



Assimilation de données pour la prédiction de paramètres hydrodynamiques et écologiques : cas de la lagune de l'Oder

Laurent Bertino

► **To cite this version:**

Laurent Bertino. Assimilation de données pour la prédiction de paramètres hydrodynamiques et écologiques : cas de la lagune de l'Oder. Mathématiques [math]. École Nationale Supérieure des Mines de Paris, 2001. Français. <tel-00005782>

HAL Id: tel-00005782

<https://pastel.archives-ouvertes.fr/tel-00005782>

Submitted on 16 Jul 2004

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

pour obtenir le grade de
Docteur de l'Ecole des Mines de Paris
Spécialité « **Géostatistique** »

présentée et soutenue publiquement

par

Laurent Bertino

le 7 Décembre 2001

Assimilation de données pour la prédiction de paramètres hydrodynamiques et écologiques : cas de la lagune de l'Oder

Directeur de thèse : Hans Wackernagel

Jury

MM. Philippe	Jamet	Président
Geir	Evensen	Rapporteur externe
Olivier	Talagrand	Rapporteur externe
Eric	Dombrowsky	Examineur
Christian	Lajaunie	Examineur
Hans	Wackernagel	Examineur

Polskie chrząszczy brzmiały w trzcinach
«Les grillons polonais bruissent dans les roseaux»
Phrase la plus difficile à prononcer pour un anglophone
(livre Guinness des records).

Remerciements

Professionnels Ce travail a été effectué dans le cadre du projet européen PIONEER (description disponible sur : <http://pioneer.geogr.ku.dk>) du programme MAST III. Le modèle TRIM3D et les données utilisées ont été fournis par PIONEER, l'argent de PIONEER a lui aussi été bien utilisé : notre studio boulevard de Belleville est calme et ensoleillé et nos guitares nous accompagnent à chaque concert. Le modèle écologique en une dimension et l'algorithme de l'EnKF ont été fournis par Geir Evensen et le NRSC.

Un directeur de thèse est à la fois un guide, un modèle (humain et professionnel), un coach, une mine de suggestions, un bon conseil pour les choix délicats et quelqu'un qui pense à l'avenir de son thésard. De ce point de vue, je recommande Hans Wackernagel en toute hâte à ceux intéressés par la place (unique !) que je laisse vacante. Je tiens aussi à vanter son ouverture culturelle, sa neutralité helvétique envers les vaines querelles, son engagement résolu dans les Grands Combats et son ironie discrète qui fait qu'on apprécie d'autant mieux les congrès houleux et les réunions épiques du projet européen.

Je me suis beaucoup inspiré des travaux de Christian Lajaunie le clairvoyant. J'ai longuement médité chacune de nos conversations pendant lesquelles il a toujours soulevé des points importants et des pistes prometteuses pour poursuivre les recherches, nul doute qu'il en a encore des tonnes en réserve.

Julien Sénégal m'a légué un boulot colossal qu'il a abattu en seulement trois mois de stage. Il est ensuite resté un acteur enthousiaste de ce travail.

Merci tous les membres de PIONNER, surtout à Thomas Wolf pour son expertise sur TRIM3D et sur les données (c'est à dire tout ce dont on a besoin pour vivre en assimilation de données), à Geir Evensen qui a une solution à tous les problèmes et à Hans von Storch, celui qui voit plus haut et plus loin grâce à son point de vue aérien et à sa grande envergure.

Nicolas Bez, Christian Lantuéjoul, Hélène, Didier et David m'ont particulièrement aidé pour la présentation orale. Il faut dire que j'en avais sérieusement besoin. Aldona m'a aussi donné ma première leçon d'acteur à cette occasion. Je remercie aussi ceux qui m'ont aidé à rédiger de ce mémoire : en plus de Christian Lajaunie et Julien que j'ai déjà remerciés, Stéphanie Even et Olivier Talagrand ont porté des remarques judicieuses. Merci à Françoise Poirier pour son œil d'aigle (redoutable).

Je voudrais remercier ceux qui m'ont si souvent sorti du borbier informatique : François et Didier au CG, Dr. Kuhn au GKSS puis ceux dont je connais pas le nom mais qui m'ont aussi autorisé à me loguer à distance et à lancer des calculs énormes sur leurs machines et enfin les utilisateurs qui ont supporté sans broncher de voir sur leur terminal le message suivant :

```
User = bertino, Process = filter_trim3d, CPU = 95% .
```

Enfin, il était bon d'avoir l'avis de Michel Poulin pour initier un matheux comme moi à l'écologie des estuaires. J'ai aussi apprécié l'enseignement des experts du centre de géostatistique. J'espère avoir saisi au

moins une petite fraction de la philosophie si riche, et pourtant unique, qui règne au CG depuis plus de trente ans.

Ce travail a été soutenu en partie par une bourse de temps de calcul du Comité Norvégien des Supercalculateurs (TRU). Une notice du modèle TRIM3D écrite par Hartmut Kapitza a été utilisée pour décrire la physique sous-jacente.

Personnels Voici une bonne occasion de parler de ceux que j'aime, qui m'ont portés et supportés à travers la longue mue que j'ai subie ces trois dernières années (perte de cheveux, voix plus grave, peau plus épaisse, etc.). Je ne saurais trop louer leur mérite et leur patience car il n'est certainement pas de tâche plus éprouvante que de pousser un jeune émotif à penser et agir par lui-même.

Aldona, cette thèse est pour toi ! Bon, je sais que tu ne vas pas la lire en entier, et que des phrases du genre «les méthodes de l'adjoint et des représentants donnent aussi des solutions exactes mais en utilisant le formalisme variationnel et ne sont pas développées» prètent plutôt à rire, pour toi qui aimes tant la poésie. Pour compenser mon manque de style et pour répondre à tes "*dlaczego*" incessants, voici un extrait de poème d'Ibn Hazm (994-1064)¹ :

L'amour que j'ai pour toi persévère en son être Absolu,
sans en rien retrancher ni grandir ;
Il n'a d'autre raison d'être que mon vouloir,
et nul ne lui connaît de cause que lui-même ;
Si une chose porte en soi sa raison d'être,
cette existence-là ne périra jamais ;
Mais quand nous lui trouvons une autre raison qu'elle,
Elle périt avec ce qui semblait sa vie.

Je tiens à signaler que c'est aussi Aldona qui m'a fait ma première proposition d'emploi, celle qui m'a fait le plus plaisir, mais aussi celle que j'ai refusé le plus vite (va comprendre).

Je suis très fier de remercier ici mes parents sans qui je ne serais pas de ce monde (certes). Ils mériteraient un ou deux chapitres de plus pour vanter leur abnégation, leur persévérance et la confiance qu'ils m'ont toujours accordé (à raison ?). Je leur épargnerai cependant cette entorse à leur modestie, elle aussi exemplaire. Je leur dois encore un gros merci pour le pot de thèse, c'était du grand art.

Merci à ma sœur Evelyne qui était encore plus stressée que moi le 7 Décembre au matin, et à tous les amis indispensables, ceux qui ont aidé ma famille à organiser le pot Polono-mauriennais : Roman (*ale autentyczny mebel ! Bardzo wartosciowy !*), JB pour la musique, David et Delphine Lecocq et ceux qui sont venus au spectacle : Agnieszka, Anne, Emmanuel et Sophie et qui se sont retenus de rire de mes herbivores. J'en oublie certainement.

Tout le petit monde des thésards de Fontainebleau, ceux qui m'ont donné le bon exemple (Bertrand, Nicolas) et ceux à qui je donne le mauvais exemple (David, Eric de O., Malek, Julien, Simon, Marta, Sarah, Mireille), puis dans le désordre des bons souvenirs les *twisting sisters* Anne et Delphine Patriarche, Yannick,

¹Traduit de l'arabe par Gabriel Martinez-Gros, Sindbad

Valérie, Léna, Bastien, Eric Gomez, Nico Flipo, Guiraude, Jean-Mi, Ben, Yoann, Thomas, Alan, Valéry (qui nous a déposés à un de nos premiers concerts), Pau, Andres, Arnaud et les skieurs chevronnés qui ont fini par admettre que la Maurienne était la plus belle vallée du monde, en particulier Abdelmajid avec qui j'ai eu le plaisir d'habiter pendant un an. Ensuite tous les anciens des Mines, et les plaisanciers de Houlgate, d'Ochotnica Dolna et de Mazurie, Adam, Edyta et *trzy male swińki* Kuba, Martyna et Jaś, Julien et Sabine, Yannick, Gilles, Ford et Elise, Gé et Gi, Christian et Hyun, Philou et Christelle, Philou et Flo, Gillue, Mikael, Seb Favre, Willock et Aurore, Mascaró, Agnieszka et Patrick, Marta, Maritza, Monika, Willy, Dorota et Nicolas, Guilhem et Franziska, Alicja, Asia, Ola, King (le pionnier de L'AD au CG) et tous les membres des projets Kazimodal, Peinture Fraîche, Fatal Error et Mandale (y'en a qui apparaissent plusieurs fois) que j'ai lâchement abandonnés pour aller jouer avec Aldona («place aux groupes qui tournent», comme disait Ronan : 150 concerts en trois ans c'est vrai que c'est dur de tenir la cadence.) et enfin les amitiés anciennes, les plus ancrées, Geez, Jean-No et Djem, Tata Soph, Manu et Axou (tous de la plus chouette vallée du monde). Mention spéciale pour les CFSG Armando, Yacqueline, Mister Bin, Fabri, Cabrolina, Jose Cuador, Arimon, Oliver, Giovanni, Roger et Rose. Un petit bisou sur les frimousses qui ont vu le jour pendant ma thèse (dans l'ordre approximatif d'apparition) : Isolde, Coralie, Viviane, Emma, Franck, Cippora, Orlane, Bianca, Mathilde, Emilio, Adèle, Nicolas, Robin, Aodréna, Wojtuś (junior), et Yohann.

Je garde enfin une pensée pour Zofia, qui est partie beaucoup trop tôt.

Just a few useful words

General :

à tâtons	by trial and error
améliorer	to improve
bateau	ship
bilan	balance
cas	case
calculer	to compute
célérité	velocity
cependant	but
chenal	channel
conditions aux limites	boundary conditions
courant	flow
données	data
donner	to give
ébauche	rough sketch (forecast)
est	east
état	state
grille	grid
hauteur	elevation
hebdomadaire	weekly
heure	hour
horaire	hourly
incertitude	uncertainty
jour, journée	day
mélanger	to mix
minuit	midnight
mise en œuvre	setup
nœud	node
ouest	west
semaine	week
sous-estimer	underestimate
sur-estimer	overestimate
temps	time
vent	wind

Applied Maths :

ACP	Principal Component Analysis
ADS	Sequential DA
ADV	Variational DA
aléatoire	random
bruit	noise
dérivée	derivative
développement limité	Taylor expansion
...au premier ordre	first order
échantillonnage	sampling
écart-type	Standard deviation
erreur quadratique moyenne	MSE
lissage, lissant	smoothing
moyenne	mean
mode propre	eigenmode
pondérer	weighting
racine carrée	square root
sensibilité	sensitivity
valeurs propres	eigenvalues
vecteurs propres	eigenvectors
vraisemblance	likelihood

Figures :

bas	bottom
bleu	blue
droite	right
gauche	left
échelle	scale
haut	top
ligne continue	solid line
ligne discontinue	dashed lines
milieu	middle
pointillés	dotted line
rouge	red

Table des matières

1	Introduction	1
1.1	La prédiction pour les systèmes estuariens	1
1.1.1	Les buts	1
1.1.2	Les moyens	2
1.2	Historique de l'assimilation de données en océanographie	5
1.2.1	AD variationnelle	5
1.2.2	AD séquentielle	5
1.2.3	Équivalences et comparaisons	6
1.2.4	Domaines d'application	7
1.3	AD et géostatistique : «fertilisation croisée»	8
1.3.1	Apports de l'AD	8
1.3.2	Apports de la géostatistique	9
1.4	Plan du manuscrit	9
I	Théorie	11
2	Le cas général	13
2.1	Modèle d'état et observations	13
2.1.1	Le modèle	14
2.1.2	Les observations	14
2.1.3	Différences de support	15
2.2	Le problème du filtrage en AD	15
2.3	Le problème du lissage	16
2.4	Contraintes fortes ou contraintes faibles ?	16
2.5	La résolution séquentielle dans le cas général	16
2.5.1	Loi de transition	16
2.5.2	Prise en compte des observations	17
3	Le cadre linéaire	19
3.1	Préliminaire sur l'estimation linéaire aux moindres carrés	20
3.1.1	Stationnarité	20

3.1.2	Estimation	20
3.2	Le problème du support	21
3.2.1	Des covariances blocs aux covariances ponctuelles	21
3.2.2	Cas stationnaire	22
3.2.3	Cas non stationnaire	22
3.3	Les méthodes primitives d'assimilation	23
3.4	Le filtre de Kalman	24
3.4.1	Modèle	24
3.4.2	Équations de l'ébauche	24
3.4.3	Correction	25
3.4.4	Remarque à propos de l'initialisation	26
3.5	Stabilité du filtre de Kalman	26
3.5.1	Types de convergence	26
3.5.2	Conditions théoriques de convergence	27
3.5.3	Causes de divergence	27
3.5.4	Effets du biais sur le filtre de Kalman	27
3.6	Estimation d'un biais constant dans le temps	28
3.7	Le filtre de Kalman avec conditions d'universalité	29
3.7.1	Équivalence du filtre de Kalman avec le krigeage	29
3.7.2	Modèle et équations, cas monovarié	29
3.7.3	Analogie avec les autres types de krigeage	31
3.7.4	Cas multivarié	32
3.8	Autres impulsions venant de la géostatistique	32
4	Le cadre non linéaire	33
4.1	Propagation non linéaire, correction linéaire	34
4.1.1	Le filtre de Kalman étendu (EKF)	34
4.1.2	Un exemple de SOS, le filtre de Kalman RRSQRT	36
4.1.3	L'ensemble Kalman filter (EnKF)	39
4.1.4	Comparaison entre le RRSQRT et l'EnKF	41
4.2	Propagation et correction non linéaires	42
4.2.1	Le cas gaussien anamorphosé	42
4.2.2	Méthodes de Monte Carlo alternatives	44
4.2.3	Effet de support dans un cadre non gaussien	47
4.3	Aspects pratiques de l'ADS multivarié	47
4.3.1	Normalisation	47
4.3.2	Corrections successives ou jointes ?	47
II	Applications	49
5	Application à un modèle écologique simple	51
5.1	Le modèle	51

5.1.1	Interactions	51
5.1.2	Conditions aux limites	52
5.1.3	<i>Run</i> de référence \mathbf{z}_n^t	52
5.1.4	Stabilité du modèle	53
5.2	Mise en œuvre de l'ADS	55
5.2.1	Mesures synthétiques	55
5.2.2	Statistiques d'erreur	55
5.3	EnKF : variations de la taille de l'ensemble	56
5.4	Comparaison de l'EnKF au RRSQRT KF	58
5.4.1	Précautions d'emploi	58
5.4.2	Stabilité	58
5.4.3	Résultats et discussion	59
5.4.4	Conclusion de la comparaison	62
5.5	EnKF en version lognormale	62
5.5.1	Limites du modèle gaussien	62
5.5.2	Des variables lognormales ?	62
5.5.3	Mise en œuvre	63
5.5.4	Comparaison de l'EnKF gaussien contre lognormal	65
6	Cas de l'hydrodynamique de la lagune de l'Oder	69
6.1	Généralités	69
6.1.1	Données géographiques	69
6.1.2	Le projet PIONEER	71
6.2	Le modèle TRIM3D	71
6.2.1	Équations	71
6.2.2	Discretisation	72
6.2.3	Implémentation	75
6.2.4	Conditions initiales et conditions aux limites	76
6.2.5	Sensibilité du modèle	76
6.3	Les données de hauteur d'eau	77
6.3.1	Dispositif expérimental	77
6.3.2	Analyse des données	77
6.3.3	Effet de support	80
6.4	Mise en œuvre du filtre de Kalman	81
6.4.1	Compatibilité du filtre de Kalman avec TRIM3D	81
6.4.2	Statistiques des erreurs	82
6.4.3	Mise en œuvre du filtre de Kalman RRSQRT	85
6.5	Résultats du filtre de Kalman RRSQRT : un <i>run</i> de validation	87
6.5.1	Aspects numériques	87
6.5.2	Performances de l'initialisation	87
6.5.3	Amélioration des résultats du modèle	88
6.5.4	Comparaison avec un estimateur trivial	89
6.5.5	Caractéristiques spatiales	89

7 Assimilation jointe de hauteurs d'eau et de salinité dans la Lagune de l'Oder	93
7.1 Généralités	93
7.1.1 Salinité et biologie de la lagune de l'Oder	93
7.1.2 La salinité : exemple de traceur passif	94
7.1.3 Le transport du sel dans TRIM3D	94
7.1.4 Sensibilité du modèle	95
7.2 Analyse des données de la campagne SYNOPTA	96
7.3 Mise en œuvre de l'EnKF	99
7.3.1 Adaptation des conditions aux limites	99
7.3.2 L'erreur initiale	100
7.3.3 Pourquoi l'EnKF ?	101
7.3.4 L'erreur de modèle	101
7.3.5 L'erreur d'observation	102
7.4 Résultats	103
7.4.1 Un <i>run</i> d'ensemble	103
7.4.2 Influence du paramètre de taille de l'ensemble	105
7.4.3 Gradients verticaux de salinité	105
7.4.4 Performances et temps de calcul	106
7.4.5 Structures non stationnaires simples et croisées	107
7.4.6 Comparaison avec les résultats du modèle	111
7.4.7 Validation avec des données indépendantes	114
7.5 Conclusion	115
Conclusions	119
Conclusion	119
Perspectives	120
Bibliographie	123

Chapitre 1

Introduction

1.1 La prédiction pour les systèmes estuariens

1.1.1 Les buts

Les systèmes estuariens en tant qu'interfaces d'échanges entre les domaines continentaux et marins jouent un rôle clef dans l'équilibre écologique des régions avoisinantes (océans, rivières et zones côtières). On a pu en saisir l'importance à travers les problèmes de raréfaction des stocks de poissons et des oiseaux migrateurs.

La richesse écologique des estuaires est en effet menacée par les grandes quantités de pollution provenant des activités urbaines et agricoles du bassin du fleuve. Stålnacke (1996) montre que dans le cas de la mer Baltique, les rivières sont les sources principales d'azote et de phosphore, loin devant les apports atmosphériques ou les rejets directs des industries et des villes côtières. L'Oder, dont il sera question plus en détails dans les applications, apporte environ 90 000 tonnes d'azote et 5 000 tonnes de phosphore par an dans son estuaire. La majorité de ces nutriments est, directement ou non, issue des activités humaines.

Le processus tenu par les biologistes pour principal responsable des dommages subis par les écosystèmes estuariens est l'eutrophisation qui résulte d'un afflux excessif en nutriments (phosphates, ammonium, nitrates et nitrites). Des algues microscopiques (phytoplancton) se nourrissent de ces nutriments et prolifèrent, ce qui nuit à la qualité des eaux côtières et leur donne une odeur pestilentielle. L'eutrophisation fait aussi apparaître des algues toxiques en Bretagne depuis deux ans. Enfin, quand ces algues meurent, la minéralisation de leur matière organique provoque une baisse de la concentration en oxygène dans l'eau, asphyxiant la flore et la faune.

L'eutrophisation a été observée et étudiée depuis la seconde guerre mondiale dans des lacs, puis au cours des années 70 dans des eaux côtières et ce n'est que dans les années 80 qu'il a été reconnu que de grandes étendues d'eau telles que la mer Baltique pouvaient elles aussi être menacées (Stålnacke, 1996). On constate actuellement que dans la plupart des estuaires des rivières d'Europe, une zone d'anoxie (zone «morte» par absence d'oxygène) se forme périodiquement et bloque le passage des poissons migrateurs.

Mais au delà des aspects écologiques, les estuaires jouent aussi un grand rôle socio-économique mêlant de multiples acteurs aux intérêts divergents. Parmi eux les moyens de transport fluviaux nécessitent souvent des aménagements (chenalisations, approfondissements . . .), les activités portuaires et industrielles cherchent des zones d'extension à proximité de l'estuaire, le tourisme demande lui aussi des aménagements de loisir et des paysages préservés¹ et on doit bien entendu compter avec les pêcheurs. L'arbitrage politique entre les intérêts de ces acteurs est un art délicat qui repose essentiellement sur la prévision des retombées d'une décision sur chacun des acteurs. Les prévisions de routine peuvent donc être nécessaires aux activités habituelles de l'estuaire (indice de qualité d'eau pour la baignade, détermination des courants pour la navigation . . .) tout comme aux situations exceptionnelles (projets d'aménagement, catastrophes naturelles, pollutions . . .)

Prenons l'exemple d'un incident ayant lieu dans une usine de produits chimiques implantée au bord d'une rivière et provoquant une pollution dangereuse pour les riverains en aval. Avant d'agir, les décideurs devront mesurer deux risques distincts : d'une part les risques de contamination de la population, d'autre part celui de prendre une mesure d'urgence pour un incident moins grave que prévu, ce qui peut parfois s'avérer aussi dangereux. Or les mesures de protection à mettre en œuvre dépendent essentiellement de la concentration de polluants qui parviendra en aval, celle-ci étant largement conditionnée par des transformations chimiques qu'ils auront subies, les conditions de transport dans le milieu aqueux et des divers effets écologiques pouvant se produire. En milieu estuarien, on doit aussi tenir compte des différents modes de transport (dissous et particulaire) et de l'effet de marée. L'information nécessaire concerne donc plusieurs variables, elle doit être quantitative et assortie d'un contenu probabiliste (au moins un intervalle d'erreur) pour prendre en compte les deux types de risque. Devant tant de complexité, on a donc recours à des simulations numériques des différentes variables afin d'aider la prise de décision en complétant l'avis des experts spécialistes du terrain.

1.1.2 Les moyens

La tendance, ces dernières décennies, est aux prédictions numériques qui utilisent deux types d'information : d'une part un modèle numérique bi- ou tridimensionnel basé sur les équations décrivant la physique, la chimie et l'écologie du système, et d'autre part des observations *in situ* utilisées intensivement *via* une procédure d'assimilation de données (AD) afin de corriger les prédictions du modèle, d'estimer l'état du système

¹Le nettoyage d'une plage souillée par les algues représente un coût très important pour une commune côtière.

et d'évaluer la précision de cette estimation. Lorsque cette estimation est faite dans le futur, on parlera alors de «prédictions». L'AD peut aussi servir des buts plus larges que la prédiction des systèmes physiques. On peut en effet utiliser ses résultats, plus précis que ceux du modèle, pour analyser les processus dynamiques mis en jeu ou profiter du cadre stochastique et chercher les zones d'influence de chaque observation. On peut ainsi qualifier la pertinence de celles-ci, mais aussi améliorer la stratégie d'échantillonnage en déterminant le meilleur emplacement des instruments de mesure, comme cela se pratique en géostatistique (Vénard *et al.*, 2001).

1.1.2.1 Les données océanographiques

Comme dans tous les domaines scientifiques ou industriels les données océanographiques sont coûteuses, donc rares, mais d'une grande diversité de nature et d'utilité.

- Les bouées dérivantes et les stations fixes munies d'appareils de mesure des propriétés physiques et chimiques d'une colonne d'eau entière (vitesse du courant, conductivité, pH, température, turbidité de l'eau, parfois aussi concentrations en oxygène et ammonium dissous) donnent une information en continu et peuvent être utilisées dans un système de prédiction opérationnel si ces données sont transmises au centre d'analyse en temps réel. C'est le cas pour les données des stations fixes dans la lagune de l'Oder pour le projet PIONEER, sur lequel j'aurai l'occasion de revenir dans les parties applications, ou par exemple celles de la bouée flottante dans le projet BTM dans les Bermudes (voir la page de l'Université de Californie, Santa Barbara <http://www.opl.ucsb.edu/>). L'installation de ces stations de mesure et l'entretien des sondes sont fort coûteux et on ne peut disposer en général de plus d'une poignée. Par contre, leurs données sont temporellement très riches et il est possible d'avoir des informations en profondeur, ce qui est rarement le cas des données de télédétection.
- Les campagnes en bateau donnent aussi une information utile et offrent un plus grand choix de variables parmi celles qui ne peuvent pas être mesurées par des sondes en continu. A. R. Robinson a montré la possibilité de donner des prédictions à bord même du navire de recherche lors de la collecte des données. Cependant, dans le cas de la lagune de l'Oder, ces campagnes en bateau ne sont qu'exceptionnelles et ne sont donc pas directement utilisables pour la surveillance au jour le jour de l'estuaire. C'est pour cela que ces données ne sont pas assimilées dans les applications de cette thèse malgré leur intérêt pour l'AD. En effet un transept donne un jeu de données à peu près synoptique² qui couvre à la fois les directions horizontales et verticales. Il est donc préférable d'utiliser ces données pour valider l'assimilation ou pour estimer les covariances spatiales *via* l'hypothèse de stationnarité. Dans le cas du projet Seine Aval les campagnes en bateau sont faites tous les jours, ce qui les rend plus intéressantes.
- Enfin les images satellites sont devenues le centre d'intérêt de la majeure partie de la communauté océanographique depuis que les données d'altimétrie de Topex-Poseidon et de ERS-2 sont disponibles. Ces données ont l'avantage de couvrir de grandes étendues du domaine d'étude et elles sont

² Voir le chapitre synopticit  du rapport PIONEER sur les donn es de salinit  de l'Oderhaff (Bertino et Wackernagel, 2000).

assimilées dans des modèles hydrodynamiques. Les potentialités des images satellites sont continuellement explorées, par exemple Zavialov a montré au 33ième colloque d'océanographie de Liège en Mai 2001 comment calculer les courants de surface à partir d'images satellites NCEP en eaux peu profondes. Leur exploitation en modélisation écologique est aussi récente, les données de température et de couleur de la surface de l'océan (notamment les SeaWiFS) sont assimilées dans les modèles biologiques puisque la température de l'eau influence les processus biologiques et la couleur de l'eau donne une indication sur la production primaire dans les océans (activité du phytoplancton). Cependant, des informations en profondeur font toujours cruellement défaut en biologie marine et les problèmes numériques posés par l'AD de millions de données conduisent à des approximations drastiques des calculs théoriques.

Dans les cas étudiés dans cette thèse, nous n'avons considéré que des données spatialement très rares et les solutions numériques proposées ne sont pas toujours adaptées à des applications aux données satellites.

1.1.2.2 Les modèles océanographiques

Une grande diversité de modèles est actuellement disponible sur le marché, nous passerons sous silence les modèles de vagues, qui ne sont pas prises en compte dans nos cas. Pour les modèles hydrodynamiques, on notera le succès du modèle MIKE du Danish Hydraulic Institute pour les eaux peu profondes et ceux du *Princeton Ocean Model* POM (Blumberg et Mellor, 1987) et du *Miami Isopycnic Coordinate Ocean Model* MICOM (Bleck, 1978) pour les zones côtières et les océans.

Les modèles de transport de sédiments et les modèles biologiques présentent des difficultés bien supérieures car le nombre de paramètres inconnus et de paramètres régionalisés (variables dans l'espace) est beaucoup plus élevé. Pour les modèles biologiques, il faut aussi ajouter que les réactions chimiques et les croissances biologiques ont des dynamiques explosives et donc très sensibles aux conditions initiales. Les modèles écologiques généralement utilisés en AD comportent peu de variables (ou «compartiments») : des nutriments (ammonium, nitrates, nitrites et phosphates), un ou deux types de phytoplancton, le zooplankton et la chlorophylle qui est la variable la plus facilement observable car liée à la couleur de l'eau. Des compartiments «oxygène» et «gaz carbonique dissous» sont occasionnellement ajoutés suivant les modèles.

Si les processus qui lient ces variables sont assez bien connus, par contre les paramètres des relations qui les décrivent (constantes de croissance, de mortalité, de réaction à la lumière . . .) sont difficiles à estimer, et de surcroît peuvent être significativement variables dans le temps et dans l'espace³.

³Evans et Parslow (1985) disent à propos des modèles écologiques «*as their complexity grows, they can behave as mysteriously as the real world*».

1.2 Historique de l'assimilation de données en océanographie

Petite sœur de l'AD météorologique qui a fait ses preuves en prédiction opérationnelle, l'AD océanographique utilise depuis les années 80 les mêmes méthodologies (analyse objective, filtre de Kalman, méthode de l'adjoint) et doit aussi surmonter deux problèmes : la non linéarité des équations du modèle et les grandes dimensions du vecteur d'état, les applications nécessitant en effet une discrétisation du système sur une grille assez fine pour que le modèle puisse reproduire des phénomènes de moyenne échelle qui contribuent au transport des matières solubles. Ceci affecte les temps de calcul dont la maîtrise est un problème crucial en AD océanographique opérationnelle puisque les résultats doivent parvenir aux experts avant que les événements ne deviennent incontrôlables. Verlaan (1998) cite le cas des lames de fond qui menacent les digues néerlandaises par temps de tempête.

On regroupe les techniques d'AD en deux catégories, ADS (séquentielle) et ADV (variationnelle).

1.2.1 AD variationnelle

La première catégorie est issue de la théorie du contrôle optimal où l'on cherche la solution qui minimise une fonction de coût prédéfinie. Elle comporte la méthode des représentants (Bennett *et al.*, 1996) et surtout celle de l'adjoint (Le Dimet et Talagrand, 1986) qui a été adoptée à la fin des années 90 dans de nombreux centres de prévisions opérationnelles en météorologie. Cette dernière est aussi appliquée en océanographie (Luong *et al.*, 1998 ; Nechaev et Yaremchuk, 1994 ; Lellouche *et al.*, 1998) mais plus rarement. On les regroupe sous la terminologie d'assimilation de données variationnelle (ADV). Ces méthodes étant plus complexes à mettre en œuvre dans les cas d'étude non linéaires (notamment le calcul du modèle adjoint quand le modèle direct est non linéaire), elles ne seront pas détaillées par la suite.

1.2.2 AD séquentielle

L'autre catégorie regroupe les techniques d'assimilation de données séquentielles (ADS) dérivées de la théorie de l'estimation statistique : l'interpolation optimale (Daley, 1991), exposée brièvement en section 3.3, qui est la plus rudimentaire mais aussi la seule qui soit largement utilisée opérationnellement (Robinson *et al.*, 1996), et le filtre de Kalman. Ce dernier correspond aussi à une solution optimale pour les systèmes linéaires. Il a été étendu en linéarisant les dynamiques (*Extended Kalman Filter*, EKF) aux systèmes non linéaires par Jazwinski (1970) mais reste encore inapplicable aux systèmes régionalisés à cause de son encombrement numérique. Une myriade de schémas sous-optimaux ont donc été déduits de l'EKF et appliqués à des systèmes météorologiques (Dee, 1991), et océanographiques (Cane *et al.*, 1996 ; Cohn et Todling, 1996 ; Fukumori, 1995). J'écarterai par la suite les méthodes utilisant une décomposition explicite du modèle physique (Cane *et al.*, 1996) afin de pouvoir traiter tous les modèles en tant que «boîtes noires».

Les méthodes les plus prometteuses pour la prédiction opérationnelle sont de ce point de vue le RRSQRT de Verlaan et Heemink (1997) appliqué aux modèles physiques des zones marines et côtières (Cañizares *et al.*, 2001 ; Verlaan et Heemink, 1997 ; Wolf *et al.*, 2001), les filtres SEEK et SEIK de Pham *et al.* (1997) appliqués au modèle océanique physique MICOM (Verron *et al.*, 1998) et à un modèle écologique de l'océan Atlantique (Carmillet *et al.*, 2001) et enfin la méthode de l'université de Harvard développée par Lermusiaux et Robinson (1999) et appliquée entre autres à l'étude physique et écologique de la baie du Massachussets. Certaines de ces méthodes sont utilisées pour des prédictions opérationnelles mais dans des versions «en régime permanent» (*steady state*) plus rapides numériquement mais sans évolution dynamique de l'erreur. Ces versions «en régime permanent» du filtre de Kalman sont donc plus proches de l'interpolation optimale dans leur principe. Une méthode apparentée au filtre de Kalman se détache de l'EKF, il s'agit de l'*Ensemble Kalman filter* (EnKF) d'Evensen (1994) qui procède par simulation de Monte Carlo tout en appliquant une étape de correction linéaire. Il a été appliqué à des modèles physiques (Evensen et van Leeuwen, 1996 ; Haugen et Evensen, 2001 ; Cañizares, 1999 ; Echevin *et al.*, 2000) et écologiques (Eknes et Evensen, 2001 ; Breuillin *et al.*, 2000 ; Natvik et Evensen, 2001). L'*Unscented Kalman filter* de Julier *et al.* (1995) semble obéir au même principe que l'EnKF pour des applications en guidage automatique de véhicules.

1.2.3 Équivalences et comparaisons

En théorie, méthodes variationnelles et séquentielles sont équivalentes dans le cadre linéaire ; cependant dans les cas réels non linéaires les techniques citées ci-dessus se compliquent considérablement et cette équivalence n'est plus garantie⁴.

En se limitant à l'ADS, quelques comparaisons entre filtres de Kalman ont été menées dans des applications représentatives des problèmes rencontrés en océanographie. Verlaan et Heemink (1999) a comparé l'EnKF à deux variantes du filtre RRSQRT et au filtre SEIK à l'aide des équations classiques de Lorenz⁵. Faisant varier le degré de non linéarité du système d'ADS, ils ont trouvé préférable l'utilisation des filtres dérivés de l'EKF (filtre SEIK et RRSQRT) pour les systèmes faiblement non linéaires car ils sont plus efficaces numériquement que l'EnKF, mais pour les systèmes fortement non linéaires ils recommandent l'utilisation de l'EnKF.

⁴Michael Ghil dira à ce propos lors de la conférence de Québec de Mai 1999 «*God is in everything, but the Devil is in the details*».

⁵Les équations de Lorenz décrivent l'évolution d'une particule dans un espace à deux puits de potentiel dont le comportement chaotique est censé être représentatif des instabilités et des non linéarités des systèmes météorologiques et océanographiques. Il est cependant discutable que toutes les non linéarités se ressemblent.

Les performances de l'EnKF et du filtre de Kalman RRSQRT ont été comparées sur un système hydrodynamique fictif à deux dimensions (Cañizares, 1999) et ont été trouvées équivalentes tant du point de vue de leur rapidité que de leur précision. Une comparaison des mêmes filtres de Kalman sur un modèle écologique simple est discutée dans Bertino *et al.* (2002) et rapportée dans le chapitre 5, les deux ne montrent pas de différence évidente.

L'EnKF a aussi été comparé au filtre SEIK dans le cas d'école des équations de Lorenz (Pham, 2000), concluant à la supériorité du filtre SEIK. Mais Pham (2000) les utilise les deux méthodes avec moins de 10 réalisations (ce qui est extrêmement peu pour des méthodes de Monte Carlo) et introduit de nombreuses complications des algorithmes ce qui rend les résultats moins éloquents. Les filtres SEIK et EnKF sont aussi comparés dans le cadre du projet DIADEM pour traiter un cas réel s'intéressant à l'hydrodynamique et à l'écologie de l'océan Atlantique nord (Evensen *et al.*, 2001), dans ce dernier cas les deux méthodes ne semblent pas donner des résultats significativement différents.

1.2.4 Domaines d'application

Une difficulté propre à la modélisation de l'hydrodynamique des océans et des zones côtières est la grande durée des temps d'initialisation (durée pendant laquelle les résultats sont dépendants des conditions initiales). Ces systèmes ont «la mémoire longue» bien plus que les systèmes météorologiques et exigent une connaissance précise des conditions initiales et des conditions aux limites alors que les mesures font défaut. De ce point de vue, le cas d'une lagune comme celle de l'Oder présente l'avantage d'avoir peu de frontières ouvertes et donc d'être plus facilement contrôlable. A notre connaissance, l'AD a été seulement appliquée à l'étang de Berre (Leredde *et al.*, 2000), mais pas à d'autres systèmes d'estuaires ou de lagunes.

Si l'efficacité de l'ADS est avérée pour l'hydrodynamique, par contre elle doit encore faire ses preuves en transport de sédiments et en écologie. Pour ces problèmes plus complexes, des approches alternatives font leur apparition pour la détermination des paramètres inconnus : réseaux neuronaux, recuit simulé, etc. On notera aussi l'émergence d'algorithmes génétiques (Athias *et al.*, 2001). En dehors de la météorologie et de l'océanographie, d'autres domaines d'applications de l'ADS sont la chimie atmosphérique (van Loon et Heemink, 1997 ; Segers *et al.*, 2000) et l'hydrogéologie. Depuis peu, l'ADS s'attaque à des problèmes plus exotiques : l'halieutique (Robinson et le groupe AFMIS, 2001 ; Grønnevik et Evensen, 2001) et l'épidémiologie (Bui, 2001).

1.3 AD et géostatistique : «fertilisation croisée»

1.3.1 Apports de l'AD

La géostatistique spatio-temporelle est une discipline en perpétuel développement abordée par Matheron (1970). On peut en trouver un *reviewing* dans Kyriakidis et Journel (1999). Cependant il est rare que l'on puisse appliquer les théories géostatistiques développées pour des systèmes statiques à l'espace-temps en quatre dimensions (3 dimensions spatiales plus l'axe des temps). En effet les anisotropies entre directions spatiales et temporelles sont bien plus complexes que celles que l'on trouve entre deux directions spatiales. Kyriakidis et Journel (1999) suggèrent une séparation des variables en une somme, ou un produit, d'une fonction de l'espace et d'une fonction du temps. Cependant cette séparation n'est possible que quand l'évolution temporelle de la variable est linéaire, l'AD permet donc d'atteindre une classe de problèmes plus large. Enfin la géostatistique se trouve aussi impuissante devant la nécessité de fournir des lois spatiales pour des variables soumises à des contraintes physiques et elle se cantonne en général aux systèmes statiques.

1.3.1.1 L'introduction de connaissances physiques du système

L'AD, en introduisant un modèle physique dans un cadre stochastique, apporte un plus par rapport au krigeage spatio-temporel, où seules des contraintes physiques simples (par exemple des contraintes linéaires (Lajaunie, 1997) ou géométriques) peuvent être prises en compte. Les connaissances des processus dynamiques complexes et non linéaires peuvent être combinées aux observations directes des données.

On trouve parfois sous la dénomination filtre de Kalman des algorithmes utilisant un modèle statistique (modèles autorégressifs ou moyennes mobiles) pour son évolution temporelle (Wikle et Cressie, 1999 ; Mardia *et al.*, 1998 ; Van Geer et Zuur, 1997). Or les modèles autorégressifs représentent un cas particulier du schéma exponentiel dans le temps et ces algorithmes sont donc équivalents à un krigeage spatio-temporel avec un modèle temporel exponentiel. L'emploi de modèles autorégressifs peut toutefois être pratique car ils sont faciles à mettre en œuvre.

1.3.1.2 Des covariances non stationnaires

En contraignant l'évolution d'un système stochastique par un système d'équations dynamiques quelconque, l'AD fournit aussi *a posteriori* les moments des variables aléatoires étudiées (en général seulement les deux premiers) et en particulier les covariances spatiales non stationnaires qui vérifient les contraintes physiques du système. Puisqu'il n'est pas facile de les calculer directement pour des équations différentielles non linéaires quelconques, l'ADS apporte une réponse à une question d'un grand intérêt dont le traitement théorique n'est pas simple (Sampson et Guttorp, 1992 ; Guillot, Senoussi, et Monestiez, 2001).

1.3.2 Apports de la géostatistique

L'interprétation de l'ADS en termes de géostatistique permet de mettre en évidence l'hypothèse de stationnarité qui est l'hypothèse fondamentale des méthodes géostatistiques. L'ADS a toujours recours à une hypothèse stationnaire pour sa mise en œuvre et on doit pouvoir observer cette stationnarité au moins sur un sous-ensemble du système étudié. La géostatistique apporte une longue expérience des différents modèles de covariances spatiales que l'ADS utilise à la fois pour modéliser l'erreur commise sur l'état initial et celle commise à chaque pas de temps par le modèle physique.

En outre, partant de l'équivalence entre le filtre de Kalman et le krigeage à moyenne connue, on peut se demander à quoi correspondent les extensions du Krigeage Simple (KS) avec d'autres conditions d'universalité (Krigeage Ordinaire (KO), Krigeage en Dérive Externe (KDE) ou Krigeage Universel (KU)), mais aussi les extensions non stationnaires ou non linéaires appliquées dans le cadre du filtre de Kalman. Ces approches sont-elles utiles ?

L'expérience accumulée depuis les débuts de la géostatistique en simulations conditionnelles et en optimisation de la répartition des mesures peut aussi être mise à profit dans le cadre de l'AD. Ces deux derniers points ne sont pas abordés dans cette thèse mais peuvent faire l'objet de développements futurs.

Enfin, les notations continues permettent de prendre conscience des problèmes de support puisque lors de l'AD on doit inévitablement comparer les résultats du modèle aux mesures prises *in situ* et que les deux sont généralement prises sur des supports différents.

1.4 Plan du manuscrit

Ce manuscrit se divise en une partie théorique et une partie applications. Puisque les applications ne portent que sur des problèmes de prédiction, on omettra la question du lissage (estimation d'un temps passé utilisant des données à venir) au profit du filtrage. Le chapitre 2 décrit l'AD dans un cadre général puis on passe au cas linéaire dans le chapitre 3, axé sur les méthodes d'ADS et le filtre de Kalman.

Des extensions inspirées de la géostatistique sont montrées dans le cadre linéaire avant de passer au cas des modèles physiques non linéaires dans le chapitre 4 où deux filtres de Kalman sont détaillés (le RRSQRT schéma sous-optimal de l'EKF et le filtre de Kalman d'ensemble (EnKF)). Ces méthodes sont répandues en ADS océanographique. Quelques perspectives pour l'estimation non linéaire sont ensuite données.

La seconde partie contient trois cas d'étude. Tout d'abord un cas fictif de modèle écologique à trois variables et une dimension géographique (chapitre 5). Il ne concerne ni un estuaire ni une lagune mais il permet de tester l'efficacité des deux méthodes dans un cadre non linéaire, régionalisé (1-D) et multivariable. Le deuxième cas d'étude du chapitre 6 est un cas réel de modèle hydrodynamique de la lagune de l'Oder

gouverné par l'influx de la rivière, les vents et les interactions avec la mer Baltique. Les mesures assimilées sont des hauteurs d'eau prises dans les stations fixes en activité durant la période de la crue d'août 1997. Le modèle utilisé est TRIM3D. On suppose que la méconnaissance des conditions limites est source d'erreurs que l'on modélise comme des fonctions aléatoires. Leurs caractéristiques statistiques spatiales sont déduites de l'analyse structurale temporelle grâce à une équation d'onde. La méthode d'ADS utilisée est le filtre de Kalman RRSQRT. On valide les résultats en retirant deux des cinq stations de mesure de l'ADS. Dans ces deux stations de validation, le système d'ADS ainsi construit donne des estimations des hauteurs d'eau plus précises que celles du modèle physique seul. Les temps de calculs sont raisonnables en vue d'applications opérationnelles. Le troisième cas d'étude (chapitre 7) est une suite logique du deuxième puisqu'il s'agit du transport de sel dans la lagune de l'Oder. TRIM3D assure la modélisation de ce transport. On assimile des mesures de salinité de deux stations conjointement avec les mesures de hauteur d'eau du cas précédent. On suppose également que les erreurs sur le transport de salinité proviennent des erreurs sur les conditions limites en plus des erreurs sur l'hydrodynamique. L'étude structurale faite sur des données d'une campagne en bateau effectuée pendant la période d'étude est utilisée pour donner les caractéristiques statistiques des erreurs sur l'état initial de salinité et sur les conditions aux limites. Ce dernier cas d'étude présente un contre-exemple où les caractéristiques du modèle font que l'ADS perd son efficacité au cours du temps.

Première partie

Théorie

Chapitre 2

Le cas général

Sommaire

Comment traiter le problème de l'assimilation de données dans un cadre probabiliste. Equation de Fokker-Planck pour l'évolution temporelle des fonctions de densité de probabilité, équation de Bayes pour leur intégrer les données. Pourquoi le problème du filtrage n'est-il pas directement soluble pour des applications en océanographie ?

2.1 Modèle d'état et observations

Le système de notations est propre à ce manuscrit. Par rapport aux autres systèmes de notation ayant cours actuellement dans le monde de l'AD, il est plus intuitif au sens du géostatisticien et j'espère qu'il n'est pas un obstacle pour les spécialistes de l'AD. \mathbf{Z} et \mathbf{Y} sont des variables aléatoires alors que \mathbf{z} et \mathbf{y} en sont des réalisations.

Le problème de l'AD est similaire à celui d'une chaîne de Markov à état caché. On appelle **l'état du système** au temps n $\mathbf{Z}_n(\mathbf{x})$ une fonction aléatoire vectorielle de k variables définies sur l'espace. Dans les applications l'état pourra regrouper différentes variables de l'estuaire, aussi bien physiques (hauteurs d'eau, courants) que biologiques (concentrations en nutriments, phytoplancton ...). On suppose connue la loi initiale $\mathcal{L}(\mathbf{Z}_0)$. On notera $\mathbf{Z}_{1:n}$ la suite $(\mathbf{Z}_1, \dots, \mathbf{Z}_n)$.

2.1.1 Le modèle

La première information dont nous disposons sur cet état est **modèle dynamique \mathbf{f}** qui assure les transitions temporelles (appelées couramment «évolutions») entre deux états successifs. Il regroupe l'ensemble des connaissances dont nous disposons sur la dynamique du système et qui sont applicables dans un code de calcul. Il est cependant entaché d'une erreur ε_n^m à chaque pas de temps, liée aussi bien à la conception du modèle (description imparfaite des processus) qu'à sa mise en œuvre pratique (calage des paramètres, résolution numérique, erreurs d'arrondi ...). On considère une forme de \mathbf{Z}_n discrétisée dans l'espace, ce qui correspond aux applications numériques. Les composantes du vecteur \mathbf{Z}_n sont alors les valeurs des différentes variables d'état en chaque nœud de grille du modèle. Ceci représente $N = k \times l$ paramètres, où l est le nombre de nœuds de grille du modèle. Dans les applications en 2 ou 3 dimensions, la taille du vecteur d'état N atteint facilement 10^5 , ce qui constitue un problème numérique important dans la suite. L'évolution de \mathbf{Z}_n est donnée par les équations du modèle physique

$$\mathbf{Z}_n = \mathbf{f}_n(\mathbf{Z}_{n-1}) + \varepsilon_n^m, \quad (2.1)$$

\mathbf{f}_n est non linéaire dans le cas général, il sera noté F_n si on le suppose linéaire.

