state and then to improve significantly the model trajectory.

An analysis is also conducted on the representation of water mass properties in EXPT4. A comparison of Figure A.7f with the other figures (particularly Figure A.7d, where the same MDT was used) shows that in EXPT4 temperature has generally been brought into better agreement with the climatology in the ocean interior. In particular, the 18°C mode water properties have been successfully modified. In mid-latitudes, the estimated temperature exhibits some small-scale colder features of less than 1°C around 500 m depth. This might be due to a limitation of the system in producing accurate water mass property corrections using a limited number of error modes (10), when the number of degrees of freedom for the problem is much greater. Finally, Figure A.9 shows the resulting RMS misfit between the XBTs or climatology and the equivalent model state, also illustrating the positive impact of the assimilation of TS profiles.

These results illustrate that the corrections to the depth of the model isopycnic layers provided by the assimilation of TS profiles are accurately estimated. A specific TS profile assimilation experiment was performed without using this protocol (not shown), and the results suffered from significant errors in the estimation of depths for water mass correction.

A.7 Summary and conclusions

The main objective of this work was to assess the sensitivity of simulated ocean state estimates to different MDT products in a multivariate assimilation context. Ongoing and future satellite missions should provide accurate measurements of the gravity field over the ocean, but it will still take several years before precise estimates of the absolute dynamic

topography are produced. In the near future, to compute this quantity we will still have to rely on estimates obtained using the different existing approaches.

In this study, we first presented three different experiments, assimilating SSH, SST, and SSS observations into a hybrid vertical coordinate model of the North and Tropical Atlantic Ocean. In each experiment, a different MDT product was used to assimilate satellite sea level anomalies into the model. Important differences between the simulations were observed in terms of ocean circulation, transport, and thermohaline fields, which can be related to the differences between the three MDT estimates. The diagnostics demonstrate that the assimilation system is sensitive to the choice of the MDT, not only in terms of surface currents, but also in terms of deep properties. It remains difficult to determine which MDT estimate is the most appropriate for assimilation. With regard to the different diagnostics presented, and particularly to the systematic comparison with in situ TS profiles, the results vary from one oceanic region to another. In conclusion, errors in MDT estimates are shown to be among the most important factors affecting simulation quality.

An additional experiment was performed to evaluate the joint assimilation of dynamic topography and TS profiles. According to the protocol proposed, the MDT information was used to control the horizontal mean sea level pattern, while TS profiles were used to correct progressively the horizontal averaged sea level elevation. Assimilation of TS profiles proved efficient in correcting the significant error in the initial model state (largely due to model drift) and in making the vertical structure of the ocean more realistic. Finally, the study also underlines the necessity of ensuring consistency between MDT and in situ profiles.

Acknowledgments. This work was conducted within the framework of the TOPAZ project (Grant EVK3-2000-00601). We would like to thank Eric Chassignet (RSMAS) for providing the model configuration and the MICOM mean sea surface, Fabrice Hernandez and Eric Greiner (CLS) for MDT1 data, Pascal Legrand (LPO/IFREMER) for MDT2 data and Catherine Maillard (SYSMER/IFREMER) for the XBT database.

Annexe B

Une paramétrisation de l'espace d'erreur en rang réduit basée sur un ensemble de simulations perturbées

B.1 Positionnement du problème

Les flux atmosphériques représentent une source d'erreur importante, si ce n'est la plus importante, dans les OGCMs actuels. C'est la raison pour laquelle il serait nécessaire pour les systèmes d'assimilation de les contrôler, comme cela a d'ailleurs été souligné par Bill Large (*Large, 2006*) lors de l'école d'été GODAE de La Londe Les Maures. Bien que ceci dépasse largement le cadre de cette thèse, ce constat nous a conduit à envisager de nouvelles méthodes pour la paramétrisation de l'espace d'erreur en rang réduit.

Bien que la tension du vent à la surface de l'océan soit fondamentale pour la modélisation de la circulation océanique, elle reste relativement mal connue. De nombreuses études ont décrit et comparé des analyses de l'intensité des vents à la surface (*Bentamy et al., 1996; Freilich et Dunbar, 1999; Smith et al., 2001*). Au vu de ces articles, il est clair qu'il existe des différences marquées entre les différents produits de forçage en vent. Ces disparités se traduisent notamment par des différences substantielles dans les simulations forcées de la dynamique océanique (*Busalacchi et al., 1990; Fu et Chao, 1997; Grima et al., 1999; Hackert et al., 2001*). Comme je l'ai dit dans le chapitre traitant de la dynamique du Pacifique Tropical (chapitre 1), la variabilité océanique dans cette région est gouvernée, de manière forte, par la variabilité de la tension du vent à la surface de l'océan, en particulier la composante zonale de cette variabilité. Les erreurs sur le forçage en vent sont ainsi reconnues comme étant une, si ce n'est la principale source d'erreur des OGCMs dans cette région (*Wittenberg, 2003*). Dans ce qui suit, nous allons donc partir de l'hypothèse que l'erreur modèle est dominé par l'erreur sur le forçage en vent. Afin d'estimer la covariance de l'erreur modèle, nous allons utiliser une méthode d'ensemble; l'ensemble étant obtenu en perturbant la tension du vent $\vec{\tau}$ à l'interface océan/atmosphère.

B.2 La paramétrisation retenue

Afin d'estimer au mieux les modes d'erreur du modèle que l'on cherche à corriger avec l'assimilation de données, nous allons utiliser un ensemble pour paramétrer l'espace d'erreur en rang réduit du filtre SEEK. Comme je l'ai dit ci-dessus, on part de l'hypothèse que la principale source d'erreur de notre modèle est l'erreur sur la tension du vent $\vec{\tau}$ utilisé pour forcer notre modèle. La tension du vent s'écrit :

$$\vec{\tau} = \rho_a C_D \|\vec{u}_{10}\| \vec{u}_{10} \quad (\text{B.1})$$

où ρ_a est la masse volumique de l'air, C_D le coefficient de traînée et \vec{u}_{10} la vitesse du vent à 10 mètres. Si la vitesse du vent à 10 mètres \vec{u}_{10} est aujourd'hui connue avec un relativement bonne précision (*Freilich et Dunbar, 1999*), il n'y a pas de consensus sur la façon d'estimer le coefficient empirique de traînée C_D . Ce dernier représente une source d'erreur importante pour le tension de vent $\vec{\tau}$ (*Large, 2006*). Nous allons donc perturber ce coefficient C_D de manière à simuler des erreurs sur la tension de vent et ainsi générer notre ensemble.

Afin d'identifier les directions principales d'erreur sur le coefficient empirique de traînée C_D , nous avons utilisé une décomposition en fonctions empiriques orthogonales (EOFs) de la norme de la tension du vent $\|\vec{\tau}\|$:

$$\|\vec{\tau}(x, y, t)\| = \sum_{k=1}^N a_k(t) \mathfrak{M}_k(x, y) \quad (\text{B.2})$$

Ceci revient à faire l'hypothèse que les erreurs sur la tension du vent (c'est-à-dire certainement sur le coefficient de traînée C_D) apparaissent préférentiellement suivant les modes de variabilité de cette même tension de vent. La figure B.1 montre la première EOF de la tension de vent ainsi que son signal temporel associé. La première EOF traduit bien l'intensité des alizés de Nord-Est et de Sud-Est avec un décalage de 6 mois entre les deux hémisphères. Les alizés de Nord-Est s'intensifient pendant l'hiver boréal et contribuent à faire descendre l'ITCZ près de l'équateur. En été, ce sont les alizés de Sud-Est qui dominent et l'ITCZ remonte au voisinage de $8^\circ N$. Au Sud, nous voyons la signature de la variabilité de la position de la SPCZ qui est déphasée de 6 mois par rapport à l'ITCZ (cf chapitre 1).

En utilisant les 30 premiers modes de variabilité de la tension du vent $\mathfrak{M}_k(x, y)$, on génère un ensemble de 30 forçages perturbés :

$$\vec{\tau}_k = \rho_a C_D [1 + \alpha \mathfrak{M}_k(x, y)] \|\vec{u}_{10}\| \vec{u}_{10} \quad (\text{B.3})$$

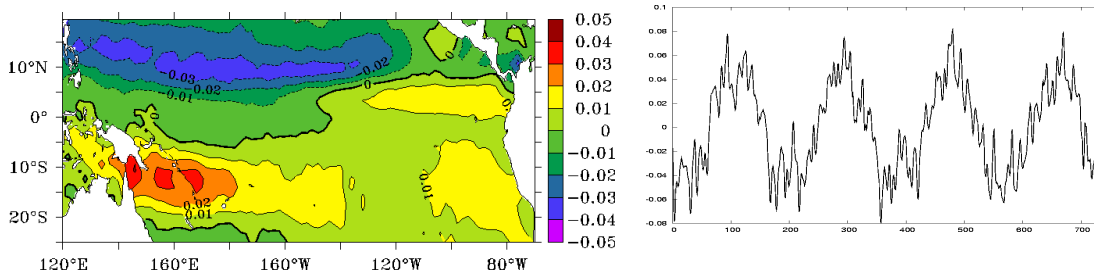


FIG. B.1 – EOF 1 de la variabilité de la tension du vent τ à la surface océan/atmosphère pour la période 1993-1996. A gauche : le signal temporel associé.

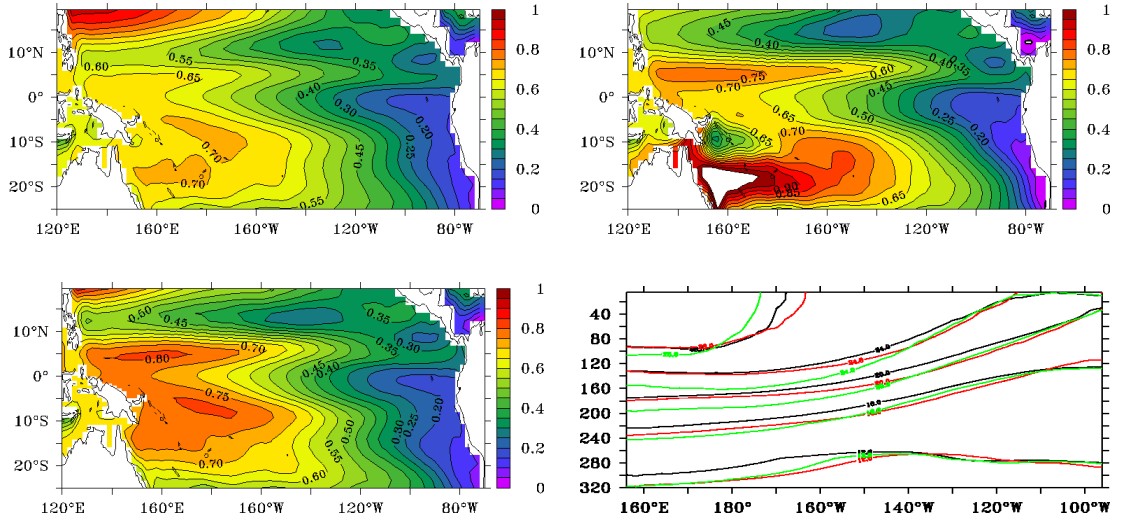


FIG. B.2 – MDT pour les trois premières simulations perturbées suivant les trois premiers modes de variabilité de la tension de vent $\vec{\tau}$ (respectivement de gauche à droite et de haut en bas) et la section de température moyenne à l'équateur pour les trois simulations (isothermes en noire pour la première simulation, rouge pour la deuxième et verte pour la troisième).

soit

$$\tau_{x_k} = [1 + \alpha \mathcal{M}_k(x, y)] \tau_x \quad (\text{B.4})$$

$$\tau_{y_k} = [1 + \alpha \mathcal{M}_k(x, y)] \tau_y \quad (\text{B.5})$$

où α est un coefficient permettant d'ajuster l'intensité de la perturbation. Dans notre cas, nous avons conservé α constant pour l'ensemble des 30 perturbations. On obtient donc des perturbations plus intenses selon les directions des premiers modes de variabilité.

Ce jeu de 30 forçages perturbés est utilisé pour forcer notre modèle. On obtient ainsi un ensemble de simulations conduisant à des structurations différentes de la circulation océanique. Comme on peut le voir sur la figure B.2 qui montre les MDTs des trois premières simulations de l'ensemble, ainsi que les sections moyennes de température à l'équateur, la réponse du modèle à la perturbation de la tension de vent $\vec{\tau}$ produit des différences substantielles sur la structure de la circulation des simulations perturbées, confirmant nos hypothèses de départ sur la sensibilité du modèle au forçage de tension de vent et sur l'importance des erreurs sur la tension de vent dans l'erreur modèle.

La matrice de covariance d'erreur initiale \mathbf{P}_0 de rang r est obtenue moyennant les hypothèses suivantes : (i) la covariance de l'ensemble peut être utilisée pour approcher la covariance d'erreur initiale, (ii) l'échantillonnage d'une série d'états instantanés de l'ensemble permet d'estimer la variance de l'ensemble, et (iii) la variance de cette ensemble peut être approchée par les r premières EOFs d'une décomposition EOF de l'ensemble. La figure B.3 montre le premier mode d'erreur 3D multivarié obtenu avec cette deuxième paramétrisation.

Un premier test qui me semble intéressant avec ce type d'approche est de vérifier qu'ajouter à la MDT du modèle libre, une combinaison linéaire des anomalies des différentes topographies dynamiques moyennes issues de l'ensemble de simulations perturbées, permet

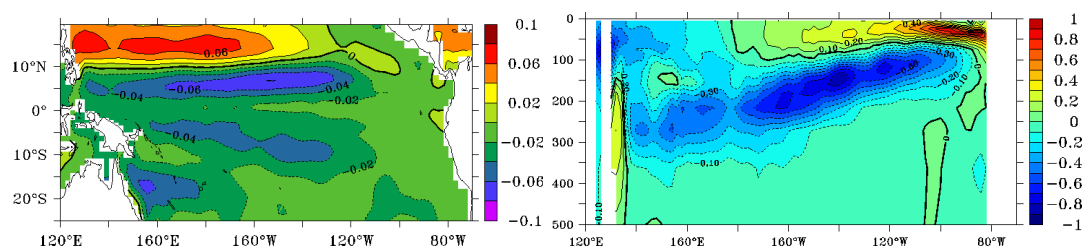


FIG. B.3 – Premier mode 3D multivarié de la covariance d'erreur initiale \mathbf{P}_0 pour la SSH (à gauche en m), la température à l'équateur (à droite en $^{\circ}\text{C}$).

de reconstruire une MDT proche de la MDT satellite observée. La figure B.4 montre le résultat d'une telle combinaison ainsi que la carte des erreurs résultantes avec la MDT observée. La différence RMS diminue ainsi de 4.3 cm pour la simulation libre non perturbée à environ 3.1 cm pour la combinaison linéaire des MDTs de l'ensemble. Ce premier résultat, même s'il ne garantit pas l'efficacité de cette paramétrisation est plutôt encourageant pour la suite. Un deuxième test intéressant mais que je n'ai pas eu le temps de mettre en place, aurait été de réutiliser les coefficients obtenus pour la combinaison linéaire permettant de minimiser l'erreur par rapport à la MDT observée afin de les utiliser pour reconstruire une tension de vents qui soit la combinaison linéaire des tensions de vents perturbées, pondérées par ces mêmes coefficients. Une simulation en mode libre forcée en utilisant cette tension de vent aurait sûrement été un test instructif. En effet, si le modèle répond de manière assez linéaire à une modification de la tension de vent, ce qui est vraisemblable autant que j'ai pu le constater, nous pourrions ainsi obtenir une solution en mode libre présentant une surface moyenne bien plus proche de la MDT GRACE.

B.3 Discussion

Afin de tester la paramétrisation de l'espace d'erreur en rang réduit décrite ci-dessus, nous avons réalisé une expérience d'un an sur l'année 1993. Dans cette expérience, l'altimétrie référencée avec la topographie dynamique moyenne GRACE et les profils de température in-situ TAO sont assimilés. Je n'ai malheureusement pas eu le temps d'analyser cette expérience de façon sérieuse. Ce que l'on peut dire après un rapide coup d'œil sur les sorties, c'est que les résultats sont mitigés. Le filtre SEEK initialisé de la sorte parvient à réduire très (trop ?) fortement les écarts par rapports aux données assimilées mais ceci au détriment de la partie non observée du vecteur d'état. Le niveau d'erreur par rapport aux données assimilées est très proche des erreurs d'observations spécifiés avec une erreur moyenne de 5.4 cm RMS par rapport la topographie dynamique et une erreur de 0.51°C RMS par rapport aux TAO. En revanche, si l'on évalue les résultats avec les données XBTs indépendantes, on observe une réduction de l'erreur dans les régions contrôlées par les TAO et une augmentation de l'erreur hors du guide d'onde équatorial échantillonné par les données TAO. Cela suggère que le filtre SEEK ne dispose pas d'assez de données avec l'altimétrie seule pour estimer les directions d'erreurs à privilégier. Le spectre d'erreur obtenu en utilisant un ensemble de simulation perturbée est certainement plus large que celui obtenu en utilisant des anomalies par rapport à une simulation de référence. La quantité d'observations nécessaires pour identifier les bonnes corrections est donc également supérieure. L'utilisation d'anomalies par rapport à une simulation de référence permet en

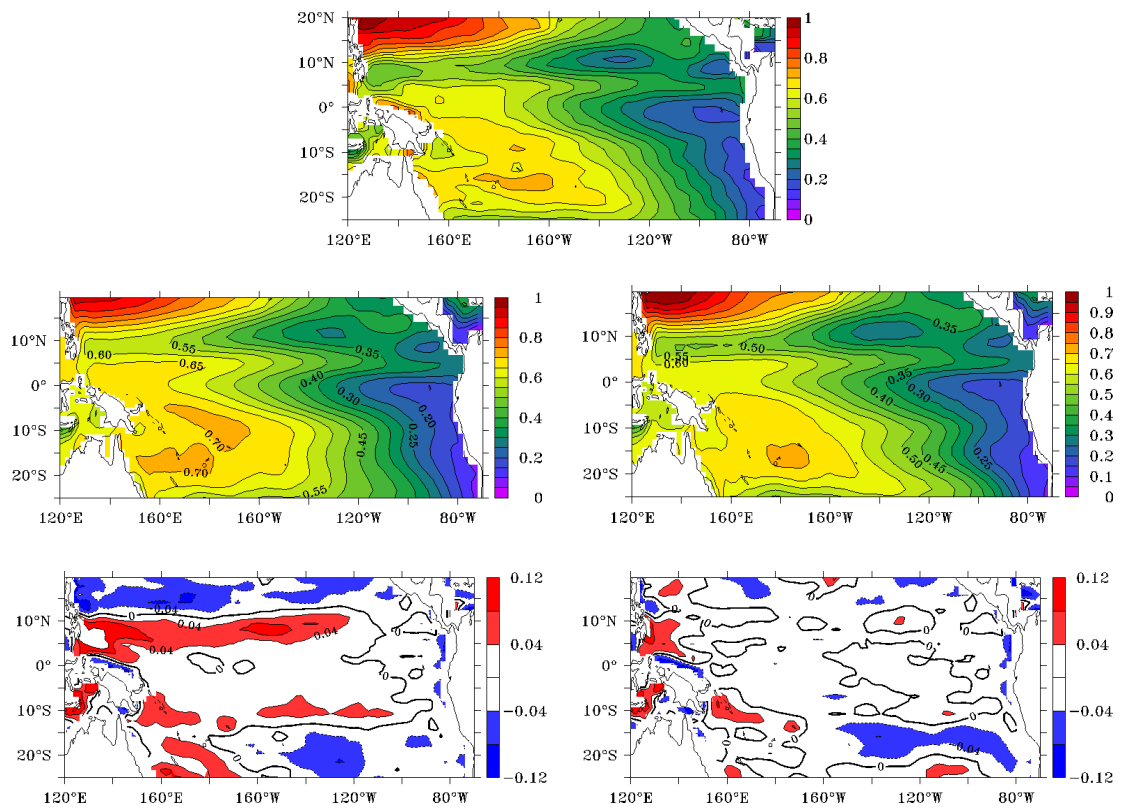


FIG. B.4 – MDT GRACE (en haut), ainsi que la MDT de la simulation non perturbée (au milieu à gauche) et la combinaison linéaire des MDTs issues de l'ensemble (au milieu à droite). En bas, carte des différences avec la MDT GRACE pour, respectivement, la MDT de la simulation non perturbée (à gauche) et la combinaison linéaire des MDTs.

