



Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique

Assimilation de données en hydraulique, maquette pour le modèle MASCARET

Andrea Piacentini, Olivier Thual, Gabriel Jonville, Chloé Ivanoff, Sophie Ricci,
Sébastien Massart et Etienne Le Pape

CERFACS, 42 avenue Gaspard Coriolis
31057, Toulouse, France

Technical Report TR/CMGC/10/16

Rapports du contrat entre le SCHAPI et le CERFACS sur l'application de l'assimilation de données à l'hydraulique

A. Piacentini, O. Thual, G. Jonville, Ch. Ivanoff, S. Massart, E. Le Pape

Rapport

"Principe de la maquette d'assimilation de données pour le modèle MASCARET"

31 octobre 2009

Rapport

"Validation de la maquette d'assimilation de données pour le modèle MASCARET"

30 novembre 2009

Rapport de stage

"Assimilation de données pour un modèle d'hydrologie fluviale"

7 septembre 2009

Note

"Expériences jumelles sur un modèle des ondes de crues"

18 janvier 2010

Guide

"Utilisation des maquettes Marne et Adour"

5 février 2010

Rapport du contrat entre le SCHAPI et le CERFACS sur l'application de l'assimilation de données à l'hydraulique

Principe de la maquette d'assimilation pour le modèle MASCARET

A. Piacentini, O. Thual, G. Jonville, Ch. Ivanoff, S. Massart, E. Le Pape

31 octobre 2009

Table des matières

Introduction.....	2
Modèle direct.....	2
Prise en main de la version F90.....	3
Amélioration du modèle direct.....	4
Critères d'évaluation des performances du modèle.....	6
Script de pilotage d'un enchaînement de runs.....	6
Rappel sur la correction a posteriori des résultats.....	7
Chaîne d'assimilation de la ligne d'eau.....	7
Principes algorithmiques.....	7
Lissage et seuillage des observations.....	9
Détermination expérimentale des longueurs de corrélation des erreurs d'ébauche.....	10
Détermination de la correction sur Q (débit).....	10
Implémentation.....	11
Évolution du fichier .cas et du dictionnaire dico.txt.....	11
Script de pilotage de l'enchaînement des runs avec assimilation.....	13
Chaîne d'assimilation pour le calage en temps réel des lois aux extrémités libres.....	13
Principes algorithmiques.....	14
Lissage et seuillage des observations.....	16
Détermination de G et GT par différences finies.....	16
Implémentation avec le coupleur PALM.....	17
Unités pour l'assimilation.....	17
Modification des sources MASCARET.....	18
Estimation des variances.....	18
Évolution du fichier .cas et du dictionnaire dico.txt.....	19
Script de pilotage de l'enchaînement des runs avec assimilation.....	21
Chaîne avec les deux types d'assimilation.....	21

Introduction

Ce document décrit la chaîne d'assimilation de données pour le modèle MASCARET comme maquette démonstrative et pédagogique dans le cadre du contrat passé entre le SCHAPI et le CERFACS sur l'application de l'assimilation de données à l'hydrologie.

Pour la construction de cette maquette nous avons retenu la configuration de MASCARET avec noyau fluvial non permanent REZO et nous avons pris comme configurations de référence les bassins « Adour Maritime » et « Marne Vallage », décrits en détail dans les rapports de calage rédigés respectivement par le SCHAPI et par le SPC SAMA (Seine Amont Marne Amont).

En se basant sur une approche d'assimilation de type BLUE (Best Linear Unbiased Estimate) nous avons mis en place deux types d'assimilation à l'intérieur de la même chaîne. Dans un cas l'assimilation modifie l'état du système au cours de son évolution temporelle de façon à se rapprocher des mesures aux stations d'observations au fur et à mesure qu'elles sont disponibles.

Avec la deuxième approche l'assimilation utilise toutes les observations sur un intervalle de temps pour corriger dynamiquement quelques paramètres ou conditions qui déterminent les résultats du modèle sur cette intervalle. Les deux approches peuvent être combinées.

Dans ce document nous expliquons les principes algorithmiques et les détails techniques pour l'implémentation des trois approches d'assimilation (ligne d'eau, calage des lois et combinaison des deux).

La chaîne a été validée sur un premier jeu d'événements tirés des épisodes illustrés dans les rapports de calage et sur un deuxième jeu provenant de l'application en temps réel pour le bassin « Adour Maritime » qui tourne en routine au SCHAPI. Les résultats de la validation font l'objet d'un deuxième rapport.

Modèle direct

Le modèle direct pour cette étude est le modèle d'hydrologie MASCARET développé par EDF/R&D et le CETMEF, dans sa version utilisée par le SCHAPI pour la chaîne de prévision qui tourne en temps réel pour le bassin « Adour Maritime ».

En particulier, parmi les différents choix possibles dans MASCARET, le noyau de calcul sélectionné est REZO, un noyau fluvial non permanent, relativement léger et bien adapté aux types de phénomènes modélisés.

Pour plus d'informations sur le modèle direct nous faisons référence aux notes de principes d'EDF/R&D qui le décrivent et au guide de prise en main de la version Windows avec interface Fudaa, rédigé par le CETMEF.

Prise en main de la version F90

Le modèle direct de Mascaret 6.1 a été livré par EDF/LNHE (Nicole Goutal) sous forme d'un ensemble de fichiers Fortran90. Aucun environnement de compilation ni aucun cas test n'ont été fournis avec le modèle.

Le style de programmation repose presque entièrement sur les modules F90 qui permettent le partage de variables entre différentes sections du code et une déclaration rigoureuse des interfaces des sous-routines. Le pilotage de l'exécution est à la charge du programme principal qui porte le nom significatif de superviseur.

Nous avons commencé par la mise en place d'un système de compilation automatique par makefile. Un makefile est consacré à la compilation de tous les modules et toutes les sous-routines et à la création d'une librairie d'objets. Le deuxième makefile compile le superviseur et effectue l'édition de liens pour créer l'exécutable `systeme_mascaret`.

Pour lancer Mascaret il faut un environnement d'exécution incluant

- un fichier ascii (extension `.cas`) de paramétrage qui définit tous les paramètres d'exécution de Mascaret, décrit la géométrie, la connectivité et la discrétisation du bassin, fixe les zones de frottement et les valeurs des coefficients de Strickler associés et indique les types des lois hydrauliques imposées aux extrémités libres et l'emplacement des fichiers qui en contiennent les valeurs
- le fichier texte `FichierCas.txt` qui indique l'emplacement du fichier `.cas`
- le fichier texte `dico.txt` qui contient le dictionnaire de l'interprète `damocles` pour la lecture du fichier `.cas`
- le fichier texte `Abaques.txt` qui décrit les pertes de charge aux confluent
- un fichier ascii (extension `.geo`) qui décrit la géométrie des profils en travers et qui doit être cohérent avec le fichier `.cas`
- un fichier ascii (extension `.lig`) qui contient la condition initiale sous forme d'une ligne d'eau (hauteur et débit) définie sur la totalité du réseau discrétisé
- un fichier ascii par extrémité libre (extension `.loi`) contenant la loi hydraulique d'évolution temporelle ou la courbe de tarage de la condition imposée à l'extrémité

Cet environnement nous a été fourni par le SCHAPI et reprend exactement la configuration utilisée pour la chaîne temps réel qui tourne pour le bassin de l'Adour et le bassin de la Marne-Vallage.

Dans cette configuration, les fichiers `.cas`, `.geo` et `.lig` ne changent pas d'une simulation à l'autre, tandis que les lois hydrauliques dépendent de l'événement modélisé. Pour cette étude nous avons pris comme cas de référence les événements utilisés pour le calage des configurations Adour et Marne. Il faut toutefois remarquer que les paramétrages de Mascaret retenus à l'issue du calage et utilisés par le SCHAPI ne permettent pas de reproduire la totalité des

événements. La base de calage est accompagnée de ses propres fichiers .cas et .geo qui varient selon les épisodes pour le modèle Adour (nombre de zones de frottement, valeurs des coefficients de Strickler, pas de planimétrage, positionnement des profils en travers) et qui permettent d'effectuer toutes les simulations, mais le parti a été pris d'utiliser seulement la configuration fournie par le SCHAPI.

Dans une exécution type de la chaîne temps réel au SCHAPI, le modèle est lancé à partir d'une condition initiale standard (fichier .lig) imposée à une date qui précède de 5 jours la temps de début de la prévision. Sur ces 5 jours le modèle est contraint aux extrémités libres par des lois hydrauliques observées. Au delà de ces 5 jours on passe en mode prévision avec lois hydrauliques constantes en temps et égales à la dernière valeur observée. Seule exception, la loi qui impose la hauteur d'eau à Convergent pour l'Adour où, pour conserver la variabilité temporelle, une prévision de marée issue du SHOM est utilisée pour prolonger dans le temps les valeurs observées.

Pour reproduire cette modalité d'exécution nous créons des fichiers .loi à partir de ceux fournis avec la base de calage en coupant les débits imposés au moment où l'on passe en mode prévision et en les prolongeant par une valeur constante. La loi de marée n'est pas coupée. On injecte donc la marée observée à Convergent (Adour) sur l'intervalle temporel de prévision, contrairement à la chaîne temps réel qui utilise la prévision du SHOM.

Amélioration du modèle direct

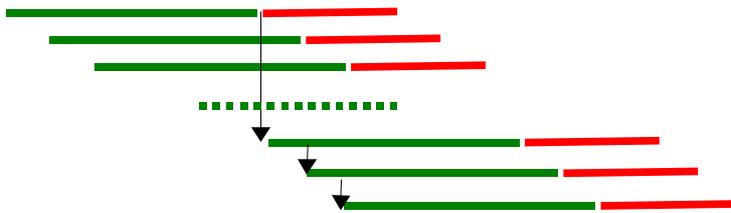
Avant d'attaquer l'implémentation des maquettes d'assimilation nous avons apporté deux corrections au modèle direct.

La première est de caractère purement informatique. Une variable non initialisée comportait, lors de l'allocation dynamique d'un tableau, le risque d'un plantage aléatoire du code. La correction a été transmise à Nicole Goutal qui l'intégrera dans les prochaines versions du code.

La deuxième, plus importante, concerne la possibilité de sauvegarder une ligne d'eau complète à un instant donné et de pouvoir la relire comme condition initiale pour une autre exécution de Mascaret. Cette amélioration permet de dépasser une limitation actuelle de la chaîne temps réel qui relance systématiquement les exécutions à partir d'une ligne d'eau standard et constante. L'utilisation d'une ligne d'eau modélisée cohérente avec les lois hydrauliques utilisées permet de diminuer le temps nécessaire à la mise à l'équilibre du modèle et diminue le risque de plantage (sections à sec ou régime torrentiel) au redémarrage. Pour la suite de l'étude, le choix de l'instant de sauvegarde de la ligne d'eau dépendra du type d'assimilation effectuée. L'utilisation d'une ligne sauvegardée pour le redémarrage nous a permis de réduire le temps de simulation avant la prévision à 3 jours pour l'Adour et à 4 jours pour la Marne.

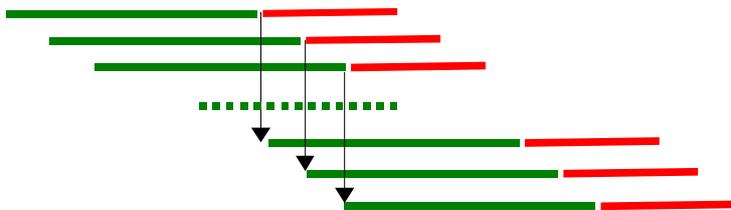
La stratégie de sauvegarde des lignes d'eau pour le redémarrage du modèle dépend du type de simulation.

Pour une simulation sans assimilation ou avec assimilation de la ligne d'eau, si on imagine de lancer le modèle toutes les n minutes à partir d'un état sauvegardé précédemment, on stockera l'état simulé au bout de n minutes. Logiquement en cas de démarrage à froid du modèle on repartira de la ligne d'eau standard et on stockera l'état modélisé à la fin de l'équilibrage. On ne stockera plus d'état jusqu'au moment où la date de début de la simulation coïncidera avec le premier état stocké.



Temps de sauvegarde des lignes d'eau
en simulation directe ou avec assimilation linéique

Pour une simulation avec contrôle des lois, avec ou sans assimilation linéique associée, on sauvegarde de façon systématique l'état prévu à la fin de l'équilibrage.



Temps de sauvegarde des lignes d'eau
en simulation avec contrôle des lois

Critères d'évaluation des performances du modèle

Pour les cas d'étude tirés de la base de calage nous disposons de mesures de hauteur d'eau à quelques stations sensibles. Pour le bassin de l'Adour il s'agit des stations à Peyrehorade, Urt et Lesseps et pour la Marne des stations à Joinville et Chamouilley.

Il est donc possible d'évaluer la qualité des résultats du modèle par comparaison aux observations. En plus de la comparaison visuelle des séries temporelles des hauteurs modélisées et observées, il est possible d'utiliser des critères quantitatifs d'évaluation de la qualité de la modélisation. Nous avons repris trois critères utilisés par le SCHAPI

- le biais sur un intervalle de temps $\frac{100}{n} \sum_{i=1}^n \frac{H_i^{mod} - H_i^{obs}}{H_i^{obs}}$
- la précision sur un intervalle de temps $\frac{100}{n} \sum_{i=1}^n \left| \frac{H_i^{mod} - H_i^{obs}}{H_i^{obs}} \right|$
- l'efficacité sur un intervalle de temps $100 \times \left[1 - \frac{\sum_{i=1}^n [H_i^{obs} - H_i^{mod}]^2}{\sum_{i=1}^n [H_i^{obs} - \overline{H^{obs}}]^2} \right]$

où n indique le nombre d'observations disponibles sur l'intervalle choisi.

Le choix de l'intervalle nous permet d'obtenir des indications différentes. Un intervalle avant le temps de début de la prévision nous permet d'estimer la qualité du modèle en mode simulation, avec lois hydrauliques observées. Des intervalles après le temps de début de la prévision nous permettent de quantifier la capacité prévisionnelle du modèle à différentes échéances avec débits amonts imposés constants.

Script de pilotage d'un enchaînement de runs

Les cas de la base de calage couvrent chacun la totalité d'un événement. Pour simuler la modalité de travail de la chaîne temps réel du SCHAPI nous avons codé un script de pilotage qui permet d'enchaîner des runs de simulation/prévision en faisant glisser la date de début de la prévision, dite « temps de base ».

Dans la pratique, pour chaque temps de base on prépare les fichiers des lois hydrauliques avec valeurs constantes au delà du temps de base, on modifie le fichier .cas pour indiquer les dates de début et de fin de la simulation, on archive les résultats au format Opthyca (fichiers .opt) et on calcule des critères quantitatifs sur la qualité des résultats (biais, précision, efficacité). Le procédé est répété en décalant le temps de base d'un incrément choisi jusqu'à couvrir tout l'événement. Au final on trace les courbes des critères quantitatifs en fonction du temps de base.

Rappel sur la correction a posteriori des résultats

Le SCHAPI a introduit dans la chaîne temps réel une modalité de correction a posteriori des résultats à partir des observations de hauteur d'eau mesurées en quelques stations sensibles. Les valeurs diffusées par le SCHAPI reprennent exactement la valeur observée pour l'instant de début de la prévision et exactement la valeur du modèle à une échéance de prévision fixée station par station. Pour les instants intermédiaires on calcule une moyenne entre la valeur prévue et la dernière observation, pondérées linéairement en fonction du temps : poids 1 à l'observation à l'instant 0 de prévision et poids 1 au modèle à l'échéance. Même si dans les documents qui décrivent la chaîne temps réel cette correction porte le nom d'assimilation, il est important de souligner qu'il s'agit seulement d'une correction des résultats qui n'a aucun impact sur l'exécution du modèle.

Chaîne d'assimilation de la ligne d'eau

La première chaîne d'assimilation implémente une méthode de type BLUE pour la correction dynamique de la ligne d'eau (hauteur et débit) aux instants où l'on dispose d'observations à Peyrehorade, Urt et Lesseps.

Les rappels algorithmiques et pratiques qui suivent n'ont pas pour but d'illustrer la méthode mais plutôt de mettre en évidence les paramètres de l'assimilation qui doivent être fixés et la façon de les estimer.

Principes algorithmiques

La méthode d'assimilation retenue implémente un algorithme d'estimation linéaire de type BLUE (Best Linear Unbiased Estimate).

Si à un instant donné le modèle a fourni une ébauche \underline{x}^b de la quantité que nous voulons contrôler et que nous disposons d'observations \underline{y}^o du système au même instant, nous corrigeons \underline{x} avec un incrément obtenu linéairement de \underline{y}^o

$$\underline{x}^a = \underline{x}^b + \underline{Kd} \quad \text{où} \quad \underline{d} = (\underline{y}^o - \underline{Hx}^b) \quad \text{et} \quad \underline{K} = \underline{B}\underline{H}^T(\underline{H}\underline{B}\underline{H}^T + \underline{R})^{-1}$$

\underline{H} représente un opérateur qui permet de passer des quantités modélisées aux quantités observées, \underline{R} est la matrice des variances/covariances des erreurs associées aux observations, \underline{B} est la matrice des variances/covariances des erreurs associées à l'ébauche.

Il s'agit donc d'une formule classique d'interpolation optimale, sans mise à jour de la matrice \underline{B} . Pour éviter de déstabiliser le système on rejète les observations qui ont des valeurs aberrantes, ou des valeurs dans un intervalle non représentable par le modèle ou qui diffèrent excessivement du modèle. L'étape d'assimilation est activée seulement si l'on dispose d'au moins une observation utile pour le pas de temps traité.

Choix du contrôle

Mascaret représente une ligne d'eau par les valeurs de la hauteur d'eau Z et du débit Q sur chacune des n sections du maillage qui discrétise le domaine modélisé. Pour corriger la ligne d'eau à un instant donné il faut donc contrôler les n valeurs de Z et les n valeurs de Q . Mathématiquement le contrôle est un vecteur de taille $2n$ $[\underline{Z}^T, \underline{Q}^T]^T$.

Vecteur d'observations

On dispose d'observations de Z en correspondance de p stations de mesure ($p=3$ pour l'Adour, $p=2$ pour la Marne) à des instants qui ne coïncident pas nécessairement avec les pas de temps du modèle.

On utilise une approximation « plus proche voisin » : soit s_k l'abscisse curviligne de la station k ($k=1,p$), nous faisons correspondre l'observation à la section i_k du modèle dont l'abscisse diffère le moins de s_k . De même, si t^o est un temps de mesure, nous faisons correspondre l'observation au pas de temps du modèle t_j le plus proche de t^o .

Mathématiquement les observations sont représentées par un vecteur \underline{y}_j^o de taille p à chaque instant t_j où l'on dispose d'observations.

Opérateur d'observation

L'opérateur d'observation effectue la sélection des valeurs de Z aux points de mesure. Il s'agit d'un opérateur linéaire, formellement d'une matrice \underline{H} de p lignes et $2n$ colonnes.

La ligne k ($k=1,p$) a un 1 en position i_k et 0 ailleurs. Les colonnes de $n+1$ à $2n$, qui concernent les débits, sont donc nulles.

Matrice \underline{R} des variances des erreurs d'observation

Les informations statistiques sur les erreurs d'observation sont contenues dans la matrice \underline{R} , de taille $p \times p$. Nous négligeons les corrélations d'erreur entre différentes stations. La matrice \underline{R} est ainsi diagonale. Elle sera donc entièrement spécifiée si l'on indique les variances des erreurs pour chaque station. Dans la pratique on rentrera les écarts type dans le fichier .cas.

Matrice \underline{B} des variances/covariances des erreurs d'ébauche

En théorie, la matrice \underline{B} est une matrice dense de taille $2n$. Nous allons distinguer pour leur significations les éléments diagonaux, les éléments extra-diagonaux des blocs diagonaux et les éléments des blocs extra-diagonaux.

Les éléments diagonaux contiennent les variances des erreurs d'ébauche. Nous avons choisi d'indiquer seulement les variances des erreurs sur Z , qui est aussi la variable observée. La partie de l'incrément d'analyse correspondant à Q sera obtenue par linéarisation de la courbe de tarage à partir de l'incrément sur Z .

Ce procédé évite la spécification des variances d'erreur pour Q.

Dans la pratique on rentrera les écarts type en Z dans le fichier .cas. On indique les écarts type absolus en mètres. A noter que d'autres choix seraient possibles : e.g. des écarts type en pourcentage ou des lois pour calculer dynamiquement les écarts type en temps réel en fonction du régime de l'écoulement.

Les éléments hors diagonaux des blocs diagonaux représentent les corrélations spatiales des erreurs d'ébauche. On fait l'hypothèse d'une corrélation de forme gaussienne en fonction de la distance curviligne le long des biefs, avec longueurs de corrélation différentes pour l'amont et pour l'aval, spécifiées à chaque station pour Z et pour Q.

Les blocs extra diagonaux représentent les corrélations entre Z et Q. Au lieu de les rentrer comme éléments de la matrice, on a choisi de déterminer la correction sur Q aux sections observées par linéarisation de la courbe de tarage à partir de l'incrément sur Z.

Dans la pratique on ne construit que les éléments de $\underline{B}\underline{H}^T$, c'est à dire seulement la partie pour Z des p colonnes de \underline{B} correspondant aux sections observées. A partir de la correction en Z aux stations d'observation on obtient la correction sur Q aux stations d'observation par linéarisation de la courbe de tarage. La correction sur Q est ensuite propagée aux autres sections par multiplication par les colonnes de \underline{B} correspondantes.

Formule d'interpolation optimale

On calcule l'incrément d'analyse par interpolation optimale, c'est à dire par multiplication matricielle $\delta\underline{x} = \underline{K} \underline{d}$ ou \underline{d} est le vecteur de taille p d'écart aux observations $\underline{d} = \underline{y}^o - \underline{H}\underline{x}^b$.

L'expression de la matrice \underline{K} , appelée aussi matrice de gain ou filtre, est $\underline{K} = \underline{B}\underline{H}^T(\underline{H}\underline{B}\underline{H}^T + \underline{R})^{-1}$

Dans la pratique on construit directement les p colonnes de $\underline{B}\underline{H}^T$ et la matrice de taille p $\underline{H}\underline{B}\underline{H}^T$. p étant petit, la matrice $(\underline{H}\underline{B}\underline{H}^T + \underline{R})$ est réellement inversée par des appels à la librairie d'algèbre linéaire lapack.

Lissage et seuillage des observations

Les observations sont disponibles avec une fréquence donnée qui ne correspond pas nécessairement au pas de temps du modèle. Pour mettre en adéquation le pas d'observation avec le pas de temps du modèle et pour filtrer le bruit des observations avec un pas de temps court (typiquement 5 minutes), on procède à un filtrage par moyenne glissante des observations. Si la longueur de la fenêtre glissante correspond à un nombre pair de pas d'observation, on procède à un deuxième pas de moyenne sur 2 valeurs pour recentrer les temps des données lissées sur les temps des observations. Afin d'utiliser toutes les observations disponibles en temps réel, on prolonge de

façon appropriée les données brutes au delà du dernier temps observé. La longueur de la fenêtre glissante et le pas d'observation sont rentrés dans le fichier .cas.

Pour éviter de prendre en compte des valeurs d'observation non représentables par le modèle, on rejète des observations en dessous d'un seuil minimal que l'on spécifie station par station dans le fichier .cas. Cette sélection est fondamentale en particulier pour le bassin de la Marne ou la tentative de se rapprocher de valeurs non représentables entraîne des instabilités du système avec la mise à sec de certaines sections.

Pour éviter de prendre en compte des valeurs d'observation aberrantes, on rejète les valeurs d'observation qui diffèrent excessivement de l'ébauche. La sélection s'effectue en comparant le misfit avec l'écart type d'erreur d'ébauche à la station d'observation. Si le misfit dépasse un multiple donné (dans le fichier .cas) de l'écart type d'erreur d'ébauche, l'observation n'est pas assimilée.

Détermination expérimentale des longueurs de corrélation des erreurs d'ébauche

Les longueurs de corrélation pour les erreurs d'ébauche ont été déterminées par expérimentation numérique au cours de la première moitié du stage de Chloé Ivanoff¹. La proportion optimale entre les longueurs de corrélation amont et aval a été estimée par Olivier Thual² avec des simulations Matlab sur un modèle conceptuel. Tous les détails se trouvent dans les documents qui décrivent ces travaux. Il est possible de rentrer dans le fichier .cas des longueurs de corrélation amont et aval différentes pour Z et Q.

Une partie importante du développement informatique dans les routines d'assimilation a été consacré au calcul des distances curvilignes entre une station d'observation et toutes les sections de calcul sur tous les biefs. Une procédure basée sur le balayage récursif du réseau a été mise en place dans le module d'assimilation. Pour le calcul des corrélations d'erreur en Z on procède par continuité d'un bief à l'autre. Pour les corrélations de Q, on assure la continuité des débits totaux par répartition additive proportionnelle aux débits d'ébauche, seulement vers les biefs qui contribuent au débit à la station d'observation.

Détermination de la correction sur Q (débit)

La correction sur Q doit être calculée à partir de la correction calculée sur la variable observée Z. Si l'on disposait de courbes de tarage valables en tout point et à tout instant on obtiendrait facilement cette relation. Dans la pratique, la deuxième moitié du stage de Chloé Ivanoff a été consacrée à la détermination par expérimentation numérique des relations qui lient les

1 Rapport "Assimilation de données pour un modèle d'hydrologie fluviale" du 07/09/2009

2 Note "Expériences jumelles sur un modèle des ondes de crues" du 18/01/2010

variation de Q aux variations de Z. Dans le cas d'un bassin où l'écoulement n'est pas influencé par la marée (e.g. Marne-Vallage), ces relations peuvent être exprimées comme puissances de Z de la forme $\delta Q = a(\delta Z)^r + b$ (une seule équation retenue par courbe de tarage). Les coefficients a, r, b sont spécifiés dans le fichier .cas, station par station. Malheureusement, l'influence de la marée empêche d'utiliser des relations aussi simples. L'étude a été ébauchée : on arrive à la détermination de relations représentées par des ellipses. Il faudrait donc utiliser la relation exprimée par le côté de l'ellipse où se situe le couple de valeurs d'ébauche (Z^b, Q^b). Pour le moment ceci n'a pas été fait et la relation entre δQ et δZ est de proportion directe, ce qui rentre dans le formalisme exponentiel avec $a=1, r=1, b=0$.

Implémentation

Toutes les routines d'assimilation (lecture des paramètres, traitement des observations, calculs d'assimilation) sont codées dans un module fortran 90 qui se trouve dans le fichier Assimilation/M_ASSIMILATION_S.f90.

Seulement les fichiers Pretrait/Pretrait.f90 et rezo/Rezo.f90 ont été modifiés pour insérer les appels aux routines d'assimilation.

Tous les paramètres utiles à l'assimilation sont rentrés dans le fichiers .cas : le code Mascaret avec assimilation linéique ne doit pas être recompilé pour redéfinir l'expérience (changement des paramètres ou changement de bassin).

Évolution du fichier .cas et du dictionnaire dico.txt

Le fichier .cas s'est enrichi d'un nombre important de lignes pour définir les paramètres de l'assimilation linéique. Le fichier dictionnaire dico.txt a évolué en parallèle. Voyons dans le détail les nouvelles lignes du fichier .cas avec exemples des valeurs tirés de la configuration Adour

ACTIVATION DE L'ASSIMILATION = VRAI

Switch principal de l'assimilation linéique. Si positionné à FAUX, on effectue un run direct Mascaret classique

TEMPS DE DEBUT DE L'ASSIMILATION = 172800

Temps de début de l'assimilation par rapport au début de la simulation (en secondes)

TEMPS DE FIN DE L'ASSIMILATION = 259200

Temps de fin de l'assimilation par rapport au début de la simulation (en secondes)

FREQUENCE D'ASSIMILATION = 300.0

Fréquence des assimilations qui ne sont pas nécessairement activées à chaque pas de temps du modèle (en secondes)

NOMBRE DE STATIONS DE MESURE = 3

BRANCHE DE LA STATION =

5;2;1

ABSCISSE DE LA STATION =

1361.0;10340.0;518.0

NOM DE LA STATION =

'Peyrehorade';'Urt';'Lesseps'

On décrit les stations de mesure avec le même format utilisé pour décrire les sections de stockage des résultats. On ajoute le nom en clair pour rendre plus lisibles les fichiers de listing.

DECALAGE DE LA STATION =

0.;0.;0.

Décalage de l'origine des observations par rapport à la côte maritime (en mètres)

SEUIL REJET OBSERVATIONS =

1.;1.;1.

Seuil de rejet des observations : valeur min en mètres

SEUIL REJET MISFIT ASSIM EN NOMBRE D'ECARTS TYPE = 8.0

Seuil de rejet des observations sur le misfit en nombre d'écart type

PREFIXE FICHIERS OBSERVATIONS = '.././Observations/Observations_stat'

Les noms complets des fichiers seront "\$PREFIXE\$NB_STAT.dat"

FREQUENCE DES DONNEES DANS LE FICHIER D'OBSERVATIONS = 300.

TAILLE DE LA FENETRE POUR LE LISSAGE DES OBS = 1500.

Paramètres pour le lissage des observations par moyenne glissante

ECART TYPE ERREUR D'OBSERVATION A LA STATION =

0.1;0.1;0.1

Racine de la diagonale de R (en mètres)

ECART TYPE ERREUR D'EBAUCHE A LA STATION =

0.3;0.25;0.25

Racine des éléments de B aux points d'observation (en mètres)

LONGUEUR DE PORTEE AMONT POUR Z =

20204.34;6613.27;34180.65

Longueur de portée de l'erreur d'ébauche pour Z en mètres (amont)

LONGUEUR DE PORTEE AVAL POUR Z =

2020.43;661.33;3418.06

Longueur de portée de l'erreur d'ébauche pour Z en mètres (aval)

LONGUEUR DE PORTEE AMONT POUR Q =

20204.34;6613.27;34180.65

Longueur de portée de l'erreur d'ébauche pour Q en mètres (amont)

LONGUEUR DE PORTEE AVAL POUR Q =

2020.43;661.33;3418.06

Longueur de portée de l'erreur d'ébauche pour Q en mètres (aval)

COEFF A COURBE DE TARAGE =

1.;1.;1.

COEFF R COURBE DE TARAGE =

1.;1.;1.

COEFF B COURBE DE TARAGE =

0.;0.;0.

Coefficients pour les courbes de tarage simplifiées de la forme $Q=A*Z^R+B$

Les lignes qui suivent décrivent la modalité de sauvegarde d'une ligne d'eau calculée pour en faire la condition initiale d'un run futur. Elles ne concernent pas seulement l'assimilation mais constituent un ajout au fichier .cas d'origine et pour cette raison on les copie ici :

```
ACTIVATION DE LA SAUVEGARDE LIGNE D'EAU INITIALE = VRAI  
TEMPS DE SAUVEGARDE LIGNE D'EAU INITIALE = 3600  
  Relatif au début de la simulation (en secondes)  
CHEMIN SAUVEGARDE LIGNE D'EAU INITIALE = '../../DonneesStat'
```

Script de pilotage de l'enchaînement des runs avec assimilation

La modalité de lancement des simulations dépend des centres d'exploitation. Nous n'avons pas repris les scripts de rapatriement de données et de lancement utilisés au SCHAPI. Nous avons plutôt codé un script de pilotage en korn shell qui permet de simuler un enchaînement de runs de type temps réel. Pour un temps de base donné, on utilise des débits amont observés pour la période d'équilibrage avant le temps de base et on prolonge les débits par valeurs constantes après le temps de base en modalité prévision. Les lois de marée ne sont pas coupées. Si des observations de hauteur d'eau sont disponibles avant le temps de base on peut les assimiler.

Le script modifie dynamiquement le fichier .cas pour imposer les temps de début et de fin de l'assimilation et pour régler les paramètres d'assimilation. Il gère aussi la modalité de sauvegarde de la ligne d'eau pour le redémarrage de la simulation.

Après chaque simulation il sauvegarde les résultats dans un répertoire différent par type de simulation (sans assimilation, avec assimilation linéique, ...) et calcule les diagnostics des performances. A la fin de l'enchaînement les résultats sont rendus graphiquement à l'écran et sauvegardés dans un fichier postscript.

Chaîne d'assimilation pour le calage en temps réel des lois aux extrémités libres

La deuxième chaîne d'assimilation implémente une méthode de type BLUE pour la correction dynamique des lois hydrauliques imposées aux extrémités libres à partir des observations aux stations de mesure sur un intervalle de temps précédant le temps de début de la prévision.

Encore une fois, les rappels algorithmiques et pratiques qui suivent n'ont pas pour but d'illustrer la méthode mais plutôt de mettre en évidence les paramètres de l'assimilation qui doivent être fixés et la façon de les estimer.

Principes algorithmiques

Par rapport à la chaîne d'assimilation linéique, ici nous ne voulons pas modifier progressivement l'état du système en y apportant systématiquement une correction additive. Il s'agit plutôt de se fixer un intervalle sur lequel comparer les résultats du modèle et les observations comme critère pour le calage de quelques paramètres du modèle.

Une étude de sensibilité et de stabilité a montré qu'avec les modèles pour les bassins Adour Maritime et Marne Vallage dans leur état actuel il est fort difficile de contrôler les coefficients de Strickler ou les conditions aux limites de type marée sans que ceci entraîne des instabilités du modèle. Ce résultat ne signifie pas que l'assimilation n'est pas utile pour ces types de calage, mais seulement qu'elle demande une plus étroite collaboration avec les développeurs du code de calcul et les modélisateurs du bassin pour être certains de bien respecter toutes les contraintes numériques et physiques qui assurent la stabilité du modèle et la pertinence des résultats. L'étude de sensibilité et de stabilité a par contre montré qu'il est possible de contrôler les lois imposées aux extrémités libres, ce que nous avons fait.

Après avoir fixé un contrôle représentatif de ces lois, il faut définir un opérateur d'observation généralisé \mathcal{G} qui relie les variations du contrôle aux variations des valeurs observées sur l'intervalle choisi. On utilisera ensuite sa version linéarisée au voisinage de l'ébauche \underline{G} . En définissant les matrices des variances/covariances d'erreur sur les observations (\underline{R}) et sur l'ébauche du contrôle (\underline{B}), le contrôle issu de l'assimilation sera

$$\underline{x}^a = \underline{x}^b + (\underline{B}^{-1} + \underline{G}^T \underline{R}^{-1} \underline{G})^{-1} \underline{G}^T \underline{R}^{-1} \underline{d}$$

où \underline{d} est le misfit sur l'intervalle choisi $\underline{d} = \underline{y}^o - \mathcal{A}(\underline{x}^b)$

Choix du contrôle

Si l'on voulait contrôler les valeurs des lois à tout instant on obtiendrait un vecteur de contrôle de taille égale au nombre des lois fois le nombre de pas de temps du modèle ce qui est décidément excessif. Pour cette raison nous avons restreint les transformations des lois à des transformations linéaires de type

$$\text{loi}_j(t) = a_j * \text{loi}_j(t - c_j) + b_j$$

Il s'agit donc de déterminer seulement au maximum 3 coefficients par loi. La taille du contrôle sera donc au plus 3 fois le nombre de lois.

Vecteur d'observations

Exactement comme pour l'assimilation linéique, on dispose d'observations de Z en correspondance de m stations de mesure (m=3 pour l'Adour, m=2 pour la Marne) à des instants qui ne coïncident pas nécessairement avec les pas de temps du modèle.

On utilise une approximation « plus proche voisin » : soit s_k l'abscisse

curviligne de la station k ($k=1,m$), nous faisons correspondre l'observation à la section i_k du modèle dont l'abscisse diffère le moins de s_k . De même, si t^o est un temps de mesure, nous faisons correspondre l'observation au pas de temps du modèle t_j le plus près de t^o . Pour cette chaîne nous prenons en compte toutes les observations sur un intervalle temporel fixé.

Mathématiquement les observations sont représentées par un vecteur \underline{y}^o de taille p égale à m fois le nombre de pas d'observation sur l'intervalle. Typiquement, sur le bassin de l'Adour, pour une expérience tirée de la chaîne temps réel nous avons $m=3$ stations et nous assimilons des observations avec un pas de 10 minutes sur une intervalle de 8 heures. La taille du vecteur d'observation est alors de $p = 3*48 = 144$.

Opérateur d'observation

L'opérateur d'observation non linéaire \mathcal{G} relie les éléments du vecteur de contrôle (donc un jeu de coefficients a_j, b_j, c_j) aux observations en correspondance des stations de mesure. Dans la pratique, appliquer l'opérateur \mathcal{G} revient à intégrer Mascaret sur toute la durée de la simulation avec lois aux extrémités libres déterminées par le contrôle et ensuite à sélectionner les valeurs de Z des sections où il y a une station de mesure, aux instants d'observation retenus pour l'assimilation.

Matrice \underline{R} des variances des erreurs d'observation

Les informations statistiques sur les erreurs d'observation sont contenues dans la matrice \underline{R} , de taille $p \times p$. Nous négligeons les corrélations d'erreur entre différentes stations et entre différents pas d'observation. La matrice \underline{R} est ainsi diagonale. Elle sera donc entièrement spécifiée si l'on indique les variances des erreurs pour chaque station. Dans la pratique on rentrera les écarts type dans le fichier .cas.

Matrice \underline{B} des variances/covariances des erreurs d'ébauche

Pour la matrice \underline{B} aussi nous négligeons les corrélations des erreurs sur les différents coefficients a, b, c . La matrice est donc diagonale et entièrement spécifiée si l'on indique les variances des erreurs pour chaque coefficient. Dans la pratique nous imposons les mêmes écarts type pour toutes les lois de type débit et d'autres pour toutes les lois de type côte (si l'on voulait par exemple contrôler la marée imposée).

Formule d'interpolation optimale

Puisque la taille du vecteur d'observation est nettement plus importante que la taille du contrôle, nous avons implémenté la formule d'interpolation optimale avec inversion dans l'espace de contrôle

$$\underline{x}^a = \underline{x}^b + (\underline{B}^{-1} + \underline{G}^T \underline{R}^{-1} \underline{G})^{-1} \underline{G}^T \underline{R}^{-1} \underline{d}$$

où d est le misfit sur l'intervalle choisi $\underline{d} = \underline{y}^o - \mathcal{A}(\underline{x}^b)$.

L'inversion de \underline{B} et \underline{R} n'est pas coûteuse car elles sont diagonales.

\underline{G} et \underline{G}^T indiquent l'opérateur linéaire tangent de \mathcal{A} et son adjoint, ou la matrice qui représente l'opérateur linéaire tangent et sa transposée. Puisque \mathcal{A} implique une intégration de Mascaret, et que nous ne disposons ni du modèle linéaire tangent ni du modèle adjoint de Mascaret, il faudra les construire par approximation.

Lissage et seuillage des observations

Le chargement des observations est partagé avec la chaîne d'assimilation linéique. Les observations sont lues avec une fréquence donnée. Elles sont ensuite lissées par moyenne glissante des observations, avec un deuxième pas de moyenne sur 2 valeurs si la longueur de la fenêtre glissante correspond à un nombre pair de pas d'observation et avec une correction sur les derniers pas afin de pouvoir utiliser toutes les observations disponibles en temps réel jusqu'au temps de base.

Pour éviter de prendre en compte des valeurs d'observation non représentables par le modèle, on rejette des observations en dessous d'un seuil minimal que l'on spécifie station par station dans le fichier .cas.

Pour éviter de prendre en compte des valeurs d'observation aberrantes, on rejette les valeurs d'observation qui diffèrent excessivement de l'ébauche. La sélection s'effectue en comparant le misfit avec la valeur de l'ébauche à la station d'observation. Si le misfit dépasse un pourcentage (spécifié dans le fichier .cas) de la valeur de l'ébauche, l'observation n'est pas assimilée.

Détermination de \underline{G} et \underline{G}^T par différences finies

Pour construire l'approximation de la matrice \underline{G} , on procède par différences finies. Chaque colonne de \underline{G} exprime la dépendance au premier ordre des variations des valeurs observées en fonction des variations du coefficient a_j ou b_j ou c_j correspondant à la colonne.

