



Assimilation de données ensembliste et couplage de modèles hydrauliques 1D-2D pour la prévision des crues en temps réel. Application au réseau hydraulique "Adour maritime"

Sébastien Barthélémy

► **To cite this version:**

Sébastien Barthélémy. Assimilation de données ensembliste et couplage de modèles hydrauliques 1D-2D pour la prévision des crues en temps réel. Application au réseau hydraulique "Adour maritime". Hydrologie. Institut National Polytechnique de Toulouse - INPT, 2015. Français. NNT : 2015INPT0047 . tel-04238349

HAL Id: tel-04238349

<https://theses.hal.science/tel-04238349>

Submitted on 12 Oct 2023

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.



Université
de Toulouse

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par :

Institut National Polytechnique de Toulouse (INP Toulouse)

Discipline ou spécialité :

Hydrologie, Hydrochimie, Sols, Environnement

Présentée et soutenue par :

M. SÉBASTIEN BARTHÉLÉMY

le mardi 12 mai 2015

Titre :

ASSIMILATION DE DONNEES ENSEMBLISTE ET COUPLAGE DE
MODELES HYDRAULIQUES 1D-2D POUR LA PREVISION DES CRUES
EN TEMPS REEL. APPLICATION AU RESEAU HYDRAULIQUE "ADOUR
MARITIME".

Ecole doctorale :

Sciences de l'Univers de l'Environnement et de l'Espace (SDUEE)

Unité de recherche :

CERFACS

Directeur(s) de Thèse :

M. OLIVIER THUAL

MME SOPHIE RICCI

Rapporteurs :

M. ERIC BLAYO, UNIVERSITE GRENOBLE 1

M. JACQUES SAINTE-MARIE, CEREMA

M. PHILIPPE SERGENT, CEREMA

Membre(s) du jury :

M. ERIC BLAYO, UNIVERSITE GRENOBLE 1, Président

Mme SOPHIE RICCI, CERFACS, Membre

M. OLIVIER THUAL, INP TOULOUSE, Membre

M. SEBASTIEN BOURBAN, HR WALLINGFORD, Membre

*A toute ma famille,
et plus particulièrement à ma grand-mère Alice Renaud*

« Pourtant, en repensant maintenant à ces trois années, je me rends compte qu'elles n'étaient nullement gaspillées. Sans même le savoir, j'ai appris alors dans la solitude ce qui fait l'essentiel du métier [...] - ce qu'aucun maître ne peut véritablement enseigner. »

Alexandre Grothendieck - *Récoltes et Semailles* - 2.2. L'importance d'être seul

« Et toute science, quand nous l'entendons non comme un instrument de pouvoir et de domination, mais comme aventure de connaissance de notre espèce à travers les âges, n'est autre chose que cette harmonie, plus ou moins vaste et plus ou moins riche d'une époque à l'autre, qui se déploie au cours des générations et des siècles, par le délicat contrepoint de tous les thèmes apparus tour à tour, comme appelés du néant, pour se joindre en elle et s'y entrelacer. »

Alexandre Grothendieck - *Récoltes et Semailles* - 2.8. La vision - ou douze thèmes pour une harmonie

Remerciements

Il convient en premier lieu de remercier mes directeurs de thèse, *Sophie Ricci* et *Olivier Thual*, qui m’ont permis d’effectuer ce travail avec l’autonomie qui correspond à mon caractère. Ils m’ont introduit dans une communauté dans laquelle je me sens à l’aise et ont toujours su se mobiliser rapidement lorsque je les ai sollicités.

En second lieu je tiens à remercier ceux sans qui rien de tout cela n’aurait été possible, à savoir les financeurs de ce travail : la région Midi-Pyrénées et le SCHAPI.

Je tiens à remercier également *Etienne Le Pape* du SCHAPI et *Laurent Diéval* du SPC GAD pour leur support et leur collaboration tout au long de ce travail mais aussi pour l’intérêt qu’ils ont manifesté pour celui-ci.

Je tiens à remercier *Nicole Goutal* (LNHE, EDF R&D) et *Olivier Pannekoucke* (CNRM, Météo-France) dont les apports scientifiques dans cette thèse ont été importants. Il est clair que sans eux cette thèse ne serait pas ce qu’elle est aujourd’hui. Ils se sont toujours montrés attentifs à mes questions, mêmes les plus idiotes, et ont toujours su trouver du temps à me consacrer lorsque j’ai eu besoin de leurs lumières.

Je tiens à remercier l’équipe des « palmistes », Thierry Morel et Florent Duchaine. En particulier Thierry qui n’a jamais rechigné à consacrer du temps pour déboguer certains de mes codes et pour corriger mes maquettes OpenPALM.

Je tiens à remercier également les personnels administratifs du CERFACS qui par leur travail nous permettent de nous consacrer pleinement au notre. Plus particulièrement je tenais à remercier les deux secrétaires de l’équipe GLOBC : *Isabelle Moity* et *Nathalie Brousset* ainsi que *Chantal Nasri* (qui malgré ses menaces et ses pressions ne s’est pas retrouvée à la tête de ces remerciements !) et *Michèle Campassens*.

Je tiens à remercier tous les membres de l’équipe GLOBC qui m’ont entouré au quotidien pendant ces trois années et qui font de cette équipe une équipe dans laquelle il est agréable de travailler. Plus particulièrement je tiens à remercier : *Olivier Titaud*, *Nabil El Moçayd*, *Yohan Ruprich-Robert*, *Paul-Arthur Monerie*, *Marie-Pierre Moine*, *Gildas Dayon*, *Elodie Fernandez*, *Thomas Oudar*, *Laure Coquart*, *Margot Bador*, *Julien Boé*, *Elizabeth Harader*, *Marie Drouard* et enfin *Johan Habert* qui a « partagé » mon bureau pendant une très grande partie de la thèse et dont l’expérience de prévisionniste a été très éclairante pour mon travail.

Je remercie toute ma famille qui, semble-t-il, a toujours cru en moi.

Je profite de ces remerciements pour saluer également toutes les personnes qui ont agrémenté, à un moment ou à un autre, mon quotidien en dehors du CERFACS durant ce séjour toulousain : *Sandra, Julien, Alex, Aurélie, François Mathieu, Flora, Greg, Léo, Marie, Pascal* et enfin *Violaine* qui m'a rejoint à la toute fin de ce travail.

Je remercie enfin les grandes figures créatrices qui, à leur manière, m'ont accompagné tout au long de ce travail à savoir *Alexandre Grothendieck, Wolfgang Amadeus Mozart* et *Franck Angelis*.

Résumé

Ce travail de thèse porte sur l'assimilation de données et le couplage de modèles hydrauliques 1D-2D avec pour objectif d'améliorer la simulation et la prévision des crues dans un contexte temps réel. Les méthodes mathématiques utilisées ici permettent de dépasser les limites de l'usage déterministe des codes d'hydraulique. Elles permettent d'une part d'étendre les échances de prévision et d'autre part de représenter la dynamique des écoulements fluviaux de manière réaliste sur l'ensemble du réseau modélisé.

Ce manuscrit est composé de trois grandes parties.

Une première partie présente le contexte de la prévision des crues et des inondations en France, ses enjeux ainsi que le rôle du SCHAPI¹ et des SPCs². On présente aussi le bassin versant de l'Adour, sous influence maritime qui sera le bassin test pour la suite de notre étude.

Les méthodes liées à la modélisation numérique des équations de Saint-Venant en hydraulique à surface libre, les algorithmes d'assimilation de données, la prévision d'ensemble et les méthodes de couplage sont présentées dans la deuxième partie. On s'intéresse particulièrement aux codes de calcul Mascaret (1D) et Telemac (2D) utilisés dans cette thèse. Dans une troisième partie nous procédons à l'exposé des développements menés dans le cadre de la thèse et des principaux résultats obtenus. Une première section présente un article en cours de review qui traite de l'émulation du filtre de Kalman d'ensemble (EnKF) sur un modèle d'onde de crue : nous montrons, sous certaines hypothèses, qu'un filtre de Kalman invariant avec une fonction de covariance d'erreur sur l'état du modèle au point d'observation stationnaire donne des résultats comparables à ceux du Filtre de Kalman d'Ensemble mais pour un coût de calcul moindre. On montre en particulier qu'une fonction de corrélation au point d'observation initialement gaussienne est transformée à travers les cycles d'assimilation en une fonction anisotrope avec une longueur de portée réduite en aval du point d'observation.

La dynamique des fonctions de covariance et de corrélation sur l'erreur de l'état du modèle avec un modèle Saint-Venant 1D (code Mascaret) et une géométrie réelle est ensuite décrite pour le réseau hydraulique « Adour maritime ». On montre en particulier la grande extension spatiale de ces fonctions et l'impact de la géométrie de la rivière sur la forme de celles-ci. On montre alors que l'extension spatiale de la correction apportée par l'EnKF combinée à une méthode d'inflation spatialisée permettant de compenser la sous-dispersivité de l'ensemble corrige l'état hydraulique du modèle sur une grande partie du réseau de manière cohérente avec les observations disponibles sur l'ensemble du réseau. L'apport de l'assimilation des données in-situ de hauteurs d'eau sur des cas synthétiques

¹Service Central d'Hydrométéorologie d'Appui à la Prévision des Inondations

²Service de Prévision des Crues

et sur des crues réelles a été démontré et permet une correction séquentielle des états spatialisés de hauteur d'eau et de débit. En conséquence, on constate que les prévisions à court terme sont améliorées. Cette étude est complétée par une estimation *a posteriori* des variances d'erreur d'observation. Enfin nous montrons que l'établissement d'un système de prévision probabiliste sur l'Adour est fortement dépendant de la connaissance que nous avons des forçages amonts.

Le modèle 1D décrit l'écoulement mono-dimensionnel sur une grande partie du bassin de l'Adour. Un modèle 2D localisé sur la confluence entre la Nive et l'Adour dans la zone de Bayonne décrit l'écoulement bi-dimensionnel sur une partie restreinte du bassin. Deux méthodes de couplage 1D-2D ont été implémentées ici : une méthode de couplage à interfaces et une méthode de superposition des modèles. Le chainage simple du modèle 1D avec le modèle 2D donne une description de l'ensemble du domaine mais n'assure pas la continuité des deux variables hydrauliques. Cette contrainte est en revanche correctement satisfaite par les deux méthodes de couplage. Le coût de calcul de la méthode par interfaces est significativement plus élevé que celui associé à la méthode de couplage par superposition, mais assure une meilleure continuité des variables. En revanche, la méthode de superposition est immédiatement compatible avec l'approche d'assimilation de données sur la zone 1D. L'apport, sur la zone 2D, de l'assimilation des observations *in situ* des hauteurs d'eau sur la zone 1D a été mis en évidence pour un fort événement de crue en 2014.

Pour finir, les conclusions et les perspectives de ce travail pour l'utilisation en opérationnel au sein du SCHAPI et des SPC sont présentées.

Abstract

This PhD thesis presents research work on data assimilation and multidimensional model coupling in the framework of real-time flood forecasting. The deterministic numerical models proposed to simulate and forecast river flow are limited because of various uncertainties and simplified physics. Data assimilation and coupling approaches allow to overcome these limits and extend the forecast lead time while providing a coherent and realistic description of the hydraulic state.

This thesis is composed of 3 parts gathering 9 chapters.

Part one presents the context of flood forecasting and its stakes, the role of SCHAPI and SPCs. A focus is made on the Adour maritime catchment.

The shallow water equations in the context of the dynamics of river flows are then presented in Part two along with numerical methods usually implemented in hydraulic codes such as Mascaret and Telemac which are used in this work. Data assimilation methods are described with a focus on ensemble approach, then model coupling methods are presented.

Results are presented in Part three. It was shown, on a flood wave propagation model, that an Emulated Ensemble Kalman filter algorithm provides results that are close to those of the Ensemble Kalman Filter but with a much reduced computational cost. The Emulated Ensemble Kalman Filter is built as a Kalman filter algorithm with an invariant covariance function at the observation point. It was also shown that an initially gaussian shaped correlation function at the observing point is transformed into an asymmetric function with a shorter correlation length scale downstream of the observation point. In a real case application on the Adour maritime catchment, it was shown that the dynamics of covariance and correlation model state error functions are strongly influenced by the geometrical characteristics of the river. It was also shown that those functions have a great spatial extent. When combined with an inflation method to circumvent under-dispersivity of the ensemble, the correction provided by the EnKF has an important spatial extent that allows to correct water level and discharge away from the observation points. It was shown on synthetic and real cases experiments that data assimilation provides a hydraulic state that is in great agreement with water level observations. As a consequence of the sequential correction of the hydraulic state over time, the forecasts were also greatly improved by data assimilation over the entire hydraulic network for both assimilated and non-assimilated variables, especially for short term forecasts. This study is combined with the *a posteriori* estimation of the observation error variance using Desroziers criterion. It was finally shown that the quality of ensemble forecast on the Adour maritime network mostly relies on the knowledge of the upstream forcings.

While the 1D model covers a great spatial domain and describes the mono-dimensional flow, the 2D model provides a focus on the Adour-Nive confluence in the Bayonne area.

Two coupling methods have been implemented in this study : a first one based on the exchange of the state variables at the liquid boundaries of the models and a second one where the models are superposed. While simple 1D or chained 1D-2D solutions provide an incomplete or discontinuous (either on water level or on discharge) description of the hydraulic state, both coupling methods provide a full and dynamically coherent description of the hydraulic state for water level and discharge over the entire 1D-2D domain. On the one hand, the interface coupling method presents a much higher computational cost than the superposition methods but the continuity is better preserved. On the other hand, the superposition methods allows to combine data assimilation of the 1D model and 1D-2D coupling. The positive impact of water level in-situ observations in the 1D domain was illustrated over the 2D domain for a flood event in 2014. Conclusions and perspective towards the operational use of both data assimilation and coupling for flood forecasting in France are given.

Table des matières

Introduction	1
I Quelques mots sur la prévision des crues	7
1 La prévision des crues	9
1.1 Le risque d'inondation	9
1.1.1 Quelques considérations sur le risque d'inondation	9
1.1.2 Ce qu'est une inondation et les différents types d'inondation	11
1.2 Quelques aspects des services de prévision des crues	12
1.2.1 Le réseau d'observation et les données d'observation	12
1.2.2 Les modèles utilisés dans le cadre de la prévision des crues	14
1.2.3 Les alertes inondations et la transmission de l'information	19
1.3 La prévision des crues en France	22
1.3.1 Les services de prévision des crues en France	22
1.3.2 Le bassin versant de l'Adour et le réseau hydraulique Adour maritime	23
II Méthodologie	31
2 Les équations de l'hydrodynamique à surface libre et les codes de calcul	33
2.1 Les équations de Navier-Stokes	34
2.1.1 Equation de conservation de la masse	34
2.1.2 Equation de conservation de la quantité de mouvement	34
2.1.3 Cas d'un fluide newtonien et incompressible	35
2.1.4 Conditions aux limites	36
2.2 Dérivation des équations de Saint-Venant	38
2.2.1 Adimensionnalisation des équations de Navier-Stokes	38
2.2.2 Passage à la limite des milieux peu profonds	39
2.2.3 Equations de Saint-Venant	41
2.3 Les modèles d'ondes de crues	42
2.3.1 Le modèle de l'approximation des ondes de crues	43
2.3.2 Le modèle de l'approximation des ondes de crues diffusantes	43
2.4 Principe de résolution des équations de Saint-Venant avec le code Mascaret	44
2.4.1 Forme conservative des équations de Saint-Venant 1D	44

2.4.2	Résolution numérique des équations de Saint-Venant avec une méthode volumes finis	46
2.5	Principe de résolution des équations de Saint-Venant avec le code Telemac-2D	51
2.5.1	Forme non-conservative des équations de Saint-Venant	51
2.5.2	La méthode des pas fractionnaires	52
2.5.3	Traitement de la propagation, de la diffusion et des termes sources – Discrétisation en temps et en espace	53
2.5.4	Formulation faible, approximation interne et mise sous forme d'un système linéaire	55
3	L'assimilation de données	61
3.1	Une introduction	62
3.2	Le vocabulaire de l'assimilation de données	62
3.3	La modélisation des erreurs et les matrices de covariance d'erreur	63
3.3.1	Remarques préliminaires sur les erreurs, les matrices de covariance et de corrélation d'erreur	64
3.3.2	Les erreurs liées à l'état du modèle	65
3.3.3	Les erreurs liées à la modélisation	66
3.3.4	Les erreurs liées aux observations	66
3.4	Filtrage bayésien	67
3.4.1	Règle de Bayes	67
3.4.2	Estimation optimale	68
3.4.3	Formule de Bayes récursive ou filtre bayésien	69
3.4.4	Considération sur le filtrage bayésien et le choix de l'estimateur optimal	70
3.5	Le Best Linear Unbiased Estimator - BLUE	70
3.6	Les méthodes variationnelles	72
3.6.1	Formulation de la fonction coût	72
3.6.2	Le 3D-Var	73
3.6.3	Le 4D-Var	74
3.6.4	Quelques variantes du 3D-Var et du 4D-Var	77
3.7	Le filtre de Kalman et ses dérivations	80
3.7.1	Le filtre de Kalman	80
3.7.2	Le filtre de Kalman Etendu ou Extended Kalman Filter (EKF)	81
3.7.3	Le filtre de Kalman de rang réduit - RRSQRT	82
3.7.4	Le filtre SEEK	83
3.7.5	Le filtre de Kalman d'ensemble - EnKF	84
3.7.6	Les versions déterministes du filtre de Kalman d'ensemble	86
3.7.7	Les problèmes d'échantillonnage et les solutions liées	88
4	Les prévisions d'ensemble	91
4.1	Principe de la prévision d'ensemble et représentation	91
4.1.1	Principe	91
4.1.2	Représentation	92
4.2	Description quantitative et évaluation des prévisions d'ensemble	94

4.2.1	Description quantitative	95
4.2.2	Evaluation des prévisions d'ensemble	96
5	Couplage de modèles hydrauliques 1D-2D	105
5.1	Couplage à interfaces	106
5.1.1	Raccordement des modèles à l'interface	107
5.1.2	Algorithme de couplage	109
5.1.3	Quelques exemples d'applications numériques de l'algorithme de couplage	111
5.2	Couplage de modèles superposés	118
5.2.1	Calcul des termes d'échange entre lit mineur et lit majeur	120
5.2.2	Equations de Saint-Venant 1D conservatives avec terme source	124
5.2.3	Algorithme de couplage	124
5.2.4	Variations autour d'une application numérique	125
III	Résultats	129
6	Etat de l'art de l'assimilation de données pour l'hydraulique fluviale	131
7	Emulation du filtre de Kalman d'ensemble sur le modèle d'ondes de crue 1D	139
7.1	Introduction	139
7.2	Article soumis dans la revue « Journal of Hydrology »	140
7.3	Conclusion	158
8	Application du filtre de Kalman d'ensemble au réseau hydraulique Adour maritime	161
8.1	Le coupleur de code OpenPALM, le réseau Adour maritime et l'implémentation des méthodes d'ensemble sur ce réseau	162
8.1.1	Le coupleur de codes OpenPALM	162
8.1.2	Rappels sur le réseau Adour maritime	164
8.1.3	Implémentation des méthodes d'ensemble sur le réseau Adour maritime	166
8.1.4	Génération des membres de l'ensemble	169
8.2	Etude des fonctions de covariance et de corrélation d'erreur d'ébauche sans assimilation	170
8.2.1	Diagnostic des fonctions de covariance et de corrélation	171
8.2.2	Les fonctions de covariance	172
8.2.3	Les fonctions de corrélation	175
8.2.4	Caractérisation et évolution temporelle de l'extension spatiale des fonctions de corrélation ZZ	177
8.3	Application de l'EnKF au réseau Adour maritime : première approche et limitations	181
8.3.1	Une expérience d'assimilation de données de hauteurs d'eau synthétiques	181

8.3.2	Résultats numériques	182
8.3.3	Discussion	189
8.4	Application des critères de Desroziers	191
8.4.1	Estimation des covariances d'erreur d'observation	191
8.4.2	Estimation des covariances d'erreur d'ébauche et inflation	195
8.4.3	Une expérience d'assimilation de données synthétiques	198
8.4.4	Expériences d'assimilation sur des données réelles	201
8.5	Vers des prévisions d'ensemble sur le réseau Adour maritime?	204
8.5.1	Cadre de l'expérience et démarche	204
8.5.2	Estimation des prévisions d'ensemble	205
8.5.3	Discussion	211
9	Couplage de modèles hydrauliques 1D-2D sur la ville de Bayonne	213
9.1	Le modèle hydraulique 2D sur Bayonne	214
9.2	Couplage à interfaces	218
9.2.1	Modèles, interfaces et raccordement	218
9.2.2	Algorithme de couplage et implémentation	219
9.2.3	Expérience numérique de couplage à interfaces sur l'épisode de 2014	222
9.3	Couplage de modèles superposés	226
9.3.1	Modèles utilisés, disposition et termes échangés entre les modèles	227
9.3.2	Algorithme de couplage et implémentation	230
9.3.3	Expérience numérique de couplage de modèles superposés sur l'épisode de 2014	231
9.4	Combinaison de l'assimilation de données ensembliste et du couplage de modèles superposés	240
9.4.1	Implémentation	241
9.4.2	Résultats numériques	241
	Conclusions et perspectives	247
A	Inondations des Gaves Réunis à Peyrehorade du 19 Juin 2013	251
B	Illustration de l'évolution temporelle des fonctions de corrélation ZZ sur le chemin 7-5-2-1	259
C	Données utilisées pour les expériences réelles	261
	Bibliographie	271
	Table des figures	279
	Liste des tableaux	287

Introduction

Crues et inondations constituent un risque naturel majeur pour la sécurité des biens et des personnes. Ces dernières années de nombreux pays ont été frappés par des crues qualifiées d'historiques par les médias : épisodes cévenols à l'automne 2014 dans le Sud de la France, inondations majeures dans les Balkans au printemps 2014, inondations en Europe Centrale avec une crue du Danube à l'été 2013... Prévoir les inondations, en anticiper l'ampleur, informer le grand public et les autorités compétentes font partie des rôles des services météorologiques et de prévision des crues. Pour mener à bien leur mission de prévision et d'anticipation, ceux-ci disposent d'un ensemble d'observations satellitaires, radar, *in situ* ainsi que de modèles numériques de prévision hydrologiques et hydrodynamiques.

En France, c'est le SCHAPI (Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations) et les 22 Services de Prévision des Crues (SPC) qui sont en charge de la surveillance du réseau hydraulique et de la prévision du risque lié aux crues. Le SCHAPI établit quotidiennement (2 mises à jour quotidiennes) une cartographie du risque de dépassement de seuil pour les cours d'eau principaux et pour les stations d'observation *in situ* maintenues. Cette information mise à disposition en ligne pour le grand public et les instances gouvernementales (<http://www.vigicrues.gouv.fr/>) permet notamment de visualiser et quantifier les hauteurs d'eau observées sur les jours passés. Un bulletin d'évolution prévue ainsi qu'un récapitulatif des prévisions à différentes échéances sont aussi mis à disposition, par tronçon, ou bien sur l'ensemble du territoire national métropolitain (voir figure 1).

Pour mener à bien cette mission, le SCHAPI et les SPC disposent d'outils de modélisation et d'un réseau d'observation *in situ* (principalement de hauteur d'eau). La qualité des simulations et des prévisions hydrologiques et hydrauliques repose d'une part sur la qualité des données géométriques et géographiques : bathymétrie de la rivière, description des ouvrages, description du réseau hydraulique (affluents, apports/prélèvements) et description des paramètres hydrauliques (coefficients de frottement). Elle dépend d'autre part en grande partie de la description de l'apport en eau sur le réseau qui provient principalement de l'observation des hauteurs d'eau à l'amont des tronçons mais aussi de précipitations prévues. Enfin, l'apport d'information pour l'observation *in situ* est crucial : cette information est utilisée pour calibrer et valider les modèles ainsi que pour aider le prévisionniste dans son analyse.

Les résultats des modèles hydrauliques sont limités par plusieurs facteurs d'incertitudes qui impactent la capacité de simulation et de prévision. L'erreur sur la géométrie des

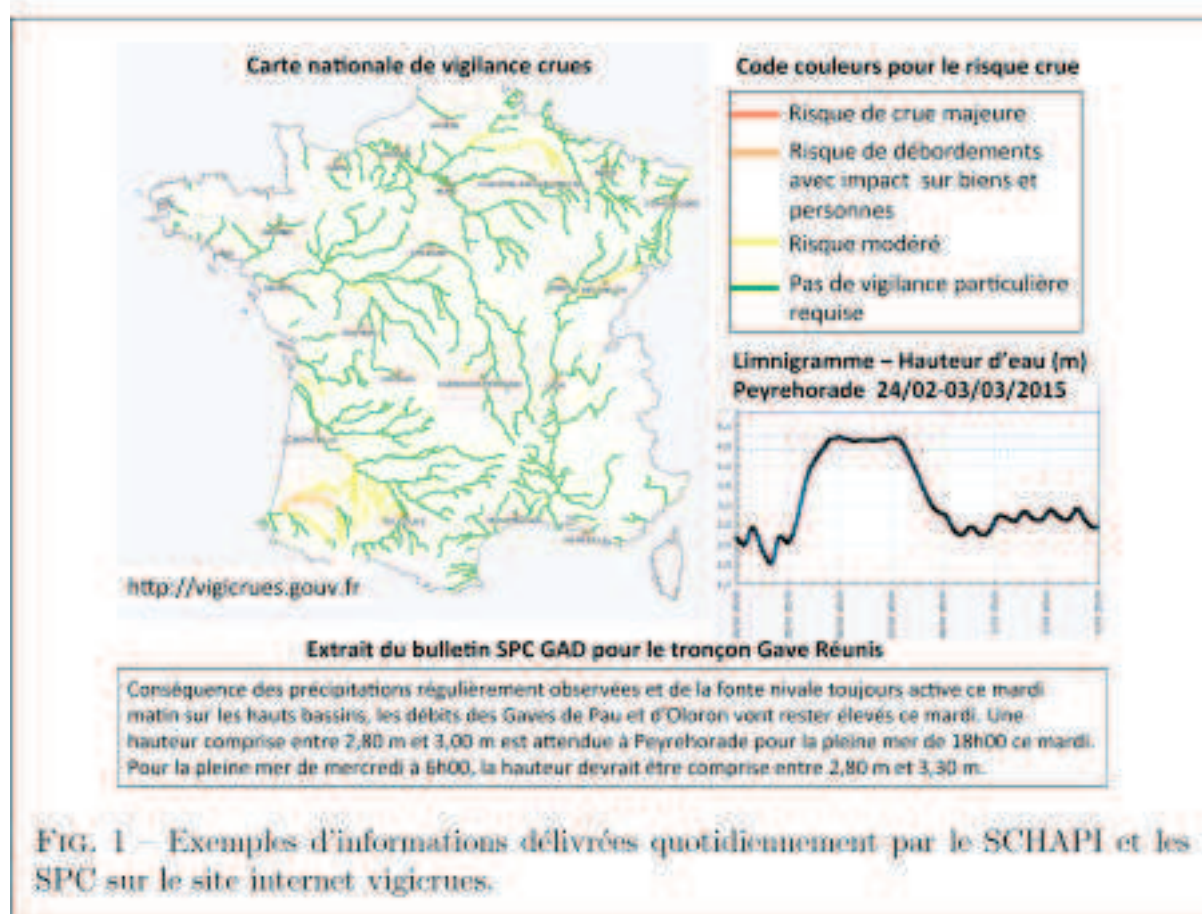


FIG. 1 – Exemples d'informations délivrées quotidiennement par le SCHAPI et les SPC sur le site internet vigicrues.

rivières et les paramètres hydrauliques se traduisent en une erreur sur l'état hydraulique simulé même lorsque les apports sont considérés comme parfaits. Il est de plus tout à fait légitime de considérer que les apports hydrologiques sont entachés d'une erreur qui peut être soit une erreur d'observation (mesure de la hauteur d'eau, transformation en débit via une courbe de tarage) soit une erreur de prévision (prévision des débits via un modèle hydrologique), ou une combinaison des deux effets. Classiquement, il est raisonnable d'admettre que le réseau hydraulique est décrit de manière partielle, par des données géométriques, hydrologiques et hydrauliques approchées et de manière imparfaites par des équations simplifiées et discrétisées au sein des codes de calcul, malgré les progrès considérables effectués dans le domaine de la modélisation et de l'observation. On peut citer à titre d'exemple les avancées dans le domaine de l'observation LIDAR qui permettent à présent d'observer la topographie à haute résolution.

Malgré ces limitations, la compréhension et la prévision de la dynamique des écoulements fluviaux est un défi crucial qu'il convient de relever en rallongeant autant que possible les échéances de prévision afin de permettre l'anticipation et la gestion des événements à risque. Pour ce faire, la communauté hydraulique s'attache à quantifier et réduire les sources d'erreurs précédemment citées. L'étape de quantification passe par une étude de sensibilité : on cherche à quantifier la perturbation sur l'état hydraulique résultant d'une perturbation sur une des entrées incertaines du modèle hydraulique. Les approches de type Monte-Carlo permettent notamment de réaliser ce type d'étude au prix d'un

coût de calcul élevé. D'autres méthodes permettent d'établir des diagnostics similaires en limitant le coût de calcul, par exemple en passant par l'utilisation d'un modèle approché (POD, Polynômes du Chaos...). Une fois l'incertitude quantifiée, il convient alors de proposer des méthodes pour la réduire et ainsi améliorer la qualité de la simulation hydraulique au temps passé, présent ou futur. C'est l'objectif des méthodes d'assimilation de données qui consistent à combiner les observations avec le modèle numérique.

Les méthodes d'assimilation de données ont été développées originellement dans le cadre de la météorologie et ont ensuite été appliquées à d'autres domaines comme par exemple l'océanographie. Elles sont également appliquées en hydraulique fluviale pour réduire l'erreur liée aux différentes sources d'incertitudes apparaissant dans les modèles. Ainsi on distingue les méthodes liées à l'estimation et/ou à la correction des forçages (amonts ou latéraux). Ces méthodes trouvent tout leur intérêt dans le cas où un ou plusieurs forçages sont calculés via des courbes de tarage ayant une erreur importante sur les événements extrêmes ou encore lorsqu'un forçage est donné par une station et que celle-ci est défaillante. Une seconde approche est basée sur la correction des paramètres physiques comme les coefficients de frottement ou la géométrie du modèle. Et enfin il faut mentionner les approches portant sur la correction de l'état hydraulique, dans ce cas on ne corrige pas la source de l'erreur mais son impact sur l'état hydraulique simulé. C'est particulièrement à cette dernière approche que nous nous sommes intéressés durant ce travail thèse.

Naturellement, le développement et l'implémentation des méthodes d'assimilation de données pour la prévision des crues doivent être guidés par les contraintes de temps inhérentes à la prévision en temps réel ainsi qu'à l'architecture de la chaîne opérationnelle. La prochaine chaîne opérationnelle française pour la prévision des crues est la POM (Plateforme Opérationnelle de Modélisation), actuellement en cours de développement au SCHAPI. A terme elle sera l'outil communautaire pour les SPC pour consulter les données, lancer des calculs et ainsi faire des prévisions. Cela suppose de gérer les flux de données, intégrer les codes de calcul, effectuer du post-traitement, délivrer des prévisions (avec potentiellement une marge d'erreur) sur le site vigicrue. En l'état actuel, la POM n'est par exemple pas conçue pour gérer des méthodes d'ensemble. Cela a donc un impact sur le choix des méthodes d'assimilation de données utilisées en opérationnel.

Ce travail de thèse s'insère dans un contexte collaboratif entre le CERFACS (URA 1875), le SCHAPI (co-financeur de la thèse aux côtés de la région Midi-Pyrénées) et EDF (Electricité De France), acteur central pour la gestion de la ressource en eau en France liée à l'activité de production d'électricité. L'objectif de la thèse est de proposer des méthodes avancées pour améliorer la simulation et la prévision de la dynamique des écoulements fluviaux, en s'intéressant particulièrement à la prévision des crues. On s'appuie pour ce travail sur les outils de simulation numérique existant au sein de la communauté hydraulique, notamment les codes de calcul Mascaret et Telemac développés par EDF et le CEREMA (Centre d'Etudes et d'Expertise sur les Risques, l'Environnement, la Mobilité et l'Aménagement). Une première approche consiste à assimiler de manière séquentielle, et avec un algorithme ensembliste, les hauteurs d'eau observées *in situ* pour améliorer l'état hydraulique. Cette approche permet de rejouer de manière plus réaliste des événements de crues et de donner une description plus fiable (au sens moins incertaine) d'un état hydraulique initial pour la prévision. Une deuxième approche propose de coupler les modèles 1D et 2D afin de décrire de manière cohérente, dans l'espace et dans le temps, la dynamique

de l'écoulement. On démontrera ici que la combinaison de l'assimilation de données et du couplage de modèles multi-dimensionnels permet d'améliorer la modélisation de l'Adour maritime, site d'intérêt pour le SPC GAD (Gironde-Adour-Dordogne).

Les développements et résultats majeurs obtenus au cours de ce travail de thèse se résument de la sorte :

- un premier axe de travail qui porte sur l'assimilation de données pour la prévision des crues en temps réel. Le fil rouge de cette partie réside dans les fonctions de covariance d'erreur sur l'état du modèle. Ces fonctions sont très importantes en assimilation de données car ce sont elles qui répartissent sur le domaine l'information apportée en un point par une observation. Nous nous sommes intéressés en premier lieu à l'application sur un modèle d'onde de crue avec géométrie simplifiée (canal rectiligne à fond plat et pente constante) d'un algorithme d'assimilation de données ensembliste : le filtre de Kalman d'ensemble ou EnKF. On montre, sous certaines hypothèses, qu'on peut simuler le comportement de cet algorithme par un algorithme d'assimilation au coût de calcul plus faible. Celui-ci utilise au point d'observation une fonction de covariance asymétrique dont l'extension spatiale est plus importante à l'amont qu'à l'aval. On justifie par la même occasion la forme asymétrique des fonctions de covariance au cours de l'analyse d'assimilation. Dans un second temps nous nous sommes intéressés à l'étude des fonctions de covariance sur un réseau avec un modèle Saint-Venant 1D et une géométrie réelle (le réseau Adour maritime), ainsi qu'à l'applicabilité de l'EnKF sur ce réseau. Nous montrons les limitations de l'EnKF dans sa version classique et montrons les avantages que l'on peut retirer des méthodes d'estimation des covariances d'erreur d'observation ainsi que des méthodes d'inflation pour compenser les éventuels problèmes de sous-dispersivité de l'ensemble. Enfin nous montrons que l'établissement d'un système de prévisions probabilistes sur le réseau Adour maritime dépend fortement de la connaissance que l'on a des forçages ;
- un second axe de travail qui porte sur le couplage de modèles hydrauliques 1D-2D. Sur le réseau Adour maritime deux modèles hydrauliques ont été développés : un modèle 1D couvrant la portion maritime de l'Adour et un modèle 2D localisé autour de Bayonne et développé pour comprendre le rôle des plaines d'inondation dans cette zone. Deux méthodes de couplage ont été implémentées dans cette thèse pour combiner ces deux modèles. Une première, assez classique, que l'on désignera par « couplage à interfaces » où le modèle 1D est décomposé en 3 sous-modèles 1D et ceux-ci sont couplés au modèle 2D au niveau des frontières liquides de celui-ci. Cette méthode permet d'établir la continuité des variables d'état entre les modèles mais avec un temps de calcul très important. Une seconde dite de couplage de « modèles superposés » où le modèle 1D est conservé en un seul morceau et superposé au modèle 2D là où les deux modèles co-existent. Dans ce cas, le modèle 1D force le modèle 2D et en retour celui-ci calcule lorsqu'il est en crue les termes sources pour les équations de conservation de la masse et de conservation de la quantité de mouvement modélisant les échanges entre lit mineur et lit majeur. Néanmoins, dans ce cas, la géométrie du modèle 1D ainsi que ses coefficients de frottement sont ajustés sur la zone de recouvrement de manière à assurer une équivalence grossière des relations hauteurs/débits entre les modèles. Le fait que le modèle 1D soit conservé d'un seul tenant et que les relations hauteurs/débits entre les modèles soient équivalentes nous a permis finalement de combiner couplage de modèles superposés et assimilation de données ensembliste.

Ce mémoire de thèse s'articule en 3 parties qui regroupent au total 9 chapitres.

La première partie inclut un unique chapitre dans lequel on présente les enjeux de la prévision des crues ainsi que certains de ses aspects les plus importants. Nous présentons succinctement la prévision des crues en France et en particulier le réseau Adour maritime sur lequel nous avons appliqué les méthodes développées dans ce travail.

La deuxième partie est dédiée à la méthodologie. Elle regroupe le chapitre 2 dans lequel on présente les principales équations de l'hydrodynamique à surface libre : Navier-Stokes, Saint-Venant, modèles d'onde de crue, ainsi que les principes de résolution des équations de Saint-Venant par les codes de calcul que nous avons utilisés : Mascaret-1D et Telemac-2D. Dans le chapitre 3 nous donnons une présentation générale du vocabulaire de l'assimilation de données et des principales méthodes existantes : méthodes du type interpolation stochastique (filtres de Kalman) et les méthodes variationnelles (4D-Var par exemple). Dans le chapitre 4 nous présentons les principes de la prévision d'ensemble ainsi que les propriétés qu'un système de prévision probabiliste doit vérifier. Enfin dans le chapitre 5 nous présentons les méthodes de couplage utilisées dans ce travail et illustrons ces méthodes avec quelques résultats tirés de ces articles.

La troisième partie est dédiée aux résultats. Celle-ci s'ouvre avec le chapitre 6 qui présente un état de l'art des méthodes d'assimilation de données pour l'hydraulique fluviale et la prévision des crues. Le chapitre 7 présente la reproduction d'un article soumis à la revue « Journal of Hydrology » et portant sur l'émulation du filtre de Kalman d'ensemble sur le modèle d'onde de crue. Le chapitre 8 traite de l'applicabilité de l'EnKF sur le réseau Adour maritime (utilisé en opérationnel par le SPC GAD) sur un modèle Saint-Venant avec une géométrie réelle. Enfin dans le chapitre 9 nous présentons l'application des méthodes de couplage, présentées au chapitre 5, entre le modèle 2D développé sur Bayonne et le modèle 1D développé sur l'Adour maritime. Nous présentons également la combinaison du couplage de modèles superposés et de l'assimilation de données ensembliste avec le modèle 1D.

Naturellement ce manuscrit se termine par une section dédiée aux conclusions de ce travail où nous évoquons l'apport des méthodes d'assimilation et de couplage développées au cours de ce travail ainsi que les limitations de celles-ci. Nous terminons enfin sur les perspectives à donner à cette étude, notamment des perspectives sur la modélisation du réseau hydraulique « Adour maritime » et les méthodes de couplage à développer à l'avenir dans l'objectif d'une implémentation pour la prévision en temps réel.

Première partie

Quelques mots sur la prévision des crues

Chapitre 1

La prévision des crues

Sommaire

1.1	Le risque d'inondation	9
1.1.1	Quelques considérations sur le risque d'inondation	9
1.1.2	Ce qu'est une inondation et les différents types d'inondation	11
1.2	Quelques aspects des services de prévision des crues	12
1.2.1	Le réseau d'observation et les données d'observation	12
1.2.2	Les modèles utilisés dans le cadre de la prévision des crues	14
1.2.3	Les alertes inondations et la transmission de l'information	19
1.3	La prévision des crues en France	22
1.3.1	Les services de prévision des crues en France	22
1.3.2	Le bassin versant de l'Adour et le réseau hydraulique Adour maritime	23

1.1 Le risque d'inondation

1.1.1 Quelques considérations sur le risque d'inondation

Le risque d'inondation est l'un des risques naturels les plus importants. Les chiffres que nous citons dans ce paragraphe et le suivant sont tirés de [World Meteorological Organization, 2011]. Sur la période 1990-2001 celui-ci représente 50% des désastres liés à l'eau loin devant les épidémies liées à l'eau (28%) et les sécheresses (11%), figure 1.1. Les chiffres sont impressionnants et donnent la mesure de l'impact des crues tant en termes de vies humaines que de coût économique. Ainsi on estime que les crues sont responsables de 15% des décès liés aux risques naturels et sur la période 1987-1997, 44% des désastres liés aux crues ont touché l'Asie emportant 228000 vies soit 93% de tous les décès liés aux crues à travers le monde. Les pertes économiques pour cette région du monde sur cette période s'élèvent à 136 milliards USD. Pendant l'été 2007 au Royaume-Uni les crues de nature pluviale ont causé des pertes s'élevant à 4 milliards de livre dont 3 milliards de pertes sur des biens assurés. En 2009 le Bangladesh Water Development estime le coût des dommages liés aux crues dans le secteur de l'eau exclusivement à 750 millions USD.

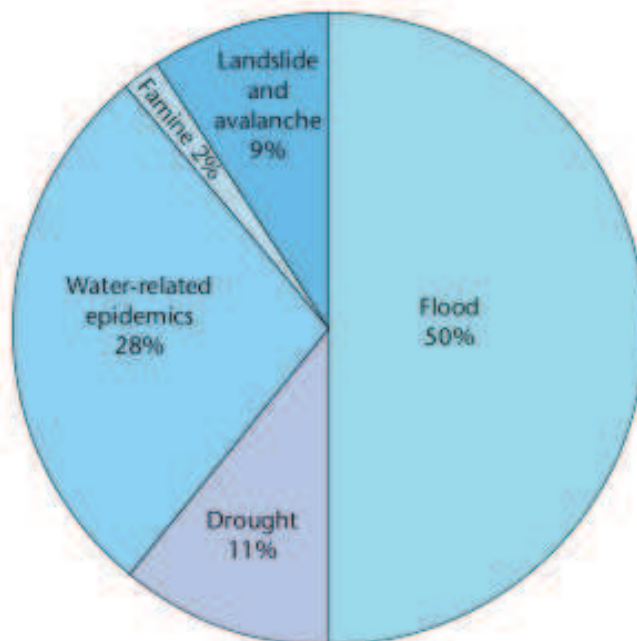


FIG. 1.1 – Types de catastrophes naturelles liées à l'eau sur la période 1990-2001.
Source : [World Meteorological Organization, 2011].

Le risque d'inondation est d'autant plus important que la proportion d'individus exposés à ce risque est importante. Par exemple plus de 12% de la population de la Grande-Bretagne et de l'Irlande du Nord vit dans des plaines d'inondations fluviales ou dans des zones littorales exposées au risque d'inondations côtières. Environ la moitié de la population des Pays-Bas vit sous le niveau de la mer et en Hongrie c'est 25% de la population qui vit dans les plaines d'inondation du Danube. De plus l'exposition des populations est d'autant plus importante que le risque d'inondation est très répandu. En effet les populations montagneuses sont menacées par le risque de crues éclair, crues aggravées en cas de fonte importante des neiges, tandis que les populations urbaines sont menacées par les pluies importantes dépassant les capacités d'évacuation des canaux et des égouts. S'ajoutent à cela les dégâts que peuvent causer en zones urbaines le déplacement de débris, de matériaux ou de véhicules. Les populations vivant dans les plaines d'inondation des rivières sont bien évidemment exposées ainsi que les populations littorales où un coefficient de marée important conjugué à de forts vents peut résulter en des inondations côtières. Selon leur implantation géographique ces populations sont également menacées par les inondations résultant de cyclones tropicaux ou encore par des tsunamis. Plus rarement des inondations surviennent du fait d'opérations de régulation d'un ou plusieurs ouvrages hydrauliques ou encore du fait de la ruptures d'un barrage. On estime que ce sont 1,5 milliard d'individus qui ont été touchés par les crues durant la dernière décennie du 20ème siècle.

Le risque d'inondation est donc un risque naturel aux conséquences potentiellement dévastatrices. Cela est d'autant plus préoccupant que les événements météorologiques importants du fait du changement climatique tendent à rendre les inondations plus extrêmes, plus répandues et plus fréquentes. Par conséquent il est légitime de se demander si les dégâts liés aux crues et leur coût tant humain qu'économique vont augmenter dans les années à venir et dans quelle mesure [Coppola et al., 2014, Dumas et al., 2013]. Il n'est certes pas possible de réduire à zéro le risque d'inondation néanmoins peut-on l'anticiper de manière à en atténuer les conséquences. C'est l'un des buts de la prévision des crues qui a plus globalement pour objet d'alerter le grand public et les autorités le plus tôt possible d'un risque de crue, de sa localisation ainsi que de son importance.

1.1.2 Ce qu'est une inondation et les différents types d'inondation

Pour établir les définitions suivantes nous nous sommes basés sur la définition qui en est donnée dans [World Meteorological Organization, 2011]. Dans la suite de ce travail nous définirons :

- ***une crue*** comme étant la hausse du niveau d'un cours d'eau jusqu'à un pic après lequel le niveau d'eau décroît plus lentement qu'il n'est monté ;
- ***une inondation*** comme étant la submersion de zones qui ne sont habituellement pas recouvertes d'eau.

On distingue selon leur origine différents types d'inondations :

- ***les crues éclair*** : celles-ci se produisent le plus souvent dans les régions montagneuses sujettes à de fréquents et sévères épisodes orageux et sont souvent la conséquence de pluies à la fois intenses et de courte durée. On peut évoquer à ce titre les épisodes cévenoles dans le Sud de la France ;
- ***les inondations fluviales*** : celles-ci se produisent lorsque l'écoulement dépasse les capacités de la rivière, déborde par dessus les digues et se répand dans les plaines d'inondations. C'est à ce type d'inondations que nous nous intéresserons plus particulièrement dans ce travail ;
- ***les inondations ponctuelles*** : ces inondations sont causées par des pluies pouvant durer de plusieurs heures à plusieurs jours sur un bassin versant et peuvent s'avérer très importantes. En général ces pluies sont accompagnées de situations météorologiques bien particulières avec des dépressions et des tempêtes ;
- ***les inondations multiples*** : ces inondations sont le fruit d'importantes chutes de pluie associées à une succession d'épisodes météo rapprochés. On citera l'exemple des inondations de la plaine indo-gangétique causées par le passage d'une succession de dépressions provenant de la baie du Bengale et suivant plus ou moins le même cheminement ;
- ***les inondations saisonnières*** : ces inondations peuvent être la conséquence de pluies saisonnières comme c'est le cas par exemple dans les régions du monde affectées par la mousson ; elles peuvent durer plusieurs semaines avec plusieurs pics de crue durant l'évènement. Ces inondations peuvent également être la conséquence d'élévation du niveau d'eau dans des lacs en amont d'un bassin versant (exemple du Lac Victoria et du fleuve Nil) ;

- **les inondations côtières** : celles-ci résultent d'une surcote coïncidant avec des vents très forts et un important coefficient de marée. La surcote elle-même résultant de l'élévation du niveau de la mer due à la faible pression atmosphérique. Les tsunamis sont aussi à l'origine d'inondations côtières spectaculaires et particulièrement destructrices ;
- **les inondations estuariennes** : comme leur dénomination l'indique ces inondations ont lieu dans les estuaires. L'interaction entre les hautes eaux fluviales et un haut niveau de marée peut provoquer des inondations. Les inondations estuariennes sont plus fréquentes mais cependant moins importantes en termes de zone inondée que les inondations liées à la surcote ;
- **les inondations urbaines** : celles-ci ont lieu dans les zones urbanisées lors d'évènements pluvieux importants résultant en des écoulements dépassant les capacités des systèmes de drainage et d'écoulement de la ville ;
- **les inondations causées par la fonte des neiges** : au printemps en altitude, l'élévation de la température cause la fonte des neiges tombées pendant l'hiver. Si celle-ci est rapide elle peut engendrer des crues se propageant en aval de la rivière. La gravité de ces crues peut être aggravée si elles sont accompagnées d'importantes chutes de pluie et que le sous-sol demeure gelé ;
- **les inondations causées par un bouchon de glace** : dans les régions subissant un dégel saisonnier des accumulations de glace faisant barrage peuvent se créer et former des poches de rétention. Lorsque le bouchon de glace se rompt sous la pression de l'eau en amont il peut être à l'origine d'une onde de crue (tout comme lors de la rupture d'un barrage) se propageant en aval et causant des dégâts importants. Mais ces phénomènes sont très difficiles à prévoir.

Distinguer ces différents types de crue n'est pas sans intérêt puisqu'elles permettent d'orienter de la manière la plus pertinente la mise en place d'un système d'alerte et de prévision des crues adapté à la zone considérée.

1.2 Quelques aspects des services de prévision des crues

Les services de prévision des crues ont pour rôle d'informer le grand public ainsi que la protection civile et les autorités concernées (préfets, maires...) de l'imminence d'un épisode de crue ainsi que de son intensité et ce suffisamment à l'avance de manière à ce que des mesures appropriées puissent être prises. Pour assurer cette mission au mieux les systèmes de prévision des crues reposent sur trois composantes majeures que nous abordons dans les grandes lignes ci-dessous : le réseau d'observation et les données d'observation, les modèles de prévision et un système d'alerte et d'information.

1.2.1 Le réseau d'observation et les données d'observation

Idéalement le réseau d'observation est une combinaison de pluviomètres et de station d'observation des hauteurs d'eau (et éventuellement des débits) le long des cours d'eau aux points d'intérêt et télétransmettant en temps réel ces données au(x) centre(s) opérationnel(s) de prévision.

Le choix des sites où implanter des stations d'observation ou des pluviomètres est guidé par l'identification des zones à risques. On reporte en figure 1.2 un tableau qui relie le risque de crue au type de zone touchée par la crue. Le risque lui-même est fortement corrélé au degré d'aménagement de la zone touchée par la crue. Par exemple une zone peu développée comme ça peut être le cas de certaines plaines d'inondation ou des forêts peuvent présenter un risque faible. Au contraire des zones fortement urbanisées avec des infrastructures importantes et avec une grande densité de population présentera un risque très fort. Ces zones peuvent nécessiter de prendre des actions ciblées comme la fermeture de barrages, l'installation de protection d'appoint démontables ou même la préparation pour l'évacuation.

Risk impacts matrix				
Flood risk	Flood impact zone			
	Undeveloped area (low)	Agricultural land (medium)	Low density urban (high)	Urban centres and key infrastructure (very high)
High	High/low	High/medium	High/high	High/very high
Medium	Medium/low	Medium/medium	Medium/high	Medium/very high
Low	Low/low	Low/medium	Low/high	Low/very high

FIG. 1.2 – Tableau des risques d'impact d'une crue en fonction de la zone considérée.
Source : [World Meteorological Organization, 2011].

Le choix des sites où implanter des stations d'observation va également fortement dépendre du temps de prévision dont on souhaite disposer en fonction des zones concernées par les crues, dans les zones urbanisées, notamment. En Angleterre où en général le temps de réponse d'un cours d'eau est assez court, l'agence pour l'environnement a fixé l'échéance minimale de prévision à deux heures. Bien sûr cette échéance de prévision minimale n'est pas nécessairement la même à travers tout le pays. Au contraire au Bangladesh où le temps de réponse des cours d'eau est nettement plus long l'échéance de prévision est fixée à 72h avec un rapport périodique sur l'état de la rivière en général toutes les 3 à 6h. Enfin un autre facteur influençant le choix des sites pour les stations d'observation est la configuration du cours d'eau. En effet il est nécessaire aux services de prévision de pouvoir suivre l'évolution de la crue tout au long de sa progression de l'amont vers l'aval. A ce titre dans les réseaux fluviaux importants concentrant de nombreux affluents il est nécessaire de ne pas négliger les affluents mineurs. Cela a pu être le cas notamment en Europe centrale pendant les crues de 2002 où certains des dégâts les plus importants ont eu lieu dans des petits bassins versants où la dangerosité de la situation n'a pas été identifiée à temps.

Aux données collectées par les pluviomètres il faut ajouter éventuellement les données transmises par des radars météo ou des images satellites. Les radars météo présentent

l'avantage évident par rapport aux pluviomètres disséminés sur un bassin versant de donner une indication spatialisée des précipitations. Néanmoins en pratique il existe d'importantes limitations liées à l'utilisation des radars, notamment la limitation de la portée de ceux-ci, l'atténuation du signal ou encore la calibration. De plus le coût de la maintenance de ces instruments est un obstacle important à leur utilisation à l'échelle nationale.

En plus des données radar il faut considérer les données satellitaires. Dans le cadre de la prévision hydrologique celles-ci sont émises par des satellites géostationnaires ou en orbite polaire. Les capacités des satellites géostationnaires s'accroissent continuellement avec une mise à jour rapide des données observées de l'ordre de la demi-heure et une résolution spatiale allant en augmentant. L'efficacité des satellites à orbite polaire est quant à elle limitée par la fréquence de passage au-dessus d'un lieu donné. Néanmoins la combinaison de données de satellites géostationnaires et de satellites à orbite polaire pourrait donner de très bonnes estimations des chutes de pluie pour la prévision des crues. Cela s'avèrerait particulièrement utile pour les cours d'eau non instrumentés. Il faut souligner qu'il existe aux Etats-Unis comme en Europe des missions satellitaires chargées d'estimer les précipitations en temps réel et mettant gratuitement à disposition de la communauté ces données. On peut également évoquer la mission franco-américaine SWOT (Surface Water and Ocean Topography) pour l'étude des surfaces d'eau océanique et continentale. Cette mission aura pour objet de mesurer les hauteurs d'eau aussi bien des surfaces continentales (fleuves, lacs, zones inondées) que des océans avec une résolution spatiale de l'ordre de 50 à 100 m et permettra, grâce à son altimètre à fauchée, de couvrir tous les lacs rivières, océans... au moins deux fois tous les 21 jours.

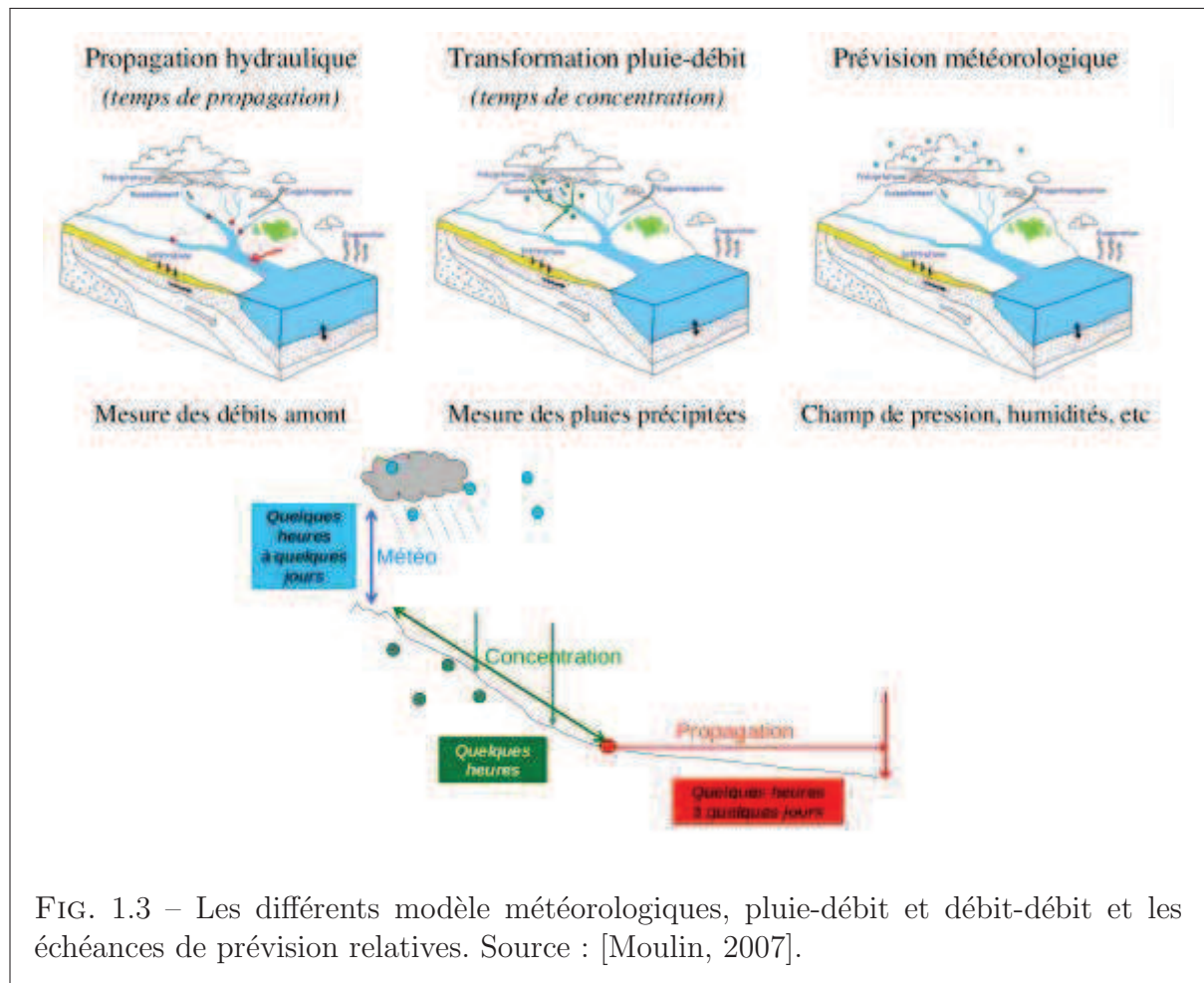
Les données collectées par les différents systèmes d'observation sont ensuite télétransmises aux centres opérationnels de prévision. Après conversion au format adéquat ces données sont utilisées comme données de forçages pour les modèles utilisés pour la prévision des crues. Idéalement cette suite de modèles est constituée de :

- d'un (ou plusieurs) modèle pluie-débit qui calcule en fonction des pluies données comme forçage les débits résultant à l'exutoire du bassin versant ;
- d'un modèle de routage qui permet de suivre l'évolution de l'onde de crue de l'amont vers l'aval du cours d'eau.

Ces modèles vont permettre d'établir des prévisions à partir desquelles les systèmes d'alerte et d'information pourront être mis en place.

1.2.2 Les modèles utilisés dans le cadre de la prévision des crues

Les modèles utilisés dans le cadre de la prévision des crues vont dépendre des phénomènes que l'on souhaite modéliser. Ainsi on distingue trois types de modèles : les modèles météorologiques dont il ne sera pas question ici et dont les sorties seront fournies par les services concernés, les modèles pluie-débit qui en fonction des précipitations données en entrée calculent les débits à l'exutoire du bassin versant et les modèles débit-débit (ou modèles de routage) qui permettent le suivi de l'onde de crues. Nous illustrons ces différents modèles ainsi que les échéances de prévision relatives sur la figure 1.3.



Les modèles débit-débit

Les modèles débit-débit sont utilisés pour suivre et estimer l'évolution d'une crue le long d'un cours d'eau. [World Meteorological Organization, 2011] distingue deux types de modèles de routage : les modèles hydrologiques et les modèles hydrauliques. L'échéance de prévision de ces modèles va, en fonction des dimensions du réseau hydraulique, de quelques heures à quelques jours mais n'excède pas le temps de propagation du forçage amont à la limite aval du domaine.

Les modèles hydrologiques sont principalement focalisés sur la relation entre les hydrogrammes amonts et des points le long du cours d'eau à l'aval. Il existe une grande diversité de modèles hydrologiques conçus pour le routage d'une crue, nous invitons le lecteur à se reporter à la figure 1.4 pour voir les différents modèles. Ici la dénomination « modèle hydrologique » peut prêter à confusion puisque selon la communauté il peut s'agir des modèles pluie-débit.

Les modèles hydrauliques vont quant à eux du simple modèle d'onde de crue aux modèles basés sur la résolution des équations de Saint-Venant et demandent une connaissance assez avancée de la bathymétrie de la rivière et de la topographie des plaines d'inondation attenantes ainsi que des coefficients de frottement à appliquer. Il peut s'agir de modèles

1D ou 2D selon que l'écoulement est uni-directionnel ou pas.

Parmi les modèles hydrauliques les plus utilisés on peut citer LISFLOOD-FP (voir [Bates and De Roo, 2000, Durand et al., 2008]) qui modélise l'écoulement dans le lit mineur avec un modèle d'onde de crues 1D et dans le lit majeur avec l'équation de Manning. Le modèle est discrétisé avec un raster dérivé d'un modèle digital d'élévation de terrain. On peut citer d'autre part HEC-RAS (pour Hydrologic Engineering Center's River Analysis System), [Knebl et al., 2005] qui est un modèle résolvant les équations de Saint-Venant dans lequel le domaine de calcul est décomposé en sections en travers. Ce principe régit plusieurs autres codes de calcul tels les codes Mike-11 [Thompson et al., 2004] mais aussi Mascaret [Goutal and Maurel, 2002], voir section chapitre 2. Dans le cas des modèles 2D on citera notamment Telemac-2D [Hervouet, 2007] ou MIKE21 [Warren and Bach, 1992] qui résolvent les équations de Saint-Venant 2D avec un schéma différences finies (MIKE21) ou un schéma éléments finis ou volumes finis (Telemac-2D).

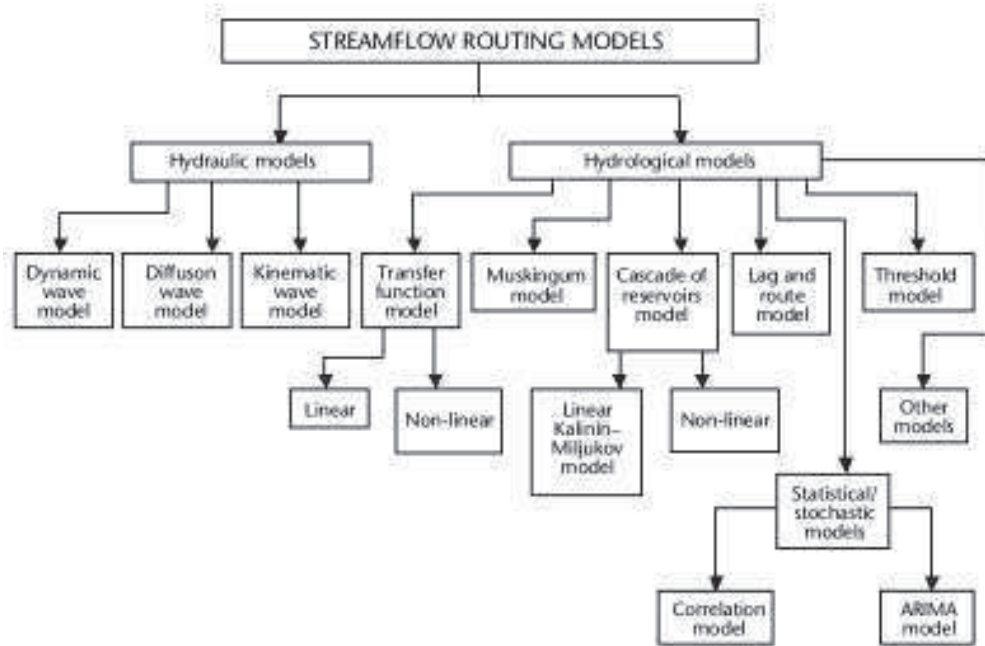


FIG. 1.4 – Les différents modèles de routage. Source : [World Meteorological Organization, 2011].

Les modèles pluie-débit

Comme leur nom l'indique ces modèles prennent en entrée des précipitations de pluie ou de neige et calculent les débits qui en résultent. Ces modèles vont du simple modèle empirique [Edijatno and Michel, 1989, Makhoulouf, 1994] à des modèles distribués hybrides physique/conceptuel. Les précipitations en entrée peuvent être des pluies observées avec

des pluviomètres ou avec des radars météo ou encore des satellites. Mais ce peut être également des sorties de modèles météorologiques. La finesse des prévisions de débits de ces modèles dépendra fortement de leur résolution spatiale. Les recherches actuelles tendent à affiner celle-ci en-deçà de deux kilomètres et elles tendent également à repousser l'échéance de prévision de ces modèles pour passer de quelques jours à 5 à 10 jours. Il apparaît alors clairement que la qualité des prévisions fournies par ces modèles dépend de la qualité des précipitations données et prévues en entrée. Néanmoins des techniques de correction en temps réel des modèles existent (voir [Tangara, 2005]), elles sont basées sur l'utilisation des débits observés jusqu'à l'origine temporelle de la prévision mais peuvent utiliser également d'autres variables comme les précipitations ou encore l'évapotranspiration. La combinaison de ces modèles avec des modèles débit-débit permet de repousser l'échéance de prévision au-delà de l'échéance de propagation de ces derniers (voir figure 1.3).

Parmi les modèles les plus utilisés par la communauté on peut citer notamment HEC-HMS (pour Hydrologic Engineering Center's Hydrologic Modeling System), [Knebl et al., 2005], développé par le corps des ingénieurs de l'armée américaine. HEC-HMS simule le processus de pluie-débit au niveau d'un bassin versant. La part de pluie s'infiltrant dans le sol et la part de pluie en excès transformée en débit sont calculées à partir d'un ensemble d'équations empiriques et utilisent l'algorithme ModClark [Bhattacharya et al., 2012]. On peut également citer le logiciel GR4J [Sauquet et al., 2007] qui utilise des données de pluviométrie et d'évapotranspiration sur tout le bassin pour estimer les débits à l'exutoire. Le calage des paramètres se fait grâce à une procédure automatique minimisant la différence entre les débits simulés et les observations. On peut citer enfin la chaîne ISBA-MODCOU [Habets et al., 2008]. ISBA (Interaction Sol-Biosphère-Atmosphère) simule l'ensemble des flux d'eau avec l'atmosphère (interception, évaporation, transpiration) et avec le sol (ruissellement des précipitations et drainage dans le sol). MODCOU est un modèle hydrologique distribué, il utilise en entrée le ruissellement d'ISBA pour calculer l'évolution des nappes et le débit des rivières.

Pour dépasser les limitations d'un seul modèle pluie-débit, dont les performances sont bonnes dans certains cas et moindres dans d'autres, il peut être avantageux d'utiliser plusieurs modèles dont les points faibles des uns sont contrebalancés par les points forts des autres. Les débits prévus seront alors l'agrégation des débits prévus par les différents modèles. On peut également établir des prévisions d'ensemble ou prévisions probabilistes. Celles-ci reposent sur la perturbation des conditions initiales ou des conditions aux limites d'un même modèle selon des observations ou des lois de probabilité. Elles permettent d'estimer la dépendance aux conditions initiales et/ou aux conditions aux limites (études de sensibilités) ainsi que l'erreur sur l'état du modèle ou les incertitudes liées à une prévision mais aussi d'estimer la probabilité d'un événement en considérant la fréquence relative de ses réalisations [Gouweleeuw et al., 2005].

Les modèles météorologiques

On ne rentrera pas dans le détail des modèles météorologiques ici néanmoins nous tenons à préciser que leur utilisation peut s'avérer utile lorsque l'on souhaite étendre l'échéance de prévision au-delà de celle donnée par l'enchaînement des modèles pluie-débit et débit-débit. Dans ce cas les prévisions hydrologiques et hydrauliques sont basées

sur des prévisions météorologiques et peuvent par conséquent être entachées d'erreurs importantes.

Les sources d'erreur dans les modèles et l'assimilation de données

Les modèles de prévision sont entachés d'erreurs. Celles-ci affectent toutes les facettes des modèles que ce soit le modèle en lui-même, les conditions initiales, les sorties... En effet un modèle, aussi sophistiqué soit-il, restera toujours une représentation simplifiée de la réalité et de fait sera toujours entaché d'erreur. En général l'erreur modèle peut être compensée par une calibration des paramètres mais alors ceux-ci peuvent avoir peu de sens physique.

La calibration [Krause et al., 2005, Bárdossy, 2006] est une étape importante de la conception d'un modèle, elle permet de réduire l'erreur modèle et les sources d'incertitude. La calibration consiste à ajuster les paramètres du modèle (par exemple les coefficients de Strickler dans un modèle hydraulique) de manière à ce que celui-ci minimise l'erreur par rapport aux observations. Ainsi le modèle simule et prédit bien les événements. En général pour cela on se base sur un jeu de crues de référence et on cherche la valeur des paramètres qui minimise certains scores comme la « root mean square error » (RMSE) ou le critère de Nash-Sutcliffe. La démarche ici pourra être une simple méthode d'essais et d'erreurs ou une méthode plus élaborée du type optimisation automatique de paramètres. Pour ce qui est de la validation d'un modèle en mode prévision la communauté hydrologique se base sur les scores classiques (RMSE, calcul du biais...) comparant sorties du modèle et observations et utilise aussi de plus en plus les critères développés dans le cadre de la météorologie [Brier, 1950, Laio and Tamea, 2006]. Ceux-ci consistent à calculer le score de Brier, la mesure du taux de fausses alertes... et permettent d'évaluer la pertinence d'un système de prévision en comparaison de ce qu'il s'est vraiment passé. Ils peuvent constituer également des outils de diagnostic des erreurs liées à la modélisation également appelées erreurs modèle.

A l'erreur modèle peut s'ajouter l'erreur sur les forçages. Par exemple dans le cas d'un modèle hydraulique forcé par un modèle pluie-débit l'erreur sur les forçages résulte de l'erreur de ce dernier. Parfois certains forçages sont simplement inconnus. Reprenons l'exemple d'un modèle hydraulique, on suppose que dans la réalité un affluent non instrumenté vient se jeter dans le cours d'eau modélisé. Lors d'épisodes de crue importants les hauteurs d'eau observées à une station en aval de la confluence en question pourront être très différents de ceux calculés avec le modèle du fait de cet apport latéral non jaugé. Pour les modèles basés sur une représentation physique il faut considérer en plus les erreurs relatives à la géométrie de la zone étudiée. Celles-ci peuvent être d'autant plus importantes que les incertitudes sont fortes sur la zone étudiée. En effet en pratique les données de bathymétrie et de topographie pour modéliser le lit d'une rivière ainsi que les plaines d'inondation attenantes sont incomplètes ou leur résolution spatiale trop faible. Une berge incorrectement modélisée pourra se traduire en la submersion de zones dans le modèle là où dans la réalité les zones en question n'ont pas été submergées ce qui peut s'avérer problématique dans les études prospectives.

L'erreur sur les conditions initiales peut affecter tous les types de modèle. Dans le cas des modèles hydrodynamiques l'erreur sur la condition initiale persistera autant de temps qu'il est nécessaire pour que la condition initiale soit évacuée. Dans le cas des modèles

pluie-débit une erreur sur la condition initiale pourra se répercuter jusqu'à la fin de la période de prévision qui peut être relativement longue. Comme on l'a vu pour ces modèles l'échéance de prévision peut aller jusqu'à plusieurs jours et cette erreur sera d'autant plus problématique si le modèle pluie-débit en question est combiné à l'aval à un modèle de routage puisque cette erreur sera ensuite transférée de celui-ci à celui-là...

Les erreurs sur les observations tiennent elles aussi une place importante dans l'utilisation des modèles. En effet les observations en ce qu'elles peuvent servir comme données de forçages pour les modèles sont également une source d'erreur. On estime par exemple (voir [?]) que des pluies moyennées spatialement peuvent avoir une erreur comprise entre 20 et 30 % tandis que les mesures de hauteurs d'eau et/ou de débit peuvent être biaisées ou affectées par un défaut de l'instrument. D'autre part des observations fausses (du fait d'erreur des instruments ou dans la chaîne d'acquisition) peuvent biaiser la calibration d'un modèle ou encore conduire à une mauvaise estimation des courbes de tarage.

Pour réduire l'impact que les erreurs et les sources d'incertitudes ont sur les résultats des modèles on peut appliquer les techniques de l'assimilation de données. Le lecteur intéressé pourra se reporter au chapitre 3 pour une présentation plus générale de l'assimilation de données. Pour un état de l'art des techniques d'assimilation de données appliquées à l'hydraulique fluviale nous invitons le lecteur à se reporter à la chapitre 6. Nous mentionnerons simplement ici que les techniques d'assimilation de données sont utilisées dans la prévision des crues en temps réel pour combiner les données d'observation aux modèles de prévision. Cela permet de corriger les variables d'état ou les paramètres du modèle en temps réel afin d'améliorer les prévisions.

1.2.3 Les alertes inondations et la transmission de l'information

Le but final d'un service de prévision des crues est d'alerter la communauté lorsqu'un risque de crue se présente, de manière à ce que les actions appropriées soient prises. Cela consiste à :

- informer le grand public du lieu et de l'heure de survenue de l'évènement ainsi que des dommages éventuels aux infrastructures de manière à ce que celui-ci puisse se préparer au mieux. Cela peut consister en la mise en place de sacs de sable comme protections, à surélever les meubles, rester confiné chez soi...
- informer les autorités (maires, préfets), les décideurs, les gestionnaires d'industries, d'ouvrages hydro-électriques...
- faire en sorte que les services d'urgence soient préparés lors de la survenue de l'évènement ;
- dans les cas les plus graves, préparer l'évacuation et appliquer les mesures d'urgence nécessaires.

Plus les alertes sont émises de manière précoce plus on peut minimiser l'impact global de la crue. Les échéances de prévision dépendent de plusieurs facteurs. Comme on l'a souligné plus haut l'échéance de prévision va dépendre du temps de réponse du cours d'eau. Celle-ci sera d'autant plus courte que le temps de réponse du cours d'eau est faible. C'est le cas par exemple des épisodes cévenols dans le sud de la France qui donnent lieu à des crues éclair dévastatrices. Pour mieux les prévoir une combinaison de modèle hydrologique/hydraulique avec une procédure d'assimilation de données ainsi que des abaques de vigilance des crues ont été mis en place [Borrell Estupina et al., 2014]. Idéalement l'échéance de prévision devrait être suffisamment longue pour que les mesures nécessaires

de protection puissent être prises. De ce point de vue les prévisions météorologiques peuvent fournir des informations importantes sur l'éventuelle survenue d'un épisode de crue à plus ou moins long terme (5 à 10 jours). Naturellement plus la prévision météo sera émise à court terme plus elle sera précise et plus elle pourra assister les autorités dans leurs prises de décisions. Ainsi même des prévisions émises à échéance de moins d'une journée peuvent permettre de prendre des mesures préventives de préparation à la crue. A plus courte échéance les radars météos peuvent fournir une information importante quant au développement et à l'évolution d'une zone d'intempéries. Les prévisions météos sont combinées à des prévisions hydrologiques de manière à augmenter l'échéance de prévision. Ces prévisions hydrologiques sont basées sur la projection dans le futur, grâce à des relations statistiques dans les cas les plus simples ou grâce à des modèles hydrauliques résolvant les équations de Saint-Venant, d'un état hydrologique observé à un instant donné.

Les prévisions sont ensuite comparées aux seuils de vigilance afin de déterminer quelles actions sont à prendre et si il y a lieu de donner l'alerte. En général les seuils d'alerte sont de deux types. On distingue d'une part les seuils liés aux quantités de précipitations. On distingue d'autre part les seuils relatifs aux niveaux d'eau. Par exemple si les hauteurs d'eau mesurées ou les variations de hauteurs d'eau dépassent un certain seuil par exemple 25 *cm* en une heure. Bien sûr les différents seuils d'alerte différeront selon les enjeux des zones concernées, ils ne seront pas les mêmes dans une zone résidentielle et dans une plaine d'inondation où il n'y a aucune activité humaine. En général les différents seuils d'alerte sont définis en fonction du niveau de vigilance requis : vigilance crue, risque de crue, risque de crue sévère. Une illustration schématique de différents seuils d'alerte en fonction de la hauteur d'eau est donnée en figure 1.5.

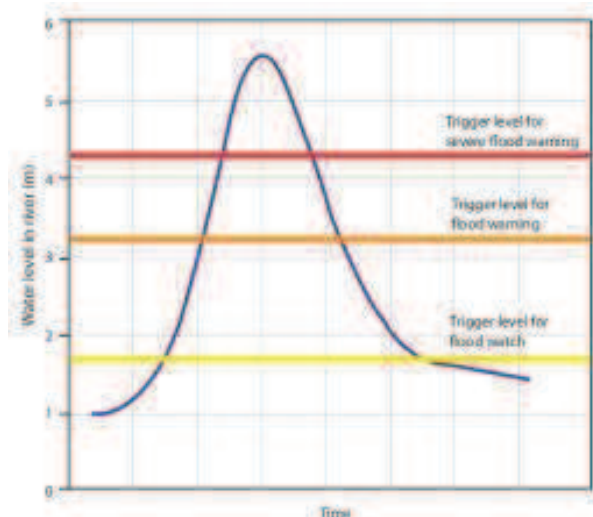


FIG. 1.5 – Illustration de différents niveaux d'alerte en fonction du niveau d'eau.
Source : [World Meteorological Organization, 2011].

Lorsqu'un niveau d'alerte est atteint l'information doit être diffusée et accessible au plus grand nombre. De fait elle doit être émise dans un langage clair et précis. Par exemple les inondations qui se sont produites en Angleterre et au Pays de Galles en 1998 ont donné lieu à un rapport mettant en lumière le fait que les symboles alors utilisés pour représenter le degré de gravité des inondations avaient été mal compris par les gens qui avaient reçu ces informations. Ces symboles ont depuis lors été remplacés par d'autres symboles montrant une petite maison plus ou moins submergée en fonction de la gravité potentielle de l'évènement. Ces symboles sont accompagnés de brefs messages décrivant les enjeux de la crue du type de celui-ci :

« ***Severe flood warning*** : *This is the warning issued when serious flooding is expected and there is imminent danger to life and property. If your warning is upgraded to this you should be prepared for your gas, electricity, water and telephone supplies being lost. You're advised to keep calm and reassure others, and cooperate with the emergency services.* »

De plus ces messages ont donné lieu à d'importantes campagnes d'information de manière à en instruire le plus efficacement le grand public. D'autre part l'utilisation d'Internet et des nouvelles technologies de l'information peut s'avérer particulièrement utile pour la dissémination de l'information. Néanmoins certaines catégories de la population comme les plus pauvres ou les personnes âgées peuvent ne pas avoir la même facilité d'accès à ces technologies. D'autant plus que lors d'épisodes de crues dévastateurs les moyens de communication peuvent être rendus inopérants du fait des dégâts occasionnés. Ainsi malgré l'apport des avancées technologiques il conviendrait de ne pas renoncer à des systèmes d'alertes plus conventionnels comme des alarmes de crue.

Selon les moyens mis à leur disposition et selon le degré de sophistication recherché les services de prévision des crues peuvent établir différents types d'alerte ou de prévision. Cela va dépendre fortement des enjeux locaux. Par exemple dans le cas d'un cours d'eau qui n'est que très peu sujet à des épisodes d'inondation des modèles hydrologiques/hydrauliques ne seront pas nécessairement développés pour suivre l'évolution des crues. Celle-ci pourra être extrapolée de manière statistique à partir des données d'observation et des archives hydrologiques dont on dispose sur le cours d'eau. Au degré de sophistication au dessus le service de prévision des crues pourra émettre des cartes de vigilance en fonction du risque de crue selon les tronçons. C'est par exemple ce qui est fait en France où une carte de vigilance est émise deux fois par jour en fonction du risque d'inondation estimé 24 heures à l'avance sur chaque tronçon. Cette carte de vigilance est accompagnée d'un état de la situation hydrologique sur le territoire ainsi que d'un commentaire sur l'évolution prévue de la situation. Aux Etats-Unis des bulletins d'information précisant la localisation et la nature du risque (crue éclair sur tel cours d'eau) accompagnés d'informations décrivant le phénomène et précisant les risques associées sont émis les services de prévision. Dans les cas les plus aboutis les services de prévision des crues pourront accompagner ces informations de cartes de risque d'inondation montrant les zones à risque comme cela se fait par exemple en Angleterre.

1.3 La prévision des crues en France

1.3.1 Les services de prévision des crues en France

En France c'est le réseau pour la prévision des crues et l'hydrométrie (le réseau PC&H) qui est en charge de la production et la diffusion de la vigilance crues et des prévisions associées, dispositif opérationnel depuis le mois de Juillet 2006. Cela s'est fait sous l'impulsion du Ministère de l'écologie et du développement durable qui a souhaité concrétiser l'évolution de l'annonce des crues vers la prévision des crues. Le réseau PC&H regroupe le service central d'hydrométéorologie et d'appui à la prévision des inondations (SCHAPI) créé en 2003, les services de prévision des crues (SPC), les unités d'hydrométrie (UH) et bientôt les cellules de veille hydrologique (CVH) en Corse et dans les DOM.

La vigilance crues permet de prévenir les autorités et le public qu'il existe un risque de crues à échéance de 24 heures, plus ou moins important selon la couleur de vigilance. Il existe quatre couleurs représentant les différents niveaux de vigilance allant de vert (pas de vigilance particulière) à rouge (risque de crue majeur) en passant par jaune et orange. Ainsi la vigilance crue est-elle en cohérence avec la vigilance météorologique mise en place par Météo-France depuis 2001, et qui couvre l'ensemble des risques d'origine météorologique, en les localisant par département. Les deux vigilances permettent aux autorités locales (notamment préfets et maires), ainsi qu'au public, de se mettre en situation de réagir de manière appropriée si le danger se précise et de gérer la situation dans les meilleures conditions.



FIG. 1.6 – Exemple de carte de vigilance émise par le SCHAPI le 17 Juin 2013 à 16h.

La vigilance crue localise le niveau de vigilance par tronçons ou ensembles de cours d'eau du réseau hydrographique surveillé par l'Etat (plus de 21000 km de cours d'eau sur le territoire métropolitain continental) aux abords desquels se situent plus des trois quarts des personnes habitant ou travaillant en zone inondable. Cela se matérialise par l'émission bi-journalière de cartes de vigilance sur lesquelles sont reportées les niveaux de vigilance pour la majorité des tronçons et cours d'eau évoqués ci-dessus. Cette vigilance est accompagnée d'informations complémentaires :

- les niveaux d'eau et les débits des cours d'eau observés en temps réel sur près de 1500 points de mesure télétransmis, ainsi que leurs variations au cours des derniers jours et des dernières heures ;
- des informations sur les évolutions prévues de niveaux d'eau et de débit, de façon de plus en plus systématique et pour des échéances pouvant aller au-delà de 24 heures.

Ces informations sont centralisées sur le site www.vigicrues.gouv.fr. Il permet de diffuser directement cette information, qui est aussi relayée par Météo-France dans le cadre de la vigilance météorologique au travers de deux volets : « pluie-inondation » et « inondation (sans pluie simultanée) », co-produits avec le SCHAPI.

En dehors de cette mission de vigilance le réseau PC&H a aussi des activités scientifiques et techniques comme le développement de nouveaux modèles hydrologiques/hydrauliques pour la prévision des crues, une meilleure anticipation des crues soudaines et une meilleure prévision des enveloppes de zones inondées potentielles. Elle assure également le maintien en service de la base nationale des données hydrométriques, Banque Hydro (www.hydro.eaufrance.fr) ainsi que la mise à disposition de ses données, 365 jours par an, avec un appui à leur utilisation, pour l'ensemble du réseau mais également les gestionnaires de la ressource en eau et des milieux aquatiques, les chercheurs et le grand public.

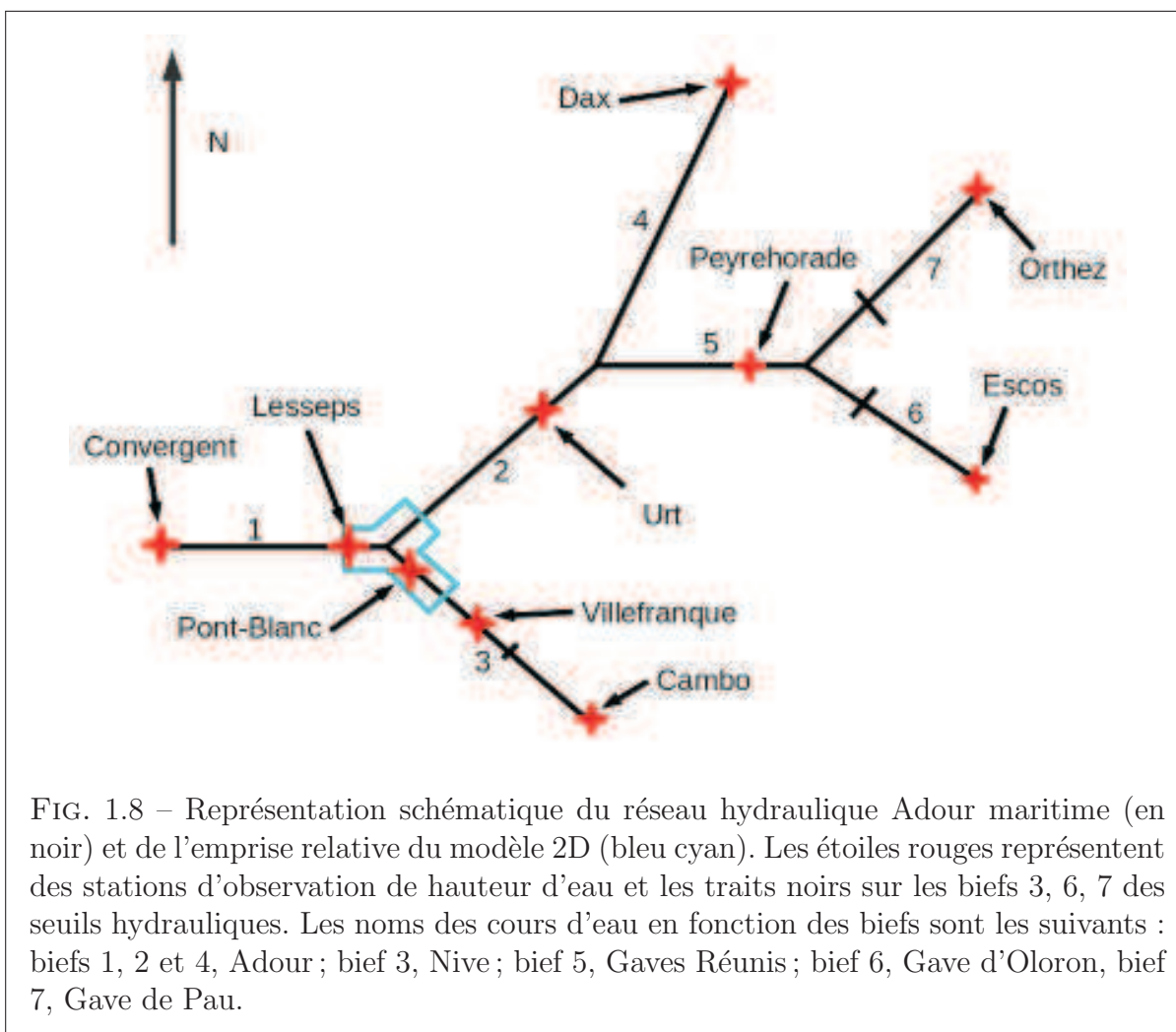
Parmi les nombreux SPC français nous avons collaboré pendant ce travail avec le SPC GAD (Gironde-Adour-Dordogne) et les méthodes numériques que nous avons développées ont été appliquées au réseau hydraulique Adour maritime. C'est pourquoi dans la prochaine section nous présentons le bassin versant de l'Adour et le réseau hydraulique Adour maritime.

1.3.2 Le bassin versant de l'Adour et le réseau hydraulique Adour maritime

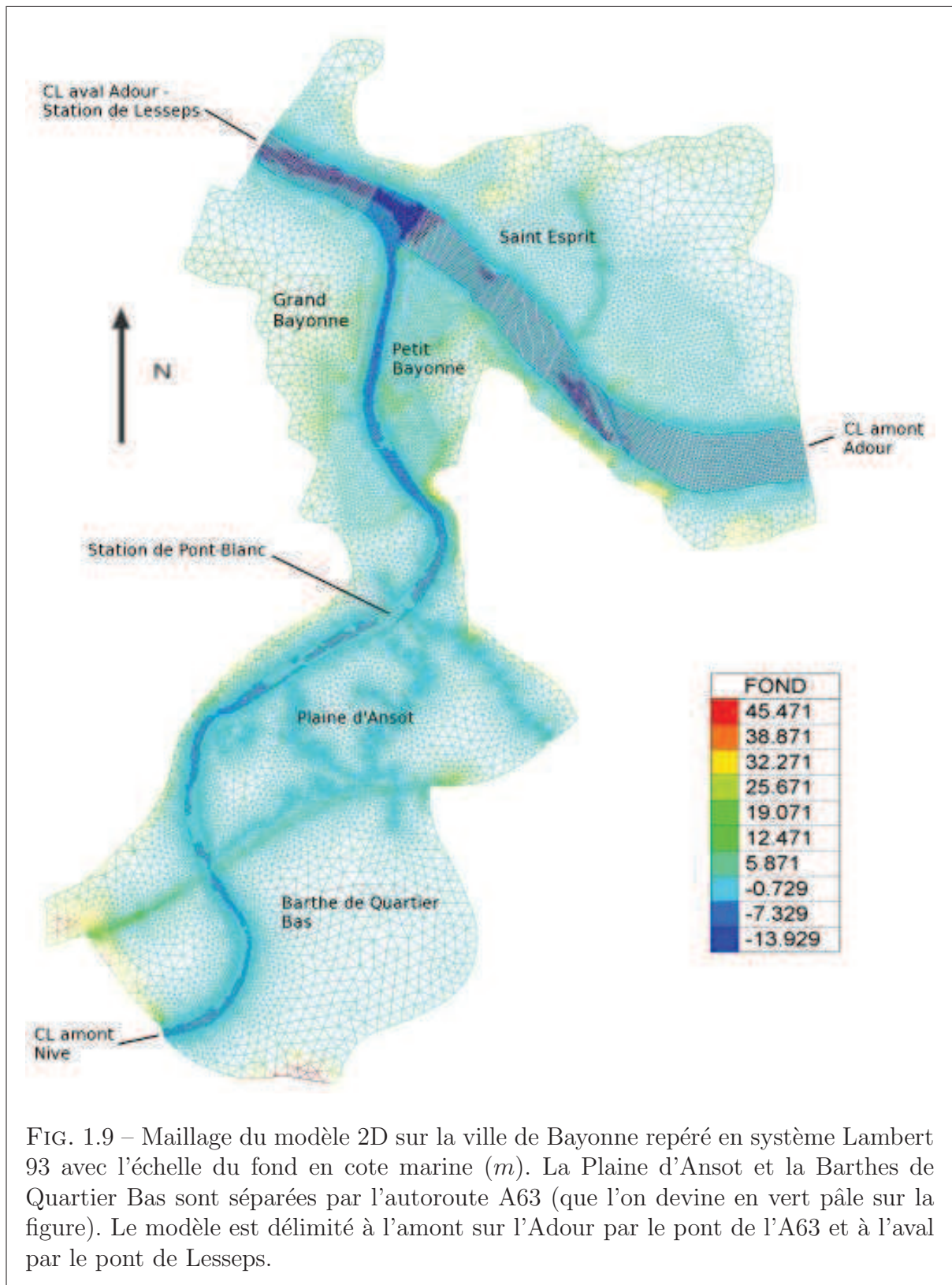
Le bassin versant de l'Adour est situé dans le Sud-Ouest de la France et s'étend du piémont pyrénéen à la côte aquitaine. Sa zone de drainage couvre une surface de 16 890 km^2 et recouvre les départements des Pyrénées Atlantiques et des Landes. L'Adour prend sa source dans les Pyrénées à l'altitude de 2600 m et se jette dans l'océan Atlantique 312 km à l'aval, à Bayonne. Le bassin versant de l'Adour est l'un des plus humides de France du fait des fortes précipitations dans sa partie amont. Le bassin versant de l'Adour se décompose en deux parties. La première située à l'aval et sous influence maritime et la seconde située à l'amont et qui est sous l'influence principalement des affluents. L'Adour a trois affluents principaux (responsables de 65% du débit total à Bayonne durant les périodes de fortes crues). On distingue d'un côté les Gave de Pau et d'Oloron drainant respectivement une surface de 5226 km^2 et 608 km^2 et qui sont souvent affectés par des

« modèle Adour maritime ». Une représentation schématique de ce modèle est donnée en figure 1.8. Celui-ci recouvre environ 160 km de cours d'eau et est composé de 7 biefs avec 3 confluences. Tout le modèle Adour maritime est sous influence maritime à l'exception des portions situées à l'amont des barrages situés sur les biefs 3, 6 et 7 et symbolisés sur la figure 1.8 par un trait noir. Les forçages à l'amont (aux stations de Dax, Orthez, Escos et Cambo) sont donnés par des hauteurs d'eau converties en débits grâce à des courbes de tarage. Le forçage à l'aval est donné par les hauteurs d'eau observées à la station de Convergent. On aurait pu également prescrire des hauteurs d'eau à l'amont et des débits à l'aval mais comme la station de Convergent est sous influence maritime on ne dispose pas de courbes de tarage permettant de convertir des hauteurs d'eau en débits. En opérationnel les hauteurs d'eau en prévision à Convergent sont fournies par le SHOM (Service Hydrographique et Océanographique de la Marine). On dispose sur ce réseau de 5 stations d'observation de hauteur d'eau (et pas de débit) situées à Peyrehorade, Urt, Lesseps, Pont-Blanc et Villefranque.

Le réseau hydraulique « Adour maritime » fait l'objet d'importants travaux de recherches orientés vers l'opérationnel et dont ce travail de thèse est l'une des facettes. Ces travaux ont pris la forme de collaborations formalisées à travers des conventions entre le CERFACS et le SCHAPI. Un des aboutissements de cette collaboration a été le développement et la mise en place à l'opérationnel sur ce réseau et sur le réseau hydraulique Marne amont [Ricci et al., 2011] d'un algorithme d'assimilation de données permettant la correction en temps réel des forçages amonts ainsi que de l'état hydraulique du cours d'eau améliorant les prévisions à moyen terme (6h-12h). Ce résultat est intéressant car il permet d'améliorer les prévisions au-delà de l'échéance égale au temps nécessaire au forçage amont pour se propager jusqu'à la station où l'on effectue la prévision, ce qui est une des limitations des modèles hydrauliques [Matgen et al., 2010].



Sur la zone de Bayonne un modèle 2D (utilisant le code de calcul Telemac-2D) a également été développé par le SCHAPI et le SPC GAD. Cette zone se décompose en deux parties bien identifiées. On distingue d’une part le secteur urbain du centre ville avec les quartiers du Petit Bayonne et Saint-Esprit qui présentent des enjeux importants et qui peuvent être touchés lors des fortes crues de la Nive. On distingue d’autre part la zone de la Plaine d’Ansot et la Barthes de Quartier Bas au sud de Bayonne qui peuvent jouer un rôle important de laminage des fortes crues de la Nive. Ainsi le modèle 2D sur Bayonne a-t-il été mis en place pour modéliser les phénomènes 2D (horizontaux) sur la zone qu’un modèle 1D ne peut par nature pas modéliser. Cela permet de représenter les phénomènes 2D au niveau de la confluence Nive/Adour, l’extension spatiale des inondations dans le centre ville et dans la Plaine d’Ansot mais aussi de comprendre le remplissage et la vidange de celle-ci. Une image du maillage de ce modèle 2D est donnée en figure 1.9. L’emprise relative de celui-ci par rapport au modèle 1D est donnée en figure 1.8.



Comme nous le verrons dans le chapitre 9, dans le cadre de cette thèse, le modèle 2D a été couplé au modèle 1D sur la zone de Bayonne de manière à disposer sur l'ensemble de

la zone Adour maritime/Bayonne d'un modèle numérique complet répondant aux besoins des services de prévision des crues.

Les épisodes de crue sur le réseau Adour maritime peuvent prendre plusieurs formes, on en distingue trois en particulier qui peuvent néanmoins se combiner au cours d'un même épisode (voir données de forçage en annexe C) :

- crue éclair sur les gaves de Pau et gave d'Oloron. Dans ce cas les stations d'Escos et d'Orthez connaissent une forte montée des débits pouvant atteindre respectivement $1800 \text{ m}^3.\text{s}^{-1}$ et $1200 \text{ m}^3.\text{s}^{-1}$ suivie d'une crue presque aussi rapide. En général les forçages à ces deux stations sont fortement corrélés ;
- crue éclair sur la Nive avec là encore une forte augmentation des débits à la station de Cambo pouvant atteindre $600 \text{ m}^3.\text{s}^{-1}$. Ces crues de la Nive peuvent être aggravées par un fort coefficient de marée. De plus, souvent (mais ce n'est pas toujours le cas), les débits à Cambo sont corrélés aux débits des stations d'Orthez et Escos ;
- crue fluviale sur l'Adour avec une lente augmentation des débits au niveau de la station de Dax suivie d'un plateau, le pic de crue, pouvant s'étendre sur plusieurs jours suivie d'une lente décrue.

On donne en figure 1.10 une illustration des forçages à l'amont du modèle pour l'épisode de Janvier 2014. Pour saisir l'ordre de grandeur de cet épisode il faut noter qu'en général les forçages à Dax, Orthez et Escos sont compris entre 50 et $100 \text{ m}^3.\text{s}^{-1}$ et ceux à Cambo sont de l'ordre de $25 \text{ m}^3.\text{s}^{-1}$. Toutefois tous les épisodes de crue sur l'Adour maritime n'ont pas cette violence.

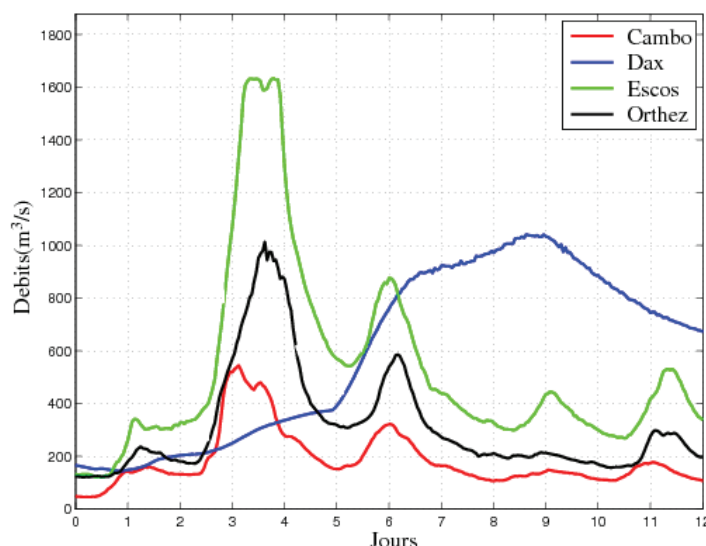


FIG. 1.10 – Illustration des forçages à l'amont du modèle Adour maritime, épisode de Janvier 2014.

Etant donné la nature campagnarde de la région les crues sur l'Adour maritime donnent lieu à l'inondation des champs, des barthes et des plaines d'inondation à proximité du cours d'eau. Voir à ce titre les images fournies par le SPC GAD en annexe (A) concernant la crue du 19 Juin 2013 à Peyrehorade, les débits aux stations d'Orthez et d'Escos avaient alors atteint entre 1000 et 1100 $m^3.s^{-1}$ chacun. Néanmoins les villes sur le réseau Adour maritime ne sont pas épargnées et notamment la ville de Dax en cas de fortes crues de l'Adour comme cela a été le cas par exemple en février 1952 où les niveaux d'eau avaient atteint 6,52 m (à l'échelle) alors qu'en temps normal les hauteurs à Dax sont plus proches de 0 m ; voir la figure 1.11-(a) ou plus récemment en Janvier 2014, voir la figure 1.11-(b). La ville de Bayonne quant à elle peut être touchée lors de crues de la Nive, notamment sur les quais du quartier Petit Bayonne comme cela est illustré par la figure 1.12-(a), ou sur les quais de l'Adour au niveau du quai de Lesseps, figure 1.12-(b).



(a)



(b)

FIG. 1.11 – (a) Photo aérienne de la crue de l'Adour à Dax, Février 1952. (b) Crue de l'Adour à Dax, Janvier 2014.



(a)



(b)

FIG. 1.12 – Episodes de crue à Bayonne, (a) quais de la Nive, quartier Saint-Espirit, (b) quai de Lesseps.

Deuxième partie

Méthodologie

Chapitre 2

Les équations de l'hydrodynamique à surface libre et les codes de calcul

Sommaire

2.1	Les équations de Navier-Stokes	34
2.1.1	Equation de conservation de la masse	34
2.1.2	Equation de conservation de la quantité de mouvement	34
2.1.3	Cas d'un fluide newtonien et incompressible	35
2.1.4	Conditions aux limites	36
2.2	Dérivation des équations de Saint-Venant	38
2.2.1	Adimensionnalisation des équations de Navier-Stokes	38
2.2.2	Passage à la limite des milieux peu profonds	39
2.2.3	Equations de Saint-Venant	41
2.3	Les modèles d'ondes de crues	42
2.3.1	Le modèle de l'approximation des ondes de crues	43
2.3.2	Le modèle de l'approximation des ondes de crues diffusantes	43
2.4	Principe de résolution des équations de Saint-Venant avec le code Mascaret	44
2.4.1	Forme conservative des équations de Saint-Venant 1D	44
2.4.2	Résolution numérique des équations de Saint-Venant avec une méthode volumes finis	46
2.5	Principe de résolution des équations de Saint-Venant avec le code Telemac-2D	51
2.5.1	Forme non-conservative des équations de Saint-Venant	51
2.5.2	La méthode des pas fractionnaires	52
2.5.3	Traitement de la propagation, de la diffusion et des termes sources – Discrétisation en temps et en espace	53
2.5.4	Formulation faible, approximation interne et mise sous forme d'un système linéaire	55

2.1 Les équations de Navier-Stokes

Les équations de Navier-Stokes sont les équations de base de la mécanique des fluides. Comme leur nom l'indique en partie, elles ont été établies par le mathématicien Henri Navier et le physicien Georges Gabriel Stokes. Elles sont couramment utilisées en sciences, en ingénierie, en aérodynamique... pour modéliser l'océan, l'atmosphère, l'écoulement d'une rivière, les écoulements industriels... Elles sont composées de trois équations : une équation de continuité ou équation de conservation de la masse, une équation de conservation de la quantité de mouvement et une équation de conservation de l'énergie. Nous ne nous intéresserons pas dans ce travail à cette dernière équation car les écoulements que nous considérons sont incompressibles et que nous ne nous intéressons pas à leur température. Pour l'établissement des équations de Navier-Stokes le lecteur intéressé pourra se reporter [?, ?].

2.1.1 Equation de conservation de la masse

Si l'on note ρ la masse volumique du fluide considéré et $\vec{U} = (u, v, w)$ le vecteur vitesse, alors la conservation de la masse d'un volume Ω de fluide que l'on suit dans son mouvement peut se traduire mathématiquement par l'égalité suivante :

$$\frac{d}{dt} \left(\int_{\Omega} \rho d\Omega \right) = 0 \quad (2.1)$$

D'après la règle de Leibnitz, le membre de gauche de l'égalité (2.1) se décompose en l'intégrale sur le volume Ω des variations de la masse volumique et le flux aux frontières du volume Ω (notées Γ), soit :

$$\frac{d}{dt} \left(\int_{\Omega} \rho d\Omega \right) = \int_{\Omega} \frac{\partial \rho}{\partial t} d\Omega + \int_{\Gamma} \rho \vec{U} \cdot \vec{n} d\Gamma \quad (2.2)$$

où $\vec{n}$ est le vecteur normal extérieur à Ω .

Or, d'après la formule de Green-Ostrogradsky, le second terme du membre de droite de l'égalité (2.2) peut s'écrire :

$$\int_{\Gamma} \rho \vec{U} \cdot \vec{n} d\Gamma = \int_{\Omega} \operatorname{div}(\rho \vec{U}) d\Omega \quad (2.3)$$

Le volume de fluide Ω étant quelconque, on déduit des égalités (2.1), (2.2) et (2.3) l'équation de conservation de la masse :

$$\boxed{\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \vec{U}) = 0} \quad (2.4)$$

2.1.2 Equation de conservation de la quantité de mouvement

L'équation de conservation de la quantité de mouvement est dérivée du principe fondamental de la dynamique qui stipule que la somme $\vec{F}$ des forces s'exerçant sur un solide est égale au produit de l'accélération du solide et de sa masse m , soit : $\vec{F} = m \frac{d\vec{U}}{dt}$. Dans le cas d'un fluide où la masse est susceptible de varier il faut écrire :

$$\vec{F} = \frac{d}{dt} \left(\int_{\Omega} \rho \vec{U} d\Omega \right) \quad (2.5)$$

Comme précédemment, le terme de droite de l'équation (2.5) peut se décomposer en une somme en utilisant la règle de Leibnitz :

$$\frac{d}{dt} \left(\int_{\Omega} \rho \vec{U} d\Omega \right) = \int_{\Omega} \frac{\partial(\rho \vec{U})}{\partial t} d\Omega + \int_{\Gamma} \rho \vec{U} \vec{U} \cdot \vec{n} d\Gamma \quad (2.6)$$

On peut réécrire le second terme du membre de droite de l'équation (2.6) :

$$\int_{\Gamma} \rho \vec{U} \vec{U} \cdot \vec{n} d\Gamma = \int_{\Omega} \operatorname{div}(\rho \vec{U} \otimes \vec{U}) d\Omega$$

où $\otimes$ désigne le produit tensoriel de deux tenseurs d'ordre 1.

D'autre part, la somme des forces $\vec{F}$ s'appliquant au fluide peut se décomposer en deux types de force :

- les forces volumiques $\int_{\Omega} \rho \vec{f} d\Omega$ où $\vec{f}$ peut désigner l'accélération de la pesanteur, la force de Coriolis...
- les forces de surface $\int_{\Gamma} \vec{\tau} d\Gamma$ où $\vec{\tau}$ désigne les contraintes surfaciques.

Les contraintes surfaciques $\vec{\tau}$ s'écrivent sous la forme $\vec{\tau} = \sigma \cdot \vec{n}$ où σ est le tenseur des contraintes. D'après le théorème de Green-Ostrogradsky les forces de surface peuvent s'écrire :

$$\int_{\Gamma} \vec{\tau} d\Gamma = \int_{\Omega} \operatorname{div}(\sigma) d\Omega$$

On peut réécrire les équations de conservation de la quantité de mouvement :

$$\int_{\Omega} \frac{\partial(\rho \vec{U})}{\partial t} + \operatorname{div}(\rho \vec{U} \otimes \vec{U}) d\Omega = \int_{\Omega} \rho \vec{f} d\Omega + \int_{\Omega} \operatorname{div}(\sigma) d\Omega$$

Le volume Ω étant quelconque, on en déduit **l'équation de conservation de la quantité de mouvement** :

$$\boxed{\frac{\partial(\rho \vec{U})}{\partial t} + \operatorname{div}(\rho \vec{U} \otimes \vec{U}) = \rho \vec{f} + \operatorname{div}(\sigma)}$$

2.1.3 Cas d'un fluide newtonien et incompressible

Dans la suite nous supposons que le fluide que l'on considère est newtonien et incompressible. Dans ce cas le tenseur des contraintes s'écrit sous la forme :

$$\sigma = -pId + 2\mu D(\vec{U}) \quad (2.7)$$

où p désigne la pression exprimée en Pascal, μ le coefficient de viscosité dynamique et $D(\vec{U})$ le tenseur des taux de déformation :

$$D(\vec{U}) = \frac{1}{2} (\nabla \vec{U} + \nabla \vec{U}^T)$$

D'autre part dans le cas incompressible l'équation (2.4) se réduit à :

$$\operatorname{div}(\vec{U}) = 0$$

Ainsi les équations de Navier-Stokes dans le cas d'un fluide newtonien et incompressible sont données par le système :

$$\left\{ \begin{array}{l} \operatorname{div}(\vec{U}) = 0 \\ \frac{\partial \vec{U}}{\partial t} + \operatorname{div}(\vec{U} \otimes \vec{U}) = \frac{1}{\rho} \operatorname{div}(\sigma) + \vec{f} \end{array} \right. \quad (2.8)$$

On remarque que :

$$\operatorname{div}(\vec{U} \otimes \vec{U}) = (\vec{U} \cdot \overrightarrow{\operatorname{grad}}) \vec{U} \quad (2.9)$$

$$\frac{1}{\rho} \operatorname{div}(\sigma) = \nu \Delta \vec{U} - \frac{1}{\rho} \overrightarrow{\operatorname{grad}} p \quad (2.10)$$

où $\nu = \frac{\mu}{\rho}$ est le coefficient de viscosité cinématique (m^2/s) et où l'opérateur différentiel $\vec{U} \cdot \overrightarrow{\operatorname{grad}}$ est l'opérateur d'advection. Alors le système (2.8) est équivalent au système suivant :

$$\left\{ \begin{array}{l} \operatorname{div}(\vec{U}) = 0 \\ \frac{\partial \vec{U}}{\partial t} + (\vec{U} \cdot \overrightarrow{\operatorname{grad}}) \vec{U} = \nu \Delta \vec{U} - \frac{1}{\rho} \overrightarrow{\operatorname{grad}} p + \vec{f} \end{array} \right.$$

Si on note $\vec{f} = (f_x, f_y, f_z)$, alors les équations de Navier-Stokes, dans le cas newtonien et incompressible, peuvent s'écrire en coordonnées cartésiennes :

$$\left\{ \begin{array}{l} \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \\ \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} = \nu \Delta u - \frac{1}{\rho} \frac{\partial p}{\partial x} + f_x \\ \frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} = \nu \Delta v - \frac{1}{\rho} \frac{\partial p}{\partial y} + f_y \\ \frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + v \frac{\partial w}{\partial y} + w \frac{\partial w}{\partial z} = \nu \Delta w - \frac{1}{\rho} \frac{\partial p}{\partial z} + f_z \end{array} \right. \quad (2.11)$$

2.1.4 Conditions aux limites

Nous nous plaçons dans ce travail dans le cas d'un fluide à surface libre seulement limité par cette surface et par le fond sur lequel il repose. Nous ne considérerons donc pas le cas de frontières ouvertes ou de parois verticales.

On considère dans un premier temps les conditions aux limites sur la surface libre puis, dans un second, les conditions aux limites sur le fond.

Conditions aux limites sur la surface libre

On établit tout d'abord la condition aux limites cinématique au niveau de la surface libre. Pour cela on considère une particule quelconque de fluide à la surface libre de Ω de coordonnées (x, y, z) et on suppose qu'elle vérifie l'équation :

$$z = \psi_s(x, y, t) \quad (2.12)$$

où ψ_s est une fonction continue. L'équation (2.12) rend impossible les configurations de type déferlement.

Si Φ est définie par $\Phi(x, y, z, t) = z - \psi_s(x, y, t)$ la condition aux limites s'écrit $\frac{d\Phi}{dt} = 0$ pour $\Phi = 0$. Si on note $\vec{U}_s = (u, v, w)$ le vecteur vitesse de la particule x alors on a l'égalité :

$$w - \frac{\partial \psi_s}{\partial t} - u \frac{\partial \psi_s}{\partial x} - v \frac{\partial \psi_s}{\partial y} = 0 \quad (2.13)$$

L'équation (2.13) est équivalente à l'équation suivante qui exprime **la condition aux limites cinématique** à la surface libre du fluide :

$$\boxed{\frac{\partial \psi_s}{\partial t} - \vec{U}_s \cdot \vec{n}_s = 0} \quad (2.14)$$

où $\vec{n}_s = \overrightarrow{\text{grad } \Phi}$ est le vecteur normal (mais non normé) extérieur à la surface libre.

La condition aux limites dynamique (ou continuité des contraintes de surface) s'exprime par la relation :

$$\boxed{\sigma \cdot \vec{n}_s = -p_a \vec{n}_s}$$

où σ est le tenseur des contraintes turbulentes défini en (2.7) et où p_a est la pression atmosphérique.

Conditions aux limites sur le fond

Sur le fond, on considère que la composante verticale de la vitesse est nulle et on impose les conditions aux limites :

$$\boxed{w = 0 \quad \text{et} \quad \begin{cases} \vec{e}_x \cdot \sigma \cdot \vec{e}_z = \tau_x \\ \vec{e}_y \cdot \sigma \cdot \vec{e}_z = \tau_y \end{cases} \quad \text{pour } z = \psi_f(x, y, t)}$$

où $\vec{\tau} = (\tau_x, \tau_y)$ représente la contrainte de cisaillement exercée par le fluide sur la paroi et où ψ_f est une fonction continue.