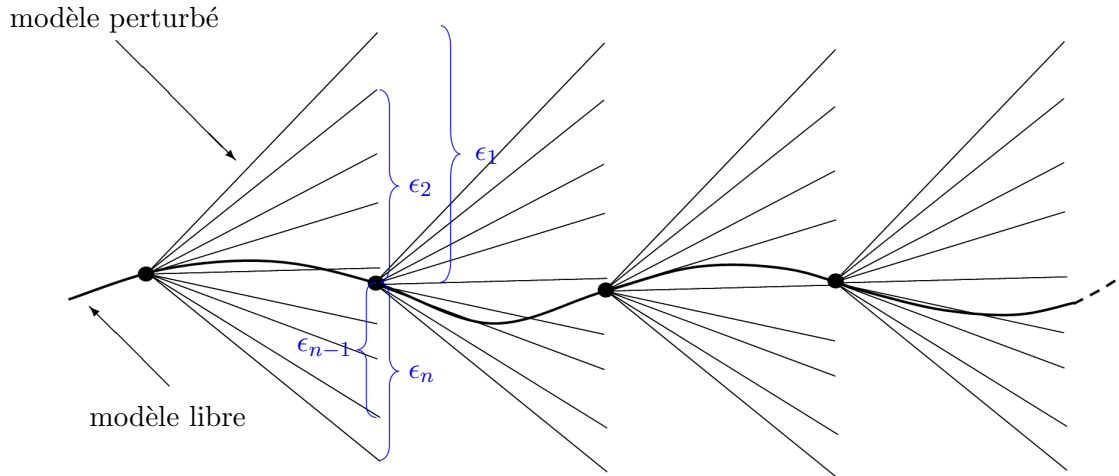


FIG. B.5 – Schéma de principe.

quelque sorte de limiter a priori les directions d'erreurs. Mais dans ce contexte, l'efficacité de la paramétrisation dépend très fortement de la validité des hypothèses faites.

Malgré les relatives mauvaises performances obtenues avec la paramétrisation détaillée ci-dessus, je reste confiant sur le fait que cette approche est intéressante et nécessite de plus amples investigations. Je n'ai malheureusement pas eu le temps d'aller plus en avant dans la cadre de ce travail de thèse. Par exemple, une piste qui me semble prometteuse serait la suivante. En conservant l'idée que l'erreur sur la tension du vent est la principale source d'erreur du modèle dans la région du Pacifique Tropical, un ensemble de forçage de vents perturbée comme expliqué ci-dessus est généré (rien n'empêche cependant de considérer que les autres jeux de forçages représentent une source d'erreur non négligeable et de perturber d'autres flux, comme par exemple le flux de chaleur). Cette ensemble de modèle perturbé est utilisé pour générer un ensemble de prévision à l'échéance du cycle d'assimilation à partir d'un ensemble de conditions initiales échantillonnées à partir de la simulation libre. La figure B.5 illustre le principe d'une telle méthode. Un ensemble d'anomalies $\epsilon_1, \epsilon_2, \dots, \epsilon_{n-1}, \epsilon_n$ est calculé pour chaque cycle en soustrayant de chaque prévision obtenue avec un modèle perturbé, la prévision obtenue en partant de la même condition initiale mais en utilisant la simulation libre nominale. La covariance de cette ensemble d'anomalies devrait fournir une estimation appropriée de la statistique de l'erreur de prévision du modèle. L'utilisation d'anomalie est une chose qui me paraît importante dans l'optique de contrôler une erreur modèle à l'aide d'une méthode d'assimilation statistique comme le SEEK. Cela permet de limiter les directions d'erreur aux seules erreurs apparaissant lors de l'intégration du modèle, c'est à dire précisément les erreurs que l'on essaye de corriger. La tâche du filtre SEEK lors de l'étape d'analyse s'en trouve facilitée et l'information contenue dans les observations est utilisée de manière plus efficace.

Une approche assez similaire a été utilisé par Sergey Skachko pour une étude menée dans le cadre du projet européen MERSEA sur le contrôle des flux océan-atmosphère par assimilation de données, et plus particulièrement des coefficients empiriques des formules aéro-dynamiques bulk régissant les échanges à l'interface océan-atmosphère dans le modèle (*Skachko et al., 2007*). Il est basé sur la configuration modèle-assimilation développée dans le cadre de cette thèse et décrite dans les chapitres 3 et 4 et se propose d'explorer la

possibilité de contrôler les flux océan-atmosphère par assimilation de données dans un cadre d'expériences jumelles. Le principe est d'utiliser un vecteur de contrôle de l'assimilation augmenté incluant les coefficients empiriques des formules aéro-dynamiques bulk que l'on veut contrôler, ici les coefficients C_E et C_H (cf équations de transfert océan-atmosphère 3.32 et 3.32 gouvernant les flux turbulents de chaleur sensible et latente du modèle). La paramétrisation de l'espace d'erreur en rang réduit du filtre SEEK est réalisée en utilisant une méthode similaire à celle introduite ci-dessus. La seule différence est qu'au lieu de perturber le coefficient de traînée C_D qui régit la tension du vent, ce sont les coefficients C_E et C_H qui sont perturbés afin d'obtenir des perturbations sur les flux de chaleur sensible et latente du modèle. Les résultats obtenus sont très encourageants et renforce l'idée que ce type de paramétrisation est intéressante et mérite de plus amples investigations. Cette étude a fait l'objet d'une publication, dont je suis co-auteur. Le papier a été soumis dans le journal Remote Sensing of Environment (annexe C).

Annexe C

A sequential data assimilation approach for estimating turbulent air-sea flux bulk parameters

S. Skachko,⁽¹⁾ J. M. Brankart⁽¹⁾, F. Castruccio⁽¹⁾, P. Brasseur⁽¹⁾, and J. Verron⁽¹⁾

⁽¹⁾Laboratoire des Ecoulements Géophysiques et Industriels, Grenoble, France

Submitted to Remote Sensing of Environment

Abstract

Bulk formulations parameterizing turbulent air-sea fluxes remain one of the main source of error in present ocean models. The objective of this study is to investigate the possibility of estimating the turbulent bulk exchange coefficients using sequential data assimilation. The method involves increasing the control vector of the assimilation scheme using the model parameters that are to be controlled. In this paper, we focus on estimating two bulk coefficients that drive the sensible heat flux, the latent heat flux and the evaporation flux of a global ocean model, by assimilating temperature and salinity profiles using a horizontal and temporal sampling similar to that to be provided by the ARGO float system. The results of twin experiments show that the method is able to correctly estimate the large-scale variations in the bulk parameters, leading to a significant improvement in the atmospheric forcing applied to the ocean model. Consequently, the method can be used directly by existing ocean operational systems to produce more realistic forecasts of the thermohaline characteristics of the mixed layer.

C.1 Introduction

Turbulent momentum, heat and fresh water fluxes at the air-sea interface (usually computed using bulk formulations) remain one of the main sources of error in present ocean models (*Large, 2006*). They strongly limit the capacity of such models to provide a realistic forecast of the thermohaline characteristics of the mixed layer and of the surface ocean currents. This is the reason why we decided to investigate ways of improving the knowledge of these fluxes through assimilation of oceanic observations.

A natural way to address this problem would be to use a 4D-VAR variational assimilation scheme, by including the fluxes in the control parameters of the model in addition to the initial condition (*Roquet et al., 1993; Stammer et al., 2004*). However, such a procedure is hard to set up and computationally expensive. In addition, many existing ocean operational systems are based on much simpler assimilation schemes, that would greatly benefit from more precise estimates of the fluxes. This is why it is useful to look for alternative schemes that are numerically less expensive and easy to implement in a wide range of existing systems. Our starting point is a reduced-order Kalman filter in which the background error covariance matrix is represented by a set of 3D error modes in the state space of the ocean model. In this paper, we will use a simplified version of the SEEK filter (*Pham et al., 1998*), in which the background error covariance matrix is kept unchanged from one assimilation cycle to the next. However, the method that we propose using to correct the fluxes is directly applicable in any similar assimilation scheme, such as ensemble optimal interpolation schemes, ensemble Kalman filters (*Evensen, 2003*), or any variant of reduced-order Kalman filters : the RRSQRT scheme (*Heemink et al., 2001*), the ESSE system (*Lermusiaux et Robinson, 1999*), the SEEK filter (*Pham et al., 1998*), etc...

The general method used to identify model parameters (like the air-sea fluxes) by means of a Kalman filter is to augment the filter control space by including these parameters in addition to the state variables. This technique, although very common in the engineering literature (*Cox, 1964; Ho et Whalen, 1963; Nelson et Stear, 1976; Ljung, 1979*), has not often been used in atmospheric or oceanographic applications of the Kalman filter (*Aksoy et al., 2006*). If the purpose of the Kalman filter is to control the air-sea fluxes, the control parameters could be the fluxes themselves or the atmospheric fields from which they are computed. But it seems preferable to include a few key parameters of the bulk formula in the control vector, (i) because they are more likely to persist over time (the aim is to improve the forecast), (ii) because they are likely to be as easy to control by ocean observations (provided that we only include parameters that are linearly linked to the value of the flux), and (iii) because they can be assumed (*Large, 2006*) to be the real source of error (even this will probably result in compensating errors on the atmospheric parameters by correcting the bulk coefficients).

In this paper, we present an example in which only the latent heat flux coefficient (C_E) and the sensible heat flux coefficient (C_H) are included in the control vector. The procedure is tested using twin assimilation experiments with a low resolution ($2^\circ \times 2^\circ$) global ocean configuration. In this context, assimilation experiments, with sequential corrections of the bulk coefficients, can be performed if we provide a background error covariance matrix in the augmented control space. This will be done using ensemble simulations characterized by various values of C_E and C_H .

The objective of the paper is to demonstrate the possibility of estimating turbulent

air-sea flux bulk coefficients by inverting oceanic observations using a sequential assimilation scheme. The synthetic observations that will be assimilated to correct the fluxes are temperature and salinity profiles with horizontal and temporal sampling comparable to that of the ARGO floats system that is expected in the near future. In particular, we will study the ability to reconstitute the bulk coefficients by assimilating these kinds of oceanic observations, and examine how correcting the parameters can help improve the quality of the air-sea fluxes and thereby the quality of temperature and salinity forecasts. Investigations will also be conducted into the behaviour of the method in the presence of systematic errors arising from inaccurate parameters.

The choice of the synthetic observation system used in this paper is obviously just an example. The estimation of the air-sea flux bulk parameters mainly depends on the availability of surface observations of temperature and salinity. Both can also be derived from remote sensing information. Thus, the method and results presented in this paper are all directly transferable to the assimilation of remote sensing observations of sea surface temperature (SST) and sea surface salinity (SSS), and can also be viewed as a demonstration of the usefulness of these data to improve the estimation of the air-sea fluxes.

The structure of the paper is as follows : section 2 and 3 present the ocean model and the assimilation methodology ; section 4 describes the experimental protocol ; section 5 presents the results of the experiments.

C.2 The ocean model

The ocean model used to perform this study is the ORCA2 configuration of the OPA model (*Madec et al.*, 1998). This is a global ocean configuration using a $2^\circ \times 2^\circ$ ORCA type horizontal grid, with a meridional grid spacing reduced to $1/2^\circ$ in the tropical regions in order to improve the representation of the equatorial dynamics. This is a free surface configuration based on the resolution of ocean dynamic primitive equations, with a z-coordinate vertical discretization. There are 31 levels along the vertical, and the vertical resolution varies from 10 m in the first 120 m to 500 m at the bottom. The model uses a turbulent kinetic energy (TKE) closure scheme to evaluate the vertical mixing of momentum and tracers, while the lateral mixing for active tracers (temperature and salinity) is parameterized along isopycnal surfaces.

The model is forced at the surface boundary with heat, freshwater and momentum fluxes. The momentum flux is derived from the ERS scatterometer wind stresses complemented by TAO derived stresses (*Menkes et al.*, 1998). The heat and fresh water turbulent fluxes are computed using bulk formulation from NCEP atmospheric interannual data provided by the NOAA-CIRES ESRL/PSD Climate Diagnostics branch, Boulder, Colorado, USA. In this model, a common way of dealing with errors in the atmospheric forcing is to add a local restoring term to relax the model solution towards SST and SSS climatological data. Since the objective of this work is to develop a method to correct for errors in the air-sea fluxes, none of these restoring terms have been applied in the experiments described below.

In this paper, the focus will be on the latent heat flux coefficient (C_E) and on the sensible heat flux coefficient (C_H), both of which are involved in the computation of the following quantities :

– the latent heat flux :

$$Q_L = \rho_a L C_E W \max(0, q_s - q_a) \quad (\text{C.1})$$

where L is the vaporization latent heat, q_s and q_a are the surface and atmospheric specific humidities ;

– the evaporation fresh water flux :

$$E = Q_L / L \quad (\text{C.2})$$

– the sensible heat flux :

$$Q_S = \rho_a C C_H W (T_w - T_a) \quad (\text{C.3})$$

where ρ_a is the air density, C the air specific heat, W the wind speed, T_w and T_a the air and sea surface temperature ;

In the standard OPA parameterization, C_E and C_H are the product of very complex parameterization depending, in particular, on the stability of the air column close to the sea surface. This complex parameterization will not be described in this paper since it is only used to produce the reference model simulation (see section C.4.1). It is only important to note that the values of the C_E and C_H parameters depend on the current state of the system, so that the C_E and C_H parameters in the reference simulation are not constant in space and time.

C.3 The assimilation scheme

The assimilation method used in this study is derived from the SEEK filter, which is a reduced-order Kalman filter introduced by *Pham et al.* (1998). This sequential method has already been used in a number of studies (*Verron et al.*, 1999; *Brasseur et al.*, 1999; *Penduff et al.*, 2002; *Birol et al.*, 2004). In the present implementation, the background error covariance matrix is kept unchanged from one assimilation cycle to the next, transforming the filter into an optimal interpolation algorithm. The main feature of the scheme is the structure of the background error covariance matrix which is parameterized in a low-order subspace of the state space. This means that error is assumed to occur only in a few known directions of the state space. We also make use of the local variant of the SEEK filter described in *Brankart et al.* (2003) and *Testut et al.* (2003), so that the analysis for each water column of the model will depend only on the observations within a specified influence bubble. In this study, the size of each influence bubble is parameterized inside a box of 10×10 grid points.

In addition to this standard procedure, we also apply the incremental scheme implemented by *Ourmières et al.* (2006). Indeed, a significant drawback of sequential methods is the time discontinuity of the solution resulting from intermittent corrections of the model state. In order to tackle this problem, *Bloom et al.* (1996) have proposed an algorithm called Incremental Analysis Update (IAU). The principle of this algorithm is to incorporate the sequential observational update ($\delta \mathbf{x}$) in a more gradual way. At each assimilation cycle, the model is restarted from the initial condition of the current assimilation cycle with a forcing term added in the model equations :

$$\dot{\mathbf{x}} = \mathcal{M}(\mathbf{x}, \mathbf{p}^0) + \lambda(t) \delta \mathbf{x} \quad \text{with} \quad \int_0^{\Delta t} \lambda(t) dt = 1 \quad (\text{C.4})$$

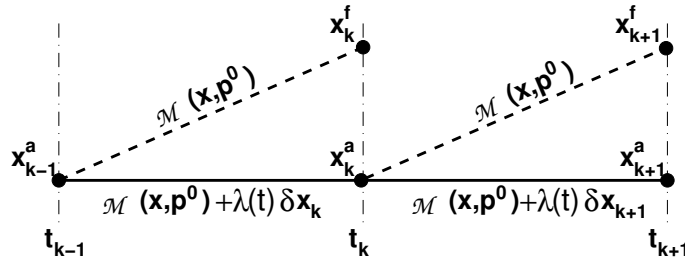


FIG. C.1 – Assimilation process of the original variant V_0 of the scheme. The dashed lines show the model forecast, and the solid line shows the analysis. The analysis is obtained by adding the observational update of time t_k , in the model operator between time t_{k-1} and t_k . The forecast is performed using the nominal values of the parameters \mathbf{p}^0 .

where \mathcal{M} is the model operator and Δt is the duration of the assimilation cycle. The state vector \mathbf{x} of our model contains four variable field : temperature (T), salinity (S), zonal and meridional velocity (U and V). The model operator \mathcal{M} also depends on a vector of parameters \mathbf{p} . Here, \mathbf{p}^0 denotes the nominal values of the parameters, without any correction through data assimilation. The procedure is illustrated in Figure C.1, where \mathbf{x}_k^f and \mathbf{x}_k^a denote respectively the forecast state vector and the analysis state vector at time t_k . \mathbf{x}_k^a is computed from \mathbf{x}_{k-1}^a by running the model \mathcal{M} as in equation C.4 with the observational increment computed at time t_k . This is the original algorithm that we use to start with, without any correction of the fluxes. It is the reference assimilation scheme, denoted as variant V_0 , with which new schemes will be compared.

In order to reach our objective of using such an assimilation scheme to correct bulk parameters, the control vector was augmented by the current values of the bulk parameters $\mathbf{p} = [C_E, C_H]$, in addition to the 4 prognostic variables of the model. The control vector is then written as $\hat{\mathbf{x}} = [T, S, U, V, C_E, C_H]$ instead of $\mathbf{x} = [T, S, U, V]$. Thus, from observations of temperature and salinity we will be able to make a correction for the bulk parameters. Obviously, the estimation will depend very much on the background error covariance between these 6 variables ; it is the pattern of the correlation between $[T, S]$ and $[C_E, C_H]$ that will govern the computation of a correction to $[C_E, C_H]$ from the diagnostics of the error in the $[T, S]$ fields. A method of building such a background error covariance matrix will be discussed in section C.4.2. Before that, however, we need to determine the best way to incorporate a correction on the model parameters into the model simulation. Three different approaches will be compared : in variant V_0 (described above), we compute corrections of the model parameters, but they are never used in the model. In variant 1, denoted as V_1 , we only correct the model parameters \mathbf{p} , and in variant 2, denoted as V_2 , we correct both the model parameters \mathbf{p} , and the model state \mathbf{x} . The remainder of this section gives the details of variants V_1 and V_2 .