Si nous indiquons par \underline{y}^b le vecteur d'observations associé à l'ébauche $\underline{x}^b = [\{a_j^b, b_j^b, c_j^b\}]$ et par \underline{y}^k le vecteur associé au vecteur de contrôle qui coïncide avec l'ébauche sauf pour un élément qui a été perturbé $x_k = x_k^b + \delta x_k$, la colonne k de \underline{G} sera $(\underline{y}^k - \underline{y}^b) / \delta x_k$.

Pour estimer la matrice \underline{G} il faudra donc effectuer un seul run de référence pour l'ébauche et autant de runs perturbés que de coefficients à contrôler.

Le run de référence fourni aussi le vecteur $\underline{y}^b = \mathcal{A}(\{a_j^b, b_j^b, c_j^b\})$ pour le calcul du misfit \underline{d} .

Afin d'éviter des plantages de Mascaret (sections à sec ou passage en régime

torrentiel) on impose des valeurs min et max pour les lois après correction. Avec l'implémentation actuelle ce seuillage intervient à la lecture des lois. Ceci a deux conséquences. La première est que le run d'ébauche (qui est aussi la référence pour le calcul de \underline{G}) ne coïncide pas avec un run du modèle direct seul, où le seuillage ne s'applique pas. La deuxième est que l'estimation par différences finies n'est pas compatible avec une opération complètement non linéaire, tel le seuillage. En effet, si le seuil s'applique pour l'ébauche et/ou pour le run perturbé, la différence finie sera faussée, car l'effet de la perturbation est caché par le seuillage. Ce stratagème, qui n'est pas en accord avec le principe de fonctionnement d'un modèle de prévision des crues, ne sera plus nécessaire quand on passera à un noyau de Mascaret plus robuste, ce qui est prévu pour un futur proche.

Implémentation avec le coupleur PALM

Cette chaîne d'assimilation demande d'effectuer un nombre important de runs Mascaret et d'en manipuler les résultats. Au lieu de gérer les exécutions par script et de récupérer les résultats dans des fichiers nous avons opté pour l'utilisation d'un coupleur. En particulier nous avons choisi le coupleur dynamique PALM développé au CERFACS dont on peut trouver une description détaillée sur le site http://www.cerfacs.fr/globc/PALM_WEB.

Ce choix permet une plus grande efficacité au niveau de l'échange des données (effectué par message passing MPI) mais surtout il permet d'exploiter de façon simple et automatique le parallélisme intrinsèque de cet algorithme : lors de l'approximation de \underline{G} , tous les runs de Mascaret sont indépendants les uns des autres. Sur une machine multiprocesseur on pourra donc exécuter les runs perturbés en parallèle. Tous simultanément si la machine a assez de processeurs ou par lots. Le coupleur PALM permet de gérer dynamiquement cet aspect.

Les calculs d'assimilation feront partie de l'application couplée et ils seront suivis d'une dernière intégration de Mascaret avec les lois modifiées.

L'application couplée est assemblée dans l'interface graphique PrePALM. Le canevas de l'application est reproduit à la fin de ce document.

Unités pour l'assimilation

Une partie des procédures d'assimilation sont codées en dehors des sources Mascaret dans des codes qui sont couplés à Mascaret sous forme d'unités PALM.

En particulier l'unité `read_damocles` lit le fichier `.cas` pour récupérer les paramètres qui configurent l'expérience d'assimilation (modalités et dimensionnements).

L'unité `H_linear` construit la matrice \underline{G} à partir des résultats des runs perturbés.

L'unité `blue` construit et inverse les matrices de la formule du BLUE et produit

les valeurs des coefficients a_j , b_j , c_j analysés.

Des scripts gnuplot commandés par PALM sont activés pour l'affichage des séries temporelles de valeurs de Z aux stations d'observation avant et après assimilation si la clé TRACER RESULTATS est positionnée à VRAI dans le fichier .cas.

Modification des sources MASCARET

L'utilisation d'un coupleur permet de réduire au minimum les interventions dans les sources de Mascaret. Cependant il a fallu introduire les appels aux primitives de couplage et quelques traitements spécifiques, en particulier pour la prise en compte des coefficients a_j , b_j , c_j , lors de la lecture des lois aux extrémités libres.

Pour garantir la possibilité de compiler le même jeu de sources pour des applications qui n'utilisent pas le couplage (runs du modèle direct seul ou avec assimilation linéique), toutes les interventions relative à l'utilisation de PALM et au calage des lois sont sous clé CPP CALAGE_PALM.

Ces modifications affectent les fichiers :

SUPERVISEUR/SUPERVISEUR.f90 (interfaces avec le coupleur, gestion des arrêts sur erreur et des sorties)

Pretrait/Pretrait.f90 (interfaces avec le coupleur, lecture des paramètres de couplage dans le fichier .cas)

Pretrait/LEC_HYDRAU.f90 et son interface modules/M_LEC_HYDRAU_I.f90 (interfaces avec le coupleur, modification des lois après lecture)

Pretrait/LEC_LOI.f90 (prise en compte de la nouvelle interface de LEC_HYDRAU)

Pretrait/LEC_LIGNE.f90 (gestion des sorties)

rezo/Rezo.f90 (envoi des résultats vers PALM et différenciation des cas avec ou sans assimilation linéique)

Estimation des variances

La matrice \underline{B} est diagonale. Elle est entièrement spécifiée par les écarts type associés aux coefficients a_j , b_j , c_j des lois imposées aux extrémités libres. Pour le moment les valeurs utilisées sont issues d'une estimation statistique de la variance des valeurs des lois utilisées dans les cas de calage. Des méthodes classiques d'estimation des variances par comparaison aux observations avant et après assimilation permettraient d'affiner ces estimations.

Évolution du fichier .cas et du dictionnaire dico.txt

Le fichier .cas s'est enrichi de quelques lignes supplémentaires pour définir les paramètres du calage des lois. Le fichier dictionnaire dico.txt a évolué en parallèle. Voyons dans le détail les nouvelles lignes du fichier .cas avec exemples de valeurs tirés de la configuration Adour. A noter que les stations d'observation sont décrites exactement comme pour l'assimilation linéique.

```
LOI 1 CALAGE = 'abc'
LOI 1 STATIONS CALAGE = FAUX;FAUX;VRAI
LOI 1 VALEUR MIN = 6.
```

/

```
LOI 2 CALAGE = 'abc'
LOI 2 STATIONS CALAGE = VRAI;VRAI;VRAI
LOI 2 VALEUR MIN = 0.
```

/

```
LOI 3 CALAGE = 'abc'
LOI 3 STATIONS CALAGE = VRAI;VRAI;VRAI
LOI 3 VALEUR MIN = 18.
```

/

```
LOI 4 CALAGE = 'abc'
LOI 4 STATIONS CALAGE = VRAI;VRAI;VRAI
LOI 4 VALEUR MIN = 4.
LOI 4 VALEUR MAX = 660.
```

/

```
LOI 5 CALAGE = ""
LOI 5 STATIONS CALAGE = FAUX;VRAI;VRAI
LOI 5 VALEUR MIN = 1.4
```

Pour chaque loi imposée comme condition limite, on indique quels coefficients on cale par assimilation. Si aucun paramètre n'est indiqué (comme pour la loi 5), la loi n'est pas calée. On indique quelles stations d'observation contribuent au calage de chaque loi (ici dans le cas de la loi 1, seulement les observations de la troisième station – Lesseps) sont pris en compte. On indique s'il faut seuiller les lois modifiées en imposant une valeur min et une valeur max (optionnelles).

```
TRACER RESULTATS = FAUX
```

Il est possible d'afficher l'évolution temporelle de la hauteur prévue aux stations de mesure avant et après le calage par assimilation. Pratique pour la mise au point, mais déconseillé en cas d'enchaînement automatique.

```
TEMPS DE DEBUT DU CALAGE PAR ASSIMILATION = 230400
```

```
TEMPS DE FIN DU CALAGE PAR ASSIMILATION = 259200
```

Délimitation (en secondes par rapport au temps de début de la simulation) de l'intervalle temporel où l'on calcule l'écart aux observations pour le calage par assimilation.

```
FREQUENCE DES OBS POUR LE CALAGE = 600.0
```

Fréquence temporelle des observations prises en compte pour le calage.

```
NOMBRE DE STATIONS DE MESURE = 3
```

BRANCHE DE LA STATION =
5;2;1

ABSCISSE DE LA STATION =
1361.0;10340.0;518.0

NOM DE LA STATION =
'Peyrehorade';'Urt';'Lesseps'

On décrit les stations de mesure avec le même format utilisé pour décrire les sections de stockage des résultats. On ajoute le nom en clair pour rendre plus lisibles les fichiers de listing.

DECALAGE DE LA STATION =
0.;0.;0.

Décalage de l'origine des observations par rapport à la côte maritime (en mètres)

SEUIL REJET OBSERVATIONS =
1.;1.;1.

Seuil de rejet des observations : valeur min en mètres

PERTURBATION PARAMETRE A DEBIT = 1.1

Perturbation du paramètre a pour les lois de type débit lors du calcul des différences finies (adimensionnel)

PERTURBATION PARAMETRE B DEBIT = 6

Perturbation du paramètre b pour les lois de type débit lors du calcul des différences finies (m³/sec)

PERTURBATION PARAMETRE C DEBIT = 7200

Perturbation du paramètre c pour les lois de type débit lors du calcul des différences finies (sec)

PERTURBATION PARAMETRE A COTE = 1.1

Perturbation du paramètre a pour les lois de type côte lors du calcul des différences finies (adimensionnel)

PERTURBATION PARAMETRE B COTE = 0.5

Perturbation du paramètre b pour les lois de type côte lors du calcul des différences finies (m)

PERTURBATION PARAMETRE C COTE = 7200

Perturbation du paramètre c pour les lois de type débit lors du calcul des différences finies (sec)

ECART TYPE PARAMETRE A DEBIT = 0.2

Ecart type d'erreur sur le paramètre a pour les lois de type débit (adim.)

ECART TYPE PARAMETRE B DEBIT = 6

Ecart type d'erreur sur le paramètre b pour les lois de type débit (m³/sec)

ECART TYPE PARAMETRE C DEBIT = 7200

Ecart type d'erreur sur le paramètre c pour les lois de type débit (sec)

ECART TYPE PARAMETRE A COTE = 0.2

Ecart type d'erreur sur le paramètre a pour les lois de type côte (adim.)

ECART TYPE PARAMETRE B COTE = 0.5

Ecart type d'erreur sur le paramètre b pour les lois de type côte (m)

ECART TYPE PARAMETRE C COTE = 7200

Ecart type d'erreur sur le paramètre c pour les lois de type côte (sec)

SEUIL REJET MISFIT CALAGE = 50.

Seuil de rejet des observations sur le misfit en pourcentage de l'ébauche

Script de pilotage de l'enchaînement des runs avec assimilation

Le script de pilotage en korn shell qui permet de simuler un enchaînement de runs de type temps réel gère aussi le lancement du cas de calage des lois.

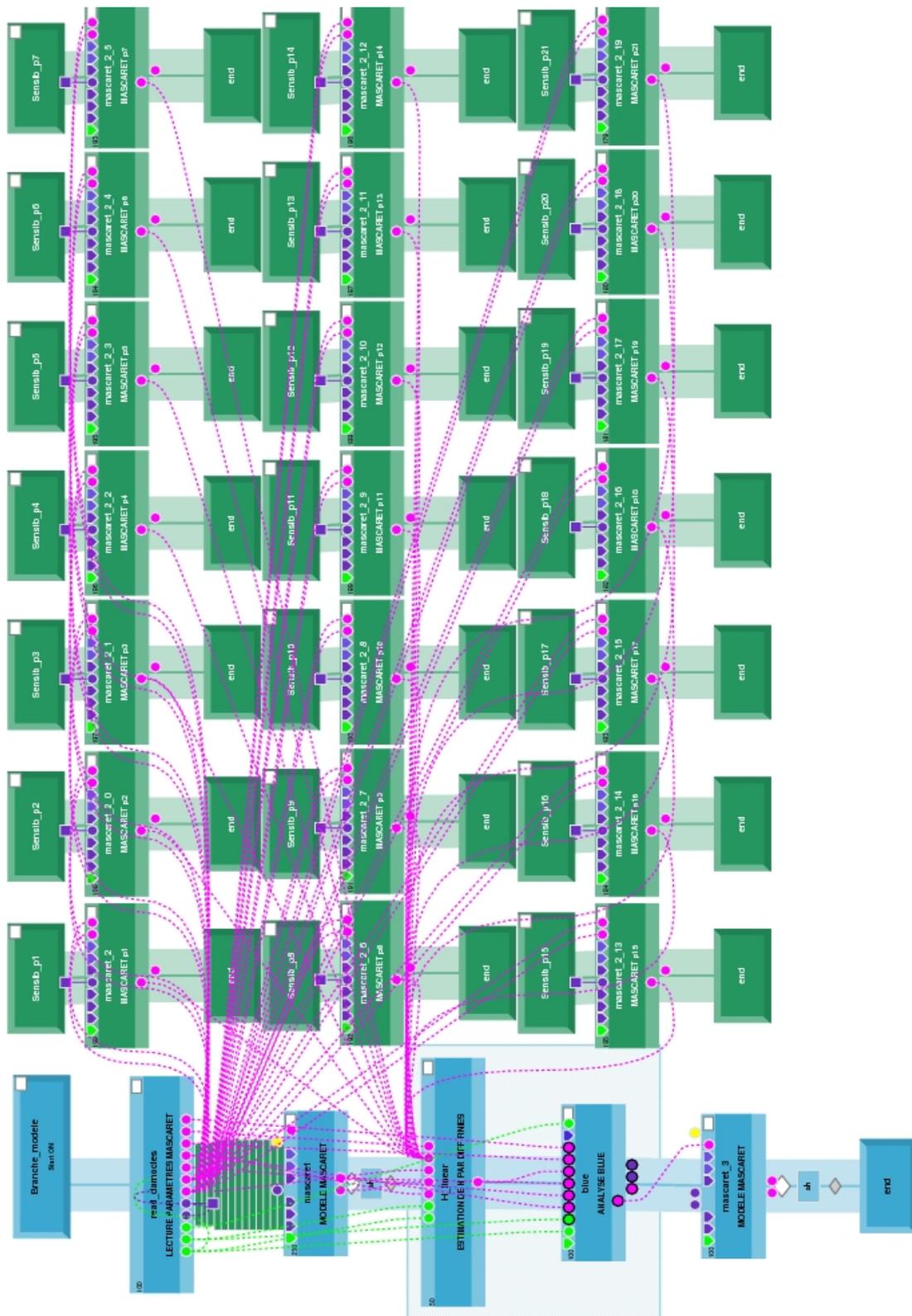
Sur PC multiprocesseur il démarre l'environnement de calcul parallèle MPI (LAM) et exécute automatiquement l'application sur le nombre de processeurs disponibles sur la machine.

Après chaque simulation il sauvegarde les résultats dans un répertoire différent par type de simulation (sans assimilation, avec assimilation linéique, ...) et calcule les diagnostics des performances. A la fin de l'enchaînement les résultats sont rendus graphiquement à l'écran et sauvegardés dans un fichier postscript.

Chaîne avec les deux types d'assimilation

Il s'agit tout simplement d'activer l'étape d'assimilation linéique au cours du run final après le calage des lois aux extrémités libres. Puisque l'implémentation des deux chaînes est compatible, tout se fait dans une seule exécution de la chaîne sous PALM. Il suffit de demander l'activation de l'assimilation linéique dans le fichier .cas.

Ce cas aussi est piloté par le script korn shell. Il modifie le fichier .cas pour mettre à VRAI la ligne qui active l'assimilation linéique. Il modifie aussi les variances d'erreur d'ébauche sur la ligne d'eau. Par rapport au cas où l'on effectue seulement l'assimilation linéique on divise les écarts type par deux pour prendre en compte la correction déjà apportée sur l'ébauche par le calage des lois.



Canevas PrePALM pour le calage des lois aux extrémités libres

Rapport du contrat entre le SCHAPI et le CERFACS sur l'application de l'assimilation de données à l'hydraulique

Validation de la maquette d'assimilation de données pour le modèle MASCARET

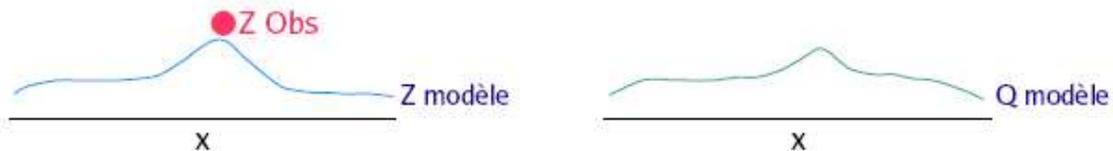
G. Jonville, A. Piacentini, O. Thual, C. Ivanoff, E. Le Pape

30 novembre 2009

Contexte

Dans le cadre de la convention passée entre le SCHAPI et le CERFACS sur l'application de l'assimilation de données à l'hydraulique, le CERFACS a réalisé une maquette de chaîne d'assimilation de données pour le code hydraulique MASCARET.

Cette chaîne implémente deux algorithmes de type BLUE (Best Linear Unbiased Estimate) pour l'assimilation de données de hauteur d'eau à deux stations pour le réseau hydraulique de Marne Vallage et à trois stations pour le réseau hydraulique Adour Maritime. Le premier algorithme, nommé assimilation linéique, a pour objectif d'améliorer directement la ligne d'eau simulée en agissant aux stations de mesure et leurs voisinages.



Le deuxième algorithme, nommé assimilation en calage, a pour objectif d'améliorer la ligne d'eau simulée en perturbant les hydrogrammes imposés aux stations limitantes amonts par calage de trois paramètres perturbateurs (a,b,c) pour chaque hydrogramme. L'algorithme est basé sur la linéarisation de la dépendance de la ligne d'eau par rapport aux paramètres (a,b,c) de chaque station amont. La linéarisation est calculée par différences finies après perturbation légère des valeurs initiales de (a,b,c). Les deux algorithmes peuvent être combinés afin de tirer parti simultanément de leurs bénéfices.

Un protocole a été défini pour la validation lors de la réunion du 19 juin 2009.

Protocole

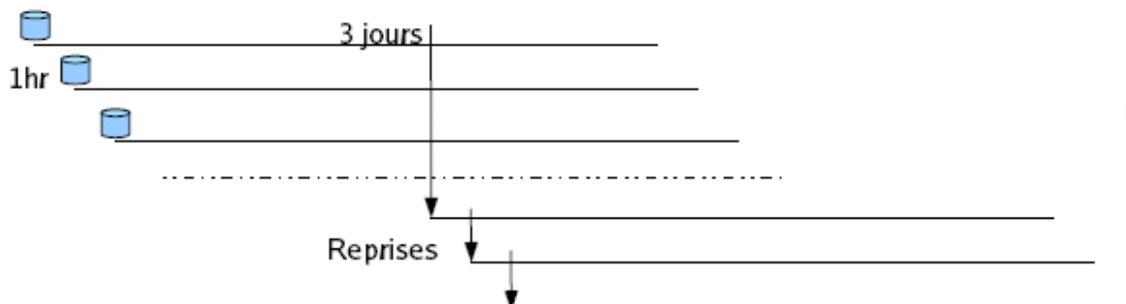
Des épisodes de crue pour les réseaux Marne Vallage et Adour Maritime ont été sélectionnés par Etienne Le Pape.

Pour le réseau Marne Vallage, 7 épisodes de novembre 2002 à mars 2007 ont un pas de temps d'observation de 1h aux stations de mesure.

Pour le réseau Adour Maritime, 8 épisodes de crue ont un pas de temps d'observation de 1h aux stations de mesure, et 2 épisodes de avril 2009 et novembre 2009 ont un pas de temps d'observation de 5mn aux stations de mesure.

Pour chaque épisode de crue, le principe de simulation de la chaîne temps réel du SCHAPI a été mis en œuvre avec la même structure de données. La première simulation s'effectue sur un intervalle de temps autour du premier temps de base T (ou temps présent) précédant suffisamment le pic de crue. Chaque simulation suivante décale cet intervalle de temps et le temps de base de +1h.

La ligne d'eau initiale est celle de la chaîne temps réel pour les premières simulations. Cependant pour chaque simulation, une ligne d'eau calculée peut être écrite dans un fichier à un temps donné grâce à un développement dans MASCARET réalisé dans le cadre de cette convention. Cette ligne d'eau est ensuite utilisée pour initialiser une simulation ultérieure. Ce mode opératoire rend plus cohérent la ligne d'eau initiale avec les conditions imposées aux stations limitantes et permet de réduire le temps d'équilibrage du modèle (à 3 jours pour l'Adour au lieu de 5) et d'augmenter la robustesse du modèle en évitant les plantages intempestifs (mise à sec d'un bief).



Reprise de ligne d'eau.

Les conditions imposées aux stations limitantes sont des débits obtenus de hauteurs d'eau mesurées et transformées par une courbe de tarage. A chaque simulation, l'hydrogramme aux stations limitantes amonts est égal aux débits mesurés jusqu'au temps de base T, puis est un prolongement continu et constant sur la période de prévision.

Les observations aux stations de mesure sur le réseau sont des hauteurs d'eau mesurées à l'échelle.

Des critères d'évaluation ont été sélectionnés pour évaluer la qualité du modèle et quantifier le gain apporté par les algorithmes d'assimilation de données. Les critères calculés sur le passé (période précédant le temps de base T) sont le biais et la précision. Les critères calculés sur la prévision sont le biais, la précision, l'efficacité et la persistance. Une comparaison graphique est aussi produite entre les hauteurs mesurées et les hauteurs prévues à échéance fixe.

Réalisations

Les simulations de chaque épisode de crue ont été réalisées à l'aide d'un pilote qui permet d'enchaîner des simulations MASCARET avec glissement de la fenêtre temporelle.

Ce pilote est un script développé durant cette convention, nommé piloteMV ou piloteAM selon le réseau Marne Vallage ou Adour Maritime. A partir d'un temps de base initial, d'un temps de base final et d'une fréquence de déplacement, le pilote lance successivement les simulations MASCARET sur la période allant du temps de base courant moins la durée d'équilibrage au temps de base courant plus la durée de prévision. Pour ce faire, le pilote va avant chaque simulation préparer les fichiers d'entrée :

- choisir la ligne d'eau initiale standard ou une ligne d'eau précédemment calculée
- modifier le fichier de paramètres Mascaret pour les paramètres temporels temps initial et temps maximum, le temps de sauvegarde de la ligne d'eau, les temps de début et de fin de l'assimilation linéique, les écarts type d'erreur d'ébauche à la station et les temps de début et de fin du calage par assimilation,
- couper les fichiers des lois hydrodynamiques amonts au temps de base courant et compléter par le dernier débit jusqu'à la fin de prévision.

Puis le pilote va exécuter le système Mascaret.

Le pilote stocke ensuite pour chaque exécution le fichier de résultats au format Ophyca et lance le calcul des critères d'évaluation du modèle sur la période simulée. La valeur de chaque critère est calculée en chaque station de mesure à partir du fichier Résultats et du fichier Observations. Les valeurs des critères associées au temps de base courant sont écrites dans un fichier Critères, unique pour toutes les exécutions Mascaret.

Après la dernière exécution Mascaret, le pilote trace en chaque station d'observation des représentations graphiques des hauteurs mesurées avec les critères d'évaluation du modèle et des hauteurs mesurées avec les hauteurs prévues à différentes échéances.

Pour chaque épisode de crue, nous avons ainsi réalisé à l'aide du pilote plusieurs simulations par MASCARET et avec différents modes de simulation afin de comparer les algorithmes d'assimilation. Le mode DRYRUN correspond à des simulations du modèle Mascaret direct. Le mode ASSIM lance Mascaret avec assimilation linéique. Le mode CALAGE lance Mascaret avec calage par assimilation des paramètres sur les lois hydrodynamiques. Le mode CALAGEASSIM combine les deux modes précédents.

Les résultats sont présentés sous forme de limnigrammes avec deux axes d'ordonnées, le premier pour les hauteurs, le deuxième pour les critères.

La première série de simulations en mode DRYRUN a fourni les "Hauteur modèle". L'autre série de simulations a fourni les "Hauteur assim". Elle est en général en mode CALAGEASSIM, sauf pour les cas où le modèle direct plante l'algorithme de l'assimilation par calage ne peut pas fonctionner. Par contre pour ces cas, l'assimilation linéique permet souvent de simuler l'épisode de crue complet. Le mode de simulation ASSIM est alors présenté.

Pour le réseau Marne Vallage, les résultats aux stations Joinville et Chamouilley sont présentés pour chaque épisode.

Pour le réseau Adour Maritime, les résultats aux stations Peyrehorade et Urt sont présentés pour chaque épisode.

Pour chaque épisode et pour chaque station, sont présentés :

1- un graphique de la précision du modèle sur la période passé, dit en mode simulation. Cette période précédant le temps de base est fixée à 48h pour la Marne et à 24h pour l'Adour. La légende est (-48h) ou (-24h). Chaque valeur de précision est issue d'une simulation Mascaret.

2- un graphique de précision du modèle sur la période de prévision usuelle de la station, dit en mode prévision. La légende est (+temps d'échéance usuel). Chaque valeur de précision est issue d'une simulation Mascaret.

3- un graphique des hauteurs prévues à l'échéance usuelle de la station. Chaque hauteur est le résultat d'une simulation Mascaret.

Les échéances usuelles des stations Joinville et Chamouilley sont respectivement 7h et 15h. Les échéances usuelles des stations Peyrehorade et Urt sont 6h et 10h.

Sur chaque graphique, la courbe en bleu indique les hauteurs observées. La courbe en pointillée indique les hauteurs observées décalées de la durée de la période d'étude, vers l'avant si la période concerne le passé, vers l'arrière si la période concerne la prévision (lag +/- temps). Elle permet de repérer précisément sur la courbe bleue pour un temps de base donné la portion de courbe des hauteurs observées sur laquelle est calculée la précision : entre les deux points à l'aplomb du temps de base.

La courbe en noire indique les précisions ou les hauteurs du modèle direct.

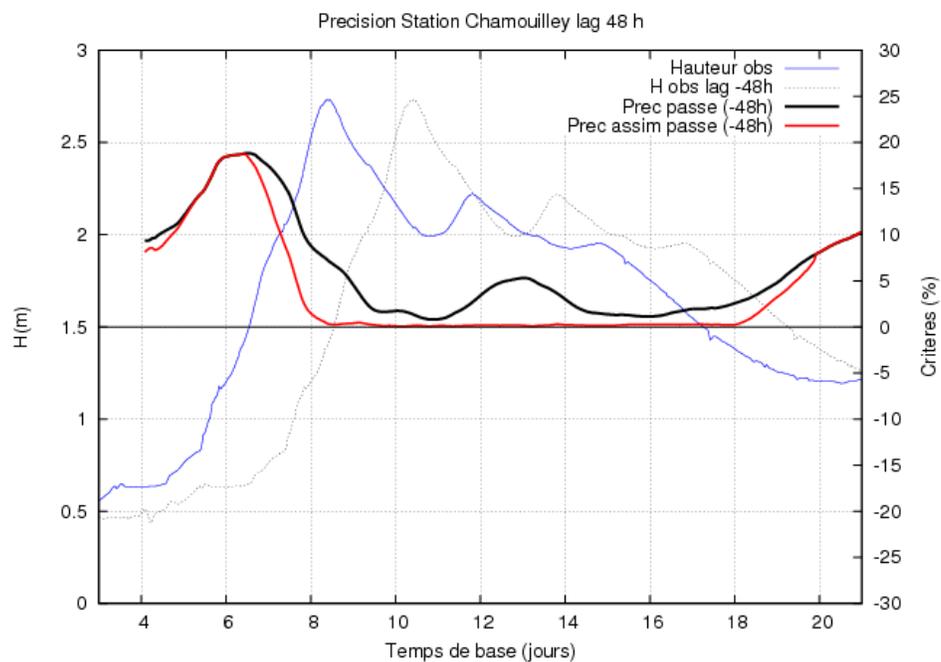
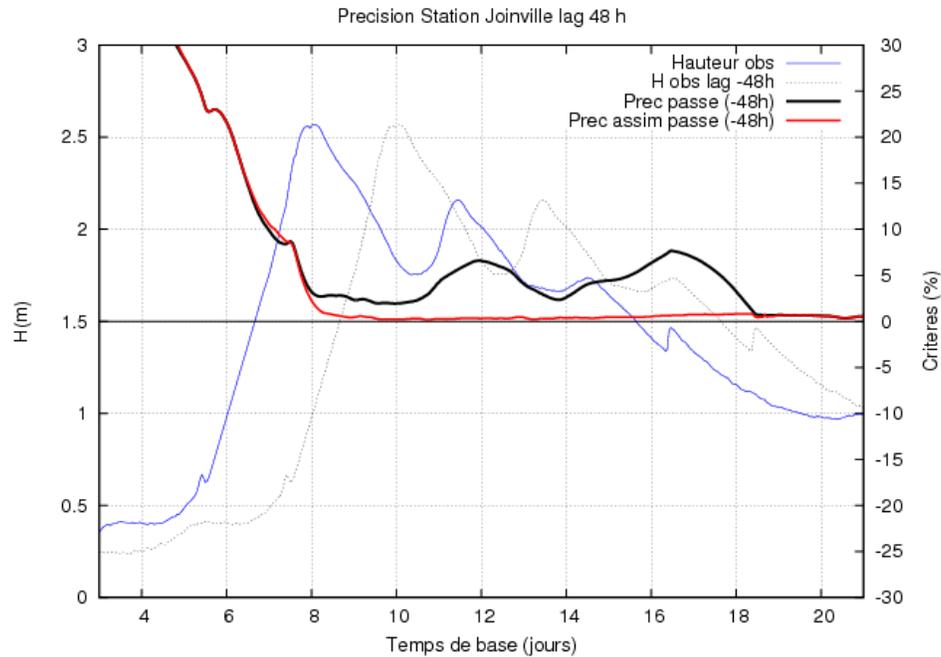
La courbe en rouge indique les précisions ou les hauteurs du modèle avec assimilation.

Episode MARNE VALLAGE 01 : janvier 2004

L'assimilation combine le calage et l'assimilation linéique (CALAGEASSIM).

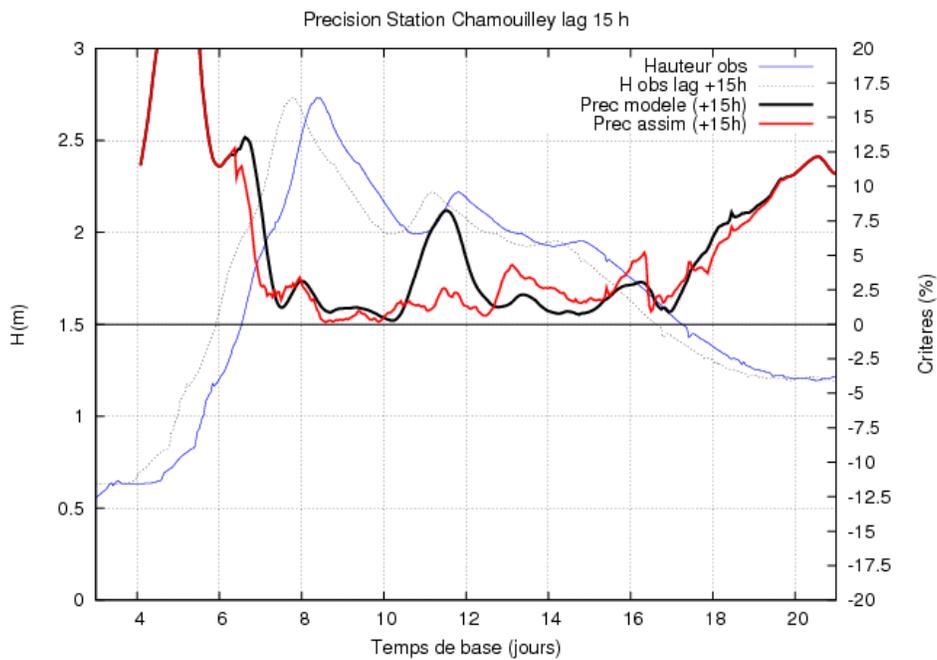
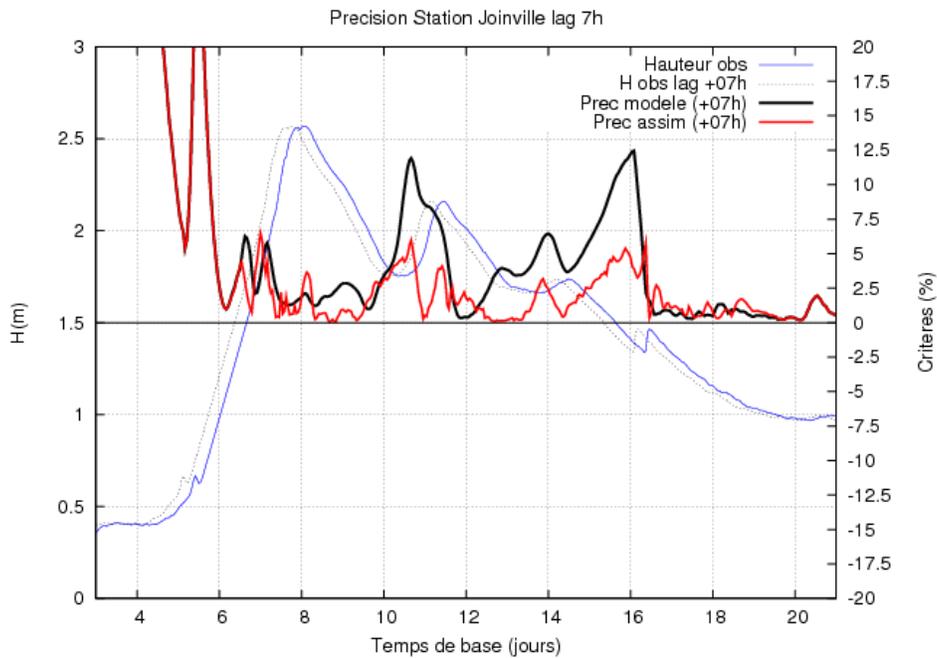
1- Précisions en mode simulation

La précision sur le passé est améliorée par l'assimilation à partir de douze heures avant le pic de crue à Joinville, et à partir de deux jours avant le pic de crue à Chamouilley.



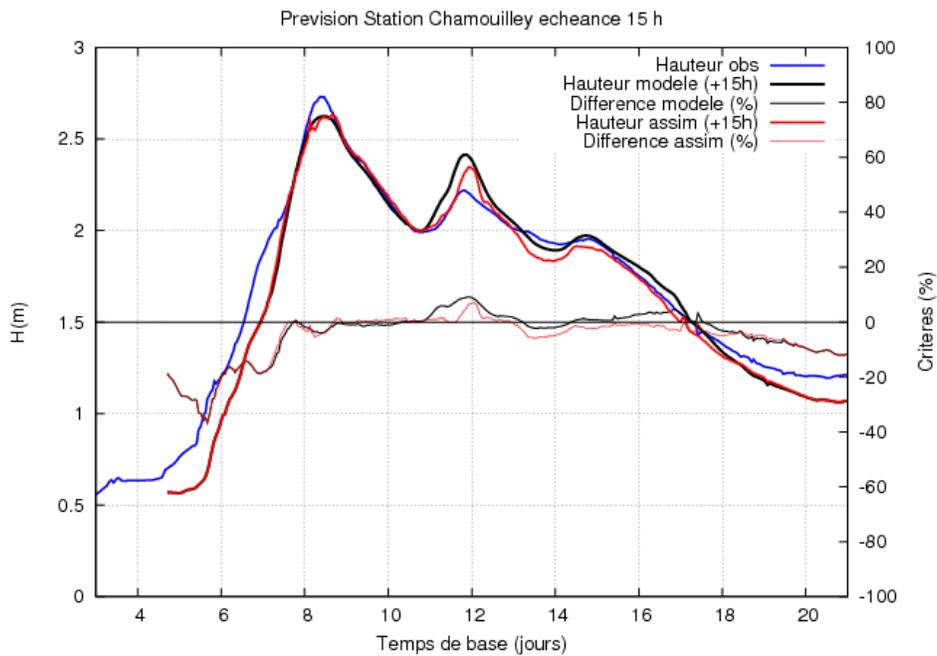
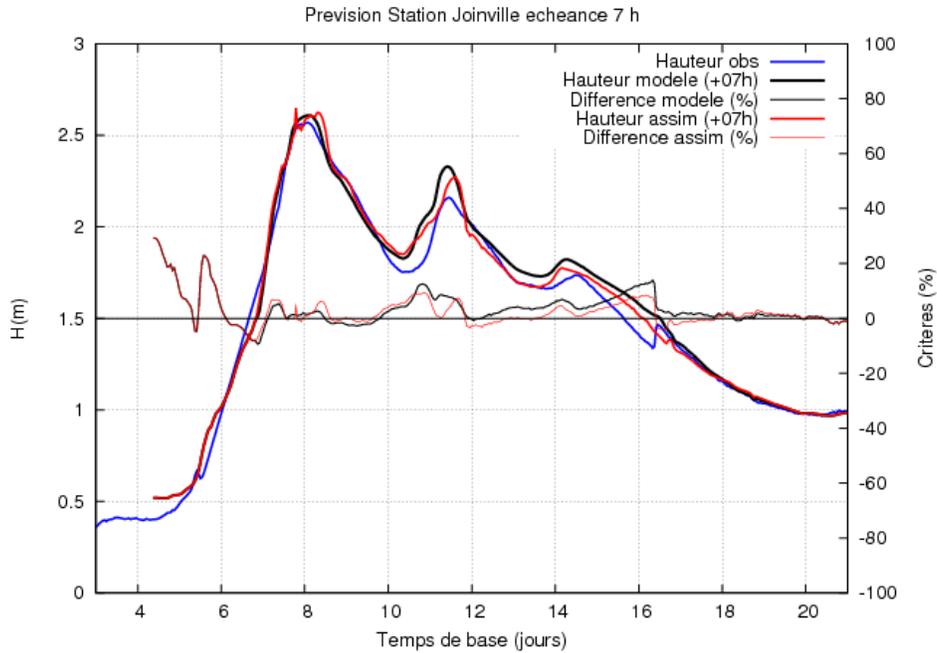
2- Précisions en mode prévision

La précision sur la prévision est légèrement améliorée par l'assimilation aux deux stations.



3- Hauteurs prévues à l'échéance usuelle

Les prévisions de hauteurs sont quasiment semblables sur l'épisode.