On suppose par la suite que l'évolution de l'état est un processus markovien d'ordre 1 (\mathbf{Z}_n ne dépend du passé que par l'intermédiaire de \mathbf{Z}_{n-1}), ce qui semble raisonnable pour les systèmes océanographiques étudiés. La loi de transition associée est $\mathcal{L}(\mathbf{Z}_n | \mathbf{Z}_{n-1})$ on suppose pour simplifier qu'elle admet la densité $\phi(\mathbf{z}_n | \mathbf{z}_{n-1})$.

2.1.2 Les observations

La seconde information disponible est un ensemble d'**observations \mathbf{y}_n** supposées correspondre aux temps discrets n du modèle, mais elles sont aussi entachées d'une erreur aléatoire ε_n^o . \mathbf{Y}_n est un vecteur de m observations liées au vecteur d'état par la fonction d'observation notée \mathbf{h}

$$\mathbf{Y}_n = \mathbf{h}(\mathbf{Z}_n) + \varepsilon_n^o, \quad (2.2)$$

ou H si elle est linéaire. En général les variables observées sont les mêmes que les variables d'état mais ceci n'est pas nécessaire dans le formalisme suivant, il suffit qu'il existe une fonction d'observation de \mathbf{Z}_n .

On suppose que les observations \mathbf{Y}_n sont indépendantes conditionnellement à l'état \mathbf{Z}_n , les lois d'observation sont donc

$$\mathcal{L}(\mathbf{Y}_n | \mathbf{Z}_{1:n}) = \mathcal{L}(\mathbf{Y}_n | \mathbf{Z}_n). \quad (2.3)$$

Cette hypothèse aussi est justifiée compte tenu de la nature des observations utilisées dans les applications. On suppose aussi l'existence de la densité $\phi(\mathbf{y}_n | \mathbf{z}_n)$.

2.1.3 Différences de support

Pour anticiper sur les **problèmes de support**, on remarque que les composantes des vecteurs \mathbf{Z}_n et \mathbf{Y}_n sont des régularisées des fonctions continues de l'espace $\mathbf{Z}_n(\mathbf{x})$ et $\mathbf{Y}_n(\mathbf{x})$ respectivement sur le volume d'un nœud de grille V et le volume, nettement plus faible, d'une cellule de mesure v :

$$\mathbf{Z}_n = \int_V \mathbf{Z}_n(\mathbf{x}) d\mathbf{x} \quad (2.4)$$

$$\mathbf{Y}_n = \int_v \mathbf{Y}_n(\mathbf{x}) d\mathbf{x}. \quad (2.5)$$

S'il est possible de moyenniser des données pour passer d'un petit support à un grand support, en revanche on ne peut pas affiner la résolution d'une information et la transformation \mathbf{h} donne en général un résultat exprimé sur le grand support V . Cette différence de support représente une difficulté pour comparer les grandeurs du modèle aux mesures prises *in situ*, on y reviendra dans la partie 3.2. On n'évoquera pas dans la partie théorique les différences de support temporel (quand le pas de temps du modèle est différent du temps d'acquisition des observations) car le problème du changement de support spatio-temporel deviendrait encore plus compliqué dans le cadre d'une loi d'évolution non linéaire.

2.2 Le problème du filtrage en AD

L'AD consiste à prendre le meilleur parti des deux informations dont nous disposons : une information de nature dynamique, le modèle, et une information statique, les données. La question du filtrage peut alors se poser ainsi : «Étant donné un jeu d'équations différentielles décrivant la dynamique d'un système et un ensemble d'observations sur ce système, quel est l'état du système le plus vraisemblable à un temps présent ou futur ?»

En termes probabilistes la définition précédente signifie qu'on cherche à chaque pas de temps n le maximum de vraisemblance de la densité conditionnelle aux données passées $\phi(\mathbf{z}_n | \mathbf{y}_{1:n})$, en se donnant la loi d'évolution $\mathcal{L}(\mathbf{Z}_{n+1} | \mathbf{Z}_n)$ (ce qui revient à donner celle de ε_n^m) et celle des observations $\mathcal{L}(\mathbf{Y}_n | \mathbf{Z}_n) = \mathcal{L}(\varepsilon_n^o)$.

Les nombreuses méthodes d'assimilation de données existantes représentent des solutions sous-optimales de ce problème général. J'utiliserai la terminologie de l'AD selon laquelle les estimations dans le futur sont appelées «prédictions».

2.3 Le problème du lissage

L'estimation d'un état passé connaissant des données ultérieures, ce qui revient à chercher la loi :

$$\mathcal{L}(\mathbf{Z}_{1:n} | \mathbf{Y}_{1:n}) \quad (2.6)$$

constitue un problème de lissage, très utile pour l'analyse des processus physiques et dont le problème de filtrage est un cas particulier (pour la loi marginale $\mathcal{L}(\mathbf{Z}_n | \mathbf{Y}_{1:n})$). L'estimation est plus précise qu'en filtrage car on utilise les données à venir, mais plus difficile à mettre en œuvre. Différents algorithmes ont été proposés utilisant un modèle inverse (Bennett, 1992) ou adjoint (Le Dimet et Talagrand, 1986). Une solution séquentielle par augmentation de l'état du modèle a été proposée (Evensen et van Leeuwen, 2000) et appliquée en océanographie à des modèles écologiques (Breuillin, Evensen, et Eknes, 2000 ; Natvik et Evensen, 2001). Nous ne considérerons par la suite que le problème du filtrage car les applications envisagées sont des prédictions.

2.4 Contraintes fortes ou contraintes faibles ?

On distingue deux types de problèmes dits en contraintes fortes ou faibles. Le problème en contraintes fortes suppose que le modèle physique est parfait, auquel cas la loi d'évolution devient $\mathbf{Z}_{n+1} = \mathbf{f}_n(\mathbf{Z}_n)$. Ce problème est certes moins général mais plus pratique que celui en contraintes faibles qui autorise à la fois le modèle et les données à comporter des erreurs. Il est fréquemment traité en météorologie par la méthode de l'adjoint (Le Dimet et Talagrand, 1986) où l'optimisation porte sur l'état initial du système. Contrairement à ce que son nom pourrait indiquer, le problème en contraintes faibles est donc plus complexe que celui en contraintes fortes.

2.5 La résolution séquentielle dans le cas général

2.5.1 Loi de transition

En toute rigueur, le fait d'ajouter une perturbation aléatoire aux équations différentielles déterministes qui régissent l'évolution de l'état du système en fait des équations différentielles stochastiques d'Itô. C'est la position adoptée par Miller *et al.* (1999). Cependant la résolution des équations d'Itô est en théorie très compliquée. Miller *et al.* (1999) se limite à l'intégration numérique d'une réalisation particulière, mais en temps continu. Si sa densité de probabilité $\phi(\mathbf{z}(t) | \mathbf{y}_{n < t})$ est définie, elle obéit alors à l'équation de Fokker-Planck, aussi appelée équation directe de Kolmogorov, notée sans conditionnement au passé $\phi(\cdot | \mathbf{y}_{n < t})$ pour

alléger :

$$\frac{\partial \phi(\mathbf{z}(t))}{\partial t} = - \sum_i \frac{\partial (f_i(\mathbf{z}(t))\phi(\mathbf{z}(t)))}{\partial z_i} + \sum_{i,j} \frac{1}{2} \Sigma_{ij}^m \frac{\partial^2 \phi(\mathbf{z}(t))}{\partial z_i \partial z_j}, \quad (2.7)$$

avec Σ^m la covariance de l'erreur de modèle. L'équation de Fokker-Planck est une équation d'advection-diffusion déterministe que l'on peut résoudre comme une équation aux dérivées partielles (Jazwinski, 1970) pour des petites dimensions de l'état $N \leq 4$. Pour des dimensions aussi élevées que dans les cas d'études en océanographie, la discrétisation devient pratiquement impossible et les contraintes des densités de probabilités (positivité et intégrale égale à un) ne rendent la résolution que plus difficile. Miller *et al.* (1999) donnent un exemple simple qui explique pourquoi ils préfèrent les méthodes de Monte Carlo pour des dimensions de l'état $N > 4$. Les méthodes d'AD utilisées dans les cas réels évitent l'intégration explicite de l'équation (2.7).

Si l'on connaît déjà la densité $\phi(\mathbf{z}_{n-1}|\mathbf{y}_{1:n-1})$, l'équation (2.7) permet alors de calculer $\phi(\mathbf{z}_n|\mathbf{y}_{1:n-1})$.

2.5.2 Prise en compte des observations

Le théorème de Bayes est la pierre angulaire de l'AD car il donne la densité de probabilité *a posteriori* connaissant une nouvelle donnée \mathbf{y}_n , ce qui permet de l'assimiler dans le système. Puisque \mathbf{Y}_n ne dépend que de \mathbf{Z}_n (2.3) on a :

$$\phi(\mathbf{z}_n|\mathbf{y}_{1:n}) = \frac{\phi(\mathbf{z}_n|\mathbf{y}_{1:n-1})\phi(\mathbf{y}_n|\mathbf{z}_n)}{\int_{\mathbb{R}^N} \phi(\mathbf{z}|\mathbf{y}_{1:n-1})\phi(\mathbf{y}_n|\mathbf{z})d\mathbf{z}}. \quad (2.8)$$

Au numérateur du membre droit, la première densité de probabilité (loi *a priori*) est le résultat de l'intégration du modèle de $n-1$ à n et la seconde est la densité des observations.

2.5.2.1 Un algorithme récursif ?

La recherche d'un algorithme récursif est motivée par le caractère répétitif de la prédiction. En effet, quand on reçoit une nouvelle donnée \mathbf{y}_n , il semble préférable de réutiliser le calcul précédent de $\phi(\mathbf{z}_{n-1}|\mathbf{y}_{1:n-1})$ plutôt que de devoir tout reprendre à zéro. La formule de Bayes (2.8) fournit en principe une solution récursive, tout comme au problème plus général du lissage :

$$\phi(\mathbf{z}_{1:n}|\mathbf{y}_{1:n}) = \phi(\mathbf{z}_{1:n-1}|\mathbf{y}_{1:n-1}) \frac{\phi(\mathbf{y}_n|\mathbf{z}_n)\phi(\mathbf{z}_n|\mathbf{z}_{n-1})}{\phi(\mathbf{y}_n|\mathbf{y}_{1:n-1})}, \quad (2.9)$$

mais la simulation de la densité *a posteriori* est très difficile, surtout à cause des dénominateurs qui requièrent l'évaluation d'intégrales dans \mathbb{R}^N . On ne trouve donc d'applications directes que pour des valeurs faibles de N , mais cela n'est pratiquement pas envisageable pour les systèmes océanographiques où N avoisine le million.

Dans la partie théorique de ce manuscrit, on étudiera d'abord le cas particulier linéaire et Gaussien où une solution exacte est donnée par le filtre de Kalman (chapitre 3). Les méthodes de l'adjoint et des représentants donnent aussi des solutions exactes mais en utilisant le formalisme variationnel et ne sont pas développées. On étudiera ensuite (chapitre 4) des approximations pour les cas non linéaires, le filtre de Kalman étendu (Jazwinski, 1970), et enfin des méthodes de Monte Carlo : l'*Ensemble Kalman Filter* (Evensen, 1994), le filtre à particules et le filtre *Bootstrap* (Doucet, de Freitas, et Gordon, 2001).

Chapitre 3

Le cadre linéaire

Sommaire

Où l'on découvre les bienfaits des hypothèses linéaires et gaussiennes. On se place dans le cadre séquentiel et gaussien, on rappelle brièvement ce qu'est le meilleur estimateur linéaire non biaisé d'où on déduit la formulation des méthodes simples d'assimilation et du filtre de Kalman ainsi que les modifications à lui apporter en cas de différences de support ou en cas de biais.

Dans le cadre séquentiel, on notera \mathbf{Z}^t l'état aléatoire «vrai» (*true*). L'estimateur *a priori* sera noté \mathbf{Z}^f comme *forecast*, et l'estimateur *a posteriori* \mathbf{Z}^a comme «analyse». On considère les équations du modèle linéaire :

$$\mathbf{Z}_n^t = F_n \mathbf{Z}_{n-1}^t + \varepsilon_n^m \quad (3.1)$$

$$\mathbf{Y}_n = H \mathbf{Z}_n^t + \varepsilon_n^o. \quad (3.2)$$

Le modèle physique F_n et la fonction d'observation H sont linéaires et les variables d'état \mathbf{Z}_n^t sont des variables aléatoires gaussiennes. Ces hypothèses ont un grand intérêt pratique : d'une part les multiplications par F_n et H conservent la loi gaussienne (par conséquent \mathbf{Y}_n doit aussi être gaussienne), d'autre part les variables d'état sont parfaitement déterminées par leur deux premiers moments : l'espérance $E(\mathbf{Z}_n^t)$ et la matrice de covariance $C_n = E[(\mathbf{Z}_n^t - E(\mathbf{Z}_n^t))(\mathbf{Z}_n^t - E(\mathbf{Z}_n^t))^T]$.

Dans les méthodes séquentielles, on distingue traditionnellement deux étapes :

1. l'étape de propagation par le modèle physique qui donne l'ébauche $\mathbf{Z}_n^f = E(\mathbf{Z}_n^t | \mathbf{Y}_{1:n-1})$, et sa covariance $C_n^f = \text{Cov}(\mathbf{Z}_n^t, (\mathbf{Z}_n^t)^\top | \mathbf{Y}_{1:n-1})$, le f de *forecast* rappelant qu'il s'agit d'un état prédit par le modèle.
2. l'étape de correction par les données qui produit l'état analysé $\mathbf{Z}_n^a = E(\mathbf{Z}_n^t | \mathbf{Y}_{1:n})$ et sa covariance $C_n^a = \text{Cov}(\mathbf{Z}_n^t, (\mathbf{Z}_n^t)^\top | \mathbf{Y}_{1:n})$.

3.1 Préliminaire sur l'estimation linéaire aux moindres carrés

3.1.1 Stationnarité

Dans le cas monovarié, pour simplifier les notations, on dira d'une fonction aléatoire $Z(x)$ qu'elle est stationnaire d'ordre 2 si ses deux premiers moments sont invariants par translations :

$$\begin{aligned} \forall \mathbf{x}, \quad E(Z(\mathbf{x})) &= z^f \\ \forall (\mathbf{x}_1, \mathbf{x}_2), \quad \text{Cov}(Z(\mathbf{x}_1), Z(\mathbf{x}_2)) &= C(|\mathbf{x}_1 - \mathbf{x}_2|). \end{aligned} \quad (3.3)$$

La fonction de covariance C doit être définie positive.

Cette hypothèse stationnaire est nécessaire pour passer à l'étape de l'analyse structurale où l'on modélise la fonction de covariance expérimentale (ou plutôt le variogramme qui est un outil plus général et plus facile à estimer) pour obtenir une fonction théorique définie positive. Dans le cadre multivarié, la forme se complique car on doit modéliser les variogrammes croisés entre les différentes variables.

3.1.2 Estimation

On cherche le meilleur estimateur d'une variable en un point \mathbf{x}_0 , $\mathbf{Z}_0 = \mathbf{Z}(\mathbf{x}_0)$ connaissant sa moyenne $E(\mathbf{Z}^t(\mathbf{x})) = \mathbf{z}^f(\mathbf{x})$ que l'on supposera uniformément nulle pour simplifier¹ et m observations ponctuelles $\mathbf{y} = (y(\mathbf{x}_i), i = 1, \dots, m)$ avec une erreur d'observation gaussienne centrée de covariance Σ^o . On considère un estimateur linéaire de \mathbf{Z}_0^t

$$\mathbf{Z}_0^a = \sum_{i=1}^m \kappa_i Y(\mathbf{x}_i) \quad (3.4)$$

Dans l'hypothèse multigaussienne, l'estimateur du maximum de vraisemblance coïncide avec l'espérance conditionnelle $E(\mathbf{Z}_0^t | \mathbf{Y})$, la géométrie des espaces de Hilbert montre ensuite que cette espérance conditionnelle est aussi l'estimateur des moindres carrés. \mathbf{Z}_0^a devra donc minimiser la variance de $(\mathbf{Z}_0^t - \mathbf{Z}_0^a)$.

¹On peut de manière équivalente considérer la variable $\mathbf{Z}^t - \mathbf{z}^f$ et lui rajouter \mathbf{z}^f après analyse

Connaissant les covariances suivantes :

- entre observations $C_{i,j} = \text{Cov}(Y(\mathbf{x}_i), Y(\mathbf{x}_j))$,
- entre les observations et la cible $C_{i,0} = \text{Cov}(Y(\mathbf{x}_i), \mathbf{Z}_0^t)$
- et C_0 la variance de \mathbf{Z}_0^t ,

la variance d'estimation s'écrit

$$J(\kappa) = \text{Var}(\mathbf{Z}_0^t - \mathbf{Z}_0^a | Y(\mathbf{x}_i)) = C_0 + 2 \sum_i \kappa_i C_{i,0} + \sum_{i,j} \kappa_i \kappa_j C_{i,j}. \quad (3.5)$$

On la minimise en annulant les dérivées de cette forme quadratique. Les pondérateurs κ_i sont les solutions du système linéaire

$$\forall i = 1, \dots, m, \quad \sum_{j=1}^{j=m} C_{i,j} \kappa_j = C_{i,0} \quad (3.6)$$

et on obtient ainsi l'estimation optimale ainsi que sa variance d'estimation en retournant aux équations (3.4) et (3.5). Dans le cadre multivariable, on peut se référer à Wackernagel (1998) pour un exposé plus développé. Si la moyenne est inconnue, ou dans certaines hypothèses de dérive, le système linéaire à résoudre se trouve augmenté de quelques lignes et colonnes correspondant à des conditions d'universalité, voir les manuels classiques de géostatistique (Chauvet, 2000).

Pour reprendre un parallèle avec l'ADS en notant H la matrice d'observation et C la covariance de \mathbf{Z}_n^t conditionnellement à $\mathbf{Z}_{1\dots n-1}$, le système devient

$$\forall i = 1, \dots, m, \quad \sum_{j=1}^m (H_i C H_j^\top + \Sigma_{i,j}^o) \kappa_j = H_i C_0. \quad (3.7)$$

3.2 Le problème du support

3.2.1 Des covariances blocs aux covariances ponctuelles

L'inconvénient de la notation matricielle est qu'elle camoufle les différences de support entre les données ponctuelles prises aux points $\mathbf{x}_{1:m}$ et les composantes de l'état définies sur les nœuds de grille $V_{1:l}$. En notations continues et en omettant le conditionnement au passé, l'élément (i, j) des matrices C , CH^\top et HCH^\top devrait s'écrire respectivement :

$$\begin{array}{lll} i = 1:l & j = 1:l & (C)_{i,j} = \text{Cov}(\mathbf{Z}_n^t(V_i), \mathbf{Z}_n^t(V_j)) \\ i = 1:l & j = 1:m & (CH^\top)_{i,j} = \text{Cov}(\mathbf{Z}_n^t(V_i), \mathbf{Y}_n(\mathbf{x}_j)) \\ i = 1:m & j = 1:m & (HCH^\top + \Sigma^o)_{i,j} = \text{Cov}(\mathbf{Y}_n(\mathbf{x}_i), \mathbf{Y}_n(\mathbf{x}_j)), \end{array} \quad (3.8)$$

cependant, l'opérateur d'observation H ne permet pas d'affiner la résolution et en général dans les équations ci-dessus les $\mathbf{Y}(\mathbf{x}_i)$ sont en fait des $\mathbf{Y}(V_i)$.

Afin de retrouver les véritables covariances, faute de pouvoir inverser les convolutions suivantes :

$$\text{Cov}(\mathbf{Z}_n^t(V_i), \mathbf{Y}_n(V_j)) = \frac{1}{|V_j|} \int_{V_j} \text{Cov}(\mathbf{Z}_n^t(V_i), \mathbf{Y}_n(\mathbf{x}_j)) d\mathbf{x}_j \quad (3.9)$$

$$\text{Cov}(\mathbf{Y}_n(V_i), \mathbf{Y}_n(V_j)) = \frac{1}{|V_i| \times |V_j|} \int \int_{V_i \times V_j} \text{Cov}(\mathbf{Y}_n(\mathbf{x}_i), \mathbf{Y}_n(\mathbf{x}_j)) d\mathbf{x}_i d\mathbf{x}_j, \quad (3.10)$$

on doit avoir recours à un modèle de changement de support qui nous donne la loi jointe

$$\mathcal{L}(\mathbf{Z}(V_1), \dots, \mathbf{Z}(V_N), \mathbf{Y}(\mathbf{x}_1), \dots, \mathbf{Y}(\mathbf{x}_m)),$$

connaissant la loi jointe (gaussienne) des blocs $\mathbf{Z}(V_i)$ et $\mathbf{Y}(V_j)$. Suivant la taille des blocs, les effets d'une régularisation (passage d'un support ponctuel au support bloc) peuvent être plus ou moins importants, voir sur les figures 7.4 et 6.5 les effets sur le variogramme. Par contre dans les cas courants ce sont les microstructures qui diffèrent le plus, alors que les structures à grande échelle sont les moins affectées.

3.2.2 Cas stationnaire

De nombreux modèles de changement de support existent pour les cas stationnaires (ce qui correspond ci-dessous aux méthodes primitives d'assimilation, mais non au filtre de Kalman) et permettent de déterminer la distribution d'une variable définie sur un bloc puis les courbes de sélectivité à différents supports ; nous ne les exposerons pas ici puisqu'ils n'ont pas été appliqués. Pour les lois gaussiennes (ou assimilées) le modèle de correction affine et le modèle gaussien discret conviennent. Ce dernier est plus général et s'applique à d'autres variables que celles de lois gaussiennes par anamorphose ; notons aussi que certains modèles s'accommodent mieux que d'autres des «effets zéro»² comme le modèle mosaïque. Une revue récente de ces modèles est donnée dans Chilès et Delfiner (1999).

3.2.3 Cas non stationnaire

Dans le cas non stationnaire le problème revient à retrouver les microstructures présentes au support ponctuel mais non au support bloc. Cette question représente à l'heure actuelle une question ouverte, comme toute la théorie de la géostatistique à la fois non stationnaire et non linéaire. On ne peut pas non plus simuler l'effet de support par un méthode de Monte Carlo puisque c'est justement le modèle spatial ponctuel qui fait défaut. On doit ainsi renoncer à toutes les méthodes exactes ou itératives car d'une part les observations,

²Accumulation de valeurs nulles

aussi nombreuses soient-elles, d'une réalité unique ne permettent pas d'inférer les covariances inconnues $\text{Cov}(\mathbf{Y}_n(\mathbf{x}_i), \mathbf{Y}_n(\mathbf{x}_j))$ et $\text{Cov}(\mathbf{Z}_n^t(V_i), \mathbf{Y}_n(\mathbf{x}_j))$ et d'autre part la dynamique du système décrite à une résolution plus grande V n'est d'aucun secours.

On est donc réduit à supposer dans le cas non stationnaire un effet de support trivial, par exemple un effet de pépite ajouté, dont on peut calculer la variance à chaque pas de temps sur la base des données assimilées si celles-ci sont assez nombreuses³. On rejoint ainsi la méthode proposée par Brasseur *et al.* (2001), à ma connaissance la seule réponse pratique à ce problème dans le cadre de l'ADS : il s'agit d'un calcul empirique de la variance d'extension, qui dans le cas étudié est possible car les données satellites utilisées sont très denses par rapport à la résolution de la grille du modèle. Une complication supplémentaire dans son cas vient du fait que la grille est définie en coordonnées polaires et la taille de la maille dépend de sa position.

En retournant à (3.9), l'effet de pépite ajouté disparaît de la covariance $\text{Cov}(\mathbf{Z}_n^t(V_i), \mathbf{Y}_n(\mathbf{x}_j))$, mais pas de $\text{Cov}(\mathbf{Y}_n(\mathbf{x}_i), \mathbf{Y}_n(\mathbf{x}_j))$ à laquelle il s'ajoute comme une erreur de mesures supplémentaire.

3.3 Les méthodes primitives d'assimilation

On trouve de nombreuses méthodes «simples» d'ADS : insertion directe, *nudging*, etc. Voir Robinson, Lermusiaux, et Sloan III (1998) pour une revue récente. On ne retiendra ici qu'une seule, «l'analyse objective» aussi appelée interpolation optimale (OI) en ADS météorologique où elle a longtemps été utilisée en prévision opérationnelle, voir Daley (1991) pour une description détaillée. Les autres méthodes simples se déduisent souvent de l'OI pour des cas particuliers d'interpolation.

En OI, les variables d'état $\mathbf{Z}_n^t(\mathbf{x})$ à chaque temps n sont supposées avoir pour moyenne l'ébauche $E(\mathbf{Z}_n^t) = \mathbf{z}_n^f$ et une covariance stationnaire, C , fixée à l'avance et indépendante de n . En notations vectorielles on a ainsi la matrice de covariances $\forall n, C_n^f = C$.

La phase d'analyse est alors un krigeage à moyenne connue des données \mathbf{Y} . C étant la matrice de covariance du vecteur d'état \mathbf{Z}_n^t , les matrices des covariances entre l'état et les données, et entre les données sont respectivement CH^\top et $HCH^\top + \Sigma^o$. On considère des erreurs de mesure $\mathbf{Y}_n - H\mathbf{Z}_n^t \sim \mathcal{N}(0, \Sigma^o)$.

Les équations de l'OI sont ainsi dans le cas linéaire :

$$\mathbf{Z}_n^f = F_n \mathbf{Z}_{n-1}^a \quad (3.11)$$

$$\mathbf{Z}_n^a = \mathbf{Z}_n^f + K(\mathbf{Y}_n - H\mathbf{Z}_n^f), \quad (3.12)$$

où la matrice K des pondérateurs de krigeage n'est autre que $K = CH^\top (HCH^\top + \Sigma^o)^{-1}$, elle est invariable

³A partir de 50 données, on peut raisonnablement tracer un variogramme expérimental et mesurer l'effet de pépite en tant qu'ordonnée à l'origine.

dans le temps. Cette méthode est donc très économique au point de vue des temps de calcul et est encore largement utilisée en prédiction opérationnelle océanographique.

3.4 Le filtre de Kalman

3.4.1 Modèle

Le filtre de Kalman apporte un durcissement des hypothèses de l'OI. En effet ce n'est plus l'état du système \mathbf{Z}_n^t qui est supposé stationnaire mais les deux erreurs du modèle $\varepsilon_n^m = \mathbf{Z}_n^t - F_n \mathbf{Z}_{n-1}^t$ et d'observation $\varepsilon_n^o = \mathbf{Y}_n - H \mathbf{Z}_n^t$.

Les hypothèses du filtre de Kalman sont alors les suivantes :

1. Erreurs de modèle et d'observations gaussiennes centrées

$$\forall n, \varepsilon_n^m \sim \mathcal{N}(0, \Sigma^m), \varepsilon_n^o \sim \mathcal{N}(0, \Sigma^o)$$

donc sans biais $\forall n, E(\varepsilon_n^m) = E(\varepsilon_n^o) = 0$. Les deux matrices de covariance Σ^o et Σ^m sont supposées parfaitement connues, ce qui n'est pas vrai dans les cas réels.

2. Bruits blancs temporels

$$\forall n, \text{Cov}(\varepsilon_n^m, \varepsilon_{n-1}^m) = 0 \quad (3.13)$$

$$\text{Cov}(\varepsilon_n^o, \varepsilon_{n-1}^o) = 0 \quad (3.14)$$

3. Erreurs indépendantes entre elles $\forall (n_1, n_2), \text{Cov}(\varepsilon_{n_1}^m, \varepsilon_{n_2}^o) = 0$

3.4.2 Équations de l'ébauche

Le modèle dynamique linéaire conserve le caractère gaussien de la variable \mathbf{Z}^t . Puisque le modèle et les observations sont sans biais à chaque pas de temps, alors on a par récurrence :

$$\mathbf{z}_n^f = E(\mathbf{Z}_n^t | \mathbf{y}_{1:n-1}) \quad (3.15)$$

$$= E(\mathbf{Z}_n^t | \mathbf{z}_{n-1}^a) \quad (3.16)$$

$$= F_n \mathbf{z}_{n-1}^a. \quad (3.17)$$

La matrice de covariance de l'ébauche de \mathbf{Z}_n^t est :

$$C_n^f = \text{Cov}(\mathbf{Z}_n^t | \mathbf{y}_{1:n-1}) \quad (3.18)$$

$$C_n^f = F_n C_{n-1}^a F_n^\top + \Sigma^m. \quad (3.19)$$

3.4.3 Correction

Suivant l'étape de propagation, la variable \mathbf{Z}_n^t est gaussienne, on connaît sa moyenne \mathbf{z}_n^f et sa matrice de covariance C_n^f . On peut donc assimiler les données \mathbf{Y}_n , aussi gaussiennes d'espérance $H\mathbf{z}_n^f$, par estimation linéaire aux moindres carrés :

$$\mathbf{z}_n^a = E(\mathbf{Z}_n^t | \mathbf{y}_{1:n}) \quad (3.20)$$

$$= E(\mathbf{Z}_n^t | \mathbf{z}_n^f, \mathbf{y}_n) \quad (3.21)$$

$$= \mathbf{z}_n^f + K_n (\mathbf{y}_n - H\mathbf{z}_n^f). \quad (3.22)$$

Le vecteur $(\mathbf{y}_n - H\mathbf{z}_n^f)$ est couramment appelé «innovation», ce qui rappelle l'interprétation de l'estimation aux moindres carrés comme projection orthogonale de \mathbf{Z}_n^t sur l'espace décrit par $\mathbf{Y}_{1:n}$: à chaque pas de temps, les nouvelles données \mathbf{y}_n se décomposent en une projection $H\mathbf{z}_n^f$ sur l'espace décrit engendré par $\mathbf{Y}_{1:n-1}$ et le résidu $\mathbf{y}_n - H\mathbf{z}_n^f$ orthogonal et innovant. L'équation (3.22) est similaire à celle de l'analyse par l'OI, cependant la matrice K_n est calculée à toutes les assimilations pour minimiser la variance de l'erreur d'estimation $\varepsilon^a = \mathbf{Z}_n^t - \mathbf{z}_n^a$ pour chaque composante du vecteur d'état et la covariance C^f n'est plus supposée stationnaire. K_n minimise la variance de l'erreur d'estimation *a posteriori* :

$$J(K_n) = E(\varepsilon_n^a \varepsilon_n^{a\top}) \quad (3.23)$$

$$= \text{Var}(\varepsilon_n^f - K_n H \varepsilon_n^f - K \varepsilon^o) \quad (3.24)$$

$$= (I - K_n H) C_n^f (I - K_n H)^\top + K_n \Sigma^o K_n^\top \quad (3.25)$$

$$= C_n^f - K_n H C_n^f - C_n^f H^\top K_n^\top + K_n (H C_n^f H^\top + \Sigma^o) K_n^\top. \quad (3.26)$$

La covariance de l'erreur de mesure Σ^o représente une structure filtrée (en général un effet de pépité). Pour les N éléments du vecteur d'état, le krigeage à moyenne connue donne alors les pondérateurs κ_i qui forment la ligne correspondante de la matrice de gain de Kalman K_n :

$$K_n = C_n^f H^\top (H C_n^f H^\top + \Sigma^o)^{-1} \quad (3.27)$$

$$C_n^a = C_n^f - K_n H C_n^f. \quad (3.28)$$

On obtient la matrice C_n^a de covariance après analyse en reportant K_n dans la variance d'estimation (3.26). Voir le calcul matriciel menant à (3.27) dans Maybeck (1979). Sur l'équation (3.28) on remarque que les valeurs des données \mathbf{y}_n n'interviennent pas dans l'expression de C_n^a . Seule l'implantation des données est prise en compte, à l'instar de la variance d'erreur d'estimation d'un krigeage. L'incertitude sur l'estimation n'est donc pas fonction des erreurs calculées par le filtre de Kalman ce qui nous fait relativiser l'adage selon lequel «le filtre de Kalman analyse ses propres erreurs». La covariance de l'erreur *a posteriori* dépend exclusivement de la covariance initiale C_0 , de la covariance de l'erreur de modèle Σ^m , du modèle F et de l'implantation des données au travers de la matrice d'observation H et de la covariance d'erreur d'observation Σ^o qui intervient *via* le gain de Kalman K_n . Ceci signifie qu'un phénomène présent dans les données mais négligé dans le modèle dynamique ne peut pas apparaître après analyse dans C_n^a . Cependant, la matrice C_n^a peut quand même être valorisée en pratique pour évaluer le système d'AD car elle montre l'influence qu'une donnée peut avoir sur les points avoisinants et d'une variable à l'autre.

3.4.4 Remarque à propos de l'initialisation

Dans les équations précédentes, l'état \mathbf{Z}_0^a et la covariance C_0^a au temps 0 sont inconnus et doivent être attribués arbitrairement en espérant que le filtre de Kalman corrige de lui-même cette imprécision. L'influence de l'initialisation peut être plus ou moins longue suivant la nature des variables et leurs dynamiques. On verra par exemple dans les applications à la lagune de l'Oder que la persistance de la variable salinité est plus longue que celle des hauteurs d'eau car le transport est un phénomène plus lent (de l'ordre de $1\text{cm}\cdot\text{s}^{-1}$) que la propagation d'ondes de surface (de célérité $\sqrt{gH} \approx 5\text{m}\cdot\text{s}^{-1}$).

Une technique à la mode pour déterminer C_0^a consiste à se baser sur une succession d'états générés par le modèle pour en donner une approximation sur un sous-espace vectoriel. Brasseur *et al.* (1999) utilisent à cette fin des EOF (*Empirical Orthogonal Functions*) de ces vecteurs d'états obtenus par analyse en composantes principales de la succession d'états. Outre les précautions usuelles d'emploi de l'ACP⁴, le fait même de n'utiliser que le modèle dynamique pour initialiser C_0^a est critiquable puisque les données n'interviennent dans les C_n^f et C_n^a successifs que par leur implantation. Les phénomènes absents du modèle mais présents dans les données ne seront donc pas représentés dans les covariances successives ce qui limite les possibilités de correction. Plus généralement ce type d'initialisation constitue pour l'ADS, censée arbitrer équitablement entre le modèle et les observations, un déséquilibre en faveur du modèle. Dans les applications où l'initialisation est importante comme au chapitre 7 on préférera donc obtenir C_0^a à partir d'une analyse de données plutôt qu'à partir d'une analyse de résultats du modèle.

3.5 Stabilité du filtre de Kalman

Toutes les questions de convergence et de sensibilité du filtre de Kalman sont traitées classiquement pour un modèle invariant dans le temps $F_n = F$. Les résultats gardent cependant une certaine généralité.

3.5.1 Types de convergence

La résolution du problème de l'AD nécessite en toute rigueur une convergence en loi de $\mathcal{L}(\mathbf{Z}_n^a)$ vers $\mathcal{L}(\mathbf{Z}_n^t)$ quand $n \rightarrow +\infty$. La théorie de l'automatique (d'Andréa Novel et de Lara, 1994) indique dans quelles conditions de commandabilité et d'observabilité cette convergence est assurée indépendamment de l'état et de la covariance initiaux. Cependant ces conditions ne sont pas remplies dans les cas usuels puisque les observations sont trop rares. De surcroît, il est impossible de vérifier empiriquement la convergence en loi dans la plupart des cas réels car l'état «vrai» est inconnu. En pratique on est donc réduit à suivre l'évolution

⁴On peut voir dans Wackernagel (1998) comment améliorer cette procédure en tenant compte des corrélations à différentes échelles.

des innovations $\mathbf{Y}_n - H\mathbf{Z}_n^f$ qui doivent tendre asymptotiquement vers zéro, ou bien à utiliser des données auxiliaires pour valider l'assimilation.

Verlaan (1998) préconise aussi de surveiller la convergence des matrices $K_n \rightarrow K_\infty$ et $C_n \rightarrow C_\infty$ qui se stabilisent à l'équilibre du filtre de Kalman (toujours dans le cas d'un modèle invariant dans le temps). Cañizares *et al.* (2001) utilisent ensuite la matrice limite K_∞ dans une procédure d'interpolation optimale nommée «filtre de Kalman en régime permanent», plus économique en temps de calcul que le filtre de Kalman brut. Ils prédisent ainsi les hauteurs d'eau sur les côtes du Danemark avec de bons résultats. Par contre dans le cas où le modèle F_n n'est pas constant dans le temps, par exemple s'il dépend des saisons, la convergence des matrices C_n et K_n n'est plus certaine. L'étude des propriétés asymptotique de ces matrices peut par contre être utile pour qualifier le système d'AD constitué (voir par exemple le chapitre 7).

3.5.2 Conditions théoriques de convergence

3.5.3 Causes de divergence

Fitzgerald (1971) distingue deux type de divergence du filtre de Kalman : la divergence «vraie» où les valeurs diagonales de C_n croissent indéfiniment avec n et la divergence «apparente» où la loi de l'estimateur ne tend pas vers celle de \mathbf{Z}_n^t . Fitzgerald (1971) s'est intéressé au comportement asymptotique des erreurs d'estimation pour les système invariants dans le temps afin de mieux comprendre le comportement des systèmes plus complexes.

Dans ces cas d'école, la divergence «vraie» intervient pour des modèles F instables (une au moins de ses valeurs propres est de module $\rho > 1$), ou si les matrices F et H sont mal spécifiées. Ni les biais de modèle et de mesure, ni la détermination des matrices Σ^o et Σ^m , ni les erreurs numériques ne constituent des raisons supplémentaires de divergence «vraie». Pour la divergence «apparente», par contre, les causes potentielles sont multiples : mauvaises spécifications de F , H , Σ^m ou Σ^o , biais de mesure, de modèle, voire erreurs numériques. Il est donc possible dans les cas réels que ces causes de divergence détériorent considérablement les performances du filtre de Kalman.

3.5.4 Effets du biais sur le filtre de Kalman

L'équation (3.22) montre que le filtre de Kalman, s'il ignore un biais, ne le réduit que partiellement. Le biais après correction est :

$$E(\varepsilon_n^a) = E(\mathbf{Z}_n^f - \mathbf{Z}_n^t) + K_n E[\varepsilon_n^o - H(\mathbf{Z}_n^f - \mathbf{Z}_n^t)] \quad (3.29)$$

$$= (I - K_n H) E(\varepsilon_n^f) + K_n E(\varepsilon_n^o). \quad (3.30)$$

Le premier terme du membre droit montre qu'une correction largement répartie sur toute l'étendue spatiale de l'état ($I - K_n H \approx 0$) peut estomper un biais d'estimation $E(\varepsilon_n^f)$ mais le second terme peut répandre tout aussi largement un éventuel biais de mesure $E(\varepsilon_n^o)$. Inversement si les corrections sont locales, le biais de modèle s'accumule d'un pas de temps sur l'autre.

3.6 Estimation d'un biais constant dans le temps

Le sujet de l'estimation d'un biais *on line* a été largement étudié depuis que la non linéarité s'est avérée être une source de biais de modèle (Jazwinski, 1970). La difficulté dans les applications consiste à ne pas aggraver la complexité des systèmes d'ADS déjà très gourmands en temps de calcul. Une approche classique consiste à probabiliser des paramètres de biais β (biais de modèle ou de mesure). Ils sont supposés constants dans le temps, on suppose de plus qu'on connaît leur liens avec les vecteurs d'état et d'observation (les matrices G_n et L_n doivent indiquer quelle est l'influence du biais) :

$$\mathbf{Z}_n^t = F_n \mathbf{Z}_{n-1}^a + G_n \beta + \varepsilon_n^m \quad (3.31)$$

$$\mathbf{Y}_n = H \mathbf{Z}_n^t + L_n \beta + \varepsilon_n^o. \quad (3.32)$$

Jazwinski (1970) inclut les β dans le vecteur d'état. Pour simplifier, Friedland (décrit et commenté dans Dee et da Silva (1998)) a montré que l'on peut de manière équivalente mener deux récursions, l'une sur l'état sans tenir compte du biais, l'autre sur les paramètres β . On peut ainsi appliquer d'abord les équations classiques du filtre de Kalman, puis estimer le biais par une procédure similaire à un second filtre de Kalman greffé sur le premier. Si l'avantage de cette approche est de permettre le calcul du biais en ligne en ajoutant peu de calculs, elle requiert par contre une connaissance *a priori* des mécanismes du biais (c'est à dire des matrices G_n et L_n). De plus l'algorithme calcule itérativement les covariances des paramètres de biais à partir de covariances de l'état C_n^f et C_n^a , ce qui rend ce calcul deux fois plus hasardeux.

Une variante de l'algorithme de Friedland (Dee et da Silva, 1998) suppose $G_n = I_n$ (matrice identité) et $L_n = 0$. Dee et da Silva (1998) introduit ainsi N paramètres de biais, autant qu'il y a de composantes dans le vecteur d'état. Ce problème comportant beaucoup plus d'inconnues (N paramètres de biais supplémentaires), il formule une hypothèse simplificatrice : le biais a une matrice de covariance proportionnelle à C_n^a covariance du vecteur d'état. Cette hypothèse est certes abusive (ne ferait-on pas mieux d'attribuer au biais une covariance constante pour simplifier ?), mais elle a une caractéristique qui satisfait l'intuition : le biais est plus incertain là où l'état est lui-même incertain. Le biais estimé qui en résulte est une moyenne temporelle des innovations interpolées $\sum_n K_n (\mathbf{y}_n - H \mathbf{z}_n^f)$. On devine que cette approche ne sera efficace que si le biais n'affecte pas les covariances ce qui réduit beaucoup la portée de l'algorithme.

Cette méthode, tout comme celle de Friedland ou celle de Jazwinski (1970), met à profit l'historique des innovations. Leurs modèles de biais aléatoire comportent cependant des paramètres trop souvent inconnus et il vaut certainement mieux disposer d'informations supplémentaires sur les sources possibles de biais.

Duflo (1997) propose un gradient itératif pour la vraisemblance que l'on pourrait apparenter à la méthode d'estimation du biais de Friedland, les paramètres évoluent dans une relation linéaire :

$$\beta_n^a = \beta_{n-1}^a + \gamma_n(\mathbf{y}_n - H\mathbf{z}_n^f). \quad (3.33)$$

On en déduit un algorithme très simple car le calcul du biais est fait par une suite γ_n arbitraire⁵. Duflo (1997) montre la stabilité de l'estimateur ainsi obtenu dans le cas où $\sum_n \gamma_n^2 < +\infty$. $\sum_n \gamma_n = +\infty$ est aussi imposée pour d'autres raisons théoriques. Dans cette méthode, l'historique des innovations est aussi capitalisé. Avec $\gamma_n = 1/n$ on peut calculer un biais de mesure qui est tout simplement la moyenne arithmétique de 1 à n des innovations.

3.7 Le filtre de Kalman avec conditions d'universalité

3.7.1 Équivalence du filtre de Kalman avec le krigeage

L'étape d'analyse (3.22) est une interpolation par Krigeage Simple (KS) (ou plutôt par cokrigeage simple dans le cas multivariable) des écarts entre le modèle et les données $\mathbf{Y} - H\mathbf{Z}^f$ sur tous les nœuds de la grille du modèle. La matrice de covariance après correction (3.28) contient aussi l'erreur d'estimation de ce krigeage sur la diagonale, les termes non diagonaux étant les covariances *a posteriori* entre deux nœuds de grille distincts, éventuellement entre deux variables différentes. Voyant l'équivalence entre le filtre de Kalman et le KS, il est tentant d'en déduire d'autres filtres correspondant aux types classiques de krigeage (KO, KU, KDE, krigeage intrinsèque d'ordre k , etc.) couramment utilisés en géostatistique linéaire (Chauvet, 2000). On retiendra ici quelques variantes parmi celles qui pourraient le plus vraisemblablement trouver une application en océanographie.

3.7.2 Modèle et équations, cas monovarié

On appelle condition d'universalité une condition de non biais qui ne porte pas sur un paramètre du modèle stochastique mais sur l'ensemble de la variable d'état \mathbf{Z}_n^t . Dans cette partie, on se place dans un premier temps dans le cas monovarié pour simplifier. On prend en compte des biais de modèle qui ne sont pas stochastiques cette fois, mais seulement des paramètres inconnus a_n , ce qui fait l'économie d'une variable aléatoire par rapport aux modèles cités ci-dessus. L'équation de propagation de l'état devient :

$$\mathbf{Z}_n^t = a_n F_n \mathbf{Z}_{n-1}^t + \varepsilon_n^m, \quad (3.34)$$

⁵D'autres méthodes plus élaborées sont aussi exposées, mais le but ici est de simplifier.

où l'on ne suppose plus les erreurs de modèles d'espérance nulle mais où l'on fait l'hypothèse d'une décomposition du biais sur une famille de fonctions fixes \mathbf{g}_i , $i = 1, \dots, m'$. En général on note \mathbf{g}_1 la constante 1, puis les \mathbf{g}_i suivantes sont des fonctions polynômes et exponentielles des coordonnées spatiales.

$$\varepsilon_n^m \sim \mathcal{N}\left(\sum_i b_n^i \mathbf{g}_i, \Sigma^m\right). \quad (3.35)$$

La portée physique de ce modèle est relative. Si on peut voir dans les fonctions de biais \mathbf{g}_i des fonctions de l'espace filtrées par le modèle physique mais présentes dans les données, par contre l'hypothèse d'un biais multiplicatif semble assez particulière : on peut imaginer un modèle de transport biaisé qui achemine des constituants à la mauvaise vitesse ou au mauvais endroit, mais dans ces cas de figure il paraît peu réaliste qu'on aboutisse à une prédiction uniformément proportionnelle à la réalité sur toute l'étendue du domaine. On doit donc chercher un modèle de biais spécifique à chaque cas d'étude.