Sur le fond on a une condition aux limites d'imperméabilité du fond similaire à celles établies pour la surface libre en (2.13) et (2.14) :

$$w - \frac{\partial \psi_f}{\partial t} - u \frac{\partial \psi_f}{\partial x} - v \frac{\partial \psi_f}{\partial y} = 0 \quad (2.15)$$

$$\boxed{\frac{\partial \psi_f}{\partial t} - \vec{U}_f \cdot \vec{n}_f = 0} \quad (2.16)$$

où $\vec{n}_f$ désigne la normale extérieure au fond et où $\vec{U}_f$ désigne les composantes horizontales de la vitesse, i-e $\vec{U}_f = (u, v)$.

2.2 Dérivation des équations de Saint-Venant

Pour dériver les équations de Saint-Venant des équations de Navier-Stokes nous allons adimensionnaliser le système d'équations (2.11). L'adimensionnalisation du système (2.11) permet d'identifier les termes que l'on va pouvoir supprimer sous l'hypothèse des milieux peu profonds. En cela nous allons suivre la démarche présentée dans [Thual, 2010]. Le lecteur intéressé par une autre démarche de dérivation des équations de Saint-Venant pourra également se reporter à [Hervouet, 2007, Perthame and Gerbeau, 2001].

2.2.1 Adimensionnalisation des équations de Navier-Stokes

Nous supposons que la seule force s'appliquant au fluide est la gravité, autrement dit : $\vec{f} = (0, 0, -g)$, où g représente l'accélération de la gravité.

Nous définissons tout d'abord le rapport d'aspect ε comme étant le quotient des grandeurs caractéristiques verticales H_0 et horizontales L_0 , soit : $\varepsilon = \frac{H_0}{L_0}$. Nous définissons ensuite les variables et paramètres sans dimension suivants :

$$(x, y) = L_0(x^*, y^*) \quad z = H_0 z^* \quad t = \frac{L_0}{U_0} t^*$$

$$(u, v) = U_0(u^*, v^*) \quad w = \varepsilon U_0 w^* \quad p = \rho g H_0 p^*$$

où U_0 est une vitesse caractéristique.

Alors on peut montrer que les équations de Navier-Stokes adimensionnalisées s'écrivent :

$$\left\{ \begin{array}{l} \frac{\partial u^*}{\partial x^*} + \frac{\partial v^*}{\partial y^*} + \frac{\partial w^*}{\partial z^*} = 0 \\ \frac{\partial u^*}{\partial t^*} + u^* \frac{\partial u^*}{\partial x^*} + v^* \frac{\partial u^*}{\partial y^*} + w^* \frac{\partial u^*}{\partial z^*} = \frac{1}{\varepsilon R t} \Delta^* u^* - \frac{1}{F r^2} \frac{\partial p^*}{\partial x^*} \\ \frac{\partial v^*}{\partial t^*} + u^* \frac{\partial v^*}{\partial x^*} + v^* \frac{\partial v^*}{\partial y^*} + w^* \frac{\partial v^*}{\partial z^*} = \frac{1}{\varepsilon R t} \Delta^* v^* - \frac{1}{F r^2} \frac{\partial p^*}{\partial y^*} \\ \varepsilon^2 \left(\frac{\partial w^*}{\partial t^*} + u^* \frac{\partial w^*}{\partial x^*} + v^* \frac{\partial w^*}{\partial y^*} + w^* \frac{\partial w^*}{\partial z^*} \right) = \frac{\varepsilon}{R t} \Delta^* w^* - \frac{1}{F r^2} \frac{\partial p^*}{\partial z^*} - \frac{1}{F r^2} \end{array} \right. \quad (2.17)$$

où $\Delta^* = \varepsilon^2 \left(\frac{\partial^2}{\partial x^{*2}} + \frac{\partial^2}{\partial y^{*2}} \right) + \frac{\partial^2}{\partial z^{*2}}$ et où $Rt = \frac{H_0 U_0}{\nu}$ et $Fr = \frac{U_0}{\sqrt{g H_0}}$ sont les nombres sans dimension de Reynolds et de Froude respectivement.

Conditions aux limites

A la surface libre, on peut montrer que la condition aux limites dynamique (ou continuité des contraintes de surface) se traduit par le système :

$$\begin{cases} \frac{1}{Rt} \frac{\partial u^*}{\partial z^*} + \frac{\varepsilon}{Fr^2} (p^* - p_a^*) \frac{\partial \psi_s^*}{\partial x^*} + \frac{\varepsilon^2}{Rt} \left(\frac{\partial w^*}{\partial x^*} - 2 \frac{\partial u^*}{\partial x^*} \frac{\partial \psi_s^*}{\partial x^*} - \frac{\partial \psi_s^*}{\partial y^*} \left(\frac{\partial u^*}{\partial y^*} + \frac{\partial v^*}{\partial x^*} \right) \right) = 0 \\ \frac{1}{Rt} \frac{\partial v^*}{\partial z^*} + \frac{\varepsilon}{Fr^2} (p^* - p_a^*) \frac{\partial \psi_s^*}{\partial y^*} + \frac{\varepsilon^2}{Rt} \left(\frac{\partial w^*}{\partial y^*} - 2 \frac{\partial v^*}{\partial y^*} \frac{\partial \psi_s^*}{\partial y^*} - \frac{\partial \psi_s^*}{\partial x^*} \left(\frac{\partial u^*}{\partial y^*} + \frac{\partial v^*}{\partial x^*} \right) \right) = 0 \\ -\frac{1}{Fr^2} (p^* - p_a^*) + \frac{\varepsilon}{Rt} \left(2 \frac{\partial w^*}{\partial z^*} - \frac{\partial \psi_s^*}{\partial x^*} \frac{\partial u^*}{\partial z^*} - \frac{\partial \psi_s^*}{\partial y^*} \frac{\partial v^*}{\partial z^*} \right) - \frac{\varepsilon^3}{Rt} \left(\frac{\partial \psi_s^*}{\partial x^*} \frac{\partial w^*}{\partial x^*} + \frac{\partial \psi_s^*}{\partial y^*} \frac{\partial w^*}{\partial y^*} \right) = 0 \end{cases} \quad (2.18)$$

La condition aux limites cinématique adimensionnée s'écrit :

$$w^* - \frac{\partial \psi_s^*}{\partial t^*} - u^* \frac{\partial \psi_s^*}{\partial x^*} - v^* \frac{\partial \psi_s^*}{\partial y^*} = 0 \quad \text{pour } z^* = \psi_s^*(x^*, y^*, t^*) \quad (2.19)$$

La condition d'imperméabilité sur le fond s'écrit :

$$w^* - \frac{\partial \psi_f^*}{\partial t^*} - u^* \frac{\partial \psi_f^*}{\partial x^*} - v^* \frac{\partial \psi_f^*}{\partial y^*} = 0 \quad \text{pour } z^* = \psi_f^*(x^*, y^*, t^*) \quad (2.20)$$

La contrainte de cisaillement s'écrit :

$$\frac{\partial u^*}{\partial z^*} + \varepsilon^2 \frac{\partial w^*}{\partial x^*} = \frac{Rt}{Fr^2} \tau_x^*, \quad \frac{\partial v^*}{\partial z^*} + \varepsilon^2 \frac{\partial w^*}{\partial y^*} = \frac{Rt}{Fr^2} \tau_y^*, \quad \text{pour } z^* = \psi_f^*(x^*, y^*, t^*)$$

où on a choisi d'adimensionnaliser la contrainte de cisaillement $\vec{\tau}$ par :
 $\vec{\tau} = \rho g H_0 \vec{\tau}^*$

2.2.2 Passage à la limite des milieux peu profonds

Dans le cas des milieux peu profonds le rapport d'aspect ε est très faible, soit $\varepsilon \ll 1$. On suppose ici que $Fr = O(1)$ et $\frac{1}{Rt} = O(\varepsilon)$. Les équations de Navier-Stokes adimensionnalisées données par le système (2.17) s'écrivent alors :

$$\left\{ \begin{array}{l} \frac{\partial u^*}{\partial x^*} + \frac{\partial v^*}{\partial y^*} + \frac{\partial w^*}{\partial z^*} = 0 \\ \varepsilon \left(\frac{\partial u^*}{\partial t^*} + u^* \frac{\partial u^*}{\partial x^*} + v^* \frac{\partial u^*}{\partial y^*} + w^* \frac{\partial u^*}{\partial z^*} \right) = \frac{1}{Rt} \frac{\partial^2 u^*}{\partial z^{*2}} - \frac{\varepsilon}{Fr^2} \frac{\partial p^*}{\partial x^*} \\ \varepsilon \left(\frac{\partial v^*}{\partial t^*} + u^* \frac{\partial v^*}{\partial x^*} + v^* \frac{\partial v^*}{\partial y^*} + w^* \frac{\partial v^*}{\partial z^*} \right) = \frac{1}{Rt} \frac{\partial^2 v^*}{\partial z^{*2}} - \frac{\varepsilon}{Fr^2} \frac{\partial p^*}{\partial y^*} \\ 0 = -\frac{1}{Fr^2} \frac{\partial p^*}{\partial z^*} - \frac{1}{Fr^2} \end{array} \right. \quad (2.21)$$

La continuité des contraintes de surface se réduit aux conditions suivantes :

$$\frac{\partial u^*}{\partial z^*} = \frac{\partial v^*}{\partial z^*} = 0, \quad p^* = p_a^*, \quad \text{pour } z^* = \psi_s^*(x^*, y^*, t^*) \quad (2.22)$$

La contrainte de cisaillement sur le fond s'écrit quant à elle :

$$\frac{\partial u^*}{\partial z^*} = \frac{Rt}{Fr^2} \tau_x^*, \quad \frac{\partial v^*}{\partial z^*} = \frac{Rt}{Fr^2} \tau_y^* \quad \text{pour } z^* = \psi_f^*(x^*, y^*, t^*)$$

La condition aux limites cinématique sur la surface libre et celle d'imperméabilité sur le fond restent identiques respectivement aux conditions données en (2.19) et (2.20).

La dernière équation du système (2.21) implique que la pression est hydrostatique. En effet si l'on revient aux variables dimensionnées on peut montrer que l'on a la relation :

$$p(x, y, z, t) = p_a + \rho g(\psi_s(x, y, t) - z) \quad (2.23)$$

L'équation (2.23) nous permet de réécrire en variables dimensionnées le système (2.21) et les conditions limites données en (2.22) :

$$\left\{ \begin{array}{l} \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \\ \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} = \nu \frac{\partial^2 u}{\partial z^2} - g \frac{\partial \psi_s}{\partial x} \\ \frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} + w \frac{\partial v}{\partial z} = \nu \frac{\partial^2 v}{\partial z^2} - g \frac{\partial \psi_s}{\partial y} \end{array} \right. \quad (2.24)$$

Les conditions aux limites sur la surface et sur le fond données en (2.14) et (2.16) restent identiques. La continuité des contraintes sur la surface libre s'écrit :

$$\frac{\partial u}{\partial z} = \frac{\partial v}{\partial z} = 0, \quad p = p_a, \quad \text{pour } z = \psi_s(x, y, t)$$

Et la contrainte de cisaillement sur le fond s'écrit :

$$\rho \nu \frac{\partial u}{\partial z} = \tau_x, \quad \rho \nu \frac{\partial v}{\partial z} = \tau_y, \quad \text{pour } z = \psi_f(x, y, t) \quad (2.25)$$

2.2.3 Equations de Saint-Venant

Pour dériver les équations de Saint-Venant du système d'équations (2.24) il faut intégrer les équations selon la verticale. Pour cela, on définit les vitesses moyennes selon la verticale U et V selon respectivement $\vec{e}_x$ et $\vec{e}_y$ comme étant :

$$U(x, y, t) = \frac{1}{h(x, y, t)} \int_{\psi_f(x, y, t)}^{\psi_s(x, y, t)} u(x, y, t) dz$$

$$V(x, y, t) = \frac{1}{h(x, y, t)} \int_{\psi_f(x, y, t)}^{\psi_s(x, y, t)} v(x, y, t) dz$$

où $h(x, y, t) = (\psi_s - \psi_f)(x, y, t)$

En intégrant selon la verticale les équations du système (2.24), on montre que les variables h , u , v , U et V vérifient le système d'équations suivant :

$$\left\{ \begin{array}{l} \frac{\partial h}{\partial t} + \frac{\partial(hU)}{\partial x} + \frac{\partial(hV)}{\partial y} = 0 \\ \frac{\partial(hU)}{\partial t} + \frac{\partial}{\partial x} \int_{\psi_f}^{\psi_s} u^2 dz + \frac{\partial}{\partial y} \int_{\psi_f}^{\psi_s} uv dz = -gh \frac{\partial \psi_s}{\partial x} - \nu \frac{\partial u}{\partial z} \Big|_{z=\psi_f} \\ \frac{\partial(hU)}{\partial t} + \frac{\partial}{\partial x} \int_{\psi_f}^{\psi_s} uv dz + \frac{\partial}{\partial y} \int_{\psi_f}^{\psi_s} v^2 dz = -gh \frac{\partial \psi_s}{\partial y} - \nu \frac{\partial v}{\partial z} \Big|_{z=\psi_f} \end{array} \right. \quad (2.26)$$

On souhaite écrire le système (2.26) seulement en fonction des champs $h(x, y, t)$, $U(x, y, t)$ et $V(x, y, t)$. Pour cela on réécrit les termes faisant intervenir une intégrale et, suivant [Thual, 2010], on utilise la modélisation suivante pour les termes intégraux (obtenue par analyse dimensionnelle) :

$$\int_{\psi_f}^{\psi_s} u^2 dz = \alpha U^2 h, \quad \int_{\psi_f}^{\psi_s} uv dz = \alpha UV h, \quad \int_{\psi_f}^{\psi_s} v^2 dz = \alpha V^2 h \quad (2.27)$$

[Thual, 2010] considère l'existence d'une couche limite sur le fond de sommet $z_{lim}(x, y, t)$. L'écoulement étant turbulent $u(z)$ et $v(z)$ sont quasiment constants sur une grande partie de la couche limite et [Thual, 2010] suppose que $\alpha = 1$.

D'autre part, d'après (2.25), on peut écrire : $\nu \frac{\partial u}{\partial z} \Big|_{z=\psi_f} = \frac{\tau_x}{\rho}$ et $\nu \frac{\partial v}{\partial z} \Big|_{z=\psi_f} = \frac{\tau_y}{\rho}$. Or, l'analyse dimensionnelle montre que $\vec{\tau}$ est de la forme :

$$\vec{\tau} = -\frac{1}{2} \rho C_f(h) \sqrt{U^2 + V^2} \vec{U}_h \quad (2.28)$$

où $\vec{U}_h = (U, V)$ et où le coefficient C_f est un coefficient de frottement sans dimension. Il existe plusieurs modélisations de ce coefficient (formule de Chézy, formule de Manning...). En hydraulique à surface libre on utilise souvent la formule de Strickler :

$$C_f(h) = \frac{2g}{K_s h^{\frac{1}{3}}} \quad (2.29)$$

où K_s est le nombre de Strickler. En utilisant les expressions données en (2.27) ainsi que les relations (2.25), (2.28) et (2.29) on peut réécrire le système (2.26) sous la forme suivante :

$$\left\{ \begin{array}{l} \frac{\partial(hU)}{\partial x} + \frac{\partial(hV)}{\partial y} + \frac{\partial h}{\partial t} = 0 \\ \frac{\partial(hU)}{\partial t} + \frac{\partial(hU^2)}{\partial x} + \frac{\partial(hUV)}{\partial y} = -gh \frac{\partial \psi_s}{\partial x} - gh S_{fx} \\ \frac{\partial(hV)}{\partial t} + \frac{\partial(hUV)}{\partial x} + \frac{\partial(hV^2)}{\partial y} = -gh \frac{\partial \psi_s}{\partial y} - gh S_{fy} \end{array} \right. \quad (2.30)$$

Le système d'équations (2.30) est le système des équations de Saint-Venant.

Dans le système d'équations (2.30) les termes S_{fx} et S_{fy} représentent les termes de frottement et sont donnés par les expressions suivantes :

$$S_{fx} = \frac{1}{K_s^2} \frac{U \sqrt{U^2 + V^2}}{h^{\frac{4}{3}}} \quad \text{et} \quad S_{fy} = \frac{1}{K_s^2} \frac{V \sqrt{U^2 + V^2}}{h^{\frac{4}{3}}} \quad (2.31)$$

2.3 Les modèles d'ondes de crues

Dans cette section on s'intéresse au cas particulier des écoulements mono-dimensionnels dans un canal à fond plat formant un angle γ avec l'horizontale, (voir figure 2.1). On montre, que sous certaines hypothèses complémentaires, on peut dériver des équations de Saint-Venant deux approximations linéaires que nous utiliserons dans le chapitre 7 portant sur l'émulation du filtre de Kalman d'ensemble sur un modèle d'onde de crues.



FIG. 2.1 – Illustration d'un écoulement mono-dimensionnel dans un canal à fond plat formant un angle γ avec l'horizontale. Source : [Thual, 2010].

2.3.1 Le modèle de l'approximation des ondes de crues

Dans le cas où la pente est forte et le frottement important on peut considérer le modèle de l'approximation des ondes de crues.

On considère que $\varepsilon \ll 1$, $\frac{1}{Rt} = O(1)$, $Fr^2 = O(1)$ et $\tan\gamma = O(1)$. On peut montrer dans ce cas particulier que les équations de Saint-Venant se réduisent au système suivant :

$$\frac{\partial(hU)}{\partial x} + \frac{\partial h}{\partial t} = 0 \quad \text{et} \quad 0 = gI - g \frac{1}{K_s^2} \frac{U|U|}{h^{\frac{4}{3}}}$$

où $I = \sin\gamma$.

Si l'on s'intéresse à des régimes où $U \geq 0$ on peut montrer que pour des petites perturbations $(\tilde{h}, \tilde{U})$ autour d'un état d'équilibre (h_n, U_n) vérifiant la relation de Strickler :

$$U(h) = K_s I^{\frac{1}{2}} h^{\frac{2}{3}} \quad (2.32)$$

on peut établir l'**approximation des ondes de crues** :

$$\boxed{\frac{\partial \tilde{h}}{\partial t} + \frac{5U_n}{3} \frac{\partial \tilde{h}}{\partial x} = 0} \quad (2.33)$$

où $(\tilde{h}, \tilde{U})$ et (h_n, U_n) vérifient les relations : $h = h_n + \tilde{h}$ et $U = U_n + \tilde{U}$.

L'équation (2.33) est une équation d'advection, autrement dit ce modèle advecte simplement de l'amont vers l'aval le signal imposé à l'amont du modèle.

2.3.2 Le modèle de l'approximation des ondes de crues diffusantes

Dans le cas où la frottement est important et la pente modérée on peut considérer le modèle de l'approximation des ondes de crues diffusantes.

On considère que $\varepsilon \ll 1$, $\frac{1}{Rt} = O(1)$, $Fr^2 = O(\varepsilon)$ et $\tan\gamma = O(\varepsilon)$. On peut montrer dans ce cas particulier que :

$$\frac{\partial(hU)}{\partial x} + \frac{\partial h}{\partial t} = 0 \quad \text{et} \quad 0 = -g' \frac{\partial h}{\partial x} + gI - g \frac{1}{K_s^2} \frac{U|U|}{h^{\frac{4}{3}}}$$

où $g' = g \cos\gamma$.

Pour une perturbation $(\tilde{h}, \tilde{U})$ autour d'un état d'équilibre (h_n, U_n) (où (h_n, U_n) vérifie la relation de Strickler (2.32)) on peut établir l'**approximation des ondes de crues diffusantes** :

$$\boxed{\frac{\partial \tilde{h}}{\partial t} + \frac{5U_n}{3} \frac{\partial \tilde{h}}{\partial x} = k_n \frac{\partial^2 \tilde{h}}{\partial x^2}} \quad (2.34)$$

où $k_n = \frac{U_n h_n}{2 \tan\gamma}$.

L'équation (2.34) est une équation d'advection-diffusion ; ce modèle va donc advecter et diffuser le signal imposé à l'amont.

2.4 Principe de résolution des équations de Saint-Venant avec le code Mascaret

La résolution des équations de Saint-Venant par le code Mascaret développé par EDF et le CEREMA est basée sur la formulation de ces équations sous forme conservative mono-dimensionnelle avec les variables de section mouillée et de débit que nous présentons dans une première partie. Dans une seconde partie nous présentons dans les grandes lignes les méthodes numériques développées pour résoudre ce système d'équations dans Mascaret.

Pour écrire cette section nous nous sommes très largement appuyés sur [Goutal, 2014] qui synthétise les principes et les méthodes numériques développées dans Mascaret pour la résolution des équations de Saint-Venant. [Goutal, 2014] donne également de nombreuses références qui ont servi au développement des méthodes numériques utilisées dans Mascaret, le lecteur intéressé pourra s'y reporter.

2.4.1 Forme conservative des équations de Saint-Venant 1D

On définit tout d'abord :

$$\mathcal{S}(x, Z) = \int_{Z_f(x)}^Z L(x, z) dz$$

où L est la largeur de la section en travers en x à la cote z .

On définit la surface mouillée $S(x, t)$ et le débit $Q(x, t)$ comme étant :

$$\begin{aligned} S(x, t) &= \mathcal{S}[x, Z_s(x, t)] \\ Q(x, t) &= \int_{Z_f(x)}^{Z(x, t)} L(x, z) u(x, t) dz \end{aligned}$$

où $Z_f(x) = \min_y \psi_f(x, y)$, $Z_s(x, t) = \psi_s(x, y, t)$.

Les équations de Saint-Venant 1D peuvent être dérivées du système d'équations (2.30) en moyennant celles-ci selon la largeur et en faisant l'hypothèse que la vitesse transverse est négligeable par rapport à la vitesse dans le sens de l'écoulement [Decoene et al., 2009, Goutal and Sainte-Marie, 2011] :

$$\begin{cases} \frac{\partial S}{\partial t} + \frac{\partial Q}{\partial x} = 0 \\ \frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(\beta \frac{Q^2}{S} \right) + gS \frac{\partial Z_s}{\partial x} = -gSS_f \end{cases}$$

où S_f est le frottement modélisé par la formule de Strickler : $S_f = \frac{Q^2}{K^2 S^2 R_h^{4/3}}$ et où β est un coefficient adimensionnel résultant des variations de la vitesse réelle de l'écoulement en une section, β est donné par l'expression suivante :

$$\beta(x, S) = \frac{S}{Q^2} \int_S V^2 dS$$

Dans le cas où le lit n'est pas composé, i-e il n'est pas composé d'un lit mineur et d'un lit majeur, nous pouvons prendre $\beta = 1$. Dans la suite nous ne considérerons que des lits non composés.

Pour obtenir la forme conservative de ces équations il faut décomposer le terme représentant les forces de pesanteur et de pression en faisant apparaître un terme de pression [Goutal, 2014]. Alors **les équations de Saint-Venant 1D sous forme conservative** peuvent s'écrire :

$$\left\{ \begin{array}{l} \frac{\partial S}{\partial t} + \frac{\partial Q}{\partial x} = 0 \\ \frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(\frac{Q^2}{S} + P \right) = g \int_0^h \left[\frac{\partial \mathcal{S}(x, z)}{\partial x} \right]_z dz - gS \frac{\partial Z_f}{\partial x} - gSS_f \end{array} \right. \quad (2.35)$$

où $h = Z - Z_f$, $P = \int_0^h \mathcal{S}(x, Z) dZ = \mathcal{P}[x, S(x, t)]$.

Un système d'équations conservatif s'écrit sous la forme :

$$\frac{\partial W}{\partial t} + \text{div} [F(W)] = B(W) \quad (2.36)$$

où W est le vecteur d'état, $F(W)$ est le flux et $B(W)$ est le terme source. Le système d'équations (2.35) est donc bien un système conservatif avec :

$$\begin{aligned} W &= \begin{pmatrix} S \\ Q \end{pmatrix} \\ F &= \begin{pmatrix} Q \\ \frac{Q^2}{S} + P \end{pmatrix} \\ B &= \begin{pmatrix} 0 \\ g \int_0^h \left[\frac{\partial \mathcal{S}(x, Z)}{\partial x} \right]_z dz - gS \frac{\partial Z_f}{\partial x} - gSS_f \end{pmatrix} \end{aligned}$$

Nature du système d'équations Saint-Venant conservatif

La matrice jacobienne de F , $\left(\frac{\partial F}{\partial W} \right)_x$, dont l'expression est donnée par :

$$\left(\frac{\partial F}{\partial W} \right)_x = \begin{pmatrix} 0 & 1 \\ -\frac{Q^2}{S^2} + \left(\frac{\partial \mathcal{P}}{\partial S} \right)_x & \frac{2Q}{S} \end{pmatrix}$$

admet, si la hauteur d'eau est positive, deux valeurs propres réelles distinctes λ^+ et λ^- dont les expressions respectives sont données par :

$$\lambda^+ = \frac{Q}{S} + c \quad \lambda^- = \frac{Q}{S} - c$$

avec $c(x, t) = \mathcal{C}[x, S(x, t)]$ et $\mathcal{C}(x, S) = \frac{\partial \mathcal{P}}{\partial S}(x, S)$.

Les invariants de Riemann du système sont :

$$R^+ = \frac{Q}{S} + K \quad R^- = \frac{Q}{S} - K$$

$$\text{où } K = \int_0^S \frac{\mathcal{C}(x, s)}{s} ds$$

Si la solution est régulière, le système (2.35) peut se réécrire sous la forme de deux équations de convection des invariants de Riemann le long des courbes caractéristiques et avec intégration des termes sources le long de ces courbes. Celles-ci sont définies par :

- la courbe caractéristique C^+ de pente $\frac{dx}{dt} = \lambda^+$;
- la courbe caractéristique C^- de pente $\frac{dx}{dt} = \lambda^-$.

Selon la pente de la caractéristique C^- le régime d'écoulement est fluvial ou torrentiel. En effet dans le cas où elle est positive les courbes caractéristiques sont toutes les deux orientées dans le sens des x croissant et l'écoulement est torrentiel. Au contraire si la pente de la caractéristique C^- est négative les courbes caractéristiques sont orientées, relativement à l'axe des x , dans des sens différents. Les courbes caractéristiques ainsi que les invariants de Riemann vont trouver tout leur intérêt pour définir les conditions aux limites selon le régime d'écoulement (voir section 2.4.2).

2.4.2 Résolution numérique des équations de Saint-Venant avec une méthode volumes finis

Dans cette section, on expose tout d'abord la résolution numérique du système d'équations de Saint-Venant sous forme conservative présenté en (2.35) dans le cadre d'un bief de géométrie quelconque mais sans lit composé. On présente ensuite, dans la perspective d'une application à une rivière réelle avec confluence, le traitement des confluences dans le code Mascaret.

Dans cette section on identifie la rivière à un intervalle $[a; b]$ et on considère celui-ci comme étant l'union d'intervalles $[x_i; x_{i+1}]$. On note $x_{i+1/2}$ le point milieu de cet intervalle. Par la suite on s'intéresse à la résolution numérique des équations de Saint-Venant avec un schéma volumes finis sur des cellules du type $[x_{i-1/2}; x_{i+1/2}]$, celle-ci repose sur le système homogène suivant :

$$\begin{cases} W = (S, Q) \\ \frac{\partial W}{\partial t} + \frac{\partial F(x, W)}{\partial x} = 0 \end{cases} \quad (x, t) \in [a; b] \times [0; T] \quad (2.37)$$

En recherchant la solution au système (2.37) au temps t_{n+1} , W_h^{n+1} , dans l'espace des fonctions constantes sur chaque cellule on se ramène à chaque interface entre deux cellules à un problème de Riemann :

$$\begin{cases} \frac{\partial W_h}{\partial t} + \frac{\partial F(x_{i+1/2}, W_h)}{\partial x} = 0 \\ W_h = \begin{cases} W_i^n & \text{si } x < x_{i+1/2} \\ W_{i+1}^n & \text{si } x > x_{i+1/2} \end{cases} \end{cases} \quad (2.38)$$

La résolution du système (2.38) est basée sur l'utilisation d'un solveur de Roe qui postule l'existence d'une matrice $\tilde{A}$ telle que :

- $F(W_g) - F(W_d) = \tilde{A}(W_g, W_d)(W_g - W_d)$;
- $\tilde{A}(W, W) = DF(W)$;

où W_g et W_d sont respectivement les états à gauche et à droite de l'interface et où la matrice de Roe $\tilde{A}(x, W_g, W_d)$ approxime $\left(\frac{\partial F}{\partial W}\right)_x$ calculée en un état moyen $(\tilde{u}, \tilde{c})$ et définie par :

$$\left(\frac{\partial F}{\partial W}\right)_x = \begin{pmatrix} 0 & 1 \\ -\tilde{u}^2 + \tilde{c} & 2\tilde{u} \end{pmatrix}$$

$$\text{où } \tilde{u} = \frac{u_g \sqrt{S_g} + u_d \sqrt{S_d}}{\sqrt{S_g} + \sqrt{S_d}}, \tilde{c} = c(\tilde{S}), u = \frac{Q}{S} \text{ et } \tilde{S} = \frac{S_g \sqrt{S_g} + S_d \sqrt{S_d}}{\sqrt{S_g} + \sqrt{S_d}}.$$

Le problème de Riemann (2.38) se ramène donc à la résolution à chaque interface d'un problème linéarisé :

$$\begin{cases} \frac{\partial W(x, t)}{\partial t} + \tilde{A}(x_{i+1/2}, W_g, W_d) \frac{\partial W(x, t)}{\partial x} = 0 \\ W(x, t^n) = \begin{cases} W_g & \text{si } x < x_{i+1/2} \\ W_d & \text{si } x > x_{i+1/2} \end{cases} \end{cases} \quad (2.39)$$

$$\text{où } \tilde{A}(x_{i+1/2}, W_g, W_d) = A(x_{i+1/2}, W).$$

Les solutions du système (2.39) sont données par trois états constants séparés par les caractéristiques λ_1 et λ_2 (on suppose $\lambda_1 < \lambda_2$) :

$$W\left(\frac{x}{t}, W_g, W_d\right) = \begin{cases} W_g = \sum_{i=1,2} \alpha_{ig}(W_g, W_d) r_i(W_g, W_d) & \text{si } \frac{x}{t} < \lambda_1 \\ W_m = \alpha_{1d}(W_g, W_d) r_1(W_g, W_d) + \alpha_{2g}(W_g, W_d) r_2(W_g, W_d) & \text{si } \lambda_1 < \frac{x}{t} < \lambda_2 \\ W_d = \sum_{i=1,2} \alpha_{id}(W_g, W_d) r_i(W_g, W_d) & \text{si } \frac{x}{t} > \lambda_2 \end{cases}$$

où les r_i sont les vecteurs propres à droite associés à la valeur propre λ_i et où les caractéristiques λ_1 et λ_2 sont données par les expressions :

$$\begin{aligned}\lambda_1(x, S, Q) &= \frac{Q}{S} - c(x, S) \\ \lambda_2(x, S, Q) &= \frac{Q}{S} + c(x, S) \\ c(x, S) &= \frac{\partial P(x, S)}{\partial x}\end{aligned}$$

Maintenant que l'on dispose de l'expression des solutions au système (2.39) on peut donner l'expression du flux numérique de Roe. Si on note A la matrice jacobienne de F diagonalisable dans la base de vecteurs propres X alors A s'écrit $X\Lambda X^{-1}$. On note de plus $|\Lambda|$ la matrice diagonale de terme générique $|\Lambda_i|$ et Λ^+ et Λ^- les matrices définies par :

$$\begin{cases} \Lambda_{i,j}^+ = \lambda_i^+ \delta_{i,j} \\ \Lambda_{i,j}^- = \lambda_i^- \delta_{i,j} \end{cases}$$

On définit de plus les matrices $A^+ = X\Lambda^+X^{-1}$, $A^- = X\Lambda^-X^{-1}$ et $|A| = X|\Lambda|X^{-1}$. Alors l'expression du flux numérique de Roe est donnée par :

$$\begin{aligned}F(x_{i+1/2}, W_g, W_d) &= \frac{1}{2}(F(x_{i+1/2}, W_g) + F(x_{i-1/2}, W_d)) \\ &\quad + \frac{1}{2} \left| \tilde{A}(x_{i+1/2}, W_g, W_d) \right| (W_d - W_g) \\ &= F(x_{i+1/2}, W_d) - \tilde{A}^+(x_{i+1/2}, W_g, W_d)(W_d - W_g) \\ &= F(x_{i+1/2}, W_d) + \tilde{A}^-(x_{i+1/2}, W_g, W_d)(W_d - W_g)\end{aligned}\tag{2.40}$$

L'expression (2.40) est une expression explicite du flux numérique, celle-ci donne de bons résultats pour des écoulements à la fois instationnaires et fluviaux [Goutal, 2014]. Mais la contrainte sur le pas de temps peut être très pénalisante dans le cas d'écoulements fluviaux. Pour lever cette contrainte le schéma volumes finis peut être implicité linéairement. Celui-ci s'écrit (dans le cas d'un système homogène pour en simplifier l'écriture) :

$$h_i(W_i^{n+1} - W_i^n) + \Delta t (F_{i+1/2}(W_i^{n+1}, W_{i+1}^{n+1}) - F_{i-1/2}(W_{i-1}^{n+1}, W_i^{n+1})) = 0$$

La fonction $F_{i+1/2}$ n'étant pas linéaire, on peut la linéariser à l'aide d'un développement limité. On se ramène alors à résoudre le système linéaire suivant :

$$\begin{aligned}
 & \frac{1}{2} \left[\frac{\partial F(W_{i-1})}{\partial W_{i-1}} + \left| \tilde{A}(x_{i-1/2}, W_i, W_{i-1}) \right| \right] \delta W_{i-1} \\
 & + \left[\frac{h_i}{\Delta t} + \frac{1}{2} \left| \tilde{A}(x_{i+1/2}, W_i, W_{i+1}) \right| + \frac{1}{2} \left| \tilde{A}(x_{i-1/2}, W_i, W_{i-1}) \right| \right] \delta W_i \\
 & \frac{1}{2} \left[\frac{\partial F(W_{i+1})}{\partial W_{i+1}} - \left| \tilde{A}(x_{i+1/2}, W_i, W_{i+1}) \right| \right] \delta W_{i+1} \\
 & = - \frac{F_{i+1/2}(W_i^n, W_{i+1}^n) - F_{i-1/2}(W_i^n, W_{i-1}^n)}{\Delta t}
 \end{aligned}$$

où : $\delta W_i = W_i^{n+1} - W_i^n$.

Traitement numérique des conditions aux limites

Pour évaluer les flux numériques il est nécessaire de connaître à chaque pas de temps les valeurs des deux variables aux limites du domaine. Des cellules fictives sont introduites aux limites du domaine et l'état dans une cellule fictive est donné par la théorie des caractéristiques et les invariants de Riemann. Selon que l'on considère une cellule à l'amont ou à l'aval du domaine et selon le régime de l'écoulement l'utilisateur sera amené à fournir une ou deux informations.

A l'amont du domaine :

- **régime fluvial** : l'utilisateur fournit le débit ou la cote et l'information manquante vient de l'invariant de Riemann porté par la caractéristique sortante ;
- **régime torrentiel** : l'utilisateur fournit la cote et le débit, les deux caractéristiques étant entrantes dans le domaine.

A l'aval du domaine :

- **régime fluvial** : l'utilisateur fournit le débit ou la cote ou une courbe de tarage et l'information manquante vient de l'invariant de Riemann porté par la caractéristique sortante ;
- **régime torrentiel** : l'utilisateur ne fournit aucune information car les deux caractéristiques sont sortantes du domaine de calcul.

Prise en compte des termes sources

Les équations de Saint-Venant présentent trois types de terme source distincts :

- les apports de débit : q_a ;
- le terme source lié à la géométrie : il contient les gradients de fond et les variations de largeur ;
- le terme de frottement.

La mise sous forme conservative des équations de Saint-Venant (2.35) peut engendrer la non préservation d'états d'équilibre si le traitement des termes sources est mauvais. Dans Mascaret, le traitement des termes sources pour une rivière réelle est dérivé des travaux de [Vázquez-Cendón, 1999]. Si on note $B(x, W)$ le terme source et $X(\tilde{W})$ la base

de vecteurs propres de la matrice de Roe linéarisée $\tilde{A}$ alors la contribution des termes sources dans le code Mascaret est évaluée dans chaque cellule par les relations :

- à gauche de l'interface, dans la moitié amont de la cellule $i+1$: $\int_{x_{i+1/2}}^{x_{i+1/2}} \Phi_g(i, i+1) dx$;
- à droite de l'interface, dans la moitié aval de la cellule i : $\int_{x_{i+1/2}}^{x_i} \Phi_d(i, i+1) dx$.

où :

$$\begin{cases} \Phi_g(x, y, \tilde{W}) = 2X(\tilde{W}) \times \left(\frac{1}{2}(I + |\Lambda(\tilde{W})|\Lambda^{-1}(\tilde{W}))X^{-1}(\tilde{W})B(x, y, W) \right) \\ \Phi_d(x, y, \tilde{W}) = 2X(\tilde{W}) \times \left(\frac{1}{2}(I - |\Lambda(\tilde{W})|\Lambda^{-1}(\tilde{W}))X^{-1}(\tilde{W})B(x, y, W) \right) \end{cases}$$

Cette méthode permet de traiter les termes de gradient de fond et de variation de largeur ainsi que la contribution du terme de frottement lorsque celle-ci est calculée de manière explicite. Néanmoins comme il est indiqué dans [Goutal, 2014] le traitement explicite du terme de frottement peut poser problème pour des faibles hauteurs d'eau avec des petits coefficients de Strickler et mener à des comportements numériques qui ne sont pas physiques. Pour éviter ce problème un traitement implicite de la contribution du terme de frottement basé sur la méthode des caractéristiques [Paquier, 1995] existe et a été implémenté dans le code Mascaret.

Traitement des zones de confluence

Le traitement des zones de confluence dans le code Mascaret repose sur le fait que localement au niveau des confluences l'écoulement n'est plus mono-dimensionnel. Pour traiter les zones de confluence des modèles 2D ont été développés au niveau de ces zones de confluence (figure 2.2-(a)). Ces modèles sont couplés à des modèles 1D décrivant l'écoulement dans les biefs en amont et en aval de la confluence. Ce couplage se fait par recouvrement des modèles 1D et 2D comme cela est représenté sur la figure 2.2-(b).

Le principe de l'algorithme de couplage 1D-2D repose sur un schéma de Roe sur chacun des sous-domaines 1D et 2D. Les états dans les cellules 2D limites du modèle 1D (cellules A, B et C sur la figure 2.2-(b)) sont calculés avec le modèle 2D : on conserve la surface libre et on projette le débit selon la normale à la frontière libre pour obtenir la valeur 1D. De manière similaire les états dans les cellules 1D limites du modèle 2D (cellules a, b et c sur la figure 2.2-(b)) sont calculés avec les modèles 1D : on conserve la surface libre et on suppose que le débit a la même direction que la normale à la frontière libre.

Cet algorithme a été mis en place sur des géométries simplifiées avec des canaux rectangulaires aussi bien que sur des géométries réelles. Dans le cas où l'on voudrait avoir une résolution plus fine de ce qui se passe dans la zone de confluence on peut avoir recours à des techniques plus élaborées du type couplage de modèles 1D-2D où le modèle 2D cette fois-ci fournit une modélisation à une résolution plus fine et peut s'étendre bien au-delà de la simple confluence.

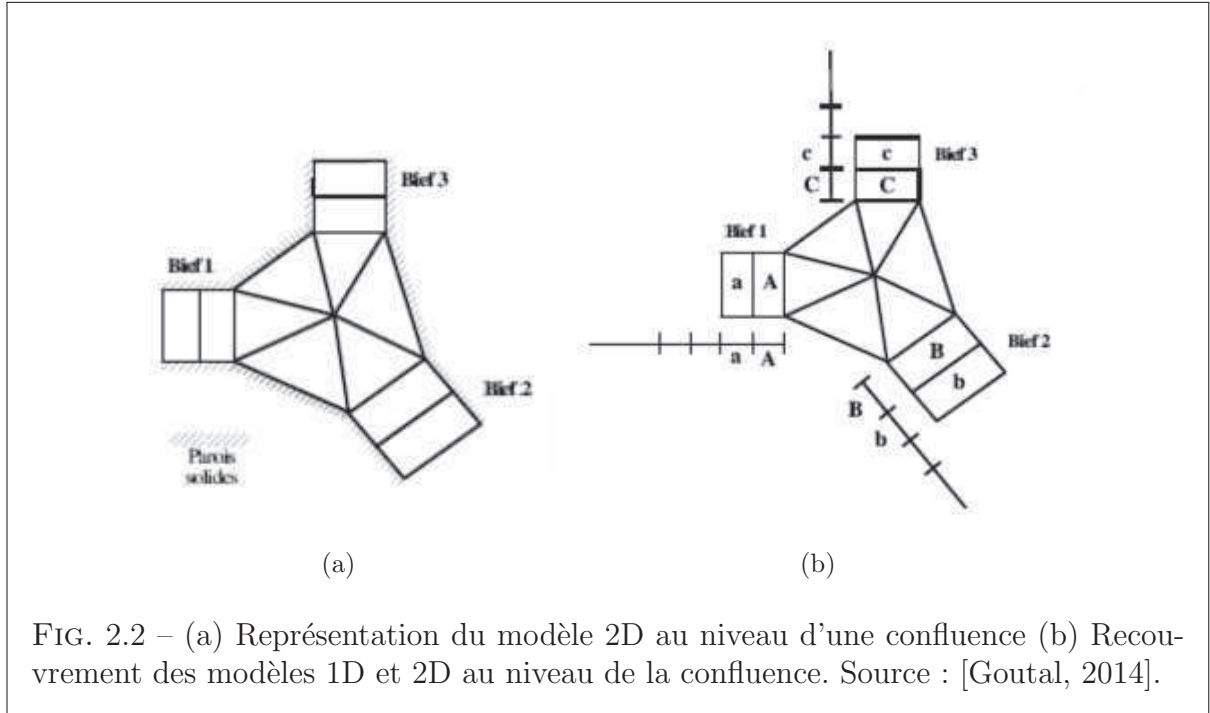


FIG. 2.2 – (a) Représentation du modèle 2D au niveau d’une confluence (b) Recouvrement des modèles 1D et 2D au niveau de la confluence. Source : [Goutal, 2014].

2.5 Principe de résolution des équations de Saint-Venant avec le code Telemac-2D

Dans le cas du système Telemac-2D c’est la méthode de résolution par éléments finis qui a été développée. Nous supposons dans la suite que le lecteur est familier de cette méthode.

Pour des raisons algorithmiques Telemac-2D résout les équations de Saint-Venant sous leur forme non-conservative. Si la méthode des caractéristiques est choisie pour traiter ces équations alors leur résolution se fait en deux étapes : une première dans laquelle on traite les termes de convection des grandeurs physiques puis une seconde dans laquelle on traite le reste des équations. Sinon, si la méthode des caractéristiques n’est pas retenue, les termes de transport sont traités dans la même étape que le traitement de la diffusion et des termes sources. Dans la suite nous présentons préalablement la forme non-conservative des équations de Saint-Venant utilisée dans Telemac-2D avant de présenter les méthodes de résolution de ces équations avec la méthode éléments finis. Pour cela nous reproduisons dans les grandes lignes l’exposé fait par Jean-Michel Hervouet dans son livre [Hervouet, 2007].

2.5.1 Forme non-conservative des équations de Saint-Venant

Dans un souci de simplicité nous reprenons dans cette section les notations utilisées dans la section (2.1) lors de la présentation des équations de Navier-Stokes et on note le vecteur vitesse : $\vec{U} = (u, v)$.

La forme non-conservative des équations de Saint-Venant telle qu’utilisée dans Telemac-2D s’énonce ainsi :

$$\begin{cases} \frac{\partial h}{\partial t} + \vec{U} \cdot \overrightarrow{\text{grad}}(h) + h \text{div}(\vec{U}) = S_{ce} \\ \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = -g \frac{\partial \psi_s}{\partial x} + F_x + \frac{1}{h} \text{div} \left(h \nu_e \overrightarrow{\text{grad}}(u) \right) \\ \frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} = -g \frac{\partial \psi_s}{\partial y} + F_y + \frac{1}{h} \text{div} \left(h \nu_e \overrightarrow{\text{grad}}(v) \right) \end{cases}$$

où ν_e est une diffusion qui prend en compte la viscosité turbulente et la dispersion, où $\vec{U} = (u, v)$ et où $\vec{F} = (F_x, F_y)$ représente la somme des forces et des termes sources s'exerçant sur le fluide et S_{ce} les termes sources pour l'équation de conservation de la masse.

2.5.2 La méthode des pas fractionnaires

Cette méthode n'est utilisée que si la méthode des caractéristiques est utilisée pour traiter l'étape de convection.

Si l'on considère une grandeur scalaire f , alors la discrétisation de la dérivée temporelle de cette grandeur selon un schéma différence finie peut s'écrire sous la forme suivante :

$$\frac{\partial f}{\partial t} = \frac{f^{n+1} - f^n}{\Delta t} \quad (2.41)$$

Il s'agit d'une discrétisation en temps d'ordre 1. La méthode des pas fractionnaires consiste alors à décomposer le terme de droite de l'équation (2.41) de manière à calculer d'abord les termes de transport :

$$\frac{\tilde{f} - f^n}{\Delta t} + \text{termes transport} = 0$$

puis les autres termes :

$$\frac{f^{n+1} - \tilde{f}}{\Delta t} + \text{autres termes} = 0$$

Dans l'étape de convection par la méthode des caractéristiques on résout les équations suivantes :

$$\begin{aligned} \frac{\tilde{h} - h^n}{\Delta t} + \vec{U} \cdot \overrightarrow{\text{grad}}(h) &= 0 \\ \frac{\tilde{u} - u^n}{\Delta t} + \vec{U} \cdot \overrightarrow{\text{grad}}(u) &= 0 \\ \frac{\tilde{v} - v^n}{\Delta t} + \vec{U} \cdot \overrightarrow{\text{grad}}(v) &= 0 \end{aligned}$$

Ce qui signifie que la dérivée lagrangienne des grandeurs h , u et v est nulle le long des courbes caractéristiques et donc que ces grandeurs restent inchangées le long de ces courbes. On renvoie le lecteur à [Hervouet, 2007] pour le détail du calcul des courbes caractéristiques.

2.5.3 Traitement de la propagation, de la diffusion et des termes sources – Discrétisation en temps et en espace

Discrétisation en temps

Cas d'une étape de convection préalable

Dans ce cas le système à résoudre est le suivant :

$$\begin{cases} \frac{h^{n+1} - \tilde{h}}{\Delta t} + h \operatorname{div}(\vec{U}) = S_{ce} \\ \frac{u^{n+1} - \tilde{u}}{\Delta t} = -g \frac{\partial \psi_s}{\partial x} + F_x + \frac{1}{h} \operatorname{div} \left(h \nu_e \overrightarrow{\operatorname{grad}}(u) \right) \\ \frac{v^{n+1} - \tilde{v}}{\Delta t} = -g \frac{\partial \psi_s}{\partial y} + F_y + \frac{1}{h} \operatorname{div} \left(h \nu_e \overrightarrow{\operatorname{grad}}(v) \right) \end{cases} \quad (2.42)$$

où $\tilde{h}$, $\tilde{u}$ et $\tilde{v}$ sont le résultat de la convection de h , u et v .

Cas sans étape de convection préalable

Dans ce cas le système d'équations à résoudre est :

$$\begin{cases} \frac{h^{n+1} - h^n}{\Delta t} + \vec{U} \cdot \overrightarrow{\operatorname{grad}}(h) + h \operatorname{div}(\vec{U}) = S_{ce} \\ \frac{u^{n+1} - u^n}{\Delta t} + \vec{U} \cdot \overrightarrow{\operatorname{grad}}(u) = -g \frac{\partial \psi_s}{\partial x} + F_x + \frac{1}{h} \operatorname{div} \left(h \nu_e \overrightarrow{\operatorname{grad}}(u) \right) \\ \frac{v^{n+1} - v^n}{\Delta t} + \vec{U} \cdot \overrightarrow{\operatorname{grad}}(v) = -g \frac{\partial \psi_s}{\partial y} + F_y + \frac{1}{h} \operatorname{div} \left(h \nu_e \overrightarrow{\operatorname{grad}}(v) \right) \end{cases} \quad (2.43)$$

Cependant dans chacun des deux systèmes (2.42) et (2.43) persistent des termes h , u et v dont il n'est pas précisé si ils sont pris au temps t^n ou au temps t^{n+1} . [Hervouet, 2007] précise que si l'on souhaite avoir un schéma d'ordre 2 en temps on pourrait écrire par exemple $h = \frac{h^{n+1} + h^n}{2}$. Or il précise que ces schémas ont tendance à diverger et propose de remplacer cette écriture par une écriture, pour une fonction scalaire f , du type :

$$f = \theta f^{n+1} + (1 - \theta) f^n \quad (2.44)$$

avec $\theta > 0.5$ tout en restant proche de 0.5. L'écriture (2.44) est appelée un θ schéma. En particulier lorsque θ vaut 0.5 on parle de schéma de Crank-Nicholson.

Sous-itérations

L'écriture de f donnée en (2.44) ne résout cependant pas tous les problèmes et des termes comme $h \operatorname{div}(\vec{U})$ peuvent faire apparaître le produit d'inconnues : $h^{n+1} \vec{U}^{n+1}$ qui

rendrait le système non linéaire et qui sans autre traitement serait donc impossible à résoudre. On utilise néanmoins une formule assez similaire à celle donnée en (2.44) :

$$h \operatorname{div}(\vec{U}) = h_{prop} \operatorname{div} \left(\theta_u \vec{U}^{n+1} + (1 - \theta_u) \vec{U}^n \right) \quad (2.45)$$

où θ_u est un coefficient d'implication des vitesses et h_{prop} est la hauteur de propagation qui, si on veut avoir un problème linéaire, doit être égale à h^n . Néanmoins on peut prendre une valeur plus proche de $h = \frac{h^{n+1} + h^n}{2}$ en résolvant de manière itérative au même pas de temps :

$$h_{prop} = \theta_h h'^{n+1} + (1 - \theta_h) h^n$$

où θ_h est le coefficient d'implication de la hauteur et h'^{n+1} est l'estimation courante de h^{n+1} .

De la même manière les termes de convection $\vec{U} \cdot \overrightarrow{\operatorname{grad}}(f)$ lorsqu'ils ne sont pas traités par la méthode des caractéristiques sont écrits sous la forme :

$$\vec{U} \cdot \overrightarrow{\operatorname{grad}}(f) = \vec{U}_{conv} \cdot \overrightarrow{\operatorname{grad}}(\theta_f f^{n+1} + (1 - \theta_f) f^n) \quad (2.46)$$

où $\vec{U}_{conv} = \theta_u \vec{U}^{m+1} + (1 - \theta_u) \vec{U}^n$ est la vitesse de convection et $\vec{U}^{m+1}$ est l'estimation courante de $\vec{U}^{n+1}$.

Termes de diffusion

Les termes de diffusion sont quant à eux semi-implicites sous la forme :

$$\frac{1}{h} \operatorname{div} \left(h \nu_e \overrightarrow{\operatorname{grad}} \left(\theta_u^d u^{n+1} + (1 - \theta_u^d) u^n \right) \right) \quad (2.47)$$

Le coefficient θ_u^d est un coefficient spécifique pour la diffusion et peut être choisi différent de θ_u et ce afin d'avoir plus de liberté dans le choix des schémas numériques.

Formulation en accroissement pour la hauteur

La fonction représentant la surface libre ψ_s est exprimée en fonction de la cote du fond ψ_f et de la hauteur h . Du fait de la semi-implication de h on écrit :

$$-\overrightarrow{\operatorname{grad}}(\psi_s) = -g \overrightarrow{\operatorname{grad}} (\theta_h h^{n+1} + (1 - \theta_h) h^n + \psi_f) \quad (2.48)$$

$$= -g \theta_h \overrightarrow{\operatorname{grad}} (h^{n+1} - h^n) - g \overrightarrow{\operatorname{grad}}(\psi_s^n) \quad (2.49)$$

où $\psi_s^n = h^n + \psi_f$. Au vu de l'expression (2.49) et de la discrétisation en temps de l'équation de continuité l'inconnue pour les hauteurs d'eau est $\delta h = h^{n+1} - h^n$ et non pas h^{n+1} . Ce choix permet d'assurer naturellement la conservation de l'état de repos pour un fond irrégulier.

Les termes de frottement

Les termes de frottement, sauf dans le cas où ils sont linéaires, sont écrits de manière à conserver une formulation linéaire et les produits $u\sqrt{u^2 + v^2}$ et $v\sqrt{u^2 + v^2}$ sont réécrits :

$$\begin{aligned} u\sqrt{u^2 + v^2} &\approx u^{n+1}\sqrt{(u^n)^2 + (v^n)^2} \\ v\sqrt{u^2 + v^2} &\approx v^{n+1}\sqrt{(u^n)^2 + (v^n)^2} \end{aligned}$$

Réécriture du système d'équations

Le système d'équations (2.43) avec les expressions (2.45), (2.46), (2.47) et (2.49) s'écrit :

$$\left\{ \begin{aligned} \frac{h^{n+1} - h^n}{\Delta t} + \vec{U}_{conv} \cdot \overrightarrow{\text{grad}}(h) + h_{prop} \text{div}(\vec{U}) &= S_{ce} \\ \frac{u^{n+1} - u^n}{\Delta t} + \vec{U}_{conv} \cdot \overrightarrow{\text{grad}}(u) &= -g\theta_h \frac{\partial(h^{n+1} - h^n)}{\partial x} - g \frac{\partial \psi_s^n}{\partial x} + F_x + \frac{1}{h} \text{div} \left(h\nu_e \overrightarrow{\text{grad}}(u) \right) \\ \frac{v^{n+1} - v^n}{\Delta t} + \vec{U}_{conv} \cdot \overrightarrow{\text{grad}}(v) &= -g\theta_h \frac{\partial(h^{n+1} - h^n)}{\partial y} - g \frac{\partial \psi_s^n}{\partial y} + F_y + \frac{1}{h} \text{div} \left(h\nu_e \overrightarrow{\text{grad}}(v) \right) \end{aligned} \right. \quad (2.50)$$

où les variables h , u et v sans indice de temps désignent les valeurs semi-implicites de la hauteur et des composantes de la vitesse évoquées précédemment dans cette section.

Discretisation en espace

En vue de la résolution du système d'équations (2.50) les fonctions scalaires apparaissant dans ce système seront par la suite discrétisées en espace. Nous donnons ici le principe d'une telle décomposition pour une fonction scalaire f quelconque. Cette décomposition est effectuée aux points de discrétisation (ou nœuds du maillage) selon une base de fonctions Ψ_i :

$$f = \sum_{i=1}^n f^i \Psi_i$$

où les f^i sont les valeurs de f aux points i et les Ψ_i sont des fonctions valant 1 au point i et 0 aux autres points. Les valeurs de la fonction f à l'intérieur d'un élément sont interpolées à partir des valeurs de cette fonction aux nœuds de l'élément.

2.5.4 Formulation faible, approximation interne et mise sous forme d'un système linéaire

Pour résoudre le système des équations de Saint-Venant on va chercher à discrétiser celui-ci et à le mettre sous forme d'un système linéaire que l'on sait résoudre. Pour cela on va, dans un premier temps, donner la formulation faible du problème. Celle-ci va nous mener à l'approximation interne qui consiste en la décomposition des fonctions inconnues

suivant les bases ce qui va nous permettre de mettre le système sous forme d'un système linéaire et ainsi de le résoudre facilement.

Formulation faible

On considère la *formulation forte* d'un problème :

$$\text{Trouver } u \text{ définie sur } \Omega \text{ vérifiant } R(u) = f \text{ en tout point de } \Omega \quad (2.51)$$

où R est un opérateur différentiel et f une fonction définie sur un domaine ouvert Ω .

La *formulation faible* (ou formulation variationnelle) du problème (2.51) est la suivante :

$$\text{Trouver } u \text{ définie sur } \Omega \text{ vérifiant } \int_{\Omega} R(u)v \, d\Omega = \int_{\Omega} f v \, d\Omega \text{ pour toute fonction } v \text{ définie sur } \Omega \quad (2.52)$$

En règle général les fonctions v sont appelées fonctions tests. Pour donner la formulation faible du système des équations de Saint-Venant on va multiplier chacune des équations de ce système par les fonctions tests et intégrer sur le domaine Ω . Dans la suite nous noterons :

- ϕ^h fonction test pour l'équation de conservation de la masse ;
- ϕ^u fonction test pour les deux équations de conservation de la quantité de mouvement.

Ces fonctions apprtiennent respectivement aux espaces $H_0^1(\Omega)$ et $L^2(\Omega)$. $L^2(\Omega)$ est l'ensemble des fonctions de carré intégrable sur Ω et $H_0^1(\Omega)$ est le sous-ensemble de $L^2(\Omega)$ des fonctions dont la dérivée est de carré intégrable sur Ω et qui s'annulent sur la frontière de Ω .

On commence par l'équation de conservation de la quantité de mouvement :

$$\int_{\Omega} \frac{h^{n+1} - h^n}{\Delta t} \phi^h \, d\Omega + \int_{\Omega} \vec{U}_{conv} \cdot \overrightarrow{\text{grad}}(h) \phi^h \, d\Omega + \int_{\Omega} h_{prop} \text{div}(\vec{U}) \phi^h \, d\Omega = \int_{\Omega} S_{ce} \phi^h \, d\Omega \quad (2.53)$$

Puis on fait la même chose pour chacune des deux équations de conservation de la quantité de mouvement :

$$\begin{aligned} \int_{\Omega} \frac{u^{n+1} - u^n}{\Delta t} \phi^u \, d\Omega &+ \int_{\Omega} \vec{U}_{conv} \cdot \overrightarrow{\text{grad}}(u) \phi^u \, d\Omega \\ &= - \int_{\Omega} g \theta_h \frac{\partial(h^{n+1} - h^n)}{\partial x} \phi^u \, d\Omega - \int_{\Omega} g \frac{\partial \psi_s^n}{\partial x} \phi^u \, d\Omega \\ &+ \int_{\Omega} F_x \phi^u \, d\Omega + \int_{\Omega} \frac{1}{h} \text{div} \left(h \nu_e \overrightarrow{\text{grad}}(u) \right) \phi^u \, d\Omega \end{aligned} \quad (2.54)$$

$$\begin{aligned} \int_{\Omega} \frac{v^{n+1} - v^n}{\Delta t} \phi^u \, d\Omega &+ \int_{\Omega} \vec{U}_{conv} \cdot \overrightarrow{\text{grad}}(v) \phi^u \, d\Omega \\ &= - \int_{\Omega} g \theta_h \frac{\partial(h^{n+1} - h^n)}{\partial y} \phi^u \, d\Omega - \int_{\Omega} g \frac{\partial \psi_s^n}{\partial y} \phi^u \, d\Omega \\ &+ \int_{\Omega} F_y \phi^u \, d\Omega + \int_{\Omega} \frac{1}{h} \text{div} \left(h \nu_e \overrightarrow{\text{grad}}(v) \right) \phi^u \, d\Omega \end{aligned} \quad (2.55)$$

Dans l'équation de conservation de la masse le terme $\int_{\Omega} h_{prop} \text{div}(\vec{U}) \phi^h d\Omega$ est intégré par parties et donne :

$$\int_{\Omega} h_{prop} \text{div}(\vec{U}) \phi^h d\Omega = \int_{\Gamma} h_{prop} \vec{U} \cdot \vec{n} \phi^h d\Gamma - \int_{\Omega} \vec{U} \cdot \overrightarrow{\text{grad}}(h_{prop} \phi^h) d\Omega \quad (2.56)$$

D'autre part les termes de diffusion $\int_{\Omega} \frac{1}{h} \text{div} \left(h \nu_e \overrightarrow{\text{grad}}(v) \right) \phi^u d\Omega$ sont simplifiés et traités sous la forme :

$$\int_{\Omega} \text{div} \left(\nu_e \overrightarrow{\text{grad}}(u) \phi^u \right) d\Omega \quad (2.57)$$

Le terme (2.57) est lui-même intégré par parties de manière à faire disparaître l'opérateur div et aboutir par la suite à une écriture qui nous permettra d'écrire ce terme comme combinaison linéaire des fonctions inconnues :

$$\int_{\Omega} \text{div} \left(\nu_e \overrightarrow{\text{grad}}(u) \phi^u \right) d\Omega = \int_{\Gamma} \phi^u \nu_e \overrightarrow{\text{grad}}(u) \cdot \vec{n} d\Gamma - \int_{\Omega} \nu_e \overrightarrow{\text{grad}}(u) \cdot \overrightarrow{\text{grad}}(\phi^u) d\Omega \quad (2.58)$$

Les termes faisant apparaître des intégrales de bord dans les expressions (2.56) et (2.58) sont les termes par lesquels sont traitées les conditions aux limites dans la résolution du système linéaire obtenu après avoir établi la formulation faible.

Approximation interne

L'approximation interne consiste à écrire les équations (2.53), (2.54) et (2.55) pour des sous-espaces de $H_0^1(\Omega)$ et $L^2(\Omega)$ de dimension finie. Pour cela nous allons décomposer les fonctions inconnues suivant les fonctions de base et utiliser les fonctions tests relatives aux points du maillage. Dans la suite nous utiliserons donc les notations suivantes :

- Ψ_i^h fonction de base relative au point i pour la hauteur ;
- ϕ_i^h fonction test relative au point i pour l'équation de conservation de la masse ;
- Ψ_i^u fonction de base relative au point i pour les composantes de la vitesse ;
- ϕ_i^u fonction test relative au point i pour les deux équations de conservation de la quantité de mouvement ;
- nph le nombre de fonctions de base portant la hauteur ;
- npu le nombre de fonctions de base portant la vitesse.

On écrit les fonctions h , u et v suivant leurs bases respectives :

$$h = \sum_{j=1}^{nph} h_j \Psi_j^h \quad u = \sum_{j=1}^{npu} u_j \Psi_j^u \quad v = \sum_{j=1}^{npu} v_j \Psi_j^u \quad (2.59)$$

On remplace alors dans les équations (2.53), (2.54) et (2.55), h , u et v par leurs expressions respectives données en (2.59).

Pour tout degré de liberté $i = 1, \dots, nph$ l'équation de conservation de la masse s'écrit :

$$\begin{aligned}
& \sum_{j=1}^{nph} \left(\frac{h_j^{n+1} - h_j^n}{\Delta t} \right) \int_{\Omega} \Psi_j^h \phi_i^h d\Omega + \sum_{j=1}^{nph} h_j \int_{\Omega} \vec{U}_{conv} \overrightarrow{\text{grad}}(\Psi_j^h) \phi_i^h d\Omega \\
& + \int_{\Gamma} h_{prop} \vec{U} \cdot \vec{n} \phi_i^h d\Gamma - \sum_{j=1}^{nph} \vec{U}_{h,j} \int_{\Omega} \Psi_j^u \overrightarrow{\text{grad}}(h_{prop} \phi_i^h) d\Omega \\
& = \sum_{j=1}^{nph} \int_{\Omega} \Psi_j^h \phi_i^h d\Omega
\end{aligned} \tag{2.60}$$

Pour tout degré de liberté $i = 1, \dots, npu$, la première équation de conservation de la quantité de mouvement s'écrit :

$$\begin{aligned}
& \sum_{j=1}^{npu} \left(\frac{u_j^{n+1} - u_j^n}{\Delta t} \right) \int_{\Omega} \Psi_j^u \phi_i^u d\Omega + \sum_{j=1}^{npu} u_j \int_{\Omega} \vec{U}_{conv} \overrightarrow{\text{grad}}(\Psi_j^u) \phi_i^u d\Omega \\
& = - \sum_{j=1}^{nph} (h_j^{n+1} - h_j^n) \int_{\Omega} \theta_h g \frac{\partial \Psi_j^h}{\partial x} \phi_i^u d\Omega - \int_{\Omega} g \frac{\partial \psi_s^n}{\partial x} \phi_i^u d\Omega + \int_{\Omega} F_x \phi_i^u d\Omega \\
& + \int_{\Gamma} \phi_i^u \nu_e \overrightarrow{\text{grad}}(u) \cdot \vec{n} d\Gamma - \sum_{j=1}^{npu} u_j \int_{\Omega} \nu_e \overrightarrow{\text{grad}}(\Psi_j^u) \cdot \overrightarrow{\text{grad}}(\phi_i^u) d\Omega
\end{aligned} \tag{2.61}$$

Pour tout degré de liberté $i = 1, \dots, npu$, la seconde équation de conservation de la quantité de mouvement s'écrit :

$$\begin{aligned}
& \sum_{j=1}^{npu} \left(\frac{v_j^{n+1} - v_j^n}{\Delta t} \right) \int_{\Omega} \Psi_j^u \phi_i^u d\Omega + \sum_{j=1}^{npu} v_j \int_{\Omega} \vec{U}_{conv} \overrightarrow{\text{grad}}(\Psi_j^u) \phi_i^u d\Omega \\
& = - \sum_{j=1}^{nph} (h_j^{n+1} - h_j^n) \int_{\Omega} \theta_h g \frac{\partial \Psi_j^h}{\partial y} \phi_i^u d\Omega - \int_{\Omega} g \frac{\partial \psi_s^n}{\partial y} \phi_i^u d\Omega + \int_{\Omega} F_y \phi_i^u d\Omega \\
& + \int_{\Gamma} \phi_i^u \nu_e \overrightarrow{\text{grad}}(v) \cdot \vec{n} d\Gamma - \sum_{j=1}^{npu} v_j \int_{\Omega} \nu_e \overrightarrow{\text{grad}}(\Psi_j^u) \cdot \overrightarrow{\text{grad}}(\phi_i^u) d\Omega
\end{aligned} \tag{2.62}$$

Traitement des conditions aux limites

Des termes de bord apparaissent dans les équations (2.60), (2.61) et (2.62).

Le terme de bord apparaissant dans l'équation de conservation de la masse est :

$$\int_{\Gamma} h_{prop} \vec{U} \cdot \vec{n} \phi_i^h d\Gamma \tag{2.63}$$

Celui-ci correspond à un flux à travers une paroi. Dans notre cas toutes les parois sont solides et alors le terme (2.63) est annulé. Toutes les parois étant solides nous ne nous intéresserons pas non plus au cas du traitement de frontières ouvertes.

Les termes de bord apparaissant dans les équations de conservation de la quantité de mouvement sont :

$$\int_{\Gamma} \phi_i^u \nu_e \overrightarrow{\text{grad}}(u) \cdot \vec{n} d\Gamma \quad (2.64)$$

$$\int_{\Gamma} \phi_i^u \nu_e \overrightarrow{\text{grad}}(v) \cdot \vec{n} d\Gamma$$

Or par définition :

$$\overrightarrow{\text{grad}}(f) \cdot \vec{n} = \frac{\partial f}{\partial n}$$

Ce terme est traité implicitement sous la forme :

$$\frac{\partial u_t}{\partial n} = a u_t$$

où u_t désigne la vitesse tangentielle, $a = -\frac{(u^*)^2}{u_t^n \nu_e}$ et $u^* = \sqrt{\frac{C_f}{2} (u^2 + v^2)}$ et C_f est le coefficient de frottement dont l'expression est donnée en (2.29).

Le terme (2.64) s'écrit alors :

$$\int_{\Gamma} \phi_i^u \nu_e a u d\Gamma$$

et il est traité implicitement.

Mise sous forme d'un système linéaire

Du fait du traitement des conditions aux limites présenté ci-dessus les équations (2.53), (2.54) et (2.55) peuvent être écrites comme une combinaison linéaire des inconnues et ainsi il est possible de mettre le système constitué par ces trois équations sous forme d'un système linéaire $AX = B$. Nous ne rentrerons pas dans le détail des matrices blocs formant les matrices A et B , le lecteur pourra se reporter à [Hervouet, 2007].

Le système linéaire $AX = B$ est résolu en utilisant un algorithme itératif de type gradient conjugué.

Chapitre 3

L'assimilation de données

Sommaire

3.1	Une introduction	62
3.2	Le vocabulaire de l'assimilation de données	62
3.3	La modélisation des erreurs et les matrices de covariance d'erreur	63
3.3.1	Remarques préliminaires sur les erreurs, les matrices de covariance et de corrélation d'erreur	64
3.3.2	Les erreurs liées à l'état du modèle	65
3.3.3	Les erreurs liées à la modélisation	66
3.3.4	Les erreurs liées aux observations	66
3.4	Filtrage bayésien	67
3.4.1	Règle de Bayes	67
3.4.2	Estimation optimale	68
3.4.3	Formule de Bayes récursive ou filtre bayésien	69
3.4.4	Considération sur le filtrage bayésien et le choix de l'estimateur optimal	70
3.5	Le Best Linear Unbiased Estimator - BLUE	70
3.6	Les méthodes variationnelles	72
3.6.1	Formulation de la fonction coût	72
3.6.2	Le 3D-Var	73
3.6.3	Le 4D-Var	74
3.6.4	Quelques variantes du 3D-Var et du 4D-Var	77
3.7	Le filtre de Kalman et ses dérivations	80
3.7.1	Le filtre de Kalman	80
3.7.2	Le filtre de Kalman Etendu ou Extended Kalman Filter (EKF)	81
3.7.3	Le filtre de Kalman de rang réduit - RRSQRT	82
3.7.4	Le filtre SEEK	83
3.7.5	Le filtre de Kalman d'ensemble - EnKF	84
3.7.6	Les versions déterministes du filtre de Kalman d'ensemble	86
3.7.7	Les problèmes d'échantillonnage et les solutions liées	88

3.1 Une introduction

Il peut être difficile de définir d'emblée ce qu'est l'assimilation de données. On peut considérer qu'il s'agit de la science des compromis. Compromis entre un modèle et des données d'observation. Le modèle représente un système physique et éventuellement son évolution dans le temps. Cela peut être un modèle d'écoulement d'une rivière modélisé par les équations de Saint-Venant, un modèle océanographique ou météorologique (domaines desquels est issue originellement l'assimilation de données) ou encore un modèle agronomique, hydrologique... Cependant par nature ces modèles ne sont pas exacts, du fait de la discrétisation du modèle, de certaines hypothèses d'approximation, du fait de certains paramètres mal connus entrant en compte dans la modélisation. D'autre part les observations représentent les données observées du système. Elles peuvent être de natures très différentes : données *in situ*, comme c'est par exemple le cas d'observations de concentrations de certains éléments chimiques dans l'air, données radars ou encore données satellitaires. Néanmoins celles-ci, de manière analogue aux modèles, peuvent être entachées d'erreurs. L'objet de l'assimilation est alors de combiner ces deux sources d'information en prenant en compte, si cela est possible, les incertitudes qui leurs sont liées pour donner la meilleure estimation possible de l'état du système (en un sens que nous préciserons ultérieurement). Cela a notamment deux applications directes. Tout d'abord la prévision, en effet donner la meilleure estimation possible de l'état du système permet d'améliorer l'état prévu du système à plus ou moins longue échéance. D'autre part cela permet la réanalyse des données passées de manière à donner la meilleure estimation de la trajectoire d'un système au cours du temps.

Parmi les grandes familles de méthodes d'assimilation de données on peut citer d'une part les méthodes basées sur l'estimation stochastique comme le filtre de Kalman et les méthodes qui en dérivent : filtre de Kalman étendu, filtre de Kalman d'ensemble... D'autre part on peut citer les méthodes variationnelles dans lesquelles la meilleure estimation de l'état du système est obtenue en minimisant une certaine fonctionnelle. Ces deux types de méthodes ont néanmoins ceci de commun qu'elles peuvent être dérivées du filtrage bayésien selon que l'on cherche à déterminer l'espérance conditionnelle d'une variable aléatoire (cas des filtres de Kalman) ou son mode (assimilation variationnelle). Nous reviendrons en détail sur chacune de ces familles plus loin dans ce manuscrit. Bien que ces familles soient différentes il existe des ponts entre elles et elles ont notamment un vocabulaire commun que nous présentons dans la section suivante.

3.2 Le vocabulaire de l'assimilation de données

On note tout d'abord $\tilde{x}$ l'état réel du système physique que l'on considère à un instant donné. Cet état réel est continu et a de fait un nombre infini de degrés de liberté. Il est à noter que cet état réel peut être multivarié et représenter par exemple la température et la salinité de l'océan à un instant donné et en tout point. Les modèles numériques qui résolvent les équations modélisant le système physique considéré dans un espace discrétisé ayant un nombre fini de points ont quant à eux un nombre fini de degré de liberté. On appelle alors état vrai et on note x^t (t pour true) la projection de $\tilde{x}$ dans l'espace discrétisé du modèle. Si on note Π cet opérateur de projection alors $\tilde{x}$ et x^t sont liés par la relation :

$$x^t = \Pi(\tilde{x})$$

On appelle vecteur d'état et on note x le vecteur qui contient l'état estimé du système à un instant donné dans l'espace du modèle. Par exemple dans le cas des équations de Saint-Venant le vecteur d'état est constitué de la discrétisation des hauteurs d'eau et des débits sur le domaine.

On définit le vecteur de contrôle comme étant l'ensemble des variables d'intérêt que l'on cherche à corriger avec l'assimilation de données. Dans l'exemple précédent des équations de Saint-Venant le vecteur de contrôle peut être le vecteur d'état. Mais on pourrait également chercher à corriger le coefficient de Strickler avec l'assimilation de données. Plus généralement le vecteur de contrôle peut être constitué des variables d'état, ou des paramètres du modèle, des conditions limites...

Dans la suite on notera $\mathcal{M}$ le modèle numérique résolvant les équations mathématiques du système physique que l'on considère. Celui-ci nous permet de décrire l'évolution temporelle de l'état du système. Si on note x_k et x_{k+1} l'estimation de l'état du système respectivement aux temps t_k et t_{k+1} alors en l'absence d'erreur modèle (voir section 3.3.3) ceux-ci sont reliés par :

$$x_{k+1} = \mathcal{M}_{k,k+1}(x_k) \tag{3.1}$$

où $\mathcal{M}_{k,k+1}$ désigne l'évolution du modèle entre les temps t_k et t_{k+1} .

On peut maintenant considérer les différents statuts du vecteur x :

- si x est l'état estimé du système consécutif à l'assimilation d'une donnée d'observation on appellera celui-ci état analysé et on le notera x^a ;
- si x est une estimation *a priori* de l'état du système avant assimilation on appelle x ébauche et on le notera x^b (b pour background). Celui-ci peut être obtenu comme la propagation par le modèle d'un état analysé du temps t_k au temps t_{k+1} , soit : $x_{k+1}^b = \mathcal{M}_{k,k+1}(x_k^a)$;
- si x est une estimation du système obtenue par la propagation à plus ou moins longue échéance par le modèle d'un état analysé alors on appellera x prévision et on notera x^f (f pour forecast).

3.3 La modélisation des erreurs et les matrices de covariance d'erreur

Comme on l'a précisé dans l'introduction l'objet de l'assimilation est de combiner les informations fournies par un modèle et celles fournies par des observations en prenant en compte les erreurs qui sont liées. Celles-ci occupent donc une place importante en assimilation de données. Nous donnons dans cette section la définition des différents types d'erreur que nous pouvons rencontrer en assimilation de données ainsi que la définition des matrices de covariance d'erreur qui leurs sont associées et qui apparaissent dans les différentes méthodes d'assimilation.

On définit tout d'abord les erreurs comme étant la différence entre la grandeur considérée et l'état vrai. Lui-même étant inconnu et accessible de manière indirecte à travers les données d'observation, elles-mêmes entachées d'erreur.

En général on distingue trois types d'erreur liées aux modèles numériques [Cohn, 1997] : l'erreur liée à l'état du modèle, l'erreur liée à la modélisation du système physique (discrétisation, choix des paramètres...) et l'erreur de représentativité du modèle. Mais avant de présenter les différents types d'erreur nous présentons quelques remarques qui nous seront utiles par la suite.

3.3.1 Remarques préliminaires sur les erreurs, les matrices de covariance et de corrélation d'erreur

Dans la suite de ce travail dans un souci de simplicité et sauf mention contraire on considérera que les différentes erreurs rencontrées suivent une loi normale de moyenne nulle. C'est la raison pour laquelle dans les sections suivantes lorsque nous présenterons les différents types d'erreurs nous ne nous intéresserons qu'aux moments d'ordre 2 de ces variables i-e les matrices de covariance.

Dans le cas d'un **problème univarié** où les composantes de l'erreur $\varepsilon = (\varepsilon_1, \dots, \varepsilon_N)$ sont des variables aléatoires indépendantes alors l'erreur admet comme matrice de covariance une matrice diagonale avec sur celle-ci les variances relatives à chaque composante de l'erreur. Dans le cas où les composantes de l'erreur ne sont pas des variables aléatoires indépendantes celle-ci admet comme matrice de covariance d'erreur $\mathbf{P}$:

$$\mathbf{P} = E[\varepsilon\varepsilon^T] = \begin{pmatrix} E[\varepsilon_1\varepsilon_1] & \dots & E[\varepsilon_1\varepsilon_N] \\ \vdots & \ddots & \vdots \\ E[\varepsilon_N\varepsilon_1] & \dots & E[\varepsilon_N\varepsilon_N] \end{pmatrix}$$

où E désigne l'espérance mathématique.

Dans le cas d'un **problème multivarié** où les variables sont indépendantes alors la matrice des covariances d'erreur se réduit à une matrice diagonale par blocs où chaque bloc décrit les covariances univariées propres à chaque variable. Si les erreurs relatives à chacune des différentes variables ne sont pas indépendantes alors la matrice des covariances d'erreur n'est pas diagonale par blocs. Dans ce cas les blocs diagonaux indiquent toujours les covariances univariées tandis que les blocs extra-diagonaux représentent les covariances entre les erreurs relatives à chaque variable. Supposons par exemple que l'on a 3 variables a , b et c de vecteurs d'erreur respectifs ε^a , ε^b et ε^c , alors la matrice des covariances d'erreur s'écrit :

$$\mathbf{P} = E[\varepsilon\varepsilon^T] = \begin{pmatrix} \mathbf{P}^a & E[\varepsilon^a\varepsilon^b] & E[\varepsilon^a\varepsilon^c] \\ E[\varepsilon^b\varepsilon^a] & \mathbf{P}^b & E[\varepsilon^b\varepsilon^c] \\ E[\varepsilon^c\varepsilon^a] & E[\varepsilon^c\varepsilon^b] & \mathbf{P}^c \end{pmatrix}$$

D'autre part si on a la connaissance des variances sur l'ensemble du domaine ainsi que de la matrice des corrélations d'erreur $\mathbf{C}$ alors la matrice des covariances d'erreur peut se déduire de la matrice des corrélations d'erreur par la formule suivante :

$$\mathbf{P} = \Sigma^{1/2} \mathbf{C} \Sigma^{T/2}$$

où Σ est une matrice diagonale avec sur celle-ci les variances des erreurs. De manière équivalente les corrélations d'erreurs peuvent se déduire des covariances d'erreur par la relation :

$$\mathbf{C} = \Sigma^{-1/2} \mathbf{P} \Sigma^{-T/2} \quad (3.2)$$

Dans ce travail nous nous sommes intéressés aux fonctions de covariances d'erreur relatives à certains points. Si on considère un point en particulier ces fonctions nous intéressent car elles décrivent les covariances entre l'erreur à ce point et les erreurs aux autres points du domaine mais aussi les erreurs relatives aux autres variables aux autres points du domaine. Lorsque l'on connaît les matrices de covariance d'erreur on connaît les fonctions de covariance, en effet celles-ci sont les colonnes et les lignes de la matrice : la fonction de covariance relative au i -ème point du domaine est contenue dans la i -ème colonne de la matrice $\mathbf{P}$ (mais aussi dans la i -ème ligne puisque par construction $\mathbf{P}$ est symétrique). Ce fait a été présenté ici pour les matrices et fonctions de covariance mais il reste valable pour les matrices et fonctions de corrélation.

3.3.2 Les erreurs liées à l'état du modèle

L'erreur liée à l'état du modèle représente la différence entre l'état estimé du modèle x et l'état vrai. On définit ainsi, selon que l'on considère l'ébauche, l'état analysé ou prévu, trois erreurs liées à l'état du modèle.

Si on considère l'erreur d'ébauche ε^b à un temps donné on notera :

$$\varepsilon^b = x^b - x^t$$

Dans la foulée on définit la matrice des covariance d'erreur d'ébauche :

$$\mathbf{B} = \mathbf{E} [(\varepsilon^b - \mathbf{E}[\varepsilon^b])(\varepsilon^b - \mathbf{E}[\varepsilon^b])^T]$$

De manière analogue on note ε^a l'erreur d'analyse à un temps donné :

$$\varepsilon^a = x^a - x^t$$

Et on définit la matrice des covariances d'erreur d'analyse :

$$\mathbf{A} = \mathbf{E} [(\varepsilon^a - \mathbf{E}[\varepsilon^a])(\varepsilon^a - \mathbf{E}[\varepsilon^a])^T]$$

Enfin on définit la matrice des covariances d'erreur de prévision :

$$\varepsilon^f = x^f - x^t$$

Et on définit la matrice des covariances d'erreur de prévision :

$$\mathbf{F} = \mathbf{E} [(\varepsilon^f - \mathbf{E}[\varepsilon^f])(\varepsilon^f - \mathbf{E}[\varepsilon^f])^T]$$

3.3.3 Les erreurs liées à la modélisation

On distingue généralement quatre grands types d'erreurs liées à la modélisation :

- les erreurs liées à certaines approximations physiques comme c'est par exemple dans les équations de Saint-Venant qui sont dérivées des équations de Navier-Stokes sous certaines hypothèses simplificatrices ;
- les erreurs liées aux choix de certains paramètres. Un bon exemple de ce type d'erreur réside dans la modélisation hydraulique avec les équations de Saint-Venant dans lesquelles le frottement peut être représenté avec un coefficient appelé coefficient de Strickler. Dans le cas de la modélisation de l'écoulement d'une rivière avec ces équations il peut être difficile de donner avec précision la valeur du coefficient de Strickler. En règle générale celui-ci est calé par les modélisateurs sur une série de crues de références. Mais il peut également être estimé au mieux en utilisant l'assimilation de données ;
- les erreurs liées à la discrétisation. Approcher un problème continu par un problème discrétisé amène inévitablement des erreurs. Celles-ci seront d'autant plus importantes que la discrétisation est lâche ;
- les erreurs sur les forçages. En effet ceux-ci peuvent être mal connus comme cela peut être par exemple le cas avec un modèle hydraulique où les apports latéraux non jaugés peuvent être très importants lors des pics de crue mais qui ne sont pas pris en compte dans la modélisation car on a *a priori* aucune information dessus.

Ces erreurs étant liées entre elles de manière indissociable on modélise les erreurs liées à la modélisation à l'instant k comme une variable aléatoire q_k telle que :

$$q_k = x_k^t - \mathcal{M}_{k-1,k}(x_{k-1}^t)$$

De manière analogue à ce que l'on a fait précédemment on définit la matrice des covariances d'erreur de modélisation $\mathbf{Q}$:

$$\mathbf{Q}_k = \mathbf{E} [(q_k - \mathbf{E}[q_k])(q_k - \mathbf{E}[q_k])^T]$$

3.3.4 Les erreurs liées aux observations

Les erreurs liées aux observations sont de deux ordres. On distingue tout d'abord les erreurs liées aux instruments de mesure. Ceux-ci sont par nature imparfaits et de fait toute mesure (ou observation) est entachée d'erreur. Si on note h l'opérateur de projection de l'état vrai continu $\tilde{x}$ dans l'espace des observations alors on définit l'erreur de mesure comme étant la différence entre le vecteur des observations y^o et $h[\tilde{x}]$ soit :

$$\varepsilon^m = y^o - h[\tilde{x}] \quad (3.3)$$

En pratique, dans le cas des observations de hauteur d'eau d'une rivière, on peut estimer de manière statistique l'erreur de mesure en faisant plusieurs relevés dans lesquels on compare la hauteur d'eau mesurée par l'instrument de mesure et la hauteur d'eau affichée à l'échelle lorsqu'une station en est équipée.

On distingue d'autre part les erreurs de représentativité. Celles-ci sont liées à la modélisation du système physique que l'on considère (en ce sens elles sont à rapprocher des erreurs de modélisation). En effet si on note $\mathcal{H}$ l'opérateur d'observation discret

qui est la projection dans l'espace des observations de l'état vrai alors on définit l'erreur de représentativité ε^r comme étant la différence :

$$\varepsilon^r = h[\tilde{x}] - \mathcal{H}(x^t) \quad (3.4)$$

$$\varepsilon^r = h[\tilde{x}] - \mathcal{H}(\Pi[\tilde{x}]) \quad (3.5)$$

En combinant (3.3) et (3.4) on peut écrire :

$$\varepsilon^m + \varepsilon^r = y^o - \mathcal{H}(x^t)$$

On définit alors l'erreur d'observation comme étant la somme des erreurs de mesure et de représentativité, soit :

$$\varepsilon^o = \varepsilon^m + \varepsilon^r$$

Et on définit la matrice des covariances d'erreur d'observation :

$$\mathbf{R} = \mathbf{E} [(\varepsilon^o - \mathbf{E}[\varepsilon^o])(\varepsilon^o - \mathbf{E}[\varepsilon^o])^T]$$

Maintenant que nous avons défini les différents types d'erreur nous allons exposer dans les grandes lignes les principes du filtrage bayésien.

3.4 Filtrage bayésien

Nous présentons dans cette section les principes du filtrage bayésien [Doucet et al., 2000] à partir duquel on va définir certains estimateurs qui sont à l'origine de la distinction entre les méthodes variationnelles et les méthodes basées sur l'interpolation stochastique.

3.4.1 Règle de Bayes

On considère deux variables aléatoires x et y dont on suppose qu'elles admettent chacune une densité de probabilité notée respectivement : $p(x)$ et $p(y)$. On peut définir les densités de probabilité conditionnelle $p(x|y)$ et $p(y|x)$ comme étant les quotients :

$$p(x|y) = \frac{p(x, y)}{p(y)} \quad (3.6)$$

$$p(y|x) = \frac{p(x, y)}{p(x)} \quad (3.7)$$

où $p(x, y)$ désigne la densité de probabilité conjointe des variables x et y . En combinant les égalités (3.6) et (3.7) on peut relier $p(x|y)$ et $p(y|x)$:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)} \quad (3.8)$$

La relation (3.8) est appelée règle de Bayes.

3.4.2 Estimation optimale

On considère maintenant l'état du modèle x comme une variable aléatoire de densité de probabilité $p(x)$ dont on suppose qu'elle est connue et une observation y comme une variable aléatoire de densité de probabilité $p(y)$. x et y sont liés par la relation :

$$y = \mathcal{H}(x) + \varepsilon^o \quad (3.9)$$

où ε^o est une variable aléatoire modélisant les erreurs d'observation dont la densité de probabilité est supposée connue.

On note tout d'abord que :

$$p(y) = \int p(x, y) dx = \int p(y|x)p(x) dx$$

Alors la règle de Bayes peut se réécrire :

$$p(x|y) = \frac{p(y|x)p(x)}{\int p(y|x)p(x) dx} \quad (3.10)$$

Donc la densité conditionnelle $p(x|y)$ est connue puisque $p(x)$ est connue ainsi que $p(y|x)$ en vertu de la relation (3.9) et du fait que les densités de x et ε^o le sont. Ainsi l'égalité (3.10) permet de connaître la densité conditionnelle $p(x|y)$ et de donner une estimation « optimale » de x connaissant une réalisation de y . L'expression « estimation optimale » est à mettre en regard du critère d'optimalité choisi, il en existe plusieurs, nous en présentons ci-après trois.

L'estimateur de variance minimum

L'estimateur de variance minimale est celui qui minimise la variance de l'état analysé sachant y . On peut montrer que minimiser la variance de l'état analysé sachant y est équivalent à résoudre l'équation suivante :

$$\frac{\partial \mathcal{J}(\hat{x})}{\partial \hat{x}} = 0$$

où $\mathcal{J}(\hat{x})$ est donné par l'expression :

$$\mathcal{J}(\hat{x}) = \int (x - \hat{x})^T (x - \hat{x}) p(x|y) dx$$

Or on peut montrer également :

$$\frac{\partial \mathcal{J}(\hat{x})}{\partial \hat{x}} = 2 \int (x - \hat{x}) p(x|y) dx \quad (3.11)$$

On déduit alors de (3.11) que l'état analysé minimisant la variance est donnée par :

$$x^a = \int x p(x|y) dx = \mathbb{E}[x|y]$$

L'estimateur du maximum *a posteriori*

L'estimateur du maximum *a posteriori* est l'état le plus probable de x sachant y , il maximise donc la densité conditionnelle $p(x|y)$ et par suite vérifie la relation :

$$\frac{\partial p(x|y)}{\partial x} = 0$$

Dans le cas où la densité de probabilité $p(x|y)$ est gaussienne alors les estimateurs de variance minimum et de maximum *a posteriori* sont équivalents.

L'estimateur du maximum de vraisemblance

L'estimateur du maximum de vraisemblance est l'état le plus probable de y sachant x , il maximise la densité conditionnelle $p(y|x)$ et par suite vérifie la relation :

$$\frac{\partial p(y|x)}{\partial x} = 0$$

Dans le cas où la densité de probabilité $p(y|x)$ est gaussienne alors les estimateurs de variance minimum et de maximum de vraisemblance sont équivalents.

3.4.3 Formule de Bayes récursive ou filtre bayésien

Les résultats exposés à la section (3.4.2) sont valables pour une seule observation. Néanmoins ils peuvent être étendus au cas temporel [Doucet et al., 2000]. En effet on considère une succession d'observations $y_0^o, y_1^o, \dots, y_k^o$ prises aux temps $t_0, t_1, \dots, t_k$ respectivement. On considère également une succession d'états du modèle $x_0, x_1, \dots, x_k$. On suppose de plus que la dépendance entre l'état du modèle x_k au temps t_k et les autres états du modèles aux temps précédents se limite à l'état du modèle x_{k-1} au temps précédent t_{k-1} et on appelle loi de transition la densité conditionnelle $p(x_k|x_{k-1})$.

La formule de Bayes récursive se propose d'estimer de manière récursive, au fur et à mesure que des observations sont disponibles, les états du modèle x_k . Pour cela le filtre bayésien se décompose en deux étapes, une première étape dite de prédiction dans laquelle on cherche à déterminer le filtre prédit c'est-à-dire la densité conditionnelle $p(x_k|y_0^o, \dots, y_{k-1}^o)$. Or sous ces hypothèses on peut montrer que $p(x_k|y_0^o, \dots, y_{k-1}^o)$ a pour expression :

$$p(x_k|y_0^o, \dots, y_{k-1}^o) = \int p(x_k|x_{k-1})p(x_{k-1}|y_0^o, \dots, y_{k-1}^o)dx_{k-1} \quad (3.12)$$

Une seconde étape dite de correction dans laquelle on estime le filtre $p(x_k|y_0^o, \dots, y_k^o)$ à partir du filtre prédit $p(x_k|y_0^o, \dots, y_{k-1}^o)$:

$$p(x_k|y_0^o, \dots, y_k^o) = \frac{p(y_k|x_k)p(x_k|y_0^o, \dots, y_{k-1}^o)}{\int p(y_k|x_k)p(x_k|y_0^o, \dots, y_{k-1}^o)dx_k} \quad (3.13)$$

3.4.4 Considération sur le filtrage bayésien et le choix de l'estimateur optimal

Comme on vient de le voir le but du filtrage bayésien est de calculer la densité de probabilité de la variable x sachant y . Dans les faits il n'est pas toujours possible de calculer la densité de $p(x|y)$. C'est pourquoi en règle générale les méthodes d'assimilation de données ne s'intéressent qu'aux deux premiers moments de la variable aléatoire représentant le système considéré i-e la moyenne et la variance. Néanmoins on peut dans une certaine mesure dépasser cette limitation en approximant la densité de probabilité $p(x|y)$, c'est l'objet des méthodes dites de filtre à particules que nous n'aborderons pas dans ce travail.

Sinon on peut chercher à estimer l'état optimal x ; celui-ci n'est pas unique et va dépendre de la définition qu'on s'en donne. Par exemple si l'on considère l'état optimal comme étant l'espérance conditionnelle de la variable x connaissant l'observation y . Dans ce cas l'état optimal est donné par l'estimateur de variance minimum. Cet estimateur est à l'origine des méthodes de type interpolation statistique comme le BLUE, le filtre de Kalman ainsi que les formes qui en sont dérivées (exposées en section 3.7.1) : filtre de Kalman de rang réduit, filtre de Kalman d'ensemble... A noter que dans le cas où les densités sont gaussiennes [Lorenc, 1986, Cohn, 1997] ont montré que le filtre de Kalman et le filtre Bayésien sont équivalents. Si on considère l'état optimal comme étant le plus probable celui-ci nous est donné par l'estimateur du maximum *a posteriori*. Cet estimateur est quant à lui à l'origine des méthodes dites variationnelles exposées en section 3.6. Il faut souligner que sous l'hypothèse que le modèle considéré est linéaire et les densités conditionnelles sont gaussiennes alors ces deux estimateurs sont équivalents [Jazwinski, 2007, Lorenc, 1986, Cohn, 1997].

3.5 Le Best Linear Unbiased Estimator - BLUE

On souhaite construire un estimateur linéaire non biaisé qui serait le meilleur estimateur au sens où il minimise la variance de l'état analysé et dans lequel celui-ci serait décrit comme la combinaison linéaire d'une estimation *a priori* de notre système, une ébauche notée x^b et d'une observation y^o [Gelb, 1974].

On suppose qu'un tel estimateur existe et que les erreurs d'ébauche et d'observation sont non biaisées (i-e $E[\varepsilon^b] = 0$ et $E[\varepsilon^o] = 0$), on l'écrit :

$$x^a = \mathbf{L}x^b + \mathbf{K}y^o \quad (3.14)$$

On suppose de plus que l'opérateur d'observation $\mathcal{H}$ est linéaire et on écrit $\mathbf{H}$. D'après les définitions des erreurs que nous avons données précédemment nous pouvons écrire :

$$x^a = x^t + \varepsilon^a \quad (3.15)$$

$$x^b = x^t + \varepsilon^b \quad (3.16)$$

$$y^t = \mathbf{H}x^t \quad (3.17)$$

$$y^o = \mathbf{H}x^t + \varepsilon^o \quad (3.18)$$

Alors en remplaçant les termes de l'équation (3.14) par les expressions données en (3.15)-(3.18) on peut écrire :

$$x^t + \varepsilon^a = \mathbf{L}x^t + \mathbf{L}\varepsilon^b + \mathbf{K}\mathbf{H}x^t + \mathbf{K}\varepsilon^o \quad (3.19)$$

Ainsi en considérant l'espérance mathématique de chacun des termes de l'équation (3.19) on peut montrer que :

$$\mathbf{L} = \mathbf{I} - \mathbf{K}\mathbf{H} \quad (3.20)$$

On peut alors réécrire (3.14) sous la forme suivante :

$$x^a = (\mathbf{I} - \mathbf{K}\mathbf{H})x^b + \mathbf{K}y^o \quad (3.21)$$

$$x^a = x^b + \mathbf{K}(y^o - \mathbf{H}x^b) \quad (3.22)$$

Cette écriture est valable quelle que soit la matrice $\mathbf{K}$ or on souhaite que cet estimateur minimise la variance de l'état analysé. Minimiser la variance de l'état analysé est équivalent à minimiser la trace de la matrice des covariances d'erreur d'analyse, i-e $\text{Tr}(\mathbf{A})$. On cherche donc la matrice $\mathbf{K}$ (également appelée matrice de gain) qui minimise $\text{Tr}(\mathbf{A})$.

D'après (3.21) l'erreur d'analyse peut s'exprimer ainsi :

$$\varepsilon^a = (\mathbf{I} - \mathbf{K}\mathbf{H})\varepsilon^b + \mathbf{K}\varepsilon^o \quad (3.23)$$

Ecriture de laquelle on peut déduire celle de $\mathbf{A}$:

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}(\mathbf{I} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T \quad (3.24)$$

puis celle de $\text{Tr}(\mathbf{A})$:

$$\text{Tr}(\mathbf{A}) = \text{Tr}(\mathbf{B}) + \text{Tr}(\mathbf{K}\mathbf{H}\mathbf{B}\mathbf{H}^T\mathbf{K}^T) - 2\text{Tr}(\mathbf{B}\mathbf{H}^T\mathbf{K}^T) + \text{Tr}(\mathbf{K}\mathbf{R}\mathbf{K}^T) \quad (3.25)$$

On s'intéresse alors à la différence $d_{\mathbf{K}}[\text{Tr}(\mathbf{A})] \mathbf{M} = \text{Tr}(\mathbf{A})(\mathbf{K} + \mathbf{M}) - \text{Tr}(\mathbf{A})(\mathbf{K})$ où $\mathbf{M}$ est une matrice quelconque. On peut montrer que (voir [?]) :

$$d_{\mathbf{K}}[\text{Tr}(\mathbf{A})] \mathbf{M} = 2\text{Tr}(\{\mathbf{K}(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}) - \mathbf{B}\mathbf{H}^T\} \mathbf{M}^T) \quad \forall \mathbf{M} \quad (3.26)$$

On déduit de (3.26) que la matrice $\mathbf{K}$ qui minimise la variance de l'état analysé est donnée par l'expression :

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} \quad (3.27)$$

En injectant (3.27) dans l'égalité (3.22) on obtient l'expression de l'état analysé pour lequel la variance est minimale :

$$x^a = x^b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(y^o - \mathbf{H}x^b)$$

Dans ce cas la matrice des covariances d'erreur de l'état analysé s'écrit :

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B} \quad (3.28)$$

On peut alors établir un lien entre l'estimateur BLUE et l'estimateur de variance minimum présenté en section (3.4.2) : si x et y sont des variables aléatoires gaussiennes alors la densité de probabilité $p(x|y)$ est une gaussienne de moyenne

$$E[x|y] = x^a = x^b + \mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} (y^o - \mathbf{H}x^b)$$

et de matrice de covariance égale à $\mathbf{A}$.

3.6 Les méthodes variationnelles

Les méthodes variationnelles, en ce qu'elles visent à établir l'état le plus probable x du système sachant l'observation y^o , sont basées sur l'estimation du maximum *a posteriori*. Dans cette section on montre dans un premier temps comment sous l'hypothèse de gaussianité des variables considérées l'estimation du maximum *a posteriori* peut se ramener à la minimisation d'une fonctionnelle $\mathcal{J}$ dont on établit l'expression ci-dessous. Ensuite on présente les différentes méthodes variationnelles parmi les plus utilisées : 3D-Var, 4D-Var ainsi que deux méthodes qui en sont dérivées à savoir : le 3D-FGAT et le 4D-Var incrémental.

3.6.1 Formulation de la fonction coût

Supposons que le vecteur d'ébauche x^b suit une loi normale de matrice de covariance $\mathbf{B}$ alors la densité de probabilité de l'ébauche vérifie la relation :

$$p(x) \propto \exp \left\{ -\frac{1}{2} (x - x^b)^T \mathbf{B}^{-1} (x - x^b) \right\}$$

D'autre part si la densité de probabilité de la vraisemblance des observations est gaussienne de matrice de covariance d'erreur $\mathbf{R}$ alors :

$$p(y^o|x) \propto \exp \left\{ -\frac{1}{2} (y^o - \mathcal{H}(x))^T \mathbf{R}^{-1} (y^o - \mathcal{H}(x)) \right\}$$

D'après la règle de Bayes établie en (3.8) il vient que la densité de probabilité conditionnelle de x sachant y^o vérifie l'expression :

$$p(x|y^o) \propto \exp \left\{ -\frac{1}{2} (x - x^b)^T \mathbf{B}^{-1} (x - x^b) - \frac{1}{2} (y^o - \mathcal{H}(x))^T \mathbf{R}^{-1} (y^o - \mathcal{H}(x)) \right\}$$

Estimer le mode de la variable x sachant y^o est équivalent à maximiser la densité conditionnelle $p(x|y^o)$, cela correspond donc à l'estimation du maximum *a posteriori* [Maybeck, 1979]. Or maximiser $p(x|y^o)$ est équivalent à un problème de minimisation, en effet :

$$\max_{x \in \mathbb{R}^n} p(x|y^o) \iff \min_{x \in \mathbb{R}^n} \{-\ln(p(x|y^o))\} = \min_{x \in \mathbb{R}^n} \mathcal{J}(x)$$

où $\mathcal{J}(x)$ est la fonctionnelle donnée par l'expression :

$$\mathcal{J}(x) = \frac{1}{2} (x - x^b)^T \mathbf{B}^{-1} (x - x^b) + \frac{1}{2} (y^o - \mathcal{H}(x))^T \mathbf{R}^{-1} (y^o - \mathcal{H}(x)) \quad (3.29)$$

Cette fonctionnelle est celle que l'on va minimiser dans le cadre du 3D-Var. Avec des arguments assez proches nous pouvons établir la formulation de la fonction coût à minimiser dans le cadre du 4D-Var et que l'on expose en section 3.6.2.

3.6.2 Le 3D-Var

Dans le 3D-Var [Sasaki et al., 1955, Sasaki, 1958] nous souhaitons estimer l'état le plus probable d'une variable aléatoire x étant donnée une observation y^o . Conformément à ce que nous venons d'exposer en section 3.6.1 nous sommes amenés à minimiser la fonctionnelle donnée par l'expression (3.29) que l'on décompose en la somme de deux fonctionnelles $\mathcal{J}_b$ et $\mathcal{J}_o$ où :

$$\mathcal{J}_b(x) = \frac{1}{2} (x - x^b)^T \mathbf{B}^{-1} (x - x^b) \quad (3.30)$$

$$\mathcal{J}_o(x) = \frac{1}{2} (y^o - \mathcal{H}(x))^T \mathbf{R}^{-1} (y^o - \mathcal{H}(x)) \quad (3.31)$$

Les fonctionnelles $\mathcal{J}_b$ et $\mathcal{J}_o$ mesurent la différence quadratique respectivement entre le terme d'ébauche x^b et le vecteur de contrôle d'une part et l'observation y^o et le vecteur de contrôle projeté dans l'espace des observations $\mathcal{H}(x)$ d'autre part. Ces différences sont pondérées par la précision que l'on accorde au modèle et aux observations c'est-à-dire l'inverse de la matrice des covariances d'erreur d'ébauche $\mathbf{B}$ et l'inverse de la matrice des covariances d'erreur d'observation $\mathbf{R}$. L'inverse des matrices $\mathbf{B}$ et $\mathbf{R}$ définissent chacune une métrique dans l'espace du modèle et dans l'espace des observations respectivement. Alors minimiser la fonctionnelle $\mathcal{J}$ revient à minimiser l'expression :

$$\|x - x^b\|_{\mathbf{B}^{-1}}^2 + \|y^o - \mathcal{H}(x)\|_{\mathbf{R}^{-1}}^2$$

Ce qui revient à chercher l'état qui minimise la somme de la distance à l'ébauche dans l'espace du modèle (muni de la métrique définie par $\mathbf{B}^{-1}$) et la distance de la projection dans l'espace des observations (muni de la métrique définie par $\mathbf{R}^{-1}$) à l'observation elle-même.

Concrètement l'état analysé x^a qui minimise $\mathcal{J}$ est alors obtenu en résolvant l'équation $\nabla \mathcal{J}(x) = 0$ où $\nabla \mathcal{J}(x)$ (le gradient de $\mathcal{J}$ pris en x) est donné par l'expression :

$$\nabla \mathcal{J}(x) = \mathbf{B}^{-1} (x - x^b) - \mathbf{H}^T \mathbf{R}^{-1} (y^o - \mathcal{H}(x)) \quad (3.32)$$

Dans le cas où l'opérateur $\mathcal{H}$ est non-linéaire l'expression (3.32) ne permet pas de donner une expression analytique de x^a et la solution de l'équation $\nabla \mathcal{J}(x) = 0$ est approchée en utilisant une méthode de gradient.

Dans le cas où l'opérateur $\mathcal{H}$ est linéaire (on écrira par la suite $\mathbf{H}$) on peut montrer que le résultat du 3D-Var est identique au résultat du BLUE. En effet dans ce cas l'équation $\nabla \mathcal{J}(x) = 0$ est équivalente à l'équation :

$$\mathbf{B}^{-1} (x - x^b) - \mathbf{H}^T \mathbf{R}^{-1} (y^o - \mathbf{H}x^b - \mathbf{H} (x - x^b)) = 0 \quad (3.33)$$

x^a est donné par l'expression :

$$x^a = x^b + (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} (y^o - \mathbf{H}x^b) \quad (3.34)$$

Or on peut montrer que :

$$\begin{aligned} (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} &= (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R}) (\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R})^{-1} \\ &= (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{H}^T) (\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R})^{-1} \\ &= (\mathbf{B}^{-1} + \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{B}^{-1}) \mathbf{B} \mathbf{H}^T (\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R})^{-1} \\ &= \mathbf{B} \mathbf{H}^T (\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R})^{-1} \end{aligned}$$

Ainsi :

$$x^a = x^b + \mathbf{B} \mathbf{H}^T (\mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R})^{-1} (y^o - \mathbf{H}x^b) \quad (3.35)$$

Et l'on retrouve bien l'égalité (3.22).

Présenté de cette manière le 3D-Var semble être une méthode destinée aux systèmes stationnaires puisque la dimension temporelle n'apparaît pas dans sa formulation. Néanmoins celui-ci a été appliqué à des systèmes dynamiques. Dans ce cas toutes les observations disponibles sur une fenêtre d'assimilation sont comparées à l'état du modèle à un temps donné t_0 (voir la représentation schématique du 3D-Var en figure 3.1. Cela n'a pas d'impact sur la fonctionnelle $\mathcal{J}_b$. Si sur la fenêtre d'assimilation en question on assimile N observations y_k^o , $k = 1, \dots, N$ alors la fonctionnelle $\mathcal{J}_o$ s'écrit :

$$\mathcal{J}_o(x) = \frac{1}{2} \sum_{k=1}^N (y_k^o - \mathcal{H}(x(t_0)))^T \mathbf{R}^{-1} (y_k^o - \mathcal{H}(x(t_0))) \quad (3.36)$$

Bien qu'inexacte par nature cette méthode présente l'avantage par rapport à d'autres méthodes prenant en compte la dimension temporelle de ne pas nécessiter le calcul de l'adjoint du modèle ni même d'intégration du modèle.

Un moyen à la fois plus rigoureux mais plus coûteux de prendre en compte la dimension temporelle du système est de considérer l'écart entre les observations et la projection dans l'espace des observations du modèle aux temps correspondant. Cela est l'objet du 4D-Var que nous présentons dans la section suivante.

3.6.3 Le 4D-Var

Le 4D-Var [Le Dimet and Talagrand, 1986] est une extension du 3D-Var prenant en compte la dimension temporelle du système que l'on considère (voir figure 3.2). Pour cela on prend en compte l'écart entre les observations et la projection dans l'espace des observations de l'état du modèle aux temps correspondants sur la fenêtre d'assimilation.

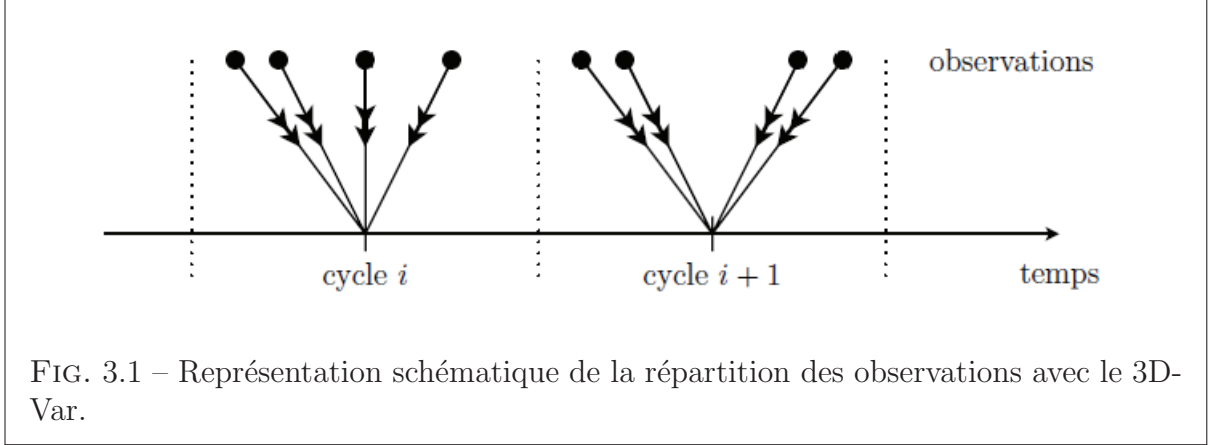


FIG. 3.1 – Représentation schématique de la répartition des observations avec le 3D-Var.