In V_1 , we assume that the system can be controlled by adjusting only the bulk parameters, i.e. by correcting only the vector \mathbf{p} . The state vector is therefore no longer corrected directly (the IAU increment $\delta\mathbf{x}$ in equation C.4 is set to zero), and the correction is applied by restarting the model from the initial condition of the current assimilation cycle using only the corrected values of the parameters \mathbf{p} . This is illustrated on the Figure C.2, where \mathbf{p}_k^a denotes the estimated value of the parameter after the observational update of time t_k . This variant of the scheme differs from V_0 only by the fact that the application of the IAU

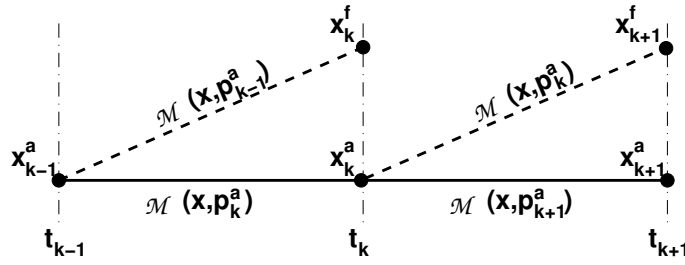


FIG. C.2 – Assimilation process of the variant V_1 of the scheme. The dashed lines show the model forecast, and the solid line show the analysis. The analysis is obtained by modifying the model operator between time t_{k-1} and t_k using the parameters \mathbf{p}_k^a estimated from the observations at time t_k . The forecast between time t_k and t_{k+1} is performed using the updated values of the parameters \mathbf{p}_k^a .

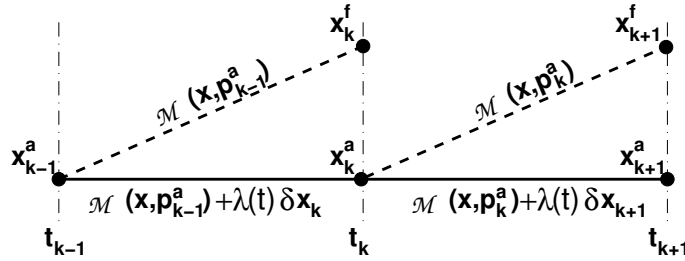


FIG. C.3 – Assimilation process of the variant V_2 of the scheme. The dashed lines show the model forecast, and the solid line show the analysis. The analysis is obtained by adding the observational update of time t_k , in the model operator between time t_{k-1} and t_k , without modifying the model parameters \mathbf{p}_{k-1}^a . The forecast between time t_k and t_{k+1} is performed using the updated values of the parameters \mathbf{p}_k^a .

increment ($\delta\mathbf{x}$) is replaced by a correction of the bulk coefficients, and that the upgraded values of the coefficients are used in the next assimilation cycle.

In V_2 , the state vector of the model \mathbf{x} is corrected as in the original scheme V_0 , but the new bulk parameters (\mathbf{p}) are taken into account in the next forecast. Again the observational update is incorporated gradually as in the original IAU scheme (Figure C.3). Since the corrected bulk parameters are used as background values for the next analysis, this variant of the scheme differs from V_0 only in that upgraded values of the coefficients are used in the next assimilation cycle.

Variant V_2 can be considered as an assimilation scheme written for the augmented state vector $\hat{\mathbf{x}}$. Up to now, we have tacitly admitted (as in most parameter estimation study using a Kalman filter) that the model equations are augmented by the equations $\dot{\mathbf{p}} = 0$, i.e. the parameter are assumed constant over time. However, in order to avoid unstable estimations of the bulk coefficients due to repeated observational updates without any feedback loop, we introduce a complementary model that governs the bulk parameters. This model computes the “forecast values” of the parameters \mathbf{p}_k^f from the analyzed values \mathbf{p}_k^a . This involves applying (i) a window mean filter to \mathbf{p} (with a window of 10 model grid points) and (ii) a relaxation toward the nominal value \mathbf{p}^0 of the parameters :

$$\mathbf{p}_k^f = \mathcal{F}(\mathbf{p}_k^a) + K [\mathbf{p}^0 - \mathcal{F}(\mathbf{p}_k^a)] \quad (\text{C.5})$$

where \mathbf{p}_k^f represents the “forecast values” of the parameters, $\mathcal{F}(\cdot)$ is the filter applied to the parameters and K is a relaxation constant. The values \mathbf{p}_k^f are used instead of \mathbf{p}_k^a in the next assimilation cycle as model parameters and as background values for the analysis. This procedure stabilizes the sequential estimation of the coefficients by preventing them drifting away from the reference value and by removing small-scale noise in the estimation. In all the experiments, the constant K is set to 0.4 for C_H (without that relaxation, the assimilation process gives unstable estimations for C_H) and to 0 for C_E (for which this proved unnecessary).

The conclusions of the paper are based on a comparison of the 3 variants of the assimilation scheme described in Figures C.1, C.2, C.3.

C.4 Design of the assimilation experiments

C.4.1 Experimental setup

The 3 variants of the assimilation scheme will be tested using twin experiments designed as follows. The reference simulation (the true ocean) is a standard ORCA2 simulation for the year 1993, with the original ORCA2 bulk formula. Figure C.4 illustrates the average value and standard deviation of C_E and C_H in this experiment. The figure shows that the average value of the coefficients is very inhomogeneous over the world ocean, varying mainly between 10^{-3} (in the equatorial band) and 2×10^{-3} (in the subpolar regions). The time variability of the coefficients is generally quite low, with standard deviation lower than 0.5×10^{-4} except in a few regions where it can be 4 to 10 times larger : south of the Antarctic Circumpolar Current, along the Western coasts of the North subpolar gyres, along the path of the Gulf Stream and of the Kuroshio, in the Eastern equatorial Pacific,...

Synthetic observations of temperature and salinity profiles are then sampled from this reference simulation to be assimilated in a perturbed simulation (the false ocean) in which the nominal values of the bulk coefficients $\mathbf{p}^0 = [C_E^0, C_H^0]$ are defined constant in space and time ($C_E^0 = 1.18 \times 10^{-3}$, $C_H^0 = 1.14 \times 10^{-3}$). These values agree with the range of bulk coefficients values determined in the most popular bulk parameterizations (e.g., *Friehe et al., 1976; Blanc, 1985; Smith, 1988; De Cosmo et al., 1996*). The initial condition of the false ocean is kept the same as the initial condition of the reference simulation, so that the only source of difference between the true ocean and the false ocean is model error due to C_E and C_H . Model error is indeed defined as the forecast error occurring if the model is started from perfect initial conditions.

From this, we can also anticipate that the model error in the false ocean due to inaccurate bulk parameters C_E and C_H will not usually be centred : the expected value of the model error (the model bias) is not zero. For instance, overestimating C_E or C_H will systematically lead to underestimate SST in the forecast. Such a situation is likely to occur very often in practice with real atmospheric forcings, so that our demonstration experiments must take this into account. In principle, Kalman filters are not designed to deal optimally with model errors that are not centred : in such a situation, a specific procedure must be added to estimate the model bias separately (*Dee et da Silva, 1998*). The way in which the method presented in this paper deals with systematic model error due to errors on the parameters will be discussed in section C.5.4.

In the assimilation experiments, we assume that we know the false ocean, the initial condition of the true ocean, and that we have temperature and salinity profiles observed

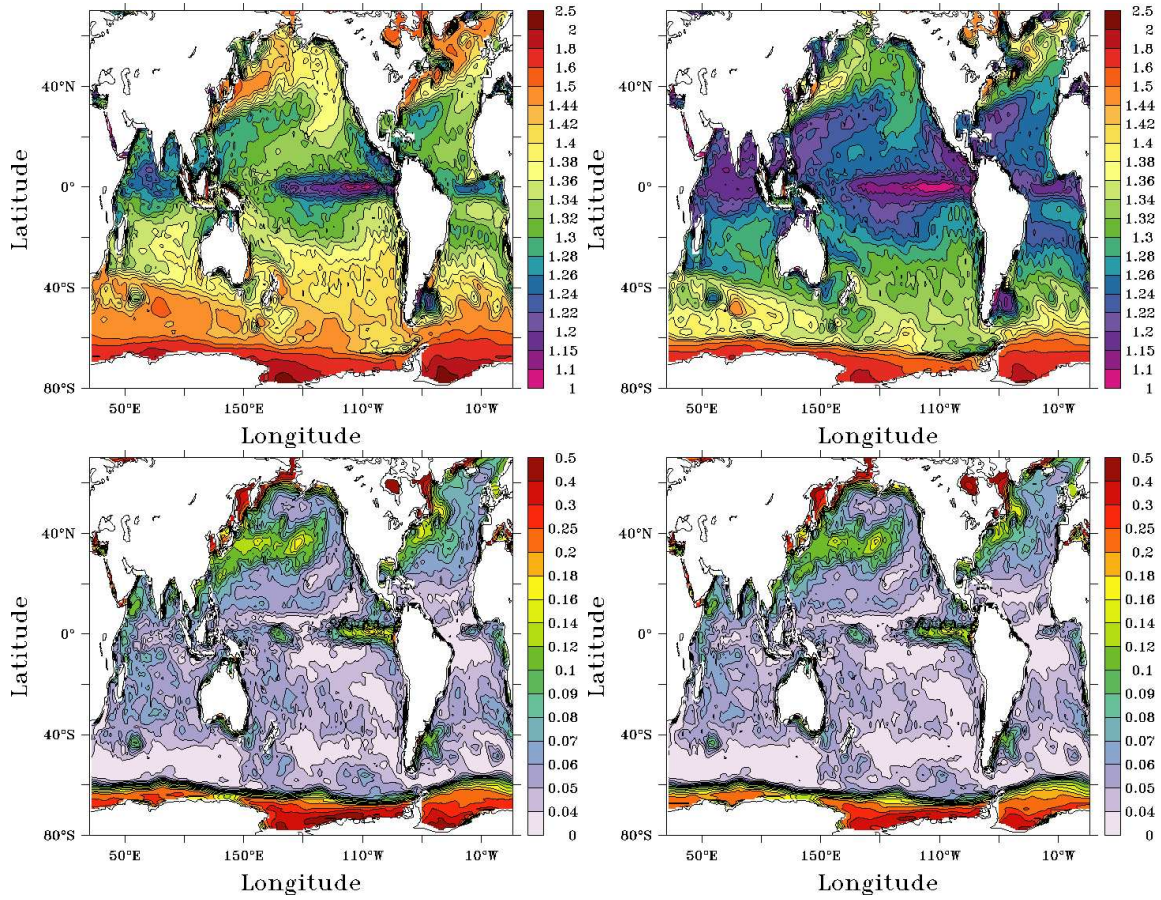


FIG. C.4 – Average values (top panels) and standard deviations (bottom panels) of the turbulent exchange coefficients C_E (left panels) and C_H (right panels) in the reference model (original ORCA2 bulk formulation). The values are scaled by a factor 10^3 .

from the true ocean every 10 days, down to a depth of 216 m (i.e., for the first 17 levels of the model), and every two nodes of the model horizontal grid. This sampling of observations is meant to roughly simulate the horizontal and temporal distribution of ARGO floats that can be expected in the near future. Note that in our experiments no observational update will be applied north of the latitude $75^\circ N$. These synthetic observations are perturbed by adding a Gaussian noise with standard deviations of 0.01 degrees and 0.002 psu for temperature and salinity respectively.

Our aim is to control the false ocean by shifting it towards the true ocean by assimilating these observations. We will therefore attempt to obtain an effective control not only of the state of the system, but also of the bulk coefficients or the air-sea heat and fresh water fluxes. The evaluation will be performed by computing errors on the model state vector and the parameters. In twin experiments, errors can indeed be directly computed as differences with respect to the reference simulation (i.e. the true ocean).

Since the experimental setup is designed in such a way that the only source of error in the system is model error due to the bulk parameters C_E and C_H , this provides us with an ideal framework to check the procedure we have designed (given that we have a priori knowledge that the origin of the error in the system can only be in the bulk

parameters). Thus, we simply have to demonstrate the possibility of tracing back the error on bulk parameters from observations of the state of the ocean. However, it can already be anticipated that this framework will only be that perfect at the beginning of the experiments. With time, imperfections in our corrections will result in initial condition error being present in the subsequent assimilation cycles. Thus, that kind of experimental setup will also allow us to evaluate the robustness of the scheme with regard to the presence of initial condition errors in the system.

Another important feature of this system is that the parameters included in the augmented state vector are not constant over time (see standard deviation in Figure C.4). Hence, this is more like a forcing function that will be estimated by the assimilation scheme. However, as stated in the introduction, we expect the bulk coefficients to be easier to estimate than the fluxes themselves, because their variations over time is relatively smaller and more gradual. For the same reasons, estimating the parameters is also expected to be more useful for improving forecasts. Nevertheless, the simple fact that they are not constant means that the model $\dot{\mathbf{p}} = 0$ for the parameters is inaccurate, there is still model error in the augmented model. If the parameters were constant, the augmented model would be free of model error because model error due to incorrect parameters (which is the only source of model error in our system) in the original system becomes initial error on the parameters in the augmented system. In such a case, the parameters would be much easier to estimate from the accumulation of observational information over time. However, parameters that vary over time are much more difficult to estimate. This difficulty partly explains why we modified the model $\dot{\mathbf{p}} = 0$ by using the more robust model of equation C.5. Based on these arguments, we may also anticipate that it will be more difficult to control the augmented system in those regions where the standard deviation of the parameters is greatest (see Figure C.4).

C.4.2 Parameterization of background error covariance

As explained in section C.3, the key element of the algorithm governing the inversion of temperature and salinity innovations used to compute correction for the bulk coefficients is the background error covariance matrix for the augmented state vector $\hat{\mathbf{x}}$. In order to generate this matrix, we built an ensemble of ocean models, characterized by different values of C_E and C_H (that are constant in space and time for each model). For every member of the ensemble, C_E and C_H values are obtained as random numbers sampled from a normal distribution with mean values $C_E^0 = 1.18 \times 10^{-3}$, $C_H^0 = 1.14 \times 10^{-3}$ equal to the nominal values of the parameters (corresponding to the false ocean) and a variance $\sigma = 0.15 \times 10^{-3}$. The dispersion of the parameters in this ensemble is chosen so as to be consistent with a priori information on parameter error covariance (see figure C.4).

This ensemble of models is used to perform a set of 10-day forecasts from a series of initial conditions distributed at 10-day intervals over the year 1993 (the period of our assimilation experiments). The 10-day interval is chosen to correspond to the frequency of observational updates in the assimilation experiments, and that series of initial conditions is used so that the covariance of the ensemble will be representative of the full period of the experiments. More specifically, a set of 5 model forecasts is performed for each of the 37 initial conditions, using different random C_E and C_H values for each of these 185 simulations. An ensemble of 185 forecast anomalies (augmented with C_E and C_H anomalies) is then obtained by subtracting from each forecast the forecast obtained from

the same initial condition, using the nominal values of the parameters.

The background error covariance matrix is then parameterized using the 40 first empirical orthogonal functions (EOFs) of this ensemble. This covariance matrix is a consistent estimator of the 10-day model error covariance, and is adequate for parameterizing the filter background error covariance, assuming that the initial error covariance remains small, cycle after cycle.

However, it should be pointed out that this ensemble parameterization is totally inadequate for representing the horizontal error correlation patterns. Indeed, using horizontally constant C_E and C_H coefficients for every member of the ensemble means that we assume that the error on the coefficients is always uniform worldwide, or similarly that the horizontal autocorrelation function of the coefficient error is equal to 1 for any couple of geographical location. Hence, using such a background error covariance matrix we would only be able to make a uniform correction to the C_E and C_H coefficients. In order to obtain a more realistic parameterization, we would need to generate an ensemble using a distribution of C_E and C_H which varies spatially. This ensemble should describe a horizontal correlation pattern for C_E and C_H that would need to be known a priori. However, it would require so a huge ensemble of worldwide fields of C_E and C_H to correctly represent such complex information, that the method would become impossible to use in practice.

In order to overcome this difficulty, we decided to rely only on the local SEEK parameterization to constrain the horizontal correlation pattern. Indeed, reducing the weight of an observation inside each influence bubble as a function of the distance from the centre of the bubble is similar to imposing a spatial correlation pattern on the background error covariance matrix. Roughly speaking, the ensemble is mainly defined to locally parameterize multivariate and vertical correlations, while horizontal correlations are driven essentially by the particular structure of the local gain algorithm. In this way, we will be able to obtain local corrections for the C_E and C_H coefficients even from an ensemble of simulations with constant coefficients worldwide.

C.5 Comparison of the 3 variants of the assimilation scheme

According to the experimental setup described in the previous section, three assimilation experiments were performed using the three different variants of the scheme presented in section C.3 and summarized in the Figures C.1, C.2, C.3.

C.5.1 Time evolution of temperature and salinity errors

In this section, the results of V_0 will be compared to those obtained using the new algorithms (V_1 and V_2), including surface flux correction.

Figure C.5 shows the evolution of error variance on SST and SSS over the world ocean (except the Northern polar zones). The black solid line shows the error corresponding to the free simulation, without data assimilation (i.e. the false ocean). The green, yellow, blue and red lines correspond to V_0 , V_0^* , V_1 and V_2 respectively. (The yellow curve shows the results for the variant V_0^* , that will be discussed in section C.5.4.) The solid lines represent the continuous analysis (see figures C.1, C.2 and C.3) and the bullets represent the 10-day model forecasts. Figure C.6 presents the same information for the Atlantic ocean only. (The error on the free simulation is not shown in figure C.6 because it exhibits the same kind of behaviour as in figure C.5.)

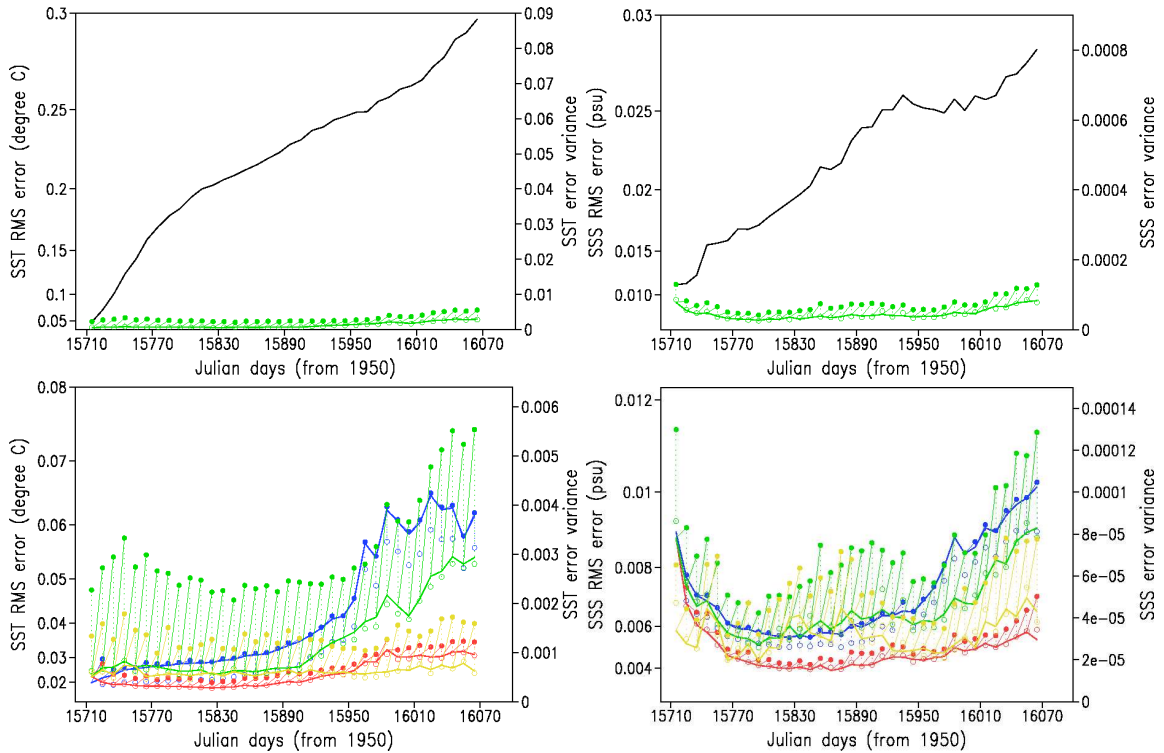


FIG. C.5 – Evolution of error variance on SST and SSS for the world ocean (except the Northern polar zones). The black solid line shows the error corresponding to the free simulation (without assimilation). The green, yellow, blue and red lines correspond to V_0 , V_0^* , V_1 and V_2 respectively. The solid lines represent the continuous analysis, while the bullets represent the 10-day model forecasts.