Les paramètres de biais a_n et b_n^i sont des inconnues que l'on estime grâce aux données lors de chaque assimilation. Si on note $\mathbf{A}_n^\top = (a_n, b_n^1, \dots, b_n^{m'})$ et $G = (F_n \mathbf{Z}_{n-1}^a, \mathbf{g}_1, \dots, \mathbf{g}_{m'})$, on obtient les espérances suivantes :

$$E(\mathbf{Z}_n^t | \mathbf{Z}_{n-1}^a) = a_n F_n \mathbf{Z}_{n-1}^a + \sum_i b_n^i \mathbf{g}_i \quad (3.36)$$

$$= G \mathbf{A}_n \quad (3.37)$$

$$E(\mathbf{Y}_n | \mathbf{Z}_{n-1}^a) = a_n H F_n \mathbf{Z}_{n-1}^a + \sum_i b_n^i H \mathbf{g}_i \quad (3.38)$$

$$= H G \mathbf{A}_n. \quad (3.39)$$

On cherche comme précédemment un estimateur linéaire,

$$\mathbf{Z}_n^a = K_n \mathbf{Y}_n, \quad (3.40)$$

mais cette fois-ci sans en connaître la moyenne. Si l'on veut alors assurer la condition d'universalité

$$E(\mathbf{Z}_n^a | \mathbf{Z}_{n-1}^a) = E(\mathbf{Z}_n^t | \mathbf{Z}_{n-1}^a) \quad (3.41)$$

$$K_n H G \mathbf{A}_n = G \mathbf{A}_n, \quad (3.42)$$

quelles que soient les valeurs des paramètres inconnus \mathbf{A}_n , le gain de Kalman K_n doit satisfaire les $m' + 1$ contraintes suivantes :

$$(K_n H - I)G = 0. \quad (3.43)$$

Selon le calcul des variations, on peut minimiser la variance de l'erreur d'estimation $\text{Var}(\mathbf{Z}_n^t - \mathbf{Z}_n^a | \mathbf{Z}_{n-1}^a)$ sous ces contraintes linéaires en introduisant $m' + 1$ paramètres de Lagrange $\lambda_n = (\lambda_n^i)$. La forme quadratique à minimiser devient alors

$$J(K_n, \lambda_n) = E(\varepsilon_n^a \varepsilon_n^{a\top}) + \lambda_n^1 (K_n H \mathbf{Z}_n^f - \mathbf{Z}_n^f) + \sum_{i=2}^{m'} \lambda_n^i (K_n H - I) \mathbf{g}_i \quad (3.44)$$

et en annulant son gradient, on obtient K_n et autant de vecteurs λ_n qu'il y a de points de \mathbf{Z}_n^t à estimer (regroupés dans la matrice Λ_n de taille $(m' + 1) \times N$). Le système linéaire suivant donne ainsi à la fois K_n et Λ_n :

$$\begin{pmatrix} HC_n^f H^\top + \Sigma^m & HG \\ G^\top H^\top & 0 \end{pmatrix} \begin{pmatrix} K_n \\ \Lambda_n \end{pmatrix} = \begin{pmatrix} HC_n^f \\ G^\top \end{pmatrix} \quad (3.45)$$

Le nombre de colonnes de G doit être inférieur au nombre de données $m' + 1 < m$ pour que l'estimation soit possible. Dans le cas contraire la première matrice au membre gauche n'est pas inversible.

On peut noter que cette modification du filtre de Kalman est exacte dans le cas gaussien, mais on peut aussi l'appliquer dans les cas non linéaires pour lesquels l'approximation linéaire reste valable (c'est-à-dire la majorité des applications courantes). Il est très simple d'ajouter ces modifications à un code de calcul existant des filtres de Kalman qui utilisent une étape de correction linéaire (EKF, EnKF, etc.).

3.7.3 Analogie avec les autres types de krigeage

En pratique le nombre $m' + 1$ de lignes et de colonnes à ajouter au système (3.45) est bien moins élevé que le nombre d'observations. Si on suppose que le modèle commet seulement un biais additif, soit l'équation (3.34) avec $a_n = 1$ et b_n^1 inconnu et $b_n^i = 0, \forall i > 1$, alors par un système analogue au KO, on estime un paramètre de moyenne et la matrice G a une seule colonne.

Si on suppose en plus du biais additif inconnu b_n^1 un biais multiplicatif a_n inconnu, alors l'interpolation est analogue à un KDE et l'estimation pour chaque variable observée des deux paramètres a_n et b_n^1 ajoute deux lignes et deux colonnes au système à inverser (3.45).

Dans les deux cas, on voit que le surcoût en temps de calcul ajouté par ces variantes du filtre de Kalman est très faible. Cependant leur utilisation dans les cas d'étude est restreinte par la nécessité d'une couverture d'observations assez dense. Les paramètres a_n et b_n^i sont en effet réestimés à chaque étape de correction sur la base des m nouvelles observations qui doivent être assez nombreuses pour que l'estimation soit stable. Dans les applications avec moins de 5 données, le résultat peut être désastreux.

On remarque par contre que l'historique des innovations n'est pas mis à profit dans l'estimation du biais car celui-ci est supposé indépendant d'un pas de temps à l'autre. On pourrait remplacer cette hypothèse d'indépendance par un modèle auto-régressif, ce qui aurait alors l'avantage de donner un estimateur d'un biais variable dans le temps mais tenant compte des innovations passées. Cependant l'étape de correction demandera alors l'utilisation des états et des covariances des pas de temps passés, ce qui signifie qu'il faut augmenter l'état du modèle. On évitera donc cette approche dans un souci d'économie pour le filtrage, mais dans les applications de lissage où l'on utilise un état augmenté, elle est envisageable.

3.7.4 Cas multivariable

Dans le cas à k variables, les paramètres a_n et b_n^i ne sont en général pas les mêmes d'une variable à l'autre. Il est en effet difficile d'imaginer qu'un modèle physique biaisé produise le même biais sur chaque variable et on supposera plutôt les biais de chaque variable complètement indépendants. Dans ce cas \mathbf{a}_n et \mathbf{b}_n^i sont des vecteurs de taille k et le nombre de colonnes de la matrice G se voit alors multiplié par k .

Dans le cas où l'introduction de ces paramètres de biais contredit une propriété physique du système (de type conservation de la masse), on peut imposer aux paramètres de biais de respecter ces contraintes (à condition qu'elles soient linéaires) par élimination dans le système linéaire. On diminue ainsi la taille de la matrice G .

3.8 Autres impulsions venant de la géostatistique

A l'ordre 1 le krigeage intrinsèque aboutit aux mêmes équations que le KO et ne pose donc pas plus de difficultés. Par contre aux ordres supérieurs, cette approche peut se justifier dans les cas où les données sont assez nombreuses et les erreurs fortement non stationnaires. On peut penser à des exemples dans l'assimilation de données satellites, mais ce domaine souffre déjà de nombreuses contraintes numériques et une telle complication du filtre de Kalman n'est certainement pas la bienvenue. En effet les efforts actuels se font plutôt vers la recherche de voisinages glissants (*local updates*) petits et adaptés à des covariances non stationnaires, ce qui est un problème sans doute aussi ardu mais plus sensible parce qu'il touche plus directement au respect des données.

Chapitre 4

Le cadre non linéaire

Sommaire

Rappel des extensions du filtre de Kalman aux lois de propagation non linéaires, Extended Kalman filter (EKF) dont on déduit le reduced rank square root Kalman filter (RRSQRT) comme algorithme simplifié et l'ensemble Kalman filter (EnKF). Quelques pistes vers une correction non linéaire avec des filtres qui n'utilisent plus une correction linéaire et qui ne sont donc plus à proprement parler des filtres de Kalman.

On considère les équations d'un modèle non linéaire particulier. Celui qui se présente habituellement en AD océanographique :

$$\mathbf{Z}_n^t = \mathbf{f}_n(\mathbf{Z}_{n-1}^t, \varepsilon_n^m) \quad (4.1)$$

$$\mathbf{Y}_n = \mathbf{h}(\mathbf{Z}_n^t) + \varepsilon_n^o. \quad (4.2)$$

La dimension de l'état \mathbf{Z} est N et celle des observations \mathbf{Y} est m . Dans ce modèle (voir par exemple le chapitre 6 pour une application), on suppose que l'erreur de modèle ε_n^m porte sur q les paramètres du modèle dynamique (conditions aux limites, champs de vent, viscosités et autres paramètres internes du modèle ...) sa matrice de covariance est Σ^m . Ce modèle est plus physique que celui où l'erreur est une fonction aléatoire stationnaire ajoutée à $\mathbf{f}_n(\mathbf{Z}_{n-1}^t)$ puisque les sources d'erreurs sont directement désignées dans le modèle physique. L'interprétation des résultats de l'AD permet alors d'améliorer le modèle.

On note $\frac{\partial \mathbf{f}}{\partial \mathbf{z}}|_{\mathbf{z}_{n-1}^a}$ et $\frac{\partial \mathbf{f}}{\partial \varepsilon^m}|_{\mathbf{z}_{n-1}^a}$ les matrices Jacobiennes des dérivées partielles de \mathbf{f}_n au point \mathbf{z}_{n-1}^a . La première est aussi appelée «modèle linéaire tangent». Sénégas *et al.* (2001) calculent la covariance de l'erreur sur l'état entier provoquée par ε^m :

$$\text{Cov}(\mathbf{Z}_n^t - \mathbf{f}(\mathbf{Z}_{n-1}^t, 0), (\mathbf{Z}_n^t - \mathbf{f}(\mathbf{Z}_{n-1}^t, 0))^\top) = \frac{\partial \mathbf{f}}{\partial \varepsilon^m}|_{\mathbf{z}_{n-1}^t} \Sigma^m \frac{\partial \mathbf{f}^\top}{\partial \varepsilon^m}|_{\mathbf{z}_{n-1}^t}. \quad (4.3)$$

L'erreur $\mathbf{Z}_n^t - \mathbf{f}(\mathbf{Z}_{n-1}^t, 0)$ est non stationnaire et dépendante de n . Ceci démontre un aspect pratique important du modèle (4.1) parce qu'il est très difficile techniquement de construire directement des matrices de covariance de grande taille $N \times N$ présentant de telles non stationnarités ou des variabilités temporelles complexes.

4.1 Propagation non linéaire, correction linéaire

On présentera dans un premier temps des solutions pour lesquelles la loi de propagation est non linéaire, mais qui conservent une étape de correction linéaire. Dans ce cas l'algorithme du filtre de Kalman donné au chapitre 3 n'est plus valable. En effet la non linéarité du modèle nous prive de l'hypothèse gaussienne puisque l'image d'une variable gaussienne par une fonction non linéaire n'est plus nécessairement gaussienne. La description de la loi de l'état par ses deux premiers moments est donc seulement approximative. On perd de surcroît la relation donnant l'ébauche de la covariance (3.19) connaissant la covariance avant propagation C_{n-1}^a .

4.1.1 Le filtre de Kalman étendu (EKF)

4.1.1.1 Application directe

La première idée utilisée pour adapter le KF au problème des modèles non linéaires a été celle d'un développement limité au premier ordre du modèle dans un voisinage du dernier état estimé, on obtient le filtre de Kalman étendu (EKF) (Jazwinski, 1970) :

1. Ébauche

$$\mathbf{Z}_n^f = \mathbf{f}_n(\mathbf{Z}_{n-1}^a, 0) \quad (4.4)$$

$$C_n^f = \frac{\partial \mathbf{f}}{\partial \mathbf{z}}|_{\mathbf{z}_{n-1}^a} C_{n-1}^a \frac{\partial \mathbf{f}^\top}{\partial \mathbf{z}}|_{\mathbf{z}_{n-1}^a} + \frac{\partial \mathbf{f}}{\partial \varepsilon^m}|_{\mathbf{z}_{n-1}^a} \Sigma^m \frac{\partial \mathbf{f}^\top}{\partial \varepsilon^m}|_{\mathbf{z}_{n-1}^a}. \quad (4.5)$$

Cette étape nécessite, $2N + 2q$ intégrations du modèle linéaire tangent, soit $2N^3 + 2Nq^2$ opérations. C'est un coût numérique exagéré pour les applications visées. De plus, il y a fort à parier que ces opérations soient redondantes et que certains calculs soient répétés un grand nombre de fois.

2. Correction

$$\mathbf{Z}_n^a = \mathbf{Z}_n^f + K_n(\mathbf{Y}_n - \mathbf{h}(\mathbf{Z}_n^f)) \quad (4.6)$$

$$K_n = C_n^f \frac{\partial \mathbf{h}^\top}{\partial \mathbf{z}} \Big|_{\mathbf{z}_n^f} \left(\frac{\partial \mathbf{h}}{\partial \mathbf{z}} \Big|_{\mathbf{z}_n^f} C_n^f \frac{\partial \mathbf{h}^\top}{\partial \mathbf{z}} \Big|_{\mathbf{z}_n^f} + \Sigma^o \right)^{-1} \quad (4.7)$$

$$C_n^a = C_n^f - K_n \frac{\partial \mathbf{h}}{\partial \mathbf{z}} \Big|_{\mathbf{z}_n^f} C_n^a. \quad (4.8)$$

Un développement limité aux ordres supérieurs donne un filtre qui a certes de meilleures propriétés face aux dynamiques non linéaires (Maybeck, 1979 ; Jazwinski, 1970) mais au prix d'une grande complication des calculs. On le trouve donc rarement appliqué.

L'algorithme de l'EKF ci-dessus a été appliqué à différents modèles océanographiques physiques (Evensen, 1992 ; Gauthier, Courtier, et Moll, 1993 ; Miller, Ghil, et Gauthiez, 1994), mais si le modèle entre dans un régime instable il diverge au sens de la «divergence vraie» de Fitzgerald (1971). En effet, la variance de l'erreur d'estimation augmente indéfiniment là où le flot est plus instable. Evensen (1992) décompose explicitement le modèle linéaire tangent d'un modèle quasi-géostrophique (pourtant presque linéaire) et identifie les termes provoquant cette divergence. Il observe le rôle de la linéarisation des équations dynamiques de propagation de l'erreur ainsi que l'amplification des oscillations à hautes fréquences dans le modèle linéaire tangent.

4.1.1.2 Schémas sous-optimaux (SOS) de l'EKF

Dans les problèmes à grandes dimensions l'application de l'EKF est en général assortie d'un schéma sous optimal (SOS comme *sub-optimal scheme*) afin de réduire son coût numérique qui varie comme le carré de la dimension de l'espace d'état. On distingue alors deux classes de SOS selon que la simplification porte sur le modèle \mathbf{f}_n ou sur la covariance d'erreur C_n . Dans la première classe, Dee (1991) suppose que l'erreur d'estimation suit des règles physiques (divergence nulle et équilibre géostrophique). Fukumori et Malanotte-Rizzoli (1995) écrivent la propagation de l'erreur sur une grille de basse résolution. Cohn et Todling (1996) décomposent en valeurs singulières le modèle linéaire tangent. Cane *et al.* (1996) utilisent une décomposition permanente du modèle et de l'état sur une base orthonormale de l'espace d'état (appelés EOFs en statistiques environnementales) calculée *a priori*. Cette approche semble moins intéressante que les précédentes car d'une part elle dépend de la décomposition initiale de l'espace d'état et d'autre part les sous-espaces propres rejetés initialement ne sont pas autorisés à réapparaître au cours de l'ADS. Elle s'applique

donc idéalement à des modèles linéaires invariants dans le temps. Le *Kriged Kalman Filter* (Wikle et Cressie, 1999 ; Mardia *et al.*, 1998) suit le même principe. Notons que sa dénomination est redondante puisque les filtres de Kalman sont tous équivalents au krigeage.

Dans la seconde classe de SOS plusieurs algorithmes basés sur une approximation en valeurs propres et vecteurs propres, mais cette fois-ci à chaque pas de temps de l'ADS, ont été proposés indépendamment et largement appliqués en océanographie. Ces algorithmes semblent mieux adaptés que celui de Cane *et al.* (1996) aux problèmes de prédictions car la base de projection peut être réorientée suivant la dynamique du système et elle est moins dépendante du choix initial de la base de projection. Il s'agit du filtre de Kalman RRSQRT, *reduced rank square root*, de Verlaan et Heemink (1997) et du filtre SEEK (ou SEIK, version avec échantillonnage de Monte Carlo, apparemment plus proche de l'EnKF exposé ci-dessous), *singular extended evolutive/interpolated Kalman filter*, de Pham *et al.* (1997). Par rapport à l'EKF, ces deux filtres négligent la partie de la variance correspondant aux modes propres de C_n^f de poids le plus faible. Ces deux méthodes étant en principe équivalentes, il m'a semblé préférable d'appliquer le RRSQRT dont l'algorithme est exposé plus clairement. Les articles traitant des filtres SEEK et SEIK (Pham *et al.*, 1997 ; Pham, 2000 ; Hoteit *et al.*, 2001) sont encombrés de nombreuses complications (*forgetting factor*, échantillonnage exact au second ordre, EOF locaux ...) nuisibles à la compréhension du fonctionnement du filtre.

4.1.2 Un exemple de SOS, le filtre de Kalman RRSQRT

Le RRSQRT n'apporte pas de grands changements conceptuels par rapport à l'EKF. On introduit la matrice S dont le carré est une approximation de la matrice de covariance C

$$S_n^f S_n^{f\top} \approx C_n^f, \quad S_n^a S_n^{a\top} \approx C_n^a.$$

Son rang r est bien inférieur à la dimension du modèle d'état (entre 50 et 100 dans la plupart des applications). Ceci permet d'alléger le calcul car on manipule alors une matrice S de dimension $N \times r$ au lieu de C de taille $N \times N$.

4.1.2.1 Ébauche

L'état et la racine carrée de la matrice de covariance sont propagées par le modèle :

$$\mathbf{z}_n^f = \mathbf{f}_n(\mathbf{z}_{n-1}^a, 0) \quad (4.9)$$

$$S_n^f = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{z}} \Big|_{\mathbf{z}_{n-1}^a} S_{n-1}^a, \frac{\partial \mathbf{f}}{\partial \varepsilon^m} \Big|_{\mathbf{z}_{n-1}^a} (\Sigma^m)^{1/2} \right]. \quad (4.10)$$

On vérifie bien que le produit $S_n^f S_n^{f\top}$ redonne l'équation (4.5).

La racine carrée de Σ^m existe puisqu'il s'agit d'une matrice de covariance définie positive¹. Le i -ème vecteur colonne de S_n sera noté \mathbf{s}_n^i (resp. $\mathbf{s}_n^{i,f}$ ou $\mathbf{s}_n^{i,a}$ après la n -ième étape de propagation ou d'analyse); on peut interpréter l'espace engendré par ces vecteurs comme l'ensemble des perturbations de l'état du système dans lequel le filtre de Kalman RRSQRT peut le recalculer. Pour éviter de calculer à chaque pas de temps le modèle linéaire tangent utilisé dans l'équation (4.5), Verlaan et Heemink (1997) utilisent une méthode de différences finies :

$$\frac{\partial \mathbf{f}}{\partial \mathbf{z} | \mathbf{z}_{n-1}^a} \mathbf{s}_{n-1}^{i,a} = \lim_{\epsilon \rightarrow 0} \frac{\mathbf{f}(\mathbf{z}_{n-1}^a + \epsilon \mathbf{s}_{n-1}^{i,a}, 0) - \mathbf{f}(\mathbf{z}_{n-1}^a, 0)}{\epsilon} \quad (4.11)$$

$$\mathbf{s}_n^{i,f} \approx \frac{\mathbf{f}(\mathbf{z}_{n-1}^a + \epsilon \mathbf{s}_{n-1}^{i,a}, 0) - \mathbf{f}(\mathbf{z}_{n-1}^a, 0)}{\epsilon}, \quad (4.12)$$

où le scalaire ϵ est censé être le plus petit possible, mais pas trop car les valeurs faibles provoquent des instabilités du filtre, voir Verlaan (1998) et les applications aux chapitres 5 et 6.

Les vecteurs colonnes de la matrice $(\Sigma^m)^{1/2}$ sont eux aussi propagés par le modèle linéaire tangent suivant une relation similaire à (4.12), en notant $\mathbf{s}^{i,m}$ le i -ème vecteur colonne de $(\Sigma^m)^{1/2}$, qui dans nos notations ne dépend pas de n :

$$\frac{\partial \mathbf{f}}{\partial \boldsymbol{\varepsilon}^m | \mathbf{z}_{n-1}^a} \mathbf{s}^{i,m} = \lim_{\epsilon \rightarrow 0} \frac{\mathbf{f}(\mathbf{z}_{n-1}^a, \epsilon \mathbf{s}^{i,m}) - \mathbf{f}(\mathbf{z}_{n-1}^a, 0)}{\epsilon} \quad (4.13)$$

$$\mathbf{s}_n^{r+i,f} \approx \frac{\mathbf{f}(\mathbf{z}_{n-1}^a, \epsilon \mathbf{s}^{i,m}) - \mathbf{f}(\mathbf{z}_{n-1}^a, 0)}{\epsilon}, \quad (4.14)$$

On peut remarquer que le choix de la racine carrée de Σ^m aura une incidence sur les résultats. On peut s'en convaincre en considérant l'opposé d'une racine carrée de Σ^m et reporter le vecteur colonne $-\mathbf{s}^{i,m}$ dans (4.14). Le filtre RRSQRT «*second order exact*»² (Verlaan et Heemink, 1999) apporte une solution partielle : on ajoute à la base propre tous les vecteurs propres orientés en sens opposé. Verlaan et Heemink (1999) obtient ainsi un filtre RRSQRT plus résistant aux non linéarités. Mais ce filtre ne semble pas encore tenir compte de l'ensemble des racines carrées possibles de Σ^m , sauf dans le cas trivial où cette matrice est diagonale. Je n'emploierai pas cette seconde version du RRSQRT dans la suite pour garder les algorithmes les plus simples.

Dans (4.10), la matrice S_n^f est la jonction des deux matrices entre crochets, de dimensions respectives $N \times r$ et $N \times q$. Sa dimension est alors augmentée ($N \times (r + q)$). Numériquement, cette étape représente $r + q + 1$ intégrations du modèle physique (les $r + q$ colonnes de S_n^f et la moyenne \mathbf{z}_n^f), ce qui réduit considérablement la charge de calcul en comparaison avec l'EKF.

¹Dans les applications cette matrice est de petite taille. Trouver une racine carrée ne pose pas de problème. Celle de Cholesky fait très bien l'affaire.

²Le terme de «second ordre» semble indiquer que le filtre utilise un développement limité au second ordre du modèle dynamique ce qui n'est pas le cas puisque son but est seulement d'améliorer le calcul du moment d'ordre 2 de \mathbf{Z}_n^t ; il prête donc à confusion.

On notera que la relation (4.11) sert à approcher l'équation (4.5) qui est elle-même issue du développement au premier ordre de (4.1). Cette étape de l'algorithme est donc la plus critiquable.

4.1.2.2 Réduction

On réduit la dimension de S_n^f de $N \times (r + q)$ à $N \times r$ par orthogonalisation en négligeant les q valeurs propres les plus faibles :

$$S_n^{f\top} S_n^f = UDU^\top \quad (4.15)$$

$$S_n^{f*} = [UD^{1/2}]_{1:N,1:r} \quad (4.16)$$

On note dans l'équation (4.15) que ce n'est pas la matrice C_n^f de taille $N \times N$ qui est diagonalisée car le produit matriciel a commuté. On diagonalise ainsi une matrice de taille plus modeste $(r + q) \times (r + q)$. Cette manipulation est autorisée car le rang de C_n^f est au plus de $r + q$ et les valeurs propres du produit commuté ne diffèrent que par $N - (r + q)$ valeurs propres nulles.

Numériquement cette étape revient donc à deux produits matriciels et une décomposition en valeurs propres et vecteurs propres d'une matrice $(r + q) \times (r + q)$. Pour cette décomposition, les routines classiques de la librairie Fortran `eispack` sont mises en œuvre : la matrice est d'abord rendue tridiagonale, puis diagonale par une méthode *QR* itérative. Le nombre d'opérations dépend donc de la précision demandée, mais il est courant de supposer que le nombre d'itérations n'augmente pas avec la taille de la matrice à diagonaliser (Verlaan, 1998). Ce nombre d'opération est alors un $\mathcal{O}((r + q)^2)$.

En tout, le coût numérique associé aux trois calculs est de $N(r + q)^2 + \mathcal{O}((r + q)^2) + 2Nr(r + q)$ opérations, en tenant compte de la symétrie de $S_n^{f\top} S_n^f$, et il varie donc comme le carré du nombre de modes retenus r .

Dans le cas multivariable la réduction pourrait rendre l'ADS dépendante des unités des différentes variables si elle coupait systématiquement les modes liés aux variables exprimées dans les unités les plus fortes. Il convient donc de diagonaliser des matrices de corrélation au lieu de matrices de covariances, comme dans Segers *et al.* (2000).

4.1.2.3 Correction

L'algorithme du RRSQRT utilise le schéma de correction scalaire de Potter, dont on trouve aussi la justification dans (Maybeck, 1979). Ce schéma suppose que la matrice $HC_n^f H^\top + \Sigma^o$ est diagonale. Si ce n'est pas le cas³, on peut procéder au préalable à une orthogonalisation de cette matrice. L'algorithme de

³C'est-à-dire si les observations sont trop rapprochées au sens de la covariance C_n^f

Potter procède au calcul du gain de Kalman et à la mise à jour de S_n^{f*} observation par observation. Si on omet l'indice de pas de temps n et si on note \mathbf{k}_i la i -ème colonne de K pour simplifier les notations, on a :

$$\text{do loop } i = 1 : m \quad (4.17)$$

$$\alpha_i = (H_{(i,\cdot)} S^{f*})^\top \quad (4.18)$$

$$\beta_i = \frac{1}{\alpha_i^\top \alpha_i + \Sigma_{i,i}^o} \quad (4.19)$$

$$\mathbf{k}_i = \beta_i S^{f*} \alpha_i \quad (4.20)$$

$$S^a = S^{f*} - \frac{\mathbf{k}_i \alpha_i^\top}{1 + \sqrt{\beta_i \Sigma_{i,i}^o}} \quad (4.21)$$

$$\text{end loop } i. \quad (4.22)$$

Ayant la matrice de gain K , on peut alors appliquer l'équation linéaire de correction de l'état du filtre de Kalman (3.22). Le coût numérique de cette étape est environ $6mrN$, auquel on ajoute un $\mathcal{O}(m^2)$ s'il faut diagonaliser la matrice $HC^f H^\top + \Sigma^o$. Dans nos applications où le nombre d'observations est très faible, ce coût est négligeable face à celui des deux autres étapes.

Pour les trois étapes du RRSQRT, il faut noter que la réduction du rang de S_n^f provoque une sous-estimation systématique de la variance d'estimation, mais dans le cadre linéaire à termes constants, il est possible de prouver la convergence asymptotique de la covariance estimée par le RRSQRT vers celle du filtre de Kalman de rang plein, lui-même convergent dans ce cas d'école (Verlaan, 1998). Dans le cas général, la convergence du filtre de Kalman RRSQRT n'est pas garantie, elle est cependant constatée dans de nombreux cas réels de grandes dimensions et souvent non linéaires.

4.1.3 L'ensemble Kalman filter (EnKF)

L'EnKF doit être mis à part des SOS ci-dessus car il n'est pas basé sur l'EKF pour l'ébauche, on doit toutefois supposer que les lois conditionnelles de l'état et des observations sont gaussiennes pour pouvoir appliquer une étape de correction linéaire. On verra plus loin des approches de Monte Carlo à la fois pour l'ébauche et la correction.

4.1.3.1 Ébauche

L'EnKF utilise une méthode de Monte Carlo pour résoudre l'équation stochastique générale de propagation d'un modèle non linéaire (4.1). La densité $\phi(\mathbf{z}_n^t)$ est approchée par un échantillon de r états simulés du système $\mathbf{z}_n^{1:r}$, dont l'évolution dans le temps est donnée par une **simulation non conditionnelle** de l'erreur de modèle :

$$\forall i = 1 : r, \mathbf{Z}_n^{i,f} = \mathbf{f}_n(\mathbf{Z}_{n-1}^{i,a}, \varepsilon_n^{i,m}), \quad (4.23)$$

ce qui nécessite r propagations dans le temps du modèle et autant de simulations des erreurs de modèle ε^m , mais sans devoir le linéariser comme en EKF. On peut calculer sur l'échantillon deux premiers moments empiriques (moyenne \mathbf{z}_n^f , covariances C_n^f) que l'on va utiliser dans la suite de l'algorithme.

4.1.3.2 Correction

On suppose ici que les densités $\phi(\mathbf{y}_n | \mathbf{z}_n^f)$ et $\phi(\mathbf{z}_n^t | \mathbf{z}_n^f)$ sont gaussiennes, comme c'était le cas dans le filtre de Kalman. Par le même raisonnement, on a de nouveau

$$\mathcal{L}(\mathbf{z}_n^t | \mathbf{y}_n, \mathbf{z}_n^f) \sim \mathcal{N}(\mathbf{z}_n^a, C_n^a), \quad (4.24)$$

avec \mathbf{Z}_n^a , K_n et C_n^a vérifiant les mêmes équations que dans le filtre de Kalman (3.22), (3.27) et (3.28). Evensen (1994), repris par Burgers *et al.* (1998), propose un algorithme de **simulations conditionnelles** de \mathbf{Z}_n^t connaissant sa moyenne \mathbf{Z}_n^f (celle de l'ensemble) et les données \mathbf{Y}_n de lois gaussiennes :

$$\mathbf{Z}_n^{i,a} = \mathbf{Z}_n^{i,f} + K_n(\mathbf{y}_n - H\mathbf{Z}_n^{i,f} + \varepsilon_n^{i,o}), \quad (4.25)$$

où chaque membre d'indice $i = 1 : r$ est corrigé indépendamment des autres, et les $\varepsilon_n^{i,o}$ sont r simulations non conditionnelles de la variable $\varepsilon^o \sim \mathcal{N}(0, \Sigma^o)$. Il est aisé de vérifier que les deux premiers moments de la variable simulée tendent bien vers \mathbf{z}_n^a et C_n^a à la limite quand r tend vers l'infini. Cet algorithme ressemble ainsi à l'algorithme classique de simulations conditionnelles utilisé en géostatistique (Lantuejoul, 2001), page 200 :

- le terme $K_n \mathbf{y}_n$ est l'estimateur du krigeage simple des données conditionnantes,
- $\mathbf{Z}_n^{i,f}$ est une simulation non conditionnelle d'une variable gaussienne de moyenne \mathbf{z}_n^f et de covariance C_n^f sur tout le domaine,
- $K_n(H\mathbf{Z}_n^{i,f} - \varepsilon_n^{i,o})$ est le krigeage simple de la simulation non conditionnelle prise aux points conditionnants (avec simulation du bruit de mesure $\varepsilon_n^{i,o}$).

4.1.4 Comparaison entre le RRSQRT et l'EnKF

4.1.4.1 Aspects théoriques

L'EnKF doit être plus performant que le RRSQRT pour l'étape d'ébauche car il est basé sur la loi de propagation non linéaire ce qui évite le développement limité mis en œuvre dans l'EKF (et en particulier dans le RRSQRT) pour propager l'erreur d'estimation. Cette particularité le rend visiblement mieux adapté aux problèmes fortement non linéaires que les autres filtres actuels (Verlaan et Heemink, 1999 ; Pham, 2000). Par contre, les deux méthodes utilisent des corrections linéaires similaires à celles du filtre de Kalman. L'algorithme de correction de Potter est une traduction exacte de celui du filtre de Kalman pour les racines carrées S des matrices de covariances C .

Pour des questions plus appliquées, l'approximation faite par le RRSQRT en tronquant les vecteurs propres de poids faible semble couper systématiquement les composantes à courte portée de la covariance (Hoteit *et al.*, 2001), à la manière d'un lissage, ce qui n'est pas le cas de l'approximation de type Monte Carlo de l'EnKF. Ces deux approximations provoquent une sous-estimation de la variance de l'erreur optimale. On peut compenser en augmentant le paramètre r (Verlaan, 1998 ; Evensen, 1994). On notera aussi qu'il est plus facile d'assurer les propriétés physiques de l'état estimé dans le cadre de l'EnKF que du RRSQRT car pour ce dernier l'étape de réduction est basée sur un critère statistique qui ne tient pas compte des contraintes physiques du système. Au contraire, les perturbations aléatoires ajoutées par l'EnKF sont directement issues du modèle d'erreur et les responsabilités en cas de rupture dans le modèle physique sont donc plus faciles à établir.

4.1.4.2 Aspects algorithmiques

Verlaan et Heemink (1999) font ressortir des similitudes entre les deux algorithmes et c'est ce que j'ai voulu faire aussi ici en notant r à la fois le nombre de vecteurs propres retenus par le RRSQRT et la dimension de l'ensemble de l'EnKF, avec une réserve sur cette notation : si r désigne dans les deux cas le nombre d'évaluations du modèle nécessaires à chaque pas de temps, on ne peut en revanche pas confondre un vecteur propre avec un membre d'ensemble. Ce sont en effet deux entités qui ont deux significations différentes en statistiques.

De ce point de vue, l'inconvénient de l'approche de Monte Carlo est que les membres de l'ensemble sont pris au hasard et ils peuvent alors transporter une information redondante (réalisations quasi-similaires), alors que dans les approches utilisant une décomposition en valeurs propres et vecteurs propres, cette redondance est en principe évitée. On peut en déduire que l'EnKF nécessitera un grand nombre de réalisations, ce qui peut le rendre plus «gourmand» numériquement que ses concurrents. Dans les applications en cas réels, la taille de l'ensemble est quasi-systématiquement mise à $r = 100$, ce qui paraît faible pour un algorithme de

Monte Carlo (l'erreur commise ne décroît qu'en $\mathcal{O}(1/\sqrt{r})$). Mais dans les applications cela semble suffisant pour les performances demandées à l'ADS.

Cañizares (1999) étudie l'évolution des temps de calculs demandés par l'EnKF et le RRSQRT sur un même cas d'étude (un modèle hydrodynamique en 2-D) : à performances égales, le paramètre r doit être plus élevé pour l'EnKF que pour le RRSQRT, mais le temps de calcul de la diagonalisation (4.15) nécessaire en RRSQRT lui fait perdre cet avantage, et au final les temps de calculs demandés par les deux méthodes sont équivalents. On trouvera aussi une comparaison des deux algorithmes dans le cas d'un modèle écologique 1-D dans le chapitre 5.

4.2 Propagation et correction non linéaires

Le filtre de Kalman est basé sur une interpolation linéaire des données vérifiant un critère de moindres carrés. Cette analyse est bel et bien optimale dans le cadre linéaire du chapitre 3 mais pas dans les cas réels. Il y a deux raisons à cela. L'une est la non linéarité du modèle physique qui ne conserve pas le caractère gaussien des variables d'état. L'autre est la non linéarité des fonctions d'observation $\mathbf{Y} = \mathbf{h}(\mathbf{Z})$ assez fréquente en océanographie (on pourrait citer le cas des observations de *sea surface height* SSH ou de chlorophylle). On esquive généralement ce second problème en incorporant les variables observées au vecteur d'état.

4.2.1 Le cas gaussien anamorphosé

4.2.1.1 Quelle anamorphose ?

On considère dans cette section un problème que l'on peut ramener au filtre de Kalman. Supposons que les variables d'état et les observations soient les images de variables gaussiennes \mathbf{U} et \mathbf{V} par des transformations non linéaires bijectives :

$$\mathbf{U}_n^t = \psi_n(\mathbf{Z}_n^t) \quad (4.26)$$

$$\mathbf{V}_n = \chi_n(\mathbf{Y}_n). \quad (4.27)$$

Ces deux fonctions (vectorielles) ne sont pas indépendantes car elles doivent vérifier la relation d'observation

$$\mathbf{Y}_n = \mathbf{h}(\mathbf{Z}_n^t),$$

et de plus on suppose pour simplifier que l'opérateur d'observation liant les variables gaussiennes $\mathbf{V}_n = \chi_n \circ \mathbf{h} \circ \psi_n^{-1}(\mathbf{U}_n^t)$ est linéaire.

En particulier, si l'opérateur d'observation est simplement une projection π de \mathbb{R}^N sur \mathbb{R}^m , alors il est naturel de choisir la même projection (aux domaines de définition près) pour lier χ_n et ψ_n :

$$\chi_n = \pi \circ \psi_n.$$

Les deux fonctions d'anamorphoses pourront alors être calculées à chaque pas de temps sur la base des histogrammes des variables d'état (ou mieux, des observations si elles sont en nombre suffisant). On pourra préférer des versions constantes dans le temps de ψ et χ pour plus de stabilité.

4.2.1.2 Estimation des moments de \mathbf{Z}_n^t

Si ces transformations sont des exponentielles (variables \mathbf{Z} et \mathbf{Y} lognormales), et si on suppose que la log-normalité des variables est conservée aux différentes échelles (ponctuelles et blocs), il est possible en théorie d'estimer les deux premiers moments de \mathbf{Z}_n^t par un krigeage lognormal (Rivoirard, 1990), et d'étendre ainsi le filtre de Kalman aux modèles lognormaux. Cependant cette idée semble compromise car le succès du krigeage lognormal est très dépendant de la covariance *a priori* (Roth, 1998) qui est plutôt mal estimée dans les applications habituelles du filtre de Kalman.

Il est donc plus sage de procéder par approximation de Monte Carlo pour obtenir un estimateur sans biais, ce qui a de plus l'avantage de s'appliquer à des anamorphoses ψ_n et χ_n quelconques mais liées par la relation d'observation. On tire alors r échantillons de la loi *a priori* :

$$\mathbf{z}_n^{f,i} \sim \mathcal{L}(\mathbf{z}^t | \mathbf{z}_n^f), \quad i = 1 : r \quad (4.28)$$

ce qui est simple pour la plupart des méthodes issues de l'EKF, et qui est déjà fait dans le cas de l'EnKF. On applique ensuite les fonctions d'anamorphose :

$$\mathbf{u}_n^{f,i} = \psi_n(\mathbf{z}_n^{f,i}) \quad (4.29)$$

$$\mathbf{v}_n = \chi_n(\mathbf{y}_n), \quad (4.30)$$

et on peut appliquer l'étape de correction linéaire du filtre de Kalman classique en utilisant les statistiques de \mathbf{U}_n^f calculées sur l'échantillon.

Inversement pour estimer les moments *a posteriori* de \mathbf{Z}_n^t , on doit disposer d'échantillons de la loi *a posteriori* de \mathbf{U} . L'étape de correction devient donc une simulation conditionnelle de $\mathcal{L}(\mathbf{u}_n^t | \mathbf{u}_n^f, \mathbf{v}_n)$ et puisque toutes ces variables sont gaussiennes, il est judicieux d'utiliser la même méthode de simulation que dans l'EnKF (4.25) appliquée cette fois aux variables \mathbf{U} et \mathbf{V} . Après anamorphose inverse on a bien :

$$\mathbf{z}_n^{a,i} = \psi_n^{-1}(\mathbf{u}_n^{a,i}) \quad (4.31)$$

$$\sim \mathcal{L}(\mathbf{z}_n^t | \mathbf{z}_n^f, \mathbf{y}_n). \quad (4.32)$$

Cette méthode peut permettre en particulier d'éviter les valeurs négatives après correction en ADS. On peut en voir une application à un cas fictif dans le chapitre 5.

4.2.2 Méthodes de Monte Carlo alternatives

Cette partie est tirée du livre édité par Doucet, de Freitas, et Gordon (2001) qui résume très clairement les avancées théoriques faites en dehors de l'hypothèse gaussienne. Ce livre fournit aussi des preuves de convergence de certains algorithmes exposés ci-dessous. Ces méthodes sont libérées de l'hypothèse gaussienne et ne font plus partie de la catégorie des filtres de Kalman. Leur principe est d'échantillonner les distributions apparaissant dans la formule de Bayes (2.8) afin d'obtenir par un calcul intégral approché les statistiques de la loi de lissage $\mathcal{L}(\mathbf{Z}_{1:n}|\mathbf{Y}_{1:n})$ et en particulier celles de la loi de filtrage $\mathcal{L}(\mathbf{Z}_n|\mathbf{Y}_{1:n})$. Ce formalisme Bayésien est certainement un cadre plus malléable que celui du filtre de Kalman pour traiter les problèmes de biais et d'estimation de paramètres inconnus car il est en théorie possible de randomiser n'importe quel paramètre additionnel.

Malheureusement, on peut rarement échantillonner directement les lois apparaissant dans la relation de Bayes et il faut s'atteler à des méthodes de Monte Carlo alternatives. En traitement du signal, l'attention s'est récemment portée sur l'échantillonnage d'importance, que l'on peut aussi rendre séquentiel (SIS comme *sequential importance sampling*). Mais l'algorithme SIS seul n'est pas convergent et on doit lui ajouter une étape supplémentaire de sélection. De nombreuses méthodes ont vu le jour, je me contenterai ici d'une description rapide et de mettre en valeur les potentialités pour l'ADS océanographique. A ce jour la seule application de ces méthodes de Monte Carlo en ADS océanographique est celle de van Leeuwen (2001) pour un modèle physique de l'océan autour de l'Afrique du Sud, aux dynamiques fortement non linéaires.

4.2.2.1 Echantillonnage d'importance

La technique de l'échantillonnage d'importance permet de calculer les statistiques d'une variable $\mathbf{Z} \sim \phi(\mathbf{z})$ à l'aide d'une densité $\pi(\mathbf{z})$, appelée «fonction d'importance», plus facile à simuler. Soit une fonction \mathcal{M} intégrable (en jouant sur \mathcal{M} on calcule n'importe quelle statistique de \mathbf{Z}), on a :

$$E_{\phi}(\mathcal{M}(\mathbf{Z})) = \int \mathcal{M}(\mathbf{z})\phi(\mathbf{z})d\mathbf{z} \quad (4.33)$$

$$= \int \mathcal{M}(\mathbf{z})\frac{\phi(\mathbf{z})}{\pi(\mathbf{z})}\pi(\mathbf{z})d\mathbf{z} \quad (4.34)$$

$$= E_{\pi}\left(\mathcal{M}(\mathbf{Z})\frac{\phi(\mathbf{Z})}{\pi(\mathbf{Z})}\right). \quad (4.35)$$

Sous certaines conditions sur π (Doucet *et al.*, 2001), les échantillons de la loi π , pondérés par les ratios ci-dessus (à une constante de normalisation près) permettent ainsi d'estimer $E_{\phi}(\mathcal{M}(\mathbf{Z}))$ sans biais et avec une vitesse de convergence indépendante de la dimension de l'espace d'état. On applique ensuite ce principe à la loi de filtrage $\phi(\mathbf{z}_n|\mathbf{y}_{1:n})$ difficile à simuler⁴.

⁴Il en va de même pour la loi de lissage $\phi(\mathbf{z}_{1:n}|\mathbf{y}_{1:n})$.

La méthode d'intégration par échantillonnage d'importance n'est pas récursive sous sa forme générale, mais elle peut devenir récursive (SIS) sous une condition supplémentaire sur π (Doucet *et al.*, 2001). Cette condition est vérifiée dans le cas particulier où la loi π est la loi *a priori* $\phi(\mathbf{z}_n | \mathbf{y}_{1:n-1})$, cette fonction d'importance est aussi celle utilisée par van Leeuwen (2001). Mais ce choix n'est pas nécessairement le plus judicieux : l'efficacité de l'algorithme est plus faible dans le cas où ces deux distributions sont très dissemblables, ce qui correspond, hélas, aux cas les plus utiles.

Selon Zaritskii *et al.*, cité par Doucet, de Freitas, et Gordon (2001), si on note $\mathbf{z}_{1:n}^i$ la i -ème trajectoire (ou «particule») simulée, la fonction d'importance optimale (celle qui limite le mieux la dégénérescence de l'algorithme) est la densité $\phi(\mathbf{z}_n | \mathbf{z}_{n-1}^i, \mathbf{y}_n)$, déjà conditionnée aux données actuelles. On ne peut donc l'utiliser seulement si l'on sait déjà simuler \mathbf{Z}_n conditionnellement à \mathbf{Y}_n . Cette fonction d'importance optimale mène aussi à un algorithme récursif, (Bui, 2001) l'utilise en épidémiologie pour la prédiction de cas de grippe.

En ADS océanographique, la seule fonction d'importance dont on dispose simplement est la loi *a priori*. On pourrait aussi simuler la densité *a posteriori* par une méthode itérative comme l'échantillonneur de Gibbs, ce qui donnerait lieu à des expérimentations intéressantes.

4.2.2.2 SIS avec ré-échantillonnage

Le phénomène de dégénérescence de l'algorithme SIS est classiquement reconnu, il est plus marqué dans les espaces à grande dimensions et plus ou moins rapide suivant le choix de la fonction d'importance. Pratiquement, les pondérateurs de l'échantillonnage d'importance deviennent de plus en plus déséquilibrés et à terme, une seule particule prend un poids d'importance non nul. Le filtre SIS ne peut donc pas reproduire la distribution *a posteriori* si l'on ne rajoute pas une étape de rééchantillonnage (on trouve alors ce SIS modifié sous la dénomination de *sequential importance resampling* (SIR). L'idée principale est de supprimer les particules de poids faible et de reproduire celles de poids fort, l'analogie avec le principe de la sélection naturelle peut aussi donner lieu à des rapprochements avec les algorithmes génétiques qui ont, eux aussi, été utilisés en océanographie (Athias *et al.*, 2001).

Il y a plusieurs manières de rééchantillonner les particules,

- la plus courante revient à un *bootstrap* : elle consiste à faire r tirages aléatoires de la distribution discrète à r particules obtenue par échantillonnage d'importance. Les particules de poids fort peuvent ainsi ressortir plusieurs fois alors que celles de poids faible peuvent ne pas ressortir du tout et «mourir». Le filtre bootstrap garde donc un nombre constant de particules, l'inconvénient pratique de cette méthode est la perte de «diversité» qui intervient quand les particules de poids relativement faible ont le malheur de ne pas survivre. Elle nécessite aussi un grand nombre de particules pour assurer que celles de poids fort seront bien représentées, alors qu'en ADS océanographiques, vus les problèmes de temps de calcul on cherchera plutôt à limiter le nombre de particules au minimum nécessaire.

Plusieurs alternatives ont ainsi vu le jour, en lien avec les algorithmes génétiques pour les problèmes dits de «diversité génétique».