La fonctionnelle $\mathcal{J}$ que l'on cherche à minimiser s'écrit cette fois-ci :

$$\mathcal{J}(x_0) = \frac{1}{2} (x_0 - x^b)^T \mathbf{B}^{-1} (x_0 - x^b) + \frac{1}{2} \sum_{k=0}^N (y_k^o - \mathcal{H}_k(x_k))^T \mathbf{R}_k^{-1} (y_k^o - \mathcal{H}_k(x_k))$$

$$\mathcal{J}(x_0) = \mathcal{J}_b(x_0) + \mathcal{J}_o(x_0)$$

où pour tout $k = 1, \dots, N$, $x_k = \mathcal{M}_{k-1,k}(x_{k-1})$. Si on note :

$$\mathcal{M}_{0 \rightarrow k} = \mathcal{M}_{k-1,k} \circ \mathcal{M}_{k-2,k-1} \circ \dots \circ \mathcal{M}_{0,1}(x_0)$$

et $\mathcal{G}_k = \mathcal{H}_k \circ \mathcal{M}_{0 \rightarrow k}$ on peut réécrire $\mathcal{J}(x_0)$:

$$\begin{aligned} \mathcal{J}(x_0) &= \frac{1}{2} (x_0 - x_b)^T \mathbf{B}^{-1} (x_0 - x_b) \\ &+ \frac{1}{2} \sum_{k=0}^N (y_k^o - \mathcal{G}_k(x_0))^T \mathbf{R}_k^{-1} (y_k^o - \mathcal{G}_k(x_0)) \end{aligned}$$

L'état analysé optimal x^a qui minimise la fonctionnelle $\mathcal{J}$ est solution de l'équation $\nabla \mathcal{J}(x_0) = 0$ où :

$$\begin{aligned} \nabla \mathcal{J}(x) &= \mathbf{B}^{-1} (x - x_b) + \sum_{k=0}^N \mathbf{M}_{0 \rightarrow k}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} (y_k^o - \mathcal{G}(x)) \\ &= \mathbf{B}^{-1} (x - x_b) + \sum_{k=0}^N \mathbf{M}_{0,1}^T \mathbf{M}_{1,2}^T \dots \mathbf{M}_{k-1,k}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} (y_k^o - \mathcal{G}_k(x)) \end{aligned}$$

Avec $\nabla \mathcal{J}$ écrit sous cette forme il peut paraître difficile de résoudre $\nabla \mathcal{J} = 0$. Néanmoins si on écrit :

$$d_k = y_k^o - \mathcal{G}_k(x) \quad \forall k = 0, \dots, N$$

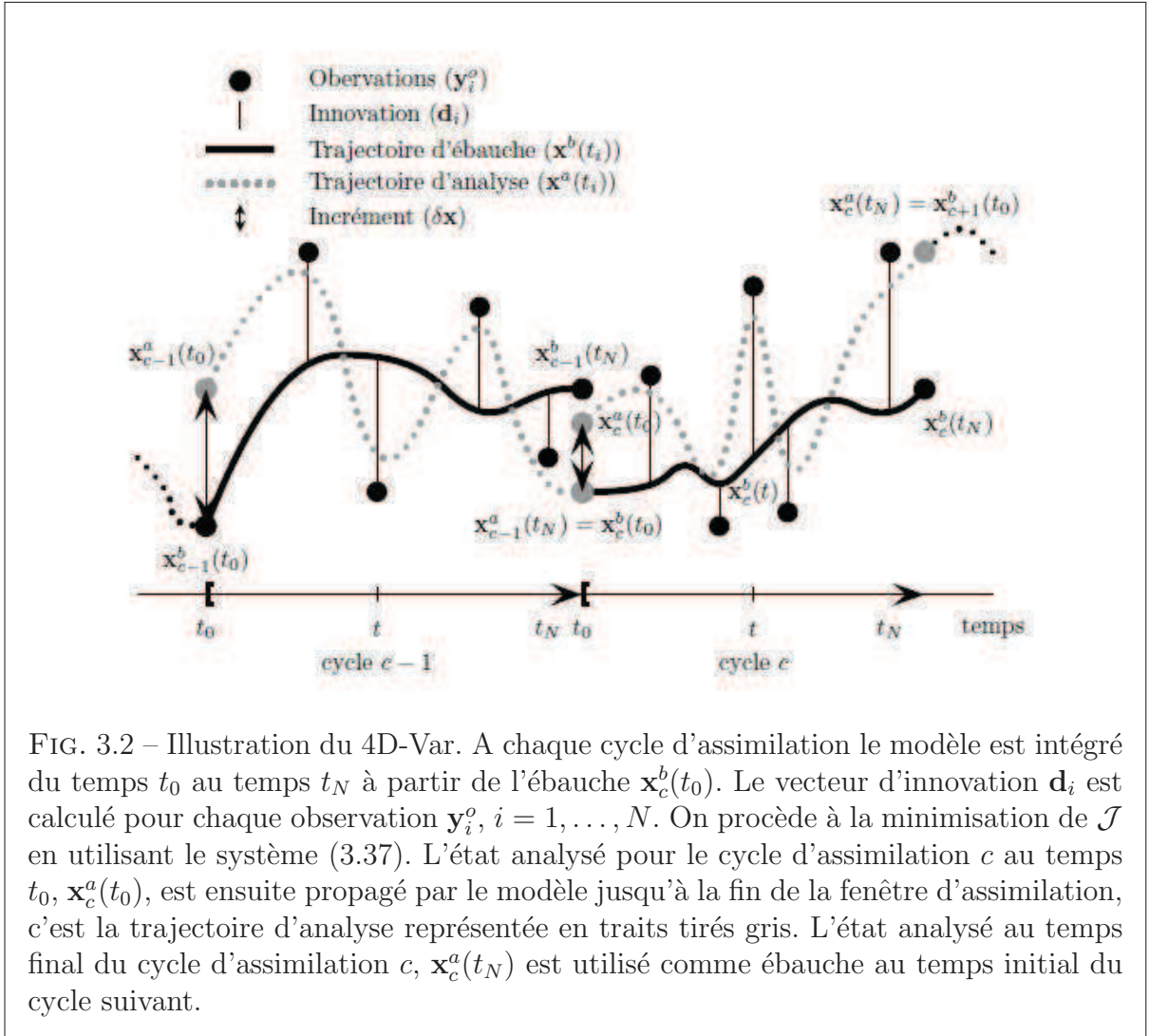
alors on peut réécrire le terme de la fonctionnelle $\nabla \mathcal{J}$ relatif aux observations sous cette forme :

$$\begin{aligned}
\nabla \mathcal{J}(x) &= \sum_{k=0}^N \mathbf{M}_{0,1}^T \mathbf{M}_{1,2}^T \dots \mathbf{M}_{k-1,k}^T \mathbf{H}_k^T \mathbf{R}_k^{-1} d_k \\
&= \mathbf{H}_0^T \mathbf{R}_0^{-1} d_0 + \mathbf{M}_{0,1}^T \mathbf{H}_1^T \mathbf{R}_1^{-1} d_1 + \mathbf{M}_{0,1}^T \mathbf{M}_{1,2}^T \mathbf{H}_2^T \mathbf{R}_2^{-1} d_2 + \dots + \\
&\quad \mathbf{M}_{0,1}^T \dots \mathbf{M}_{N-1,N}^T \mathbf{H}_N^T \mathbf{R}_N^{-1} d_N \\
&= \mathbf{H}_0^T \mathbf{R}_0^{-1} d_0 \\
&\quad + \mathbf{M}_{0,1}^T (\mathbf{H}_1^T \mathbf{R}_1^{-1} d_1 + \mathbf{M}_{1,2}^T (\mathbf{H}_2^T \mathbf{R}_2^{-1} d_2 + \dots + \mathbf{M}_{N-1,N}^T \mathbf{H}_N^T \mathbf{R}_N^{-1} d_N) \dots)
\end{aligned}$$

Cette écriture permet de calculer $\nabla \mathcal{J}(x)$ avec seulement une intégration du modèle direct sur la fenêtre d'assimilation et une du modèle adjoint. En effet $\nabla \mathcal{J}$ se calcule avec :

$$\begin{cases} x_N^* &= \mathbf{H}_N^T \mathbf{R}_N^{-1} d_N \\ x_{k-1}^* &= \mathbf{H}_k^T \mathbf{R}_k^{-1} d_k + \mathbf{M}_{k-1,k}^T x_k^* \end{cases} \quad k = N-1, \dots, 1 \quad (3.37)$$

Le calcul de l'état optimal x^a minimisant $\mathcal{J}$ sur une fenêtre d'assimilation se fait de manière itérative en utilisant une méthode d'optimisation basée sur l'usage du gradient.



D'autre part comme on vient de le voir le calcul de $\nabla \mathcal{J}$ nécessite l'adjoint du modèle. On peut calculer celui-ci en suivant les règles d'adjointisation d'un code comme cela est exposé dans [Giering and Kaminski, 1998]. Il existe également des différentiateurs automatiques de code tels TAPENADE développé par l'INRIA Sophia-Antipolis (Hascoet et Pascal, 2013).

3.6.4 Quelques variantes du 3D-Var et du 4D-Var

On expose dans cette section deux méthodes dérivées des méthodes 3D-Var et 4D-Var.

Le 4D-Var incrémental La première méthode que l'on présente est la méthode dite 4D-Var incrémental. Elle a été développée pour éviter d'avoir à minimiser une fonctionnelle $\mathcal{J}$ non convexe lorsque les opérateurs $\mathcal{M}$ et $\mathcal{H}$ ne sont pas linéaires. Cette méthode repose sur l'hypothèse que, à chaque pas de temps, le modèle et/ou l'opérateur d'observation sont faiblement non linéaires, c'est ce qu'on appelle l'hypothèse du linéaire tangent.

Autrement dit, on a :

$$\mathcal{M}_{0 \rightarrow k}(x) - \mathcal{M}_{0 \rightarrow k}(x^b) \approx \mathbf{M}_{0 \rightarrow k}(x - x^b) \quad (3.38)$$

$$\mathcal{H}_k(x_k) - \mathcal{H}_k(x_k^b) \approx \mathbf{H}_k(x_k - x_k^b) \quad (3.39)$$

où $x_k = \mathbf{M}_{0 \rightarrow k}(x)$ et $x_k^b = \mathbf{M}_{0 \rightarrow k}(x^b)$

En utilisant les approximations données en (3.38) et (3.39) on peut approcher la fonction coût du 4D-Var par une fonctionnelle quadratique (et donc facile à minimiser) $\tilde{\mathcal{J}}$ dont l'expression est donnée par :

$$\begin{aligned} \tilde{\mathcal{J}}(\delta x) &= \delta x^T \mathbf{B}^{-1} \delta x \\ &+ \sum_{k=1}^N (d_k - \mathbf{H}_k \mathbf{M}_{k-1,k} \dots \mathbf{M}_{0,1} \delta x)^T \mathbf{R}_k^{-1} (d_k - \mathbf{H}_k \mathbf{M}_{k-1,k} \dots \mathbf{M}_{0,1} \delta x) \end{aligned}$$

où $d_k = y_k^o - \mathcal{H}_k \mathcal{M}_{0 \rightarrow k}(x^b)$ est l'innovation et où δx est un incrément défini par $\delta x = x - x^b$.

L'incrément qui minimise $\tilde{\mathcal{J}}$, δx^a , est obtenu en résolvant l'équation $\nabla \tilde{\mathcal{J}}(\delta x) = 0$. On définit un nouvel état de référence x^r par :

$$x^r = x^b + \delta x^a$$

dans le voisinage duquel on estime le linéaire tangent des opérateurs $\mathcal{M}$ et $\mathcal{H}$ et ainsi de suite jusqu'à convergence de l'algorithme.

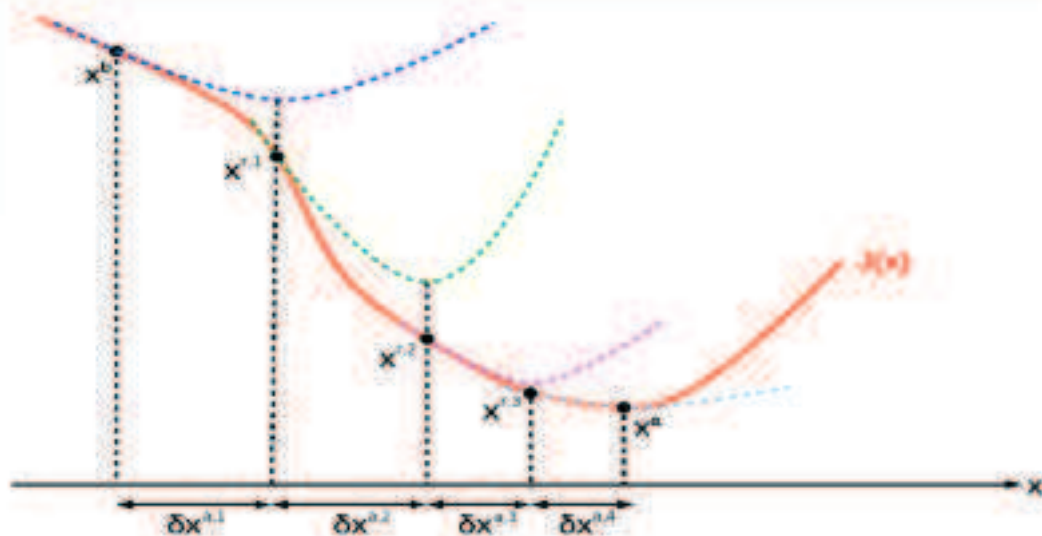
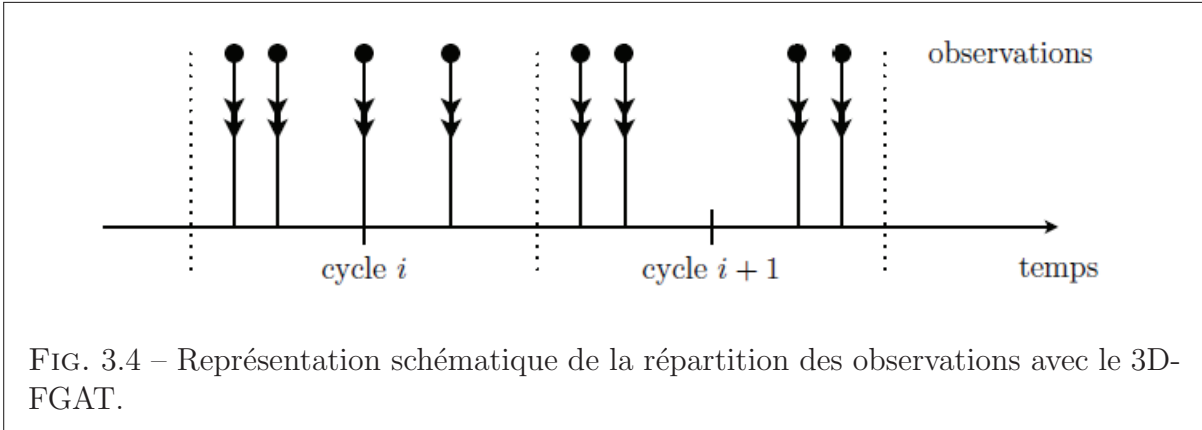


FIG. 3.3 – Représentation schématique du 4D-Var incrémental. Dans un premier temps on définit une fonctionnelle quadratique $\tilde{J}_1$ approchant $\mathcal{J}$ (courbe rouge) dans le voisinage de x^b (parabole bleue). L'incrément qui minimise $\tilde{J}_1$, $\delta x^{a,1}$, est obtenu en résolvant $\nabla \tilde{J}_1(\delta x) = 0$. On définit un nouvel état de référence $x^{r,1}$ par $x^{r,1} = x^b + \delta x^{a,1}$. On itère le processus jusqu'à ce que les $x^{r,k}$ convergent vers le minimum de la fonctionnelle $\mathcal{J}$, x^a . Source : [Isabelle Mirouze, 2010].

Le 3D-FGAT La seconde méthode est appelée 3D-FGAT pour First Guess At Appropriate Time. Elle se situe à mi-chemin du 3D-Var et du 4D-Var. Plus précisément il s'agit d'une approximation du 4D-Var incrémental dans laquelle le linéaire tangent du modèle est considéré comme stationnaire sur la fenêtre d'assimilation i-e :

$$\mathcal{M}_{0 \rightarrow k}(x) - \mathcal{M}_{0 \rightarrow k}(x^b) \approx x - x^b$$

Partant de là on peut appliquer l'algorithme du 4D-Var incrémental avec $\mathbf{M}_{0 \rightarrow k} = \mathbf{I}$ et $\mathbf{M}_{0 \rightarrow k}^T = \mathbf{I}$. Ainsi cette méthode ne nécessite pas de calculer le linéaire tangent du modèle ni son adjoint. De plus cette méthode donne de meilleurs résultats que le 3D-Var car dans ce cas les vecteurs d'innovation sont calculés au temps exact d'observation.



3.7 Le filtre de Kalman et ses dérivations

Contrairement aux méthodes variationnelles qui comme on l'a vu sont des méthodes liées à la théorie de l'optimisation le filtre de Kalman est une méthode d'interpolation statistique dans laquelle les observations sont assimilées de manière séquentielle. Nous présentons tout d'abord le filtre de Kalman proprement dit puis nous présenterons les nombreuses méthodes qui en sont dérivées.

3.7.1 Le filtre de Kalman

Le filtre de Kalman [Kalman, 1960, Cohn, 1997] est une extension sous certaines hypothèses de la méthode BLUE aux problèmes d'évolution temporelle. Ces hypothèses s'énoncent en quatre points :

- les erreurs de modélisation suivent une loi gaussienne de moyenne nulle et de matrice de covariance $\mathbf{Q}$. On suppose de plus qu'elles sont décorrélées en temps, i-e $E[q_k q_l^T] = 0, k \neq l$;
- les erreurs d'observation suivent une loi gaussienne de moyenne nulle et de matrice de covariance $\mathbf{R}$. On suppose de plus qu'elles sont décorrélées en temps i-e : $E[\varepsilon_p^o \varepsilon_q^{oT}] = 0, p \neq q$;
- les erreurs d'ébauche suivent une loi gaussienne de matrice de covariance $\mathbf{B}$;
- les erreurs sont mutuellement décorrélées :

$$E[q_p(\varepsilon_p^b)^T] = E[\varepsilon_p^o(\varepsilon_p^b)^T] = E[q_p(\varepsilon_p^o)^T] = 0.$$

Tout comme dans le BLUE l'opérateur d'observation est supposé linéaire et on suppose également que le modèle dynamique aussi est linéaire.

On peut alors décrire le filtre de Kalman. Celui-ci se décompose en deux étapes. Une première étape dite étape d'analyse dans laquelle on met à jour l'état du système. On suppose que l'on dispose d'une ébauche au temps t_k notée x_k^b pour laquelle on connaît sa matrice de covariance d'erreur notée $\mathbf{B}_k$. D'après les équations du BLUE la matrice de gain est donnée par :

$$\mathbf{K}_k = \mathbf{B}_k \mathbf{H}_k^T (\mathbf{H}_k \mathbf{B}_k \mathbf{H}_k^T + \mathbf{R}_k)^{-1}$$

Alors l'état analysé au temps t_k noté x_k^a est donné par l'expression :

$$x_k^a = x_k^b + \mathbf{B}_k \mathbf{H}_k^T (\mathbf{H}_k \mathbf{B}_k \mathbf{H}_k^T + \mathbf{R}_k)^{-1} (y_k^o - \mathbf{H}_k x_k^b) \quad (3.40)$$

où $\mathbf{H}_k$ l'opérateur d'observation au temps t_k . Suivant les équations du BLUE la matrice des covariances d'erreur d'analyse $\mathbf{A}_k$ est donnée par l'expression :

$$\mathbf{A}_k = (\mathbf{I}_k - \mathbf{K}_k \mathbf{H}_k) \mathbf{B}_k$$

L'étape de propagation consiste à propager avec le modèle l'état analysé du temps t_k au temps t_{k+1} pour servir d'ébauche au temps t_{k+1} , soit :

$$x_{k+1}^b = \mathbf{M}_{k,k+1} x_k^a \quad (3.41)$$

Et de manière corollaire durant l'étape de propagation on propage également la matrice des covariances d'erreur d'analyse qui servira de matrice des covariances d'erreur d'ébauche au temps t_{k+1} . En effet par définition :

$$\begin{aligned} \varepsilon_{k+1}^b &= \mathbf{M}_{k,k+1} x_k^a - \mathbf{M}_{k,k+1} x_k^t - q_k \\ \varepsilon_{k+1}^b &= \mathbf{M}_{k,k+1} (x_k^a - x_k^t) - q_k \\ \varepsilon_{k+1}^b &= \mathbf{M}_{k,k+1} \varepsilon_k^a - q_k \end{aligned}$$

D'où on déduit :

$$\begin{aligned} \mathbf{B}_{k+1} &= \mathbf{E} [\varepsilon_{k+1}^b (\varepsilon_{k+1}^b)^T] \\ \mathbf{B}_{k+1} &= \mathbf{M}_{k,k+1} \mathbf{E} [\varepsilon_k^a (\varepsilon_k^a)^T] \mathbf{M}_{k,k+1}^T - \mathbf{E} [q_k q_k^T] \\ \mathbf{B}_{k+1} &= \mathbf{M}_{k,k+1} \mathbf{A}_k \mathbf{M}_{k,k+1}^T + \mathbf{Q}_k \end{aligned}$$

A travers la matrice des covariances d'erreur d'analyse $\mathbf{A}_k$ le filtre de Kalman donne une estimation de la qualité de l'état analysé à chaque temps t_k sur une fenêtre d'assimilation tandis que le 4D-Var ne donne qu'une estimation de la qualité de l'estimation de l'état initial ou encore de l'ébauche.

3.7.2 Le filtre de Kalman Étendu ou Extended Kalman Filter (EKF)

Le filtre de Kalman étendu consiste en l'extension du filtre de Kalman aux modèles et/ou opérateurs d'observation faiblement non-linéaires. Dans ce cas seules les équations du filtre de Kalman (3.40) et (3.41) sont modifiées. L'équation (3.40) s'écrit alors :

$$x_k^a = x_k^b + \mathbf{B}_k \mathbf{H}_k^T (\mathbf{H}_k \mathbf{B}_k \mathbf{H}_k^T + \mathbf{R}_k)^{-1} (y_k^o - \mathcal{H}_k x_k^b)$$

où $\mathcal{H}_k$ et $\mathbf{H}_k$ désignent respectivement l'opérateur d'observation et son linéaire tangent au temps t_k .

L'équation (??) devient :

$$x_{k+1}^b = \mathcal{M}_{k,k+1} x_k^a$$

où $\mathcal{M}_{k,k+1}$ désigne l'évolution temporelle du modèle entre les temps t_k et t_{k+1} .

Il faut souligner dans ce cas que dans les autres équations du filtre de Kalman $\mathbf{M}_{k,k+1}$ désigne le linéaire tangent de $\mathcal{M}_{k,k+1}$.

Lorsqu'au moins l'un des opérateurs $\mathcal{H}$ ou $\mathcal{M}$ n'est pas linéaire l'état analysé n'est plus optimal au sens où l'EKF ne minimise pas la variance de l'état analysé.

Dans le cas où la taille du vecteur d'état est très grande comme cela peut être le cas en météorologie ou en océanographie les approches directes de type filtre de Kalman et filtre de Kalman étendu ne sont pas applicables du fait de la grande taille des matrices à manipuler. Pour contourner ce problème et approcher le filtre de Kalman des méthodes alternatives ont été développées. Nous présentons en particulier deux méthodes, le filtre de Kalman de rang réduit, autrement désigné par l'acronyme RRSQRT pour Reduced Rank Square Root ainsi que le filtre SEEK (pour Singular Evolutive Extended Kalman Filter).

3.7.3 Le filtre de Kalman de rang réduit - RRSQRT

Le filtre RRSQRT a été introduit par [Verlaan and Heemink, 1995]. Dans ce filtre on cherche à résoudre les équations du filtre de Kalman dans un espace dont la dimension est nettement moindre que la taille du vecteur d'état.

Pour cela on approche les matrices de covariance d'erreur modèle, d'ébauche et d'analyse par le produit de matrices racine carré (d'où la dénomination) contenant les n principaux modes de ces matrices.

Considérons l'ébauche du système x_0^b au temps t_0 et sa matrice de covariance d'erreur associé $\mathbf{B}_0$. On sélectionne les n premiers modes de la matrice $\mathbf{B}_0$:

$$\mathbf{B}_0 \approx \mathbf{S}_0^b (\mathbf{S}_0^b)^T$$

ainsi que les $n_q < n$ premiers modes des matrices des covariances d'erreur modèle :

$$\mathbf{Q}_0 \approx \tilde{\mathbf{Q}}_0^{1/2} \left(\tilde{\mathbf{Q}}_0^{1/2} \right)^T$$

Considérons maintenant l'étape d'analyse au temps t_k . La projection des n premiers modes de la matrice de covariance d'erreur d'ébauche dans l'espace des observations s'écrit :

$$\Psi = (\mathbf{H}_k \mathbf{S}_k^b)^T$$

La matrice de gain approchée s'écrit quant à elle :

$$\tilde{\mathbf{K}}_k = \mathbf{S}_k^b \Psi (\Psi^T \Psi + \mathbf{R}_k)^{-1}$$

Le calcul de l'état analysé se fait par une expression similaire à celle utilisée dans le Filtre de Kalman classique :

$$x_k^a = x_k^b + \tilde{\mathbf{K}}_k (y_k^o - \mathcal{H}_k(x_k^b))^{-1}$$

Le calcul approché des n premiers modes de la matrice des covariances d'erreur d'analyse est donné par l'expression :

$$\mathbf{S}_k^b = \mathbf{S}_k^a \left(\mathbf{I}_n - \Psi (\Psi^T \Psi + \mathbf{R}_k)^{-1} \Psi^T \right)^{1/2}$$

La racine carrée est facile à calculer étant de taille $n \times n$.

Dans le Filtre de Kalman de rang réduit la phase de prévision est précédée par une phase de réduction de la matrice réduite des covariances d'erreur d'analyse à ses $n - n_q$ premiers modes. Pour cela on procède tout d'abord à la diagonalisation de la matrice $(\mathbf{S}_k^a)^T \mathbf{S}_k^a$, celle-ci étant symétrique définie positive on peut l'écrire sous forme d'un produit : $\mathbf{V} \mathbf{\Lambda} \mathbf{V}^T = (\mathbf{S}_k^a)^T \mathbf{S}_k^a$. Alors si on considère les $n - n_q$ vecteurs propres contenus dans $\tilde{\mathbf{V}}$ (que l'on note $\tilde{\mathbf{V}}$) et associés aux $n - n_q$ plus grandes valeurs propres de $\mathbf{\Lambda}$, on écrit : $\tilde{\mathbf{S}}_k^a = \mathbf{S}_k^a \tilde{\mathbf{V}}$.

L'ébauche au temps t_{k+1} est calculée en propageant l'état analysé au temps t_k du temps t_k au temps t_{k+1} :

$$x_{k+1}^b = \mathbf{M}_{k+1} x_k^a$$

L'évaluation des n premiers modes de la matrice des covariances d'erreur d'ébauche est obtenue par concaténation des $n - n_q$ premiers modes de $\tilde{\mathbf{S}}_k^a$ propagés par le modèle linéaire tangent et des n_q premiers modes de $\mathbf{Q}_k$ contenus dans $\tilde{\mathbf{Q}}_k^{1/2}$:

$$\mathbf{S}_{k+1}^b = \left[\mathbf{M}_k \tilde{\mathbf{S}}_k^a \mathbf{M}_k^T, \tilde{\mathbf{Q}}_k^{1/2} \right]$$

3.7.4 Le filtre SEEK

Le filtre SEEK (pour Singular Evolutive Extended Kalman filter), introduit par Din Tuan Pham [Pham et al., 1998], repose sur l'hypothèse de la stagnation ou de la décroissance du rang des matrices de covariances d'erreur d'ébauche et d'analyse. Cette hypothèse est vraie dans le cas où le modèle est parfait mais elle ne l'est plus dans le cas où le modèle n'est pas parfait. Dans ce dernier cas on peut néanmoins projeter la matrice des covariances d'erreur modèle sur un espace « bien choisi ».

Il est à noter que le filtre SEEK est un filtre réduit déduit du Filtre de Kalman Etendu mais dans un souci de simplicité on se restreindra dans ce travail au cas du Filtre de Kalman linéaire.

Considérons l'ébauche du système x_0^b au temps t_0 et sa matrice de covariance d'erreur associé $\mathbf{B}_0$. $\mathbf{B}_0$ étant une matrice symétrique définie positive on peut la décomposer sous la forme d'un produit :

$$\mathbf{B}_0 \approx \mathbf{L}_0 \tilde{\mathbf{B}}_0 \mathbf{L}_0^T$$

où $\tilde{\mathbf{B}}_0$ est diagonale et contient les r plus grandes valeurs propres de $\mathbf{B}_0$, il s'agit de la matrice des covariance d'erreur d'ébauche réduite. $\mathbf{L}_0$ contient les vecteurs propres associés ou directions principales associées des erreurs.

Par la matrice réduite $\tilde{\mathbf{B}}_k$ des covariances d'erreur d'ébauche au temps t_k n'est pas nécessairement diagonale car elle évolue au cours du temps.

On considère alors l'étape d'analyse au temps t_k . Si $\tilde{\mathbf{B}}_k$ est inversible ce qui est le cas sauf si le rang de $\mathbf{B}_k$ a été réduit (mais alors dans ce cas on peut la décomposer à nouveau) la matrice réduite des covariances d'erreur d'analyse au temps t_k , $\tilde{\mathbf{A}}_k$ se calcule avec l'expression :

$$\left(\tilde{\mathbf{A}}_k\right)^{-1} = \left(\tilde{\mathbf{B}}_k\right)^{-1} + \mathbf{L}_k^T \mathbf{H}_k^T \mathbf{R}_k^{-1} \mathbf{H}_k \mathbf{L}_k$$

L'expression de la matrice de gain approchée est donnée par :

$$\tilde{\mathbf{K}}_k = \mathbf{L}_k \tilde{\mathbf{A}}_k \mathbf{L}_k^T \mathbf{H}_k^T \mathbf{R}_k^{-1}$$

Le calcul de l'état analysé se fait avec une expression très proche de l'expression classique :

$$x_k^a = x_k^b + \tilde{\mathbf{K}}_k (y_k^o - \mathcal{H}(x_k^b))$$

L'état analysé au temps t_k est propagé par le modèle pour servir d'ébauche au temps t_{k+1} :

$$x_{k+1}^b = \mathbf{M}_{k+1} x_k^a$$

Les vecteurs engendrant l'espace des directions principales sont également propagés par le modèle :

$$\mathbf{L}_{k+1} = \mathbf{M}_{k+1} \mathbf{L}_k$$

Dans le cas où le modèle est parfait la matrice réduite des covariances d'erreur d'analyse au temps t_k sert comme matrice réduite des covariances d'erreur d'ébauche au temps t_{k+1} , soit :

$$\tilde{\mathbf{B}}_{k+1} = \tilde{\mathbf{A}}_k$$

Dans le cas avec erreur modèle, du fait de la relation $\mathbf{B}_{k+1} = \mathbf{M}_{k+1} \mathbf{A}_k \mathbf{M}_{k+1}^T + \mathbf{Q}_k$, le rang de $\mathbf{B}_{k+1}$ peut être supérieur à celui de $\mathbf{A}_k$. Dans ce cas pour réduire le rang on projette $\mathbf{Q}_k$ sur l'espace engendré par les vecteurs colonnes de $\mathbf{L}_k$ où la projection s'écrit :

$$\Pi_k = \mathbf{L}_k (\mathbf{L}_k^T \mathbf{L}_k)^{-1} \mathbf{L}_k^T$$

Dans le cas avec erreur modèle la matrice réduite des covariances d'erreur d'ébauche s'écrit :

$$\tilde{\mathbf{B}}_{k+1} = \tilde{\mathbf{A}}_k + (\mathbf{L}_{k+1}^T \mathbf{L}_{k+1})^{-1} \mathbf{L}_{k+1}^T \mathbf{Q}_k \mathbf{L}_{k+1} (\mathbf{L}_{k+1}^T \mathbf{L}_{k+1})^{-1}$$

3.7.5 Le filtre de Kalman d'ensemble - EnKF

Le filtre de Kalman d'ensemble ou EnKF pour Ensemble Kalman Filter a été introduit en 1994 par Evensen, [Evensen, 1994]. Il est basé sur l'estimation statistique de l'état du modèle et des matrices de covariances d'erreur. Celui-ci ne fait pas d'hypothèse sur la linéarité du modèle contrairement au filtre de Kalman. Néanmoins pour donner une

estimation satisfaisante des matrices de covariance d'erreur l'EnKF peut nécessiter la réalisation d'un grand nombre de réalisations perturbées du modèle. C'est en effet un inconvénient reconnu de l'EnKF qu'il converge lentement vers le filtre de Kalman. De plus il ne converge vers la distribution du filtrage bayésien que dans le cas où le modèle est linéaire et la loi de probabilité de l'état considéré est gaussienne. Cependant sa relative facilité de mise en place en fait un outil très répandu dans le domaine de l'assimilation de données.

On considère N états perturbés (ou membres) du modèle. Ces états perturbés peuvent être générés en perturbant par exemple les forçages du modèle. Et on note ces états perturbés au temps t_k :

$$x_k^1, x_k^2, \dots, x_k^N$$

Dans l'EnKF l'état du modèle au temps t_k est estimé comme étant la moyenne des membres, ainsi :

$$\bar{x}_k = \frac{1}{N} \sum_{i=1}^N x_k^i$$

On rappelle que la matrice des covariances d'erreur d'ébauche est définie par l'égalité :

$$\mathbf{B}_k = \mathbf{E} \left[\varepsilon_k^b (\varepsilon_k^b)^T \right]$$

L'erreur d'ébauche pour un membre pouvant être approchée par :

$$\tilde{\varepsilon}^{b,i} = x_k^{b,i} - \bar{x}_k^b$$

on estime naturellement la matrice des covariances d'erreur d'ébauche par la formule suivante :

$$\mathbf{B}_{e,k} = \frac{1}{N-1} \sum_{i=1}^N \left(x_k^{b,i} - \bar{x}_k^b \right) \left(x_k^{b,i} - \bar{x}_k^b \right)^T \quad (3.42)$$

Tout comme le filtre de Kalman l'EnKF se décompose en deux phases : une phase d'analyse et une phase de propagation.

Dans la phase d'analyse de l'EnKF nous appliquons simplement les équations du filtre de Kalman à chaque membre de l'ensemble en remplaçant la matrice des covariances d'erreur d'ébauche $\mathbf{B}_k$ par son approximation donnée en (3.42). Tout d'abord la matrice de gain au temps t_k est donnée par :

$$\mathbf{K}_k = \mathbf{B}_{e,k} \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{B}_{e,k} \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1}$$

Ensuite on applique l'équation d'analyse du filtre de Kalman à chaque membre de l'ensemble :

$$x_k^{a,i} = x_k^{b,i} + \mathbf{K}_k \left(y_k^{o,i} - \mathbf{H}_k x_k^{b,i} \right) \quad (3.43)$$

où $y_k^{o,i} = y_k^o + \varepsilon_k^i$ et où les ε_k^i suivent une loi normale de moyenne nulle et de matrice de covariance $\mathbf{R}_k$. Dans la version initiale du filtre de Kalman d'Ensemble proposée par

Evensen l'observation assimilée par chaque membre était la même et c'était celle observée au temps t_k , à savoir y_k^o . Cela présentait l'inconvénient de sous-estimer la variance de l'ensemble des membres analysés, voire même dans certains cas de faire diverger le filtre. C'est à [Burgers et al., 1998] que l'on doit d'introduire l'assimilation d'observations perturbées (de moyenne y_k^o et de matrice de covariance $\mathbf{R}$). Cela permet de conserver la variance de l'ensemble et d'éviter que le filtre ne diverge. On peut montrer dans ce cas que la moyenne des états analysés et l'estimation de la matrice des covariances d'erreur d'analyse tendent respectivement, quand N croît vers l'estimation de l'état analysé et la matrice des covariances d'erreur d'analyse donnés par le BLUE dans le cas non linéaire. Ces méthodes basées sur la perturbation des observations sont appelées stochastiques. Nous évoquerons plus loin en quoi consistent les versions dites déterministes.

La phase de prévision consiste quant à elle en la propagation de chacun des membres analysés par le modèle pour servir d'ébauche au cycle d'assimilation suivant, i-e si le modèle est parfait :

$$x_{k+1}^{b,i} = \mathcal{M}_{k,k+1}(x_k^{a,i}) \quad i = 1, \dots, N$$

Dans le cas où le modèle n'est pas parfait, si la loi de probabilité de l'erreur modèle est connue on peut générer de manière aléatoire un ensemble de vecteurs d'erreur modèle $\{\eta_k^1, \dots, \eta_k^N\}$ (voir [Bonan, 2013]). L'ébauche au temps t_{k+1} pour chaque membre est donnée par la relation :

$$x_{k+1}^{b,i} = \mathcal{M}_{k,k+1}(x_k^{a,i}) + \eta_k^i \quad i = 1, \dots, N \quad (3.44)$$

De manière assez similaire si nous ne connaissons que la matrice des covariances d'erreur modèle $\mathbf{Q}_k$ nous pouvons supposer que l'erreur modèle suit une loi normale de moyenne nulle et de matrice de covariance $\mathbf{Q}_k$. Comme précédemment nous pouvons générer de manière aléatoire un ensemble de vecteurs d'erreur modèle tel que l'ébauche au temps t_{k+1} pour chaque membre est donnée par la relation (3.44).

3.7.6 Les versions déterministes du filtre de Kalman d'ensemble

Les filtres de Kalman d'ensemble déterministes sont une dérivation du filtre de Kalman d'ensemble qui ne reposent pas sur la perturbation des observations.

Dans la suite, par un souci de simplicité d'écriture nous noterons :

$\mathbf{X}_k = [x_k^1 - \bar{x}_k, \dots, x_k^N - \bar{x}_k]$ et $\bar{\mathbf{X}}^b$ la moyenne de cet ensemble.

Dans les versions déterministes du filtre de Kalman d'ensemble on approche la matrice des covariances d'erreur d'ébauche par un produit de matrices racines carrées comme dans la relation (3.42). La phase d'analyse consiste en la mise à jour de l'ensemble des anomalies $\mathbf{X}^a$, du calcul de la moyenne des anomalies et ajout de celle-ci à chaque élément de $\mathbf{X}^a$.

Nous exposons ici quelques uns des filtres déterministes parmi les plus connus.

Ensemble Transform Kalman Filter - ETKF

L'ETKF a été introduit par [Bishop et al., 2001]. Ce filtre est basé sur le principe que si la matrice des covariances d'erreur d'ébauche $\mathbf{B}_e$ s'écrit $\mathbf{B}_e = \frac{1}{N-1} \mathbf{X}^b \mathbf{X}^{bT}$ alors le

produit $\mathbf{X}^a \mathbf{X}^{aT}$ peut s'écrire sous la forme :

$$\mathbf{X}^a \mathbf{X}^{aT} = \mathbf{X}^b \left(\mathbf{I}_N + \frac{1}{N-1} \mathbf{X}^{bT} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{X}^b \right)^{-1} \mathbf{X}^{bT} \quad (3.45)$$

$$\mathbf{X}^a \mathbf{X}^{aT} = \mathbf{X}^b \mathbf{T} \mathbf{T}^T \mathbf{X}^{bT} \quad (3.46)$$

Par identification cette relation est satisfaite si $\mathbf{X}^a = \mathbf{X}^b \mathbf{T}$.

Il est à noter que la matrice $\mathbf{T}$ dans l'équation (3.46) n'est pas unique, on peut remplacer en effet $\mathbf{T}$ par $\mathbf{T}\mathbf{U}$ où $\mathbf{U}$ est une matrice orthogonale quelconque.

Dans la suite on note : $\mathbf{Y} = \mathbf{H}\mathbf{X}^b$. (Bishop *et al.*, 2001) propose de décomposer en éléments propres $\mathbf{I}_N + \frac{1}{N-1} \mathbf{Y}^T \mathbf{R}^{-1} \mathbf{Y} = \mathbf{C}\mathbf{T}\mathbf{C}^T$ et de prendre $\mathbf{T} = \mathbf{C}\mathbf{T}^{-1/2}$. Ce choix présente certains inconvénients [Wang and Bishop, 2003, Lawson and Hansen, 2004] comme celui de ne pas préserver la moyenne des états analysés et de mal répartir les membres autour de la moyenne. Pour remédier à ces problèmes [Wang et al., 2004, Ott et al., 2002] proposent de choisir $\mathbf{T} = \mathbf{C}\mathbf{T}^{-1/2} \mathbf{C}^T \mathbf{U}$ où $\mathbf{U}$ est une matrice orthogonale préservant la moyenne. L'expression de la matrice d'anomalies des états analysés est donnée par :

$$\mathbf{X}^a = \sqrt{N-1} \mathbf{X}^b \mathbf{C}\mathbf{T}^{-1/2} \mathbf{C}^T \mathbf{U}$$

Le calcul de la moyenne de l'ensemble des états analysés se fait avec la formule :

$$\bar{x}^a = \bar{x}^b + \mathbf{X}^b \mathbf{C}\mathbf{T}^{-1} \mathbf{C}^T (\mathbf{Y}^b)^T \mathbf{R}^{-1} (y^o - \bar{y}^b)$$

où $\bar{y}^b$ est la moyenne de l'ensemble des ébauches projetées dans l'espace des observations i-e : $\{y^{b,i} = \mathcal{H}(x^{b,i}); i = 1, \dots, N\}$

Le calcul de l'état analysé pour chacun des membres se fait en additionnant moyenne et anomalie, soit :

$$x^{a,i} = \bar{x}^a + \mathbf{X}^{a,i}$$

Il faut souligner que l'ETKF peut se réécrire sous forme variationnelle [Hunt et al., 2007, Harlim and Hunt, 2007].

Singular Evolutive Interpolated Kalman

Il s'agit du premier filtre de Kalman d'ensemble déterministe inventé [Pham, 1996] et est en quelque sorte dérivé du filtre SEEK. On évalue les directions principales du filtre SEEK et la matrice des covariances d'erreur d'ébauche avec les formules suivantes :

$$\mathbf{L} = [x^{b,1} \dots x^{b,N}] \mathbf{W} \quad (3.47)$$

$$\mathbf{U}^b = (r+1) \mathbf{W}^T \mathbf{W} \quad (3.48)$$

où :

$$\mathbf{W} = \begin{bmatrix} & \mathbf{I}_r & \\ 0 & \dots & 0 \end{bmatrix} - \frac{1}{r+1} \begin{pmatrix} 1 & \dots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \dots & 1 \end{pmatrix}$$

Connaissant $\mathbf{L}$ et $\mathbf{U}^b$ on peut calculer $\mathbf{U}^a$ et par suite la matrice d'anomalies $\mathbf{X}^a = \sqrt{r+1} \mathbf{L} \mathbf{U}^{a1/2}$.

Ensemble Square Root Filter - EnSRF

L'EnSRF a été introduit par [Whitaker and Hamill, 2002] et repose sur l'idée de la mise à jour des anomalies sous la forme suivante : $\mathbf{X}^a = (\mathbf{I} - \tilde{\mathbf{K}}\mathbf{H})\mathbf{X}^b$. On peut montrer que cela implique que la matrice de gain $\tilde{\mathbf{K}}$ vérifie :

$$\mathbf{A}_e = (\mathbf{I} - \tilde{\mathbf{K}}\mathbf{H})\mathbf{B}_e = (\mathbf{I} - \tilde{\mathbf{K}}\mathbf{H})\mathbf{B}_e(\mathbf{I} - \tilde{\mathbf{K}}\mathbf{H})^T$$

Dans ce cas on montre que [Andrews, 1968] :

$$\tilde{\mathbf{K}} = \mathbf{B}_e \mathbf{H}^T ((\mathbf{H}\mathbf{B}_e \mathbf{H}^T)^{-1/2})^T ((\mathbf{H}\mathbf{B}_e \mathbf{H}^T)^{1/2} + \mathbf{R}^{1/2})^{-1}$$

Ensemble Adjustment Kalman Filter - EAKF

L'EAKF a été introduit par [Anderson, 2001] et utilise deux décompositions en valeurs singulières :

$$\mathbf{B}_e = \mathbf{L}\mathbf{\Sigma}\mathbf{L}^T \quad \text{et} \quad \mathbf{I}_N + \frac{1}{N-1} \mathbf{X}^{bT} \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \mathbf{X}^b = \mathbf{C}\mathbf{T}\mathbf{C}^T$$

Les matrices d'anomalie sont liées par la relation $\mathbf{X}^a = \tilde{\mathbf{T}}\mathbf{X}^b$ où $\tilde{\mathbf{T}} = \mathbf{X}^b \mathbf{C}\mathbf{T}^{-1/2} \mathbf{\Sigma}^{-1/2} \mathbf{L}$

3.7.7 Les problèmes d'échantillonnage et les solutions liées

L'approximation des matrices de covariance d'erreur d'ébauche et d'analyse par un nombre fini de membres conduit à certaines erreurs d'échantillonnage. On sait par exemple d'après le théorème central limite que la convergence de la matrice $\mathbf{B}_e$ vers la matrice des covariances d'erreur d'ébauche $\mathbf{B}$ calculée se fait en $1/\sqrt{N}$. Ce qui implique par exemple que même avec 10000 membres l'erreur d'estimation de la matrice $\mathbf{B}_e$ est en $O(10^{-2})$.

[Bocquet, 2011] à la suite de [Mallat et al., 1998] ou encore [Raynaud et al., 2009] s'est intéressé à l'erreur commise lorsque l'on souhaite estimer une matrice de covariance $\mathbf{B}$ avec un nombre N de tirages aléatoires d'une loi gaussienne ayant $\mathbf{B}$ pour matrice de covariance. On note $\mathbf{B}_e$ la matrice ainsi calculée. Celle-ci diffère de la matrice $\mathbf{B}$ et on peut montrer avec le théorème de Wick [Wick, 1950] que l'on a :

$$\mathbb{E} [(\mathbf{B}_{e,ij} - \mathbf{B}_{ij})^2] = \frac{1}{N} (\mathbf{B}_{ii}\mathbf{B}_{jj} + \mathbf{B}_{ij}^2) \quad (3.49)$$

Dans le cas particulier des variances il découle de l'égalité (3.49) que :

$$\mathbb{E} [(\mathbf{B}_{e,ii} - \mathbf{B}_{ii})^2] = \frac{2}{N} \mathbf{B}_{ii}^2 \quad (3.50)$$

Dans le cas particulier des covariances longues distances $\mathbf{B}_{ij}$ est souvent négligeable dans les systèmes géophysiques et il découle de l'égalité (3.49) :

$$\mathbb{E} [(\mathbf{B}_{e,ij} - \mathbf{B}_{ij})^2] = \frac{1}{N} \mathbf{B}_{ii} \mathbf{B}_{jj} \quad (3.51)$$

Il vient de l'égalité (3.50) que lorsque le nombre de membres utilisé pour estimer la matrice $\mathbf{B}_e$ est faible les variances peuvent être sous-estimées. En effet un petit nombre de membres ne sera pas suffisant pour estimer correctement les variances de la matrice $\mathbf{B}$. D'autre part l'égalité (3.51) suppose l'existence de covariances/corrélations longues distances lorsque le nombre de membres est faible. Ces phénomènes sont bien connus de la communauté des utilisateurs de filtre de Kalman d'ensemble, voir [van Leeuwen, 1999] ou encore [Sacher and Bartello, 2008]. Pour remédier à ces problèmes de sous-échantillonnage des méthodes on été développées ; nous présentons plus particulièrement ci-après l'inflation et la localisation.

L'inflation

L'inflation est une méthode qui a été introduite par [Anderson and Anderson, 1999]. Elle a pour objet de compenser la sous-estimation des variances de l'ensemble analysé par l'augmentation (ou inflation) des anomalies des membres de l'ensemble en utilisant un coefficient $\lambda > 1$ tel que :

$$\tilde{x}^k = \lambda(x^k - \bar{x}) + \bar{x} \quad k = 1, \dots, N \quad (3.52)$$

λ étant un coefficient strictement positif il apparait alors clairement que la variance des membres de l'ensemble est accrue tandis que la moyenne ainsi que les corrélations restent inchangées. En augmentant la variance des membres on réduit le rapport des variances d'erreur d'observation et de l'ébauche aux points d'observation et donnons par suite plus de poids à celles-ci dans la procédure d'assimilation. Mécaniquement les résultats de l'assimilation en termes de différence entre l'état analysé et les observations sont meilleurs.

Le coefficient λ peut être réglé « à la main » de manière empirique en fonction des valeurs qui donnent les résultats les plus satisfaisants. Néanmoins des méthodes ont été développées pour estimer mathématiquement ce coefficient. D'une part on a les méthodes dérivées des travaux de [Dee, 1995] basées sur la maximisation de la densité de probabilité conditionnelle $p(\lambda|y^o)$. On pourra se reporter aux travaux de [Mitchell and Houtekamer, 2000] ou encore [Anderson, 2007]. D'un autre côté on a les méthodes dérivées des critères de consistance de Desroziers, [Desroziers et al., 2005], et qui sont illustrées par le travail de [Li et al., 2009]. Bien que différentes dans le principe ces méthodes donnent des résultats comparables sur un cas scalaire (voir [Miyoshi, 2011]).

La méthode d'inflation présentée en (3.52) est dite multiplicative pour la raison évidente que l'on multiplie les anomalies par le coefficient λ . Il existe d'autres méthodes d'inflation dites additives car plutôt que de multiplier les anomalies on ajoute une perturbation aléatoire aux membres de l'ensemble (voir [Mitchell and Houtekamer, 2000, Corazza et al., 2003])

La localisation

La localisation a été introduite par [Houtekamer and Mitchell, 2001] et [Hamill et al., 2001] et a pour but de gommer les covariances longue distance induites par le sous-échantillonnage en multipliant par un produit de Schur (qui n'est rien de plus qu'un produit terme à terme de deux matrices) les matrices des covariances d'erreur d'ébauche et d'analyse par une matrice $\mathbf{L}$ contenant des fonctions à support compact.

En général on prend :

$$\mathbf{L}_{ij} = \rho(|p_i - p_j|)$$

où p_i et p_j désignent respectivement les points de l'espace correspondant à la i -ème ligne et à la j -ème colonne de la matrice $\mathbf{B}$ et où ρ est la fonction donnée par [Gaspari and Cohn, 1999] et définie par :

$$\rho(r, c) = \begin{cases} -\frac{1}{4} \left(\frac{r}{c}\right)^5 + \frac{1}{2} \left(\frac{r}{c}\right)^4 + \frac{5}{8} \left(\frac{r}{c}\right)^3 - \frac{5}{3} \left(\frac{r}{c}\right)^2 + 1 & 0 \leq r \leq c \\ \frac{1}{12} \left(\frac{r}{c}\right)^5 - \frac{1}{2} \left(\frac{r}{c}\right)^4 + \frac{5}{8} \left(\frac{r}{c}\right)^3 + \frac{5}{3} \left(\frac{r}{c}\right)^2 - 5 \left(\frac{r}{c}\right) + 4 - \frac{2}{3} \left(\frac{r}{c}\right)^{-1} & c < r \leq 2c \\ 0 & 2c < r \end{cases} \quad (3.53)$$

La fonction ρ dont l'expression est donnée en (3.53) est représentée sur la figure 3.5.

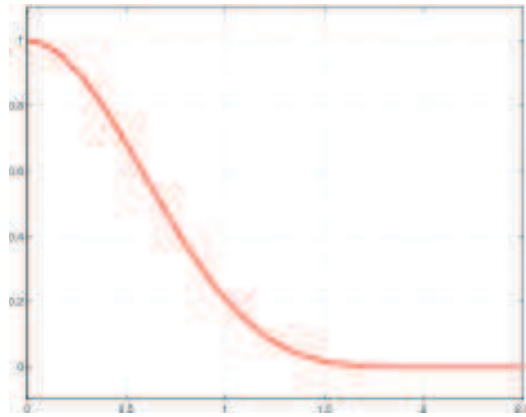


FIG. 3.5 – Représentation graphique de la fonction ρ avec $c = 1$

Chapitre 4

Les prévisions d'ensemble

Sommaire

4.1	Principe de la prévision d'ensemble et représentation	91
4.1.1	Principe	91
4.1.2	Représentation	92
4.2	Description quantitative et évaluation des prévisions d'ensemble	94
4.2.1	Description quantitative	95
4.2.2	Evaluation des prévisions d'ensemble	96

Dans ce travail de thèse l'algorithme d'assimilation de données que nous avons utilisé est basé sur une méthode ensembliste. Comme on le verra au chapitre 8 ce choix a été fait entre autres pour pouvoir étudier facilement les fonctions de covariance et de corrélation dans le cadre de l'assimilation de données pour l'hydraulique fluviale mais aussi de disposer de fonctions de covariance dynamiques et consistantes avec la physique du modèle (écoulement, géométrie de la rivière...). Mais ce n'est pas là le seul intérêt des méthodes ensemblistes, elles permettent en effet d'établir, comme leur nom l'indique, des prévisions dites ensemblistes, ou probabilistes. C'est à ce dernier point auquel on va s'intéresser dans ce chapitre et en particulier on va se pencher sur l'estimation des prévisions d'ensemble. Les covariances d'erreur d'ébauche estimées dans les méthodes d'ensemble étant liées à l'ensemble lui-même il est important de savoir si celui-ci est représentatif des erreurs réelles.

4.1 Principe de la prévision d'ensemble et représentation

4.1.1 Principe

Le principe de la prévision d'ensemble est assez simple, il repose sur une collection de plusieurs réalisations perturbées du modèle à un instant donné. Une prévision d'ensemble consiste alors à estimer à partir de ces réalisations perturbées la densité de probabilité (ou

encore la fonction de répartition ou la fonction de répartition complémentaire) associée au modèle au fil du temps comme cela est représenté sur la figure 4.1. En pratique les différents membres sont issus de différents runs déterministes dont les conditions initiales sont des réalisations de la variable aléatoire au temps initial de prévision.

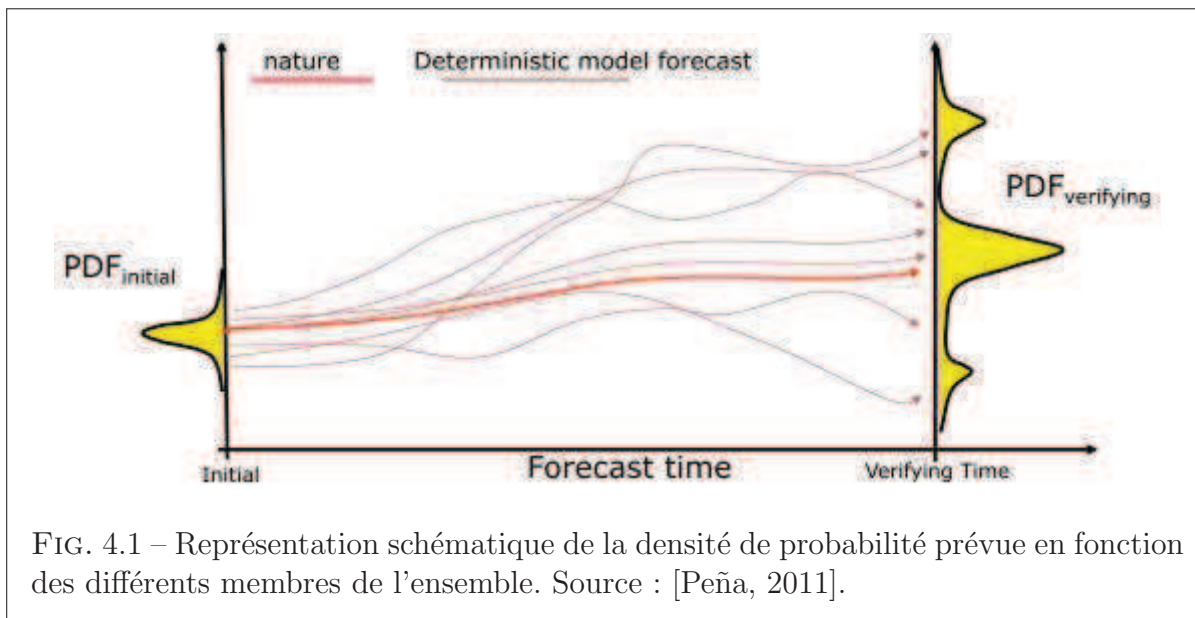


FIG. 4.1 – Représentation schématique de la densité de probabilité prévue en fonction des différents membres de l'ensemble. Source : [Peña, 2011].

4.1.2 Représentation

La représentation d'une prévision d'ensemble (Neal *et al.*, 2007 ; Vrugt *et al.*, 2005 ; De Roo *et al.* 2003) doit permettre de visualiser facilement l'évolution temporelle de l'incertitude liée à la prévision. On considérera ici trois types de représentations communément utilisés dans la littérature et que l'on illustre sur des prévisions de hauteurs d'eau :

- la première représentation est une représentation simple de l'évolution temporelle des membres qui permet d'avoir une idée qualitative de l'incertitude et de l'évolution globale de l'ensemble (voir figure 4.2) ;
- la représentation basée sur l'utilisation de boîtes à moustache (voir figure 4.3). Celle-ci est assez synthétique et représente la médiane, les valeurs extrémales de l'ensemble ainsi que les premier et troisième quartiles ;
- la représentation basée sur la moyenne ainsi qu'une barre d'incertitude correspondant généralement à plus ou moins un écart-type (voir figure 4.4).

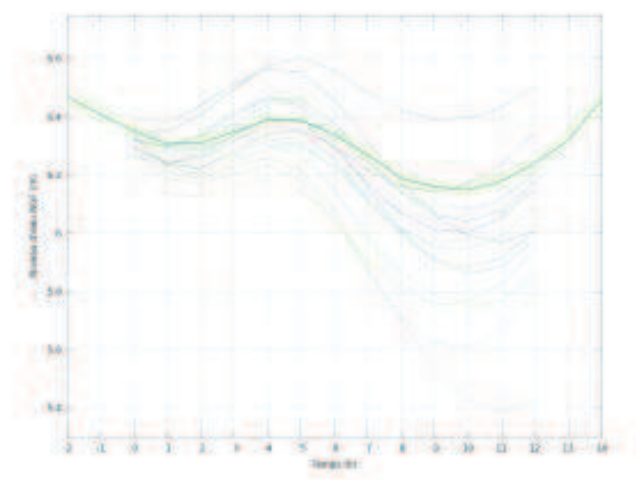


FIG. 4.2 – Série temporelle des observations (ligne verte épaisse) et évolution temporelle des membres (traits fins) lors d’une prévision du temps 0h (temps d’analyse) au temps 12h.

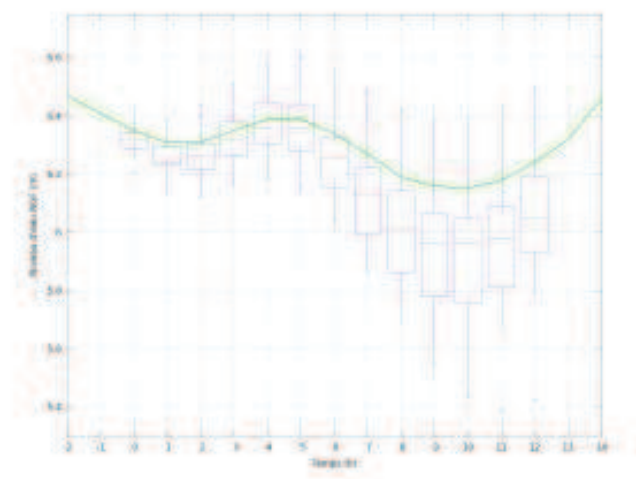


FIG. 4.3 – Série temporelle des observations (ligne verte épaisse) et boîtes à moustache (en bleu) lors d’une prévision du temps 0h (temps d’analyse) au temps 12h.

L’intérêt des figures 4.3 et 4.4 est de synthétiser les informations données par l’ensemble. Les boîtes à moustache représentent les valeurs extrémales de l’ensemble, la

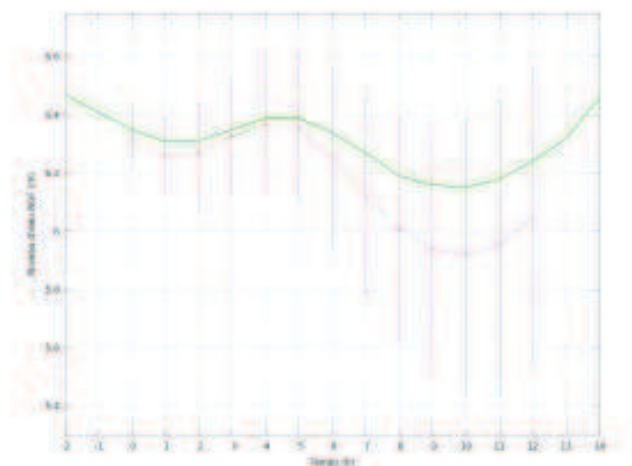


FIG. 4.4 – Série temporelle des observations (ligne verte épaisse) et de la moyenne des membres de l'ensemble plus ou moins deux écarts-types (traits rouges) lors d'une prévision du temps 0h (temps d'analyse) au temps 12h.

médiane ainsi que le premier quartile q_1 et le troisième quartiles q_3 et les outliers (représentés par une croix rouge sur la figure 4.3). Les outliers étant les membres de l'ensemble plus grands (ou plus petits) que $q_3 + w(q_3 - q_1)$ (que $q_1 - w(q_3 - q_1)$) avec $w = 1.5$ ce qui correspond à plus ou moins 2.7 écart-type soit une couverture de 99.3% des valeurs de l'ensemble si celui-ci suit une loi normale. La moyenne plus ou moins deux écart-types représente une couverture d'environ 95% des valeurs de l'ensemble si celui-ci suit une loi normale.

4.2 Description quantitative et évaluation des prévisions d'ensemble

La description quantitative est une information numérique qui synthétise tout ou partie de l'information contenue dans l'ensemble. Bien que cela puisse paraître contradictoire avec l'idée d'ensemble on distingue d'une part le point de vue déterministe dans lequel on se ramène à quelques informations scalaires (voir [Peña, 2011]) comme la moyenne et l'écart-type. Dans ce cas la validation des prévisions se fait en comparant la moyenne des membres (ou encore la médiane ou le mode) et les observations en utilisant des scores standards tels que le biais et/ou la RMSE. Nous ne reviendrons pas dans ce travail sur l'évaluation de ce type de prévisions. On distingue d'un autre côté le point de vue probabiliste où l'on s'intéresse à la probabilité d'occurrence de certains événements. Dans ce cas la validation des prévisions est plus délicate puisqu'il s'agit de déterminer si l'ensemble estime correctement les probabilités d'occurrence des événements (entendus au sens probabiliste du terme). Pour répondre à cette interrogation un certain nombre de critères ont été développés dans le cadre des prévisions d'ensemble météorologiques. Dans cette

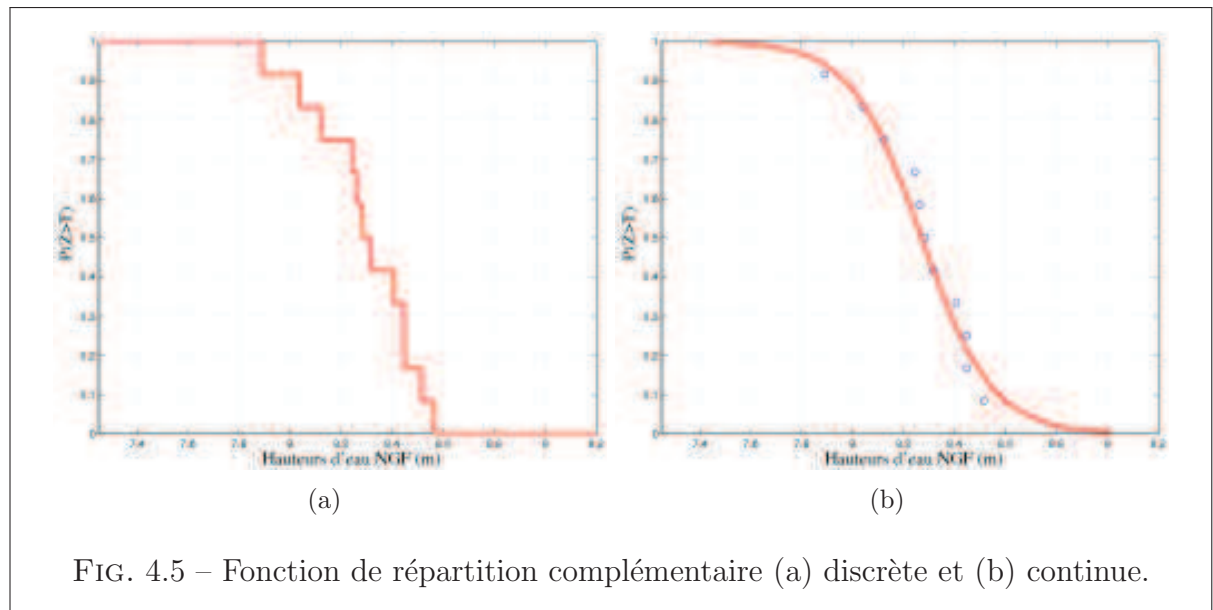
section nous nous intéressons dans un premier temps à la description quantitative des prévisions probabilistes puis dans un second temps nous nous intéresserons dans le détail aux critères pour l'évaluation des prévisions d'ensemble.

4.2.1 Description quantitative

Dans le point de vue probabiliste on considère que chaque membre est une réalisation indépendante et équi-probable de la même variable aléatoire. Si on note N le nombre de membres de l'ensemble et N_E le nombre de membres de l'ensemble vérifiant un évènement E alors on peut estimer la probabilité que l'évènement se réalise avec la formule :

$$P(E) \approx \frac{N_E}{N} \quad (4.1)$$

Dans ce travail nous nous intéresserons principalement à des évènements du type $\{Z > T\}$ où Z représente des hauteurs d'eau prévues et T un seuil (T pour Threshold en anglais). C'est pourquoi pour représenter de manière claire les probabilités calculées par l'ensemble on peut calculer la fonction de répartition complémentaire empirique en calculant pour chaque valeur T la probabilité de l'évènement $\{Z > T\}$ en utilisant la formule (4.1). C'est ce que l'on a représenté en figure 4.5-(a) où l'on a reporté un exemple de fonction de répartition complémentaire empirique discrète



Le fait que la fonction de répartition complémentaire ainsi calculée soit décroissante, comprise entre 0 et 1 et telle que ses limites en $-\infty$ et $+\infty$ soient 1 et 0 respectivement amène à approcher celle-ci par une fonction continue représentée en figure 4.5-(b) (courbe rouge) et dont l'expression (donnée par l'égalité (4.2)) est à rapprocher de celles des fonctions logistiques :

$$\frac{1}{1 + ae^{rx}} \quad (4.2)$$

Une fonction de répartition complémentaire continue permet, par rapport à une fonction de répartition discrète, d'estimer plus finement la probabilité d'un événement, mais aussi de comparer plus facilement des ensembles avec des nombres de membres différents.

4.2.2 Evaluation des prévisions d'ensemble

Avant de chercher à évaluer les prévisions d'ensemble il faut d'abord cerner quelles sont les propriétés qu'un ensemble est censé vérifier. [Talagrand et al., 1997] distingue un certain nombre de propriétés parmi lesquelles nous retiendrons :

- **la consistance** : est-ce que les membres et les observations suivent la même variable aléatoire ?
- **la fiabilité** : la probabilité prédite *a priori* coïncide-t-elle avec la fréquence observée *a posteriori* ?
- **la précision** : la prévision était-elle proche de ce qui s'est passé ?
- **l'habileté** : dans quelle mesure une prévision d'ensemble est-elle meilleure qu'une prévision de référence ?
- **la discrimination** : l'ensemble est-il capable de distinguer les situations menant à l'occurrence d'un événement de celles menant à la non-occurrence du même événement ?

Pour évaluer chacune de ces propriétés il existe un critère que nous détaillerons par la suite. Dans un souci de clarté nous avons regroupé dans un tableau les propriétés mentionnées ci-dessus ainsi que les critères correspondants :

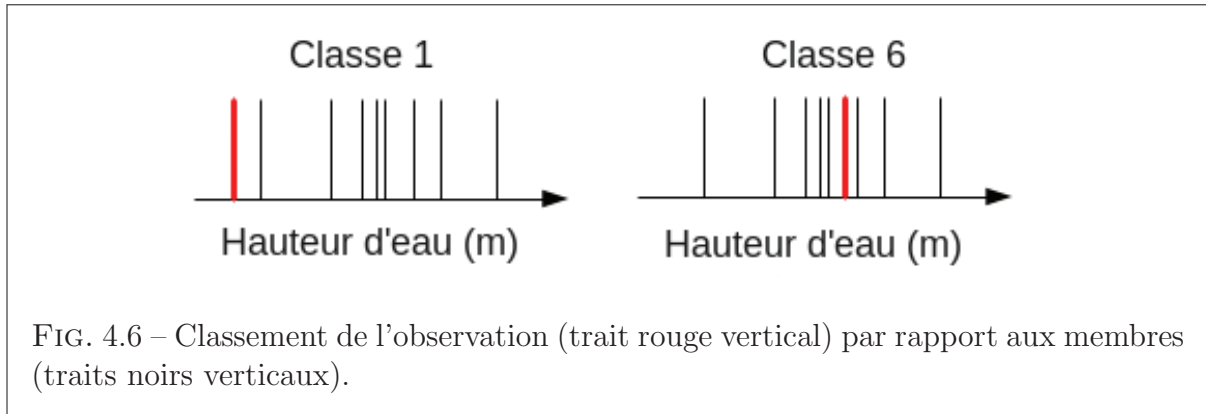
Propriétés	Critères
Consistance	Diagramme de rang (ou de Talagrand)
Fiabilité	Diagramme de fiabilité
Précision	Brier Score
Habileté	Brier Skill Score (BSS)
Précision intégrée sur toute la PDF	Discrete and Continuous Ranked Probability Score (DRPS & CRPS)
Habileté intégrée sur toute la PDF	Discrete and Continuous Ranked Probability Skill Score (DRPSS & CRPSS)
Discrimination	Courbe de ROC (Relative Operating Characteristic)

TAB. 4.1 – Propriétés et critères correspondants.

Consistance et diagramme de rang

Un ensemble vérifie la propriété de consistance si les membres de l'ensemble et les observations suivent la même variable aléatoire. Pour établir si l'ensemble vérifie cette propriété on établit un diagramme de rang (ou encore appelé diagramme de Talagrand)

(voir [Talagrand et al., 1997, Anderson, 1995]). Pour chaque observation on classe par ordre croissant les membres créant ainsi $N + 1$ classes (où N est le nombre de membres de l'ensemble) et on range ensuite l'observation dans l'une de ces classes comme cela est représenté sur la figure 4.6.



Idéalement si les membres et les observations suivent la même variable aléatoire le diagramme tracé doit être plat (voir figure 4.7). Néanmoins si le diagramme n'est pas plat celui-ci peut nous fournir une information sur la dispersivité de l'ensemble ou sur le biais. Un ensemble sur-dispersif aura une grande occurrence d'observations parmi les classes du centre du diagramme, celui-ci formera alors une « bosse » en son centre. Au contraire un ensemble sous-dispersif aura une grande occurrence d'observations parmi les classes aux extrémités du diagramme. Celui-ci présentera de grandes barres en ses extrémités et un creux en son centre. Enfin si l'ensemble est biaisé certaines classes sont privilégiées en termes d'occurrence d'observation et le diagramme correspondant présente une asymétrie. Ces différentes configurations de diagramme sont regroupées dans la figure 4.7.

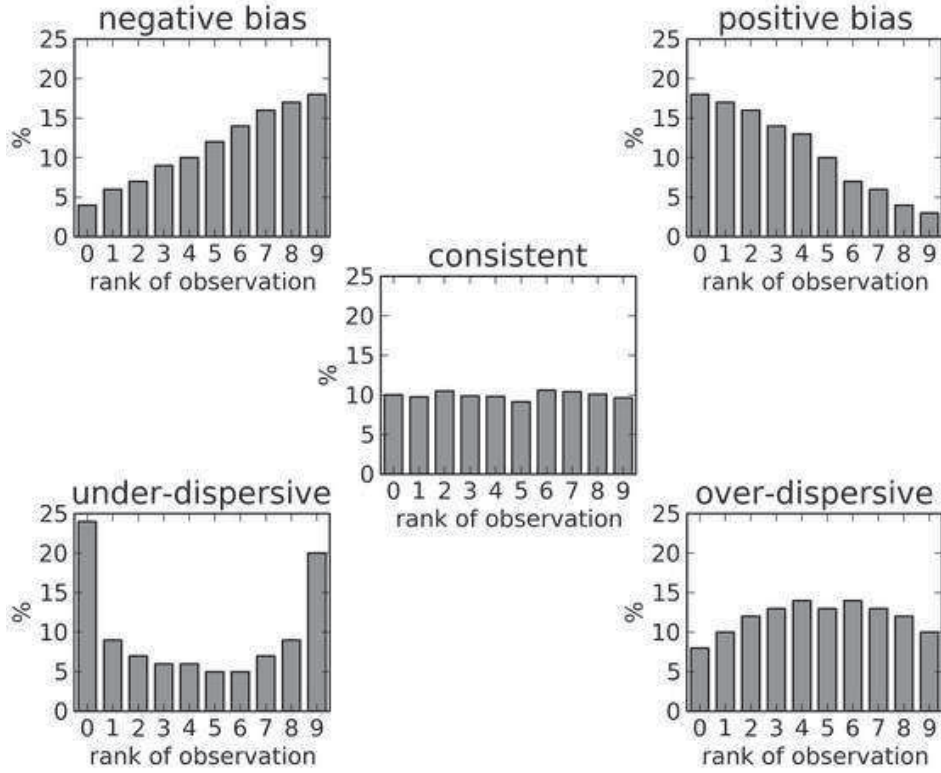


FIG. 4.7 – Différentes configurations pour le diagramme de rang. Source : [Callado et al., 2013].

L'information apportée par ces diagrammes peut être complétée par le calcul de la RMSD entre le diagramme considéré et le diagramme plat correspondant à un ensemble parfaitement consistant :

$$RMSD = \sqrt{\frac{1}{N+1} \sum_{k=1}^{N+1} \left(S_k - \frac{M}{N+1} \right)^2}$$

où N est le nombre de membres, S_k le nombre d'occurrences dans la k -ème classe, et M est le nombre total d'échantillons avec lequel est établi le diagramme.

Fiabilité et diagramme de fiabilité

Dans cette section on considère la fiabilité comme la fiabilité relative à un événement probabiliste du type $\{Z > T\}$. Un ensemble est fiable si les probabilités prédites *a priori* coïncident avec les fréquences observées *a posteriori*. Pour vérifier si l'ensemble est fiable on peut établir un diagramme de fiabilité. Celui-ci admet en abscisse les probabilités prédites et en ordonnées les fréquences relatives d'observation. Pour chaque probabilité prédite on va calculer la fréquence relative d'observation et placer le point correspondant dans un repère et ainsi tracer une courbe. Idéalement si l'ensemble est fiable la courbe

doit être très proche de la diagonale. Si le diagramme de fiabilité n'est pas proche de la diagonale on peut néanmoins en tirer des informations sur l'ensemble : s'il est sous-prédictif ou de faible fiabilité... Les cas les plus standards sont illustrés en figure 4.8. Il peut également subvenir des combinaisons de ces différentes configurations : un ensemble peut par exemple être sous-prédictif pour des probabilités inférieures à 0.3 et sur-prédictif au-delà ; ou un ensemble peut être très fiable pour des probabilités inférieures à 0.5 et être tout à fait non fiable au-delà.

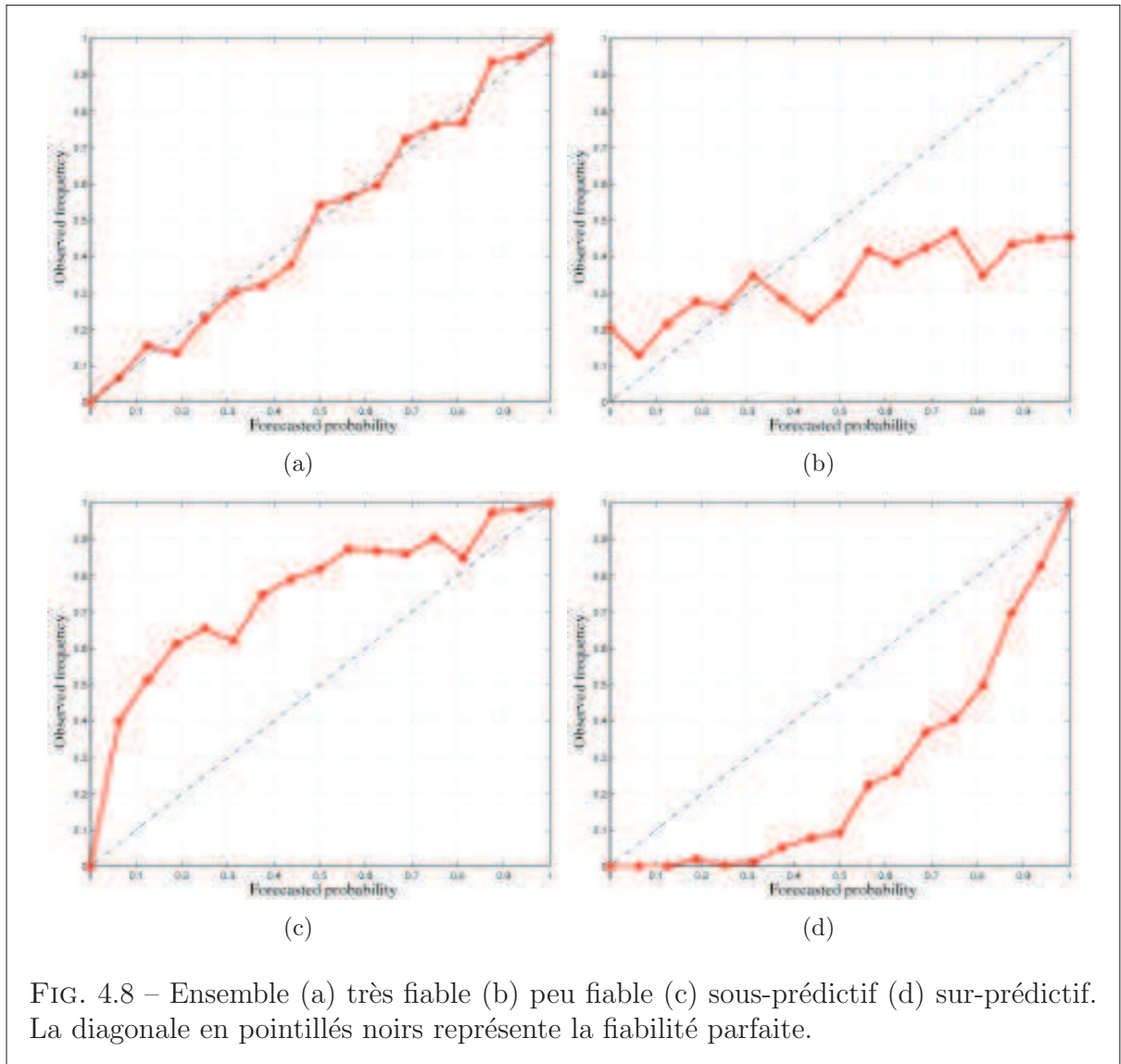


FIG. 4.8 – Ensemble (a) très fiable (b) peu fiable (c) sous-prédictif (d) sur-prédictif. La diagonale en pointillés noirs représente la fiabilité parfaite.

Précision, score de Brier et habileté

La précision de la prédiction d'un évènement du type $\{Z > T\}$ peut se mesurer avec le score de Brier [Brier, 1950, Murphy, 1973] dont l'expression est donnée en (4.3) :

$$BS = \frac{1}{T} \sum_{i=1}^T (p_i - o_i)^2 \quad (4.3)$$

où pour chaque temps i , p_i est la probabilité prédite, o_i vaut 0 ou 1 selon que respectivement l'évènement s'est produit ou pas. Le score de Brier confronte donc les probabilités prédites d'occurrence d'un évènement avec son occurrence ou non-occurrence effective. Il apparait donc que la valeur du score de Brier est comprise entre 0 et 1. Si le système de prévision prévoit avec une grande exactitude l'occurrence ou la non-occurrence de l'évènement alors le score de Brier sera proche de 0 tandis que dans le cas contraire le score de Brier sera proche de 1.

On peut montrer [Murphy, 1973] que le score de Brier peut être décomposé en trois termes : la fiabilité (qui correspond à la propriété de fiabilité évoquée plus haut), la résolution (qui correspond à la propriété de discrimination évoquée plus haut) et l'incertitude inhérente à l'évènement considéré (qui est sa variance).

$$BS = \underbrace{\frac{1}{T} \sum_{k=1}^{N+1} n_k (p_k - \bar{o}_k)^2}_{\text{Fiabilité}} - \underbrace{\frac{1}{T} \sum_{k=1}^{N+1} n_k (\bar{o}_k - \bar{o})^2}_{\text{Résolution}} + \underbrace{\bar{o}(1 - \bar{o})}_{\text{Incertitude}} \quad (4.4)$$

où N est le nombre de membre de l'ensemble, $p_k = \frac{k-1}{N}$ est la probabilité associée à chaque classe de l'ensemble, $\bar{o}_k$ la fréquence d'observation étant donnée la probabilité p_k et $\bar{o} = \sum_{i=1}^T \frac{o_i}{T}$.

Il existe donc un lien entre la fiabilité et le score de Brier à travers la formule (4.4). D'autre part le terme d'incertitude évoqué ici n'a rien à voir avec la qualité de la prévision mais correspond simplement à la variance de l'évènement considéré, il est maximal lorsque $\bar{o} = 0.5$.

Le terme d'incertitude $\bar{o}(1 - \bar{o})$ ne permet pas de comparer le score de Brier sur différents échantillons. Pour dépasser cette limitation le Brier Skill Score (BSS) a été introduit. Celui-ci permet d'estimer dans quelle mesure un système de prévision permet d'améliorer la précision par rapport à un système dit de référence, c'est la propriété d'habileté évoquée précédemment. Le BSS est donné par l'expression (4.5) :

$$BSS = -\frac{BS - BS_{ref}}{BS_{ref}} = 1 - \frac{BS}{BS_{ref}} \quad (4.5)$$

Le BSS est compris entre $-\infty$ et 1. Si le BSS est négatif alors $BS \geq BS_{ref}$ et donc le système de prévision n'améliore pas les prévisions en comparaison du système de référence. Au contraire si le BSS vaut 1 alors le Score de Brier du système de prévision est nul et donc améliore parfaitement les prévisions par rapport au système de prévision de référence.

Discrete and Continuous Ranked Probability Score

Le Score de Brier s'applique à un évènement, on peut cependant l'étendre à un ensemble d'évènements E_k , $k \in \{1, \dots, K\}$ [Weigel et al., 2007]. On peut associer à chaque

évènement une probabilité prédite $p_t^{|k}$ et une observation de l'évènement $o_t^{|k} \in \{0, 1\}$. On peut alors évaluer la distribution complète :

$$DRPS = \frac{1}{TK} \sum_{t=1}^T \sum_{k=1}^K \left(p_t^{|k} - o_t^{|k} \right)^2 \quad (4.6)$$

De la même manière que le Score de Brier, le $DRPS$ est compris entre 0 et 1, il est proche de 0 si le système prédit avec une grande exactitude l'occurrence ou la non-occurrence des évènements et dans le cas contraire il sera proche de 1.

De la même manière que pour le Brier Skill Score il existe le Discrete Ranked Probability Skill Score :

$$DRPSS = -\frac{DRPS - DRPS_{ref}}{DRPS_{ref}} = 1 - \frac{DRPS}{DRPS_{ref}}$$

Pour le $DRPS$ on considèrerait un ensemble discret d'évènements, pour étendre le Score de Brier à un ensemble d'évènements on peut également considérer un ensemble continu d'évènements, c'est l'objet du Continuous Ranked Probability Score ($CRPS$). Si on note $F_p(z)$ la fonction de répartition associée à l'ensemble et $F_o(z) = \mathbb{1}_{[y; +\infty[}(z)$ la fonction de répartition associée à l'observation alors l'expression du $CRPS$ est donnée par :

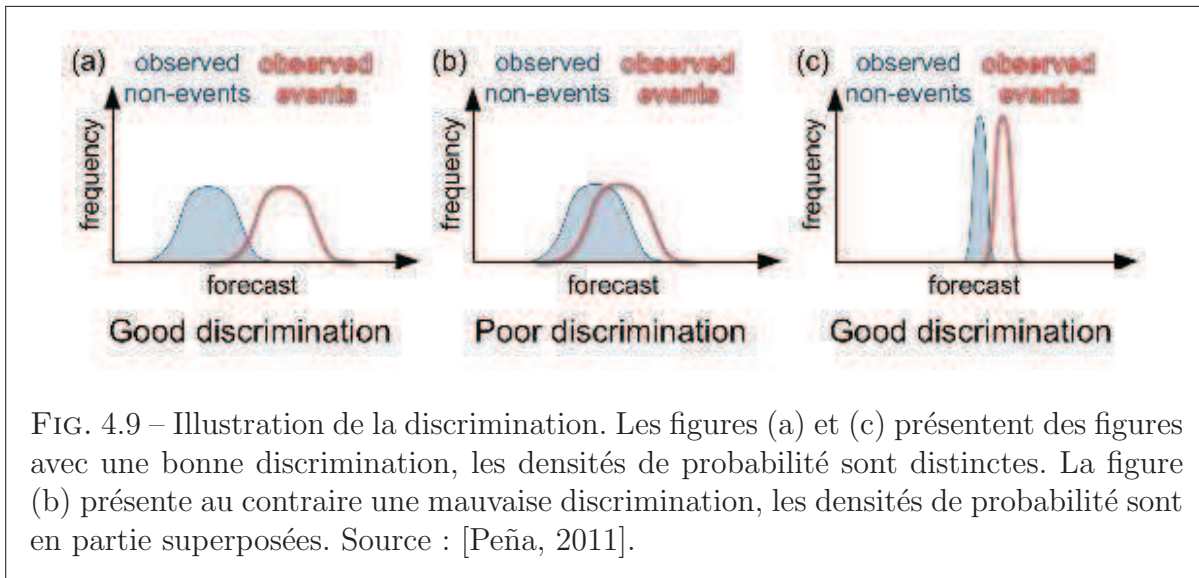
$$CRPS = \frac{1}{T} \sum_{t=1}^T \int_{\mathbb{R}} (F_p(z) - F_o(z))^2 dz \quad (4.7)$$

Là encore il existe le Continuous Ranked Probability Skill Score qui permet de comparer le système de prévision à un système de référence :

$$CRPSS = -\frac{CRPS - CRPS_{ref}}{CRPS_{ref}} = 1 - \frac{CRPS}{CRPS_{ref}}$$

Table de contingence, discrimination et courbe de ROC

La discrimination est la capacité du système de prévision à distinguer clairement les situations menant à l'occurrence d'un évènement des situations menant à la non-occurrence de cet évènement (voir [Wilks, 2011]). Cela se traduit par le fait que les densités de probabilités respectives sont « à support disjoints ». On a reporté en figure 4.9 des images illustrant la propriété de discrimination.



Dans la pratique pour établir si un système vérifie la propriété de discrimination on peut établir la courbe de ROC (Relative Operating Characteristic). Cette courbe est établie grâce à certains scores qui eux-mêmes sont établis à partir de la table de contingence relative à l'évènement considéré. Dans la suite nous présentons donc dans un premier temps ce qu'est une table de contingence ainsi que quelques scores associés parmi lesquels ceux que nous utiliserons pour tracer la courbe de ROC.

Dans le cadre d'un système de prévision une table de contingence est un tableau qui pour un évènement confronte d'une part les états prévus (prédit/non prédit) aux états observés (observé/non observé) (voir tableau 4.2). On considérera que l'évènement a été prédit si la probabilité prédite par l'ensemble est supérieure à la fréquence d'observation de l'évènement.

	Prédit	Non prédit
Observé	Correct positif <i>Succès</i>	Faux négatif <i>Echec</i>
Non observé	Faux positif <i>Fausse alerte</i>	Correct négatif

TAB. 4.2 – Table de contingence.

A partir du renseignement de cette table on peut dériver un certain nombre de scores relatifs à un évènement (voir table 4.3) qui nous renseignent sur la capacité du système à prédire l'évènement. Nous en présentons ici quelques-uns sans pour autant en donner le sens, ceux-ci s'interprétant assez facilement.

Scores	Formules
Probabilité d'occurrence	$\frac{Succès}{Succès + Fausse\ alerte}$
Précision	$\frac{Succès + Correct\ négatif}{Total}$
Taux de réussite	$\frac{Succès}{Succès + Echec}$
Taux de corrects négatifs	$\frac{Correct\ négatif}{Correct\ négatif + Fausse\ alerte}$
Menace	$\frac{Succès}{Succès + Echec + Fausse\ alerte}$
Taux de faux positifs	$\frac{Fausse\ alerte}{Fausse\ alerte + Correct\ négatif}$

TAB. 4.3 – Quelques scores associés à la table de contingence.

Chaque point de la courbe de ROC est constitué d'un couple (taux de faux positifs ; taux de réussite) (qui correspondent respectivement à l'abscisse et à l'ordonnée du point) obtenu pour un seuil de probabilité différent. Alors si pour chaque seuil de probabilité testé le taux de faux positifs est égal au taux de réussite cela signifie que le système est incapable de discriminer entre deux situations différentes. Dans ce cas la courbe de ROC est superposée à la diagonale comme on peut le voir sur la figure 4.10-(b) et l'aire sous la courbe vaut 0.5. Au contraire si quel que soit le seuil de probabilité utilisé le taux de réussite est toujours égal à 1 alors cela signifie que le système de prévision arrive parfaitement à discriminer entre deux situations. Dans ce cas la courbe de ROC est confondue avec le segment $[0;1] \times 1$ comme on peut le voir sur la figure 4.10-(a) et l'aire sous la courbe vaut 1. De manière plus générale plus l'aire sous la courbe de ROC est proche de 1 plus le système de prévision est capable de faire la distinction entre deux situations. De cette réflexion sur les aires sous la courbe on peut dériver le ROC Skill Score (*ROCSS*) qui détermine dans quelle mesure le système de prévision peut discriminer entre deux événements. Si on note A l'aire sous la courbe de ROC alors l'expression de ROCSS est donnée par (4.8).

$$ROC_{SS} = 2A - 1 \quad (4.8)$$

Plus ROC_{SS} est proche de 0 et moins le système de prévision est capable de discriminer entre deux évènements. Au contraire plus ROC_{SS} est proche de 1 et plus le système de prévision est discriminant.

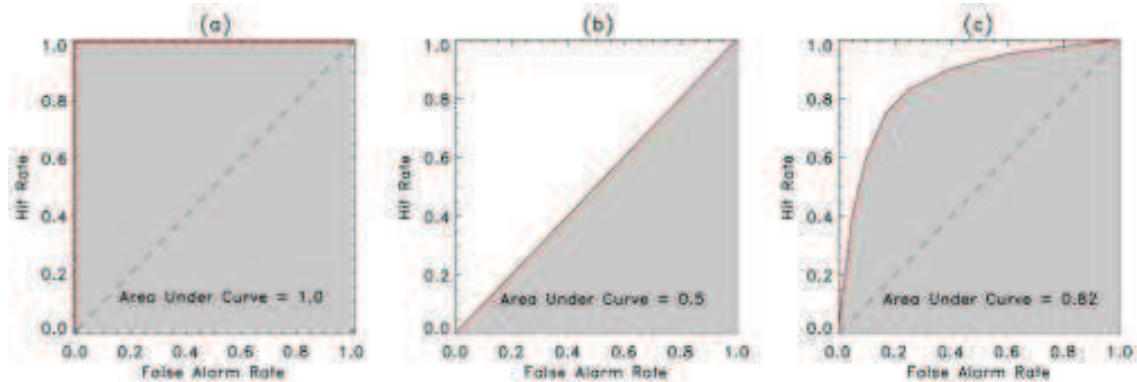


FIG. 4.10 – Courbes de ROC pour un système de prévisions (a) parfait (b) incapable de discriminer (c) réaliste.

Chapitre 5

Couplage de modèles hydrauliques 1D-2D

Sommaire

5.1	Couplage à interfaces	106
5.1.1	Raccordement des modèles à l'interface	107
5.1.2	Algorithme de couplage	109
5.1.3	Quelques exemples d'applications numériques de l'algorithme de couplage	111
5.2	Couplage de modèles superposés	118
5.2.1	Calcul des termes d'échange entre lit mineur et lit majeur	120
5.2.2	Equations de Saint-Venant 1D conservatives avec terme source	124
5.2.3	Algorithme de couplage	124
5.2.4	Variations autour d'une application numérique	125

Dans le cadre de la modélisation hydraulique des cours d'eau les modèles 2D ont un réel apport en comparaison des modèles 1D en ce qu'ils fournissent une description précise et potentiellement très fine des phénomènes modélisés. Néanmoins le développement et l'utilisation dans la pratique de modèles 2D à l'échelle d'un réseau hydraulique n'est pas toujours possible du fait des données que cela suppose mais aussi et surtout du fait des coûts de calcul que cela implique. D'autant plus que l'utilisation d'un modèle 2D n'est pleinement justifiée que là où l'écoulement est bi-dimensionnel ce qui est loin d'être toujours le cas dans le cadre de l'écoulement d'une rivière. D'autre part certains modèles 2D sont développés sur des zones d'intérêt (zones urbaines, centrale nucléaire...) couvrant en partie une zone où il existe déjà un modèle 1D calibré et validé et dans lequel beaucoup d'expertise a été investie. Ainsi pour limiter les temps de calcul tout en tirant pleinement profit de modèles 1D et 2D là où ceux-ci ont de l'intérêt et pour assurer la continuité physique entre les différents modèles il est judicieux d'avoir recours à des méthodes de couplage de modèles.

Dans la suite de ce travail on s'intéressera au couplage de modèles Saint-Venant 1D et 2D. Ce thème de recherche a fait l'objet de nombreuses études ces dernières années démontrant à la fois l'intérêt et la faisabilité de la démarche aussi bien sur des cas d'étude académiques que sur des cas tests réels. Parmi les méthodes ainsi développées on dis-

tingue :

- le couplage de modèles avec interfaces aux frontières des modèles. Ici le terme de frontières désigne les frontières liquides dans le cas d'un modèle 2D et les frontières amont et/ou aval pour un modèle 1D. Dans ce cas les variables échangées à l'interface peuvent être le débit, la section mouillée, les invariants de Riemann...
- le couplage de modèles superposés. Dans ce cas le modèle 1D force le modèle 2D et ce dernier calcule les flux et les termes de quantité de mouvement échangés entre les lits mineur et majeur. Ces termes sont ensuite injectés dans le modèle 1D pour modéliser les échanges entre lit mineur et lit majeur ;
- le couplage latéral. Cette méthode est dans le principe assez proche de la précédente mais diffère de celle-ci en ce que les modèles ne sont pas superposés : un ou plusieurs modèles 2D modélisent l'écoulement dans les plaines d'inondations et un modèle 1D modélise l'écoulement dans le lit de la rivière. Le lecteur intéressé pourra se reporter à [Goutal et al., 2014].

Ces deux dernières méthodes sont plus particulièrement indiquées lorsque l'on souhaite modéliser les écoulements dans les plaines d'inondation adjacentes au cours d'eau. Dans la suite nous ne nous intéresserons qu'aux deux premières méthodes de couplage citées : le couplage à interfaces et le couplage de modèles superposés.

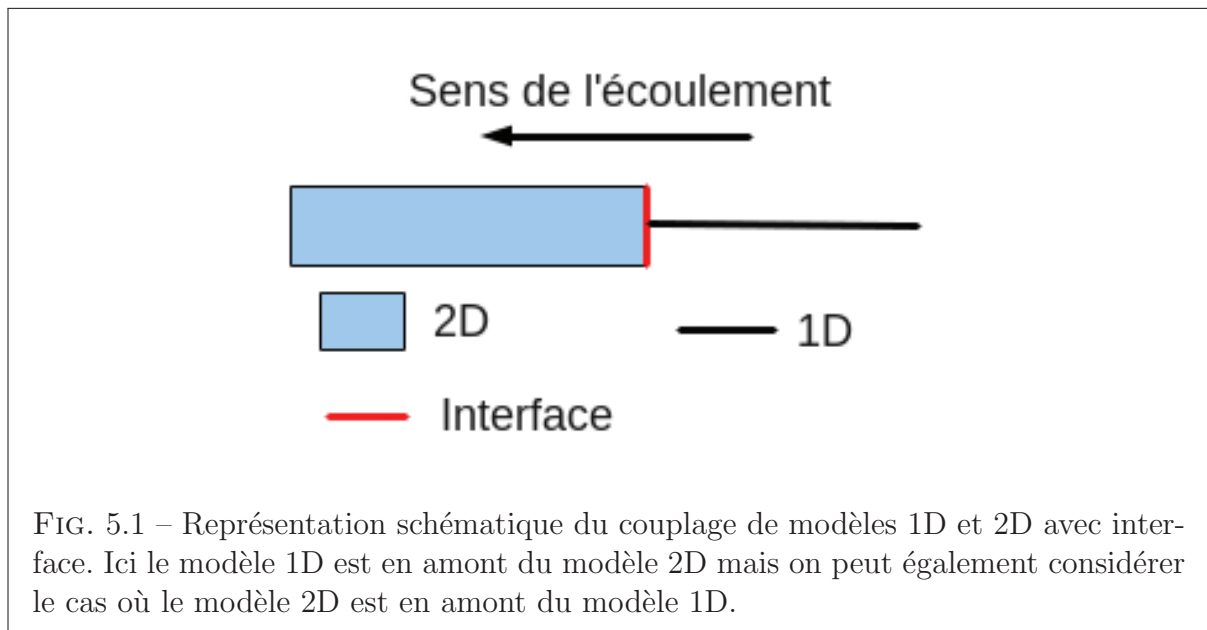
5.1 Couplage à interfaces

Nicole Goutal dans sa thèse de HDR [Goutal, 2014] a fait un travail de synthèse très clair sur les différents aspects intervenant dans le couplage de modèles de Saint-Venant 1D et 2D. Pour rédiger cette section ainsi que pour implémenter la méthode de couplage à interfaces (voir chapitre 9, section 9.2) nous nous sommes largement appuyés sur ce travail et sur les travaux exposés dans [Miglio et al., 2005, Malleron et al., 2011].

Le couplage de modèles à interfaces peut être mis en œuvre lorsque plusieurs modèles sont utilisés séquentiellement le long d'un cours d'eau comme cela est illustré schématiquement sur la figure 5.1. Dans la suite, dans un souci de simplicité mais sans perte de généralité, nous considérerons que le modèle 1D est situé en amont du modèle 2D mais dans les faits nous pouvons également considérer des configurations où le 2D est en amont du 1D ou encore des configurations du type 1D-2D-1D.

Usuellement, pour démontrer la pertinence du couplage on met en place sur la zone d'étude un modèle 2D complet qui recouvre les domaines couverts par les modèles 1D et 2D. Le but de la démarche est de coupler les modèles tout en minimisant l'erreur à l'interface entre d'un côté les modèles couplés et d'autre part le modèle 2D complet qui sert de référence. Si on dispose d'observations on pourra également comparer le modèle couplé aux observations. La méthode de couplage à interfaces repose sur un algorithme itératif (que l'on détaille plus loin) et qui doit permettre (à convergence) de vérifier les contraintes suivantes [Goutal, 2014] :

- la continuité des grandeurs physiques transmises d'un modèle à l'autre doit être assurée. Néanmoins la variable transmise ne sera pas nécessairement la condition limite imposée à l'autre modèle. Par exemple dans le cas où le modèle 1D transmet



un débit au modèle 2D alors les vitesses imposées à celui-ci en condition limite sont calculées de telle manière que les débits aux limites des modèles sont égaux mais les vitesses calculées ne le sont pas forcément. C'est le cas en particulier si les sections en travers à l'interface de chacun des modèles sont différentes ;

- les conditions limites doivent être telles que les problèmes sur chaque sous-domaine sont bien posés et la vitesse de convergence de l'algorithme maximale ;
- à l'interface le modèle le plus grossier doit être valide. Ainsi l'interface entre les modèles doit être positionnée là où l'écoulement est mono-dimensionnel, dans ce cas les deux modèles sont valides au niveau de l'interface et on minimise l'erreur de couplage.

Afin de s'assurer que les contraintes ci-dessus sont bien respectées il faut préciser comment est opéré le raccordement entre les modèles à l'interface.

5.1.1 Raccordement des modèles à l'interface

Le raccordement des modèles à l'interface doit se faire de telle manière que les problèmes sur chacun des sous-domaines 1D et 2D soient bien posés mais aussi que la continuité entre les grandeurs physiques à l'interface soit assurée.

Problèmes bien posés pour les sous-modèles Saint-Venant 1D et 2D

Pour que les problèmes pour les sous-modèles Saint-Venant 1D et 2D soient bien posés il faut que les conditions aux limites de ces modèles soient en conformité avec la théorie des caractéristiques que l'on a évoquée à la section 2.4.1 dans le cas d'un modèle 1D. Celles-ci vont donc dépendre du régime d'écoulement mais aussi du modèle (1D ou 2D) ainsi que de la position relative des modèles.

Pour le **modèle 1D** les conditions limites à imposer au niveau de l'interface sont :

- en **régime fluvial** : la section mouillée ou le débit ou l'invariant de Riemann lié à la caractéristique entrante ;
- en **régime torrentiel** : il n'y a aucune condition à imposer puisqu'il n'y a pas de caractéristique entrante à l'aval du domaine 1D.

Pour le **modèle 2D** les conditions limites à imposer au niveau de l'interface sont :

- en **régime fluvial** : la hauteur d'eau en chaque point de l'interface ou la vitesse en chaque point de l'interface ou l'invariant de Riemann lié à la caractéristique entrante ;
- en **régime torrentiel** : on impose soit la hauteur et la vitesse soit les deux invariants de Riemann portés par les deux caractéristiques entrant dans le domaine.

Continuité des grandeurs physiques au niveau de l'interface

Les conditions limites à imposer aux modèles au niveau de l'interface impliquent la continuité de certaines variables physiques au niveau de cette interface.

• Continuité des variables 1D

Les variables du modèle 1D sont la section mouillée et le débit, il convient donc à convergence de l'algorithme de couplage d'assurer la continuité de ces deux variables ou la continuité des invariants de Riemann 1D :

- continuité de la section mouillée : afin que le problème 2D soit bien posé la continuité de la section mouillée au niveau de l'interface se traduit dans le modèle 2D par une hauteur d'eau qui servira de condition limite. Si au niveau de l'interface les géométries des modèles 1D et 2D diffèrent alors la continuité de la section mouillée résultera en la discontinuité des cotes d'eau. Il faut préciser ici que les sections en travers de la rivière sont décrites dans la géométrie du modèle 1D, ainsi cela a du sens de comparer les géométries des modèles 1D et 2D à l'interface entre ces deux modèles ;
- continuité du débit : afin que le problème 2D soit bien posé un profil de vitesse (constant ou parabolique) est calculé à partir des débits imposés par le modèle 1D.
- continuité des invariants de Riemann 1D.

• Continuité des variables 2D

Les variables du modèle 2D sont les vitesses horizontales et la hauteur d'eau. A convergence il convient d'assurer la continuité de ces deux variables :

- continuité des cotes d'eau : si l'interface de couplage est située dans une zone où l'écoulement est mono-dimensionnel alors la cote d'eau sur la section en travers du modèle 2D devrait être constante. Sinon on définit la cote d'eau du modèle 2D au niveau de l'interface comme étant la moyenne des cotes d'eau du modèle 2D prise sur la longueur de la section en travers à l'interface ;
- continuité des vitesses : si l'interface de couplage est située dans une zone où l'écoulement est mono-dimensionnel alors la vitesse 2D est égale à la vitesse 1D avec un profil uniforme. Par contre, de la même manière que pour la continuité de la section mouillée évoquée plus haut, si les géométries des sections en travers des modèles 1D et 2D au niveau de l'interface de couplage ne sont pas identiques la

continuité des débits ne sera pas assurée.

5.1.2 Algorithme de couplage

Dans le cadre du couplage à interface l'algorithme de couplage généralement utilisé est un algorithme de type Schwarz sans recouvrement [Lions, 1996]. Ces méthodes de couplage ont été initialement élaborées dans un cadre purement mathématiques par H. A. Schwarz [Schwarz, 1869] au 19ème siècle afin d'établir l'existence de fonctions harmoniques sur des domaines de formes particulières comme par exemple l'union d'un cercle et d'un rectangle. Dans la suite nous présenterons tout d'abord les versions locale et globale en temps de l'algorithme de Schwarz dans un cadre général avant de donner leur formulation dans le cadre de modèles Saint-Venant.

L'algorithme de Schwarz

On considère deux ouverts de $\mathbb{R}^n$ (n étant un entier quelconque) disjoints Ω_1 et Ω_2 mais dont l'intersection de leurs frontières notées $\partial\Omega_1$ et $\partial\Omega_2$ n'est pas vide ; on note $\Gamma = \partial\Omega_1 \cap \partial\Omega_2$. On considère sur ces ouverts deux opérateurs aux dérivées partielles $\mathcal{L}_1$ et $\mathcal{L}_2$ représentant deux systèmes physiques. On considère alors l'ensemble des deux problèmes suivants :

$$\left\{ \begin{array}{ll} \mathcal{L}_1 u = f_1 & \text{sur } \Omega_1 \\ B_1 u = g_1 & \text{sur } \partial\Omega_1 \setminus \Gamma \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{ll} \mathcal{L}_2 v = f_2 & \text{sur } \Omega_2 \\ B_2 v = g_2 & \text{sur } \partial\Omega_2 \setminus \Gamma \end{array} \right. \quad (5.1)$$

où u et v sont les variables associées aux opérateurs $\mathcal{L}_1$ et $\mathcal{L}_2$ (à ne pas confondre avec les composantes horizontales de la vitesse utilisées dans les équations de Saint-Venant 2D), f_1 , g_1 , f_2 et g_2 sont des fonctions données, B_1 et B_2 sont des opérateurs aux limites du domaine définis par la physique du problème et où u et v doivent satisfaire à la contrainte suivante au niveau de l'interface Γ :

$$C_1 u = C_2 v$$

où C_1 et C_2 sont des opérateurs à l'interface.

Pour résoudre le problème (5.1) l'algorithme de Schwarz local en temps consiste à résoudre de manière itérative à chaque pas de temps i , si on suppose connus u_i^0 et v_i^0 sur Γ , le problème suivant :

$$\left\{ \begin{array}{ll} \mathcal{L}_1 u_i^{k+1} = f_1 & \text{sur } \Omega_1 \\ B_1 u_i^{k+1} = g_1 & \text{sur } \partial\Omega_1 \setminus \Gamma \\ C_1 u_i^{k+1} = C_2 v_i^k & \text{sur } \Gamma \end{array} \right. \quad \text{et} \quad \left\{ \begin{array}{ll} \mathcal{L}_2 v_i^{k+1} = f_2 & \text{sur } \Omega_2 \\ B_2 v_i^{k+1} = g_2 & \text{sur } \partial\Omega_2 \setminus \Gamma \\ C_2 v_i^{k+1} = C_1 u_i^k & \text{sur } \Gamma \end{array} \right. \quad (5.2)$$

En règle générale les valeurs u_i^0 et v_i^0 sont les solutions du problème (5.2) au pas de temps précédent $i - 1$. Cette version est appelée version « additive » de l'algorithme de Schwarz et est du fait des conditions sur l'interface Γ naturellement parallélisable.

Il existe également une version dite « multiplicative » de l'algorithme de Schwarz qui se formule de la manière suivante, si on suppose connus u_i^0 et v_i^0 sur Γ alors on résout de manière itérative le problème suivant :

$$\left\{ \begin{array}{lll} \mathcal{L}_1 u_i^{k+1} = f_1 & \text{sur} & \Omega_1 \\ B_1 u_i^{k+1} = g_1 & \text{sur} & \partial\Omega_1 \setminus \Gamma \\ C_1 u_i^{k+1} = C_2 v_i^k & \text{sur} & \Gamma \end{array} \right. \text{ et } \left\{ \begin{array}{lll} \mathcal{L}_2 v_i^{k+1} = f_2 & \text{sur} & \Omega_2 \\ B_2 v_i^{k+1} = g_2 & \text{sur} & \partial\Omega_2 \setminus \Gamma \\ C_2 v_i^{k+1} = C_1 u_i^{k+1} & \text{sur} & \Gamma \end{array} \right. \quad (5.3)$$

Du fait que dans cette méthode on résolve les problèmes successivement sur chacun des sous-domaines Ω_1 et Ω_2 son temps de calcul est en général plus long que pour l'algorithme de Schwarz additif.

On peut introduire une dimension temporelle dans l'algorithme de Schwarz en prenant en compte le temps, dans ce cas on va chercher à résoudre un problème du type (5.3) ou (5.2) mais sur un intervalle de temps $[0; T]$. On suppose connus u^0 et v^0 sur Γ sur l'intervalle de temps $[0; T]$ alors l'algorithme de Schwarz additif global en temps consiste à résoudre de manière itérative le problème suivant :

$$\left\{ \begin{array}{lll} \mathcal{L}_1 u^{k+1} = f_1 & \text{sur} & \Omega_1 \times [0; T] \\ B_1 u^{k+1} = g_1 & \text{sur} & (\partial\Omega_1 \setminus \Gamma) \times [0; T] \\ C_1 u^{k+1} = C_2 v^k & \text{sur} & \Gamma \times [0; T] \end{array} \right. \text{ et } \left\{ \begin{array}{lll} \mathcal{L}_2 v^{k+1} = f_2 & \text{sur} & \Omega_2 \times [0; T] \\ B_2 v^{k+1} = g_2 & \text{sur} & (\partial\Omega_2 \setminus \Gamma) \times [0; T] \\ C_2 v^{k+1} = C_1 u^k & \text{sur} & \Gamma \times [0; T] \end{array} \right. \quad (5.4)$$

Bien sûr on peut également concevoir un algorithme de Schwarz multiplicatif global en temps.

Ces méthodes s'étendent naturellement au cas de plusieurs sous-domaines.

Le choix des opérateurs C_1 et C_2 apparaissant dans les problèmes (5.2), (5.3) et (5.4) est d'une grande importance puisqu'ils vont influencer sur le nombre d'itérations nécessaire pour converger et donc sur le temps de calcul. On remarquera en particulier que dans le cas où $C_1 = C_2 = Id$ l'algorithme de Schwarz ne converge pas car les valeurs de u^0 et v^0 à l'interface n'évoluent pas. Aussi en pratique il est d'intérêt de chercher à définir des opérateurs C_1 et C_2 à l'interface permettant d'optimiser la vitesse convergence de l'algorithme de Schwarz, de nombreuses études traitent de ce sujet dont par exemple [Gander, 2006].

Formulation dans le cadre du couplage 1D-2D des équations de Saint-Venant

Dans le cadre du couplage 1D-2D des équations de Saint-Venant la formulation de l'algorithme de Schwarz est légèrement différente de celle exposée précédemment. Tout d'abord les opérateurs $\mathcal{L}_1$ et $\mathcal{L}_2$ désignent respectivement l'opérateur discrétisé de Saint-Venant sur les domaines 1D et 2D. D'autre part on note au temps i :

- $(S, Q)_i$ la section mouillée et le débit dans le domaine 1D ;
- $(h, u, v)_i$ la hauteur et les composantes de la vitesse dans le domaine 2D.

Dans la suite on définit Γ_1 et Γ_2 comme étant les interfaces de raccordement des modèles 1D et 2D respectivement. Si au temps i on connaît Q_i^0 et $h_{2D,i}^0$ sur Γ_1 et Γ_2 respectivement alors l'algorithme de Schwarz additif local en temps dans le cadre du

couplage des équations Saint-Venant 1D-2D se ramène à résoudre de manière itérative les problèmes suivants :

$$\begin{cases} \mathcal{L}_1(S, Q)_i^{k+1} = f_1 & \text{sur } \Omega_1 \\ \text{Conditions limites classiques} & \text{sur } \partial\Omega_1 \setminus \Gamma_1 \\ S_i^{k+1} = C_{21}(h_{2D,i}^k) & \text{sur } \Gamma_1 \end{cases} \quad (5.5)$$

et

$$\begin{cases} \mathcal{L}_2(h, u, v)_i^{k+1} = f_2 & \text{sur } \Omega_2 \\ \text{Conditions limites classiques} & \text{sur } \partial\Omega_2 \setminus \Gamma_2 \\ (u, v)_i^{k+1} = C_{12}(Q_i^k) & \text{sur } \Gamma_2 \end{cases} \quad (5.6)$$

où Ω_s et $\partial\Omega_s$ désignent respectivement le domaine et la frontière des modèles 1D et 2D et Γ désigne l'interface entre les modèles.

Là encore on peut donner une formulation globale en temps de l'algorithme de Schwarz. Si sur l'intervalle de temps $[0; T]$ on connaît sur Γ_1 et Γ_2 , h_{2D}^0 et Q^0 alors l'algorithme de Schwarz additif global en temps pour les équations Saint-Venant 1D-2D se ramène à résoudre de manière itérative les problèmes suivants :

$$\begin{cases} \mathcal{L}_1(S, Q)^{k+1} = f_1 & \text{sur } \Omega_1 \times [0; T] \\ \text{Conditions limites classiques} & \text{sur } (\partial\Omega_1 \setminus \Gamma_1) \times [0; T] \\ S^{k+1} = C_{21}(h_{2D}^k) & \text{sur } \Gamma_1 \times [0; T] \end{cases} \quad (5.7)$$

et

$$\begin{cases} \mathcal{L}_2(h, u, v)^{k+1} = f_2 & \text{sur } \Omega_2 \times [0; T] \\ \text{Conditions limites classiques} & \text{sur } (\partial\Omega_2 \setminus \Gamma_2) \times [0; T] \\ (u, v)^{k+1} = C_{12}(Q^k) & \text{sur } \Gamma_2 \times [0; T] \end{cases} \quad (5.8)$$

Pour que ces algorithmes soient bien définis il reste à préciser ce que sont dans ce cas les opérateurs C_{12} et C_{21} :

- l'opérateur C_{12} calcule le champ de vitesse à imposer au modèle 2D comme condition limite amont à l'interface en fonction du débit calculé par le modèle 1D à l'interface.
- l'opérateur C_{21} calcule la cote d'eau à imposer en condition limite aval du modèle 1D en prenant la moyenne des cotes d'eau du modèle 2D sur la section en travers au niveau de l'interface.

5.1.3 Quelques exemples d'applications numériques de l'algorithme de couplage

Dans cette section nous nous intéressons à quelques exemples d'application d'un algorithme itératif de couplage de modèles Saint-Venant 1D-2D tirés de la littérature. Dans un premier temps nous nous intéressons à l'étude réalisée dans [Miglio et al., 2005] portant sur des exemples de couplage dans des cas académiques. Dans un second temps nous nous intéressons aux résultats présentés par [Malleron et al., 2011] et dans lequel des cas d'applications réels ont été utilisés : la rupture du barrage de Malpasset et le couplage

de modèles sur le Rhin. Ces travaux ont été choisis pour leur pertinence illustrative dans le cadre de cette section. On peut citer également les travaux présentés dans [?] portant sur le couplage de modèles hydrauliques 1D et 3D réalisés avec les codes hydrauliques Mascaret-1D et Telemac-3D respectivement.