First, we observe that V_0 is already able to control the model error due to inaccurate bulk parameters : the large SST and SSS drift of the false ocean with respect to the true ocean is easily cancelled by the assimilation of temperature and salinity profiles. The analysis is stable over time on both variables, and the error variances are generally 3 to 6 times lower than the corresponding values of the free run. This means that the reduced space, in which the error covariance matrix is parameterized, has been correctly defined. Most of the temperature and salinity innovation has been captured by the reduced order analysis and interpreted as a correction on the full model state vector.

Second, it may be observed that the forecasts obtained from V_1 and V_2 are always better than those of V_0 . The poor quality of the V_0 forecast is easy to understand because it is always performed using the false model operator $\mathcal{M}(\mathbf{x}, \mathbf{p}^0)$ (i.e. without correction of the bulk parameters). The smaller error increase in the forecast using V_1 and V_2 means that the correction of the bulk parameters has a positive impact on the subsequent model forecast of SST and SSS. The model operators modified using the bulk parameters estimated by the assimilation scheme are thus more accurate than the false model operator used to forecast SST and SSS. A large part of the 10-day model error standard deviation on SST and SSS has been directly eliminated by correcting the model parameters.

Third, the analyses and forecasts obtained using V_2 rapidly become better than those obtained using V_0 . This is an indirect consequence of the improved forecasts, as they

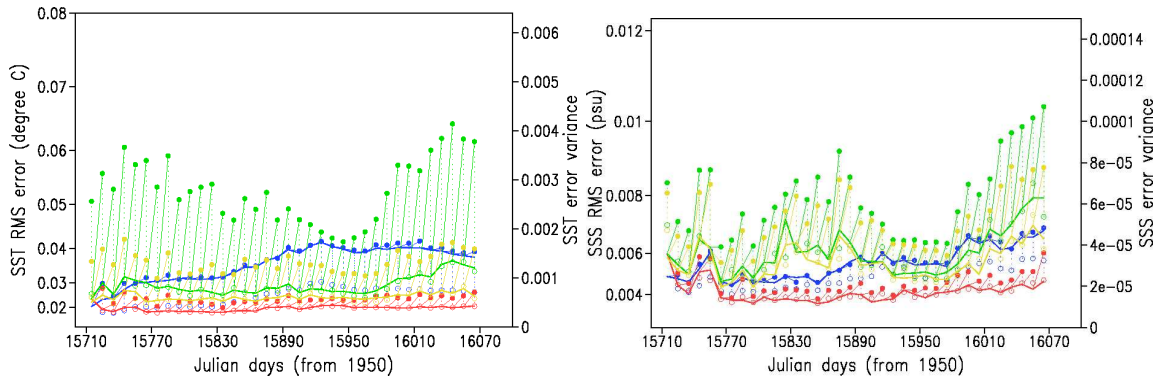


FIG. C.6 – Evolution of error variance on SST and SSS for the Atlantic ocean only. The green, yellow, blue and red lines correspond to V_0 , V_0^* , V_1 and V_2 respectively. The solid lines represent the continuous analysis and the bullets represent the 10-day model forecasts.

lead to smaller innovations that can be interpreted more easily by statistical analysis. A smaller model error on SST and SSS means a smaller error accumulating from one cycle to the next and a better analysis of the evolution of the system. V_1 behaves differently. Correcting the bulk parameters only is sufficient to stabilize the error standard deviation at values that are much smaller than those of the free simulation, although the results are not as accurate as those obtained with V_2 . Moreover, in figure C.5, the SST and SSS error variance of V_1 exhibits a large increase over the last 4 months of the experiments. This increase is localized in the Western Tropical Pacific ocean (see figure C.7, described in section C.5.2), which is why it is absent in figure C.6. The reason for these differences between V_1 and V_2 is that correcting only the parameters quickly becomes insufficient as soon as there are initial errors present at the beginning of the assimilation cycles. These initial errors accumulate from the slightly inaccurate corrections of the parameters in the previous assimilation cycles. In order to deal with these errors, state corrections in V_2 are necessary.

C.5.2 Distribution of temperature and salinity errors

Figure C.7 shows maps of standard deviation error on SST and SSS 10-day forecasts for the 3 variants of the scheme. The maps show the spatial distribution of the error instead of its evolution over time as in figure C.5. Again, the standard deviation error computed for V_0 is larger everywhere than the corresponding error in V_1 and V_2 . As expected, the largest remaining errors are found where the time variance of the parameters is the largest (compare to figure C.4) : in the Gulf Stream or Kuroshio regions, in the Confluence region, . . . Generally speaking, however, the standard deviation error on SST and SSS 10-day forecast is significantly reduced by the correction of the parameters especially in regions where the error was large in V_0 , making the resulting error maps more homogeneous over the world ocean.

Finally, in order to provide an idea of the vertical structure of the standard deviation error, on temperature and salinity 10-day forecasts, figure C.8 shows a meridional section across the Atlantic at 30°W longitude. The upper plots (for V_0) show the error on the 10-day forecast without corrections on the parameters. At this time scale (10 days), errors

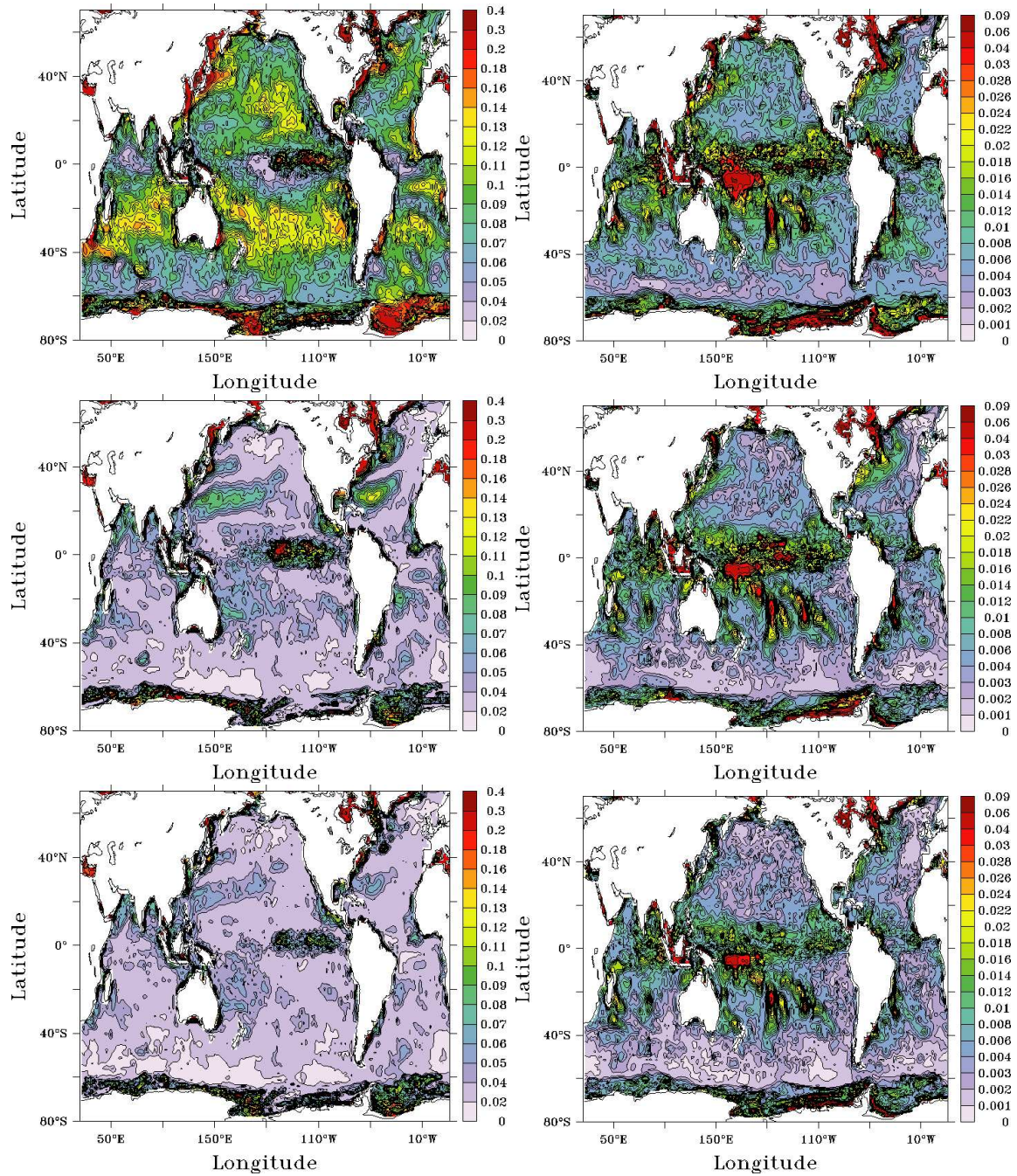


FIG. C.7 – Spatial distribution of standard deviation error on SST (left-hand column) and SSS (right-hand column) 10-day forecast for V_0 (upper plots), V_1 (middle plots) and V_2 (bottom plots) of the assimilation scheme.

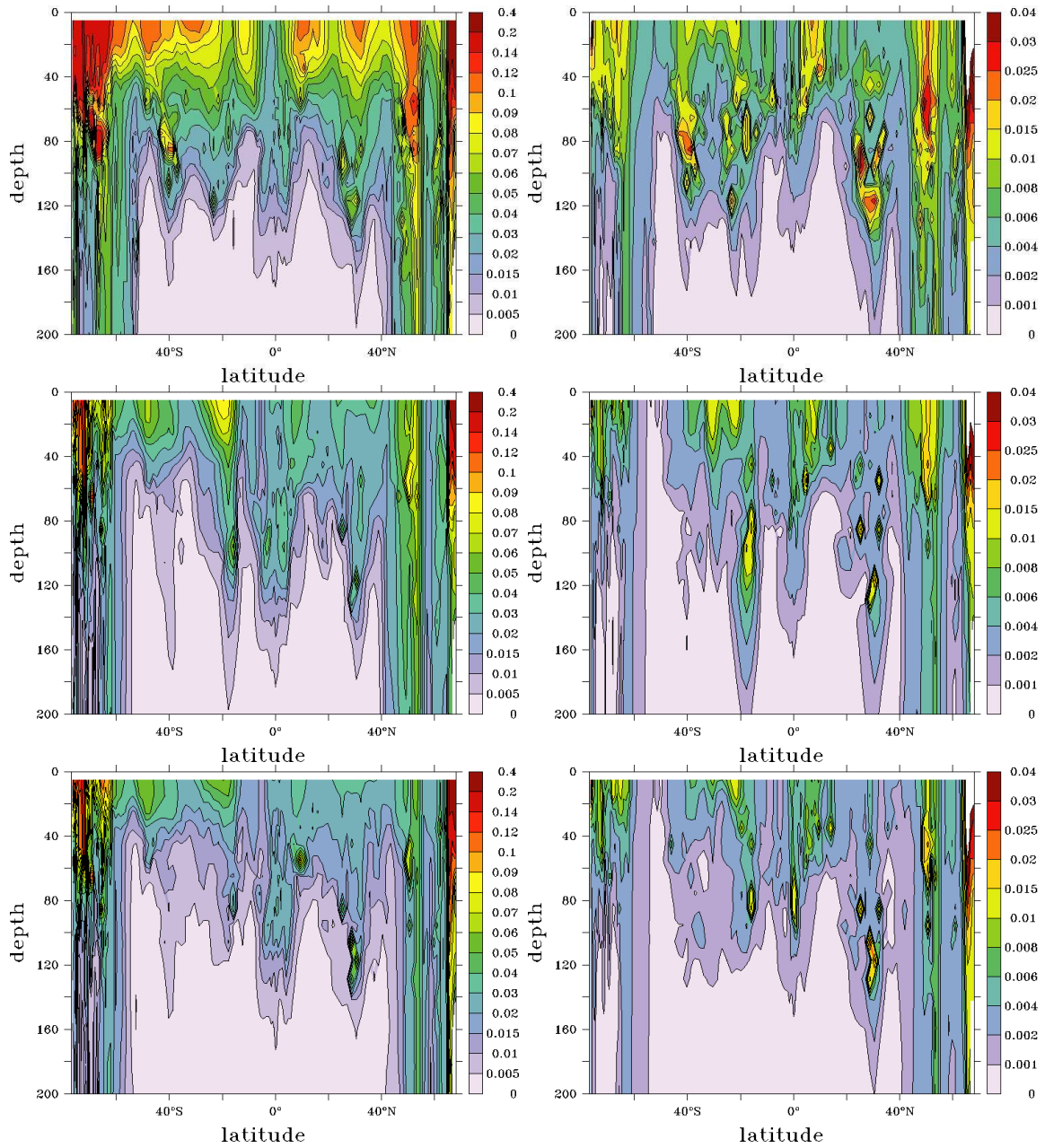


FIG. C.8 – Standard deviation error values on temperature (left-hand column) and salinity (right-hand column) 10-day forecast profiles for V_0 (upper plots), V_1 (middle plots) and V_2 (bottom plots) of the assimilation scheme

on the atmospheric forcing can only have a significant impact inside the ocean surface mixed layer. In this simulation, the largest errors occur in the North Atlantic and South Atlantic subpolar gyres. As the figure indicates, both parameter correction algorithm (V_1 and V_2) are able to significantly reduce the error with respect to V_0 . This error reduction is effective over the whole water column, meaning that a correct history of heat and salt vertical transfers is restored in the water column when the bulk parameters are controlled. The figure also reveals that V_2 produces better results than V_1 , although this does not prevent patches of residual errors remaining, especially in the subpolar gyres.

C.5.3 Quality of the parameter estimates

In the previous section, the 3 variants of the assimilation scheme were compared according to their effect on temperature and salinity errors. However, good results can sometimes be obtained for the wrong reasons. Therefore we have to analyse the method's capacity to produce reliable estimates of the model parameters. It is only if the parameters are improved by assimilation that the real source of error has been properly identified, and that the reduction in the error on temperature and salinity values has occurred for the right reasons.

Figure C.9 shows maps of standard deviation errors on the C_E and C_H coefficients (i.e. the root mean square difference with respect to the C_E and C_H values in the true ocean) for the 3 variants of the scheme. Remember that in V_0 , the coefficients used to produce the forecasts are the constant values defining the false ocean. It is thus the standard deviation error for these constant values that is presented in the figure. Clearly, the standard deviation error on the C_E coefficient calculated from V_0 is generally much larger than the corresponding error values for V_1 and V_2 . This is also true for the standard deviation error on the C_H coefficient except in the Western Tropical Pacific (for V_1 and V_2) and along the path of the Gulf Stream and Kuroshio (for V_1). As can be seen in these figures, there are large discrepancies between the constant parameters of the false model and the parameters of the reference simulation almost everywhere in the world ocean, the error values being especially large in the subpolar regions. This type of error is undoubtedly the reason of the poor quality of the forecasts on SST and SSS described in the previous section.

Variants V_1 and V_2 of our system demonstrate that they are able to correctly reproduce the spatial distribution of the true coefficients from their constant values in most regions of the world ocean for the entire period of our experiments. The overall quality of the bulk coefficients for these two methods is quite similar, though V_2 provides more reliable results for the C_H coefficient.

Note that the possibility of controlling these two coefficients through temperature and salinity observations could be anticipated. We chose *two* coefficients that *linearly* govern the value of two important contributions to the heat flux and one important contribution to the fresh water flux. These *two* fluxes also have a roughly *linear* impact (assuming that the mixed layer depth remains fairly constant during the 10-day forecast) on the surface value of *two* ocean variables : the temperature and the salinity. It may therefore be expected that, by observing these *two* variables, it would be possible to trace back a correction to the bulk coefficients.

Fortunately, these two coefficients are used to parameterize the turbulent fluxes at the air-sea interface, which are also those expected to be affected by the largest error.

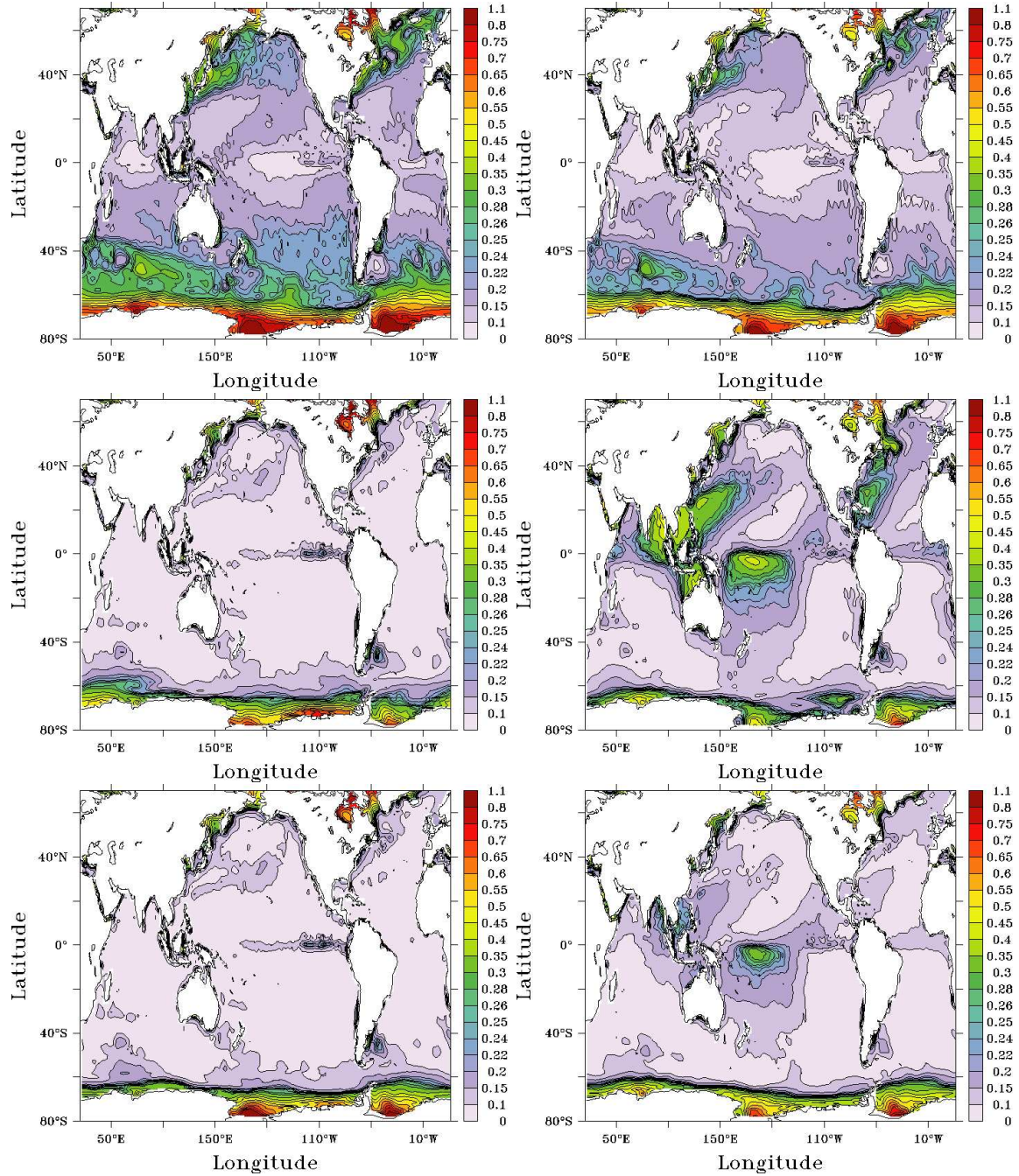


FIG. C.9 – Standard deviation error on the latent heat flux exchange coefficient C_E (left-hand column) and on the sensible heat flux exchange coefficient C_H (right-hand column), for V_0 (upper plots), V_1 (middle plots) and V_2 (bottom plots) of the assimilation scheme.