- L'une, celle utilisée par van Leeuwen (2001) pour limiter le nombre de particules et le temps de calcul, peut être qualifiée d'«élitiste». Elle consiste à reproduire r fois le poids de chaque particule, et tirer aléatoirement la partie décimale restante : si⁵ on a 100 particules, ce qui est peu pour une méthode de Monte Carlo, une particule de poids 0.115 aura un nouveau poids de $100 \times 0,112 = 11,2$. Elle sera retirée obligatoirement 11 fois et les 0,2 restant seront rééchantillonnés. Elle aura donc une chance sur cinq d'être reproduite 12 fois et quatre chances sur cinq d'être reproduite 11 fois. De même, une autre particule de poids 0,005 aura un nouveau poids de 0,5 et une chance sur deux de survivre. Cette procédure semble donc plus égalitaire car elle assure la bonne reproduction des particules de poids fort alors qu'avec l'algorithme classique utilisant seulement 100 particules, il se peut que la particule de poids fort soit sous-représentée. Au total le nombre de particules peut augmenter pendant cette phase de rééchantillonnage, ce qui peut alourdir la charge de calcul et il semble préférable pour des applications en ADS d'éviter ceci par quelque modification supplémentaire.
- D'autres variantes du filtre bootstrap existent en perturbant aléatoirement la distribution sortante par une méthode de *fixed kernel*, cette dernière pouvant se rapprocher à la fois du filtre à particules (del Moral et Salut, 1995) et du filtre de Bayes (Miller *et al.*, 1999), aussi utilisé par Liabaud (2001). Les extensions du filtre bootstrap suscitent actuellement un grand engouement. Elles sont exposées dans Doucet *et al.* (2001), chapitres 4 à 14. La plus efficace semble être un technique arborescente dotée de nombreuses propriétés favorables (conservation du nombre de particules, respect de la moyenne de l'ensemble).

Les avantages des méthodes de Monte Carlo SIR sont nombreux : elles évitent les approximations gaussiennes, utilisent directement la formule de Bayes (2.8) et sont donc plus générales que les filtres de Kalman qui reposent essentiellement sur l'hypothèse gaussienne. En particulier on est libre de choisir pour les erreurs de mesure une loi avec des queues de distribution plus longues que la loi gaussienne, on obtient ainsi un filtre plus robuste aux valeurs aberrantes. Par contre l'efficacité des algorithmes SIR dépend de deux choix : celui d'une fonction d'importance assez bien conditionnée aux données et celui d'une procédure de rééchantillonnage adéquate. La comparaison entre filtres de Kalman et SIR devrait être instructive mais à ma connaissance elle n'a pas encore été faite sur des systèmes régionalisés.

En comparaison avec les algorithmes génétiques, le SIR comporte les processus de «mort-reproduction» avec le rééchantillonnage (multiplication ou élimination des individus par rapport à un critère d'adaptation : dans notre cas ce critère est la valeur de la densité $\phi(\mathbf{y}_n | \mathbf{z}_n^i)$). Il comporte aussi un aspect «mutation» avec l'erreur de modèle aléatoire, mais pas de «croisement» car les individus ne s'échangent pas d'information. Ce dernier point semble difficile à mettre en œuvre dans le cadre de l'ADS océanographique car le fait d'échanger abruptement des portions de mers ou de lagune entre deux états du système peut créer des discontinuités et des artifices gênants. On peut imaginer que les croisements portent seulement sur les paramètres non dimensionnels mal connus du système (coefficients de viscosité, de friction, ...).

⁵En reprenant l'exemple de van Leeuwen

4.2.3 Effet de support dans un cadre non gaussien

Les modèles de changement de support ont déjà été proposés et appliqués à des variables non gaussiennes (notamment la loi gamma dans Hu (1998)). On trouvera une revue de ces méthodes dans Chilès et Delfiner (1999). La difficulté particulière à l'ADS vient de la non stationnarité des covariances utilisées $C_n^f = (\text{Cov}(\mathbf{Z}_n^f(V_i), (\mathbf{Z}_n^f(V_j))))_{i,j}$ qui nous ramène au cas non linéaire et non stationnaire, *terra incognita* de la géostatistique. Lajaunie (communication personnelle) suggère de procéder à une anamorphose gaussienne et de traiter la variable transformée comme une variable stationnaire.

4.3 Aspects pratiques de l'ADS multivariable

4.3.1 Normalisation

Lorsque le calcul du gain de Kalman est fait par une méthode approchée (décomposition en valeurs propres et vecteurs propres principalement) et que les variables observées sont exprimées dans des unités différentes, il convient de normaliser chaque variable pour éviter de discriminer systématiquement les variables exprimées dans les unités les plus fortes. La norme choisie usuellement est quadratique : on ôte à chaque variable sa moyenne et on la divise par son écart-type expérimental.

4.3.2 Corrections successives ou jointes ?

On trouve fréquemment dans les cas pratiques d'ADS multivariable (Evensen *et al.*, 2001 ; Natvik et Evensen, 2001) des approches en corrections successives (une variable observée à la fois) qui se basent sur l'indépendance conditionnelle des mesures faites par des instruments différents. Cette indépendance conditionnelle peut s'exprimer $\phi(\mathbf{y}_1|\mathbf{z}^t, \mathbf{y}_2) = \phi(\mathbf{y}_1|\mathbf{z}^t)$ et on peut écrire l'équation de Bayes comme suit :

$$\phi(\mathbf{z}^t|\mathbf{y}_1, \mathbf{y}_2, \mathbf{z}^f) = \frac{\phi(\mathbf{y}_1|\mathbf{z}^t)}{\int_{\mathbb{R}^N} \phi(\mathbf{y}_1|\mathbf{z})\phi(\mathbf{z}|\mathbf{y}_2, \mathbf{z}^f) d\mathbf{z}} \phi(\mathbf{z}^t|\mathbf{y}_2, \mathbf{z}^f). \quad (4.36)$$

Si l'on est capable d'obtenir la loi conditionnelle à \mathbf{Y}_1 on peut pratiquer une ADS multivariable en appliquant les corrections monovariabiles successivement, ce qui revient à les appliquer conjointement. Si au sein d'une même variable les erreurs de mesure sont conditionnellement indépendantes, on peut de même prendre en compte chaque mesure une par une.

Pour un filtre de Kalman, procéder par corrections successives peut s'écrire dans le cas bivariable :

$$\mathbf{z}_n^{a,1} = \mathbf{z}_n^f + K_n^1(\mathbf{y}_{1,n} - H_{(1,)}\mathbf{z}_n^f) \quad (4.37)$$

$$K_n^1 = C_n^f H_{(1,)}^\top (H_{(1,)} C_n^f H_{(1,)}^\top + \Sigma_{(1,1)}^o)^{-1} \quad (4.38)$$

$$\mathbf{z}_n^a = \mathbf{z}_n^{a,1} + K_n^2(\mathbf{y}_{2,n} - H_{(2,)}\mathbf{z}_n^{a,1}) \quad (4.39)$$

$$K_n^2 = C_n^{a,1} H_{(2,)}^\top (H_{(2,)} C_n^{a,1} H_{(2,)}^\top + \Sigma_{(2,2)}^o)^{-1}. \quad (4.40)$$

On limite ainsi la taille des matrices à inverser, ce qui peut s'avérer utile lorsque les mesures sont nombreuses. Par contre il ne faut pas en espérer des économies de temps de calcul miraculeuses car on doit corriger à la fois l'état et la covariance et les corrections successives de la covariance sont des opérations coûteuses.

Deuxième partie

Applications

Chapitre 5

Application à un modèle écologique simple

Sommaire

On teste deux filtres de Kalman sur un modèle jouet et on les trouve tous deux efficaces et équivalents. L'EnKF fait l'objet d'une extension lognormale mieux adaptée aux cas des variables positives.

5.1 Le modèle

5.1.1 Interactions

Evans et Parslow (1985) ont décrit les cycles annuels du plancton et leur floraison printanière (*spring bloom*) par un modèle écologique très simple sans dimensions. Trois variables interagissent dans des équations différentielles, les nutriments (N), le phytoplancton (P) et les herbivores (H). Les trois concentrations seront exprimées dans la même unité, leur équivalent en millimole d'azote par mètre cube (mMol N.m^{-3}). Les herbivores broutent le phytoplancton et sont supposés avoir une mortalité constante. Les nutriments sont assimilés par le phytoplancton lors de la photosynthèse et la croissance du phytoplancton est fonction de son exposition au soleil. Les conditions d'ensoleillement et la profondeur de la couche bien mélangée varient au cours de l'année mais leurs variations sont identiques d'une année sur l'autre. Tous les paramètres ont été calibrés dans (Evans et Parslow, 1985) de manière à reproduire des données écologiques mesurées au large de Terre-Neuve dans l'océan Atlantique. La discrétisation temporelle du modèle découpe l'année en

400 pas de temps égaux.

Ce modèle à été étendu à une dimension spatiale par Eknes et Evensen (2001) et des termes de diffusion verticale ont été ajoutés aux équations du modèle. Les caractéristiques de ce modèle sont conservées ici : 20 cellules verticales de 10 m de hauteur, ce qui donne une colonne d'eau de 200 m de profondeur. Breuillin, Evensen, et Eknes (2000) lui a appliqué l'EnKS (*Ensemble Kalman Smoother*) et a appliqué l'EnKF et l'EnKS à un modèle 1-D moins rudimentaire avec 7 variables.

5.1.2 Conditions aux limites

Les conditions aux limites spécifient qu'il n'y a aucun échange de matière à la surface de l'eau et qu'il entre par le fond à chaque pas de temps 10 mMol N.m⁻³ de nutriments, ce qui simule un influx idéal constant venant des sédiments. Les concentrations en phytoplancton et en herbivores au fond de l'eau sont mises à zéro. Ce modèle fait scandale parmi les biologistes car il n'impose aucun bilan de matière et il comporte trop peu de variables pour décrire avec réalisme un écosystème. Cependant il comporte déjà quelques inconvénients caractéristiques des «vrais » modèles écologiques¹ (instabilité, sensibilité aux paramètres) et certains résultats pourront être étendus aux modèles écologiques en général.

5.1.3 Run de référence \mathbf{z}_n^t

Une année de référence a été simulée par ce modèle en commençant par des conditions initiales de 10 mMol N.m⁻³ pour les nutriments et de 0.1 mMol N.m⁻³ à la fois pour le phytoplancton et les herbivores. Le modèle a ensuite simulé cinq ans d'activité écologique de la colonne d'eau (*spin up*). C'est à partir de ces cinq ans que les cycles annuels se reproduisent très précisément. On notera \mathbf{z}_n^t les «vraies » valeurs de référence par la suite.

Sur la figure 5.1 on voit la colonne d'eau remplie de nutriments au début de l'année, puis à partir d'Avril (100-ième jour) sous l'influence de l'éclairement solaire, c'est la floraison printanière : les nutriments sont consommés par le phytoplancton qui commence à croître jusqu'à ce qu'il n'en reste plus qu'au fond de la colonne d'eau, les herbivores viennent ensuite brouter le phytoplancton avec un mois de retard environ, ce qui régule leur croissance. A la fin de l'année le système retourne à son état initial. Les deux espèces (P et H) restent cantonnées dans la moitié supérieure de la colonne d'eau, la plus exposée au soleil.

¹Evans et Parslow (1985) diront à propos des modèle écologiques *as their complexity grows, they can behave as mysteriously as the real world.*

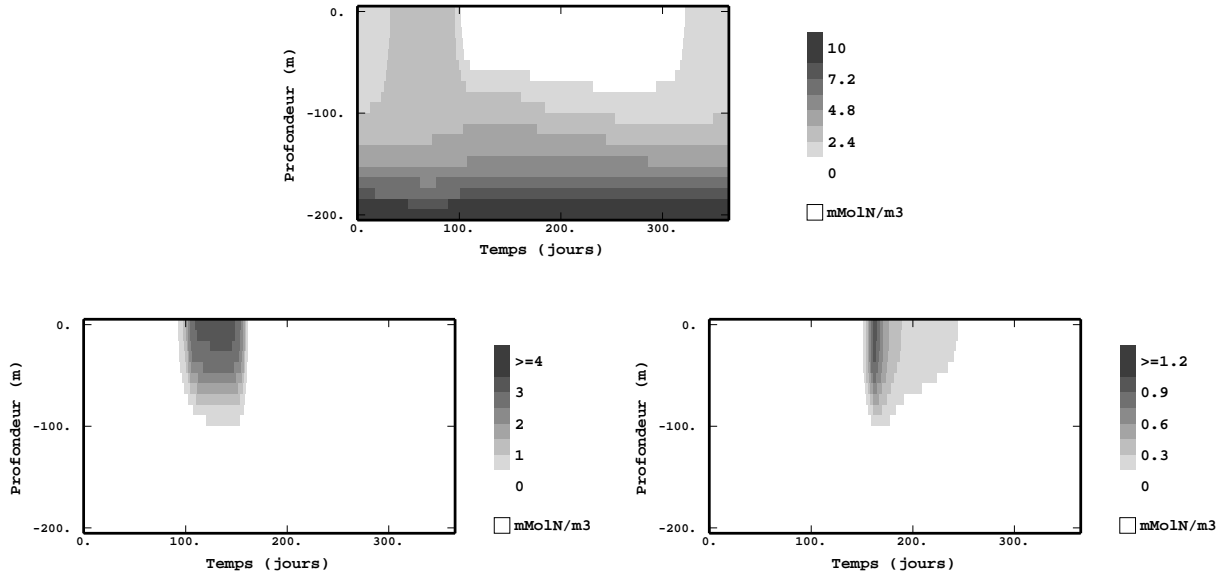


FIG. 5.1 – Cycle annuel de référence : évolution temporelle des variables Nutriments (en haut), Phytoplancton (en bas à gauche) et Herbivores (en bas à droite)

En considérant ce *run* de référence comme l'état réel du système, Eknes et Evensen ont ensuite assimilé des données fictives et ont montré qu'il est possible de contrôler ce système sous différentes conditions d'observation. Le résultat majeur étant le succès de l'assimilation de données de phytoplancton en surface, ce qui laisse présager la possibilité d'utiliser des données satellites de couleur de l'océan (qui sont, hélas, des données de surface) pour corriger les modèles écologiques.

5.1.4 Stabilité du modèle

Il est un aspect commun à la plupart des modèles écologiques : les constantes cinématiques chimiques et biologiques sont modélisées par l'hyperbole de Michaelis. Dans le cas non saturé (i.e. pour la croissance du phytoplancton, le cas où l'eau n'est pas saturée de nutriments) ces constantes sont directement proportionnelles à la concentration de substrat. Le modèle produit alors des croissances exponentielles de population et les dynamiques peuvent être fortement non linéaires.

Le modèle écologique utilisé ici est aussi très sensible aux conditions initiales. Des erreurs aléatoires gaussiennes de variance $0.01 \text{ mMol N.m}^{-3}$ ajoutées aux conditions initiales des trois variables peuvent produire des écarts considérables par rapport à la solution de référence : jusqu'à 1 mMol N.m^{-3} pour les trois variables, ce qui représente 100% d'erreur en variation relative ! Les écarts les plus forts interviennent

N	P	H	total
0.11631	0.34932	0.10933	0.38406

TAB. 5.1 – Erreurs RMS exprimées en mMol N.m^{-3} du *run* de modèle partant de conditions initiales erronées.

pendant la floraison printanière, comme on le voit à la figure 5.2 dans le cas de la variable nutriments. Les concentrations de phytoplancton et d’herbivores divergent aussi largement du cas de référence pendant cette floraison, voir le tableau 5.1 donnant la moyenne sur l’année de ces erreurs. Est-ce que le RRSQRT KF et l’EnKF peuvent corriger cette erreur commise dans l’état initial ?

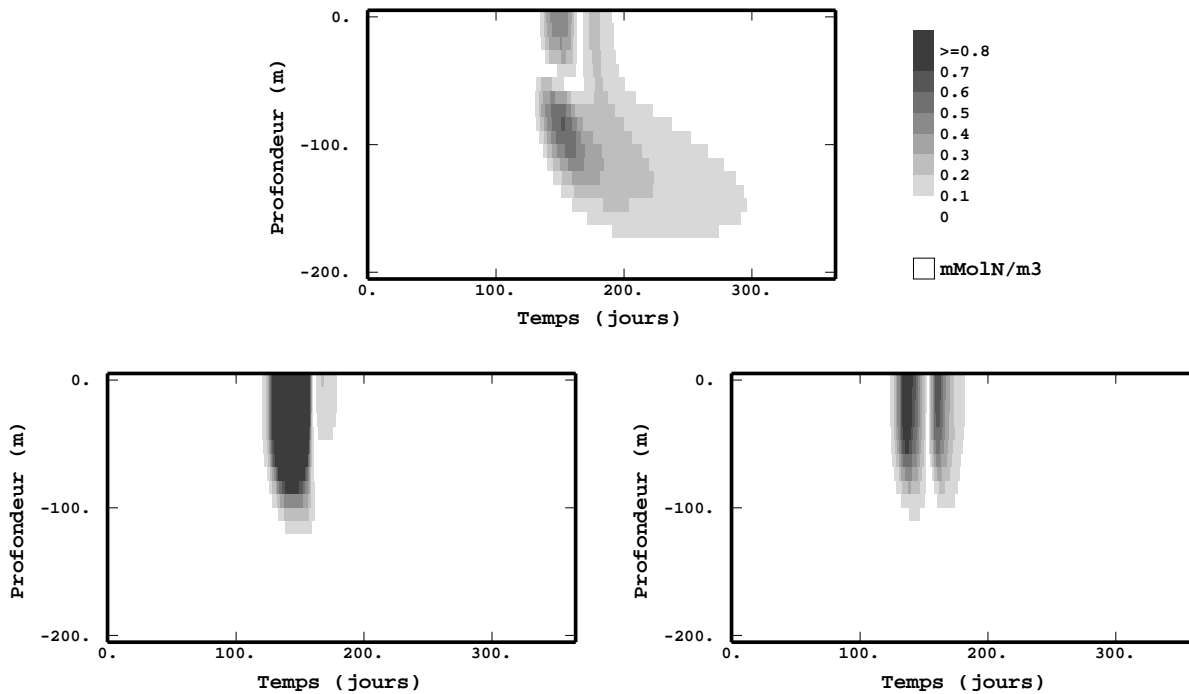


FIG. 5.2 – Valeur absolue des différences entre les concentrations du cas de référence et celles du *run* du modèle pour les trois variables NPH (même échelle) avec une erreur dans les conditions initiales telle que dans le tableau 5.2, sans ADS. Le maximum de l’erreur intervient pendant la floraison printanière. Remarquer que l’échelle est 20 fois plus grande que celle des erreurs d’estimation de l’ADS en figure 5.5. Les erreurs moyennes sur un an sont reportées dans le tableau 5.1.

variable	N	P	H
initiale	6.25e-4	1e-4	1e-4
modèle	1e-5	1e-5	1e-5
lim. haut	0	0	0
lim. fond	1e-5	1e-5	1e-5
mesures	6.25e-4	1e-4	1e-4

TAB. 5.2 – Statistiques d’erreurs utilisées pour l’ADS, les variances sont exprimées en $(\text{mMol N.m}^{-3})^2$.

5.2 Mise en œuvre de l’ADS

5.2.1 Mesures synthétiques

Dans le cas d’étude présent, des données synthétiques sont tirées du *run* de référence ; des échantillons sont pris régulièrement tous les 10 pas de temps du modèle, soit à peu près tous les 10 jours, et un processus de bruit blanc gaussien (centré, dont la variance constante sera ensuite prise comme variance d’erreur de mesure dans le tableau 5.2) leur est ajouté pour simuler la variabilité des mesures. Les deux filtres de Kalman assimilent les mêmes mesures et sont initialisés avec l’état initial entaché d’erreur décrit ci-dessus. De cet échantillonnage on peut remarquer que le support spatial et temporel des mesures est identique à celui des résultats du modèle ; on fait ainsi l’économie d’un modèle de changement de support au contraire des cas réels pour lesquels ces supports sont différents.

5.2.2 Statistiques d’erreur

Dans ce cas synthétique, les statistiques d’erreur ont été gardées telles que dans l’article d’Eknes et Evensen (2001). Elles retracent les caractéristiques du modèle écologique avec réalisme. L’erreur de mesure n’est que de 10% des valeurs moyennes et le modèle est lui aussi supposé commettre une erreur (problème en contraintes faibles) bien qu’il ait lui-même fourni la solution de référence. La variance de l’erreur du modèle a été fixée à tâtons par Eknes et Evensen, elle semble plus faible que celle de l’erreur de mesure dans le tableau 5.2 mais puisque les données sont assimilées tous les 10 pas du modèle, on peut dire que le modèle et les mesures sont équivalents du point de vue de leur précision. Les deux méthodes comparées ci-dessous (RRSQRT et EnKF) utilisent les mêmes statistiques d’erreurs.

5.2.2.1 Erreur initiale

La matrice de covariance initiale C_0^a est celle d'une fonction d'autocovariance gaussienne de portée 50 m (un quart de la profondeur totale de la colonne d'eau). Les covariances croisées sont égales à zéro ce qui signifie que les trois variables sont supposées initialement mutuellement indépendantes.

5.2.2.2 Erreur du modèle

L'erreur de modèle est faite d'une part d'une erreur sur l'état complet du système, c'est-à-dire en chaque nœud de grille pour toutes les variables, avec la même structure spatiale que l'erreur initiale. L'autre part est une erreur sur le forçage du modèle, c'est-à-dire aux conditions aux limites (voir tableau 5.2). Le modèle est supposé exact à la surface de l'eau, mais erroné à la limite inférieure. Cette erreur reflète un manque de connaissance des interactions entre phases solides et liquides au fond. C'est souvent le cas en modèles d'estuaires car elles sont méconnues et difficiles à modéliser. Il s'agit pourtant d'une source importante de nutriments, plus importante que les échanges avec l'atmosphère.

5.2.2.3 Erreur d'observation

L'erreur d'observation est un simple effet de pépité puisqu'il semble peu réaliste que deux cellules de mesures situées à des endroits différents ou mesurant des variables différentes produisent des erreurs cohérentes. La variance de cette erreur de mesure est aussi celle utilisée lors de l'échantillonnage des données synthétiques.

5.3 EnKF : variations de la taille de l'ensemble

On essaie l'EnKF avec trois tailles différentes de l'ensemble : 36, 100 et 1000 membres. Les utilisateurs de l'EnKF utilisent en général 100 membres. Dans le cas de ce modèle écologique simple, on voit sur les erreurs RMS et les écarts-types des figures 5.3 et 5.4 que le choix de ce paramètre importe peu car il n'y a quasiment pas de différences visibles entre les trois courbes. Ceci tient certainement aux faibles dimensions du vecteur d'état ; l'espace à couvrir étant restreint, il suffit de peu de membres pour bien représenter l'espace des erreurs. La haute fréquence des données par rapport aux processus mis en jeu peut aussi en être responsable.

On voit sur les premiers jours de l'expérience que l'EnKF sous-estime la variance de l'erreur (ce qui est prévisible pour une méthode de Monte Carlo), mais qu'en l'espace d'un mois l'erreur initiale décroît jusqu'à

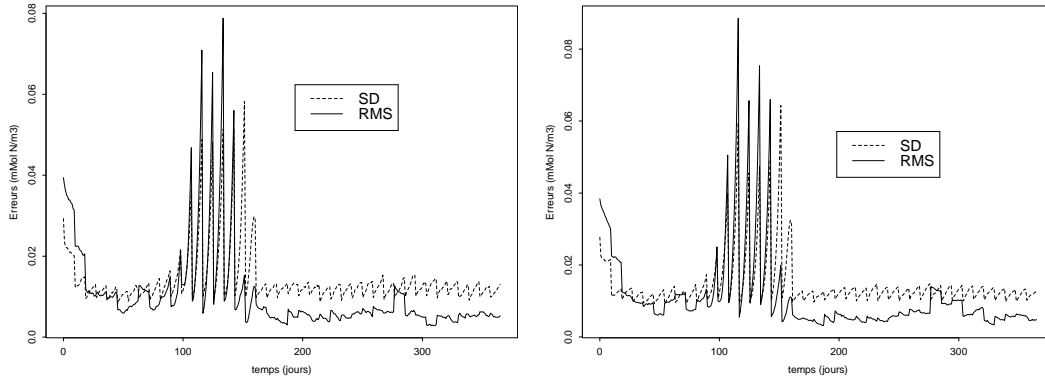


FIG. 5.3 – Erreurs absolues (RMS) commises par l'EnKF avec 36 membres (gauche) et 100 membres (droite) et écarts-types (SD) prédits pour le total des trois variables N, P et H en moyenne sur la colonne d'eau. Dans les trois cas, l'erreur prédite SD surestime l'erreur effective RMS.

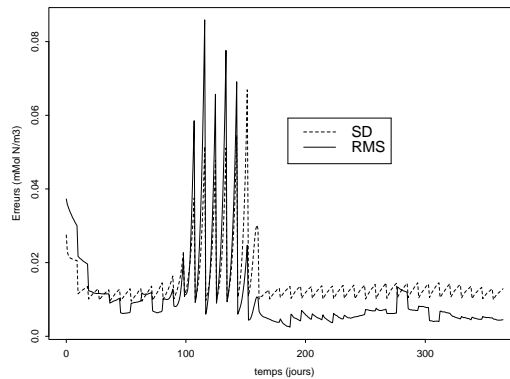


FIG. 5.4 – Idem avec 1000 membres.

un niveau plus faible où l'erreur estimée est égale à l'erreur commise. On reconnaît sur la courbe des SD les dents de scie caractéristiques du filtre de Kalman : l'erreur estimée augmente à chaque propagation du modèle (3.19) et diminue à chaque correction (3.28). Vient ensuite la période de la floraison printanière où le modèle a un comportement plus instable et provoque plus d'erreurs. L'EnKF estime bien ces erreurs, ce qui est beaucoup plus facile pour les modèles fictifs que dans les cas réels puisque c'est le même modèle qui produit la «réalité» et son estimation. Breuillin, Evensen, et Eknes (2000) montre sur cet exemple que l'*Ensemble Kalman Smoother* EnKS peut limiter les erreurs dans cette période d'instabilité. Pendant la deuxième moitié de l'année, l'EnKF surestime ensuite l'erreur commise.

5.4 Comparaison de l'EnKF au RRSQRT KF

Pour les deux procédures d'ADS on compare leur capacité à recréer le *run* de référence et à estimer leur propre erreur.

5.4.1 Précautions d'emploi

Puisque les données ont subi l'ajout d'une erreur de mesure aléatoire, on ne peut tirer de conclusions sur des résultats obtenus avec une seule de ses réalisations. Trois expériences ont donc été menées à bien en utilisant trois germes différents pour générer les nombres aléatoires. Les résultats du tableau 5.4 montrent que l'influence de ces nombres aléatoires est négligeable.

L'utilisation du RRSQRT en ADS multivariable nécessite de normaliser la covariance d'erreur avant de procéder à la décomposition en valeurs propres dans un souci d'équité. Cette normalisation évite en effet de couper systématiquement les modes associés aux variables qui ont des valeurs faibles ou celles exprimées dans les unités les plus fortes. Dans le cas présent la covariance n'est pas normalisée, ce qui ne doit pas avoir de conséquences puisque les trois variables sont du même ordre de grandeur.

5.4.2 Stabilité

Le filtre de Kalman RRSQRT manque de stabilité pour des raisons difficiles à cerner mais dont il faut sans doute chercher l'origine dans l'équation de linéarisation (4.11) utilisée pour la propagation de la covariance d'erreur. Le paramètre - arbitraire - de linéarisation ϵ est censé être petit pour approcher au plus près le modèle linéaire tangent, cependant on ne peut lui attribuer une valeur faible (inférieure à 1 dans le cas présent) sans voir le filtre diverger : la variance d'estimation augmente brusquement, tend vers l'infini et le processus se termine dans des erreurs numériques. Martin Verlaan (communication personnelle) suggère d'utiliser des valeurs du paramètre ϵ suffisamment grandes et nous indique que dans ses propres applications, la valeur $\epsilon = 1$ est satisfaisante. Dans le cas du modèle écologique cette valeur 1 joue le rôle de seuil critique en dessous duquel la divergence est assurée et au delà duquel les résultats semblent stables et reproductibles, du moins pendant un an d'ADS, ce qui ne garantit aucunement la stabilité du RRSQRT KF sur des périodes plus longues. Ainsi la valeur finalement attribuée à ϵ a été fixée à 7 afin d'assurer une certaine marge de sécurité.

5.4.3 Résultats et discussion

5.4.3.1 Aspects numériques

Le temps de calcul pris par un schéma d'ADS est *grosso modo* une fonction linéaire du nombre d'intégrations du modèle direct nécessaires dans l'étape de propagation.

La dimension du vecteur d'état est de 60 dans le cas du modèle écologique (3 variables distribuées sur 20 nœuds verticaux), le nombre de modes propres retenus pour le RRSQRT KF est fixé arbitrairement à 36 pour ne pas trop amputer la variance du système. Malgré cette réduction, le filtre de Kalman RRSQRT nécessite plus de 36 intégrations du modèle écologique. L'ébauche de la matrice de covariance (4.10) propage en plus 60 vecteurs associés à l'erreur de modèle en chaque nœud vertical de chacune des 3 variables, et l'erreur dans les conditions aux limites au fond de la colonne d'eau fait encore ajouter 3 vecteurs. Si l'on compte aussi l'ébauche du vecteur d'état estimé \mathbf{Z}^f dans (3.17), on obtient 100 intégrations du modèle par pas de temps.

Pour s'abstraire de la discussion sur l'influence du paramètre r , on dote l'ensemble de 1000 membres ce qui est certes bien supérieur aux applications habituelles de l'EnKF, mais les praticiens des MCMC (*Monte Carlo Markov Chains*) prennent en général bien plus de réalisations pour décrire correctement des covariances.

Du point de vue des temps de calcul, le filtre de Kalman RRSQRT propage 10 fois moins de vecteurs que l'EnKF, mais le rapport des temps de calcul de l'un à l'autre n'est que de 5 environ : en effet le RRSQRT dans son étape de réduction décompose en valeurs propres une matrice de taille 99×99 (par la routine classique `rsm` de la librairie `eispack`), ce qui prend aussi beaucoup de temps. Cañizares (1999) montre les courbes donnant les temps de calculs associés aux deux méthodes en fonction du nombre de modes retenus et du nombre de membres de l'ensemble dans le cas d'un modèle hydrodynamique côtier en 2-D.

5.4.3.2 Capacité à estimer ses erreurs

Les graphiques du haut de la figure 5.5 montrent l'écart-type d'estimation (SD) donné par l'EnKF (à gauche) et le RRSQRT (à droite). Ceux du bas montrent les erreurs absolues (RMS) entre les estimations et la solution de référence « vraie ». Seule la variable nutriments est représentée mais les deux autres variables phytoplancton et herbivores présentent les mêmes aspects pour ce qui nous importe. Il semble à première vue que les deux méthodes soient équivalentes : leur erreur estimée SD surestime l'erreur effective RMS sur toute la hauteur de la colonne d'eau et pendant toute l'année. Sur les valeurs moyennes de l'écart-type (tableau 5.3) et de l'erreur absolue (tableau 5.4), on voit que les deux filtres de Kalman surestiment l'erreur absolue commise pour les trois variables N, P et H. On a ainsi une indication que la modélisation de l'erreur est efficace, ce qui ne relève pas de l'exploit dans ce cas d'étude synthétique puisque les mécanismes à l'origine

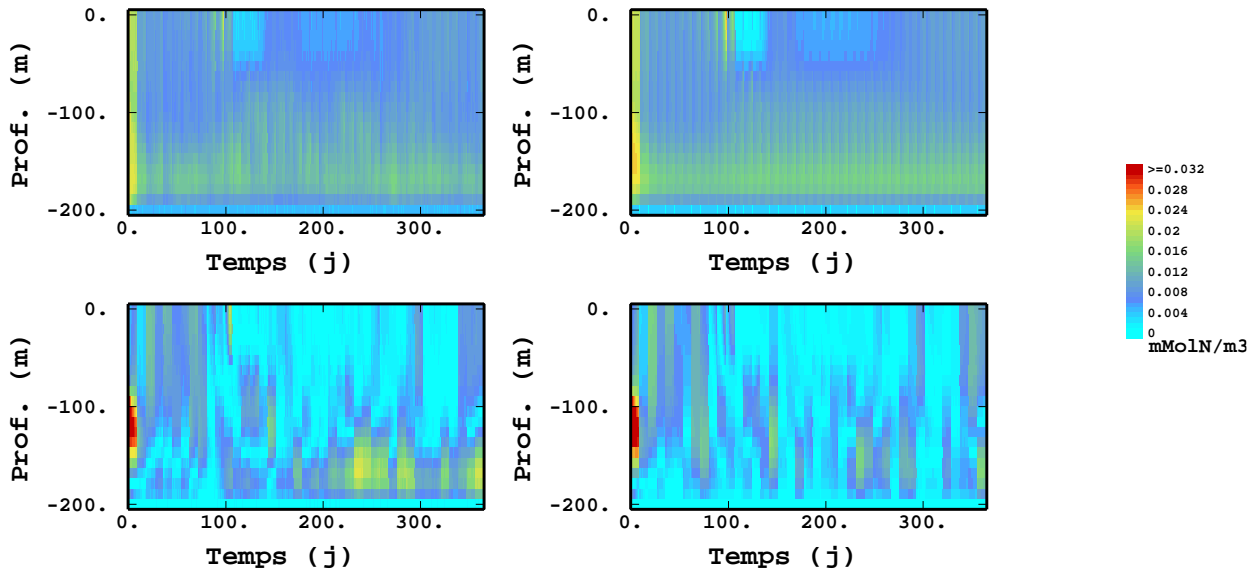


FIG. 5.5 – EnKF (gauche) contre RRSQRT KF (droite) : écarts-types prédits par le filtre de Kalman (en haut) et erreurs absolues observées (en bas) pour les nutriments (N).

de l'erreur sont connus. De plus cette surestimation de l'erreur valide les choix faits pour les paramètres de nombre de modes propres pour le RRSQRT et taille de l'ensemble pour l'EnKF. En effet quand ces deux paramètres sont trop faibles, l'erreur absolue est sous-estimée comme cela a déjà été montré par Cañizares (1999).

Par comparaison avec les résultats du modèle «faux » utilisé sans assimilation (tableau 5.1), on voit que les erreurs provenant d'un changement dans l'état initial sont bel et bien corrigées par l'ADS, quelle que soit la variable. Sur la figure 5.5 des traits verticaux apparaissent sur les cartes d'écart-type d'estimation. Ce sont les traces des deux étapes du filtre de Kalman : d'une part l'ébauche augmente l'écart-type de l'erreur d'estimation par ajout de la covariance de l'erreur de modèle, et d'autre part la correction réduit cet écart-type par interpolation aux moindres carrés des mesures.

On observe aussi que les erreurs sur les concentrations de nutriments sont plus faibles dans la moitié supérieure de la colonne d'eau et pendant la période d'Avril à Octobre (jours 100 à 300 approximativement). Ceci correspond au lieu et à l'époque où l'activité biologique est la plus forte, car les conditions d'ensoleillement sont les meilleures. Pendant le reste de l'année et dans la partie inférieure de la colonne d'eau, les erreurs sont semblables aux erreurs de modèle indiquées dans le tableau 5.2. Cette observation est confirmée dans les résultats d'Eknes et Evensen (2001) dans des conditions d'observation plus contraignantes (une seule observation en surface comme dans le cas de données satellites). On voit ainsi que la dynamique

EnKF	N	P	H	total	RRSQRT	N	P	H	total
n.1	0.01012	0.00888	0.00588	0.01469	n.1	0.01054	0.01067	0.00690	0.01651
n.2	0.01015	0.00865	0.00576	0.01453	n.2	0.01054	0.01084	0.00695	0.01664
n.3	0.01007	0.00891	0.00592	0.01469	n.3	0.01054	0.01085	0.00697	0.01665

TAB. 5.3 – Écarts-types (SD) en mMol N.m^{-3} pour 3 *runs* d'AD avec 3 germes différents, donc 3 jeux de données différents.

EnKF	N	P	H	total	RRSQRT	N	P	H	total
n.1	0.00735	0.00730	0.00444	0.01127	n.1	0.00653	0.00398	0.00229	0.00798
n.2	0.00738	0.00626	0.00389	0.01043	n.2	0.00670	0.00329	0.00205	0.00774
n.3	0.00704	0.00727	0.00447	0.01106	n.3	0.00674	0.00418	0.00254	0.00833

TAB. 5.4 – Erreurs absolues (RMS) en mMol N.m^{-3} pour des *runs* d'AD utilisant 3 réalisations différentes des données partant de 3 germes aléatoires différents.

du modèle joue un rôle important dans l'analyse de l'erreur faite par le filtre de Kalman. Cette performance est d'autant plus frappante qu'on a noté sur la figure 5.2 que la période principale d'activité biologique est celle où le modèle est le plus instable. Cette contradiction apparente est sans doute liée au fait qu'on utilise le modèle qui a lui-même produit la solution de référence, le bon comportement de l'ADS peut donc être seulement artificiel.

Si on regarde plus attentivement les graphiques du bas de la figure 5.5, on voit que les erreurs absolues sont plus faibles du côté du RRSQRT. Le tableau 5.4 confirme cette observation surprenante pour les trois variables. En effet il n'y a aucune raison théorique d'avoir de meilleurs résultats avec le RRSQRT si les deux méthodes utilisent les mêmes statistiques d'erreur et si l'ensemble de l'EnKF est assez large. Ce résultat contredit aussi ceux de Cañizares (1999) où les deux filtres de Kalman s'étaient avérés équivalents.

Une augmentation de la taille de l'ensemble devrait mettre l'EnKF et le RRSQRT sur un même pied d'égalité, mais aux dépens de l'efficacité du calcul. A l'inverse, on a été surpris de constater qu'une diminution de la taille de l'ensemble de 1000 à 100, puis à 36 n'affecte que faiblement les erreurs absolues, ce qui indique par ailleurs que cet ensemble est assez fourni et qu'il ne constitue peut-être pas le paramètre le plus sensible en l'occurrence. Bien entendu cette observation ne peut pas être généralisée à des systèmes réels en grandes dimensions qui peuvent être plus exigeants.

De la même manière, on constate dans le tableau 5.3 que les écarts-types moyens du RRSQRT sont plus élevés que ceux de l'EnKF, alors que les erreurs absolues sont plus faibles, sans plus de justification théorique. Cette différence est toutefois un peu trop faible pour qu'on puisse l'interpréter.

5.4.4 Conclusion de la comparaison

En conclusion les performances des deux méthodes sont équivalentes pour l'ADS dans le modèle écologique 1-D présenté. Les faibles différences observées entre le RRSQRT et l'EnKF dépendent peut-être du système étudié et il serait hardi de les généraliser à d'autres situations.

5.5 EnKF en version lognormale

5.5.1 Limites du modèle gaussien

On étudie dans cette partie la possibilité d'appliquer aux variables N, P et H un modèle lognormal et on le compare au modèle gaussien. La loi lognormale (parmi d'autres lois positives gamma, beta ...) peut être appliquée en sciences de la nature car une majorité des variables sont des concentrations et sont donc à valeurs positives. Les estimateurs basés sur un modèle gaussien imposent aux distributions *a posteriori* une forme gaussienne (voir Liabaud (2001) pour le cas de l'EnKF), qui est au contraire symétrique et non minorée. En particulier ils peuvent estimer des concentrations négatives, très embarrassantes dans la pratique. Dans le cadre d'une ADS, ces valeurs négatives réintroduites dans le modèle physique n'ont aucun sens et provoquent des comportements aberrants. Des méthodes existent pour supprimer les valeurs négatives, par exemple dans cette application et celle du chapitre 7, on affecte la valeur 0 aux valeurs négatives. Cette troncature est gênante d'un point de vue statistique car elle biaise l'estimateur, mais aussi du point de vue physique car elle fausse aussi les bilans ². Plus généralement le problème n'est pas vraiment celui des valeurs négatives, mais plutôt l'application d'un modèle gaussien à des variables qui ne le sont pas.

5.5.2 Des variables lognormales ?

Pour corriger l'état par interpolation linéaire, le filtre de Kalman suppose les lois conditionnelles $\mathcal{L}(\mathbf{Z}_n | \mathbf{Y}_{1:n-1})$ et $\mathcal{L}(\mathbf{Y}_n | \mathbf{Z}_n)$ multigaussiennes. Dans le cas présent cette hypothèse très exigeante n'est évidemment pas respectée. Pour s'en convaincre on considère les histogrammes des variables brutes de référence³ (figure 5.6). D'après cette figure, aucune des trois variables ne peut raisonnablement être tenue pour gaussienne puisque tous les histogrammes sont penchés à gauche. On constate que la borne inférieure de chaque variable est bien 0 et il ne sera donc pas nécessaire de translater le logarithme. L'hypothèse lognormale n'est pas pour autant la meilleure, car les histogrammes des logarithmes (figure 5.7) montrent

²Dans les applications présentées ici, le non respect du bilan de matière ne me dérange pas car le bilan initial est imprécis et ne mérite pas d'être conservé.

³A temps fixé, le nombre d'informations, 20 par variable, est trop faible pour tracer un histogramme significatif, on trace donc l'histogramme en confondant tous les pas de temps de l'année.

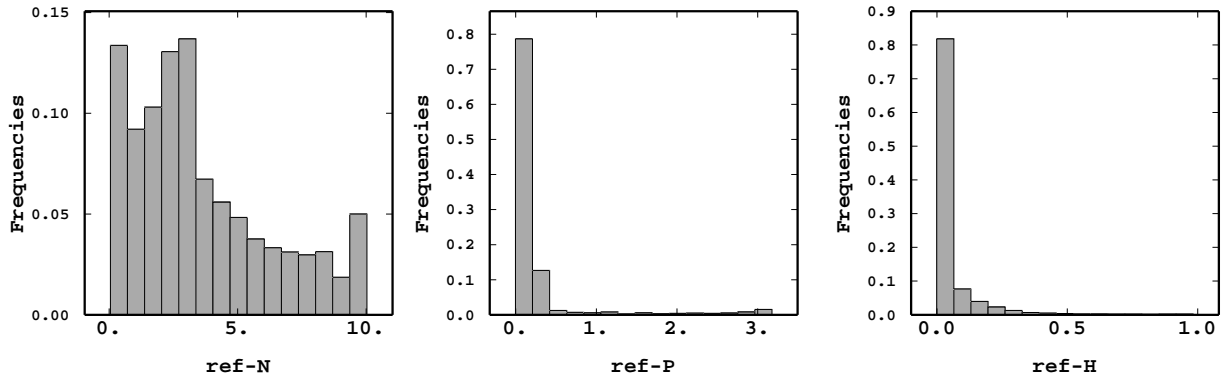


FIG. 5.6 – Histogrammes des variables Nutriments, Herbivores et Phytoplancton au cours d’un cycle annuel de référence.

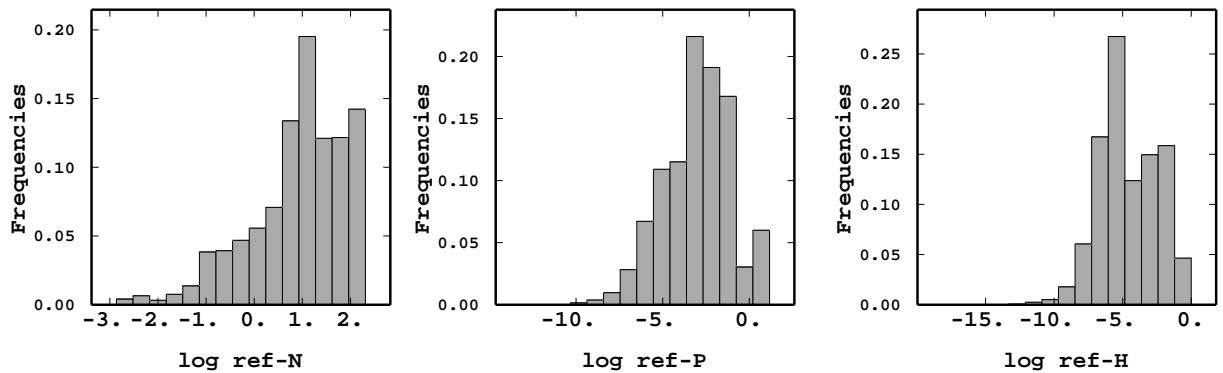


FIG. 5.7 – Histogrammes des logarithmes correspondants.

que l’histogramme de $\log(N)$ (et dans une moindre mesure ceux de $\log(P)$ et $\log(H)$) est penché à droite alors que dans le cas lognormal il devrait être gaussien et symétrique. On pourrait affiner l’anamorphose gaussienne en modélisant cet histogramme puisque d’après la section 4.2.1 cette approche est valable pour toute fonction d’anamorphose.

5.5.3 Mise en œuvre

On utilise 100 membres d’ensemble pour les deux versions. L’EnKF procédant déjà par échantillonnage de la loi de \mathbf{Z} et de \mathbf{Y} , il est très simple de l’adapter au modèle d’anamorphose de la section 4.2.1 avec des fonctions ψ et χ toutes deux exponentielles. Avec le filtre de Kalman RRSQRT il aurait fallu diagonaliser

variable	N	P	H
initiale	4e-6	1e-4	1e-4
modèle	1e-7	1e-5	1e-5
lim. haut	0	0	0
lim. fond	1e-7	1e-5	1e-5
mesures	4e-6	1e-4	1e-4

TAB. 5.5 – Variances d’erreurs lognormales, pas d’unités.

variable	N	P	H
initiale	6.25e-4	1.0e-4	1.0e-4
modèle	1.0e-5	1.0e-5	1.0e-5
lim. haut	0	0	0
lim. fond	1.0e-5	1.0e-5	1.0e-5
mesures	3.0e-5	3.0e-3	6.0e-3

TAB. 5.6 – Statistiques d’erreurs utilisées pour l’AD normale pour comparaison, les variances sont exprimées en $(\text{mMol N.m}^{-3})^2$.

la matrice de covariance du logarithme de \mathbf{Z} . La manipulation est certainement possible mais plus rebutante car le code de calcul du RRSQRT est en lui-même plus compliqué que celui de l’EnKF.

Le bruit de mesure a été simulé par un modèle lognormal

$$\mathbf{Y}_n = H\mathbf{Z}_n^t \times e^{\varepsilon_n^o}, \quad \varepsilon_n^o \sim \mathcal{N}(0, \Sigma^o) \quad (5.1)$$

plutôt que gaussien pour plus de réalisme : les mesures les plus fortes sont entachées d’une erreur plus élevée, ce qui semble logique pour certains paramètres (concentrations ...). Les trois variances d’erreur de mesure de Σ^o ont été attribuées par tâtonnements en s’assurant que les valeurs perturbées $z \exp(\varepsilon^o)$ restent réalistes dans toutes les situations. Les variances d’erreurs initiales sont égales à celles d’erreurs de mesure. Les variances d’erreurs de modèle ont été attribuées arbitrairement, aussi par tâtonnement (tableau 5.5).