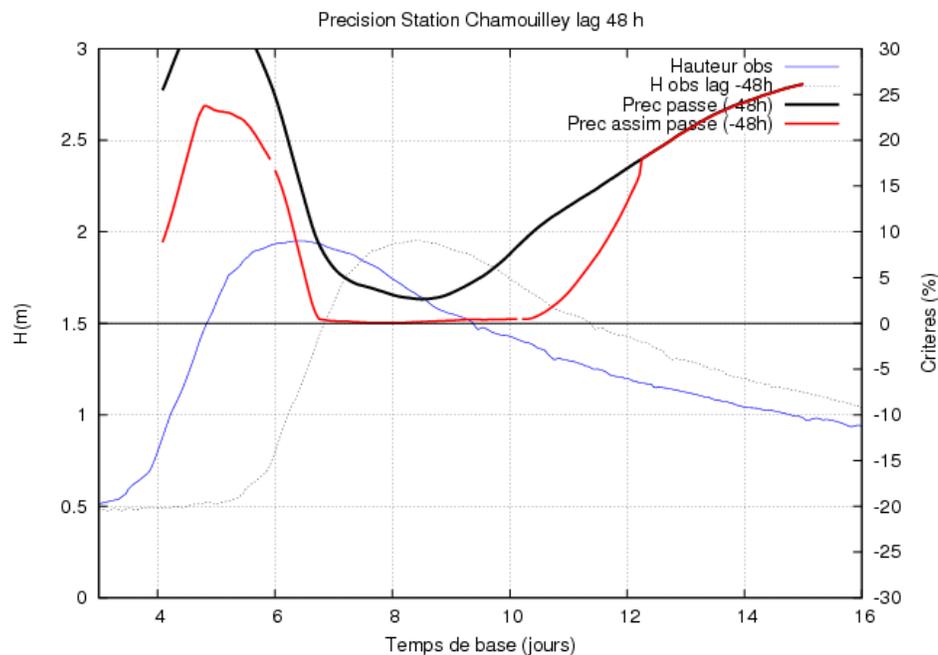
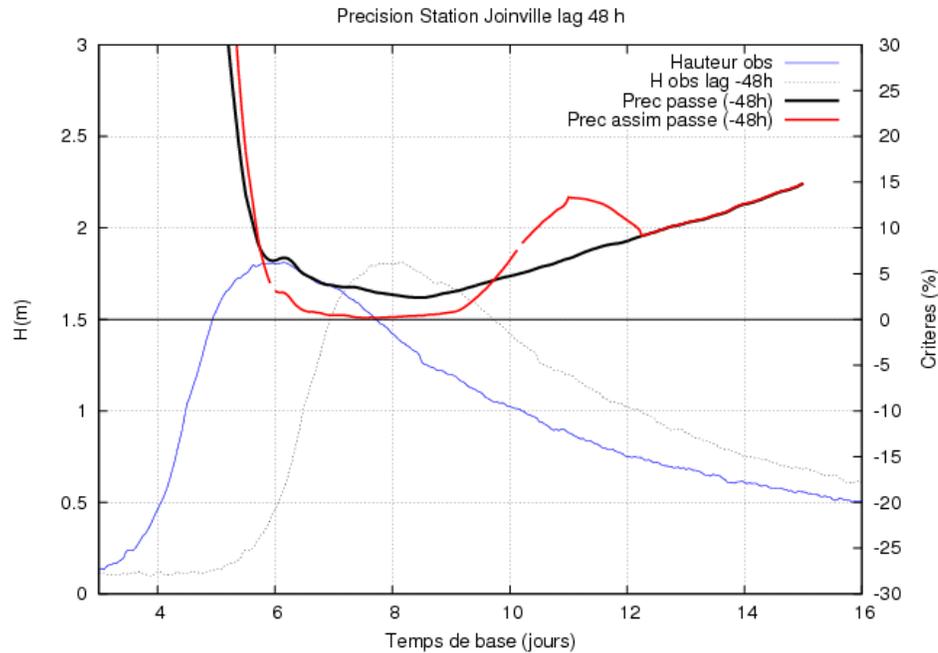
L'assimilation en mode CALAGEASSIM a apporté peu d'amélioration au modèle.

Episode MARNE VALLAGE 02 : février 2006

Le modèle direct fonctionne seulement avec des lignes d'eau initiales pré-simulées.
L'assimilation combine le calage et l'assimilation linéique (CALAGEASSIM).

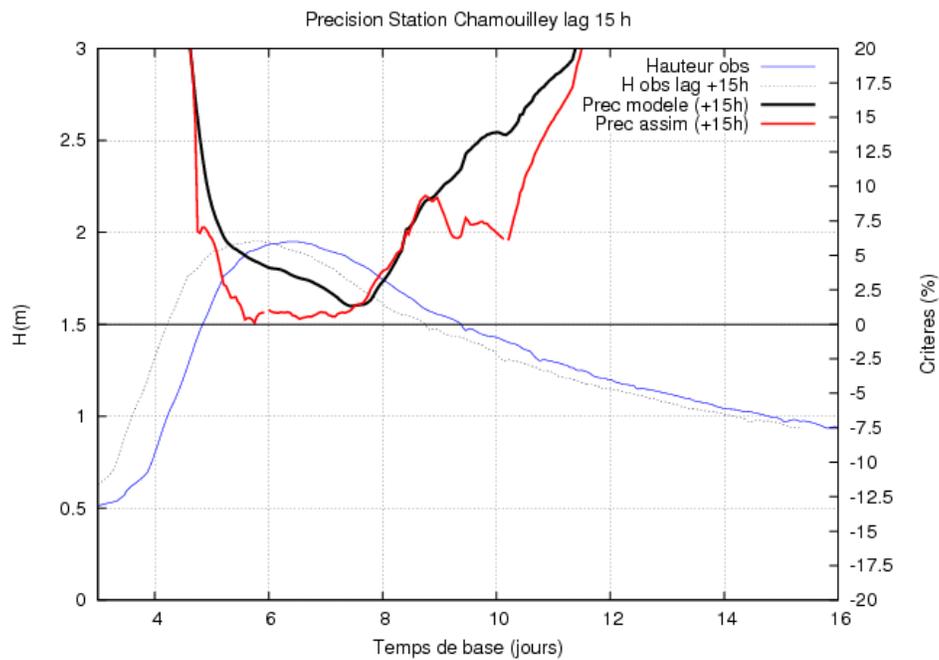
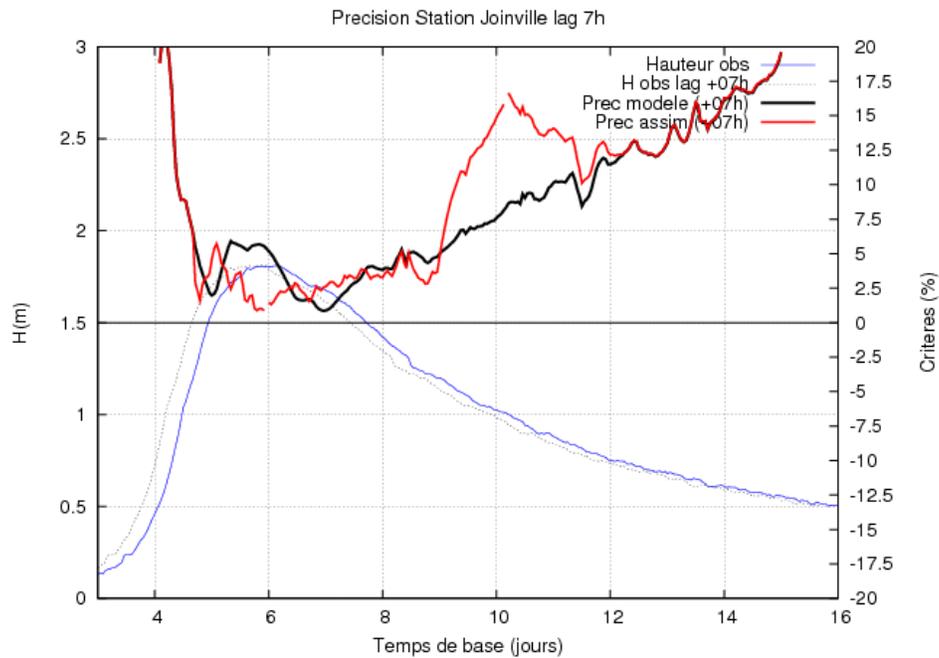
1- Précisions en mode simulation

La précision sur le passé est améliorée par l'assimilation à partir de 6 heures avant le pic de crue à Joinville, et sur tout l'épisode à Chamouilley.



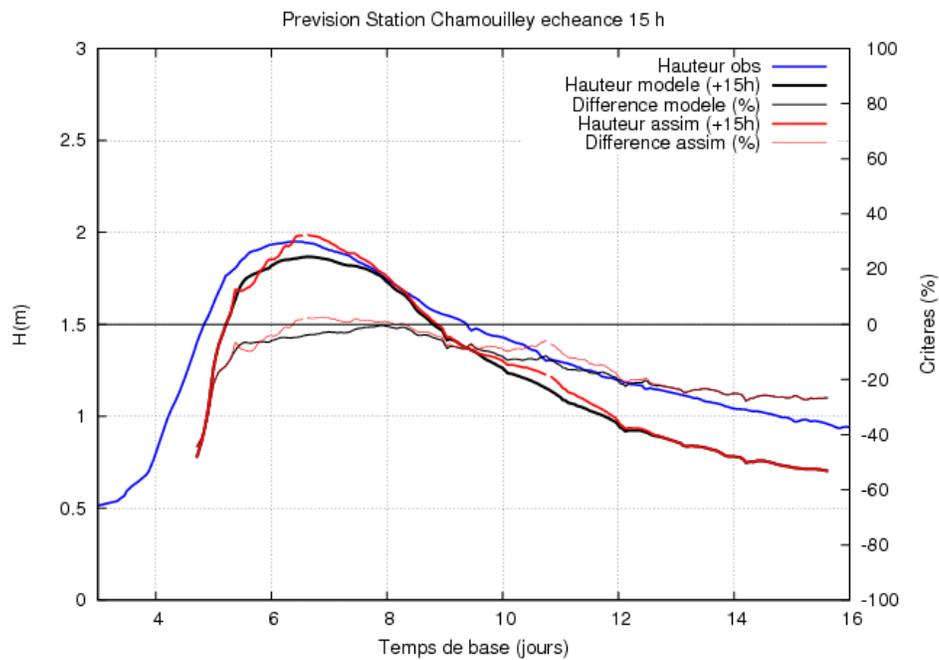
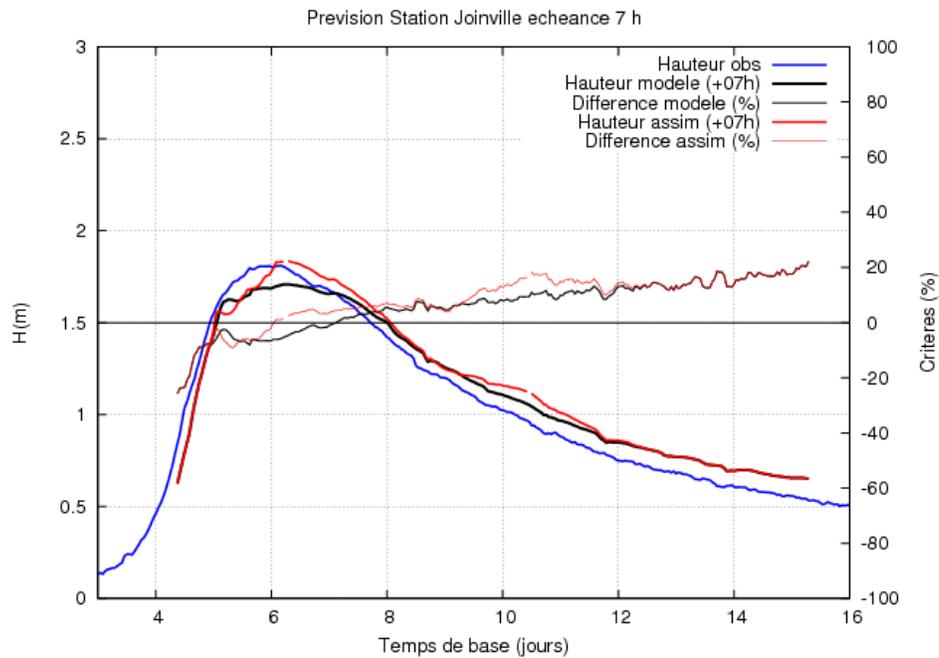
2- Précisions en mode prévision

La précision sur la prévision est améliorée par l'assimilation autour du pic de crue à Joinville, et sur tout l'épisode à Chamouilley.



3- Hauteurs prévues à l'échéance usuelle

Les prévisions de hauteurs avec assimilation sont meilleures pour le pic de crue aux deux stations.



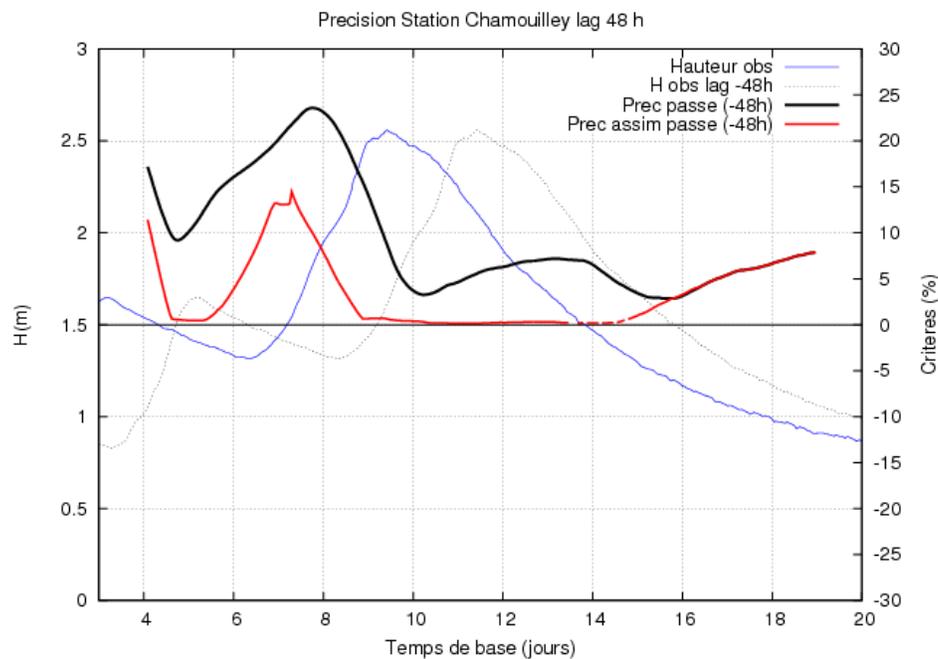
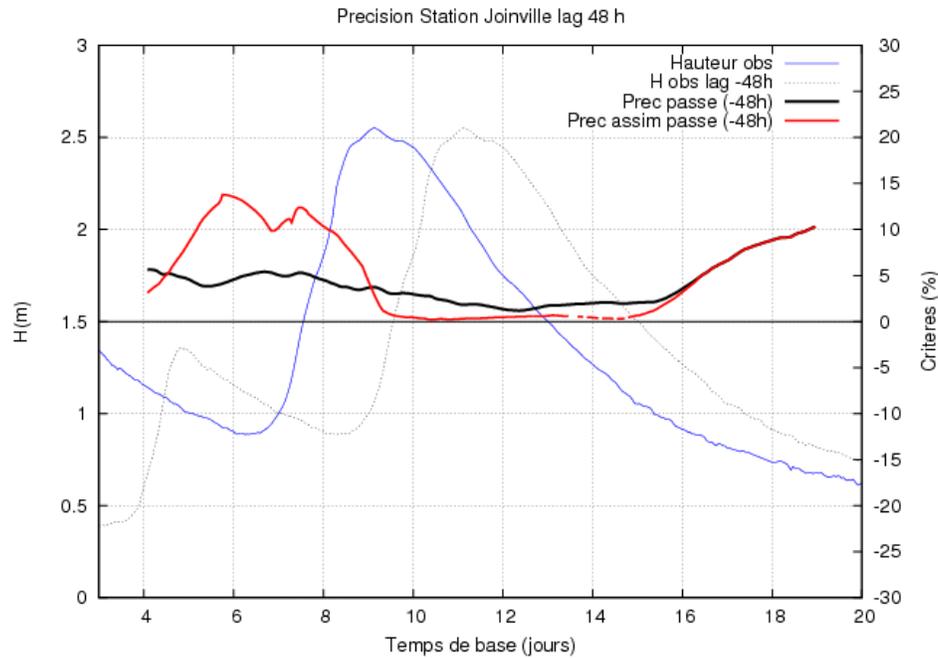
L'assimilation en mode CALAGEASSIM a amélioré la prévision de la crue aux deux stations.

Episode MARNE VALLAGE 03 : mars 2006

L'assimilation combine le calage et l'assimilation linéique (CALAGEASSIM).

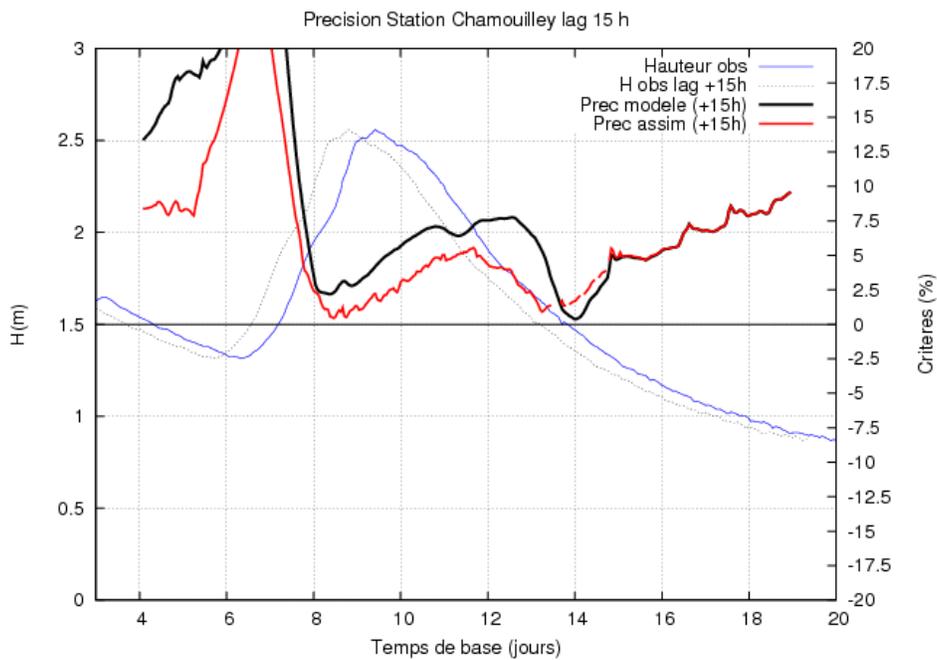
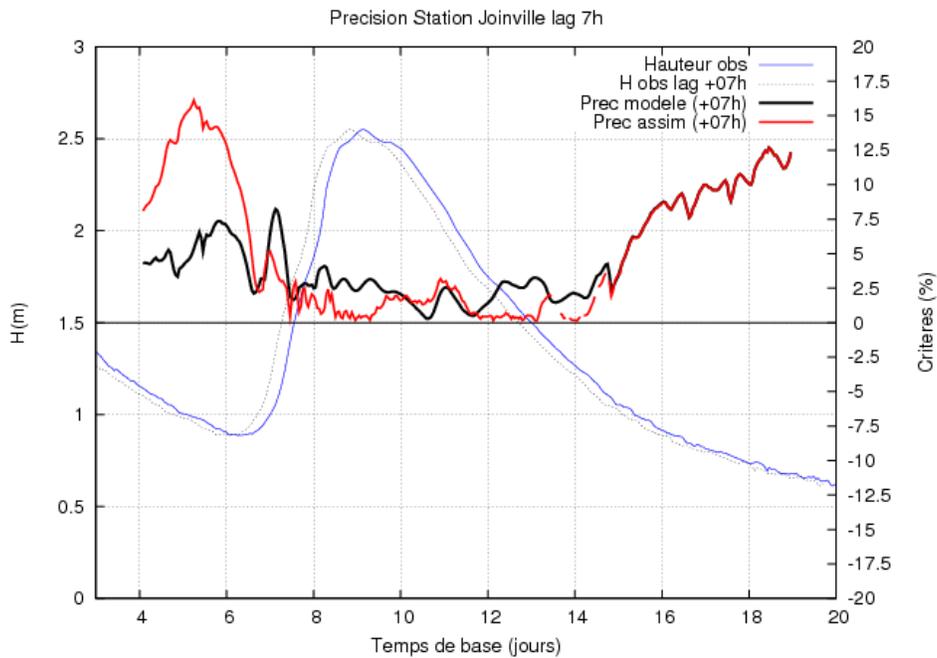
1- Précisions en mode simulation

La précision sur le passé est améliorée par l'assimilation à partir du pic de crue à Joinville, et sur tout l'épisode à Chamouilley.



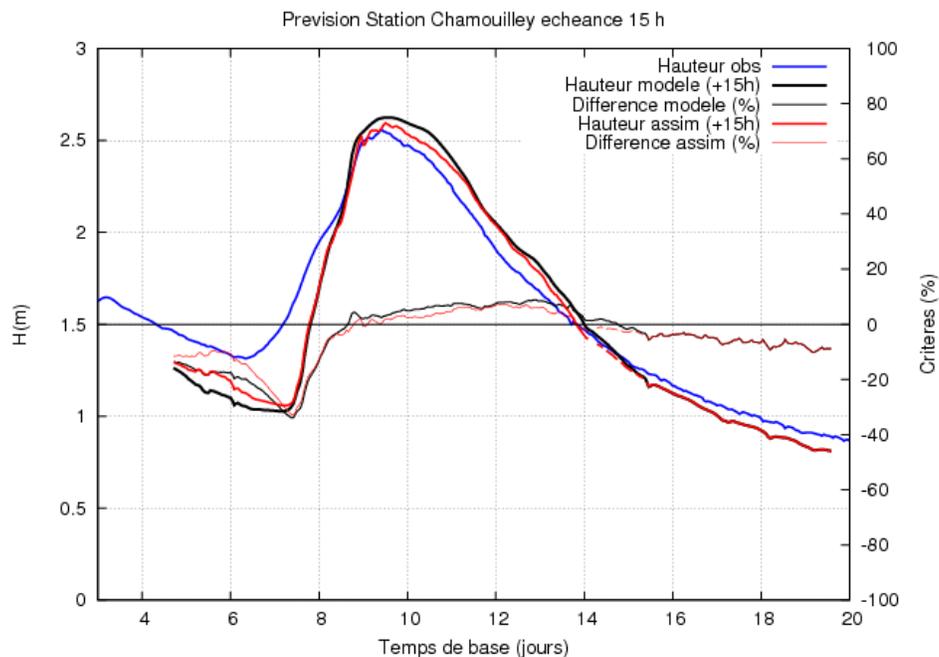
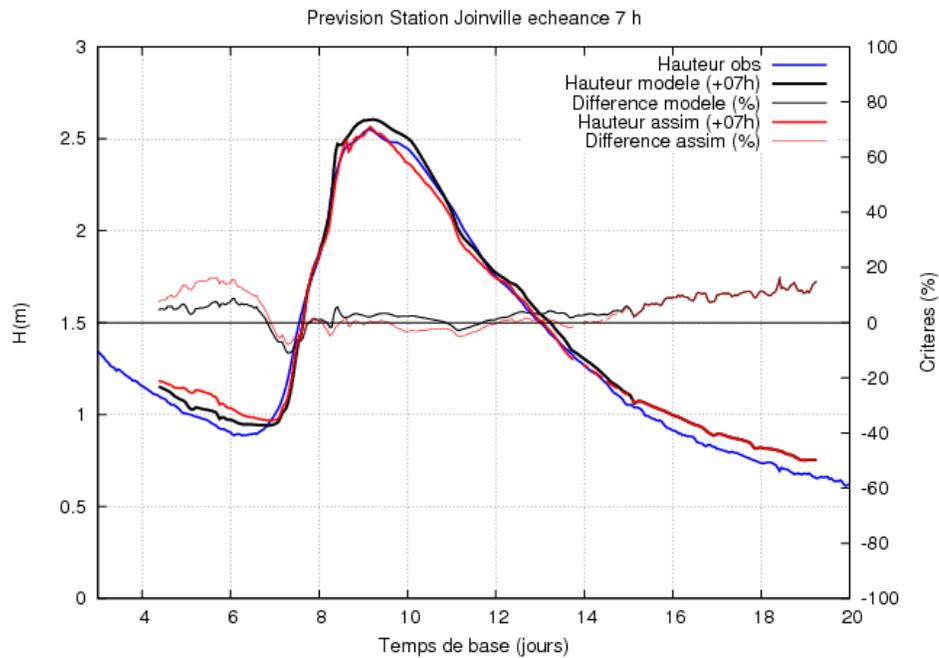
2- Précisions en mode prévision

La précision sur la prévision est améliorée par l'assimilation sur quasiment tout l'épisode de crue aux deux stations.



3- Hauteurs prévues à l'échéance usuelle

Les prévisions de hauteurs avec assimilation sont sensiblement meilleures pour le pic de crue aux deux stations.



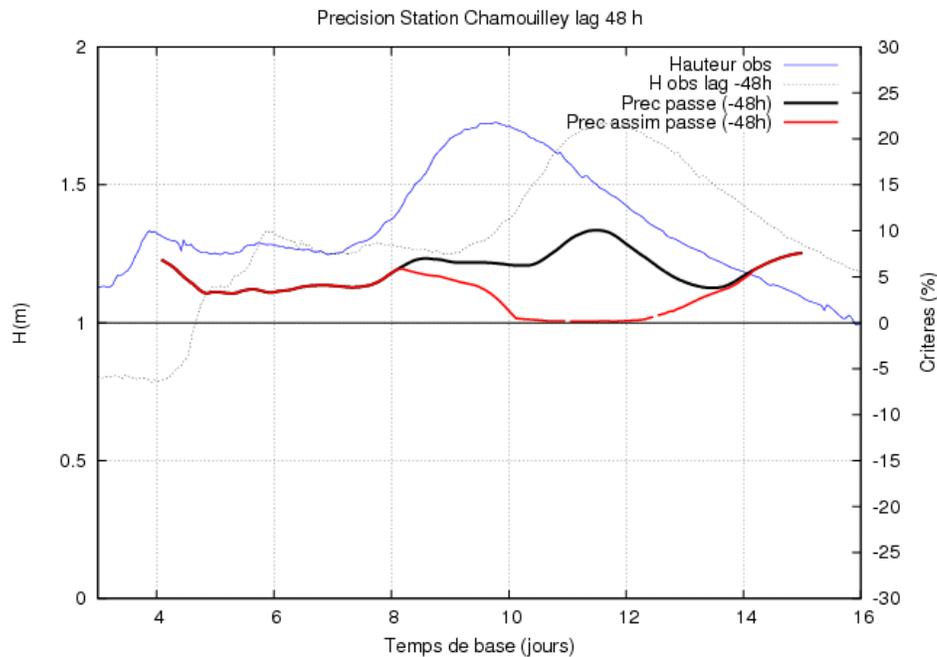
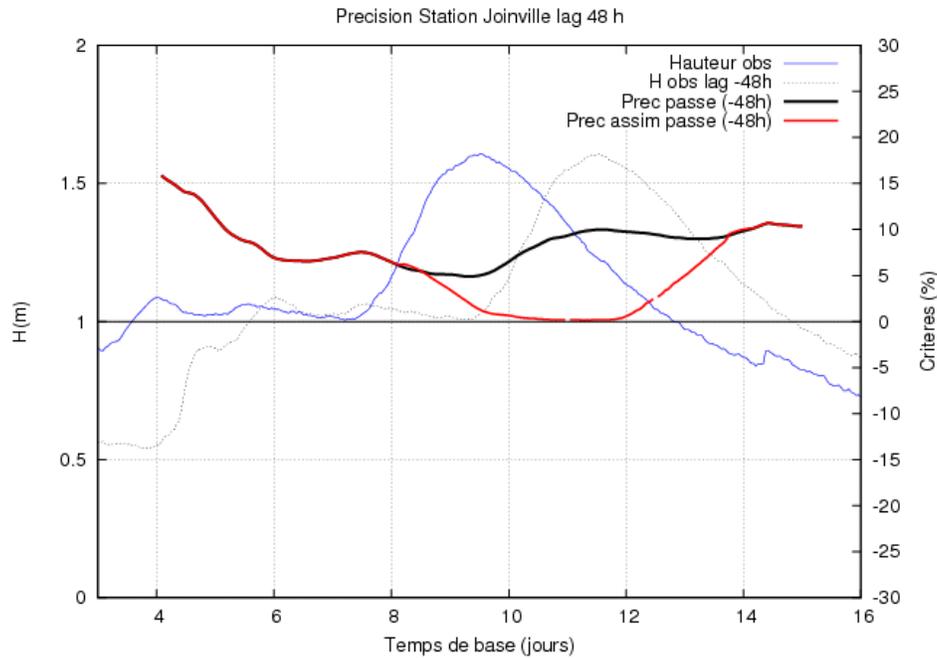
L'assimilation en mode CALAGEASSIM a amélioré la prévision de la crue aux deux stations.

Episode MARNE VALLAGE 05 : avril 2006

L'assimilation combine le calage et l'assimilation linéique (CALAGEASSIM).

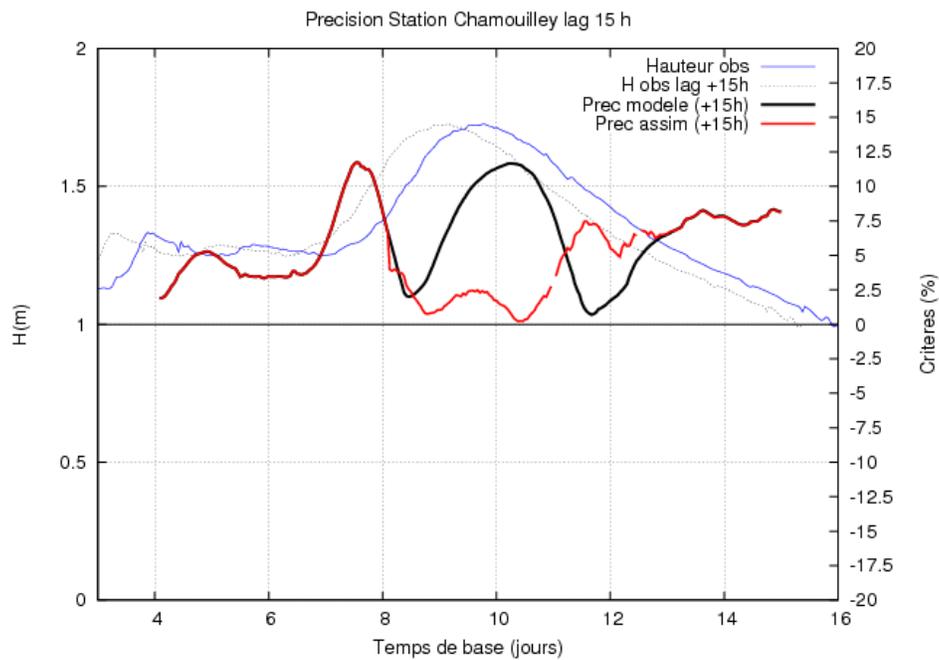
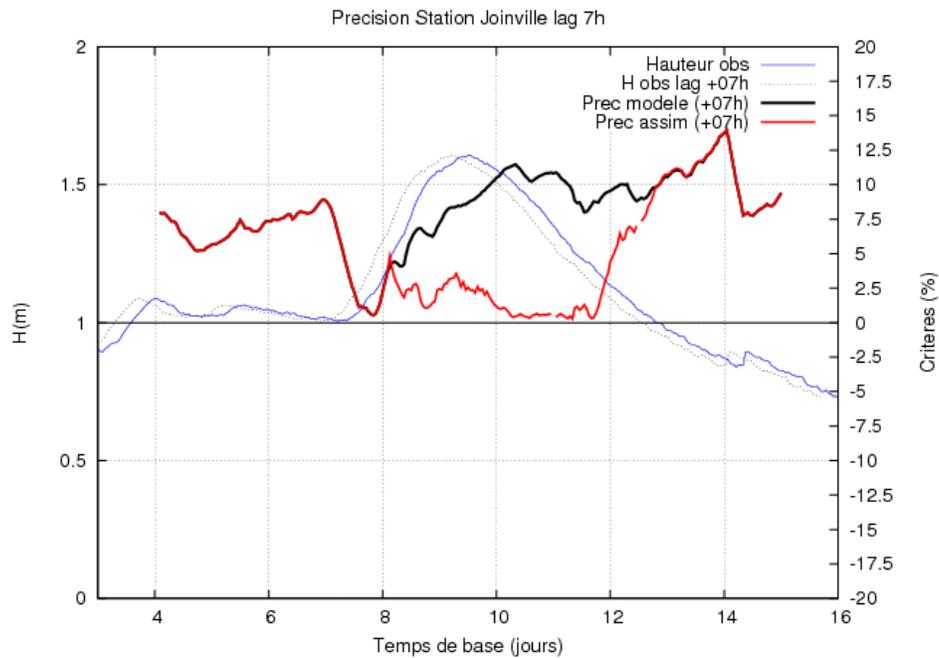
1- Précisions en mode simulation

La précision sur le passé est améliorée par l'assimilation à partir d'un jour avant le pic de crue aux deux stations.



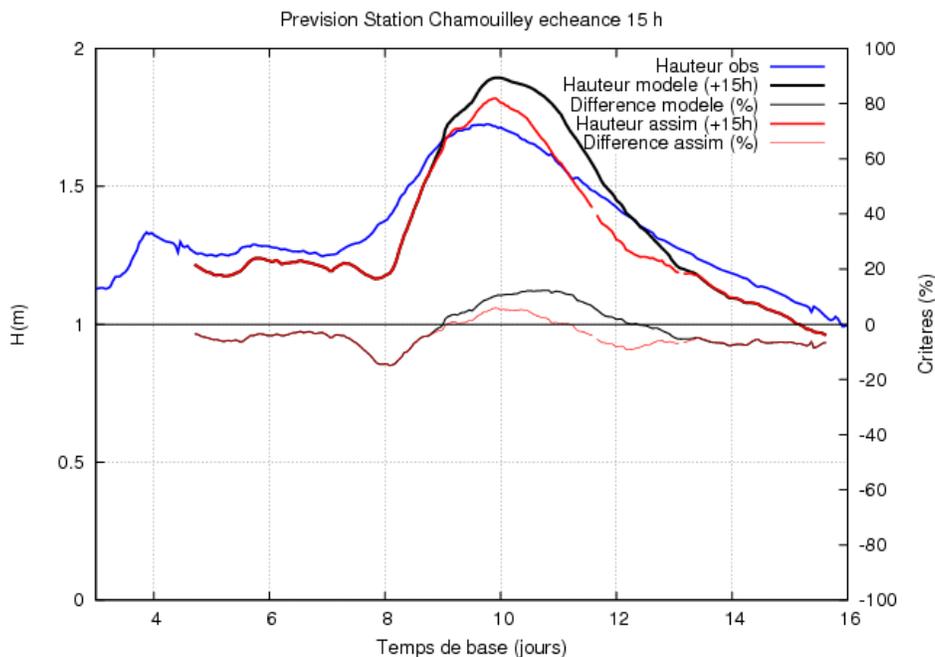
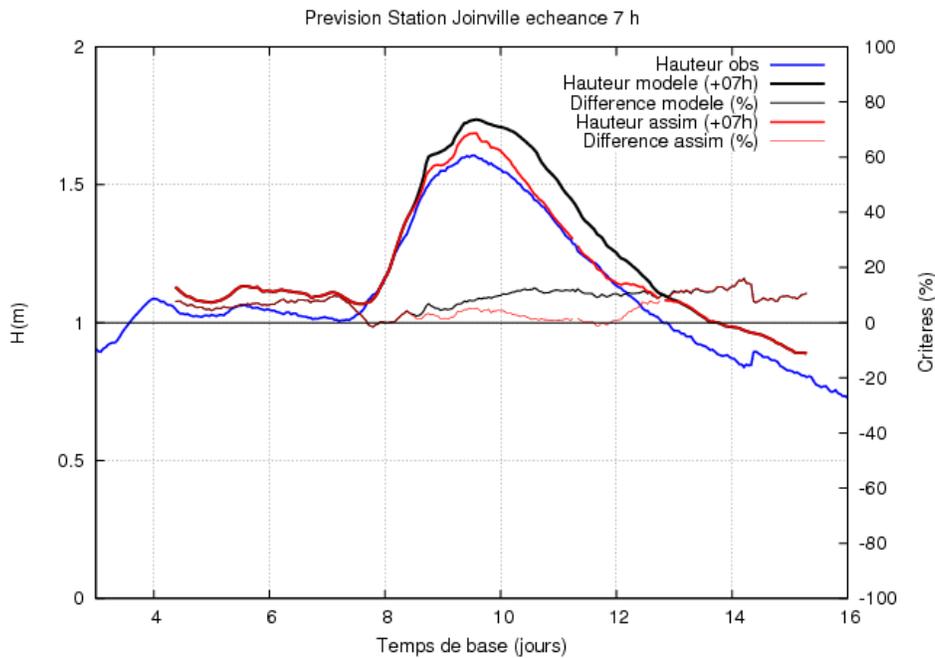
2- Précisions en mode prévision

La précision sur la prévision avec assimilation est égale ou meilleure sur tout l'épisode de crue aux deux stations, sauf en milieu de décrue à Chamouilley.



3- Hauteurs prévues à l'échéance usuelle

Les prévisions de hauteurs avec assimilation sont nettement meilleures pour le pic de crue aux deux stations.



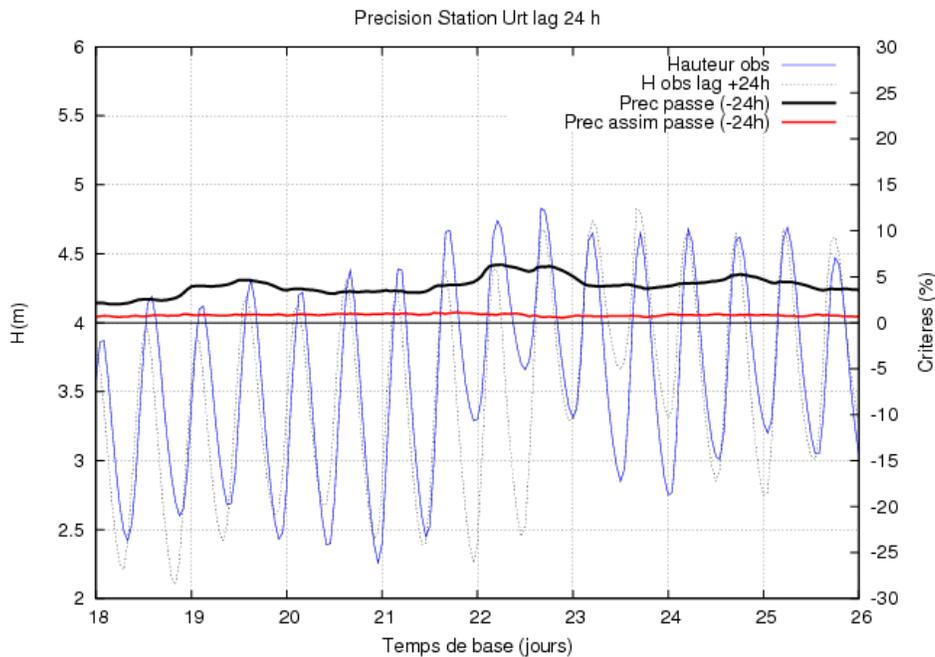
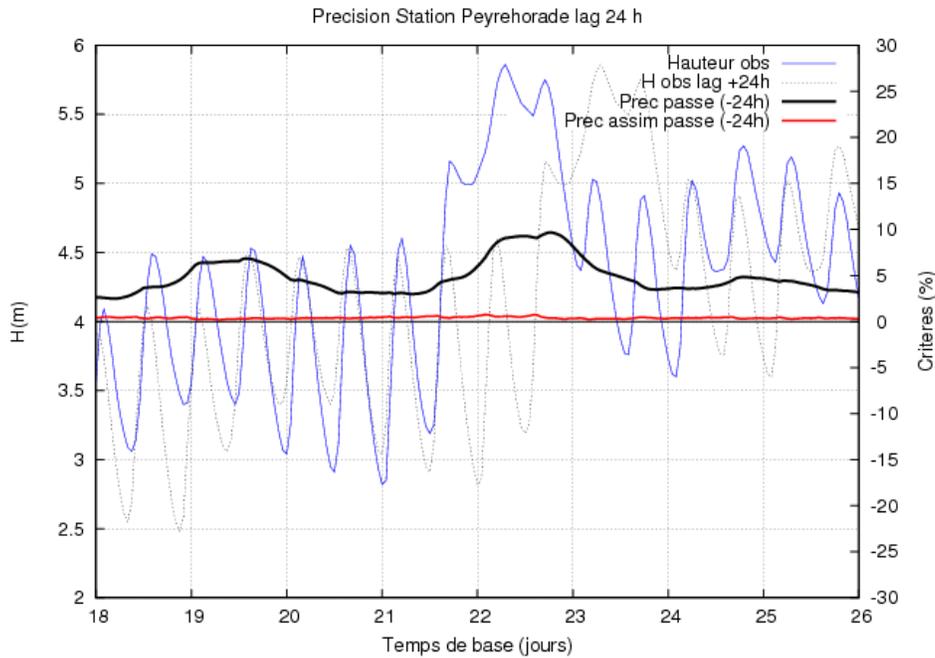
Le modèle direct surestime la crue. L'assimilation en mode CALAGEASSIM a amélioré nettement la prévision avec une légère surestimation du pic.

Episode ADOUR MARITIME 01

L'assimilation combine le calage et l'assimilation linéique (CALAGEASSIM).