Postulating that the largest errors occur in the values of these coefficients is consequently not without justification.

C.5.4 Correction of systematic errors

The improvements of the parameters has been presented in the previous section, but we need to analyse in more detail how the system actually works. How do the parameters evolve over time compared to the true ocean? What are the statistics of temperature and salinity innovations, and the corresponding statistics of parameter corrections? Figure C.10 (first column) shows the evolution of the C_E parameter at two specific locations in the Atlantic ($2^\circ\text{W}16^\circ\text{S}$ and $60^\circ\text{W}36^\circ\text{N}$) for the true ocean, the false ocean, and for V_1 and V_2 . (Parameter values in V_0 are constant and equal to those of the false ocean.) The first thing that we can see in the figure is that the parameter value in the false ocean is very different from the averaged parameter value in the true ocean. This implies that, at this ocean location, the model error on the false ocean model $\dot{\mathbf{x}} = \mathcal{M}(\mathbf{x}, \mathbf{p}^0)$ is biased. This is also obvious in figure C.10 (second column) showing the surface temperature innovations at the same location. The average innovation in V_0 , using the model $\dot{\mathbf{x}} = \mathcal{M}(\mathbf{x}, \mathbf{p}^0)$ is not close to zero, but systematically smaller given that C_E is systematically smaller than in the true ocean. In V_0 , the model is biased due to a systematic error on the parameters \mathbf{p}^0 .

On the other hand, we can see that in V_1 and V_2 the average coefficient value is similar to that of the true ocean, the temperature innovations are centred, and the parameter corrections are also centred (see figure C.10). How can this new behaviour be interpreted? The reason for it is that variants V_1 and V_2 of the assimilation scheme work in the augmented state space, where the model $\dot{\mathbf{x}} = \mathcal{M}(\mathbf{x}, \mathbf{p}), \dot{\mathbf{p}} = \mathbf{0}$ is unbiased. We stated in section C.4.1 that model error remains in the model for the parameters $\dot{\mathbf{p}} = \mathbf{0}$ since the parameters are not constant, but it is easy to see that this model is unbiased. Indeed, unless there is a drift in the parameter value (which is unlikely for C_E and C_H), the distribution of model error for the model $\dot{\mathbf{p}} = \mathbf{0}$ is obviously centred. Thus, the augmented state vector methodology has transformed biased model error in the model state space into unbiased model error in the augmented state space, together with a (possibly large) initial error on the parameter value : $\mathbf{p}(t = 0) = \mathbf{p}^0$.

Thus, we have been able to show that the augmented state vector method is appropriate to deal with model biases that are due to inaccurate model parameters. However, we have to admit that, since the model $\dot{\mathbf{x}} = \mathcal{M}(\mathbf{x}, \mathbf{p})^0$ is biased, V_0 cannot be considered as an optimal data assimilation scheme. One could argue that a classical method for correcting the bias (for instance, the one proposed by *Dee et da Silva* (1998)) could have provided a better solution than V_0 , and that V_1 and V_2 should have been compared with such a solution. There are two possible answers to this question. First, most ocean assimilation systems drive models with biased atmospheric forcing, using a scheme of the V_0 kind. It is this kind of assimilation system that our method is intended to improve. Second, correcting the model parameters that are known to be the source of the bias is certainly more effective than using a general bias identification technique that would be much more difficult to implement and to parameterize. (Identifying the error on a few parameters is undoubtedly much easier than identifying a bias on the full state vector.)

Anyway, in order to illustrate this argument, the two new techniques V_1 and V_2 will now be compared to a modified V_0 scheme, denoted as V_0^* , in which the model bias is explicitly corrected. Since we are in twin experiments, a “perfect” value of the bias can be

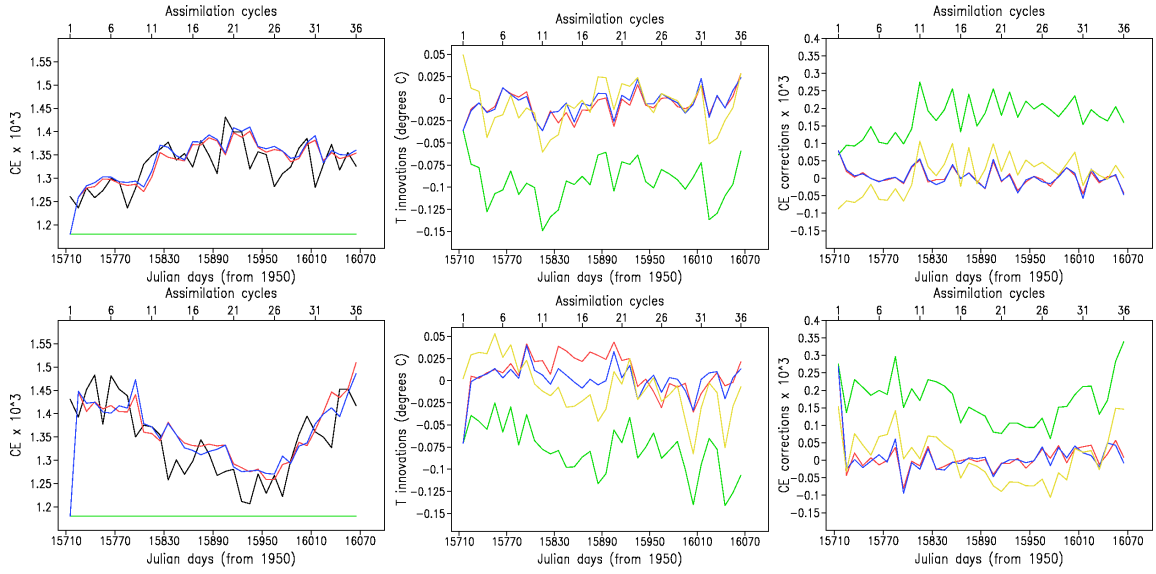


FIG. C.10 – Detailed statistics for two specific locations of the world ocean : $2^{\circ}\text{W}16^{\circ}\text{S}$ (top panels) and $60^{\circ}\text{W}36^{\circ}\text{N}$ (bottom panels) : C_E parameter evolution in time (left panels), temperature innovations (centre panels) and C_E parameter corrections (right panels). The black dashed line stands for the true ocean, and the green, yellow, blue and red lines correspond to V_0 , V_0^* , V_1 and V_2 respectively. (The green curve in the right panels displays the correction on the parameters that is computed in V_0 but never applied.)

estimated directly in the following way. We compute a series of 10-day forecast using the false (biased) model from a series of the true ocean states as initial conditions (distributed in time every 10 days in 1993). The 10-day forecast bias is then estimated as the time average of the differences between these 10-day forecasts and the true ocean. It is this “perfect” value of the 10-day forecast bias that is subtracted from each 10-day forecast in V_0^* , so that V_0^* is just V_0 with a “perfect” forecast bias correction in the model state space (better than any possible bias identification scheme). Despite of this, the results show that V_2 is most often preferable to V_0^* . This is illustrated in figures C.5 and C.6 showing that the error variance on SST and SSS 10-day forecast (with bias correction in V_0^*) is always twice larger in V_0^* than in V_2 . Figure C.10 (central column) shows the temperature innovations corresponding to V_0^* . Unlike V_0 , the innovations are centred in V_0^* because of the bias correction. The innovation values of V_0^* remain however usually larger than in V_1 and V_2 , consistently with figures C.5 and C.6.

In addition, as we can gather from figure C.10, the method is not only able to identify a correction on the average value of the parameters; the low-frequency variations of the parameters are also correctly captured by the assimilation scheme. For instance, at the first location ($2^{\circ}\text{W}16^{\circ}\text{S}$), the C_E coefficient increases in the reference simulation from 1.25×10^{-3} in January to 1.4×10^{-3} in July and then decreases to 1.35×10^{-3} between July and December. At the second location ($60^{\circ}\text{W}36^{\circ}\text{N}$), on the other hand, the C_E coefficient decreases in the reference simulation from 1.45×10^{-3} in January to 1.25×10^{-3} in August and then increases to 1.45×10^{-3} between August and December. We can see that these seasonal variations of the parameters are correctly reproduced by the assimilation of temperature and salinity data.

C.5.5 Quality of the flux estimates

Since SST and SSS are improved in V_1 and V_2 with respect to V_0 , and since the bulk parameters are also improved, we can almost be certain that the fluxes themselves computed using formula C.1 and C.3 will also be improved. This is illustrated in figure C.11 showing the spatial distributions of the standard deviation error values on the latent (upper plots) and sensible (lower plots) heat fluxes respectively for V_0 (left-hand column), V_1 (centre column) and V_2 (right-hand column). The evaporation flux is linked to the latent heat flux by equation C.2, so that the results for evaporation are obvious from the latent heat flux results. The fluxes errors corresponding to the false ocean and to V_0^* are not represented in the figure because they are similar to those of V_0 .

Figure C.11 also reveals that, except in the Western Equatorial Pacific for the sensible heat flux, the standard deviation error corresponding to V_0 is everywhere larger than the corresponding error in V_1 and V_2 . As expected, the largest errors occur where the parameter errors are the largest, and the greatest improvement in the fluxes also occurs in these regions. We can see that both new methods provide results of similar quality overall so that, using this criterion only, it is not easy to say which of the two variants is the best. The conclusion would not be the same in every region of the ocean.

C.5.6 Quality of the model forecasts

As stated previously, one objective of the correction of the model parameters is to improve the model forecast. The evaluation of the forecast that has been presented up to now has dealt with the 10-day forecast, i.e. the model forecast that provides the background state for the next observational update. The purpose of this section is to analyse the improvements in forecast quality over longer time scales. Figure C.12 shows the evolution in the error variance for a 3-months forecast initialized from the analysis on April 4, 1993, four months after the beginning of the assimilation experiments. The forecast is performed for each of the 3 variants of the scheme, using the analysis from the corresponding simulation as initial condition, and the parameters that would have been used to produce the 10-day forecast in that variant as the parameters (i.e. the false parameters for V_0 , and the analysed values of the parameters in V_1 and V_2). These experiments amount to extending the 10-day forecast of that assimilation cycle (the one beginning on April 4, 1993) to a 3 months forecast.

The results show that the error variances for V_0 , V_1 and V_2 increase monotonically during the entire period of the 3-month forecast. V_1 and V_2 provide a significantly better quality of model forecast, with error values one order of magnitude smaller than in V_0 . As expected, better parameters also improve the long term forecast, as a consequence of their relatively small and slow time variations. The same 3-month error evolution cannot be shown for V_0^* because the forecast bias has only been computed for 10-day forecast and not for other forecast periods. This illustrates another advantage of our technique over bias estimation methods working in the state space which make it not straightforward to estimate the forecast bias for several forecast period.

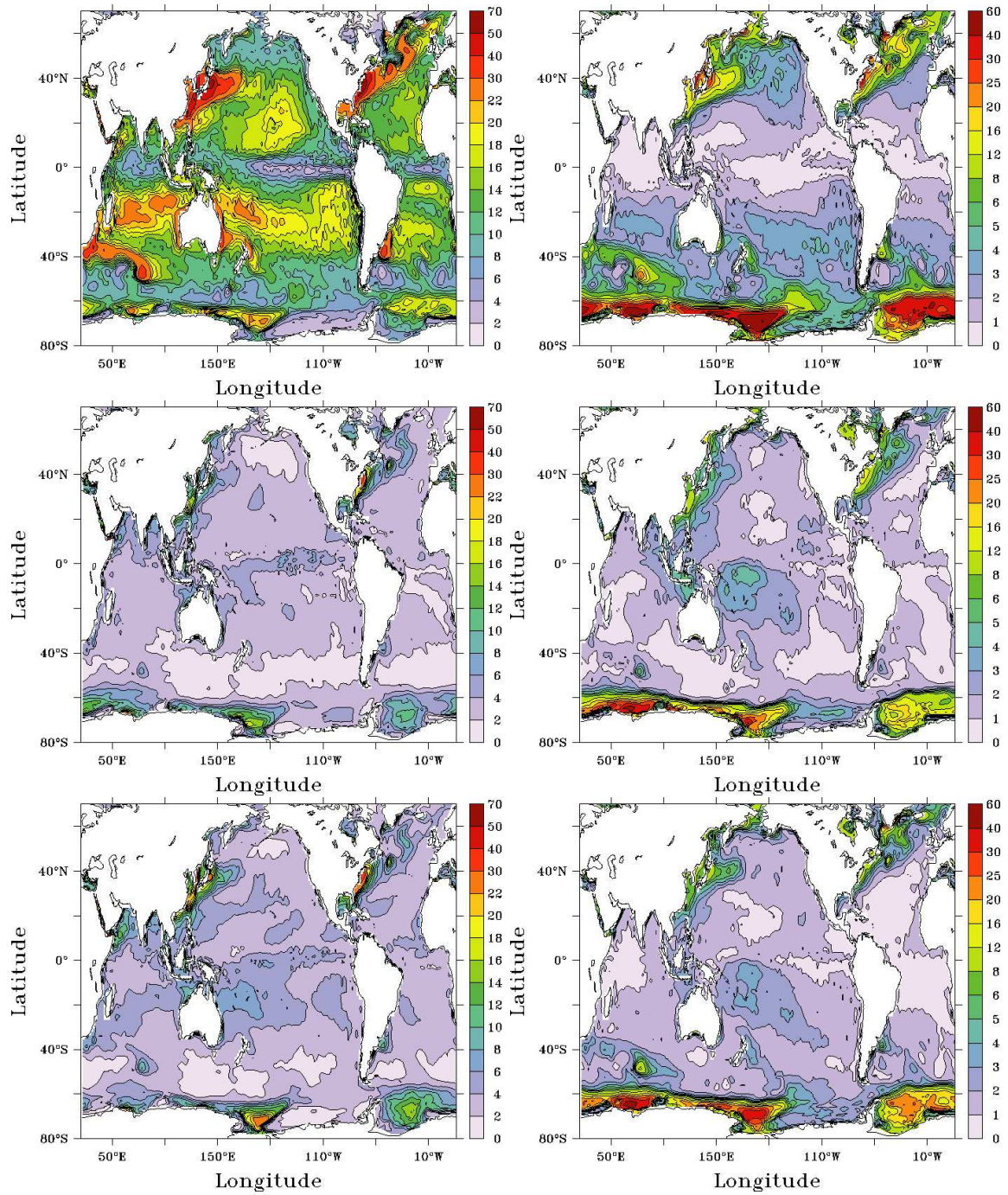


FIG. C.11 – Spatial distribution of standard deviation error values on the latent heat flux (left-hand column) and the sensible heat flux (right-hand column) for V_0 (upper plots), V_1 (middle plots) and V_2 (bottom plots).

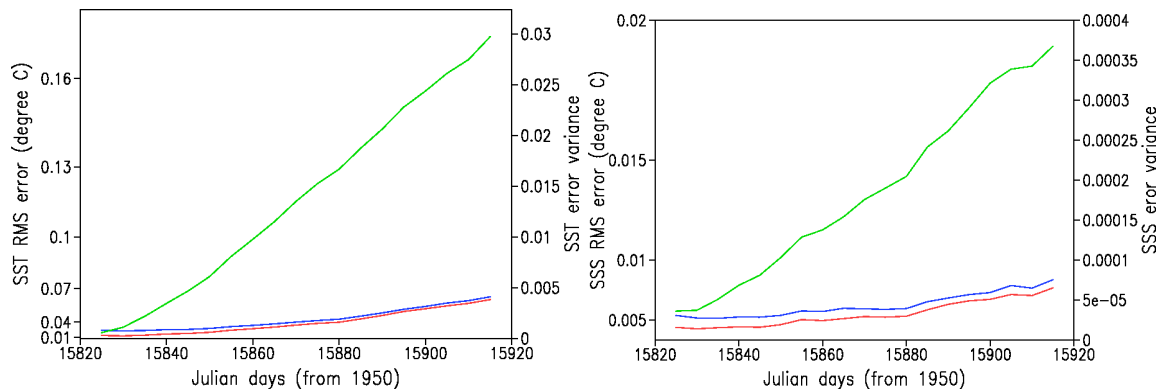


FIG. C.12 – Evolution of the error variance values for a 3 months forecast corresponding to V_0 , V_1 and V_2 of the assimilation scheme (green, blue and red curves respectively). The forecasts start on April 4, 1993 (4 months after the beginning of the experiments).

C.6 Conclusions

Two procedures for estimating turbulent air-sea flux bulk parameters by using Kalman filtering techniques have been proposed in this paper. Both are based on the augmentation of the control vector by parameters that must be estimated, with covariance matrices in the augmented control space computed using ensemble experiments. The first procedure (V_1) focusses on the correction of the model parameters, without any possibility of correcting other sources of errors. It should be useful for ocean modellers who want to improve their simulations corrupted by systematic errors in the air-sea fluxes, and can be viewed as an optimal flux correction procedure using ocean observations (like SST and SSS). It would be advantageous to use it to replace the classical Newtonian relaxation to surface temperature and salinity data, without introducing any non-physical term in the model equations. The second procedure (V_2) combines a correction of the model parameters with a correction of the full model state vector. This was designed as a complement to an existing assimilation scheme, thus making it possible to deal with other sources of errors, controlled by other kinds of observations.

In addition, errors in model parameters most often lead to a model bias. It was thus natural to study a problem in which a model bias is present. The conclusion is that the method is perfectly able to deal optimally with such model bias, transforming biased model error in the model state space into unbiased errors in the augmented state space. The procedure even gives better results than an ideal bias identification procedure working in the state space.

The results of twin experiments presented in the paper show that both new methods lead to accurate estimations of the parameters, and to a significant improvement in the analysis and forecasting of the mixed layer thermohaline characteristics. Better results are obtained if both the state vector and the parameters are corrected; correcting the parameters alone quickly becomes not sufficient as soon as initial errors, resulting from slightly inaccurate corrections in the previous assimilation cycles, are appearing in the system. These experiments were conducted with synthetic observations of temperature and salinity profiles that roughly simulate the horizontal and temporal distribution of ARGO floats expected in the near future. The positive results gives us a background

to future work in which this method will be applied to the more realistic problem of assimilating real data profiles.

Our experiments, however, are ideal in the sense that the only source of error is due to two bulk coefficients C_E and C_H : the atmosphere and the ocean initial condition were both perfectly known. To go further, several questions still need to be answered. How many bulk parameters is it possible to control using only oceanic observations ? What is the impact of initial error in the ocean state ? What perspectives may be provided by the complementarity of space and *in situ* data ? In other words, is it possible to control the full system, i.e. the ocean state and the flux parameters, with the available ocean observation system ?

Acknowledgments. This work was conducted as part of the MERSEA project funded by the E.U. (Contract No. AIP3-CT-2003-502885). The calculations were performed with the support of IDRIS/CNRS.