Miglio *et al.*, 2005-a

L'article de [Miglio et al., 2005] porte sur le couplage de modèles Saint-Venant de dimensions hétérogènes (i-e de dimensions différentes). Il est à noter qu'il existe une seconde partie à cet article, [Miglio et al., 2005], dans laquelle il est question du couplage de modèles hydrauliques à surface libre dont la physique est hétérogène mais les dimensions homogènes. Il ne sera pas question de ce dernier travail ici.

Dans ce travail l'auteur s'intéresse en particulier aux conditions aux limites à imposer à l'interface entre les modèles avant d'exposer l'algorithme itératif utilisé pour coupler les modèles. Celui-ci diffère légèrement des problèmes itératifs présentés en (5.5) et (5.6) en section 5.1.2 au niveau des variables échangées à l'interface. Dans ce cas la variable transmise du 2D au 1D est l'invariant de Riemann porté par la caractéristique entrante dans le domaine 1D à l'amont et la variable transmise par le 1D au 2D est la section mouillée. En dehors de ça le principe de l'algorithme itératif de couplage reste le même.

[Miglio et al., 2005] présente tout d'abord un cas académique avec un modèle 2D situé au niveau d'une confluence couplé à deux modèles 1D situés en aval de celle-ci (figure 5.2).

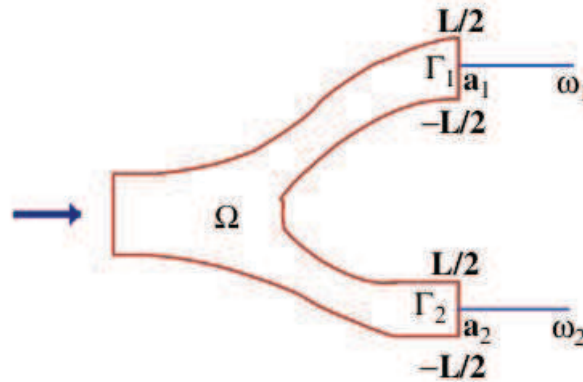
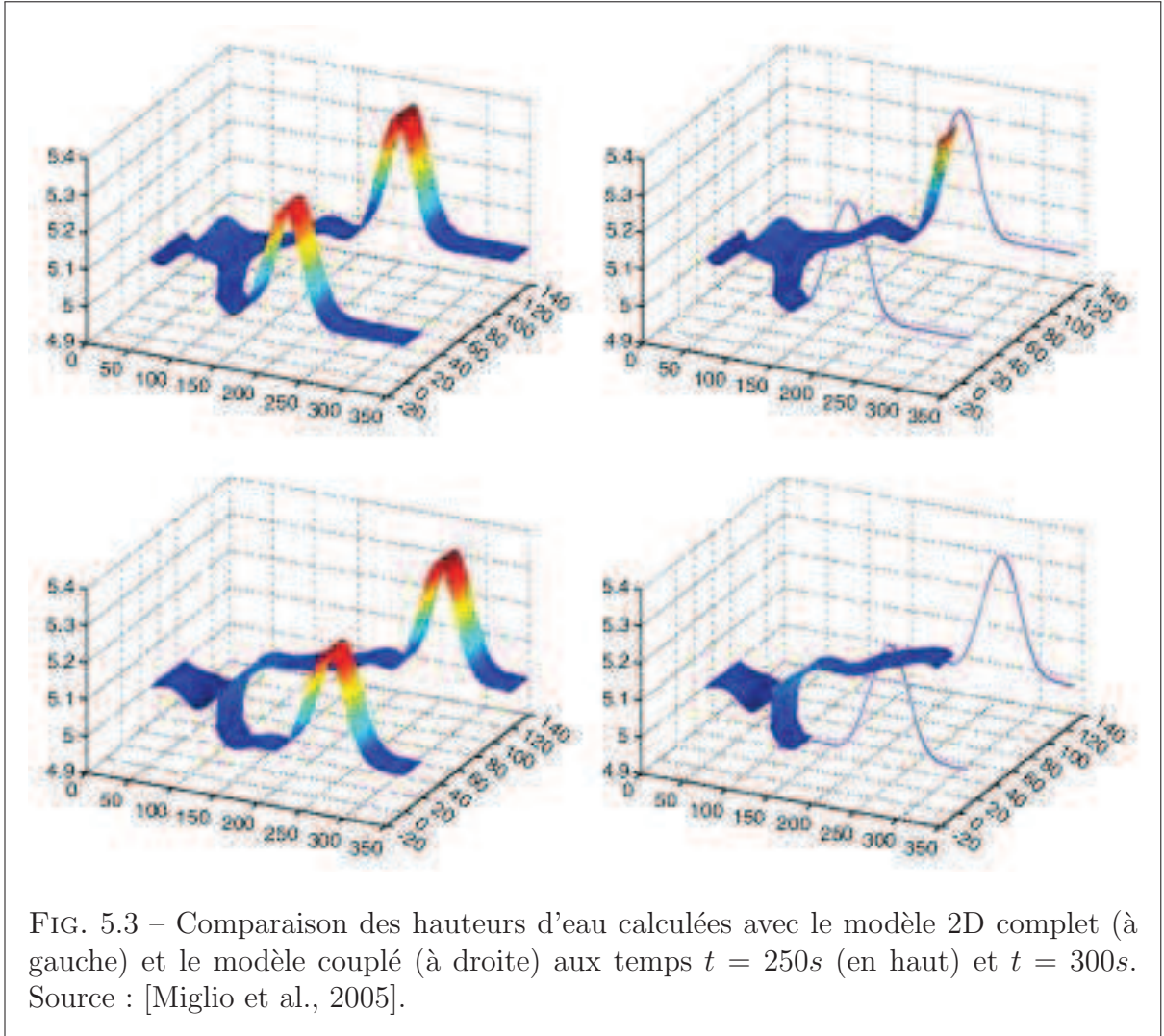


FIG. 5.2 – Le domaine délimité en rouge représente le modèle 2D et les modèles 1D sont représentés en bleu. Source : [Miglio et al., 2005].

Il est montré que le modèle couplé 1D-2D donne des résultats en termes de hauteurs d'eau modélisées très proches de ceux donnés par le modèle 2D complet sur la zone comme on peut le voir sur la figure 5.4. Le forçage à l'amont du modèle 2D est donné par la fonction h donnée en (5.9), ce qui explique cette forme d'onde qui se propage.

Malheureusement le gain en termes de coût de calcul ainsi que l'erreur entre les modèles couplé et 2D complet au niveau de l'interface n'est pas précisé.

$$h(t) = \begin{cases} 5 + 0.5 \sin\left(\frac{2\pi t}{10}\right) & t \leq 5 \\ 5 & t > 5 \end{cases} \quad (5.9)$$



[Miglio et al., 2005] présente ensuite un second cas académique, il s'agit d'un couplage 1D-2D avec un obstacle au milieu du modèle 2D. Dans cette expérience le modèle 1D est en amont du modèle 2D et l'objectif est d'étudier de manière qualitative l'influence la position de l'interface entre les modèles sur les résultats. Comme on peut l'intuiter les résultats du modèle couplé seront d'autant plus proches du modèle 2D complet que l'interface est en amont de l'obstacle (voir figure 5.4).

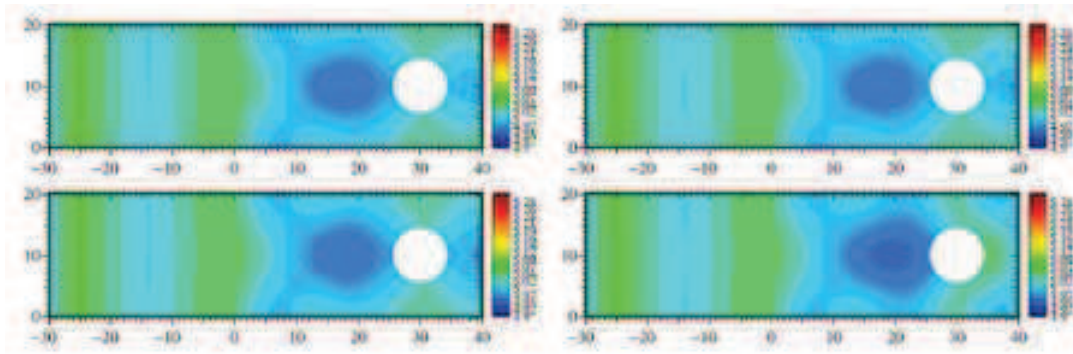


FIG. 5.4 – Hauteurs d'eau calculées avec le modèle 2D avec obstacle. De gauche à droite et de haut en bas : 2D complet, modèle couplé avec l'interface positionnée à l'abscisse $x = 70$, $x = 50$ et $x = 40$. Source : [Miglio et al., 2005].

Malleron *et al.*, 2011

Dans un second temps nous nous intéressons aux travaux présentés dans [Malleron et al., 2011]. Ceux-ci portent sur l'implémentation de l'algorithme itératif présenté en section 5.1.2 avec le coupleur de codes OpenPALM développé par le CERFACS et l'ONERA (voir [Buis et al., 2006]). Dans une première partie les auteurs se livrent à une discussion sur les avantages et les apports du coupleur de codes OpenPALM par rapport à d'autres coupleurs. Ils présentent également les codes de calcul utilisés dans cette étude, il s'agit des codes Mascaret et Telemac-2D présentés au chapitre 2. Dans une seconde partie les auteurs testent l'algorithme sur deux cas réels.

Le premier cas présenté est le cas de la rupture du barrage de Malpasset qui a eu lieu de 2 Décembre 1959. Ce barrage était situé dans le sud de la France sur la rivière Reyran. Nous ne disposons pas d'observations de hauteurs d'eau relatives à cet événement, néanmoins des transformateurs électriques étaient situés à l'aval du barrage et l'on connaît approximativement l'heure à laquelle ils ont cessé de fonctionner. Ces données ont permis de calibrer une reproduction à l'échelle 1/400 de la vallée où a eu lieu la catastrophe et qui a été agrémenté de 14 instruments de mesure. Les données produites avec cette reproduction à l'échelle sont en accord avec les marques laissées par l'eau sur certains bâtiments en aval du barrage. Au niveau du barrage et dans la partie amont de la vallée située à l'aval de celui-ci un modèle 1D a été utilisé. Là où la vallée s'élargit un modèle 2D a été utilisé (voir figure 5.5).

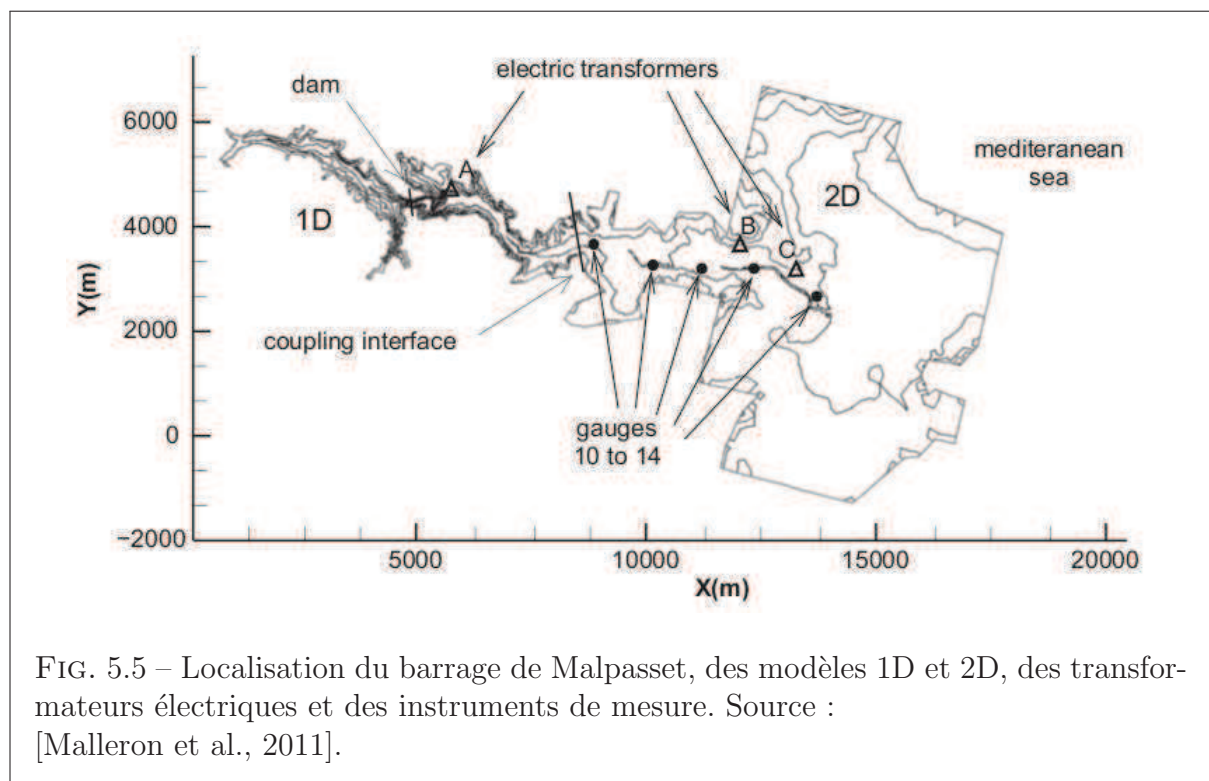
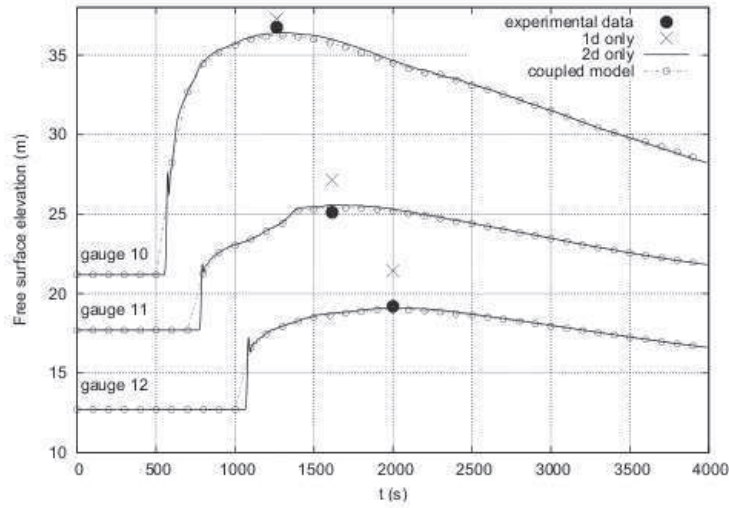
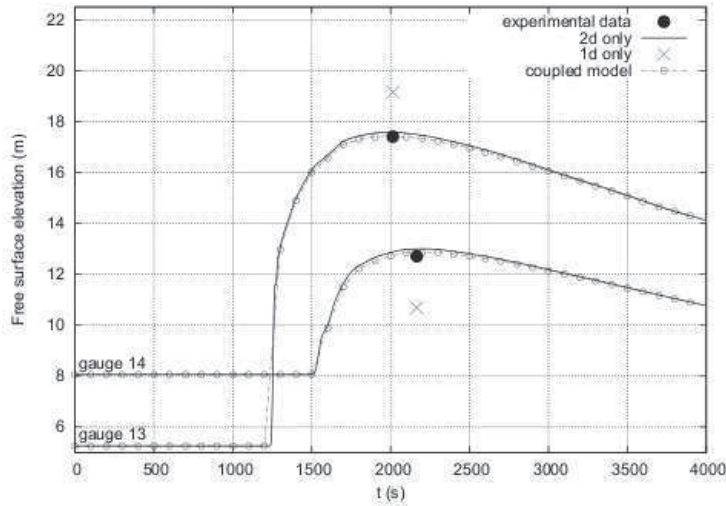


FIG. 5.5 – Localisation du barrage de Malpasset, des modèles 1D et 2D, des transformateurs électriques et des instruments de mesure. Source : [Malleron et al., 2011].

Dans cette expérience ce sont les invariants de Riemann qui ont été échangés à l'interface. Les résultats montrent que le modèle couplé et le modèle 2D complet arrivent bien à reproduire les hauteurs d'eau simulées avec la maquette aux stations 10 à 14. Ce n'est pas le cas du modèle 1D complet dont la différence avec les observations s'accroît avec l'abscisse, voir figure 5.6.



(a)

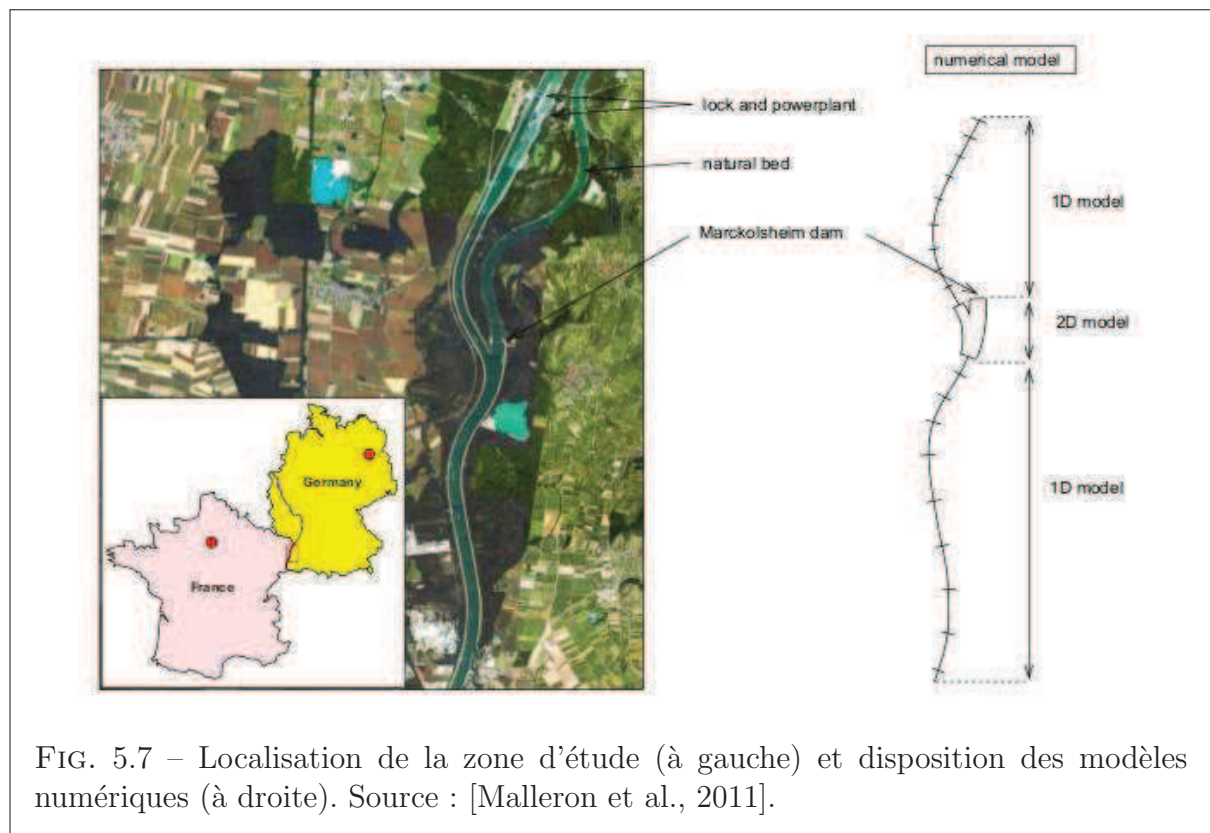


(b)

FIG. 5.6 – Comparaison des données expérimentales et des cotes d'eau calculées par les modèles 1D et 2D complets et le modèle couplé 1D-2D aux stations de mesure (a) 10, 11, 12 et (b) 13 et 14. Source : [Malleron et al., 2011].

Dans cet exemple le temps simulé est de 4000s soit un peu plus d'une heure, le temps de calcul du modèle couplé est équivalent à celui du modèle 2D complet sur la zone (47s contre 40s respectivement), il reste néanmoins nettement inférieur au coût d'un couplage par échange de fichiers (4h40) montrant tout l'intérêt que les méthodes de couplage de modèles peuvent retirer d'un coupleur de codes comme OpenPALM.

Le second cas présenté dans l'étude de (Malleron *et al.*, 2011) porte sur le couplage de modèles au niveau de la diffluence du barrage de Marckolsheim situé sur le Rhin. Cette diffluence marque la séparation entre le lit naturel du Rhin d'un côté et un canal de diversion qui alimente la centrale hydroélectrique de Marckolsheim (voir figure 5.7).



L'étude ne présente pas de données d'observation pour ce cas, par contre l'étude présente les hauteurs d'eau calculées par les modèles couplé et 2D complet au niveau de l'interface à l'amont de la diffluence pendant l'épisode qui a eu lieu du 9 au 11 Août 2007. Le maximum des débits à l'amont du modèle a été de $4513 \text{ m}^3 \cdot \text{s}^{-1}$ mais n'a pas donné lieu à des débordements par dessus les berges. Les résultats obtenus pour la cote de la surface libre par le modèle couplé et le 2D complet sont très proches tout au long de l'épisode hormis une légère différence de quelques centimètres pendant le pic de crue, voir figure 5.8.

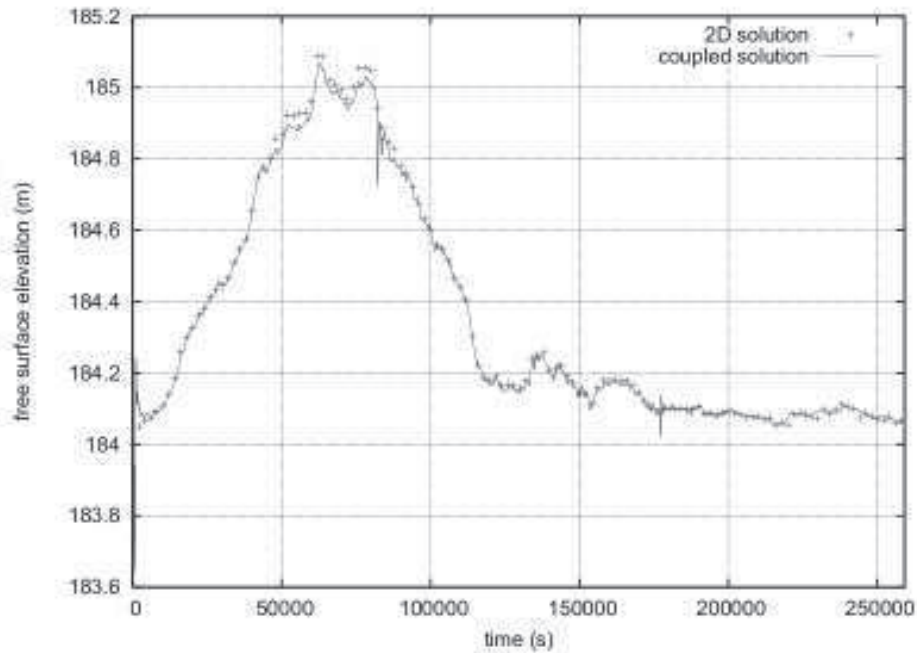


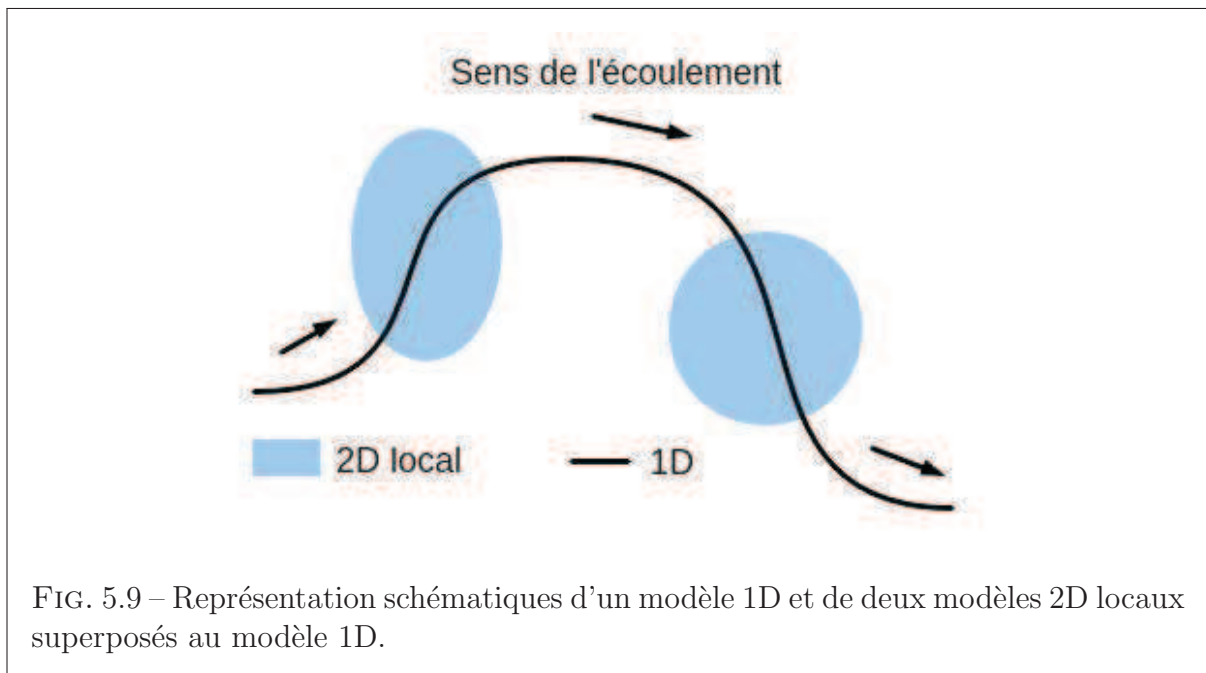
FIG. 5.8 – Séries temporelles de la cote de la surface libre calculée par le modèle 2D complet et par le modèle couplé. Source : [Malleron et al., 2011].

Bien que l'épisode porte sur plusieurs jour les données disponibles dans l'article relativement au temps de calcul pour ce cas porte sur un temps simulé de 20000s (soit environ 5h30). Pour un temps simulé de 20000s le temps de calcul avec le modèle 2D complet est de 11m10s tandis que le temps de calcul pour le modèle couplé est de 4m20s montrant la réduction potentielle du temps de calcul avec un modèle couplé.

5.2 Couplage de modèles superposés

Une autre approche pour le couplage de modèles réside dans le couplage de modèles superposés. Les modèles 1D ont donné la preuve de leur pertinence dans le cadre d'écoulements fluviaux ou torrentiels lorsque l'écoulement reste mono-dimensionnel. Néanmoins en cas de submersion des berges et si l'écoulement devient bi-dimensionnel les modèles 1D ne sont plus pertinents. Pour palier à ce problème des modèles à casiers dits « 1.5 D » ont été développés. Les casiers sont censés représenter les plaines d'inondation adjacentes au lit de la rivière. Ceux-ci peuvent offrir de bons résultats, meilleurs qu'un modèle 1D avec lits composés et comparables à ceux d'un modèle 2D lorsque l'on ne s'intéresse qu'à la cote maximale sans regarder plus précisément les effets dynamiques (voir [Besnard and Goutal, 2011]). D'autre part les modèles à casiers trouvent leurs limitations dans le fait que les échanges entre ceux-ci et le lit de la rivière sont difficiles à calibrer, sont basés sur des lois empiriques et offrent de fait une modélisation qui n'est pas très physique. Pour palier à ce problème on peut envisager de modéliser avec un modèle

2D couvrant à la fois les lit majeur et mineur les zones d'intérêt et coupler celui-ci au modèle 1D sur la zone de recouvrement des deux modèles comme cela est représenté schématiquement sur la figure 5.9. Cette disposition présente l'avantage de ne pas avoir à décomposer un modèle 1D existant calibré et validé en plusieurs sous-modèles comme cela peut être le cas avec le couplage de modèles à interfaces.



Bien que la finalité reste la même que dans le cas du couplage à interface exposé en section 5.1.2, à savoir disposer d'un modèle 2D là où l'écoulement n'est plus mono-dimensionnel, la différence ici réside dans le fait que l'on souhaite modéliser non plus les échanges entre des modèles disposés séquentiellement le long de la rivière mais les échanges entre les lits mineur et majeur de la rivière pour les utiliser comme termes sources pour le modèle 1D. A cette fin plusieurs études menées dans le cadre de méthodes aux volumes finis ont été réalisées ces dernières années, il convient de citer [Finaud-Guyot et al., 2011, Gejadze and Monnier, 2007, Marin and Monnier, 2009, Fernández-Nieto et al., 2010]

Dans ce travail nous ne nous intéresserons pas aux méthodes volumes finies utilisées dans ces travaux. Le lecteur intéressé pourra se reporter à ces articles. Par contre nous nous sommes intéressés au calcul des termes d'échanges entre lit majeur et lit mineur (qui servent comme termes sources au modèle 1D pour ce couplage) et qui est exposé dans [Marin and Monnier, 2009, Fernández-Nieto et al., 2010]. Dans la suite nous exposons le calcul de ces termes sources ainsi que l'algorithme itératif de couplage proposé par [Fernández-Nieto et al., 2010]. D'autre part dans [Gejadze and Monnier, 2007, Marin and Monnier, 2009] les auteurs ont combiné leur méthode de couplage de modèles hydrauliques superposés avec une méthode d'assimilation de données variationnelle sur le domaine 2D. Ces développements sortant du cadre de ce travail de thèse nous ne nous y intéresserons pas ici.

5.2.1 Calcul des termes d'échange entre lit mineur et lit majeur

Dans le cadre du couplage de modèles superposés les termes d'échange entre lit mineur et lit majeur servent comme termes d'apport latéraux pour le modèle 1D. Certaines lois empiriques ont été proposées de manière à estimer les flux à l'interface entre lit mineur et majeur. C'est le cas par exemple de la loi de seuil proposée par [Cunge et al., 1980] qui estime les débits à l'interface entre les lits en fonction de la différence entre les hauteurs d'eau calculées dans chacun des deux lits à proximité de l'interface. Cependant l'usage de cette loi est limité par l'estimation d'un paramètre, ce que l'on voudrait éviter. [Marin and Monnier, 2009] propose un calcul des termes sources pour les équations de conservation de la masse et de conservation de la quantité de mouvement 1D en intégrant les équations de Navier-Stokes 3D sur la section mouillée. Ce sont ces termes sources que nous avons utilisé dans le cadre du couplage des modèles développés sur la zone de Bayonne. Nous exposons en suivant la démonstration de [Marin and Monnier, 2009] le calcul de ces termes dans les sections suivantes.

Pour le calcul des termes sources pour les équations de Saint-Venant on considère une section mouillée représentée sur la figure 5.10-(b). Dans toute cette section on considère que la rivière est orientée selon l'axe des x et que les variations de largeur de celle-ci ainsi que les variations des berges sont faibles comme cela est représenté sur la figure 5.10-(a).

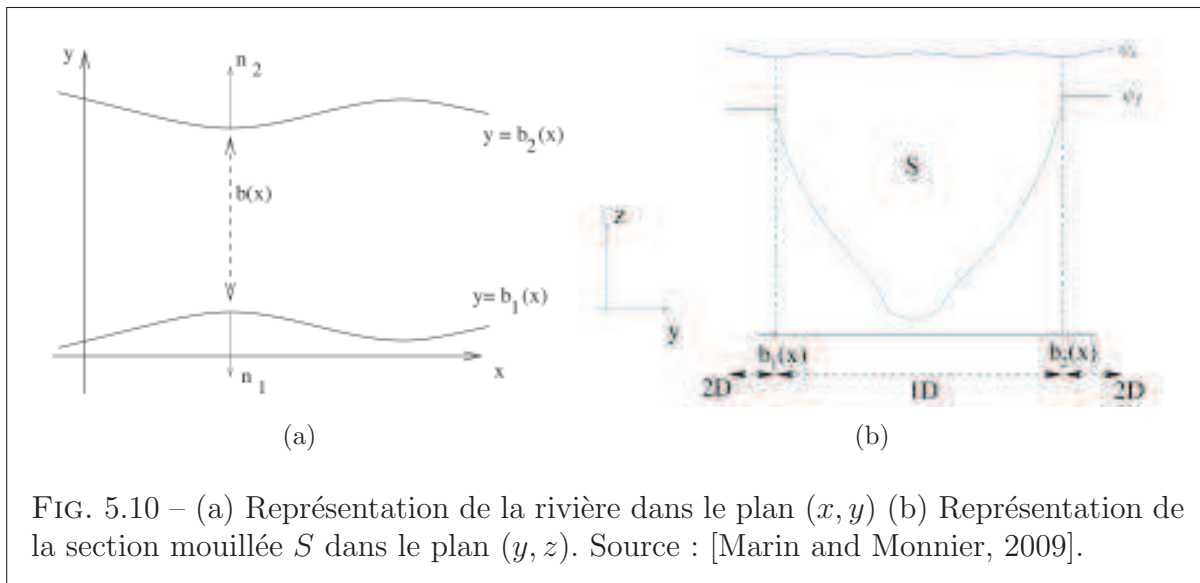


FIG. 5.10 – (a) Représentation de la rivière dans le plan (x, y) (b) Représentation de la section mouillée S dans le plan (y, z) . Source : [Marin and Monnier, 2009].

Terme source pour l'équation de conservation de la masse

Selon le système des équations de Navier-Stokes (2.11) l'équation de conservation de la masse s'écrit : $\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0$. Par intégration de cette équation sur la section mouillée on peut écrire :

$$\underbrace{\int_{b_1}^{b_2} \int_{\psi_f}^{\psi_s} \frac{\partial u}{\partial x} dz dy}_A + \underbrace{\int_{b_1}^{b_2} \int_{\psi_f}^{\psi_s} \frac{\partial v}{\partial y} dz dy}_B + \underbrace{\int_{b_1}^{b_2} \int_{\psi_f}^{\psi_s} \frac{\partial w}{\partial z} dz dy}_C = 0 \quad (5.10)$$

Or en utilisant la formule de Leibnitz et l'hypothèse que la vitesse est nulle sur le fond (voir section 2.1.4) on peut réécrire les termes A , B et C de l'équation (5.10) :

$$\begin{aligned} A &= \int_{b_1}^{b_2} \left(\frac{\partial}{\partial x} \int_{\psi_f}^{\psi_s} u dz - u_s \frac{\partial z_s}{\partial x} \right) dy \\ &= \frac{\partial}{\partial x} \int_{b_1}^{b_2} \int_{\psi_f}^{\psi_s} u dz dy - \left[\int_{\psi_f}^{\psi_s} u dz \right]_{b_2} \frac{\partial b_2}{\partial x} + \left[\int_{\psi_f}^{\psi_s} u dz \right]_{b_1} \frac{\partial b_1}{\partial x} - \underbrace{\int_{b_1}^{b_2} u_s \frac{\partial z_s}{\partial y} dy}_{CL1} \\ B &= \int_{b_1}^{b_2} \left(\frac{\partial}{\partial y} \int_{\psi_f}^{\psi_s} v dz - v_s \frac{\partial z_s}{\partial y} \right) dy \\ &= \left[\int_{\psi_f}^{\psi_s} v dz \right]_{b_2} - \left[\int_{\psi_f}^{\psi_s} v dz \right]_{b_1} - \underbrace{\int_{b_1}^{b_2} v_s \frac{\partial z_s}{\partial y} dy}_{CL2} \\ C &= \underbrace{\int_{b_1}^{b_2} w_s dy}_{CL3} \end{aligned}$$

Or d'après la condition aux limites sur la surface libre (2.13) la somme des termes $CL1$, $CL2$ et $CL3$ est égale à :

$$CL1 + CL2 + CL3 = \int_{b_1}^{b_2} \frac{\partial z_s}{\partial t} dy$$

Or par définition de la section mouillée on a :

$$\int_{b_1}^{b_2} \frac{\partial z_s}{\partial t} dy = \frac{\partial}{\partial t} \int_{b_1}^{b_2} z_s dy = \frac{\partial S}{\partial t} \quad (5.11)$$

D'autre part par définition du débit on a :

$$\frac{\partial}{\partial x} \int_{b_1}^{b_2} \int_{\psi_f}^{\psi_s} u dz dy = \frac{\partial Q}{\partial x} \quad (5.12)$$

Ainsi il vient des expressions de A , B et C et des égalités (5.11) et (5.12) que l'équation (5.10) peut se réécrire :

$$\frac{\partial S}{\partial t} + \frac{\partial Q}{\partial x} = K_1$$

où :

$$K_1 = - \left(\left[\int_{\psi_f}^{\psi_s} v dz \right]_{b_2} - \left[\int_{\psi_f}^{\psi_s} v dz \right]_{b_1} - \left[\int_{\psi_f}^{\psi_s} u dz \right]_{b_2} \frac{\partial b_2}{\partial x} + \left[\int_{\psi_f}^{\psi_s} u dz \right]_{b_1} \frac{\partial b_1}{\partial x} \right)$$

Il apparait que si l'écoulement ne déborde par les berges de la rivière alors le terme source K_1 s'annule, en effet dans ce cas $\psi_f = \psi_s$ pour $y = b_1, b_2$, et on retrouve l'équation de conservation de la masse sans terme source apparaissant dans le système Saint-Venant 1D (2.35).

Si la surface libre s'élève au-dessus du niveau des berges le terme K_1 n'est pas nul. Dans ce cas on écrit :

$$K_1 = -(\vec{q}_1 \cdot \delta_1 \vec{n}_1 + \vec{q}_2 \cdot \delta_2 \vec{n}_2) \quad (5.13)$$

où $\vec{q}_k$ est le vecteur des débits linéiques à l'interface entre lit mineur et lit majeur, soit :

$$\vec{q}_k = \left(\left[\int_{\psi_f}^{\psi_s} u dz \right]_{b_k}, \left[\int_{\psi_f}^{\psi_s} v dz \right]_{b_k} \right) \quad k = 1, 2$$

Si les variations de largeur de la rivière sont faibles alors $\delta_k \approx 1$, $k = 1, 2$ et dans ce cas l'expression du terme K_1 donnée en (5.13) est approché par l'opposé de la somme des débits linéiques normaux aux berges de la rivière, soit :

$$K_1 \approx -(\vec{q}_1 \cdot \vec{n}_1 + \vec{q}_2 \cdot \vec{n}_2) \quad (5.14)$$

Dans la suite on note : $q_k = \vec{q}_k \cdot \vec{n}_k$, $k = 1, 2$. Si on suppose que les variations de la largeur de la rivière sont faibles alors l'équation de conservation de la masse dans le cas où la rivière déborde s'écrit :

$$\boxed{\frac{\partial S}{\partial t} + \frac{\partial Q}{\partial x} = -(q_1 + q_2)} \quad (5.15)$$

où q_1 et q_2 sont calculés avec un modèle 2D. Et donc le terme source pour l'équation de conservation de la masse est $-(q_1 + q_2)$.

Terme source pour l'équation de conservation de la quantité de mouvement

On considère la première des équations de conservation de la quantité de mouvement du système d'équations de Navier-Stokes (voir (2.11)) :

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} + w \frac{\partial u}{\partial z} + \frac{1}{\rho} \frac{\partial p}{\partial x} = \nu \Delta u \quad \text{que l'on réécrit :}$$

$$\frac{\partial u}{\partial t} + \frac{\partial u^2}{\partial x} + \frac{\partial uv}{\partial y} + \frac{\partial uw}{\partial z} + \frac{1}{\rho} \frac{\partial p}{\partial x} = \nu \Delta u \quad (5.16)$$

Par intégration sur la surface mouillée S et en supposant que les variations de largeur du canal sont faibles [Marin and Monnier, 2009] montre que l'on a l'équation suivante :

$$\frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(\beta \frac{Q^2}{S} \right) + gS \frac{\partial (H + \mathcal{Z}_f)}{\partial x} = K_2 - gSS_f \quad (5.17)$$

où H et $\mathcal{Z}_b$ sont respectivement la hauteur d'eau et la topographie moyennes sur la section en travers et où S_f est le terme de frottement modélisé par la loi de Manning-Strickler (voir (2.31)), β est le coefficient de Boussinesq défini par :

$$\beta = \frac{S}{Q^2} \int_{b_1}^{b_2} \int_{\psi_f}^{\psi_s} u^2 dz dy$$

Le terme K_2 est donné par l'expression :

$$K_2 = - \left[\int_{\psi_f}^{\psi_s} uv dz \right]_{b_2} + \left[\int_{\psi_f}^{\psi_s} uv dz \right]_{b_1} - \left[\int_{\psi_f}^{\psi_s} u^2 dz \right]_{b_1} \frac{\partial b_1}{\partial x} + \left[\int_{\psi_f}^{\psi_s} u^2 dz \right]_{b_2} \frac{\partial b_2}{\partial x}$$

Il apparait que si il n'y a pas submersion des berges $\psi_f = \psi_s$ le terme K_2 s'annule et l'équation (5.17) se réduit à l'équation de conservation de la quantité de mouvement de Saint-Venant 1D sans terme source, voir (2.35).

[Marin and Monnier, 2009] introduit alors les deux coefficients :

$$\gamma_k = \frac{\left[\int_{\psi_f}^{\psi_s} uv dz \right]_{b_k}}{\left[\int_{\psi_f}^{\psi_s} u dz \right]_{b_k} \left[\int_{\psi_f}^{\psi_s} v dz \right]_{b_k}}, \quad \alpha_k = \frac{\left[\int_{\psi_f}^{\psi_s} u^2 dz \right]_{b_k}}{\left(\left[\int_{\psi_f}^{\psi_s} u dz \right]_{b_k} \right)^2}, \quad k = 1, 2$$

Dans la suite on suppose que (u, v) ne dépend pas de z sur les frontières latérales b_k (ce qui est une hypothèse raisonnable dans le cadre d'un écoulement avec un petit rapport d'aspect), dans ce cas $\gamma_k = \alpha_k = 1/h_k, k = 1, 2$. On peut écrire :

$$\begin{aligned} & \left[\int_{\psi_f}^{\psi_s} uv dz \right]_{b_1} - \left[\int_{\psi_f}^{\psi_s} u^2 dz \right]_{b_1} \frac{\partial b_1}{\partial x} \\ &= \left[\int_{\psi_f}^{\psi_s} u dz \right]_{b_1} \left(\gamma_1 \left[\int_{\psi_f}^{\psi_s} v dz \right]_{b_1} - \alpha_1 \left[\int_{\psi_f}^{\psi_s} u dz \right]_{b_1} \frac{\partial b_1}{\partial x} \right) \\ &= -u|_1 \vec{q}_1 \cdot \vec{n}_1 = -u|_1 q_1 \end{aligned}$$

De la même manière on peut montrer :

$$- \left[\int_{\psi_f}^{\psi_s} uv dz \right]_{b_2} + \left[\int_{\psi_f}^{\psi_s} u^2 dz \right]_{b_2} \frac{\partial b_2}{\partial x} = -u|_2 q_2$$

Si l'on suppose également que u est constant dans la section en travers alors $\beta = 1$ et si l'on suppose que les variations de largeur du canal sont petites on obtient l'équation suivante de conservation de la quantité de mouvement :

$$\frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(\frac{Q^2}{S} \right) + gS \frac{\partial (H + \mathcal{Z}_f)}{\partial x} = -(u|_1 q_1 + u|_2 q_2) - gSS_f \quad (5.18)$$

Le terme source pour l'équation de conservation de la quantité de mouvement est : $-(u|_1 q_1 + u|_2 q_2)$.

5.2.2 Equations de Saint-Venant 1D conservatives avec terme source

En considérant les équations (5.15) et (5.18) et la mise sous forme conservative des équations de Saint-Venant 1D donnée en (2.35) on écrit le système d'équations de Saint-Venant 1D sous forme conservative avec terme source :

$$\begin{cases} \frac{\partial S}{\partial t} + \frac{\partial Q}{\partial x} = -(q_1 + q_2) \\ \frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(\frac{Q^2}{S} + P \right) = g \int_0^h \left[\frac{\partial \mathcal{S}(x, z)}{\partial x} \right]_z dz - gS \frac{\partial Z_f}{\partial x} - gSS_f - (u|_1 q_1 + u|_2 q_2) \end{cases} \quad (5.19)$$

Le système (5.19) peut être mis sous forme d'un système conservatif (voir (2.36)) avec pour terme source Ψ :

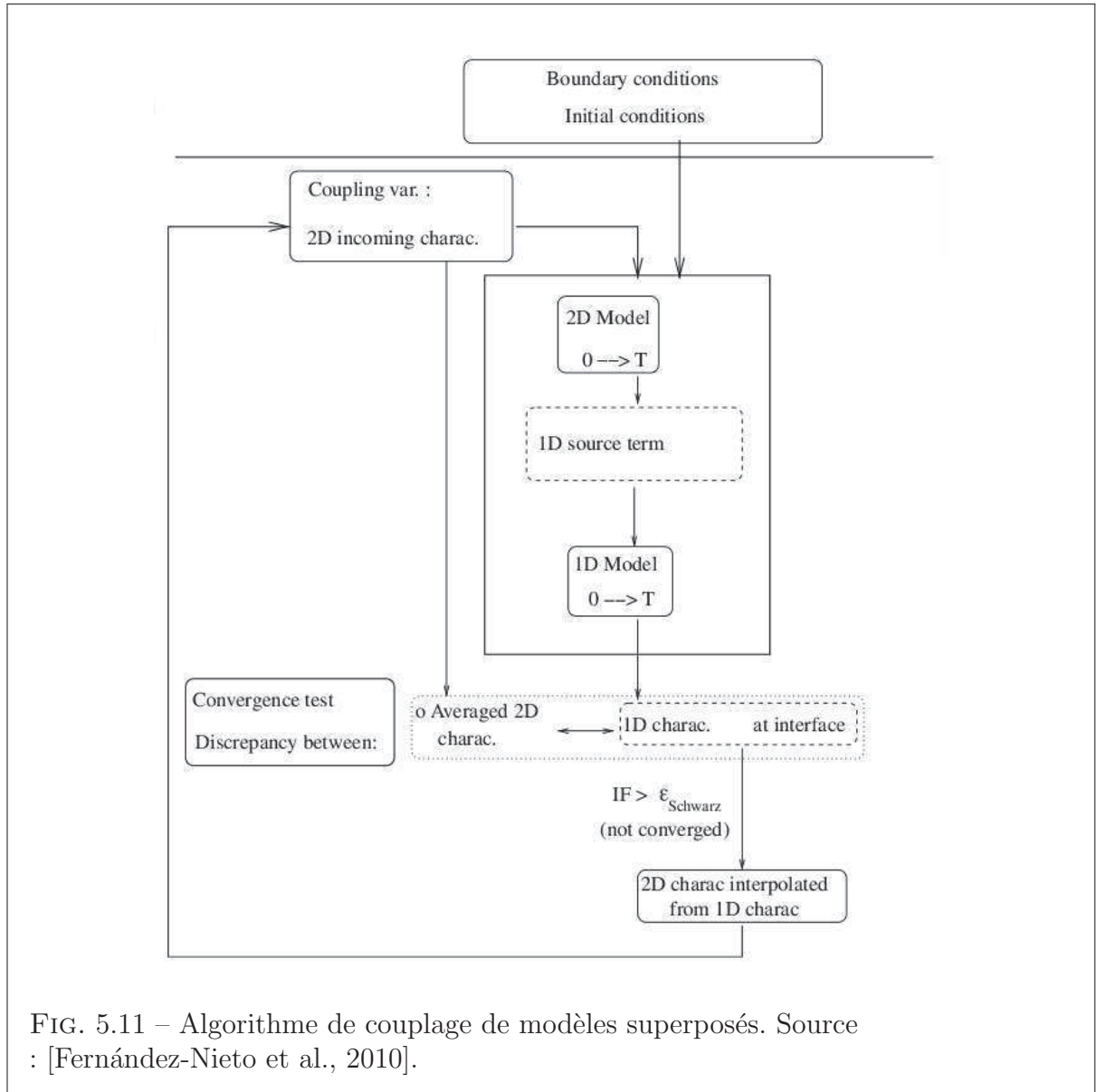
$$\frac{\partial W}{\partial t} + \text{div} [F(W)] = B(W) + \Psi \quad (5.20)$$

où $\Psi = -(q_1 + q_2, u|_1 q_1 + u|_2 q_2)^T$. Le lecteur se reportera à la section 2.4.1 pour le détail de W , F et B . C'est le terme Ψ qui va servir de terme source pour le modèle 1D dans le cadre du couplage de modèles superposés.

5.2.3 Algorithme de couplage

[Fernández-Nieto et al., 2010] propose un algorithme de couplage assez proche dans le principe à celui exposé dans la section 5.1.2. Il y a néanmoins deux grandes différences entre ces deux méthodes. D'une part il y a les termes échangés entre les deux modèles, dans [Fernández-Nieto et al., 2010] le modèle 1D force le modèle 2D avec les variables caractéristiques et le modèle 2D calcule les termes sources pour le modèle 1D dont le calcul est donné en section 5.2.1.

Le modèle 2D est donc forcé avec des variables caractéristiques calculées *a priori* ou extrapolées à partir du modèle 1D selon qu'il s'agisse de la première itération de couplage ou pas. Sur l'intervalle de temps $[0; T]$ le modèle 2D calcule les termes sources à transmettre au modèle 1D. On effectue un calcul avec le modèle 1D en utilisant ces termes sources et à l'issue du calcul on compare les variables caractéristiques moyennées avec lesquelles on a forcé le modèle 2D à celles calculées par le modèle 1D au temps T . Si la différence en valeur absolue est inférieure au critère de convergence fixé *a priori* alors on arrête le calcul ici. Sinon on extrapole de nouvelles variables caractéristiques 2D à partir de celles calculées avec le modèle 1D puis on répète l'opération avec ces nouveaux forçages pour le modèle 2D. Le principe de cet algorithme est très bien synthétisé dans la figure 5.11 tirée de [Fernández-Nieto et al., 2010].



5.2.4 Variations autour d'une application numérique

Comme on l'a déjà évoqué [Gejadze and Monnier, 2007, Marin and Monnier, 2009] proposent des exemples de couplage de modèles superposés mais ceux-ci sont combinés avec des méthodes d'assimilation variationnelle de données du domaine 2D. Comme ces exemple sortent du cadre de cette thèse nous ne traiterons pas de ceux-ci dans ce travail. Nous nous intéresserons néanmoins à l'application donnée par [Fernández-Nieto et al., 2010].

[Fernández-Nieto et al., 2010] donne un exemple d'application sur un cas académique. Le domaine 1D est un canal rectiligne de 200m de long et de 2m de large tandis que le domaine 2D est un domaine circulaire de 100m de diamètre avec un replat légèrement surélevé au centre, la bathymétrie est donnée par l'expression (5.21). Une illustration de la bathymétrie utilisée est donnée en figure 5.12.

$$Z_b(x, y) = \begin{cases} 0.5 & \text{si } 80 \leq x \leq 120 \\ 0 & \text{sinon} \end{cases} \quad (5.21)$$

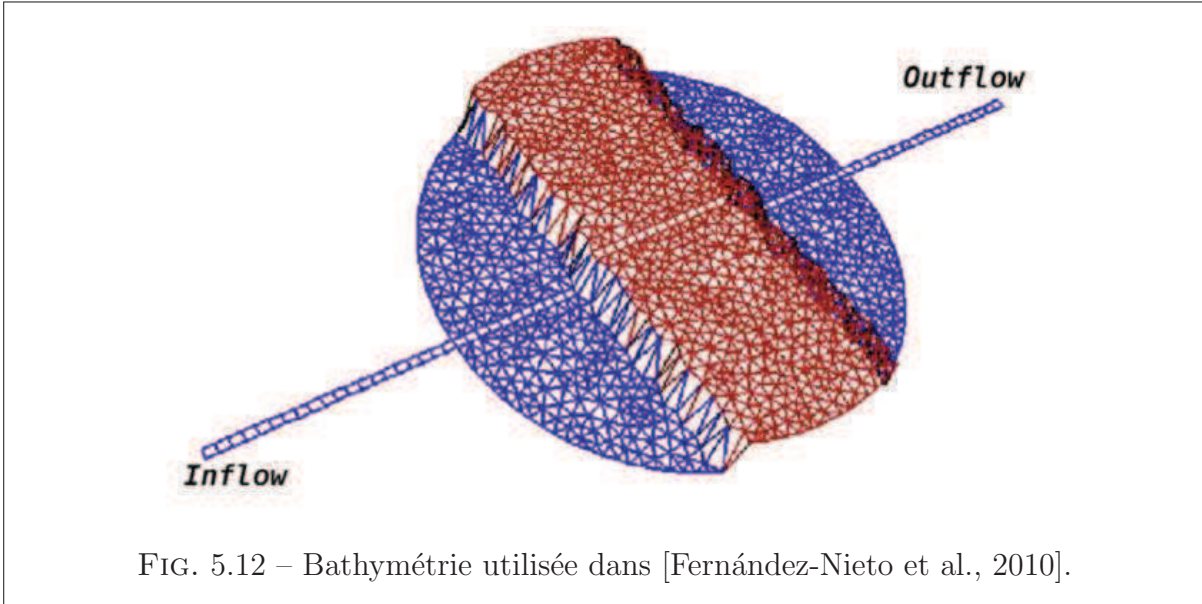
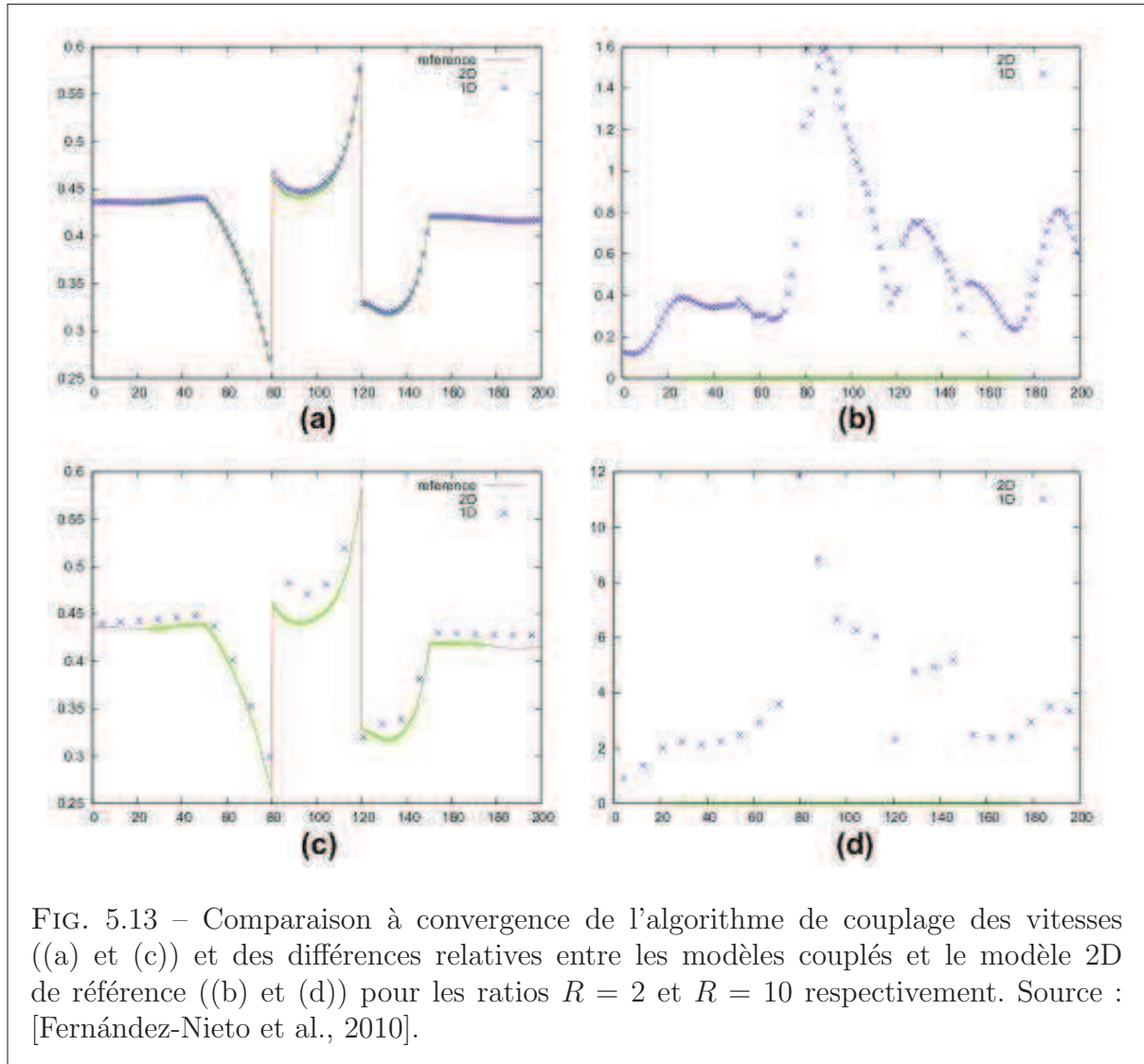


FIG. 5.12 – Bathymétrie utilisée dans [Fernández-Nieto et al., 2010].

Dans un premier temps les auteurs valident dans le cadre de leur expérimentation numérique le terme source dont on a exposé le calcul en section 5.2.1 en construisant tout d'abord une solution stationnaire avec le modèle 2D complet dont ils extraient une condition initiale pour les modèles 1D et 2D. Il faut préciser que dans ce cas les maillages des deux modèles sont parfaitement superposés. Après une itération de couplage les résultats en termes de vitesse des modèles 1D et 2D couplés ainsi que ceux du modèle 2D complet sont identiques.

Ensuite les auteurs se livrent à la même expérience de couplage mais cette fois-ci les maillages des modèles 1D et 2D sur la zone de recouvrement ne sont pas identiques, le maillage 1D étant moins fin que le maillage 2D. Là encore les conditions initiales pour les modèles 1D et 2D couplés sont calculées à partir de la solution calculée avec le modèle 2D complet. Les résultats en termes de vitesse sont moins bons que précédemment du fait d'une erreur sur les termes sources pour le modèle 1D à cause justement du fait que les maillages ne sont pas superposés. Tandis que les pas de temps des modèles 1D et 2D sont identiques les auteurs utilisent des mailles de tailles différentes pour les deux modèles et illustrent les résultats obtenus pour deux ratios R de taille de maille différents : $R = \frac{\Delta x_{1D}}{\Delta x_{2D}} = 2$ et $R = \frac{\Delta x_{1D}}{\Delta x_{2D}} = 10$. Les résultats de ces expériences sont reportés en figure 5.13. Et c'est assez naturellement que les auteurs montrent que les résultats en termes de différence de vitesses entre le modèle 1D couplé et le modèle 2D complet sont moins bons lorsque le rapport R est plus grand avec une différence relative pouvant s'élever jusqu'à 12%. On notera que la différence relative entre le modèle 2D couplé et le modèle 2D référence est nulle dans les deux cas.



Ces expériences numériques sont intéressantes car elles montrent la difficulté à coupler des modèles de dimensions différentes lorsque ceux-ci ont des maillages différents, ce qui peut être le cas en pratique lorsqu’un modèle 1D est développé sur un cours d’eau et qu’un modèle 2D est ensuite développé pour modéliser plus finement l’écoulement dans des zones d’intérêt. Dans cet ordre d’idée les auteurs se sont également livrés à une autre expérience dans laquelle ils développent un modèle 2D à partir d’un maillage 1D de telle sorte que dans la zone de recouvrement les maillages sont superposés. Cela implique que le maillage du modèle 2D ainsi développé est nettement moins fin dans la zone de recouvrement que le modèle 2D de référence. Comme on peut s’y attendre les résultats en termes de vitesse du modèle 2D couplé dans cette configuration sont nettement moins bons que ceux du modèle 2D de référence du fait que le maillage de ce dernier est beaucoup plus fin dans la zone de recouvrement.

Troisième partie

Résultats

Chapitre 6

Etat de l'art de l'assimilation de données pour l'hydraulique fluviale

Avant de rentrer dans le cœur des résultats obtenus dans cette thèse nous devons présenter un état de l'art de l'assimilation de données pour l'hydraulique fluviale afin de mieux situer le contexte global dans lequel ils s'inscrivent.

L'assimilation de données pour l'hydraulique fluviale a fait l'objet d'importants travaux ces dernières années notamment avec l'utilisation de méthodes de type filtrage séquentielle principalement mais aussi avec des travaux utilisant des méthodes variationnelles.

On peut regrouper les différents types d'incertitudes dans les modèles hydrauliques en trois grandes catégories (voir [Refsgaard, 1997, Neal et al., 2007]) : incertitudes sur les forçages aussi bien aux limites du domaine (avec des incertitudes liées par exemple aux courbes de tarage) que sur les apports latéraux (cas des affluents non jaugés et/ou non pris en compte dans la modélisation). On distingue également les incertitudes sur la modélisation physique au sens large : coefficients de frottement, bathymétrie et topographie. Et on distingue enfin les incertitudes sur l'état du modèle. Ainsi pour réduire les différents sources d'incertitudes dans les modèles on distingue différentes approches pour l'assimilation de données : correction de la paramétrisation, correction des forçages et correction de l'état hydraulique.

Dans un premier temps nous nous intéressons à **la correction des forçages**. Ceux-ci peuvent en effet être entachés d'erreur si l'on force par exemple un modèle avec des débits calculés à partir de hauteurs d'eau avec une courbe de tarage. Ou lorsque plusieurs modèles s'enchainent par exemple un modèle pluie-débit et un modèle hydrodynamique alors les forçages de ce dernier sont entachés des erreurs héritées du modèle pluie-débit. Parmi les méthodes corrigeant les forçages on peut citer [Jean-Baptiste et al., 2011, Ricci et al., 2011].

[Jean-Baptiste et al., 2011] dans une étude menée sur une portion du Rhône inclut les forçages dans le vecteur d'état (voir figure 6.1), en plus des hauteurs d'eau et des débits, pour les corriger avec l'assimilation de données en utilisant un filtre de Kalman et un filtre particulaire. Dans les deux cas les résultats sont bons, le filtre de Kalman se distinguant par son coût de calcul relativement faible tandis que le filtre particulaire est en mesure de prendre en compte les non-linéarités du modèle.

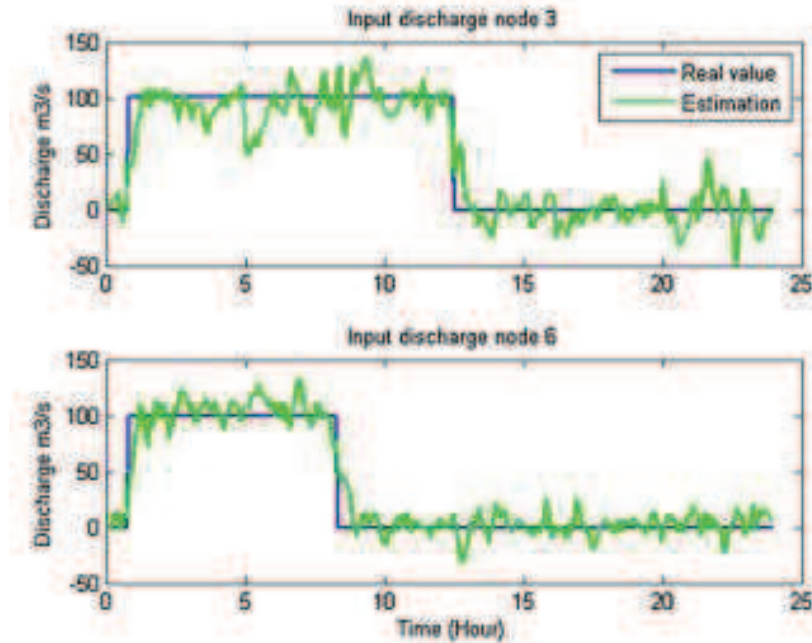


FIG. 6.1 – Estimation *a posteriori* de forçages latéraux dans [Jean-Baptiste et al., 2011].

D'un autre côté [Ricci et al., 2011] utilise un filtre de Kalman étendu pour corriger les forçages amont de deux modèles hydrauliques situés en France combiné avec une procédure de correction de l'état hydraulique utilisant un filtre de Kalman où la matrice des covariances d'erreur d'ébauche est invariante au cours du temps. Cette procédure en deux temps permet d'améliorer les hauteurs d'eau prévues le long du cours d'eau à moyenne échéance (6h-12h) (voir figure 6.2).

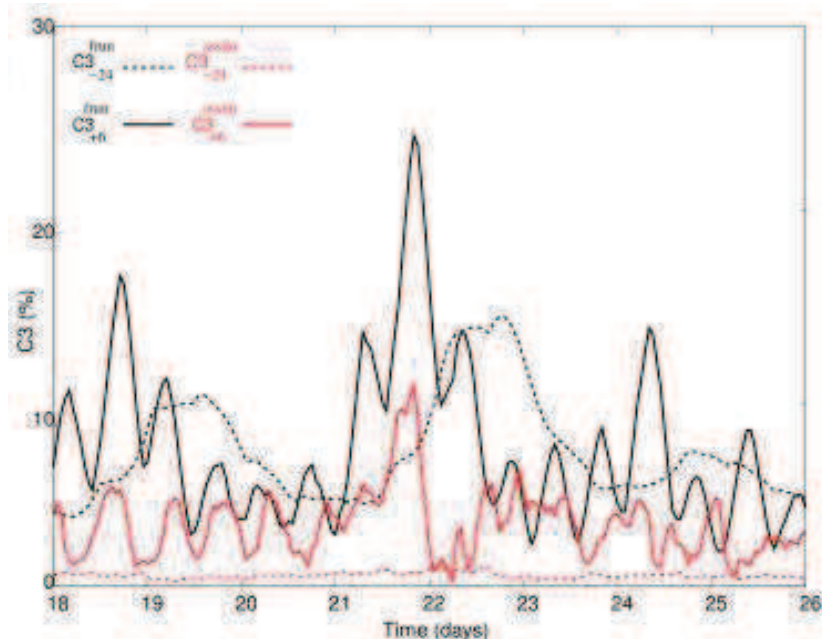


FIG. 6.2 – Pourcentage d’amélioration de l’erreur relative moyenne en valeur absolue sur une période de 26 jours avec la méthode de correction des forçages amonts présentée dans [Ricci et al., 2011].

[Matgen et al., 2010, Giustarini et al., 2011] dans le cadre du couplage de modèles hydrologique/hydraulique corrigent les forçages amont du modèle hydraulique calculés par le modèle hydrologique en ajoutant un terme d’erreur relative sur les débits. Celui-ci étant calculé en fonction des débits observés et calculés précédemment. On peut également citer l’étude de [Lai and Monnier, 2009] qui assimile avec un algorithme 4D-Var des observations satellitaires de hauteurs d’eau pour corriger (entre autres) les apports latéraux dans un modèle 2D.

La correction de l’état hydraulique est aussi une approche très importante en assimilation de données pour l’hydraulique fluviale. La correction de l’état hydraulique consiste en la correction des hauteurs d’eau et/ou des débits en fonction de données d’observations *in situ* ou de données satellitaires. Ces erreurs sur la modélisation de la ligne d’eau peuvent être la résultante de plusieurs facteurs comme des paramètres et/ou des forçages entachés d’erreur ou encore la résultante des erreurs liées à la modélisation (discretisation, description du réseau hydraulique...). Dans ce domaine les méthodes vont des plus simples basées sur l’utilisation d’un filtre de Kalman invariant [Madsen and Skotner, 2005, Ricci et al., 2011] à des méthodes plus élaborées du type filtre particulière [Matgen et al., 2010] ou filtre de Kalman d’ensemble [Neal et al., 2007]. Dans le cas de [Madsen and Skotner, 2005, Ricci et al., 2011] un filtre de Kalman invariant utilisant une fonction de covariance aux points d’observation stationnaire est appliqué. Le choix de cette méthode est justifié par la recherche d’un coût de calcul minimal en vue de l’application de ces méthodes dans un contexte opérationnel. Dans [Madsen and Skotner, 2005] cette méthode est combinée à un système de prévision

de l'erreur qui permet d'améliorer les résultats à plus long terme en mode prévision (voir figure 6.3). [Jean-Baptiste et al., 2011] utilise un filtre de Kalman pour corriger l'état hydraulique tandis que [Shiiba et al., 2000] utilise un filtre de Kalman de rang réduit pour approximer la matrice des covariances d'erreur d'ébauche et réduire le temps de calcul. Quant à [Neal et al., 2007], ils appliquent un filtre de Kalman d'ensemble sur la rivière Nith située en Ecosse et montrent que l'EnKF permet d'augmenter la précision des prévisions de l'ordre de 50% à 70 %. On peut citer également l'étude de [Gejadze and Monnier, 2007] dans laquelle les auteurs utilisent un algorithme d'assimilation de données variationnel pour corriger l'état hydraulique d'un modèle 2D. Cette procédure d'assimilation est combinée avec une méthode de couplage de modèles superposés 1D-2D (voir section 5.2).

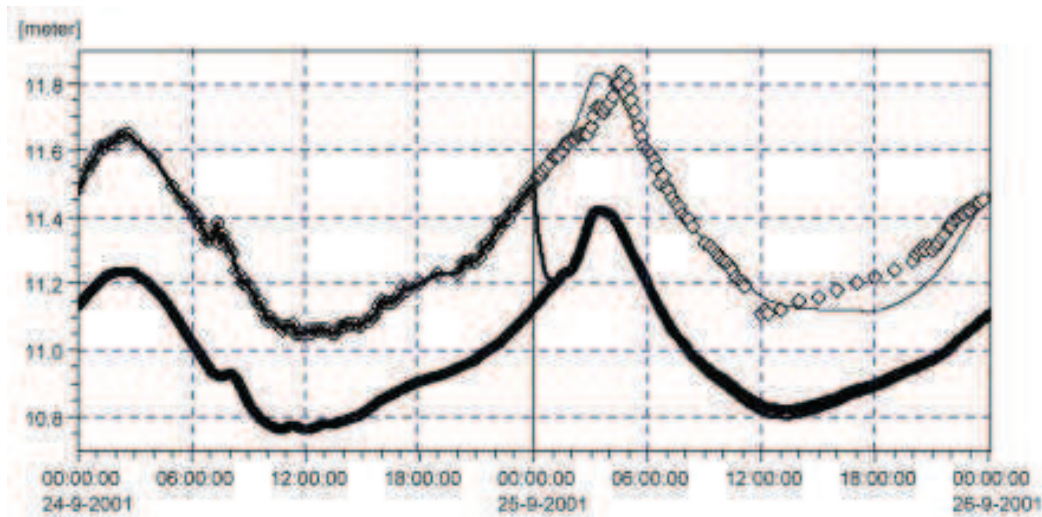


FIG. 6.3 – Hauteurs d'eau calculées dans [Madsen and Skotner, 2005] avec : combinaison de la correction de l'état hydraulique et d'un modèle de prévision de l'erreur (ligne fine continue), correction de l'état hydraulique seule (ligne continue moyenne) et absence de correction (ligne continue épaisse). Les observations sont représentées par les carrés inclinés.

A noter que ces études sont basées sur l'assimilation de données *in situ*, néanmoins de très nombreuses études explorent l'assimilation de données satellitaires pour exploiter les résultats des missions existantes (ENVISAT par exemple) ou dans l'optique de la mission SWOT (Surface Water and Ocean Topography) et dont le lancement devrait avoir lieu en 2019. Ainsi beaucoup de travaux basés sur l'assimilation de données synthétiques SAR pour Synthetic Aperture Radar (qui seront le type de données produites par la mission SWOT) ont été publiés ces dernières années dans le cadre de la correction de l'état hydraulique. [Biancamaria et al., 2011] applique un Kalman smoother d'ensemble local pour corriger les hauteurs d'eau dans le cadre d'un couplage de modèles hydraulique/hydrologique

sur la partie aval de la rivière Ob située en Sibérie occidentale et qui est non instrumentée. [Biancamaria et al., 2011] montre une amélioration de la RMSE par rapport au modèle sans assimilation. Dans le même ordre d’idée [Andreadis et al., 2007] a appliqué l’EnKF sur la rivière Ohio située aux Etats-Unis et montre également une amélioration des résultats par rapport au modèle sans assimilation (voir figure 6.4). [Andreadis et al., 2007] montre également que la réduction de RMSE dépend de la fréquence d’observation et qu’elle sera d’autant plus importante que la fréquence d’observation est grande. [Giustarini et al., 2011, Matgen et al., 2010] ont appliqué chacun un filtre à particules sur un modèle couplé hydraulique/hydrologique combiné avec une méthode de correction des débits calculés par le modèle hydrologique et servant de forçages pour le modèle hydraulique. Ces études ont été réalisées sur la rivière Alzette située dans le Grand Duché du Luxembourg. [Giustarini et al., 2011, Matgen et al., 2010] ont montré que le filtre particulaire réduisait l’incertitude sur les hauteurs d’eau modélisées qui se traduit par une réduction de la RMSE en comparaison du modèle seul. Enfin [Neal et al., 2007] a appliqué un EnKF sur la rivière Alzette en utilisant un modèle couplé hydrologique/hydraulique avec un modèle digital de terrain. Là encore il a été mis en évidence une réduction des incertitudes sur les hauteurs d’eau et les débits modélisés. Malgré ces très bons résultats il faut souligner le fait que les données satellitaires ont une résolution minimale en-dessous de laquelle elles ne sont d’aucun intérêt pour les cours d’eau plus modestes. Par exemple dans le cas de la mission SWOT la résolution spatiale est en règle générale de 50m. D’autre part les satellites ne passent au-dessus d’une zone que de manière intermittente, il peut donc se passer plusieurs jours entre deux passages successifs. L’apport des données satellitaires permettra donc de corriger l’état hydraulique modélisé sur une partie du cours d’eau et sera donc d’intérêt pour les cours d’eau longs. Ainsi les données satellitaires ont un grand intérêt pour surveiller les cours d’eau non instrumentés ayant des dimensions suffisamment importantes pour être pris en compte et ne sauraient être mis en concurrence avec les données *in situ* qui seront produites plus régulièrement.

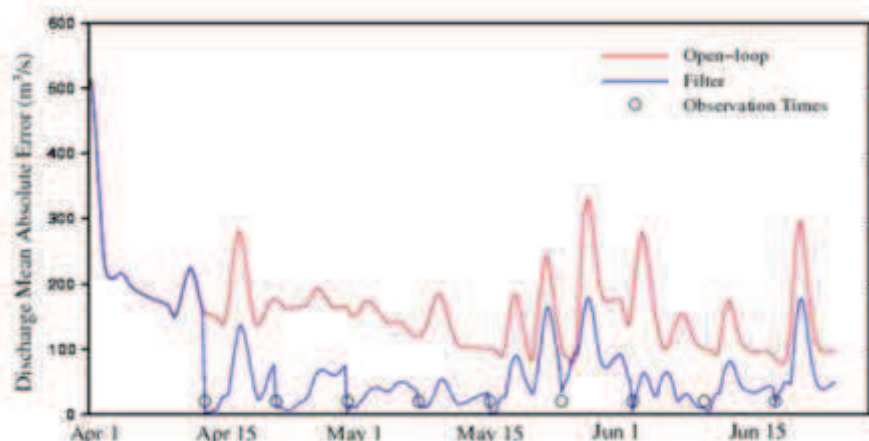


FIG. 6.4 – Erreur moyenne en valeur absolue sur les débits sans et avec la méthode d’assimilation utilisée dans [Andreadis et al., 2007].

De manière connexe à la correction de l'état hydraulique on peut évoquer les méthodes de correction et/ou de prévision de l'erreur ; en général il s'agit de modèles auto-régressifs d'ordre 1 ou 2 basés sur la différence entre les prévisions passées et les états qui ont été observés de manière à corriger les prévisions actuelles. Cette méthode a par exemple été appliquée avec succès par [Madsen and Skotner, 2005] dans le cadre d'un modèle opérationnel de prévision des crues développé sur la région de Metro Manila, aux Philippines (voir figure 6.3). Néanmoins pour que cette démarche soit valide ces méthodes supposent que le comportement de l'erreur varie peu ce qui n'est pas nécessairement le cas. Une méthode assez proche a été appliquée par [Neal et al., 2007] pour corriger les erreurs sur le forçage aval de leur modèle.

On distingue dans une dernière partie les approches pour **corriger les incertitudes liées à la modélisation physique du site**. On peut évoquer tout d'abord la correction de paramètres, par exemple dans le cas d'une fonction de transfert permettant de prévoir les débits à l'aval d'un cours d'eau en fonction des hauteurs d'eau observées à l'amont. [Lees et al., 1994] en donne un exemple sur la rivière Nith à Dumfries en Ecosse. [Pedinotti et al., 2014] réalise une étude prospective sur l'assimilation de données SWOT synthétiques pour corriger les coefficients de frottement de modèles hydrologiques sur le bassin du Niger.

Il faut également citer les travaux portant sur l'assimilation de données en hydraulique pour **la correction de la bathymétrie**. On peut citer les travaux de [Durand et al., 2008, Yoon et al., 2012]. Dans ces deux études les auteurs se sont intéressés à l'assimilation de données altimétriques synthétiques pour la correction de la bathymétrie d'un modèle hydrodynamique sur les fleuves Amazone et Ohio respectivement. Ils ont effectué pour cela des expériences jumelles et ont montré la capacité de l'assimilation de données à estimer des bathymétries synthétiques (voir figure 6.5). Cela résulte en une amélioration des hauteurs d'eau et des débits simulés. Ces travaux sont intéressants car ils laissent entrevoir la possibilité d'estimer la bathymétrie de cours d'eau pour lesquelles on a peu de données bathymétriques ou pour des cours d'eau dont le lit est susceptible d'avoir été modifié après un épisode de fortes crues ou encore des cours d'eau sous influence maritime dont le lit peut être modifié lors de grandes marées.

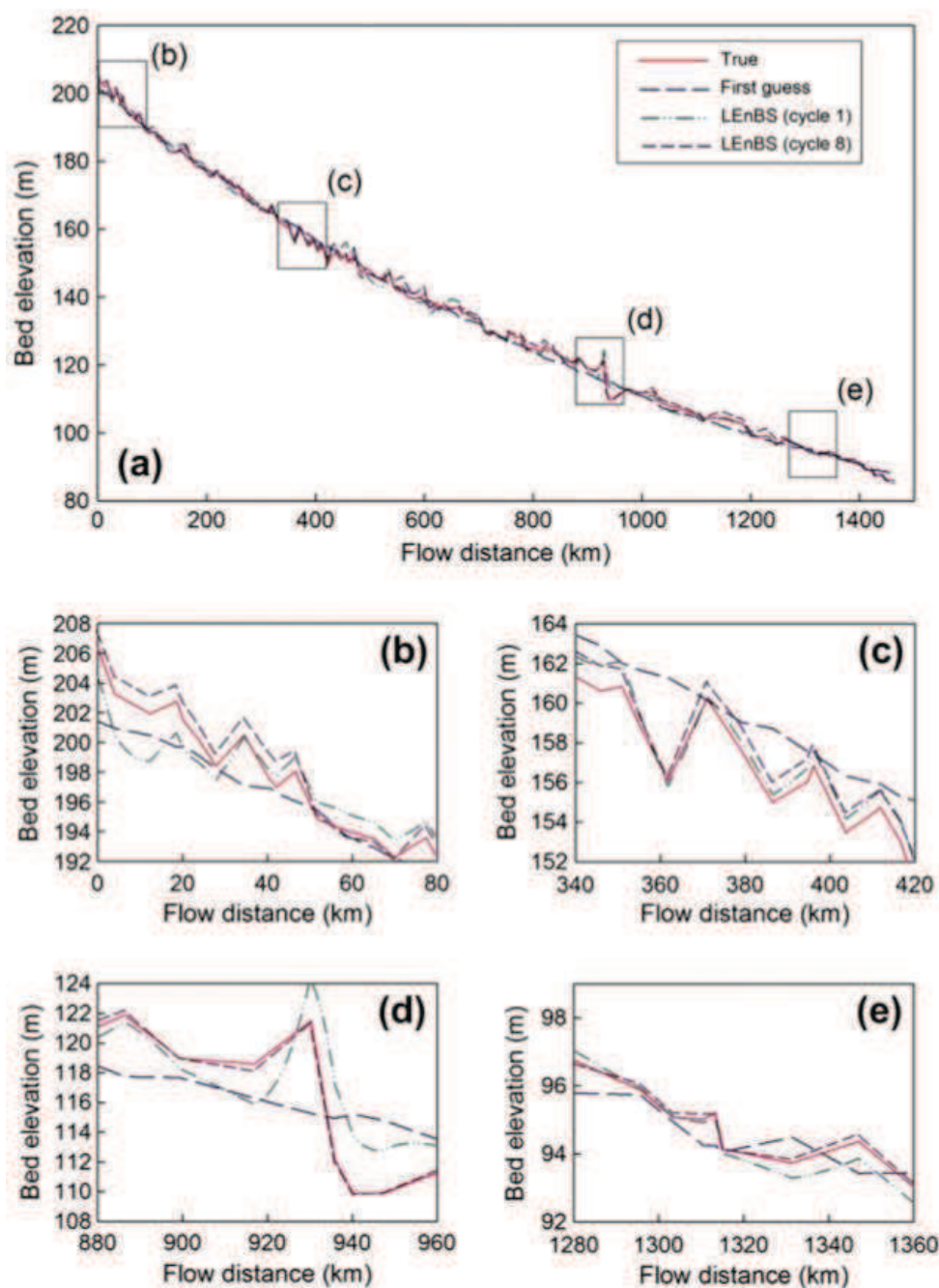


FIG. 6.5 – Estimation itérative de la bathymétrie dans [Yoon et al., 2012]. Etat vrai (ligne rouge), ébauche (ligne bleue tiretée), assimilation aux cycles 1 et 8 (lignes vertes et violettes tiretées respectivement).

Après avoir présenté dans les grandes lignes certains des travaux les plus récents en assimilation de données pour l'hydraulique devons-nous nous positionner par rapport à ceux-ci et présenter l'apport de nos propres travaux. Parmi les travaux que

nous venons de mentionner très peu s'intéressent à la forme des fonctions de covariance intervenant dans le processus d'assimilation de données. Pourtant ces fonctions tiennent un rôle central car ce sont elles qui répartissent sur le domaine l'information apportée par les observations. [Shiiba et al., 2000] propose une approximation de la matrice des covariances d'erreur d'ébauche en utilisant un filtre de Kalman de rang réduit. [Madsen and Skotner, 2005] propose un panel de fonctions stationnaires aux points d'observations tandis que [Ricci et al., 2011] utilise en ces points une fonction de covariance stationnaire évoquant une gaussienne asymétrique. Ces deux derniers travaux permettent l'implémentation d'un filtre de Kalman stationnaire pour un coût de calcul réduit puisqu'il ne nécessite pas de calculer le modèle linéaire tangent ni de propager la matrice des covariances d'erreur d'ébauche. Néanmoins le contexte mathématique sur lequel s'appuient ces travaux pour justifier le choix des fonctions de covariance aux points d'observation est limité. Dans une première partie nous nous sommes intéressés (voir chapitre 7) à la forme de la fonction de covariance au point d'observation lorsque nous appliquons un EnKF sur un modèle d'onde de crues et donnons une paramétrisation complète de ces fonctions en fonction du rapport des variances d'ébauche et d'observation au point d'observation. On justifie ainsi pleinement le choix des fonctions de covariance utilisées dans [Madsen and Skotner, 2005, Ricci et al., 2011] au point d'observation et permettant d'émuler l'EnKF mais pour un coût de calcul nettement moindre.

Dans un second temps nous avons appliqué un filtre de Kalman d'ensemble sur le réseau hydraulique « Adour maritime » présenté en section 1.3.2. Nous nous sommes intéressés plus particulièrement à l'étude qualitative des fonctions de covariance et corrélation ainsi qu'à leur évolution temporelle. (voir section 8.2). Nous avons également montré comment l'application des critères de [Desroziers et al., 2005] permet d'améliorer les résultats de l'EnKF par rapport à sa version originale non seulement au point d'observation où l'assimilation est réalisée mais aussi ailleurs sur le réseau et ce sur une expérience d'assimilation de données synthétiques (voir sections 8.3 et 8.4). Dans le cas de l'assimilation de données réelles cette méthode semble limitée par l'absence de modélisation du lit majeur avec dans certains cas une dégradation des résultats à certaines stations par rapport au modèle seul. Cette application de l'EnKF au réseau Adour a été complétée par une discussion sur la possibilité d'établir des prévisions d'ensemble sur le réseau (voir section 8.5).

Enfin dans une dernière partie nous avons montré que la longue extension spatiale des fonctions de covariance permet de corriger les forçages d'un modèle 2D dans le cadre du couplage de modèles superposés sur la zone de Bayonne (section 9.3).

Chapitre 7

Emulation du filtre de Kalman d'ensemble sur le modèle d'ondes de crue 1D

Sommaire

7.1	Introduction	139
7.2	Article soumis dans la revue « Journal of Hydrology » . . .	140
7.3	Conclusion	158

7.1 Introduction

L'étude présentée dans ce chapitre fait suite aux premiers travaux sur l'assimilation de données pour l'hydraulique fluviale menés au CERFACS. Ceux-ci, menés en collaboration avec le SCHAPI, portaient sur l'assimilation de données pour la correction des forçages et de l'état hydraulique en temps réel. Ils ont donné lieu à l'article [Ricci et al., 2011] d'une part et se sont concrétisés par la mise en place à l'opérationnel de cette méthodologie pour la prévision des crues en temps réel dans les SPC Gironde-Adour-Dordogne et Seine Amont-Marne Amont ainsi qu'au sein de la POM (Plateforme Opérationnelle Multi-modèles) au SCHAPI.

[Ricci et al., 2011] combine une procédure de correction des forçages amonts avec l'utilisation d'un filtre de Kalman invariant pour la correction de l'état hydraulique sur les réseaux « Adour maritime » et « Marne amont » et montre que cette démarche permet d'améliorer les prévisions de hauteurs d'eau à moyenne échéance, 6 à 12 heures. Dans cette étude [Ricci et al., 2011] utilise des fonctions de covariance aux points d'observation qui sont stationnaires et asymétriques avec une longueur de portée réduite en aval des stations d'observation. Dans cette étude, le choix de la forme asymétrique de ces fonctions ainsi que la valeur des longueurs de portée ne sont pas rigoureusement justifiés. [Ricci et al., 2011] montre sur un modèle d'onde de crues 1D que l'utilisation de fonctions de covariance symétriques donne des résultats moins bons que ceux obtenus avec une fonction de covariance asymétrique avec une longueur de portée aval plus petite que la longueur de portée amont. Cela s'explique par le fait qu'une fonction de cova-

riance symétrique sur-corrige les états analysés à l'aval de la station d'observation et donc dégrade les résultats de l'assimilation.

Dans l'étude présentée dans ce chapitre nous nous sommes attachés à justifier l'asymétrie des fonctions de covariance utilisées dans [Ricci et al., 2011]. Pour cela nous avons implémenté la version stochastique du filtre de Kalman d'ensemble sur le modèle d'ondes de crues 1D. Le choix de l'EnKF plutôt que celui du filtre de Kalman classique pour mener cette étude s'explique par deux points :

- l'EnKF ne nécessite pas de connaître l'adjoint du modèle ;
- l'EnKF permet de prescrire une erreur sur le forçage amont dont on estime qu'il s'agit de la principale source d'incertitude dans le modèle.

Ainsi nous avons étudié la dynamique des fonctions de covariance et de corrélation sur le modèle d'onde de crue 1D avec et sans assimilation. Les fonctions de corrélation, à travers le diagnostic de la longueur de portée, portent une information sur l'extension spatiale des fonctions de covariance et donc sur la répartition de l'information apportée par les observations sur le domaine ailleurs qu'au point d'observation. Cette étude montre qu'une fonction de corrélation au point d'observation initialement gaussienne est transformée à travers les cycles d'assimilation en une fonction asymétrique avec une longueur de portée plus petite à l'aval qu'à l'amont.

D'autre part cette étude montre, sous certaines hypothèses, qu'après un certain nombre de cycles d'assimilation les fonctions de covariances diagnostiquées par l'EnKF sont stationnaires. Alors un filtre de Kalman invariant utilisant ces fonctions de covariance donne des résultats comparables à ceux de l'EnKF mais pour un coût de calcul nettement moindre ; argument de taille pour la mise en opérationnel de l'assimilation pour la prévision en temps réel des crues.

7.2 Article soumis dans la revue « Journal of Hydrology »



Emulation of a Kalman filter algorithm on a flood wave propagation model

Sébastien Barthélémy^a, Sophie Ricci^a, Olivier Pannekoucke^{b,a}, Olivier Thual^{c,a},
Pierre-Olivier Malaterre^d

^aURA 1875/CERFACS, Toulouse, France

^bCNRM-GAME, UMR 3589, Toulouse, France

^cINPT, CNRS, IMFT, Toulouse, France

^dUMR G-EAU, Irstea, Montpellier, France

Abstract

This study describes the emulation of an Ensemble Kalman Filter (EnKF) algorithm on a 1D flood wave propagation model. This advection-diffusion model is forced at the upstream boundary through a random variable with stationary gaussian statistics and a correlation function in time with gaussian shape. When the advection time is large compared to the diffusion time, an analytic study, confirmed by numerical experiments, shows that the covariance functions of the propagated signal anomaly keeps its gaussian shape in the absence of observation. In the case of a one observation point assimilation where synthetical observations are generated by adding an error to a true state, the dynamics of the background error covariance functions is no longer simple and a numerical approach using an EnKF algorithm is preferred. It is shown that the application of a Best Linear Unbiased Estimator (BLUE) algorithm, using the background error covariance matrix converged from an EnKF algorithm, provides the same results as the EnKF. Numerical experiments lead to the construction of a parametrized matrix that emulates the background error covariance matrix converged from the EnKF. This method is then applied to a case with two different observation points with different error statistics providing results close to those of the EnKF, with a significantly reduced computational cost that allows for the use of data assimilation in the context of real time flood forecasting.

© 2014 Published by Elsevier Ltd.

Keywords: Data assimilation, Kalman Filter, Flood wave propagation, Flood forecasting

1. Introduction

With flood frequency likely to increase as a result of altered precipitation patterns triggered by climate change [6] there is a growing need for improved flood modeling. While significant advances have been made in recent years in hydraulic data assimilation (DA) for water level and discharge prediction [28, 2], as well as for parameters correction [26, 7], using in-situ as well as remote sensing data [1, 23] they are yet to be fully taken advantage of in the operational forecast of flood and inondation areas. Recent studies have shown the benefit hydrology and hydraulics can draw from the progress of DA approaches using either variational inverse problem [33], particle filtering [17, 12], Extended Kalman Filter [32], Ensemble Kalman Filter (EnKF) for state updating [21, 37], or for dual state-parameter estimation [20, 13]. In the field of hydraulic, amongst the numerous research studies in DA that aims at overcoming the limitations of the hydraulics models, only a few are formulated in an operational setting and demonstrate the performance gained from DA [14, 16, 38, 11, 27].

For most DA algorithms, the description of the background error covariance matrix is essential but fastidious. Indeed the background error covariance matrix plays a keyrole in DA as this matrix spreads the information brought by the observations over the domain and between the state variables. These covariances must be modeled as approximations of the true covariances of background errors. In sequential DA algorithms such as the Kalman Filter (KF) algorithm this matrix is propagated by the dynamic of the model and is updated each time an observation is available [3, 4], namely for each cycle of analysis. The KF algorithm decomposes in two steps. A first step where the state vector and the analysis error covariance matrix are updated using information from the observations. In the second step, the analyzed state vector and the analysis error covariance matrix are propagated from the current assimilation cycle to the next to respectively describe the new background state vector and the new background error covariance matrix. This last step requires the computation of the tangent linear of the dynamical model and computation of the matrix products. As stochastic methods [8, 19, 21] offer an alternative method to these issues, they require a large number of model integrations and should often be complemented with cost reduction methods such as localization [31, 30]. For large dimension problem and also when the computationnel time is an issue (for operational purpose for instance), the explicit formulation and propagation of the covariances is not possible and the covariance matrix should be described using a model. This is done for instance by parametrizing the covariances using smoothing functions and balance relationships [5, 35] that are formulated with a limited number of parameters such as correlation length scales and standard deviations [34, 24]. Reference studies from [9, 10] describe how the asymptotic behaviour of the Kalman Filter algorithm allows to identify an invariant background error covariance matrix under conditions of observability and controlability. Assuming the statistics of the errors in the system remain stationnary, this asymptotic matrix can thus be used in a simplified KF algorithm, at a much reduced computational cost. In the framework of data assimilation applied to hydraulics, numerous study present the results of the implementation of a KF or EnKF. Only a few studies address the modelling of the covariance matrix itself and the evolution of the background error statistics. [29] approximates the background error covariance matrix in a KF algorithm with a lower rank matrix using an eigen vector decomposition. [14] proposed an invariant formulation of the KF using a panel of simple covariance functions at the observation point whereas [27] uses at the observation points a steady covariance functions. While these studies offer a practical solution for the implementation a reduced-cost data assimilation algorithm, the mathematical framework that justifies the choice of the correlation functions at the observation point is quite limited.

In this paper, a parametrization for the background error covariance matrix in a sequential filter is proposed as a reduced cost alternative of the EnKF in order to allow for the use of DA in the context of operational flood forecasting and the form of the parametrization is fully justified. The first step of the study stands in the implementation of an EnKF from which the background error covariance model parametrization is derived and validated. It is shown that the parametrized filter successfully emulates the EnKF at a much reduced cost and that the parametrization can be extended for a different observation network. In the context of operational flood-forecasting, running an ensemble of 10 or 20 integrations of a 1D or 2D shallow-water models is not compatible with real-time constraints, computational cost but also with current operational chain implementation. For that reason, the study is carried out on a simplified 1D diffusive flood wave propagation model that approximates the Saint-Venant equations usually derived in 1D and 2D hydraulic models used for flood forecasting. With this simplified model, a large number of members for the EnKF are used, so that we obtain less than 1% accuracy compared to the asymptotic limit of an infinite number of members. The uncertainty in the model are supposed to be due to uncertainties in the upstream forcing that are not easy to take into account in the formulation of the KF algorithm which is why an ensemble method is favored.

It is first shown that without assimilation, the evolution of the water level anomaly (WLA) covariances, initially prescribed as gaussian, can be described analytically and validated with an ensemble approach via the computation of the covariance matrix $\mathbf{B}_e$. When the diffusion is small, the covariances shape remain gaussian with an increasing correlation length scale and decreasing variance as the signal propagates. Then DA experiments are carried out in the framework of Observing System Experiment (OSE) with a steady observation network. It is shown that an initial correlation function of gaussian shape turns into an anisotropic function at the observation point, with a shorter correlation length scale downstream of the observation point than upstream (which is one of the functions proposed in Madsen et al. 2005), and that the error variance of the state is significantly reduced downstream of the observation point. The resulting converged matrix $\mathbf{B}_{\text{EnKF}}$ can be used as an invariant background error covariance matrix with the deterministic DA algorithm BLUE to emulate the EnKF (EEnKF) with a significantly reduced computational cost. A parametrized model for background error correlation length scale $L_p(x)$ and variance $\sigma^2(x)$ in $\mathbf{B}_{\text{EnKF}}$ is finally established over the whole domain. Along with a diffusion operator [25, 18, 36], this model allows to emulate the

EnKF for observing networks where the number of observations and the observation error statistics vary.

The outline of the paper is as follows : Section 2 describes the diffusive flood wave propagation model. It also provides theoretical proof for the generation of a signal with a Gaussian spatial covariance function and for the evolution of the correlation function and length scale without assimilation. The numerical validation, with an ensemble approach, for these theoretical results is presented as well. A brief description of the ensemble based DA algorithms used in the paper and of the diffusion operator are given in Section 3. In Section 4 the results of the EnKF and EEnKF algorithms regarding the evolution of water level and its error statistics are outlined. The parametrization of the reduction of the background error correlation length scale and variance at the observation point as a function of the observation and background error statistics is also presented and used when the observing network is modified. Some conclusive remarks and perspectives are given in Section 5.

2. Dynamic of the flood wave propagation model

2.1. The flood wave propagation model

2.1.1. Model equations

The shallow-water equations can be approximated by the diffusive flood wave propagation equation when the river slope is important:

$$\frac{\partial h}{\partial t} + c \frac{\partial h}{\partial x} = \kappa \frac{\partial^2 h}{\partial x^2}, \quad (1)$$

where h is the Water Level Anomaly (WLA), namely a perturbation to the equilibrium state (h_m, U_m) (where U_m denotes the velocity) such that $U_m = K_s(\sin \gamma)^{1/2} h_m^{2/3}$, with K_s the Strickler coefficient. Equation (1) is a classical advection-diffusion equation where $c = \frac{5U_m}{3}$ is the advection speed and $\kappa = \frac{U_m h_m}{2 \tan \gamma}$ is the diffusion coefficient, with a constant slope γ on a 1D domain defined for $x \in [0, L]$, with $L = 200$ km, discretised in $N = 200$ points. An open boundary condition is imposed downstream with $\frac{\partial h}{\partial t}(L, t) + c \frac{\partial h}{\partial x}(L, t) = 0$ and the upstream boundary condition h_{up} is described in Sect. 2.1.2.

A fourth order Runge-Kutta (RK4) scheme is used in place of an Euler first order temporal scheme allowing for a proper diffusion in the numerical resolution of Eq. (1). Still, it should be noted that for high frequency signals the RK4 scheme can also lead to spurious dispersion thus implying a lower limit for the choice of the initial correlation length (this limit is estimated numerically). This effect is small compared to an explicit Euler scheme (see Annexe 1).

2.1.2. The upstream forcing

The upstream boundary condition is imposed by $h(0, t) = h_{up}(t)$, where h_{up} is characteristic of a flow up to a multiplicative constant. Here h_{up} is modeled as a stationary Gaussian random process characterised by a temporal auto-covariance function in time $\rho_t(\delta t) = \langle h_{up}(t) h_{up}^*(t + \delta t) \rangle$ that has a gaussian shape of correlation time scale τ , $\rho_t(\delta t) = q_m^2 e^{-\frac{\delta t^2}{2\tau^2}}$. The Gaussian hypothesis on the random process h_{up} implies that it can be fully described by the first and second order moments of its distribution and the Kalman filter equations can be used. This forcing translates into a WLA signal with a spatial covariance function that has a gaussian shape as shown in Sect. 2.2.1. This choice is made in order to prescribe a known covariance function for the WLA signal and study how it is evolved by the flood wave propagation model.

For the purpose of the theoretical derivation we now switch to a spectral representation of h_{up} . The construction of the upstream forcing requires the formulation of $h_{up}(t)$ using Fourier transform, $h_{up}(t)$ can be written as a sum of harmonic signals

$$h_{up}(t) = \int_{\mathbb{R}} h_{up,\omega} e^{-i\omega t} d\omega. \quad (2)$$

The identification of the $h_{up,\omega}$ coefficients relies on the knowledge of the covariance function $\rho(\delta t)$. Due to the stationarity of the random process, the complex amplitudes $h_{up,\omega}$ are uncorrelated so that $\langle h_{up,\omega} h_{up,\omega'}^* \rangle = \rho_\omega \delta(\omega - \omega')$ where δ is the Dirac distribution and where the energy spectrum $\rho_\omega = \frac{\tau}{\sqrt{2\pi}} q_m^2 e^{-\frac{\omega^2 \tau^2}{2}}$ is the Fourier transform of

$$\rho_t(\delta t) = \int_{\mathbb{R}} \rho_\omega e^{-i\omega \delta t} d\omega \quad (3)$$

from the *Wiener-Khintchine* theorem.

Since the random process is Gaussian, $h_{up,\omega}$ in Eq. (2) can be written as

$$h_{up,\omega} = \zeta_\omega l_\omega \quad (4)$$

where ζ_ω is a complex gaussian random variable whose module has zero mean and standard deviation 1 and $\zeta_\omega = \zeta_{-\omega}^*$ and the l_ω are complex number of arbitrary phases. Using the definition of the correlation function in time of h_{up} one can show that

$$\rho_t(\delta t) = \int_{\mathbb{R}} |l_\omega|^2 e^{-i\omega\delta t} d\omega. \quad (5)$$

By identification of the two expressions of ρ in Eq. (3) and Eq. (5) it comes that $|l_\omega| = \sqrt{\rho_\omega}$. Therefore, from a numerical point of view the upstream forcing h_{up} can be built as the inverse Fourier transform of $h_{up,\omega} = \zeta_\omega l_\omega$.

2.2. Study of the covariances dynamic

Given an upstream forcing with a known temporal covariance function for the propagation model, the description of the WLA covariance function is first described analytically in Sect. 2.2.1 and corroborated with an numerical ensemble approach in Sect. 2.2.2 when no data are assimilated. The study of the covariance dynamics when data assimilation is applied is only studied with a numerical approach in Sect. 3 since no analytical solution is available.

2.2.1. Analytical study of the covariances dynamic

Knowing the characteristics of the temporal covariance of the boundary condition flow $h_{up}(t)$, the spatial covariance of the WLA state can be derived. Given the linearity of the problem, the solution $h(x, t)$ can be formulated as the superposition of modal solutions. Assuming that the forcing is a sinusoidal function $h_{up}(t) = q_\omega e^{-i\omega t}$, a modal solution for Eq. (1) is of the form

$$h(x, t) = q_\omega e^{-i\omega t} h_\omega(x) \quad (6)$$

where q_ω is the magnitude of the mode that in the particular case of the upstream forcing can be identified to $\zeta_\omega |l_\omega|$ in Eq. (4). For any forcing $h_{up}(t)$, the general solution reads

$$h(x, t) = \int_{\mathbb{R}} q_\omega e^{-i\omega t} h_\omega(x) d\omega. \quad (7)$$

In the case of advection only ($\kappa = 0$), the solution of the form given in Eq. (6) is $h_\omega(x) = e^{i\omega \frac{x}{c}}$. Thus the general solution reads

$$h(x, t) = \int_{\mathbb{R}} q_\omega e^{-i\omega t} e^{i\omega \frac{x}{c}} d\omega. \quad (8)$$

of which the spatial covariance function ρ is a gaussian, defined as:

$$\begin{aligned} \rho(x, x + \delta x) &= \langle h(x, t) h^*(x + \delta x, t) \rangle \\ &= \int_{\mathbb{R}} |q_\omega|^2 e^{i\omega \frac{\delta x}{c}} d\omega = \rho\left(\frac{\delta x}{c}\right) \end{aligned} \quad (9)$$

In summary, in the case of advection only, a forcing signal with a gaussian temporal covariance function translates into a WLA signal with a gaussian spatial covariance function of constant length scale $L_0 = c\tau$ and constant variance $\sigma_0^2 = q_\omega^2$.

In the case of advection and small diffusion, which reads $\kappa \ll cx$, a straightforward expansion leads to,

$$h_\omega(x) = e^{\left(\frac{c}{2\kappa} - \frac{\sqrt{c^2 - 4i\omega\kappa}}{2\kappa}\right)x} \approx e^{i\omega \frac{x}{c} - \frac{\omega^2 \kappa}{c^3} x} \quad (10)$$

and a more elaborated asymptotic analysis shows that ρ can locally be approximated by

$$\rho(x, x + \delta x) \approx \int_{\mathbb{R}} \frac{q_m^2}{\sqrt{2\pi}} L_p(0) e^{-\frac{\omega^2 L_p^2(x)}{2}} e^{-i\omega \delta x} d\omega = q_m^2 \frac{L_p(0)}{L_p(x)} e^{-\frac{\delta x^2}{2L_p(x)^2}}, \quad (11)$$

that is a locally Gaussian covariance function of correlation length scale :

$$L_p(x) = \sqrt{L_0^2 + 4\kappa \frac{x}{c}} \quad (12)$$

and variance :

$$\sigma^2(x) = \sigma^2(0) \frac{L_p(0)}{L_p(x)}. \quad (13)$$

In summary, in the case of advection and small diffusion a forcing signal with a gaussian temporal covariance function translates into a WLA signal with a spatial covariance function that can be approximated by a gaussian of length scale $L_p(x)$.

2.2.2. Validation with an ensemble approach

These theoretical results are validated computing the covariance matrix $\mathbf{B}_e$ of an ensemble of N_e WLA states $\mathbf{x}^k = (h_{1,k}, \dots, h_{N,k})$ on the 1D domain $[0, L]$ where N is the number of grid point, generated with different forcings $h_{up,k}(t)$ with $k \in [1, N_e]$ that follows the statistics described in Sect. 2.1.2. The correlation length scale L_p is computed by the gaussian based approximation [24]:

$$L_p(x) = \frac{\delta x}{\sqrt{2(1 - \rho(x, x + \delta x))}} \quad (14)$$

where $\rho(x, x + \delta x)$ is the correlation estimated from data or analytical formula such as Eq. (11).

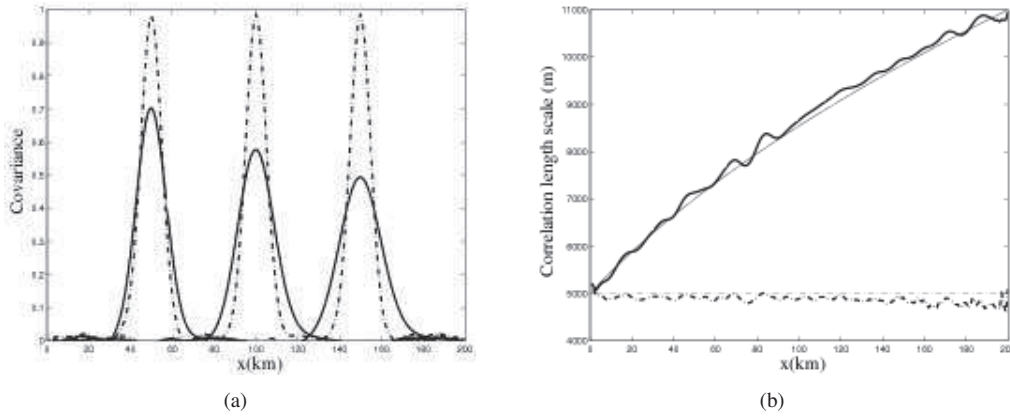


Figure 1. (a) Covariance function from $\mathbf{B}_e$ for $x = 50, 100, 150$, (b) Correlation length scale over the domain. The case of advection only is represented with dashed lines and the case of advection-diffusion is represented with solid lines. In (b), the results from the theoretical analysis are represented with thin lines and the results from the numerical analysis are represented with thick lines.

Fig. 1-a displays the covariance function at 3 different points of the 1D domain for the advection only case (dashed lines) and for the advection-diffusion case (solid lines), for $N_e = 10000$ (so that the sampling noise is less than $10^{-2} m^2$), $\sigma^2(0) = 1 m^2$ and $\tau = 5 \times 10^3 s$. In the first case, the initially gaussian function is advected. The characteristics of the covariance function remain unchanged as illustrated in Fig. 1-b where L_p , estimated from Eq. (14) is constant (dashed thick line), and in agreement with the theoretical value L_0 (dashed thin line). When diffusion occurs, the

covariance function of the WLA state is diffused as shown in Fig.1-a and L_p increases with x (solid thick line in Fig. 1-b), still in agreement with the theoretical value from Eq. (12) (solid thin line in Fig. 1-b). The WLA variance represents the maximum amplitude of the covariance functions in Fig. 1-a. The initial value prescribed at 1 m² remains constant for the advective case (dashed lines) and decreases with x for the advective-diffusive (solid lines) case in agreement with the theoretical results in Eq. (13).

3. Data assimilation algorithms

The classical equations for EnKF are presented in Sect. 3.1 while the EEnKF is presented in Sect. 3.2. For these algorithms, the background error covariance matrix is characterized by the correlation length scale $L_p(x)$ and the variance $\sigma^2(x)$ for which a parametrized model is presented in Sect. (4.3). Given these information only, the diffusion operator described in Sect. (3.3) allows to fully describe a covariance matrix to be used in the EEnKF for various observing networks.

3.1. The Ensemble Kalman Filter algorithm

The EnKF algorithm [8] is implemented on Eq. (1), using an OSE (Observing System Experiment) framework. A reference run is integrated using a given forcing $h_{up}^{true}(t)$, to simulate the *true* WLA $h^{true}(x, t)$. The observation

$$h^{obs}(x_{obs}, t) = h^{true}(x_{obs}, t) + \epsilon^o(t) \quad (15)$$

is then calculated in the middle of the 1D domain $x_{obs} = \frac{L}{2}$ where $\epsilon^o(t)$ is a Gaussian noise defined by its standard deviation σ^o ($\sigma^o = 0.2354$ m in the following), thus defining the observation vector $\mathbf{y}^o$. The background trajectories $h_k^b(x, t)$ for the ensemble approach are integrated using a perturbed set of forcing $h_{up,k}(t)$ with $k \in [1, N_e]$, defining the background vectors $\mathbf{x}^{b,k}(t)$ for the DA analysis at time t . The observation frequency is set to 3 model time steps. In the following DA is applied over a cycle between two observation times $t = i$ and $t = i + 1$ (assimilation cycle $i + 1$).

As illustrated on Fig. 2 for the assimilation cycle $i + 1$, the ensemble of previously analyzed states $\mathbf{x}_i^{a,k}$ are propagated by the diffusive flood wave model $M_{i,i+1}$ from the observation time i to $i + 1$ to provide the background states $\mathbf{x}_{i+1}^{b,k} = M_{i,i+1}(\mathbf{x}_i^{a,k})$ over which the background error covariance matrix $\mathbf{B}_{EnKF,i+1}$ is computed.

$$\mathbf{B}_{EnKF,i+1} = \frac{1}{N_e - 1} \sum_{k=1}^{N_e} (\mathbf{x}_{i+1}^{b,k} - \bar{\mathbf{x}}_{i+1}) (\mathbf{x}_{i+1}^{b,k} - \bar{\mathbf{x}}_{i+1})^T. \quad (16)$$

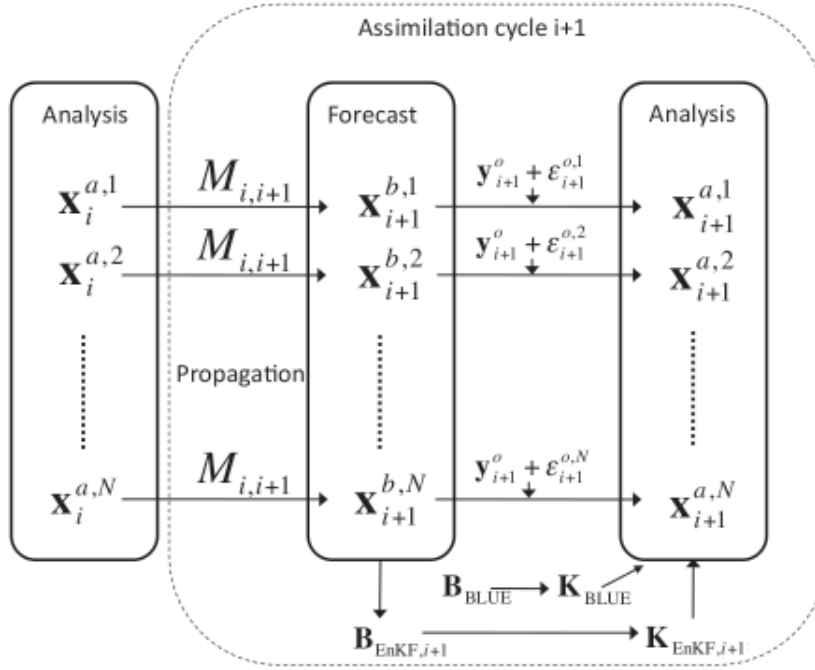
where $\bar{\mathbf{x}}_{i+1} = \frac{1}{N_e} \sum_{k=1}^{N_e} \mathbf{x}_{i+1}^{b,k}$ and T stands for the transposition operator.