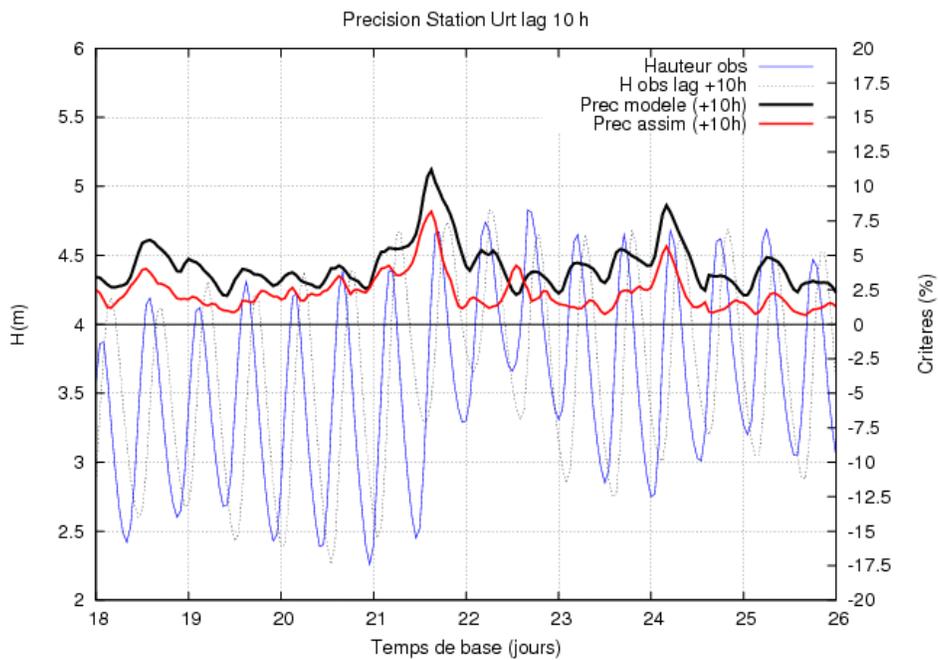
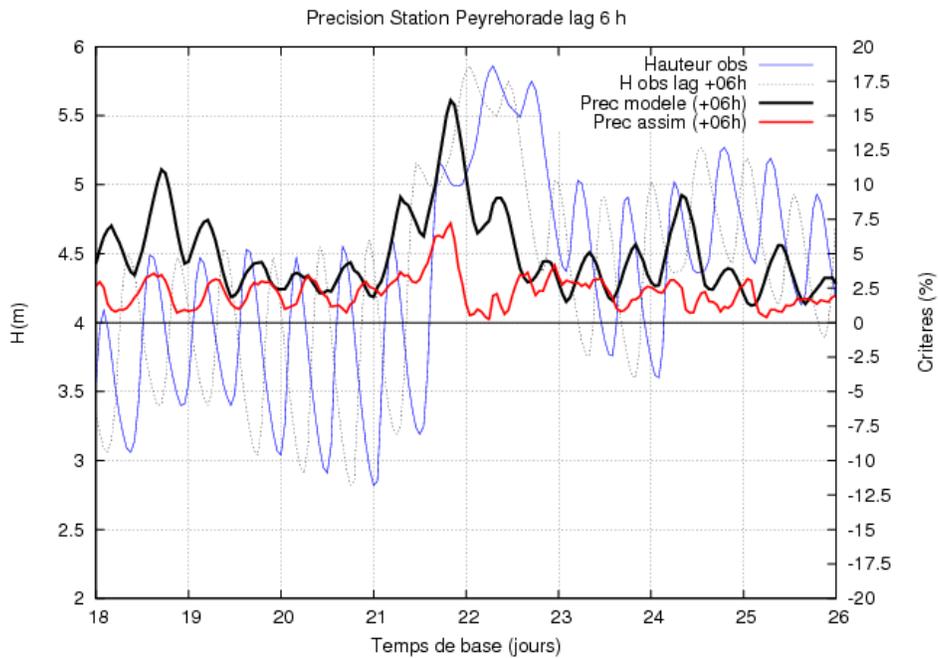
1- Précisions en mode simulation

La précision sur le passé est clairement améliorée par l'assimilation sur tout l'épisode aux deux stations.



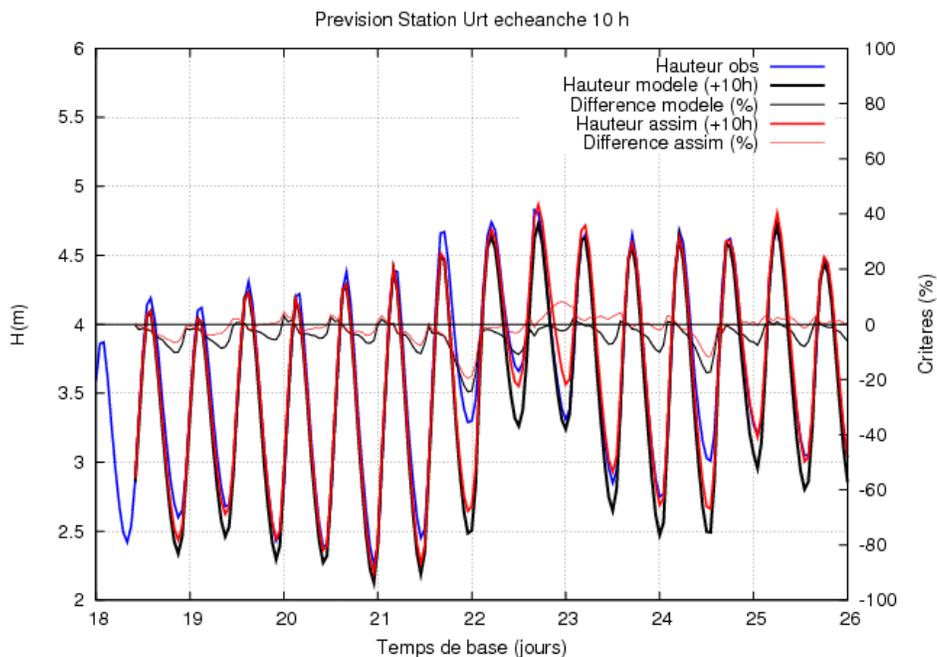
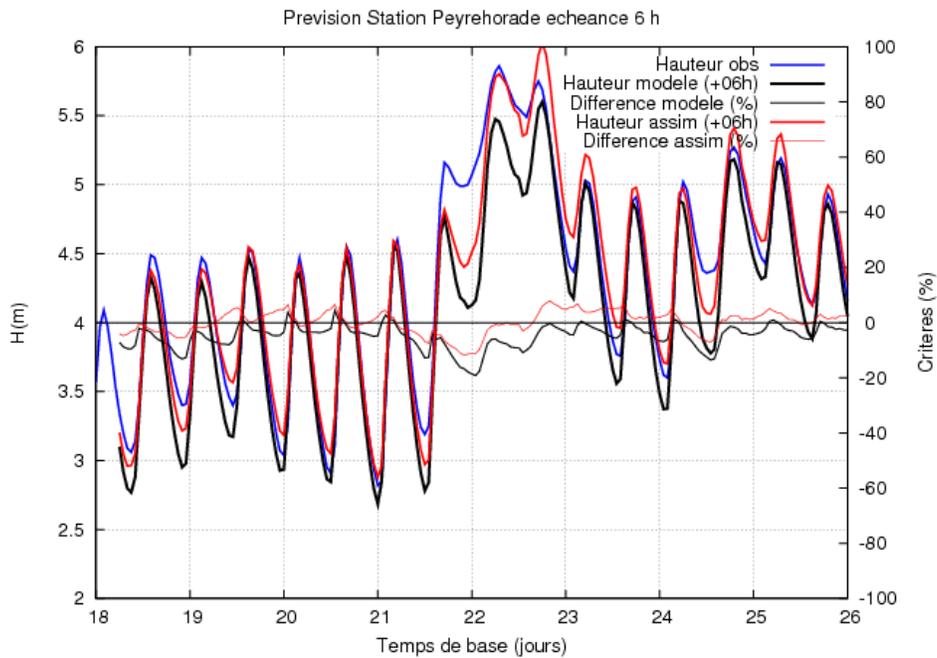
2- Précisions en mode prévision

La précision sur la prévision est améliorée par l'assimilation sur quasiment tout l'épisode aux deux stations.



3- Hauteurs prévues à l'échéance usuelle

Les prévisions de hauteurs avec assimilation sont nettement meilleures pour le premier pic de crue à la station Peyrehorade.



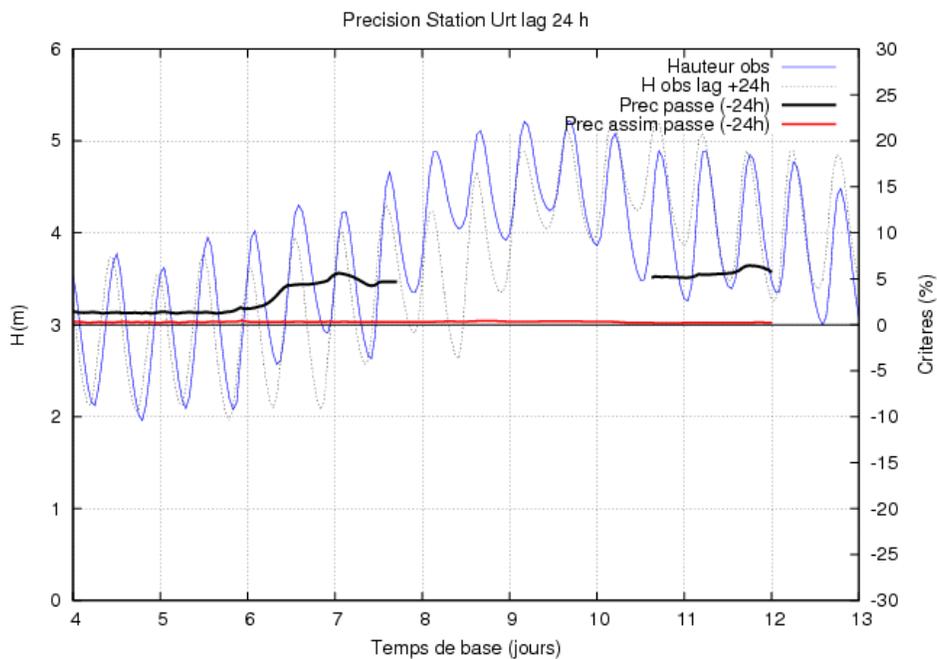
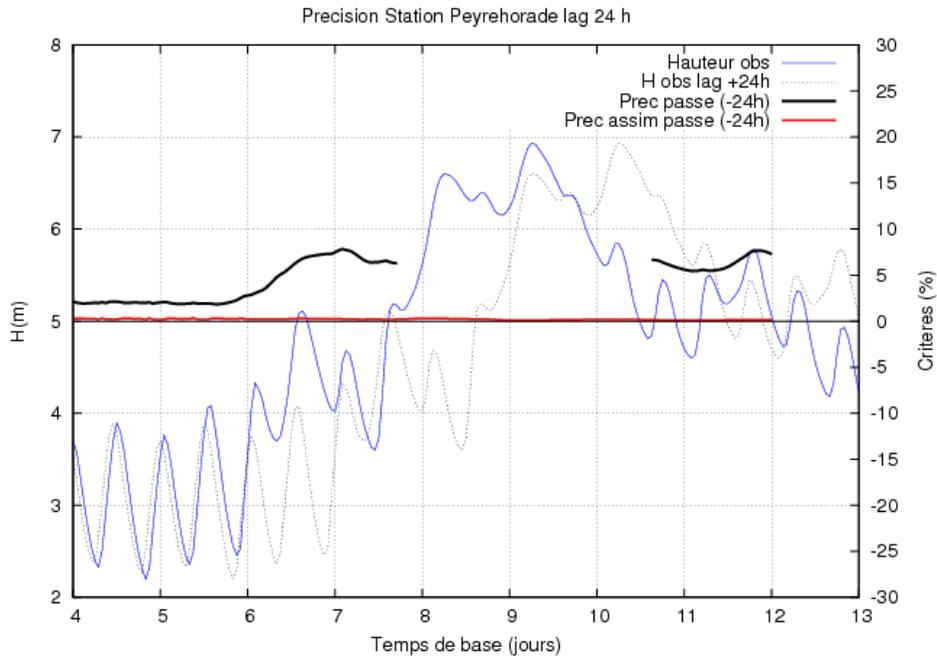
L'assimilation en mode CALAGEASSIM a amélioré nettement la prévision de crue aux deux stations.

Episode ADOUR MARITIME 03

Le modèle direct ne donne pas de résultats après le jour 8. Le calage par assimilation ne peut pas être utilisé. Par contre le modèle fonctionne avec l'assimilation linéique, seule à être activée ici (ASSIM).

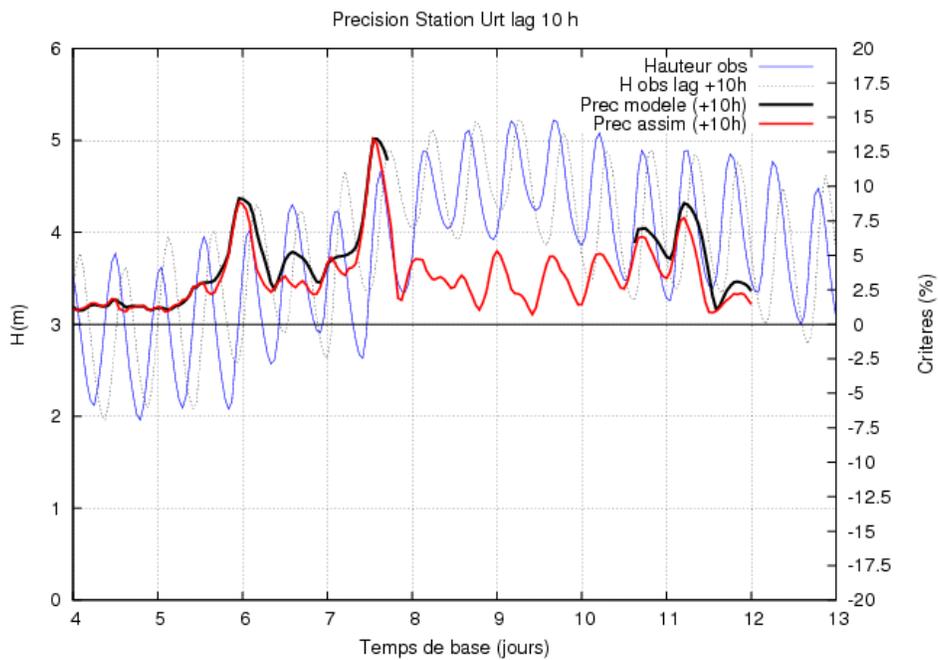
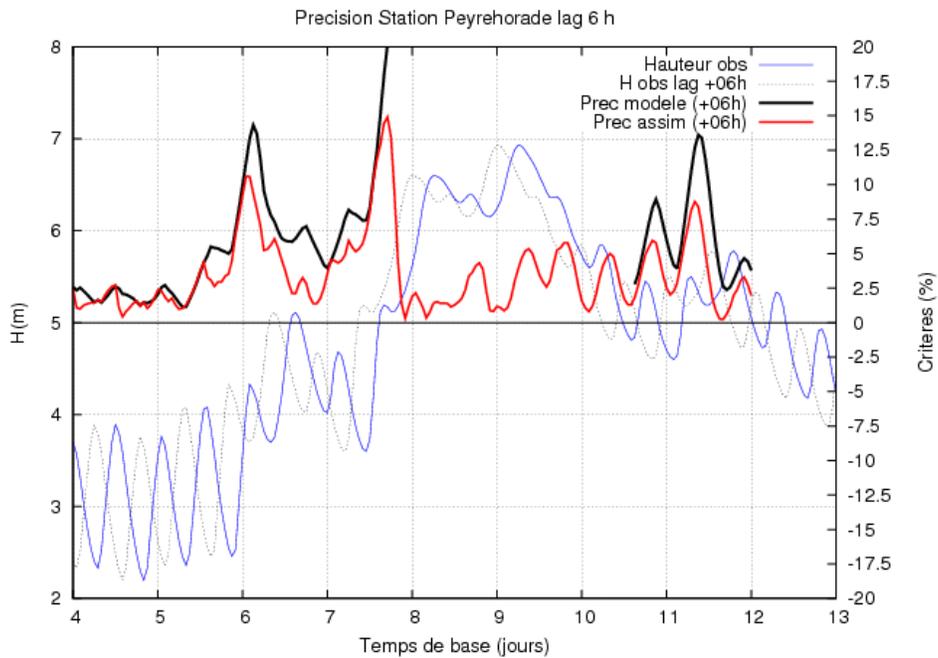
1- Précisions en mode simulation

La précision sur le passé est clairement améliorée par l'assimilation sur tout l'épisode aux deux stations.



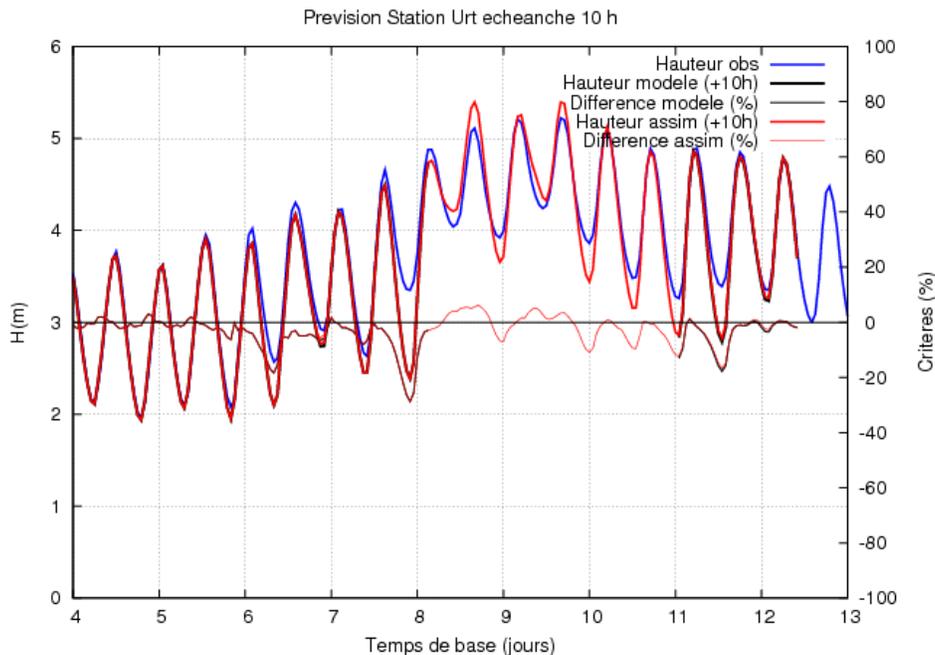
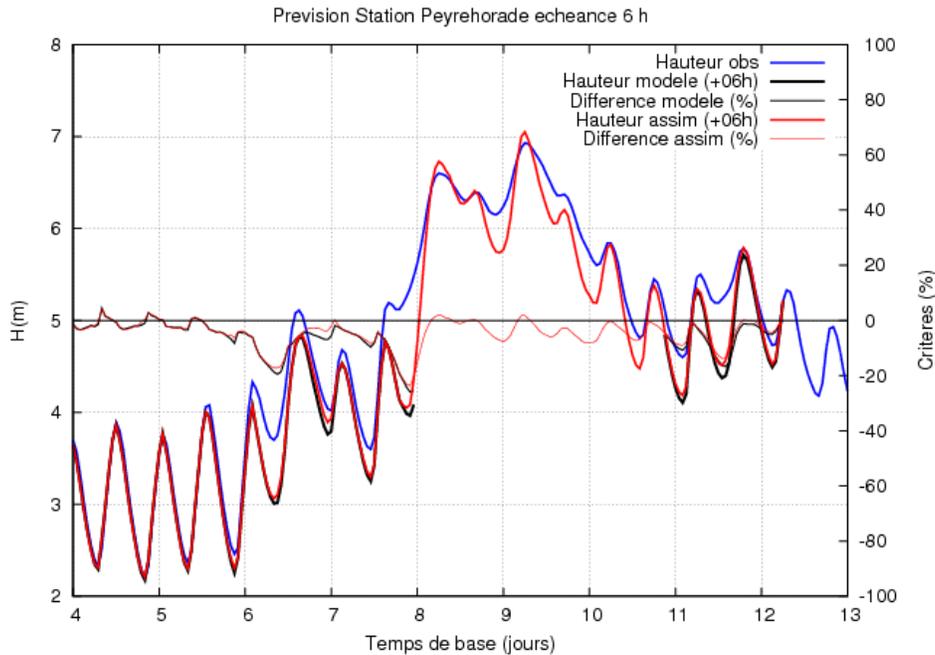
2- Précisions en mode prévision

La précision sur la prévision avec assimilation est bonne (inférieure à 5%) au pic et autour du pic de crue pour les deux stations.



3- Hauteurs prévues à l'échéance usuelle

L'assimilation linéique donne de bonnes prévisions de hauteurs pour l'enchainement des deux pics de crue, alors que le modèle direct ne fonctionne pas. Les prévisions sont très bonnes aux pics de la station Peyrehorade.



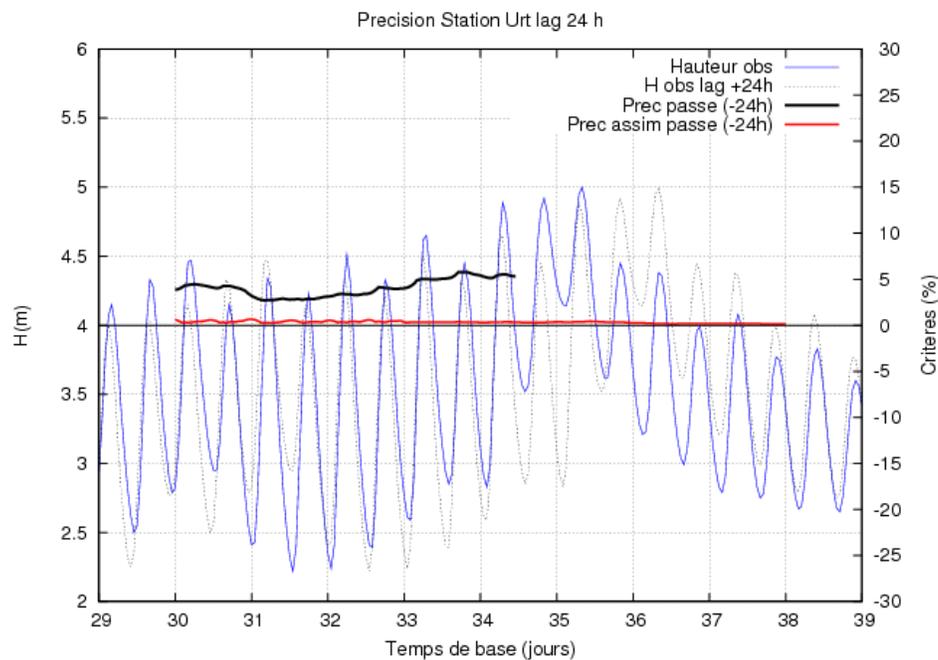
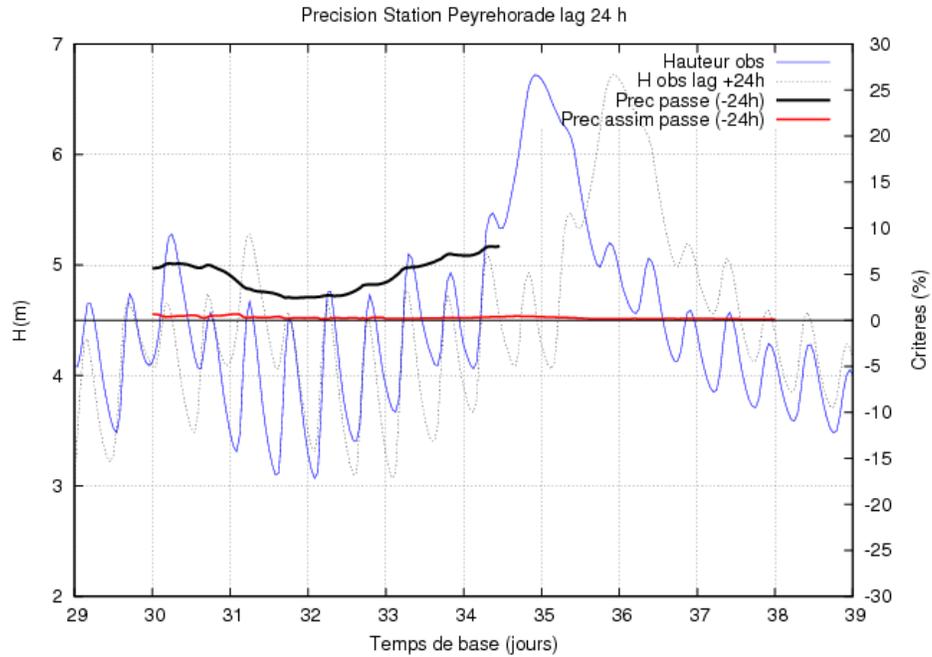
L'assimilation en mode ASSIM a permis de simuler l'épisode et donne de bons résultats de prévision de cette double crue.

Episode ADOUR MARITIME 04

Le modèle direct ne donne pas de résultats après le jour 35. Le calage par assimilation ne peut pas être utilisé. Par contre le modèle fonctionne avec l'assimilation linéique, seule à être activée ici (ASSIM).

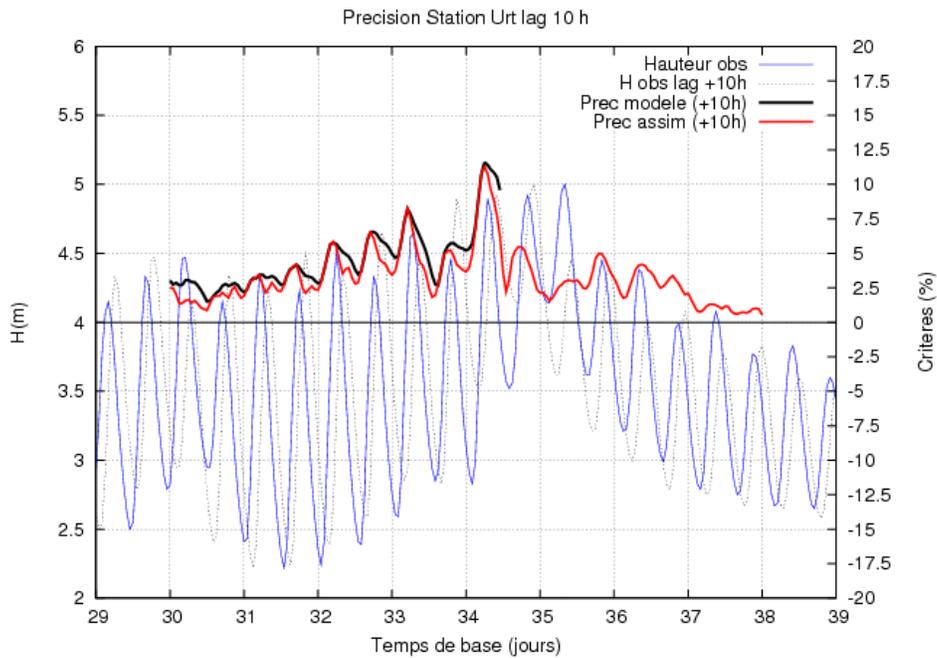
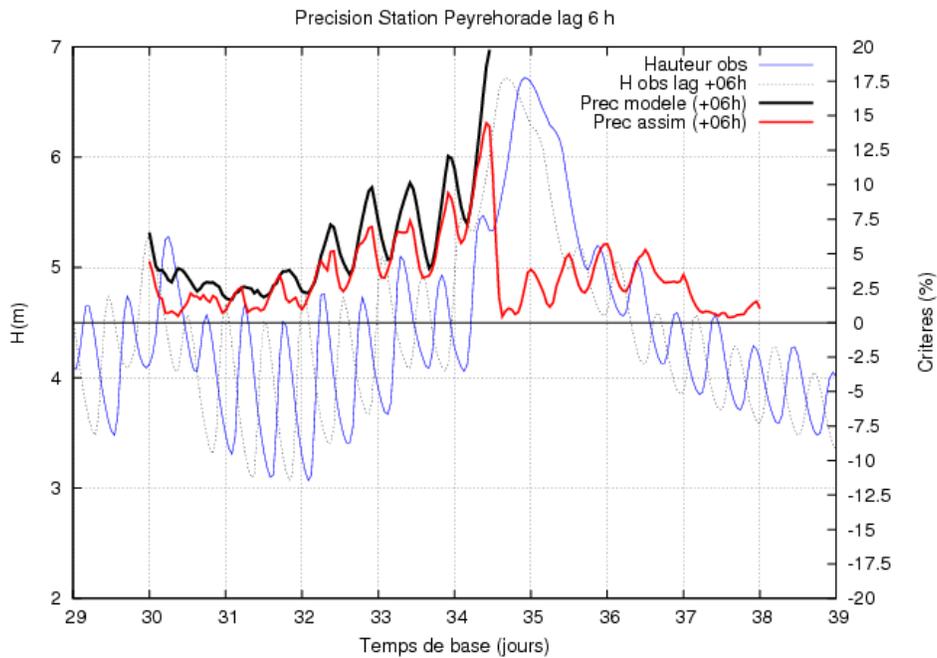
1- Précisions en mode simulation

La précision sur le passé est clairement améliorée par l'assimilation sur tout l'épisode aux deux stations.



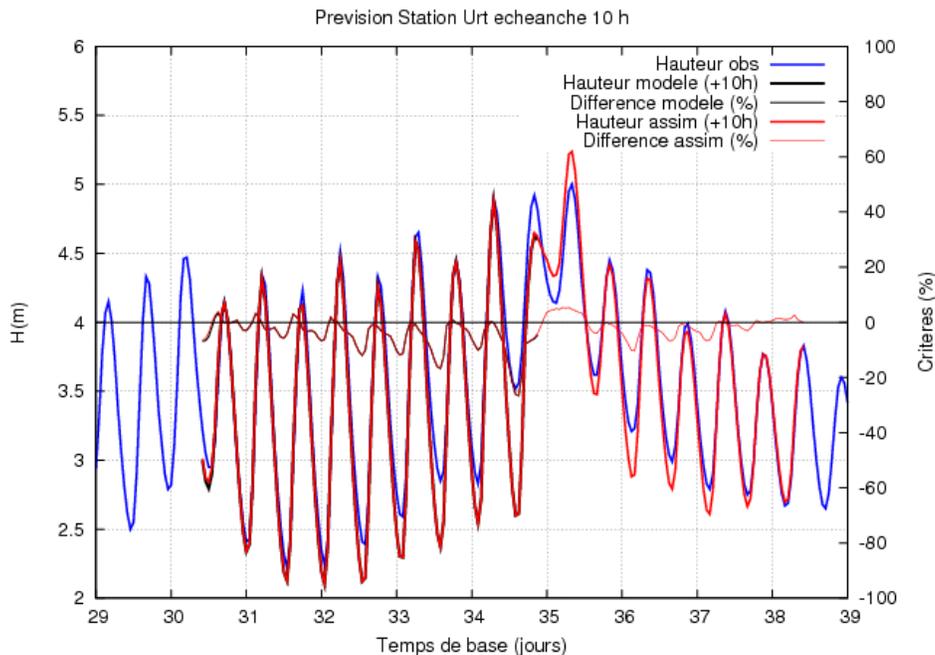
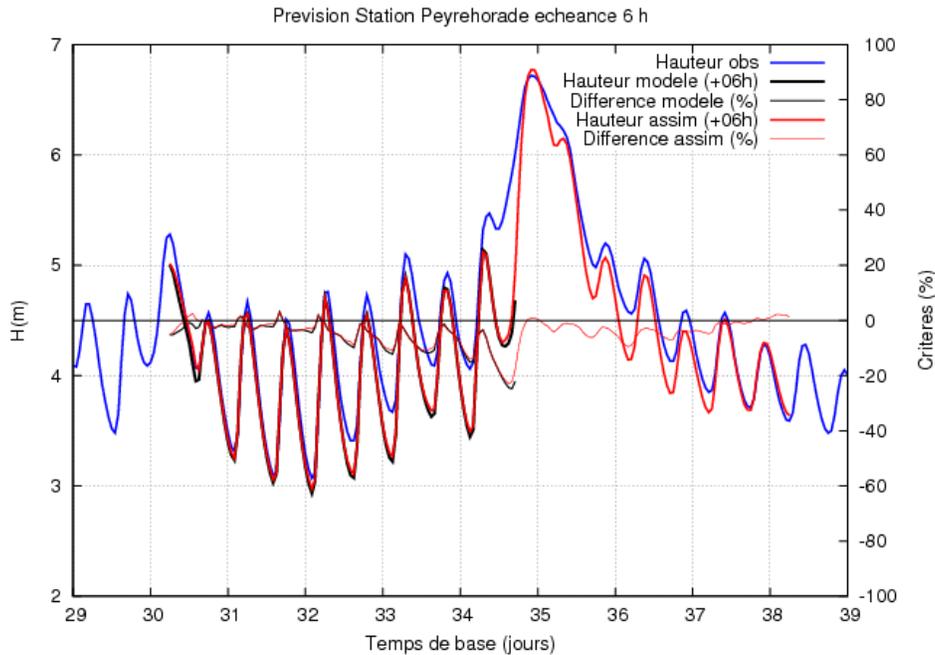
2- Précisions en mode prévision

La précision sur la prévision avec assimilation est bonne (inférieure à 5%) au pic et autour du pic de crue pour les deux stations.



3- Hauteurs prévues à l'échéance usuelle

L'assimilation linéique donne de bonnes prévisions de hauteurs pour cet épisode, alors que le modèle direct ne fonctionne pas. Les prévisions sont très bonnes au pic de la station Peyrehorade, un peu moins bien à Urt.



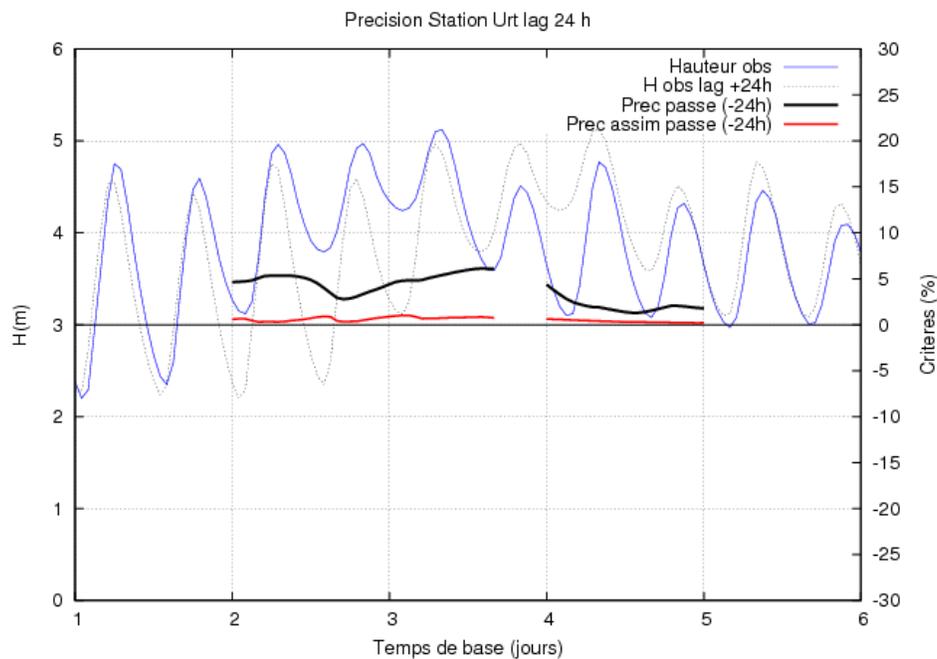
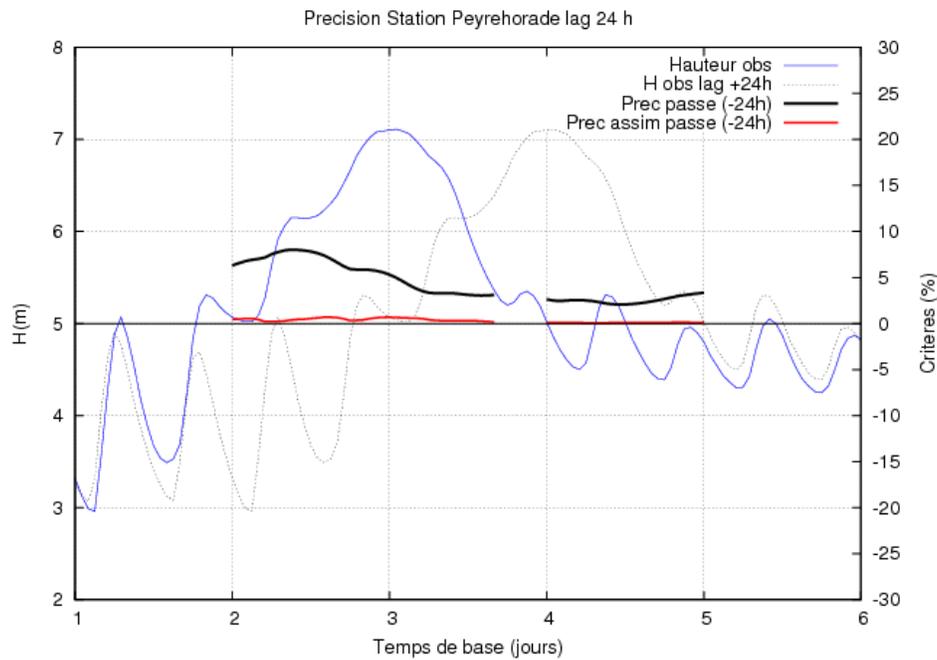
L'assimilation en mode ASSIM a permis de simuler l'épisode et a donné de bons résultats.

Episode ADOUR MARITIME 05

L'assimilation activée sur cet épisode est l'assimilation linéïque seule (ASSIM).

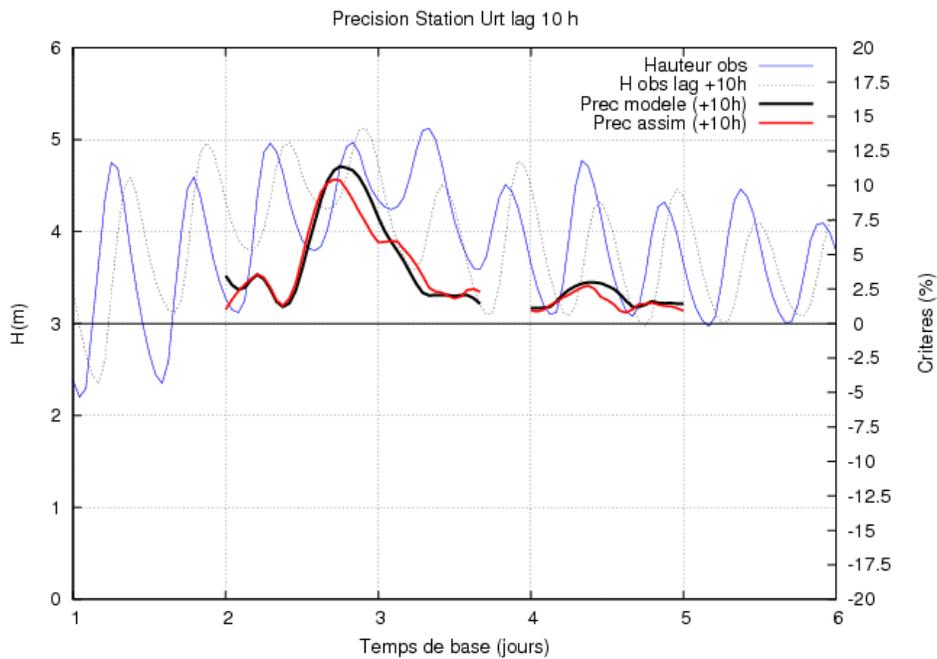
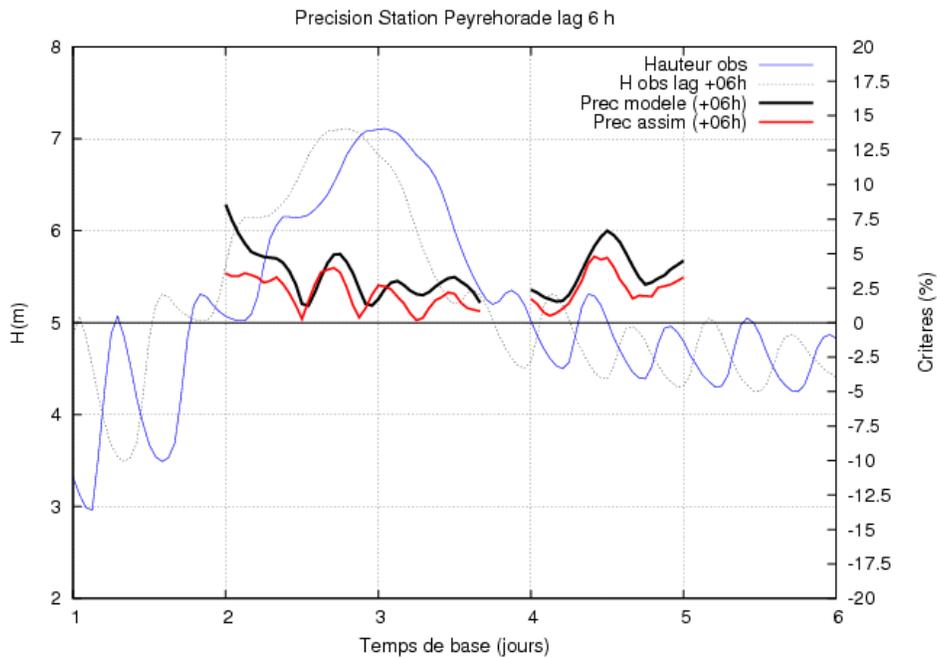
1- Précisions en mode simulation

La précision sur le passé est clairement améliorée par l'assimilation autour du pic de crue aux deux stations.