Bibliographie

- Aksoy, A., F. Zhang, et J. W. Nielsen-Gammon, Ensemble-based simultaneous state and parameter estimation with MM5, *Geophysical Research Letters*, *33*, 12,801, 2006.
- Alory, G., C. Cravatte, T. Izumo, et K. B. Rodgers, Validation of a decadal OGCM simulation for the tropical Pacific, *Ocean Modelling*, *10*, 272–282, 2005.
- Alves, O., M. A. Balmaseda, D. Anderson, et T. Stockdale, Sensitivity of dynamical seasonal forecasts to ocean initial conditions, *Quarterly Journal of the Royal Meteorological Society*, *130*, 647–667, 2004.
- Arakawa, A., et V. R. Lamb, Computational design of the basic dynamical processes of the UCLA general circulation model, *Methods in Comput. Phys.*, *17*, 173–265, 1977.
- Arnoux, D., et J. Blum, Back and forth nudging algorithm for data assimilation problems, *C. R. Acad. Sci. Paris, Ser. I*, 873–878, 2005.
- Asselin, R., Frequency filter for time integrations, *Monthly Weather Review*, *100*, 487–490, 1972.
- Ballabrera-Poy, J., P. Brasseur, et J. Verron, Dynamical evolution of the error statistics with the SEEK filter to assimilate altimetric data in eddy-resolving ocean models, . *J. R. Meteorol. Soc.*, *127*, 233–253, 2001.
- Barnier, B., *Forcing the Ocean*, vol. Ocean Modeling and Parameterization, E. P. Chassignet and J. Verron ed., Kluwer Academic Publishers, The Netherlands, 1998.
- Barnier, B., L. Siefridt, et P. Marchesiello, Thermal forcing for a global ocean circulation model using a three-year climatology of ECMWF analysis, *Journal of Marine Systems*, *6*, 363–380, 1995.
- Barnier, B., G. Madec, T. Penduff, J.-M. Molines, A.-M. Treguier, A. Beckmann, A. Bias-toch, C. Böning, J. Dengg, S. Gulev, J. Le Sommer, E. Remy, C. Talandier, S. Theetten, M. Maltrud, et A. Coward, Impact of partial steps and momentum advection schemes in a global ocean circulation model at eddy-permitting resolution, *Ocean Dynamics*, *56*, 543–567, 2006.
- Beckmann, A., et R. Döscher, A method for improved representation of dense water spreading over topography in geopotential-coordinate models, *Journal of Physical Oceanography*, *27*, 581–591, 1997.

- Bentamy, A., Y. Quilfen, F. Gohin, N. Grima, M. Lenaour, et J. Servain, Determination and validation of average wind field from ERS-1 scatterometer measurements, *Global Atmos. Ocean Syst.*, *4*, 1–29, 1996.
- Berliand, M., et T. Berliand, Determining the net long-wave radiation of the earth with consideration of the effect of cloudiness, *Isv. Akad. Nauk. SSSR Ser. Geophys.*, 1952.
- Berline, L., Assimilation de données dans un modèle couplé physique-biogéochimie de l'océan Atlantique Nord, Ph.D. thesis, Université Joseph Fourier - Grenoble I, 2005.
- Berline, L., C. E. Testut, P. Brasseur, et J. Verron, Variability of the Gulf Stream position and transport between 1992 and 1999 : a reanalysis based on a data assimilation experiment, *Journal of Remote Sensing*, *27*, 417–432, 2003.
- Biancale, R., G. Balmino, J.-M. Lemoine, J.-C. Marty, B. Moynot, F. Barlier, P. Exertier, O. Laurain, P. Gegout, P. Schwintzer, C. Reigber, A. Bode, R. König, F.-H. Massmann, J.-C. Raimondo, R. Schmidt, et S. Zhu, A new global Earth's gravity field model from satellite orbit perturbations : GRIM5-S1, *Geophysical Research Letters*, *27*, 3611–3614, 2000.
- Birol, F., J.-M. Brankart, F. Castruccio, P. Brasseur, et J. Verron, Impact of ocean mean dynamic topography on satellite data assimilation, *Journal of Marine Geodesy*, *27*, 59–78, 2004.
- Birol, F., J.-M. Brankart, J.-M. Lemoine, P. Brasseur, et J. Verron, Assimilation of satellite altimetry referenced to the new GRACE geoid estimate, *Geophysical Research Letters*, *32*, 2005.
- Bjerknes, J., Atmospheric teleconnections from the equatorial Pacific, *Monthly Weather Review*, *97*, 163–172, 1969.
- Blanc, T., Variation of bulk-derived surface flux, stability, and roughness results due to the use of different transfer coefficient schemes, *Journal of Physical Oceanography*, *15*, 650–669, 1985.
- Blanke, B., et P. Delecluse, Variability of the tropical atlantic ocean simulated by a general circulation model with two different mixed-layer physics, *Journal of Physical Oceanography*, *23*, 1363–1388, 1993.
- Blayo, E., J. Verron, et J. M. Molines, Assimilation of TOPEX/Poseidon altimeter data into a circulation model of the North Atlantic, *Journal of Geophysical Research*, *99*, 24,691–24,705, 1994.
- Bleck, R., An oceanic general circulation model framed in hybrid isopycnic-cartesian coordinates, *Ocean Modelling*, *4*, 55–88, 2002.
- Bleck, R., et D. Boudra, Initial Testing of a Numerical Ocean Circulation Model Using a Hybrid (Quasi-Isopycnic) Vertical Coordinate, *Journal of Physical Oceanography*, *11*, 755–770, 1981.
- Bloom, S. C., L. L. Takacs, A. M. Da Silva, et D. Ledvina, Data assimilation using incremental analysis updates, *Monthly Weather Review*, *124*, 1256–1271, 1996.

- Brankart, J.-M., C.-E. Testut, P. Brasseur, et J. Verron, Implementation of a multivariate data assimilation scheme for isopycnic coordinate ocean models : Application to a 1993-96 hindcast of the North Atlantic Ocean circulation, *Journal of Geophysical Research*, *108*, 1–20, 2003.
- Brasseur, P., *Ocean data assimilation using sequential methods based on the Kalman filter*, vol. Ocean Weather Forecasting, An Integrated View of Oceanography, E. P. Chassignet and J. Verron ed., Springer, 2006.
- Brasseur, P., et J. Verron, The SEEK filter method for data assimilation in oceanography : a synthesis, *Ocean Dynamics*, *56*, 650–661, 2006.
- Brasseur, P., J. Ballabrera-Poy, et J. Verron, Assimilation of altimetric data in the mid-latitude oceans using the SEEK filter with an eddy-resolving primitive equation model, *Journal of Marine Systems*, *22*, 269–294, 1999.
- Broquet, G., Caractérisation des erreurs de modélisation pour l’assimilation de données dans un modèle océanique régional du Golfe de Gascogne, Ph.D. thesis, Université Joseph Fourier - Grenoble I, 2007.
- Busalacchi, A. J., M. J. McPhaden, J. Picaut, et S. R. Springer, Sensivity of wind-driven tropical pacific ocean simulations on seasonal and interannual time scales, *Journal of Marine Systems*, *1*, 119–154, 1990.
- Cane, M. A., A. Kaplan, R. N. Miller, B. Tang, E. C. Hackert, et A. J. Busalacchi, Mapping tropical Pacific sea level : Data assimilation via a reduced state space Kalman filter, *Journal of Geophysical Research*, *101*, 22,599–22,617, 1996.
- Carton, J., G. Chepurin, X. Cao, et B. Giese, A simple ocean data assimilation analysis of the global upper ocean 1950-95. Part I : Methodology, *Journal of Physical Oceanography*, *30*, 294–309, 2000.
- Castruccio, F., J. Verron, L. Gourdeau, J. M. Brankart, et P. Brasseur, On the role of the GRACE mission in the joint assimilation of altimetric and TAO data in the tropical Pacific Ocean model, *Geophysical Research Letters*, *33*, 2006.
- Chassignet, E., et Z. Garraffo, Viscosity parameterization and the Gulf Stream separation, in *From Stirring to Mixing in a Stratified Ocean*, edited by P. Muller et H. D., vol. "Aha Huliko" a Hawaiian Winter Workshop, 2001.
- Chassignet, E. P., et J. Verron, *Ocean Modeling and Parameterization*, vol. 516 of *Mathematical and Physical Sciences*, Kluwer Academic Publishers, The Netherlands, 1998.
- Chassignet, E. P., L. T. Smith, G. R. Halliwell, et R. Bleck, North Atlantic simulation with the HYbrid Coordinate Ocean Model (HYCOM) : Impact of the vertical coordinate choice, reference density, and thermobaricity, *Journal of Physical Oceanography*, *33*, 2504–2526, 2003.
- Cooper, M., et K. Haines, Altimetric assimilation with property conservation, *Journal of Geophysical Research*, *101*, 1059–1077, 1996.

- Cox, H., On the estimation of state variables and parameters for noisy dynamic systems, *IEEE Tran. Automat. Contr.*, 9, 5–12, 1964.
- Cuny, J., P. B. Rhines, P. P. Niiler, et S. Bacon, Labrador sea boundary currents and the fate of the Irminger Sea Water, *Journal of Physical Oceanography*, 32, 627–647, 2002.
- de Boyer Montégut, C., J. Vialard, S. S. C. Shenoi, D. Shankar, F. Durand, C. Ethé, et G. Madec, Simulated seasonal and interannual variability of mixed layer heat budget in the northern Indian Ocean, *J. Climate*, 2005.
- De Cosmo, J., K. Katsaros, S. Smith, R. Anderson, W. Oost, K. Bumke, et H. Chadwick, Air-sea exchange of water vapour and sensible heat : The humidity exchange over the sea (HEXOS) results, *Journal of Geophysical Research*, 101, 12,001–12,016, 1996.
- Debost, F., Etude de nouveaux scénarios d'altimétrie satellitaire pour la reconstruction de la circulation méso-échelle par assimilation de données altimétriques, Ph.D. thesis, Université Joseph Fourier - Grenoble I, 2004.
- Dee, D., et A. da Silva, Data assimilation in the presence of forecast bias, *Quarterly Journal of the Royal Meteorological Society*, 124, 269–295, 1998.
- Delcroix, T., G. Eldin, M. H. Radenac, J. Toole, et E. Firing, Variations of the western equatorial Pacific Ocean, 1986-1988, *Journal of Geophysical Research*, 97, 5423–5445, 1992.
- Delcroix, T., J.-P. Boulanger, F. Masia, et C. Menkes, Geosat-derived sea level and surface current anomalies in the equatorial pacific during the 1986-1989 El Niño and La Niña, *Journal of Geophysical Research*, 99, 25,093–25,107, 1994.
- DeWeaver, E., et S. Nigam, Dynamics of zonal-mean assimilation implications for the winter circulation anomalies, *Journal of Atmospheric Science*, 54, 1758–1775, 1997.
- Dewitte, B., S. Illig, L. Parent, Y. du Penhoat, L. Gourdeau, et J. Verron, Tropical Pacific baroclinic mode contribution and associated lang waves for the 1994-1999 period from an assimilation experiment with altimetric data, *Journal of Geophysical Research*, 108, 3121, 2003.
- Dobricic, S., New mean dynamic topography of the Mediterranean calculated from assimilation system diagnostics, *Geophysical Research Letters*, 32, 2005.
- Douglas, B. C., et R. E. Cheney, Geosat : Beginning a new era in satellite oceanography, *Journal of Geophysical Research*, 95, 2833–2836, 1990.
- Durand, F., Contrôle des couches superficielles des océans tropicaux par assimilation de données satellite : application au Pacifique tropical, Ph.D. thesis, Université Paul Sabatier - Toulouse III, 2003.
- Durand, F., L. Gourdeau, T. Delcroix, et J. Verron, A singular evolutive extended Kalman filter to assimilate ocean color data in a coupled physical-biochemical model of the North Atlantic ocean, *Journal of Geophysical Research*, 107, 8010, 2002.

- Durand, F., L. Gourdeau, T. Delcroix, et J. Verron, Can we improve the representation of modeled ocean mixed layer by assimilating surface-only satellite-derived data? A case study for the Tropical Pacific during the 1997-1998 El Niño, *Journal of Geophysical Research*, *108*, 3200, 2003.
- Durbiano, S., Vecteurs caractéristiques de modèles océaniques pour la réduction d'ordre en assimilation de données, Ph.D. thesis, Université Joseph Fourier - Grenoble I, 2001.
- Evensen, G., Sequential data assimilation with a non linear QG model using monte carlo methods to forecast error statistics, *Journal of Geophysical Research*, *17*, 10,143–10,162, 1994.
- Evensen, G., The ensemble Kalman filter : Theoretical formulation and practical implementation, *Ocean Dynamics*, *53*, 343–367, 2003.
- Fratantoni, D. M., North atlantic surface circulation during the 1990's observed with satellitetracked drifters, *Journal of Geophysical Research*, *106*, 22,067–22,093, 2001.
- Freilich, M. H., et R. S. Dunbar, The accuracy of the NSCAT 1 vector winds : Comparisons with the National Data Buoy Center buoys, *Journal of Geophysical Research*, *104*, 11,231–11,246, 1999.
- Friehe, C., et K. Schmitt, Parameterization of air-sea interface fluxes of sensible heat and moisture by the bulk aerodynamic formulas, *Journal of Physical Oceanography*, *6*, 801–809, 1976.
- Fu, L.-L., et A. Cazenave, *Satellite Altimetry and Earth Sciences : A Handbook of Techniques and Applications*, vol. 69 of *International Geophysics Series*, Academic Press, San Diego, CA, 2001.
- Fu, L.-L., et Y. Chao, The sensitivity of a model to wind forcing : A test using sea level and wind observation from satellites and operational wind analysis, *Geophysical Research Letters*, *24*, 1783–1786, 1997.
- Fukumori, I., R. Raghunath, F. L.-L., et Y. Chao, Assimilation of TOPEX/Poseidon altimeter data into a global ocean circulation model : How good are the results ?, *Journal of Geophysical Research*, *104*, 25,647–25,665, 1999.
- Gelb, A., *Applied optimal estimation*, M.I.T. press, 1974.
- Gent, P. R., et J. C. McWilliams, Isopycnal Mixing in Ocean Circulation Models, *Journal of Physical Oceanography*, *20*, 150–155, 1990.
- Gill, A. E., *Atmospher-Ocean dynamics*, Academic Press, New York, 1982.
- Goosse, H., J. M. Campin, E. Deleersnijder, T. Fichefet, P. P. Mathieu, M. A. Morales Maqueda, et B. Tartinville, Description of the CLIO model version 3.0, *Tech. rep.*, Institut d'Astronomie et de Géophysique G. Lemaître, Université catholique de Louvain, 2 Chemin du Cyclotron, B-1348 Louvain-la-Neuve, Belgique, 1999.

- Gourdeau, L., J. Verron, T. Delcroix, A. J. Busalacchi, et R. Murtugudde, Assimilation of Topex/Poseidon altimeter data in a primitive equation model of the Tropical Pacific ocean, during the 1992-1996 ENSO period, *Journal of Geophysical Research*, *105*, 8473–8488, 2000.
- Gourdeau, L., J.-M. Lemoine, M. H. Rio, et F. Hernandez, Estimating mean dynamic topography in the tropical Pacific ocean from gravity and altimetry satellites, *Geophysical Research Letters*, *30*, 2062, 2003.
- Grima, N., A. Bentamy, K. Katsaros, et Y. Quilfen, Sensitivity of an oceanic general circulation model forced by satellite wind stress fields, *Journal of Geophysical Research*, *104*, 7967–7989, 1999.
- Hackert, E. C., A. J. Busalacchi, et R. Murtugudde, A wind comparison study using an ocean general circulation model for the 1997-1998 El Niño, *Journal of Geophysical Research*, *106*, 2345–2362, 2001.
- Halliwell, G. R., Evaluation of vertical coordinate and vertical mixing algorithms in the HYbrid Coordinate Ocean Model (HYCOM), *Ocean Modelling*, *7*, 285–322, 2004.
- Halpern, D., Observations of annual and El Niño thermal and flow variations at 0°, 110°W and 0°, 95°W during 1980-1985, *Journal of Geophysical Research*, *92*, 8197–9212, 1987.
- Heemink, A. W., M. Verlaan, et A. J. Segers, Variance reduced ensemble Kalman Filtering, *Monthly Weather Review*, *129*, 1718–1728, 2001.
- Hernandez, F., M. Schaeffer, M. H. Calvez, J. Dorandeu, Y. Faugre, et F. Mertz, Surface moyenne océanique : Support scientifique à la mission altimétrique JASON-1, et à une missions micro-satellite altimétrique, *Tech. rep.*, Contract SSALTO 2945 - Lot 2 - A.1, Rapport n° CLS/DOS/NT/00.341, 2001.
- Ho, Y.-C., et B. Whalen, An approach to the identification and control of linear dynamics systems with unknown parameters, *IEEE Tran. Automat. Contr.*, *8*, 255–256, 1963.
- Huang, B., J. L. Kinter, et P. S. Schopf, Ocean data assimilation using intermittent analyses and continuous model error correction, *Advances in Atmospheric Sciences*, *19*, 965–993, 2002.
- Ide, K., P. Courtier, M. Ghil, et A. C. Lorenc, Unified Notation for Data Assimilation : Operational, Sequential and Variational, *Journal of the Meteorological Society of Japan*, *75*, No.1B, 1997.
- Izumo, T., The equatorial undercurrent, meridional overturning circulation, and their roles in mass and heat exchanges during El Niño events in the tropical Pacific ocean, *Ocean Dynamics*, *55*, 110–123, 2005.
- Jazwinski, A. H., *Stochastic processes and filtering theory*, Academic Press, NY, 1970.
- Jin, F. F., An equatorial ocean recharge paradigm for ENSO. Part I : Conceptual model, *Journal of Atmospheric Science*, *54*, 830–847, 1997a.