Pour utiliser l’EnKF (version gaussienne) dans des conditions habituelles, les biais de mesure $E(1 - e^{\varepsilon_n^o}) = 1 - e^{-\frac{\sigma_o^2}{2}}$ (de l’ordre des variances du tableau 5.5) ont été négligés mais les variances d’erreur de mesure sont ajustées empiriquement à $\text{Var}(\mathbf{y}_n - H\mathbf{z}_n^t)$. Les autres statistiques d’erreur ont été laissées telles que dans les expériences précédentes (tableau 5.6). Puisque les deux méthodes ne peuvent pas travailler à statistiques égales nous commenterons des aspects indépendants des valeurs ajustées des variances. De plus la portée des résultats se limite au cas d’erreurs de distributions lognormales, mais peut sans doute être étendue aux autres distributions dissymétriques.

5.5.4 Comparaison de l'EnKF gaussien contre lognormal

De visu, les résultats des deux assimilations sont les mêmes : il faut tracer les erreurs RMS (figure 5.8) et les nuages de corrélations (figure 5.9) de \mathbf{z}_n^f contre \mathbf{z}_n^t pour analyser le comportement de chaque filtre. L'écart-type de ces erreurs ne peut pas être pris en compte car on peut le réduire ou l'augmenter à volonté en jouant sur les variances d'erreurs (tableaux 5.5 et 5.6).

Le modèle gaussien suppose que la variance des erreurs est la même pour les valeurs fortes et pour les valeurs faibles. En l'occurrence cette hypothèse est fautive : d'après l'équation (5.1) elle est proportionnelle ici à la valeur de \mathbf{Z} . La figure 5.8 montre que l'estimateur gaussien produit plus d'erreur pour les variables P (au milieu à gauche) et H (en bas à gauche) que son homologue lognormal (à droite), et particulièrement lors de la floraison printanière. On note aussi que les erreurs d'estimation de P sont de signe opposé à celles de H (cette observation est reproductible quelques soient les germes aléatoires et les valeurs des variances d'erreur), ce qui suggère l'explication suivante : l'estimateur gaussien est plus imprécis pour les valeurs faibles de P et H car la variance d'erreur de mesure reste aussi élevée que pour les valeurs fortes ; de plus il est légèrement biaisé par la troncature des valeurs négatives (remises à zéro car sans signification physique) qui provoque une surestimation systématique. Si on attribue à la variable H (herbivores) en début de floraison printanière une valeur même légèrement trop forte, alors la croissance de H et le broutage du phytoplancton vont démarrer en avance, ce qui explique la surestimation de H et la sous-estimation de P, visibles sur la partie gauche de la figure 5.9. Tous les 10 pas de temps, le filtre de Kalman recalcule l'estimation sur la valeur observée, traçant ainsi des rayures verticales qui correspondent à chacun des « faux départs » de croissance de H. On les évite en général par estimation lognormale où ce sont plutôt les valeurs fortes qui sont mal estimées (voir figure 5.9 à droite).

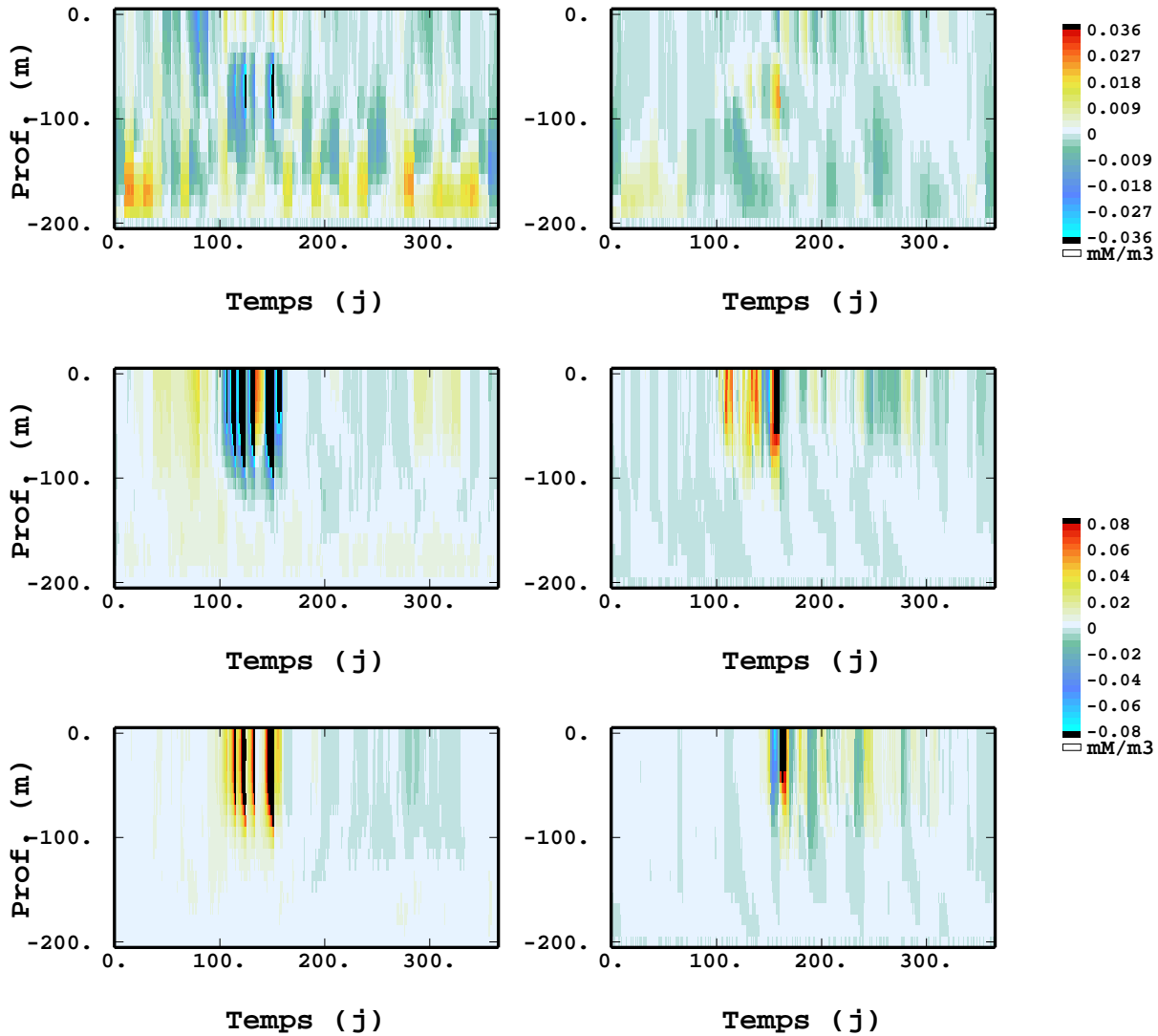


FIG. 5.8 – Erreurs observées ($z_n^f - z_n^t$) pour les trois variables : N en haut, P au milieu et H en bas. A gauche les résultats de l'estimateur gaussien, à droite ceux de l'estimateur anamorphosé lognormal. L'échelle de couleurs est affinée pour la variable N pour montrer que les erreurs de H et P sont liées entre elles, mais pas avec les erreurs de N.

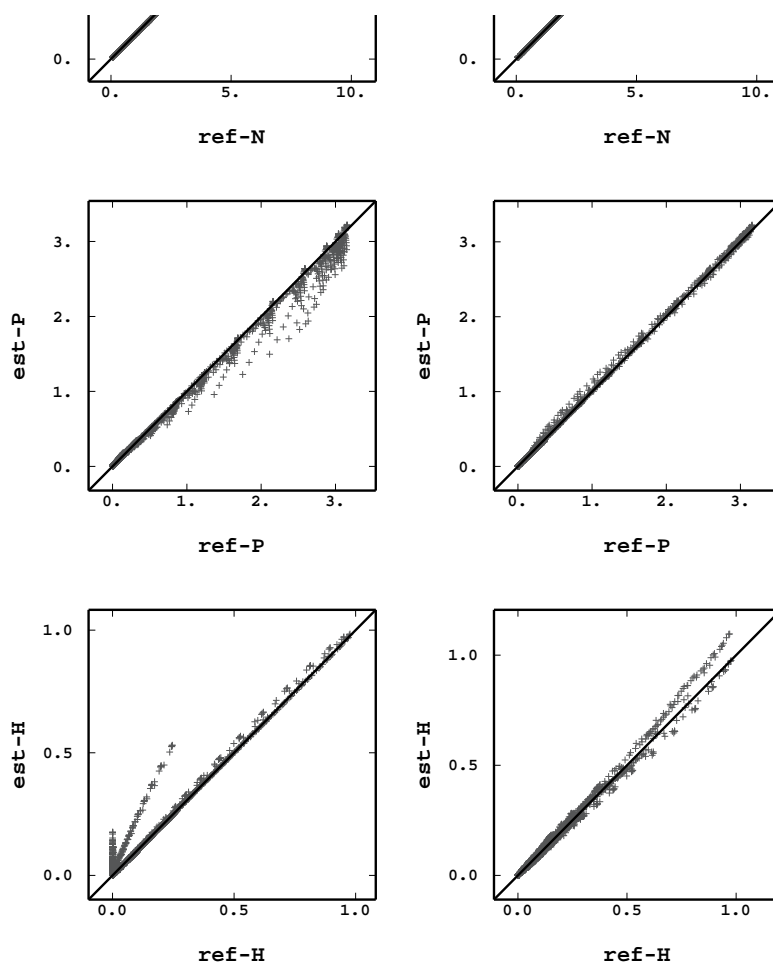


FIG. 5.9 – Nuages de corrélations entre variables estimées et variables de référence pour l’EnKF (version normale à gauche, lognormale à droite) appliqué à des observations de loi lognormale, la première bissectrice est tracée.

Chapitre 6

Cas de l'hydrodynamique de la lagune de l'Oder

Sommaire

On applique le filtre de Kalman RRSQRT au modèle TRIM3D et on assimile des données de hauteurs d'eau, on utilise un modèle de propagation d'ondes pour calculer les caractéristiques spatiales de l'erreur de modèle. On procède à un run de validation en gardant deux stations de côté, l'une au centre de la région observée, l'autre à l'extérieur. Les résultats sont bons, meilleurs que ceux obtenus par interpolation des données, mais l'assimilation sous-estime l'erreur commise.

6.1 Généralités

6.1.1 Données géographiques

La lagune de l'Oder est située à la frontière entre l'Allemagne et la Pologne. Il s'agit d'un bassin de 50 km de large et de très faible profondeur (4 m en moyenne). Il représente un volume d'eau d'environ 9 km³. L'eau douce entre par plusieurs rivières dont l'Oder est, de loin, la plus importante. En fonction des variations du niveau de la mer Baltique, l'eau de mer s'introduit par intermittence en petites quantités dans la lagune *via* les trois chenaux qui relient la lagune à la mer (voir la carte en figure 6.1). Ces intrusions



FIG. 6.1 – Carte de la lagune de l’Oder avec les 5 stations de mesure Odh1, Odh2, Ücker Münde, Karnin et Wolgast. La carte montre aussi les trois embouchures de la lagune Peenestrom (1), Świna (2) et Dźwina (3) ainsi que celle de l’Oder (4).

d’eau salée ont un impact très important sur l’ensemble de l’écosystème et sont difficiles à prévoir sans l’aide d’un modèle numérique. Un canal de navigation profond de 20m a été creusé afin de faciliter le transport commercial de Świnoujście à Szczecin (de (4) à (2) sur la figure 6.1). Il constitue aussi un chemin préférentiel pour l’écoulement de l’eau. Rosenthal *et al.* (1998) montrent que 70% des échanges passent par la Świna (2), seulement 20% par le Peene (1) et 10% par la Dźwina (3).

L’écologie de ce bassin est avant tout contrôlée par le transport des nutriments venant de la rivière et les échanges avec la Baltique. La modélisation d’un tel système requiert donc une estimation très précise de l’état hydrodynamique de la lagune.

6.1.2 Le projet PIONEER

PIONEER (*Preparation and Integration of analysis tools towards Operational forecast of Nutrients in Estuaries of European Rivers*) est un projet européen ayant débuté en Septembre 1998 et ayant pris fin en Septembre 2001 (von Storch, 2000). Son objectif a été de développer des outils de routine d'analyse et de prédiction des nutriments. Deux cas ont été étudiés : l'estuaire de l'Ebre en Espagne et la lagune de l'Oder dont nous présenterons ici les aspects physiques. L'AD physique doit donc donner les meilleures prédictions possibles des concentrations de nutriments qui seront les paramètres d'entrée d'un modèle écologique (dans le cas de la lagune de l'Oder, il s'agit du modèle ERSEM). Les prédictions du modèle écologique sont elles aussi améliorées par l'AD et doivent servir d'outil d'aide à la décision pour la gestion de l'estuaire. La simulation se situe en période de crue de l'Oder, survenue fin juillet 1997, jusqu'au premier octobre.

6.2 Le modèle TRIM3D

6.2.1 Équations

TRIM3D est un modèle hydrodynamique tridimensionnel écrit par Casulli et Cattani (1994). Il résout les équations de Navier-Stokes (6.1-6.3) pour l'écoulement d'un fluide incompressible à surface libre sous l'hypothèse hydrostatique. u , v et w sont les trois composantes du vecteur vitesse, η est la hauteur d'eau par rapport à une altitude de référence, p la pression normalisée, g la constante de gravitation, ν un coefficient de turbulence. Les notations ci-dessous sont indépendantes de celles adoptées dans les chapitres 2 à 4 pour l'ADS.

La conservation de l'énergie donne les équations ponctuelles :

$$\begin{aligned} \partial_t u + \partial_x(uu) + \partial_y(uv) + \partial_z(uw) &= -\partial_x p + \partial_x(\nu \partial_x u) \\ &+ \partial_y(\nu \partial_y u) + \partial_z(\nu \partial_z u) \end{aligned} \quad (6.1)$$

$$\begin{aligned} \partial_t v + \partial_x(uv) + \partial_y(vv) + \partial_z(vw) &= -\partial_y p + \partial_x(\nu \partial_x v) \\ &+ \partial_y(\nu \partial_y v) + \partial_z(\nu \partial_z v) \end{aligned} \quad (6.2)$$

$$\begin{aligned} \partial_t w + \partial_x(uw) + \partial_y(vw) + \partial_z(ww) &= -\partial_z p + \partial_x(\nu \partial_x w) \\ &+ \partial_y(\nu \partial_y w) + \partial_z(\nu \partial_z w) \end{aligned} \quad (6.3)$$

Seule la partie hydrostatique de la pression p est prise en considération, elle se décompose en un terme barotropique et un terme baroclinique, fonctions des gradients de hauteur d'eau et de densité. La conservation de la masse d'un fluide incompressible donne :

$$\partial_x u + \partial_y v + \partial_z w = 0 \quad (6.4)$$

On suppose maintenant que la hauteur d'eau totale est une fonction de l'espace et du temps. Les cotes du fond et de la surface libre sont respectivement :

$$\begin{aligned} z &= -H(x, y) \\ z &= \eta(x, y, t). \end{aligned}$$

où $-H(x, y)$ est négative par rapport à la surface de référence et η est la surélévation de la surface libre par rapport à la référence (qu'on appellera aussi « hauteur d'eau » par abus de langage). On suppose que la bathymétrie $H(x, y)$ est parfaitement connue. En intégrant l'équation (6.4) sur la verticale on obtient :

$$\partial_t \eta + \partial_x \left[\int_{-H}^{\eta} u dz \right] + \partial_y \left[\int_{-H}^{\eta} v dz \right] = 0 \quad (6.5)$$

A la surface libre, les gradients de vitesse sont imposés par le vent (τ_x^w, τ_y^w) :

$$\nu \partial_z u = \tau_x^w \quad \nu \partial_z v = \tau_y^w \quad (6.6)$$

Les conditions aux limites à l'interface eau-sédiments sont données par une formule de Manning-Chezy :

$$\nu \partial_z u = \gamma u \quad \nu \partial_z v = \gamma v \quad (6.7)$$

où $\gamma = \frac{g\sqrt{u^2+v^2}}{C_z^2}$ et C_z est le coefficient de friction de Chezy

Les équations (6.1-6.3) et (6.5) avec les conditions limites (6.6-6.7) constituent le modèle d'écoulement 3-D en eau peu profonde.

6.2.2 Discrétisation

Ces équations sont discrétisées selon un schéma de différences finies semi-implicites sur une grille d'Arakawa : les hauteurs d'eau η sont définies aux positions i, j , alors que les vitesses u, v et w sont décalées d'une demi-maille en avant. C'est-à-dire qu'elles sont définies aux positions $(i + \frac{1}{2}, j, k)$, $(i, j + \frac{1}{2}, k)$ et $(i, j, k + \frac{1}{2})$ respectivement (voir sur la figure 6.2 en vue de dessus). On doit alors connaître la topographie en double résolution : la profondeur H doit être donnée aux positions (i, j) , $(i + \frac{1}{2}, j)$ et $(i, j + \frac{1}{2})$. Les dérivées des hauteurs d'eau (6.1-6.2) et des vitesses (6.5) sont discrétisées par un schéma de différences finies semi-implicite (dite « méthode θ ») et les termes de mélange verticaux, ainsi que tous les autres termes, sont discrétisés explicitement. On note N_x et N_y le nombre de nœuds de grille actifs dans les directions x et y . N_z est le nombre de couches dans la direction verticale, il dépend donc de la position (i, j) mais on utilise cette notation simplifiée par commodité. Le système discrétisé peut alors s'écrire :

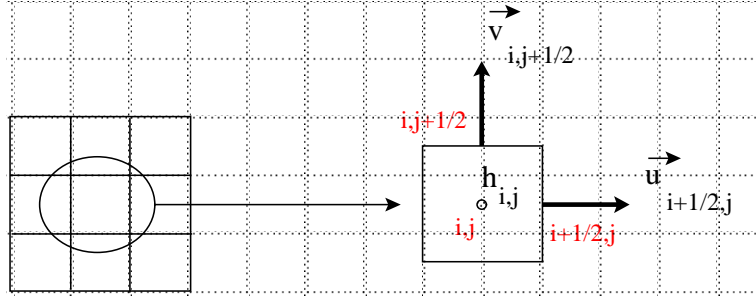


FIG. 6.2 – Définition des hauteurs d'eau et des vitesses sur une grille d'Arakawa

$$\begin{aligned}
 u_{i+\frac{1}{2},j,k}^{n+1} &= F u_{i+\frac{1}{2},j,k}^n - g \frac{\Delta t}{\Delta x} \left[\theta \left(\eta_{i+1,j}^{n+1} - \eta_{i,j}^{n+1} \right) + (1 - \theta) \left(\eta_{i+1,j}^n - \eta_{i,j}^n \right) \right] \\
 &\quad + \Delta t \frac{\nu_{k+\frac{1}{2}} \frac{u_{i+\frac{1}{2},j,k+1}^{n+1} - u_{i+\frac{1}{2},j,k}^{n+1}}{\Delta z_{i+\frac{1}{2},j,k+\frac{1}{2}}} - \nu_{k-\frac{1}{2}} \frac{u_{i+\frac{1}{2},j,k}^{n+1} - u_{i+\frac{1}{2},j,k-1}^{n+1}}{\Delta z_{i+\frac{1}{2},j,k-\frac{1}{2}}}}{\Delta z_{i+\frac{1}{2},j,k}}
 \end{aligned} \tag{6.8}$$

$$\begin{aligned}
 v_{i,j+\frac{1}{2},k}^{n+1} &= F v_{i,j+\frac{1}{2},k}^n - g \frac{\Delta t}{\Delta y} \left[\theta \left(\eta_{i,j+1}^{n+1} - \eta_{i,j}^{n+1} \right) + (1 - \theta) \left(\eta_{i,j+1}^n - \eta_{i,j}^n \right) \right] \\
 &\quad + \Delta t \frac{\nu_{k+\frac{1}{2}} \frac{v_{i,j+\frac{1}{2},k+1}^{n+1} - v_{i,j+\frac{1}{2},k}^{n+1}}{\Delta z_{i,j+\frac{1}{2},k+\frac{1}{2}}} - \nu_{k-\frac{1}{2}} \frac{v_{i,j+\frac{1}{2},k}^{n+1} - v_{i,j+\frac{1}{2},k-1}^{n+1}}{\Delta z_{i,j+\frac{1}{2},k-\frac{1}{2}}}}{\Delta z_{i,j+\frac{1}{2},k}}
 \end{aligned} \tag{6.9}$$

$$\begin{aligned}
 \eta_{i,j}^{n+1} &= \eta_{i,j}^n - \frac{\Delta t}{\Delta x} (1 - \theta) \left[\sum_{k=m}^{k=M} \Delta z_{i+\frac{1}{2},j,k} u_{i+\frac{1}{2},j,k}^n - \sum_{k=m}^{k=M} \Delta z_{i-\frac{1}{2},j,k} u_{i-\frac{1}{2},j,k}^n \right] \\
 &\quad - \frac{\Delta t}{\Delta y} (1 - \theta) \left[\sum_{k=m}^{k=M} \Delta z_{i,j+\frac{1}{2},k} v_{i,j+\frac{1}{2},k}^n - \sum_{k=m}^{k=M} \Delta z_{i,j-\frac{1}{2},k} v_{i,j-\frac{1}{2},k}^n \right] \\
 &\quad - \frac{\Delta t}{\Delta x} \theta \left[\sum_{k=m}^{k=M} \Delta z_{i+\frac{1}{2},j,k} u_{i+\frac{1}{2},j,k}^{n+1} - \sum_{k=m}^{k=M} \Delta z_{i-\frac{1}{2},j,k} u_{i-\frac{1}{2},j,k}^{n+1} \right] \\
 &\quad - \frac{\Delta t}{\Delta y} \theta \left[\sum_{k=m}^{k=M} \Delta z_{i,j+\frac{1}{2},k} v_{i,j+\frac{1}{2},k}^{n+1} - \sum_{k=m}^{k=M} \Delta z_{i,j-\frac{1}{2},k} v_{i,j-\frac{1}{2},k}^{n+1} \right],
 \end{aligned} \tag{6.10}$$

où Δx , Δy , Δz et Δt sont les dimensions de la maille et le pas de temps (Δz dépend de la position car la profondeur est variable). m et M correspondent respectivement aux couches du fond et de la surface. F est l'opérateur linéaire explicite venant de la discrétisation des équations (6.1) et (6.2).

Le principe de l'algorithme de résolution est le suivant : les équations (6.8-6.10) constituent un système linéaire dont les inconnues sont η^{n+1} , u^{n+1} et v^{n+1} (w est discrétisé explicitement et il est connu à cette étape de la résolution). Les équations (6.8) et (6.9) permettent d'exprimer u^{n+1} et v^{n+1} en fonction de η^{n+1} . Il est alors possible de substituer à u^{n+1} et v^{n+1} leur nouvelle expression dans l'équation (6.10). On calcule alors le nouveau champ de hauteur d'eau, puis les champs de vitesse en les injectant dans (6.11) et (6.12). Enfin, les vitesses verticales sont calculées dans l'équation (6.4) discrétisée.

Ainsi, en notant $\mathbf{u}_{i+\frac{1}{2},j}$ et $\mathbf{v}_{i,j+\frac{1}{2}}$ les vecteurs vitesses ($u_{i+\frac{1}{2},j,k}$ et $v_{i,j+\frac{1}{2},k}$, $\forall k$) et $\Delta \mathbf{z}_{i,j}$ le vecteur des épaisseurs de couches sur toute la verticale ($\Delta z_{i,j,k}$, $\forall k$), on peut écrire le système algébrique :

$$A_{i+\frac{1}{2},j}^n \mathbf{u}_{i+\frac{1}{2},j}^{n+1} = \mathbf{g}_{i+\frac{1}{2},j}^n - g \frac{\Delta t}{\Delta x} \left(\eta_{i+\frac{1}{2},j}^{n+1} - \eta_{i,j}^{n+1} \right) \Delta \mathbf{z}_{i+\frac{1}{2},j}^n \quad (6.11)$$

$$A_{i,j+\frac{1}{2}}^n \mathbf{v}_{i,j+\frac{1}{2}}^{n+1} = \mathbf{g}_{i,j+\frac{1}{2}}^n - g \frac{\Delta t}{\Delta y} \left(\eta_{i,j+\frac{1}{2}}^{n+1} - \eta_{i,j}^{n+1} \right) \Delta \mathbf{z}_{i,j+\frac{1}{2}}^n \quad (6.12)$$

$$\begin{aligned} \eta_{i,j}^{n+1} = \eta_{i,j}^n & - \frac{\Delta t}{\Delta x} \left[\left(\Delta \mathbf{z}_{i+\frac{1}{2},j}^n \right)^\top \mathbf{u}_{i+\frac{1}{2},j}^{n+1} - \left(\Delta \mathbf{z}_{i-\frac{1}{2},j}^n \right)^\top \mathbf{u}_{i-\frac{1}{2},j}^{n+1} \right] \\ & - \frac{\Delta t}{\Delta y} \left[\left(\Delta \mathbf{z}_{i,j+\frac{1}{2}}^n \right)^\top \mathbf{v}_{i,j+\frac{1}{2}}^{n+1} - \left(\Delta \mathbf{z}_{i,j-\frac{1}{2}}^n \right)^\top \mathbf{v}_{i,j-\frac{1}{2}}^{n+1} \right]. \end{aligned} \quad (6.13)$$

$A_{i+\frac{1}{2},j}^n$, $A_{i,j+\frac{1}{2}}^n$ sont des matrices définies positives de taille $N_z \times N_z$, donc inversibles, et $\mathbf{g}_{i+\frac{1}{2},j}^n$, $\mathbf{g}_{i,j+\frac{1}{2}}^n$ sont des vecteurs de taille N_z contenant les termes explicites. On remarque que les équations (6.11 et 6.12) représentent $N_x \times N_y$ systèmes linéaires de N_z équations, alors que l'équation (6.13) est un système linéaire de $N_x \times N_y$ équations. On remarque de plus que les matrices $A_{i+\frac{1}{2},j}^n$ et $A_{i,j+\frac{1}{2}}^n$ sont tridiagonales alors que le système (6.13) a cinq diagonales non nulles.

Ayant remplacé u^{n+1} et v^{n+1} dans l'équation (6.13) par leur expression en fonction de η^{n+1} (6.11 et 6.12), le système est pratiquement résolu par une méthode de gradient conjugué (Casulli et Cattani, 1994). Les solutions $\eta_{i,j}^{n+1}$ sont alors réinjectées dans les $2N_x N_y$ systèmes (6.11 et 6.12), en utilisant les conditions aux limites suivantes :

$$\nu_{i+\frac{1}{2},j,M+\frac{1}{2}} \frac{u_{i+\frac{1}{2},j,M+1}^{n+1} - u_{i+\frac{1}{2},j,M}^{n+1}}{\Delta \mathbf{z}_{i+\frac{1}{2},j,M+\frac{1}{2}}} = \tau_x^w \quad (6.14)$$

$$\nu_{i,j+\frac{1}{2},M+\frac{1}{2}} \frac{v_{i,j+\frac{1}{2},M+1}^{n+1} - v_{i,j+\frac{1}{2},M}^{n+1}}{\Delta \mathbf{z}_{i,j+\frac{1}{2},M+\frac{1}{2}}} = \tau_y^w \quad (6.15)$$

et

$$v_{i+\frac{1}{2},j,m-\frac{1}{2}} \frac{u_{i+\frac{1}{2},j,m}^{n+1} - u_{i+\frac{1}{2},j,m-1}^{n+1}}{\Delta z_{i+\frac{1}{2},j,m-\frac{1}{2}}} = \gamma_{i+\frac{1}{2},j,m}^{n+\frac{1}{2}} u_{i+\frac{1}{2},j,m}^{n+1} \quad (6.16)$$

$$v_{i,j+\frac{1}{2},m-\frac{1}{2}} \frac{v_{i,j+\frac{1}{2},m}^{n+1} - v_{i,j+\frac{1}{2},m-1}^{n+1}}{\Delta z_{i,j+\frac{1}{2},m-\frac{1}{2}}} = \gamma_{i,j+\frac{1}{2},m}^{n+\frac{1}{2}} v_{i,j+\frac{1}{2},m}^{n+1}. \quad (6.17)$$

Enfin, les vitesses verticales sont calculées dans la forme discrétisée de l'équation (6.4) :

$$w_{i,j,k+\frac{1}{2}}^{n+1} = w_{i,j,k-\frac{1}{2}}^{n+1} - \frac{\Delta z_{i+\frac{1}{2},j,k}^n u_{i+\frac{1}{2},j,k}^{n+1} - \Delta z_{i-\frac{1}{2},j,k}^n u_{i-\frac{1}{2},j,k}^{n+1}}{\Delta x} - \frac{\Delta z_{i,j+\frac{1}{2},k}^n v_{i,j+\frac{1}{2},k}^{n+1} - \Delta z_{i,j-\frac{1}{2},k}^n v_{i,j-\frac{1}{2},k}^{n+1}}{\Delta x}. \quad (6.18)$$

La résolution du système $N_x \times N_y$ est l'étape de l'algorithme la plus coûteuse numériquement. Par contre l'ajout de couches horizontales supplémentaires n'augmente pas beaucoup les temps de calcul. On note de plus qu'après chaque itération les épaisseurs de couches sont recalculées afin de prendre en compte les points mis «à sec».

6.2.3 Implémentation

Dans le cas d'étude suivant, le pas de temps du modèle Δt est de 5 min, la grille rectangulaire suit la topographie de la lagune. Sa résolution horizontale Δx est de 250 m, elle s'étend sur 357×259 nœuds carrés dont 16053 se trouvent à l'intérieur de la lagune. La colonne d'eau est divisée en 3 couches verticales (les cotes limites sont -3 m et -6 m) et on obtient au final 24455 nœuds «mouillés». On peut voir dans Rosenthal *et al.* (1998) et Wolf *et al.* (2001) des simulations numériques de la lagune de l'Oder avec TRIM3D dans cette configuration.

On pourrait réduire la dimension du vecteur d'état, et ainsi la charge calculatoire de l'assimilation, en augmentant la taille du nœud de grille ou la durée du pas de temps. Cependant le modèle perdrait sa capacité à reproduire les processus de moyenne et petite échelle que l'on trouve dans l'hydrodynamique de la lagune de l'Oder. Dans cette application on économise le temps de calcul au cours de l'ADS en fonction de critères statistiques de sorte que les processus négligés sont ceux qui induisent le moins de variations possibles, indépendamment de leur ordre de grandeur.

On peut aussi remarquer que pour l'AD de hauteurs d'eau à seules fins physiques, un modèle 2-D aurait suffi. Mais en vue d'applications écologiques, l'implémentation sur le modèle 3-D est indispensable pour calculer les vitesses et le transport des nutriments. En estuaire, les intrusions salines sont aussi un problème 3-D car la langue d'eau salée, plus lourde que l'eau douce, entre par le fond.

6.2.4 Conditions initiales et conditions aux limites

Puisque nous ne disposons pas d'information globale sur l'état entier du système, l'état initial est tel que toutes les variables hydrodynamiques (hauteurs d'eau, courants horizontaux et verticaux) sont mises à zéro. Ainsi à l'état initial, la lagune de l'Oder est supposée uniformément immobile et au niveau zéro. Seules les conditions aux limites mettent cette masse d'eau en mouvement et vont faire oublier cet état initial arbitraire au bout d'un temps d'initialisation qui est de l'ordre de la semaine pour notre application, soit ≈ 150 heures comme on le voit sur la figure 6.7. La dynamique du système est contrôlée principalement par trois variables :

1. à l'interface avec la mer Baltique, les débits sont libres (il s'agit donc d'une frontière ouverte) mais les hauteurs d'eau sont imposées par les oscillations propres de la mer Baltique¹. Les cotes aux trois points de contact de la lagune avec la mer sont imposées, égales aux mesures de la station Koserow (moyennes horaires) sur la côte Baltique (voir figure 6.1). Comme les hauteurs d'eau doivent changer sur les 70 km de côte qui séparent l'embouchure du Peene (1) de celle de la Dźwina (3), cette approximation sera prise en compte *a posteriori* par l'ADS grâce aux mesures de hauteurs d'eau prises sur toute l'étendue de la lagune,
2. à l'interface atmosphérique le vent influence fortement le comportement hydrodynamique de la lagune car sa profondeur est très faible. Les mesures de vent ne sont disponibles qu'en une seule station (à Ückeründe) qui délivre des moyennes horaires. Le modèle TRIM3D attribue ces mesures ponctuelles à la lagune entière c'est-à-dire jusqu'à 60 km de l'endroit où est mesuré le vent et on peut s'attendre à ce que les mesures ne correspondent pas au champ de vent réel même si les moyennes horaires ont déjà filtré un certain nombre de phénomènes transitoires locaux. On compte là aussi sur l'ADS pour corriger cette erreur,
3. l'embouchure de la rivière Oder où le débit est imposé, mais la hauteur d'eau est libre. Le débit est mesuré en amont à la station Police, les dates de ces mesures ont été retardées pour tenir compte du décalage entre Police et l'embouchure. Les mesures sont prises toutes les 4 heures (le débit varie très peu en 4 heures) et le modèle les utilise à chaque pas de temps *via* une interpolation linéaire. Cette procédure est satisfaisante parce que d'une part les variations de débit sont lentes et d'autre part parce que les hauteurs d'eau sont plus sensibles aux deux autres conditions aux limites.

6.2.5 Sensibilité du modèle

Sénégas (1999) a étudié la sensibilité du modèle aux conditions de vent et aux conditions aux limites avec la mer Baltique. Il montre que le modèle est assez peu sensible à ces paramètres pris séparément, mais que la combinaison des deux a un effet plus important, supérieur à la somme des deux effets. Cette synergie

¹Les marées sont plus faibles que ces oscillations propres.

est possible grâce aux non linéarités du modèle. La sensibilité du modèle aux conditions initiales est assez forte (il faut plus de 5 jours d'initialisation en général).

6.3 Les données de hauteur d'eau

6.3.1 Dispositif expérimental

Les données de hauteur d'eau sont prises dans des stations fixes. Les stations Ückermünde, Karnin et Wolgast sont gérées par les autorités allemandes de l'eau alors que les stations Odh1 et Odh2 ont été ajoutées par le GKSS. Le dispositif expérimental est le suivant : un ballon en flottaison est enfermé dans un cylindre vertical où il monte et descend librement. Sa position est repérée par des électro-aimants situés à différentes hauteurs du cylindre dont on mesure les caractéristiques électriques.

Une question importante est celle du calibrage de ces mesures de hauteur d'eau. Les différentes stations se doivent en effet d'avoir le même zéro. Si les trois stations officielles des autorités allemandes de l'eau sont calibrées au millimètre près, un problème se pose en revanche pour les deux stations du GKSS, Odh1 et Odh2, dont les mesures ne subissent pas la procédure de normalisation des trois stations des autorités officielles. Un biais de mesure, que l'on peut supposer constant après traitement approprié des séries temporelles, doit donc être corrigé pour ne pas biaiser l'assimilation.

6.3.2 Analyse des données

On dispose de séries temporelles de hauteurs d'eau pour la période de la crue de l'Oder de juillet 1997 jusqu'au 30 octobre aux 5 stations dans la partie ouest de la lagune plus une station dans la Baltique (voir carte 6.1). Une station polonaise (Karsibór) située dans la Świna a aussi mesuré les hauteurs d'eau, mais seulement pendant les mois de juillet et août 97. Les deux stations «au large» Odh1 et Odh2 mesurent aussi les courants mais ces données ne sont pas exploitables dans notre situation car les courants de la lagune ont trop de variations locales pour que leur assimilation en seulement deux points soit bénéfique.

Les variogrammes des 5 séries temporelles de hauteurs d'eau ont des allures semblables : ils sont bornés, ce qui indique une stationnarité apparente, et atteignent leur palier pour une portée assez courte (≈ 500 heures soit trois semaines). La valeur du palier est commune à tous les variogrammes (voir deux exemples en figure 6.3). Il était certes prévisible que des séries temporelles de hauteurs d'eau d'une lagune aient un caractère stationnaire puisqu'elles sont naturellement rappelées vers le zéro du niveau de la mer. Mais à l'échelle du mois on aurait pu se trouver confronté à une non stationnarité apparente due à des tendances saisonnières.

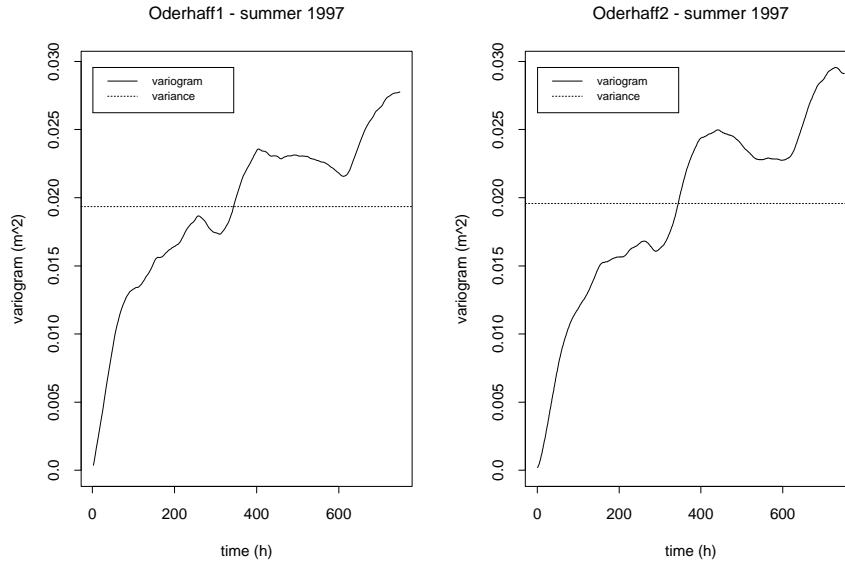


FIG. 6.3 – Variogrammes expérimentaux des séries temporelles de hauteur d'eau à Odh1 et Odh2, la ligne pointillée est la variance expérimentale.

L'analyse en covariances croisées de ces séries temporelles fait ressortir des décalages croissants en fonction de l'éloignement des stations. On note que les décalages sont très faibles entre les stations situées dans les bassins principaux mais plus marqués avec les stations situées dans les chenaux ou avec Koserow dans la Baltique, voir un exemple typique sur la figure 6.4 où la covariance croisée (en pointillés à droite) est translatée de 15 heures par rapport à l'autocovariance et Karsibór est en retard.

A l'inverse les corrélations sont plus fortes entre les stations des bassins et elles diminuent quand on avance dans les chenaux vers la mer. On peut donc en déduire que la variable hauteur d'eau est assez uniforme dans les bassins et se propage presque simultanément, mais que les variations sont plus accentuées dans les chenaux et leur propagation est plus lente.

Si on interprète ces décalages comme le fait d'ondes de surface, il est logique que ces ondes se propagent plus vite dans les grands bassins que dans les chenaux étroits et sinueux où elles sont soumises à de nombreuses réflexions. On devine aussi qu'une description stationnaire de la variable η est erronée puisque les variations sont plus fortes dans les chenaux. Dans l'expérience d'ADS qui va suivre, en utilisant l'équation (4.1), nous laisserons le modèle physique construire les covariances spatiales non stationnaires entre les hauteurs d'eau.

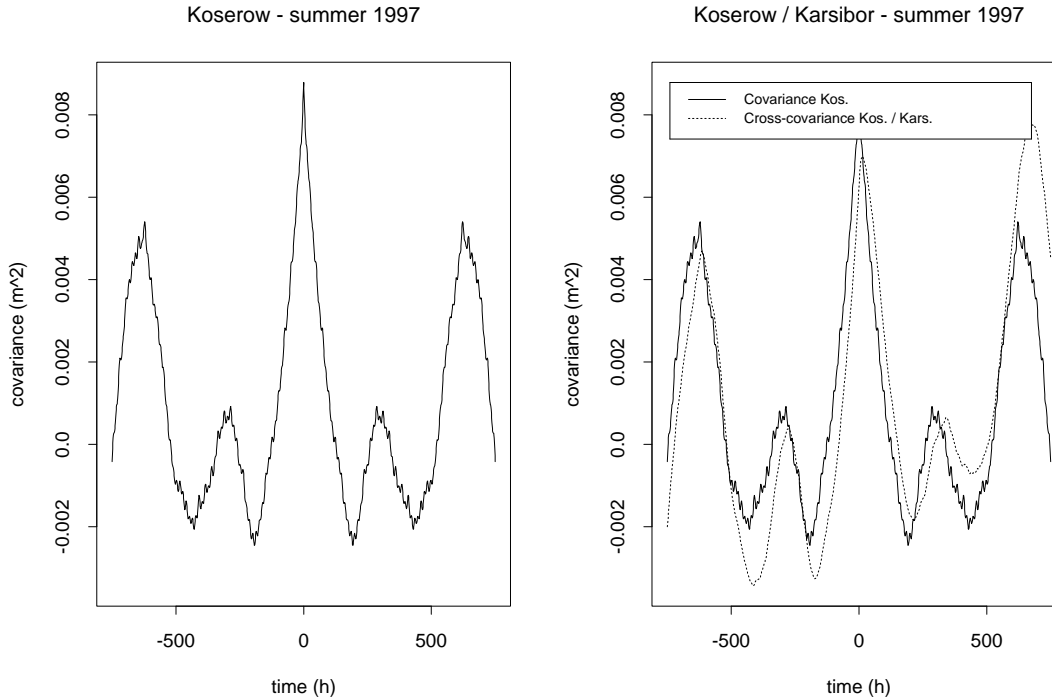


FIG. 6.4 – A gauche l'autocovariance expérimentale des séries temporelles de hauteur d'eau à Koserow, à droite la covariance croisée expérimentale entre Koserow et Karsibór.

L'étude approfondie des décalages temporels ne laisse pas transparaître de caractéristiques physiques générales. En effet ces décalages varient suivant la période d'analyse (par exemple de 8 à 20 heures entre Koserow et Karsibór) et leurs signes s'inversent car ils dépendent certainement du débit des différents canaux². Dans un souci de généralité et pour ne pas compliquer le processus, on se contente donc d'utiliser des décalages moyens sur un grand intervalle de temps et d'en tirer un seul paramètre de célérité c , sans signification physique, utilisé ultérieurement pour convertir les structures temporelles de l'erreur du modèle en structures spatiales. Des données de stations plus éloignées dans la mer Baltique prises dans le cadre du projet GOAP n'apportent pas non plus d'éléments de compréhension supplémentaires.

On modélise les variogrammes temporels par deux structures sphériques de portées 90 et 500 heures (resp. 4 jours et 3 semaines) et un léger effet de pépité. L'ajustement est manuel.

²Marek Stawarz, spécialiste de la physique de la lagune de l'Oder au GKSS, suggère aussi un lien avec la pression atmosphérique (communication personnelle) ce qui constitue une piste pour approfondir l'étude.

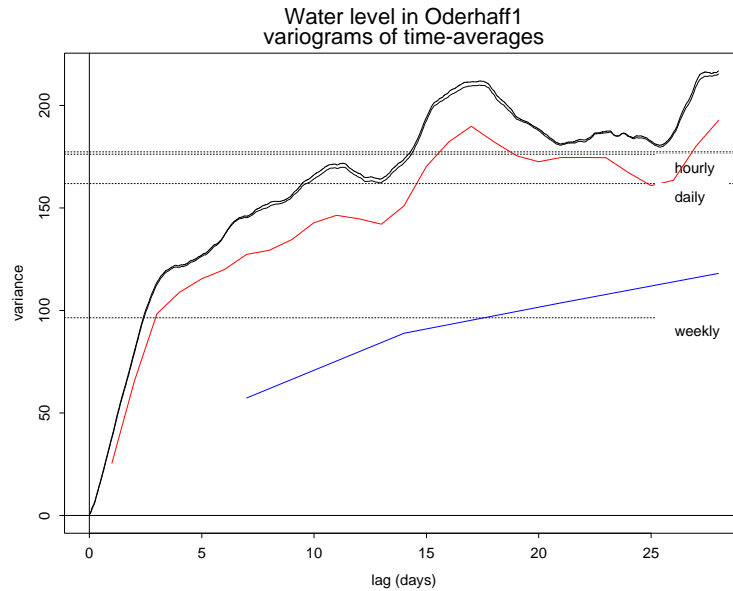


FIG. 6.5 – Variogrammes expérimentaux des séries temporelles de hauteur d'eau à Odh1 pour des mesures moyennées sur différents supports temporels (10 minutes, une heure, un jour, une semaine), les lignes pointillées horizontales sont les variances aux différents supports.

6.3.3 Effet de support

On a vu dans la partie 3.2 que la différence de support entre l'état observé $H\mathbf{Z}$ (identique à celui de \mathbf{Z}) et les observations \mathbf{Y} peut enlever tout sens à la différence $\mathbf{Y} - H\mathbf{Z}$, pourtant capitale en AD. Pour évaluer l'influence du support, on trace les variogrammes temporels des mesures de hauteurs d'eau moyennes (figure 6.5). L'état \mathbf{Z} est donné sur un support temporel de 5 minutes par le modèle TRIM3D alors que les mesures sont disponibles toutes les 10 minutes. On constate que les variogrammes des hauteurs d'eau aux supports 10 minutes et 1 heure sont quasiment confondus. On en déduit que les changements de support de 5 minutes à 10 minutes, voire à 1 heure sont sans effet sur la variabilité des hauteurs d'eau, ce qui justifie l'AD de moyennes horaires. Les variogrammes des moyennes journalières et hebdomadaires sont par contre fortement altérés par le changement de support. On remarque aussi que le variogramme des moyennes hebdomadaires reste structuré, ce qui indique la présence de structures de portées supérieures à une semaine ou de non stationnarités.

Pour ce qui est de l'effet de support spatial, il n'est pas possible de le quantifier expérimentalement par manque de données, mais étant donné la structure très lisse des hauteurs d'eau, les moyennes horaires ponctuelles constituent certainement une information valable à l'échelle d'un bloc de 250 m par 250 m. Ultérieurement, un modèle spatio-temporel (6.19) sera décrit mettant en relation les structures spatiales avec les structures temporelles. Ce modèle confirme aussi l'idée des hauteurs d'eau très lisses. On peut donc négliger également le changement de support spatial.