The assimilation step at $i + 1$ consists in assimilating a perturbed observation vector $\mathbf{y}_{i+1}^o + \epsilon_{i+1}^{o,k}$ (Burgers et al., 1998) to correct the background vector $\mathbf{x}_{i+1}^{b,k}$, using the Kalman Filter gain matrix $\mathbf{K}_{EnKF,i+1}$:

$$\mathbf{x}_{i+1}^{a,k} = \mathbf{x}_{i+1}^{b,k} + \mathbf{K}_{EnKF,i+1} (\mathbf{y}_{i+1}^o + \epsilon_{i+1}^{o,k} - \mathbf{H}(\mathbf{x}_{i+1}^{b,k})) \quad \text{with} \quad (17)$$

$$\mathbf{K}_{EnKF,i+1} = \mathbf{B}_{EnKF,i+1} \mathbf{H}^T (\mathbf{H} \mathbf{B}_{EnKF,i+1} \mathbf{H}^T + \mathbf{R})^{-1}. \quad (18)$$

Assuming that the observation network remains the same, after 1000 assimilation cycles $\mathbf{B}_{EnKF,i}$ converges to a steady matrix denoted $\mathbf{B}_{EnKF}$ (its associated correlation matrix is $\mathbf{C}_{EnKF}$, $\mathbf{C}_{EnKF}$ is computed with $\Sigma \mathbf{B}_{EnKF} \Sigma^T$ where Σ is the diagonal matrix with the inverse of the standard deviation of $\mathbf{B}_{EnKF}$ on its diagonal). [15] showed that ensemble errors due to the Monte Carlo sampling in EnKF can be dominant compared to other errors (numerical or model errors) and that in order to estimate converged statistics in $\mathbf{B}_{EnKF,i+1}$ a large number of members N_e is required. Indeed it can be shown that $N_e = 10000$ members are necessary to reduce the sampling noise on the background error covariance matrix to less than $10^{-2} m^2$.

Figure 2. Ensemble data assimilation algorithms, assimilation cycle $i+1$.

3.2. Emulation of the EnKF algorithm

The BLUE (Best Linear Unbiased Estimator, Bouttier and Courtier 1999) algorithm can be viewed as a simplification of the Kalman Filter in which the background error covariance matrix is not propagated over the cycles. The analysis equation Eq. (17) is applied sequentially with a constant matrix $\mathbf{B}_{BLUE}$:

$$\mathbf{K}_{BLUE} = \mathbf{B}_{BLUE} \mathbf{H}^T (\mathbf{H} \mathbf{B}_{BLUE} \mathbf{H}^T + \mathbf{R})^{-1}. \quad (19)$$

If $\mathbf{B}_{BLUE} = \mathbf{B}_e$ (where $\mathbf{B}_e$ is the covariance matrix computed without assimilation in Sect. 2.2.2), one misses the fact that the background error covariance matrix should be impacted by the previous assimilation steps. When the BLUE algorithm is applied with $\mathbf{B}_{BLUE} = \mathbf{B}_{EnKF}$, the algorithm is called the Emulated EnKF and is denoted by EEnKF. Once $\mathbf{B}_{EnKF}$ is computed, this algorithm only requires the integration of a single member (with a single upstream forcing h_{up}) and its computational cost is thus significantly lower than the EnKF.

Still, the BLUE algorithm can of course be applied to an ensemble of forcings $h_{up,k}$ so that a comparison of the correlation length scale and variance of the EnKF and this ensemble of BLUE algorithms can be made, as pictured in Fig. 2. In the following, the ensemble of BLUE analysis using $\mathbf{B}_{BLUE} = \mathbf{B}_e$ or $\mathbf{B}_{BLUE} = \mathbf{B}_{EnKF}$ are called respectively EnBLUE $_{B_e}$ and EnBLUE $_{B_{EnKF}}$. The correlation length scale and variance of the covariance matrix computed over the members for both algorithm are compared to those of $\mathbf{B}_{EnKF}$ in Sect. 4.

3.3. The diffusion operator

The columns and lines of the matrix $\mathbf{B}_{EnKF}$ contain the discretization of the background error covariance functions for each grid point of the domain. In this section we present how the covariance functions in $\mathbf{B}_{EnKF}$ can be fully described simply using the diagnosed correlation length scale and variance when the EnKF is converged. This can be done using the diffusion operator [34] that comes down to formulating the matrix, vector product $\mathbf{B}\mathbf{x}$ instead of formulating the matrix $\mathbf{B}$. The diffusion operator is a mathematical tool that allows to model covariances and

correlations for DA algorithm. It is widely used in meteorology and oceanography, where the dimension of the control vector is large and the covariance matrix should be formulated as an operator (applied to a vector) rather than as a matrix. The solution of a 1D pseudo-diffusion equation on an infinite domain for a pseudo-time T reads

$$\eta(x, T) = \frac{1}{\sqrt{4\pi\tilde{\kappa}T}} \int_{-\infty}^{+\infty} e^{-\frac{(x-x')^2}{4\tilde{\kappa}T}} \eta_0(x') dx' \quad (20)$$

where $\tilde{\kappa}$ is the constant pseudo-diffusion coefficient, η_0 is the initial condition and η vanishes as $x \rightarrow \pm\infty$. Equation (20) is the convolution product of $g(x) = \frac{1}{\sqrt{4\pi\tilde{\kappa}T}} e^{-\frac{x^2}{4\tilde{\kappa}T}}$ and $\eta_0(x)$. Since g is positive definite, Eq. (20) describes a covariance operator correlation length scale $L_p = \sqrt{2\tilde{\kappa}T}$ applied to η_0 . The corresponding correlation operator is obtained with a normalization factor $\lambda = \sqrt{4\pi\tilde{\kappa}T}$. The multiplication of the desired variance σ^2 finally describes the expected covariance operator. This result can be extended to heterogeneous diffusion tensor $\tilde{\kappa}(x) = \frac{L_p^2(x)}{2T}$ [25]. In the following, $L_p(x)$ and $\sigma^2(x)$ are specified with the analytical equations Eq. (12) and Eq. (13) away from the observation points and with the abacus from EnKF experiments (presented in Sect. 4) at the observation point. From a numerical point of view, the pseudo-diffusion model is applied at each grid point to a Dirac function. The result is then normalized and multiplied by the background error variance to provide the background error covariance function at the grid point and thus the complete background error covariance matrix with correlation length scale $L_p(x)$ and variance $\sigma^2(x)$.

4. Results

4.1. Comparison of EnKF and EEnKF results

The EnKF is applied for the observing network described in Sect. 3.1 with the initial background error covariance matrix $\mathbf{B}_e$ computed over the integrated members without DA (Sect. 2.2.2) - the associated correlation function is noted $\mathbf{C}_e$. Figure 3-a illustrates how the initially isotropic correlation function in $\mathbf{C}_e$ (dashed line) at the observation point ($x_{obs} = 100$) is modified by the analysis and propagation steps of the EnKF algorithm, at the end of the assimilation procedure. Considering a steady observation network, the shape of the correlation function in $\mathbf{C}_{EnKF}$ (solid line) converges towards an anisotropic function with a shorter correlation length scale downstream of the observation point than upstream. The correlation between the observation point and its neighbors is reduced since information at the observation point is introduced at this location by the analysis procedure through the observation vector at the previous analysis cycles.

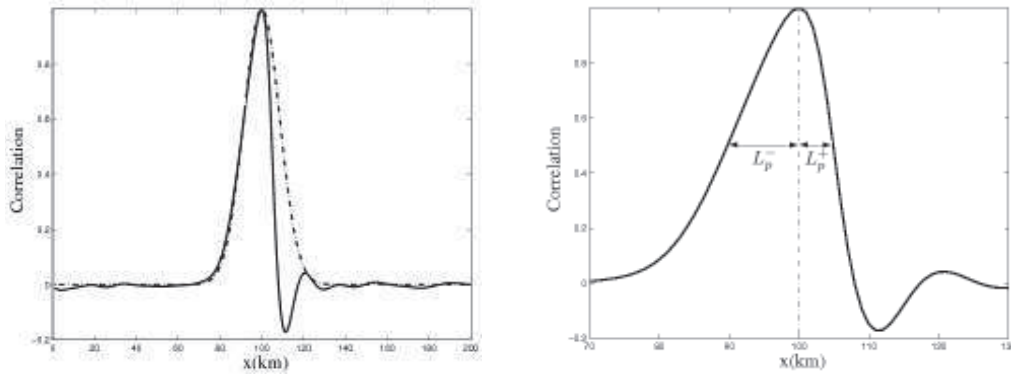


Figure 3. Background error correlation function at the observation point, for the initial correlation matrix $\mathbf{C}_e$ (dashed line) and for $\mathbf{C}_{EnKF}$ (solid line) a- on the whole domain, $x \in [0; 200]$ and b- on the domain $x \in [70; 130]$ with the upstream and downstream correlation length scale, L_p^- and L_p^+ respectively.

The correlation length scale L_p and variance σ^2 are computed for the EnKF as well as for the ensemble of BLUE analysis using either the initial covariance matrix $\mathbf{B}_e$ (EnBLUE_{B_e} algorithm) or $\mathbf{B}_{\text{EnKF}}$ (EnBLUE_{B_{EnKF}} algorithm). Equation (14) is used to estimate the correlation length scale L_p at each point away from the observation point. Fig. 4-a shows that the evolution of the background error correlation length scale for the EnKF (thick solid line) follows the theory (thin solid line) upstream of the observation point. At the observation point, where a discontinuity occurs, the upstream and downstream correlation length scales (respectively L_p^- and L_p^+) differ. For the EnKF algorithm (thick solid line), the reduction of the correlation length scale spreads over the entire domain downstream of the observation point. This result is well reproduced by the EnBLUE_{B_{EnKF}} (thick dashed line that overlaps the thick solid one) algorithm, whereas for EnBLUE_{B_e} (thin dashed line), the reduction is local. The upstream correlation length scale at the observation point L_p^- is approximated by Eq. (12) with $x = 100\text{km}$, meaning that the assimilation has no impact on the correlation function upstream of the observation point, which is consistent with the flood wave approximation in the propagation model. The downstream correlation length scale at the observation point L_p^+ should be extrapolated from the ensemble based estimate function for $x > 100\text{km}$ that is of the form given in Eq. (12). L_p^+ can thus be used in place of L_0 in Eq. (12) to derive the analytical function in Eq. (23) for the rest of the domain.

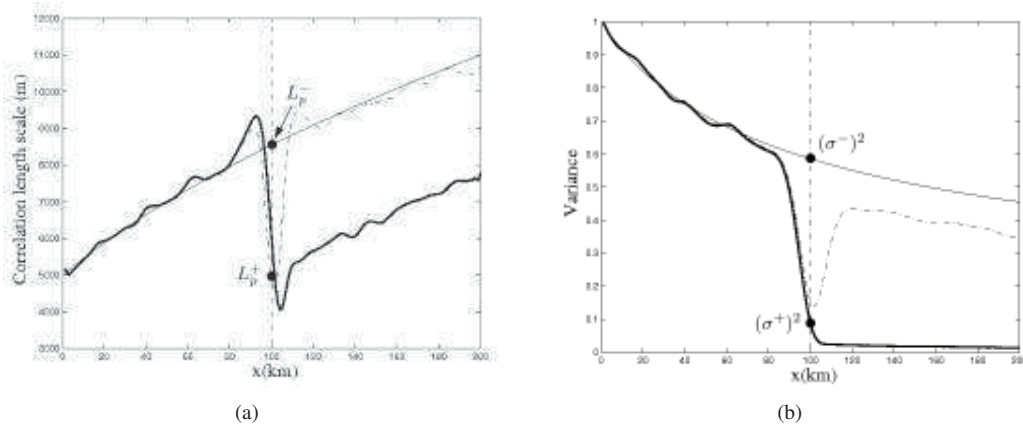


Figure 4. a- Background error correlation length scale $L_p(x)$ and b- Variance σ^2 , in theory without assimilation (thin solid line), for $\mathbf{B}_{\text{EnKF}}$ (thick solid line), for EnBLUE_{B_e} (thin dashed line) and for EnBLUE_{B_{EnKF}} (thick dashed line). As $\mathbf{B}_{\text{EnKF}}$ and EnBLUE_{B_{EnKF}} provide the same results for the variance and the correlation length scales the corresponding curves overlap.

The variances of the background error covariance matrix are presented in Fig. 4-b. The standard deviations at the observation point before assimilation and after assimilation are respectively $\sigma^- = \sqrt{\mathbf{B}_e(x_{\text{obs}}, x_{\text{obs}})}$ and $\sigma^+ = \sqrt{\mathbf{B}_{\text{EnKF}}(x_{\text{obs}}, x_{\text{obs}})}$. The error variance is significantly reduced at the observation point and beyond with the EnKF (thick solid line), compared to the initially prescribed variances (thin solid line). However, when $\mathbf{B}$ is kept invariant and isotropic (EnBLUE_{B_e}), the reduction of the variance is only located in the close neighboring of the observation point (thin dashed line); the invariant matrix is not optimal. In this case, the merits of using a DA algorithm that evolves the background error statistics with the dynamics are demonstrated; the shorten length scale of $\mathbf{B}_{\text{EnKF}}$ prevents from overcorrecting downstream of the observation once information from the observation is taken into account. The EnBLUE_{B_{EnKF}} algorithm shows the same results (thick dashed line that overlaps the thick solid line) as the EnKF.

Since the analysis for the different members in the EnBLUE_{B_{EnKF}} are independent, these results demonstrate that the computation of the EnKF converged background covariance matrix can be achieved at first and then used with the EEnKF with a single analysis algorithm such as BLUE with a much reduced computational cost, thus emulating the EnKF. This result is of particular interest in the framework of real-time forecasting where a single analysis is usually carried out instead of an ensemble of analysis (what we are looking for is the forecasted WLA state and not its covariance matrix). Assuming that an EnKF analysis has previously been carried out, the real-time data assimilation procedure can be achieved with a non expensive EEnKF algorithm, using $\mathbf{B}_{\text{EnKF}}$ as the invariant background error covariance matrix for a single EEnKF analysis. Indeed the emulated EnKF presented here reduces the computational

cost of the EnKF by a factor ranging from $O(100)$ to $O(10000)$. This factor depends on the expected accuracy (that closely relates to the number of members used to run the EnKF) and the number of processors used to run the EnKF. In this study 10000 members are used for the EnKF to ensure that the sampling noise of the background error covariance matrix is less than $10^{-2}m^2$, see Sect. (2.2.2).

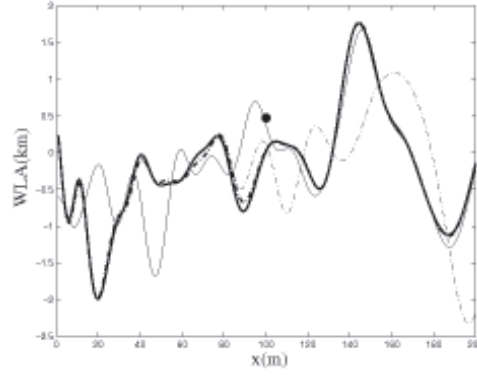


Figure 5. WLA for the EnKF analysis (thick solid line for h^a_{EnKF}), the single BLUE analysis with $\mathbf{B}_e$ (thin dashed line for $h^a_{\text{BLUE}, \mathbf{B}_e}$) and the single EEnKF analysis with $\mathbf{B}_{\text{EnKF}}$ (thick solid line for $h^a_{\text{BLUE}, \mathbf{B}_{\text{EnKF}}}$). The observation is denoted by a black dot, the true state h^{true} by a thin solid line.

This approach is illustrated in Fig. 5 where the improvement of the WLA for a single BLUE analysis (with $\mathbf{B} = \mathbf{B}_e$) and for a single EEnKF is shown: the analyzed state h^a for the EEnKF (thick dashed line) is significantly closer to the true state h^{true} (thin solid line) than the analyzed state for the BLUE using $\mathbf{B}_e$ (thin dashed line).

4.2. Influence of the observation error standard deviation on the correlation length scale and the variance

Both the reduction of the background error variance and the correlation length scale depend on the ratio $r = \frac{\sigma^-}{\sigma_o}$. In section (4.1) σ_o has been chosen so that $r = 3$. In the following σ^- is assumed to be fixed and σ_o varies to represent different observation error statistics so that r ranges from 0.5 to 4. Fig. 6 shows the diagnosed correlation length scale (a), correlation functions at x_{obs} (b) and variance over the domain (c) with the EnKF for different values of the ratio r . When $r = 0.5$, the observation error standard deviation is large (the observations are not reliable), the correlation function at the observation point is close to the initial isotropic one ($\frac{L_p^+}{L_p^-} \simeq \frac{1}{2}$), the reduction of variance is small, hence the analyzed WLA remains close to the background. On the contrary when $r = 4$ the observation error standard deviation is small (the observations are reliable), the correlation function at the observation point evolves into an anisotropic function ($\frac{L_p^+}{L_p^-} \ll 1$), the reduction of variance at the observation point and downstream of that point is significant, hence the assimilation provides good results and the analyzed WLA is brought closer to the true state. Figures 6-(a) and (b) show that when the observation error decreases (r increases), the ratio $\frac{L_p^+}{L_p^-}$ decreases (the anisotropy of the correlation function at the observation point increases). Similarly, Fig. 6-(c) shows that when the observation error decreases (r increases), the ratio $\frac{\sigma^+}{\sigma^-}$ decreases. It should be noted that the data assimilation algorithm leads to a reduction of the variance inside the interval $\mathbf{I}_\varepsilon = [x_{\text{obs}} - \varepsilon^-; x_{\text{obs}} + \varepsilon^+]$ (where $\varepsilon^- = 2.L_p^-$ and $\varepsilon^+ = 2 \times \max_{r \in [0.5; 4]} \{L_p^+\}$). In order to describe the variance outside this interval downstream of the observation point with Eq. (13), the variance in $x_{\text{obs}} + \varepsilon^+$ should be estimated and used in place of $\sigma^2(0)$. In the following, for $x \in \mathbf{I}_\varepsilon$, the ratio $\frac{\sigma^+(x)}{\sigma^-(x)}$ is defined with $\sigma^-(x) = \sqrt{\mathbf{B}_e(x, x)}$ and $\sigma^+(x) = \sqrt{\mathbf{B}_{\text{EnKF}}(x, x)}$.

A set of EnKF experiments is achieved for $r \in [0.5; 4]$ by 0.5 increments of r . It is shown that the relation between r and $\frac{L_p^+}{L_p^-}$ at the observation point can be described with an abacus built from a linear regression in logarithmic scales, represented in dashed lines on Fig. 7-(a), while the results of the EnKF are represented with thick solid lines:

$$\ln\left(\frac{L_p^+}{L_p^-}\right) = \alpha \ln(r) + \beta. \quad (21)$$

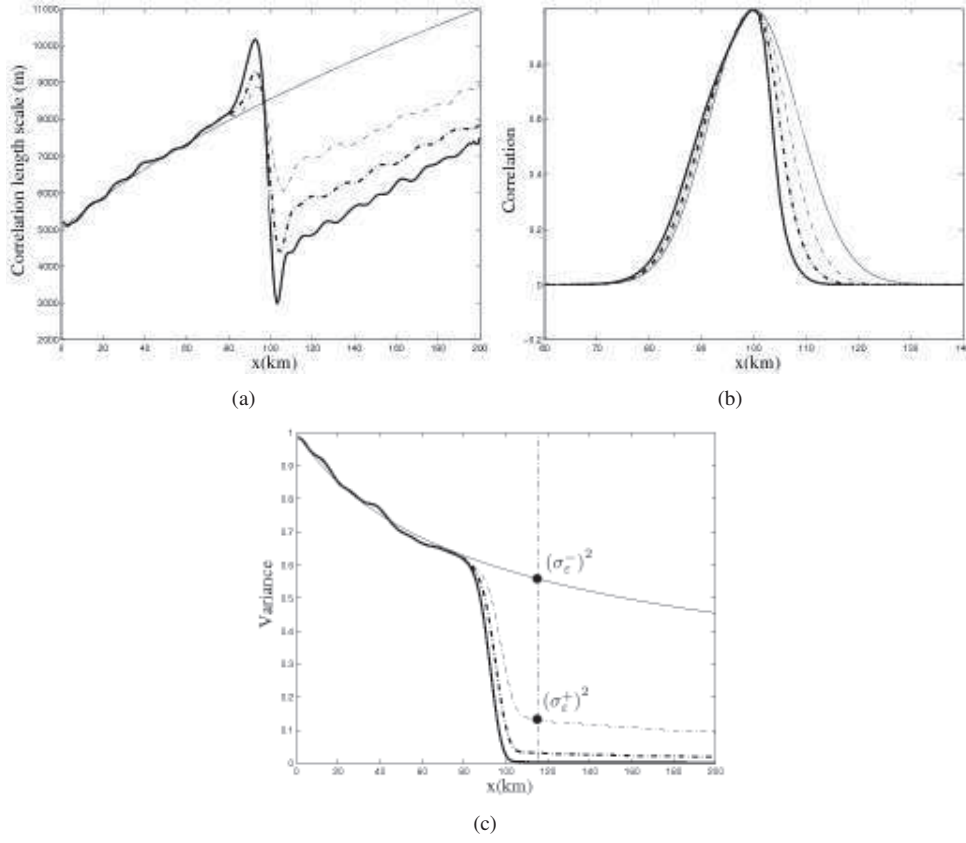


Figure 6. a- Correlation length scale $L_p(x)$ b- Correlation function at x_{obs} c- Variance σ^2 in theory without assimilation (thin solid line), for $r = 0.5$ (thin dashed line), for $r = 1.75$ (thick dashed line) and for $r = 4$ (thick solid line).

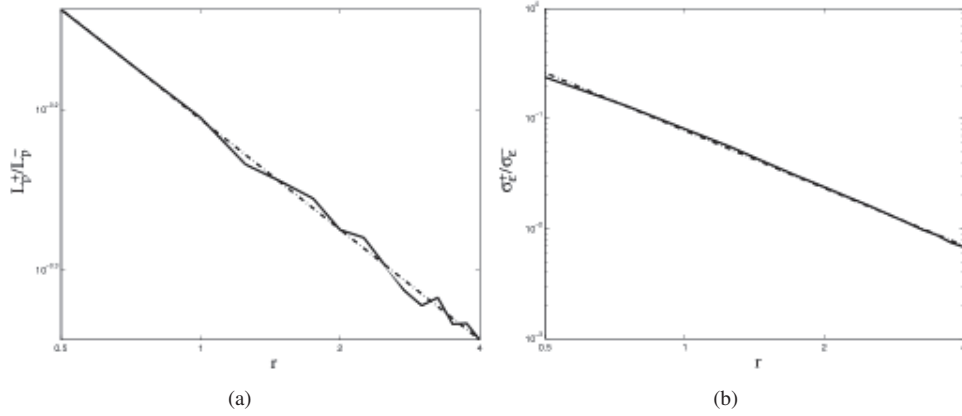


Figure 7. Abacus for (a) $\frac{L_p^+}{L_p^-}$ and (b) $\frac{\sigma_\varepsilon^+}{\sigma_\varepsilon^-}$ as function of the ratio r in solid lines and linear regression in logarithmic scales in dashed lines.

Similarly, for any $x \in \mathbf{I}_\varepsilon$, the relation between r and $\frac{\sigma_\varepsilon^+(x)}{\sigma_\varepsilon^-(x)}$ can be described with an abacus built from a linear regression in logarithmic scales using the same set of EnKF experiments. The linear regression is shown in Fig. 7-b

with a thick dashed line, for $x = x_{obs} + \varepsilon^+$. In the following, $\sigma^-(x_{obs} + \varepsilon^+)$ and $\sigma^+(x_{obs} + \varepsilon^+)$ are respectively noted σ_ε^+ and σ_ε^- and are represented in Fig. 6-(c) for $r = 0.5$:

$$\ln\left(\frac{\sigma_\varepsilon^+}{\sigma_\varepsilon^-}\right) = \gamma \ln(r) + \delta. \quad (22)$$

Abacus in Eqs. (21-22) have been established for $x_{obs} = 100$ but numerical experiments show that the coefficients α, β, γ and δ do not depend on the position of the observation point x_{obs} . These equations allow for the quantification of the impact of the assimilation and the dynamics on the background correlation length scale and variance. More importantly, they lead to the parametrization of the correlation length over the whole domain (see Sect. (4.3)) allowing for the modeling of the converged background error covariance matrix from the EnKF. Using this parametrization, the EnKF can be emulated for any observation network described by the number of observations, their locations and the variance of their respective error (see Sect. (4.4)).

4.3. Parametrized model for correlation length scale and variance reduction at the observation point

This section presents how Eqs. (21-22-23-24) provide a parametrized model for the background error correlation length scale $L_p(x)$ and variance $\sigma^2(x)$, for any observation error variance σ_o^2 . These information are used to model the converged background error covariance matrix from the EnKF using the diffusion operator presented in Sect. (3.3) in order to integrate the EEnKF. While the methodology is applicable for any $r \in [0.5; 4]$, illustrations are given here for $r = 0.75$.

Away from the observation point the expression of the correlation length scale $L_p(x)$ is derived from the analytical expression in Eq. (12):

$$L_p(x) = \sqrt{L_p^2(0) + 4\kappa \frac{x}{c}}, \quad x \leq x_{obs} \quad \text{and} \quad L_p(x) = \sqrt{(L_p^+)^2 + 4\kappa \frac{x - x_{obs}}{c}}, \quad x > x_{obs} \quad (23)$$

where L_p^+ is the downstream correlation length scale at the observation point computed using the abacus from Eq. (21) with $L_p^- = \sqrt{L_p^2(0) + 4\kappa \frac{x_{obs}}{c}}$. Figure 8-(a) illustrates how the correlation length scale obtained from the EnKF (thick solid line) compares with the parametrized one (thick dashed line). It should be noted that here, the EnKF is run for validation purpose only.

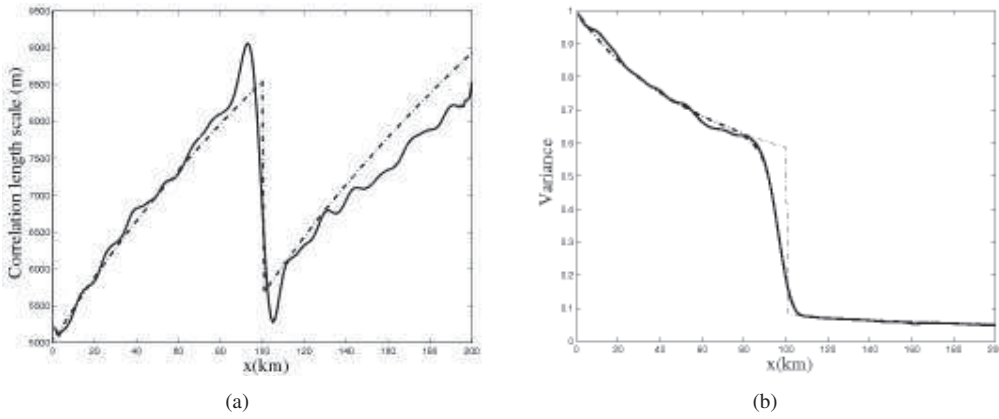


Figure 8. (a) Background error correlation length scale and (b) variance computed with the EnKF (thick solid lines) and with the parametrized model for $r = 0.75$, using 25 points in $\mathbf{I}_\varepsilon$ (thick dashed line) or 2 points (thin dashed line) for the description of the variance reduction

The variance $\sigma^2(x)$ away from the observation point is derived from the analytical expression in Eq. (13):

$$\begin{aligned}\sigma^2(x) &= \sigma(0)^2 \frac{L_p(0)}{\sqrt{L_p^2(0) + 4\kappa \frac{x}{c}}}, & x \leq x_{obs} - \varepsilon^- \\ \sigma^2(x) &= (\sigma_\varepsilon^+)^2 \frac{L_p(x_{obs} + \varepsilon^+)}{\sqrt{L_p^2(x_{obs} + \varepsilon^+) + 4\kappa \frac{x - (x_{obs} + \varepsilon^+)}{c}}}, & x > x_{obs} + \varepsilon^+\end{aligned}\quad (24)$$

where $L_p(x_{obs} + \varepsilon^+)$ is computed with Eq. (23) and $(\sigma_\varepsilon^+)^2$ is the background error variance at $x_{obs} + \varepsilon^+$ computed using the abacus from Eq. (22). In the neighboring of the observation, the reduction of variance is computed from the application of the Eq. (22) for $x \in \mathbf{I}_\varepsilon$. Depending on how many points in $\mathbf{I}_\varepsilon$ are used to derive linear regressions as shown in Eq. (22), the reduction of variance is more or less finely described. Figure 8-(b) illustrates how the variance obtained from the EnKF compares with the parametrized variance using only 2 points in $\mathbf{I}_\varepsilon$ (thin dashed line) or 25 points (thick dashed line that almost overlaps the thick solid line). It is shown that when a crude approximation for the variance in the neighboring of the observation point is used, the results of the EEnKF are degraded.

4.4. Application to any observation network

The parametrized model for background error correlation length scale and variance is validated for DA experiments with a single observation point and various observation error standard deviations. Here the results are extended to an observing network with two observation points $x_{obs,1} = 50$ and $x_{obs,2} = 150$ with respective observation error variance $\sigma_{o,1}^2$ and $\sigma_{o,2}^2$. The distance between $x_{obs,1}$ and $x_{obs,2}$ is bigger than the background error correlation length scale diagnosed previously. In order to estimate the correlation length scale and variance in the neighboring of $x_{obs,2}$ and downstream, the impact of the assimilation at $x_{obs,1}$ should be taken into account.

The parametrized correlation length scale $L_p(x)$ for this observation network is given by:

$$\begin{aligned}L_p(x) &= \sqrt{L_p^2(0) + 4\kappa \frac{x}{c}}, & x \leq x_{obs,1} \\ L_p(x) &= \sqrt{(L_{p1}^+)^2 + 4\kappa \frac{x - x_{obs,1}}{c}}, & x_{obs,1} < x \leq x_{obs,2} \\ L_p(x) &= \sqrt{(L_{p2}^+)^2 + 4\kappa \frac{x - x_{obs,2}}{c}}, & x > x_{obs,2}\end{aligned}\quad (25)$$

where L_{p1}^+ and L_{p2}^+ are the downstream correlation length scales at point $x_{obs,1}$ and $x_{obs,2}$ respectively. L_{p1}^+ and L_{p2}^+ are computed using the abacus from Eq. (21) with the upstream correlation length scales $L_{p1}^- = \sqrt{L_p^2(0) + 4\kappa \frac{x_{obs,1}}{c}}$ and $L_{p2}^- = \sqrt{(L_{p1}^+)^2 + 4\kappa \frac{x_{obs,2} - x_{obs,1}}{c}}$.

The parametrized variance $\sigma^2(x)$ is given by:

$$\begin{aligned}\sigma^2(x) &= \sigma(0)^2 \frac{L_p(0)}{\sqrt{L_p^2(0) + 4\kappa \frac{x}{c}}}, & x \leq x_{obs,1} - \varepsilon^- \\ \sigma^2(x) &= (\sigma_{\varepsilon,1}^+)^2 \frac{L_p(x_{obs,1} + \varepsilon^+)}{\sqrt{L_p^2(x_{obs,1} + \varepsilon^+) + 4\kappa \frac{x - (x_{obs,1} + \varepsilon^+)}{c}}}, & x_{obs,1} + \varepsilon^+ < x \leq x_{obs,2} - \varepsilon^- \\ \sigma^2(x) &= (\sigma_{\varepsilon,2}^+)^2 \frac{L_p(x_{obs,2} + \varepsilon^+)}{\sqrt{L_p^2(x_{obs,2} + \varepsilon^+) + 4\kappa \frac{x - (x_{obs,2} + \varepsilon^+)}{c}}}, & x > x_{obs,2} + \varepsilon^+\end{aligned}\quad (26)$$

where $L_p(x_{obs,1} + \varepsilon^+)$ and $L_p(x_{obs,2} + \varepsilon^+)$ are computed using the abacus Eq. (25) and where $\sigma_{\varepsilon,1}^+$ and $\sigma_{\varepsilon,2}^+$ are computed using Eq. (22) with respectively $\sigma_{\varepsilon,1}^-$ and $\sigma_{\varepsilon,2}^-$ described from Eq. (13).

Thus the background error correlation length scale and variance are fully parameterized as shown in Fig. 9 and are very close to the correlation length scale and variance diagnosed with the EnKF (here again, run is validation purpose only). The parametrized $L_p(x)$ and $\sigma^2(x)$ are used with the diffusion operator to model the converged background error covariance matrix from the EnKF with two observation points, $\mathbf{B}_{\text{EnKF}}$. The analyzed WLA for the EEnKF algorithm

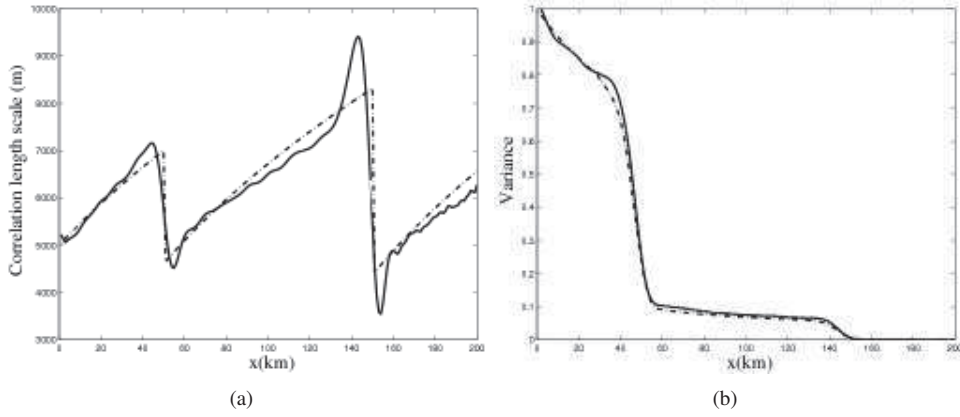


Figure 9. a- Background error correlation length scale $L_p(x)$ and b- variance $\sigma^2(x)$ for the EnKF (solid line) and the parametrized model (dashed line).

using the constant $\tilde{\mathbf{B}}_{\text{EnKF}}$ matrix (only one member to integrate) compares very well with that of the EnKF algorithm as shown in Fig. 10, for a much reduced computational cost, which is compatible with real-time flood-forecasting constraint.

This method can be extended to any observation network building on the characterization of the correlation length scale and variance reduction that established in Eqs. (21-22) in the neighboring of an observation point and the analytical solutions in Eqs. (12-13) away from the observation point.

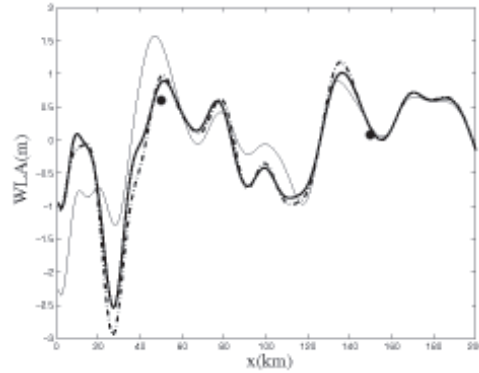


Figure 10. Analyzed WLA for the EnKF (thick solid line) and the EEnKF (thick dashed line). The true state is represented by the thin solid line and the observations by the black dots.

5. Summary and conclusions

This study describes the evolution of the background error covariance matrix with an EnKF algorithm in the framework of OSE, for a steady observation network, meaning that the observation frequency and locations remain the same through the assimilation cycles. It was shown that the filter converges to an optimal and invariant matrix, characterized, at the observation point by an anisotropic correlation function with a shorter correlation length scale downstream of the observation point and a reduction of the error variance. As the model is forced at its upstream boundary with a random variable characterized by a gaussian correlation function over time, the background error correlation length scale and the background error variance away from the observation points are described by analytical equations. The correlation length scale and variance reduction at the observation point are described by abacus as

functions of the observation error variance. The construction of such abacus requires the integration of a very small set of EnKF experiments and can be used for any observation point. Thus a parametrization of the background error correlation length scale and variance is proposed over the entire simulation domain for any observation network given the number of observations, their locations and their respective error variance.

The parametrized model was then used to build the invariant matrix using a diffusion operator, which is a convenient tool especially for large dimension problems. This methods allows to emulate the EnKF at a much reduced computational cost with a deterministic BLUE algorithm where the background error covariance matrix does not evolve in time. It was shown that the resulting algorithm, denoted by EEnKF (for Emulated EnKF), leads to similar results to the EnKF allowing for the use of DA for real-time flood forecasting.

A perspective for this work is to study how the background error statistics evolve with the full shallow-water equations instead of the flood wave propagation model. This would give a closer idea of what to expect with an operational hydraulic model such as MASCARET, MIKE or LISFLOOD. In this context, it is expected that the impact of the assimilation would also spread upstream of the observation points thus leading to the reduction of the background error correlation length scale and variance on both side of the observing stations. Whether the resulting correlation would be isotropic or not, and to what extent still need to be investigated.

Appendix A

Let us consider the advection-diffusion equation :

$$\begin{cases} \partial_t h + c \partial_x h = \kappa \partial_x^2 h & (x, t) \in [0, L] \times \mathbb{R}^+ \\ h(x, 0) = h_0(x) & x > 0 \\ h(0, t) = h_{up}(t) & t > 0 \\ \partial_t h + c \partial_x h = 0 & t > 0 \end{cases} \quad (.1)$$

Let us denote $\Delta x = \frac{L}{N}$ the space step and Δt the time step. Let us denote also h_j^i the value of the discrete solution of (.1) at point $(i\Delta t, j\Delta x)$. Using the finite difference method one can write the following numerical scheme for equation (.1) :

$$\begin{cases} \frac{h_j^{i+1} - h_j^i}{\Delta t} + c \frac{h_{j+1}^i - h_{j-1}^i}{2\Delta x} = \kappa \frac{h_{j+1}^i - 2h_j^i + h_{j-1}^i}{\Delta x^2} & j = 2, \dots, N-1 \\ h_j^0 = h_0(j\Delta x) & j = 1, \dots, N \\ h_j^i = q(i\Delta t) & j = 1 \\ \frac{h_j^{i+1} - h_j^i}{\Delta t} + c \frac{h_j^i - h_{j-1}^i}{\Delta x} = 0 & j = N \end{cases} \quad (.2)$$

The scheme (.2) allows for the numerical resolution of the advection-diffusion equation with an error. Now let us demonstrate why the numerical solution of (.2) is a solution of the equation (.5). Using Taylor expansion one can write :

$$\begin{aligned} \frac{u_j^{i+1} - u_j^i}{\Delta t} &= (\partial_t u)_j^i + \frac{\Delta t}{2} (\partial_t^2 u)_j^i + O(\Delta t^2) \\ \frac{u_{j+1}^i - u_{j-1}^i}{2\Delta x} &= (\partial_x u)_j^i + \frac{\Delta x^2}{6} (\partial_x^3 u)_j^i + O(\Delta x^3) \\ \frac{u_{j+1}^i - 2u_j^i + u_{j-1}^i}{\Delta x^2} &= (\partial_x^2 u)_j^i + \frac{\Delta x^2}{12} (\partial_x^4 u)_j^i + O(\Delta x^3) \end{aligned} \quad (.3)$$

Thereafter to lighten the notations we note u instead of u_j^i the value of u at point $(i\Delta t, j\Delta x)$. It comes from (.2) and (.3) that u solves the following equation :

$$\partial_t u + c \partial_x u - \kappa \partial_x^2 u + \underbrace{\frac{\Delta t}{2} \partial_t^2 u + \Delta x^2 \left(\frac{c}{6} \partial_x^3 u - \frac{\kappa}{12} \partial_x^4 u \right)}_{\varepsilon} + O(\Delta t^2, \Delta x^3) = 0 \quad (.4)$$

Using the derivation of the advection-diffusion equation with respect to time one can write the expression of the temporal derivatives $\partial_t^2 u$ with respect to the spatial derivatives : $\partial_t^2 u = c^2 \partial_x^2 u - 2ck \partial_x^3 u + \kappa^2 \partial_x^4 u$ that allows for, using (.4), writing :

$$\frac{\partial v}{\partial t} + c \frac{\partial v}{\partial x} + \underbrace{\left(\frac{c \Delta x^2}{6} - c \kappa \Delta t \right)}_{\mu} \frac{\partial^3 v}{\partial x^3} = \underbrace{\left(\kappa - \frac{c^2 \Delta t}{2} \right)}_{\kappa'} \frac{\partial^2 v}{\partial x^2} - \left(\frac{\kappa^2 \Delta t}{2} - \frac{\kappa \Delta x^2}{12} \right) \frac{\partial^4 v}{\partial x^4} + O(\Delta t^2, \Delta x^3) \quad (.5)$$

Practically, the numerical model solves Eq. (.5) instead of Eq. (.2) inducing spurious dispersion due to the term μ (particularly for high frequencies) and spurious diffusion due to the term κ' (as neither c nor Δt is equal to zero, the numerical diffusion κ' is not equal to the physical diffusion κ). Numerical experiments highlighted that the numerical model can under or over estimate the diffusion and for $\kappa' < 0$ the scheme is unstable because the CFL (Courant-Friedrichs-Lewy) condition is not verified (Quarteroni et al. [12]).

References

- [1] Andreadis, K.M., Clark, E.A., Lettenmaier, D.P. and Alsdorf, D.E.: Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model, *Geophys. Res. Lett.*, 34, L10403, doi:10.1029/2007GL029721, 2007.
- [2] Biancamaria, S., Durand, M., Andreadis, K.M., Bates, P.D., Boone, A., Mognard, N.M., Rodríguez, E., Alsdorf, D.E., Lettenmaier, D.P. and Clark, E.A.: Assimilation of virtual wide swath altimetry to improve Arctic river modeling, *Remote Sens. Environ.*, vol. 115, no. 2, pp. 373–381, 2011.
- [3] Bouttier, F.: The dynamics of error covariances in an assimilation system. *Tellus A*, 45, 408–423, 1993.
- [4] Bouttier, F.: A dynamical estimation of error covariances in an assimilation system. *Mon. Wea. Rev.*, 122, 2376–2390, 1994.
- [5] Daley, R.: *Atmospheric data analysis*. Cambridge atmospheric and space science series. Cambridge University Press, 1991.
- [6] Drogue, G., Pfister, L., Leviandier, T., El Idrissi, A., Iffly, J.-L., Matgen, P., Humbert, J., and Hoffmann, L.: Simulating the spatio-temporal variability of streamflow response to climate change scenarios in a mesoscale basin, *J. Hydrol.*, 293, 255–269, 2004.
- [7] Durand, M., Lee-Lang, F., Lettenmaier, D.P., Alsdorf, D.E., Rodriguez, E., and Esteban-Fernandez, D.: The Surface Water and Ocean Topography Mission: Observing Terrestrial Surface Water and Oceanic Submesoscale Eddies, *IEEE*, 98, 766–779, 2010.
- [8] Evensen, G.: *Data Assimilation - The Ensemble Kalman Filter*, Springer, 2009.
- [9] Fukumori, I., Benveniste, J., Wunsch, C., and Haidvogel, D. B.: Assimilation of sea surface topography into an ocean circulation model using a steady-state smoother. *J. Phys. Oceanogr.*, Vol. 23, Issue 8, 1831–1855, 1993.
- [10] Gelb, A.: *Applied optimal estimation*, The MIT press, 1974.
- [11] Jean-Baptiste, N., Malaterre, P.-O., Dorée, C. and Sau, J.: Data assimilation for real-time estimation of hydraulic states and unmeasured perturbations in a 1D hydrodynamic model, *Journal of Mathematics and Computers in Simulation*, Vol. 81, Issue 10, 2201–2214, 2011.
- [12] Giustarini, L., Matgen, P., Hostache, R., Montanari, M., Plaza, D., Pauwels, V. R. N., De Lannoy, G. J. M., De Keyser, R., Pfister, L., Hoffmann, L., and Savenije, H. H.G.: Assimilation SAR-derived water level data into a hydraulic model: a case study, *Hydrol. Earth Syst. Sci.*, 8, 2103–2144, 2011.
- [13] Hendricks, F.H.J., and Kinzelbach, W.: Real-time groundwater flow modeling with the Ensemble Kalman Filter: Joint estimation of states and parameters and the filter inbreeding problem, *Water Resour. Res.*, 44, W09408, doi:10.1029/2007WR006505, 2008.
- [14] Madsen, H., and Skotner, C.: Adaptive state updating in real-time river flow forecasting - a combined filtering and error forecasting procedure, *J. Hydrol.*, 308, 302–312, 2005.
- [15] Li, J., and Xiu, D.: A generalized polynomial chaos based ensemble Kalman filter with high accuracy, *J. Comput. Phys.*, 228, 5454–5469, 2009.
- [16] Malaterre, P.-O., Baume, J.-P., Jean-Baptiste, N., and Sau, J.: Calibration of open channel flow models: a system analysis and control engineering approach, *SimHydro 2010: Hydraulic modeling and uncertainty*, 2–4 June 2010, Sophia Antipolis.
- [17] Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Guingla, D.P., Pauwels, V., De Lannoy, G., De Keyser, R., and Savenije, H.H.G.: Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the Particle Filter: proof of concept, *Hydrol. Earth Syst. Sci.*, 14, 1773–1785, 2010.
- [18] Mirouze, I., and Weaver, A.T.: Representation of correlation functions in variational assimilation using an implicit diffusion operator, *Q. J. R. Meteorol. Soc.*, 136, 1421–1443, DOI:10.1002/qj.643, 2010.
- [19] Moradkhani, H., Hsu, K.-L., Gupta, H.V., and Sorooshian, S.: Improved streamflow forecasting using self-organizing radial basis function artificial neural networks, *J. Hydro.*, 295, 246–262, 2004.
- [20] Moradkhani, H., Sorooshian, S., Gupta, H.V., and Houser, P.-R.: Dual state-parameter estimation of hydrological models using ensemble Kalman filter, *Advances in Water Resources*, 28, 135–147, 2005a.
- [21] Moradkhani, H., Hsu, K.-L., Gupta, H.V., and Sorooshian, S.: Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, *Water Resour. Res.*, 41, W05012, doi:10.1029/2004WR003604, 2005b.
- [22] Neal, J. C., Atkinson, P. M., and Hutton, C. W.: Flood inundation model updating using an ensemble Kalman filter and spatially distributed measurements, *J. Hydro.*, Volume 336, Issue 3, 401–415, 2007.
- [23] Neal, J., Schumann, G., Bates, P., Buytaert, W., Matgen, P., and Pappenberger, F.: A data assimilation approach to discharge estimation from space, *Hydrol. Process.*, 23, 3641–3649. doi: 10.1002/hyp.7518, 2009.
- [24] Pannekoucke O., Berre, L., and Desroziers, G.: Background-error correlation length scale estimates and their sampling statistics, *Q. J. R. Meteorol. Soc.*, 134, 497–508, 2008.
- [25] Pannekoucke, O., and Massart, S.: Estimation of the local diffusion tensor and normalization for heterogeneous correlation modelling using a diffusion equation, *Q. J. R. Meteorol. Soc.*, 134, 1425–1438, 2008.
- [26] Pappenberger, F., Bevena, K., Horritt, M., and Blazkovic, S.: Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations, *J. Hydrol.*, 302, 46–49, 2005.

- [27] Ricci, S., Piacentini, A., Thual, O., Le Pape, E., and Jonville, G.: Correction of upstream flow and hydraulics state with data assimilation in the context of flood forecasting, *Hydrol. Earth Syst. Sci.*, Vol. 15, 1–21, 2011.
- [28] Schumann, G., Bates, P.D., Horritt, M.S., Matgen, P., and Pappenberger, F.: Progress in integration of remote sensing-derived flood extent and stage data and hydraulic models, *Rev. Geophys.*, 47, RG4001, doi:10.1029/2008RG000274, 2009.
- [29] Shiiba, M., Laurenson, X., and Tachikawa, Y.: Real-time stage and discharge estimation by a stochastic-dynamic flood routing model. *Hydrol. Process.*, Vol. 14, Issue 3, 481–495, 2000.
- [30] Szunyogh, I., Kostelich, E., Gyarmati, G., Kalnay, E., Hunt, B., Ott, E., Satterfield, E., and Yorke, J.: A local ensemble transform Kalman filter data assimilation system for the NCEP global model, *Tellus A*, 60, 113–130, 2008.
- [31] Tippett, M.K., Anderson, J.L., Bishop, C. H., Hamill, T.M., and Whitaker, J. S.: Ensemble Square Root Filters, *Mon. Wea. Rev.*, 131, 1485–1490, 2003.
- [32] Thirel, G., Martin, E., Mahfouf, J.-F., Massart, S., Ricci, S., and Habets, F.: A past discharges assimilation system for ensemble streamflow forecasts over France, *Hydrol. Earth Syst. Sci.*, 14, 1623–1637, 2010.
- [33] Valstar, J.R., McLaughlin, D.B., te Stroet, C.B.M., and van Geer, F.C.: A representer-based inverse method for groundwater flow and transport applications, *Water Resour. Res.*, 40, W05116, doi:10.1029/2003WR002922, 2004.
- [34] Weaver, A., and Courtier, P.: Correlation modelling on the sphere using a generalized diffusion equation, *Q. J. R. Meteorom. Soc.*, 127, 1815–1846, 2001.
- [35] Weaver, A.T., Deltel, C., Machu, E. , Ricci, S., and Daget, N.: A multivariate balance operator for variational ocean data assimilation, *Q. J. R. Meteorol. Soc.*, 131, 3605–3625, 2005.
- [36] Weaver, A.T., and Mirouze, I.: On the diffusion equation and its application to isotropic and anisotropic correlation modelling in variational assimilation, *Q. J. R. Meteorol. Soc.*, DOI:10.1002/qj.1955, 2012.
- [37] Weerts, A.H., and El Serafy, G.Y.H.: Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models, *Water Resour. Res.*, 42, W09403, doi:10.1029/2005WR004093., 2006.
- [38] Weerts, A.H., El Serafy, G.Y. , Hummel, S., Dhondia, J., and Gerritsen, H.: Application of generic data assimilation tools (DATools) for flood forecasting purposes, *Computer & Geosciences*, Vol. 36, Issue 4, 453–463, 2010.

7.3 Conclusion

Dans ce travail nous nous sommes livrés à l'étude des fonctions de covariance et de corrélation sur un modèle d'onde de crues 1D avec et sans assimilation. Les fonctions de corrélation à travers le diagnostic des longueurs de portée fournissent une information sur l'extension spatiale des fonctions de covariance.

Dans le cas sans assimilation un signal imposé à l'amont avec une fonction d'auto-corrélation temporelle gaussienne est propagé le long du domaine en un signal avec une fonction de corrélation spatiale gaussienne.

Dans le cas avec assimilation, on montre que celle-ci, cycle après cycle, impacte le signal de telle façon qu'une fonction de corrélation au point d'observation initialement gaussienne est transformée en une fonction asymétrique avec une longueur de portée plus petite à l'aval qu'à l'amont justifiant l'asymétrie des fonctions de covariance utilisées par [Ricci et al., 2011]. L'utilisation des fonctions de covariance stationnaires diagnostiquées par l'EnKF et utilisées dans un filtre de Kalman invariant montre que ces deux algorithmes donnent des résultats équivalents mais pour un coût de calcul moindre dans le cas du filtre de Kalman invariant.

L'étude de l'impact de l'assimilation de données sur les longueurs de portée et la variance ont permis d'établir une paramétrisation des fonctions de covariance diagnostiquées par l'EnKF au point d'observation. Ces fonctions permettent d'émuler l'EnKF avec un filtre de Kalman invariant mais sans avoir à rejouer l'EnKF. Cela a un intérêt notamment pour la prévision des crues en temps réel où la contrainte sur les temps de calcul est forte et où la chaîne opérationnelle n'a pas été conçue pour gérer des méthodes d'ensemble.

Néanmoins une des limitations de ce travail est l'hypothèse de stationnarité de la variance du forçage amont. Hypothèse qui n'est pas réaliste puisque dans le cas de fortes crues la variance des forçages amont connaît justement une importante variabilité. Pour contourner ce problème une solution serait de mettre à jour, par exemple toutes les heures, les fonctions de covariance aux stations d'observation en fonction de la variance des forçages amonts lors d'un pic de crue. On pourrait également envisager, au début d'un pic de crue, de déterminer une fonction de covariance en fonction de l'intensité supposée de l'épisode.

D'autre part des expériences numériques qui n'ont pas été reportées dans ce travail montrent que des fonctions de covariance symétriques sur un modèle Saint-Venant 1D avec une géométrie réelle fournissent des résultats moins bons que des fonctions de covariance asymétriques. Ce fait suggère que c'est la propriété d'asymétrie des fonctions de covariance qui joue un rôle important plus qu'une paramétrisation très fine de celles-ci en fonction des données du problème. D'ailleurs dans leur utilisation opérationnelle ces fonctions de covariance restent constantes et ne sont pas mises à jour en fonction des variations des forçages.

Enfin il faut rappeler que dans cette étude nous nous sommes placés dans le cas particulier du modèle d'onde crue qui est une approximation des équations de Saint-Venant (voir section 2.3), avec une variance du forçage amont qu'on a supposé constante et sur une géométrie simplifiée : un canal rectiligne à fond plat et pente constante. Or il est fort possible que dans le cas des équations de Saint-Venant avec une géométrie réelle, les fonctions de covariance et de corrélation d'erreur ne soient pas les mêmes que celles

diagnostiquées sur le modèle d'onde de crue et qu'elles aient une dynamique totalement différente. C'est à ce travail et à l'application de l'EnKF sur le réseau hydraulique Adour maritime auxquels nous nous livrons dans le prochain chapitre.

Chapitre 8

Application du filtre de Kalman d'ensemble au réseau hydraulique Adour maritime

Sommaire

8.1	Le coupleur de code OpenPALM, le réseau Adour maritime et l'implémentation des méthodes d'ensemble sur ce réseau	162
8.1.1	Le coupleur de codes OpenPALM	162
8.1.2	Rappels sur le réseau Adour maritime	164
8.1.3	Implémentation des méthodes d'ensemble sur le réseau Adour maritime	166
8.1.4	Génération des membres de l'ensemble	169
8.2	Etude des fonctions de covariance et de corrélation d'erreur d'ébauche sans assimilation	170
8.2.1	Diagnostic des fonctions de covariance et de corrélation	171
8.2.2	Les fonctions de covariance	172
8.2.3	Les fonctions de corrélation	175
8.2.4	Caractérisation et évolution temporelle de l'extension spatiale des fonctions de corrélation ZZ	177
8.3	Application de l'EnKF au réseau Adour maritime : première approche et limitations	181
8.3.1	Une expérience d'assimilation de données de hauteurs d'eau synthétiques	181
8.3.2	Résultats numériques	182
8.3.3	Discussion	189
8.4	Application des critères de Desroziers	191
8.4.1	Estimation des covariances d'erreur d'observation	191
8.4.2	Estimation des covariances d'erreur d'ébauche et inflation	195
8.4.3	Une expérience d'assimilation de données synthétiques	198
8.4.4	Expériences d'assimilation sur des données réelles	201
8.5	Vers des prévisions d'ensemble sur le réseau Adour maritime ?	204

8.5.1	Cadre de l'expérience et démarche	204
8.5.2	Estimation des prévisions d'ensemble	205
8.5.3	Discussion	211

Dans ce chapitre nous nous intéresserons à la mise en place du filtre de Kalman d'ensemble sur le réseau hydraulique Adour maritime. Dans un premier temps présenterons le coupleur de codes OpenPALM que nous avons utilisé pour implémenter les méthodes d'ensemble sur l'Adour avant de donner quelques rappels sur ce réseau et de présenter l'implémentation des méthodes d'ensemble sur ce réseau. Puisqu'elles sont au cœur de tout système d'assimilation de données nous nous intéresserons dans un second temps à l'étude qualitative des fonctions de covariance et de corrélation dans le cadre de l'hydraulique fluviale sans assimilation. Le modèle hydraulique considéré ici étant le modèle des équations de Saint-Venant 1D, nous mettrons ainsi en évidence la différence de forme et la sensibilité à la géométrie de ces fonctions par rapport à l'étude menée au chapitre 7 dans le cadre du modèle d'onde de crue. Ensuite nous nous pencherons sur la mise en place proprement dite du filtre de Kalman d'ensemble sur l'Adour et en montrerons les limitations. Cela nous amènera alors à étudier les solutions à notre disposition pour dépasser ces limitations. Enfin nous nous pencherons sur la question de la validation des prévisions d'ensemble sur l'Adour maritime.

8.1 Le coupleur de code OpenPALM, le réseau Adour maritime et l'implémentation des méthodes d'ensemble sur ce réseau

8.1.1 Le coupleur de codes OpenPALM

Généralités

OpenPALM [Buis et al., 2006] est un coupleur dynamique de codes parallèles peu intrusif développé par le CERFACS et l'ONERA. Cet outil, distribué sous licence LGPL, est utilisé dans de nombreux domaines allant des systèmes opérationnels d'assimilation de données aux couplages de codes massivement parallèles pour la modélisation multiphysique. OpenPALM permet de lancer (de manière simultanée ou séquentielle) et de coupler différents codes de calcul en échangeant différentes informations entre ceux-ci.

Les composantes d'OpenPALM

OpenPALM est composée de quatre parties complémentaires : la librairie PALM, la librairie CWINPI, la fonctionnalité PARASOL et l'interface graphique PrePALM.

La librairie PALM est une librairie permettant l'échange de données entre composants élémentaires via des primitives de haut niveau s'appuyant elles-mêmes sur le standard de communication MPI. Les composants peuvent être aussi bien des codes de calcul industriels comme Mascaret que des sous-routines fortran, C ou C++ développées par l'utilisateur pour ses propres besoins. Cette librairie permet le lancement dynamique (en boucle ou sous condition) des différents composants. Cette librairie facilite l'échange

de données pour les codes parallèles en proposant une redistribution automatique des données entre les codes couplés, elle permet la séparation des composantes modélisant la physique du problème de celles dédiées aux opérations algébriques. Elle offre également la possibilité de combiner en un seul exécutable les différents composants impliqués.

La librairie CWIPI (Coupling With Interpolation Parallel Interface) offre à OpenPALM une fonctionnalité d'interpolation spatiale pour des champs s'appuyant sur des maillages non structurés de type éléments finis. Elle permet la communication parallèle proc à proc de ces données interpolées entre les codes couplés.

La fonctionnalité PARASOL permet de lancer en parallèles plusieurs instances d'un même code de calcul. PARASOL gère l'exécution de ces différentes instances en fonction des ressources de la machine et permet de gérer via des communications PALM des vecteurs de données d'entrée et des vecteurs de résultats des différentes instances de ce code.

PrePALM est l'interface graphique d'OpenPALM. C'est dans PrePALM que l'on peut définir le nombre de composants utilisés, si ceux-ci sont lancés de manière séquentielle ou parallèle. Dans la terminologie PrePALM les composants sont des unités, celles-ci sont lancées via des branches. Plusieurs unités disposées à la suite sur une même branche sont lancées séquentiellement, des unités disposées sur des branches différentes sont lancées en parallèle. Les communications entre les unités se font via des connections mettant en relation les données produites par un code avec celle consommées par un autre code. Lors de la création de l'exécutable OpenPALM les unités sont encapsulées comme sous-programme d'un programme généré par PrePALM. On donne en illustration 8.1-(a) et 8.1-(b) deux figures de l'interface PrePALM.

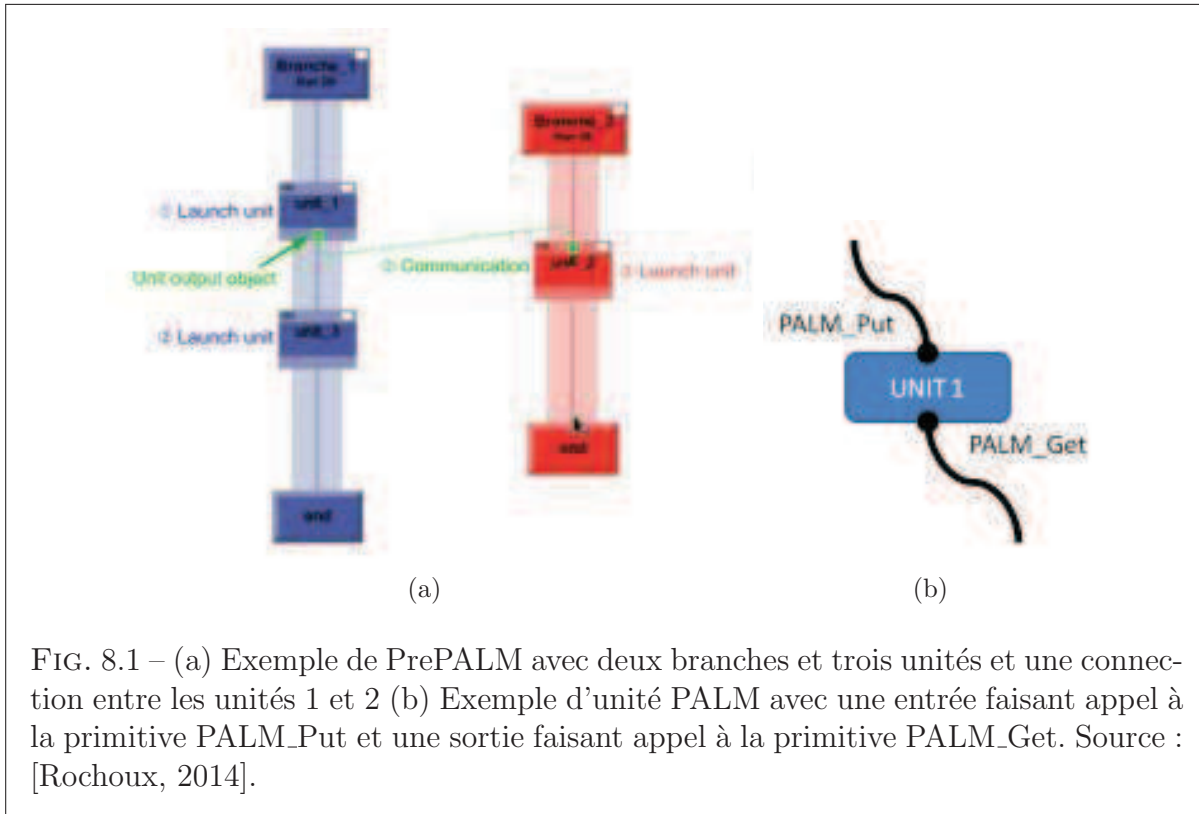
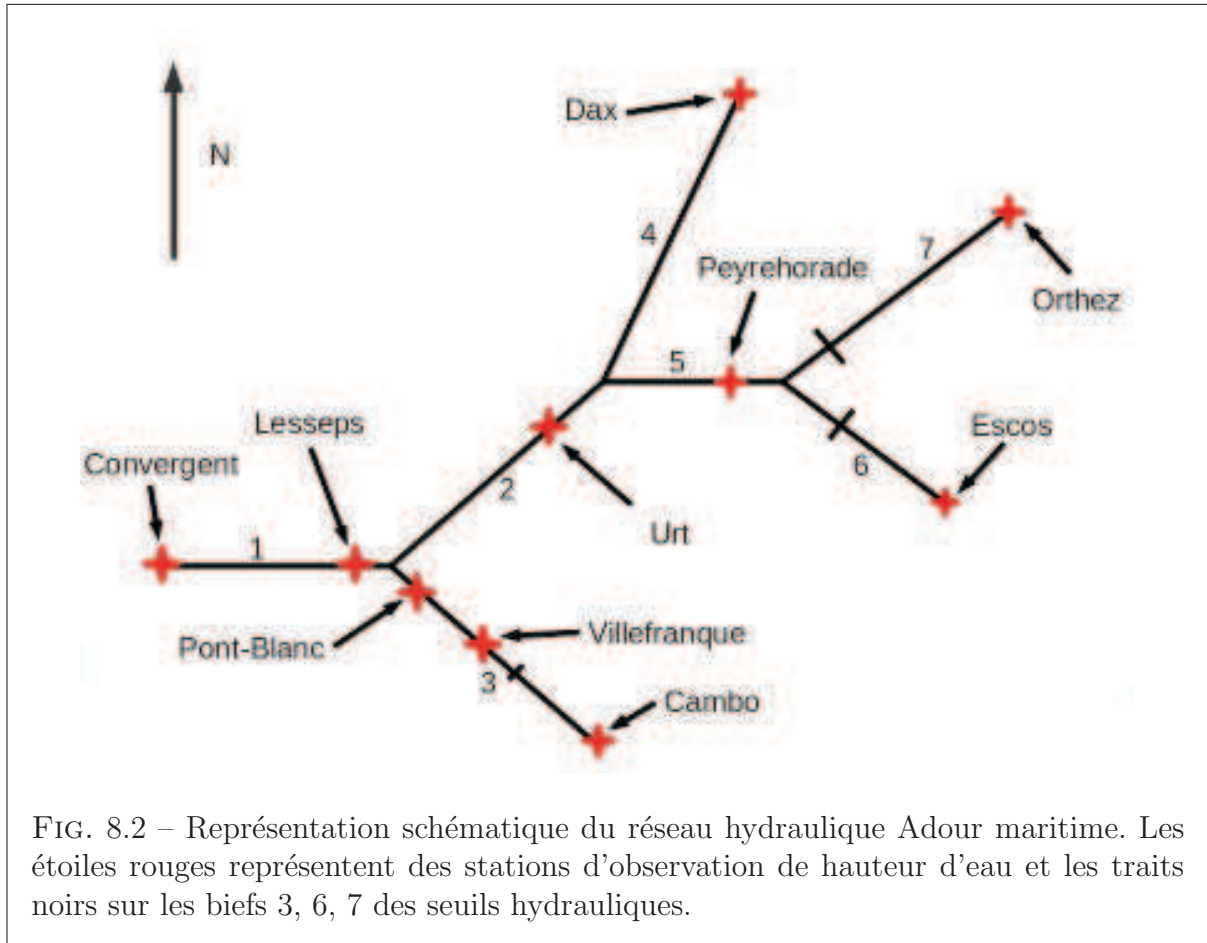


FIG. 8.1 – (a) Exemple de PrePALM avec deux branches et trois unités et une connexion entre les unités 1 et 2 (b) Exemple d'unité PALM avec une entrée faisant appel à la primitive PALM_Put et une sortie faisant appel à la primitive PALM_Get. Source : [Rochoux, 2014].

8.1.2 Rappels sur le réseau Adour maritime

Bien que l'on ait déjà évoqué le réseau hydraulique Adour maritime à la section 1.3.2 il n'est pas inutile de revenir dessus à ce point de notre exposé.

Comme on peut le voir sur la figure 8.2 le réseau Adour maritime est composé de sept biefs avec trois confluences. L'extrémité Ouest du bief 1 correspond à l'embouchure de l'Adour sur l'océan Atlantique. Comme son nom l'indique tout le modèle Adour maritime est sous influence maritime à l'exception des zones situées à l'amont des seuils hydrauliques situés sur les biefs 3, 6 et 7. En effet ces seuils ont été conçus et disposés pour endiguer l'influence maritime sur les deux Gave (biefs 6 et 7) et la Nive (bief 3). Les hauteurs observées à l'amont (aux stations de Dax, Orthez, Escos et Cambo) sont converties en débits via des courbes de tarage pour servir de forçages pour le modèle 1D. Les forçages à l'aval sont donnés par les observations de hauteurs d'eau à la station de Convergent. Sur le réseau Adour nous disposons de cinq stations d'observation (Peyrehorade, Urt, Lesseps, Pont-Blanc et Villefranque) mais seulement celles de Peyrehorade, Urt, Lesseps et Pont-Blanc sont utilisées pour les prévisions opérationnelles. La quasi-totalité du réseau Adour est modélisé sans lit majeur hormis quelques sections de calcul à proximité de la station de Peyrehorade et à l'amont du bief 4, ailleurs le réseau est modélisé avec des berges infiniment hautes. Enfin il faut préciser que dans les conditions opérationnelles il n'y a pas de données prévues sur les forçages amonts, c'est pourquoi



pendant la phase de prévision les forçages amonts sont fixés constants égaux aux dernières valeurs connues. Dans le cas du forçage aval en mode prévision celui-ci est donné par les prévisions du SHOM étant donné que la station de Convergents est à l'embouchure de l'Adour sur l'Atlantique.

Dans la suite de ce travail on s'intéressera à certaines grandeurs comme par exemple les débits, les corrélations spatiales d'erreur sur les hauteurs d'eau... La configuration complexe du réseau Adour ne permet pas de représenter les grandeurs considérées de manière simple sur l'ensemble du domaine. Pour représenter celles-ci nous nous restreindrons à les afficher sur des successions de biefs que l'on appellera « chemin » et qui relie chacun des forçages amonts au forçage aval. On distingue donc quatre chemins composés des biefs suivants : 3-1, 4-2-1, 7-5-2-1, 6-5-2-1. Une illustration des hauteurs d'eau et des débits calculés sur différents chemins est donnée en figure 8.3.

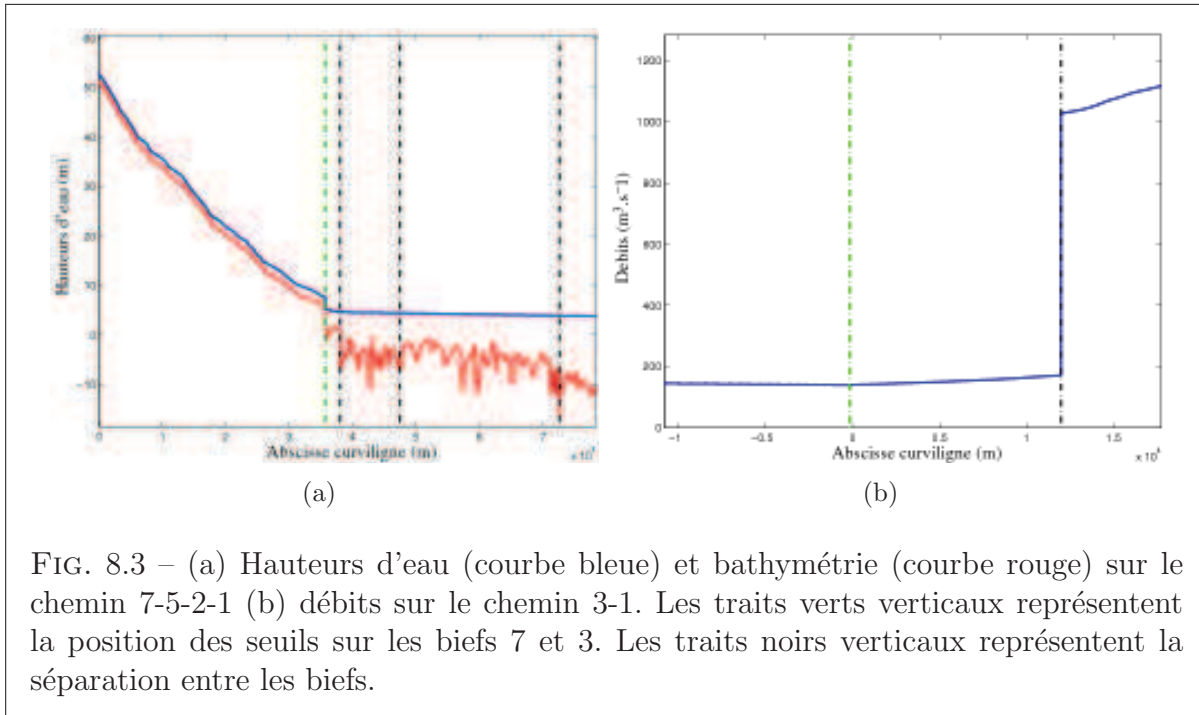


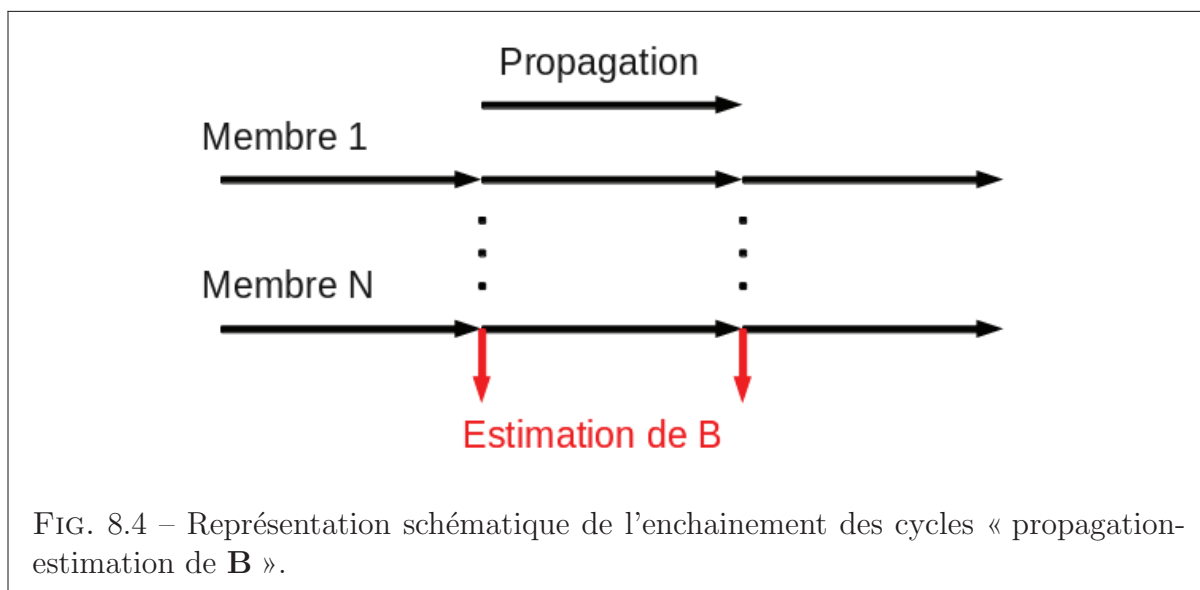
FIG. 8.3 – (a) Hauteurs d'eau (courbe bleue) et bathymétrie (courbe rouge) sur le chemin 7-5-2-1 (b) débits sur le chemin 3-1. Les traits verts verticaux représentent la position des seuils sur les biefs 7 et 3. Les traits noirs verticaux représentent la séparation entre les biefs.

8.1.3 Implémentation des méthodes d'ensemble sur le réseau Adour maritime

Dans ce travail nous avons implémenté une approche ensembliste, que ce soit pour diagnostiquer les fonctions de covariance et de corrélation sans assimilation comme cela est exposé en section 8.2 ou pour corriger hauteurs d'eau et débits avec l'assimilation de données comme cela est exposé en section 8.3 ou 8.4.4.

Avant toute chose nous précisons un terme utilisé dans la suite : nous appellerons *Temps de base* le temps de début de simulation et *Temps de prévision* la durée de la prévision. Dans le cas où nous procédons simplement à l'estimation des fonctions de covariance et corrélation l'approche ensembliste consiste à appliquer itérativement le cycle constitué des étapes suivantes (voir figure 8.4) :

- **propagation** des membres de l'ensemble par le modèle hydraulique Mascaret de *Temps de base* à *Temps de base*+1h;
- **sauvegarde** des membres à *Temps de base*+1h pour le prochain cycle;
- **estimation statistique** de la matrice des covariances d'erreur d'ébauche **B**.



Dans le cas où nous procédons à l'assimilation de données l'approche consiste à appliquer itérativement le cycle constitué des étapes suivantes (voir figure 8.5) :

- **estimation statistique** de la matrice des covariances d'erreur d'ébauche **B** ;
- **assimilation** des observations disponibles ;
- **prévision** : propagation des membres corrigés avec l'assimilation de *Temps de base* à *Temps de base + Temps de prévision* ;
- **sauvegarde** des membres à *Temps de base + 1h* pour le prochain cycle.

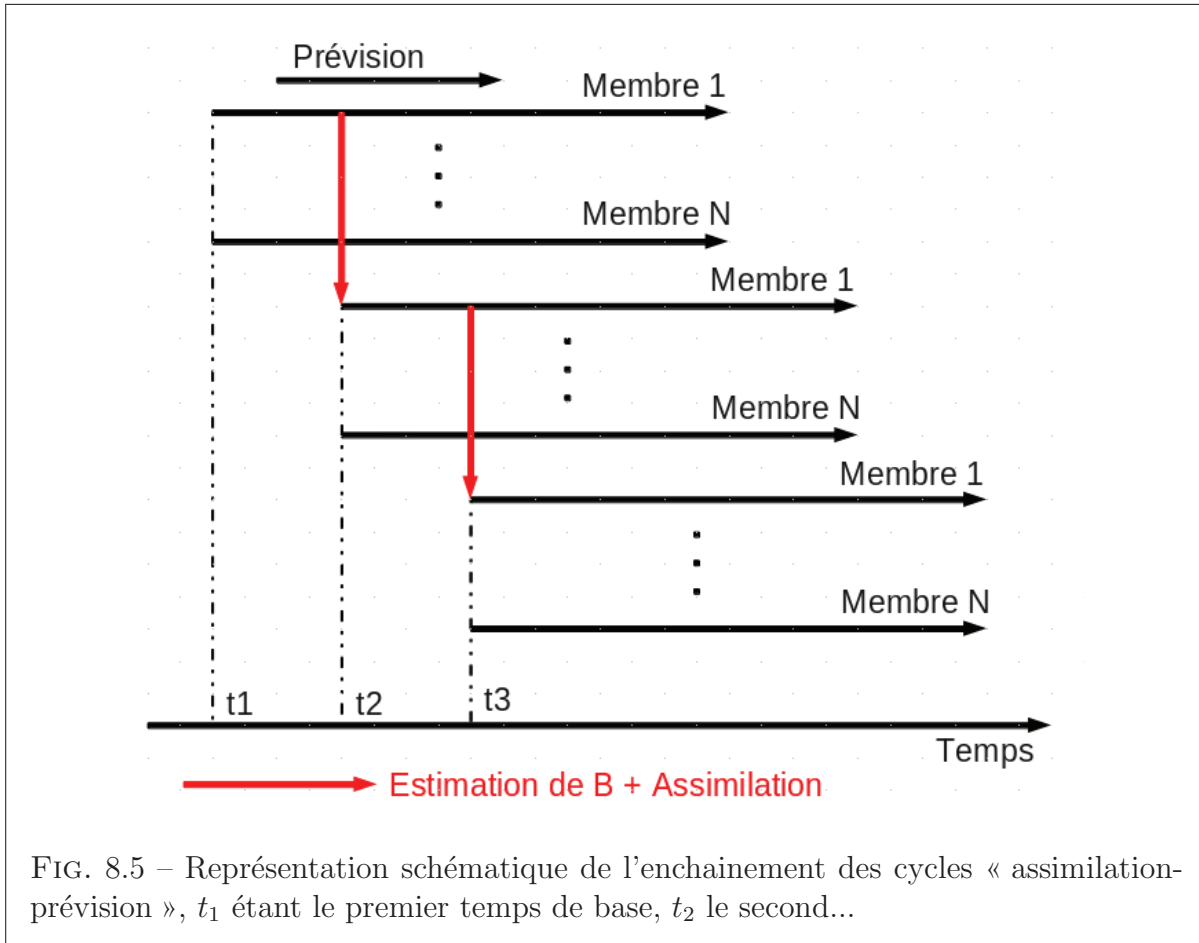


FIG. 8.5 – Représentation schématique de l'enchaînement des cycles « assimilation-prévision », t_1 étant le premier temps de base, t_2 le second...

Du point de vue du code un cycle « propagation-estimation de $\mathbf{B}$ » ou un « assimilation-prévision » est géré par l'exécutible PALM correspondant qui se charge de lancer les différentes unités (unité de lecture des fichiers de données, unité Mascaret de propagation des membres, unités relatives à l'assimilation...) et d'échanger entre elles les données nécessaires (par exemple transmission de la matrice $\mathbf{B}$ de l'unité chargée de la calculer à celle procédant à l'assimilation sur chacun des membres). On a reporté en figure 8.6 une image du schéma PrePALM correspondant à l'EnKF. Pour reproduire l'enchaînement des cycles l'exécutible PALM est lancé de manière itérative par un pilote. Il s'agit d'un script KornShell qui prend en arguments le premier et le dernier temps de base, le temps de prévision ainsi que la fréquence d'assimilation. Le pilote modifie les fichiers de données pour Mascaret et l'assimilation en fonction de ces entrées à chaque itération puis lance l'exécutible PALM. C'est également le pilote qui s'occupe de la création et de la gestion des fichiers et des données de sortie.

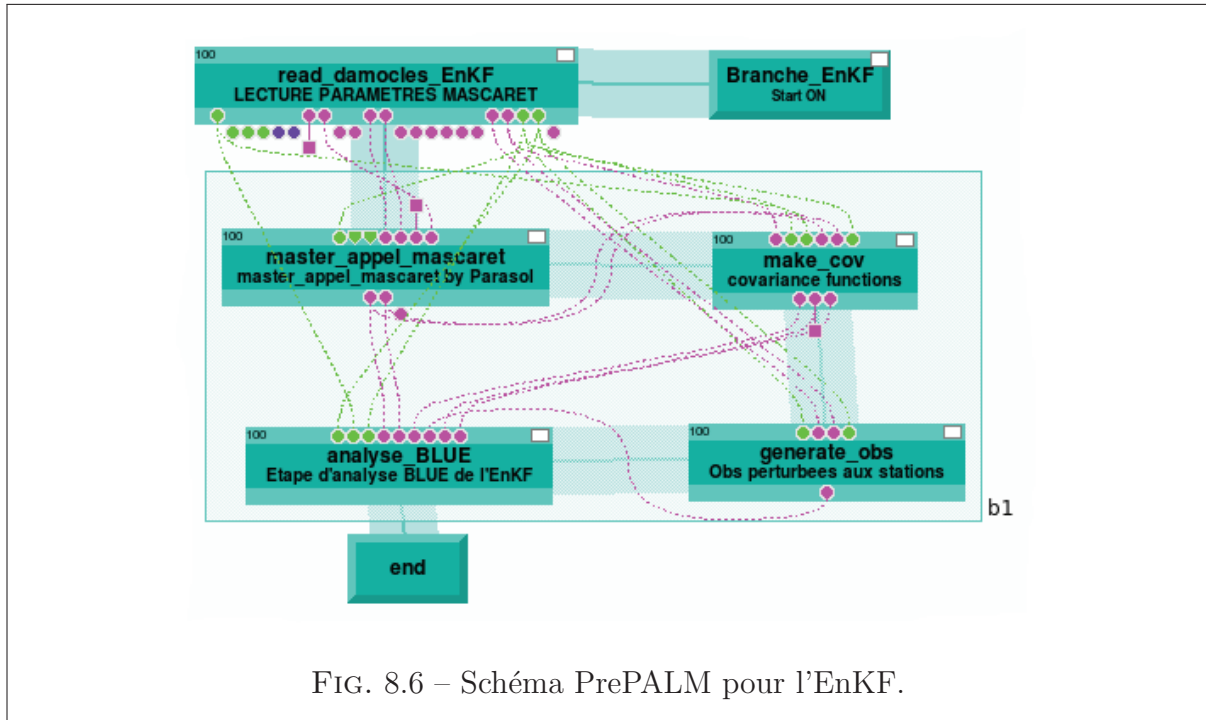


FIG. 8.6 – Schéma PrePALM pour l'EnKF.

8.1.4 Génération des membres de l'ensemble

Dans ce travail nous considérons que la principale source d'incertitude dans le modèle réside dans les forçages (amont comme aval). Les autres sources d'incertitude que l'on pourrait considérer sont liées à la bathymétrie et au fortament. Nous générons donc les membres de l'ensemble en ajoutant un bruit aléatoire à chaque forçage. Ce bruit est généré d'une manière très proche de la méthode présentée au chapitre 7 pour générer les membres de l'EnKF sur le modèle d'ondes de crues. La différence réside dans le fait que cette fois-ci nous multiplions à chaque temps le bruit ajouté aux forçages par une fonction dépendant de l'écart-type des forçages calculé sur une période de 36 h précédant le temps considéré. L'idée étant que l'incertitude sur les forçages est plus importante lorsque ceux-ci ont une importante variabilité et donc la variance du bruit ajouté aux forçages doit être plus grande dans ce cas. Pour ce qui est de la perturbation ajoutée au forçage aval nous utilisons la méthode exposée au chapitre 7 avec une variance constante et égale à 0.1 m dans un premier temps. Dans la section 8.4.1 nous estimons la variance d'erreur d'observation à 0.051 m . C'est la valeur que nous utilisons pour la variance du bruit sur le forçage aval dans les sections suivant la section 8.4.1. Enfin le choix du temps caractéristique des perturbations ajoutées aux forçages a été fait de manière empirique en retenant celui qui donnait les résultats les plus consistants, il est de 4 heures pour les forçages amonts et de 6 heures pour le forçage aval. On rappelle d'autre part que les bruits ajoutés aux forçages ont des erreurs corrélées en temps de manière à ce que les erreurs sur les hauteurs d'eau soient corrélées en espace.

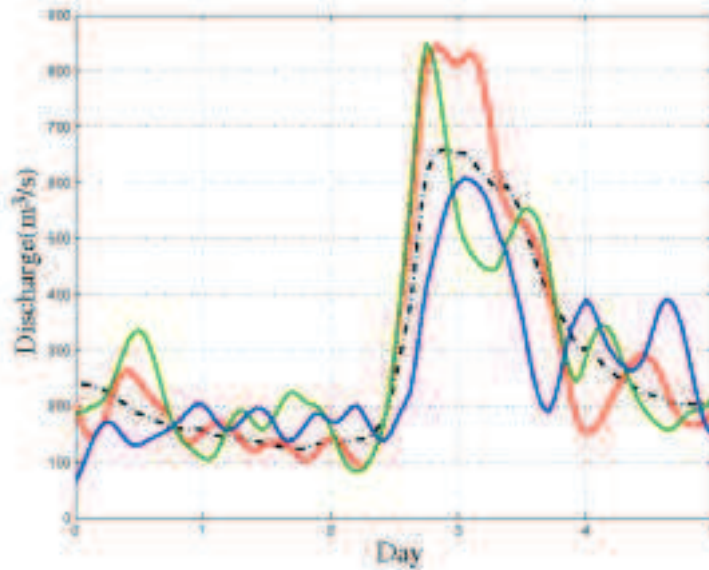


FIG. 8.7 – Illustration de l'ajout d'un bruit sur le forçage à la station d'Escos lors d'un pic de crue. La ligne noire en pointillé représente le forçage initial et les lignes rouge, verte et bleue représentent le forçage initial auquel on a ajouté une perturbation.

8.2 Etude des fonctions de covariance et de corrélation d'erreur d'ébauche sans assimilation

Dans la continuité du chapitre 7 on s'intéresse dans cette section à l'étude des fonctions de covariance et de corrélation d'erreur d'ébauche (ou covariance et corrélation d'erreur d'état du modèle) dans le cas sans assimilation sur le réseau Adour maritime avec le code hydraulique Mascaret-1D. Matrice et fonctions de covariance/corrélation d'erreur d'ébauche sont au cœur de la plupart des systèmes d'assimilation de données. Ainsi il est intéressant de connaître les propriétés de ces fonctions et de mettre en lumière comment elles sont reliées à la géométrie du problème ainsi qu'aux forçages et notamment le forçage maritime qui joue un rôle important sur le réseau Adour. Dans cette section on présente comment nous établissons le diagnostic des fonctions de covariance et de corrélation dans notre cas avant de nous pencher plus en détail sur les propriétés de ces fonctions.

Précision préliminaire sur le vocabulaire : dans cette section nous traitons des fonctions de covariance et de corrélation d'erreur d'ébauche de hauteurs d'eau et/ou débits/hauteurs d'eau. Il a pu arriver par abus de langage que nous parlions de fonctions de covariance et de corrélation de hauteurs d'eau ou débits/hauteurs d'eau mais dans tous les cas il s'agit bien de fonctions de covariance et de corrélation d'erreur d'ébauche de hauteurs d'eau et/ou débits/hauteurs d'eau.

8.2.1 Diagnostic des fonctions de covariance et de corrélation

Comme son nom l'indique l'EnKF est une méthode d'ensemble, elle repose sur plusieurs réalisations perturbées du système que l'on appelle membres (ou aussi lignes d'eau dans le cadre de l'hydraulique). Comme on l'a vu dans la section 8.1 ces membres sont générés en ajoutant un bruit aux forçages et en propageant ceux-ci avec le modèle hydraulique. Le modèle hydraulique résolvant les équations de Saint-Venant 1D les variables d'état sont les hauteurs d'eau (ou cotes) en mètres, en cote marine, et les débits (en $m^3.s^{-1}$). Chaque membre consiste donc en la discrétisation sur la grille de calcul 1D des variables d'état, on écrit pour chaque membre X^k :

$$x^k = [z^k, q^k]^T = [z_1^k, \dots, z_n^k, q_1^k, \dots, q_n^k]^T \quad (8.1)$$

où n est le nombre de nœuds de la grille de calcul.

La matrice des covariances d'erreur d'ébauche $\mathbf{B}_e$ est estimée statistiquement à partir des membres avec la formule (3.42). Etant donné la décomposition selon les variables d'état de x^k que l'on donne en (8.1) on peut décomposer la matrice $\mathbf{B}_e$ en quatre blocs (voir section 3.3.1) :

$$\mathbf{B}_e = \begin{pmatrix} \mathbf{B}_{ZZ} & \mathbf{B}_{ZQ} \\ \mathbf{B}_{QZ} & \mathbf{B}_{QQ} \end{pmatrix}$$

où les matrices $\mathbf{B}_{ZZ}$ et $\mathbf{B}_{QQ}$ décrivent respectivement les covariances spatiales et univariées d'erreur d'ébauche de hauteurs d'eau et de débits et les matrices $\mathbf{B}_{QZ}$ et $\mathbf{B}_{ZQ}$ décrivent les covariances spatiales multivariées (parfois appelées covariances croisées) d'erreur de hauteurs d'eau et de débits. $\mathbf{B}_e$ étant symétrique par construction, on a : $\mathbf{B}_{QZ} = \mathbf{B}_{ZQ}^T$. Il faut souligner ici l'importance particulière dans le cadre de l'Adour des covariances d'erreurs multivariées de débits et de hauteurs d'eau $\mathbf{B}_{QZ}$. En effet sur le réseau Adour maritime les seules observations disponibles sont des observations de hauteurs d'eau, la correction sur les débits est obtenue en traduisant l'information sur les hauteurs d'eau contenue dans l'innovation en débits via les covariances d'erreur de débits et de hauteurs d'eau $\mathbf{B}_{QZ}$. Cela nous permet de conserver un état hydraulique cohérent lors de l'assimilation.

Connaissant la matrice $\mathbf{B}_e$ on peut alors calculer la matrice $\mathbf{C}$ des corrélations d'erreur d'ébauche avec la formule (3.2). De manière similaire à $\mathbf{B}_e$ on peut décomposer $\mathbf{C}$ en quatre blocs décrivant chacun les corrélations entre les erreurs sur les différentes variables d'état :

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{ZZ} & \mathbf{C}_{ZQ} \\ \mathbf{C}_{QZ} & \mathbf{C}_{QQ} \end{pmatrix}$$

Dans la suite, dans un souci de simplicité, nous parlerons de fonctions de covariance (de corrélation) d'erreur ZZ pour évoquer les fonctions de covariance (de corrélation) d'erreurs sur les hauteurs d'eau et de fonctions de covariance (de corrélation) d'erreur QZ pour évoquer les fonctions de covariance (de corrélation) d'erreurs de débits et de hauteurs d'eau.

Enfin il faut noter que dans la pratique on ne calcule que rarement les matrices $\mathbf{B}_e$ et $\mathbf{C}$, on se contente en général de calculer les colonnes correspondant de celles-ci

relatives aux stations d'observation. Dans notre cas les stations d'observation sont des stations *in situ*, l'opérateur d'observation $\mathbf{H}$ est un opérateur linéaire nul partout sauf aux coordonnées correspondant à une station : par exemple si la première station est positionnée au nœud de calcul i alors $\mathbf{H}(j, 1) = \delta_{ij}$. Ainsi le produit $\mathbf{B}\mathbf{H}^T$ qui apparaît dans l'équation d'analyse du filtre de Kalman consiste simplement à extraire de $\mathbf{B}$ les colonnes correspondant aux stations d'observation. Et c'est la raison pour laquelle dans la suite nous ne nous intéresserons qu'aux fonctions de covariance et/ou de corrélation relatives aux stations d'observation.

8.2.2 Les fonctions de covariance

Comme cela a déjà été évoqué en section 1.3.2 nous ne disposons sur le réseau Adour maritime que d'observations de hauteurs d'eau. Celles-ci permettent de corriger les hauteurs d'eau mais aussi, à travers les fonctions de covariance QZ de corriger les débits. Ainsi en ne disposant que de hauteurs d'eau nous sommes en mesure de corriger les deux variables d'état. C'est aussi pourquoi dans la suite nous nous intéresserons seulement aux fonctions de covariance ZZ et QZ . D'autant plus que les fonctions de covariance QQ présentent des propriétés similaires aux fonctions de covariance QZ .

Dans le chapitre 7, on a montré que dans le cas du modèle d'ondes de crues si on impose un forçage à l'amont suivant une variable aléatoire gaussienne et de fonction d'autocorrélation temporelle gaussienne alors les fonctions de covariance d'erreur d'anomalie de hauteur d'eau du signal propagé le long du domaine sont gaussiennes. Dans le cas des équations de Saint-Venant avec la géométrie du modèle Adour maritime cette propriété n'est pas préservée, les fonctions de covariance d'erreur ZZ et QZ ne sont pas gaussiennes. Celles-ci présentent en effet une importante variabilité temporelle, elles exhibent néanmoins certains motifs invariants que l'on peut mettre en relation avec la géométrie du réseau hydraulique. On a représenté sur la figure 8.8 une fonction de covariance ZZ et une fonction de covariance QZ (figures 8.8-(a) et 8.8-(b) respectivement) ainsi que la bathymétrie (figures 8.8-(c)) le long du chemin 6-5-2-1. On a choisi de représenter les fonctions de covariance le long de ce chemin mais les propriétés mises en évidence sur cet exemple restent vraies sur les autres chemins.

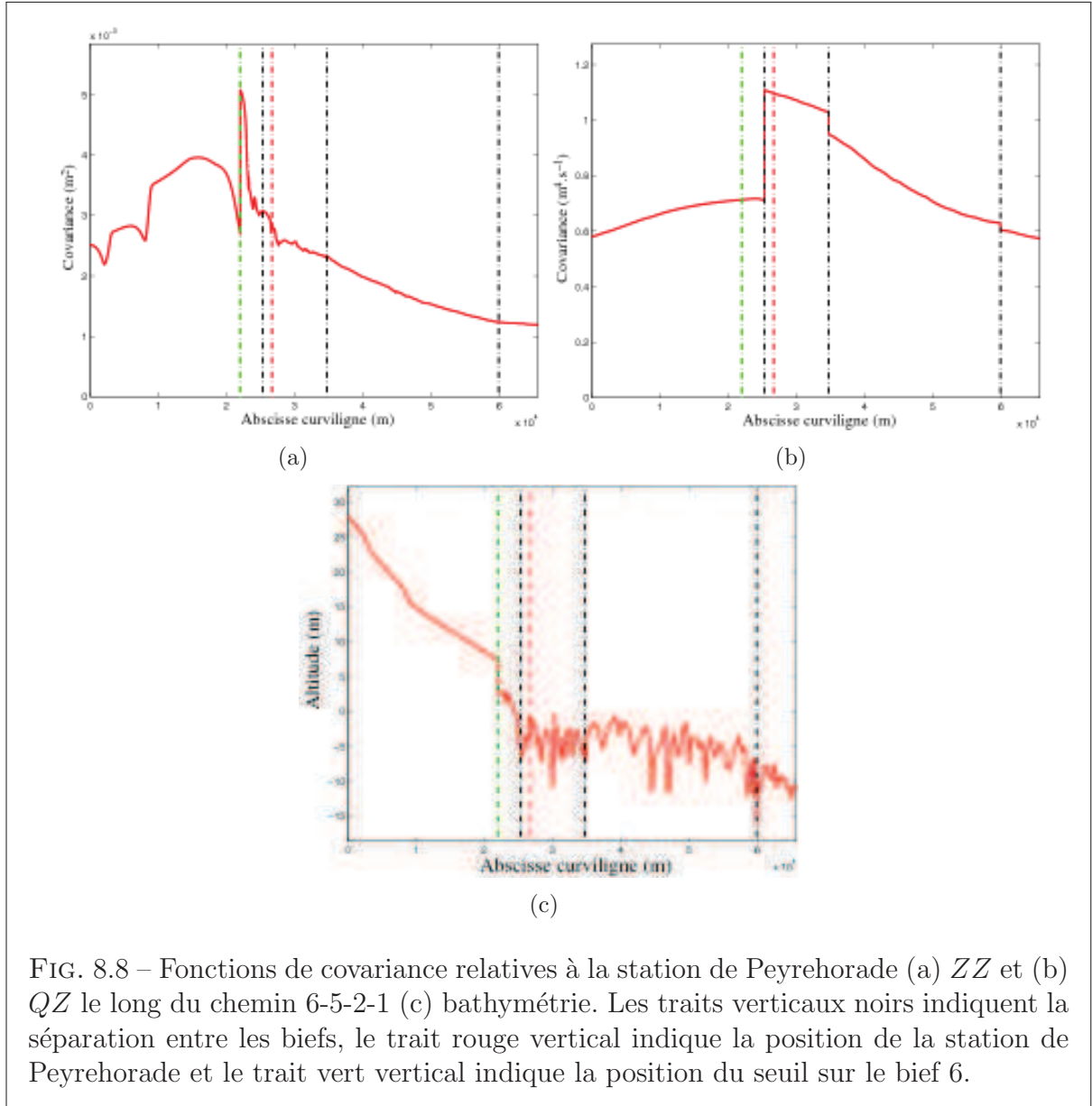


FIG. 8.8 – Fonctions de covariance relatives à la station de Peyrehorade (a) ZZ et (b) QZ le long du chemin 6-5-2-1 (c) bathymétrie. Les traits verticaux noirs indiquent la séparation entre les biefs, le trait rouge vertical indique la position de la station de Peyrehorade et le trait vert vertical indique la position du seuil sur le bief 6.

La fonction de covariance ZZ relative à la station de Peyrehorade représentée en figure 8.8-(a) présente deux parties bien distinctes. Celles-ci sont séparées par le seuil hydraulique du bief 6 dont la position est indiquée sur la figure 8.8 par un trait vert vertical. On rappelle que la partie en amont du seuil sur le bief 6 n'est pas sous influence maritime. On remarque qu'en amont du seuil et au niveau du seuil la fonction de covariance ZZ présente des irrégularités dont l'abscisse curviligne est corrélée à l'abscisse curviligne des irrégularités de la bathymétrie et notamment du seuil. En aval du seuil les covariances décroissent lentement jusqu'à l'extrémité aval du réseau hydraulique et ne présentent pas d'irrégularité. La forme de la fonction de covariance ZZ s'explique assez bien par le fait que les hauteurs d'eau en amont du seuil sont liées à la bathymétrie et en particulier à la pente, celle-ci étant relativement importante le bruit imposé à l'amont est advecté le long du domaine à l'aval. En aval du seuil la profondeur étant plus importante et la

penne globalement plus faible la diffusion est plus importante la forme de la fonction de covariance ZZ est plus régulière et sera d'autant plus influencée par le bruit ajouté au forçage aval qu'on se rapproche de celui-ci, ceci est illustré par la figure 8.9-(a).

La fonction de covariance QZ relative à la station de Peyrehorade représentée en figure 8.8-(b) ne présente pas la même sensibilité à la bathymétrie. Par contre on note des discontinuités au niveau des confluences. Ces deux constats s'expliquent très bien par le fait que le débit est une variable conservative qui n'est donc pas sensible aux variations de géométrie le long d'un bief (voir figure 8.9-(b)). D'autre part le débit étant une variable additive les bruits ajoutés aux forçages s'additionnent au niveau des confluences ce qui explique la discontinuité des fonctions de covariance QZ au niveau des confluences (qui correspondent également à la séparation entre les biefs indiquées par les traits noirs verticaux).

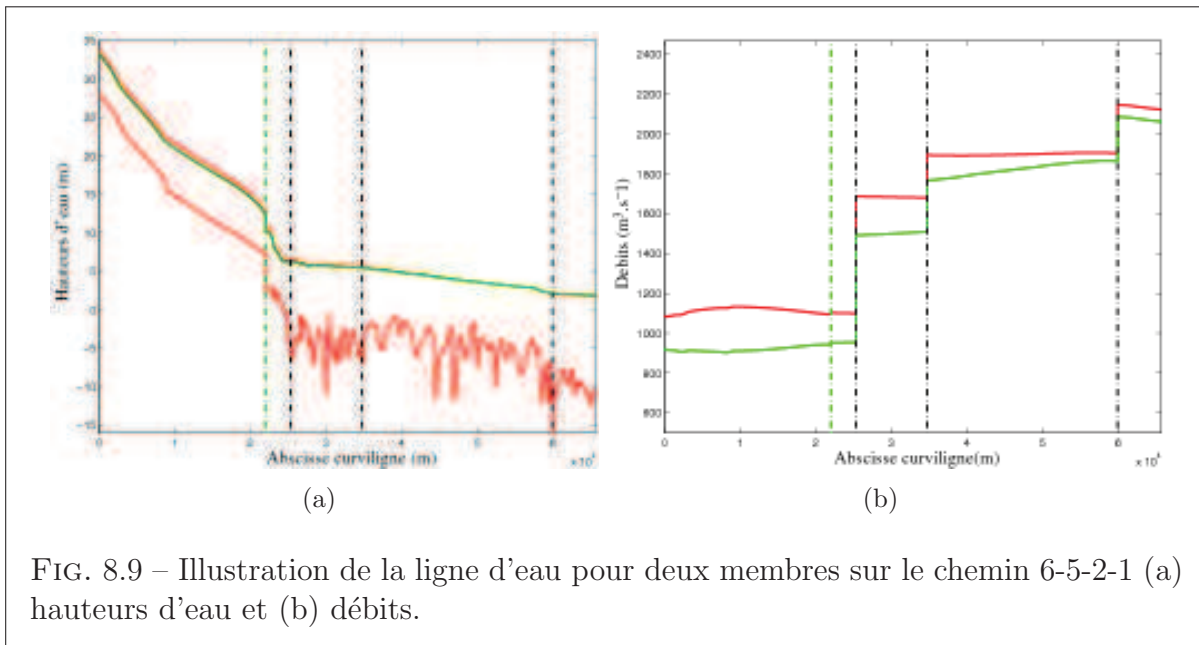


FIG. 8.9 – Illustration de la ligne d'eau pour deux membres sur le chemin 6-5-2-1 (a) hauteurs d'eau et (b) débits.

Les fonctions de covariance présentent une grande variabilité temporelle. Celle-ci est due à l'influence de la marée au niveau du forçage aval ainsi qu'au bruit ajouté aux forçages amonts. Pour illustrer ce fait nous avons regroupé en figure 8.10 les fonctions de covariance ZZ et QZ calculées à deux temps différents.

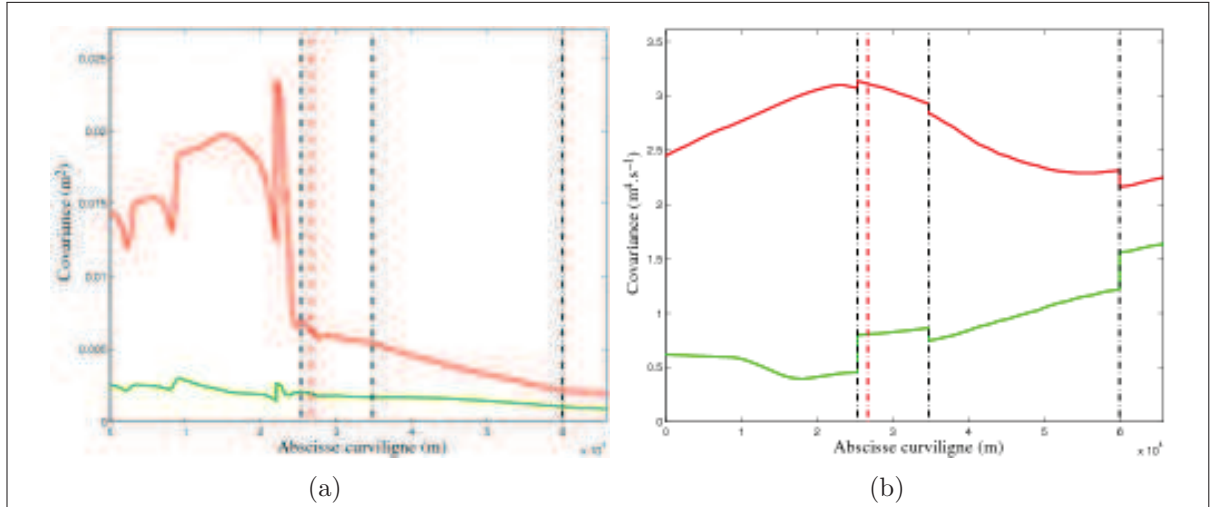


FIG. 8.10 – Fonctions de covariance (a) ZZ et (b) QZ sur le chemin 6-5-2-1 à deux temps différents. Il faut souligner que les temps auxquels on présente ces figures sont différents du temps de la figure 8.8.

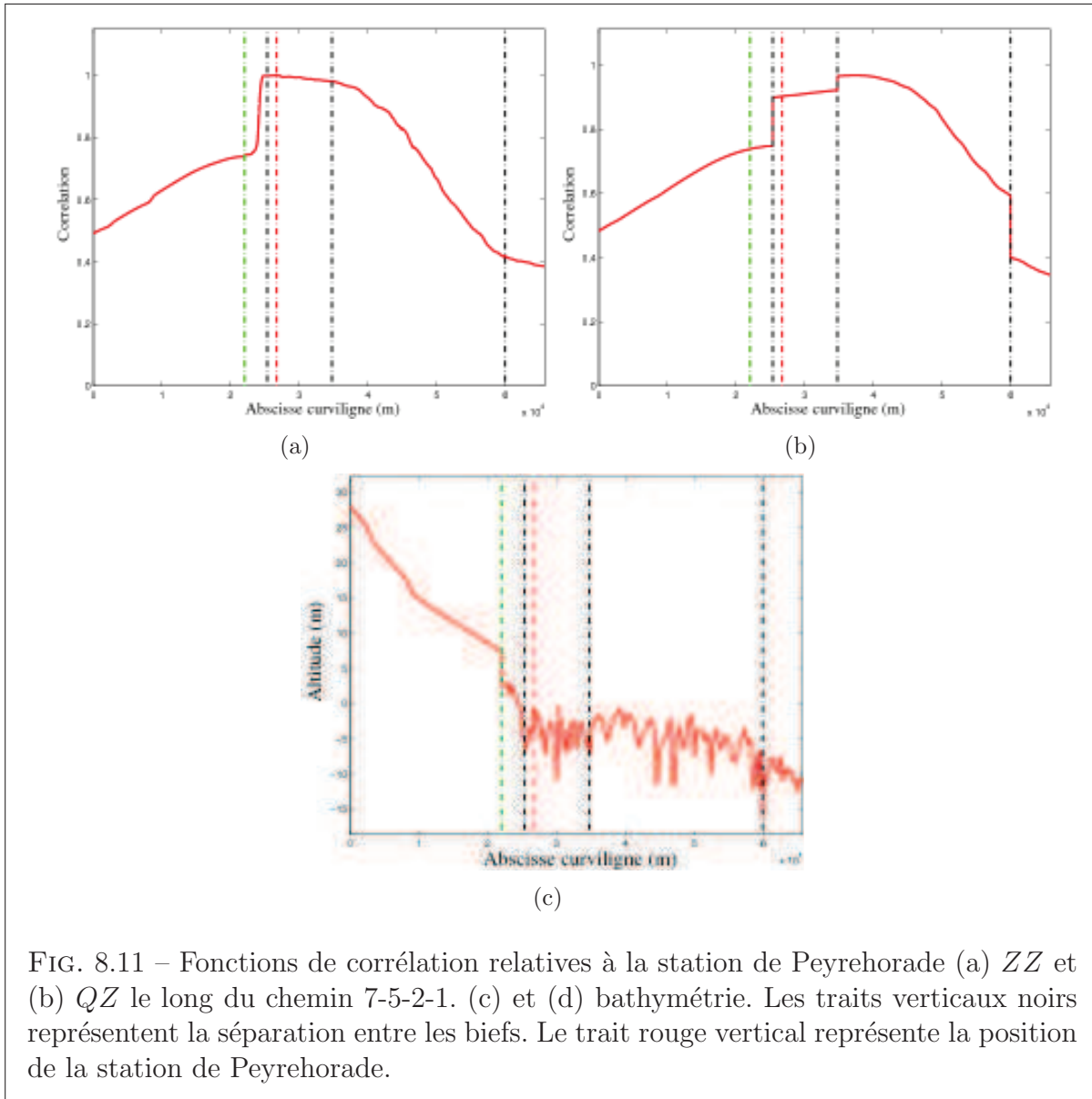
8.2.3 Les fonctions de corrélation

Les fonctions de corrélation permettent de caractériser la corrélation de l'erreur en un point avec les erreurs aux points voisins [Daley, 1991, Pannekoucke et al., 2008] mais aussi d'estimer la distance de corrélation significative à travers le diagnostic de la longueur de portée. Ainsi les fonctions de corrélation portent une information sur la structure du signal et permettent d'estimer l'extension spatiale des fonctions de corrélation mais aussi de covariance et de la correction apportée par l'assimilation.

On a vu dans le chapitre 7 que les fonctions de corrélation sur le modèle d'ondes de crues sont des gaussiennes. On peut dans ce cas estimer les longueurs de portée des fonctions de corrélation grâce à la formule proposée par [Pannekoucke et al., 2008], celle-ci étant basée sur l'approximation des fonctions de corrélation par une fonction parabolique bien choisie. Néanmoins comme nous le verrons plus loin il s'avère que les fonctions de corrélation sur le réseau Adour maritime n'ont pas une forme de gaussienne et il ne nous sera pas possible d'utiliser la formule établie par [Pannekoucke et al., 2008] pour estimer les longueurs de portée.

D'autre part dans ce travail nous avons fait un usage important des fonctions de corrélation dans le cadre des méthodes d'inflation (évoquées en section 3.7.7) et appliquées au réseau Adour maritime. En effet comme nous le verrons plus loin nous souffrons d'un problème de sous-dispersivité de l'ensemble et pour compenser ce problème nous avons eu recours à ces méthodes. Dans ce cadre l'utilisation des fonctions de corrélation a permis de calculer des facteurs d'inflation adaptés spatialement à notre réseau. Nous reviendrons sur ce point plus en détail en section (8.4.2).

Nous avons reporté sur les figures 8.11-(a) et 8.11-(b) les fonctions de corrélation ZZ et QZ relatives à la station de Peyrehorade (dont la position est indiquée par le trait



rouge vertical) le long du chemin 6-5-2-1 (prises au même temps que les fonctions de covariance présentées en figure 8.8). La fonction de corrélation ZZ est moins sensible à la bathymétrie en amont du seuil (trait vert vertical) que ne l'est la fonction de covariance ZZ représentée en figure 8.8-(a). D'autre part au niveau des seuils les fonctions de corrélation ZZ présentent une discontinuité, qui est plus ou moins prononcée en fonction du temps. On note également que la fonction de corrélation ZZ présente une très grande extension spatiale à l'aval du seuil puisque celle-ci s'étend jusqu'à l'extrémité aval du domaine. Comme on le verra plus loin l'évolution temporelle de l'extension spatiale des fonctions de corrélation semble liée à l'évolution temporelle des variables d'état aux stations. Néanmoins nous n'avons pas pu établir clairement de lien de cause à effet dans ce sens. D'autre part, de la même manière que la fonction de covariance QZ la fonction de corrélation QZ n'est pas affectée par la bathymétrie et présente des discontinuités

au niveau de la séparation entre les biefs (qui correspondent également à la position des confluences).

Les fonctions de corrélation tout comme les fonctions de covariance présentent une très grande variabilité temporelle. Celle-ci est due à la fois à l'influence de la marée à l'aval et à l'importance de la perturbation ajoutée aux forçages amonts qui varient avec le temps. Pour illustrer cela nous avons reporté en figure 8.12 des fonctions de corrélation ZZ et QZ calculées à deux temps différents.