- Jin, F. F., An equatorial ocean recharge paradigm for ENSO. Part II : A stripped-down coupled model, *54*, 830–847, 1997b.
- Johnson, G. C., M. J. McPhaden, G. Dail Rowe, et K. E. McTaggart, Upper equatorial Pacific Ocean current and salinity variability during the 1996-1998 El Nino-La Nina cycle, *Journal of Geophysical Research*, *105*, 1037–1053, 2000.
- Johnson, G. C., B. M. Sloyan, W. S. Kessler, et K. E. McTaggart, Direct measurements of upper ocean currents and water properties across the tropical Pacific during the 1990s, *Progress in Oceanography*, *52*, 31–61, 2002.
- Kalman, R. E., A new approach to linear filter and prediction problems, *J. Basic. Eng.*, *82*, 35–45, 1960.
- Kessler, W. S., Observations of long waves in the Northern Tropical Pacific, *Journal of Geophysical Research*, *95*, 5183–5217, 1990.
- Kessler, W. S., et M. J. McPhaden, The 1991-93 El Niño in the central Pacific, *Deep-Sea Research*, *42*, 295–333, 1995.
- Killworth, P. D., C. Dieterich, C. Le Provost, A. Oschlies, et J. Willebrand, Assimilation of altimetric data and mean sea surface height into an eddy-permitting model of the North Atlantic, *Progress in Oceanography*, *48*, 312–335, 2001.
- Large, W., *Surface fluxes for practitioners of global ocean data assimilation*, vol. Ocean Weather Forecasting, An Integrated View of Oceanography, E. P. Chassignet and J. Verron ed., Springer, 2006.
- Large, W. G., et S. Pond, Sensible and latent heat flux measurements over the ocean, *Journal of Physical Oceanography*, *12*, 464–482, 1982.
- Large, W. G., J. C. McWilliams, et S. C. Doney, Oceanic vertical mixing : A review and a model with a nonlocal boundary layer parameterization, *Rev. of Geophys.*, *32*, 363–403, 1994.
- Large, W. G., G. Danabasoglu, S. C. Doney, et J. C. McWilliams, Sensitivity to surface forcing and boundary layer in a global ocean model : Annual-mean climatology, *Journal of Physical Oceanography*, *27*, 2418–2447, 1997.
- Le Dimet, F. X., et O. Talagrand, Variational algorithms for analysis and assimilation of meteorological observations : theoretical aspects, *Tellus*, *38*, 1986.
- Le Grand, P., An inverse modelling estimate of the geoid height in the North Atlantic, *Tech. Rep. 9 :1-2*, CERSAT, 1998.
- Le Provost, C., et M. Brémond, Resolution needed for an adequate determination of the mean ocean circulation from altimetry and an improved geoid, *Space Sci. Rev.*, *108*, 163–178, 2003.
- Lemoine, F., S. Kenyon, J. Factor, R. Trimmer, N. Pavlis, D. Chinn, C. Cox, S. Klosko, S. Luthcke, M. Torrence, Y. Wang, R. Williamson, E. Pavlis, R. Rapp, et T. Olson, The development of the joint NASA GSFC and the National Imagery and Mapping Agency

- (NIMA) Geopotential model EGM96, *Tech. Rep. NASA/TP-1998-206861*, NASA Technical Paper, Goddard Space Flight Center, Greenbelt, 1998.
- Lengaigne, M., J.-P. Boulanger, C. Menkes, C. Masson, G. Madec, et P. Delecluse, Ocean reponse to the March 1997 Westerly Wind Event, *Journal of Geophysical Research*, *107*, 8015, 2002.
- Lengaigne, M., G. Madec, et C. Menkes, Impact of isopycnal mixing on the tropical ocean circulation, *Journal of Geophysical Research*, *108*, 3345, 2003.
- Lermusiaux, P. F. J., et A. R. Robinson, Data assimilation via error subspace statistical estimation. Part I : Theory and schemes, *Monthly Weather Review*, *127*, 1385–1407, 1999.
- Levitus, S., World ocean atlas 1998 data, *Tech. rep.*, NOAA-CIRES Clim. Diag. Cent., Boudler, Colorado, 1998.
- Li, J.-G., K. Killworth, et D. Smeed, Response of an eddy-permitting ocean model to the assimilation of sparse in situ data, *Journal of Geophysical Research*, *108*, 3111, 2003.
- Lions, J. L., *Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles*, Dunod, Gauthier-Villars, Paris, 1968.
- Lions, J. L., O. P. Manley, T. R., et W. S., Physical interpretation of the attractor dimension for the primitive equations of atmospheric circulation, *Journal of Atmospheric Science*, *54*, 1137–1143, 1997.
- Ljung, L., Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems, *IEEE Tran. Automat. Contr.*, *24*, 36–50, 1979.
- Luong, B., Techniques de contrôle optimal pour un modèle quasi-géostrophique de circulation océanique. Application à l'assimilation variationnelle des données altimétriques satellitaires, Ph.D. thesis, Université Joseph Fourier - Grenoble I, 1995.
- Luong, B., J. Blum, et J. Verron, A variational method for the resolution of a data assimilation problem in oceanography, *J. of Inverse Problems*, *14*, 979–997, 1998.
- Madec, G., et M. Imbard, A global ocean mesh to overcome the North Pole singularity, *Clim. Dyn.*, *12*, 381–388, 1996.
- Madec, G., P. Delecluse, M. Imbard, et C. Levy, OPA 8.1 general circulation model reference manual, *Tech. rep.*, Notes de l'IPSL, Université P. et M. Curie, B102 T15-E5, 4 place Jussieu, Paris cedex 5, France, 1998.
- Maes, C., Estimating the influence of salinity on sea level anomaly in the ocean., *Geophysical Research Letters*, *25*, 3551–554, 1998.
- Maes, C., M. J. McPhaden, et D. Behringer, Signatures of salinity variability in tropical Pacific Ocean dynamic height anomalies, *Journal of Geophysical Research*, *107*, 8012, 2002.
- McCreary, J. P., A model of tropical ocean-atmosphere interaction, *Monthly Weather Review*, *111*, 370–387, 1983.

- McPhaden, M., T. Delcroix, K. Hanawa, Y. Kuroda, G. Meyers, P. J., et S. M., *The El Niño/Southern Oscillation (ENSO) Observing System*, vol. Observing the Ocean in the 21st Century, Australian Bureau of Meteorology, Melbourne, Australia, 2001.
- McPhaden, M. J., et S. P. Hayes, Variability in the eastern Pacific equatorial during 1986-1988, *Journal of Geophysical Research*, *95*, 13,195–13,208, 1990.
- McPhaden, M. J., A. J. Busalacchi, R. Cheney, J. R. Donguy, K. S. Gage, D. Halpern, M. Ji, P. Julian, G. Meyers, G. T. Mitchum, P. P. Niiler, J. Picaut, R. W. Reynolds, N. Smith, et T. K., The Tropical Ocean-Global Atmosphere observing system : A decade of progress, *Journal of Geophysical Research*, *103*, 14,169–14,240, 1998.
- Menkes, C., J. Boulanger, A. Busalacchi, J. Vialard, P. Delecluse, M. J. McPhaden, E. Hackert, et N. Grima, Impact of TAO vs. ERS wind stresses onto simulations of the tropical Pacific Ocean during the 1993-1998 period by the OPA OGCM, in *Climate Impact of Scale Interaction for the Tropical Ocean-Atmosphere System*, vol. 13 of *Euroclivar Workshop Report*, 1998.
- Mercier, H., Determining the general circulation of the ocean : a non-linear inverse problem, *Journal of Geophysical Research*, *91*, 5103–5109, 1986.
- Meyers, G., H. Phillips, N. Smith, et S. J., Space and time scales for optimal interpolation of temperature-tropical Pacific Ocean, *Progress in Oceanography*, *28*, 189–218, 1991.
- Miller, R. N., et M. A. Cane, A Kalman Filter analysis of sea level height in the tropical Pacific, *Journal of Physical Oceanography*, *19*, 773–790, 1989.
- Mitchell, J. L., J. M. Dastugue, W. J. Teague, et Z. R. Hallock, The estimation of geoid profiles in the Northwest Atlantic from simultaneous satellite altimetry and airborne expendable bathythermograph sections, *Journal of Geophysical Research*, *91*, 5103–5109, 1990.
- Nelson, L., et E. Stear, The simultaneous on-line estimation of parameters and states in linear systems, *IEEE Tran. Automat. Contr.*, *21*, 94–98, 1976.
- Nerem, R. S., C. Jekeli, et W. M. Kaula, Gravity field determination and characteristics retrospective and prospective, *Journal of Geophysical Research*, *100*, 1995.
- Niiler, P., *The world ocean surface circulation*, vol. Ocean circulation and climate of *International Geophysics Series 77*, siedler, g. and church, j. and gould, j. ed., Academic Press, 2001.
- Niiler, P. P., N. A. Maximenko, et J. McWilliams, Dynamically balanced absolute sea level of the global ocean derived from near-surface velocity observations, *Geophysical Research Letters*, *30*, 2164, 2003.
- Ourmières, Y., J.-M. Brankart, L. Berline, P. Brasseur, et J. Verron, Incremental Analysis Update implementation into a sequential ocean data assimilation system, *Journal of Atmospheric and Oceanic Technology*, 2006.
- Parent, L., Assimilation de données dans l'Océan Pacifique Tropical. Reconstruction de la période 94-98., Ph.D. thesis, Université Joseph Fourier - Grenoble I, 2000.

- Parent, L., C.-E. Testut, J.-M. Brankart, J. Verron, P. Brasseur, et L. Gourdeau, Comparative assimilation of Topex/Poseidon and ERS altimeter data and of TAO temperature data in the Tropical Pacific Ocean during 1994-1998, and the mean sea-surface height issue, *Journal of Marine Systems*, 40, 2003.
- Penduff, T., P. Brasseur, C. Testut, B. Barnier, et J. Verron, A four-year eddy-permitting assimilation of the sea surface temperature and altimetric data in the South Atlantic Ocean, *Journal of Marine Research*, 60, 805–833, 2002.
- Pham, D. T., J. Verron, et R. M. C., Singular evolutive extended Kalman filter with EOF initialization for data assimilation in oceanography, *Journal of Marine Systems*, 16, 323–340, 1998.
- Philander, S. G., *El Niño, La Niña, and the Southern Oscillation*, vol. 46 of *International Geophysics Series*, Academic Press, San Diego, CA, 1990.
- Picaut, J., F. Masia, et Y. du Penhoat, An advective-reflective conceptual model for the oscillatory nature of the ENSO, *Science*, 277, 663–666, 1997.
- Pinardi, N., A. Rosati, et R. Pacanowski, The sea surface pressure formulation of rigid lid models, *Journal of Marine Systems*, 6, 109–119, 1995.
- Rapp, R. H., Equational radius estimates from topex altimeter data, *Tech. rep.*, Institute of Geodesy and Navigation, Festschrift Erwin Groten, University FAF, Munich, 1995.
- Reigber, C., R. Schmidt, F. Flechtner, R. König, U. Meyer, K.-H. Neumayer, P. Schwintzer, et S. Yuan Zhu, An Earth gravity field model complete to degree and order 150 from GRACE : EIGEN-GRACE02S, *Journal of Geodynamics*, 39, 1–10, 2005.
- Reverdin, G., P. P. Niiler, et H. Valdimarsson, North Atlantic Ocean surface currents, *Journal of Geophysical Research*, 108, 2003.
- Reynaud, T. H., W. A. J., et R. J. Greatbatch, Summer mean circulation of the northwestern Atlantic Ocean, *Journal of Geophysical Research*, 100, 779–816, 1995.
- Reynolds, R. W., et T. M. Smith, Improved global sea surface temperature analyses using optimum interpolation, *J. Climate*, 7, 929–948, 1994.
- Rio, M. H., et F. Hernandez, Estimation d'une topographie dynamique moyenne sur l'Atlantique Nord et Tropical, *Tech. Rep. n CLS/DOS/NT/02.515*, CLS, Ramonville St Agne, 2002.
- Robert, C., Développement et comparaison de méthodes d'assimilation de données de rang réduit dans une modèle de circulation océanique : application à l'océan Pacifique Tropical, Ph.D. thesis, Université Joseph Fourier - Grenoble I, 2004.
- Robert, C., E. Blayo, et J. Verron, Comparison of reduced-order, sequential and variational, data assimilation methods in the Tropical Pacific ocean, 2006.
- Roquet, H., S. Planton, et P. Gaspar, Determination of ocean surface heat fluxes by a variational method, *Journal of Geophysical Research*, 98, 10,211–10,221, 1993.

- Roulet, G., et G. Madec, Salt conservation, free surface, and varying levels : a new formulation for ocean general circulation models, *Journal of Geophysical Research*, *105*, 23,927–23,942, 2000.
- Skachko, S., J. M. Brankart, F. Castruccio, P. Brasseur, et J. Verron, Estimating the turbulent air-sea flux bulk parameters by sequential data assimilation, *Remote Sensing of Environment*, 2007.
- Smith, S., Coefficients for sea surface wind stress, heat flux and wind profiles as a function of wind speed and temperature, *Journal of Geophysical Research*, *93*, 15,467–15,472, 1988.
- Smith, S. R., D. M. Legler, et V. K. V., Quantifying uncertainties in NCEP reanalyses using high-quality research vessel observations, *Journal of Climate*, *14*, 4062–4072, 2001.
- Stammer, D., K. Ueyoshi, A. Köhl, W. G. Large, S. A. Josey, et C. Wunsch, Estimating air-sea fluxes of heat, freshwater, and momentum through global ocean data assimilation, *Journal of Geophysical Research*, *109*, 2004.
- Suarez, M. J., et P. S. Schopf, A delayed action oscillator for ENSO, *Journal of Atmospheric Science*, *45*, 3283–3287, 1988.
- Taft, B. A., et W. S. Kessler, Variations of zonal currents in the central tropical Pacific during the 1970 to 1989 : Sea level and dynamic height measurements, *Journal of Geophysical Research*, *96*, 12,599–12,618, 1991.
- Tapley, B. D., D. P. Chambers, S. Bettadpur, et J. C. Ries, Large scale ocean circulation from GRACE GGM01, *Geophysical Research Letters*, *30*, 2163, 2003.
- Testut, C. E., Assimilation de données satellites avec un filtre de Kalman de rang réduit dans un modèle aux équations primitives de l’océan atlantique, Ph.D. thesis, Université Joseph Fourier - Grenoble I, 2000.
- Testut, C.-E., P. Brasseur, J.-M. Brankart, et J. Verron, Assimilation of sea-surface temperature and altimetric observations during 1992–1993 into an eddy permitting primitive equation model of the North Atlantic Ocean, *Journal of Marine Systems*, *40*, 2003.
- Trenberth, K. E., et D. J. Shea, On the evolution of the Southern Oscillation, *Monthly Weather Review*, *115*, 3078–3096, 1987.
- van Leeuwen, P. J., et G. Evensen, Data assimilation and inverse methods in terms of a probabilistic formulation, *Monthly Weather Review*, *124*, 2898–2913, 1996.
- Verron, J., Nudging satellite data into quasi-geostrophic ocean models, *Journal of Geophysical Research*, *97*, 7479–7491, 1992.
- Verron, J., et W. R. Holland, Impacts de données d’altimétrie satellitaire sur les simulations numériques des circulations océaniques aux latitudes moyennes, *Annales Geophysicae*, *71*, 31–46, 1989.

- Verron, J., L. Gourdeau, D. T. Pham, R. Murtugudde, et A. J. Busalacchi, An extended Kalman filter to assimilate satellite altimeter data into a non-linear numerical model of the Tropical Pacific : method and validation, *Journal of Geophysical Research*, 104, 5441–5458, 1999.
- Vialard, J., C. Menkes, J. P. Boulanger, P. Delecluse, E. Guilyardi, M. J. McPhaden, et G. Madec, A model study of oceanic mechanisms affecting equatorial pacific sea surface temperature during the 1997-1998 El niño, *Journal of Geophysical Research*, 31, 1649–1675, 2001.
- Vidard, P. A., Vers une prise en compte des erreurs modèles en assimilation de données 4D-Variationnelle, Ph.D. thesis, Université Joseph Fourier - Grenoble I, 2001.
- Wang, C., et J. Picaut, *Understanding ENSO-A Review*, vol. Earth's Climate. The Ocean-Atmosphere Interaction, wang, c. and xie,s.-p. and carton, j. a. ed., 2003.
- Weisberg, R. H., et C. Wang, A western Pacific oscillator paradigm for the El Niño-Southern Oscillation, *Geophysical Research Letters*, 24, 779–782, 1997.
- Wittenberg, A. T., On extended wind stress analyses for ENSO, *Journal of Climate*, 17, 2526–2540, 2003.
- Xie, P., et P. Arkin, Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions, *Journal of Climate*, 2, 840–858, 1996.
- Zhu, Y., R. Guo, J. Cohn, I. M. Navon, et Y. Yang, The goes-3 restrospective data assimilation system : the 6-hour lag case, *Monthly Weather Review*, 131, 2129–2150, 2003.
- Zillmann, J., A study of some aspects of the radiation and the heat budgets of the southern hemisphere oceans, *Meteorol. Stud.*, 26, 1972.

Table des figures

1.1	<i>Modification globale du climat pendant un évènement ENSO.</i>	12
1.2	<i>Climatologie COADS de la vitesse du vent et de la position moyenne de l'ITCZ et de la SPCZ.</i>	13
1.3	<i>SST climatologique (Levitus, 1998).</i>	14
1.4	<i>Circulation de surface et subsurface de l'océan Pacifique Tropical (source Philander (1990)).</i>	15
1.5	<i>Section de température à l'équateur (climatologie Levitus (1998)).</i>	16
1.6	<i>Climatologie saisonnière pour la SST (Levitus, 1998).</i>	18
1.7	<i>Distribution spatiale des coefficients de corrélation ($\times 10$) entre les pressions atmosphériques de surface et la pression atmosphérique à Darwin, Australie. (source Trenberth et Shea (1987)).</i>	19
1.8	<i>L'indice SOI et l'anomalie de SST pour la région Niño-3.4.</i>	20
1.9	<i>Représentation schématique de l'oscillateur retardé présentant la croissance et la terminaison d'une anomalie chaude.</i>	21
1.10	<i>Représentation schématique du couplage océan-atmosphère en situation normale (en haut), en situation El Niño (au milieu) et en situation La Niña (en bas).</i>	22
2.1	<i>Position des flotteurs ayant délivré des données dans les 30 derniers jours le 7 août 2006 (source site ARGO).</i>	27
2.2	<i>Surface moyenne(MSSH) en mètres mesurée par altimétrie satellite (Hernandez et al., 2001).</i>	29
2.3	<i>Principe de la mesure altimétrique.</i>	31
2.4	<i>Amplitude du signal et de l'erreur par degré pour différents modèles de géoïde (source Reigber et al. (2005)).</i>	32
2.5	<i>Erreur en fonction de la résolution spatiale pour différents modèles de géoïde (source Reigber et al. (2005)).</i>	33
2.6	<i>Topographie dynamique moyenne (MDT) en mètres référencé au géoïde EIGEN-GRACE02S.</i>	34
2.7	<i>Anomalies de surface le long des traces pour l'analyse du 15 novembre 1997.</i>	35
2.8	<i>Position des mouillages TAO le long du rail équatorial.</i>	36
2.9	<i>Position des profils XBT sur la période 1993-1998.</i>	37
3.1	<i>Disposition des variables sur la grille C. T indique le point où les grandeurs scalaires sont calculées. (u, v, w) les points où les grandeurs vectorielles sont calculées.</i>	46
3.2	<i>Le domaine du modèle et la topographie de fond de ORCA2.</i>	51

3.3	<i>Maillage et masque des points T en surface dans le Pacifique.</i>	52
3.4	<i>Caractéristiques du maillage vertical.</i>	53
4.1	<i>Système d'assimilation de données.</i>	61
4.2	<i>Représentation vectorielle des erreurs d'analyse et de prévision dans l'espace d'état (Source (Brasseur, 2006)).</i>	68
4.3	<i>Représentation conceptuelle de la séquence d'assimilation du filtre de Kalman (Source Brasseur (2006)).</i>	71
4.4	<i>Schéma de l'assimilation séquentielle classique.</i>	79
4.5	<i>Schéma IAU.</i>	81
4.6	<i>Coefficient appliqué à l'incrément d'analyse. Isocontours tous les 0.2.</i>	82
5.1	<i>Topographie dynamique absolue le long des traces pour l'analyse du 22 septembre 1995. La fenêtre temporelle d'observation est de 5 jours.</i>	90
5.2	<i>Profils in-situ de température TAO pour l'analyse du 22 septembre 1995. La fenêtre temporelle d'observation est de 5 jours.</i>	91
5.3	<i>Estimation de l'erreur à l'échéance de la prévision.</i>	94
5.4	<i>Premier mode 3D multivarié de la covariance d'erreur initiale \mathbf{P}_0 pour la SSH (à gauche en m), la température à l'équateur (à droite en $^{\circ}\text{C}$).</i>	96
6.1	<i>MDT déduite des données satellite GRACE (à gauche) et de la climatologie Levitus (à droite). Isocontours tous les 5 cm. En bas : vitesse zonale géostrophique moyenne à la surface déduite respectivement de la MDT GRACE (à gauche) et de la MDT Levitus (à droite). Isocontours tous les 10 cm^{-1}.</i>	101
6.2	<i>Représentation schématique d'une coupe méridienne de topographie dynamique.</i>	102
6.3	<i>Différence pour la période 1993-2001 entre la hauteur dynamique moyenne TAO entre 0-500 dbar et, respectivement, la MDT GRACE (à gauche) et MDT déduite de la climatologie Levitus (à droite). Isocontours tous les 2 cm^{-1}.</i>	103
6.4	<i>Différence RMS pour la période 1993-2001 entre la hauteur dynamique TAO entre 0-500 dbar et le signal altimétrique de TOPEX/Poseidon référencé, respectivement, à la MDT GRACE (à gauche) et à la MDT Levitus (à droite). Isocontours tous les 2 cm^{-1}.</i>	104
6.5	<i>Ecart RMS (en cm) par rapport à la topographie dynamique (SLA + MDT) pour la simulation libre (en vert) et les expériences avec assimilation de données EXP1 (en cyan) et EXP2 (en bleu).</i>	106
6.6	<i>Ecart RMS (en $^{\circ}\text{C}$) par rapport aux profils de température TAO indépendant pour la simulation libre (en vert) et les expériences avec assimilation de données EXP1 (en cyan) et EXP2 (en bleu).</i>	106
6.7	<i>MDT (à gauche en m) et écart type de la DT (à droite en m) observé, de la simulation en mode libre, de la simulation avec assimilation de SLA EXP1 et de la simulation avec assimilation de DT EXP2, respectivement de haut en bas.</i>	107