6.4 Mise en œuvre du filtre de Kalman

6.4.1 Compatibilité du filtre de Kalman avec TRIM3D

6.4.1.1 La composition du vecteur d'état

Le vecteur d'état devrait en théorie contenir toutes les variables du modèle susceptibles d'être influencées par l'ADS. En pratique, certaines sont connues pour être faiblement influencées et elles sont écartées *a priori* pour délester l'algorithme d'une charge numérique superflue. Dans notre cas le vecteur d'état inclut donc naturellement les hauteurs d'eau corrigées par l'ADS, mais aussi les composantes horizontales du vecteur vitesse du courant (u et v). Par contre la composante verticale de la vitesse w et les épaisseurs des couches verticales Δz sont entièrement remises à jour à chaque début de cycle à partir de u , v et η . Il n'y a donc aucune raison de conserver dans le vecteur d'état les 50000 paramètres de ces deux variables 3-D. Le vecteur d'état contient donc les 16053 valeurs de η (défini en surface), 24455 valeurs des vitesses u et v (tous deux définis sur les 3 couches verticales). Au total, on a près de 65000 paramètres dans le vecteur d'état.

6.4.1.2 Un modèle semi-implicite en trois temps

La nature semi-implicite du schéma de différences finies de TRIM3D (6.10) pose un autre problème pratique à l'implémentation du filtre de Kalman RRSQRT. En effet le calcul implicite des hauteurs d'eau et des courants au temps n nécessite l'inversion d'une matrice à 5 diagonales. Une descente de gradient initiée par les hauteurs d'eau au temps $n - 1$ assure cette inversion avec efficacité. Les concepteurs de TRIM3D supposent que les états du système aux temps n et $n - 1$ sont peu différents et ils utilisent ainsi l'état précédent \mathbf{z}_{n-1} comme première approximation de \mathbf{z}_n pour initialiser la descente de gradient. Cependant le filtre de Kalman RRSQRT n'intègre pas les vecteurs par ordre chronologique dans l'équation (4.10), mais par valeurs propres décroissantes. Si on initialise la descente par un état trop peu ressemblant, on risque de devoir faire calculer plus d'itérations, ce qui allonge les temps de calcul du modèle³. Plutôt que de stocker

³On ne risque pas de converger vers un minimum local puisqu'on résout une équation linéaire.

tous les vecteurs propres du temps précédent, ce qui doublerait la taille du processus aux dépens de la performance de l'ADS, on suppose que les vecteurs de l'ensemble sont suffisamment ressemblants les uns aux autres pour que la descente de gradient initialisée sur les vecteurs $(\mathbf{Z}_{n-1}^a + \epsilon \Delta \mathbf{Z}_{i,n-1}^a)$ et $(\mathbf{Z}_n^f + \epsilon \mathbf{s}_n^{i-1,f})$ converge vers la même solution (ce qui est raisonnable dans notre cas car les différences entre ces deux vecteurs se situent surtout à proximité des frontières du domaine). L'expérience montre que dans des cas extrêmes (essais avec erreurs de modèle très élevées), l'algorithme de descente peut ne pas converger, mais dans les cas courants la convergence est seulement plus lente. On passe en moyenne de 7 à 11 itérations, apparemment sans affecter les résultats.

La remarque est valable pour la plupart des versions du filtre de Kalman qui doivent utiliser le modèle plusieurs fois à chaque pas de temps, y compris l'EnKF.

6.4.2 Statistiques des erreurs

Il est nécessaire de connaître les covariances de l'erreur d'observation et de l'erreur du modèle, du moins leurs matrices Σ^o et Σ^m , pour mettre au point un filtre de Kalman. Cependant les sources d'erreurs de mesure sont multiples et le traitement des données est très compliqué (c'est un métier à part entière). Vouloir en tirer un modèle probabiliste d'erreur de mesure relève du casse-tête et on préfère le déduire des statistiques des mesures. On ne peut pas non plus avoir une connaissance objective de l'erreur du modèle, comme toujours dans les cas réels, car l'état et les vraies dynamiques du système sont inconnus.

6.4.2.1 Erreurs d'observation

Les erreurs d'observation peuvent raisonnablement être dépourvues de structure spatiale car elles représentent un manque de précision du protocole de mesure et elles sont rarement corrélées d'un appareil de mesure à l'autre. On utilise le variogramme expérimental des séries temporelles des mesures de hauteur d'eau : la valeur de son effet de pépité fournit la variance de l'erreur de mesure. La composante «effet de pépité» du variogramme reflète en effet des variations non structurées de hauteur d'eau, ce qui contient en général les erreurs de mesure. *A posteriori* l'écart-type correspondant est de 3 mm de hauteur d'eau, ce qui est physiquement satisfaisant pour une moyenne horaire de hauteurs d'eau.

6.4.2.2 Erreurs de modèle

Rosenthal *et al.* (1998) ont utilisé TRIM3D dans la lagune de l'Oder et ont montré que les dynamiques sont bien décrites dans le modèle. Selon eux, l'erreur principale provient de la détermination des conditions aux limites, tout spécialement le champ de vent et les hauteurs d'eau à l'interface avec la mer Baltique, pour les raisons déjà exposées dans la partie 6.2.4 «conditions aux limites».

On considère l'erreur sur le champ de vent comme une variable uniforme sur toute la largeur de la lagune. Il aurait été plus correct de tenir compte de l'éloignement par rapport à la station Ückermünde où le champ de vent est mesuré, mais on verra sur les résultats que cette complication n'en valait pas la peine. Dans la dimension temporelle, les variogrammes expérimentaux mettent en évidence un effet de pépite, l'erreur sur le champ de vent est donc supposée être un bruit blanc, d'écart-type $0,5 \text{ m.s}^{-1}$.

L'erreur sur les hauteurs d'eau dans la Baltique a une structure spatiale. Cependant les données disponibles sont trop pauvres spatialement (6 points de données dont un seul dans la Baltique) pour pouvoir mener une analyse structurale rigoureuse. On va donc modéliser une erreur stationnaire seulement sur un sous-ensemble des paramètres étudiés pour lequel on connaît les structures spatiales, comme dans l'équation (4.1). Une hypothèse faite dans Wolf *et al.* (2001) permet de mettre à profit la structure temporelle contenue dans les longues séries de mesures des hauteurs d'eau : l'hypothèse la plus simple reliant les variations spatiales et temporelles de η est celle d'une équation d'onde plane, introduisant une variable auxiliaire de l'espace $e(\mathbf{x})$:

$$\eta(\mathbf{x}, t) = e(\mathbf{x} - \mathbf{c}t). \quad (6.19)$$

La hauteur d'eau est constante sur le front d'onde défini par le vecteur de célérité locale \mathbf{c} dont la norme est donnée par la profondeur de l'eau $c = \sqrt{gH(\mathbf{x})}$. Pour simplifier, cette célérité sera tenue constante et uniforme et on va supposer que tous les points de contact avec la Baltique sont alignés. Ces hypothèses représentent une vision très idéalisée de la côte Baltique, mais la manœuvre consiste à donner une description raisonnable de l'erreur commise aux conditions aux limites sans avoir besoin à ce niveau de réécrire un second modèle hydrodynamique. On considère alors l'erreur sur les hauteurs d'eau :

$$\varepsilon(x, t) = \eta(x_0, t) - \eta(x, t) \quad (6.20)$$

entre la station Koserow où les hauteurs d'eau sont enregistrées (notée x_0 ci-dessus) et le point de contact effectif entre la lagune et la mer (noté x). Puisque x_0 est fixé, les variations spatiales de $\varepsilon(x, t)$ sont les mêmes que celles de $\eta(x, t)$. Si on suppose de plus la stationnarité de la variable $e(x - ct)$ définie sur \mathbb{R} . On peut utiliser la relation (6.19) pour déduire le variogramme spatial γ_x du variogramme temporel γ_t :

$$\gamma_x(h) = \frac{1}{2} \text{Var}(e(x + h - ct) - e(x - ct)) \quad (6.21)$$

$$\gamma_x(h) = \frac{1}{2} \text{Var}(e(x - c(t - h/c)) - e(x - ct)) \quad (6.22)$$

$$= \gamma_t(h/c). \quad (6.23)$$

Il suffit donc de disposer de la valeur de c pour assurer la transition du variogramme spatial au variogramme temporel.

Si on ne suppose pas les points de contacts alignés, le variogramme ci-dessus n'est plus isotrope puisqu'il dépend de l'orientation du vecteurs \mathbf{h} par rapport à \mathbf{c} . On devrait alors se doter de deux modèles de variogramme de \mathbf{h} suivant ses deux composantes colinéaire et orthogonale à \mathbf{c} , $\mathbf{h} = h_{//}\mathbf{c} + h_{\perp}\mathbf{c}^{\perp}$. Il y a cependant peu à gagner dans cette complication car elle introduit des paramètres inconnus bien difficiles à estimer avec seulement 6 points de mesure. De plus les structures spatiales très lisses des hauteurs d'eau n'y auraient pas été très sensibles.

L'hypothèse de l'équation d'ondes (6.19) entraîne deux conséquences : d'une part le variogramme temporel γ_t est invariant dans l'espace, d'autre part les covariances croisées entre deux séries temporelles s'obtiennent par translation des covariances simples. En effet si on se donne deux points x_1 et x_2 dans le domaine d'étude, on a d'une part :

$$\gamma_t(\tau) = \frac{1}{2}\text{Var}(\eta(x_1, t + \tau) - \eta(x_1, t)) \quad (6.24)$$

$$= \frac{1}{2}\text{Var}(e(x_1 - ct - c\tau) - e(x_1 - ct)) \quad (6.25)$$

$$= \frac{1}{2}\text{Var}(e(x_2 - ct - c\tau) - e(x_2 - ct)), \quad (6.26)$$

et d'autre part les covariances croisées entre les deux séries temporelles :

$$C_t^{x_1, x_2}(\tau) = \text{Cov}(\eta(x_1, t), \eta(x_2, t + \tau)) \quad (6.27)$$

$$= \text{Cov}(e(x_1 - ct), e(x_2 - c(t + \tau))) \quad (6.28)$$

$$= \text{Cov}(e(x_1 - ct), e(x_2 - c(t + \tau + \frac{x_1 - x_2}{c}))) \quad (6.29)$$

$$= C_t^{x_1, x_1}(\tau + \frac{x_1 - x_2}{c}). \quad (6.30)$$

Ces dernières s'obtiennent ainsi par translation des covariances simples. Ces deux résultats font aussi intervenir la stationnarité de $e(x - ct)$. Leur vérification expérimentale est possible en traçant les variogrammes et covariances croisées expérimentaux des séries temporelles de mesures de hauteurs d'eau en différents points de la lagune. Les variogrammes expérimentaux (voir figure 6.3) sont bien indépendants de x , et les covariances croisées se déduisent bien des covariances simples par une translation comme sur la figure 6.4. Cette dernière a par ailleurs servi au calcul du paramètre c représentant une célérité moyenne sur les mois de juillet à octobre 97 de l'onde fictive qui relie les hauteurs d'eau des deux stations les plus proches de la mer (de Koserow à Karsibór). On obtient ainsi $c = 36 \text{ km.h}^{-1}$, ce qui correspond à $H = 10 \text{ m}$ de profondeur dans $c = \sqrt{gH}$. Cette hauteur d'eau est physiquement raisonnable mais il est difficile de donner un sens à ce paramètre car il s'agit d'une hauteur d'eau typique pour la propagation d'une onde au bord de la mer Baltique, valable pour une période donnée.

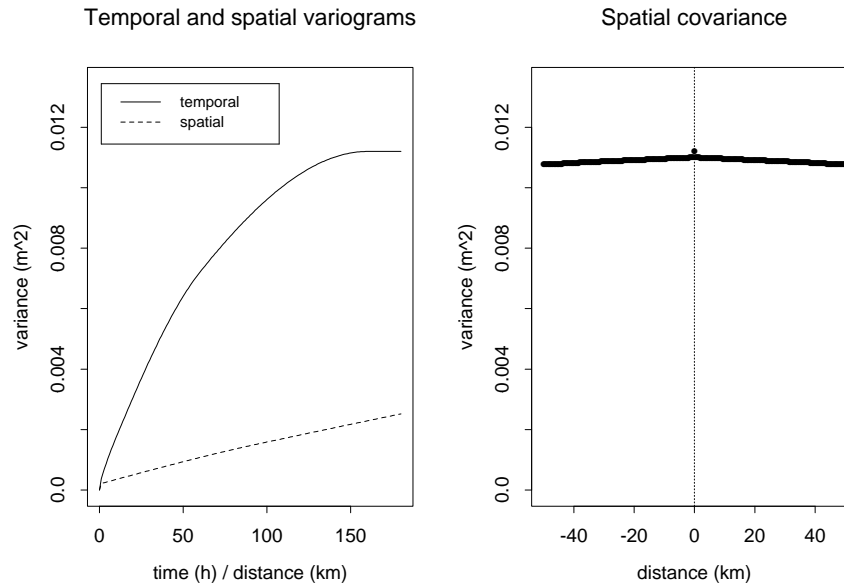


FIG. 6.6 – À gauche les variogrammes spatiaux et temporels modélisés, à droite la fonction de covariance spatiale associée.

Le paramètre $c = 36 \text{ km.h}^{-1}$ est utilisé dans le calcul du variogramme spatial dans (6.21) et les portées temporelles de 90 et 500 h deviennent respectivement 3000 km et 18000 km. Ces portées extrêmement longues génèrent un variogramme spatial quasiment plat (figure 6.6) pratiquement équivalent à un variogramme linéaire car la longueur des côtes ne dépasse pas 70 km. La signification de ce modèle spatial est que les hauteurs d'eau à l'interface Baltique sont très fortement liées, ce qui correspond bien à l'aspect général des hauteurs d'eau observé sur les résultats du modèle TRIM : une variable très lisse. Les résultats ultérieurs semblent valider les hypothèses simples faites ici, qui suffisent donc à rendre compte des structures spatiales.

6.4.3 Mise en œuvre du filtre de Kalman RRSQRT

6.4.3.1 Covariance initiale

A l'instant initial, la covariance est supposée nulle bien qu'on soit loin de connaître à la perfection l'état initial de la lagune. Il est possible d'introduire un modèle de covariance initiale dans le RRSQRT en diagonalisant une matrice de covariance stationnaire, par exemple en utilisant un algorithme de Lanczos. Cependant dans le cas d'une matrice de covariance de grande taille ($N \times N = 24455 \times 24455$), la procé-

ture peut s'avérer compliquée. En effet ces matrices de covariances n'ont pas nécessairement leurs valeurs propres bien séparées, comme on peut le constater pour la matrice d'un effet de pépite (une seule valeurs propre d'ordre N); et les algorithmes de recherche de valeurs propres et vecteurs propres de matrices de grandes tailles s'accommodent mal des valeurs propres multiples⁴. Compte tenu de la structure spatiale très lisse de la variable η , on se contente de prendre une covariance initiale nulle. On compte sur les erreurs de modèle pour compenser cet «excès de certitude» à l'état initial. Comme les hauteurs d'eau ont une dynamique rapide, les incertitudes introduites aux conditions aux limites gagneront rapidement l'ensemble du système.

6.4.3.2 Paramètres spécifiques au RRSQRT

Le nombre de modes propres retenus a été fixé à 50 sur les conseils de Verlaan, en tenant compte des grandes dimensions du système. Ce nombre semble suffisant *a posteriori* car chaque étape de réduction laisse un très faible pourcentage de la variance totale au rebut (moins de 0,2 %).

Le paramètre de linéarisation ϵ utilisé dans (4.10) a été ajusté à tâtons. Les valeurs inférieures à 0, 1 ont donné un filtre divergent avec croissance infinie de l'erreur d'estimation alors que les valeurs 1 et 10 laissent le filtre stable et donnent des résultats identiques sur une journée à condition toutefois que les variances d'erreurs de modèle et de mesure soient réalistes. La valeur la plus faible $\epsilon = 1$ a été choisie, l'intervalle des valeurs possibles de ϵ correspond aux autres cas d'études de Verlaan (comm. pers.).

La matrice des covariances des erreurs de modèle étant définie positive et de taille faible⁵, sa racine carrée a été calculée par une décomposition de Cholesky. Les composantes des vecteurs colonnes sont ensuite introduites dans les sous-routines de TRIM qui tiennent compte des interactions avec le vent et avec la mer Baltique aux conditions aux limites. Comme ces sous-routines sont suivies de transformations non linéaires pour atteindre l'état entier du système, on se trouve dans le cas d'une perturbation non linéaire du modèle $\mathbf{Z}_n^t = \mathbf{f}_n(\mathbf{Z}_{n-1}^a, \epsilon_n^m)$.

Numériquement l'étape la plus coûteuse est celle de propagation qui nécessite d'appliquer le modèle TRIM successivement à la moyenne \mathbf{z}_{n-1}^a , aux 50 perturbations correspondant aux modes propres et aux 16 perturbations par les conditions aux limites, soit 67 fois en tout. En comparaison, la décomposition en valeurs propres et vecteurs propres d'une matrice 66×66 représente un temps de calcul secondaire.

⁴Ce sont souvent des algorithmes qui recherchent itérativement la direction propre associée à la valeur propre dominante, en cas de valeur propre multiple les algorithmes simples ne peuvent pas converger.

⁵14 nœuds de grille en contact avec la Baltique sur les trois chenaux et les 2 composantes du vecteur vent, $\Sigma^m \in \mathbb{R}^{16 \times 16}$

6.5 Résultats du filtre de Kalman RRSQRT : un *run* de validation

6.5.1 Aspects numériques

La période du 4 Août au 13 Septembre 1997 a été simulée et seules les données de trois stations ont été assimilées (Üeckermünde, Karnin et Wolgast). Les calculs ont pris 10 jours sur une station Sun SPARC et on pourrait certainement les réduire sans affecter significativement les résultats par réduction du nombre de modes propres retenus et par regroupement les points de contact avec la mer Baltique chenal par chenal (comme lors de l'ADS jointe du sel et des hauteurs d'eau dans le chapitre 7). Par ailleurs les calculs les plus coûteux (ceux de l'ébauche) sont parallélisables car on peut propager un état perturbé sur un processeur sans tenir compte de la propagation des autres perturbations. Un calcul sur plusieurs processeurs peut alors réduire significativement les délais, mais pas exactement en proportion du nombre de processeurs car l'étape de réduction constitue un point commun à chaque pas de temps. Du point de vue numérique, le filtre de Kalman RRSQRT peut donc être utilisé en prédiction opérationnelle suivant la procédure suivante :

- on dispose d'un état \mathbf{Z}_n^a et de la racine carrée de la covariance de l'erreur S_n^a , tous deux issus d'un *run* précédent d'ADS,
- on propage l'état et la racine carrée de la covariance (éventuellement sur plusieurs processeurs) qui donneront la moyenne et les écarts-types de prédiction au temps $n + 1$,
- à chaque pas de temps du modèle on réduit S_{n+1}^f .

On doit par ailleurs disposer à temps des prédictions des conditions aux limites du modèle TRIM (débit de la rivière, vents et hauteurs d'eau de la Baltique).

6.5.2 Performances de l'initialisation

L'ADS réduit l'impact des conditions initiales comme on peut le constater dans la station de validation Odh1 (figure 6.7). Là où le modèle seul met entre 5 et 7 jours pour rejoindre le niveau observé (ce qui correspond à un temps «de remplissage» de la lagune par les eaux de la Baltique et de l'Oder), il suffit de 12 heures pour caler le filtre de Kalman RRSQRT. La première correction est inopérante car la covariance initiale est nulle, mais à chaque pas de temps, les perturbations introduites sur les hauteurs d'eau à la frontière avec la mer Baltique et sur les vents sont propagées par TRIM sur l'ensemble du domaine. Ainsi, puisque l'étape de correction du filtre de Kalman est une interpolation spatiale suivant la covariance décrite par ces perturbations, les points centraux comme Odh1 bénéficient rapidement de l'assimilation des données côtières en Üeckermünde, Karnin et Wolgast.

6.5.3 Amélioration des résultats du modèle

Les statistiques du tableau 6.1 sont calculées sur les séries temporelles des hauteurs d'eau en Odh1⁶ toutes les heures pendant les 40 jours. Le biais est une erreur systématique du modèle sur toute la période d'étude. On le mesure par la moyenne $1/n \sum_n (\mathbf{Z}_n^a - \mathbf{Z}_n^t)$. Les erreurs RMS sont les écarts-types observés et le total est l'erreur RMS moyenne non centrée.

A la ligne TRIM3d, on voit que le modèle sans ADS sous-estime sur la période étudiée les hauteurs d'eau en Odh1 de 3 cm. L'erreur RMS produite est de 4 cm. Ceci montre que le modèle est déjà capable de bien reproduire les hauteurs d'eau. Cependant l'utilisation de l'ADS permet de réduire de 86% le biais et de 34% l'écart-type. On peut donc envisager grâce à l'ADS des applications du modèle à des problèmes demandant une plus grande précision tels que des bilans de matière et la modélisation écologique.

Odh1	Biais	RMS	Total
TRIM3d	+0.035	0.041	0.054
Moy. mesures	+0.018	0.026	0.032
RRSQRT KF	+0.005	0.027	0.027

TAB. 6.1 – Erreurs moyennes, RMS et racine carrée des écarts quadratiques moyens exprimés en m de trois estimateurs des hauteurs d'eau en Odh1 : le modèle TRIM3D sans ADS, la moyenne des mesures des trois stations assimilées, et le filtre de Kalman RRSQRT.

Odh2	Biais	RMS	Total
TRIM3d	+0.031	0.042	0.052
Moy. mesures	+0.018	0.045	0.047
RRSQRT KF	+0.002	0.031	0.031

TAB. 6.2 – Cf. tableau précédent pour la station de validation Odh2

Dans la station Odh2, la calibration des mesures n'a pas pu être faite par des moyens physiques et l'*offset* a été calculé sur ce même *run* en calant la moyenne des mesures sur celle des hauteurs d'eau estimées par l'ADS pendant 39 jours (le premier jour a été retiré pour éviter le temps d'initialisation). Le biais d'estimation est quasiment annulé pour l'ADS et on doit voir dans les 2 mm restants l'effet de l'initialisation. Si l'on compare les résultats des tableaux 6.1 et 6.2 à ceux obtenus dans une autre application côtière comparable (Cañizares *et al.*, 2001), tant du point de vue des mesures de hauteurs d'eau utilisées (fréquence et résolution spatiale) que du modèle hydrodynamique (même résolution de grille), les résultats du modèle et de l'ADS sont légèrement plus précis dans le cas présent. On pouvait s'attendre à ce que le modèle sans ADS soit plus précis dans le cas de la lagune de l'Oder pour deux raisons : d'une part une lagune est un bassin

⁶Ou plutôt du nœud de grille où Odh1 est situé

quasiment fermé donc plus favorable que les autres systèmes côtiers bordés de grandes frontières ouvertes avec les mers et océans alentours et difficiles à observer. D'autre part on peut aussi attribuer cette différence au calme relatif des eaux de la lagune de l'Oder en cette période consécutive à l'inondation puisqu'il est plus facile de modéliser des eaux calmes que des eaux turbulentes. On peut donc être agréablement surpris par les apports conséquents de l'ADS dans un cas où le modèle dynamique est déjà très performant à lui seul.

Les tableaux 6.1 et 6.2 sont dressés à partir des différences entre résultats du modèle et mesures *in situ*. On a pu les calculer sans tenir compte de l'effet de support car la variable hauteur d'eau est suffisamment lisse pour que les moyennes horaires des mesures quasi ponctuelles soient tenues pour représentatives de nœuds de grille entiers.

6.5.4 Comparaison avec un estimateur trivial

L'interpolation faite par le filtre de Kalman au prix d'énormes calculs est-elle plus efficace qu'une interpolation triviale des données (qui nécessiterait environ 67 fois moins de calculs) ? Les données de hauteur d'eau étant spatialement très fortement corrélées (voir le modèle en figure 6.6), un krigeage à moyenne inconnue aurait été très proche de la moyenne arithmétique des données des trois stations de mesure côtières. On compare donc les résultats du modèle TRIM et de l'ADS à la moyenne des mesures.

On remarque sur le tableau 6.1 que cet estimateur trivial est meilleur que le modèle TRIM3d dans le cas de la station Odh1, ce qui était prévisible car Odh1 est située au milieu des trois stations assimilées. Par contre dans le cas de la station Odh2 qui est à l'extérieur de la région bien observée, ils donnent des résultats équivalents. La moyenne arithmétique semble aussi être un estimateur biaisé qui, en moyenne sur quarante jours, sous-estime de 1,8 cm les hauteurs d'eau en Odh1 et Odh2.

On se trouve donc dans une situation où à la fois les données et le modèle sont très informatifs : les structures spatiales étant très lisses les mesures prises sur un bord de la lagune sont de bons estimateurs des hauteurs d'eau au centre ; le modèle est lui aussi dans un cas favorable pour estimer les hauteurs d'eau, comme on l'a vu précédemment. Il est rassurant que la combinaison de ces deux informations puisse encore accroître la précision des estimations.

6.5.5 Caractéristiques spatiales

Sur la carte 6.8 à gauche, l'estimation des hauteurs d'eau est lisse, il y a moins de 10 cm de différences de hauteurs d'eau entre l'est et l'ouest de la lagune et le filtre de Kalman RRSQRT donne ainsi une image réaliste de l'état du système. Le filtre de Kalman applique des corrections à l'ensemble de la lagune et non pas seulement dans le voisinage des stations.

Les écarts-types d'estimation à droite sur la figure 6.8 montrent que l'erreur estimée par le filtre est de l'ordre de quelques millimètres, ce qui est inférieur aux erreurs qu'il commet (voir tableaux 6.1 et 6.2). La sous-estimation des écarts-types d'erreur revient tout au long de l'ADS par le RRSQRT, d'avantage sur les bassins principaux que dans les chenaux, comme sur la figure 6.8. On peut l'imputer aux propriétés «lissantes» du modèle hydrodynamique. Ces écarts-type sont maximaux (1,6 cm) à la limite près de la mer Baltique, ce qui peut s'expliquer par l'erreur de modèle, d'écart-type 1,5 cm, qui est introduite à chaque pas de temps aux embouchures des chenaux. Pour ce qui est de l'erreur au centre du bassin, il semble, au vu de la sensibilité du modèle à l'influence jointe des champs de vents et de hauteurs d'eau dans la Baltique, que celle-ci soit une combinaison de ces deux sources d'erreur. On notera que dans le fonctionnement du RRSQRT, ces deux erreurs sont introduites initialement sur des vecteurs ε_i^m distincts (4.1), mais qu'elles sont mélangées à chaque pas de temps du modèle lors de l'étape de réduction (4.15).

La carte des écarts-types d'estimation permet d'analyser les réseaux d'observation en faisant ressortir les «zones d'ombre » où l'information apportée par les mesures ne passe pas. Dans le cas présent, si la connaissance des hauteurs d'eau de la lagune avait été l'unique but du projet, cette carte n'aurait pas motivé l'implantation des stations de mesure Odh1 et Odh2, là où l'écart-type d'estimation est déjà très faible, mais aurait plutôt désigné les sorties vers la Baltique comme l'emplacement idéal pour placer les stations de mesure et parfaire nos connaissances des hauteurs d'eau de la lagune. Ces emplacements seraient en outre de bons observatoires des échanges avec la Baltique. La connaissance des hauteurs d'eau n'est cependant pas le but ultime de l'étude, mais la méthode reste valable pour les autres variables étudiées.

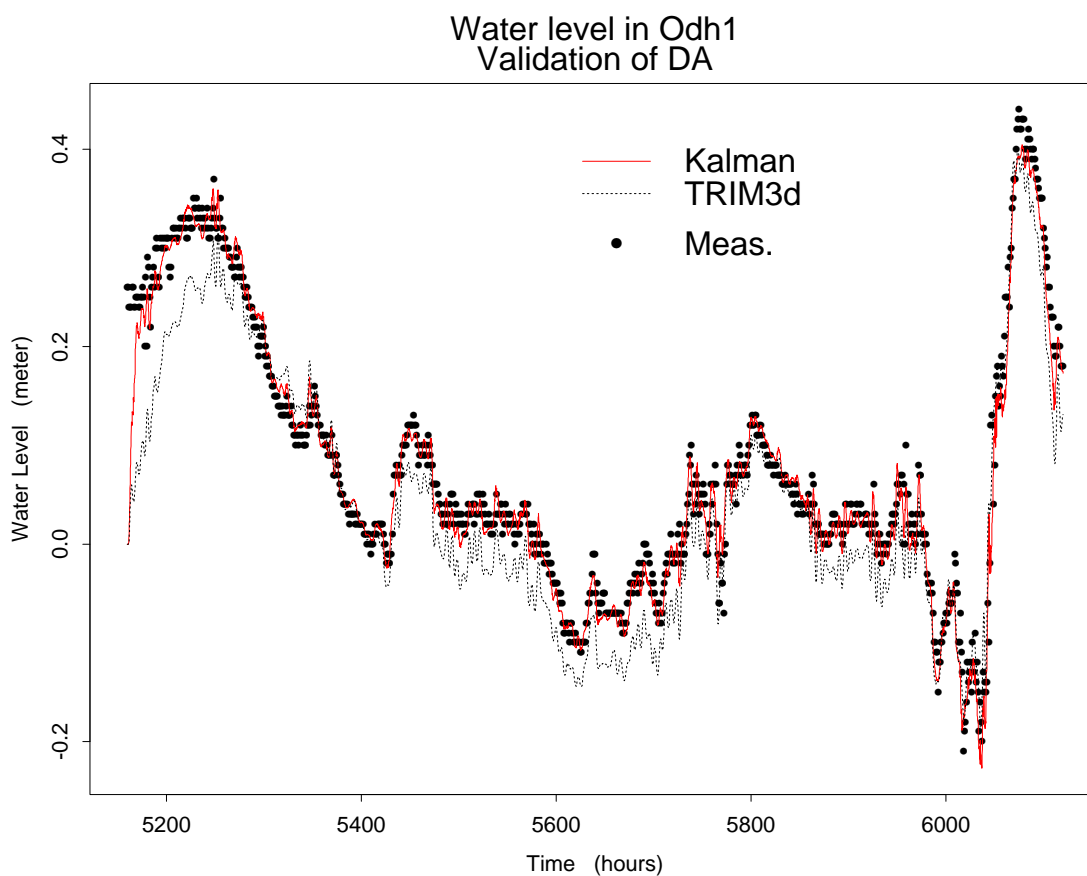


FIG. 6.7 – Séries temporelles des hauteurs d'eau mesurées (points), modélisées par le modèle TRIM3D (pointillés) et analysées par l'ADS (ligne continue rouge), les stations Odh1 et Odh2 n'ont pas été assimilées. On note que les erreurs sont globalement réduites et que le temps d'initialisation (≈ 150 h soit une petite semaine) est réduit à quelques heures. Pour superposer les courbes théoriques et expérimentales, on a dû supposer que les hauteurs d'eau mesurées en Odh1 sont valables pour le nœud de grille $250\text{m} \times 250\text{m}$ entier. En temps de calcul, la simulation des 40 jours a pris environ 2 heures pour le modèle TRIM sans ADS et 10 jours avec le filtre de Kalman RRSQRT.

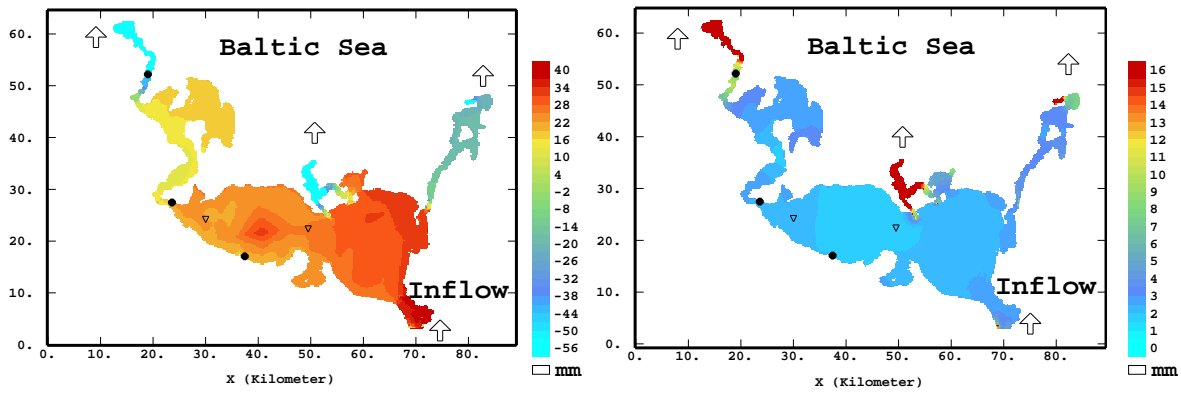


FIG. 6.8 – RRSQRT KF : cartes des hauteurs d'eau assimilées (gauche) et des écarts-types d'estimation (droite) après 10 jours et 240 cycles d'ADS. L'embouchure de l'Oder est en bas à droite de la carte et les trois chenaux vers la Baltique sont en haut à gauche, au milieu et à droite de la figure. Les stations de mesures assimilées sont marquées par un ● et les deux stations de validation (Odh1 et Odh2) par un ▽.

Chapitre 7

Assimilation jointe de hauteurs d'eau et de salinité dans la Lagune de l'Oder

Sommaire

On applique l'EnKF au modèle TRIM3D et on assimile conjointement des données de hauteurs d'eau et de salinité. Les covariances d'erreurs liées à la variable salinité sont tirées de l'analyse des mesures d'une campagne en bateau faite pendant la crue de l'Oder. On assimile dans le modèle TRIM3D les hauteurs d'eau de cinq stations fixes et les salinités de deux stations pour la période de la crue de l'Oder de juillet 1997. Les mesures prises en bateau étant gardées pour valider l'ADS. Les résultats montrent que l'ADS perd son efficacité au cours du temps car les processus physiques «lissants» du modèle créent un attracteur.

7.1 Généralités

7.1.1 Salinité et biologie de la lagune de l'Oder

La salinité d'une solution est, à un facteur 35 près, le rapport de sa conductivité avec celle d'une solution de KCl de référence. Les deux solutions doivent avoir une température et une pression fixée. Pour éviter toute confusion, la salinité est communément exprimée en psu (*practical salinity unit*) ce qui n'a aucun sens puisqu'il s'agit d'un rapport. On peut calculer aisément la salinité à partir de la conductivité par une relation

tenant compte de la température et de la pression.

Selon Fenske (1999), la salinité est le facteur le plus important pour la distribution de la communauté macrozoobenthique de l'estuaire. En effet l'eau de la lagune de l'Oder est une eau saumâtre qui abrite à la fois des espèces marines, des espèces d'eaux douces et des espèces spécifiques d'eaux saumâtres. Le seuil de salinité qui assure la transition entre espèces d'eaux douces et espèces marines est habituellement situé entre 2 et 4 psu. Or la salinité de la lagune peut varier de 0.5 à 7 psu suivant les conditions hydrodynamiques. De plus, de forts gradients verticaux peuvent aussi se présenter (mais apparemment cette stratification n'apparaît pas dans la période d'étude). Ces variations de salinité et d'autres facteurs tels que le déficit en oxygène ou la présence d'H₂S rendent le régime physiologique de la lagune très difficile. Très peu d'espèces sont adaptées à ce régime ce qui explique que l'écosystème de la lagune soit fragile. Dans ce contexte la prédiction du paramètre salinité est indispensable à celle de tous les paramètres écologiques de la lagune de l'Oder.

7.1.2 La salinité : exemple de traceur passif

On a vu l'importance des distributions de nutriments dans l'étude de la biologie d'un système côtier en section 1.1. Dans le cas de la lagune de l'Oder, la majorité des nutriments apportés par l'Oder transitent par le bassin est (appelé *Zalew Wielki* ou *Großes Haff*) via la mer Baltique, et à long terme seuls 2 à 5 % des nutriments sont retenus dans la lagune (Fenske, 1999). Cependant dans le bassin ouest (ou *Kleines Haff*) les courants sont plus faibles et le temps de résidence est plus long. On a ainsi des processus de renouvellement internes (*turnover*) plus importants. L'étude du paramètre salinité permet alors d'évaluer notre capacité à reproduire le transport des composés en solution à travers la lagune de l'Oder. La dynamique des nutriments comporte cependant un degré de complexité supérieur car ceux-ci sont aussi liés aux particules en suspension dont on connaît moins bien les dynamiques de transport.

7.1.3 Le transport du sel dans TRIM3D

La propagation du sel est assurée dans le modèle TRIM3D¹, en ajoutant aux équations décrites dans la section 6.2 la résolution numérique des équations d'advection-diffusion. Un solveur semi-lagrangien empêche le transport à travers les bords du domaine. Les deux coefficients de diffusion turbulente (*eddy viscosity*), horizontale et verticale, sont prédéfinis par l'utilisateur. Ils sont supposés constants dans le temps et uniformes, bien qu'idéalement la valeur de ces coefficients doive évoluer suivant la nature plus ou moins turbulente de l'écoulement. On verra cependant que l'écoulement dans la lagune de l'Oder est suffisamment calme, même pendant la crue de l'été 1997, pour que cette approximation soit valable. On montre aussi que le choix de ce paramètre importe peu. La salinité n'est pas à proprement parler un traceur passif car

¹Comme n'importe quel autre constituant de l'eau que l'on peut ajouter librement au modèle.

le surplus de densité dû à la constitution locale de l'eau change le gradient de pression baroclinique dans TRIM3D.

La salinité est imposée à l'interface avec la mer Baltique à la valeur 7,5 psu (d'après Morholz, spécialiste de la salinité à l'Institut de Recherche sur la mer Baltique de Warnemünde, comm. pers.). On suppose que l'eau sortant dans la Baltique est parfaitement mélangée à l'eau de mer, celle-ci reste alors constamment à 7,5 psu. La salinité de l'eau de l'Oder est supposée nulle et celle des autres affluents aussi. On reviendra sur ces hypothèses lors de la mise en œuvre de l'ADS.

7.1.4 Sensibilité du modèle

7.1.4.1 Choix du paramètre de viscosité

La figure 7.1 montre les distributions de salinité de la lagune au bout d'un *run* de quatre jours (du 16 au 20 juillet 1997, cette période ayant été fixée arbitrairement afin d'observer un état transitoire de la lagune) modélisés par TRIM3D en partant d'un état initial de salinité uniformément égale à 1,4 psu. On voit l'extension de la langue d'eau douce venant du fleuve et celle de l'eau salée venant de la mer dans les chenaux. Le paramètre viscosité prend trois valeurs : une limite inférieure 1,5, la valeur 3 utilisée couramment dans l'étude de la lagune (Rosenthal *et al.*, 1998 ; Wolf *et al.*, 2001 ; Sénégal, 1999) et la limite supérieure 5. On devine que, conformément au principe de la diffusion, la langue d'eau douce est plus ramassée quand la viscosité est plus forte mais cet effet est à peine visible. Dans la suite de l'étude, ce paramètre sera donc fixé à 3 comme lors des études précédentes.

7.1.4.2 Initialisation

Dans l'expérience suivante (figure 7.2) le modèle TRIM3D simule l'arrivée d'eau douce dans la lagune de l'Oder en commençant à deux dates initiales différentes : dans la première simulation l'état initial est uniforme (1,4 psu) au 16 juillet 1997, dans la seconde il est initié au 20 juillet 1997, cette différence est importante car le bilan de sel est changé par l'introduction d'une grande quantité d'eau douce (4 jours de débit de l'Oder en crue). Par ailleurs toutes les conditions aux limites sont maintenues identiques pour simuler le mois de la crue. On voit sur la figure 7.2 que cette différence d'état initial induit un biais de salinité à la station Odh1 alors que cette différence initiale est atténuée au bout d'un mois en Odh2. On peut en déduire que les salinités sont sensibles à des erreurs initiales à grande échelle et que les effets de ces erreurs seront plus durables dans le *Kleines Haff* à l'ouest que dans le *Zalew Wielki* à l'est car les temps de résidence y sont plus longs.

Puisque les temps d'initialisation sont bien plus longs pour la distribution du paramètre salinité que pour

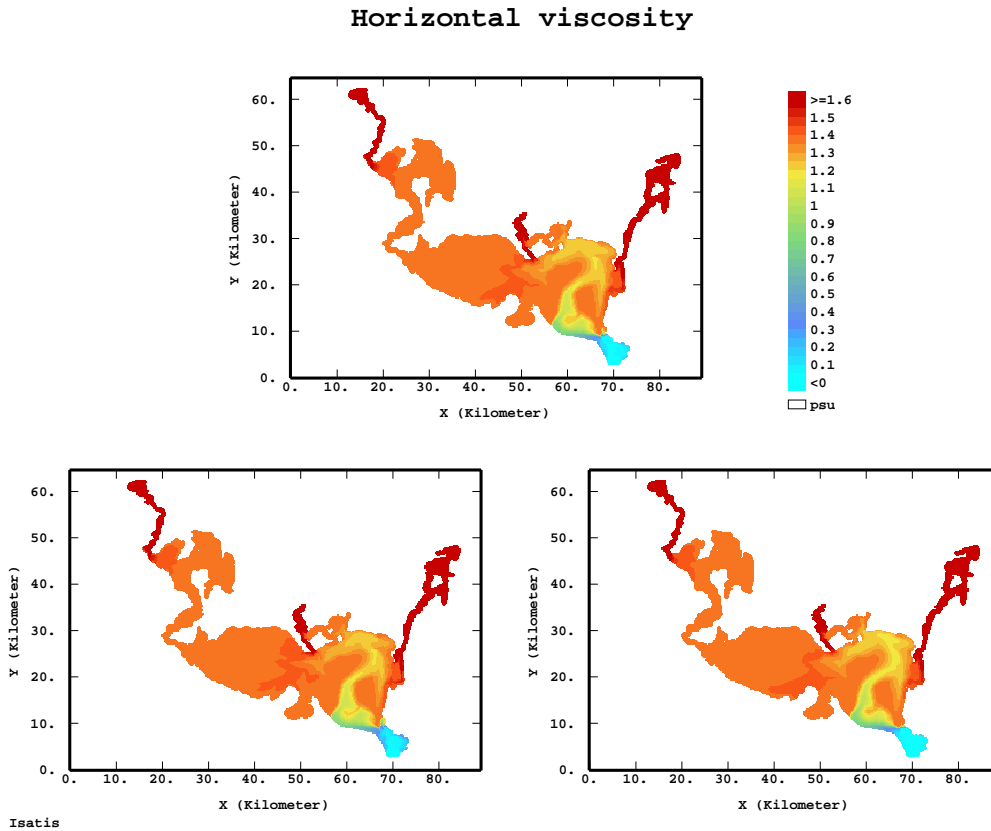


FIG. 7.1 – Distributions de salinité, avec la même échelle de couleurs : viscosité = 3 en haut, 1,5 en bas à gauche et 5 en bas à droite

les variables hydrodynamiques, une première approximation de l'état initial du système entier (variables hydrodynamiques η , u et v ainsi que salinité) est extraite d'une simulation de l'année entière par TRIM3D guidée par les conditions aux limites mais sans AD.

7.2 Analyse des données de la campagne SYNOPTA

Une campagne de mesure à bord du navire de recherche «Ludwig Prandtl» du GKSS a été menée pendant la crue de l'Oder du 22 juillet au 6 août 1997, quasiment tous les jours. Le bateau a longé le canal Świna (le canal de navigation qui va de (4) à (2) sur le plan de la figure 6.1) et trois profils verticaux ont été effectués tous les 4 km en enregistrant divers paramètres physiques (entre autres la conductivité, convertie en salinité par une relation classique tenant compte de la pression et de la température). La campagne de mesure est

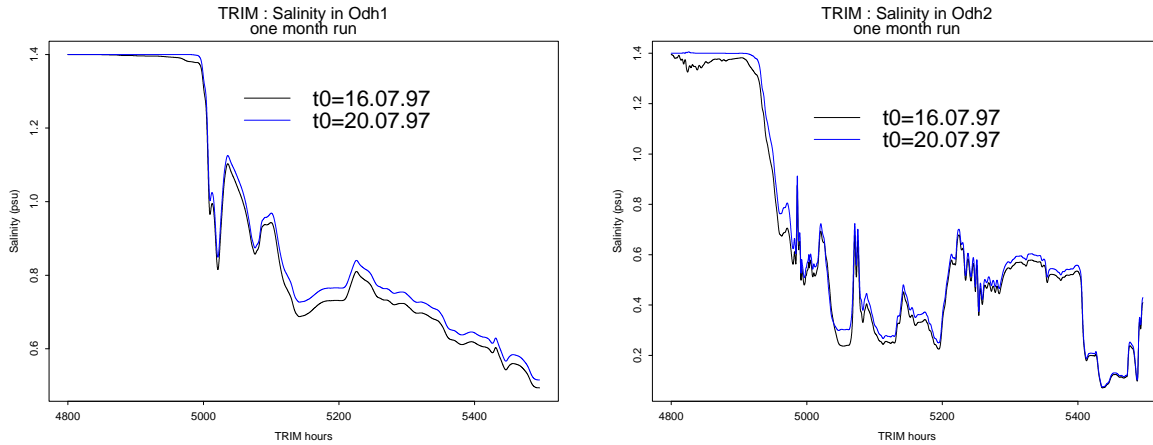


FIG. 7.2 – Salinités simulées pendant un mois par le modèle TRIM3D sans AD. Ligne rouge : départ d'un état initial uniforme le 16 juillet 97, ligne bleue : idem le 20 juillet 97. Le décalage entre les deux courbes montre la sensibilité à l'état initial, à gauche en Odh1 (*Kleines Haff*), à droite en Odh2 (*Zalew Wielki*).

nommée SYNOPTA, mais contrairement à ce que son nom indique, les données ne sont pas synoptiques². En effet les variations journalières de salinité, dues à l'évaporation, sont visibles aux points où le bateau est passé deux fois dans la même journée (Bertino et Wackernagel, 2000).

L'assimilation des données bateau aurait certes été plus efficace que celle des données des deux stations fixes, essentiellement parce qu'elles contiennent une information en profondeur que les stations fixes n'ont pas, mais ces campagnes en bateau sont exceptionnelles et leurs observations ne correspondent pas aux objectifs du projet PIONEER : évaluer les possibilités de surveillance de routine utilisant les sources d'observations *continuellement* disponibles. On assimile donc seulement les mesures des stations Odh1 et Odh2.

En faisant momentanément abstraction du modèle physique et de l'ADS, on peut utiliser l'information spatiale contenue dans ce jeu de données pour étudier les structures spatiales de la variable salinité qui ne sont pas accessibles directement par les données des deux stations fixes.