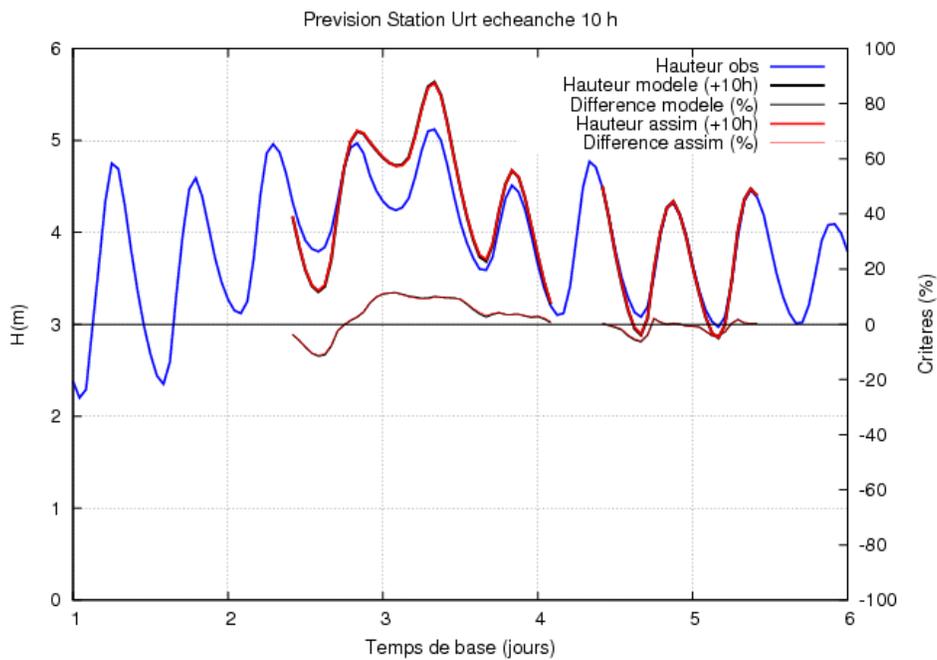
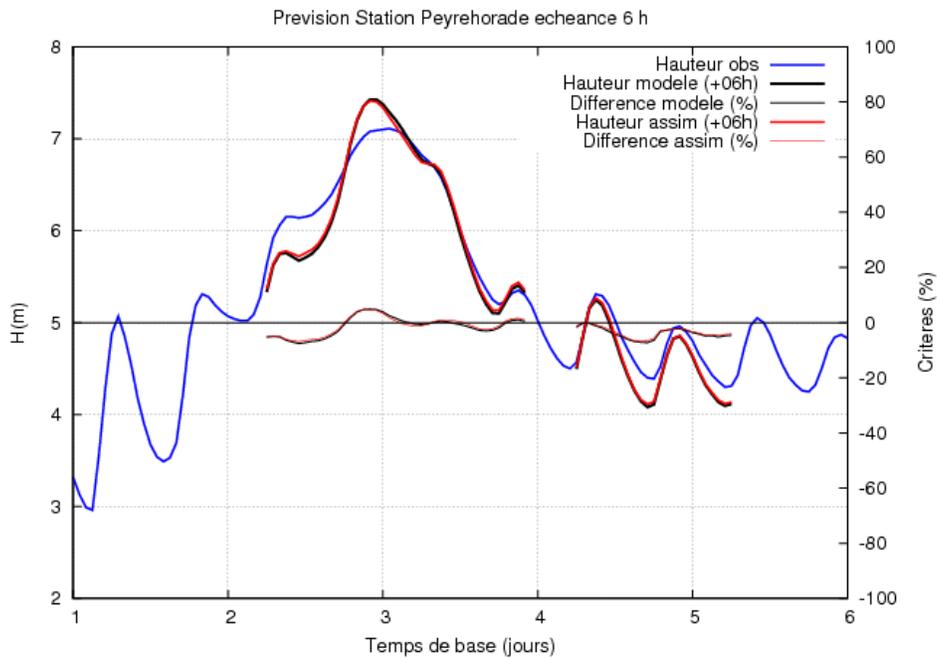
2- Précisions en mode prévision

La précision sur la prévision est légèrement améliorée par l'assimilation aux deux stations.



3- Hauteurs prévues à l'échéance usuelle

Les prévisions de hauteurs avec assimilation sont identiques à celles du modèle direct aux deux stations.



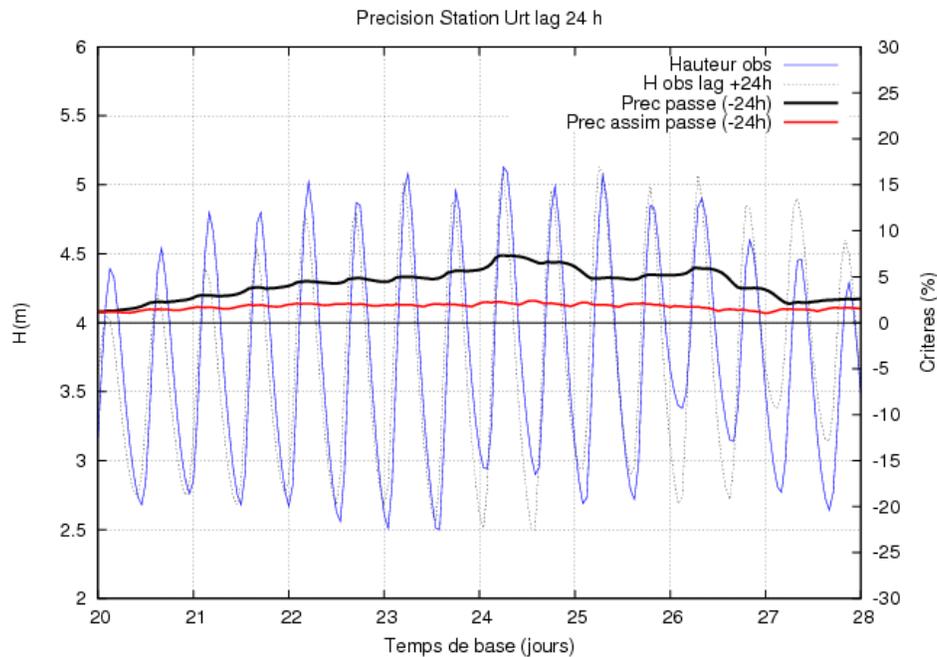
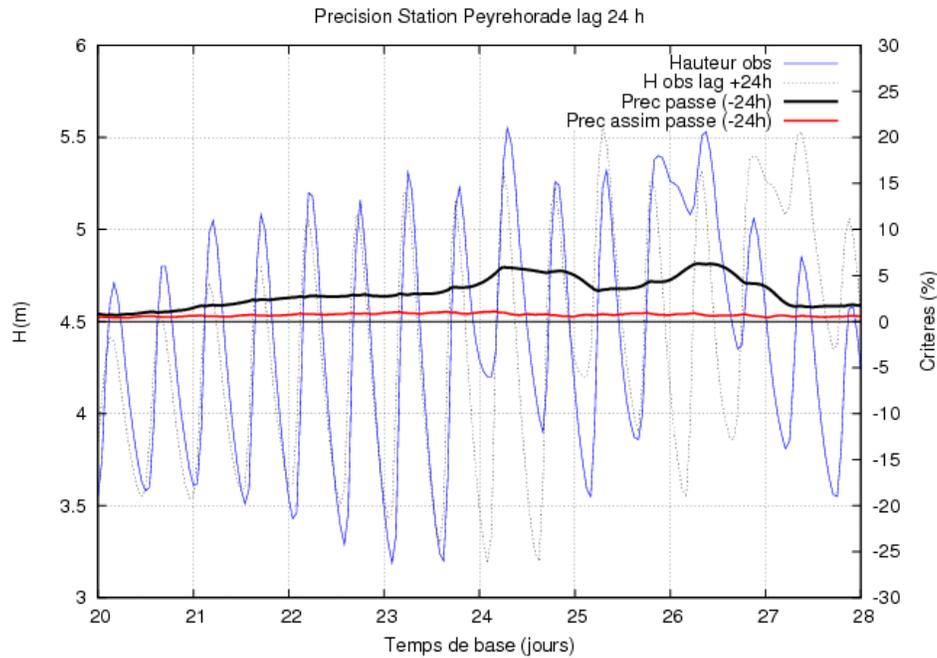
L'assimilation en mode ASSIM a apporté peu d'amélioration au modèle direct.

Episode ADOUR MARITIME 07

L'assimilation combine le calage et l'assimilation linéique (CALAGEASSIM).

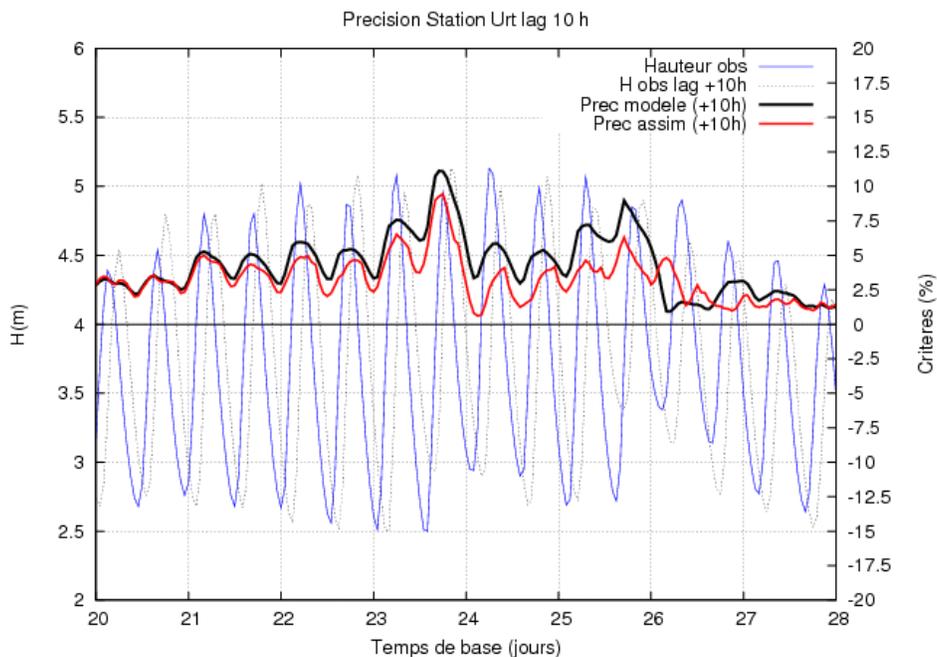
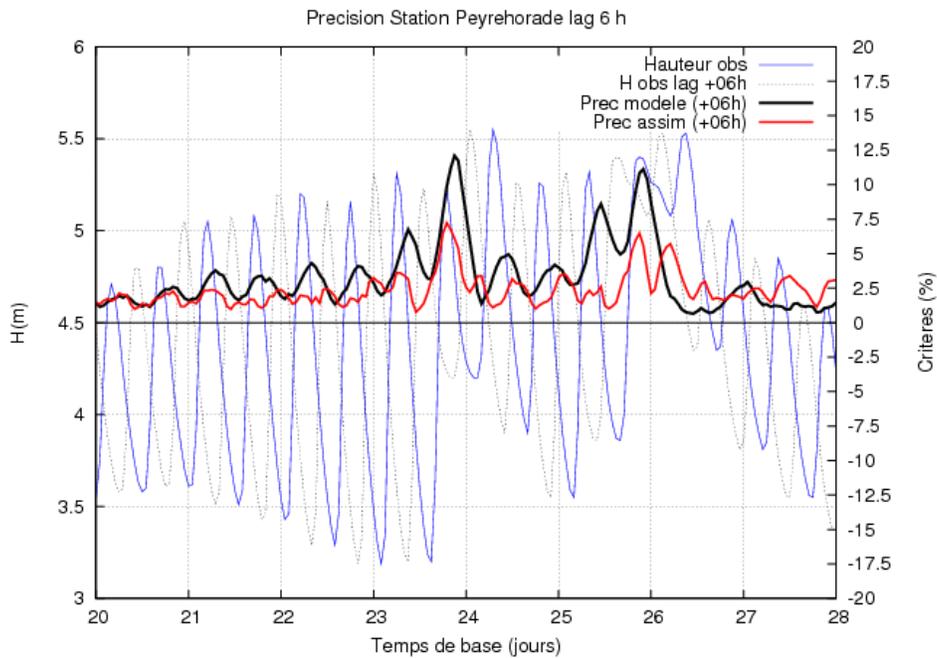
1- Précisions en mode simulation

La précision sur le passé est améliorée par l'assimilation sur tout l'épisode aux deux stations.



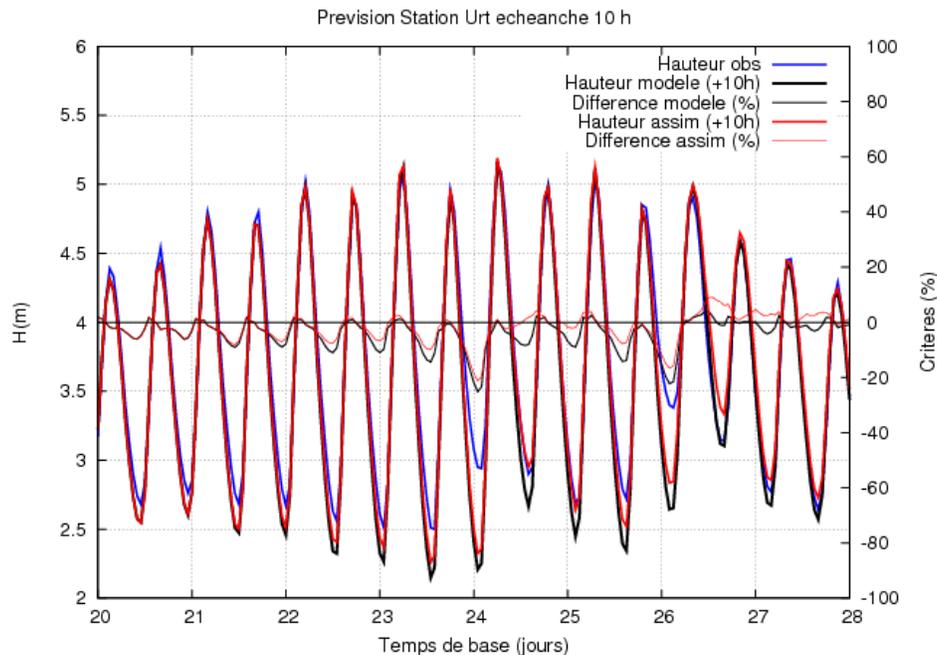
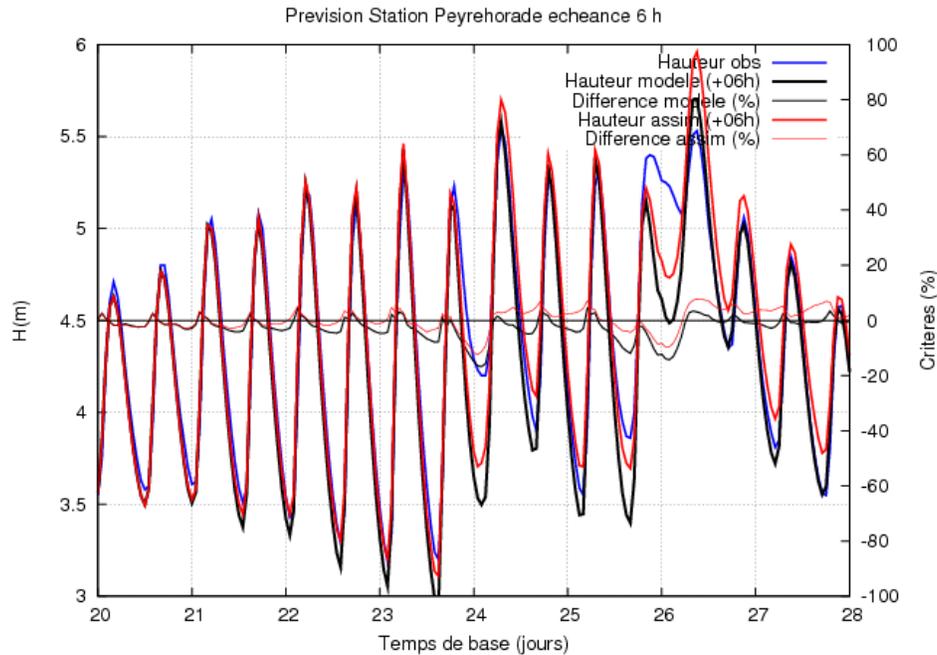
2- Précisions en mode prévision

La précision sur la prévision est améliorée par l'assimilation aux deux stations, sauf pour le deuxième pic de crue à la station Peyrehorade.



3- Hauteurs prévues à l'échéance usuelle

Les prévisions de hauteurs avec assimilation sont améliorées jusqu'au dernier pic pour les deux stations. Le pic de crue et la décrue sont surestimés à la station Peyrehorade.



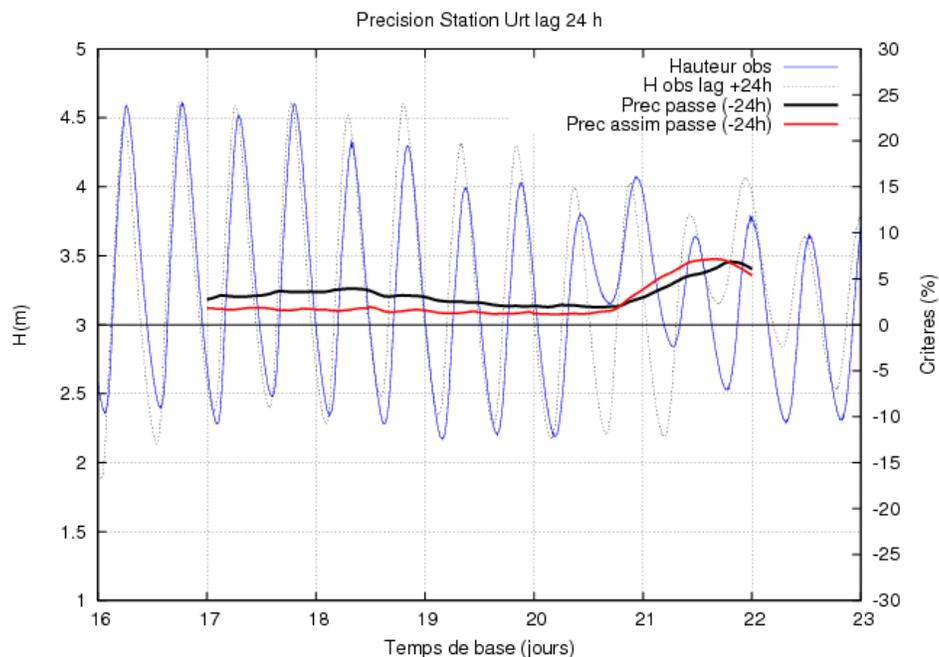
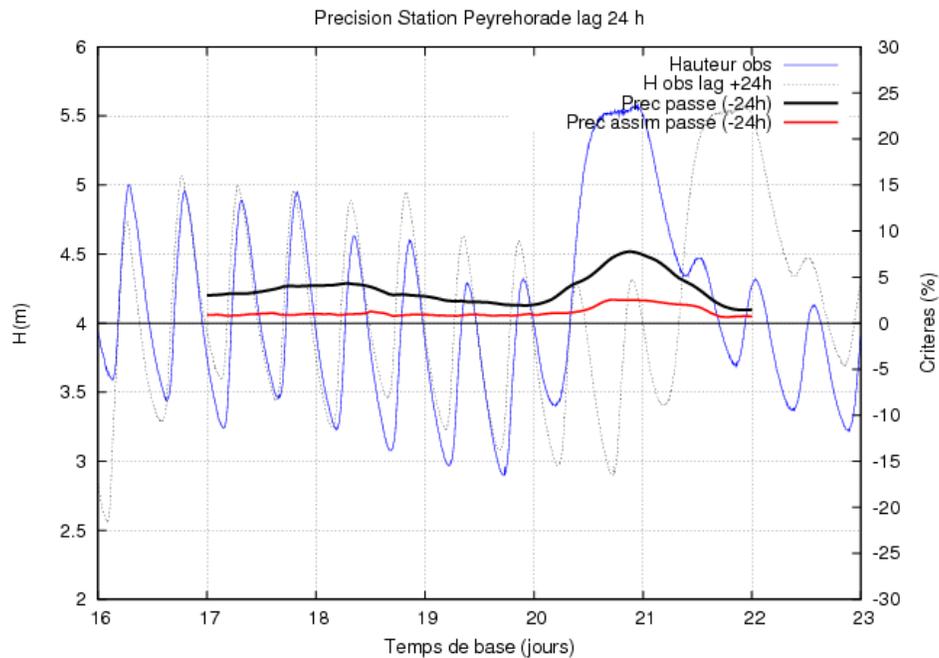
L'assimilation en mode CALAGEASSIM a globalement amélioré la prévision de l'épisode. Le dernier pic est néanmoins surestimé.

Episode ADOUR MARITIME 09 : avril 2009

L'assimilation activée sur cet épisode est l'assimilation linéïque seule (ASSIM).
Les observations ont ici un pas de temps de 5mn ; elles sont moyennées sur une fenêtre glissante de 25mn.

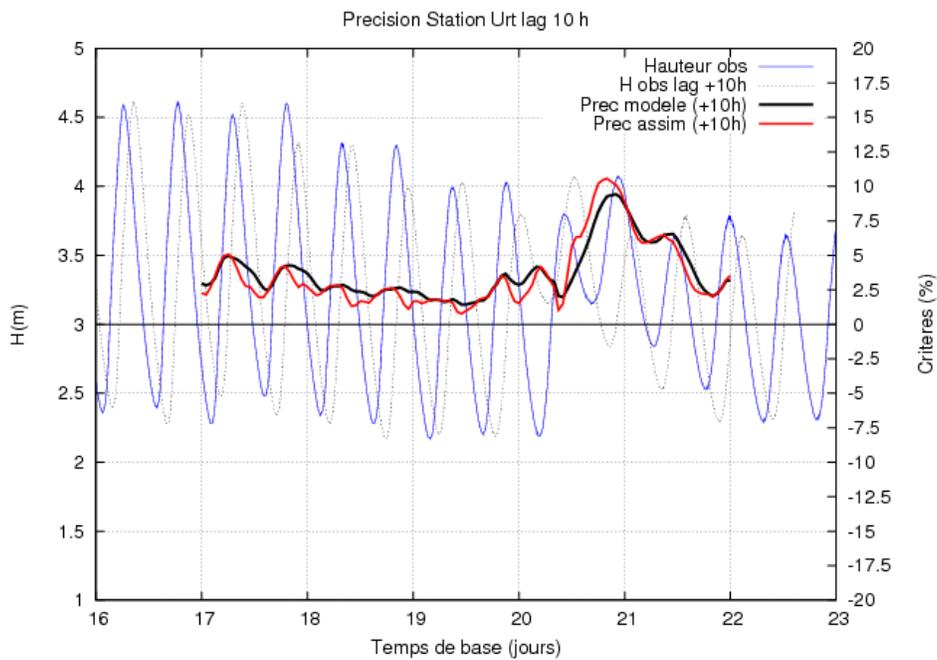
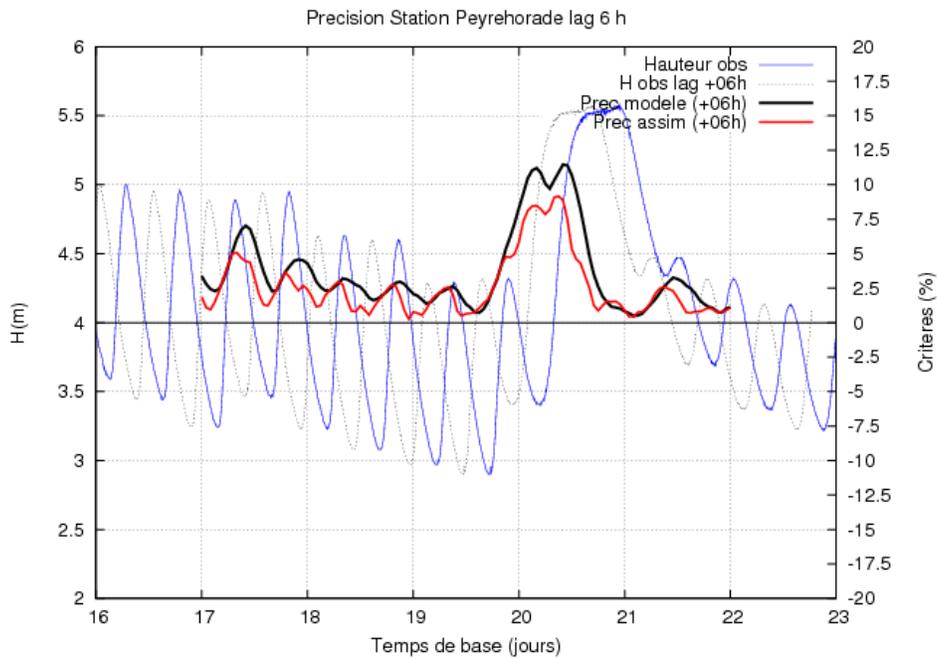
1- Précisions en mode simulation

La précision sur le passé est améliorée par l'assimilation sur quasiment tout l'épisode aux deux stations.



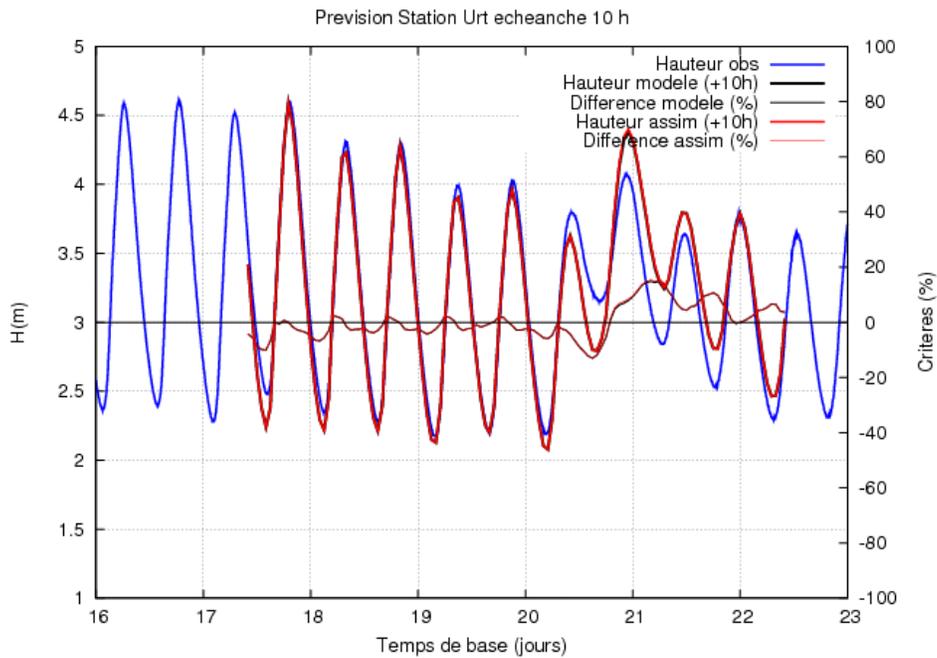
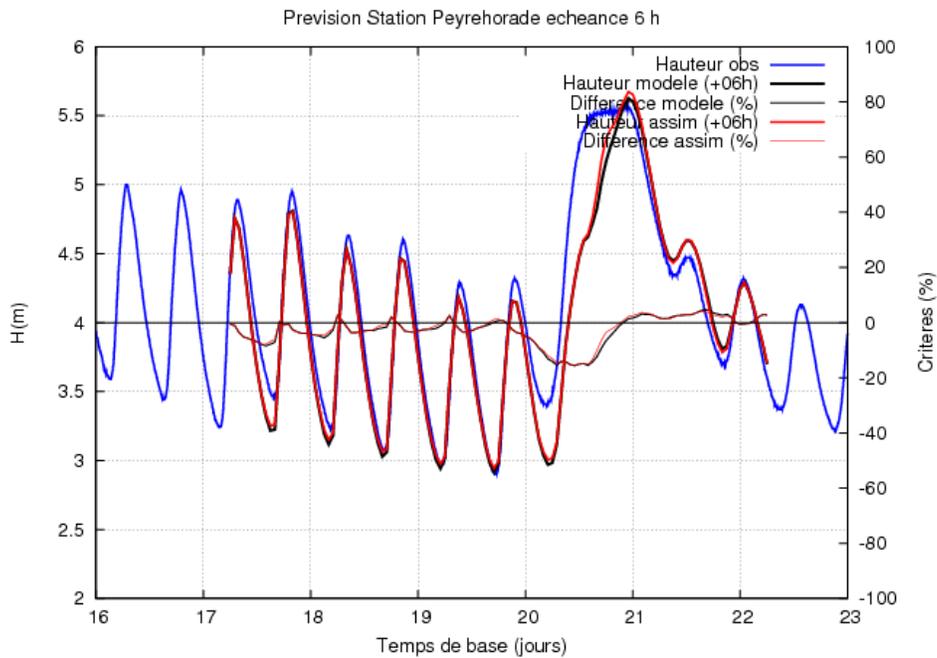
2- Précisions en mode prévision

La précision sur la prévision est légèrement améliorée par l'assimilation aux deux stations.



3- Hauteurs prévues à l'échéance usuelle

Les prévisions de hauteurs avec assimilation sont quasiment identiques à celles du modèle direct aux deux stations.



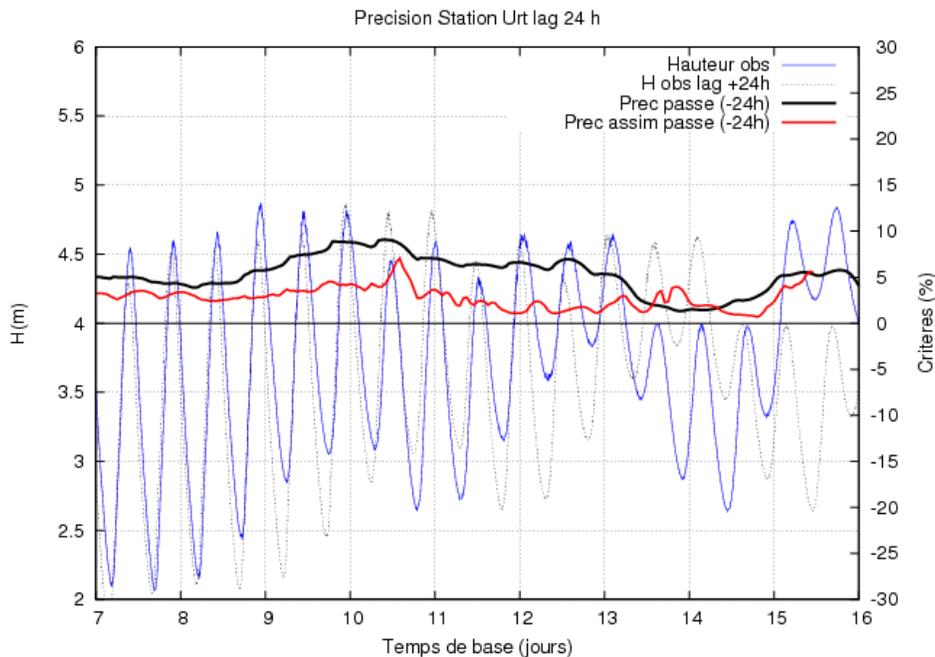
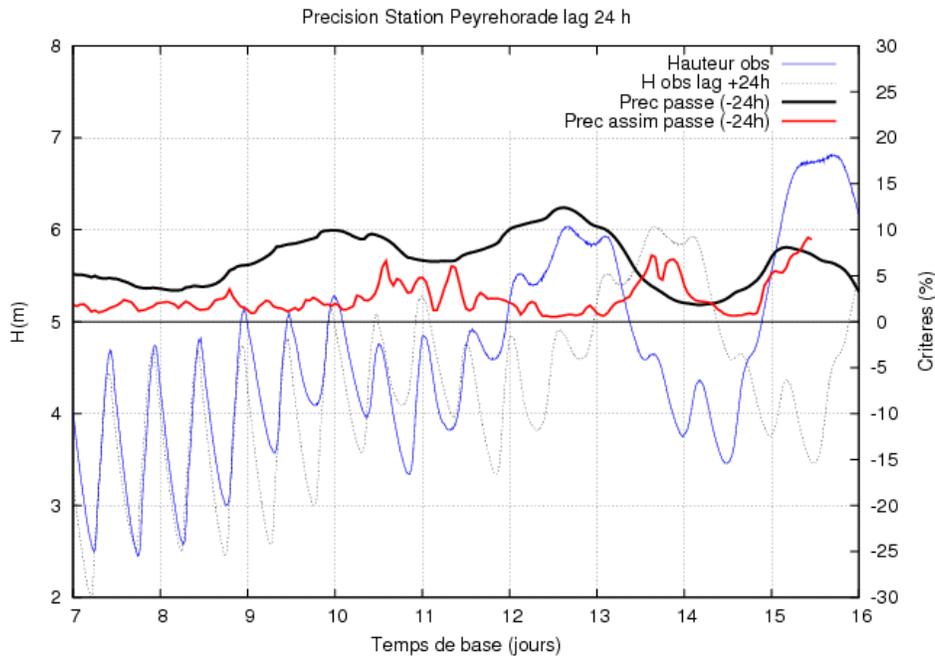
L'assimilation en mode ASSIM a apporté peu d'amélioration au modèle direct.

Episode ADOUR MARITIME 10 : novembre 2009

Le modèle direct fonctionne seulement avec des lignes d'eau initiales pré-simulées. Les observations ont ici un pas de temps de 5mn, moyennées sur une fenêtre de 25mn. L'assimilation combine le calage et l'assimilation linéique (CALAGEASSIM), où la loi hydrodynamique 1030 (à Escos) a dû être limitée supérieurement à $660\text{m}^3/\text{s}$.

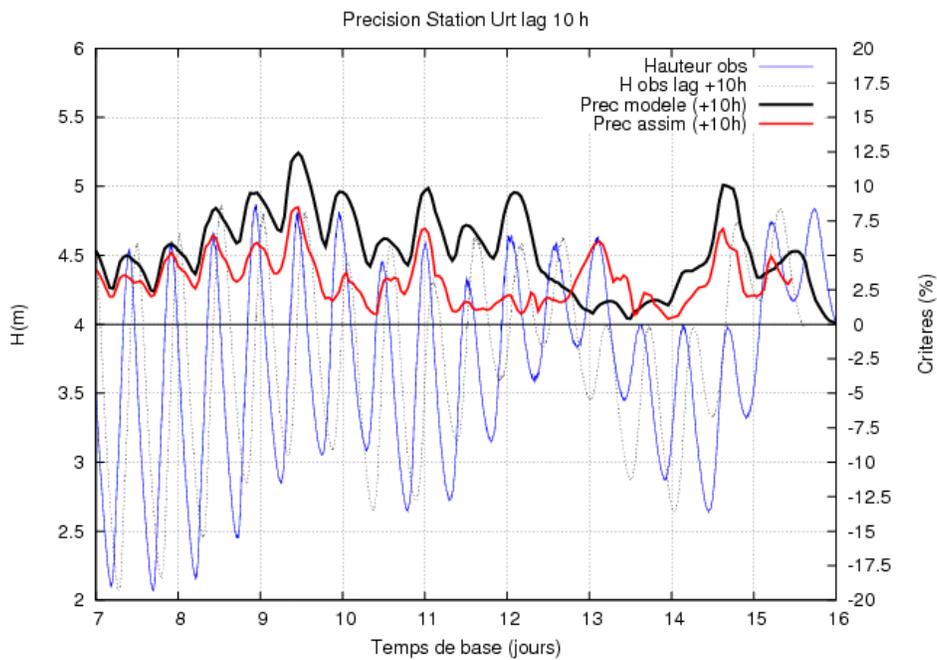
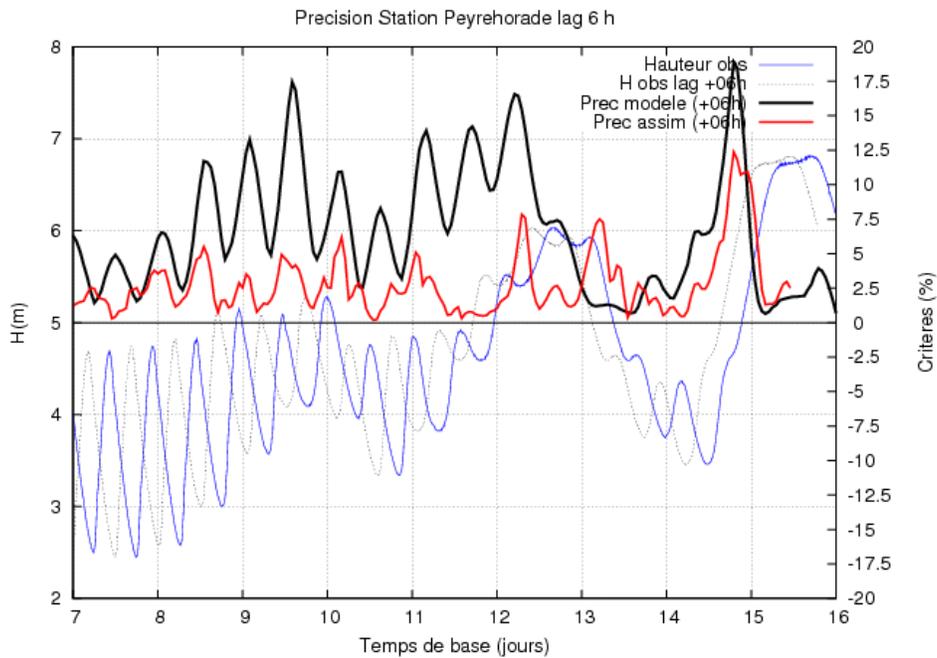
1- Précisions en mode simulation

La précision sur le passé est améliorée par l'assimilation sur quasiment tout l'épisode aux deux stations.