6.8	<i>Vitesse zonale moyenne à 15 m de profondeur (en ms^{-1}). Respectivement de gauche à droite et de haut en bas, la climatologie proposée par Niiler (2001), la moyenne pour l'année 1993 pour la simulation libre, l'EXP1 et l'EXP2.</i>	109
6.9	<i>Section de MDT à 140° E pour la MDT GRACE (en tiret), la MDT modèle (en tiret mixte) et la MDT de la simulation libre sur l'année 1993 (en vert).</i>	110
6.10	<i>Profondeur moyenne pour l'année 1993 des isothermes 12, 16, 20, 24 et 28° C le long de l'équateur pour la simulation libre (en vert), les expériences avec assimilation de données EXP1 (en cyan) et EXP2 (en bleu) et les données indépendantes TAO (en magenta).</i>	111
6.11	<i>Section de température moyenne à 140° E pour la simulation libre (en vert), et les expériences avec assimilation de données EXP1 (en cyan) et EXP2 (en bleu).</i>	112
7.1	<i>Ecarts RMS (en cm) par rapport à la topographie dynamique (SLA + MDT) pour la simulation libre (en vert) et les expériences avec assimilation de données EXP2 (en bleu) et EXP3 (en rouge) et EXP4 (en noir).</i>	117
7.2	<i>Ecarts RMS (en °C) par rapport aux profils de température TAO pour la simulation libre (en vert) et les expériences avec assimilation de données EXP2 (en bleu), EXP3 (en rouge) et EXP4 (en noir).</i>	118
7.3	<i>MDT (à gauche en m) et écart type de la DT (à droite en m) pour les simulations EXP3 (en haut) avec assimilation des profils de température TAO et EXP4 (en bas) avec assimilation conjointe de la DT et des profils de température TAO.</i>	119
7.4	<i>Evolution temporelle de la DT à l'équateur à 156° E de longitude de la simulation assimilant la topographie dynamique (en bleu), de la simulation assimilant les TAO (en rouge) et de la simulation assimilant conjointement la topographie dynamique et les TAO (en noir). L'évolution temporelle de la hauteur dynamique déduite des TAO au même point est tracée (en magenta). Pour la hauteur dynamique, le niveau de référence est arbitraire.</i>	120
7.5	<i>Profil de température (à gauche) et de salinité (à droite) moyen pour l'année 1993 au point 0° N-156° E. Les profils des observations TAO (en magenta), de la simulation assimilant la topographie dynamique (en bleu), de la simulation assimilant les TAO (en rouge) et de la simulation assimilant conjointement la topographie dynamique et les TAO (en noir), sont tracés.</i>	120
7.6	<i>Vitesse zonale moyenne à 15 m de profondeur (en ms^{-1}) pour les simulation EXP3 (en haut) avec assimilation des profils de température TAO et de la simulation EXP4 (en bas) avec assimilation conjointe de la DT et des profils de température TAO.</i>	122
7.7	<i>Profondeur moyenne pour l'année 1993 des isothermes 12, 16, 20, 24 et 28° C le long de l'équateur pour la simulation libre (en vert), les expériences avec assimilation de données EXP3 (en rouge) et EXP4 (en noir) et les données indépendantes TAO (en magenta).</i>	123
7.8	<i>GRACE MDT (in meter).</i>	126
7.9	<i>Difference for the 1993-2001 period between the mean 0/500 dbar TAO dynamic height and the model MDT (a) and the GRACE MDT (b). Contour interval is 2 cm. TAO moorings are located at crosses.</i>	128

7.10	<i>Mean zonal velocity for year 1993 at 25 m depth for the free run (a), for EXP1 (b) and for EXP2 (c). Contour interval is 0.2 ms^{-1}.</i>	129
7.11	<i>Difference maps between the simulated MDT for year 1993 and the satellite observed MDT for the free run (a), for EXP1 (b) and for EXP2 (c). Contour interval is 5 cm.</i>	130
7.12	<i>Temperature RMSD with respect to independent XBT data for the 25°S and 20°N strip for year 1993 for the free run (green), for EXP1 (blue) and for EXP2 (red).</i>	132
8.1	<i>Ecart RMS (en cm) par rapport à la topographie dynamique (SLA T/P+ERS + MDT) pour la simulation libre (en vert) et pour la simulation assimilée (en noir).</i>	137
8.2	<i>Ecart RMS (en $^{\circ}\text{C}$) par rapport aux profils de température TAO pour la simulation libre (en vert) et pour la simulation assimilée (en noir).</i>	137
8.3	<i>Profils d'écarts RMS (en $^{\circ}\text{C}$) sur la verticale par rapport aux profils de température TAO (à gauche) et par rapport aux profils de température XBTs indépendantes collectés par les navire VOS (à droite) pour tout le Pacifique Tropical. Les RMSD pour la simulation libre (en vert) et pour la simulation assimilée (en noir) sont tracées.</i>	139
8.4	<i>Identique à la figure 8.3 mais pour la boîte Niño 3.4 (5°S-5°N, 170°W-120°W).</i>	139
8.5	<i>MDT (à gauche en m) et écart type de la DT (à droite en m) observés, de la simulation en mode libre, de la simulation avec assimilation de données (respectivement de haut en bas).</i>	141
8.6	<i>Diagramme longitude-temps à l'équateur de la topographie dynamique (en haut, en m) et de la SLA (en bas, en m). L'évolution de la position de l'isotherme 29°C est superposée sur chaque diagramme (pour les observations, la SST NCEP Reynolds hebdomadaire proposée par Reynolds et Smith (1994) est utilisée). Observation, à gauche. Simulation en mode libre, au milieu. Simulation avec assimilation de données, à droite.</i>	142
8.7	<i>Vitesse zonale moyenne pour la période 1993-1998 pour la simulation libre (à gauche) et pour la simulation assimilée (à droite).</i>	144
8.8	<i>Section moyenne de vitesse zonale à l'équateur pour le simulation libre (à gauche) et la simulation avec assimilation de données (à droite).</i>	146
8.9	<i>Profils moyens de vitesse zonale le long de l'équateur pour les longitudes 156°E, 165°E, 170°W, 140°W et 110°W, respectivement de droite à gauche et de haut en bas. Profils pour la simulation libre (en vert), la simulation avec assimilation de données (en noir) et les observations ADCP TAO (en magenta).</i>	146
8.10	<i>Ecart type de température le long de l'équateur et pour une section méridienne à 155°W pour les observations TAO (en haut), la simulation libre (au milieu) et la simulation avec assimilation de données (en bas). Les isocontours du champ moyen de température sont superposés.</i>	148
8.11	<i>Position des rails West Pacific, Central Pacific et East Pacific échantillonné de manière régulière par les navire VOS. Ces rails correspondent respectivement aux lignes maritimes Tokyo-Auckland, San Francisco-Auckland et Panama-Auckland.</i>	149

8.12	<i>Position des profils collectés le long des rails West P., Central P. et East P. respectivement de haut en bas en fonction du temps.</i>	150
8.13	<i>Différences de température moyenne sur la période 1993-1998 entre les données XBTs et la simulation libre (à gauche) et entre les données XBTs et la simulation avec assimilation de données (à droite) le long des rails West Pacific, Central P. et East P. définis sur la figure 8.11. Les isocontours du champ moyen de température observé par les XBTs sont superposés.</i>	151
8.14	<i>Diagramme longitude-temps à l'équateur de la profondeur de l'isotherme 20° C. L'évolution de la position de l'isotherme 29° C est superposée.</i>	153
8.15	<i>Carte des différences entre la MDT simulée pour la période 1993-1998 et la MDT observée par satellite pour la simulation libre (à gauche) et pour la simulation avec assimilation de données (à droite).</i>	154
8.16	<i>MDT le long des longitude 140° W (à gauche) et 145° W (à droite) pour la simulation libre (en vert), la simulation avec assimilation de données (en noir), et la MDT satellite observée (en tiret). Sur la section à 140° W, les hauteurs dynamiques issues des mouillages TAO sont également tracées (triangle magenta).</i>	156
9.1	<i>Section moyenne de vitesse zonale sur la période 1993-1998 à l'équateur. Les isopycnes $\sigma_\theta = 24 \text{ kgm}^{-3}$ et $\sigma_\theta = 26.25 \text{ kgm}^{-3}$ sont superposées.</i>	162
9.2	<i>Transport zonal de masse ($10^6 \text{ m}^3 \text{ s}^{-1}$) pour la couche de surface avec $\sigma_\theta < 24 \text{ kgm}^{-3}$. Les vecteurs sont proportionnels au transport moyen dans la classe de densité pour chaque intervalle de 1° de latitude. Pour chaque année, le printemps boréal (en haut) et l'automne boréal (en bas).</i>	163
9.2	<i>(suite) Transport zonal de masse ($10^6 \text{ m}^3 \text{ s}^{-1}$) pour la couche de surface avec $\sigma_\theta < 24 \text{ kgm}^{-3}$. Les vecteurs sont proportionnels au transport moyen dans la classe de densité pour chaque intervalle de 1° de latitude. Pour chaque année, le printemps boréal (en haut) et l'automne boréal (en bas).</i>	164
9.2	<i>(suite) Transport zonal de masse ($10^6 \text{ m}^3 \text{ s}^{-1}$) pour la couche de surface avec $\sigma_\theta < 24 \text{ kgm}^{-3}$. Les vecteurs sont proportionnels au transport moyen dans la classe de densité pour chaque intervalle de 1° de latitude. Pour chaque année, le printemps boréal (en haut) et l'automne boréal (en bas).</i>	165
9.3	<i>Diagramme longitude-temps à l'équateur de l'anomalie de la tension zonale de vent (à gauche) et de la SST (à droite) pour la simulation avec assimilation de données.</i>	166
9.4	<i>Transport zonal de masse ($10^6 \text{ m}^3 \text{ s}^{-1}$) pour la couche intermédiaire avec $24 \text{ kg m}^{-3} < \sigma_\theta < 26.25 \text{ kg m}^{-3}$. Les vecteurs sont proportionnels au transport moyen dans la classe de densité pour chaque intervalle de 1° de latitude. Pour chaque année, le printemps boréal (en haut) et l'automne boréal (en bas).</i>	168
9.4	<i>Transport zonal de masse ($10^6 \text{ m}^3 \text{ s}^{-1}$) pour la couche intermédiaire avec $24 \text{ kg m}^{-3} < \sigma_\theta < 26.25 \text{ kg m}^{-3}$. Les vecteurs sont proportionnels au transport moyen dans la classe de densité pour chaque intervalle de 1° de latitude. Pour chaque année, le printemps boréal (en haut) et l'automne boréal (en bas).</i>	169

9.4	<i>Transport zonal de masse ($10^6 m^3 s^{-1}$) pour la couche intermédiaire avec $24 kg m^{-3} < \sigma_\theta < 26.25 kg m^{-3}$. Les vecteurs sont proportionnels au transport moyen dans la classe de densité pour chaque intervalle de 1° de latitude. Pour chaque année, le printemps boréal (en haut) et l'automne boréal (en bas).</i>	170
A.1	<i>Sketch of the HYCOM model configuration.</i>	188
A.2	<i>Mean Dynamic Topography (in meters). (a) MDT0, (b) MDT1, (c) MDT2 and (d) MDT3.</i>	190
A.3	<i>Section of the different mean dynamic topographies (in meters) along (a) $72^\circ W$ and (b) along $55^\circ W$. MDT1 is in green, MDT2 is in red and MDT3 is in blue.</i>	191
A.4	<i>Definition of height measurements.</i>	194
A.5	<i>Mean surface current velocity for (a) the control run EXPT0, and the (b) EXPT1, (c) EXPT2 and (d) EXPT3 assimilation experiment.</i>	196
A.6	<i>Mean barotropic streamfunction for (a) the control run EXPT0, and the (b) EXPT1, (c) EXPT2 and (d) EXPT3 assimilation experiment.</i>	197
A.7	<i>(a) Temperature section along $55^\circ W$ from Levitus climatology. Anomaly between the mean temperature from (b) the control run, (c) EXPT1, (d) EXPT2, (e) EXPT3, (f) EXPT4 and the climatological temperature field along $55^\circ W$.</i>	199
A.8	<i>Horizontal distribution of XBT profiles available during 1993 (SISMER).</i>	200
A.9	<i>Temperature ($^\circ C$) and salinity (PSU) RMS misfit with respect to TS climatology and to XBT profiles in (a) the Gulf Stream region and in (b) the Northeast region, down to 700 m and averaged for 1993. The figure shows plots of the free run (black curve), the EXPT1 (turquoise), EXPT2 (green), EXPT3 (red) and EXPT4 (yellow) solutions. For each assimilation experiment, the 3-day forecast is shown by the continuous-line curve and the analyses by the dashed-line curves.</i>	201
A.10	<i>Time evolution of the sea surface elevation averaged over the model domain for EXPT0 (black), EXPT2 (red) and EXPT4 (green).</i>	203
B.1	<i>EOF 1 de la variabilité de la tension du vent τ à la surface océan/atmosphère pour la période 1993-1996. A gauche : le signal temporel associé.</i>	206
B.2	<i>MDT pour les trois premières simulations perturbées suivant les trois premiers modes de variabilité de la tension de vent $\vec{\tau}$ (respectivement de gauche à droite et de haut en bas) et la section de température moyenne à l'équateur pour les trois simulations (isothermes en noire pour la première simulation, rouge pour la deuxième et verte pour la troisième).</i>	207
B.3	<i>Premier mode 3D multivarié de la covariance d'erreur initiale \mathbf{P}_0 pour la SSH (à gauche en m), la température à l'équateur (à droite en $^\circ C$).</i>	208
B.4	<i>MDT GRACE (en haut), ainsi que la MDT de la simulation non perturbée (au milieu à gauche) et la combinaison linéaire des MDTs issues de l'ensemble (au milieu à droite). En bas, carte des différences avec la MDT GRACE pour, respectivement, la MDT de la simulation non perturbée (à gauche) et la combinaison linéaire des MDTs.</i>	209
B.5	<i>Schéma de principe.</i>	210

C.1	Assimilation process of the original variant V_0 of the scheme. The dashed lines show the model forecast, and the solid line shows the analysis. The analysis is obtained by adding the observational update of time t_k , in the model operator between time t_{k-1} and t_k . The forecast is performed using the nominal values of the parameters \mathbf{p}^0	217
C.2	Assimilation process of the variant V_1 of the scheme. The dashed lines show the model forecast, and the solid line show the analysis. The analysis is obtained by modifying the model operator between time t_{k-1} and t_k using the parameters \mathbf{p}_k^a estimated from the observations at time t_k . The forecast between time t_k and t_{k+1} is performed using the updated values of the parameters \mathbf{p}_k^a	218
C.3	Assimilation process of the variant V_2 of the scheme. The dashed lines show the model forecast, and the solid line show the analysis. The analysis is obtained by adding the observational update of time t_k , in the model operator between time t_{k-1} and t_k , without modifying the model parameters \mathbf{p}_{k-1}^a . The forecast between time t_k and t_{k+1} is performed using the updated values of the parameters \mathbf{p}_k^a	218
C.4	Average values (top panels) and standard deviations (bottom panels) of the turbulent exchange coefficients C_E (left panels) and C_H (right panels) in the reference model (original ORCA2 bulk formulation). The values are scaled by a factor 10^3	220
C.5	Evolution of error variance on SST and SSS for the world ocean (except the Northern polar zones). The black solid line shows the error corresponding to the free simulation (without assimilation). The green, yellow, blue and red lines correspond to V_0 , V_0^* , V_1 and V_2 respectively. The solid lines represent the continuous analysis, while the bullets represent the 10-day model forecasts.	223
C.6	Evolution of error variance on SST and SSS for the Atlantic ocean only. The green, yellow, blue and red lines correspond to V_0 , V_0^* , V_1 and V_2 respectively. The solid lines represent the continuous analysis and the bullets represent the 10-day model forecasts.	224
C.7	Spatial distribution of standard deviation error on SST (left-hand column) and SSS (right-hand column) 10-day forecast for V_0 (upper plots), V_1 (middle plots) and V_2 (bottom plots) of the assimilation scheme.	225
C.8	Standard deviation error values on temperature (left-hand column) and salinity (right-hand column) 10-day forecast profiles for V_0 (upper plots), V_1 (middle plots) and V_2 (bottom plots) of the assimilation scheme	226
C.9	Standard deviation error on the latent heat flux exchange coefficient C_E (left-hand column) and on the sensible heat flux exchange coefficient C_H (right-hand column), for V_0 (upper plots), V_1 (middle plots) and V_2 (bottom plots) of the assimilation scheme.	228
C.10	Detailed statistics for two specific locations of the world ocean : $2^\circ\text{W}16^\circ\text{S}$ (top panels) and $60^\circ\text{W}36^\circ\text{N}$ (bottom panels) : C_E parameter evolution in time (left panels), temperature innovations (centre panels) and C_E parameter corrections (right panels). The black dashed line stands for the true ocean, and the green, yellow, blue and red lines correspond to V_0 , V_0^* , V_1 and V_2 respectively. (The green curve in the right panels displays the correction on the parameters that is computed in V_0 but never applied.) . . .	230

C.11 Spatial distribution of standard deviation error values on the latent heat flux (left-hand column) and the sensible heat flux (right-hand column) for V_0 (upper plots), V_1 (middle plots) and V_2 (bottom plots).	232
C.12 Evolution of the error variance values for a 3 months forecast corresponding to V_0 , V_1 and V_2 of the assimilation scheme (green, blue and red curves respectively). The forecasts start on April 4, 1993 (4 months after the beginning of the experiments).	233

Liste des tableaux

3.1	<i>Principales caractéristiques de la configuration utilisée.</i>	57
4.1	<i>Les différentes étapes de l'algorithme SEEK.</i>	75
5.1	<i>Les différentes expériences réalisées.</i>	97
7.1	Standard deviation and percentage error variance reduction of the RMS misfit with respect to the SSH referenced to GRACE MDT and the TAO and XBT temperature profiles. EXP2* excludes the Western Equatorial region (see text).	131
9.1	<i>Moyenne saisonnière et écart type du transport zonal pour le SEC et le NECC dans la couche de surface ($\sigma_\theta < 26.25 \text{ kg m}^{-3}$) et pour l'EUC dans la couche intermédiaire ($24 \text{ kg m}^{-3} < \sigma_\theta < 26.25 \text{ kg m}^{-3}$).</i>	172