L'analyse de ces données, détaillée dans Bertino et Wackernagel (2000), montre que l'arrivée de l'eau de la crue dans la lagune forme un front d'eau douce et crée une dérive (voir les cartes de salinité krigées en figures 7.13, 7.15 et 7.17). Selon la procédure du filtre de Kalman, il revient au modèle TRIM3D de modéliser l'introduction d'eau douce dans la lagune (la dérive), et l'analyse structurale doit déterminer les structures de trois erreurs résiduelles que l'on suppose stationnaires : les erreurs à l'instant initial ε_0 , erreurs

²On appelle «synoptiques» des données qui photographient le système entier en un seul instant t , comme si on pouvait l'observer en un seul coup d'œil.

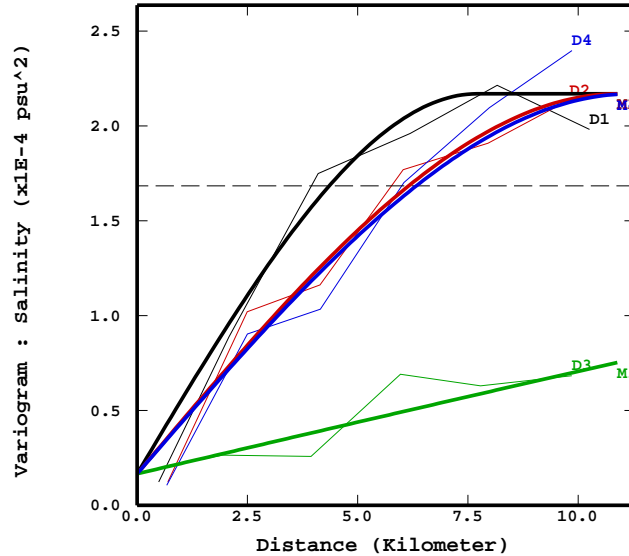


FIG. 7.3 – Variogrammes expérimentaux et théoriques pour les données de salinité du 5 août 1997 anisotropes dans le plan $(x - z)$ le long du canal Świna, D1 est l'axe horizontal du canal, D3 est la verticale, D2 et D4 sont les axes intermédiaires à 45° . L'échelle verticale a été modifiée (1km correspond à 1m de profondeur) par commodité.

de modèle ε_n^m et erreurs de mesure ε_n^o . On doit donc s'intéresser à la structure sous-jacente des données, masquée par la dérive. On considérera donc les données du 5 août 1997, journée bien informée car les profils sont faits tous les 2 km au lieu de 4 km et où l'on retrouve la stationnarité spatiale recherchée car le front d'eau douce est déjà avancé au delà de la moitié de la lagune. De plus l'eau étant plus douce après la crue, les effets de l'évaporation sont moins forts et on peut supposer que ces données sont quasiment simultanées (ou «synoptiques»). Le variogramme expérimental dans le plan $(x - z)$ est modélisé par une structure sphérique anisotrope (figure 7.3), de portée 8 km dans la direction parallèle au canal Świna et 20 m dans la verticale. Comparée aux variogrammes expérimentaux des données des autres transepts, cette structure peut raisonnablement passer pour la structure sous-jacente des données non stationnaires en ôtant la part du variogramme due à la dérive.

On ne peut pas mettre à profit les covariances croisées entre les séries temporelles de salinité comme on l'a fait pour la variable hauteur d'eau en 6.4.2.2. En effet, il existe bien un décalage temporel de 15 jours entre les séries temporelles des stations Odh2 et Odh1 qui correspond en moyenne à la différence des temps d'arrivée de l'eau de la crue de l'Oder en chacun des deux points. Si par analogie avec le raisonnement fait en 6.4.2.2 on suppose un écoulement laminaire entre les stations Odh2 et Odh1, distantes d'environ 30 km, on aboutit à une vitesse de transport de 2 km par jour, ce qui est bien inférieur au rapport entre la portée spatiale horizontale calculée ci-dessus (8 km) et la portée du variogramme temporel (1 jour, non montré ici).

Le fait est que l'écoulement dans la lagune n'est pas laminaire mais turbulent et le front d'eau douce suit un parcours tortueux, bien plus long que les 30 km qui séparent les deux points.

7.3 Mise en œuvre de l'EnKF

On assimile conjointement les données horaires de hauteur d'eau des cinq stations fixes de la lagune et les séries temporelles de salinité des deux stations du GKSS, Odh1 et Odh2 (voir la carte 6.1) dont la résolution temporelle est de 10 min.

7.3.1 Adaptation des conditions aux limites

L'ADS de hauteurs d'eau dans le modèle TRIM3D provoque une discontinuité de la variable η aux conditions limites. En effet les hauteurs d'eau à l'interface avec la mer Baltique sont imposées par les mesures de la station Koserow alors que les hauteurs d'eau des nœuds de grille adjacents sont réévalués par l'ADS. Si cette discontinuité n'a pas d'effets notables sur les paramètres hydrodynamiques (au pas de temps suivant, la discontinuité est effacée par advection), par contre le traitement de la salinité dans TRIM3D n'y est pas adapté et génère des valeurs aberrantes de salinité. Pour éviter cela, on «ramollit» les conditions aux limites (astuce soufflée par Geir Evensen) : la hauteur d'eau à un point \mathbf{x} à l'interface avec la mer Baltique n'est pas directement celle de la station Koserow η_0 , bruitée par l'erreur de modèle $\varepsilon^m(\mathbf{x})$, mais une estimation $\eta^*(\mathbf{x})$ qui y tend progressivement suivant l'équation de relaxation :

$$\frac{d\eta^*(\mathbf{x})}{dt} = -\frac{1}{\lambda}(\eta^*(\mathbf{x}) - \eta_0) + \varepsilon_t^m(\mathbf{x}), \quad (7.1)$$

avec λ la constante de temps qui donne la vitesse de convergence de l'estimation $\eta^*(\mathbf{x})$ de sa valeur initiale (donnée par l'ADS) vers la valeur spécifiée η_0 . L'équation (7.1) est discrétisée sur le pas de temps de TRIM3D et le paramètre λ est ajusté à tâtons à sa valeur la plus faible possible, en l'occurrence 25 min. La convergence devrait donc être quasiment obtenue en 5 pas de temps du modèle et l'estimation $\eta^*(\mathbf{x})$ est donc faite à 80% de la valeur courante donnée par l'ADS et à 20% par la valeur spécifiée η_0 . On ne supprime pas pour autant le conditionnement aux limites car les variations de hauteur d'eau dans la Baltique sont suffisamment lentes pour s'imposer à la longue³.

³l'échelle typique de variation des hauteurs d'eau est de l'ordre de quelques jours alors qu'un pas de temps du modèle est de 5 min

7.3.2 L'erreur initiale

On suppose que la structure spatiale anisotrope modélisée précédemment (figure 7.3) est représentative de celle des erreurs initiales ε_0 le 17 juillet 1997. Puisque les mesures du 5 août 1997 utilisées pour l'analyse structurale ont été prises le long du canal Świna dans l'eau de la crue, on applique une correction affine du variogramme pour qu'elles correspondent à la situation initiale où l'eau de la lagune était plus salée. Le palier du variogramme est ajusté arbitrairement à $0,02 \text{ psu}^2$ au lieu de $2 \cdot 10^{-4} \text{ psu}^2$. Disposant d'un modèle de covariance, on peut simuler une variable gaussienne suivant ce modèle. La méthode utilisée ici est celle des bandes tournantes (Lantuejoul, 2001). On tire autant de réalisations qu'il y a de membres dans l'ensemble et on leur ajoute l'état initial du 17 juillet 1997 pour obtenir l'ensemble initial de simulations de l'état \mathbf{Z}_0^t , gaussien de moyenne égale à l'état initial et de covariance stationnaire. La première étape de propagation de chacune des réalisations dans le modèle dynamique enlèvera le caractère stationnaire de l'erreur d'estimation.

On remarque aussi que l'initialisation par une covariance initiale non nulle est moins aisée dans un filtre dérivé de l'EKF que par une approche de Monte Carlo comme ci-dessus ; premièrement parce que la matrice de covariance initiale de taille $N \times N$ ne peut pas être écrite explicitement pour des raisons d'encombrement ; ensuite parce qu'une description simplifiée entraîne d'autres difficultés pratiques. Dans le cas présent, deux méthodes classiques ont été testées pour obtenir les valeurs propres et vecteurs propres dominants de la matrice de covariance stationnaire ci-dessus : une transformée de Fourier et l'algorithme de Lanczos.

La première est envisageable car la covariance est définie sur une grille régulière et la matrice est donc contenue dans une matrice de Toeplitz que l'on obtient par symétrie du domaine par rapport aux trois plans de base. La matrice résultante est alors obtenue entièrement par permutations du premier vecteur. Suite à l'extension de la grille du modèle à une grille parallélépipédique puis aux trois symétries, la taille de ce vecteur dépasse toutefois le million. Sa transformée de Fourier fait essentiellement ressortir le terme constant et toutes les autres puissances oscillent autour de 0. On découvre ainsi que la positivité de la matrice de covariance est perdue lors de la symétrie et que la transformée de Fourier n'est donc pas exploitable. Cette méthode ne semble pas adaptée aux matrices de grandes tailles, mais pour des matrices plus petites pour lesquelles le caractère défini positif de la matrice de covariance a plus de chances d'être conservé.

L'algorithme de Lanczos, basé sur des multiplications itératives d'un vecteur par la matrice à diagonaliser, a fait l'objet de nombreux développements et il est très efficace dans le cas où les premières valeurs propres se détachent clairement des autres (le vecteur trouve alors naturellement les directions propres au fil des itérations). Cependant, une implémentation de l'algorithme primitif montre après 100 puis 1000 itérations que les premières valeurs propres sont multiples et que leur degré de multiplicité est inconnu. Cette matrice présente donc un cas difficile pour l'algorithme de Lanczos et sans doute aussi pour d'autres algorithmes itératifs.

7.3.3 Pourquoi l'EnKF ?

C'est pour identifier le problème de stabilité du filtre de Kalman, mais aussi pour la simplicité de l'initialisation du filtre que j'ai utilisé l'EnKF au lieu du filtre de Kalman RRSQRT. En effet ce dernier ayant lui-même un problème de stabilité interne (discuté en section 5.4.2), l'origine des valeurs aberrantes est plus difficile à rechercher quand le paramètre de linéarisation ϵ de (4.12) doit aussi être ajusté. A l'inverse, l'EnKF est connu pour sa stabilité et il ne requiert pas d'autre paramètre interne que la taille de l'ensemble. Dans une approche de Monte Carlo, il est aussi beaucoup plus facile de trouver des algorithmes généraux pour résoudre le problème de l'initialisation dans les systèmes à grandes dimensions.

7.3.4 L'erreur de modèle

Tout comme dans la section 6.4.2.2, on suppose une erreur aux conditions aux limites du système (hauteurs d'eau dans la mer Baltique, champs de vent); leurs caractéristiques restant inchangées. De plus les salinités de l'eau de la Baltique et de l'Oder sont aussi affublées d'une erreur.

On suppose que la salinité en chaque point de contact avec la mer Baltique, imposée à 7,5 psu, est connue à une erreur ϵ_n^m près, gaussienne de moyenne nulle et dont la covariance spatiale est proportionnelle à la covariance théorique modélisée sur les données bateau du 5 août 97 (figure 7.3). Compte tenu des portées horizontales (8 km) et verticales (20 m) de ces erreurs, il apparaît que cette erreur est indépendante d'un chenal à l'autre et qu'au sein d'un même chenal les seules variations significatives seront verticales (d'une couche à l'autre) mais non pas horizontales car l'extrémité de chaque chenal est trop étroite. On a donc regroupé les points d'interface appartenant à la même couche et à l'extrémité du même chenal, ce qui économise le temps de calcul sans approximation excessive. La variance de cette erreur est fixée à 2 psu^2 car selon Morholz, la salinité peut varier entre 7 psu et 8,5 psu. Temporellement un modèle auto-régressif aurait été plus judicieux qu'un bruit blanc car il est peu vraisemblable que les salinités de la mer Baltique puissent changer de quelques psu toutes les cinq minutes. Cependant cette modification n'aurait pas changé fondamentalement les résultats ci-dessous.

Aux points d'interface avec l'Oder, comme la salinité est faible (inférieure à 0,1 psu selon les données de la campagne SYNOPTA), on suppose que l'erreur suit un modèle lognormal dont on a vu en section 5.5.4 qu'il évitait les défauts du modèle gaussien pour les variables positives aux valeurs faibles. La salinité d'entrée est donc l'exponentielle d'une variable gaussienne de moyenne -4 et de variance 3. Ces paramètres ont été fixés arbitrairement afin que la probabilité que salinité de l'Oder reste inférieure à 0,1 psu soit de 95%.

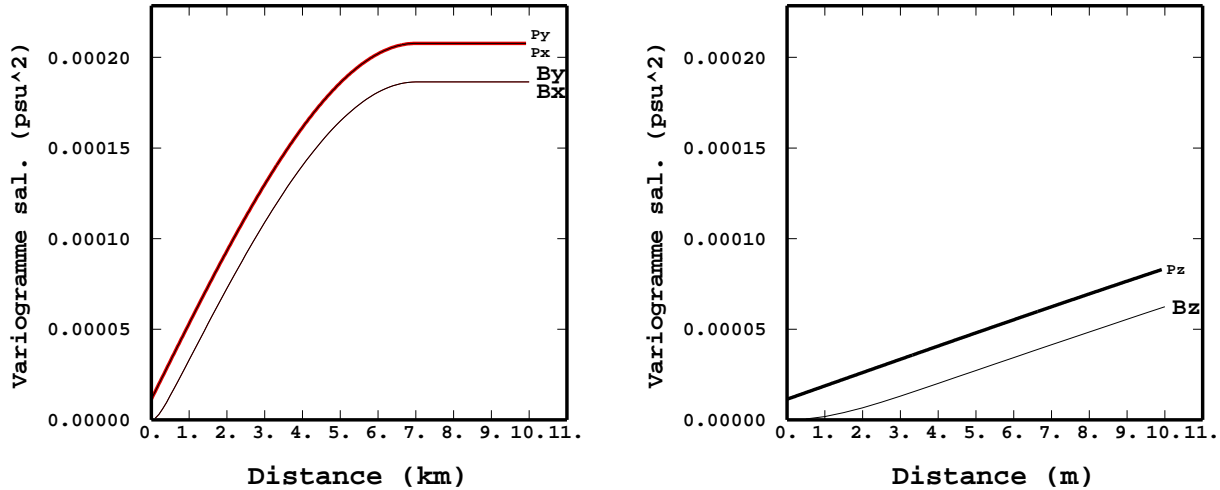


FIG. 7.4 – Variogrammes théoriques ponctuels (P) et régularisés par blocs (B) pour les données de salinité du 5 août 1997 anisotropes. A gauche les directions horizontales x et y , à droite la direction verticale. Les directions intermédiaires ne sont pas représentées.

7.3.5 L'erreur d'observation

7.3.5.1 Effet de support

Les difficultés liées à l'effet de support en filtre de Kalman ont été discutées en section 3.2. Afin de visualiser dans le cas présent l'influence du changement de support, on choisit un jeu de données de salinité dont on suppose que les caractéristiques spatiales sont assez représentatives des erreurs d'estimation non stationnaires du filtre de Kalman : celles de la campagne en bateau du 5 août. Puis on compare le modèle de variogramme ponctuel (figure 7.3) au variogramme de sa régularisée sur la grille du modèle TRIM3D (de maille v 250 m par 250 m par 3 m de profondeur) :

$$\gamma_v(\mathbf{x}) = \int \int \int_v \gamma(\mathbf{x} + \mathbf{u}) d\mathbf{u} - \int \int \int_v \gamma(0 + \mathbf{u}) d\mathbf{u}. \quad (7.2)$$

Sur la figure 7.4 une différence nette apparaît entre les variogrammes ponctuels et régularisés. Le comportement du variogramme régularisé est parabolique aux petites distances (inférieures à 500 m en horizontal et 2 m en vertical) alors que celui du variogramme ponctuel comporte un effet de pépite (son ordonnée à l'origine) et une pente non nulle à l'origine. Cependant comme ces échelles sont voisines de la taille de la grille, le comportement parabolique du variogramme régularisé ne concerne que les couples de voisins immédiats. On peut donc décrire la différence par une translation verticale (ce qui revient à supposer que les

micro-structures filtrées par le changement de support ne sont que l'effet de pépité de variance 10^{-5} psu², soit 0.003 psu d'écart-type). Conformément à la discussion faite en section 3.2, on va ajouter cet effet de pépité à l'erreur de mesure, il sera donc filtré lors de l'estimation des valeurs de blocs \mathbf{Z}_n^a , en supposant de plus qu'il soit représentatif de l'effet de support pour les covariances (3.9) non stationnaires à chaque pas de temps. Compte tenu du faible nombre de données assimilées (deux), il semble difficile d'estimer à chaque pas de temps l'effet de pépité filtré.

7.3.5.2 Calcul des paramètres

Les erreurs de mesure de salinité, aux deux stations Odh1 et Odh2, sont gaussiennes et purement pépitiennes. Leur écart-type est calculé statistiquement sur les données de la campagne SYNOPTA avec un facteur de correction prenant en compte le degré de salissement des conductimètres des stations fixes (voir les effets de ce salissement dans Bertino et Wackernagel (2000)). On aboutit au final à un écart-type de 0,037 psu. On lui a ajouté l'effet de pépité provenant du changement de support (d'écart-type 0,003 psu) qui est marginal en comparaison.

7.4 Résultats

7.4.1 Un *run* d'ensemble

Le processus de diffusion du modèle de transport de sel a une propriété «lissante». En effet la diffusion tend à homogénéiser les distributions de sel. Ceci peut se constater en utilisant une méthode de Monte Carlo, ou «*run* d'ensemble» selon la terminologie d'Evensen (1994). Il s'agit d'une batterie de tests parallèles du modèle, différant par leurs conditions initiales d'un bruit aléatoire ε_0^a (dont l'écart-type est réduit à 0,02 psu au lieu de 0,14 psu pour ce *run* d'ensemble et pour la comparaison suivante en 7.4.2) et par leurs conditions aux limites d'un bruit aléatoire de modèle ε_n^m comme décrit ci-dessus. Ainsi la figure 7.5 montre à gauche l'évolution temporelle des écarts-types de salinité calculés sur un ensemble de 50 réalisations en trois points du *Kleines Haff*. Ces écarts-types sont décroissants pendant les 4 premiers jours, puis se stabilisent à un niveau bas. Cette décroissance de la variance dans les premiers temps est le signe d'une perte de diversité initiale des 50 réalisations de la lagune qui convergent vers un *attracteur* commun. On peut expliquer ainsi cette perte de diversité : la variabilité à petite échelle introduite dans les simulations de l'erreur initiale de salinité ε_0^a est lissée par le processus de diffusion et les 50 réalisations se ressemblent beaucoup plus après quelques heures de fonctionnement du système. La variance résiduelle après plusieurs jours peut résulter des composantes à plus grande échelle de l'erreur initiale entretenues par les erreurs de modèle simulées aux conditions aux limites (hauteur d'eau et salinité dans la Baltique, champ de vent). A droite de la figure 7.5, les écarts-types des trois stations plus proches de la mer Baltique laissent deviner des comportements

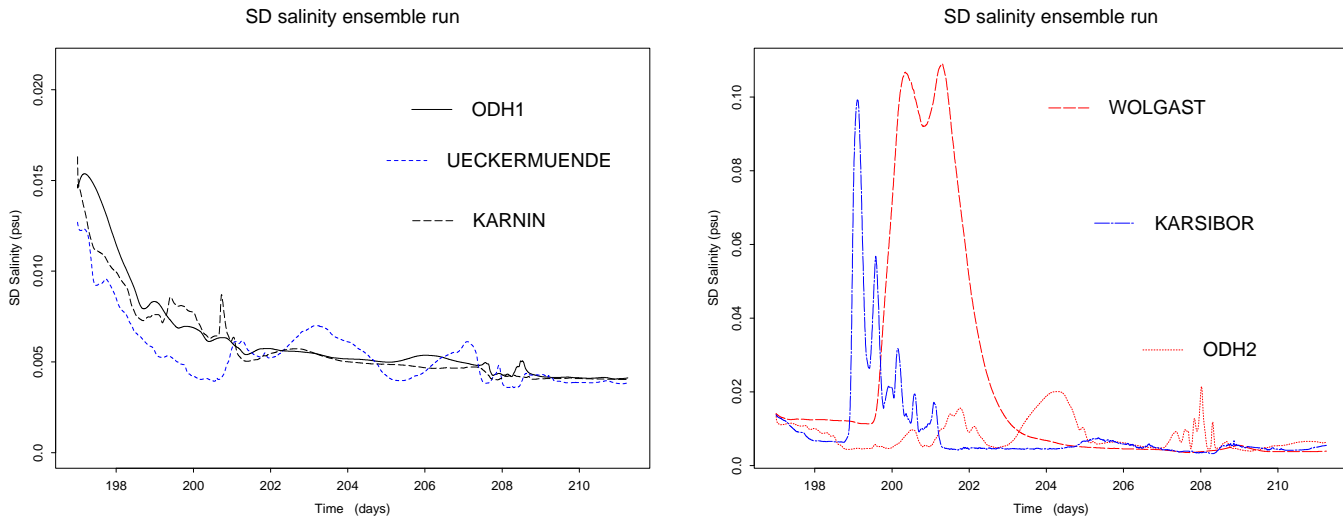


FIG. 7.5 – Écarts-types de salinité pour 50 réalisations d'un *run* d'ensemble (sans AD) en trois nœuds de grille correspondant à trois stations du *Kleines Haff* à gauche et trois nœuds plus proches de la mer Baltique à droite où les erreurs de modèle aux conditions limites provoquent des instabilités.

plus instables, dus à l'introduction d'erreurs de modèle sur la salinité à l'interface avec la mer Baltique.

Ce résultat contredit le test de sensibilité en section 7.1.4. En effet la décroissance de l'écart-type de salinité qui apparaît lors du *run* d'ensemble ne signifie pas que la salinité de la lagune est indépendante des conditions initiales mais indique plutôt la présence d'un attracteur pour tous les états de l'ensemble initial qui sont trop ressemblants (leurs différences sont locales, de portée 8 km, globalement les bilans de matière sont les mêmes pour tous les membres). Cet attracteur n'est donc certainement pas unique et dépend des modèles donnés à ε_0^a et ε_n^m . On peut en déduire que l'erreur sur l'ensemble initial (et les conditions aux limites) n'est pas la plus appropriée puisqu'elle ne fait pas ressortir la sensibilité du modèle aux structures larges. Pour parfaire la modélisation de l'erreur initiale on pourrait revenir sur l'hypothèse de stationnarité et, tout en écartant volontairement l'utilisation du modèle qui nous mettra inéluctablement sur la voie de son propre attracteur, se baser sur l'analyse structurale d'autres données non stationnaires et plus représentatives de l'état initial au 17 juillet 1997, ce qui permettrait en particulier de générer des simulations initiales avec des moyennes différentes (donc des bilans de sel différents).

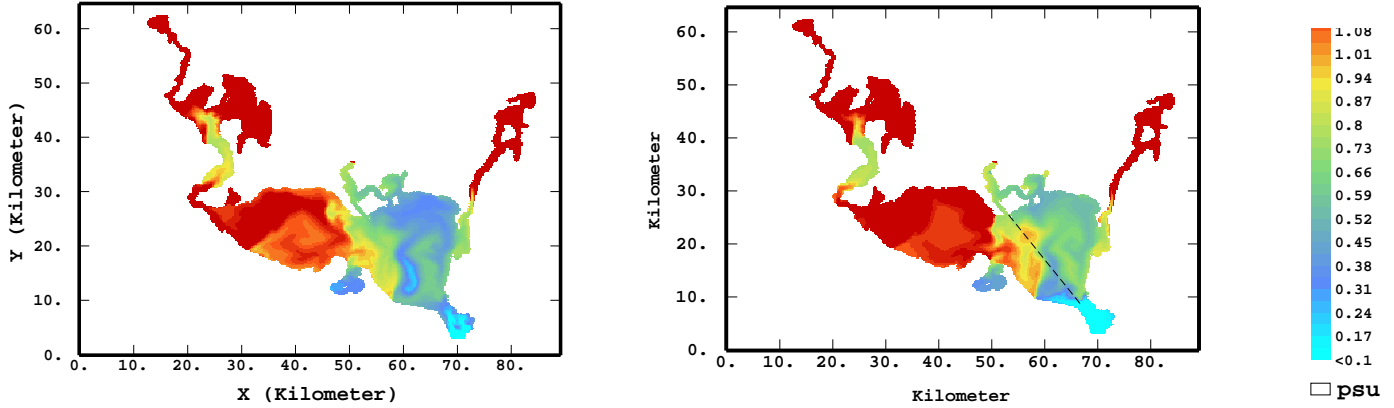


FIG. 7.6 – Salinités assimilées le 22 juillet 1997 avec l’EnKF 50 membres à gauche contre 100 membres à droite, les salinités sont moins élevées dans la partie est avec 50 membres seulement.

7.4.2 Influence du paramètre de taille de l’ensemble

Dans une expérience préliminaire, on applique l’EnKF avec deux dimensions différentes de l’ensemble (50 et 100 membres, les tailles supérieures n’ont pas été explorées pour limiter le temps de calcul). Après 5 jours d’assimilation jointe de hauteurs d’eau et de salinité, on constate des différences au niveau des salinités du bassin est, le moins informé des deux, et marginalement dans le bassin ouest. Il est possible, contrairement au cas du chapitre 5, que la taille nécessaire à cet état de grandes dimensions soit supérieure à 100.

7.4.3 Gradients verticaux de salinité

Etrangement, le modèle TRIM3D sans ADS ne simule pas de gradients verticaux de salinité, quel que soit le nombre de couches verticales utilisées dans la modélisation et le gradient vertical de salinité imposé aux conditions aux limites dans la mer Baltique. Thomas Wolf propose l’utilisation de la version non hydrostatique de TRIM pour parvenir à générer ces gradients.

L’ADS offre par contre la possibilité d’imposer un gradient vertical arbitraire en introduisant une dérive verticale lors de la correction comme décrit dans la section 3.7. Cette possibilité n’a pas été explorée ici car durant la période de la crue de l’Oder de 1997, les gradients verticaux sont faibles par rapport aux gradients horizontaux (voir figures 7.13 et 7.15) et leur forme est mal connue.

On peut aussi espérer que l'assimilation de données de surface affecte inégalement les nœuds de grille suivant leur hauteur : ceux proches de la surface seront plus influencés que ceux du fond. Cependant dans le cas présent, malgré la portée verticale de seulement 20 m de la covariance initiale et de la covariance de l'erreur aux conditions limites, aucun gradient vertical n'apparaît sur les résultats de l'ADS. Il semble donc que le modèle TRIM3D efface les gradients verticaux sur chaque membre de l'ensemble à chaque étape d'ébauche, peut-être par diffusion verticale. Dans le cas présent les données de surface sont insuffisantes pour recréer une caractéristique absente du modèle et des observations tout au long de la verticale seraient certainement plus efficaces pour simuler un gradient vertical lors de l'étape de correction. Mais il est probable que les propagations suivantes dans le modèle rendent la salinité uniforme sur toute la colonne d'eau car l'ADS considérerait les gradients verticaux observés comme une erreur de mesure. On devrait donc revenir sur la mise en œuvre du modèle TRIM3D lui-même pour corriger ce défaut, ce qui prouve aussi que l'ADS faite sur un modèle imparfaitement calibré ne peut pas faire de miracles puisque c'est le modèle physique qui transforme les covariances d'erreur à sa guise.

7.4.4 Performances et temps de calcul

La période de simulation couvre la crue de l'Oder à partir de l'instant où des données de salinité ont été disponibles (c'est-à-dire du 17 juillet 1997 à minuit) jusqu'au 13 août 1997 à minuit. Le ratio temps réel sur temps machine $T_{\text{réel}}/T_{\text{CPU}} = 3,95$ sur une station de travail Sun SPARC ce qui signifie qu'une prédiction à l'échelle de 15 jours peut être faite en 4 jours de calculs, ce qui laisse 11 jours pour réagir. Certaines mesures préventives nécessiteront certainement des délais plus grands, mais dans ce cas on peut obtenir des réductions plus importantes du temps de calcul en rétrogradant TRIM3D vers un modèle en deux dimensions au lieu de trois (puisque les gradients verticaux sont cachés) et en ayant recours à des calculs parallèles, ce qui est très facile pour l'EnKF, naturellement parallélisable en tant que méthode de Monte Carlo. Par contre paralléliser l'algorithme du filtre de Kalman RRSQRT serait moins efficace à cause du temps de calcul de l'étape commune de réduction. Dans les applications où le temps de réponse doit être immédiat, l'ADS peut être utilisée pour générer les champs initiaux et le modèle sans ADS, donc plus rapide, pour donner une prédiction simple sans erreur aléatoire.

7.4.5 Structures non stationnaires simples et croisées

7.4.5.1 Décroissance des variances-covariances d'estimation

La figure 7.7 montre sur la colonne de gauche les écarts-types d'estimation de la salinité les trois premiers jours de l'assimilation après 1, 145 et 289 étapes de correction, et à droite les covariances entre l'erreur de salinité d'un point typique de la lagune (Odh2) et les autres nœuds de grille en surface. Au premier pas de temps, l'ADS pratique un krigeage simple (ou interpolation optimale) de la salinité en Odh2 (la mesure en Odh1 n'était pas disponible le 17 juillet à minuit). La carte d'écart-type d'estimation à gauche est donc pratiquement celle d'un krigeage classique, à l'exception de l'embouchure de l'Oder où les salinités sont proches de 0 et la troncature des valeurs négatives réduit la variance de l'ensemble (voir discussion en section 5.5.4). On voit autour du point Odh2 au centre de la lagune une zone d'influence où l'écart-type d'estimation est plus faible, de diamètre correspondant à la portée de la covariance initiale (8 km). En haut à droite, la carte de la covariance entre la salinité à Odh2 et tous les autres points ne correspond pas aux écarts-types à gauche car la covariance bloc à bloc est traitée de manière non stationnaire. Celle représentée a une forme complexe puisqu'elle a été calculée à partir de 100 simulations de Monte Carlo (ce qui est peu pour estimer une covariance), puis mise à jour par l'assimilation d'une donnée de la station Odh2.

Les deux jours suivants (au milieu et en bas de la figure 7.7), puis les deux semaines suivantes (figure 7.8), les zones d'influence des stations Odh1 et Odh2 prennent des formes plus effilées en suivant les lignes d'écoulement de l'eau de la lagune. La variance globale de la salinité décroît à un niveau bien inférieur à celui de la variance de l'erreur initiale $\sqrt{\text{Var}(\varepsilon_0^a)} = 0,14$ psu, comme c'était le cas pour le *run* d'ensemble dans la région du bassin ouest en figure 7.5 à gauche. Près des embouchures, l'erreur de modèle est introduite et entretient un écart-type d'estimation plus élevé. Cet effondrement est explicable comme lors du *run* d'ensemble par les processus d'advection et de diffusion horizontale du modèle TRIM3D qui ont pour effet de gommer les variations à petite échelle de l'ensemble initial. Par rapport au *run* d'ensemble, cet effet du modèle est encore plus marqué car l'étape de correction du filtre de Kalman, en tant que krigeage, a elle aussi un effet lissant. Les cartes de covariances entre Odh2 au centre de la lagune et les autres points montrent (à droite sur les figures 7.7 et 7.8) que cet effet a aussi tendance à uniformiser les covariances. Les structures initiales à petite échelle visibles en haut à droite de la figure 7.7 ont disparu en moins de 24h et les structures à grande échelle disparaissent elles aussi jusqu'à ce que la covariance tende uniformément vers 0.

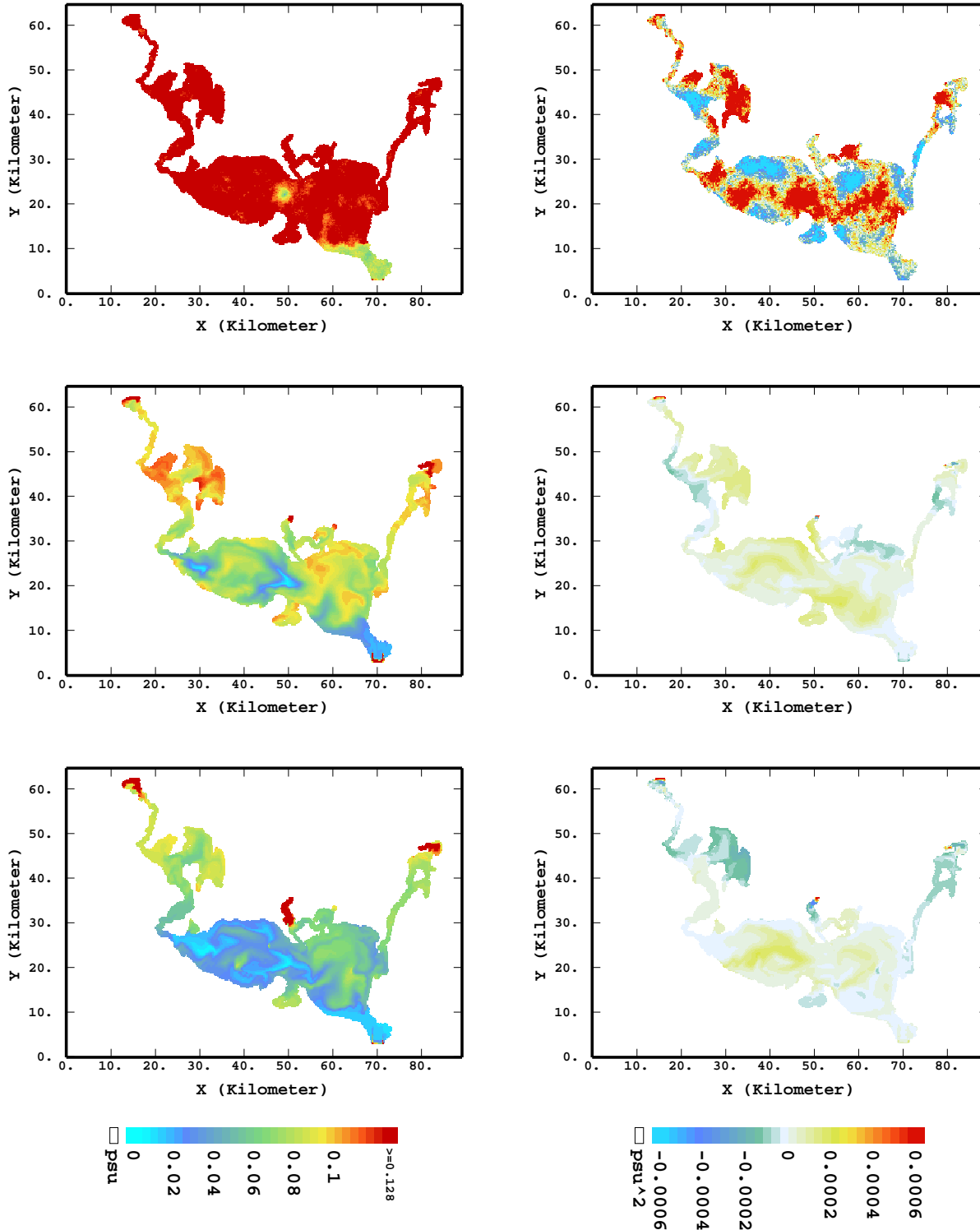


FIG. 7.7 – Initialisation de l'EnKF : à gauche = écarts-types d'estimation de la salinité $\sqrt{\text{Var}(s_n^a)}$, à droite = covariances entre la salinité au nœud Odh2 et les salinités aux autres nœuds de grille $\text{Cov}(s_n^a(\text{Odh2}), s_n^a(\mathbf{x}))$, après la première assimilation (17 juillet à minuit), 1 jour (18 juillet) et 2 jours (19 juillet).

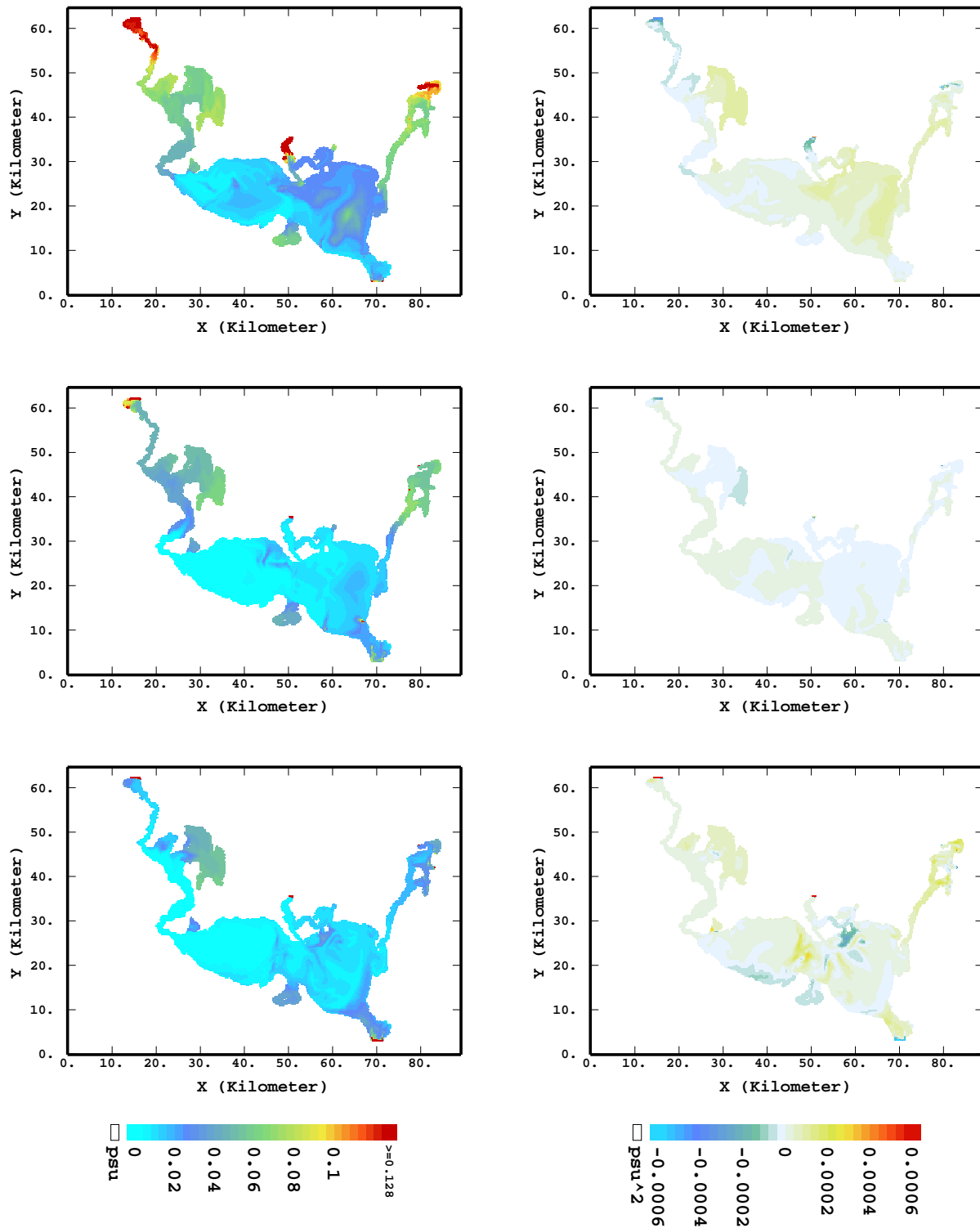


FIG. 7.8 – EnKF : à gauche = $\sqrt{\text{Var}(s_n^a)}$, à droite = $\text{Cov}(s_n^a(\text{Odh2}), s_n^a(\mathbf{x}))$, après 3 jours (20 juillet), 6 jours (23 juillet) et 11 jours (28 juillet).

7.4.5.2 Conséquences sur l'ADS

D'après les équations du filtre de Kalman (3.27) et (3.22) on voit que les corrections qu'il apporte sont importantes pendant son temps d'initialisation, puis s'estompent progressivement au fur et à mesure que l'erreur d'estimation a une covariance plus faible, ce qui rend les corrections de l'ADS de moins en moins efficaces. La décroissance des variances d'estimation laisse croire que le système n'a aucune «zone d'ombre». Elle signifie plutôt que le filtre reconnaît de moins en moins ses erreurs et donne des résultats avec de plus en plus de certitude. On pourrait alors le qualifier «d'optimiste» puisque ce filtre analyse ses propres erreurs et les considère comme très faibles alors que les erreurs de mesures restent constantes. En termes d'équilibre entre modèle et observation, on voit ici que l'ADS donne systématiquement raison au modèle plutôt qu'aux observations puisque la variance de l'erreur d'estimation devient beaucoup plus faible que celle de l'erreur de mesure.

7.4.5.3 Corrélations *a posteriori* entre une salinité et les hauteurs d'eau

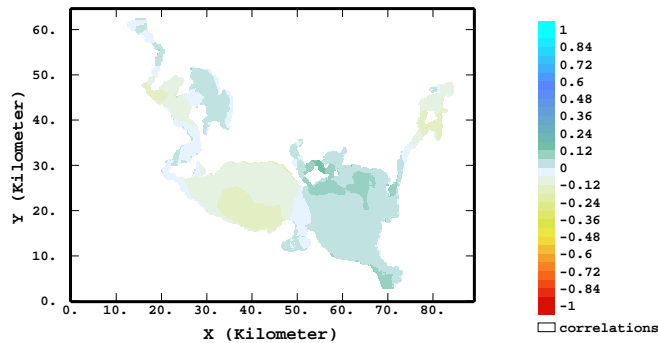


FIG. 7.9 – Corrélations entre la salinité du nœud de grille de Odh2 (entre les deux bassins) et les hauteurs d'eau des autres nœuds de grille $\text{Corr}(s_{\text{Odh2}}, \eta_x)$. 18 juillet 1997, 24h d'assimilation de données. Pour plus de clarté les corrélations ont été représentées plutôt que les covariances car les écarts-types d'estimation de hauteur d'eau sont très contrastés entre les bassins et les chenaux comme à la figure 6.8.

A l'instant initial, les variables salinité et hauteur d'eau sont supposées indépendantes. Or on voit sur la figure 7.9 qu'après un jour d'ADS (soit 145 étapes de correction), des corrélations croisées sont créées, qui ne sont pas seulement le fait d'une variabilité à la Monte Carlo. En effet ces corrélations, bien que très faibles (inférieures à 0,2 en valeur absolue) présentent une structure nette : la salinité au point Odh2 (limite entre les deux bassins) est corrélée positivement avec les hauteurs d'eau du bassin est et négativement avec celles du bassin ouest. On peut aussi voir un découpage similaire du petit bassin situé au nord-ouest dans le chenal du Peenestrom, ce qui suggère une interprétation de ce résultat à l'aide d'une troisième variable :

le vent. Un vent venant de l'est peut à la fois provoquer une différence de hauteur d'eau d'est en ouest et apporter de l'eau douce de l'embouchure de l'Oder vers Odh2, qui fait baisser la salinité de l'eau. La salinité en Odh2 est donc corrélée positivement avec la hauteur d'eau des régions où l'eau descend (à l'est des bassins principaux de la lagune et à l'est du bassin mineur du Peenestrom) et corrélée négativement avec les hauteurs d'eau des régions où l'eau monte. Inversement, un vent d'ouest ferait augmenter la salinité de l'eau en Odh2 en apportant de l'eau plus salée alors que les hauteurs d'eau varieraient en sens inverse. Cet effet du vent peut expliquer pourquoi les variables hauteur d'eau et salinité sont corrélées alors que le lien physique entre les deux n'est pas immédiat. La corrélation est faible et son influence dans les résultats suivants est certainement très faible aussi.

7.4.6 Comparaison avec les résultats du modèle

Sur les cartes produites par le modèle TRIM3D sans ADS et contrôlé par les conditions aux limites depuis le mois de janvier 1997 (colonne de gauche sur les figures 7.10 et 7.11), le front d'eau douce est absent des bassins principaux et les salinités prédites sont inférieures d'environ 0,8 psu aux salinités mesurées dans les stations Odh1 et Odh2. Le modèle TRIM3D doit vraisemblablement surestimer le transport pour avoir ainsi rempli le *Kleines Haff* d'eau douce avant même le début de la crue de l'Oder. Par contre l'ADS permet de recréer ce front et de le faire balayer la lagune d'est en ouest en deux semaines, conformément à ce qui avait été constaté sur les séries temporelles des mesures aux stations Odh1 et Odh2 dans Rosenthal *et al.* (1998).

Si on compare les distributions de salinité de l'EnKF à celles du modèle TRIM3D sans ADS, on voit dans les premiers temps (journée du 18 juillet 1997) la marque des deux stations de données Odh1 et Odh2, dont les salinités sont plus élevées aux alentours de 1.4 psu. L'assimilation de ces données crée deux panaches introduisant du sel dans le bassin ouest de la lagune de l'Oder, la forme de ces panaches correspondant aussi à la zone d'influence des deux mesures de salinité visible sur la carte d'écart-type d'estimation (au milieu à gauche de la figure 7.7). Après 6 jours d'ADS (le 23 juillet 1997) le panache a quasiment rempli toute la lagune de l'Oder, mais on voit encore la limite de la zone d'influence de l'ADS dans le chenal ouest (Peenestrom) où il reste quelques kilomètres d'eau douce.

Après 11 jours d'assimilation (cartes du 28 juillet puis des 2, 7 et 12 août 1997), la marque des positions des stations Odh1 et Odh2 a disparu et les estimations de salinité semblent plus réalistes. Les cartes fournies par TRIM3D et l'EnKF exhibent des structures similaires (mêmes positions des tourbillons, des poches de sel), seulement la salinité est globalement plus élevée avec que sans ADS. Cependant d'après la section précédente, les données de salinité des stations Odh1 et Odh2 ne sont assimilées efficacement que dans les premiers jours de l'expérience, ensuite les erreurs d'estimations sont atténuées par l'effet d'un attracteur du système et puisqu'elles sont bien inférieures à l'erreur de mesure, la correction est inopérante et l'ADS agit pratiquement comme le modèle seul. Ainsi les différences entre les cartes fournies par le modèle TRIM3D et l'ADS sont essentiellement dues à l'ADS de salinité durant les premiers jours de l'expérience.