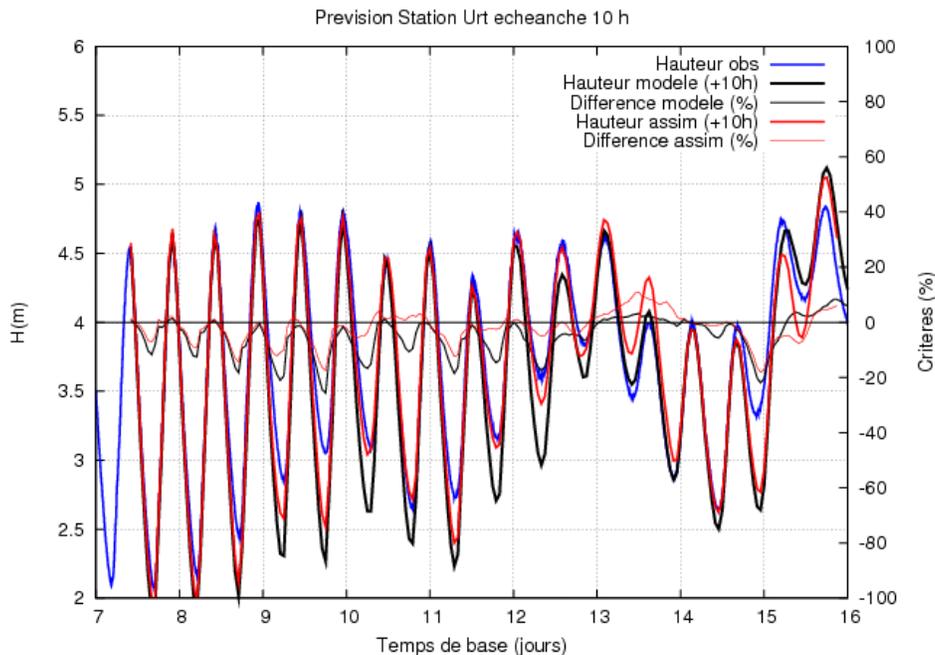
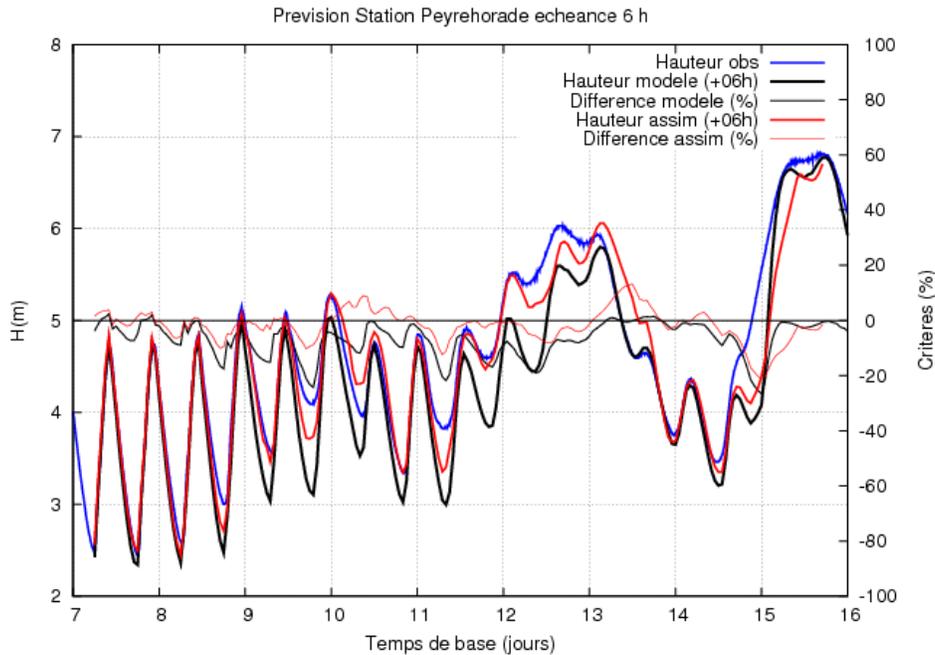
2- Précisions en mode prévision

La précision sur la prévision est nettement améliorée par l'assimilation aux deux stations pour la montée au premier pic de crue.



3- Hauteurs prévues à l'échéance usuelle

Les prévisions de hauteurs avec assimilation sont nettement meilleures pour le premier pic de crue aux deux stations. Le deuxième pic est néanmoins moins bien simulé à la station Peyrehorade.



L'assimilation en mode CALAGEASSIM a amélioré nettement la prévision du premier pic.

Conclusions

La combinaison des deux algorithmes d'assimilation donne de bons résultats en tirant les bénéfices de chacun des algorithmes. Si la qualité du modèle direct est satisfaisante, l'assimilation ne dégrade pas la prévision. Quand le modèle direct s'écarte des observations, l'assimilation apporte un gain certain de qualité sur la prévision à l'échéance.

Sur le modèle Marne Vallage, les essais d'améliorations apportées par l'assimilation ont mis en évidence un comportement contradictoire des stations Joinville et Chamouilley. Ce point a d'ailleurs été identifié par le SPC SAMA et une refonte de ce modèle a entre temps été réalisée.

Le modèle Adour Maritime gagne en robustesse d'une part avec l'initialisation par une ligne d'eau pré-calculée, permettant au modèle direct seul de sortir des résultats. D'autre part, pour les épisodes où le modèle direct fait néanmoins des plantages, l'assimilation linéique permet alors de simuler l'épisode complet avec de bons résultats. L'apport de l'assimilation est très significatif à la station Peyrehorade pour les épisodes 01, 03, 04 et 10.

Sur l'épisode 10 de Novembre-2009 un problème de plantage du modèle direct a été remarqué quand le débit de la loi 1030 passe au-delà de $1000\text{m}^3/\text{s}$. Ce plantage doit correspondre à un passage en torrentiel du régime de l'écoulement au niveau local.

Un passage en revue avec Etienne Le Pape de la base de données des résultats des runs AdourMaritime a permis d'identifier un seuil de débit à $660\text{m}^3/\text{s}$ au-delà duquel les runs ont planté. Ce seuil est donc imposé comme borne supérieure des débits pour la loi '1030'. Néanmoins les débits à cette station amont atteignent pour cet épisode la valeur de $1400\text{m}^3/\text{s}$ au passage du deuxième pic.

Une perspective d'amélioration du modèle pour éviter ces plantages sans borner l'hydrogramme amont serait d'activer localement le noyau transcritique permettant de simuler les régimes torrentiels en utilisant la formulation peu couteuse implicite (recommandation de N. Goutal, EDF/LNHE, Journée modélisation Mascaret du 30/11/2009).

Une perspective d'amélioration de la maquette d'assimilation de données serait de mettre en place une réutilisation des lois calées d'une simulation à l'autre pour améliorer en particulier l'assimilation des observations à fréquence de 5mn.



**Master 2 Professionnel Ingénierie
Mathématique à Toulouse**

Rapport de Stage

***Assimilation de données pour un
modèle d'hydrologie fluviale***

Chloé IVANOFF

Stage effectué sous la direction de :
Olivier THUAL, Project Leader

Toulouse, le 7 Septembre 2009

Remerciements

Je voudrais tout d'abord exprimer ma profonde reconnaissance à mon tuteur Mr Olivier THUAL pour ses nombreuses explications et le temps qu'il m'a accordé malgré son emploi du temps chargé. Ses conseils et ses commentaires m'ont permis de surmonter mes difficultés et de mener à bien ce projet.

Je tiens également à remercier Messieurs Andrea PIACENTINI et Gabriel JONVILLE pour leurs explications et leur grande disponibilité (et cela malgré la distance!) qui m'ont été d'une aide très précieuse tout au long de ce stage.

Je voudrais aussi exprimer mes remerciements à Étienne LE PAPE du SCHAPI, mais aussi à Sophie RICCI et aux "professeurs" de la formation PALM.

Je tiens tout particulièrement à remercier l'équipe CSG du CERFACS sans qui je n'aurais jamais rien pu faire!

Enfin à l'ensemble des membres du CERFACS et tout particulièrement à l'équipe "Modélisation du Climat et de son Changement Global" (GlobC), un grand merci pour l'accueil chaleureux dont ils ont fait preuve tout au long de ces six mois.

Table des matières

<i>Remerciement</i>	3
1 Introduction	9
2 Le CERFACS : Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique	10
3 Contexte général	11
3.1 L'hydrologie	11
3.1.1 Les équations de Saint Venant : cas parfait	11
3.1.2 Cas réel : forme non-conservative des équations de Saint Venant .	12
3.1.3 Forme conservative des équations de Saint Venant	13
3.2 L'Assimilation de Données (AD)	14
3.2.1 Notations	14
3.2.2 Méthode du BLUE	16
3.3 Problématique et objectifs	18
4 Expérience d'Assimilation de Données	20
4.1 Présentation du modèle	20
4.1.1 Résolution numérique	21
4.1.2 Application	21
4.2 Mise en place d'une méthode d'Assimilation de donnée sur la condition initiale	22
4.2.1 À un temps fixe	22
4.2.2 Tout les N pas de temps	24
4.2.3 Application : Assimilation de Données avec un débit amont aléatoire	27
4.3 Bilan	30
5 Étude de MASCARET	31
5.1 Présentation du logiciel	31
5.2 Présentation des cours d'eau servant de support	32
5.2.1 La Marne-Vallage	32
5.2.2 L'Adour-Maritime	33
5.3 Détermination des courbes de tarages	35
5.3.1 Pour la Marne	35
5.3.2 Pour l'Adour	37
5.4 Étude de la diffusion	39
5.4.1 Approche théorique de la diffusion	40

5.4.2	Expérience sur la Marne	42
5.4.3	Expérience sur l'Adour	44
6	Conclusion et perspectives	48
	<i>Références</i>	51
	<i>Annexe</i>	53

Table des figures

1	Le cycle de l'eau. USGS http://www.usgs.gov	11
2	Géométrie idéalisée 2D des équations de Saint Venant.	12
3	a) Vue longitudinale, b) Profil en travers.	13
4	Illustration du schéma centré; en bleu la condition initiale, en rouge la solution exacte issue du modèle.	22
5	Expérience d'AD sur la solution initiale avec $\sigma_{erreur_B} = \sigma_{erreur_R} = 50\%$	23
6	Expérience d'AD sur la solution initiale a) avec $\sigma_{erreur_B} = 20\%$, et $\sigma_{erreur_R} = 50\%$, b) avec $\sigma_{erreur_B} = 50\%$, et $\sigma_{erreur_R} = 20\%$	23
7	Expérience d'AD sur la solution initiale avec $\sigma_{erreur_B} = 10\%$, et $\sigma_{erreur_R} = 7\%$	24
8	Expérience d'AD sur la solution initiale avec correction à chaque pas de temps avec $\sigma_{erreur_B} = 50\%$, et $\sigma_{erreur_R} = 50\%$	24
9	Expérience d'AD sur la solution initiale avec CL aval.	25
10	Évolution de l'expérience d'AD sur la solution initiale avec correction à chaque pas de temps, $\sigma_{erreur_B} = \sigma_{erreur_R} = 50\%$: a) N itérations, b) $2 N$ itérations	25
11	Évolution de l'expérience d'AD sur la solution initiale avec correction à chaque pas de temps, $\sigma_{erreur_B} = \sigma_{erreur_R} = 50\%$: c) $3 N$ itérations, d) $4 N$ itérations.	26
12	Expérience d'AD sur la solution initiale avec correction à chaque pas de temps a) avec $\sigma_{erreur_B} = 20\%$, et $\sigma_{erreur_R} = 50\%$, b) avec $\sigma_{erreur_B} = 50\%$, et $\sigma_{erreur_R} = 20\%$	26
13	Évolution de l'expérience d'AD sur la solution initiale avec correction à chaque pas de temps, $\sigma_{erreur_B} = 7\%$ $\sigma_{erreur_R} = 6\%$: a) $t = N$, b) $t = 2 * N$, c) $t = 3 * N$ itérations, d) $t = 4 * N$ itérations.	27
14	Représentation du signal implanté comme condition amont.	28
15	schéma centré explicite avec CI nulle et CL.	28
16	Solution $\tilde{h}(x, t)$ pour $\alpha = \beta = 1$. Signal temporel en a) $x \approx L/16$, b) $x \approx L/5$	29
17	Évolution de l'expérience d'AD sur la solution initiale nulle avec correction à chaque pas de temps et CL, $\sigma_{erreur_B} = \sigma_{erreur_R} = 50\%$: a) N itérations, b) $2 N$ itérations.	29
18	Évolution de l'expérience d'AD sur la solution initiale nulle avec correction à chaque pas de temps et CL, $\sigma_{erreur_B} = \sigma_{erreur_R} = 50\%$: c) $3 N$ itérations, d) $4 N$ itérations.	29
19	courbe de tarage imposée en sortie.	33
20	Représentation de la marée imposée en aval.	35
21	Expérience effectuée sur la Marne.	35
22	Courbe de tarage expérimentale pour la Marne : rouge pour Joinville, bleu pour Chamouilley.	36
23	Comparaison courbe de tarage imposée (bleu) / courbe de tarage simulée (rouge).	36
24	Expérience effectuée sur l'Adour.	37
25	Relation hauteur débit pour l'expérience 1 : a) à Lesseps, b) à Urt.	37

26	Relation hauteur débit à Peyrehorade pour l'expérience 1.	38
27	Relation Z/Q pour deux cas de débit amont constant : expérience (1) et (11).	38
28	Courbe de tarage expérimentale : a) à Lesseps, b) à Urt.	39
29	Courbe de tarage expérimentale à Peyrehorade.	39
30	Graphique de la fonction erf.	41
31	Illustration de l'augmentation du débit à x_s	41
32	Évolution du débit / au temps a) à Joinville , b) à Chamouilley.	42
33	Zoom sur le débit simulé a) à Joinville, b) à Chamouilley.	43
34	Tableau récapitulatif des temps de la Marne.	44
35	Résultats du calcul de la diffusion pour la Marne.	44
36	Différence entre le débit simulé du jeu A et celui du jeu AB à Lesseps.	44
37	Différence entre le débit simulé du jeu A et celui du jeu AB à Lesseps.	45
38	Différence entre le débit simulé du jeu B et celui du jeu AB à Lesseps.	46
39	Tableau des temps relevés pour l'Adour.	46
40	Résultats du calcul de la diffusion pour l'Adour.	47
41	Différence entre les débits simulés à Urt pour l'exp 2 a) entre A - AB, b) entre B - AB.	47
42	Situation globale.	55
43	Limite de la zone d'étude pour la Marne.	55
44	Schéma représentant la Marne sous Mascaret.	56
45	Localisation de l'Adour	57
46	Schéma représentant l'Adour.	57

1 Introduction

Les modèles géophysiques actuels (atmosphère, océan, ...) sont à présent à même de pouvoir donner une représentation des processus considérés, que ce soit la dynamique (météorologie par exemple) ou les processus réactifs (chimie atmosphérique, biogéochimie). Ces modèles sont ensuite discrétisés et implémentés numériquement sous forme de “codes” permettant de délivrer des prévisions relativement exactes.

Pour réaliser de telles prévisions, ces modèles se nourrissent des données d’observations délivrées par les réseaux de mesures (que ce soit les réseaux météorologiques, les réseaux d’observations de la qualité de l’air, les satellites d’observation de la terre ...).

Les chaînes modernes de simulation sont fondées sur un couplage entre les modèles numériques et ces données observées afin notamment de mieux contraindre les modèles. Ces dernières comportent de nombreuses incertitudes, que ce soit sur les conditions initiales (dont l’impact est clé en météorologie), sur les paramétrisations (liées aux changements d’échelle ou à des processus mal connus) ou sur des données physiques.

En effet, les équations qui régissent les phénomènes physiques, même les plus simples, n’ont, en règle générale, pas de solution analytique rigoureuse du point de vue mathématique. Pour les résoudre, au prix d’approximations qu’il convient de savoir borner, on a généralement recours à deux techniques : soit en simplifier l’expression mathématique pour la ramener à des formulations plus simples, soit discrétiser les équations pour les résoudre par des méthodes de type “éléments finis”, soit bien souvent combiner les deux méthodes.

La discrétisation “brutale” effectuée mène souvent à des calculs non envisageables même par les moyens les plus puissants ; un premier niveau d’approximation permet alors fréquemment de rendre le résultat de la discrétisation concrètement calculable, au prix d’une puissance informatique souvent considérable. C’est dire que la solution de problèmes concrets très complexes, tels que la compréhension et la prévision des phénomènes climatiques, météorologiques, océaniques, même à faible résolution spatio-temporelle, nécessite une approche multidisciplinaire “fine” entre physiciens, spécialistes de mathématiques appliquées et informaticiens. Une réelle surenchère existe entre les besoins scientifiques, théoriques ou appliqués, et les progrès technologiques considérables de ces dernières années.

Pour répondre à ce besoin croissant du monde scientifique, une méthode mathématique a été développée : l’assimilation de données. Encore peu utilisée dans certains domaines, cette méthode a déjà fourni ses preuves dans le domaine de la météorologie.

L’objectif de ce stage va donc être, dans un premier temps d’illustrer les atouts de l’assimilation de données dans le cas de l’hydrologie sur un modèle académique, puis, dans un deuxième temps, de contribuer à la mise en place d’une chaîne d’assimilation dans un modèle réaliste.

2 Le CERFACS : Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique

Créé à Toulouse à la fin des années 80, le **CERFACS** est un organisme de recherche développe des méthodes avancées pour la simulation numérique et la résolution algorithmique. Faisant parti des meneurs mondiaux dans son domaine, il traite de vastes problèmes scientifiques et technologiques intéressant aussi bien la recherche que l'industrie, et nécessitant de ce fait l'accès aux ordinateurs les plus puissants existant actuellement.

Comportant sept actionnaires – le **CNES**, **EADS**, **EDF**, **Météo-France**, l'**ONERA**, **TOTAL** et **SNECMA** – le **CERFACS** accueille des équipes interdisciplinaires, tant pour la recherche que pour la formation avancée, qui sont composées de physiciens, mathématiciens appliqués, numériciens et encore ingénieurs en logiciel.

Une part de l'activité de recherche du **CERFACS** est associée au CNRS comme une "Unité de Recherche Associée" (SUC, URA 1875). Le **CERFACS** participe aussi au projet TVE (Terre Vivante et Espace) et est également membre du RTRA / STAE (Réseau Thématique de Recherche Avancée "Sciences et Technologies pour l'Aéronautique et l'Espace"). Il contribue de plus aux activités de AESE ("Pôle de compétitivité Aéronautique, Espace et Systèmes Embarqués").

Environ 100 personnes, dont 90 chercheurs et ingénieurs de tous horizons, travaillent au **CERFACS** sur des projets dans six grands domaines de recherche : l'algorithmique parallèle, l'aérodynamique, la combustion, le climat et l'environnement, l'assimilation de données et enfin l'électromagnétisme.

Ces domaines sont regroupés en quatre équipes : la modélisation du climat et du changement global, le traitement de l'image et du signal, les calculs de dynamique des fluides, l'électromagnétisme et le contrôle.

Le stage effectué se déroule au sein de l'équipe GLOBC (Global Change), équipe en charge de la modélisation du climat et du changement global. Créée en 1990, cette équipe jouit aujourd'hui d'une renommée internationale dans la modélisation du climat, la prévision saisonnière et l'assimilation de données.

3 Contexte général

3.1 L'hydrologie

L'hydrologie est la science de la terre qui s'intéresse au cycle de l'eau, c'est-à-dire aux échanges entre l'atmosphère, la surface terrestre et son sous-sol. La description et la compréhension du cycle de l'eau est très complexe et nécessite d'avoir recours à de nombreuses disciplines ou sous-disciplines scientifiques telles que la mécanique du solide, la mécanique des fluides ... etc

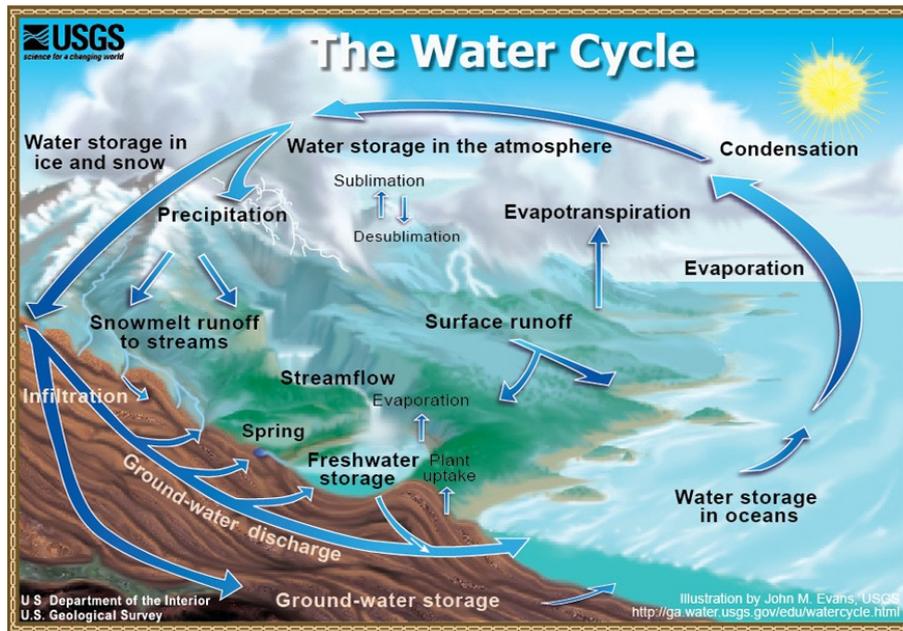


FIG. 1 – Le cycle de l'eau. USGS <http://www.usgs.gov>

Nous nous intéresserons ici à ce que l'on appelle l'hydrologie de surface. Cette science étudie en détail les phénomènes de ruissellement, les écoulements des cours d'eau et les inondations. La modélisation de ces milieux repose principalement sur les modèles de mécanique des fluides. La modélisation 3D, avec prise en compte de la turbulence, est la plus précise, mais requiert cependant beaucoup de temps de calcul pour obtenir des résultats corrects. C'est pourquoi on a recours à un modèle 1D qui permet de faire évoluer les champs moyen sur une section (vitesse débit ...). L'écoulement à surface libre des eaux observé dans les rivières est régi par le système d'équations différentielles non linéaires de Barré de Saint Venant (1871).

Dans de nombreuses situations, les conditions de l'hydrologie permettent de reformuler en les simplifiant ces équations de l'hydraulique en cherchant notamment à les linéariser.

3.1.1 Les équations de Saint Venant : cas parfait

Le modèle mathématique de Saint-Venant découle de l'intégration des équations de Navier-Stokes en supposant que la profondeur du cours d'eau soit faible devant l'échelle de variation des phénomènes dans la direction de l'écoulement. On appelle également ce modèle "modèle des

eaux peu profondes” ou encore “équations d’ondes longues” (*shallow water equations*).

Nous allons ici présenter de manière rapide les notions qui seront utiles à notre étude.

Pour une géométrie idéalisée (cf Figure 2), les équations de Saint Venant s’écrivent sous la forme :

$$\left\{ \begin{array}{l} \frac{\partial h}{\partial t} + U \frac{\partial h}{\partial x} = -h \frac{\partial U}{\partial x}, \\ \frac{\partial U}{\partial t} + U \frac{\partial U}{\partial x} = -g' \frac{\partial h}{\partial x} + gI - \frac{C_f U |U|}{2h}, \end{array} \right. \quad (1)$$

où $I = \sin \gamma$ et $g' = g \cos \gamma$; γ étant l’angle entre l’axe x et l’horizontal.

De plus :

- $g' \frac{\partial h}{\partial x}$ représente le gradient de pression en x ,
- gI la force de pesanteur en \mathbf{x} ,
- $C_f(h, U)$ modélise le frottement sur la paroi.

La formule de Manning Strickler souvent utilisée en hydraulique à surface libre s’écrit :

$$C_f(h) = \frac{2g}{K_s^2 h^{\frac{1}{3}}} = \frac{\Phi_{MS}}{4} \left(\frac{k_s}{h} \right)^{\frac{1}{3}}, \quad (2)$$

où K_s est le nombre de Strickler (en $m^{\frac{1}{3}} s^{-1}$), k_s la rugosité du fond et Φ_{MS} un nombre sans dimension (on peut le choisir de l’ordre de $\Phi_{MS} = 0.2$).

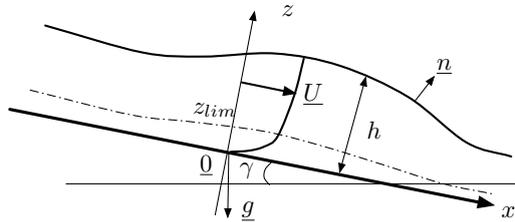


FIG. 2 – Géométrie idéalisée 2D des équations de Saint Venant.

3.1.2 Cas réel : forme non-conservative des équations de Saint Venant

L’écoulement dans une rivière peut généralement être considéré comme monodimensionnelle, c’est-à-dire comme fonction d’une seule variable d’espace, l’abscisse curviligne x du lit de la rivière. Pour une géométrie réaliste (cf Figure 3), on peut exprimer les équations de Saint Venant 1D à l’aide des variables Q , le débit de la rivière et S , la section mouillée. Ces paramètres sont liés par la relation suivante : $Q = US$, où U est la vitesse.

Hypothèses choisies :

On va se placer dans un cas où la pression du fluide est hydrostatique et où les effets visqueux internes sont négligeables devant les frottements sur les parois du lit de la rivière.

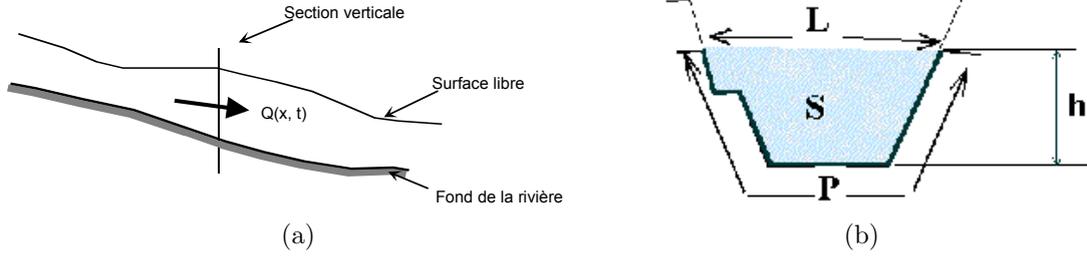


FIG. 3 – a) Vue longitudinale, b) Profil en travers.

Les équations de Saint Venant sont alors équivalentes à :

$$\begin{cases} \frac{\partial S}{\partial t} + \frac{\partial Q}{\partial x} = q_a, \\ \frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(\frac{Q^2}{S} \right) + gS \frac{\partial Z}{\partial x} = gSJ, \end{cases} \quad (3)$$

où $Z = Z(x, S)$ est la cote de la surface libre, $S(x, t)$ la surface mouillée, g la gravité, J la perte de charge linéaire par frottement et $Q(x, t)$ le débit. Le terme q_a représente ici les apports de débit. J est évalué par la formule de Strickler :

$$J = \frac{Q^2}{K^2 S^2 R_h^{\frac{4}{3}}}. \quad (4)$$

K étant le coefficient de Strickler et R_h le rayon hydraulique tel que $R_h = \frac{S}{P}$, où P est le périmètre mouillé.

Caractéristique

Ces équations peuvent s'obtenir à partir des invariants de Riemann le long de courbes caractéristiques. Cependant, leur calcul est long et complexe et présente peu d'intérêt pour notre étude actuelle.

3.1.3 Forme conservative des équations de Saint Venant

Sous forme conservative les équations de Saint Venant s'écrivent de la manière suivante :

$$\begin{cases} \frac{\partial S}{\partial t} + \frac{\partial Q}{\partial x} = 0, \\ \frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(\frac{Q^2}{S} + P \right) = g \int_0^{\mathcal{H}(x,S)} \left(\frac{\partial S}{\partial x} \right)_h dh - gS \frac{\partial Z_f}{\partial x} - gSJ(S, Q), \end{cases} \quad (5)$$

où $P = g \int_0^{\mathcal{H}(x,S)} S(x, h) dh$ le terme de pression, $S = S(x, h)$ et $H = \mathcal{H}(x, S)$. Z_f désigne la cote du fond (de la rivière).

On a Q et S fonctions de x et t (pour simplifier on les laissera tels quels). où S désigne non plus la section mouillée mais la section géométrique. De plus, h représente la section verticale (la hauteur).

On va poser : $g \int_0^{\mathcal{H}(S)} \left(\frac{\partial S}{\partial x} \right)_h [x, \mathcal{H}(x, S)] dh - g S \frac{\partial Z_f}{\partial x} - g S J(S, Q) = E(S, Q)$.

Ce système d'équation peut encore, pour une dimension d'espace, se mettre sous la forme :

$$\frac{\partial \underline{W}}{\partial t} + \frac{\partial \underline{F}}{\partial x} = \underline{B}, \quad (6)$$

où $\underline{W} = \begin{pmatrix} S \\ Q \end{pmatrix}$ est le vecteur d'état, $\underline{F} = \begin{pmatrix} Q \\ \frac{Q^2}{S} + P \end{pmatrix}$ le flux, et $\underline{B} = \begin{pmatrix} 0 \\ E(S, Q) \end{pmatrix}$ le terme source.

Cette nouvelle formulation, quoique complexe à obtenir, est utilisée par le code numérique MAS-CARET.

3.2 L'Assimilation de Données (AD)

L'assimilation de données est une technique d'analyse mathématique qui consiste à réunir les observations du milieu avec les modèles qui permettent de le décrire. Il s'agit de déterminer l'état instantané du milieu modélisé et, très souvent, d'en prédire l'évolution. Cette technique utilise les équations discrétisées de la physique ainsi que les observations sur le système afin d'en donner une représentation aussi fidèle et cohérente que possible.

En effet, aux erreurs causées par la modélisation mathématique s'ajoute une erreur non négligeable associée à la méconnaissance du jeu des conditions initiales, des conditions aux limites et des paramètres initiaux. Aussi, on constate que la simulation numérique fournit fréquemment des résultats éloignés de la réalité physique du système. L'objectif de l'assimilation de données va donc être d'exploiter toutes les sources d'informations concernant le système afin de diminuer les incertitudes des sorties des simulations numériques et de pouvoir améliorer notre connaissance sur le système physique. De ce fait, la théorie de l'assimilation de données est fortement liée à la théorie d'estimation linéaire optimale.

3.2.1 Notations

- On souhaite effectuer la description optimale de l'état du système physique étudié. Pour cela on va poser :

- \underline{x} vecteur de dimension \mathbf{n} : vecteur de contrôle, encore appelé **vecteur d'état** ou champ analysé. Ce vecteur peut contenir l'état d'un modèle de simulation numérique, ses paramètres de calages ou les deux à la fois.
- \underline{y}^0 vecteur de dimension \mathbf{p} : ensembles des **observations** disponibles (mesures...).

On suppose que l'on dispose d'une estimation a priori des variables qui décrivent le système étudié que l'on notera \underline{x}^b . On appelle encore \underline{x}^b l'**ébauche** du système, de même dimension que le vecteur d'état.

On considère \underline{x}^t l'**état vrai** du système.

On va chercher à déterminer par le processus d'assimilation de données le meilleur estimateur de \underline{x}^t appelé l'**analyse** et notée \underline{x}^a .

- Afin de corriger la valeur des variables de contrôle, on va introduire un **opérateur d'observation**, à priori non linéaire, que l'on va noter \mathcal{H} , qui fait passer du vecteur de contrôle \underline{x} aux observations simulées $y = \mathcal{H}(x)$. Dans le cas linéaire, on le notera \mathbf{H} . Cet opérateur permet de relier l'ébauche aux observations.

On peut donc écrire \underline{y}^0 sous la forme :

$$\underline{y}^0 = \mathcal{H}(\underline{x}^b), \quad (7)$$

où $\mathcal{H}(\underline{x}^b)$ est la projection de l'état \underline{x}^b sur l'espace des observations.

On considère que l'ébauche et les observations ne sont pas parfaites, on va donc leur associer une erreur (ou incertitude).

- Le but étant de réduire l'écart entre l'état du système et l'état vrai, on va donc introduire deux matrices qui vont contenir respectivement les variances et les covariances des erreurs d'ébauches et d'observations.

– Notons $\underline{\underline{B}}$ la **matrice de covariance des erreurs d'ébauche**, de dimension $\mathbf{n} \times \mathbf{n}$:

$$\underline{\underline{B}} = \begin{bmatrix} VAR(\epsilon_1^b) & \dots & COV(\epsilon_1^b, \epsilon_n^b) \\ \vdots & VAR(\epsilon_j^b) & \vdots \\ COV(\epsilon_n^b, \epsilon_1^b) & \dots & Var(\epsilon_n^b) \end{bmatrix} \quad (8)$$

$\underline{\epsilon}^b = \underline{x} - \underline{x}^b$ est l'erreur associée à l'ébauche et $\underline{\epsilon}_j^b$ ses composantes.

Ici les éléments diagonaux de la matrice correspondent aux variances de l'erreur liée aux variables de contrôle, et les termes extra-diagonaux représentent les covariances entre erreurs liées à deux variables de contrôles distinctes.

– Soit $\underline{\underline{R}}$ la **matrice des variances-covariances des erreurs d'observations** \underline{y}^0 , liée à la précision des instruments, de dimension $\mathbf{p} \times \mathbf{p}$:

$$\underline{\underline{R}} = \begin{bmatrix} VAR(\epsilon_1^o) & \dots & COV(\epsilon_1^o, \epsilon_p^o) \\ \vdots & VAR(\epsilon_j^o) & \vdots \\ COV(\epsilon_p^o, \epsilon_1^o) & \dots & Var(\epsilon_p^o) \end{bmatrix} \quad (9)$$

$\underline{\epsilon}^0 = \underline{y} - \underline{y}^0$ est l'erreur liée aux observations et $\underline{\epsilon}_j^0$ ses composantes.

Les termes diagonaux correspondent aux variances de l'erreur liées aux observations, et les termes extra-diagonaux représentent les covariances entre les erreurs liées à des observations distinctes.

- De même, on va associer à l'analyse une erreur définie comme l'écart entre l'analyse \underline{x}^a et l'état vrai \underline{x}^t . On appellera cette erreur l'**erreur d'analyse**.

$$\underline{\epsilon}^a = \underline{x}^a - \underline{x}^t. \quad (10)$$

C'est cette erreur que l'on va essayer de réduire via l'assimilation de données.

- **L'analyse**

L'étape d'analyse consiste à mettre à jour toutes les données représentatives du système physique d'étude pour s'approcher des observations disponibles en prenant en compte toutes les sources d'erreur. On va chercher à identifier une solution optimale pour la description du système en utilisant au mieux le modèle mathématique, les observations, l'ébauche et les erreurs sur ces informations.

3.2.2 Méthode du BLUE

On va choisir ici une méthode de type séquentielle qui consiste à réajuster le modèle à chaque fois qu'une observation est disponible.

L'algorithme d'assimilation identifie, à l'instant d'analyse t_i , un état analysé \underline{x}^a de la variable aléatoire \underline{x} dépendant linéairement de l'ébauche \underline{x}^b et des observations \underline{y}^0 à cet instant.

$$\underline{x}^a = \underline{L} \underline{x}^b + \underline{K} \underline{y}^0. \quad (11)$$

On cherche \underline{L} et \underline{K} de manière à trouver une analyse optimale. L'état analysé \underline{x}^a approxime l'état vrai \underline{x}^t . On a

$$\begin{aligned} \underline{x}^b &= \underline{x}^t + \underline{\epsilon}^b, \\ \underline{y}^t &= \underline{H} \underline{x}^t, \\ \underline{y}^0 &= \underline{y}^t + \underline{\epsilon}^0, \\ \underline{x}^a &= \underline{x}^t + \underline{\epsilon}^a. \end{aligned} \quad (12)$$

Supposons à présent que les erreurs liées à l'ébauche et aux observations sont non-biaisées, c'est-à-dire $\mathbf{E}(\underline{\epsilon}^0) = \mathbf{0}$ et $\mathbf{E}(\underline{\epsilon}^b) = \mathbf{0}$. On recherche de plus un estimateur non-biaisé donc $\mathbf{E}(\underline{\epsilon}^a) = \mathbf{0}$.

En substituant les équations (14) et (15), on obtient :

$$\begin{aligned} \underline{x}^t + \underline{\epsilon}^a &= \underline{L} \underline{x}^t + \underline{L} \underline{\epsilon}^b + \underline{K} \underline{y}^t + \underline{K} \underline{\epsilon}^0, \\ &= \underline{L} \underline{x}^t + \underline{L} \underline{\epsilon}^b + \underline{K} \underline{H} \underline{x}^t + \underline{K} \underline{\epsilon}^0. \end{aligned} \quad (13)$$

Prenons l'espérance mathématique de (16)

$$E[\underline{x}^t] = \underline{L}E[\underline{x}^t] + \underline{K} \underline{H}E[\underline{x}^t], \quad (14)$$

qui conduit à

$$\begin{aligned} \underline{I} &= \underline{L} + \underline{K} \underline{H}, \\ \underline{L} &= \underline{I} - \underline{K} \underline{H}. \end{aligned} \quad (15)$$

On peut donc écrire en utilisant (18) dans (14) :

$$\underline{x}^a = \underline{x}^b + \underline{K} [y^0 - \underline{H} \underline{x}^b]. \quad (16)$$

Pour tout choix de l'opérateur linéaire \underline{K} .

La matrice de covariance d'erreur d'analyse \underline{A} est obtenue en développant $E[\underline{\epsilon}^a \underline{\epsilon}^{aT}]$ (Bouttier et Courtier, 1999) :

$$\underline{A} = (\underline{I} - \underline{K} \underline{H}) \underline{B} (\underline{I} - \underline{K} \underline{H})^T + \underline{K} \underline{R} \underline{K}^T. \quad (17)$$

La trace de la matrice $\underline{A} = E[(\underline{x}^a - \underline{x}^t)(\underline{x}^a - \underline{x}^t)^T] = E[\underline{\epsilon}^a \underline{\epsilon}^{aT}]$ représente la variance d'erreur de l'analyse. On cherche \underline{x}^a le plus près possible de l'état vrai \underline{x}^t au sens optimal, c'est-à-dire que \underline{x}^a minimise la trace de \underline{A} donnée par :

$$Tr(\underline{A}) = Tr(\underline{B}) + Tr(\underline{K} \underline{H} \underline{B} \underline{H}^T \underline{K}^T) - 2Tr(\underline{B} \underline{H}^T \underline{K}^T) + Tr(\underline{K} \underline{R} \underline{K}^T). \quad (18)$$

Cette fonction scalaire est une fonction continue dérivable des coefficients de \underline{K} (un polynôme de degré 2) ; on peut calculer et annuler la dérivée de $Tr(\underline{A})$ par rapport à \underline{K} pour trouver (Bouttier et Courtier, 1999) :

$$\underline{K} = \underline{B} \underline{H}^T (\underline{H} \underline{B} \underline{H}^T + \underline{R})^{-1}, \quad (19)$$

soit en remplaçant dans (19)

$$\underline{x}^a = \underline{x}^b + \underline{B} \underline{H}^T (\underline{H} \underline{B} \underline{H}^T + \underline{R})^{-1} [y^0 - \underline{H}(\underline{x}^b)]. \quad (20)$$

Pour cette valeur de \underline{K} , la matrice de covariance d'erreur d'analyse \underline{A} s'écrit (avec (20) et (22)) :

$$\underline{A} = (\underline{I} - \underline{K} \underline{H}) \underline{B}. \quad (21)$$

On appelle \underline{K} la **matrice de gain**. Le vecteur \underline{x}^a ainsi défini est appelé le **Meilleur Estimateur Linéaire Non Biisé (Best Linear Unbiased Estimate : BLUE)**. On remarque que l'analyse \underline{x}^a est la somme de l'ébauche et d'un terme correctif que l'on va appeler **incrément d'analyse** ou encore **vecteur d'innovation**. On le note $\underline{d} = y^0 - \underline{H}(\underline{x}^b)$.

3.3 Problématique et objectifs

Ce stage vient s'inscrire dans une optique de collaboration entre le CERFACS et le SCHAPI (Service Central d'Hydrométéorologie et d'Appui à la Prévention des Inondations).