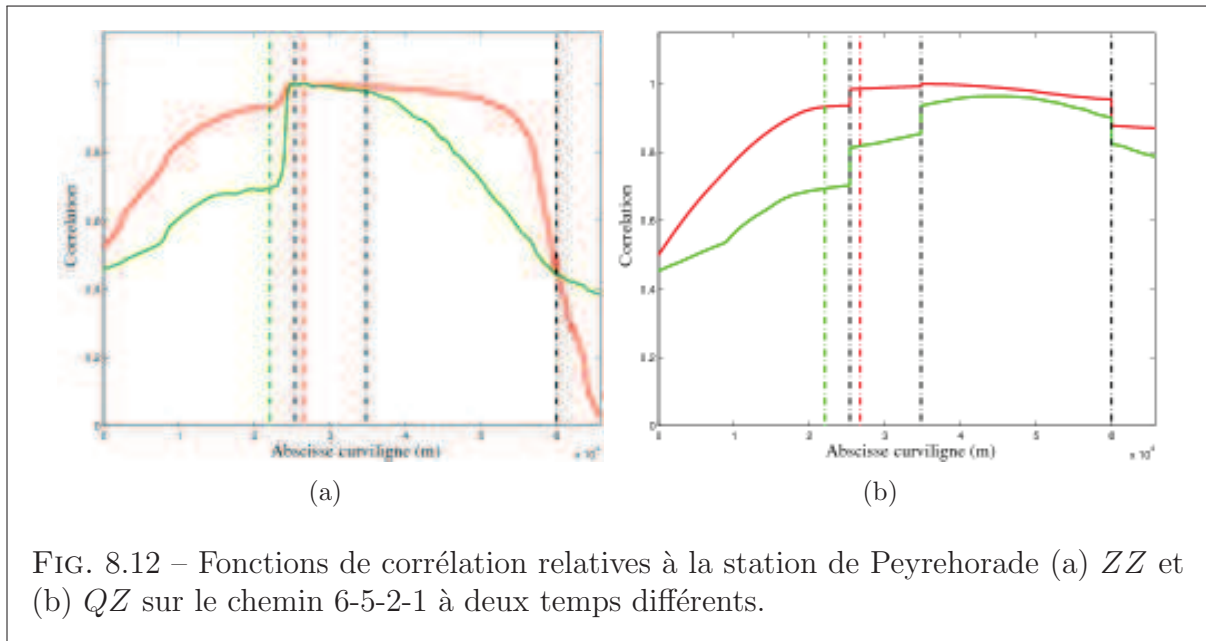


FIG. 8.12 – Fonctions de corrélation relatives à la station de Peyrehorade (a) ZZ et (b) QZ sur le chemin 6-5-2-1 à deux temps différents.

Comme on peut le voir sur la figure 8.12-(a) l'extension spatiale des fonctions de corrélation ZZ peut présenter une très grande variabilité temporelle. En effet la fonction de corrélation représentée en rouge est plus étendue à l'aval que ne l'est celle représentée en vert. La variabilité temporelle de l'extension spatiale des fonctions de corrélation semble liée au forçage aval; voir annexe (B) où on a représenté l'évolution des fonctions de corrélation relatives à la station de Peyrehorade sur le chemin 7-5-2-1 sur la durée d'un cycle de marée. Mais elle semble également liée à la variabilité temporelle des variables d'état à la station relative à la fonction considérée. Néanmoins ce fait reste à apprécier quantitativement. Pour cela il faut définir l'analogie d'une longueur de portée pour notre réseau hydraulique et étudier sa variabilité temporelle aux stations du réseau, c'est l'objet de la section suivante.

8.2.4 Caractérisation et évolution temporelle de l'extension spatiale des fonctions de corrélation ZZ

Comme on a pu le constater à la section précédente les fonctions de corrélation ZZ mises en évidence sur le réseau hydraulique Adour maritime ne sont pas gaussiennes.

Il ne nous est donc pas possible d'utiliser le calcul de la longueur de portée proposé par [Pannekoucke et al., 2008] et que l'on a utilisé précédemment au chapitre 7 pour caractériser l'extension spatiale des fonctions de corrélation. A cela s'ajoute la difficulté de la configuration du réseau Adour sur lequel les fonctions de corrélation ne sont pas à proprement à parler 1D puisqu'elles peuvent s'étendre sur plusieurs biefs qui ne sont pas disposés séquentiellement de l'amont vers l'aval de la rivière.

On peut néanmoins définir une forme de longueur de portée pour les fonctions de corrélation ZZ sur le réseau Adour par analogie avec celle proposée par [Pannekoucke et al., 2008]. Celle-ci correspond à la distance telle que la parabole qui approxime la fonction de corrélation gaussienne soit égale à 0.5 et renseigne donc sur l'extension spatiale de la fonction de corrélation correspondante. Dans ce travail nous définirons la longueur de portée aval (amont) de la fonction de corrélation ZZ associée à un point x , ρ_x , sur le réseau Adour maritime, comme étant la somme de la longueur des zones sur chacun des biefs à l'aval (à l'amont) du point considéré telles que $\rho_x \geq 0.65$. Cette définition présente l'avantage d'être indépendante de la forme des fonctions de corrélation considérées mais est par contre fortement contrainte par les faibles dimensions du domaine et il pourra se révéler impossible dans certains cas de calculer des longueurs de portée amont et/ou aval qui reflètent fidèlement l'extension spatiale de la fonction de corrélation considérée. C'est par exemple le cas lorsque sur un chemin à l'aval du point considéré la fonction de corrélation est décroissante mais à l'extrémité aval du domaine reste supérieure à 0.65. Dans ce cas la longueur de portée aval est simplement égale à la distance entre le point considéré et la limite aval du domaine alors que dans les faits elle est plus grande.

Nous avons illustré le calcul de la longueur de portée sur la figure 8.13. Cette définition des longueurs de portée amont et aval vise à rendre compte de l'extension spatiale des fonctions de corrélation à l'amont ou à l'aval d'un point considéré sur le réseau. Mais on peut également considérer l'extension spatiale de la fonction de corrélation sur l'ensemble du réseau, pour cela on peut définir une longueur de portée moyenne comme la moyenne des longueurs de portée amont et aval.

Cette définition des longueurs de portée est une définition très générale puisqu'elle vise à établir l'extension spatiale des fonctions de corrélation sur l'ensemble du réseau. On peut également définir une notion de longueur de portée par chemin qui est simplement la restriction de la définition ci-dessus à un chemin en particulier, par exemple le chemin 4-2-1. Cela est illustré sur la figure 8.13-(c) où la longueur du trait rouge horizontal représente la longueur de portée amont sur le chemin 7-5-2-1 de la fonction de corrélation relative à la station de Peyrehorade.

Les longueurs de portée présentent une importante variabilité temporelle ainsi qu'une certaine cyclicité qui n'est pas sans rappeler l'évolution temporelle des hauteurs d'eau ainsi que des débits au niveau des points correspondants comme on peut le voir sur la figure 8.14. Même si les séries temporelles des hauteurs d'eau/débits d'une part et celles des longueur de portée d'autre part semblent présenter des liens il est néanmoins difficile d'établir des corrélations entre les deux. En effet les limites du domaine ainsi que les seuils parasitent le calcul des longueurs de portée tel qu'on l'a proposé ici. Cependant dans les

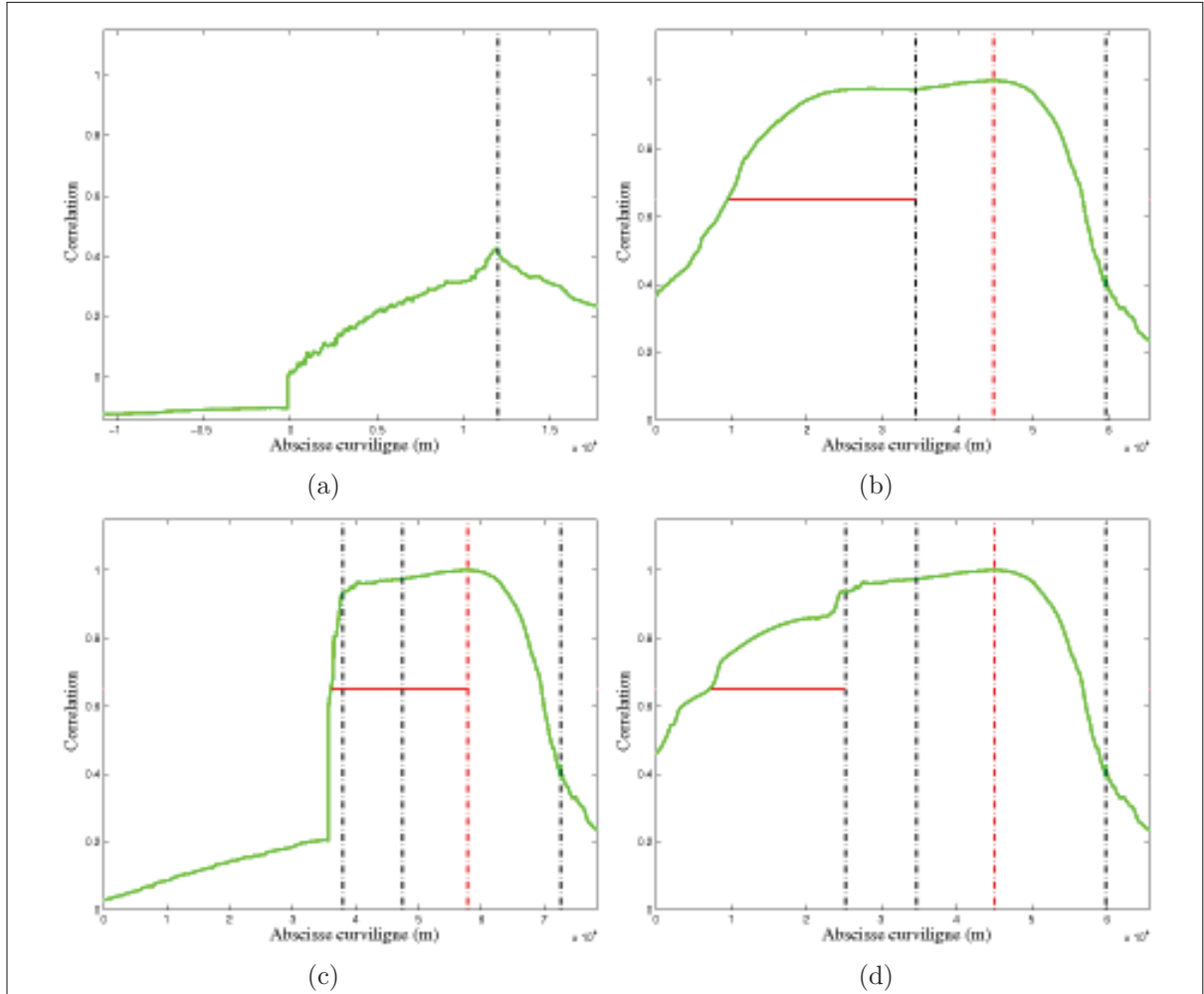


FIG. 8.13 – Fonction de corrélation ZZ relative à la station de Urt (trait rouge vertical) sur les chemins (a) 3-1 (b) 4-2-1 (c) 7-5-2-1 (d) 6-5-2-1. La longueur de portée amont est la somme des longueurs des zones représentées par le trait rouge horizontal.

cas favorables où les limites du domaine tout comme les seuils ne perturbent pas le diagnostic des longueurs de portée les corrélations calculées entre hauteurs d'eau/débits et longueur de portée sont généralement relativement correctes et sont comprises (en valeur absolue) entre 0.6 et 0.9. Nous présentons ci-dessous en figures 8.14 et 8.15 deux exemples où les corrélations entre hauteurs d'eau/débits d'une part et longueur de portée d'autre part sont les plus importantes.

Sur la figure 8.14 nous présentons l'évolution temporelle des longueurs de portée (amont, aval et moyenne) et des hauteurs d'eau calculées à la station de Peyrehorade en basses eaux. Les séries temporelles des longueurs de portées moyennes et aval et celles des hauteurs d'eau présentent un coefficient de corrélation de l'ordre de respectivement 0.72 et 0.66. On note en particulier que celles-ci présentent un temps caractéristique qui semble très proche et que les temps où les extremums locaux sont atteints sont eux aussi

très proches. Ainsi en basses eaux lorsque les hauteurs d'eau décroissent les longueurs de portée moyenne et aval décroissent et au contraire elles croissent lorsque les hauteurs d'eau croissent. Par contre pour ce qui est de la corrélation des débits et des longueurs de portée elles sont assez faibles puisque comprises en valeur absolue entre 0.1 et 0.36.

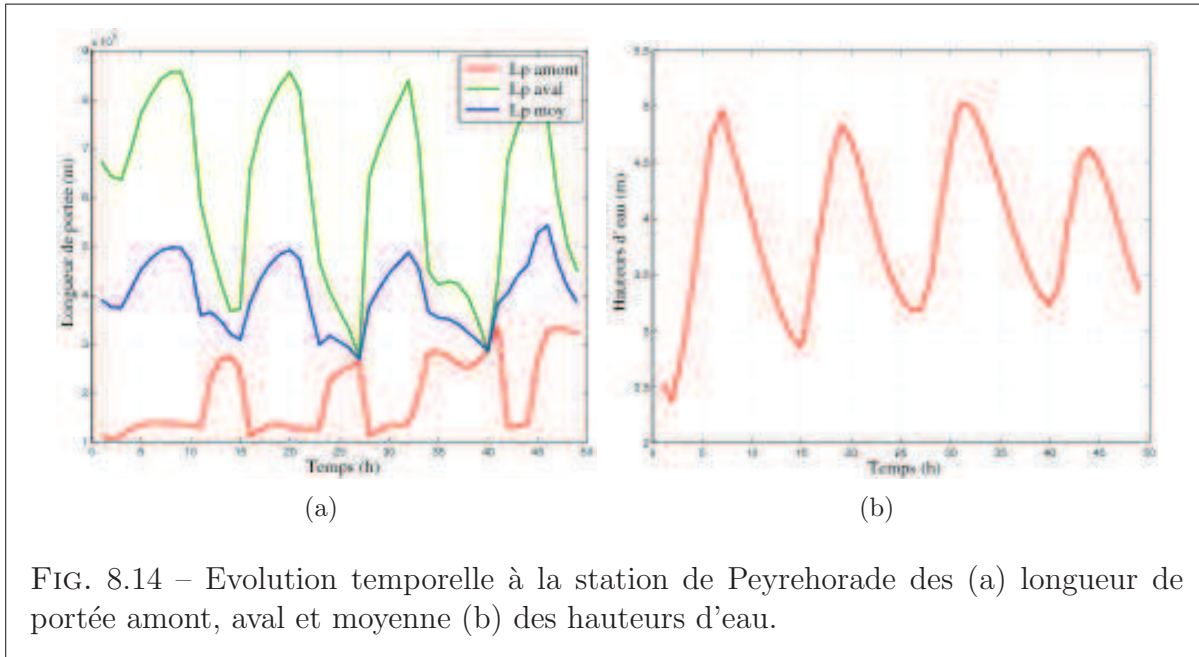


FIG. 8.14 – Evolution temporelle à la station de Peyrehorade des (a) longueur de portée amont, aval et moyenne (b) des hauteurs d'eau.

Sur la figure 8.15 nous présentons l'évolution temporelle des longueurs de portée (amont et aval) et des débits calculés à la station de Urt lors d'un pic de crue. Les longueurs de portée amont semblent avoir des variations similaires à celles des débits, le coefficient de corrélation liant les deux séries est de l'ordre de 0.92. Par contre pour ce qui est des longueurs de portée aval celles-ci semblent se réduire lors du pic de crue, les séries temporelles des débits et des longueurs de portée présentent un coefficient de corrélation de -0.64 . Ainsi l'extension spatiale de la fonction de corrélation relative à Urt présente-t-elle une forme d'asymétrie pendant le pic de crue avec une longueur de portée aval nettement plus petite que la longueur de portée amont. A noter que dans cet exemple les corrélations entre longueurs de portée amont et aval et hauteurs d'eau sont de l'ordre de 0.72 et 0.6.

Pour conclure cette section nous dirons que les figures 8.14 et 8.15 suggèrent que l'évolution temporelle des longueurs de portée amont et aval semble liée à l'évolution temporelle des hauteurs d'eau ou des débits selon le moment considéré (pic de crue ou basses eaux). D'autre part les figures de l'annexe (B) suggèrent un lien entre les longueurs de portée aval. Néanmoins malgré ces quelques rapprochements il ne nous a pas été possible d'établir clairement un lien entre longueurs de portée et variables d'état.

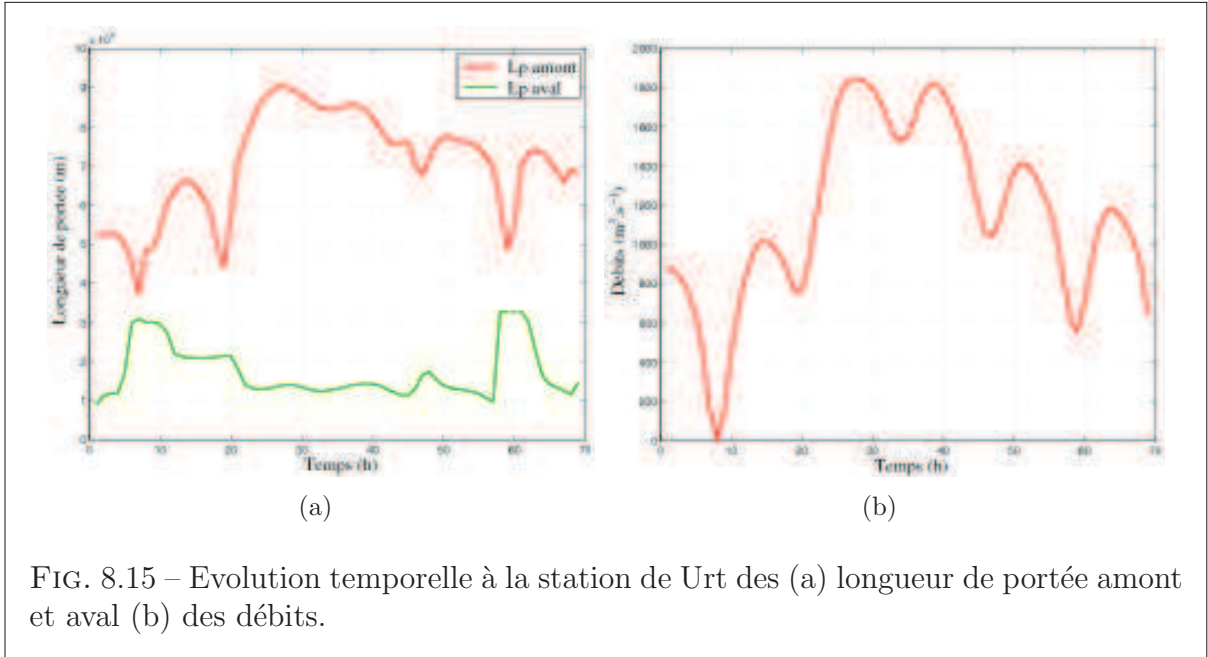


FIG. 8.15 – Evolution temporelle à la station de Urt des (a) longueur de portée amont et aval (b) des débits.

8.3 Application de l'EnKF au réseau Adour maritime : première approche et limitations

Dans cette section nous nous intéressons à l'application de l'EnKF au réseau Adour maritime dans sa forme classique telle qu'elle a été présentée en section 3.7.5. Dans la section 8.2 nous avons mis en évidence la grande extension spatiale des fonctions de covariance et de corrélation diagnostiquées sur ce réseau avec une méthode d'ensemble. Nous voulons dans cette section montrer l'intérêt pour l'assimilation que présentent ces fonctions. Pour cela nous nous plaçons dans le prolongement des travaux exposés au chapitre 7 et dans l'article de [Ricci et al., 2011]. Dans le chapitre 7 nous avons appliqué le filtre de Kalman d'ensemble sur le modèle d'ondes de crues 1D (qui est, on le rappelle, une approximation des équations de Saint-Venant 1D sous certaines hypothèses de pente et de frottement). Dans l'article de [Ricci et al., 2011] un filtre de Kalman utilisant des fonctions de covariance stationnaires aux points d'observation avait été mis en place sur le réseau Adour. Dans la suite nous désignerons ce filtre de Kalman par l'expression « filtre de Kalman invariant » ou encore « KF invariant ». Ainsi nous montrons qu'à variance équivalente au point d'observation l'EnKF donne des résultats meilleurs que ceux du KF invariant utilisé dans [Ricci et al., 2011] et montrons les limitations de l'application de l'EnKF sur le réseau Adour maritime.

8.3.1 Une expérience d'assimilation de données de hauteurs d'eau synthétiques

Pour montrer l'apport de l'EnKF par rapport à un KF invariant nous procédons à une expérience d'assimilation de données synthétiques de hauteurs d'eau à la station de

Peyrehorade et comparons les résultats de ces deux algorithmes d'assimilation de données à Peyrehorade mais aussi aux autres stations où nous n'avons pas procédé à l'assimilation et notamment à Urt. On souhaite ainsi mettre en évidence l'apport de l'EnKF en terme d'extension spatiale de la correction par rapport au KF invariant tout en montrant les limitations. Le choix de la station de Peyrehorade pour assimiler les observations n'est pas anodin. En effet c'est la station où en général l'erreur du modèle par rapport aux observations est la plus importante.

Pour cette expérience nous avons généré des observations de hauteurs d'eau et de débits en ajoutant aux forçages donnés sur un épisode de crue un apport latéral juste en aval du seuil sur le bief 6. Cet apport latéral est corrélé aux forçages à Orthez et Escos et les débits maximaux de cet apport ont été choisis de telle manière que les hauteurs d'eau maximales à Peyrehorade correspondent aux hauteurs d'eau maximales observées à cette station lors de l'épisode de crue correspondant aux forçages utilisés. Nous avons choisi ce protocole expérimental de manière à simuler un épisode de crue dans lequel les apports latéraux non instrumentés (et donc pour lesquels nous n'avons aucune données d'observations) jouent un rôle important lors de la crue (ce qui est souvent le cas sur l'Adour) et pour lequel nous disposerons d'observations de hauteurs d'eau et de débits. Ce dernier point est important car dans la réalité nous ne disposons que de données de hauteurs d'eau.

Choix des écarts-types d'erreur d'observation et d'erreur d'ébauche

L'écart-type d'erreur d'observation a été fixé à 0.1 m ce qui correspond à la valeur utilisée en prévision opérationnelle sur l'Adour. Pour l'EnKF l'écart-type d'erreur d'ébauche au niveau de la station de Peyrehorade va évoluer au cours du temps et est estimé à travers le calcul des covariances d'erreur d'ébauche donc on ne le fixe pas. De manière à ce que les résultats de l'EnKF et du filtre Kalman invariant soient comparables à la station de Peyrehorade l'écart-type d'erreur d'ébauche à cette station (pour le KF invariant) est fixé égal à la moyenne des écarts-types d'erreur d'ébauche calculés par l'EnKF sur l'épisode de crue.

8.3.2 Résultats numériques

Dans cette section nous nous intéressons dans un premier temps à la comparaison de la spatialisation de la correction apportée par l'EnKF et le KF invariant avant de nous intéresser dans un deuxième temps aux séries temporelles des états analysés aux stations de Peyrehorade et de Urt. Dans un troisième temps nous nous intéresserons aux résultats de ces deux méthodes en mode prévision.

Spatialisation de la correction

Dans le KF invariant les fonctions de covariance relatives aux hauteurs d'eau aux points d'observation consistent en deux demi-gaussiennes avec une longueur de portée plus petite en aval qu'en amont (voir figure 8.16). En effet dans l'étude menée au chapitre 7 on a montré que sur le modèle d'ondes de crues les fonctions de corrélation ont une forme de gaussienne asymétrique avec une longueur de portée plus petite à l'aval qu'à l'amont

justifiant l'usage d'une fonction de covariance avec une extension spatiale plus importante à l'amont qu'à l'aval. D'autre part il faut souligner que des expériences numériques que l'on n'a pas reportées ici montrent que dans le cas du réseau Adour maritime l'utilisation de fonctions de covariance gaussiennes et symétriques donne des résultats moins bons que ceux obtenus avec des fonctions de covariance gaussiennes asymétriques.

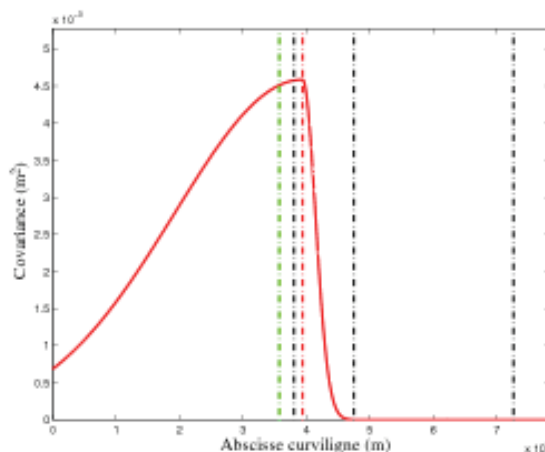


FIG. 8.16 – Fonction de covariance relative à la station de Peyrehorade (trait rouge vertical) sur le chemin 6-5-2-1.

Un point important mis en lumière en section 8.2 est la grande extension spatiale des fonctions de covariance et de corrélation. Cela se traduit naturellement par une grande extension de la correction des lignes d'eau lors de l'étape d'analyse de l'EnKF, que ce soit en termes de hauteurs d'eau ou de débits comme on peut le voir sur la figure 8.17 où l'on a représenté à un temps l'incrément d'analyse moyen sur le chemin 6-5-2-1 pour l'EnKF (ligne verte) et l'incrément d'analyse de hauteur d'eau pour le KF invariant (ligne rouge).

Le choix d'une fonction de covariance au point d'observation asymétrique pour le KF invariant se traduit par un incrément d'analyse (figure 8.17-(a) courbe rouge) en forme de deux demi-gaussiennes avec une extension spatiale plus importante à l'amont qu'à l'aval. Il apparaît que l'incrément d'analyse en Z à l'aval et à l'amont de Peyrehorade pour l'EnKF a une extension spatiale plus importante que celle du KF invariant permettant de corriger les hauteurs d'eau sur une très grande partie du réseau et notamment à l'amont de cette station.

Pour ce qui est des fonctions de covariance du KF invariant relatives aux débits là encore il s'agit de deux demi-gaussiennes avec une longueur de portée à l'amont plus grande qu'à l'aval mais l'incrément d'analyse des débits est calculé de manière à respecter l'additivité des débits au niveau des confluences d'où la discontinuité de l'incrément

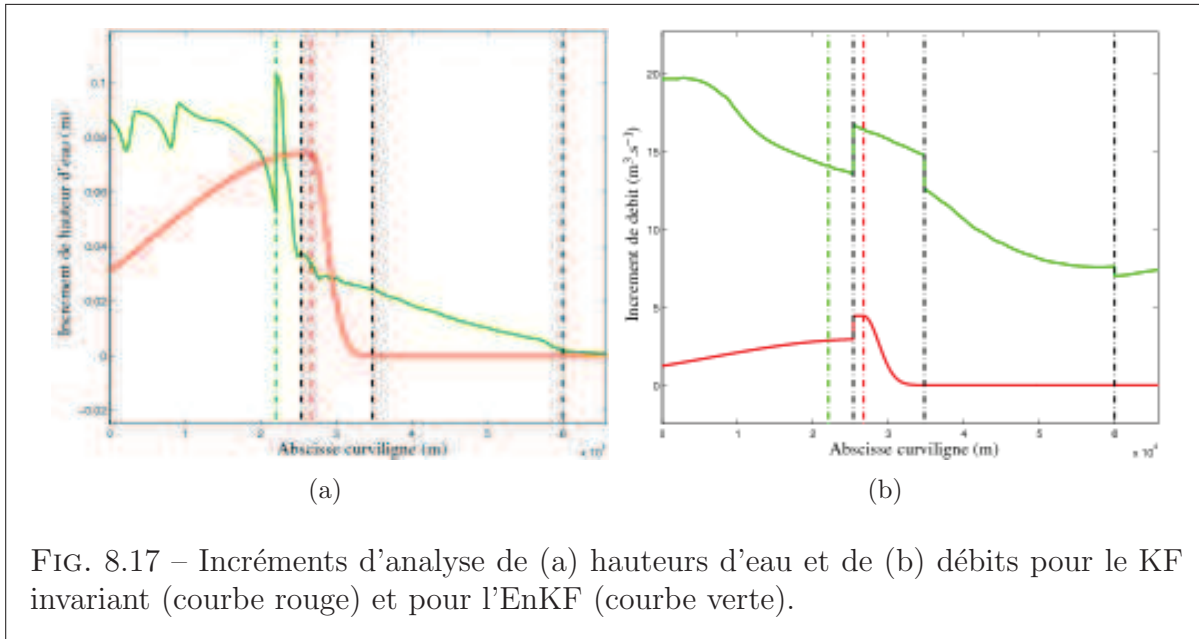


FIG. 8.17 – Increments d'analyse de (a) hauteurs d'eau et de (b) débits pour le KF invariant (courbe rouge) et pour l'EnKF (courbe verte).

observée à la confluence des biefs 6 et 7 sur la figure 8.17-(b) (courbe rouge). Dans le KF invariant l'incrément d'analyse des débits à Peyrehorade est proportionnel à l'incrément d'analyse des hauteurs d'eau et le coefficient de proportionnalité est le quotient des débits et des hauteurs d'eau à la station tandis que dans l'EnKF l'incrément d'analyse des débits à la station et sur le reste du domaine est calculé avec les covariances QZ . Là encore la correction des débits est nettement plus étendue avec l'EnKF qu'avec le KF invariant. Et comme on peut le voir l'incrément d'analyse des débits de l'EnKF présente des discontinuités au niveau des confluences et ce de manière consistante avec les fonctions de covariance QZ . D'autre part il faut noter que l'incrément de débit est nettement plus important pour l'EnKF que pour le KF invariant sur toute la longueur du chemin 6-5-2-1. Cela montre que les deux algorithmes ne traduisent pas du tout de la même manière l'information apportée par les observations de hauteurs d'eau en débits. Ceci explique en partie pourquoi à variance équivalente l'EnKF donne de meilleurs résultats que le KF invariant.

L'évolution temporelle des fonctions de covariance se traduit par une évolution temporelle des incréments d'analyse de hauteurs d'eau et de débits comme on peut le voir sur la figure 8.18 où on les a reporté sur le chemin 7-5-2-1 pour deux temps différents. Cette évolution temporelle de l'incrément d'analyse permet d'adapter la correction apportée par l'assimilation tout au long de l'épisode de crue en fonction des covariances diagnostiquées par l'EnKF.

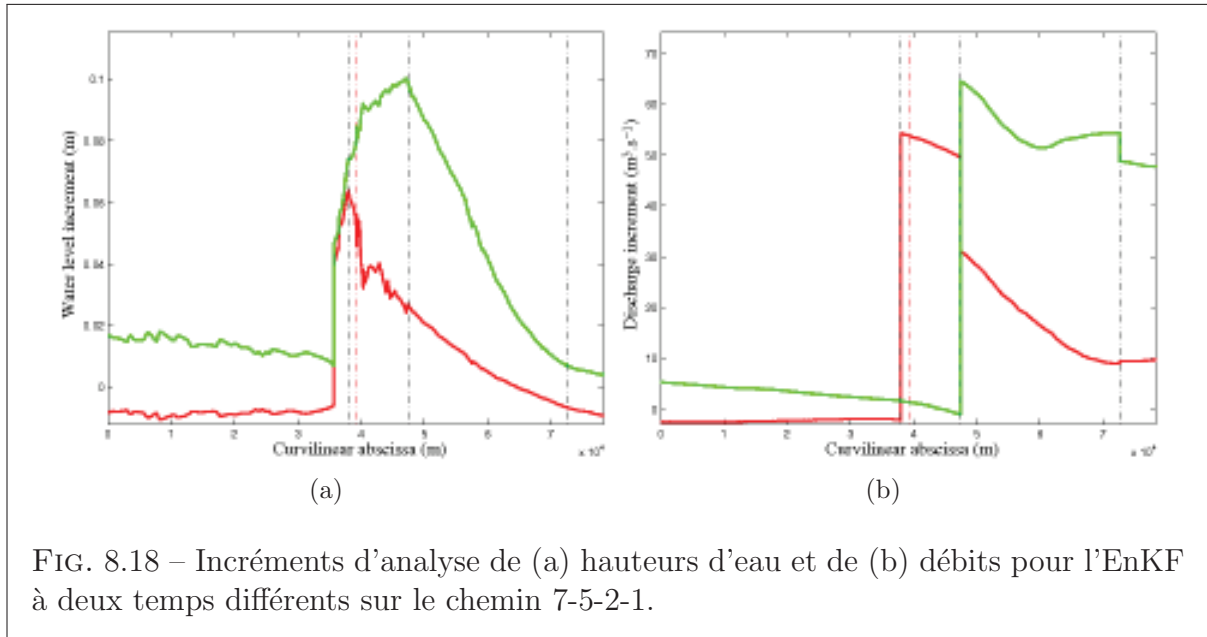


FIG. 8.18 – Incréments d'analyse de (a) hauteurs d'eau et de (b) débits pour l'EnKF à deux temps différents sur le chemin 7-5-2-1.

Evolution temporelle des états analysés

Dans la suite, sauf mention contraire, l'expression « états analysés » pour les méthodes d'ensemble désignera la moyenne de l'ensemble des états calculés après assimilation, que ce soient les hauteurs d'eau ou les débits.

On rappelle également que dans cette expérience on a assimilé des hauteurs d'eau à la station de Peyrehorade et nous regardons les résultats à cette station mais aussi à celle de Urt en termes de hauteurs d'eau dans un premier temps puis en termes de débits.

L'évolution temporelle des états analysés à Peyrehorade et aux autres stations montre des résultats relativement corrects en termes de hauteurs d'eau de l'EnKF par rapport au modèle sans assimilation (voir figure 8.19) et aussi par rapport au KF invariant. Celui-ci a une variance constante qui du fait de la manière avec laquelle ont été construits les bruits ajoutés aux forçages pour l'EnKF est plus faible que celle de l'EnKF pendant les pics de crue, ce qui explique que le KF invariant donne de moins bons résultats pendant les pics de crue que l'EnKF. Dans tous les cas les résultats sont nettement meilleurs que ceux calculés avec le modèle seul i-e sans assimilation. Ceci illustre l'intérêt des fonctions de covariance (et de leur grande extension spatiale) diagnostiquées par l'EnKF qui permettent de corriger débits et hauteurs d'eau au niveau de la station d'observation utilisée mais aussi ailleurs sur le domaine.

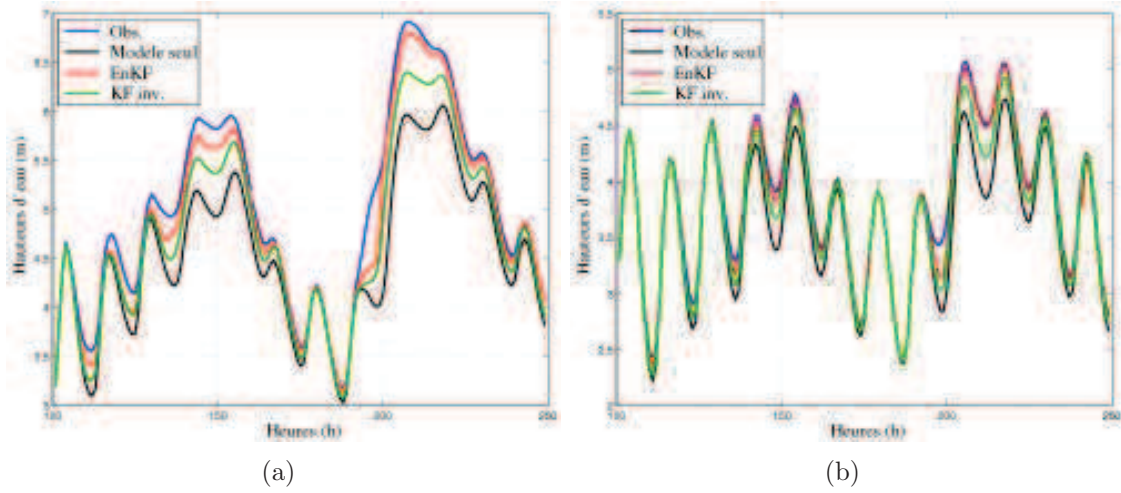


FIG. 8.19 – Séries temporelles des hauteurs d'eau analysées à (a) Peyrehorade et (b) Urt.

Pour ce qui est des débits les conclusions sont similaires à celles tirées pour les hauteurs d'eau : les résultats sont relativement corrects à la station d'observation ce qui met en lumière l'intérêt des covariances QZ en ce point mais aussi aux autres stations du fait de la grande extension spatiale des fonctions de covariance QZ comme on peut le voir sur la figure 8.20.

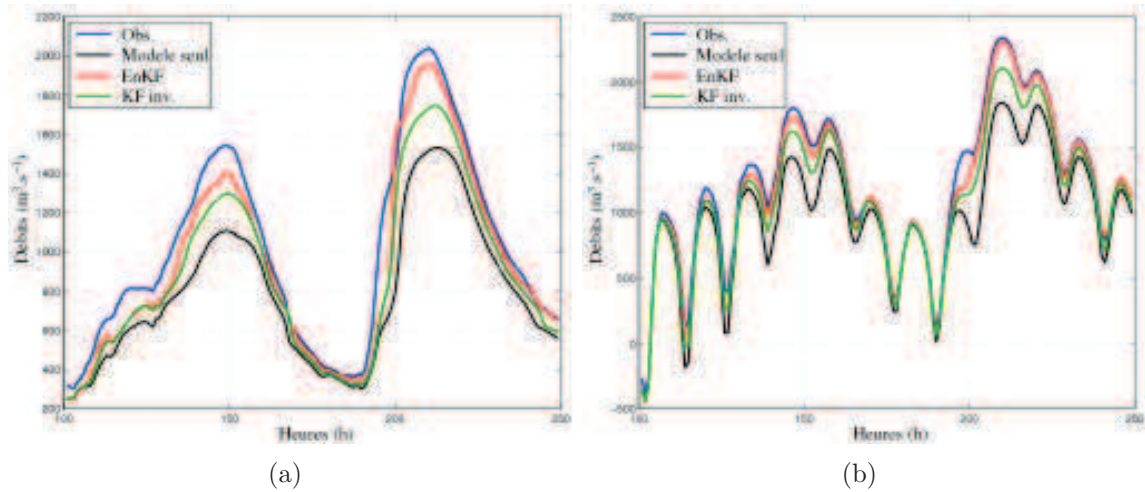


FIG. 8.20 – Séries temporelles des débits analysés à (a) Peyrehorade et (b) Urt.

Afin d'estimer quantitativement les performances de l'EnKF en comparaison du KF invariant et du modèle sans assimilation nous avons reporté dans les tableaux 8.1 et 8.2

les scores suivants calculés après assimilation sur les hauteurs d'eau et les débits (sur la durée de l'épisode considéré ici soit 150 heures) : RMSE, corrélation, biais et critère de Nash.

Station	Algo.	RMSE	Corrélation	Biais	Nash
Peyrehorade	EnKF	0.169	0.992	0.117	0.97
	KF inv.	0.211	0.993	0.172	0.954
	Modèle	0.549	0.959	0.448	0.735
Urt	EnKF	0.067	0.996	0.038	0.99
	KF inv.	0.082	0.997	0.062	0.985
	Modèle	0.278	0.963	0.216	0.836

TAB. 8.1 – Scores sur les hauteurs d'eau.

Station	Algo.	RMSE	Corrélation	Biais	Nash
Peyrehorade	EnKF	120.06	0.985	83.759	0.941
	KF inv.	129.83	0.99	103.70	0.932
	Modèle seul	287.80	0.968	235.94	0.719
Urt	EnKF	64.124	0.997	44.965	0.987
	KF inv.	76.389	0.997	64.139	0.981
	Modèle seul	286.17	0.962	234.46	0.775

TAB. 8.2 – Scores sur les débits.

Les scores calculés dans les tableaux 8.1 et 8.2 sont relativement bons notamment pour la corrélation et le critère de Nash, que ce soit avec les algorithmes d'assimilation de données ou alors avec le modèle seul sans assimilation. Pour le biais et la RMSE ces scores sont assez élevés pour les deux algorithmes d'assimilation de données.

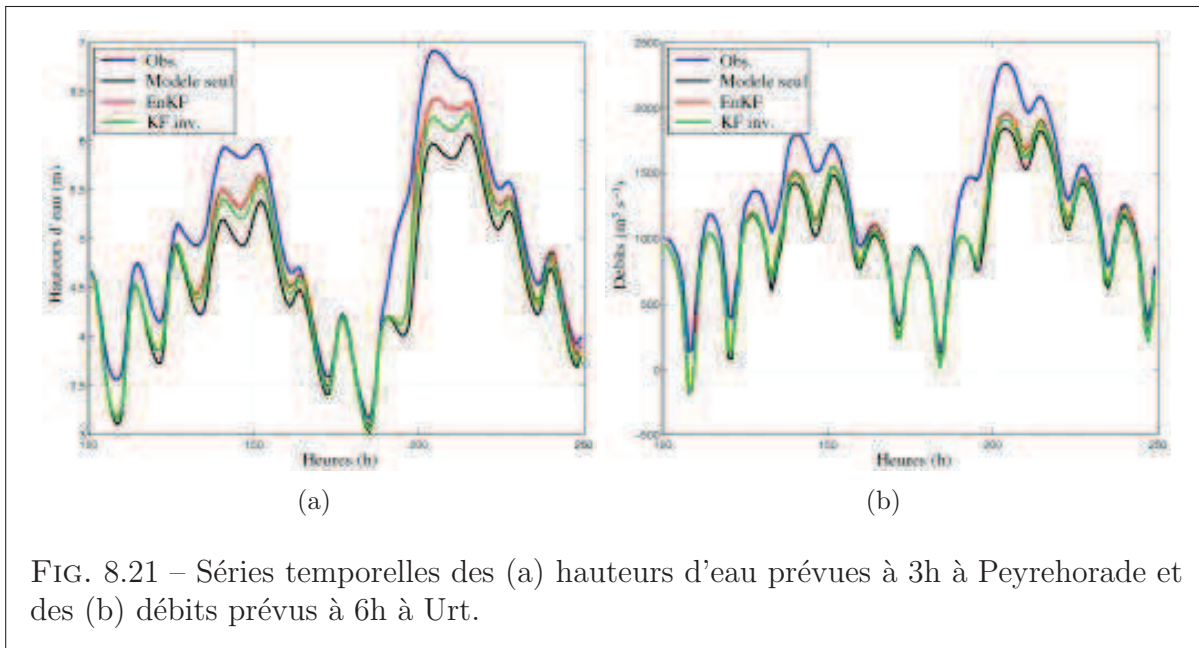
On notera également que les résultats de l'assimilation sont meilleurs à Urt qu'à Peyrehorade qui est pourtant la station où on procède à l'assimilation, cela s'explique par le fait que le modèle seul est meilleur à cette station qu'à Peyrehorade. Conformément aux courbes présentées en figure 8.19 et 8.20 l'EnKF réalise pratiquement dans tous les cas le meilleur score et surtout les deux algorithmes d'assimilation de données réalisent tous deux des scores nettement meilleurs que le modèle seul sans assimilation.

Comparaison des résultats en mode prévision

Il convient de rappeler que le cyclage des d'assimilation-prévision pour l'EnKF a été décrit en section 8.1.3 avec la figure 8.5. Pour le KF invariant un schéma similaire reste valable mais avec un seul membre cette-fois-ci.

De la même manière qu'en prévision opérationnelle sur l'Adour durant la phase de prévision les forçages amont sont fixés constants égaux aux derniers forçages disponibles dans cette expérience les forçages amonts sont fixés constants durant la phase de prévision. En mode prévision la correction apportée par l'assimilation va être propagée de l'amont vers l'aval et par suite les résultats des hauteurs d'eau et des débits vont se dégrader pour se rapprocher des résultats du modèle seul comme cela est illustré sur la figure 8.21. Là encore les résultats calculés avec l'EnKF sont meilleurs que ceux calculés avec le KF invariant et ce quelle que soit l'échéance de prévision, cela est tout à fait normal puisqu'au temps d'analyse l'EnKF a déjà de meilleurs résultats que le KF invariant. Nous ne reporterons pas ici les scores calculés pour les algorithmes d'assimilation et pour le modèle seul en phase de prévision, ceux-ci en effet se dégradent durant toute la phase de prévision pour se rapprocher des scores du modèle seul qui sont reportés dans les tableaux 8.1 et 8.2.

Le fait que les résultats se dégradent en mode prévision est dû au fait qu'avec ces deux méthodes d'assimilation de données on corrige la ligne d'eau et donc la condition initiale de la prévision alors que l'erreur vient des forçages latéraux qui ne sont pas connus. Or la condition initiale est propagée vers l'aval par Mascaret et donc « évacuée » au fil du temps. Des méthodes d'assimilation de données corrigeant les conditions limites (voir [Jean-Baptiste et al., 2011, Ricci et al., 2011]) permettraient d'améliorer les résultats en mode prévision (voir [Ricci et al., 2011]).



Impact de l'assimilation de données sur les longueurs de portée

Comme on l'a vu dans le chapitre 7 portant sur l'émulation de l'EnKF l'assimilation de données dans ce cas a une influence sur les longueurs de portée des fonctions de

corrélation avec notamment une longueur de portée aval plus petite que la longueur de portée amont. Il est donc normal de se poser la question de savoir si l'assimilation de données a un impact comparable sur l'Adour avec un modèle Saint-Venant. Comme on peut le voir sur la figure 8.22 l'assimilation introduit une plus grande variabilité des longueurs de portée (amont comme aval) cassant le rapport entre l'évolution temporelle des hauteurs d'eau et/ou des débits à la station d'observation et l'évolution temporelle des longueurs de portée diagnostiquées à la même station. On observe néanmoins que les longueurs de portée aval sont quelque peu réduites pendant le pic de crue avec l'EnKF en comparaison des longueurs de portée aval calculées sans assimilation. Ce résultat est à rapprocher des résultats mis en évidence dans le chapitre 7 où l'assimilation de données avait pour effet de réduire la longueur de portée aval des fonctions de corrélation.

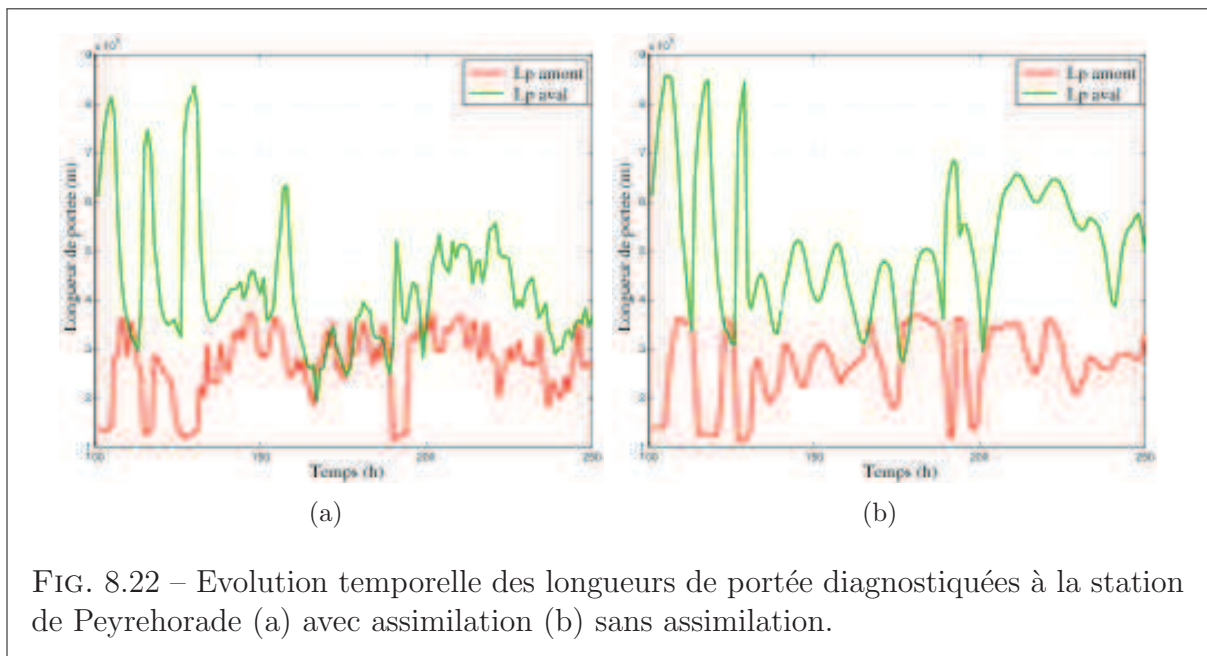


FIG. 8.22 – Evolution temporelle des longueurs de portée diagnostiquées à la station de Peyrehorade (a) avec assimilation (b) sans assimilation.

8.3.3 Discussion

Cette expérience avait pour objectif de mettre en lumière l'intérêt pour la correction des lignes d'eau de la grande extension spatiale des fonctions de covariance diagnostiquées avec l'EnKF en comparaison de la correction de celles-ci avec des fonctions de covariance stationnaires comme celles utilisées dans l'article [Ricci et al., 2011]. A ce titre les résultats sont assez intéressants. Mais on a mis également en lumière la pertinence des fonctions de covariance QZ qui permettent de traduire une information sur les hauteurs d'eau, apportée par les observations de hauteurs d'eau, en une information sur les débits puisque comme on l'a vu la correction sur les débits à la station de Peyrehorade mais aussi ailleurs sur le domaine bien que biaisée est relativement bonne. Ces résultats sont donc intéressants puisqu'ils montrent qu'on peut avoir une bonne connaissance de la ligne d'eau aux stations d'observations mais aussi ailleurs sur le réseau.

On pourra s'étonner que les résultats du KF invariant ne soient pas meilleurs pour un algorithme d'assimilation de données utilisé pour la prévision opérationnelle.

Cela s'explique très bien par le fait que l'écart-type d'erreur d'ébauche choisi est égal à la moyenne des écarts-types d'erreur d'ébauche calculés avec l'EnKF sur la durée de l'épisode de crue ainsi les résultats du KF sont moins bons que ceux de l'EnKF pendant les pics de crue. En prévision opérationnelle l'écart-type d'erreur d'ébauche est fixé à 0.25 tandis que pour cette expérience il a été fixé à 0.067, les résultats en opérationnel sont donc bien meilleurs que ceux exposés ici pour le KF invariant.

Sur cet exemple les états analysés avec l'EnKF sont relativement corrects que ce soit pour les débits comme pour les hauteurs d'eau or selon les épisodes de crue considérés les états analysés avec l'EnKF peuvent être entachés d'une erreur plus importante. Cela est gênant car d'une part même si ces résultats peuvent paraître corrects une erreur de l'ordre de quelques centimètres dans la connaissance des hauteurs d'eau peut se traduire par une erreur sur la connaissance des zones inondées. L'assimilation devrait donc viser à donner les meilleurs résultats possibles pour avoir la meilleure connaissance possible de l'état hydraulique et des zones inondées. C'est pourquoi on ne peut pas se contenter d'avoir une erreur de 10 cm sur les hauteurs d'eau analysées et qu'il faut réduire au maximum cette erreur. Cela est gênant d'autre part car les états analysés forment les conditions initiales pour les prévisions. Ce qui signifie que des états analysés entachés d'erreur seront à l'origine de moins bonnes prévisions. Or les états analysés par l'EnKF sont entachés d'erreur justement lorsque l'assimilation est biaisée (au sens où en valeur absolue la moyenne des innovations est au moins de l'ordre de grandeur de l'écart-type des innovations, [Trudel et al., 2014]) comme on peut le voir sur la figure 8.23.

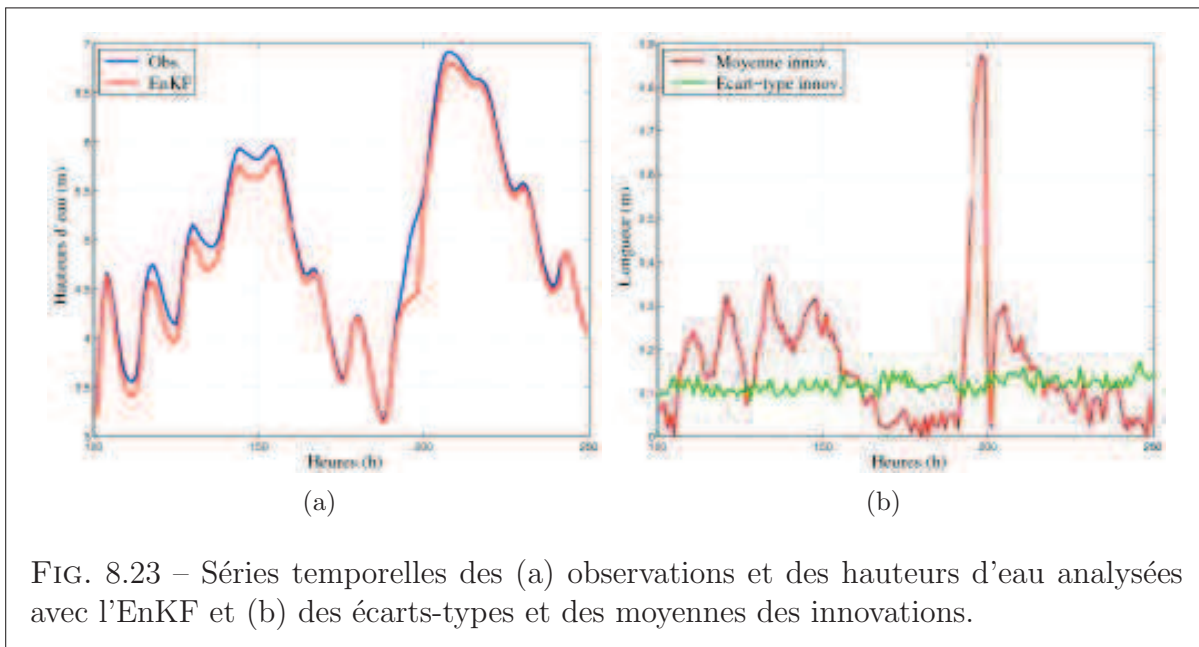


FIG. 8.23 – Séries temporelles des (a) observations et des hauteurs d'eau analysées avec l'EnKF et (b) des écarts-types et des moyennes des innovations.

L'ensemble peut donc être sous-dispersif. Une solution *a priori* évidente mais trompeuse consisterait à augmenter de façon significative la variance des forçages amonts de

manière à augmenter la variance d'erreur d'ébauche au niveau des observations et ainsi améliorer les résultats de l'assimilation. Cette solution n'est pas viable car si la variance des forçages amonts est trop importante par rapport au forçage lui-même on risque soit de créer des instabilités numériques qui mèneront à un « plantage » du code Mascaret soit de se rapprocher de la frontière entre régime fluvial et torrentiel, or dans le noyau de calcul qu'on utilise Mascaret ne gère pas les changements de régime ce qui se traduit là encore par un plantage de Mascaret. Une solution consiste donc à compenser la sous-dispersivité de l'ensemble en ayant recours aux critères de [Desroziers et al., 2005] et aux méthodes d'inflation associées tout en procédant à quelques adaptations de la méthode au cas de l'Adour. Nous utiliserons également les critères de [Desroziers et al., 2005] pour estimer les covariances d'erreur d'observation. Celles-ci ont en effet été fixées de manière arbitraire à la valeur de 0.1 m jusqu'à maintenant. C'est ce à quoi nous nous intéresserons dans la prochaine section.

D'autre part, comme on l'a vu précédemment, du fait de la propagation de l'état analysé par le modèle tout au long de la phase de prévision on perd le bénéfice de la correction apportée par l'assimilation de données, comme cela est illustré en figure 8.21. Ne disposant pas de modèles permettant de prévoir les débits à l'amont du réseau Adour une manière d'améliorer les résultats en prévision serait de combiner la méthode de correction des forçages amonts exposée dans [Ricci et al., 2011] avec l'EnKF. Néanmoins au-delà du temps de transfert des forçages amonts jusqu'aux stations l'amélioration apportée tient à la correction des forçages amonts et non à l'EnKF c'est pourquoi nous ne reviendrons pas sur cette méthode dans la suite.

8.4 Application des critères de Desroziers

Dans cette section nous nous intéressons à l'application des critères de [Desroziers et al., 2005] au réseau Adour pour améliorer les résultats de l'assimilation en termes d'états analysés. Pour cela nous suivons en partie la démarche suivie par [Li et al., 2009] pour utiliser les critères de [Desroziers et al., 2005] dans le contexte d'un filtre de Kalman d'ensemble déterministe. Les critères de [Desroziers et al., 2005] sont en fait des diagnostics *a posteriori* de consistance des matrices de covariance d'erreur d'ébauche, d'observation, d'analyse... Tout comme [Li et al., 2009] parmi ces critères nous ne retiendrons que que les diagnostics portant sur les matrices de covariance d'erreur d'ébauche et d'observation.

8.4.1 Estimation des covariances d'erreur d'observation

Le critère de Desroziers sur les covariances d'erreur d'observation

Le but de cette section est d'estimer les covariances d'erreur d'observation en utilisant le diagnostic de [Desroziers et al., 2005] sur les covariances d'erreur d'observation. [Desroziers et al., 2005] ont en effet montré que si les matrices $\mathbf{B}$ et $\mathbf{R}$ spécifiées dans l'égalité :

$$\mathbf{H}\mathbf{K} = \mathbf{H}\mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}$$

sont les vraies matrices de covariance d'erreur d'ébauche et d'observation alors l'égalité (8.2) est vérifiée :

$$\mathbf{R} = \langle \mathbf{d}_{o-a} \mathbf{d}_{o-b}^T \rangle \quad (8.2)$$

où $\langle . \rangle$ désigne l'espérance mathématique, $\mathbf{d}_{o-a} = (y^o - \mathbf{H}x^{a,1}, \dots, y^o - \mathbf{H}x^{a,N})$ et $\mathbf{d}_{o-b} = (y^o - \mathbf{H}x^{b,1}, \dots, y^o - \mathbf{H}x^{b,N})$.

L'expression (8.2) permet donc d'établir un diagnostic des covariances d'erreur d'observation en fonction des observations, des états analysés et des ébauches. Dans ce travail nous avons suivi la démarche utilisée par [Li et al., 2009] qui estime la variances d'observation avec l'expression (8.2) et qui complète cette estimation avec une méthode de lissage temporelle que nous exposons ci-dessous.

Démarche

La démarche suivie ici est celle proposée par [Miyoshi, 2005] et suivie par [Li et al., 2009]. On procède à une suite de plusieurs cycles d'assimilation et après chaque cycle on estime la matrice des covariances d'erreur d'observation $\mathbf{R}$ avec la formule (8.2). Dans notre cas les stations d'observation sont des stations de mesure *in situ* donc la matrice $\mathbf{R}$ est une matrice diagonale et on suppose que la variance d'erreur d'observation est identique à toutes les stations ce qui nous permet d'écrire : $\mathbf{R} = \sigma_o \mathbf{Id}$ où $\mathbf{Id}$ est la matrice identité et σ_o la variance d'erreur d'observation. Dans la suite nous cherchons donc simplement à estimer σ_o .

Pour palier aux problèmes de sous-échantillonnage liés à l'estimation statistique de σ_o nous avons suivi la démarche de lissage temporel des variances d'erreur d'observation estimées suivie par [Li et al., 2009] basée sur une approche scalaire du filtre de Kalman classique. On considère les variances estimées à un cycle d'assimilation en utilisant l'expression (8.2) σ_o comme une « observation » des vraies variances d'erreur d'observation. On combine alors cette valeur avec l'ancienne estimation des variances d'erreur d'observation que l'on note σ^f selon les équations du filtre de Kalman qui nous donne la nouvelle estimation des variances d'erreur d'observation σ^a :

$$\alpha^a = \frac{\nu^o \alpha^f + \nu^f \sigma_o}{\nu^o + \nu^f}$$

où ν^o et ν^f désignent respectivement la variance de l'erreur associée à σ_o et α^f .

Suivant les équations du filtre de Kalman la variance d'erreur d'analyse est donnée par :

$$\nu^a = \left(1 - \frac{\nu^f}{\nu^f + \nu^o} \right) \nu^f$$

Si en suivant [Li et al., 2009] on suppose une persistance du modèle (d'estimation des variances d'erreur d'observation) et une erreur dans la persistance des variances d'erreur modèle alors on écrit :

$$\alpha_{t+1}^f = \alpha_t^a \quad \nu_{t+1}^f = \kappa \nu_t^a$$

où $\kappa > 1.0$ est un paramètre d'oubli.

[Miyoshi, 2005] montre que l'estimation des variances d'erreur d'observation ne dépend pas des valeurs choisies pour ν^o et κ . Pour le choix des paramètres ν^o et κ [Li et al., 2009] a suivi [Miyoshi, 2005] et a pris pour valeurs : $\nu^o = 1$ et $\kappa = 1.03$. Dans ce travail nous avons choisi : $\nu^o = 0.85$ et $\kappa = 1.03$. Le choix de $\nu^o = 0.85$ a été fait de manière à donner plus de poids dans l'évaluation courante de la variance d'erreur d'observation α^a à la dernière estimation σ_o que dans les travaux de [Li et al., 2009, Miyoshi, 2005]. On a noté en effet que l'algorithme semblait converger légèrement plus vite avec cette valeur de ν^o .

Il faut enfin préciser que pour être fidèle à [Li et al., 2009] l'algorithme d'assimilation de données ensembliste utilisé dans le cadre de l'estimation des variances d'observation n'est pas la version stochastique de l'EnKF présentée en section 3.7.5 mais une version déterministe, l'ETKF présenté en section 3.7.6.

Des expériences d'assimilation de données pour estimer les variances d'erreur d'observation

Dans notre cas les observations étant *in situ* on considère que la matrice $\mathbf{R}$ est diagonale et que tous les termes diagonaux sont égaux ; c'est pourquoi dans la suite pour estimer les variances d'erreur d'observation nous ne ferons le diagnostic que sur une seule station d'observation.

Pour montrer que le critère de Desroziers donné par l'expression (8.2) et la démarche exposée ci-dessus nous permettent d'estimer les variances d'erreur d'observation nous nous livrons dans un premier temps à des expériences simples d'assimilation de données d'observations synthétiques de hauteurs d'eau. Pour chaque expérience nous propageons des forçages le long du réseau Adour et ajoutons un bruit suivant une loi centrée et de variance prescrite *a priori*. Nous procédons ensuite à l'assimilation de ces données et estimons les variances d'erreur d'observation avec l'expression (8.2) et en appliquant la démarche ci-dessus. Dans notre cas, les variances d'observation étant constantes l'algorithme utilisé finit par converger vers la bonne valeur comme cela est illustré sur la figure 8.24.

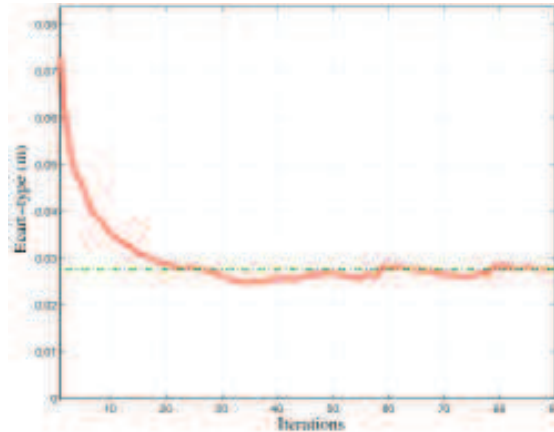


FIG. 8.24 – Evolution des écarts-types d'erreur d'observation estimés (courbe rouge) et prescrits (ligne verte). Chaque itération correspond en fait à un cycle d'assimilation.

Les résultats de ces expériences sont reportés dans le tableau 8.3.

Ecart-type prescrit (m)	0.025	0.075	0.125
Ecart-type estimé (m)	0.0269	0.0759	0.1297
Erreur relative	7.6 %	1.2 %	3.8 %

TAB. 8.3 – Ecarts-types prescrits et estimés avec l'assimilation.

Le tableau 8.3 montre que l'on est en mesure d'estimer avec une relative précision les variances d'erreur d'observation. Il ne nous reste plus qu'à appliquer la méthode utilisée ici sur des données d'observations réelles pour avoir une estimation des variances d'erreur d'observation sur l'Adour.

Dans les expériences précédentes, par construction, les forçages n'étaient pas biaisés et l'erreur sur les hauteurs d'eau à la station où on a procédé à l'assimilation correspondait donc à l'erreur d'observation. Pour estimer les variances d'erreur d'observation sur des données réelles on choisit donc un épisode de basses eaux et une station (celle de Urt) pour lesquels on sait que les forçages sont pas ou très peu biaisés et on suppose que l'erreur entre le modèle et les observations correspond à l'erreur d'observation. On applique la démarche présentée ci-dessus, à convergence l'écart-type d'erreur d'observation estimé est de 0.051 m . Elle est donc nettement plus petite que celle initialement prescrite sur l'Adour et qui était, on le rappelle, de 0.1 m . Dans la suite nous utiliserons donc la valeur de 0.051 m dans nos algorithmes d'assimilation de données, ce qui devrait permettre d'améliorer les résultats de l'EnKF pendant les pics de crue en comparaison de ceux présentés en section 8.3.

8.4.2 Estimation des covariances d'erreur d'ébauche et inflation

Le critère de Desrozier sur les covariances d'erreur d'ébauche

Dans cette section nous nous intéressons à l'utilisation du critère de Desrozier sur les covariances d'erreur d'ébauche pour compenser les problèmes de sous-dispersivité évoqués en fin de section 8.3. Le critère de Desrozier sur les covariances d'erreur d'ébauche postule que si les matrices de covariance d'erreur d'ébauche et d'observation, $\mathbf{B}$ et $\mathbf{R}$, sont correctement spécifiées alors la relation (8.3) est vérifiée :

$$\langle \mathbf{d}_{o-b} \mathbf{d}_{o-b}^T \rangle = \mathbf{H} \mathbf{B} \mathbf{H}^T + \mathbf{R} \quad (8.3)$$

C'est sur ce critère que s'appuient une partie des méthodes dites d'inflation multiplicative évoquées en section 3.7.7 et que l'on présente plus en détail dans la section suivante.

Inflation

Si l'ensemble est sous-dispersif alors les covariances d'erreur d'ébauche sont sous-estimées et l'égalité (8.3) n'est pas vérifiée. Pour palier à ce problème des méthodes dites d'inflation multiplicative permettant d'augmenter les covariances d'erreur d'ébauche ont été développées (Anderson and Anderson, 1999). Il existe également des méthodes d'inflation dites additives mais nous ne nous intéresserons pas à celles-ci dans ce travail. L'inflation multiplicative (par la suite nous dirons seulement inflation) consiste à multiplier l'ensemble des anomalies par un facteur $\lambda > 1$ (d'où la dénomination d'inflation multiplicative) de manière à augmenter les covariances d'erreur d'ébauche :

$$\tilde{x}^k - \bar{x} = \lambda(x^k - \bar{x}) \quad (8.4)$$

où $\tilde{x}^k$ est le membre correspondant au membre x^k obtenu par inflation.

Du fait de la relation (3.42) la matrice des covariances d'erreur d'ébauche après inflation des membres (que l'on note $\tilde{\mathbf{B}}$) est égale à $\lambda^2 \mathbf{B}$. En effet :

$$\tilde{\mathbf{B}} = \frac{1}{N-1} \sum_{i=1}^N (\tilde{x}_k^i - \bar{x}_k) (\tilde{x}_k^i - \bar{x}_k)^T = \frac{\lambda^2}{N-1} \sum_{i=1}^N (x_k^i - \bar{x}_k) (x_k^i - \bar{x}_k)^T = \lambda^2 \mathbf{B}$$

En augmentant les covariances d'erreur d'observation on donne plus de poids aux observations dans le processus d'assimilation et mécaniquement l'inflation permet d'améliorer les résultats analysés en comparaison de l'assimilation sans inflation.

Le facteur d'inflation λ ne doit pas être choisi au hasard mais de manière à réduire la différence entre les membres de l'égalité (8.3) si celle-ci n'est pas vérifiée. Dans le cas où l'on ne dispose que d'une seule station d'observation les matrices $\mathbf{B}$ et $\mathbf{R}$ se réduisent à des scalaires, respectivement σ_b^2 et σ_o^2 . On peut alors facilement déduire l'expression de λ^2 de l'égalité (8.3) :

$$\lambda^2 = \frac{\langle \mathbf{d}_{o-b} \mathbf{d}_{o-b}^T \rangle - \sigma_b^2}{\sigma_o^2} \quad (8.5)$$

Dans le cas où l'on a plusieurs stations d'observation [Wang and Bishop, 2003] proposent d'estimer λ^2 en ne considérant que les termes diagonaux de l'égalité (8.3), soit :

$$\lambda^2 = \frac{\mathbf{d}_{o-b}^T \mathbf{d}_{o-b} - \text{Tr}(\mathbf{H}\mathbf{B}\mathbf{H}^T)}{\text{Tr}(\mathbf{R})} \quad (8.6)$$

Cependant dans le cadre de l'application de l'inflation à l'hydraulique il peut être impossible d'appliquer le facteur λ à chaque membre et de manière uniforme spatialement comme cela est suggéré par l'expression (8.4). C'est le cas en particulier sur l'Adour où des expériences numériques ont montré qu'un facteur λ trop grand (calculé par exemple à une station où l'erreur entre le modèle et l'observation à un temps donné est important) pouvait entraîner des instabilités numériques aux limites du domaine (du fait de la discontinuité entre les débits aux limites du domaine et les forçages) ou des changements de régime (passage de fluvial à torrentiel) déjà évoqués précédemment et qui conduisent à des plantages du code Mascaret. Pour contourner ce problème il faut pouvoir spatialiser l'inflation de manière consistante avec le modèle physique, c'est ce que nous exposons dans le paragraphe suivant.

Spatialisation de l'inflation

Pour éviter les potentiels plantages du code Mascaret liés à l'utilisation d'un facteur d'inflation identique sur tout le domaine il faut spatialiser l'inflation de manière consistante avec le modèle. Pour cela on va utiliser les fonctions de corrélation que nous avons présentées en section 8.2. On considère à un temps donné le facteur λ diagnostiqué à une station d'observation par l'expression (8.5) et la fonction de corrélation C relative à cette station. On suppose que $C(x) \xrightarrow{x \rightarrow \pm\infty} 0$. Alors plutôt que d'utiliser la formule (8.4) pour multiplier les anomalies on utilisera la formule suivante :

$$\tilde{x}_i^k - \bar{x}_i = (1 + (\lambda - 1)|C_i|) \cdot (x_i^k - \bar{x}_i) \quad (8.7)$$

où i désigne le i -ème nœud de calcul du domaine.

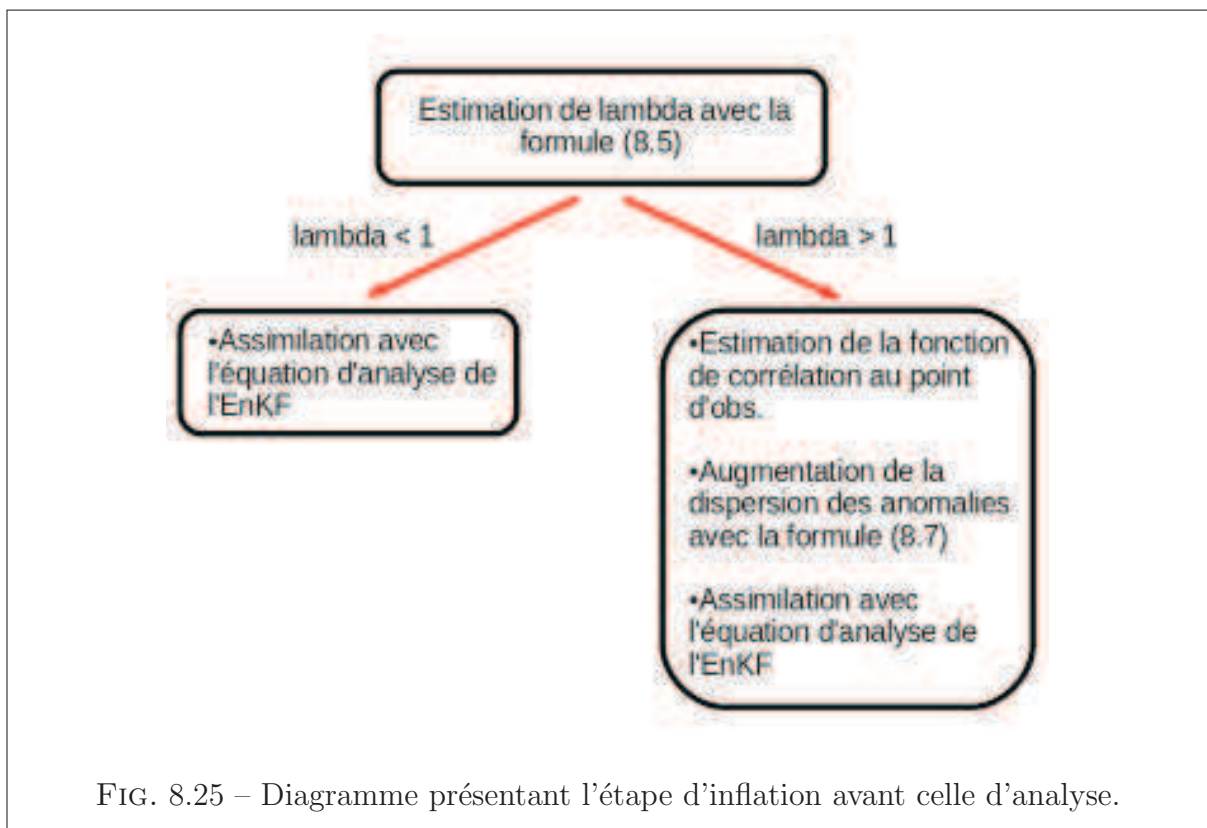
Si on note i_o le numéro du nœud de calcul où se situe la station avec laquelle on a calculé λ alors $C_{i_o} = 1$, $\tilde{x}_{i_o}^k - \bar{x}_{i_o} = \lambda(x_{i_o}^k - \bar{x}_{i_o})$ et on retrouve bien l'égalité (8.5). Au contraire si i est tel que $C_i \approx 0$ alors $\tilde{x}_i^k - \bar{x}_i = x_i^k - \bar{x}_i$. Ce qui signifie qu'en dehors de la zone de corrélation relative à la station où l'on a calculé le facteur λ les anomalies sont préservées. Dans cette zone le facteur d'inflation $1 + (\lambda - 1)|C_i|$ dépend des corrélations et est donc consistant avec le modèle physique.

L'écriture (8.7) présente deux autres propriétés intéressantes. Tout d'abord elle préserve la moyenne des membres et n'introduit donc pas de biais dans l'ensemble. D'autre part comme $|C_i| \leq 1, \forall i$ on peut montrer que le facteur $1 + (\lambda - 1)|C_i|$ est positif et donc l'écriture (8.7) préserve les corrélations spatiales entre les membres mais aussi les corrélations entre variables d'état.

Dans la suite l'EnKF combiné avec une méthode d'inflation sera désigné par l'acronyme IEnKF pour Inflated Ensemble Kalman Filter.

Démarche pour appliquer l'inflation

Dans la pratique à chaque cycle d'assimilation nous estimons le facteur λ calculé avec la formule (8.5). Celui-ci peut être inférieur à 1 dans ce cas nous ne procédons pas à l'augmentation de la dispersion des anomalies, dans le cas contraire on applique les étapes présentées ci-dessus. Pour la clarté du propos nous avons regroupé les différentes étapes dans le diagramme donné en figure 8.25.



Contrairement aux covariances d'erreur d'observation dont on a supposé qu'elles étaient constantes le facteur λ estimé peut avoir une importante variabilité temporelle. Celui-ci dépend en effet du terme $\langle \mathbf{d}_{o-b} \mathbf{d}_{o-b}^T \rangle$ dans l'expression (8.5) et est donc fortement dépendant de l'erreur sur les forçages. Aussi lorsque celle-ci n'est pas importante l'inflation n'est pas nécessaire. Au contraire, pendant les épisodes de forte crue l'erreur sur les forçages peut être importante du fait des apports latéraux non jaugés qui ne sont pas pris en compte dans le modélisation. Dans ce cas l'inflation est nécessaire pour réduire l'impact de l'erreur des forçages sur les états analysés.

Enfin, sauf mention du contraire, nous avons toujours calculé le facteur λ à la station de Peyrehorade puisque c'est en règle général la station où la différence entre le modèle et les observations est la plus importante.

8.4.3 Une expérience d'assimilation de données synthétiques

Pour attester de l'intérêt des méthodes d'estimation des covariances d'erreur d'observation et d'inflation présentées ci-dessus nous nous livrons à la même expérience d'assimilation de données synthétiques de hauteurs d'eau que celle présentée en section 8.3. C'est-à-dire qu'on assimile les observations de hauteurs d'eau à la station de Peyrehorade puis on compare les résultats aux observations aux stations de Peyrehorade et de Urt pour les hauteurs d'eau et pour les débits.

Pour cet épisode les séries temporelles des hauteurs d'eau ainsi que des débits analysés aux stations de Peyrehorade et de Urt sont reportés en figure 8.26 et 8.27.

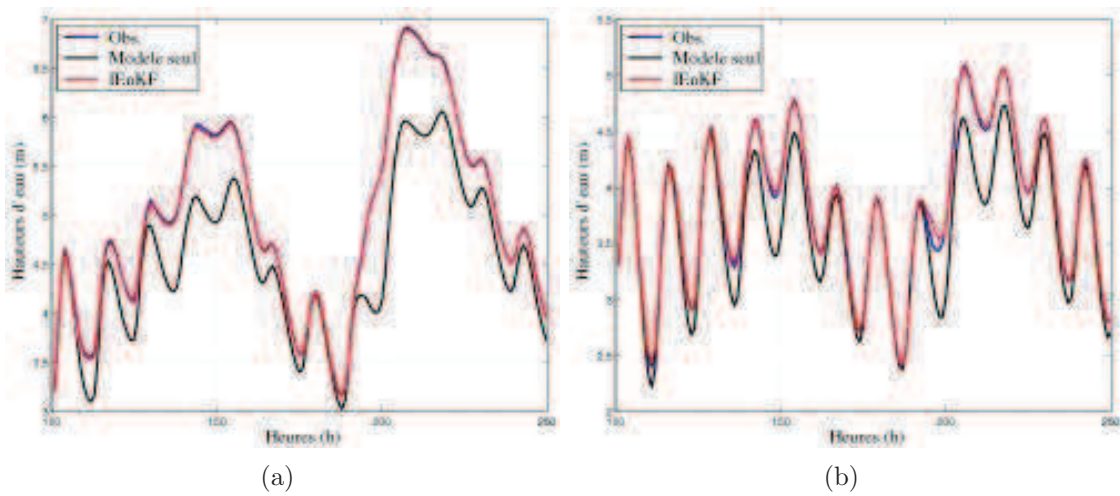
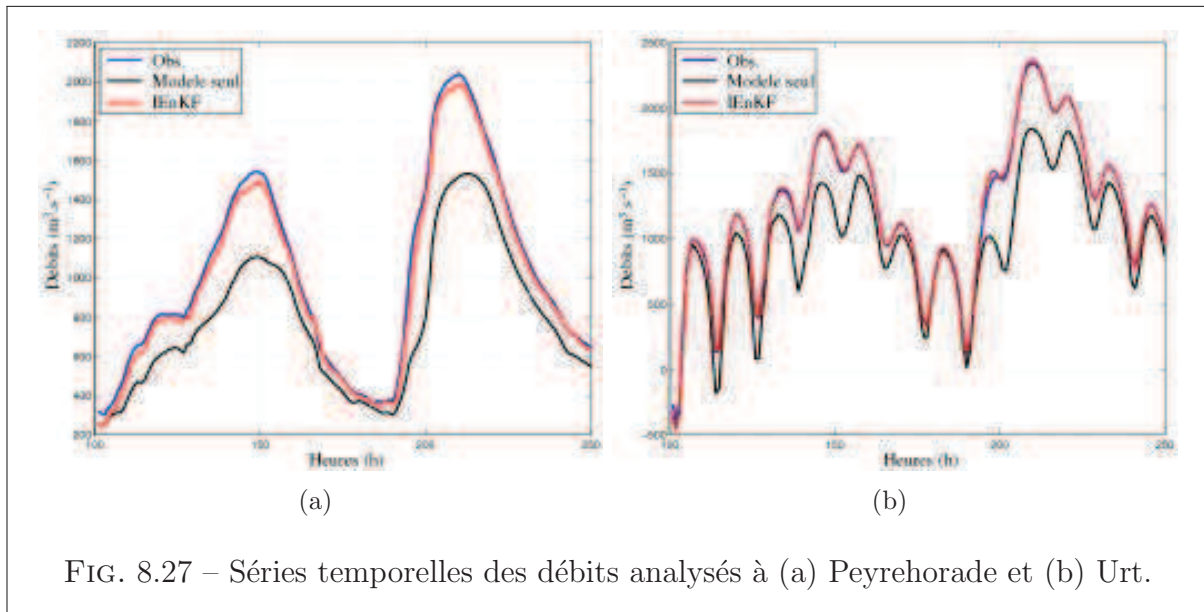


FIG. 8.26 – Séries temporelles des hauteurs d'eau analysées à (a) Peyrehorade et (b) Urt.



Comme on peut le constater visuellement sur les figures 8.26 et 8.27 l'IEEnKF donne de meilleurs résultats en termes d'états analysés que l'EnKF que ce soit à la station de Peyrehorade où on a procédé à l'assimilation ou à la station de Urt. Ce constat est conforté par le calcul des scores que l'on reporte dans les tableaux 8.4 et 8.5.

Station	Algo.	RMSE	Corrélation	Biais	Nash
Peyrehorade	IEEnKF	0.025	0.999	0.021	0.999
	EnKF	0.169	0.992	0.117	0.97
	Modèle seul	0.549	0.959	0.448	0.735
Urt	IEEnKF	0.032	0.999	-0.014	0.998
	EnKF	0.067	0.996	0.038	0.989
	Modèle seul	0.28	0.963	0.216	0.836

TAB. 8.4 – Scores sur les hauteurs d'eau.

De la même manière que ce qui a été évoqué en section 8.3, en mode prévision, la ligne d'eau calculée avec l'assimilation est propagée par le modèle et au fil du temps on perd le bénéfice de l'assimilation et les hauteurs d'eau et les débits prédits s'approchent de plus en plus de ceux calculés par le modèle seul sans assimilation.

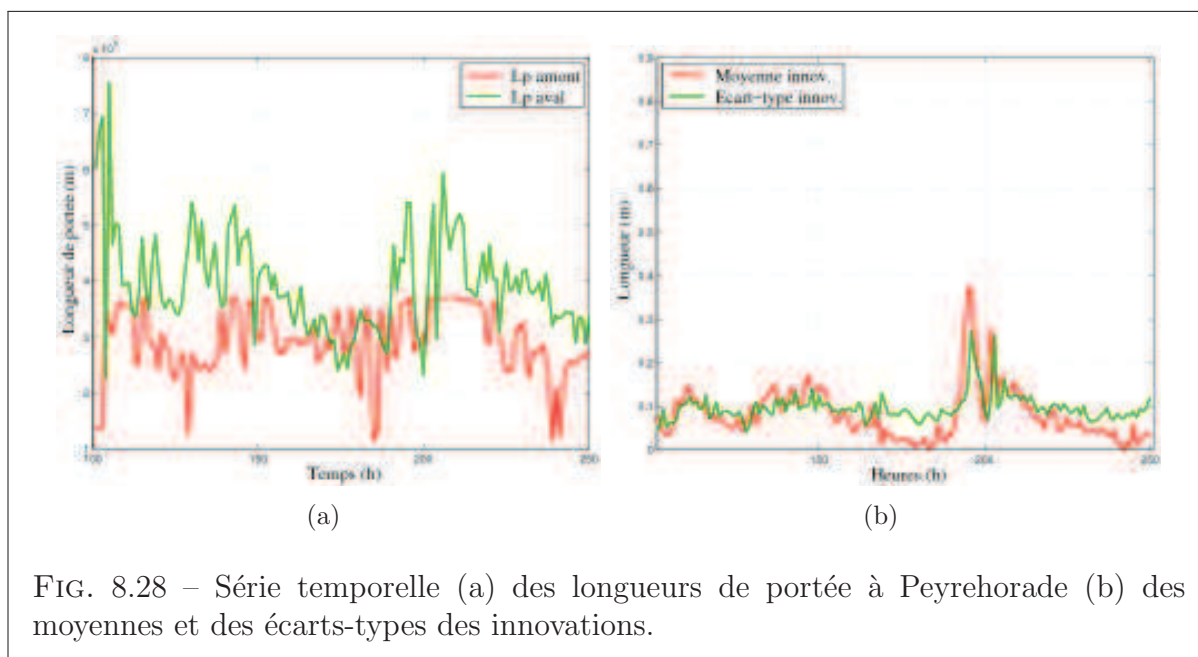
D'autre part l'inflation ne semble pas avoir d'impact significatif sur les longueurs de portée par rapport à l'EnKF comme on peut le constater sur les figures 8.28-(a) et 8.22-(a) sauf au début de l'épisode de crue où les longueurs de portée aval sont réduites par rapport à celles diagnostiquées avec l'EnKF. Enfin on constate grâce à la figure 8.28-(b) que l'IEEnKF permet de réduire le « biais d'assimilation » au sens où la différence entre la moyenne des innovations et l'écart-type des innovations est grandement réduite par rapport à la figure 8.23-(b).

Station	Algo.	RMSE	Corrélation	Biais	Nash
Peyrehorade	IEnKF	44.589	0.998	33.596	0.992
	EnKF	120.06	0.985	83.759	0.941
	Modèle seul	287.80	0.968	235.94	0.719
Urt	IEnKF	24.374	0.999	-8.507	0.998
	EnKF	64.124	0.997	44.965	0.987
	Modèle seul	286.17	0.962	234.46	0.775

TAB. 8.5 – Scores sur les débits.

Ces résultats sont bons mais il faut néanmoins les tempérer. Ils ont en effet été obtenus dans le cadre d'une expérience d'assimilation de données synthétiques, les corrélations spatiales mais aussi multivariées des observations sont produites par le modèle hydraulique Mascaret et ne correspondent donc pas à un modèle physique réel (comme cela aurait été le cas de véritables observations). Il est donc normal que l'assimilation nous permette de reproduire ces corrélations et ainsi d'obtenir les résultats que l'on a exhibés dans cette section.

A ce stade nous n'avons pas encore procédé à des expériences sur des observations réelles mais on peut d'ores et déjà supposer que les covariances et corrélations d'erreur d'ébauche estimées par le modèle sont fausses, le réseau Adour maritime étant modélisé sans lit majeur, ce qui ne va pas être sans poser problème pendant les forts pics de crue. Pour conforter les points forts de notre démarche mais aussi mettre en lumière les limites de celle-ci nous devons procéder à des expériences d'assimilation de donnée réelles. C'est l'objet de la prochaine section.



8.4.4 Expériences d'assimilation sur des données réelles

Dans cette section nous nous intéressons à l'application de l'IEnKF sur des jeux de données réelles qui nous ont été fournies par le SCHAPI et le SPC GAD. Il s'agit de crues relativement récentes puisque toutes ces crues sauf une ont eu lieu après la mise en place de la station d'observation de Bayonne Pont-Blanc en 2009. Nous avons reporté en annexe (C) les données utilisées pour ces épisodes, données de forçages et d'observations de hauteurs d'eau.

Dans ces expériences nous avons fait le choix de ne pas utiliser toutes les stations pour pratiquer l'assimilation. En effet de manière similaire aux expériences jumelles menées dans les sections 8.3 et 8.4.3 nous souhaitons tester la capacité de l'IEnKF à corriger les hauteurs d'eau aux stations d'observation où on assimile et aux autres stations sur le domaine. Nous avons décidé de pratiquer l'assimilation à des stations présentant une faible corrélation des erreurs d'état du modèle (ou erreur d'ébauche), nous avons retenu la station de Peyrehorade sur l'Adour car c'est la station où l'erreur est la plus importante pendant les pics de crue, c'est aussi la station sur laquelle nous allons pratiquer l'inflation. Et nous avons retenu la station de Pont-Blanc sur la Nive.

Il faut rappeler que pour valider ces expériences nous ne disposons que d'observations de hauteurs d'eau. En effet ni le SCHAPI, ni le SPC GAD ne disposent de séries temporelles (même très restreintes) d'observations de débits sur le réseau Adour. Néanmoins on peut d'ores et déjà supposer que les débits calculés pendant des pics de crue pour lesquels l'Adour a débordé sont faux, en effet le réseau Adour étant modélisé sans lit majeur et avec des berges infiniment hautes la relation hauteurs d'eau/débits calculée par le modèle hydraulique Mascaret ne correspondra pas à la réalité. Nous validerons donc les résultats seulement sur les observations de hauteurs d'eau.

Enfin pour chacun de ces épisodes le cycle d'assimilation est suivie d'une phase de prévision de 12h. C'est en effet le temps nécessaire à un forçage amont pour être propagé

jusqu'à la station aval de Convergent. D'autre part le temps de transfert des forçages amonts jusqu'aux stations de Peyrehorade, Urt et Pont-Blanc est de respectivement au plus 6 heures, 10 heures et 6 heures. C'est pourquoi nous ne nous intéresserons pas aux scores de prévision au-delà de ces échéances car ceux-ci ne représenteraient pas la capacité de l'IEnKF à améliorer les résultats par rapport au modèle seul.

Nous synthétisons les résultats de ces expériences par station et par échéance en reportant dans le tableau 8.6 les pourcentages d'amélioration des scores de l'IEnKF en comparaison du modèle seul.

Station	Echéance (h)	RMSE	Corrélation	Biais	Nash
Peyrehorade	0	87.91	2.66	94.31	25.44
	1	74.18	2.54	77.03	24.23
	3	40.59	1.86	41.8	16.47
	6	12.08	0.63	13.15	5.49
Urt	0	-3.4	1.11	-422.23	6.14
	2	5.9	0.97	-225.14	7.58
	5	5.61	0.58	-94.87	4.6
	10	-0.99	0.032	-38.26	0.12
Lesseps	0	30.33	-0.0064	62.17	1.53
	2	11.01	-0.0378	19.53	0.29
	6	0.77	-0.0261	3.64	-0.0265
	12	-1.36	-0.0372	2.19	-0.093
Pont-Blanc	0	70.95	1.64	77.61	17.65
	1	45.44	1.41	46.45	13.14
	3	10.2	0.23	11.19	3.01
	6	3.8	0.007	4.86	0.89

TAB. 8.6 – Pourcentage d'amélioration des scores de l'IEnKF par rapport au modèle seul. Lorsque les scores sont négatifs cela signifie que l'assimilation a dégradé les résultats par rapport au modèle seul.

Les résultats sur les états analysés (à échéance de 0h) en comparaison du modèle seul sont globalement bons et sont en cohérence avec ceux présentés dans les tableaux 8.1 et 8.4 et notamment aux stations de Peyrehorade et de Pont-Blanc où l'on a procédé à l'assimilation. Le modèle seul sans assimilation étant relativement bon à Lesseps, l'apport de l'assimilation à cette station reste faible mais tout de même non négligeable puisque l'on note par exemple une amélioration de la RMSE de l'IEnKF par rapport au modèle seul de l'ordre de 30 %. Dans tous les cas la corrélation n'est que faiblement impactée par l'IEnKF, c'est normal puisque la corrélation des résultats entre le modèle seul et les observations est déjà bonne. Cependant à la station de Urt le constat est nettement plus contrasté et on peut noter une importante dégradation du biais par rapport au modèle seul. Cela s'explique très bien par le fait que les covariances spatiales diagnostiquées par le modèle à la station de Peyrehorade où l'on a procédé à l'assimilation ne correspondent pas à la réalité. En effet le modèle ne disposant pas de lit majeur et les berges modélisées

étant infiniment grandes les covariances calculées lors des pics de crues sont fausses lors des pics de crue importants ayant donné lieu dans la réalité à des débordements que nous ne pouvons pas modéliser. Cela est illustré sur la figure 8.29 tirée de l'épisode de 2013 pour lequel le biais est le plus important et les hauteurs d'eau analysées à Urt sont sur-estimées tout au long de l'épisode. Les travaux de modélisation en cours sur le réseau Adour maritime avec ajout d'un lit majeur sur l'ensemble du réseau devraient à l'avenir permettre de corriger ce problème mais au moment où nous rédigeons ce travail nous n'avons pas pu tester ce modèle amélioré pour valider cette hypothèse.

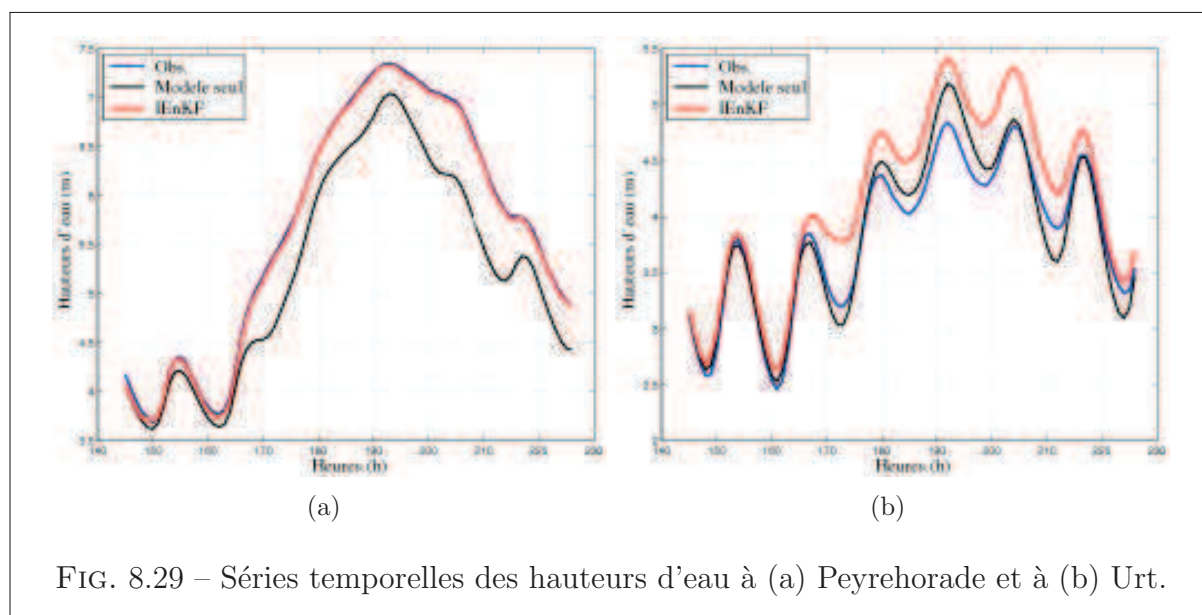


FIG. 8.29 – Séries temporelles des hauteurs d'eau à (a) Peyrehorade et à (b) Urt.

Néanmoins comme on peut le voir sur la figure 8.29 en mode prévision l'IEnKF a des résultats plus contrastés. Pour les échéances de prévision très courtes les résultats sont relativement corrects par rapport au modèle seul. Mais en général au-delà de 3 heures de prévision on perd le bénéfice de l'assimilation et les scores convergent vers ceux du modèle seul sans assimilation car la correction apportée par l'assimilation est propagée à l'aval par le modèle au fil de la prévision. Ce fait est très probablement aggravé par la présence des seuils sur les biefs 3, 6 et 7 qui limite la correction à l'amont des stations.

D'autre part il faut souligner que là encore la situation est légèrement différente à la station de Urt. En effet sur ces épisodes le modèle sans assimilation est relativement bon à Urt, les hauteurs d'eau prévues avec l'IEnKF vont au fil de la prévision converger vers celles calculées avec le modèle sans assimilation et ainsi les résultats vont s'améliorer au cours de la prévision !

Ainsi l'IEnKF a permis de montrer certaines des limitations du modèle Adour maritime et notamment le besoin d'ajout d'un lit majeur sur le réseau ainsi que l'extension du modèle à l'amont de ses limites actuelles pour augmenter le temps de prévision. Comme on l'a évoqué plus haut des travaux dans ce sens sont en cours au SPC GAD en lien avec le SCHAPI. Un des autres intérêts de l'IEnKF est de pouvoir faire des réanalyses

assez bonnes (dans la limite des seuils et du modèle qui pour le moment est mauvais sur certaines zones) sur le réseau. Cela a un intérêt pour des études statistiques, de compréhension de la dynamique de l'écoulement...

D'autre part il convient d'évoquer le temps de calcul de l'IEEnKF sur l'Adour. Celui-ci est dû principalement à la multiplicité des membres qu'il faut intégrer dans le temps pour les prévisions. Dans les expériences réalisées dans cette section et dans les sections 8.3.2 et 8.4.3 nous avons utilisé un ensemble de 80 membres. La machine que nous avons utilisé pour effectuer ces calculs dispose de 8 processeurs sur lesquels les membres sont lancés (gestion assurée par la fonctionnalité PARASOL d'OpenPALM) en parallèle par groupes de 8 membres. Ainsi un cycle assimilation et prévision à échéance de 12 heures plus gestion des fichiers de sortie, de reprise... dure environ 30 secondes. Ce temps de calcul est donc relativement faible. D'après le SCHAPI pour un basculement il faut compter une multiplication du temps de calcul par un facteur 2 ou 3 (du fait par exemple du temps d'acquisition des données observées *in situ* et de conversion au format adéquat) soit au pire un temps de calcul pour un cycle assimilation et prévision à 12 heures d'au plus 1 minute 30 secondes. Ce temps de calcul reste donc compatible avec les contraintes opérationnelles. Néanmoins la structure de la chaîne opérationnelle, la POM¹ n'est pas faite pour gérer des méthodes d'ensemble. Même si il est prévu que celles-ci soient intégrées à terme dans les outils de prévision du SCHAPI et des SPC.

Enfin, les prévisions faites dans cette section et dont les résultats sont présentés dans le tableau 8.6 sont des prévisions déterministes au sens où on prévoit des hauteurs d'eau mais on ne prévoit pas la probabilité de la réalisation d'un événement comme par exemple quelle est la probabilité que les hauteurs d'eau dépassent tel seuil à telle échéance. Nous traitons de la question de l'applicabilité sur l'Adour maritime des prévisions d'ensemble, ou prévisions probabilistes, dans la section suivante.

8.5 Vers des prévisions d'ensemble sur le réseau Adour maritime ?

Dans cette section nous nous interrogeons sur la possibilité d'établir des prévisions d'ensemble sur le réseau Adour maritime. Par l'information qu'elles contiennent les prévisions d'ensemble présentent un intérêt indéniable en comparaison des prévisions déterministes. Encore faut-il que l'ensemble, ou système de prévision, présente certaines « bonnes » propriétés. Les propriétés en question dites de consistance, de fiabilité... ont été présentées au chapitre 4. Dans cette section nous montrons sur une expérience d'assimilation de données synthétiques que la possibilité de définir un système de prévision ensembliste pertinent au regard des propriétés évoquées au chapitre 4 est fortement dépendante de la connaissance que nous avons des forçages amonts.

8.5.1 Cadre de l'expérience et démarche

L'expérience à laquelle nous nous livrons consiste à assimiler des observations de hauteurs d'eau synthétiques. Ces observations sont générées en propageant le long du

¹POM pour Pate-forme Opérationnelle de Modélisation.

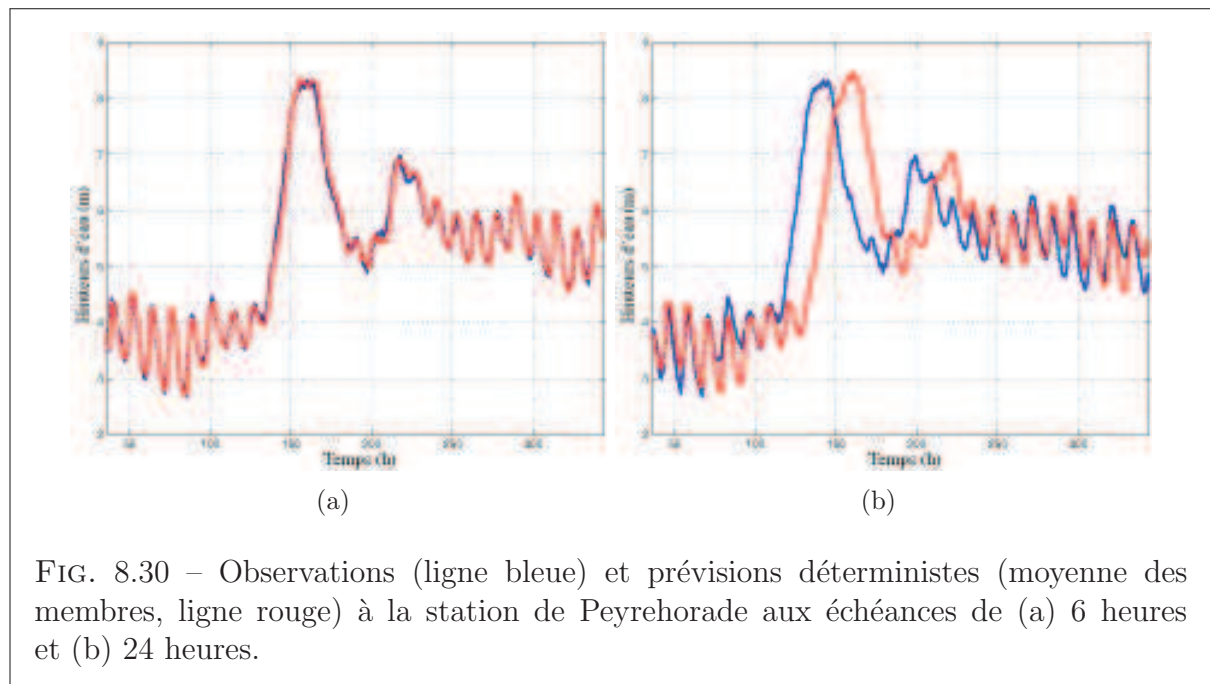
domaine les forçages de l'épisode de 2014 et en ajoutant aux hauteurs d'eau calculées un bruit de moyenne nulle et d'écart-type égal à 0.051 m . Cette valeur correspond à l'écart-type d'erreur d'observation diagnostiqué à la section 8.4.1 sur l'Adour.

Durant l'expérience d'assimilation nous utilisons les mêmes forçages que ceux que nous avons utilisés pour la génération des observations synthétiques hormis bien sûr pendant les phases de prévision où pour nous rapprocher du contexte opérationnel les forçages amonts sont fixés constants et égaux aux derniers forçages connus. Ainsi pour chaque station les prévisions à une échéance inférieure au temps de transfert des forçages amonts jusqu'à cette station sont réalisées avec une connaissance parfaite des forçages et au-delà du temps de transfert (des forçages jusqu'aux stations) les prévisions sont entachées d'erreur car les forçages eux-mêmes sont entachés d'erreur, erreur d'autant plus importante que l'échéance de prévision est grande. On montre alors que la pertinence du système de prévision ensembliste dépend fortement de la connaissance des forçages.

8.5.2 Estimation des prévisions d'ensemble

Prévisions déterministes

Dans un premier temps nous nous intéressons aux prévisions déterministes à la station de Peyrehorade sur un épisode à deux échéances (voir figure 8.30) : une première échéance égale au temps de transfert de 6 heures et une seconde échéance de 24 heures qui est bien supérieure au temps de transfert et qui correspond à l'échéance à partir de laquelle les niveaux de vigilance sont définis par les services de prévision des crues. Même si cela peut sembler une trivialité l'idée est de montrer ici l'importance de l'erreur engendrée sur les prévisions par des forçages inexacts.



Consistance

L'établissement des diagrammes de rang pour chacune de ces échéances de prévision montre un diagramme de rang relativement proche du diagramme plat pour une échéance de prévision égale au temps de transfert (figure 8.31-(a)). Cela montre que les membres de l'ensemble et les observations suivent des variables aléatoire « proches » et donc nous pensons que cela a du sens d'établir des prévisions auxquelles on associe des marges d'incertitude ou des diagrammes de type « boîte à moustaches », voir figures 4.4 et 4.3 respectivement. Néanmoins la différence entre le diagramme diagnostiqué et le diagramme plat est plus importante que la marge d'incertitude qui lui est liée, cela devrait à terme nous amener à revoir la manière avec laquelle l'ensemble est généré de manière à disposer d'un ensemble reflétant au mieux les incertitudes. Pour l'échéance de 24 heures (figure 8.31-(b)), le diagramme de rang montre que l'ensemble est clairement sous-dispersif. Naturellement cela est dû à l'erreur importante sur les hauteurs d'eau prévues à cette échéance (voir figure 8.30-(b)). Pour les autres échéances comprises entre le temps de transfert et 24 heures les diagrammes de rang montrent un ensemble dont la sous-dispersivité s'accroît avec l'échéance de prévision.

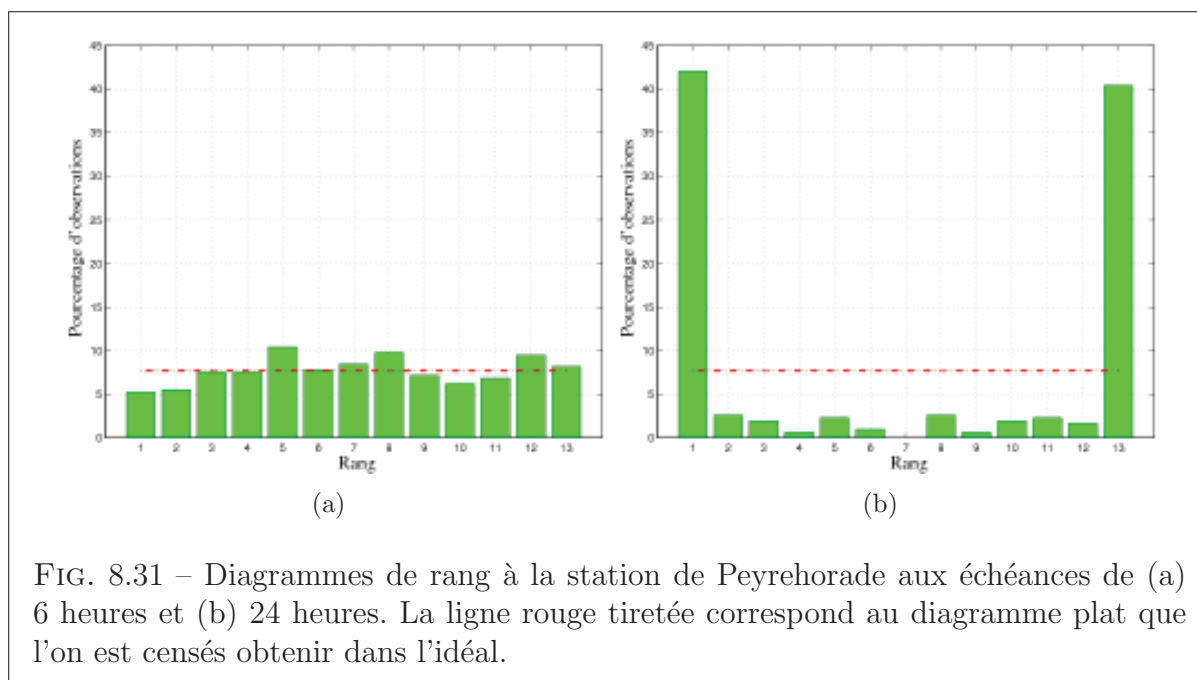


FIG. 8.31 – Diagrammes de rang à la station de Peyrehorade aux échéances de (a) 6 heures et (b) 24 heures. La ligne rouge tiretée correspond au diagramme plat que l'on est censés obtenir dans l'idéal.

Comme on l'a évoqué plus haut pour une échéance de prévision inférieure ou égale au temps de transfert les membres et les observations suivent des variables aléatoires relativement proches, cela a alors du sens d'associer à la prévision une barre d'incertitude ou un diagramme de type « boîte à moustaches » comme cela est illustré, sur une portion de l'épisode considéré, sur les figures 8.32-(a) et 8.32-(b) respectivement.

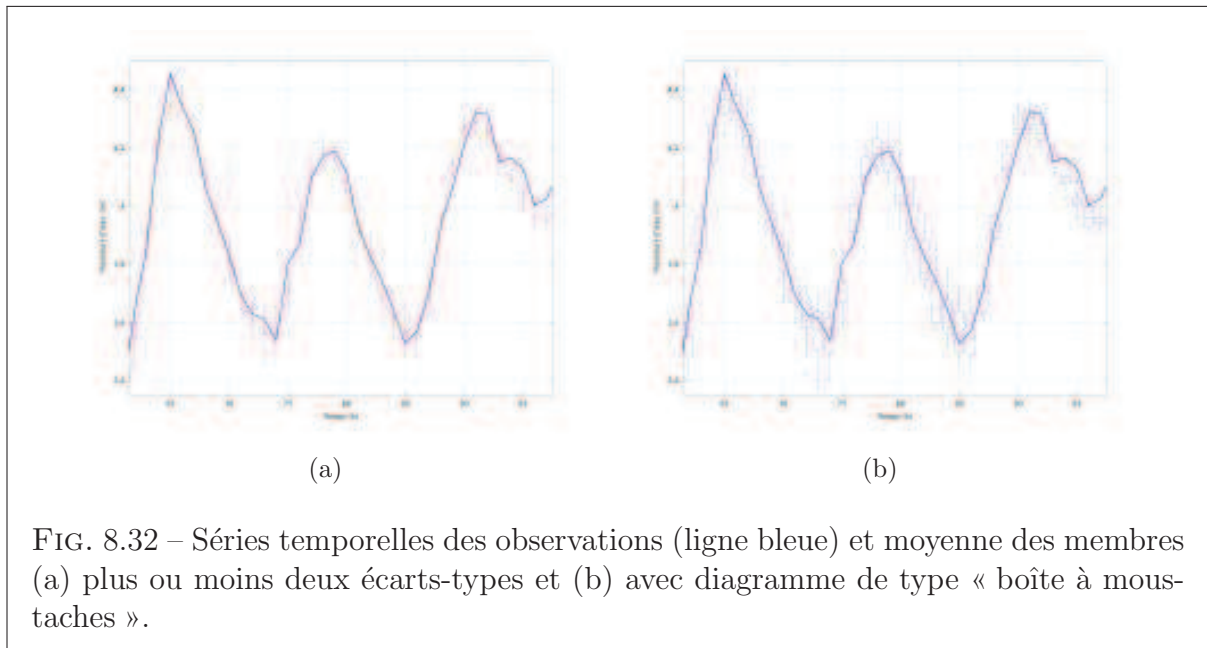
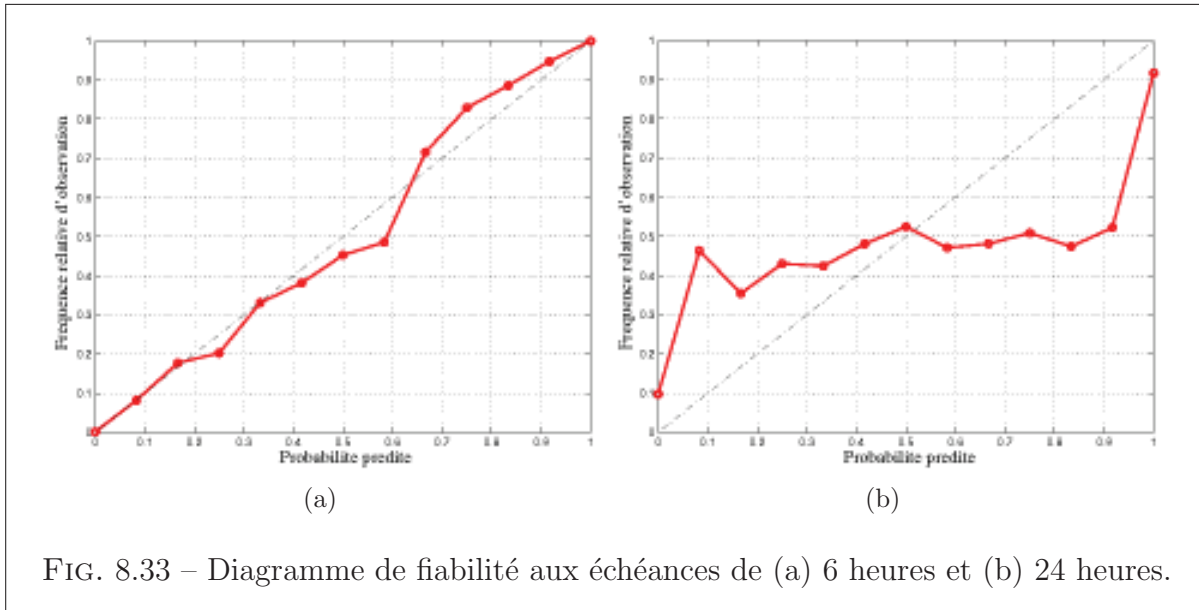


FIG. 8.32 – Séries temporelles des observations (ligne bleue) et moyenne des membres (a) plus ou moins deux écarts-types et (b) avec diagramme de type « boîte à moustaches ».