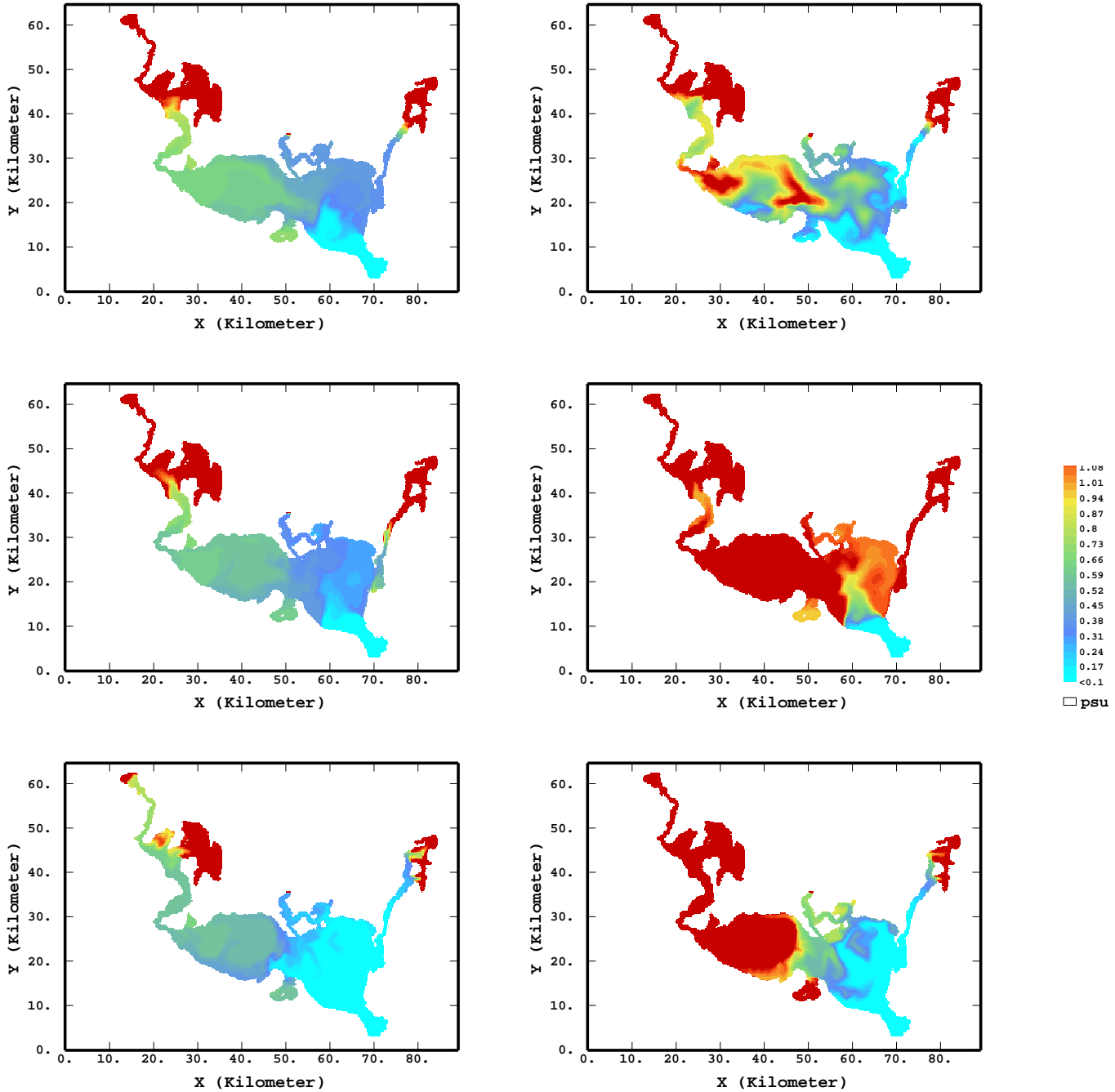


FIG. 7.10 – Salinités après 1 jour, 6 jours et 11 jours (18, 23 et 28 juillet 1997) avec le modèle TRIM3D sans ADS à gauche contre l'EnKF 100 membres à droite

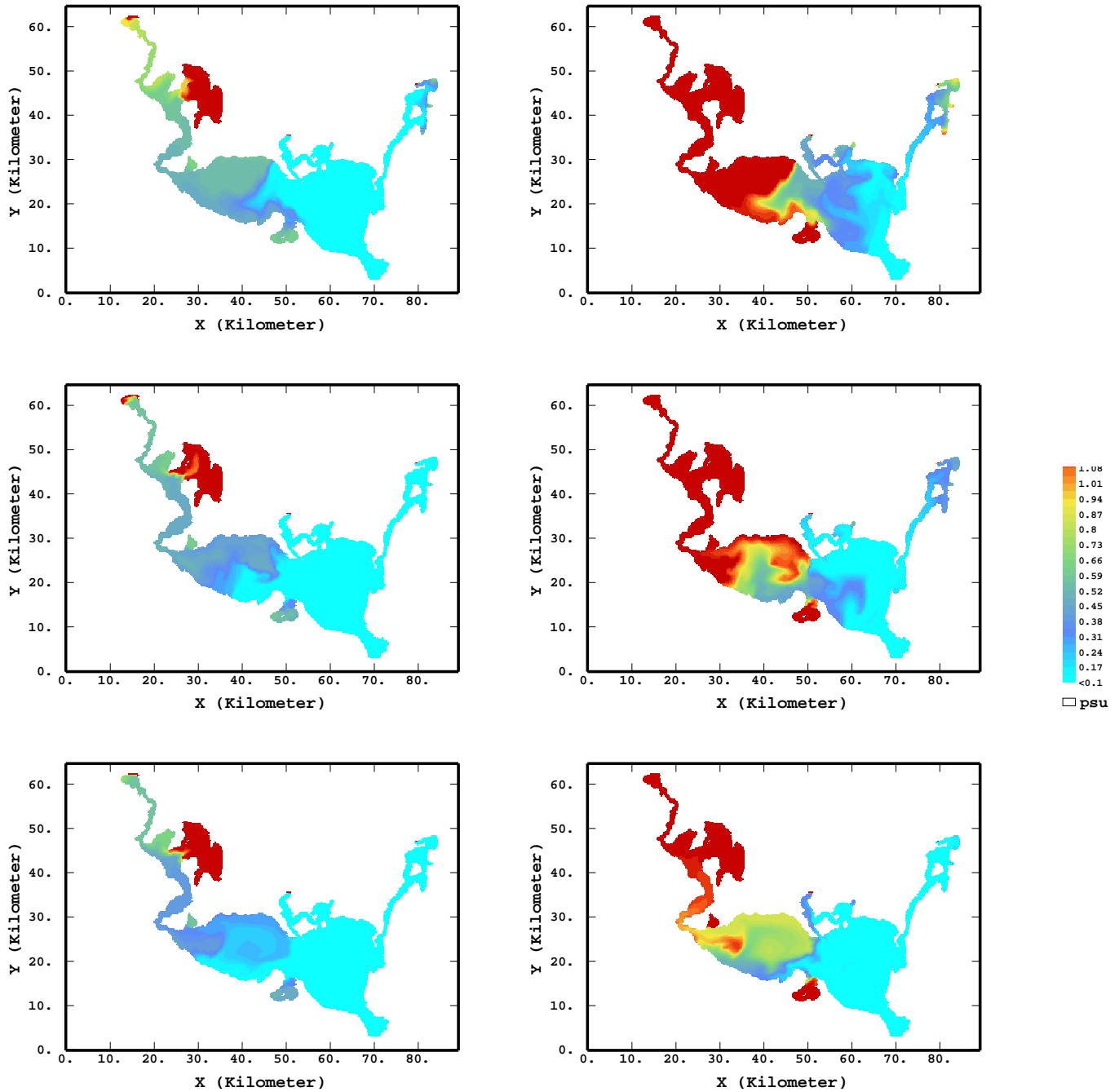


FIG. 7.11 – Salinités après 16, 21 et 26 jours (les 2, 7 et 12 août 1997) avec le modèle TRIM3D sans ADS à gauche contre l'EnKF 100 membres à droite

On peut aussi remarquer que la salinité est aussi corrigée dans le bassin est de la lagune, bien que cette partie soit, par rapport à l'écoulement dans la lagune, en amont des stations assimilées. L'ADS peut ainsi corriger les salinités en remontant le courant, mais quelle est la validité de ces corrections en amont ?

7.4.7 Validation avec des données indépendantes

On dispose des données de la campagne en bateau, les 22 et 23 juillet ainsi que le 5 août. Le «Ludwig Prandtl» a traversé le bassin est de la lagune au matin. Puisque les différences entre les résultats du modèle, de l'ADS et les mesures sont assez flagrantes, je me contenterai de montrer les cartes $(x - y)$ de ces premiers et de les comparer qualitativement aux sections $(x - z)$ interpolées soit par krigeage universel (dans le cas des données des 22 et 23 juillet car le front d'eau douce est présent dans les données) soit par krigeage ordinaire (dans le cas du 5 août où il n'y a plus que de l'eau douce). Il est évident en comparant l'état au 22 et au 23 juillet que les variations journalières sont assez lentes pour qu'on puisse comparer des simulations faites à minuit au données prises le matin. Pour plus de détails sur ces interpolations voir Bertino et Wackernagel (2000).

Notons que le bateau a traversé le bassin est alors que les deux mesures assimilées sont situées l'une à l'extrémité du bassin ouest (Odh1) et l'autre à la frontière entre les deux bassins (Odh2). La comparaison n'est donc pas des plus favorables à l'ADS. Dans les trois situations, figures 7.12 à 7.17, le modèle sans assimilation de données sous-estime les salinités observées lors de la campagne en bateau, certainement parce que le transport de sel dans les deux bassins de la lagune est surestimé.

L'ADS, au contraire, surestime les concentrations en sel tout au long du canal Świna les 22 et 23 juillet 1997 (figures 7.12 et 7.14). Le front d'eau douce est simulé par l'ADS, mais il est situé trop au sud. L'erreur faite par l'ADS (environ 0,3 psu de surestimation) est bien supérieure aux écarts-types prévus sur la carte 7.8 au milieu à gauche, ce qui signifie que l'optimisme affiché par le filtre de Kalman n'est pas approprié dans le cas présent.

Les résultats du 5 août sont encore plus marqués dans ce sens (figures 7.16 et 7.17). Ni le modèle TRIM3D guidé par ses conditions aux limites, ni l'ADS ne permettent de restituer le faible gradient de salinité (de 0.15 psu à 0.19 psu) observé lors de la campagne en bateau. De plus, le modèle sans ADS prédit un front d'eau douce beaucoup plus net et déplacé vers l'ouest. L'ADS replace ce front d'eau douce plus à l'est, mais celui-ci reste toujours trop net et décrit une recirculation d'eau salée non observée lors de la campagne en bateau.

L'écart de salinité avec les observations du bateau est réduit par l'utilisation de l'ADS, mais il reste considérablement plus important que l'erreur estimée par le filtre de Kalman. Nous dirons donc que le filtre de Kalman a été détourné vers un attracteur qui ne correspond pas aux observations de la salinité de la lagune de l'Oder. Notons que cet aspect de l'ADS n'est pas propre à l'EnKF mais à sa mise en œuvre. L'emploi du

filtre de Kalman RRSQRT ou de tout autre filtre de Kalman aurait entraîné la même situation.

7.5 Conclusion

Ce cas d'étude présente un contre-exemple où l'ADS perd son efficacité en quelques jours d'intégration, et se comporte peu à peu comme le modèle dynamique sans tenir compte des observations.

Si des améliorations peuvent être apportées à la mise en œuvre de l'ADS (introduction d'une erreur initiale et d'une erreur de modèle plus appropriées) et au réseau d'observations (deux mesures ne sont pas suffisantes pour couvrir efficacement la distribution de salinité de la lagune entière), l'ADS paye en premier lieu les imprécisions du modèle physique et son aspect «lissant». Une recalibration de TRIM3D serait donc nécessaire pour améliorer les résultats de l'ADS, lesquels apparaissent donc comme un outil d'analyse des atouts et des défauts du modèle dynamique.

On aurait pu traiter entièrement la variable salinité dans un modèle gaussien anamorphosé (par exemple lognormal puisque la salinité est une variable positive), mais l'ADS aurait sans doute présenté les mêmes défauts du fait de la calibration.

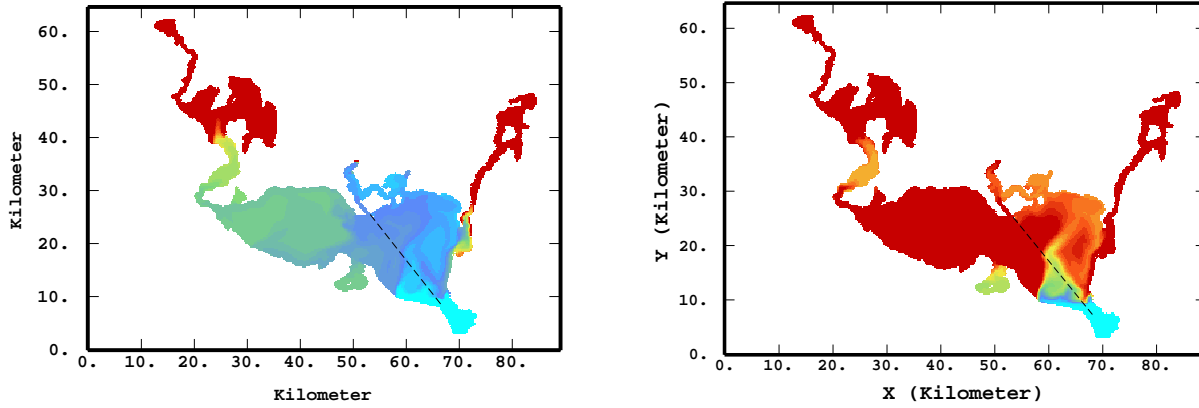


FIG. 7.12 – Distributions de sel simulées au 22 juillet 1997 (5-ième jour) à gauche = TRIM3D sans ADS, à droite = TRIM3D avec EnKF, la trace du bateau est symbolisée par la ligne discontinue, même échelle de couleurs que dans la figure 7.13.

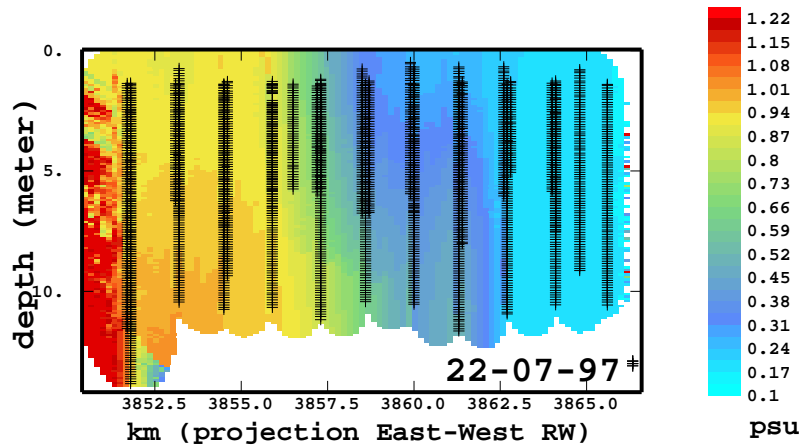


FIG. 7.13 – Salinité (section $x - z$) mesurée pendant le transept du 22 juillet (SYNOPTA2) et interpolée par krigeage avec une dérive linéaire orientée à 15 degrés de la verticale. Les valeurs fortes à l'ouest sont un *artefact* d'extrapolation.

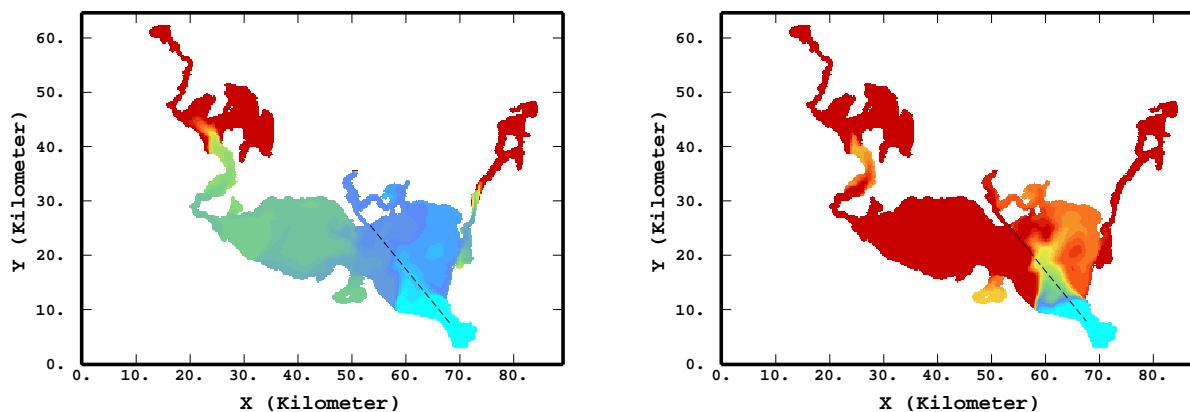


FIG. 7.14 – Distributions de sel simulées au 23 juillet 1997 (6-ième jour) à gauche = TRIM3D sans ADS, à droite = TRIM3D avec EnKF

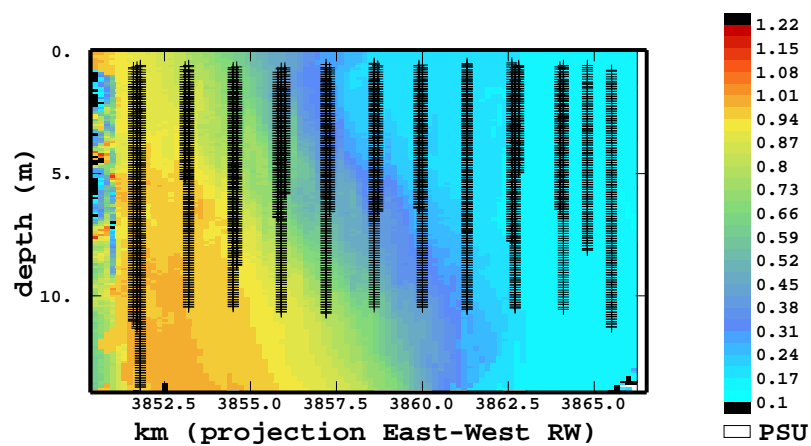


FIG. 7.15 – Salinité (section $x - z$) mesurée pendant le transect du 23 juillet (SYNOPTA2) et interpolée par krigeage avec une dérive linéaire orientée à 30 degrés de la verticale.

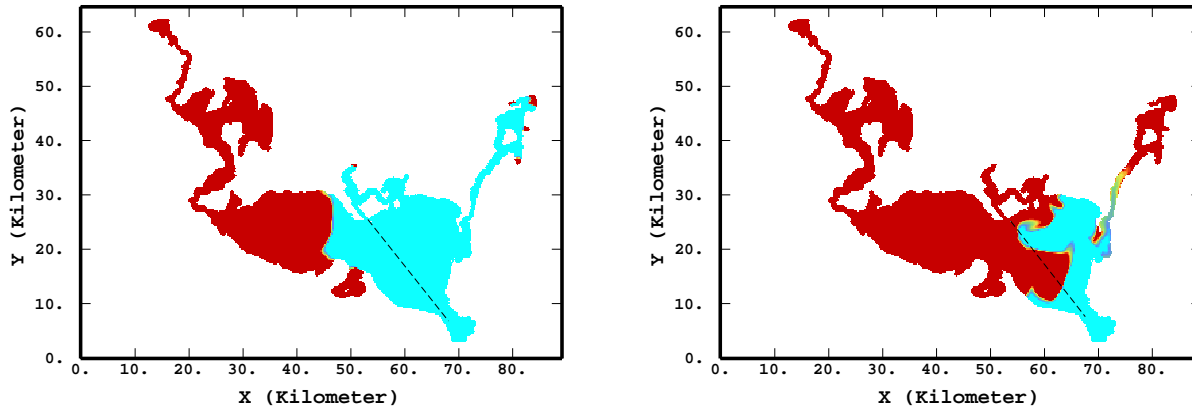


FIG. 7.16 – Distributions de sel simulées au 5 août 1997 (19-ième jour) à gauche = TRIM3D sans ADS, à droite = TRIM3D avec EnKF. Attention ! L'échelle n'est plus celle des figures 7.10 à 7.15, mais celle de la figure 7.17.

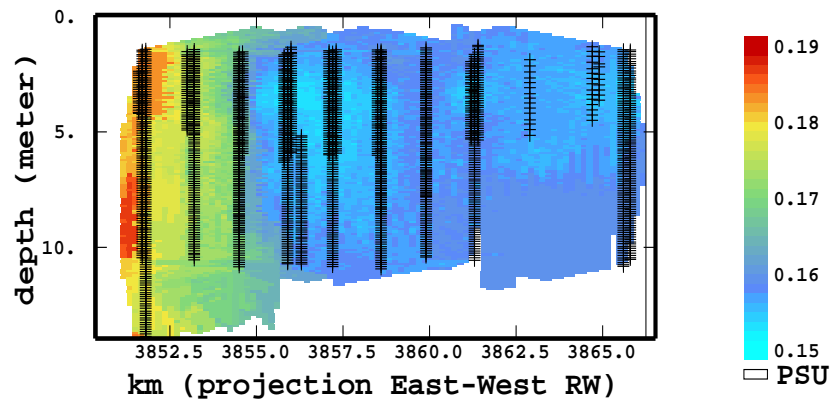


FIG. 7.17 – Salinité (section $x - z$) mesurée pendant le transept du 5 août (SYNOPTA2) et interpolée par krigeage ordinaire.

Conclusions

L'AD permet d'intégrer un modèle numérique et des données de terrain dans un même cadre probabiliste. Cette thèse montre d'abord les difficultés que ce problème représente dans le cas général lorsqu'on considère un problème défini sur un système régionalisé tel qu'en océanographie (chap. 2). Dans le cas linéaire (chap. 3) le filtre de Kalman donne une solution gaussienne. En raison de son équivalence avec le krigeage, on lui ajoute des extensions inspirées des méthodes classiques de géostatistique linéaire (KO, KDE ou KU). Dans le chapitre 4, on décrit les méthodes qui ont adapté le filtre de Kalman aux dynamiques non linéaires soit par développement limité au premier ordre (*Extended Kalman filter*, EKF et ses schémas sous-optimaux dont on a retenu le *Reduced Rank square root* RRSQRT) soit par une approximation de type Monte Carlo (*Ensemble Kalman filter* EnKF), ce qui est plus approprié pour les dynamiques non linéaires. On compare les deux dans le cas fictif d'un modèle écologique 1-D (chap. 5), et leurs performances sont quasiment équivalentes. Par contre les deux approches appliquent des corrections linéaires tout comme le filtre de Kalman et ne donnent pas la solution optimale. On propose alors un modèle d'anamorphose gaussienne, appliqué dans le même cas fictif (chap. 5).

Le chapitre 6 montre une application du RRSQRT au modèle hydrodynamique tridimensionnel TRIM3D, assimilant des données réelles de hauteurs d'eau de la lagune de l'Oder. L'erreur du modèle est imputée à la méconnaissance des champs de vents et des conditions aux limites avec la Baltique ; la structure spatiale de cette dernière erreur, inaccessible par les données disponibles, est déduite de la structure temporelle des séries de hauteurs d'eau mesurées *via* une équation simple de propagation d'onde. Les résultats du modèle TRIM3D sont améliorés sur toute l'étendue de la lagune par l'assimilation des données de seulement trois stations côtières ; on notera aussi la forte réduction du temps d'initialisation. L'ADS donne de meilleurs résultats qu'une interpolation simple des données, ce qui justifie l'emploi d'un modèle hydrodynamique par rapport à un modèle uniquement statistique. On montre aussi l'intérêt du filtre de Kalman RRSQRT en tant qu'outil de prédiction *a priori* utilisable opérationnellement. Dans ces deux premiers cas d'étude, les filtres de Kalman RRSQRT et EnKF surmontent donc les deux difficultés que posent les systèmes estuariens et océanographiques en général, les grandes dimensions de l'espace d'état et la non linéarité des dynamiques du système.

Dans le chapitre 7, l'application de l'EnKF au cas de la salinité de la lagune de l'Oder, mesurée en

deux stations, se heurte aux imperfections du modèle physique. Celui-ci lisse la salinité et entraîne le filtre de Kalman vers un attracteur qui le rend inopérant, ce qui rappelle que l'AD n'est pas une alternative à la calibration des modèles. On peut améliorer l'efficacité de l'ADS en perfectionnant les modèles d'erreurs initiales et d'erreurs de modèle, mais surtout en recalibrant des paramètres du modèle TRIM3D.

Perspectives

Le point de vue de la géostatistique offre de nombreuses perspectives théoriques à l'ADS. D'une part le traitement des problèmes de biais qui nécessite des solutions particulières à chaque problème (chapitre 3). En effet une longue expérience a été accumulée dans les domaines traités par la géostatistique pour les problèmes de biais et de non stationnarité en général. La question de savoir estimer des paramètres de biais corrélés dans le temps tout en évitant d'augmenter l'état semble toutefois assez épineuse. D'autre part on peut remarquer que le traitement du cas non linéaire n'en est qu'à ses débuts en ADS. Mis à part van Leeuwen (2001), personne n'a abordé l'équations de Bayes en dehors du cadre gaussien pour applications aux cas réels en océanographie. L'utilisation d'un algorithme séquentiel basé sur un échantillonnage d'importance et un rééchantillonnage SIR (Doucet, de Freitas, et Gordon, 2001) semble être une piste d'avenir, supérieure à l'idée de l'anamorphose gaussienne puisqu'aucune hypothèse sur les lois des variables n'est faite. On peut aussi compléter l'approche en SIR par d'autres algorithmes de simulations conditionnelles utilisés couramment en géostatistique, en particulier l'échantillonneur de Gibbs. La modélisation de l'effet de support dans le cas non stationnaire est une autre question ouverte de la théorie de l'ADS. On peut aussi greffer sur les sorties de l'ADS d'autres extensions inspirées de la géostatistique non linéaire telles que le traitement de variables de risque et le tracé de cartes de probabilités de dépassement de seuils (utiles pour quantifier les risques d'exposition d'une population). De plus les idées et les résultats de cette thèse ont été appliqués à des problèmes de filtrage, mais la plupart d'entre eux sont facilement transposables au lissage qui offre un cadre plus général.

Les améliorations apportées par l'ADS à la simulation de l'hydrodynamique et du transport de sel et de nutriments dans la lagune de l'Oder permettront d'améliorer l'estimation des bilans de matières entre les différents compartiments du modèle écologique ERSEM appliqué à la lagune de l'Oder.

Enfin les méthodes d'ADS telles que l'EnKF ou le filtre de Kalman RRSQRT ont fait leurs preuves dans des applications océanographiques physiques, chimiques et biologiques pour lesquelles il existe un besoin pressant de prédictions opérationnelles et d'analyse de processus. Cependant on peut remarquer que ces deux méthodes utilisent le modèle comme une boîte noire et que la problématique résolue a une portée plus générale. Tous les systèmes régionalisés et obéissant à des dynamiques non linéaires sont susceptibles de trouver dans l'ADS une méthodologie adaptée.

Bibliographie

Références

- Athias, V., P. Mazzega, et C. Jeandel (2001). Estimating the aluminium cycling parameters in the tropical NE Atlantic (oligotrophic EUMELI site) using a genetic algorithm. *Journal of Marine Research*. En révision.
- Bennett, A. F. (1992). *Inverse methods in physical oceanography*, Cambridge : Cambridge University Press.
- Bennett, A. F., B. S. Chua, et L. M. Leslie (1996). Generalized inversion of a global numerical weather prediction model. *Meteorol. Atmos. Phys.* 60, 165–178.
- Bertino, L., G. Evensen, et H. Wackernagel (2002). Combining geostatistics and Kalman filtering for data assimilation in an estuarine system. *Inverse Problems* 18, 1–23.
- Bertino, L. et H. Wackernagel (2000). The oderhaff salinity data : their use for data assimilation. Rapport technique N-9/00/G, ENSMP, Centre de Géostatistique, Fontainebleau.
- Bleck, R. (1978). Simulation of coastal upwelling frontogenesis with an isopycnic coordinate model. *Journal of Geophysical Research* 83(C12), 6163–6172.
- Blumberg, A. et G. Mellor (1987). A description of a three-dimensional coastal ocean circulation model. *Three-dimensional Coastal ocean models*, 1–16.
- Brasseur, P., J. Ballabrera-Poy, et J. Verron (1999). Assimilation of altimetric data in the mid-latitude oceans using the Singular Evolutive Extended Kalman filter with an eddy-resolving, primitive equation model. *Journal of Marine Systems* 22, 269–294.
- Brasseur, P., V. Carmillet, J. Brankart, et J. Verron (2001). A singular evolutive extended Kalman filter to assimilate ocean colour data in a coupled physical-ibiochemical model of the North Atlantic. J. Nihoul (Ed.), *International Liège Colloquium on Ocean Hydrodynamics*.
- Breuilin, C., G. Evensen, et M. Eknes (2000). Data assimilation methods applied to marine ecosystem models. Rapport de stage d'option 187, ENSMP, Centre de Géostatistique, Fontainebleau. 64p.
- Bui, V. A. P. (2001). Modélisation par chaînes de markov cachées d'une épidémie de grippe. Rapport de stage d'option, ENSMP, Centre de Géostatistique, Fontainebleau. 35p.
- Burgers, G., P. J. van Leeuwen, et G. Evensen (1998). On the analysis scheme in the ensemble Kalman filter. *Monthly Weather Review* (126), 1719–1724.

- Cañizares, R. (1999). *On the Application of Data Assimilation in Regional Coastal Models*. Thèse, TU Delft, Pays-Bas, Rotterdam : Balkema.
- Cañizares, R., D. Madsen, H. Jensen, et H. J. Vested (2001). Developments in operational shelf sea modelling in danish waters. *Estuarine and coastal shelf science* 53(4), 595–605.
- Cane, M., A. Kaplan, R. N. Miller, B. Tang, E. Hackert, et A. Busalacchi (1996). Mapping tropical Pacific sea level : data assimilation via a reduced state space Kalman filter. *Journal of Geophysical Research* 101(C10), 22599–22617.
- Carmillet, V., J.-M. Brankart, P. Brasseur, H. Drange, et G. Evensen (2001). A singular evolutive extended kalman filter to assimilate ocean color data in a coupled physical-biochemical model of the north atlantic. *Ocean Modelling*. A paraître.
- Casulli, V. et E. Cattani (1994). Stability, accuracy and efficiency of a semi-implicit method for three-dimensional shallow water flow. *Computers Math. Applic.* 27(4), 99–112.
- Chauvet, P. (2000). *Aide-mémoire de géostatistique linéaire*, Paris : Presses de l’Ecole des Mines.
- Chilès, J. P. et P. Delfiner (1999). *Geostatistics : Modeling Spatial Uncertainty*, New York : Wiley. Probability and statistics.
- Cohn, S. et R. Todling (1996). Appropriate data assimilation schemes for stable and unstable dynamics. *Journal of the Meteorological Society of Japan* 74(1), 63–75.
- Daley, R. (1991). *Atmospheric Data Analysis*, Cambridge : Cambridge University Press.
- d’Andréa Novel, B. et M. C. de Lara (1994). Modélisation analyse simulation commande. *Commande linéaire des systèmes dynamiques*. Paris : Masson.
- Dee, D. P. (1991). Simplification of the Kalman filter for meteorological data assimilation. *Q. J. R. Meteorol. Soc.* (117), 365–384.
- Dee, D. P. et A. M. da Silva (1998). Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.* (124), 269–295.
- del Moral, P. et G. Salut (1995). Filtrage non-linéaire : résolution particulière à la Monte-Carlo. *CRAS320 Série I*, 1147–1152.
- Doucet, A., N. de Freitas, et N. Gordon (Eds.) (2001). *Sequential Monte Carlo methods in practice*, New York : Springer. Statistics for engineering and information science.
- Duflo, M. (1997). *Random Iterative methods*, Berlin : Springer. Application of Mathematics 34.
- Echevin, V., P. De Mey, et G. Evensen (2000). Horizontal and vertical structure of the representer sections for sea level measurements in a coastal model. *Journal of Physical Oceanography* 30(2), 2627–2635.
- Eknes, M. et G. Evensen (2001). An ensemble Kalman filter with a 1-D marine ecosystem model. *Journal of Marine Systems*. Accepté.
- Evans, G. et J. Parslow (1985). A model for annual plankton cycles. *Biol. Oceanogr.* 3, 327–347.

- Evensen, G. (1992). Using the extended Kalman filter with a multilayer quasi-geostrophic ocean model. *Journal of Geophysical Research* 97(C11), 17905–17924.
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *Journal of Geophysical Research* 99(C5), 10143–10162.
- Evensen, G. et P. J. van Leeuwen (1996). Assimilation of Geosat altimeter data for the agulhas current using the ensemble Kalman filter with a quasi-geostrophic model. *Monthly Weather Review* 124, 85–96.
- Evensen, G. et P. J. van Leeuwen (2000). An ensemble Kalman smoother for nonlinear dynamics. *Monthly Weather Review* 128, 1852–1867.
- Evensen, G., J. Verron, P. J. van Leeuwen, et DIADEM (2001). Results from the DIADEM data assimilation experiment with an OGCM for the North Atlantic. J. Nihoul (Ed.), *International Liège Colloquium on Ocean Hydrodynamics*.
- Fenske, C. (1999). How has the GOAP project changed our understanding of the Odra estuary. I. Morski (Ed.), *5th international scientific conference : hydrodynamic and ecological aspects of nutrient forecasting for Odra and Ebro estuaries*, Szczecin, Pologne, p. 23–26.
- Fitzgerald, R. J. (1971). Divergence of the Kalman filter. *IEEE transactions on automatic control AC-16*(6), 736–747.
- Fukumori, I. (1995). Assimilation of Topex sea level measurements with a reduced-gravity, shallow water model of the tropical pacific ocean. *Journal of Geophysical Research* 100(C12), 25027–39.
- Fukumori, I. et P. Malanotte-Rizzoli (1995). An appropriate Kalman filter for ocean data assimilation : an example with an idealized Gulf Stream model. *Journal of Geophysical Research* 100(C4), 6777–6793.
- Gauthier, P., P. Courtier, et P. Moll (1993). Assimilation of simulated wind lidar data with a Kalman filter. *Monthly Weather Review* 121, 1803–1820.
- Grønnevik, R. et G. Evensen (2001). Application of ensemble-based techniques in fish stock assessment. *Sarsia*. A paraître.
- Guillot, G., R. Senoussi, et P. Monestiez (2001). A positive definite estimator of the nonstationary covariance of random fields. Monestier, Allard, et Froidevaux (Eds.), *GeoENV III : Geostatistics for Environmental Applications*, Amsterdam, p. 333–344. Kluwer Academic.
- Haugen, V. E. J. et G. Evensen (2001). Assimilation of SST and SLA data into an OGCM for the Indian Ocean. *Ocean Modelling*. Soumis.
- Hoteit, I., D. T. Pham, et J. Blum (2001). A semi-evolutive partially local filter for data assimilation. *Marine Pollution Bulletin* 43((7-12)), 164–174.
- Hu, L. Y. (1998). *Mise en oeuvre du modèle gamma pour l'estimation des distributions spatiales*. Thèse de doctorat en géostatistique, ENSMP, Centre de Géostatistique, Fontainebleau.

- Jazwinski, A. A. (1970). *Stochastic and Filtering Theory*, New York : Academic Press. Mathematics in Sciences and Engineering 64.
- Julier, S. J., J. K. Uhlmann, et H. F. Durrant-Whyte (1995). A new approach for filtering nonlinear systems. *The 1995 American Control Conference, Seattle, Washington*, p. 1628–1632.
- Kyriakidis, P. C. et A. G. Journel (1999). Geostatistical space-time models : a review. *Mathematical geology* 31(6), 651–684.
- Lajaunie, C. (1997). Kriging with mass balance. E. Baafi et N. Schofield (Eds.), *Geostatistics Wollongong '96*, Volume 1, Dordrecht, p. 80–91. Kluwer.
- Lantuejoul, C. (2001). *Geostatistical Simulation : Models and Algorithms*, Berlin : Springer-Verlag.
- Le Dimet, F. X. et O. Talagrand (1986). Variational algorithms for analysis and assimilation of meteorological observations : theoretical aspects. *Tellus* (38A), 97–110.
- Lellouche, J.-M., J.-L. Devenon, et I. Dekeyser (1998). Data assimilation by optimal control method in a 3d coastal oceanic model : The problem of discretization. *Journal of Atmospheric and Oceanic Technology* 15(2), 470–481.
- Leredde, Y., I. Dekeyser, et J.-L. Devenon (2000). T-S data assimilation to optimize turbulent viscosity. an application to the Berre lagoon hydrodynamics. *Journal of Coastal Research*. Soumis.
- Lermusiaux, P. F. J. et A. R. Robinson (1999). Data assimilation via error subspace statistical estimation, part i : Theory and schemes. *Monthly Weather Review* 127(8), 1385–1407.
- Liabaud, R. (2001). Comportement des filtres d'ensemble sur des systèmes fortement non-linéaires en assimilation de données météorologiques. Rapport de stage d'option, LMD, Ecole Normale Supérieure, 24 rue Lhomond, Paris.
- Luong, B., J. Blum, et J. Verron (1998). A variational method for the resolution of a data assimilation problem in oceanography. *Inverse Problems* 14, 979–997.
- Mardia, K., C. Goodall, E. Redfern, et F. Alonso (1998). The kriged Kalman filter. *Test* 7(2), 217–285.
- Matheron, G. (1970). *La théorie des variables régionalisées et ses applications*, Fontainebleau : ENSMP, cahiers du centre de morphologie mathématique, Fasc. 5. 210 p.
- Maybeck, P. (1979). *Stochastic models, estimation, and control*, New York : Academic Press. Mathematics in Science and Engineering 141-1.
- Miller, R. N., E. F. Carter, Jr., et S. T. Blue (1999). Data assimilation into nonlinear stochastic models. *Tellus* 51A, 167–194.
- Miller, R. N., M. Ghil, et F. Gauthiez (1994). Advanced data assimilation in strongly nonlinear dynamical systems. *Journal of Atmos. Sci.* 51, 1037–1056.
- Natvik, Lars, J. et G. Evensen (2001). Assimilation of ocean colour data into a biochemical model of the North Atlantic. Part 1. Data assimilation experiments. *Ocean Modelling*. Submitted.

- Nechaev, V. et M. Yaremchuk (1994). Applications of the adjoint technique to processing of a standard section data set : world ocean circulation experiment section S4 along 67 deg S in the Pacific ocean. *Journal of Geophysical Research* 100(C1), 875–79.
- Pham, D., J. Verron, et M. Roubaud (1997). A singular evolutive extended Kalman filter for data assimilation in oceanography. *Journal of Marine Systems* 16(3-4), 323–340.
- Pham, D. T. (2000). Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Monthly Weather Review* 129(5), 1194–1207.
- Rivoirard, J. (1990). A review of lognormal estimators for in situ reserves. *Mathematical Geology* 22, 213–221.
- Robinson, A. et le groupe AFMIS (2001). The development and demonstration of an advanced fisheries management information system. *Preprint Volume of the Seventeenth International Conference on Interactive Information and Processing Systems (IIPS), Albuquerque, New Mexico*, Boston. Americal Meteorological Society.
- Robinson, A. R., H. Arango, A. Miller, A. Warms-Varnas, P. Poulain, et W. Leslie (1996). Real-time operational forecasting on shipboard of the iceland-faeroe frontal variability. *Bulletin of the American Meteorological Society* 77(2), 243–259.
- Robinson, A. R., P. F. J. Lermusiaux, et N. Q. Sloan III (1998). Data assimilation. A. R. K.H. Brink (Ed.), *The Global Coastal Ocean. Processes and Methods. The Sea*, Volume 10, p. 541–593. New York : John Wiley & Sons, Inc.
- Rosenthal, W., T. Wolf, G. Witte, W. Buchholz, et P. Rybczok (1998). Measured and modelled water transport in the Odra estuary for the flood period july/august 1997. *German Journal of Hydrography* 50(2-3), 215–230.
- Roth, C. (1998). Is lognormal kriging suitable for local estimation ? *Mathematical Geology* 30, 999–1009.
- Sampson, P. et P. Guttorp (1992). Nonparametric estimation of nonstationary spatial covariance structure. *Journal of the American Statistical Association* 87(417), 108–119.
- Segers, A., A. Heemink, M. Verlaan, et M. van Loon (2000). A modified rrsqrt-filter for assimilating data in atmospheric chemistry models. *Environmental Modelling and Software* 15(6-7), 663–671.
- Sénégas, J. (1999). Hydrodynamical modeling and data assimilation within the Odra estuary. Rapport externe 99/E/42, GKSS, D-21 494 Geesthacht, Allemagne.
- Sénégas, J., H. Wackernagel, W. Rosenthal, et T. Wolf (2001). Error covariance modeling in sequential data assimilation. *Stochastic Environmental Research and Risk Assessment* 15(1), 65–86.
- Stålnacke, P. (1996). *Nutrient loads to the Baltic sea*. Linköping studies for art and science 146.
- Van Geer, F. et A. Zuur (1997). An extension of the Box-Jenkins transfer-noise model for spatial interpolation of groundwater head series. *Journal of Hydrology* 192, 65–80.
- van Leeuwen, P. J. (2001). A truly variance minimizing ensemble filter and smoother. *Nonlinear Proc. Geophys.*. Soumis.

- van Loon, M. et A. Heemink (1997). Kalman filtering for nonlinear atmospheric chemistry models : first experiences. Technical report MAS-R9711, CWI, TU Delft, Pays-Bas.
- Verlaan, M. (1998). *Efficient Kalman filtering algorithms for hydrodynamic models*. Thèse, TU Delft, Pays-Bas.
- Verlaan, M. et A. W. Heemink (1997). Tidal flow forecasting using reduced rank square root filters. *Stochastic Hydrology and Hydraulics* 11(5), 349–368.
- Verlaan, M. et A. W. Heemink (1999). Non-linearity in data assimilation applications : a practical method for analysis. *Monthly Weather Review*. Soumis.
- Verron, J., L. Gourdeau, D. Pham, R. Murtugudde, et A. Busalacchi (1998). 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*. Soumis.
- Vénard, C., A. Aulinger, C. Lajaunie, et H. Wackernagel (2001). Geostatistical evaluation of a monitoring strategy for water quality assessment. Monestier, Allard, et Froidevaux (Eds.), *GeoENV III : Geostatistics for Environmental Applications*, Amsterdam, p. 149–156. Kluwer Academic.
- von Storch, H. (2000). The PIONEER project as an example of operational coastal analysis. *4th Symposium on Integrated Observing Systems*, Boston, p. 215–220. American Meteorological Society.
- Wackernagel, H. (1998). *Multivariate Geostatistics*, Berlin : Springer-Verlag.
- Wikle, C. K. et N. Cressie (1999). A dimension-reduced approach to space-time Kalman filtering. *Biometrika* 86(4), 815–829.
- Wolf, T., J. Sénégas, L. Bertino, et H. Wackernagel (2001). Application of data assimilation to three-dimensional hydrodynamics : the case of the Odra lagoon. Monestier, Allard, et Froidevaux (Eds.), *GeoENV III : Geostatistics for Environmental Applications*, Amsterdam, p. 157–168. Kluwer Academic.

Summary :

The management of ocean and coastal systems requires short term predictions of their physical and ecological state. Oceanographical data assimilation improves the results of numerical models and has been recently applied to operational forecasting, in spite of difficulties both theoretical (the dynamics are non linear) and practical (the regionalized systems represent more than a hundred thousand parameters).

This work focuses on sequential data assimilation (i.e. the Kalman filter and its extensions to non linear dynamics), in order to introduce dynamical equations into spatio-temporal statistical analysis. A thorough analysis of different aspects of the Kalman filter from a geostatistical point of view is proposed ; in particular the equivalence between the model update step of the Kalman filter and simple kriging led us to explore several modifications of the classical formulation. They were derived from linear geostatistics for on-line bias estimation (kriging with unknown mean and external drift) and from nonlinear geostatistics to account for deviations from the Gaussian framework.

Three case studies are presented, first a twin experiment using a 1-D ecological model of an ocean water column, that helps us in comparing two widely used Kalman filtering methods : the RRSQRT KF (based on a first order expansion of the model followed by an approximation on dominant eigenmodes) and the EnKF (based on Monte Carlo sampling). Secondly, water levels measured at three pile stations in the Odra lagoon are assimilated in the hydrodynamical model TRIM3D by the RRSQRT KF. Lastly, the EnKF is applied to joint assimilation of water levels and salinity of the Odra lagoon in TRIM3D.

Keywords :

Data assimilation, Kalman filter, geostatistics, estuary, hydrodynamics, ecological modeling.

Résumé :

La gestion des milieux océaniques et côtiers nécessite des prédictions à court terme de leurs caractéristiques physiques et écologiques. L'assimilation de données océanographiques améliore les résultats des modèles numériques et l'océanographie opérationnelle fait ainsi ses débuts, malgré des obstacles théoriques (les dynamiques sont non linéaires) et pratiques (les systèmes régionalisés étudiés comportent plus de cent mille paramètres).

Ce travail se focalise sur les méthodes séquentielles d'assimilation de données (le filtre de Kalman et ses approximations aux dynamiques non linéaires) qui introduisent les équations dynamiques du système dans son analyse statistique spatio-temporelle. Le filtre de Kalman est analysé sous ses aspects géostatistiques ; en particulier l'équivalence de son étape de correction à un krigeage simple nous a guidé vers plusieurs modifications de sa formulation classique. Nous nous sommes donc inspirés des estimateurs utilisés en géostatistique pour tenir compte des biais (krigeages à moyenne inconnue et avec dérive externe) et des non linéarités (anamorphose gaussienne).

Trois cas d'études sont présentés, d'abord un cas fictif de modèle écologique 1-D d'une colonne d'eau océanique qui nous permet de comparer deux filtres de Kalman courants en océanographie : le RRSQRT (basé sur un développement limité des équations du modèle et une approximation en valeurs propres dominantes) et l'EnKF (basé sur un échantillonnage de Monte Carlo). On assimile ensuite les hauteurs d'eau prises en trois stations fixes de la lagune de l'Oder dans le modèle hydrodynamique TRIM3D avec le filtre de Kalman RRSQRT. Enfin l'EnKF est appliqué à l'assimilation de données de hauteurs d'eau et de salinité de la lagune de l'Oder dans TRIM3D.

Mots clefs :

Assimilation de données, géostatistique, filtre de Kalman, estuaire, hydrodynamique, modèle écologique.