Depuis de nombreuses années le SCHAPI utilise un modèle 1D nommé MASCARET afin de simuler le comportement des écoulements (crues, inondations, ...). Des améliorations du potentiel de prévision à l'aide de ce logiciel sont envisageables pour l'analyse des phénomènes étudiés. Les ingénieurs du SCHAPI effectuent donc des recherches sur divers points dans le but d'une évolution de Mascaret :

- permettre une meilleure prise en compte des observations afin de pouvoir corriger le modèle,
- diminuer les incertitudes existantes sur les variables,
- améliorer les prévisions (surtout au niveau des pics de crues).

Face à ces attentes, l'assimilation de données se place comme une solution intéressante.

Au travers de cette collaboration avec le SCHAPI, le CERFACS va chercher à montrer les atouts que peut avoir de l'assimilation de données en hydrologie, afin de permettre à long terme une généralisation de ces méthodes.

On va donc contruire une maquette d'assimilation de données sur le logiciel Mascaret en s'appuyant sur l'étude de deux cours d'eau : la Marne et l'Adour.

Après une étude approfondie du logiciel et tout particulièrement de son code, une méthode d'assimilation de type BLUE incrémentale a été implémentée au coeur du noyau de Mascaret.

Puis, par un souci de performance, la maquette ainsi contruite a été installée sous le coupleur PALM, parfaitement adapté à l'implémentation d'une chaîne d'assimilation de données. En effet, une application en assimilation de données peut être conçue comme un couplage dynamique d'un modèle et de quelques unités de traitement des observations et des statistiques d'erreur, plus des traitements algébriques. Les différentes tâches d'une méthode d'assimilation de données (intégrer le modèle de prévision, appliquer l'opérateur d'observation, calculer l'écart aux observations, estimer les statistiques d'erreur, inverser des matrices, minimiser une fonction coût, etc) sont vues comme des morceaux indépendants de code qui vont être assemblés dans la structure logicielle portable, flexible et efficace qu'est PALM.

La difficulté de ce travail va être de pouvoir renseigner avec justesse les coefficients des matrices \underline{B} et \underline{R} c'est-à-dire de pouvoir déterminer la covariance existante entre les variables du vecteur de contrôle ainsi que leur covariance spatiale.

Ce stage va s'articuler autour de deux objectifs.

Dans un premier temps nous allons nous livrer à un travail expérimental qui mettra en valeur les apports qu'apporte une méthode d'assimilation de données sur un modèle simple de mécanique des

fluides.

Puis dans un second temps, nous allons effectuer une étude approfondie des cours d'eau considérés pour la construction de la maquette Mascaret-Palm afin d'être à même de pouvoir déterminer par la suite les éléments nécessaires à l'Assimilation de Données.

4 Expérience d'Assimilation de Données

Afin de démontrer l'intérêt de l'assimilation de données en hydrologie, nous allons nous livrer dans cette partie à l'étude suivante. À partir des équations de Saint Venant, base de l'écoulement surface libre des eaux, nous allons définir un modèle 1D. Nous appliquerons sur ce modèle une méthode d'assimilation de données de type BLUE (cf §3.2.2).

Puis nous effectuerons divers expériences sur la maquette ainsi construite dans le but de mettre en évidence le rôle prédominant des paramètres caractéristiques des matrices \underline{B} et \underline{R} .

4.1 Présentation du modèle

Reprenons les équations de Saint Venant présentées plus haut pour une géométrie idéalisée :

$$\left\{ \begin{array}{l} \frac{\partial h}{\partial t} + U \frac{\partial h}{\partial x} = -h \frac{\partial U}{\partial x} \\ \frac{\partial U}{\partial t} + U \frac{\partial U}{\partial x} = -g' \frac{\partial h}{\partial x} + gI - \frac{C_f U |U|}{2h} \end{array} \right. \quad (22)$$

Nous allons nous placer dans le cas d'un régime stationnaire, i.e. $U > 0$.

L'équilibre (h_n, U_n) vérifie alors :

$$U_n = K_s I^{1/2} h_n^{2/3}. \quad (23)$$

Lorsque la pente ainsi que la vitesse de la rivière deviennent plus faible, on va considérer le modèle de "l'approximation des ondes de crues diffusantes" suivant (O. Thual, 2009) :

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(U h) = 0 \quad \text{et} \quad 0 = -g' \frac{\partial h}{\partial x} + gI - \frac{C_f U |U|}{2h}. \quad (24)$$

On pose :

$$h = h_n + \tilde{h} \quad \text{et} \quad U = U_n + \tilde{U}, \quad (25)$$

\tilde{h} et \tilde{U} étant des petites perturbations de l'équilibre (h_n, U_n) .

En linéarisant et en éliminant \tilde{U} , \tilde{h} vérifie

$$\frac{\partial \tilde{h}}{\partial t} + \frac{5 U_n}{3} \frac{\partial \tilde{h}}{\partial x} = k_n \frac{\partial^2 \tilde{h}}{\partial x^2}, \quad (26)$$

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

L'effet du terme supplémentaire va être de diminuer l'amplitude de la crue en la diffusant. Si la condition initiale $\tilde{h}(x, 0)$ est une gaussienne de variance l_0 et d'amplitude maximale \tilde{h}_m , la solution est la gaussienne

$$\tilde{h}(x, t) = \tilde{h}_m \frac{l_0}{l(t)} \exp \left[-\frac{(x - \frac{5U_n}{3}t)^2}{2l^2(t)} \right] \quad \text{avec} \quad l^2(t) = l_0^2 + 2 k_n t, \quad (27)$$

de variance $l(t)$ croissante et d'amplitude maximale décroissante.

En posant $\alpha = \frac{5 U_n}{3}$ et $\beta = k_n$, l'équation devient :

$$\frac{\partial \tilde{h}}{\partial t} + \alpha \frac{\partial \tilde{h}}{\partial x} - \beta \frac{\partial^2 \tilde{h}}{\partial x^2} = 0. \quad (28)$$

On appelle aussi cette équation l'équation d'advection-diffusion.

En prenant comme hypothèses $\beta > 0$, $x \in \mathbb{R}$ et $t > 0$ et en supposant comme condition initiale $\tilde{h}(x, 0) = \tilde{h}_0(x)$ une fonction régulière, on va se placer dans le cadre d'une équation d'évolution du 2^{nd} ordre.

4.1.1 Résolution numérique

Afin de résoudre ce problème de façon numérique, on va le discrétiser en temps et en espace. On va donc introduire un pas spatial δx , c'est-à-dire que l'on va chercher à calculer notre solution \tilde{h} aux points d'abscisses $x = 0, \delta x, 2 \delta x, \dots$

De plus, on fixe un pas temporel δt et l'on cherchera à calculer \tilde{h} aux instants $t = \delta t, 2 \delta t, 3 \delta t \dots$

On notera \tilde{h}_j^n la valeur calculée de la solution \tilde{h} à l'instant $n\delta t$ ($n > 0$) et au point d'abscisse $j\delta x$ ($0 \geq j \geq \frac{1}{\delta x}$).

On va appliquer ici le schéma explicite centré suivant :

$$\frac{\tilde{h}_j^{n+1} - \tilde{h}_j^n}{\delta t} + \alpha \frac{\tilde{h}_{j+1}^n - \tilde{h}_{j-1}^n}{2 \delta x} - \beta \frac{\tilde{h}_{j+1}^n - 2 \tilde{h}_j^n + \tilde{h}_{j-1}^n}{\delta x^2} = 0, \quad (29)$$

encore équivalent à

$$\tilde{h}_j^{n+1} = \tilde{h}_j^n - \alpha \frac{\delta t}{2 \delta x} (\tilde{h}_{j+1}^n - \tilde{h}_{j-1}^n) + \beta \frac{\delta t}{\delta x^2} (\tilde{h}_{j+1}^n - 2 \tilde{h}_j^n + \tilde{h}_{j-1}^n). \quad (30)$$

Ce schéma numérique est L^∞ stable sous les conditions suivantes (cf. annexe) :

$$\beta \frac{\delta t}{\delta x^2} \leq \frac{1}{2} \quad \text{et} \quad \alpha \frac{\delta t}{\delta x} \leq 1. \quad (31)$$

4.1.2 Application

On se place sur le domaine $[a, b]$ où $a = -20$, $b = 20$. On notera L la longueur du domaine ($L = 40$). On discrétise avec le pas d'espace $dx = 0.05$ et le pas de temps dt respectant les conditions de stabilité énoncée précédemment. On aura donc un maillage de taille $M = L/dx + 1 = 801$.

On va prendre pour conditions aux limites :

$$\tilde{h}(a, t) = \tilde{h}(b, t) = 0. \quad (32)$$

La condition initiale sera sous forme de gaussienne centrée en zéro et d'écart-type $l_0 = 5$. On note $h_m = 0.5$ le maximum de la condition initiale.

On choisit pour l'instant de poser $\alpha = \beta = 1$. On supposera pour la suite que la solution fournie par le schéma numérique est exacte. On effectuera les calculs jusqu'à $t_f = 10$.

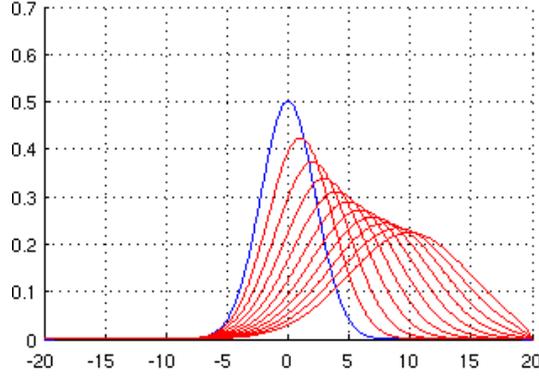


FIG. 4 – Illustration du schéma centré; en bleu la condition initiale, en rouge la solution exacte issue du modèle.

4.2 Mise en place d'une méthode d'Assimilation de donnée sur la condition initiale

On effectue une expérience d'assimilation de donnée sur le problème précédent en se basant sur le principe des expériences jumelles.

4.2.1 À un temps fixe

Dans un premier temps on choisit ici d'appliquer une méthode d'assimilation de données sur la condition initiale $\tilde{h}_0(x)$ du schéma centré défini précédemment au temps $t = 0$.

Notre vecteur de contrôle est donc être $\underline{x} = [\tilde{h}_0(x_1), \dots, \tilde{h}_0(x_M)]$ où (x_1, x_2, \dots, x_M) définissent le maillage. Soit $\underline{x}^t = [\tilde{h}_0^t(x_i)]_{i=1,M}$ la gaussienne choisie comme condition initiale (CI) pour le schéma explicite centré exposé précédemment. On effectue 3 observations de \tilde{h}_0 aux points d'abscisse -10 , 0 et 10 du domaine. On simule ces observations en ajoutant à la solution exacte du problème un bruit blanc gaussien de variance 0.1 et d'espérance 0 . On dispose d'une ébauche \underline{x}^b . Pour notre expérience on va choisir cette ébauche en ajoutant à \underline{x}^t une erreur légèrement bruitée (Figure 5 : courbe verte)

On prendra les erreurs d'ébauche et d'observation sous forme de Gaussienne spatiale centrée aux observations et de variance $\sigma = 5$. On choisit ici d'implanter \underline{B} avec des écarts types d'erreurs égaux à $\sigma_{erreur_B} = 50\%$ de la valeur de l'ébauche \tilde{x}_b , et \underline{R} avec des écarts types d'erreur tous égaux à $\sigma_{erreur_R} = 50\%$.

Résultats obtenus

On va dans un premier temps regarder ce qu'il se passe uniquement pour la solution initiale.

On a le graphique suivant :

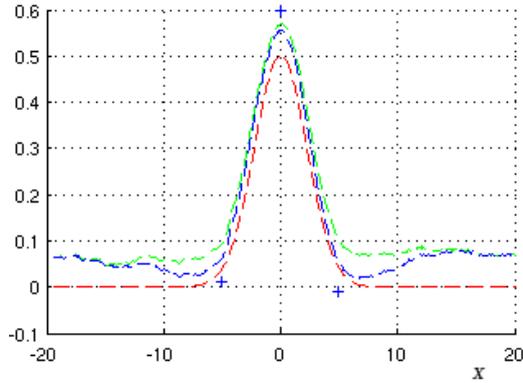


FIG. 5 – Expérience d’AD sur la solution initiale avec $\sigma_{erreur_B} = \sigma_{erreur_R} = 50\%$.

L’ébauche (**en vert**) est la courbe la plus basse obtenue. En la comparant avec la courbe de la solution h_0 supposée exacte (**en rouge**) qui est une gaussienne, on constate qu’il y a une différence assez conséquente. La courbe **bleu** représente la solution calculée par l’assimilation de données. On observe bien qu’une légère correction est apportée.

En jouant sur la précision des observations ainsi que sur les caractéristiques des matrices $\underline{\underline{B}}$ et $\underline{\underline{R}}$ on peut ajuster au mieux notre analyse et ainsi apporter la meilleure correction possible. Ainsi on observe que plus on diminue l’écart type d’erreurs de la matrice $\underline{\underline{B}}$ plus l’analyse fera confiance à l’ébauche. (Figure 6 (a)) De même, plus on diminuera l’écart type d’erreurs de la matrice $\underline{\underline{R}}$, plus on accordera de l’importance aux observations (Figure 6 (b)).

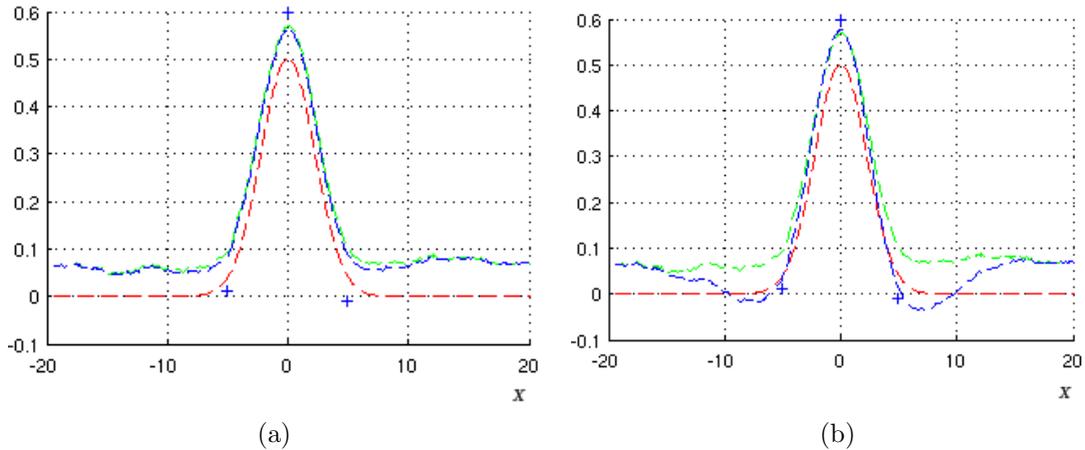


FIG. 6 – Expérience d’AD sur la solution initiale a) avec $\sigma_{erreur_B} = 20\%$, et $\sigma_{erreur_R} = 50\%$, b) avec $\sigma_{erreur_B} = 50\%$, et $\sigma_{erreur_R} = 20\%$.

Outre nous montrer l’importance de renseigner fidèlement les matrices $\underline{\underline{B}}$ et $\underline{\underline{R}}$ on remarque que l’expérience qui nous donne la figure (6) nous donne de meilleur résultat celle qui donne la figure (7). On peut donc dire que l’on a choisi des observations plus fiable que l’ébauche. Pour que l’assimilation fournisse le meilleur résultat possible, il aurait fallu implanter $\sigma_{erreur_R} = 7\%$ et

$\sigma_{erreur_B} = 10\%$. On aurait alors obtenu :

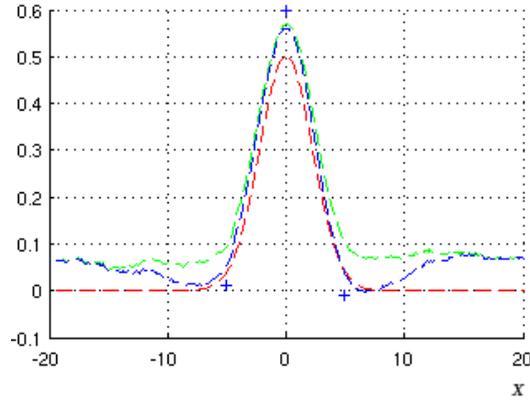


FIG. 7 – Expérience d’AD sur la solution initiale avec $\sigma_{erreur_B} = 10\%$, et $\sigma_{erreur_R} = 7\%$.

Bien sûr dans la réalité on ne dispose pas aussi précisément des erreurs que l’on a sur notre ébauche et nos observations.

4.2.2 Tout les N pas de temps

On va maintenant vouloir corriger le modèle tout les $N = 2,5$ pas de temps (i.e. toute les 2500 itérations). Pour cela on va procéder de la façon suivante : on va générer des observations tout les N pas de temps et dès que l’on se trouvera en présence de l’une d’entre elles, on appliquera la méthode d’assimilation de données décrite précédemment. La solution corrigée du modèle ainsi déterminée servira à la fois de nouvelle condition initiale pour l’itération suivante et de nouvelle ébauche pour l’assimilation suivante. On conserve pour cette expérience les mêmes paramètres que précédemment ainsi que le même code pour les légendes. Afin de mieux observer nos résultats, on va choisir de centrer la solution initiale à $1/3$ du maillage.

Résultats obtenus

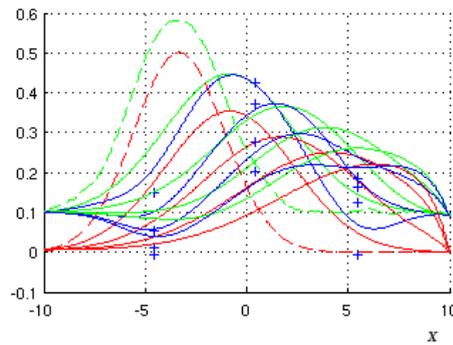


FIG. 8 – Expérience d’AD sur la solution initiale avec correction à chaque pas de temps avec $\sigma_{erreur_B} = 50\%$, et $\sigma_{erreur_R} = 50\%$.

Légende :

- rouge - - : solution initiale exacte
- vert - - : ébauche de la solution initiale
- bleu - - : analyse de la solution exacte
- + : observations
- rouge : modèle exacte
- vert : ébauche du modèle
- bleu : analyse du modèle.

Ces codes de couleurs seront conservés pour la suite des expériences.

On se trouve avec l'apparition d'un chevauchement en aval.

Pour permettre de passer outre cela, on va alors implanter une condition aux limites ouverte à droite :

$$\frac{\partial \tilde{h}}{\partial t} = -\alpha \frac{\partial \tilde{h}}{\partial x}. \quad (33)$$

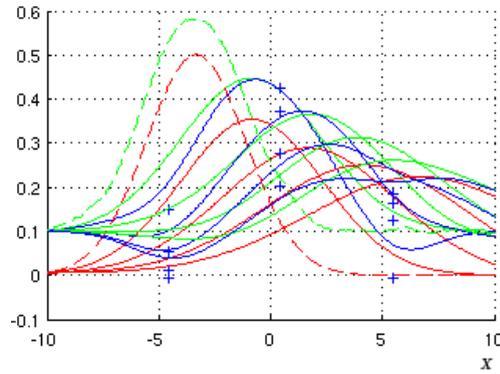


FIG. 9 – Expérience d'AD sur la solution initiale avec CL aval.

Pour mieux visualiser l'évolution au cours du temps, on va afficher la solution calculée pour chaque pas de temps :

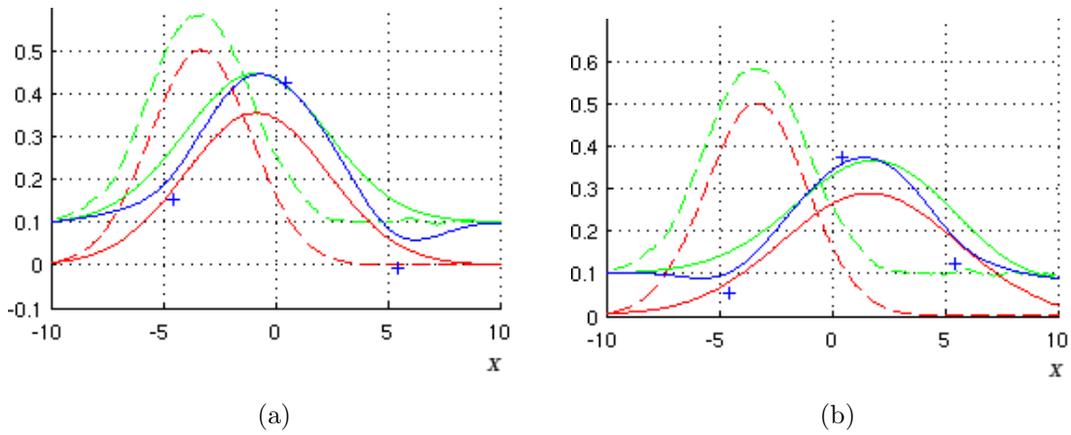


FIG. 10 – Évolution de l'expérience d'AD sur la solution initiale avec correction à chaque pas de temps, $\sigma_{erreur_B} = \sigma_{erreur_R} = 50\%$: a) N itérations, b) $2N$ itérations

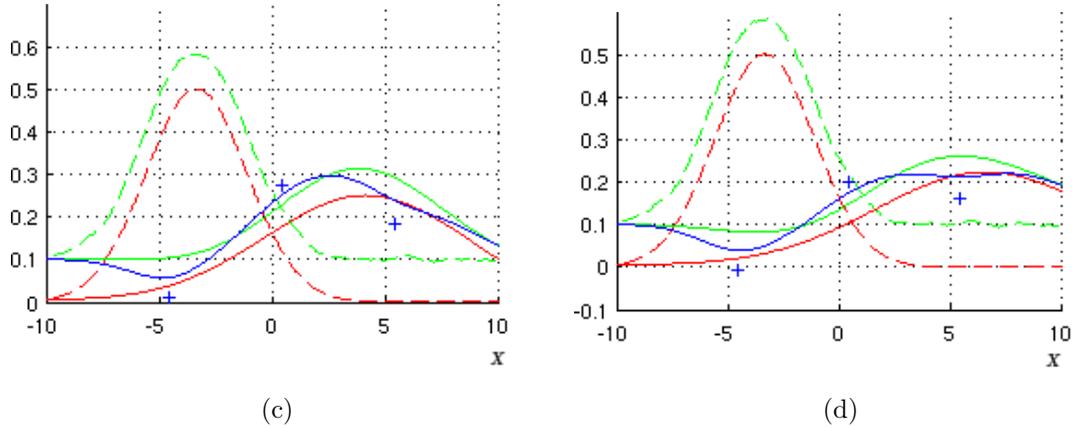


FIG. 11 – Évolution de l'expérience d'AD sur la solution initiale avec correction à chaque pas de temps, $\sigma_{erreur_B} = \sigma_{erreur_R} = 50\%$: c) 3 N itérations, d) 4 N itérations.

On constate à chaque pas de temps que l'analyse est meilleure que l'ébauche. On remarque que les résultats obtenus sont nettement meilleurs avec cette méthode que ceux obtenus pour l'expérience d'avant.

Si on accorde plus d'importance à la matrice \underline{R} (i.e en diminuant l'erreur sur les observations) on observe que l'analyse va faire plus confiance aux observations (Figure 12 (a)). Si au contraire on ne dispose pas d'une observation proche, l'analyse sera plus proche de l'ébauche. (Figure 12 (b)). On retrouve les résultats obtenus pour l'expérience à temps fixe.

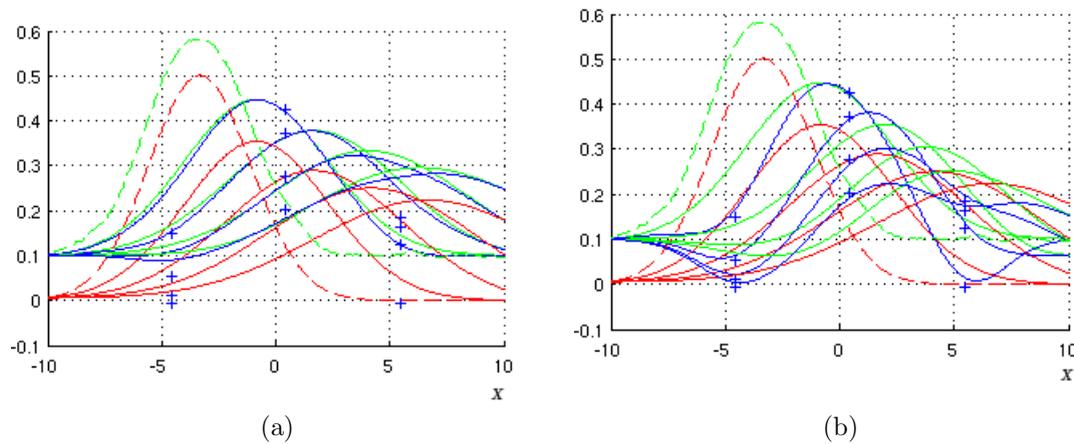


FIG. 12 – Expérience d'AD sur la solution initiale avec correction à chaque pas de temps a) avec $\sigma_{erreur_B} = 20\%$, et $\sigma_{erreur_R} = 50\%$, b) avec $\sigma_{erreur_B} = 50\%$, et $\sigma_{erreur_R} = 20\%$.

Jusqu'ici, on a généré les observations à partir de la solution exacte du modèle avec une erreur de 3% environ. Or on se situe sur une échelle très petite et cette erreur en devient importante. On va donc effectuer une nouvelle expérience en ajoutant à nos observations un bruit blanc gaussien de variance 0.01 On va directement visualiser ce qu'il se passe à chaque pas de temps.

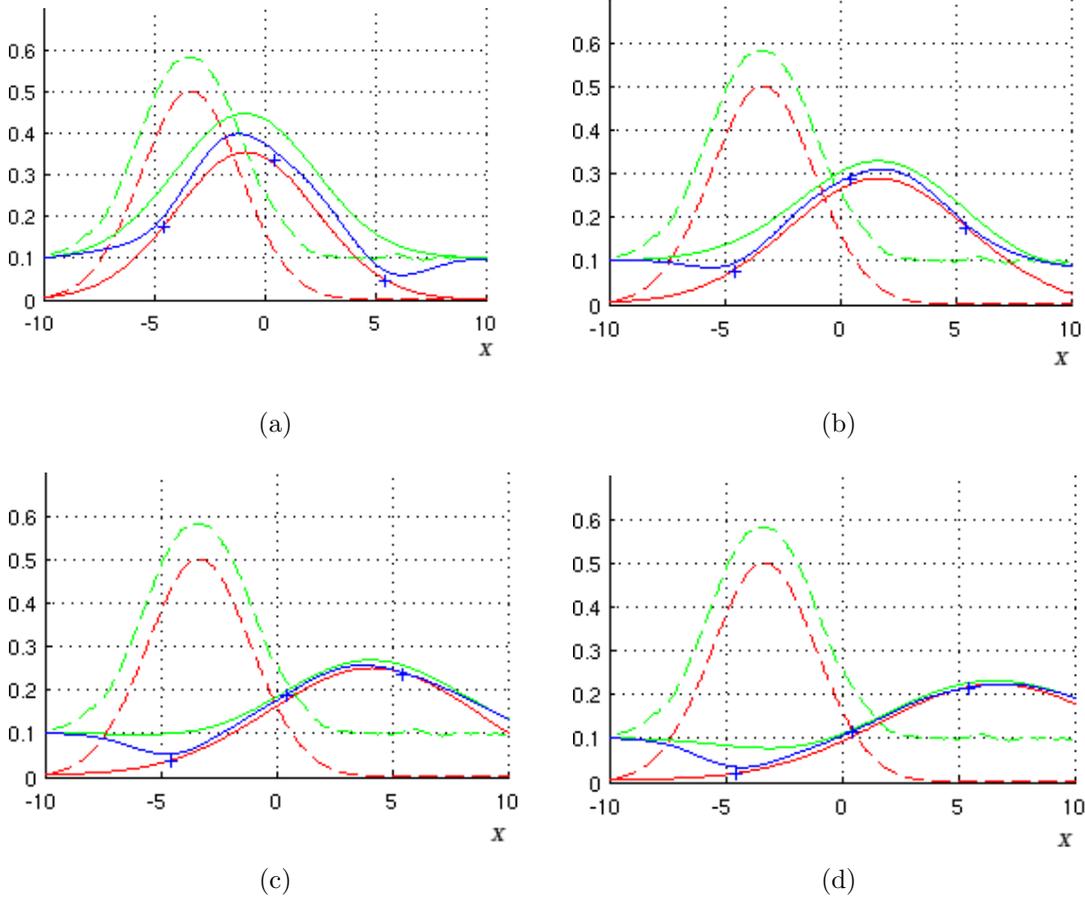


FIG. 13 – Évolution de l'expérience d'AD sur la solution initiale avec correction à chaque pas de temps, $\sigma_{erreur_B} = 7\%$ $\sigma_{erreur_R} = 6\%$: a) $t = N$, b) $t = 2 * N$, c) $t = 3 * N$ itérations, d) $t = 4 * N$ itérations.

On obtient ici une très bonne analyse.

4.2.3 Application : Assimilation de Données avec un débit amont aléatoire

On souhaite que notre modèle représente au mieux un cours d'eau, aussi on va maintenant implanter dans l'expérience précédente une condition limite à gauche (condition amont – CA). On va la définir sous la forme d'un fonction aléatoire de type gaussienne de variance u_0^2 fixée et de moyenne nulle dont la fonction d'autocorrélation sera comme précédemment :

$$R(T) = u_0^2 e^{-\frac{T^2}{2\sigma^2}}, \quad (34)$$

avec $u_0 = 5$ et $\sigma = 1$.

Le signal généré sera de la forme :

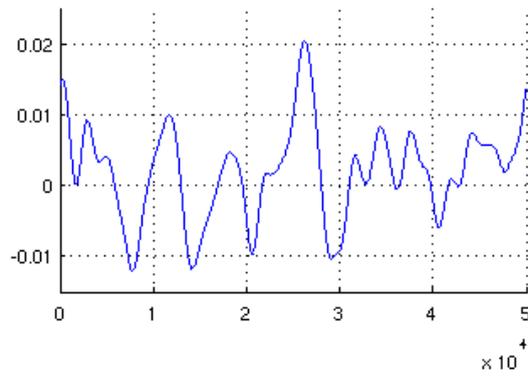


FIG. 14 – Représentation du signal implanté comme condition amont.

Résultats obtenus

On effectue cette expérience avec une condition initiale nulle afin de pouvoir observer la diffusion de la condition amont. De plus on augmente la taille de notre domaine ($L = 80$) et on s'arrête désormais au temps $t_f = 50$. On va pour l'instant observer juste ce qu'il se passe sans assimilation de données.

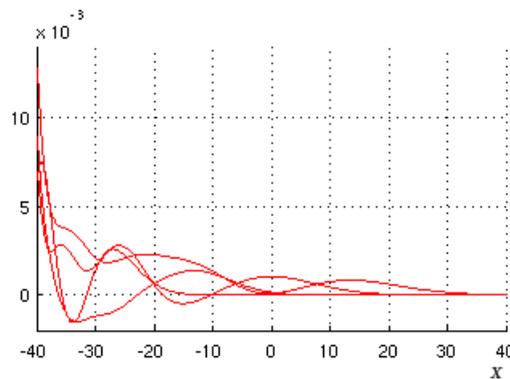


FIG. 15 – schéma centré explicite avec CI nulle et CL.

On observe que le signal imposé en amont va se diffuser vers la droite en s'atténuant plus on s'éloigne vers l'aval. On voit de plus sur la figure (2b) que la longueur de corrélation de la solution au temps final est plus petite en amont que en aval.

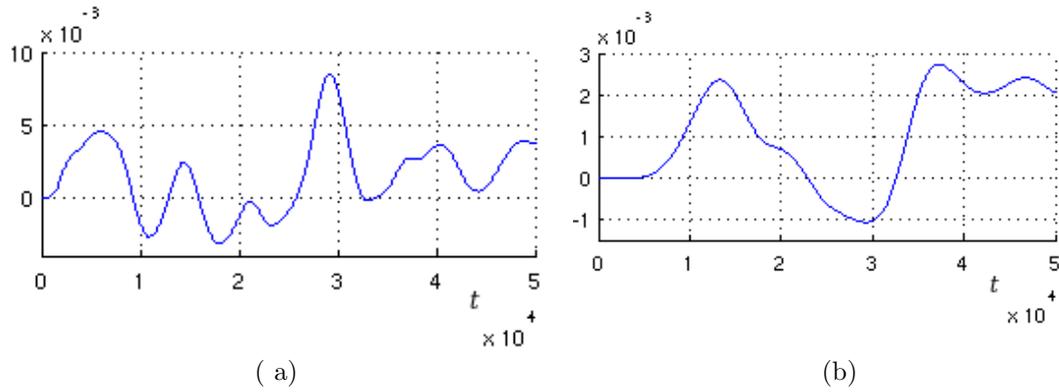


FIG. 16 – Solution $\tilde{h}(x, t)$ pour $\alpha = \beta = 1$. Signal temporel en a) $x \approx L/16$, b) $x \approx L/5$.

On va de nouveau appliquer une méthode d'assimilation de données tout les N pas de temps sur la condition initiale. On suivra la même démarche que précédemment. On va choisir ici de mettre les observations plus fiables que l'ébauche. On a à chaque pas de temps :

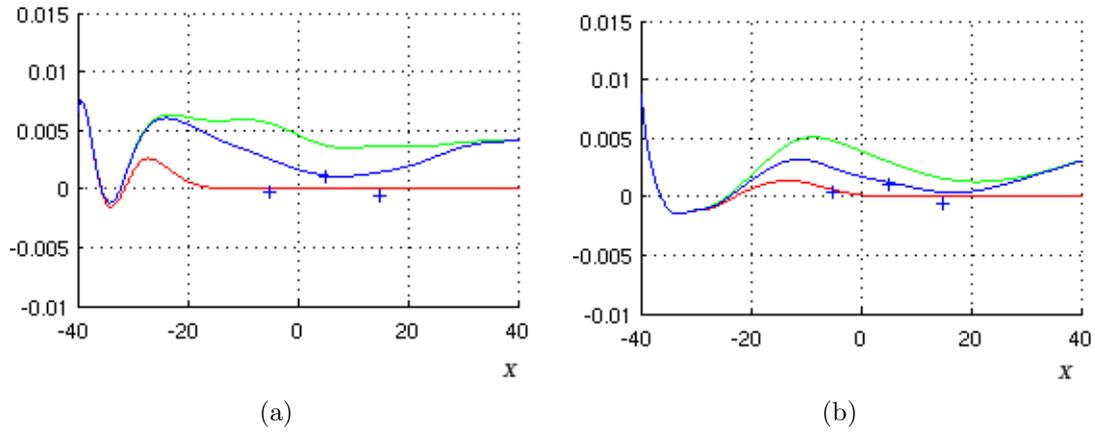


FIG. 17 – Évolution de l'expérience d'AD sur la solution initiale nulle avec correction à chaque pas de temps et CL, $\sigma_{erreur_B} = \sigma_{erreur_R} = 50\%$: a) N itérations, b) $2N$ itérations.

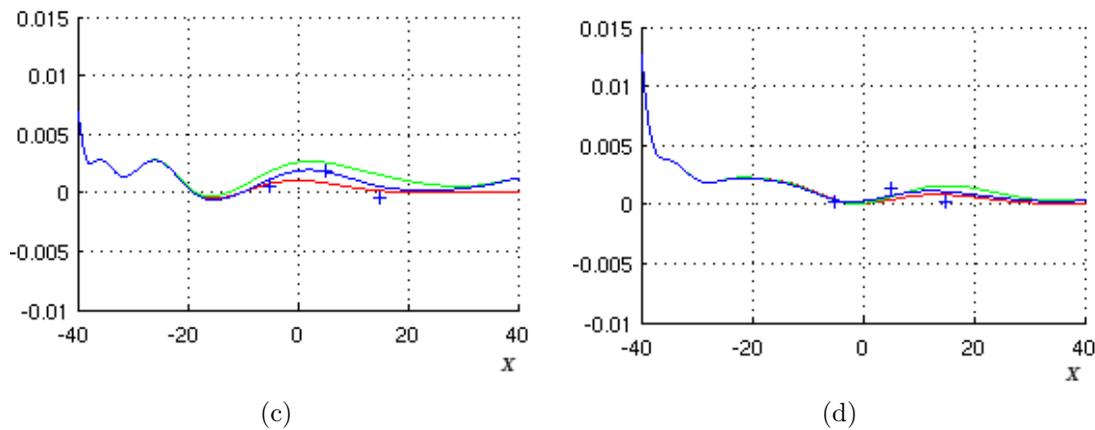


FIG. 18 – Évolution de l'expérience d'AD sur la solution initiale nulle avec correction à chaque pas de temps et CL, $\sigma_{erreur_B} = \sigma_{erreur_R} = 50\%$: c) $3N$ itérations, d) $4N$ itérations.

Les figures (17) et (18) montre que l'analyse permet de reconstituer de manière précise la solution exacte $\tilde{h}(x, t)$.

4.3 Bilan

Au travers ces expériences on a pu constater que l'Assimilation de Données apportait bien une correction non négligeable sur le modèle simple étudié. Nous avons ainsi mis en valeur la nécessité de renseigner avec précision les matrices $\underline{\underline{R}}$ et $\underline{\underline{B}}$. De plus on a montré avec l'expérience effectuée au §4.3.2, que la longueur de corrélation de la matrice B peut être estimée à partir du temps de corrélation des débits amonts à l'aide de la loi de diffusion qui fait croître cette longueur en allant vers l'aval. Cette étude souligne bien le fait que cette méthode mathématique demande des connaissances précises du phénomène physique modélisé ainsi que des observations disponibles, même pour un modèle simplifié comme le notre.

Nous allons donc exposer dans suite de ce rapport les différentes expériences qui furent effectuées pour remplir $\underline{\underline{R}}$ et $\underline{\underline{B}}$ pour la maquette Mascaret.

5 Étude de MASCARET

5.1 Présentation du logiciel

Le Code MASCARET est un logiciel libre de simulation numérique en mécanique des fluides, développé par le Laboratoire National d'Hydraulique et Environnement (LNHE) du service R&D d'EDF. MASCARET est un logiciel complexe, dont la réalisation a débuté il y a plus de vingt ans pour répondre aux besoins internes d'EDF et du CETMEF (Centre d'études techniques maritimes et fluviales). Le logiciel MASCARET est un modèle monodimensionnel à surface libre, basé sur les équations de Saint-Venant (les variables d'écoulements sont le débit et la section mouillée) dont le coeur du système est codé en Fortran 90. Il permet d'effectuer des calculs pour des écoulements aussi bien fluviaux que transcritiques, en régime permanent ou non permanent. Ainsi son champ d'application couvre de nombreux cas d'études comme les propagations de crues la modélisation des champs d'inondations, les ondes de submersion résultant de rupture d'ouvrages hydrauliques, les régulations de rivières ou encore les propagations d'ondes dans les canaux (intumescences, éclusées, mise en eau). Il permet de simuler les écoulements dans le lit mineur des rivières mais également dans le lit majeur de façon relativement réaliste. Ce modèle filiaire (en 1 dimension) suppose que tous les écoulements se font dans l'axe de la rivière.

Nous allons nous intéresser ici à la prévision des crues. Pour cela, on souhaite améliorer les résultats fournis par MASCARET, i.e. le débit Q et la hauteur Z calculés en sortie. Les cours d'eau sur lesquels on va effectuer une simulation étant des zones surveillées, on dispose généralement de nombreuses observations. Il paraît donc logique d'utiliser une méthode d'assimilation de données. On cherche à obtenir ici un débit et une hauteur les plus exactes possibles. Pour cela les ingénieurs du CERFACS ont implémenté dans le noyau du logiciel MASCARET un BLUE incrémental. Aussi l'espace de contrôle va être $\underline{x} = (Z, Q)$ où Z (que l'on note aussi h) et Q sont de la dimension du maillage illustrant notre rivière. Le modèle n'est corrigé qu'aux points où l'on dispose d'observations i.e. aux stations d'observations.