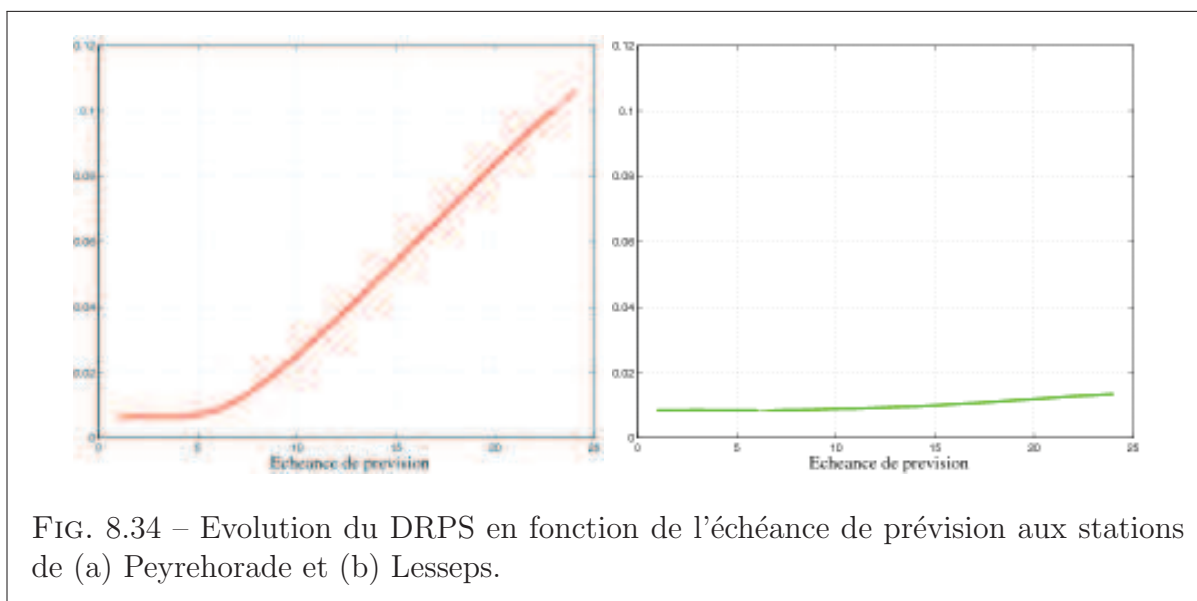
Fiabilité

Au chapitre 4 nous avons présenté la notion de fiabilité qui permettait de juger de la correspondance entre d'une part les probabilités prédites *a priori* et d'autre part les fréquences observées *a posteriori* pour un évènement probabiliste du type $\{Z \geq T\}$. Le fait que nous travaillons avec des échantillons relativement restreints ne nous a pas permis d'établir si notre ensemble vérifiait la propriété de fiabilité. Néanmoins nous proposons ici d'étendre cette notion à un ensemble d'évènements $\{Z \geq T_k\}$, les T_k couvrant l'ensemble des valeurs prises par les observations sur la durée de l'épisode. L'idée étant toujours de juger de la correspondance entre les probabilités prédites *a priori* et les fréquences observées *a posteriori*. Dans le cas de la figure 8.33-(a) avec une prévision à échéance inférieure au temps de transfert le diagramme de fiabilité est très bon puisque celui-ci est très proche de la diagonale. Ainsi globalement il y a une bonne correspondance entre les probabilités prédites *a priori* et les fréquences observées *a posteriori*. Par contre dans le cas d'une prévision à échéance supérieure au temps de transfert le diagramme est nettement moins bon que précédemment, en effet cela tient à l'erreur commise sur les hauteurs d'eau prévues.



Précision

Pour juger de la précision des prévisions ensemblistes nous avons choisi d'utiliser le Discrete Ranked Probability Score (DRPS) plutôt que le Brier Score car le DRPS évalue la distribution complète et donne donc une information synthétique. Pour chaque échéance de prévision nous avons calculé le DRPS et avons synthétisé les résultats dans un graphique présenté en figure 8.34. On constate dans le cas de Peyrehorade que pour les échéances de prévision inférieures au temps de transfert (on rappelle que le temps de transfert à cette station est de 6 heures) le DRPS est relativement faible puis au-delà de celui-ci le DRPS croît de manière affine du fait des forçages amonts constants. Cela montre que le système de prévision est très précis pour des échéances de prévision inférieures au temps de transfert et que la précision se dégrade au-delà. Bien que pouvant paraître trivial ce fait n'en est pas moins intéressant car il montre que malgré une parfaite connaissance des forçages au moment de la prévision le fait d'utiliser des forçages constants à l'amont n'est pas compatible avec l'établissement d'un système de prévision ensembliste. Cependant à la station de Lesseps qui est proche du forçage aval qui n'est pas maintenu constant pendant la prévision on constate que l'erreur bien que croissante après le temps de transfert (12 heures à cette station) reste très faible.



Une propriété corollaire à la précision est l'habileté que l'on avait évoquée au chapitre 4. Celle-ci établit dans quelle mesure une prévision est meilleure au sens de la précision qu'une prévision de référence. Dans le cadre de cette expérience d'assimilation de données synthétiques où les forçages sont parfaitement connus nous pensons que cela n'a pas de sens de juger de l'habileté puisque dans ce cas les résultats de tout système de prévision sont équivalents.

Table de contingence, discrimination et courbe de ROC

Enfin nous nous intéressons aux scores liés à la table de contingence (voir tables 4.2 et 4.3). Parmi les scores relatifs à la table de contingence nous avons décidé d'en retenir six qui nous ont paru refléter des propriétés essentielles d'un système de prévision : le taux de réussite (Tdr), la précision (Pre), le taux de corrects négatifs (Tcn), le taux de faux positifs (Tfp), la probabilité d'occurrence (Pdo) et la menace (Men). On rappelle que l'on considère qu'un évènement a été prévu par l'ensemble si la probabilité prédite est supérieure à la fréquence d'observation de l'évènement. Nous avons synthétisé les résultats à deux échéances de prévision dans le tableau 8.7 pour différents niveaux d'eau : 4.71 m, 5.11 m, 6.71 m et 7.51 m qui correspondent respectivement à la cote de débordement, au seuil d'alerte jaune, au seuil d'alerte orange et au seuil d'alerte rouge à la station de Peyrehorade.

On note que pour toutes les cotes pour lesquelles nous avons calculé les scores à l'échéance de 6 heures ceux-ci sont relativement bons ce qui est normal puisque 6 heures est le temps de transfert pour la station de Peyrehorade. Plus étonnant est le fait que ces scores restent relativement bons à une échéance de 24 heures pour les cotes de 4.71 m et 5.11 m alors que les forçages sont faux à cette échéance de prévision. Cela s'explique par le fait que avant et après le pic de crue les hauteurs d'eau observées oscillent autour de valeurs pour lesquelles les forçages bien que faux restent dans une certaine mesure pertinents. Pour ce qui est des prévisions à 24 heures pour les cotes de 6.71 m et 7.51

Cote	Echéance	Tdr	Pre	Tcn	Tfp	Pdo	Men
4.71	6	99.5	99.35	99.04	0.96	99.51	99.02
	24	90.37	91.2	93.26	6.74	97.04	87.95
5.11	6	96.13	97.72	100	0	100	96.13
	24	82.9	82.41	81.58	18.42	88.4	74.77
6.71	6	96.88	95.77	95.64	4.36	72.09	70.46
	24	34.37	82.09	87.64	12.36	24.44	16.67
7.51	6	100	99.02	98.96	1.04	86.36	86.36
	24	21.05	87.62	92.01	7.99	14.82	9.52

TAB. 8.7 – Tableau du taux de réussite (Tdr), de la précision (Pre), du taux de corrects négatifs (Tcn), du taux de faux positifs (Tfp), de la probabilité d'occurrence (Pdo) et de la menace (Men) aux échéances de 6h et 24h pour les cotes 4.71 m, 5.11 m, 6.71 m et 7.51 m.

m les scores sont nettement dégradés à l'exception de la précision et du taux de corrects négatifs. Cela est normal car ces deux scores sont basés sur le calcul des corrects négatifs que l'ensemble prévoit assez bien malgré l'échéance de prévision et c'est la même raison pour laquelle le taux de faux positifs reste relativement faible.

Pour ce qui est du taux de réussite à échéance de 24 heures pour les cotes de 6.71 m et 7.51 m celui-ci est relativement faible, en effet un taux de réussite de 33 % équivaut à deux échecs pour un succès or d'après [World Meteorological Organization, 2011] ce taux signifie que le système de prévision est totalement inutile pour la prévision des crues. En ce qui concerne la probabilité d'occurrence là aussi les résultats sont faibles, par exemple pour la cote 6.71 m la probabilité d'occurrence est 24.44 % ce qui signifie que pour 10 événements prédits par le système de prévision seulement 2 ou 3 auront réellement lieu. Cela est gênant car la crédibilité des services de prévision auprès des autorités ainsi que du grand public se joue en partie sur la concordance entre événements prévus et événements effectivement réalisés. [World Meteorological Organization, 2011] considère qu'une probabilité d'occurrence inférieure à 10 % est synonyme d'un système de prévision inefficace. Enfin on considère le score de menace, celui-ci en écartant les corrects négatifs mesure la capacité générale du système de prévision à prédire des événements extrêmes. Il apparaît donc d'après ce score que le système de prévision a beaucoup de mal à prévoir des événements pour lesquels les hauteurs d'eau dépassent les cotes de 6.71 m et 7.51 m. Pour cette dernière cote cela apparaissait déjà clairement sur la figure 8.30-(b).

Enfin nous nous intéressons à la propriété de discrimination qui permet de juger de la capacité du système de prévision à juger de la réalisation ou non d'un événement. Comme on l'avait évoqué au chapitre 4 la discrimination peut être estimée avec la courbe de ROC et le score associé, le ROCSS. Bien que disposant d'un échantillon de données relativement restreint nous avons tenu à faire figurer ici à titre illustratif les courbes de ROC établies pour l'échéance de prévision de 24 heures pour les cotes 5.11 m et 6.71 m, voir figures 8.35-(a) et 8.35-(b) respectivement. Il faut préciser que ces courbes ont été interpolées à partir des quelques points dont nous disposons. Le ROCSS pour la cote 5.11

m vaut 0.75 et pour la cote 6.71 m il vaut 0.42. La figure 8.35 montre que le système de prévision est plus discriminant pour la cote 5.11 m que pour la cote 6.71 m .

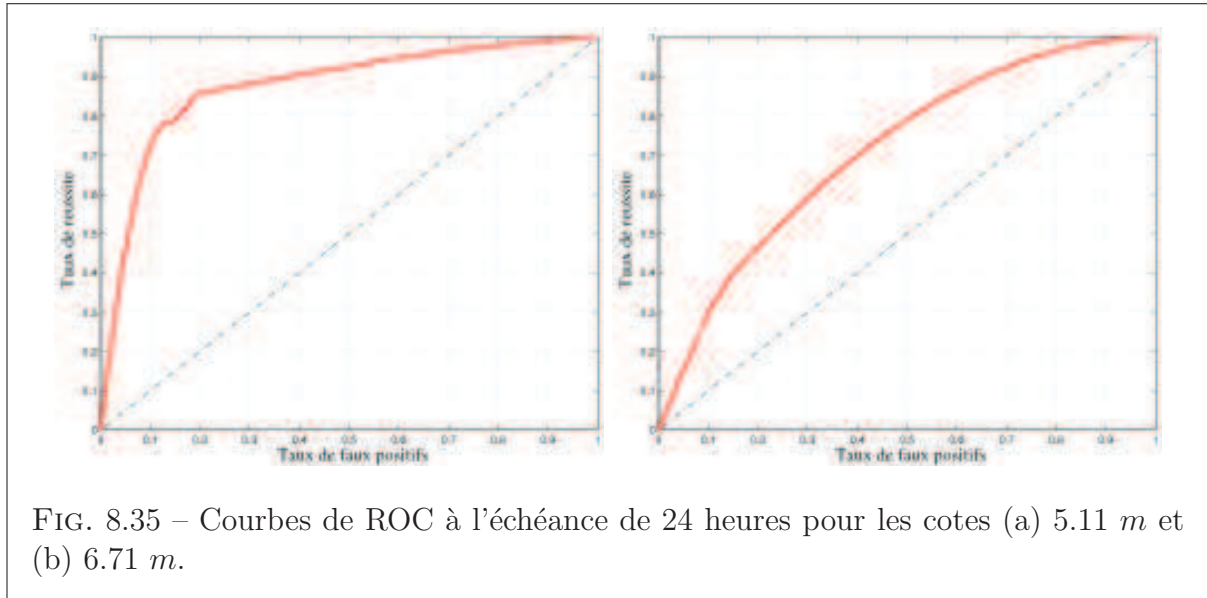


FIG. 8.35 – Courbes de ROC à l'échéance de 24 heures pour les cotes (a) 5.11 m et (b) 6.71 m .

D'autre part on note que pour les échéances de prévision de 6 heures le taux de faux positifs est très faible par rapport au taux de réussite. Dans ce cas nous n'avons pris la peine de tracer la courbe de ROC car l'ensemble est clairement discriminant et sa courbe de ROC est à rapprocher de celle donnée au chapitre 4 en figure 4.10-(a).

8.5.3 Discussion

Cette section est une section exploratoire sur la possibilité d'appliquer des prévisions d'ensemble sur le réseau Adour maritime, nous sommes pleinement conscients que dans l'idéal il aurait fallu éprouver les critères testés ici sur des volumes de données plus importants pour réaliser une étude statistique plus propre. Néanmoins il ressort de cette section quelques points intéressants qu'il convient de mettre en avant.

On s'intéresse tout d'abord au DRPS (voir figure 8.34). Celui-ci montre que du point de vue de la précision les résultats aux stations sont bons pour les échéances de prévision inférieures au temps de transfert et montre qu'au-delà les résultats sont d'autant plus mauvais que la station est sensible aux forçages fixés constants. C'est le cas par exemple de Peyrehorade (voir figure 8.34-(a)) où le temps de transfert est de 6 heures. Par contre à la station de Lesseps (voir figure 8.34-(b)) où le modèle est piloté par le forçage aval qui n'est pas fixé constant en mode prévision la dégradation du CRPS au fil du temps est moins importante.

Le constat fait ici relativement au DRPS avant et après échéance de prévision égale au temps de transfert peut être fait également avec les autres scores ou diagrammes. C'est le cas du diagramme de rang qui est relativement plat pour les échéances inférieures au

temps de transfert (voir figure 8.31) mais qui reste néanmoins au-delà de la marge d'erreur liée à cette expérience. Celui-ci ne reflète donc pas parfaitement les incertitudes puisque de fait observations et membres de l'ensemble ne suivent pas la même variable aléatoire. C'est pourquoi il est nécessaire pour établir un système de prévision ensembliste sur l'Adour de déterminer une manière de générer les membres de l'ensemble de manière à ce que celui-ci reflète mieux les incertitudes. Cependant il faut souligner que les observations pour des prévisions établies à une échéance inférieure ou égale au temps de transfert sont globalement comprises dans la marge d'incertitude établie par l'ensemble, voir figure 8.32. Au-delà du temps de transfert le diagramme de rang montre un ensemble sous-dispersif (figure 8.31-(b)) dont la distance au diagramme plat s'accroît avec l'échéance de prévision. De la même manière la fiabilité de l'ensemble diminue au fil de la prévision (figure 8.33) pour devenir totalement non fiable.

D'autre part l'étude des scores associés à la table de contingence montrent que ceux-ci sont bons pour des échéances inférieures au temps de transfert et se dégradent au-delà. En particulier pour des échéances de 24 heures et pour les cotes de 6.71 m et 7.51 m les taux de réussite et de faux positifs, la probabilité d'occurrence, le score de menace ainsi que l'ébauche de courbe de ROC pour la cote 6.71 m montrent que à longue échéance et pour des cotes importantes le système de prévision est complètement inutile.

Même si ce constat peut paraître trivial cette expérience montre que l'établissement d'un système de prévision ensembliste sur l'Adour est fortement dépendant de la connaissance des forçages avant et pendant la prévision et que des forçages constants en mode prévision ne permettent pas d'établir des prévisions ensemblistes valides du point de vue des propriétés éprouvées ici. D'autre part des expériences numériques sur des données réelles, qui n'ont pas été reportées ici, ont montré que malgré l'assimilation de données les prévisions d'ensemble dans ce cas donnaient des résultats très discutables, notamment du point de vue de la consistance et de la fiabilité.

Pour améliorer la qualité des prévisions d'ensemble au-delà du temps de transfert (quand on a une bonne connaissance des forçages) il faudrait disposer de débits prévus aux stations amont, par exemple avec des extrapolations des débits connus ou encore avec un modèle hydrologique. C'est ce qui est fait par les prévisionnistes du SPC Seine Amont-Marne Amont qui combinent modèles hydraulique et hydrologique et qui associent à leurs prévisions déterministes une marge d'incertitude. L'échéance de prévision pourrait encore être repoussée en combinant modèle hydrologique et prévisions de pluie fournies par des modèles météorologiques. Mais des travaux dans ce sens ne sont pas prévus à l'heure actuelle sur l'Adour. Cette expérience a montré également que l'ensemble ne reflète pas exactement les incertitudes et cela même pour des échéances de prévision où nous avons une connaissance parfaite des forçages. Il conviendrait donc à l'avenir de déterminer une meilleure méthode pour générer les membres. D'autre part ceux-ci ont été générés en perturbant les forçages dont on a fait l'hypothèse dans ce travail qu'ils constituaient la principale source d'erreur sur l'Adour maritime. Mais la bathymétrie ainsi que les coefficients de frottement sont également une source importante d'incertitudes. Il conviendrait donc également d'explorer la génération des membres de l'ensemble en perturbant les paramètres géométriques du modèle.

Chapitre 9

Couplage de modèles hydrauliques 1D-2D sur la ville de Bayonne

Sommaire

9.1	Le modèle hydraulique 2D sur Bayonne	214
9.2	Couplage à interfaces	218
9.2.1	Modèles, interfaces et raccordement	218
9.2.2	Algorithme de couplage et implémentation	219
9.2.3	Expérience numérique de couplage à interfaces sur l'épisode de 2014	222
9.3	Couplage de modèles superposés	226
9.3.1	Modèles utilisés, disposition et termes échangés entre les modèles	227
9.3.2	Algorithme de couplage et implémentation	230
9.3.3	Expérience numérique de couplage de modèles superposés sur l'épisode de 2014	231
9.4	Combinaison de l'assimilation de données ensembliste et du couplage de modèles superposés	240
9.4.1	Implémentation	241
9.4.2	Résultats numériques	241

Dans ce chapitre nous nous intéressons au couplage de modèles hydrauliques sur le réseau Adour maritime et plus particulièrement sur la zone de Bayonne. Sur la zone de Bayonne un modèle 2D a été développé pour améliorer la modélisation hydraulique localement et par suite la prévision des crues sur cette zone. Sur la même zone a déjà été développé le modèle 1D Adour maritime que l'on a présenté en section 8.1.2. Le modèle 2D ne disposant pas de stations d'observations à ses limites (hormis à l'aval avec la station de Bayonne-Lesseps) il est impératif de forcer celui-ci avec les sorties du modèle 1D. En effet afin d'assurer une certaine forme de continuité entre les modèles il est nécessaire de les coupler et donc de mettre en œuvre les méthodes présentées au chapitre 5.

Dans un premier temps nous donnerons donc quelques rappels sur le modèle hydraulique 2D développé sur la zone de Bayonne avant de nous intéresser au couplage à interfaces puis au couplage de modèles superposés sur cette zone.

9.1 Le modèle hydraulique 2D sur Bayonne

Le modèle 2D sur Bayonne a été développé lors du stage de Manuel Martin (Martin, 2013) au SCHAPI du mois d’Avril au mois de Septembre 2013 et encadré conjointement par le SPC GAD et le SCHAPI. Ce modèle se situe à la confluence de l’Adour et de la Nive et est donc situé par rapport au modèle Adour maritime 1D à la confluence des biefs 1, 2 et 3. Une image du maillage du modèle 2D a été reportée en figure 9.1, elle permet de voir l’emprise du modèle par rapport à la ville de Bayonne et ses alentours.

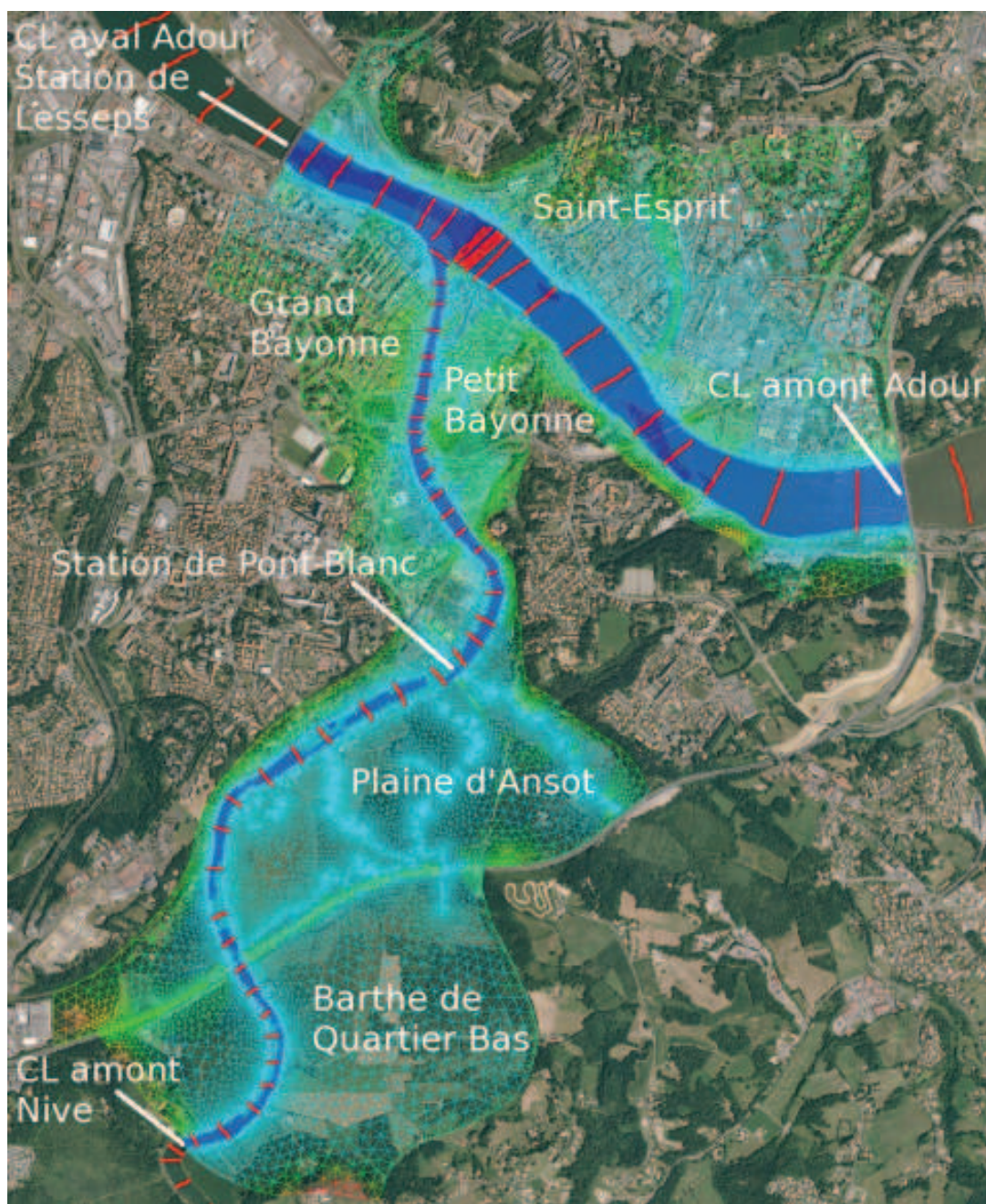


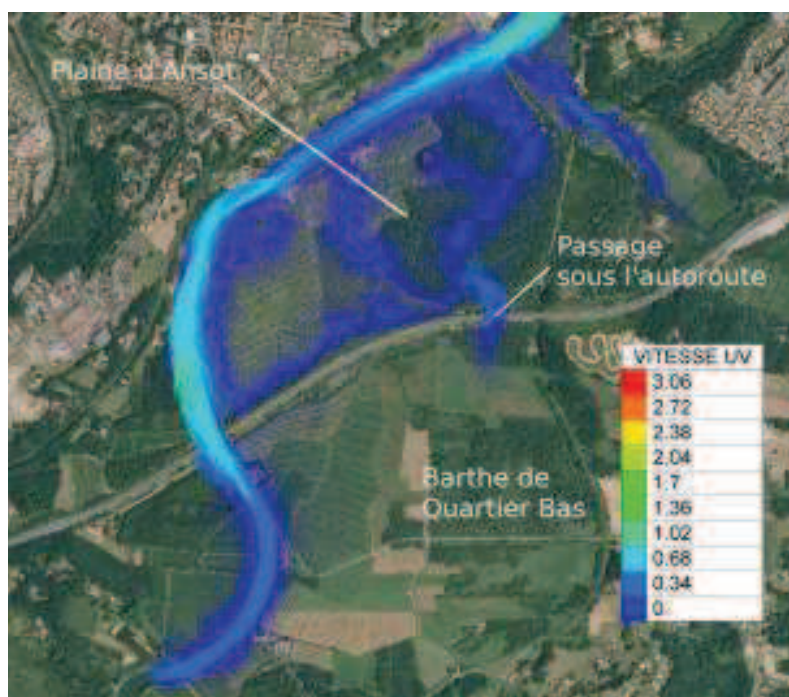
FIG. 9.1 – Superposition d’une image Google Maps de la ville de Bayonne, du maillage 2D et de la couche SIG ayant servi pour établir les sections en travers du modèle Adour maritime (traits rouges).

La zone recouverte par le maillage 2D se décompose en deux zones bien distinctes. On distingue d’une part une zone très urbanisée dans le centre de Bayonne avec les quartiers du Petit et du Grand Bayonne ainsi que le quartier de Saint-Esprit et qui comprend la confluence Nive/Adour (voir figure 1.12). Du fait de sa forte urbanisation cette zone

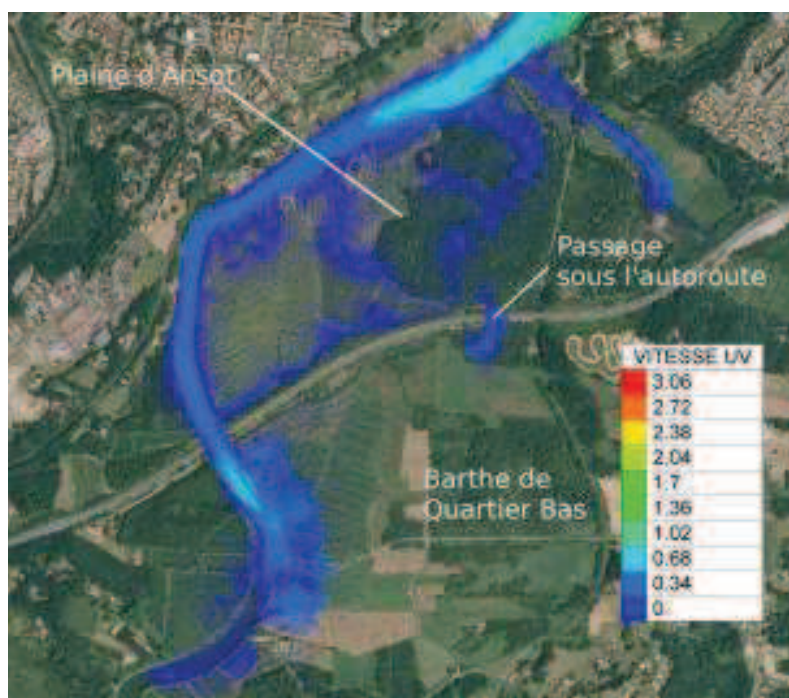
présente d'importants enjeux en termes de sécurité des biens et des personnes, il était donc important pour le SCHAPI de disposer d'un modèle 2D décrivant finement les vitesses ainsi que les hauteurs d'eau sur cette zone. On distingue d'autre part la zone située au Sud de Bayonne et constituée de la plaine d'Ansot et de la barthe de Quartier Bas qui sont deux plaines d'inondations séparées par l'autoroute A63 (qui contourne Bayonne par le Sud-Est et que l'on devine sur la figure 9.1 par la courbe verte séparant ces deux zones). Ces deux plaines d'inondation jouent un rôle majeur dans les épisodes de fortes crues de la Nive ou lors d'importantes marées car non seulement elles échangent de l'eau par un passage sous l'A63 (voir figure 9.2) mais elles en stockent et en restituent lorsque la Nive est en crue. L'écoulement dans cette zone, qui en termes de surface représente presque la moitié du modèle 2D, n'est donc pas du tout mono-dimensionnel lors des épisodes de crue et justifie pleinement la mise en place d'un modèle 2D sur cette zone.

On a reporté en figure 9.2 deux illustrations (l'une à marée descendante 9.2-(a) et l'autre à marée montante 9.2-(b)) des champs de vitesse modélisés par le modèle 2D sur la plaine d'Ansot et la barthe de Quartier Bas lors d'un épisode de crue de la Nive. Il faut noter en particulier comment l'écoulement se fait à travers la plaine d'Ansot mais aussi de manière préférentielle selon la topographie du terrain. D'autre part il ressort clairement sur chacune de ces images les échanges entre la plaine d'Ansot et la barthe de Quartier Bas sous l'autoroute A63. Il faut noter également comment la Nive déborde dans la barthe de Quartier Bas en amont de l'A63 lorsque la marée monte.

Comme on l'a souligné plus haut, le modèle 2D, hormis à sa condition limite aval, ne dispose pas au niveau de ses frontières mouillées de station d'observations dont les données pourraient forcer le modèle. Et quand bien même ce serait le cas les stations d'observation du réseau Adour ne fournissent pas d'observations de débits et du fait de l'influence maritime sur la zone de Bayonne il ne serait pas possible d'établir des courbes de tarage à ces stations. Le modèle 2D doit donc impérativement être forcé à ses frontières mouillées par les sorties du modèle 1D Adour maritime. Néanmoins si l'on désire assurer une certaine forme de continuité entre les modèles il est nécessaire d'utiliser les méthodes couplage exposées au chapitre 5. Dans la suite nous nous intéresserons donc à l'application de ces méthodes au cas des modèles 1D Adour maritime et 2D sur Bayonne. Dans un premier temps nous étudierons l'application du couplage à interfaces à nos deux modèles. Dans un second temps nous étudierons l'application des méthodes de couplage de modèles superposés à notre cas.



(a)



(b)

FIG. 9.2 – Champs de vitesse (unité $m.s^{-1}$) modélisés avec le modèle 2D sur la plaine d'Ansot et la barthe de Quartier Bas à (a) marée descendante et (b) marée montante.

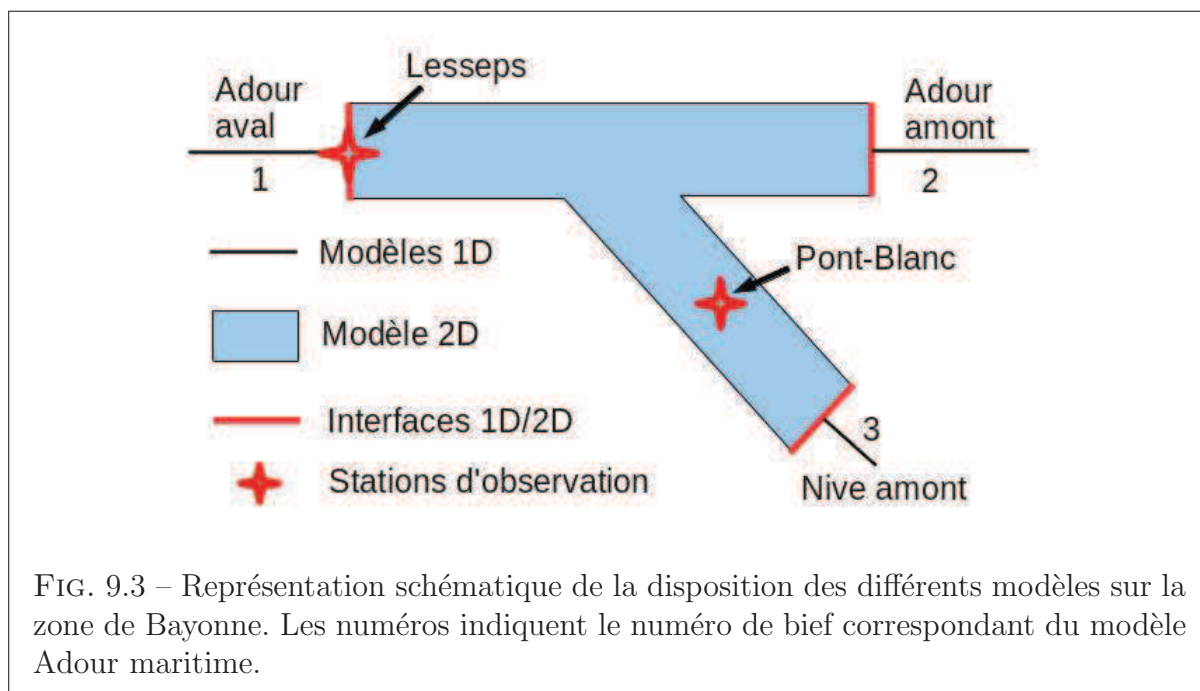
9.2 Couplage à interfaces

Pour développer le couplage à interfaces sur l'Adour maritime nous nous sommes appuyés sur les travaux précédemment cités au chapitre 5, à savoir [Miglio et al., 2005, Malleron et al., 2011].

9.2.1 Modèles, interfaces et raccordement

Modèles utilisés

Comme on l'a vu en section 5.1 la méthode de couplage à interface présentée dans [Miglio et al., 2005, Malleron et al., 2011] suppose que les modèles 1D et 2D sont disposés séquentiellement le long du cours d'eau sur lequel on travaille. Or le modèle 1D Adour maritime présenté en figure 8.2 et le modèle 2D présenté en figure 9.1 sont superposés sur la zone de Bayonne. Pour pouvoir mettre en œuvre la méthode de couplage à interfaces dans notre cas il a donc fallu décomposer le modèle 1D Adour maritime en trois sous-modèles 1D dénommés « Adour aval », « Adour amont » et « Nive amont » et qui en dehors de la zone du modèle 2D recouvrent respectivement le bief 1, les biefs 2, 4, 5, 6 et 7 et le bief 3 du modèle Adour maritime. La disposition relative des modèles est illustrée sur la figure 9.3.



Interfaces entre les modèles

Aux interfaces entre les différents modèles nous avons voulu assurer la continuité des grandeurs physiques échangées entre ceux-ci (voir section 5.1.1). Dans cette optique nous

avons donc modifié les sections en travers des modèles 1D qui sont à l'interface avec le modèle 2D de manière à ce que les sections en travers des différents modèles 1D soient identiques à celles correspondantes du modèle 2D.

Comme on l'a vu au chapitre 5 pour assurer une certaine continuité physique entre les modèles et pour assurer le bon déroulement de l'algorithme de couplage les différentes interfaces entre les modèles doivent être positionnées dans des zones où l'écoulement est mono-dimensionnel. Or dans notre cas les frontières mouillées du modèle 2D (et donc les interfaces entre modèles 1D et 2D) ont été choisies de manière à délimiter la zone de Bayonne pour l'intérêt des phénomènes hydrauliques qui s'y déroulent et donc indépendamment des considérations algorithmiques évoquées ici. Comme on le verra plus loin cela ne va pas aller sans poser des problèmes. En effet si l'Adour sort très peu de son lit au niveau de la ville de Bayonne et plus particulièrement au niveau des frontières mouillées du modèle 2D c'est loin d'être le cas de la Nive. La Nive a en effet tendance à déborder souvent au niveau de la frontière mouillée correspondante du modèle 2D même lors de faibles pics de crue perturbant la convergence de l'algorithme de couplage.

Raccordement des modèles à l'interface

Comme on l'a vu à la section 5.1.1 le raccordement entre les modèles aux interfaces et donc le choix des grandeurs physiques à échanger entre les modèles se fait de manière à ce que les problèmes sur chacun des sous-modèles 1D et 2D soient bien posés. Dans notre cas la situation est assez transparente, chacun des modèles est en régime fluvial il faut donc imposer à chaque limite du domaine la section mouillée ou le débit ou l'invariant de Riemann lié à la caractéristique entrante. D'un côté le système Telemac-2D permet d'imposer aux frontières mouillées du domaine différents types de conditions : des débits, une onde incidente, une courbe de tarage... Parmi toutes ces possibilités nous ne retiendrons que celles susceptibles de nous intéresser dans ce travail à savoir les cotes, les débits et les vitesses. D'un autre côté Mascaret offre également la possibilité d'imposer différents types de conditions aux limites du modèle (hydrogramme, courbe de tarage, seuil...) parmi lesquelles nous retiendrons là aussi celles qui nous intéressent : débits et cotes. Pour chacun des modèles nous imposerons à l'amont des débits, que ceux-ci soient fournis par des données de forçage ou par un modèle, et nous imposerons à l'aval des hauteurs d'eau, que celles-ci soient des données d'observation (par exemple en aval du modèle « Adour aval » avec la station de Convergent) ou calculées avec un modèle.

On notera que dans la section 5.1.1 on avait cité parmi les conditions limites que l'on peut imposer à l'interface d'un modèle 1D la section mouillée. Or comme nous avons fait en sorte que les sections en travers à l'interface des modèles soient identiques alors il est équivalent de transmettre entre les modèles la hauteur d'eau ou la section mouillée.

9.2.2 Algorithme de couplage et implémentation

Algorithme de couplage

On rappelle ici les notations utilisées pour écrire l'algorithme de couplage. On note respectivement $\mathcal{L}_1$ et $\mathcal{L}_2$ les opérateurs discrétisés de Saint-Venant 1D et 2D. Dans un souci de simplicité des notations on désignera dans la suite par 1 le modèle 1D « Adour

aval », par 2 le modèle 1D « Adour amont » et par 3 le modèle 1D « Nive amont » ; cette numérotation des modèles correspond à celle des biefs reportée sur la figure 9.3. D'autre part on note :

- $(h_{1D}, S, Q)_j$ la hauteur d'eau, la section mouillée et le débit dans le j -ème domaine 1D ;
- (h_{2D}, u, v) la hauteur et les composantes de la vitesse sur le domaine 2D.

Dans ce travail on a appliqué la version multiplicative et globale en temps de l'algorithme de Schwarz exposée en section 5.1.2 sur une succession d'intervalles de temps recouvrant les épisodes de crue que l'on a souhaité traiter. On note d'autre part Γ_j , $j = 1, 2, 3$ les interfaces entre le modèle 2D d'une part et les sous-modèles 1D 1, 2 et 3. Si sur l'intervalle de temps $[0; T]$ on connaît $\bar{h}_{2D,j}^0$ sur les interfaces Γ_j ($j = 2, 3$) et si on connaît $(u, v)_j^0$ sur Γ_1 alors l'algorithme de Schwarz multiplicatif et global en temps consiste à résoudre de manière itérative les problèmes suivants :

$$\left\{ \begin{array}{ll} \mathcal{L}_1(S, Q)_j^{k+1} = 0 & \text{sur } \Omega_{1,j} \times [0; T] \\ \text{Conditions limites classiques} & \text{sur } \partial\Omega_{1,j} \times [0; T] \\ h_{1D,j}^{k+1} = \bar{h}_{2D,j}^k & \text{sur } \Gamma_j \times [0; T] \quad j = 2, 3 \\ Q_{i,j}^{k+1} = A_{21}(u, v)_j^k & \text{sur } \Gamma_j \times [0; T] \quad j = 1 \end{array} \right. \quad (9.1)$$

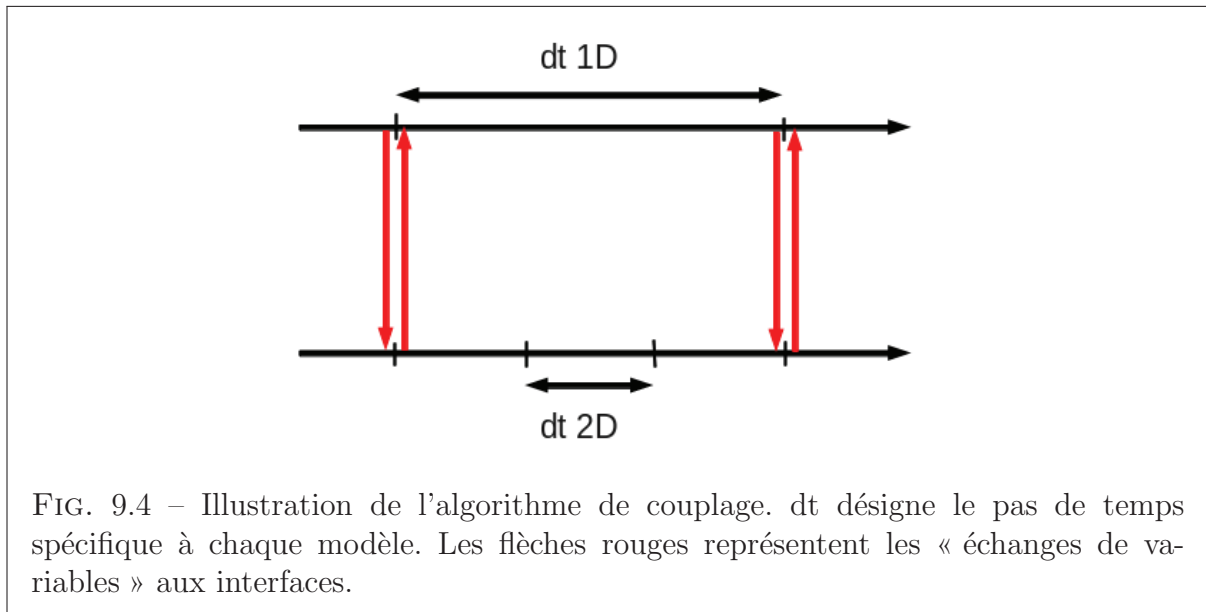
et

$$\left\{ \begin{array}{ll} \mathcal{L}_2(h, u, v)^{k+1} = 0 & \text{sur } \Omega_2 \times [0; T] \\ \text{Conditions limites classiques} & \text{sur } \partial\Omega_2 \times [0; T] \\ (u, v)_j^{k+1} = B_{12}(Q_j^{k+1}) & \text{sur } \Gamma_j \times [0; T] \quad j = 2, 3 \\ h_{2D,j}^{k+1} = h_{1D,j}^{k+1} & \text{sur } \Gamma_j \times [0; T] \quad j = 1 \end{array} \right. \quad (9.2)$$

où Ω_s et $\partial\Omega_s$, $s = 1, 2$ désignent respectivement le domaine et la frontière des modèles 1D et 2D et Γ_j désigne l'interface entre le j -ème modèle 1D et le modèle 2D et les opérateurs A_{21} et B_{12} désignent les opérateurs de transfert des variables 1D vers le 2D et réciproquement.

Dans la pratique on choisira comme valeur pour $\bar{h}_{2D,j}^0$ et $(u, v)_j^0$ la dernière valeur calculée sur les interfaces correspondantes. On notera que dans la section 5.1.2 la surface mouillée 1D apparaissait dans les conditions limites à l'interface entre les modèles. Ici comme les sections en travers sont identiques aux interfaces il est équivalent d'avoir une condition aux limites qui porte sur les hauteurs d'eau plutôt que sur la surface mouillée ce qui est plus facile à implémenter puisque Mascaret accepte les limnigrammes en condition aux limites et pas les données de surface mouillée.

L'algorithme de couplage est illustré en figure 9.4



Critère d’arrêt

Le critère d’arrêt pour l’algorithme de couplage est basé sur la continuité des invariants de Riemann à l’interface entre les modèles. Le critère d’arrêt a été fixé à une différence relative en valeur absolue entre les modèles inférieure à 1%. Dans le cas où l’algorithme de couplage ne permet pas de réduire la différence entre les invariants de Riemann le nombre maximum d’itérations de l’algorithme a été fixé à 3. Dans les faits, à la condition d’avoir des lignes d’eau initiales pour les modèles 1D et 2D présentant une cohérence physique (voir ci-dessus), la convergence se fait en au plus 3 itérations. On observe en pratique qu’au-delà de 3 itérations le couplage ne permet pas de réduire la différence entre les modèles en termes d’invariants de Riemann.

Implémentation de l’algorithme de couplage

Tout comme pour l’implémentation des méthodes d’ensemble sur l’Adour présentée en section 8.1.3 nous avons implémenté l’algorithme de couplage en utilisant le coupleur de codes OpenPALM [Buis et al., 2006]. Cette maquette PALM permet de résoudre itérativement les deux problèmes (9.1) et (9.2) jusqu’à convergence de l’algorithme. La maquette PALM est lancée de manière itérative à travers un pilote. Celui-ci est un script KornShell qui prend en entrée le premier et le dernier pas de temps de couplage ainsi que la fréquence à laquelle celui-ci a lieu. En plus de lancer l’exécutable PALM le pilote modifie les fichiers de données pour Mascaret et Telemac en fonction des paramètres donnés en entrée, il gère les fichiers de sortie et de reprise...

Il faut préciser que cette implémentation de l’algorithme de couplage est loin d’être optimale et n’exploite pas à fond les possibilités offertes par PALM. En effet comme on résout les problèmes (9.1) et (9.2) sur une succession d’intervalles de temps de l’ordre de 8 à 10 secondes qui recouvrent la durée de l’épisode de crue la fréquence de couplage est très élevée. Sur chacun de ces intervalles de temps on va lancer la maquette PALM qui

gère le couplage et donc à chaque fois les codes de calcul vont relire les fichiers de données dont les fichiers de géométrie et de frottement dont la taille peut être très importante, notamment dans le cas du modèle 2D. Plutôt que d'intégrer la boucle en temps dans le pilote on aurait pu l'intégrer dans la maquette PALM, lire les données d'entrée qui ne varient à la toute première itération, les conserver en mémoire et les transmettre aux unités PALM lorsque celles-ci en ont besoin.

9.2.3 Expérience numérique de couplage à interfaces sur l'épisode de 2014

Dans cette section nous procédons à la mise en place de l'algorithme de couplage sur l'épisode de crue de 2014 (voir figure C.8). Nous avons choisi cet épisode car c'est à la fois le plus long et le plus important épisode de crue dont on dispose sur la Nive. Pour cet épisode on dispose des observations à la station de Pont-Blanc. Il y a également l'épisode 2009-1 qui présente un important pic de crue sur la Nive mais on ne dispose pas des observations à la station de Pont-Blanc pour cet épisode.

Construction des lignes d'eau initiales pour les sous-modèles 1D et le modèle 2D

Dans le cas de Bayonne l'algorithme de couplage de modèle à interfaces tel que présenté en section 9.2.2 se caractérise par une difficulté importante à converger. Aussi les conditions initiales pour chacun des modèles ne peuvent pas être choisies au hasard sous peine de voir l'algorithme diverger, il faut qu'il y ait une cohérence physique entre les modèles. Par exemple Telemac-2D offre la possibilité de choisir comme condition initiale sur le domaine 2D une ligne d'eau avec cote constante mais les nombreuses expériences numériques menées pendant ce travail ont montré que lorsqu'on choisit cette ligne d'eau comme condition initiale pour le modèle 2D alors l'algorithme diverge.

Aussi, on construit une condition initiale pour le modèle 2D en faisant tourner le modèle Adour maritime 1D complet sur une période d'environ 24 heures (que l'on appelle dans la suite période d'équilibrage) précédant le début de la période où on va appliquer l'algorithme de couplage. Les débits et hauteurs d'eau calculées aux limites du modèle 2D avec le modèle Adour maritime 1D servent de forçage au modèle 2D sur la période d'équilibrage. La ligne d'eau calculée avec le modèle 2D à l'issue de cette période d'équilibrage va servir comme condition initiale au modèle 2D couplé.

Les conditions initiales sur les sous-modèles 1D sont simplement la restriction à chaque sous-domaine de la ligne d'eau calculée avec le modèle Adour maritime à l'issue de la période d'équilibrage.

Les lignes d'eau ainsi construites pour chacun des modèles présentent une cohérence physique qui va permettre à l'algorithme de converger.

Fréquence de couplage et temps de calcul

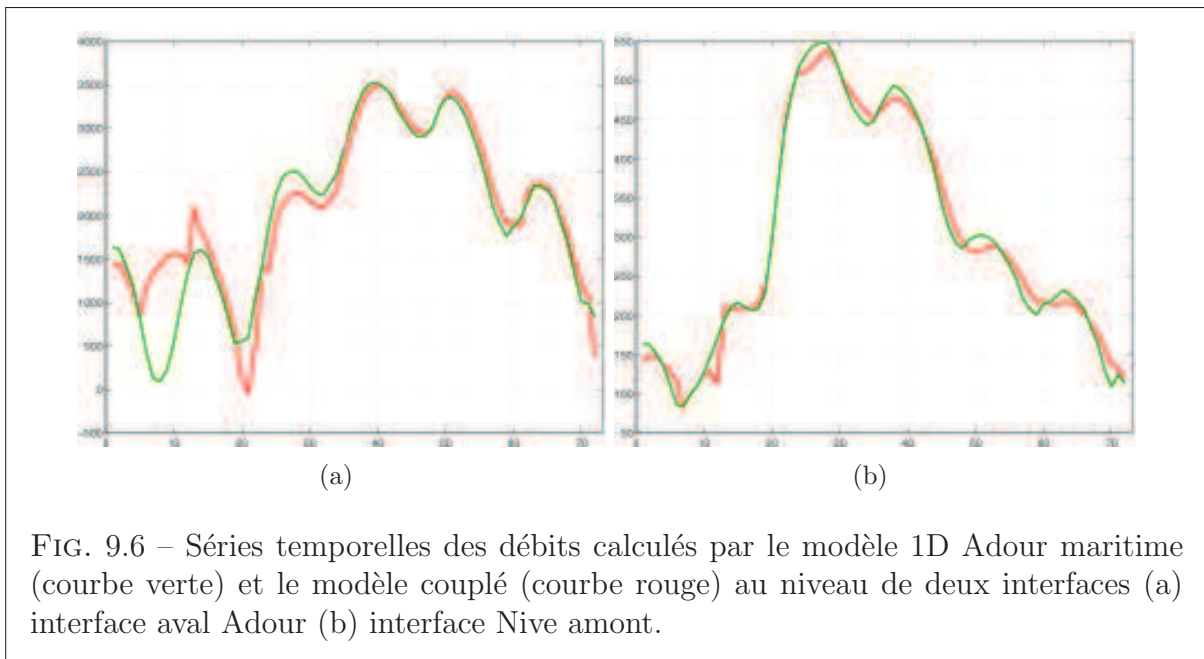
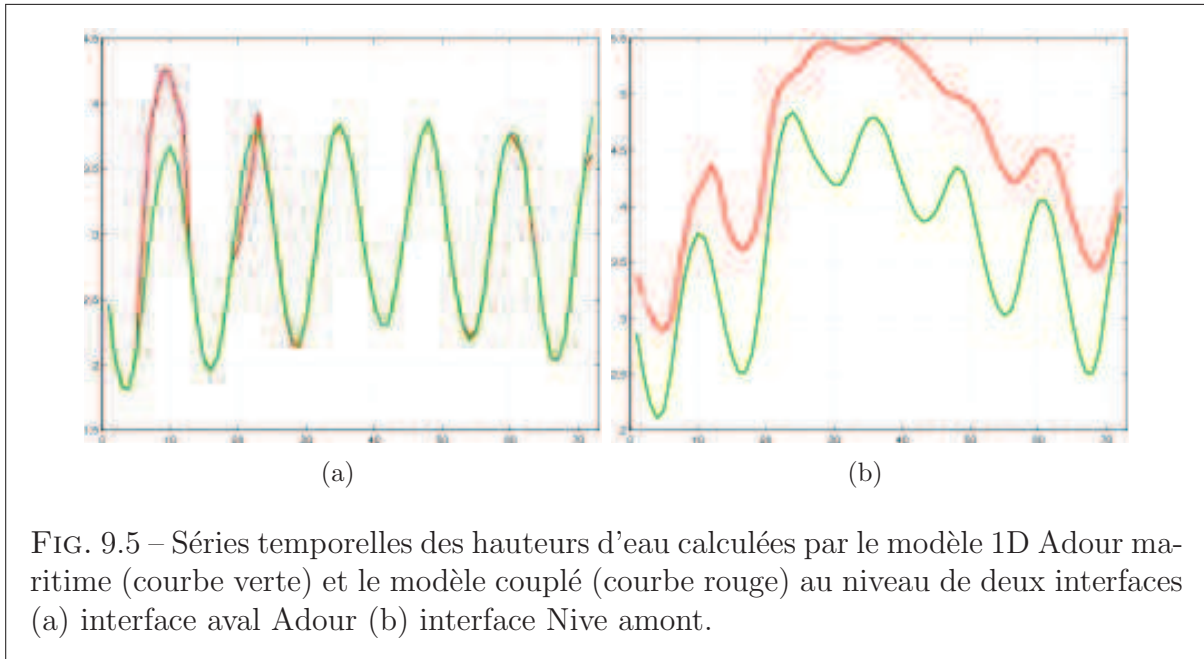
Comme on l'a déjà évoqué plus haut l'algorithme de couplage à interface utilisé dans ce travail présente des difficultés à converger ce qui se traduit une contrainte importante sur le choix des pas de temps des modèles ainsi que sur la durée des intervalles de temps

sur lesquels on applique l'algorithme de Schwarz global en temps. Dans ce travail nous avons été amenés à appliquer l'algorithme de Schwarz global en temps sur des intervalles de temps de 8 secondes sur lesquels le pas de temps des modèles 1D était lui-même de 8 secondes et le pas de temps du modèle 2D de 4 secondes. Nous avons testé de nombreuses autres configurations avec des pas de temps plus importants comme par exemple un algorithme de Schwarz global en temps sur des intervalles d'une heure avec un pas de temps du modèle 1D de 300 secondes et un pas de temps du modèle 2D de 4 secondes mais aucune de ces configurations n'a donné satisfaction hormis celles avec des intervalles et des pas de temps plus petits que ceux que nous avons utilisé et mentionné.

Cette fréquence de couplage importante se traduit par un temps de calcul très important puisque le rapport $\frac{\text{Temps de calcul}}{\text{Temps simulé}}$ est de l'ordre de 0.16. Néanmoins nous sommes convaincus que ce rapport pourrait encore être réduit si la boucle en temps sur laquelle est faite le couplage était intégrée directement dans la maquette de couplage PALM. Les fichiers de données (en particulier celui de la géométrie du modèle 2D) seraient lus lors de la première itération puis récupérés en mémoire à chaque appel des codes Mascaret et Telemac ce qui permettrait d'éviter de relire ces fichiers lors de l'appel de ces codes.

Résultats numériques

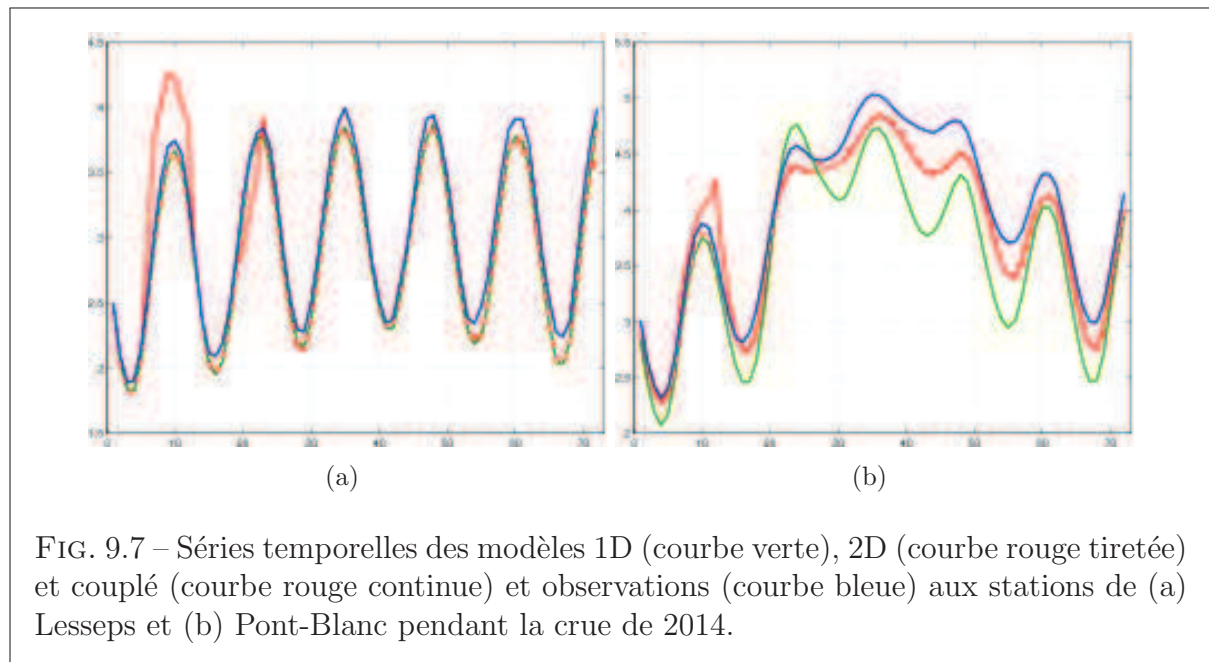
Usuellement pour pouvoir estimer l'erreur commise par le modèle couplé au niveau des interfaces 1D/2D on compare (aux interfaces) celui-ci à un modèle 2D de référence recouvrant l'union des domaines 1D et 2D. Sur l'Adour nous ne disposons pas d'un tel modèle 2D de référence par contre nous disposons du modèle 1D Adour maritime qui recouvre l'ensemble de la zone d'étude et nous pouvons donc comparer le modèle couplé et le modèle 1D au niveau des interfaces. C'est l'objet des figures 9.5 et 9.6 qui représentent respectivement les hauteurs d'eau et les débits aux interfaces sur l'Adour à la condition limite aval (station de Lesseps) et à la condition limite amont sur la Nive.



On note tout d’abord une discontinuité entre les résultats du modèle 1D et ceux du modèle couplé. C’est un phénomène qui se produit en début de calcul en général mais dont nous ne nous expliquons pas bien l’apparition. Nous faisons l’hypothèse que les interfaces ne sont pas absorbantes et que des ondes peuvent être réfléchies entraînant ces discontinuités. En dehors de ces discontinuités on note que les hauteurs d’eau ainsi que les débits calculés par les modèles au niveau de l’interface aval sont très proches. Au niveau de la condition limite amont par contre les hauteurs d’eau sont très différentes tandis que les débits sont proches. Cela illustre la différence de relation hauteur/débit qu’il peut

exister entre les modèles alors même que globalement les coefficients de frottement du modèle Adour maritime sur la zone de Bayonne sont plus faibles (et donc le frottement est plus important) que celui du 2D. Ainsi le modèle couplé modifie fortement la dynamique de l'écoulement sur la zone de Bayonne par rapport au modèle 1D Adour maritime. On n'a pas présenté ici la comparaison entre les hauteurs d'eau et les débits au niveau de l'interface en amont de l'Adour, mais les résultats à cette interface vont dans le même sens que ceux présentés sur l'interface amont sur la Nive.

Un autre moyen d'estimer l'efficacité du couplage est de comparer les résultats du modèle couplé aux observations faites au niveau des stations sur le domaine. Comme cela est représenté sur la figure 9.3) on dispose de deux stations sur le modèle 2D, l'une à Pont-Blanc et l'autre à Lesseps. Nous avons donc comparé les sorties du modèle couplé à ces stations aux observations ainsi qu'aux modèle 1D Adour maritime et modèle 2D (forcé simplement par le 1D). Les résultats sont reportés en figure 9.7.



On note que le modèle 2D et le modèle couplé reproduisent assez bien les hauteurs d'eau aux stations de Pont-Blanc et Lesseps pendant l'épisode de crue et sont tous les deux très proches dans les résultats calculés. Cette dernière observation est normale car comme on l'a vu en figure 9.6 le couplage ne modifie que très peu les débits à l'amont du modèle 2D et pas du tout les hauteurs d'eau à l'aval de celui-ci ainsi modèle 2D et modèle couplé donnent presque les mêmes résultats. Néanmoins lors du pic de crue à Pont-Blanc modèles 2D et couplé sous-estiment les hauteurs d'eau, de fait il pourrait être intéressant de combiner couplage à interfaces et assimilation de données pour corriger ce biais mais cela n'a pas fait partie des objectifs de ce travail. Enfin on note que le modèle 1D sous-estime quant à lui nettement plus les hauteurs d'eau que le modèle 2D.

Les résultats exposés ci-dessus en figure 9.7 sont à mettre néanmoins en regard des temps de calcul sur cet épisode pour chacun des modèles. On a reporté dans le tableau 9.1 les temps de calcul sur cette crue pour chacun des modèles 1D, 2D et couplé.

Modèle	1D	2D	Couplé
Temps de calcul	84	3430	39396

TAB. 9.1 – Temps de calcul (s.) des modèles 1D, 2D et couplé sur la crue de Janvier 2014.

Au vu du tableau 9.1 on pourrait être tenté de repousser d’emblée le modèle couplé pour son important coût de calcul. Néanmoins le choix du modèle à utiliser sur la zone de Bayonne va fortement dépendre du contexte et de ce que l’on souhaite modéliser. Dans un cadre opérationnel il est clair que les temps de calcul des modèles 2D et couplé sont prohibitifs et le modèle 1D sera plus indiqué. Maintenant si l’on s’intéresse à l’écoulement sur toute la zone de Bayonne il est clair qu’il faut privilégier le modèle 2D ou le modèle couplé. Enfin, étant donné les différences de hauteurs d’eau entre le modèle couplé et le modèle 1D (dont les sorties servent à forcer le modèle 2D non couplé), il apparait que si l’on souhaite avoir une modélisation qui soit physique sur tout le domaine il faut utiliser le modèle couplé. Cela peut être utile si l’on souhaite rejouer des crues *a posteriori*. Nous tenons à rappeler une fois encore que le temps de calcul du modèle couplé pourrait être réduit si l’implémentation du modèle couplé était différente. En effet si l’on intégrait la boucle en temps dans la maquette PALM avec stockage des fichiers de données en mémoire on pourrait éviter la lecture de ces fichiers à chaque appel aux codes Telemac et Mascaret et ainsi gagner en temps de calcul.

9.3 Couplage de modèles superposés

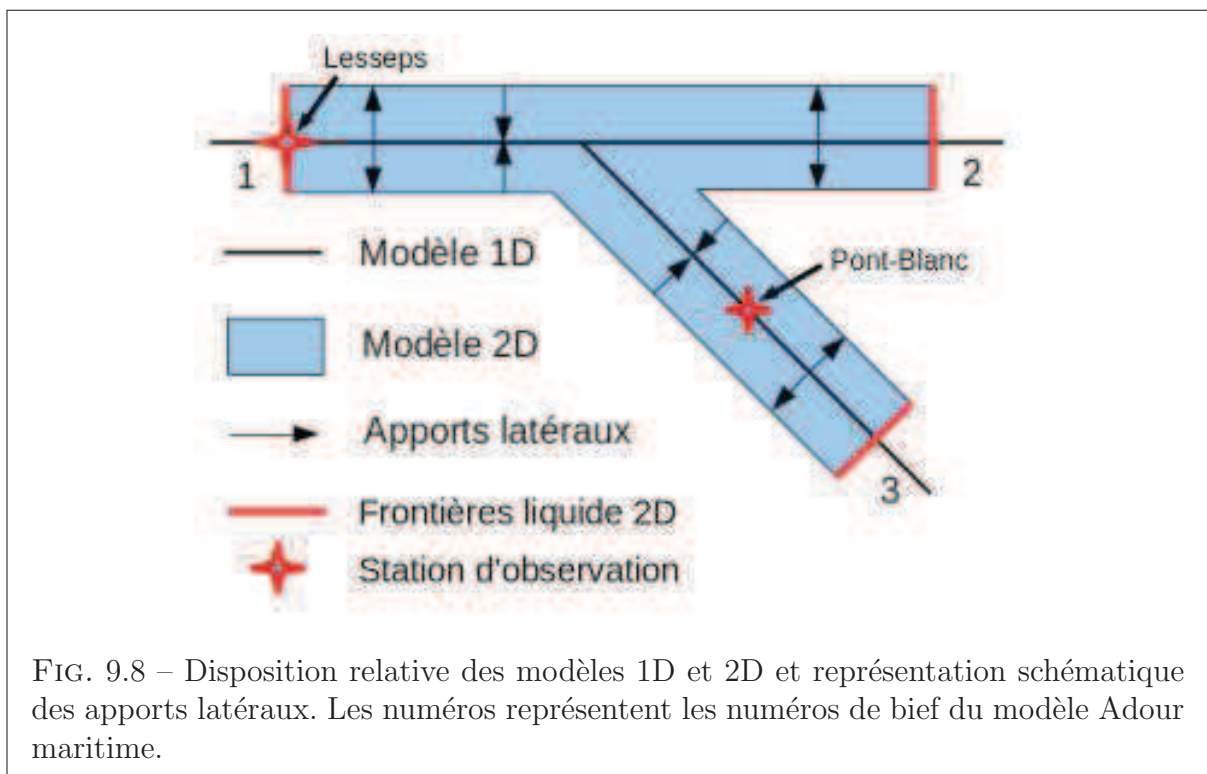
Nous avons vu dans la section précédente les limitations du couplage à interfaces sur la ville de Bayonne. Dans cette section nous abordons la problématique du couplage de modèles sur cette zone d’un point de vue totalement différent puisque nous nous intéressons à la problématique du couplage de modèles superposés. Cette configuration bien que pouvant sembler assez originale *a priori* apparait en fait de manière assez naturelle dans les faits. Comme cela est le cas avec le réseau Adour maritime et Bayonne il peut arriver que les services de prévision des crues développent un modèle 1D sur un réseau hydraulique puis dans certaines zones d’intérêt développent ensuite un modèle 2D local mais ne souhaitent pas modifier le modèle 1D existant. Dans ce cas on ne peut pas utiliser les méthodes de couplage à interfaces exposées aux sections 5.1 et 9.2 et il faut alors utiliser les méthodes de couplage de modèles superposés exposées en section 5.2.

Comme cela a été exposé à la section 5.2 cette méthode repose sur le calcul avec le modèle 2D de termes d’apports latéraux pour le modèle 1D qui décrivent les échanges entre lit mineur et lit majeur lorsque le modèle 2D est en crue. C’est pourquoi dans un premier temps nous présentons le calcul de ces termes d’apports avant d’explicitier l’algorithme de couplage utilisé puis de présenter l’application de la méthode au cas de l’épisode de crue de 2014.

9.3.1 Modèles utilisés, disposition et termes échangés entre les modèles

Modèles utilisés et disposition

Les modèles utilisés sont le modèle 1D Adour maritime présenté à la section 8.1.2 mais que l'on a « ajusté » sur la zone de recouvrement (voir section suivante) et le modèle 2D sur Bayonne présenté en section 9.1. Comme on l'a précisé ces modèles sont superposés sur la zone de Bayonne. On illustre cette configuration avec la figure 9.8.



Ajustement de la géométrie et des coefficients de frottement du modèle 1D sur la zone de recouvrement

Le modèle 1D Adour maritime et le modèle 2D sur Bayonne ont été conçus (sur la zone de recouvrement) chacun avec des jeux de données différents et ont été calés avec des jeux de coefficients de Strickler différents. Aussi il s'est révélé nécessaire pour que le couplage ait un sens physique de redéfinir la géométrie du modèle 1D sur la zone de recouvrement à partir de la géométrie du modèle 2D de manière à ce que les géométries des deux modèles sur la zone soient proches.

D'autre part il a fallu ajuster légèrement les coefficients de Strickler de manière à ce que les modèles se comportent d'une manière assez proche en termes de hauteurs d'eau. Sur la zone de recouvrement nous avons donc fixé les coefficients de Strickler du

modèle 1D légèrement inférieurs à ceux du modèle 2D de manière à réduire la différence de comportement des modèles.

Termes échangés entre les modèles

Dans le couplage de modèles superposés le modèle 1D force le modèle 2D au niveau de ses frontières mouillées. Le régime d'écoulement de chacun des modèles sur la zone de recouvrement est toujours fluvial c'est pourquoi, de la même manière que ce qui avait été fait pour le modèle 2D en section 9.2.1, celui-ci est forcé à l'amont par des débits et à l'aval par des hauteurs d'eau calculés par le modèle 1D aux sections en travers correspondantes.

Pour ce qui est des termes d'apports latéraux pour le modèle 1D nous avons implémenté la méthode présentée dans (Marin et Monnier, 2009) et exposée en section 5.2.1. Ces termes d'apports latéraux sont nuls lorsque le modèle 2D n'est pas en crue. Dans la suite nous considérerons donc que le modèle 2D est en crue et donc que ces termes ne sont pas nuls. Pour appliquer la méthode présentée dans (Marin et Monnier, 2009) nous avons décomposé la zone constituée du lit mineur (délimitée sur la figure 9.9-(a) par les lignes rouges) en plusieurs zones que nous appellerons par la suite sections, celles-ci étant délimitées entre elles par des traits noirs (figure 9.9-(a)). Sur chacune de ces sections nous calculons le débit à l'interface entre lit mineur et lit majeur. La connaissance de ces débits nous permettra en effet de calculer le terme source :

$$\Psi = -(q_1 + q_2, u|_1 q_1 + u|_2 q_2)$$

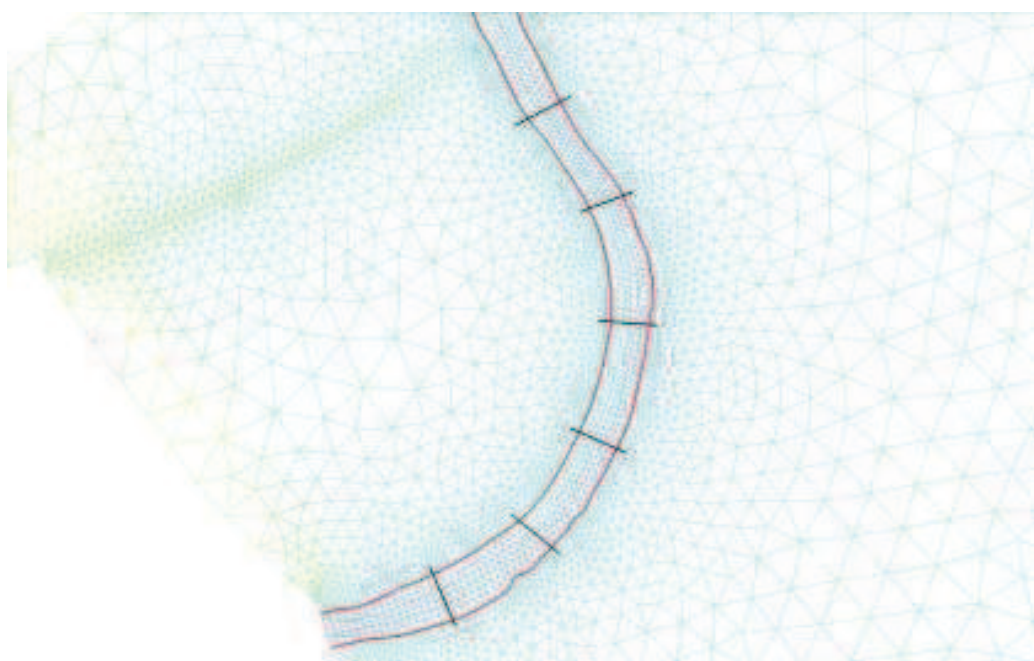
pour les équations de Saint-Venant 1D. Ici les indices 1 et 2 désignent le numéro de berge, voir figure 5.10. Dans la suite on considère que composante de la vitesse selon l'axe des x , u , est constante sur la largeur de la section en travers, ainsi $u|_1 = u|_2 = u_{1D}$ et

$$\Psi = -(q_1 + q_2, u_{1D}(q_1 + q_2))$$

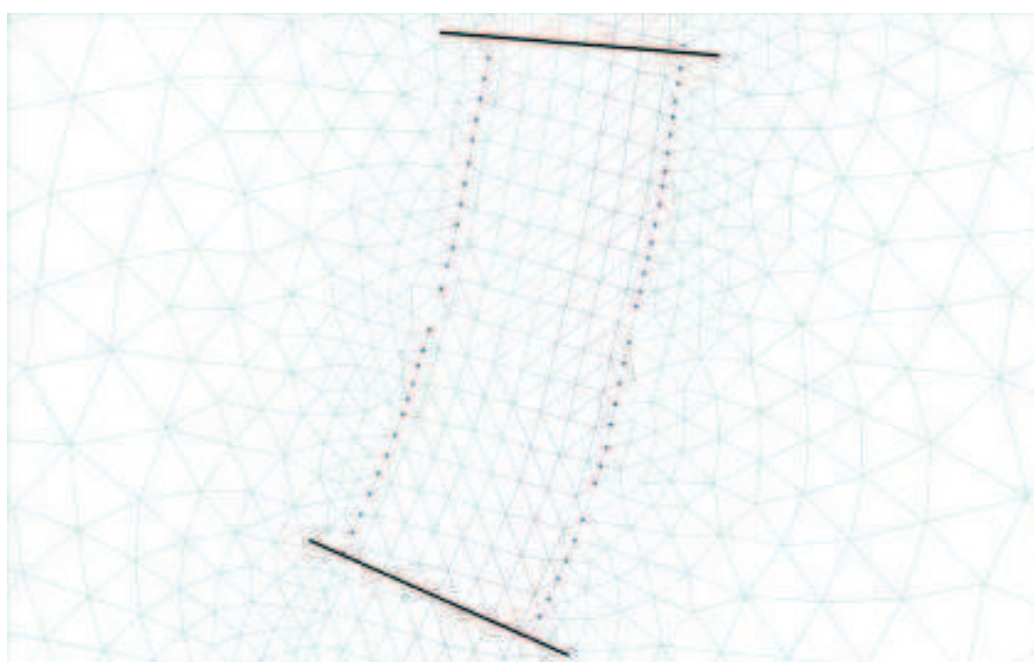
On rappelle que le terme Ψ représente les apports pour une section en travers ou dit autrement en un point du modèle 1D.

L'interface entre lit mineur et lit majeur sur une section est représentée sur la figure 9.9-(b) par les points rouges et les arêtes du maillage 2D qui relient ces points. Le débit sur chaque berge est calculé comme étant la somme des débits sur chacune des arêtes constituant la berge. A noter qu'il s'agit d'un débit signé : le débit est négatif si de l'eau sort de la rivière et il est positif sinon. On fait alors la somme des débits calculés sur chacune des rives de la section et on divise par la longueur de la section, on obtient un débit d'apport linéique que l'on identifie au terme $-(q_1 + q_2)$ dans l'expression de Ψ . Ce débit d'apport linéique (qui est aussi le débit moyenné sur la longueur de la section) est ensuite transmis au code Mascaret-1D pour calculer Ψ . Le terme Ψ est appliqué comme terme source linéique par Mascaret-1D sur la longueur de la section. Cette quantité bien que transmise par PALM est ensuite injectée dans les équations de Mascaret de la même manière que cela est fait pour apports latéraux classiques spécifiés dans des fichiers.

Comme on l'a déjà dit ce terme source représente les échanges, en termes de conservation de la masse et de la quantité de mouvement, entre lit mineur et lit majeur. Il faut



(a)



(b)

FIG. 9.9 – (a) Découpage du lit mineur (délimité par les bandes rouges) en sections sur une portion à l'amont de la Nive (b) Interfaces entre lit mineur et lit majeur sur une section (délimitée par les traits noirs), il s'agit des lignes formées par les arêtes reliant les points rouges.

donc souligner ici la différence de traitement des échanges (entre lit mineur et lit majeur) que l'implémentation des débits d'apport dans Mascaret implique entre Mascaret-1D (auquel on applique le terme source Ψ) et Telemac-2D. Dans Telemac-2D les échanges entre lit mineur et lit majeur se font en quelque sorte de manière implicite à travers la résolution des équations de Saint-Venant sur l'ensemble du domaine. Tandis que dans Mascaret-1D les termes sources sont appliqués de manière linéaire. Ainsi si dans Telemac-2D le débit linéique calculé le long d'une berge peut varier, le terme source correspondant appliqué à Mascaret sur la même section sera quant à lui constant puisqu'il s'agira de la moyenne des débits linéiques calculés avec Telemac-2D sur la même zone. Hormis la différence de dimensions entre les modèles nous verrons que cet état de fait introduit une différence notamment sur les débits calculés dans le lit mineur par chacun des modèles.

9.3.2 Algorithme de couplage et implémentation

Algorithme de couplage

L'algorithme de couplage que nous avons mis en place dans le cadre de modèles superposés est un algorithme de type Schwarz global en temps. Les notations sont celles utilisées précédemment à la section 9.2.2. On suppose que sur l'intervalle de temps $[T_1; T_2]$ on connaît le terme source pour le modèle 1D Ψ^0 là où les modèles sont superposés (i-e sur $(\Omega_1 \cap \Omega_2)$). Si on ne le connaît pas on peut néanmoins choisir Ψ^0 constant égal à la dernière valeur calculée par le modèle 2D.

$$\left\{ \begin{array}{ll} \mathcal{L}_1(S, Q)^{k+1} = 0 & \text{sur } (\Omega_1 \setminus \Omega_1 \cap \Omega_2) \times [T_1; T_2] \\ \text{Conditions limites classiques} & \text{sur } \partial\Omega_1 \times [T_1; T_2] \\ \mathcal{L}_1(S, Q)^{k+1} = \Psi^k & \text{sur } (\Omega_1 \cap \Omega_2) \times [T_1; T_2] \end{array} \right. \quad (9.3)$$

et

$$\left\{ \begin{array}{lll} \mathcal{L}_2(h, u, v)^{k+1} = 0 & \text{sur } \Omega_2 \times [T_1; T_2] \\ \text{Conditions limites classiques} & \text{sur } \partial\Omega_2 \times [T_1; T_2] \\ (u, v)_j^{k+1} = B_{12}(Q_j^{k+1}) & \text{sur } \Gamma_j \times [T_1; T_2] & j = 2, 3 \\ h_{2D,j}^{k+1} = h_{1D,j}^{k+1} & \text{sur } \Gamma_j \times [T_1; T_2] & j = 1 \end{array} \right. \quad (9.4)$$

Implémentation

Là encore l'algorithme de couplage a été implémenté en utilisant le coupleur de codes OpenPALM qui va résoudre successivement et jusqu'à convergence les problèmes (9.3) et (9.4). La maquette OpenPALM est lancée toutes les heures (ce qui correspond à la fréquence de couplage) par un pilote de manière itérative pendant la durée d'un épisode de crue. Le choix de lancer la maquette OpenPALM toutes les heures n'a pas été fait au hasard, en effet cela correspond à la fréquence à laquelle on assimile les observations dans la maquette d'assimilation de l'IEnKF. Ainsi la maquette de couplage de modèles superposés a-t-elle été conçue, entre autres choses, dans l'optique de combiner couplage de modèles superposés et assimilation de données sur le modèle 1D comme on le verra dans la section 9.4.

9.3.3 Expérience numérique de couplage de modèles superposés sur l'épisode de 2014

Construction des lignes d'eau initiales

La construction des lignes d'eau initiales pour chacun des modèles n'est pas aussi contrainte que dans le cas du couplage à interface puisque dans ce cas on n'a pas besoin que les modèles présentent une cohérence physique *a priori* au commencement du couplage. Pour le modèle 1D on construit une ligne d'eau initiale en propageant une ligne d'eau quelconque sur une période de 24 heures précédant le début du couplage (qu'on appelle période d'équilibrage) de manière ce qu'au début du couplage la ligne d'eau initiale pour le modèle 1D soit cohérente avec les conditions limites à ce temps. Pour ce qui est du modèle 2D, Telemac permet de choisir comme condition initiale une ligne d'eau de cote constante sur tout le domaine. C'est le choix que nous faisons et nous fixons la cote égale à celle calculée par le modèle 1D à l'aval.

Fréquence de couplage, itération de couplage et critère d'arrêt

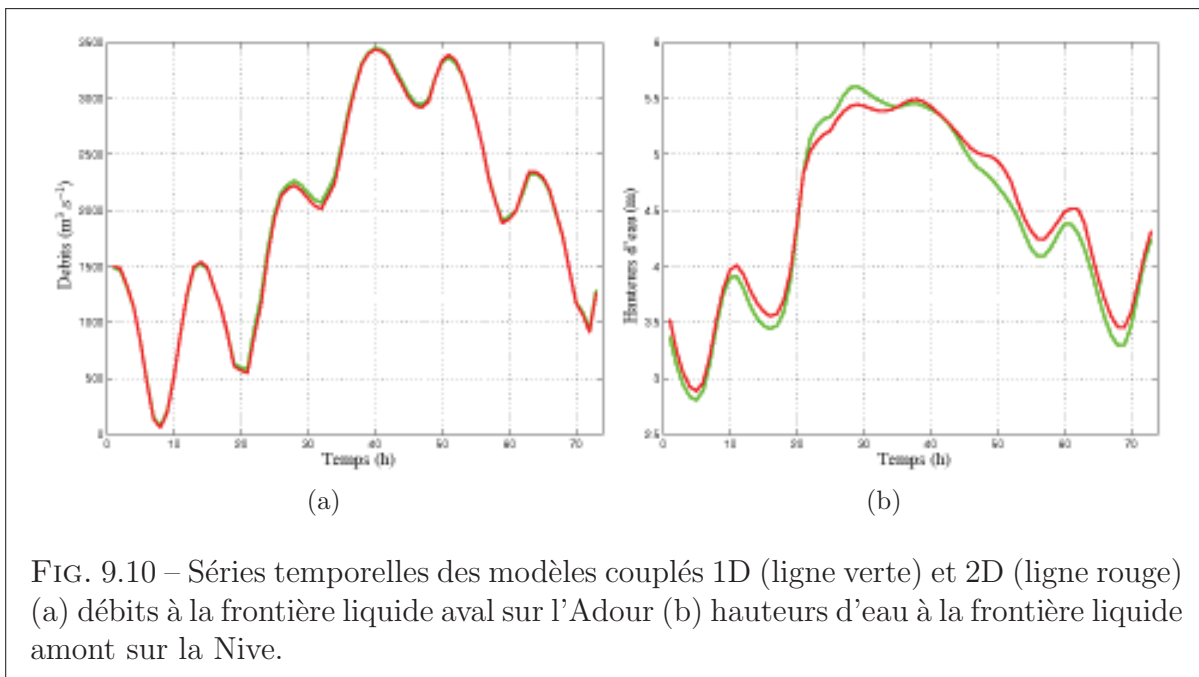
Nous appliquons l'algorithme de couplage présenté dans la section 9.3.2 sur une succession d'intervalles de temps $I_n = [T_{n-1}; T_n]$ d'une heure couvrant la durée de l'épisode. Cependant comme nous le verrons dans la section sur les résultats numériques il n'est guère possible avec cette méthode de couplage de fixer un critère d'arrêt aussi fin que dans le cas du couplage de modèles à interface. Néanmoins il apparaît que dès lors que le modèle 2D est en crue le couplage des modèles a un impact sur la dynamique de chacun d'eux par rapport à chaque modèle pris séparément. Aussi plutôt que de fixer un critère d'arrêt sur les invariants de Riemann ou sur les variables d'état aux limites du modèle qui ne pourrait pas être vérifié nous avons pris le parti, lorsque le modèle 2D est en crue, d'effectuer deux itérations de couplage. C'est-à-dire qu'on résout successivement deux fois les problèmes (9.3) et (9.4). Une première fois où le terme source pour le modèle 1D, Ψ^0 , est choisi constant égal à la dernière valeur calculée par le modèle 2D (i.e $\Psi^0 \equiv \Psi(T_n)$ sur $[T_n; T_{n+1}]$). Puis une seconde fois où le terme source pour le modèle 1D, Ψ^1 , est le terme source calculé précédemment par le modèle 2D, Ψ , sur l'intervalle de temps $[T_n; T_{n+1}]$, soit : $\Psi^1 = \Psi([T_n; T_{n+1}])$.

Résultats numériques

Dans cette section nous nous intéressons aux résultats numériques de l'application de la méthode de couplage de modèles superposés aux cas de Bayonne sur l'épisode de crue de la Nive de 2014 entre les jours 5 et 8 (voir figure C.8-(a)). Dans un premier temps nous justifions l'emploi du terme de couplage pour cette méthode au vue des résultats numériques. Dans un second temps nous comparons les résultats numériques obtenus aux stations disponibles sur le domaine de recouvrement (stations de Lesseps et Pont-Blanc, voir figures 9.1 et 9.8). Enfin nous nous intéresserons à travers l'étude des termes sources pour le modèle 1D calculés sur la zone Plaine d'Ansot/Barthes de Quartier Bas à la dynamique de l'écoulement sur cette zone.

Dans un premier temps nous nous intéressons à la question de la continuité, en termes de variable d'état, des modèles aux limites du modèle 2D. Comme nous l'avons précisé à

la section 9.3.1 un ajustement de la géométrie et des coefficients de Strickler du modèle 1D sur la zone de recouvrement a été nécessaire de manière à ce que le modèle 1D reproduise, lors du couplage, de manière la moins grossière possible la relation hauteurs d'eau/débits du modèle 2D et ainsi assurer une certaine continuité entre les modèles aux limites du modèle 2D. Comme on peut le voir sur la figure 9.10 aux limites du modèle 2D la relation hauteurs d'eau/débits est selon les frontières liquides considérées plus ou moins bien reproduite pour les variables d'état qui ne sont pas imposées aux frontières en question. Par exemple si l'on considère la frontière liquide aval sur l'Adour où les cotes sont imposées aux modèles 2D par le 1D on constate que les débits calculés en cet endroit par chacun des modèles semblent très proches. Par contre si l'on considère la frontière liquide amont sur la Nive où les débits sont imposés alors on remarque que les hauteurs d'eau de chacun des modèles restent relativement éloignées. D'autre part on remarque qu'après deux itérations l'algorithme de couplage n'apporte pas d'amélioration en termes de débits ou de hauteurs d'eau c'est pourquoi nous avons choisi de ne pas fixer pour celui-ci de critère d'arrêt sur les hauteurs d'eau et/ou les débits car en pratique celui-ci n'est presque jamais respecté.



Nous pouvons également considérer la différence en valeur absolue entre les invariants de Riemann de chacun des modèles aux limites du modèle 2D. Les invariants de Riemann étant une combinaison non linéaire des variables d'état des modèles ils permettent eux aussi d'estimer la continuité entre les modèles aux limites du modèle 2D mais aussi de savoir à quelle frontière liquide la continuité entre les modèles est la plus assurée. En toute logique la différence entre les invariants de Riemann est relativement importante et notamment pendant le pic de crue à la condition limite amont sur la Nive.

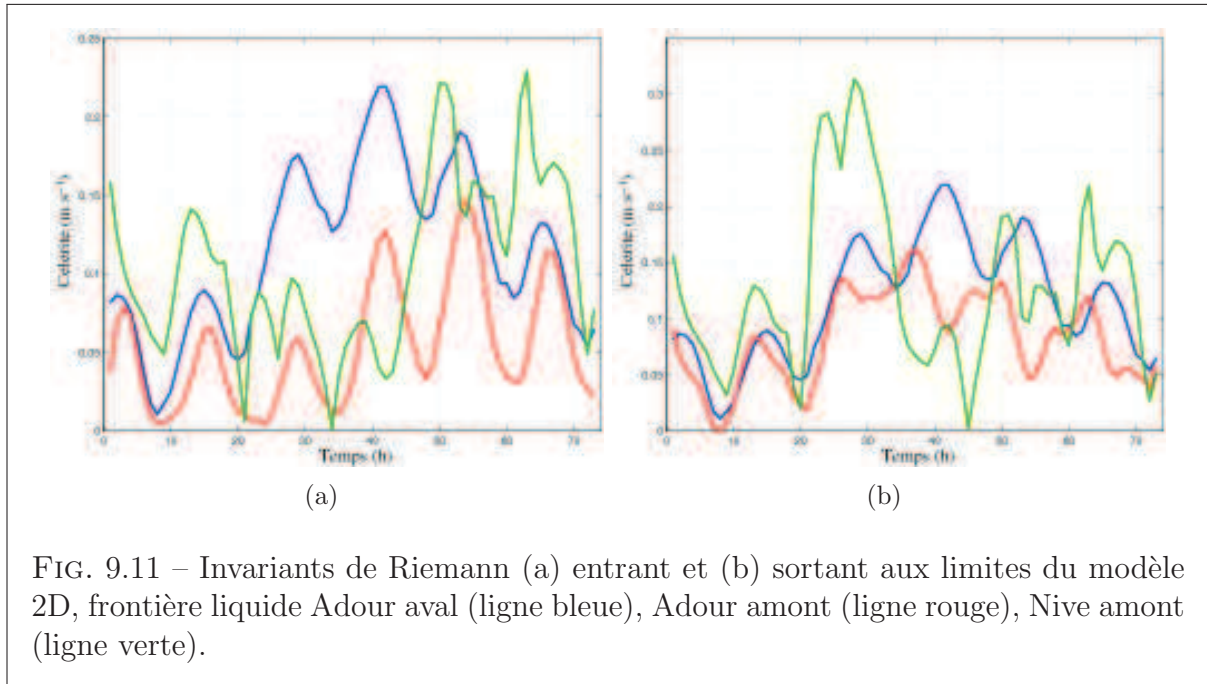


FIG. 9.11 – Invariants de Riemann (a) entrant et (b) sortant aux limites du modèle 2D, frontière liquide Adour aval (ligne bleue), Adour amont (ligne rouge), Nive amont (ligne verte).

Néanmoins malgré ces différences aux limites du modèle 2D nous pensons que cela a du sens de parler de modèle couplé. En effet le modèle 1D force le modèle 2D et celui-ci en retour calcule les termes sources correspondant aux échanges entre lit mineur et lit majeur à appliquer au modèle 1D et les deux modèles évoluent ainsi de manière cohérente comme nous le verrons plus loin.

Mais tout d'abord nous nous penchons sur l'influence du couplage sur les modèles couplés par rapport aux modèles non couplés. Ainsi on s'intéresse à l'influence du couplage sur les grandeurs imposées aux limites du modèle 2D par le modèle 1D. On considère par exemple les hauteurs d'eau imposées à l'aval du modèle 2D et les débits imposés à l'amont du modèle 2D sur la Nive par le modèle 1D avec et sans couplage (figure 9.12). On n'a pas reporté les débits calculés au niveau de la condition limite amont sur l'Adour car les conclusions sont les mêmes. Comme on peut le constater l'influence du couplage sur les grandeurs imposées aux limites du modèle 2D est presque nulle. En effet en dehors du pic de crue sur la Nive où on peut observer une légère différence entre les modèles 1D couplé et non couplé les grandeurs imposées aux limites du modèle 2D sont les mêmes. Or l'impact de cette différence en termes de débits sur la condition limite amont de la Nive est négligeable comme on peut le voir sur la figure 9.12.

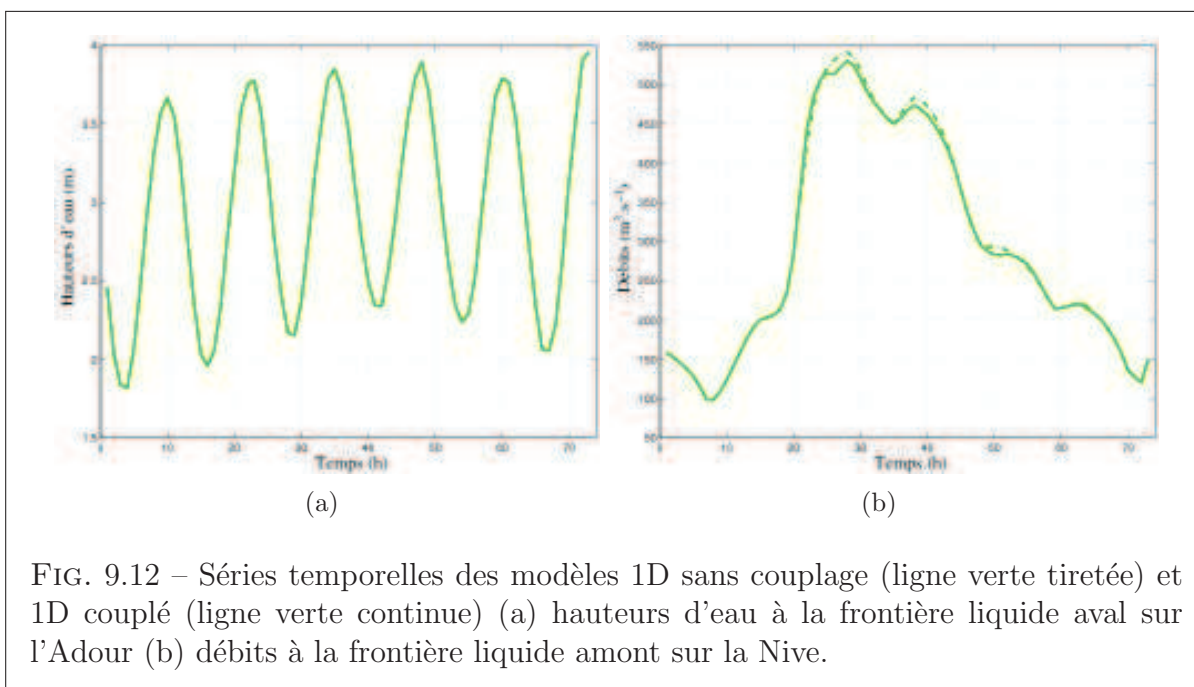
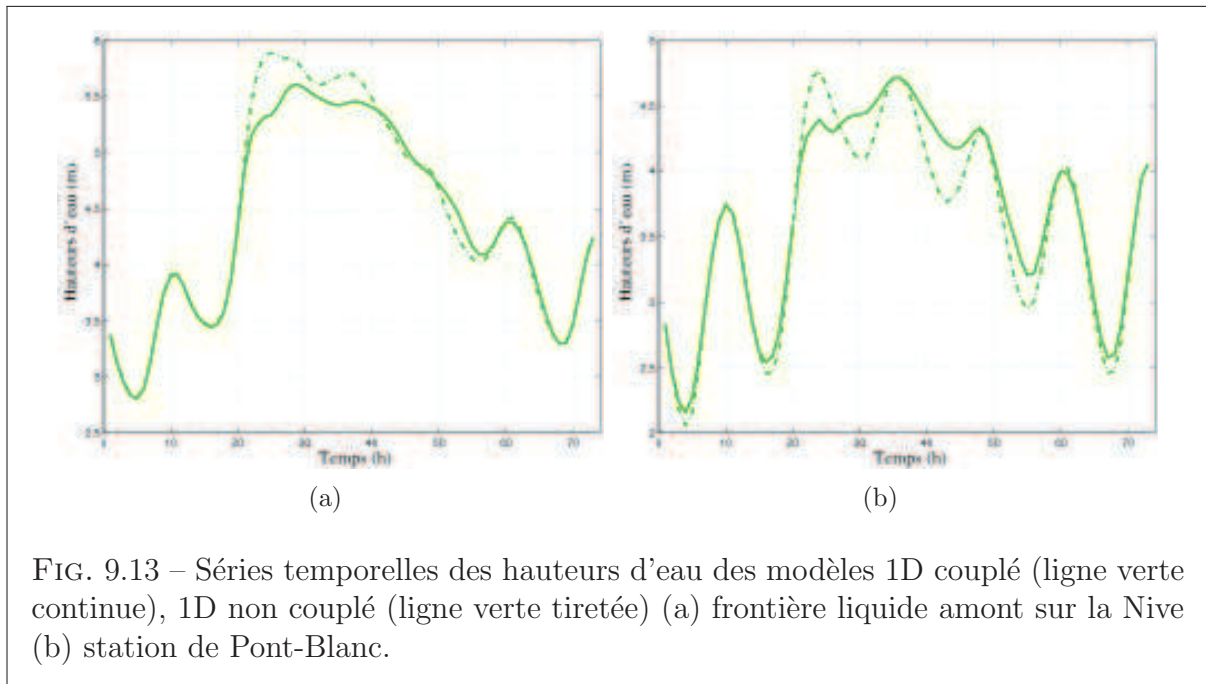


FIG. 9.12 – Séries temporelles des modèles 1D sans couplage (ligne verte tiretée) et 1D couplé (ligne verte continue) (a) hauteurs d'eau à la frontière liquide aval sur l'Adour (b) débits à la frontière liquide amont sur la Nive.

Le couplage a plus d'influence sur le modèle 1D là où les échanges entre lit mineur et lit majeur sont les plus importants c'est-à-dire sur la zone Plaine d'Ansot/Barthes de Quartier Bas. Tout d'abord au niveau de la condition limite amont sur la Nive cela a un impact sur les hauteurs d'eau du modèle 1D (figure 9.13-(a)) mais également au niveau de la station d'observation de Pont-Blanc (figure 9.13-(b)) où l'on peut constater que la dynamique du modèle 1D est totalement modifiée.



Le couplage a également une influence sur les débits du modèle 1D comme on peut le voir sur la figure 9.14 en réduisant les débits du modèle 1D couplé avant le pic de crue du fait des termes sources injectés et au contraire en augmentant les débits après le pic de crue. On explique la différence en termes de débits entre les modèles 1D et 2D couplés par les approximations faites sur le calcul du débit avec le modèle 2D. D'autre part comme on peut le voir sur la figure 9.14-(b) le couplage permet de réduire la différence en termes de hauteurs d'eau entre les modèles 1D et 2D (couplés). Nous ne reportons pas sur cette figure les hauteurs d'eau du modèle 2D non couplé car comme les forçages sont quasi-identiques aux limites du domaine 2D les hauteurs d'eau modélisées dans le domaine sont identiques. Le couplage permet donc de corriger la relation hauteurs d'eau/débits du modèle 1D pour qu'elle se rapproche de celle du modèle 2D. Dans la suite nous considérerons que cette relation dans chacun des modèles est équivalente et c'est pourquoi nous utiliserons l'assimilation des observations de hauteurs d'eau de la station Pont-Blanc pour corriger la ligne d'eau du modèle 1D.

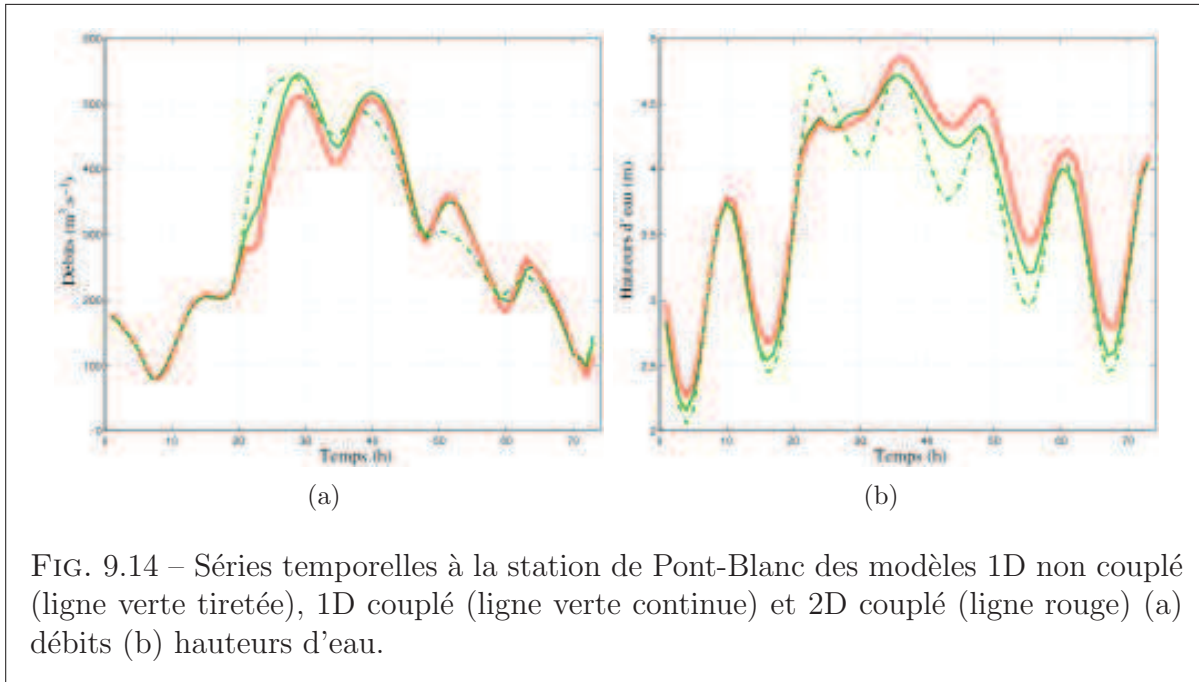
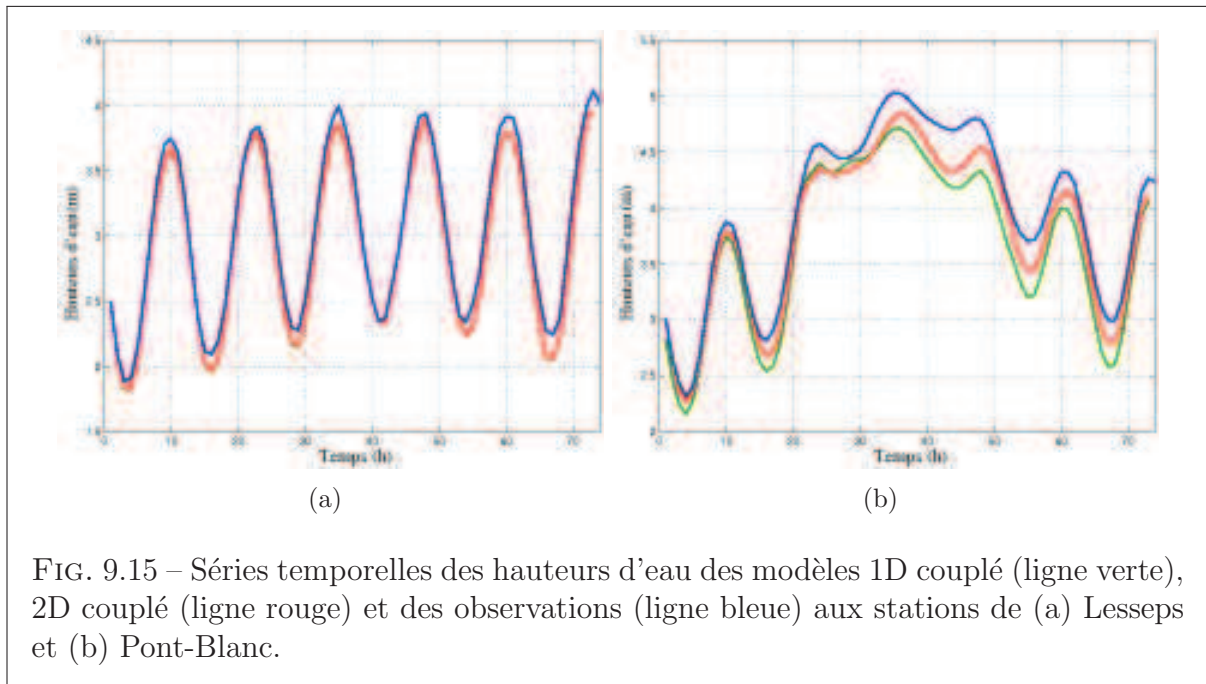


FIG. 9.14 – Séries temporelles à la station de Pont-Blanc des modèles 1D non couplé (ligne verte tiretée), 1D couplé (ligne verte continue) et 2D couplé (ligne rouge) (a) débits (b) hauteurs d'eau.

D'autre part il convient de comparer le modèle couplé aux stations d'observations disponibles sur le domaine. A la station de Lesseps où le modèle de base est déjà assez bon et où le couplage n'a pas d'influence sur les hauteurs d'eau modélisées les résultats du modèle couplé sont très bons (figure 9.15-(a)). Pour ce qui est de la station de Pont-Blanc les résultats sont légèrement différents que l'on considère le modèle 1D couplé ou le modèle 2D couplé (voir figure 9.15-(b)). En effet les résultats du modèle 2D couplé sont meilleurs que ceux du 1D couplé mais néanmoins identiques à ceux du modèle 2D non couplé. Ainsi le couplage ne dégrade ni n'améliore les résultats du modèle 2D (les forçages étant identiques avec ou sans couplage), par contre il améliore les résultats du modèle 1D là où les échanges entre lit mineur et lit majeur sont intégrés comme termes sources pour le modèle 1D.



Nous nous penchons maintenant sur les échanges de débits à l’interface entre lit mineur et lit majeur et qui sont utilisés dans les termes d’apports latéraux injectés dans Mascaret.

Nous considérons dans un premier temps les échanges entre la Nive et la Barthes de Quartier Bas. Il faut tout d’abord noter que lorsque les débits indiqués sur la figure 9.16-(a) sont négatifs cela signifie que la Nive déborde dans le lit majeur. Au contraire lorsqu’ils sont positifs cela signifie que de l’eau passe du lit majeur au lit mineur. Il apparaît qu’il y a sur cette portion de la Nive deux sections fortement débordantes indiquées par les pastilles rouge et jaune sur la figure 9.16-(b). Il s’agit de débordements importants car ramenés aux débits imposés en amont à ce moment-là ils représentent environ de l’ordre de 15 à 20% de ces débits.

Tandis que sur la zone marquée par la pastille noire l’eau de la Barthes de Quartier Bas se jette dans la Nive. Cela est dû à la présence de l’autoroute A63 (qui passe en au de la figure 9.16-(b)) qui contraint l’eau à regagner son lit à cet endroit.

Cela illustre assez bien la dynamique de l’écoulement dans cette portion du domaine 2D : l’eau déborde principalement dans les sections rouges et jaunes puis se jette dans la section noire. Nous ne l’avons pas reporté ici mais une partie de l’eau qui déborde dans la Barthes de Quartier Bas rejoint directement la Plaine d’Ansot par le passage sous l’autoroute (voir figure 9.2). Ces échanges sont relativement importants puisqu’ils peuvent représenter jusqu’à 20% des débits à l’interface entre lit mineur et lit majeur sur la Barthes de Quartier Bas.

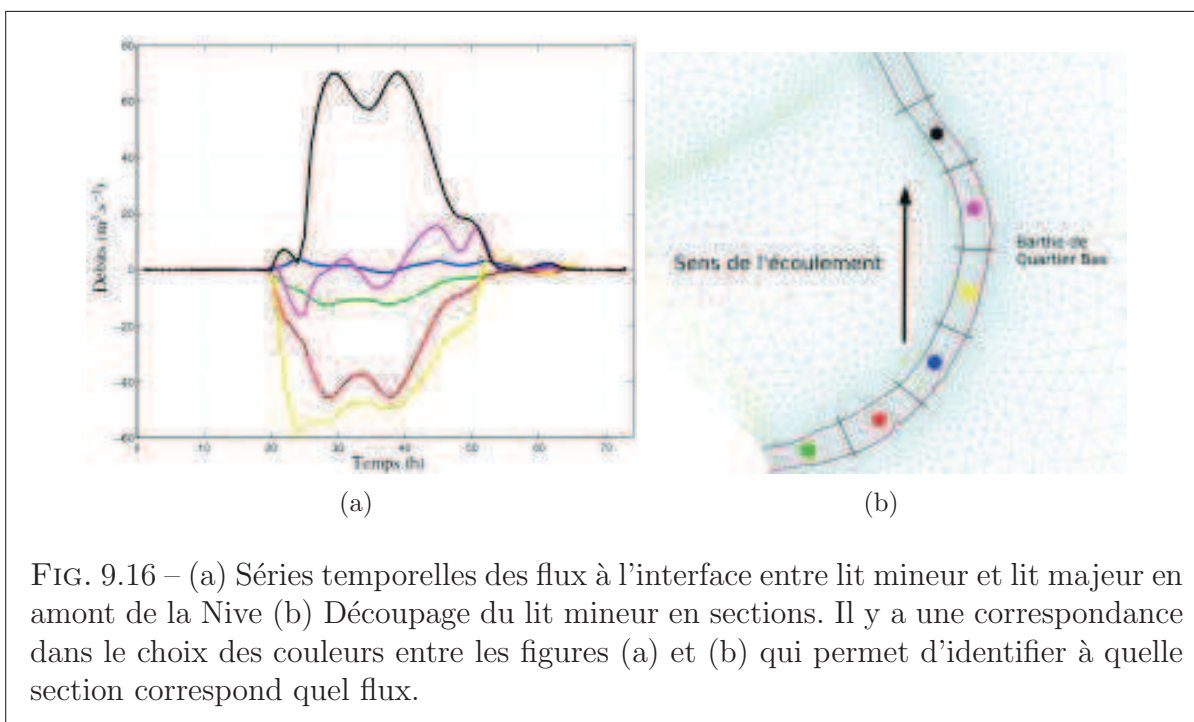
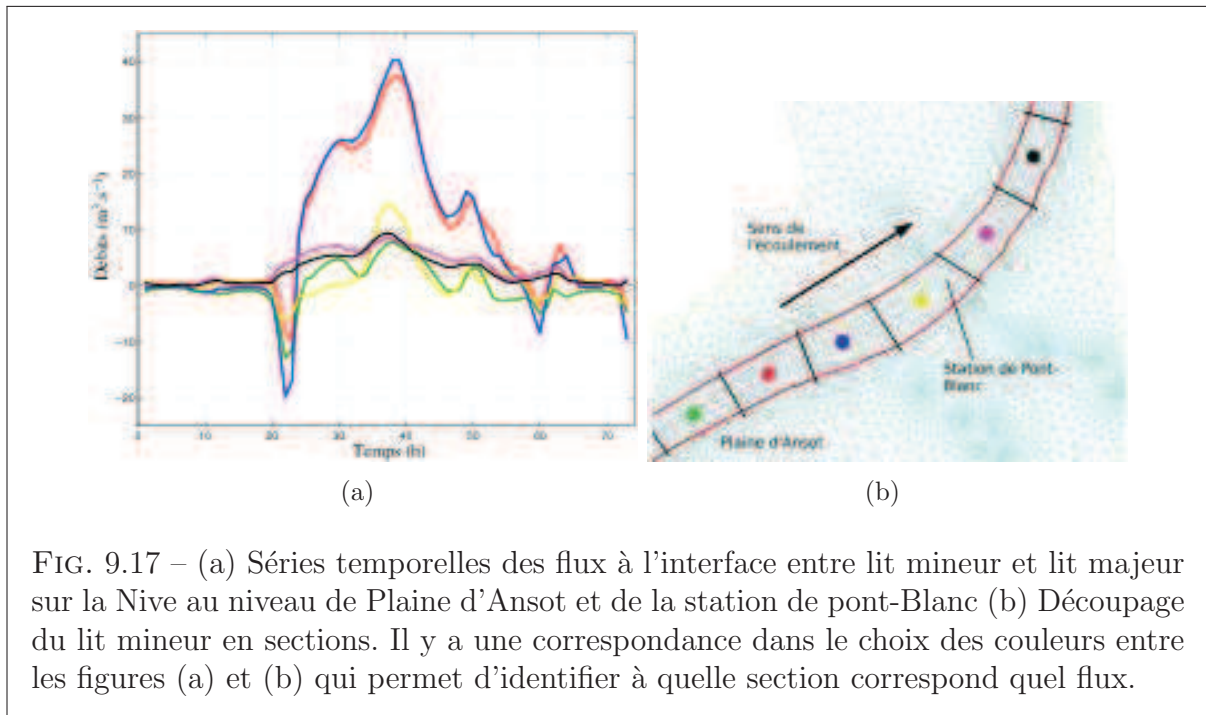


FIG. 9.16 – (a) Séries temporelles des flux à l'interface entre lit mineur et lit majeur en amont de la Nive (b) Découpage du lit mineur en sections. Il y a une correspondance dans le choix des couleurs entre les figures (a) et (b) qui permet d'identifier à quelle section correspond quel flux.

On a reporté en figure 9.17 les débits à l'interface entre lit mineur et lit majeur plus en aval sur la Nive au niveau de la Plaine d'Ansot et de la station de Pont-Blanc. On note que dans cette partie du modèle 2D les flux sont plutôt orientés dans le sens lit majeur/lit mineur, l'eau qui était sortie du lit plus en amont (et notamment après le passage de l'A63) et l'eau passée sous le pont de l'autoroute viennent se jeter dans la Nive principalement au niveau des sections identifiées par les pastilles rouge et bleu. Comparativement aux débits simulés dans le lit de la rivière ceux-ci sont relativement importants puisqu'en ce qui concerne les sections rouge et bleu la somme des débits sur ces deux sections lors du pic de crue correspond à environ 15-20 % des débits calculés dans le lit mineur.

Pour ce qui est des autres sections colorées représentées sur la figure 9.17-(b) on note qu'en valeur absolue les débits correspondants sont globalement moins importants que ceux calculés sur les sections rouge et bleu. Cela est particulièrement vrai pour la section verte, ce qui signifie que la plus grande partie de l'eau qui a débordé dans la Plaine d'Ansot se jette à nouveau dans la Nive au niveau des sections rouge et bleu. Cela est confirmé par l'étude des séries temporelles des flux entre lit mineur et lit majeur sur les autres sections et que nous n'avons pas reportées ici.



Enfin il faut souligner que sur cet épisode les flux entre lits mineur et majeur sur l'Adour (au niveau de Bayonne) sont quasi-nulles, tout au plus 1 à 2 $m^3.s^{-1}$ ce qui est négligeable par rapport aux débits calculés dans le lit de l'Adour sur cet épisode.

D'autre part lors de ce pic de crue de la Nive le coefficient de marée a été relativement faible et les échanges entre le lit de la Nive, la Plaine d'Ansot et la Barthes de Quartier Bas sont principalement dûs à l'augmentation des débits de la Nive. Or cet événement est suivi d'une augmentation importante du coefficient de marée comme on peut le voir sur la figure C.8-(b) donnant lieu à des débordements parfois plus importants et se produisant de manière cyclique comme on peut le voir sur la figure 9.18 qui représente les flux lit mineur/lit majeur sur la zone illustrée en figure 9.17-(b). De la même manière plus en amont sur la Nive on peut observer des échanges comparables entre la Nive et la Barthes de Quartier Bas. Il faut également souligner que les échanges illustrés précédemment sont potentiellement amplifiés en cas de fortes crues.

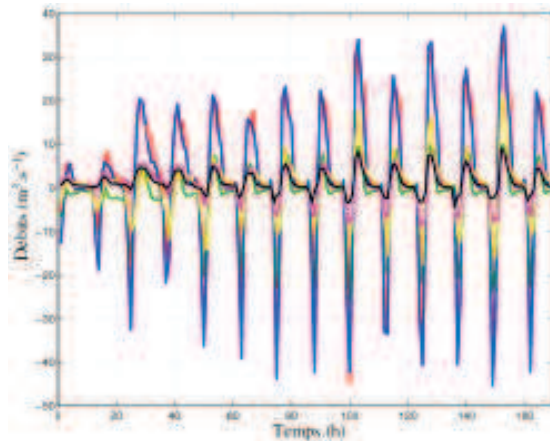


FIG. 9.18 – Série temporelle des flux à l'interface entre lit mineur et lit majeur sur la même zone celle présentée en figure 9.17 lorsque le forçage maritime aval est prépondérant.

Enfin il faut comparer le temps de calcul de cette méthode de couplage sur le pic de crue de la Nive avec la méthode de couplage présentée en section 9.2. Le temps de calcul est nettement inférieur à celui du couplage à interfaces et les résultats en termes de hauteurs d'eau du modèle 2D aux stations de Pont-Blanc et de Lesseps sont identiques. Cette méthode présente en plus l'avantage de pouvoir être combinée à l'assimilation des observations aux stations du réseau Adour maritime (en particulier à celle de Pont-Blanc). Néanmoins comme on l'a vu la continuité entre les modèles 1D et 2D aux limites de ce dernier est grossière.

Modèle	1D	2D	Couplé à interfaces	Couplé superposés
Temps de calcul	84	3430	39396	7255

TAB. 9.2 – Temps de calcul (s.) des modèles 1D, 2D, couplé à interfaces et couplé superposé sur la crue de Janvier 2014.

9.4 Combinaison de l'assimilation de données ensembliste et du couplage de modèles superposés

Comme on l'a vu sur la figure 9.15 à la station de Pont-Blanc lors du pic de crue le modèle couplé sous-estime d'environ une trentaine de centimètres les hauteurs d'eau observées. Pour réduire cette différence on peut avoir recours à l'assimilation de données avec le modèle 1D de manière à corriger la ligne d'eau de ce modèle et par la suite corriger les hauteurs d'eau modélisées par le modèle 2D. En effet étant donné la grande extension spatiale des fonctions de covariance l'assimilation des observations de Pont-Blanc permet

de corriger les forçages amonts du modèle 2D sur la Nive et par suite d'améliorer les résultats à cette station dans la modélisation 2D.

Dans cette section nous nous pencherons donc sur la question de la combinaison du couplage de modèles superposés et de l'assimilation de données ensembliste que nous désignerons dorénavant dans un souci de simplicité par l'expression « combinaison couplage-assim ». Dans un premier temps nous présenterons dans les grandes lignes l'implémentation de cette méthode avant d'exposer quelques résultats numériques obtenus sur l'épisode de 2014 sur lequel on a déjà travaillé précédemment.

9.4.1 Implémentation

L'implémentation de l'assimilation de données ensembliste et l'implémentation du couplage de modèles superposés ont été faites chacune avec une maquette PALM utilisant parfois les mêmes sous-routines fortran 90 mais compilées avec des clés de pré-compilation différentes. C'est pourquoi il n'a pas été possible de fondre ces deux maquettes en une seule maquette PALM qui aurait procédé à la fois à l'assimilation ensembliste et au couplage. Une telle implémentation est néanmoins envisageable mais n'a pas pu être réalisée dans le temps imparti pour ce travail de thèse.

Pour implémenter la combinaison couplage-assim on a donc utilisé les maquettes existantes que l'on lance là encore de manière itérative à travers un pilote. A chaque itération la maquette de couplage permet de calculer les termes sources pour le 1D si le modèle 2D est en crue. Ces termes sources sont sauvegardés dans des fichiers et sont ensuite lus par la maquette d'assimilation et utilisés comme termes d'apports latéraux pour les membres lors de l'assimilation ensembliste. Après l'étape d'analyse la moyenne des membres est sauvegardée dans un fichier qui sera lu comme ligne d'eau initiale par le modèle 1D par la maquette PALM de couplage.

Cette démarche permet ainsi de corriger la ligne d'eau du modèle 1D utilisée pour le couplage et par suite de corriger les forçages amonts sur la Nive. Néanmoins elle présente deux inconvénients importants que nous nous devons de souligner :

- le terme source est le même pour chaque membre de l'ensemble, aussi la relation entre cote et terme source peut ne pas être consistante. En effet le terme source pourra être négatif sur certaines zones (et ainsi le membre se voit retirer des débits et de la quantité de mouvement) alors même que le membre en question n'est pas débordant sur ces zones. Idéalement pour circonvier à cet inconvénient il aurait fallu utiliser un filtre de Kalman d'ensemble de modèles couplés ;
- la relation entre les cotes de la ligne d'eau calculée comme la moyenne des membres après assimilation et le terme source pour le modèle 1D peut ne pas être consistante. En effet l'assimilation va tendre à gommer le biais entre état analysé moyen et pas de biais négatif par exemple le terme source sera donc sous-estimé par rapport à ce qu'il devrait être. Néanmoins nous pensons que ce problème a tendance à s'équilibrer à travers les itérations de couplages et les résultats numériques présentés en section.

9.4.2 Résultats numériques

Dans cette section nous nous intéressons à l'application sur l'épisode de 2014 de la combinaison couplage-assim. Pour cela nous procédons à l'application itérative de la

combinaison « couplage-assim » suivie d’une période de prévision de 6 heures. 6 heures étant le temps maximum pour qu’un signal imposé à Cambo (condition limite amont sur la Nive) soit propagé jusqu’à Bayonne. En règle générale ce temps de transfert est compris entre 3 et 6 heures, cela dépend de l’intensité de la crue : plus celle-ci est importante, plus il y a de débordements et plus il y a de frottement s’appliquant sur l’onde de crue et ralentissant celle-ci en conséquence. D’autre part il faut préciser que dans ce cas les prévisions sont faites en mode déterministe : la ligne d’eau 1D calculée comme la moyenne des membres lors de l’étape d’assimilation est propagée dans le temps tout en étant couplée au modèle 2D. Enfin nous rappelons qu’en mode prévision sur l’Adour, à défaut de prévisions sur les conditions limites amonts celles-ci sont fixées constantes égales aux derniers débits imposés avant prévision.

Avant de présenter les résultats en termes d’amélioration des hauteurs d’eau calculées à la station de Pont-Blanc nous établissons d’abord l’impact de l’assimilation sur les débits au niveau de la condition limite amont de la Nive. Comme nous pouvons le voir sur la figure 9.19 l’assimilation de données à la station de Pont-Blanc a pour influence d’augmenter globalement les débits à l’amont de la Nive sur la durée de l’épisode avec notamment une décroissance des débits postérieurement au pic de crue plus lente.

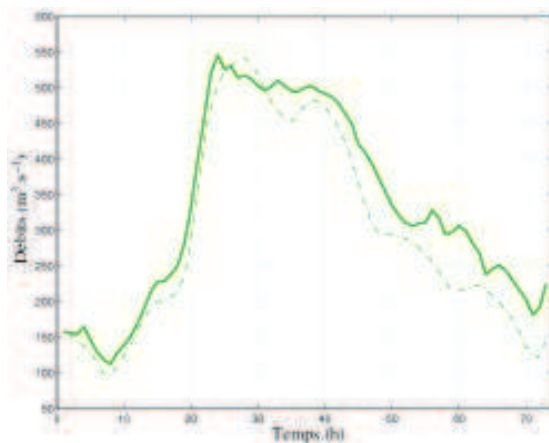
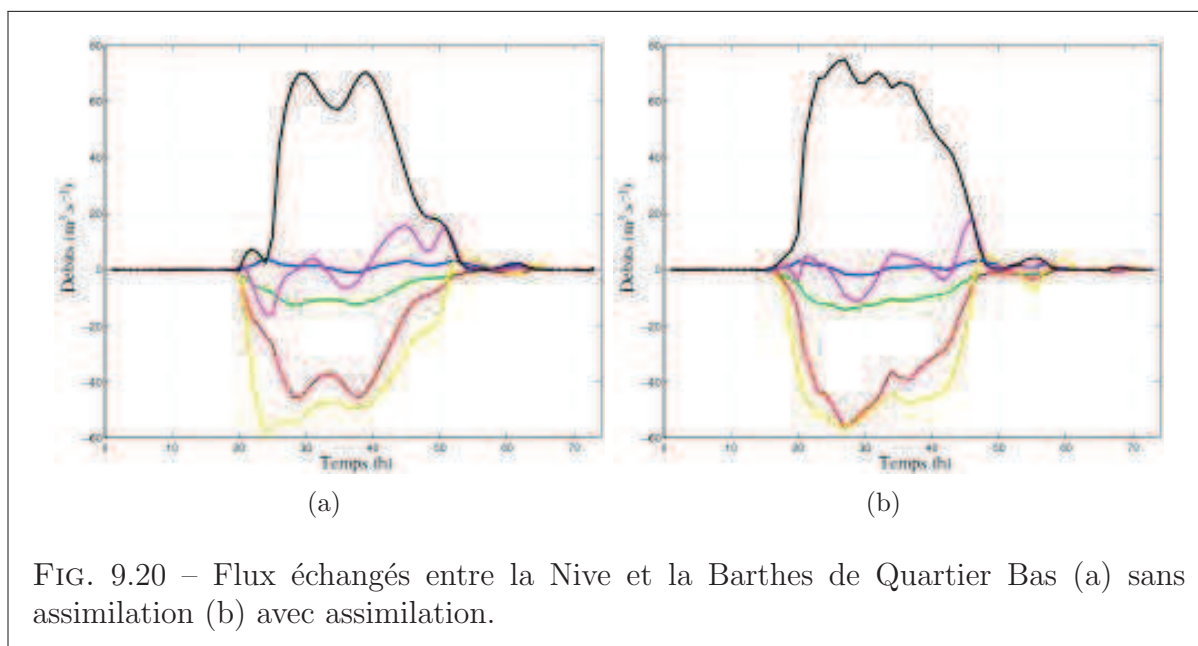


FIG. 9.19 – Débits imposés à l’amont du modèle 2D sur la Nive, sans assimilation (ligne verte tiretée), avec assimilation (ligne verte continue).

Cette augmentation des débits à l’amont de la Nive se traduit par une très légère modification des flux à l’interface entre lit mineur et lit majeur comme on peut le voir sur la figure 9.20 où on a reporté les débits échangés entre la Nive et la Barthes de Qaurtier Bas avec (figure 9.20-(b)) et sans assimilation (figure 9.20-(a)).



Nous présentons maintenant en figure 9.21 les hauteurs d'eau calculées à la station de Pont-Blanc à différentes échéances de prévision et nous avons synthétisé dans le tableau 9.3 l'évolution pendant la prévision à cette même station de la RMSE, de la corrélation et du biais.

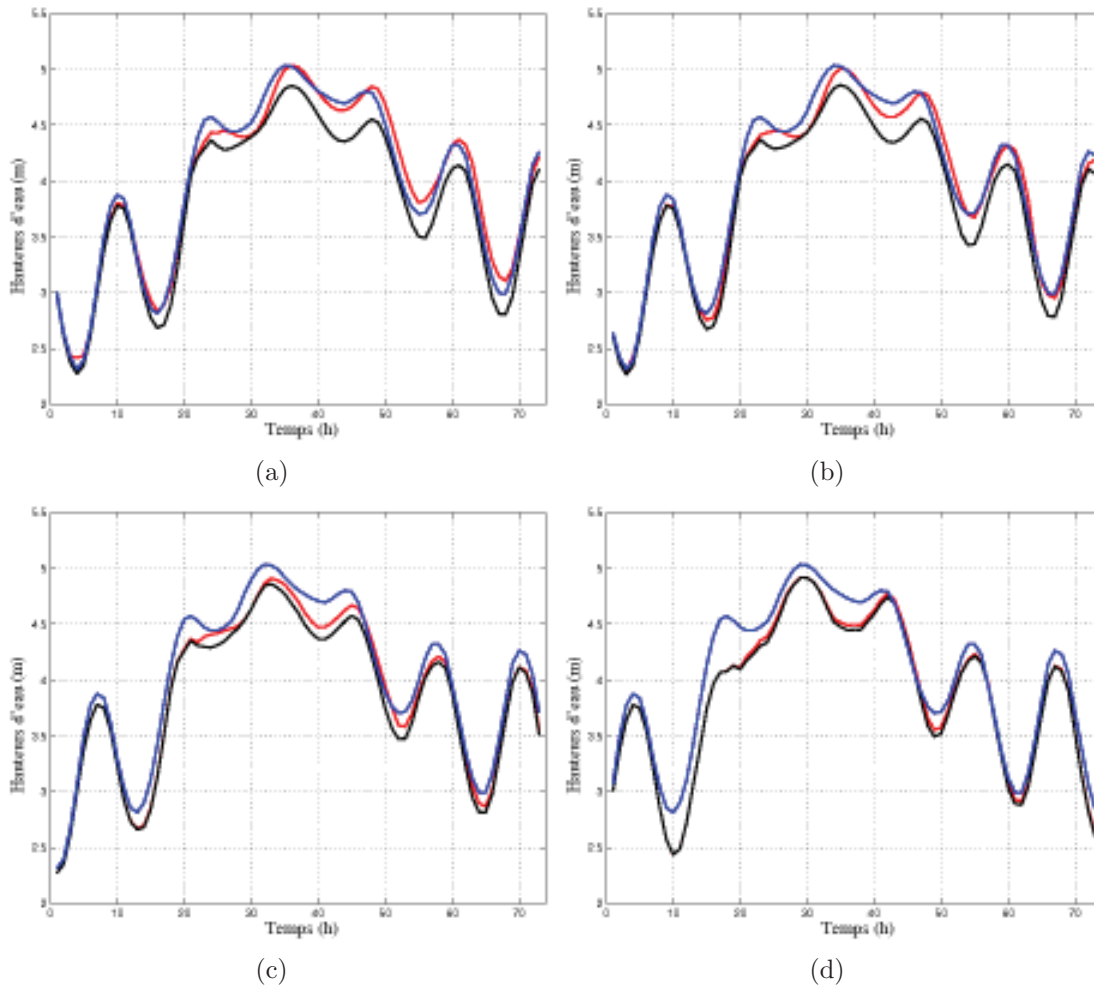


FIG. 9.21 – Observations (courbe bleue) et hauteurs d'eau calculées avec le modèle 2D avec assimilation (courbe rouge) et sans assimilation (courbe noire) à la station de Pont-Blanc (a) après assimilation (b) à 1 heure de prévision (c) à 3 heures de prévision (d) à 6 heures de prévision.

En mode prévision la correction apportée par l'assimilation est propagée à l'aval et donc tout au long de la prévision la différence entre le modèle couplé avec assimilation et le modèle couplé sans assimilation (courbes rouge et noire respectivement sur la figure 9.21) se réduit. Cela se traduit en termes de scores et si l'on compare la différence relative entre le modèle couplé avec assimilation et le modèle couplé sans assimilation pour la RMSE, la corrélation et le biais alors celle-ci (hormis pour la corrélation) décroît au fil de la prévision pour se rapprocher de 0 comme on peut le voir dans le tableau 9.3.

Echéance	RMSE	Corrélation	Biais
0 h	44.9	0.46	104.4
1 h	46.4	0.28	72.7
2 h	33.6	0.23	47.1
3 h	23.1	0.23	31.2
4 h	15.5	0.22	21.7
5 h	9.9	0.17	15.5
6 h	5.7	0.15	10.9

TAB. 9.3 – Différence relative (exprimée en pourcentage) des hauteurs d'eau entre les modèles couplés avec et sans assimilation pour la RMSE, la corrélation et le biais à la station de Pont-Blanc en fonction de l'échéance de prévision.

Conclusions et perspectives

Conclusions

La modélisation numérique est un outil incontournable de la prévision des crues en temps réel. Or les modèles étant par nature imparfaits, l'assimilation de données est devenue pour les prévisionnistes un outil complémentaire à l'utilisation des modèles numériques. En effet l'assimilation de données, en combinant l'information apportée par les observations à celle apportée par le modèle, permet dans une certaine mesure de parer aux lacunes de ces derniers et donc d'améliorer les prévisions à plus ou moins long terme.

D'autre part la multiplicité éventuelle des modèles numériques de dimensions différentes développés sur un même réseau hydraulique nécessite de développer des méthodes de couplage permettant d'assurer une certaine continuité physique entre ces modèles.

Ce travail de thèse a porté sur ces deux aspects de la modélisation et de la prévision des crues et leur application au réseau hydraulique Adour maritime.

Tout d'abord, nous avons montré que l'on pouvait émuler le filtre de Kalman d'ensemble sur un modèle d'onde de crue par un filtre de Kalman invariant mais pour un coût de calcul nettement moins important avec une fonction de covariance au point d'observation asymétrique dont l'extension spatiale à l'aval de ce point est plus faible que celle à l'amont. Cela a un intérêt pour l'application opérationnelle car, au contraire de l'EnKF, elle est compatible avec l'implémentation de la chaîne opérationnelle. Celle-ci n'est actuellement pas prévue pour gérer des méthodes ensemblistes.

L'étude des fonctions de covariance sur le réseau Adour maritime montre que celles-ci ont une très grande extension spatiale et que leur forme est liée à la géométrie du réseau.

L'application de l'EnKF au réseau hydraulique Adour maritime a montré que pour que cet algorithme soit efficient il devait être complété par une stratégie d'estimation des variances d'erreur d'observation, ainsi que par une méthode d'inflation spatialisée de manière à remédier à la sous-dispersivité de l'ensemble de façon consistante avec la physique et la géométrie du modèle. Ainsi on rappelle que le pourcentage d'amélioration de la RMSE de l'EnKF par rapport au modèle seul sur une expérience jumelle menée ici est de l'ordre de 69%, tandis que celui de l'IEnKF est de l'ordre de 95%. Cette étude a été complétée par une estimation de la validité des prévisions d'ensemble sur ce réseau. Celle-ci a montré que l'établissement d'un système de prévision probabiliste sur l'Adour était fortement lié à la connaissance que nous avons des forçages, et que dans l'état actuel des choses celle-ci était trop lacunaire. Pour étendre le temps de prévision il faut étendre le vecteur de contrôle aux forçages et aux paramètres hydrauliques.

D'autre part, sur l'Adour, deux modèles hydrauliques de dimensions différentes ont été développés : un modèle 1D appelé « Adour maritime » couvrant, comme son nom

l'indique, la portion maritime de l'Adour, et un modèle 2D local sur la zone de Bayonne modélisant la confluence Nive/Adour dans le centre-ville et certaines plaines d'inondation adjacentes à la Nive en amont de Bayonne et dans lesquelles l'écoulement n'est plus mono-dimensionnel. Le développement d'un modèle 2D local plutôt que global est justifié à la fois par la nature de l'écoulement dans cette zone, par le fait que sur de nombreuses portions l'écoulement est mono-dimensionnel et qu'un modèle 2D global serait de toute façon bien trop coûteux en temps de calcul.

Pour assurer la continuité relative des variables d'état entre ces modèles, deux méthodes de couplage ont été implémentées durant cette thèse. Une première dans laquelle le modèle 1D a été décomposé en plusieurs sous-modèles 1D, et le couplage est assuré par l'échange des variables d'état à l'interface entre les modèles. Cette méthode permet d'assurer la continuité des variables d'état aux interfaces entre les modèles mais son implémentation actuelle en fait une méthode relativement coûteuse en termes de temps de calcul : environ 12 heures de calcul pour 3 jours simulés.

Dans une seconde méthode, le modèle 1D est conservé d'un seul tenant et les deux modèles sont superposés dans la zone de recouvrement. Dans cette configuration, le modèle 1D force le modèle 2D à ses limites. En retour, lorsque celui-ci est en crue, il calcule les termes sources à fournir au modèle 1D pour les équations de conservation de la masse et de la quantité de mouvement, modélisant ainsi les échanges entre lit mineur et lit majeur dans le modèle 1D alors même que celui-ci ne dispose pas lit majeur. Il est à noter que ces échanges en termes de débits peuvent être importants et de l'ordre de 10 à 15% des débits calculés dans le lit de la rivière. Cette méthode ne permet d'assurer la continuité des variables que de manière très relative puisque après application de l'algorithme de couplage une différence de hauteurs d'eau entre les modèles de l'ordre de 20 à 25 *cm* dans les cas les plus défavorables peut subsister. Ce qui, ramené à la profondeur d'eau locale, correspond à une différence relative d'environ 3 %. Ici, le coût de calcul est nettement moindre que dans le cas précédent puisqu'il est d'environ 2 heures de calcul pour 3 jours simulés.

Le fait que dans cette dernière méthode le modèle 1D ne soit pas décomposé en plusieurs sous-modèles nous permet de combiner le couplage de modèles superposés et les méthodes d'assimilation ensemblistes développées sur l'Adour. Cela nous permet, sur la zone de recouvrement, d'améliorer la modélisation de l'état hydraulique du modèle 2D et les prévisions à court terme sur la zone de recouvrement. En effet, l'assimilation améliore les hauteurs d'eau modélisées à la station de Pont-Blanc après assimilation de l'ordre de 25 à 30 *cm*. En mode prévision, les résultats se dégradent avec l'échéance de prévision, néanmoins le pourcentage d'amélioration de la RMSE du modèle couplé avec assimilation par rapport au modèle couplé sans assimilation à l'échéance de 3 heures est par exemple de 23.1 %, et de 15.5 % à l'échéance de 4 heures.

Perspectives

A l'issue de ce travail de thèse, on retiendra naturellement deux grands axes pour les perspectives : l'un portant sur la modélisation du réseau Adour maritime et l'assimilation de données, l'autre portant sur le couplage de modèles hydrauliques 1D-2D.

La modélisation du réseau Adour maritime et l'assimilation de données

On a pu constater tout au long de ce travail à la fois l'apport de l'assimilation de données et en même temps ses limitations. En effet, on a vu sur des expériences utilisant des observations synthétiques que l'assimilation de données permet d'améliorer les résultats du modèle : que ce soit pour la hauteur d'eau qui est la variable d'état observée et assimilée, mais aussi pour les débits à travers les covariances entre hauteurs d'eau et débits. Aussi, pour juger de la pertinence de la méthode d'assimilation de données utilisée ici, et plus particulièrement des covariances hauteurs d'eau/débits diagnostiquées, il faudrait disposer d'observations réelles de débits à d'autres stations que celles où on procède à l'assimilation, notamment pendant les épisodes de crue. Or ni le SCHAPI ni le SPC GAD ne disposent de telles observations, même sur des épisodes très courts.

D'autre part nous avons montré que l'EnKF combiné avec des méthodes d'inflation permet d'avoir une meilleure estimation de l'état hydraulique non seulement au point d'observation mais aussi ailleurs sur le domaine mais dans la limite du modèle « Adour maritime ». En effet, dans l'état actuel des choses celui-ci ne dispose pas de lit majeur sauf en quelques sections en travers. Nous supposons en effet (mais nous n'avons pas pu l'établir clairement) que cela est à l'origine de la sur-évaluation des hauteurs d'eau à la station de Urt lorsque nous assimilons les hauteurs d'eau à Peyrehorade sur des épisodes de fortes crues. L'ajout d'un lit majeur sur tout le réseau Adour maritime permettrait probablement d'améliorer la modélisation hydraulique et par suite les résultats de l'assimilation. Des travaux dans ce sens sont actuellement en cours au SCHAPI et au SPC GAD mais nous n'avons pas pu pendant cette thèse tester ce réseau amélioré.

D'autre part, pour améliorer les résultats en mode prévision il faudrait pouvoir disposer de débits prévus aux stations amonts ou alors étendre le réseau au-delà de ses frontières actuelles. Des travaux dans ce sens sont également en cours au SCHAPI et au SPC GAD pour étendre le réseau en amont du bief 6 jusqu'à la station de Artiguelouve sur le Gave de Pau et ainsi disposer des données d'observation de la station d'Orthez. Là encore nous n'avons pu pendant ce travail de thèse tester le réseau avec cette nouvelle configuration.

Le couplage de modèles hydrauliques 1D-2D

En ce qui concerne le couplage de modèles hydrauliques, l'un des premiers travaux à effectuer serait de remanier l'implémentation du couplage de modèles à interfaces. En effet l'implémentation de la version présentée dans ce travail est bien trop coûteuse en temps et ne présente guère d'intérêt pour la prévision des crues en temps réel. Elle peut néanmoins en présenter pour rejouer des événements après-coup.

La seconde méthode présentée, celle du couplage de modèles superposés, présente un coût de calcul nettement moindre mais n'assure que grossièrement la continuité des variables d'état entre les modèles. Il serait intéressant d'étudier en profondeur les raisons de cette différence et par suite les moyens d'y remédier.

Néanmoins des travaux récents portant un cas d'étude académique ont montré des résultats très prometteurs dans le cas d'une troisième méthode de couplage. Celle-ci consiste à modéliser l'écoulement dans le lit mineur avec un modèle 1D et l'écoulement

dans le lit majeur avec un ou plusieurs modèles 2D. Ainsi les modèles 1D et 2D ne sont pas superposés. Cette méthode présente trois intérêts importants :

- le modèle 2D ne résout pas les équations de Saint-Venant dans le lit mineur qui constitue la zone où en général le maillage est le plus dense, supposant donc une réduction importante du coût de calcul ;
- cette méthode ne nécessite pas de modifier la géométrie du modèle 1D ;
- la conservation du modèle 1D d'un seul tenant permet d'utiliser les méthodes d'assimilation de données développées sur ce modèle.

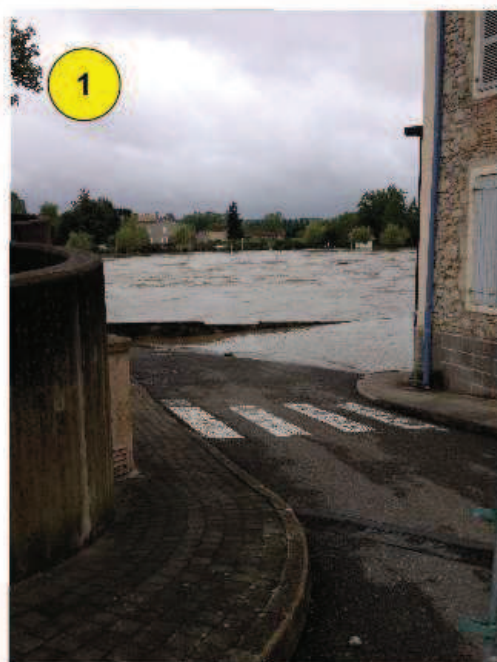
Cette dernière méthode de couplage devrait à terme être implémentée, avec l'appui d'EDF, sur la zone de Bayonne dans le cadre de la collaboration qui lie le CERFACS et le SCHAPI.

Annexe A

Inondations des Gaves Réunis à Peyrehorade du 19 Juin 2013



FIG. A.1 – Laisse de crue du 19 Juin 2013 à Peyrehorade.



(a)



(b)

FIG. A.2 – (a)-(b) Berge Nord de Peyrehorade.



(a)



(b)

FIG. A.3 – (a) Berge Sud de Peyrehorade (b) Berge Nord de Peyrehorade.



(a)



(b)

FIG. A.4 – (a) Route de Hastingue-Guiche, berge Sud de Peyrehorade (b) Route de Oeyregave, berge Sud de Peyrehorade.



(a)



(b)

FIG. A.5 – (a) Route d'Oloron, berge Sud de Peyrehorade (b) Route de Sorde l'Abbaye, berge Nord de Peyrehorade.



(a)



(b)

FIG. A.6 – (a) Gave d'Oloron à Sorde l'Abbaye (b) Gave d'Oloron à Sorde l'Abbaye.



FIG. A.7 – Route départementale de Sorde l'Abbaye.

Annexe B

Illustration de l'évolution temporelle des fonctions de corrélation ZZ sur le chemin 7-5-2-1

Nous présentons dans cette annexe un exemple d'évolution temporelle d'une fonction de corrélation ZZ le long du chemin 7-5-2-1. Tandis que l'extension spatiale de cette fonction sur ce chemin à l'amont reste relativement stationnaire son extension spatiale à l'aval est soumise à une variabilité nettement plus importante. En effet comme on peut le voir la fonction de corrélation est d'abord très étendue à l'aval puis elle a tendance à se « contracter », son extension spatiale à l'aval se réduit avant de croître brutalement (figure B.1)-(h).

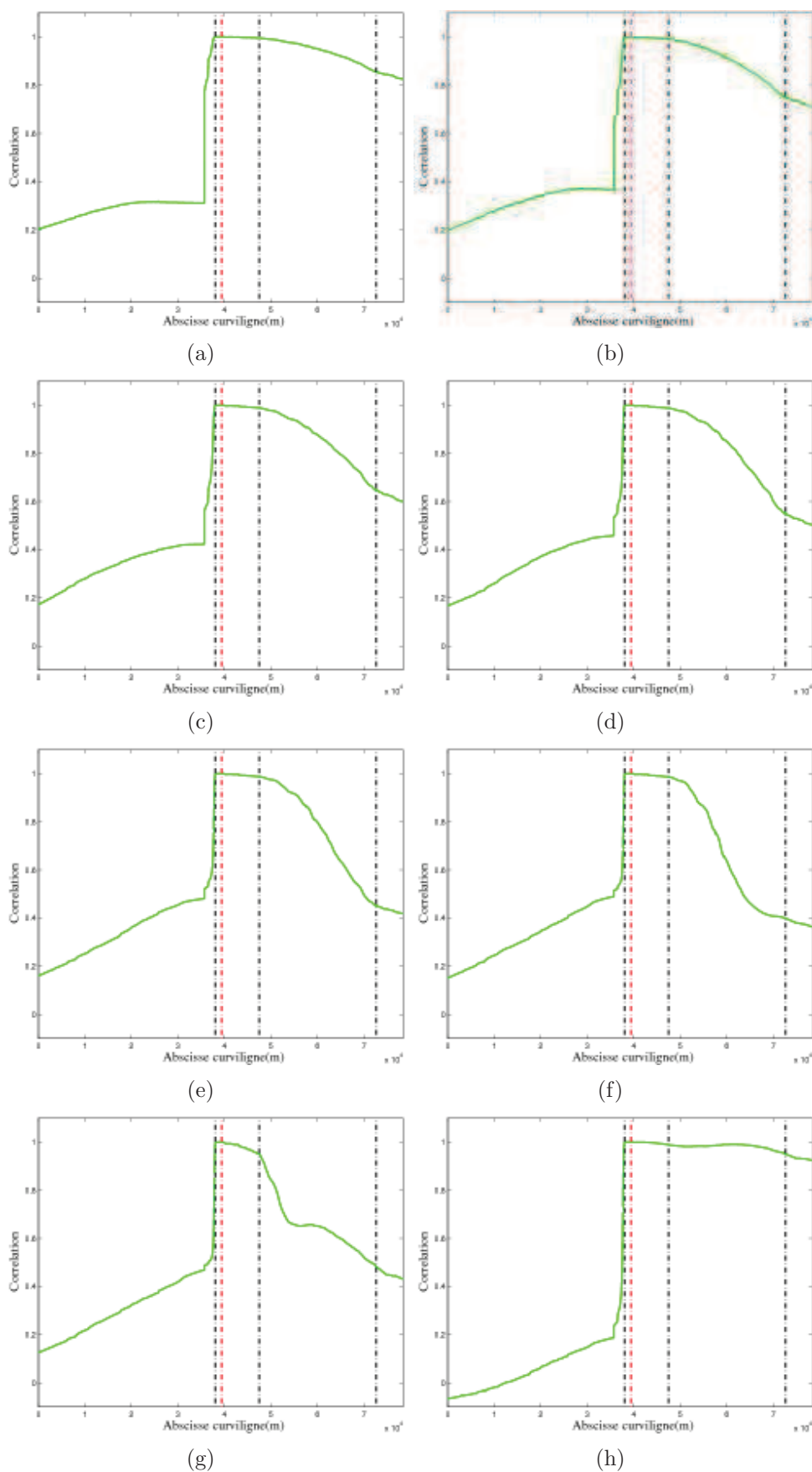
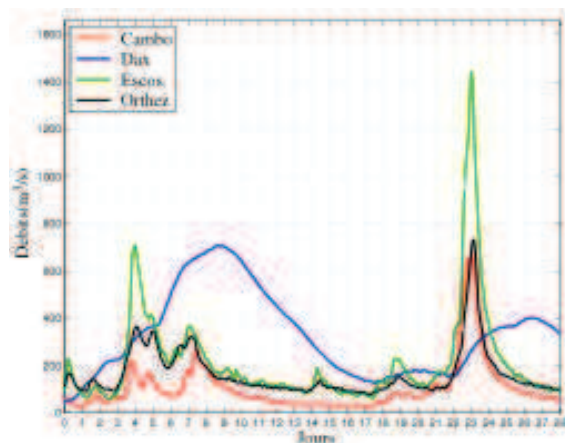


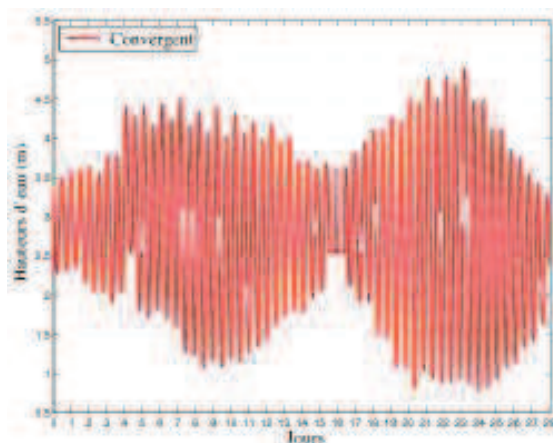
FIG. B.1 – Evolution temporelle de la fonction de corrélation ZZ relative à la station de Peyrehorade sur le chemin 7-5-2-1 du temps t au temps $t + 7h$.

Annexe C

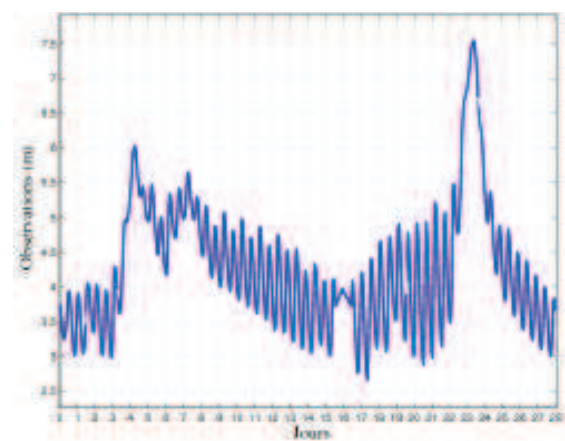
Données utilisées pour les expériences réelles



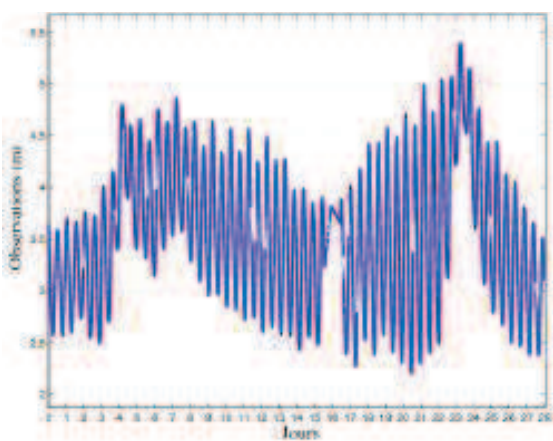
(a)



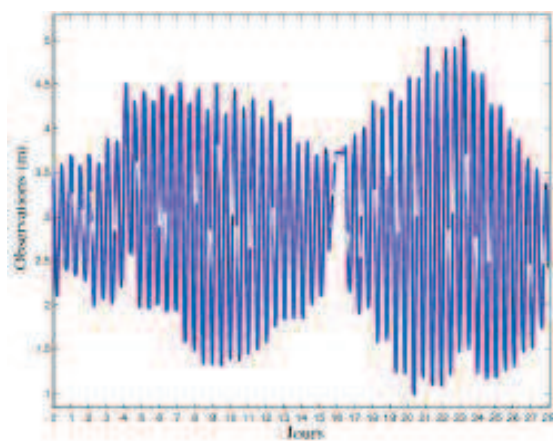
(b)



(c)



(d)



(e)

FIG. C.1 – Episode 2009-1. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps.

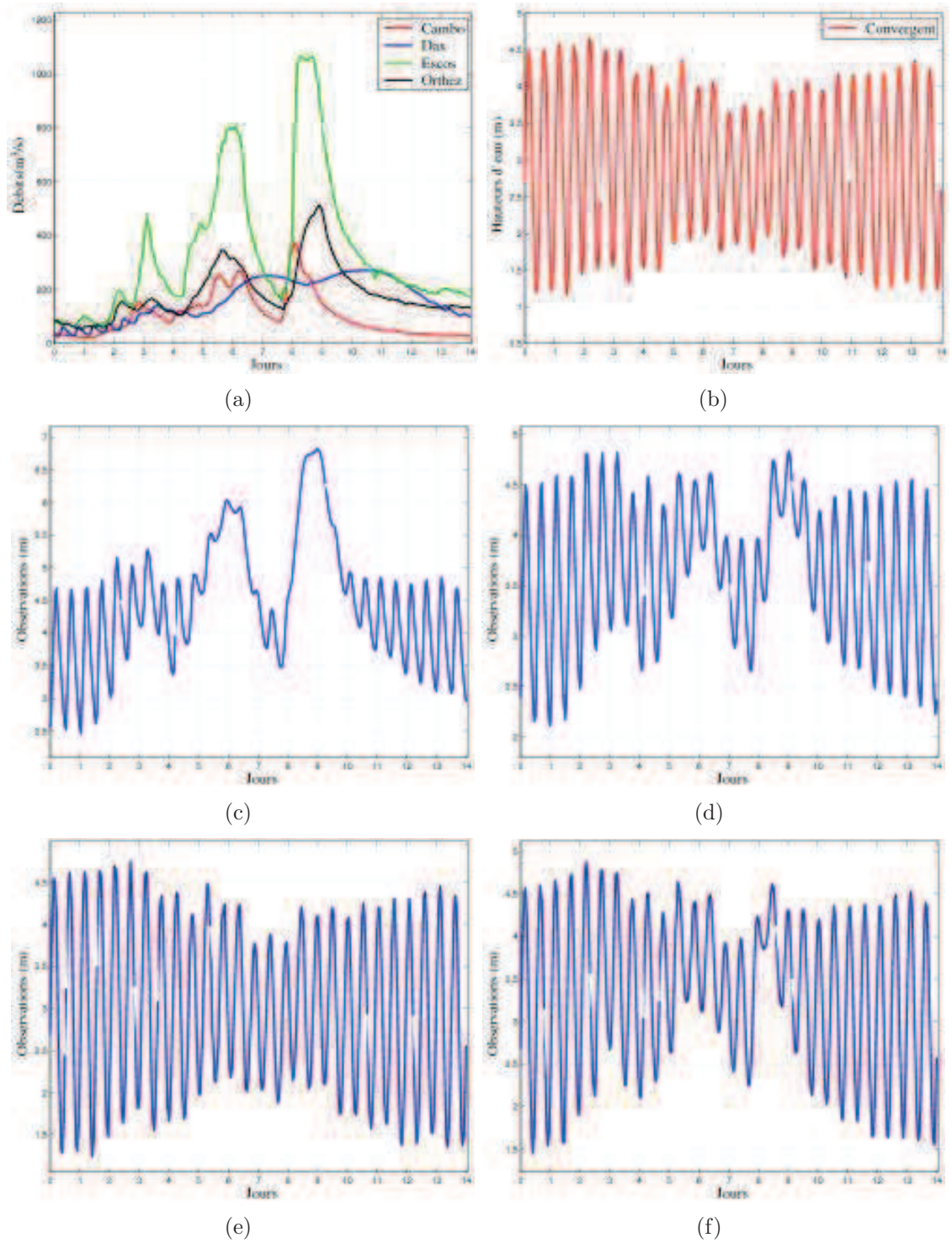


FIG. C.2 – Episode 2009-2. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.

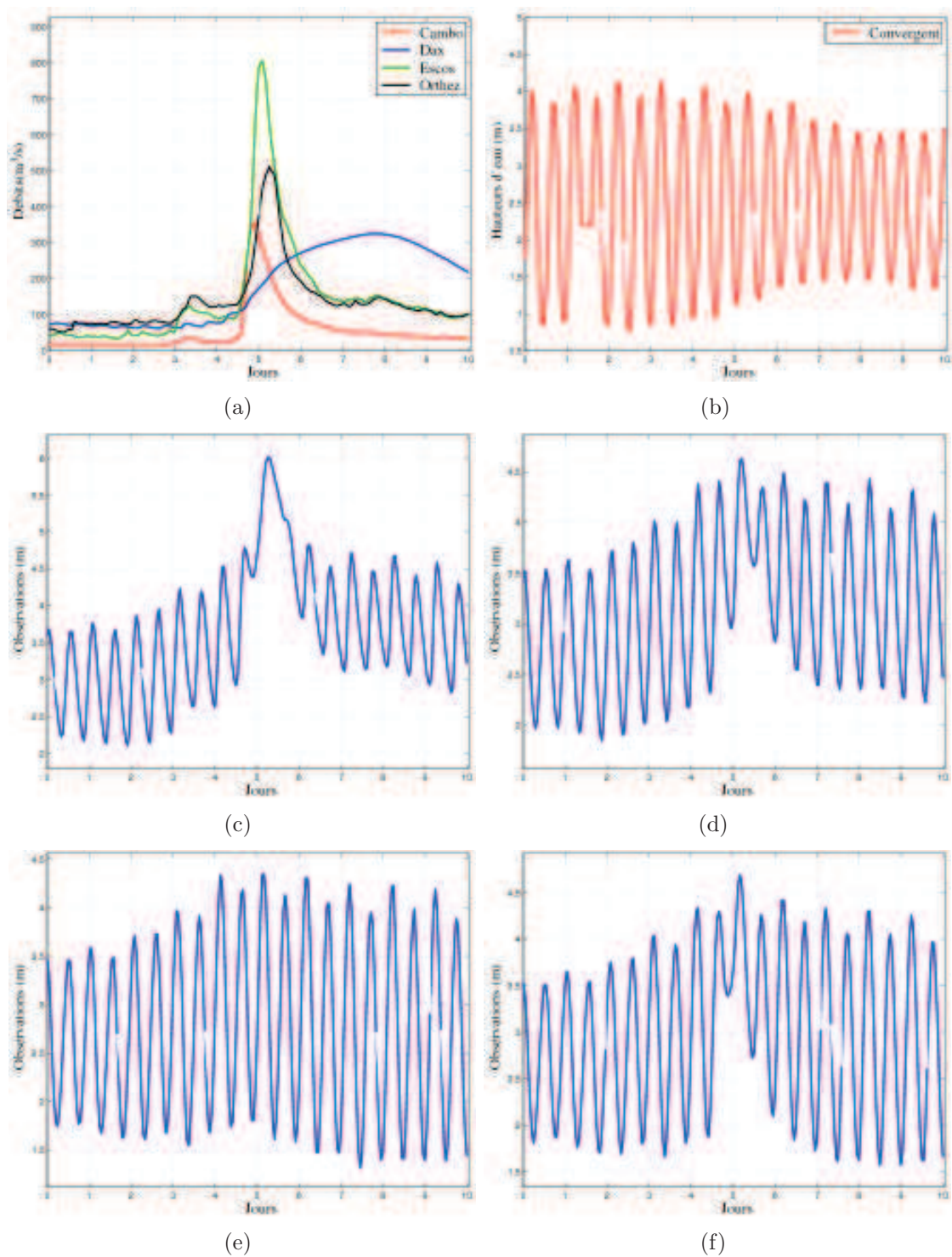


FIG. C.3 – Episode 2010. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.

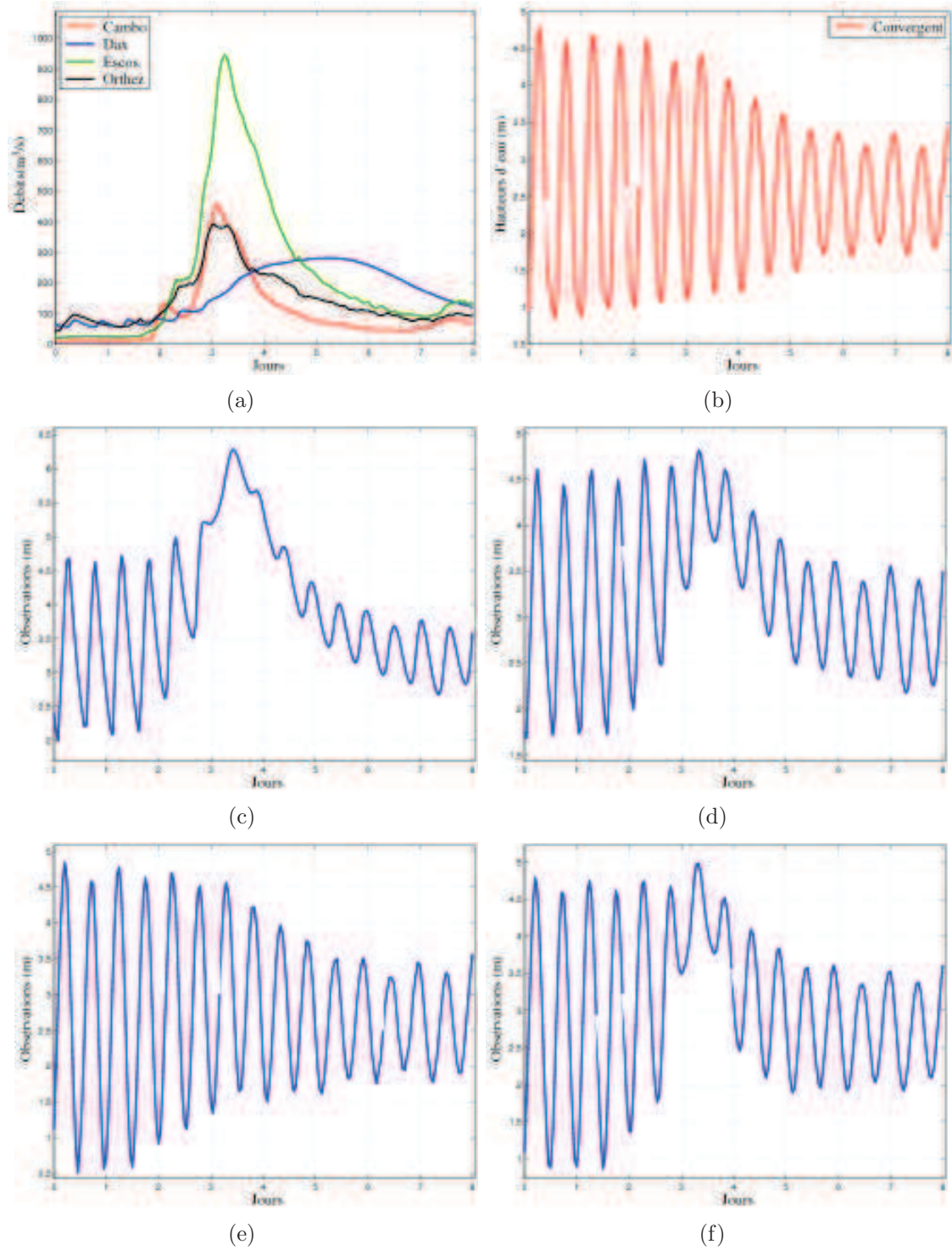


FIG. C.4 – Episode 2011-1. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.

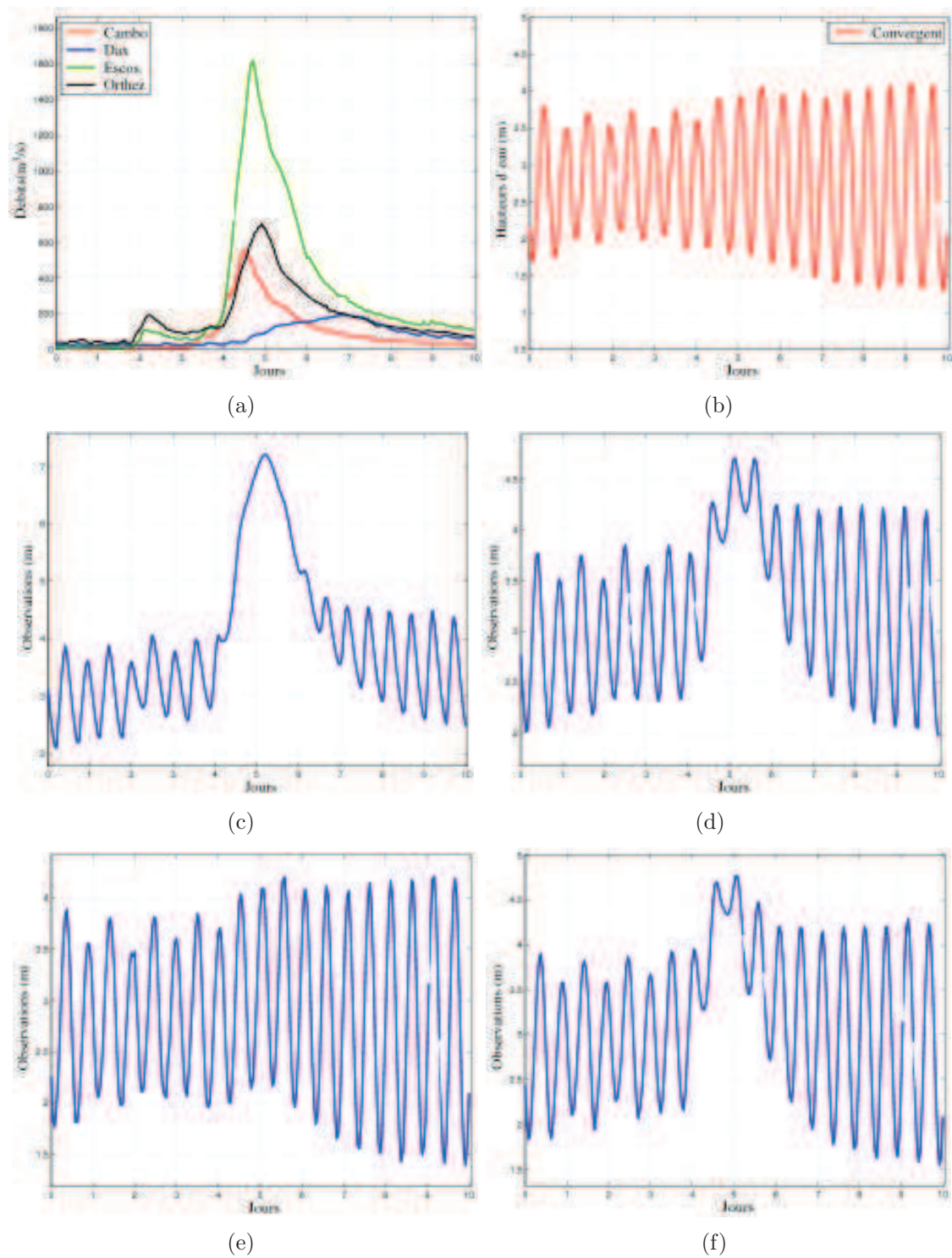


FIG. C.5 – Episode 2011-2. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.

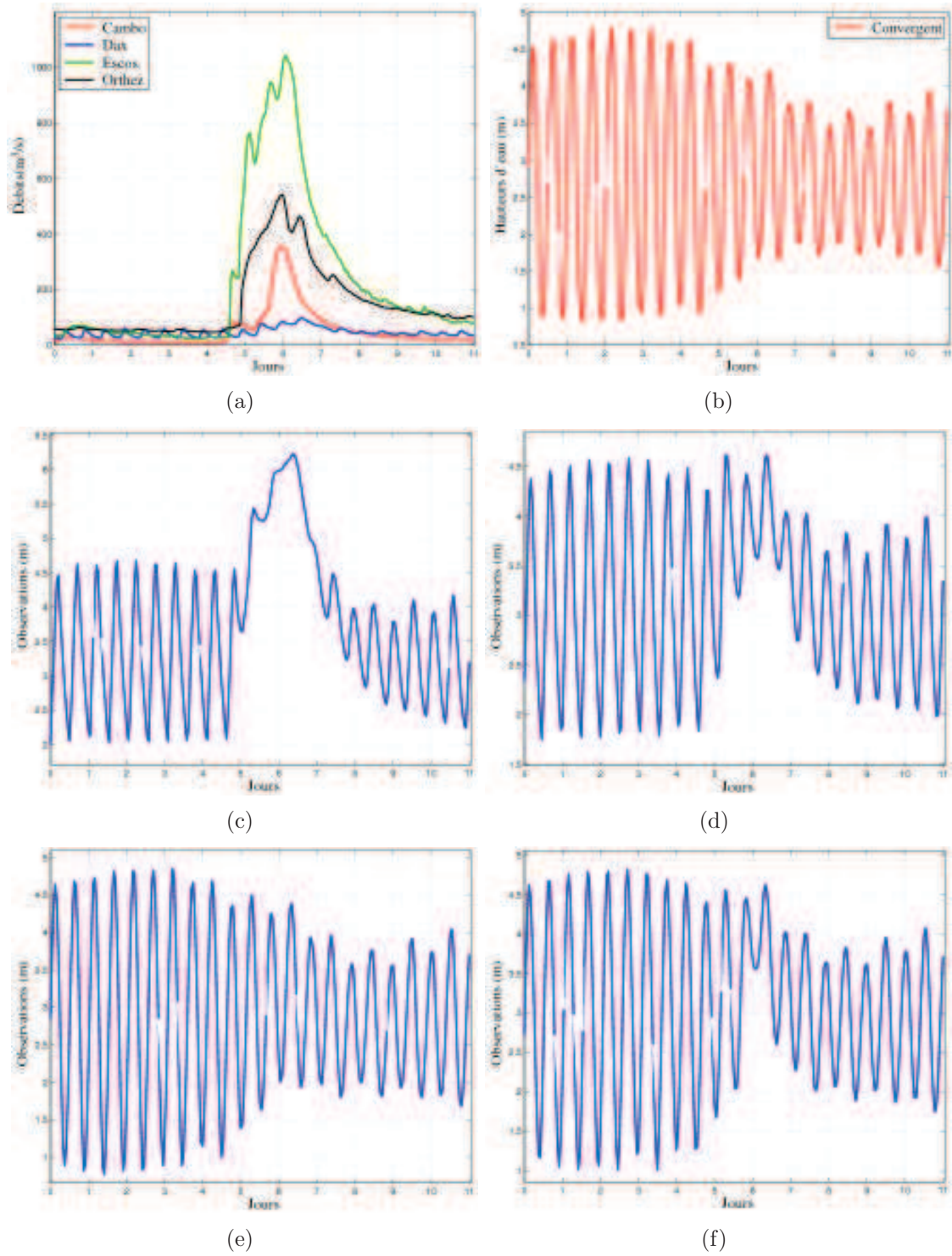


FIG. C.6 – Episode 2012. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.

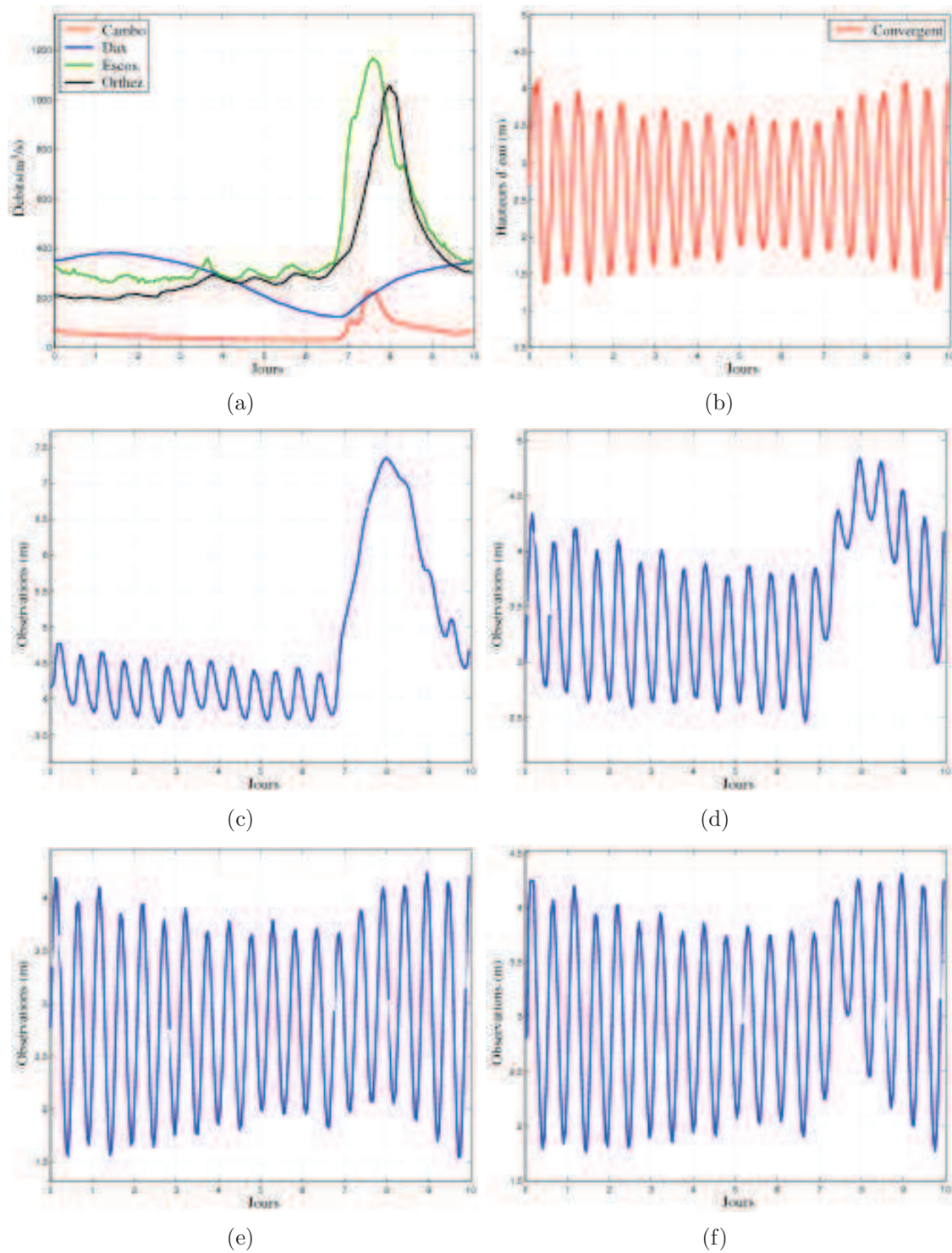


FIG. C.7 – Episode 2013. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.

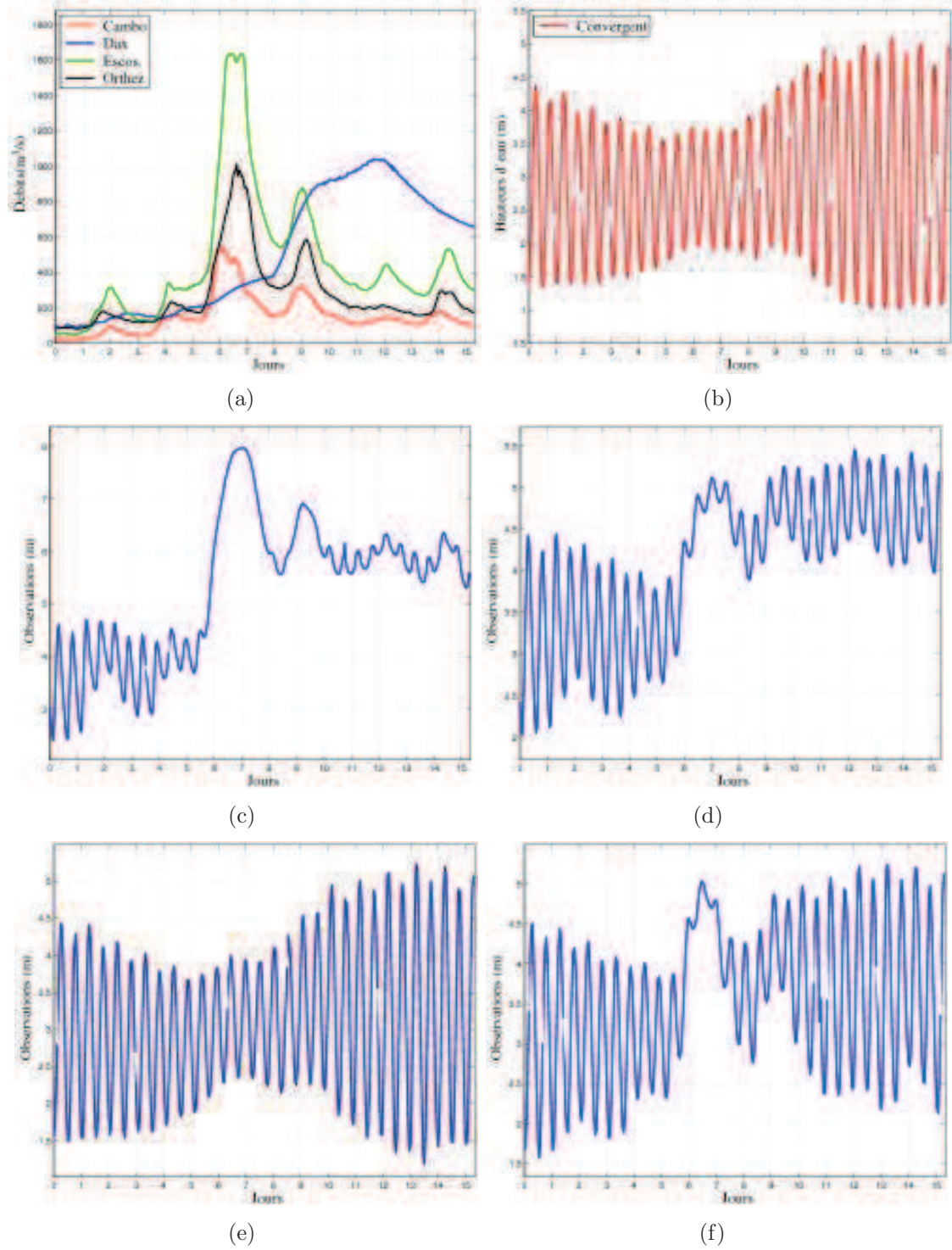


FIG. C.8 – Episode 2014. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.

Bibliographie

- [Anderson, 1995] Anderson, J. (1995). A Method for Producing and Evaluating Probabilistic Forecast from Ensemble Model Integration. *Journal of climate*, 9 :1518–1530.
- [Anderson, 2001] Anderson, J. L. (2001). An Ensemble Adjustment Kalman Filter for Data Assimilation. *Monthly Weather Review*, 129 :2884–2903.
- [Anderson, 2007] Anderson, J. L. (2007). An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus, Series A : Dynamic Meteorology and Oceanography*, 59 :210–224.
- [Anderson and Anderson, 1999] Anderson, J. L. and Anderson, S. L. (1999). A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts. *Monthly Weather Review*, 127 :2741–2758.
- [Andreadis et al., 2007] Andreadis, K. M., Clark, E. a., Lettenmaier, D. P., and Alsdorf, D. E. (2007). Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. *Geophysical Research Letters*, 34 :1–5.
- [Andrews, 1968] Andrews, A. (1968). A square root formulation of the Kalman covariance equations. *AIAA Journal*, 6(6) :1165–1166.
- [Bárdossy, 2006] Bárdossy, A. (2006). Calibration of hydrological model parameters for ungauged catchments. *Hydrology and Earth System Sciences*, 3 :1105–1124.
- [Bates and De Roo, 2000] Bates, P. D. and De Roo, a. P. J. (2000). A simple raster based model for flood inundation simulation. *Journal of Hydrology*, 236 :54–77.
- [Besnard and Goutal, 2011] Besnard, A. and Goutal, N. (2011). Comparaison de modèles 1D à casiers et 2D pour la modélisation hydraulique d’une plaine d’inondation—Cas de la Garonne entre Tonneins et La Réole. *La Houille Blanche*, (3) :42–47.
- [Bhattacharya et al., 2012] Bhattacharya, A. K., McEnroe, B. M., Zhao, H., Kumar, D., and Shinde, S. (2012). Modclark Model : Improvement and Application. *International Journal of Civil Engineering (IJCE)*, 2(7) :100–118.
- [Biancamaria et al., 2011] Biancamaria, S., Durand, M., Andreadis, K. M., Bates, P. D., Boone, a., Mognard, N. M., Rodríguez, E., Alsdorf, D. E., Lettenmaier, D. P., and Clark, E. a. (2011). Assimilation of virtual wide swath altimetry to improve Arctic river modeling. *Remote Sensing of Environment*, 115(2) :373–381.
- [Bishop et al., 2001] Bishop, C. H., Etherton, B. J., and Majumdar, S. J. (2001). Adaptive Sampling with the Ensemble Transform Kalman Filter. Part I : Theoretical Aspects. *Monthly Weather Review*, 129(3) :420–436.

- [Bocquet, 2011] Bocquet, M. (2011). Ensemble Kalman filtering without the intrinsic need for inflation. *Nonlinear Processes in Geophysics*, 18(5) :735–750.
- [Bonan, 2013] Bonan, B. (2013). *Assimilation de données pour l'initialisation et l'estimation de paramètres d'un modèle d'évolution de calotte polaire*. PhD thesis.
- [Borrell Estupina et al., 2014] Borrell Estupina, V. B., Malaterre, P.-O., Ricci, S., Fleury, P., Jay-Allemand, M., Coustau, M., Harader, E., Darras, T., Bouvier, C., Thual, O., Guilhalmenc, M., Marchandise, A., and Marechal, J.-C. (2014). Crues partie II : Genèse, propagation et prévision des crues à Montpellier Flood part II : Genesis, propagation and forecasting of flooding at Montpellier city. *Karstologia*.
- [Brier, 1950] Brier, G. W. (1950). Cerification of forecasts expressed in terms of probaility. *Monthly Weather Review*, 78(1) :1–3.
- [Buis et al., 2006] Buis, S., Piacentini, A., and Déclat, D. (2006). PALM : a computational framework for assembling high-performance computing applications. *Concurrency and Computation : Practice and experience*, 18(2) :231–245.
- [Burgers et al., 1998] Burgers, G., Jan van Leeuwen, P., and Evensen, G. (1998). Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review*, 126(6) :1719–1724.
- [Callado et al., 2013] Callado, A., Escribà, P., García-moya, J. A., Montero, J., Santos, C., Santos-mu noz, D., and Simarro, J. (2013). Ensemble Forecasting. *Climate Change and Regional/Local Responses*, pages 3–58.
- [Cohn, 1997] Cohn, S. E. (1997). An Introduction to Estimation Theory. *Sciences-New York*.
- [Coppola et al., 2014] Coppola, E., Verdecchia, M., Giorgi, F., Colaiuda, V., Tomassetti, B., and Lombardi, A. (2014). Changing hydrological conditions in the Po basin under global warming. *Science of the Total Environment*, 493 :1183–1196.
- [Corazza et al., 2003] Corazza, M., Kalnay, E., Patil, D. J., Yang, S.-C., Morss, R., Cai, M., Szunyogh, I., Hunt, B. R., and Yorke, J. a. (2003). Use of the breeding technique to estimate the structure of the analysis "errors of the day". *Nonlinear Processes in Geophysics*, 10 :233–243.
- [Cunge et al., 1980] Cunge, J. A., Holly, F. M., and Verwey, A. (1980). Practical aspects of computational river hydraulics.
- [Daley, 1991] Daley, R. (1991). Atmospheric data analysis, Cambridge atmospheric and space science series. *Cambridge University Press*, 6966 :25.
- [Decoene et al., 2009] Decoene, A., Bonaventura, L., Miglio, E., and Saleri, F. (2009). Asymptotic Derivation of the Section-Averaged Shallow Water Equations for Natural River Hydraulics. *Mathematical Models and Methods in Applied Sciences*, 19(3) :387–417.
- [Dee, 1995] Dee, D. P. (1995). Online Estimation of Error Covariance Parameters for Atmospheric Data Assimilation. *Monthly Weather Review*, 123 :1128–1145.
- [Desroziers et al., 2005] Desroziers, G., Berre, L., Chapnik, B., and Poli, P. (2005). Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society*, 131(2005) :3385–3396.

- [Doucet et al., 2000] Doucet, A., Godsill, S., and Andrieu, C. (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10 :197–208.
- [Dumas et al., 2013] Dumas, P., Hallegatte, S., Quintana-Seguí, P., and Martin, E. (2013). The influence of climate change on flood risks in France – first estimates and uncertainty analysis. *Natural Hazards and Earth System Science*, 13 :809–821.
- [Durand et al., 2008] Durand, M., Andreadis, K. M., Alsdorf, D. E., Lettenmaier, D. P., Moller, D., and Wilson, M. (2008). Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model. *Geophysical Research Letters*, 35 :1–5.
- [Edijatno and Michel, 1989] Edijatno and Michel, C. (1989). Un modèle pluie-débit journalier à trois paramètres. *Houille Blanche Revue Internationale De L'Eau*, 2 :113–121.
- [Evensen, 1994] 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.
- [Fernández-Nieto et al., 2010] Fernández-Nieto, E. D., Marin, J., and Monnier, J. (2010). Coupling superposed 1D and 2D shallow-water models : Source terms in finite volume schemes. *Computers and Fluids*, 39(6) :1070–1082.
- [Finaud-Guyot et al., 2011] Finaud-Guyot, P., Delenne, C., Guinot, V., and Llovel, C. (2011). 1D–2D coupling for river flow modeling. *Comptes Rendus Mécanique*, 339(4) :226–234.
- [Gander, 2006] Gander, M. J. (2006). Article Optimized Schwarz Methods Reference. *SIAM Journal On Numerical Analysis*, 44(2) :699–731.
- [Gaspari and Cohn, 1999] Gaspari, G. and Cohn, S. E. (1999). Construction of Correlation Functions in Two and Three Dimensions. Technical Report February 1996.
- [Gejadze and Monnier, 2007] Gejadze, I. Y. and Monnier, J. (2007). On a 2D 'zoom' for the 1D shallow water model : Coupling and data assimilation. *Computer Methods in Applied Mechanics and Engineering*, 196(March) :4628–4643.
- [Gelb, 1974] Gelb, A. (1974). *Applied optimal estimation*. MIT press.
- [Giering and Kaminski, 1998] Giering, R. and Kaminski, T. (1998). Recipes for adjoint code construction. *ACM Transactions on Mathematical Software*, 24(4) :437–474.
- [Giustarini et al., 2011] Giustarini, L., Matgen, P., Hostache, R., Montanari, M., Plaza, D., Pauwels, V. R. N., De Lannoy, G. J. M., De Keyser, R., Pfister, L., Hoffmann, L., and Savenije, H. H. G. (2011). Assimilating SAR-derived water level data into a hydraulic model : A case study. *Hydrology and Earth System Sciences*, 15 :2349–2365.
- [Goutal, 2014] Goutal, N. (2014). *Développements autour de la simulation des écoulements à surface libre en rivière. Equations de Saint-Venant - Couplage de modèles - Incertitudes - Application aux ondes de submersion*. PhD thesis.
- [Goutal and Maurel, 2002] Goutal, N. and Maurel, F. (2002). A finite volume solver for 1D shallow-water equations applied to an actual river. *International Journal for Numerical Methods in Fluids*, 38(1) :1–19.

- [Goutal et al., 2014] Goutal, N., Parisot, M., and Zaoui, F. (2014). A 2D reconstruction for the transverse coupling of shallow water models. *Numerical methods in fluids*, 75(11) :775–799.
- [Goutal and Sainte-Marie, 2011] Goutal, N. and Sainte-Marie, J. (2011). A kinetic interpretation of the section-averaged Saint-Venant system for natural river hydraulics. *International Journal for Numerical Methods in Fluids*, 67(7) :914–938.
- [Gouweleeuw et al., 2005] Gouweleeuw, B. T., Thielen, J., Franchello, G., De Roo, a. P. J., and Buizza, R. (2005). Flood forecasting using medium-range probabilistic weather prediction. *Hydrology and Earth System Sciences*, 9(4) :365–380.
- [Habets et al., 2008] Habets, F., Boone, A., Champeaux, J. L., Etchevers, P., Franchistéguy, L., Leblois, E., Ledoux, E., Le Moigne, P., Martin, E., Morel, S., Noilhan, J., Seguí, P. Q., Rousset-Regimbeau, F., and Viennot, P. (2008). The SAFRAN-ISBA-MODCOU hydrometeorological model applied over France. *Journal of Geophysical Research : Atmospheres*, 113(December 2004) :1–18.
- [Hamill et al., 2001] Hamill, T. M., Whitaker, J. S., and Snyder, C. (2001). Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter. *Monthly Weather Review*, 129 :2776–2790.
- [Harlim and Hunt, 2007] Harlim, J. and Hunt, B. R. (2007). A non-Gaussian ensemble filter for assimilating infrequent noisy observations. *Tellus, Series A : Dynamic Meteorology and Oceanography*, 59 :225–237.
- [Hervouet, 2007] Hervouet, J.-M. (2007). *Hydrodynamics of free surface flows : modelling with the finite element method*. John Wiley & Sons.
- [Houtekamer and Mitchell, 2001] Houtekamer, P. L. and Mitchell, H. L. (2001). A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review*, 129 :123–137.
- [Hunt et al., 2007] Hunt, B., Kostelich, E. J., and Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos : a local transform Kalman filter. *Physica D : Nonlinear Phenomena*, 230(1) :112–126.
- [Isabelle Mirouze, 2010] Isabelle Mirouze (2010). *Régularisation de problèmes inverses à l’aide de l’équation de diffusion, avec application à l’assimilation variationnelle de données océaniques*. PhD thesis.
- [Jazwinski, 2007] Jazwinski, A. H. (2007). *Stochastic processes and filtering theory*. Courier Corporation.
- [Jean-Baptiste et al., 2011] Jean-Baptiste, N., Malaterre, P. O., Dorée, C., and Sau, J. (2011). Data assimilation for real-time estimation of hydraulic states and unmeasured perturbations in a 1D hydrodynamic model. *Mathematics and Computers in Simulation*, 81(10) :2201–2214.
- [Kalman, 1960] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems.
- [Knebl et al., 2005] Knebl, M. R., Yang, Z. L., Hutchison, K., and Maidment, D. R. (2005). Regional scale flood modeling using NEXRAD rainfall, GIS, and HEC-HMS/RAS : A case study for the San Antonio River Basin Summer 2002 storm event. *Journal of Environmental Management*, 75 :325–336.

- [Krause et al., 2005] Krause, P., Boyle, D. P., and Bäse, F. (2005). Advances in Geosciences Comparison of different efficiency criteria for hydrological model assessment. *Advances In Geosciences*, 5 :89–97.
- [Lai and Monnier, 2009] Lai, X. and Monnier, J. (2009). Assimilation of spatially distributed water levels into a shallow-water flood model. Part I : Mathematical method and test case. *Journal of Hydrology*, 377 :1–11.
- [Laio and Tamea, 2006] Laio, F. and Tamea, S. (2006). Verification tools for probabilistic forecasts of continuous hydrological variables. *Hydrology and Earth System Sciences Discussions*, 3 :2145–2173.
- [Lawson and Hansen, 2004] Lawson, W. G. and Hansen, J. a. (2004). Implications of Stochastic and Deterministic Filters as Ensemble-Based Data Assimilation Methods in Varying Regimes of Error Growth. *Monthly Weather Review*, 132 :1966–1981.
- [Le Dimet and Talagrand, 1986] Le Dimet, F. a.-X. and Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations : theoretical aspects. *Tellus A*.
- [Lees et al., 1994] Lees, M., Young, P. C., Ferguson, S., Beven, K. J., and Burns, J. (1994). An adaptive flood warning scheme for the River Nith at Dumfries. *River Flood Hydraulics*. Wiley, Chichester, pages 65–75.
- [Li et al., 2009] Li, H., Kalnay, E., and Miyoshi, T. (2009). Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter. *Quarterly Journal of the Royal ...*, 135(February) :523–533.
- [Lions, 1996] Lions, P.-L. (1996). Mathematical Topics in Fluid Mechanics : Volume 1 : Incompressible Models.
- [Lorenc, 1986] Lorenc, A. (1986). Analysis methods for numerical weather prediction. *Quarterly Journal of the Royal Meteorological Society*, 112 :1177–1194.
- [Madsen and Skotner, 2005] Madsen, H. and Skotner, C. (2005). Adaptive state updating in real-time river flow forecasting - A combined filtering and error forecasting procedure. *Journal of Hydrology*, 308 :302–312.
- [Makhlouf, 1994] Makhlouf, Z. (1994). *Compléments sur le modèle pluie-débit gr4j et essai d'estimation de ses paramètres*. PhD thesis.
- [Mallat et al., 1998] Mallat, S., Papanicolaou, G., and Zhifeng, Z. (1998). Adaptive Covariance Estimation of Locally Stationary Processes. *The annals of statistics*, 26(1) :1–47.
- [Malleron et al., 2011] Malleron, N., Zaoui, F., Goutal, N., and Morel, T. (2011). On the use of a high-performance framework for efficient model coupling in hydroinformatics. *Environmental Modelling and Software*, 26(12) :1747–1758.
- [Marin and Monnier, 2009] Marin, J. and Monnier, J. (2009). Superposition of local zoom models and simultaneous calibration for 1D-2D shallow water flows. *Mathematics and Computers in Simulation*, 80 :547–560.
- [Matgen et al., 2010] Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Plaza, D., Pauwels, V. R. N., De Lannoy, G. J. M., De Keyser, R., and Savenije, H. H. G. (2010). Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the Particle Filter : Proof of concept. *Hydrology and Earth System Sciences*, 14 :1773–1785.

- [Maybeck, 1979] Maybeck, P. S. (1979). Stochastic Models, Estimation, and Control. Vol. 1.
- [Miglio et al., 2005] Miglio, E., Perotto, S., and Saleri, F. (2005). Model coupling techniques for free-surface flow problems : Part I. *Nonlinear Analysis, Theory, Methods and Applications*, 63 :1885–1896.
- [Mitchell and Houtekamer, 2000] Mitchell, H. L. and Houtekamer, P. L. (2000). An adaptive ensemble Kalman filter. *Monthly Weather Review*, 128(1994) :416–433.
- [Miyoshi, 2005] Miyoshi, T. (2005). Ensemble Kalman Filter Experiments With a Primitive-Equation Global Model. *Doctor*, (2002).
- [Miyoshi, 2011] Miyoshi, T. (2011). The Gaussian Approach to Adaptive Covariance Inflation and Its Implementation with the Local Ensemble Transform Kalman Filter. *Monthly Weather Review*, 139 :1519–1535.
- [Moulin, 2007] Moulin, L. (2007). *Prévision des crues rapides avec des modèles hydrologiques globaux Application aux bassins opérationnels de la Loire supérieure : évaluation des modélisations , prise en compte des incertitudes sur les précipitations moyennes*. PhD thesis.
- [Murphy, 1973] Murphy, A. H. (1973). A New Vector Partition of the Probability Score.
- [Neal et al., 2007] Neal, J. C., Atkinson, P. M., and Hutton, C. W. (2007). Flood inundation model updating using an ensemble Kalman filter and spatially distributed measurements. *Journal of Hydrology*, 336 :401–415.
- [Ott et al., 2002] Ott, E., Hunt, B. R., Szunyogh, I., Zimin, A. V., Kostelich, E. J., Corazza, M., Kalnay, E., Patil, D. J., and Yorke, J. a. (2002). A Local Ensemble Kalman Filter for Atmospheric Data Assimilation. *Tellus A*, 56(5) :415–428.
- [Pannekoucke et al., 2008] Pannekoucke, O., Berre, L., and Desroziers, G. (2008). Background error correlation length scale estimates and their sampling statistics. *Quarterly Journal of the Royal ...*, 134(March) :497–508.
- [Paquier, 1995] Paquier, A. (1995). *Modélisation et simulation de la propagation de l’onde de rupture de barrage*. PhD thesis.
- [Peña, 2011] Peña, M. (2011). Ensemble Forecasts and their Verification Expanded. page 74.
- [Pedinotti et al., 2014] Pedinotti, V., Boone, A., Ricci, S., Biancamaria, S., and Mognard, N. (2014). Assimilation of satellite data to optimize large scale hydrological model parameters : a case study for the SWOT mission. *Hydrology and Earth System Sciences Discussions*, 11 :4477–4530.
- [Perthame and Gerbeau, 2001] Perthame, B. and Gerbeau, J.-F. (2001). Derivation of viscous Saint-Venant system for laminar shallow water ; Numerical validation. *Discrete and Continuous Dynamical Systems - Series B*, 1 :89–102.
- [Pham, 1996] Pham, D. (1996). A singular evolutive interpolated Kalman filter for data assimilation in oceanography. *Rapport technique RT*, pages 1–22.
- [Pham et al., 1998] Pham, D. T., Verron, J., and Roubaud, M. C. (1998). A singular evolutive extended Kalman filter for data assimilation in oceanography. *Journal of Marine Systems*, 16 :323–340.

- [Raynaud et al., 2009] Raynaud, L., Berre, L., and Desroziers, G. (2009). Objective filtering of ensemble-based background-error variances. *Quarterly Journal of the Royal Meteorological Society*, 135 :1177–1199.
- [Refsgaard, 1997] Refsgaard, J. C. (1997). Parameterisation, calibration and validation of distributed hydrological models. *Journal of Hydrology*, 198 :69–97.
- [Ricci et al., 2011] Ricci, S., Piacentini, a., Thual, O., Le Pape, E., and Jonville, G. (2011). Correction of upstream flow and hydraulic state with data assimilation in the context of flood forecasting. *Hydrology and Earth System Sciences*, 15 :3555–3575.
- [Rochoux, 2014] Rochoux, M. (2014). *Vers une meilleure prévision de la propagation d'incendies de forêt : évaluation de modèles et assimilation de données*. PhD thesis.
- [Sacher and Bartello, 2008] Sacher, W. and Bartello, P. (2008). Sampling Errors in Ensemble Kalman Filtering. Part I : Theory. *Monthly Weather Review*, 136(Evensen 1992) :3035–3049.
- [Sasaki, 1958] Sasaki, Y. (1958). An objective analysis based on the variational method. *J. Meteor. Soc. Japan*, 36(3) :77–88.
- [Sasaki et al., 1955] Sasaki, Y., Gu, P., and Yan, L. (1955). A fundamental study of the numerical prediction based on the variational principle. *J. Meteor. Soc. Japan*, 33(6) :262–275.
- [Sauquet et al., 2007] Sauquet, E., Erat, J., and Prudhomme, C. (2007). La prévision hydro-météorologique à 3-6 mois. Etat des connaissances et applications. *Congrès SHF- Gestion sociale et économique de l'eau - Comment agir sur la demande ? Paris, 17-18 octobre 2007*, pages 133–140.
- [Schwarz, 1869] Schwarz, H. A. (1869). Ueber einige Abbildungsaufgaben. *Journal für die reine und angewandte Mathematik*, 70 :105–120.
- [Shiiba et al., 2000] Shiiba, M., Laurenson, X., and Tachikawa, Y. (2000). Real-time stage and discharge estimation by a stochastic-dynamic flood routing model. *Hydrological Processes*, 14(February 1999) :481–495.
- [Talagrand et al., 1997] Talagrand, O., Vautard, R., and Strauss, B. (1997). Evaluation of probabilistic prediction systems. In *Proc. ECMWF Workshop on Predictability*, pages 1–25.
- [Tangara, 2005] Tangara, M. (2005). *Nouvelle méthode de prévision de crue utilisant un modèle pluie-débit global*. PhD thesis.
- [Thompson et al., 2004] Thompson, J. R., Sørensen, H. R., Gavin, H., and Refsgaard, A. (2004). Application of the coupled MIKE SHE/MIKE 11 modelling system to a lowland wet grassland in southeast England. *Journal of Hydrology*, 293(1) :151–179.
- [Thual, 2010] Thual, O. (2010). *Hydrodynamique de l'environnement*. Ecole polytechnique.
- [Trudel et al., 2014] Trudel, M., Leconte, R., and Paniconi, C. (2014). Analysis of the hydrological response of a distributed physically-based model using post-assimilation (EnKF) diagnostics of streamflow and in situ soil moisture observations. *Journal of Hydrology*, 514 :192–201.

- [van Leeuwen, 1999] van Leeuwen, P. J. (1999). Comment on "Data assimilation using an ensemble Kalman filter technique". *Monthly Weather Review*, 127(8) :1374–1377.
- [Vázquez-Cendón, 1999] Vázquez-Cendón, M. E. (1999). Improved Treatment of Source Terms in Upwind Schemes for the Shallow Water Equations in Channels with Irregular Geometry. *Journal of Computational Physics*, 148 :497–526.
- [Verlaan and Heemink, 1995] Verlaan, M. and Heemink, A. W. (1995). Reduced rank square root filters for large scale data assimilation problems. In *Second International Symposium on Assimilation of Observations in Meteorology and Oceanography*, volume 1730. World Meteorological Organisation WMO.
- [Wang and Bishop, 2003] Wang, X. and Bishop, C. H. (2003). A Comparison of Breeding and Ensemble Transform Kalman Filter Ensemble Forecast Schemes. *Journal of the Atmospheric Sciences*, 60 :1140–1158.
- [Wang et al., 2004] Wang, X., Bishop, C. H., and Julier, S. J. (2004). Which Is Better, an Ensemble of positive-Negative Pairs or a Centered Spherical Simplex Ensemble? *Bulletin of the American Meteorological Society*, 132 :1590–1605.
- [Warren and Bach, 1992] Warren, I. R. and Bach, H. (1992). MIKE 21 : a modelling system for estuaries, coastal waters and seas. *Environmental Software*, 7(4) :229–240.
- [Weigel et al., 2007] Weigel, A. P., Liniger, M. a., and Appenzeller, C. (2007). Generalization of the Discrete Brier and Ranked Probability Skill Scores for Weighted Multimodel Ensemble Forecasts. *Monthly Weather Review*, 135(Wilks 1995) :2778–2785.
- [Whitaker and Hamill, 2002] Whitaker, J. S. and Hamill, T. M. (2002). Ensemble Data Assimilation without Perturbed Observations. *Monthly Weather Review*, 130 :1913–1924.
- [Wick, 1950] Wick, G.-C. (1950). The evaluation of the collision matrix. *Physical review*, 80(2) :268.
- [Wilks, 2011] Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*, volume 100. Academic press.
- [World Meteorological Organization, 2011] World Meteorological Organization (2011). *Manual on flood forecasting and warning*. Number 1072.
- [Yoon et al., 2012] Yoon, Y., Durand, M., Merry, C. J., Clark, E. a., Andreadis, K. M., and Alsdorf, D. E. (2012). Estimating river bathymetry from data assimilation of synthetic SWOT measurements. *Journal of Hydrology*, 464-465 :363–375.

Table des figures

1	Exemples d'informations délivrées quotidiennement par le SCHAPI et les SPC sur le site internet vigicrues.	2
1.1	Types de catastrophes naturelles liées à l'eau sur la période 1990-2001. Source : [World Meteorological Organization, 2011].	10
1.2	Tableau des risques d'impact d'une crue en fonction de la zone considérée. Source : [World Meteorological Organization, 2011].	13
1.3	Les différents modèle météorologiques, pluie-débit et débit-débit et les échéances de prévision relatives. Source : [Moulin, 2007].	15
1.4	Les différents modèles de routage. Source : [World Meteorological Organization, 2011].	16
1.5	Illustration de différents niveaux d'alerte en fonction du niveau d'eau. Source : [World Meteorological Organization, 2011].	20
1.6	Exemple de carte de vigilance émise par le SCHAPI le 17 Juin 2013 à 16h.	22
1.7	Le bassin versant de l'Adour et les principaux affluents de l'Adour.	24
1.8	Représentation schématique du réseau hydraulique Adour maritime (en noir) et de l'emprise relative du modèle 2D (bleu cyan). Les étoiles rouges représentent des stations d'observation de hauteur d'eau et les traits noirs sur les biefs 3, 6, 7 des seuils hydrauliques. Les noms des cours d'eau en fonction des biefs sont les suivants : biefs 1, 2 et 4, Adour; bief 3, Nive; bief 5, Gaves Réunis; bief 6, Gave d'Oloron, bief 7, Gave de Pau.	26
1.9	Maillage du modèle 2D sur la ville de Bayonne repéré en système Lambert 93 avec l'échelle du fond en cote marine (m). La Plaine d'Ansot et la Barthes de Quartier Bas sont séparées par l'autoroute A63 (que l'on devine en vert pâle sur la figure). Le modèle est délimité à l'amont sur l'Adour par le pont de l'A63 et à l'aval par le pont de Lesseps.	27
1.10	Illustration des forçages à l'amont du modèle Adour maritime, épisode de Janvier 2014.	28
1.11	(a) Photo aérienne de la crue de l'Adour à Dax, Février 1952. (b) Crue de l'Adour à Dax, Janvier 2014.	29
1.12	Episodes de crue à Bayonne, (a) quais de la Nive, quartier Saint-Esprit, (b) quai de Lesseps.	30
2.1	Illustration d'un écoulement mono-dimensionnel dans un canal à fond plat formant un angle γ avec l'horizontale. Source : [Thual, 2010].	42

2.2	(a) Représentation du modèle 2D au niveau d'une confluence (b) Recouvrement des modèles 1D et 2D au niveau de la confluence. Source : [Goutal, 2014].	51
3.1	Représentation schématique de la répartition des observations avec le 3D-Var.	75
3.2	Illustration du 4D-Var. A chaque cycle d'assimilation le modèle est intégré du temps t_0 au temps t_N à partir de l'ébauche $\mathbf{x}_c^b(t_0)$. Le vecteur d'innovation $\mathbf{d}_i$ est calculé pour chaque observation $\mathbf{y}_i^o$, $i = 1, \dots, N$. On procède à la minimisation de $\mathcal{J}$ en utilisant le système (3.37). L'état analysé pour le cycle d'assimilation c au temps t_0 , $\mathbf{x}_c^a(t_0)$, est ensuite propagé par le modèle jusqu'à la fin de la fenêtre d'assimilation, c'est la trajectoire d'analyse représentée en traits tirés gris. L'état analysé au temps final du cycle d'assimilation c , $\mathbf{x}_c^a(t_N)$ est utilisé comme ébauche au temps initial du cycle suivant.	77
3.3	Représentation schématique du 4D-Var incrémental. Dans un premier temps on définit une fonctionnelle quadratique $\tilde{\mathcal{J}}_1$ approchant $\mathcal{J}$ (courbe rouge) dans le voisinage de x^b (parabole bleue). L'incrément qui minimise $\tilde{\mathcal{J}}_1$, $\delta x^{a,1}$, est obtenu en résolvant $\nabla \tilde{\mathcal{J}}_1(\delta x) = 0$. On définit un nouvel état de référence $x^{r,1}$ par $x^{r,1} = x^b + \delta x^{a,1}$. On itère le processus jusqu'à ce que les $x^{r,k}$ convergent vers le minimum de la fonctionnelle $\mathcal{J}$, x^a . Source : [Isabelle Mirouze, 2010].	79
3.4	Représentation schématique de la répartition des observations avec le 3D-FGAT.	80
3.5	Représentation graphique de la fonction ρ avec $c = 1$	90
4.1	Représentation schématique de la densité de probabilité prévue en fonction des différents membres de l'ensemble. Source : [Peña, 2011].	92
4.2	Série temporelle des observations (ligne verte épaisse) et évolution temporelle des membres (traits fins) lors d'une prévision du temps $0h$ (temps d'analyse) au temps $12h$	93
4.3	Série temporelle des observations (ligne verte épaisse) et boîtes à moustache (en bleu) lors d'une prévision du temps $0h$ (temps d'analyse) au temps $12h$	93
4.4	Série temporelle des observations (ligne verte épaisse) et de la moyenne des membres de l'ensemble plus ou moins deux écarts-types (traits rouges) lors d'une prévision du temps $0h$ (temps d'analyse) au temps $12h$	94
4.5	Fonction de répartition complémentaire (a) discrète et (b) continue.	95
4.6	Classement de l'observation (trait rouge vertical) par rapport aux membres (traits noirs verticaux).	97
4.7	Différentes configurations pour le diagramme de rang. Source : [Callado et al., 2013].	98
4.8	Ensemble (a) très fiable (b) peu fiable (c) sous-prédictif (d) sur-prédictif. La diagonale en pointillés noirs représente la fiabilité parfaite.	99

4.9	Illustration de la discrimination. Les figures (a) et (c) présentent des figures avec une bonne discrimination, les densités de probabilité sont distinctes. La figure (b) présente au contraire une mauvaise discrimination, les densités de probabilité sont en partie superposées. Source : [Peña, 2011].	102
4.10	Courbes de ROC pour un système de prévisions (a) parfait (b) incapable de discriminer (c) réaliste.	104
5.1	Représentation schématique du couplage de modèles 1D et 2D avec interface. Ici le modèle 1D est en amont du modèle 2D mais on peut également considérer le cas où le modèle 2D est en amont du modèle 1D.	107
5.2	Le domaine délimité en rouge représente le modèle 2D et les modèles 1D sont représentés en bleu. Source : [Miglio et al., 2005].	112
5.3	Comparaison des hauteurs d'eau calculées avec le modèle 2D complet (à gauche) et le modèle couplé (à droite) aux temps $t = 250s$ (en haut) et $t = 300s$. Source : [Miglio et al., 2005].	113
5.4	Hauteurs d'eau calculées avec le modèle 2D avec obstacle. De gauche à droite et de haut en bas : 2D complet, modèle couplé avec l'interface positionnée à l'abscisse $x = 70$, $x = 50$ et $x = 40$. Source : [Miglio et al., 2005].	114
5.5	Localisation du barrage de Malpasset, des modèles 1D et 2D, des transformateurs électriques et des instruments de mesure. Source : [Malleron et al., 2011].	115
5.6	Comparaison des données expérimentales et des cotes d'eau calculées par les modèles 1D et 2D complets et le modèle couplé 1D-2D aux stations de mesure (a) 10, 11, 12 et (b) 13 et 14. Source : [Malleron et al., 2011].	116
5.7	Localisation de la zone d'étude (à gauche) et disposition des modèles numériques (à droite). Source : [Malleron et al., 2011].	117
5.8	Séries temporelles de la cote de la surface libre calculée par le modèle 2D complet et par le modèle couplé. Source : [Malleron et al., 2011].	118
5.9	Représentation schématiques d'un modèle 1D et de deux modèles 2D locaux superposés au modèle 1D.	119
5.10	(a) Représentation de la rivière dans le plan (x, y) (b) Représentation de la section mouillée S dans le plan (y, z) . Source : [Marin and Monnier, 2009].	120
5.11	Algorithme de couplage de modèles superposés. Source : [Fernández-Nieto et al., 2010].	125
5.12	Bathymétrie utilisée dans [Fernández-Nieto et al., 2010].	126
5.13	Comparaison à convergence de l'algorithme de couplage des vitesses ((a) et (c)) et des différences relatives entre les modèles couplés et le modèle 2D de référence ((b) et (d)) pour les ratios $R = 2$ et $R = 10$ respectivement. Source : [Fernández-Nieto et al., 2010].	127
6.1	Estimation <i>a posteriori</i> de forçages latéraux dans [Jean-Baptiste et al., 2011].	132
6.2	Pourcentage d'amélioration de l'erreur relative moyenne en valeur absolue sur une période de 26 jours avec la méthode de correction des forçages amonts présentée dans [Ricci et al., 2011].	133

6.3	Hauteurs d'eau calculées dans [Madsen and Skotner, 2005] avec : combinaison de la correction de l'état hydraulique et d'un modèle de prévision de l'erreur (ligne fine continue), correction de l'état hydraulique seule (ligne continue moyenne) et absence de correction (ligne continue épaisse). Les observations sont représentées par les carrés inclinés.	134
6.4	Erreur moyenne en valeur absolue sur les débits sans et avec la méthode d'assimilation utilisée dans [Andreadis et al., 2007].	135
6.5	Estimation itérative de la bathymétrie dans [Yoon et al., 2012]. Etat vrai (ligne rouge), ébauche (ligne bleue tiretée), assimilation aux cycles 1 et 8 (lignes vertes et violettes tiretées respectivement).	137
8.1	(a) Exemple de PrePALM avec deux branches et trois unités et une connexion entre les unités 1 et 2 (b) Exemple d'unité PALM avec une entrée faisant appel à la primitive PALM.Put et une sortie faisant appel à la primitive PALM.Get. Source : [Rochoux, 2014].	164
8.2	Représentation schématique du réseau hydraulique Adour maritime. Les étoiles rouges représentent des stations d'observation de hauteur d'eau et les traits noirs sur les biefs 3, 6, 7 des seuils hydrauliques.	165
8.3	(a) Hauteurs d'eau (courbe bleue) et bathymétrie (courbe rouge) sur le chemin 7-5-2-1 (b) débits sur le chemin 3-1. Les traits verts verticaux représentent la position des seuils sur les biefs 7 et 3. Les traits noirs verticaux représentent la séparation entre les biefs.	166
8.4	Représentation schématique de l'enchaînement des cycles « propagation-estimation de $\mathbf{B}$ ».	167
8.5	Représentation schématique de l'enchaînement des cycles « assimilation-prévision », t_1 étant le premier temps de base, t_2 le second...	168
8.6	Schéma PrePALM pour l'EnKF.	169
8.7	Illustration de l'ajout d'un bruit sur le forçage à la station d'Escos lors d'un pic de crue. La ligne noire en pointillé représente le forçage initial et les lignes rouge, verte et bleue représentent le forçage initial auquel on a ajouté une perturbation.	170
8.8	Fonctions de covariance relatives à la station de Peyrehorade (a) ZZ et (b) QZ le long du chemin 6-5-2-1 (c) bathymétrie. Les traits verticaux noirs indiquent la séparation entre les biefs, le trait rouge vertical indique la position de la station de Peyrehorade et le trait vert vertical indique la position du seuil sur le bief 6.	173
8.9	Illustration de la ligne d'eau pour deux membres sur le chemin 6-5-2-1 (a) hauteurs d'eau et (b) débits.	174
8.10	Fonctions de covariance (a) ZZ et (b) QZ sur le chemin 6-5-2-1 à deux temps différents. Il faut souligner que les temps auxquels on présente ces figures sont différents du temps de la figure 8.8.	175
8.11	Fonctions de corrélation relatives à la station de Peyrehorade (a) ZZ et (b) QZ le long du chemin 7-5-2-1. (c) et (d) bathymétrie. Les traits verticaux noirs représentent la séparation entre les biefs. Le trait rouge vertical représente la position de la station de Peyrehorade.	176

8.12	Fonctions de corrélation relatives à la station de Peyrehorade (a) ZZ et (b) QZ sur le chemin 6-5-2-1 à deux temps différents.	177
8.13	Fonction de corrélation ZZ relative à la station de Urt (trait rouge vertical) sur les chemins (a) 3-1 (b) 4-2-1 (c) 7-5-2-1 (d) 6-5-2-1. La longueur de portée amont est la somme des longueurs des zones représentées par le trait rouge horizontal.	179
8.14	Evolution temporelle à la station de Peyrehorade des (a) longueur de portée amont, aval et moyenne (b) des hauteurs d'eau.	180
8.15	Evolution temporelle à la station de Urt des (a) longueur de portée amont et aval (b) des débits.	181
8.16	Fonction de covariance relative à la station de Peyrehorade (trait rouge vertical) sur le chemin 6-5-2-1.	183
8.17	Incréments d'analyse de (a) hauteurs d'eau et de (b) débits pour le KF invariant (courbe rouge) et pour l'EnKF (courbe verte).	184
8.18	Incréments d'analyse de (a) hauteurs d'eau et de (b) débits pour l'EnKF à deux temps différents sur le chemin 7-5-2-1.	185
8.19	Séries temporelles des hauteurs d'eau analysées à (a) Peyrehorade et (b) Urt.	186
8.20	Séries temporelles des débits analysés à (a) Peyrehorade et (b) Urt.	186
8.21	Séries temporelles des (a) hauteurs d'eau prévues à 3h à Peyrehorade et des (b) débits prévus à 6h à Urt.	188
8.22	Evolution temporelle des longueurs de portée diagnostiquées à la station de Peyrehorade (a) avec assimilation (b) sans assimilation.	189
8.23	Séries temporelles des (a) observations et des hauteurs d'eau analysées avec l'EnKF et (b) des écarts-types et des moyennes des innovations.	190
8.24	Evolution des écarts-types d'erreur d'observation estimés (courbe rouge) et prescrits (ligne verte). Chaque itération correspond en fait à un cycle d'assimilation.	194
8.25	Diagramme présentant l'étape d'inflation avant celle d'analyse.	197
8.26	Séries temporelles des hauteurs d'eau analysées à (a) Peyrehorade et (b) Urt.	198
8.27	Séries temporelles des débits analysés à (a) Peyrehorade et (b) Urt.	199
8.28	Série temporelle (a) des longueurs de portée à Peyrehorade (b) des moyennes et des écarts-types des innovations.	201
8.29	Séries temporelles des hauteurs d'eau à (a) Peyrehorade et à (b) Urt.	203
8.30	Observations (ligne bleue) et prévisions déterministes (moyenne des membres, ligne rouge) à la station de Peyrehorade aux échéances de (a) 6 heures et (b) 24 heures.	205
8.31	Diagrammes de rang à la station de Peyrehorade aux échéances de (a) 6 heures et (b) 24 heures. La ligne rouge tiretée correspond au diagramme plat que l'on est censés obtenir dans l'idéal.	206
8.32	Séries temporelles des observations (ligne bleue) et moyenne des membres (a) plus ou moins deux écarts-types et (b) avec diagramme de type « boîte à moustaches ».	207
8.33	Diagramme de fiabilité aux échéances de (a) 6 heures et (b) 24 heures.	208

8.34	Evolution du DRPS en fonction de l'échéance de prévision aux stations de (a) Peyrehorade et (b) Lesseps.	209
8.35	Courbes de ROC à l'échéance de 24 heures pour les cotes (a) 5.11 m et (b) 6.71 m.	211
9.1	Superposition d'une image Google Maps de la ville de Bayonne, du maillage 2D et de la couche SIG ayant servi pour établir les sections en travers du modèle Adour maritime (traits rouges).	215
9.2	Champs de vitesse (unité $m.s^{-1}$) modélisés avec le modèle 2D sur la plaine d'Ansot et la barthe de Quartier Bas à (a) marée descendante et (b) marée montante.	217
9.3	Représentation schématique de la disposition des différents modèles sur la zone de Bayonne. Les numéros indiquent le numéro de bief correspondant du modèle Adour maritime.	218
9.4	Illustration de l'algorithme de couplage. dt désigne le pas de temps spécifique à chaque modèle. Les flèches rouges représentent les « échanges de variables » aux interfaces.	221
9.5	Séries temporelles des hauteurs d'eau calculées par le modèle 1D Adour maritime (courbe verte) et le modèle couplé (courbe rouge) au niveau de deux interfaces (a) interface aval Adour (b) interface Nive amont.	224
9.6	Séries temporelles des débits calculés par le modèle 1D Adour maritime (courbe verte) et le modèle couplé (courbe rouge) au niveau de deux interfaces (a) interface aval Adour (b) interface Nive amont.	224
9.7	Séries temporelles des modèles 1D (courbe verte), 2D (courbe rouge tiretée) et couplé (courbe rouge continue) et observations (courbe bleue) aux stations de (a) Lesseps et (b) Pont-Blanc pendant la crue de 2014.	225
9.8	Disposition relative des modèles 1D et 2D et représentation schématique des apports latéraux. Les numéros représentent les numéros de bief du modèle Adour maritime.	227
9.9	(a) Découpage du lit mineur (délimité par les bandes rouges) en sections sur une portion à l'amont de la Nive (b) Interfaces entre lit mineur et lit majeur sur une section (délimitée par les traits noirs), il s'agit des lignes formées par les arêtes reliant les points rouges.	229
9.10	Séries temporelles des modèles couplés 1D (ligne verte) et 2D (ligne rouge) (a) débits à la frontière liquide aval sur l'Adour (b) hauteurs d'eau à la frontière liquide amont sur la Nive.	232
9.11	Invariants de Riemann (a) entrant et (b) sortant aux limites du modèle 2D, frontière liquide Adour aval (ligne bleue), Adour amont (ligne rouge), Nive amont (ligne verte).	233
9.12	Séries temporelles des modèles 1D sans couplage (ligne verte tiretée) et 1D couplé (ligne verte continue) (a) hauteurs d'eau à la frontière liquide aval sur l'Adour (b) débits à la frontière liquide amont sur la Nive.	234
9.13	Séries temporelles des hauteurs d'eau des modèles 1D couplé (ligne verte continue), 1D non couplé (ligne verte tiretée) (a) frontière liquide amont sur la Nive (b) station de Pont-Blanc.	235

9.14	Séries temporelles à la station de Pont-Blanc des modèles 1D non couplé (ligne verte tiretée), 1D couplé (ligne verte continue) et 2D couplé (ligne rouge) (a) débits (b) hauteurs d'eau.	236
9.15	Séries temporelles des hauteurs d'eau des modèles 1D couplé (ligne verte), 2D couplé (ligne rouge) et des observations (ligne bleue) aux stations de (a) Lesseps et (b) Pont-Blanc.	237
9.16	(a) Séries temporelles des flux à l'interface entre lit mineur et lit majeur en amont de la Nive (b) Découpage du lit mineur en sections. Il y a une correspondance dans le choix des couleurs entre les figures (a) et (b) qui permet d'identifier à quelle section correspond quel flux.	238
9.17	(a) Séries temporelles des flux à l'interface entre lit mineur et lit majeur sur la Nive au niveau de Plaine d'Ansot et de la station de pont-Blanc (b) Découpage du lit mineur en sections. Il y a une correspondance dans le choix des couleurs entre les figures (a) et (b) qui permet d'identifier à quelle section correspond quel flux.	239
9.18	Série temporelle des flux à l'interface entre lit mineur et lit majeur sur la même zone celle présentée en figure 9.17 lorsque le forçage maritime aval est prépondérant.	240
9.19	Débits imposés à l'amont du modèle 2D sur la Nive, sans assimilation (ligne verte tiretée), avec assimilation (ligne verte continue).	242
9.20	Flux échangés entre la Nive et la Barthes de Quartier Bas (a) sans assimilation (b) avec assimilation.	243
9.21	Observations (courbe bleue) et hauteurs d'eau calculées avec le modèle 2D avec assimilation (courbe rouge) et sans assimilation (courbe noire) à la station de Pont-Blanc (a) après assimilation (b) à 1 heure de prévision (c) à 3 heures de prévision (d) à 6 heures de prévision.	244
A.1	Laisse de crue du 19 Juin 2013 à Peyrehorade.	251
A.2	(a)-(b) Berge Nord de Peyrehorade.	252
A.3	(a) Berge Sud de Peyrehorade (b) Berge Nord de Peyrehorade.	253
A.4	(a) Route de Hastingue-Guiche, berge Sud de Peyrehorade (b) Route de Oeyregave, berge Sud de Peyrehorade.	254
A.5	(a) Route d'Oloron, berge Sud de Peyrehorade (b) Route de Sorde l'Abbaye, berge Nord de Peyrehorade.	255
A.6	(a) Gave d'Oloron à Sorde l'Abbaye (b) Gave d'Oloron à Sorde l'Abbaye.	256
A.7	Route départementale de Sorde l'Abbaye.	257
B.1	Evolution temporelle de la fonction de corrélation ZZ relative à la station de Peyrehorade sur le chemin 7-5-2-1 du temps t au temps $t + 7h$	260
C.1	Episode 2009-1. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps.	262
C.2	Episode 2009-2. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.	263
C.3	Episode 2010. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.	264

C.4	Episode 2011-1. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.	265
C.5	Episode 2011-2. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.	266
C.6	Episode 2012. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.	267
C.7	Episode 2013. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.	268
C.8	Episode 2014. Forçages (a) amonts (b) aval et observations à (c) Peyrehorade (d) Urt (e) Lesseps (f) Pont-Blanc.	269

Liste des tableaux

4.1	Propriétés et critères correspondants.	96
4.2	Table de contingence.	102
4.3	Quelques scores associés à la table de contingence.	103
8.1	Scores sur les hauteurs d'eau.	187
8.2	Scores sur les débits.	187
8.3	Ecart-types prescrits et estimés avec l'assimilation.	194
8.4	Scores sur les hauteurs d'eau.	199
8.5	Scores sur les débits.	200
8.6	Pourcentage d'amélioration des scores de l'IEEnKF par rapport au modèle seul. Lorsque les scores sont négatifs cela signifie que l'assimilation a dégradé les résultats par rapport au modèle seul.	202
8.7	Tableau du taux de réussite (Tdr), de la précision (Pre), du taux de corrects négatifs (Tcn), du taux de faux positifs (Tfp), de la probabilité d'occurrence (Pdo) et de la menace (Men) aux échéances de 6h et 24h pour les cotes 4.71 m, 5.11 m, 6.71 m et 7.51 m.	210
9.1	Temps de calcul (s.) des modèles 1D, 2D et couplé sur la crue de Janvier 2014.	226
9.2	Temps de calcul (s.) des modèles 1D, 2D, couplé à interfaces et couplé superposé sur la crue de Janvier 2014.	240
9.3	Différence relative (exprimée en pourcentage) des hauteurs d'eau entre les modèles couplés avec et sans assimilation pour la RMSE, la corrélation et le biais à la station de Pont-Blanc en fonction de l'échéance de prévision.	245