La démarche de l'assimilation mise en place par le CERFACS est la suivante : à chaque pas de temps, si on se trouve en présence d'observations, on détermine une correction, notée δZ , à apporter à Z (en calculant le BLUE). Puis, à partir de cette correction (et d'une relation entre ces deux variables), on calcule la correction δQ qu'il faut appliquer sur Q . Une fois ces corrections connues, on les appliquera sur les paramètres de contrôle. Les nouveaux Z et Q obtenus serviront alors d'ébauche pour le pas de temps suivant. Afin de rester le plus conforme à la réalité, il a été supposé que les erreurs relatives à l'ébauche et aux observations se propagent en amont et en aval du cours d'eau en suivant des lois normales gaussiennes centrées aux stations d'observations dont on spécifiera les variances après étude de la rivière. On va donc avoir besoin de connaître en détail les cas simulés afin de pouvoir remplir correctement les matrices B et R et de pouvoir établir une

relation entre les paramètres Z et Q . C'est ici que notre étude commence.

Pour modéliser un cours d'eau de façon réaliste, le logiciel MASCARET nécessite qu'on lui passe en entrée plusieurs paramètres :

- la géométrie du cours d'eau : nombre de bief, confluent, obstacles, ponts ... Ces informations forment un maillage et sont contenues dans les fichiers *.geo*.
- une ligne d'eau initiale. Cette ligne d'eau sera la même quelque soit le cours d'eau et quelque soit le débit imposé par l'utilisateur. Elle ne sert que pour lancer la simulation et est vite remplacée par une ligne d'eau adapté. On peut la visualiser via le fichier *.lig*.
- les conditions aux limites (amont et aval) situés dans les fichiers *.loi*.
- les observations en temps réel de notre cours d'eau, cf. fichiers *.csv*,
- les coefficients de frottements (coefficient de Strickler ...), les apports latéraux ... Ces derniers sont renseignés dans le fichier *.cas*.

Toutes ces données sont renseignées au logiciel via son interface Fudaa-Mascaret qui permet aussi la visualisation des résultats de calcul.

Nous allons nous concentrer ici sur l'étude de deux cours d'eau assez différents : Marne-Vallage et l'Adour Maritime. Dans un premier temps nous allons nous intéresser à la géométrie et aux caractéristiques de nos cas d'étude, puis dans un second temps on essaiera de déterminer expérimentalement une relation entre Z et Q aux stations de mesures considérées.

5.2 Présentation des cours d'eau servant de support

5.2.1 La Marne-Vallage

- **Géométrie**

Principal affluent de la Seine, la Marne, située à l'Est du bassin parisien, est la plus longue rivière française (525 km). Dans le cas de notre étude nous allons nous restreindre à une petite partie de ce cours d'eau située en amont sur le bassin de la Marne haut-marnaise. Le secteur d'étude débute en aval à Condes et s'achève plus bas, ≈ 84 km, à Chamouilley. (\approx annexe 2)

On se trouve donc en présence d'un fleuve en **régime non permanent** composé de trois biefs et d'une confluence (cf annexe) :

- **Bief 1** : Condes – amont confluence Rognon/Marne,
- **Bief 2** : Aval confluence Rognon/Marne – Chamouilley,
- **Bief 3** : Saucourt – amont confluence Rognon/Marne,
- **Confluence** : Marne/Rognon.

Afin de mieux pouvoir étudier ce cours d'eau, Mascaret va discrétiser la Marne en pas d'espace plus ou moins égaux. On définit ainsi un maillage dont l'abscisse de début sera le début du bief 1 et l'abscisse de fin sera la fin du bief 3.

- **Observations**

On dispose ici de deux stations de mesure : la première se situe à Joinville à 8820 m en aval du confluent, et la deuxième se situe à Chamouilley, à 34940 m en aval du confluent. Pour la Marne nous disposons de 6 jeux de données caractérisant des crues plus où moins importantes :

- crue de novembre 2002,
- crue de janvier 2004,
- crue de février 2006,
- crue de mars 2006,
- crue d’avril 2006,
- crue d’octobre 2006,
- crue de mars 2007,
- non crue de janvier 2007.

Les données (hauteurs et/ou débits) sont horodatées en heure TU et au pas de temps horaire.

- **Caractéristique**

Le modèle se compose de deux entrées (en amont) et d’une sortie (en aval). On aura donc autant de conditions aux limites :

- en amont : on sera en présence de condition de type débit imposée, hydrogramme $Q(t)$,
- en aval : on aura ici une courbe de tarage. (Figure 19).

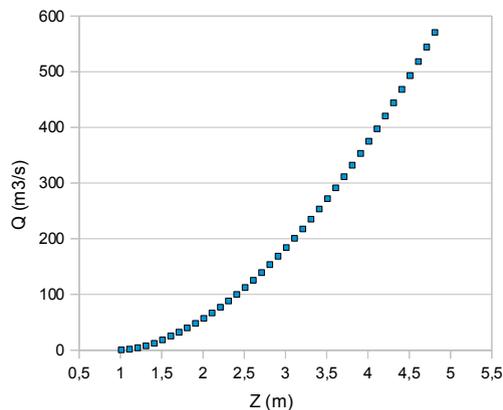


FIG. 19 – courbe de tarage imposée en sortie.

5.2.2 L’Adour-Maritime

- **Géométrie**

L’Adour est un fleuve du bassin aquitain dans le sud-ouest de la France. D’une longueur de ≈ 335 km, il prend sa source dans le massif pyrénéen du Pic du Midi de Bigorre, au col du Tourmalet, (Hautes-Pyrénées) et se jette dans l’Océan Atlantique à Bayonne. (cf annexe 3)

Derrière les “grands fleuves” français, l’Adour est le fleuve côtier le plus important par son débit. Ce fleuve de type montagnard, de régime pluvio-nival, est particulièrement redouté pour ses crues, avec des débits journaliers jusqu’à $2000 \text{ m}^3/\text{s}$ et des étiages jusqu’à $30 \text{ m}^3/\text{s}$.

Les frontières de zone d’étude choisies sont définies par rapport aux limites maximales de la marée. Le réseau hydraulique est composé de 10 biefs (7 réels et 3 “virtuels”) et trois confluences (cf annexe).

Comme précédemment, nous allons discrétiser ce fleuve en un maillage assez régulier dont le début sera l’amont du bief 1, et la fin sera l’aval du bief 10.

• Observations

Sur ce cours d’eau, on va considérer trois stations d’observations qui seront toutes exposées à différents degrés à la marée :

- Peyrehorade : située sur le bief 5, cette station est la moins sujette à l’influence de la marée ;
- Urt : localisée sur le bief 2 ;
- Lesseps : cette station est située sur le bief 1 à l’embouchure de Bayonne, elle subira le plus le phénomène de marée.

On dispose aussi de jeux de données illustrant plusieurs épisodes de crues importants :

- Crue de Novembre-Décembre 2000,
- Crue de Juin 2000,
- Crue de Novembre-Décembre 2002,
- Crue de Janvier-Février 2003,
- Crue de Janvier-Février 2004,
- Crue de Décembre 2004 - Février 2005,
- Crue de Février-Mars 2001,
- Crue de Mars 2006.

Ces données récoltées sont sous forme de débits et/ou de hauteur en TU+1 l’hiver, et en TU+2 l’été et seront renseignées au logiciel au pas de temps horaire.

• Caractéristique

Les conditions limites sont introduites sous forme d’hydrogramme $Q(t)$ en amont des biefs 4,8,9 et 10, et sous forme de limnigramme $H(t)$ à l’aval du bief 1. (Figure 20)

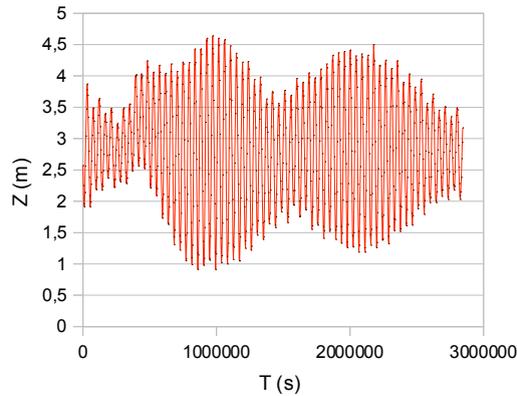


FIG. 20 – Représentation de la marée imposée en aval.

5.3 Détermination des courbes de tarages

5.3.1 Pour la Marne

Afin de déterminer une relation hauteur/débit aux stations Joinville et Chamouilley nous allons nous livrer à une série d'expériences. Nous allons faire tourner le logiciel Mascaret en imposant comme entrée à Condes et Saucourt un débit constant (ce débit est renseigné au travers des fichiers *.loi* que l'on trouve dans le répertoire *DonneeDyn*).

Le logiciel Mascaret commence sa simulation en utilisant une ligne d'eau fixe (qui est la même quelque soit le débit imposé en amont). Il faut donc attendre quelques heures pour que le débit simulé prenne en compte le débit imposé en amont.

La condition limite aval étant de type loi de tarage on aura ainsi une stabilisation du débit et de la hauteur simulés aux stations au bout d'un certain temps. On obtient de cette façon un couple (Z, Q) . Nous allons effectuer cette démarche 16 fois avec des débits d'entrées différents.

	Débit imposé à Saucourt (m ³ /s)	Débit imposé à Condes (m ³ /s)
Exp n°1	5	5
Exp n°2	5	7
Exp n°3	5	10
Exp n°4	5	12
Exp n°5	5	15
Exp n°6	7	7
Exp n°7	7	10
Exp n°8	10	10
Exp n°9	10	12
Exp n°10	10	15
Exp n°11	12	12
Exp n°12	12	15
Exp n°13	15	15
Exp n°14	15	25
Exp n°15	25	50
Exp n°16	50	75

FIG. 21 – Expérience effectuée sur la Marne.

Dans un repère, on trace les points provenant des couples obtenus, puis avec le logiciel Excel, nous allons représenter la courbe de tendance qui sera la plus exacte.

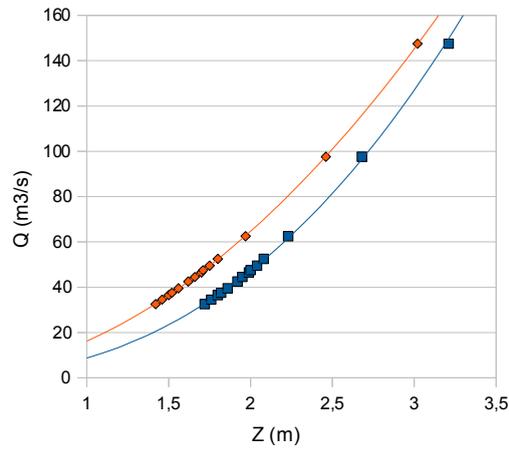


FIG. 22 – Courbe de tarage expérimentale pour la Marne : rouge pour Joinville, bleu pour Chamouilley.

On détermine pour chaque station une courbe de tarage de type puissance :

$$\begin{aligned} Q &= 8,72 Z^{2,44} && \text{en Joinville,} \\ Q &= 16,21 Z^2 && \text{en Chamouilley.} \end{aligned} \quad (35)$$

Cependant, nous a ici une condition aval de type loi de tarage, nous va donc vérifier que le débit et la hauteur simulés à la fin du bief 2 correspondent bien à la cette relation.

Pour cela on relève Z et Q calculés par le modèle en sortie du bief 2 pour la simulation de la crue de Novembre 2002, et on les compare à ceux imposés de la courbe de tarage.

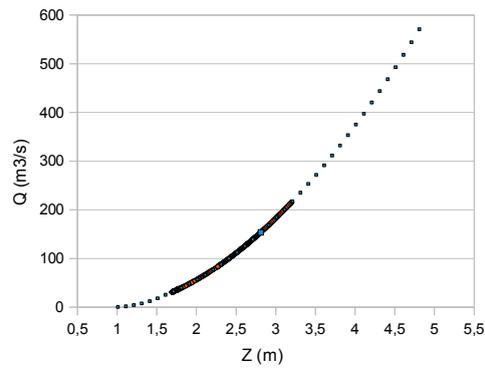


FIG. 23 – Comparaison courbe de tarage imposée (bleu) / courbe de tarage simulée (rouge).

On remarque que les points se superposent parfaitement ; aussi le débit et la hauteur simulés suivent bien la loi imposée.

5.3.2 Pour l'Adour

Nous procédons comme précédemment : nous lançons le logiciel Mascaret sur une simulation de l'Adour en imposant en amont des biefs 4, 8, 9 et 10 un débit constant.

On effectue ici 21 simulations :

	Débit imposé à Dax (m ³ /s)	Débit imposé à Cambo-les-Bains (m ³ /s)	Débit imposé à Escos (m ³ /s)	Débit imposé à Orthez (m ³ /s)
Exp n°1	7	10	12	22
Exp n°2	7	10	12	44
Exp n°3	7	10	24	22
Exp n°4	7	12	16	30
Exp n°5	7	14	16	35
Exp n°6	7	20	12	22
Exp n°7	10	20	22	44
Exp n°8	12	24	30	52
Exp n°9	14	10	12	22
Exp n°10	14	20	24	44
Exp n°11	14	30	30	75
Exp n°12	16	30	35	75
Exp n°13	20	44	52	75
Exp n°14	22	44	60	100
Exp n°15	24	52	75	125
Exp n°16	30	44	52	100
Exp n°17	30	60	100	125
Exp n°18	35	60	75	150
Exp n°19	35	75	90	150
Exp n°20	44	90	100	150
Exp n°21	52	100	125	150

FIG. 24 – Expérience effectuée sur l'Adour.

On obtient comme graphique déduit des couples (Z, Q) simulés des nuages de point sous forme d'ellipses.

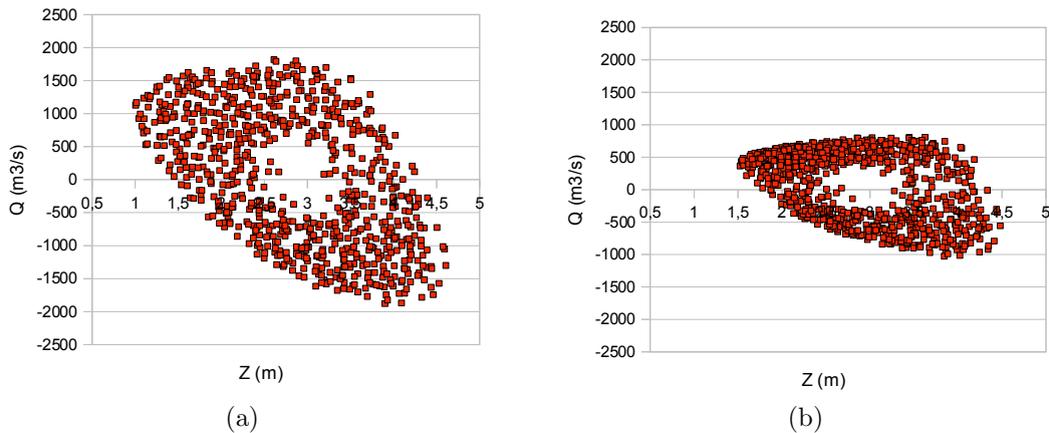


FIG. 25 – Relation hauteur débit pour l'expérience 1 : a) à Lesseps, b) à Urt.

Ces ellipses obtenues sont dues à l'influence de la marée. En effet l'estuaire de l'Adour est soumis à des marées du type semi-diurne, caractérisées par des cycles de périodes moyenne de 12 h25. Un autre cycle bi-hebdomadaire vient se superposer avec des périodes de vives-eaux et de mortes-eaux. Ces phénomènes ne permettent pas la stabilisation des débits pour nos stations d'études.

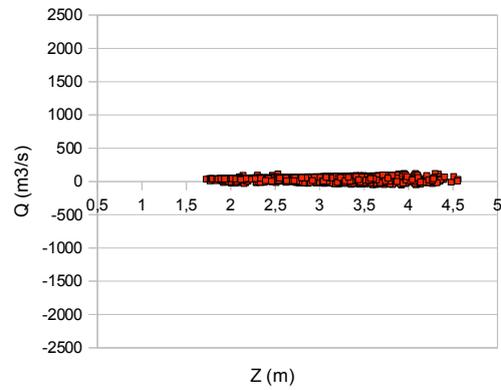


FIG. 26 – Relation hauteur débit à Peyrehorade pour l’expérience 1.

On remarque que le nuage de points correspondant à la station Peyrehorade est beaucoup plus aplati que celui obtenu pour la station Lesseps. Cela s’explique par le fait que Peyrehorade est la station la plus éloignée de la côte (donc la moins influencée par la marée). De plus, contrairement aux deux autres stations, le débit perçu à Peyrehorade dépend de seulement deux biefs (6 et 7), il est donc moins important que celui reçu par Urt et Lesseps. Par ailleurs il aussi faut prendre en compte le fait que la largeur de la rivière est plus faible à Peyrehorade.

En faisant varier les débits imposés en amont on retrouve à chaque fois une ellipse composée d’un nuage de point (Figure 27). Ces résultats ne nous permettent donc pas de pouvoir établir une relation entre le débit et la hauteur. On observe cependant que selon les débits imposés en amont l’ellipse obtenue ne sera pas la même.

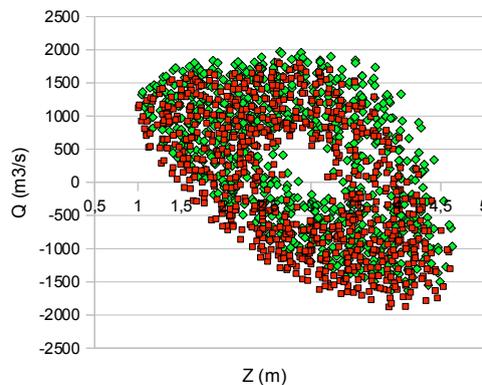


FIG. 27 – Relation Z/Q pour deux cas de débit amont constant : expérience (1) et (11).

On refait tourner nos 21 expériences et récupérer pour chaque station les nuages de points correspondants. Puis, pour chaque cas, on va déterminer le centre de l’ellipse associée. On aura donc 21 centres d’ellipse pour chaque station. On se place sur un même graphique tout les centres

obtenus. En traçant la courbe de tendance correspondante on remarque que l'on obtient pour chaque station une courbe.

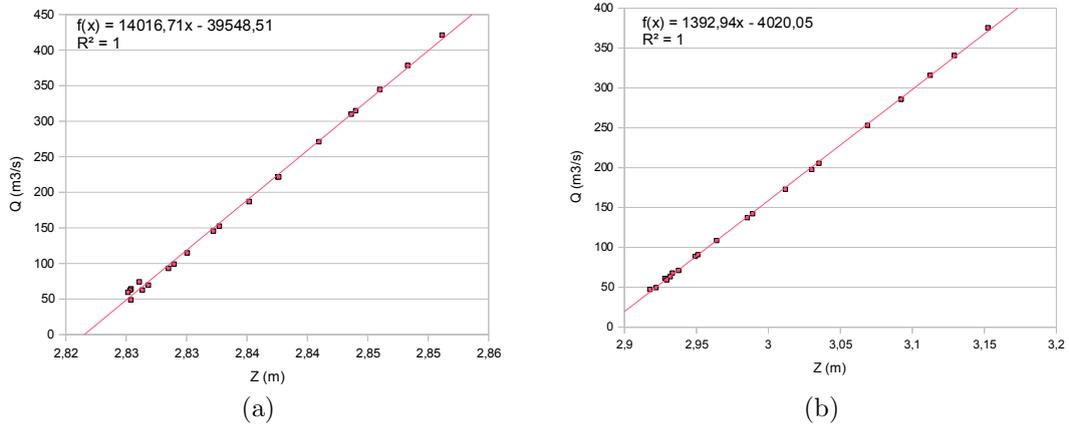


FIG. 28 – Courbe de tarage expérimentale : a) à Lesseps, b) à Urt.

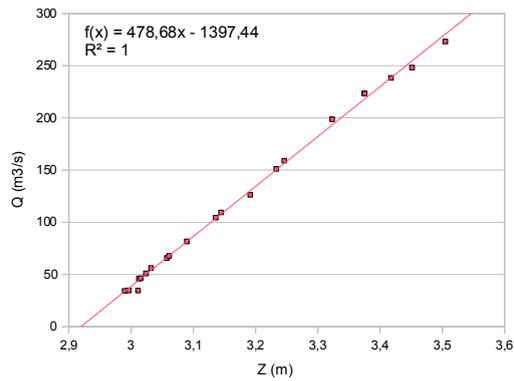


FIG. 29 – Courbe de tarage expérimentale à Peyrehorade.

Ces courbes représentent une approximation des courbes de tarages aux stations Lesseps, Urt et Peyrehorade. On a ainsi pour l'Adour les relations suivantes :

$$\begin{aligned}
 Q &= 141016.71S - 39548.51 && \text{à Lesseps,} \\
 Q &= 1392.94S - 4020.05 && \text{à Urt,} \\
 Q &= 478.68S - 1397.44 && \text{à Lesseps.}
 \end{aligned}
 \tag{36}$$

5.4 Étude de la diffusion

Une crue peut être considérée comme une onde se propageant dans une rivière. Comme toute onde elle sera donc caractérisée par une vitesse de propagation, une amplitude, une longueur d'onde, une fréquence et un taux d'amortissement. On va chercher à déterminer le coefficient de diffusion propre à chaque cours d'eau.

5.4.1 Approche théorique de la diffusion

L'écoulement à surface libre des eaux observé dans les rivières est régi par le système d'équations différentielles non linéaires de Barré de Saint Venant (1871).

Très souvent, le régime de l'écoulement est tel que l'on peut simplifier ces équations de l'hydraulique, en cherchant notamment à les linéariser.

Reprenons l'équation issue dans le cadre de l'approximation des ondes de crues des équations de Saint Venant présentée ci-dessus :

$$\frac{\partial \tilde{h}}{\partial t} + \frac{5 U_n}{3} \frac{\partial \tilde{h}}{\partial x} = k_n \frac{\partial^2 \tilde{h}}{\partial x^2}, \quad (37)$$

avec $k_n = \frac{U_n h_n}{2 \tan \gamma}$ le coefficient de diffusion. On considère ici que $\tan \gamma \approx I$ où I tel que $I = \sin \gamma$, γ étant la pente de la rivière. On a U_n d'après la relation de Strickler $U_n = K_s I^{1/2} h_n^{2/3}$ avec $h = h_n + \tilde{h}$ et $U = U_n + \tilde{U}$ où \tilde{h} et \tilde{U} des petites perturbations de l'équilibre (h_n, U_n) . On pose $\alpha = \frac{5 \tilde{U}_n}{3}$ et $\beta = k_n$.

On va considérer ici la solution initiale suivante :

$$\tilde{h}_0(x, t) = \int_0^x \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{s^2}{2\sigma_0^2}} ds + \frac{1}{2}, \quad (38)$$

où σ_0 est la variance. Lorsque $\sigma_0 \rightarrow 0$ cette fonction a la forme d'une fonction de type Heavyside. En faisant le changement de variable $T = \frac{s}{\sqrt{2}\sigma_0}$ on peut encore écrire cette solution sous la forme :

$$\tilde{h}_0(x, t) = \frac{1}{2} \operatorname{erf}(x \sqrt{2} \sigma_0) + \frac{1}{2}, \quad (39)$$

avec $\operatorname{erf}(t)$ la fonction d'erreur de Gauss définie comme suit :

$$\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-s^2} ds. \quad (40)$$

La solution générale de l'équation (37) va être de la forme :

$$\tilde{h}(x, t) = \frac{1}{\sqrt{2\pi\sigma^2(t)}} \int_0^x e^{-\frac{(s-\alpha t)^2}{2\sigma(t)^2}} ds + \frac{1}{2}, \quad (41)$$

avec $\sigma^2(t) = \sigma_0^2 + 2 \beta t$.

De nouveau en faisant le changement de variable $T = \frac{s}{\sqrt{2}\sigma_0}$ on a :

$$\tilde{h}(x, t) = \frac{1}{\sqrt{2\pi\sigma^2(t)}} \frac{\sqrt{\pi}}{2} \left(\frac{2}{\sqrt{\pi}} \int_0^{2\sigma\sqrt{2}} e^{-T^2} \frac{dT}{\sigma\sqrt{2}} \sigma\sqrt{2} \right) + \frac{1}{2}, \quad (42)$$

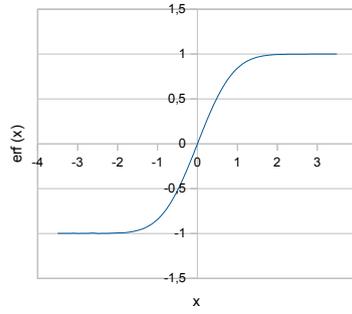


FIG. 30 – Graphique de la fonction erf.

en simplifiant on a

$$\begin{aligned} \tilde{h}(x,t) &= \frac{1}{2} \operatorname{erf} \left(\frac{x - \alpha t}{\sqrt{2}\sigma} \right) + \frac{1}{2} \\ &= \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{x - \alpha t}{\sigma\sqrt{2}} \right) \right]. \end{aligned} \quad (43)$$

On va considérer $\sigma_0 \approx 0$, d'où $\sigma^2(t) = 2 \beta t$.

*

Pour déterminer β (qui est en fait k_n) on va se livrer à une expérience. Prenons un cours d'eau de débit stable de valeur $n = 2 \text{ m}^3/\text{s}$ par exemple. À partir d'un temps t_d , injectons en amont un débit égal à $2 n \text{ m}^3/\text{s}$, puis attendons que tout se stabilise de nouveau. On notera le t_s le temps où le débit redevient stable. Observons ce qui se passe à un endroit x_s en aval :

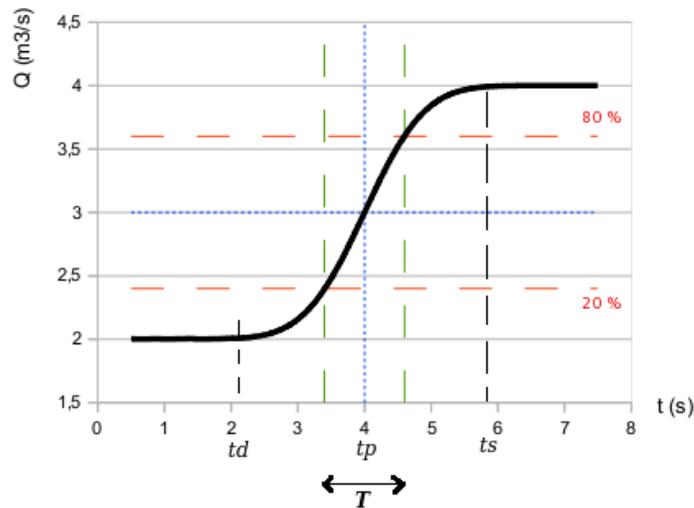


FIG. 31 – Illustration de l'augmentation du débit à x_s .

On va noter l_r la distance entre l'amont du cours d'eau et x_s , t_r le temps en secondes écoulé entre t_d et t_p le temps correspondant au milieu palier observé. On suppose $\sigma(t) \approx \sigma(t_p)$ pour $t \in [t_d, t_s]$. On aura $\alpha = \frac{l_r}{t_r}$ la vitesse de propagation de l'onde au point x_s .

On identifie cette courbe comme une fonction erf aussi on pose $T = \frac{\sigma}{5/3 U_n}$. T correspond graphiquement au temps qu'il faut pour passer de 20% de Q à 80% de Q . Or $\frac{5}{3} U_n \approx \alpha \Rightarrow \sigma = \alpha T$, et $\sigma^2 = 2\beta t_r$.

$$\text{D'où } (\alpha T)^2 = 2\beta t_r \Leftrightarrow \beta = \frac{\alpha^2 T^2}{2 t_r} = \frac{l_r^2 / t_r^2 T^2}{2 t_r} = \frac{l_r^2 T^2}{2 t_r^3}.$$

On suppose connu l_r , T , et t_r aussi on peut désormais calculer $\beta = k_n$. On en déduira σ au point x_s .

Pour notre méthode d'assimilation de données les coefficients σ correspondront aux portées des fonctions gaussiennes spatiales implantées dans les matrices \underline{B} et \underline{R} pour la station située au point d'abscisse x_s .

5.4.2 Expérience sur la Marne

Afin de déterminer k_n pour la Marne nous déterminons ici l_r , T , et t_r .

Pour cela, nous appliquons la même démarche expérimentale que pour le petit exemple précédent : nous allons faire tourner le logiciel MASCARET sur la Marne en lui imposant en amont en Condes et Saucourt un débit constant $Q = 10\text{m}^3/\text{s}$. Puis à un temps donné $t_d = 1000800$ secondes, soit 278 heures, nous allons ajouter au débit injecté en amont 50% de sa valeur i.e. $Q = 15\text{m}^3/\text{s}$.

Nous regardons ce qui se passe aux stations Joinville et Chamouilley.

On obtient ainsi :

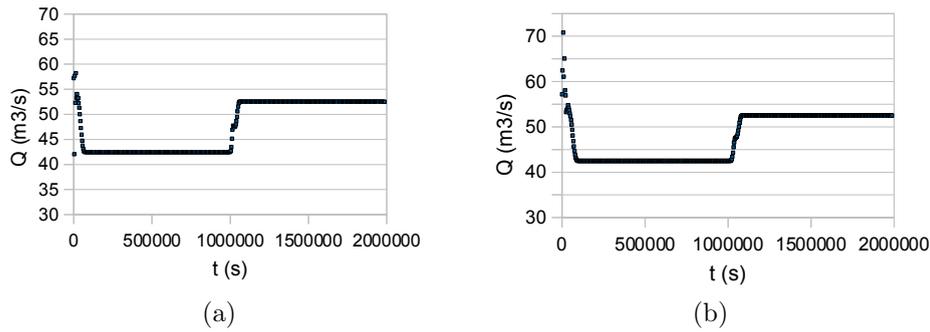


FIG. 32 – Évolution du débit / au temps a) à Joinville , b) à Chamouilley.

Les variations observées avant le temps $t = 79200\text{s}$ correspondent à l'adaptation de la ligne d'eau initiale de MASCARET pour le cas simulé, elles ne vont pas nous intéresser.

On remarque dans un premier temps que le débit simulé va passer de $42.5\text{m}^3/\text{s}$ à $52.5\text{m}^3/\text{s}$ (après stabilisation).

En observant les correspondance dans le tableau des résultats **.opt**, on remarque aussi que la hau-

teur va varier :

- à Joinville elle va passer de 1.92m à 2.08m, soit une augmentation de 0.16m pour une augmentation de débit de $10\text{m}^3/\text{s}$;
- à Chamouilley, Z va passer de 1.62m à 1.8m, soit une augmentation de 0.18m pour une augmentation de débit de $10\text{m}^3/\text{s}$.

On observe sur les courbes la présence de deux paliers. Pour mieux les visualiser, on va effectuer un zoom sur ces parties des graphiques.

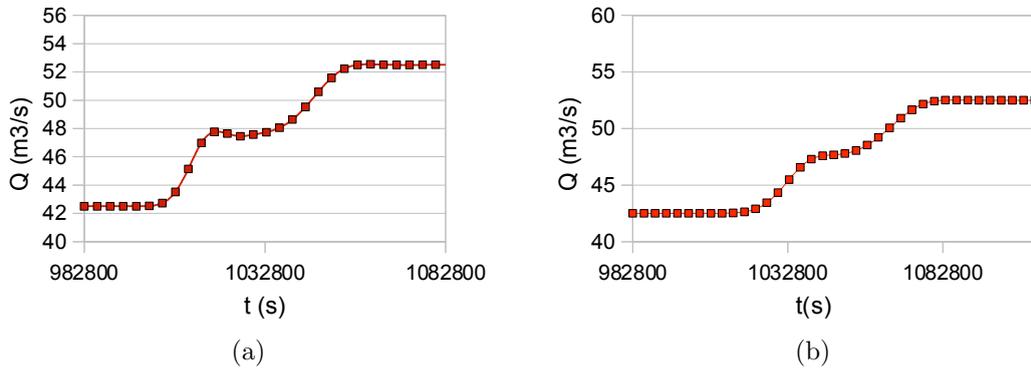


FIG. 33 – Zoom sur le débit simulé a) à Joinville, b) à Chamouilley.

La présence de ces paliers s'expliquent par la différence de longueur de bief que parcourt l'eau. En effet, le bief 3 étant nettement plus court que le bief 1, il est logique que les stations reçoivent en premier le débit issu du bief 3. Le deuxième palier correspond donc approximativement à l'arrivée du débit du bief 1.

On constate que l'apparition de l'augmentation du débit se fait d'abord en Joinville puis en Chamouilley, ce phénomène s'explique par la différence de distance entre les deux stations.

On va noter :

- $l_{r1.1} = 53830\text{m}$ la distance entre le début du bief 1 et la première station Joinville ;
- $l_{r1.2} = 79950\text{m}$ la distance entre le début du bief 1 et la deuxième station Chamouilley ;
- $l_{r3.1} = 13865\text{m}$ la distance entre le début du bief 3 et la première station Joinville ;
- $l_{r3.2} = 40985\text{m}$ la distance entre le début du bief 3 et la deuxième station Chamouilley ;
- t_{r1} le temps écoulé entre t_d et le milieu du premier palier ;
- t_{r2} le temps écoulé entre t_d et le milieu du second palier ;

En regardant de plus prêt les résultats fournis par Mascaret on observe :

	ts1 (s)	ts1(h)	ts2 (s)	ts2 (h)
Joinville	1022400	284	1069200	297
Chamouilley	1047600	291	1094400	304

FIG. 34 – Tableau récapitulatif des temps de la Marne.

où :

- $ts1$ est le temps où le débit se restabilise : c'est le premier palier ;
- $ts2$ est le temps où le débit se restabilise : c'est le second palier.

On peut alors calculer nos paramètres l_r , T , et t_r et en déduire k_s . On a donc :

		tr (s)	T (s)	lr	beta m2/s	alpha m/s
/ au bief 1	Joinville	10800	7200	53830	59622.82	4.98
	Chamouilley	34400	10800	79950	9157.55	2.32
/ au bief 3	Joinville	25200	37800	13865	8582.06	0.55
	Chamouilley	43200	37800	40985	14885.12	0.95

FIG. 35 – Résultats du calcul de la diffusion pour la Marne.

5.4.3 Expérience sur l'Adour

Pour l'Adour il va nous falloir composer avec la marée.

En effet l'influence de ce phénomène va nous empêcher d'avoir la stabilisation du débit malgré un débit amont imposé constant et donc de visualiser de façon claire nos résultats.

On cherche de nouveau à déterminer l_r , T et t_r . Ici nous sommes en présence de 4 biefs "d'entrées" aussi nous allons changer quelque peu la démarche expérimentale de l'exemple : nous allons faire tourner le logiciel Mascaret avec des débit amont imposés constant, et à un temps $t_d = 1000800s$, nous allons augmenter seulement **un seul** des 4 débits d'entrés, puis nous allons comparer les résultats que l'on aura au résultats que l'on aurait eu si on avait fait aucune modification.

On effectuera au total 4 expériences :

	Bief 4	Bief 8	Bief 9	Bief 10
exp n° 1	10 puis 20 m3/s	7 m3/s	12 m3/s	22 m3/s
exp n° 2	10 m3/s	7 puis 14 m3/s	12 m3/s	22 m3/s
exp n° 4	10 m3/s	7 m3/s	12 puis 24 m3/s	22 m3/s
exp n° 5	10 m3/s	7 m3/s	12 m3/s	22 puis 44 m3/s

FIG. 36 – Différence entre le débit simulé du jeu **A** et celui du jeu **AB** à Lesseps.

Pour mieux comprendre le principe, on va décrire ce qu'il se passe à la station Lesseps lors de l'expérience 1.

Imposons en entré un débit constant en temps tel que :

- à Cambo-les-Bains, bief 4 : $Q_{imp} = 10m^3/s$,
- à Dax, bief 8 : $Q_{imp} = 7m^3/s$,

- à Escos, bief 10 : $Q_{imp} = 12\text{m}^3/s$,
- à Orthez, bief 9 : $Q_{imp} = 22\text{m}^3/s$.

Faisons tourner le logiciel avec ces débits, on obtient un jeu de données résultat que l'on va noter **A**. Maintenant nous allons faire tourner de nouveau Mascaret, simplement à t_d , nous allons injecter à Cambo-les-Bains (bief 4) $Q_{imp} = 20\text{m}^3/s$. On récupère les résultats obtenus que l'on va nommer **AB**.

Pour pouvoir effectuer une comparaison, on va effectuer une troisième simulation, mais cette fois ci avec les débits suivants :

- à Cambo-les-Bains, bief 4 : $Q_{imp} = 20\text{m}^3/s$,
- à Dax, bief 8 : $Q_{imp} = 7\text{m}^3/s$,
- à Escos, bief 10 : $Q_{imp} = 12\text{m}^3/s$,
- à Orthez, bief 9 : $Q_{imp} = 22\text{m}^3/s$.

On notera **B** les résultats issus de cette simulation.

Ainsi pour voir apparaître les modifications apportées par l'augmentation brusque du débit en t_d , on va comparer le jeu de données **AB** aux jeux de données **A** et **B**.

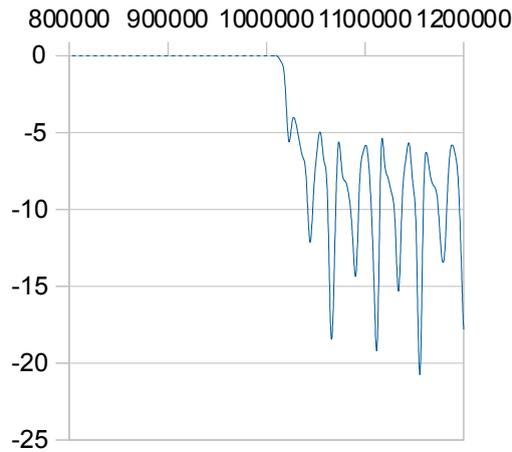


FIG. 37 – Différence entre le débit simulé du jeu **A** et celui du jeu **AB** à Lesseps.

Tant que le jeu **AB** sera confondu avec le jeu **A** aucun changement ne sera donc apparu. On pourra ainsi déterminer en superposant les deux jeux, les temps “de changement” noté t_c . De même, pour observer à partir de quand le débit se “stabilisera” de nouveau, on aura qu’à comparer le jeu **AB** au jeu **B**.

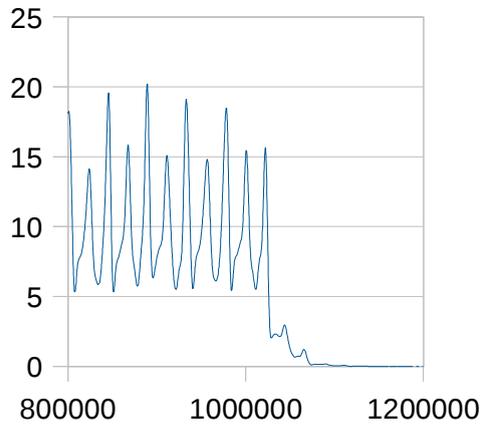


FIG. 38 – Différence entre le débit simulé du jeu **B** et celui du jeu **AB** à Lesseps.

On va procéder de même pour les stations Urt et Lesseps. On détermine ainsi :

		d (m)	ts (s)	ts (h)
exp n°1	Lesseps	66088	1141200	317
	Urt	50720	1141200	317
	Peyrehorade	48450	1119600	311
exp n°2	Lesseps	23317	1072800	298
	Urt	37649	1072800	298
	Peyrehorade	56059	1054800	293
exp n°3	Lesseps	73167	1159200	322
	Urt	57798	1159200	322
	Peyrehorade	39389	1119600	311
exp n°4	Lesseps	60446	1137600	316
	Urt	45078	1141200	317
	Peyrehorade	26668	1116000	310

FIG. 39 – Tableau des temps relevés pour l'Adour.

où on note **d** la distance entre le début du bief étudié et la station traitée. Ici $d = l_r$. On va identifier les courbes obtenues à des fonctions *erf* aussi on va pouvoir se servir des graphiques pour obtenir l_r , T , et t_r .

*

On va effectuer les 4 expériences et en sortir les paramètres qui nous intéressent :

		tr (s)	T (s)	lr (m)	Beta m2/s	alpha m/s
exp n°1	Lesseps	70200	30600	66088	5910.81	0.94
	Urt	70200	16200	50720	975.77	0.72
	Peyrhorade			48450		
exp n°2	Lesseps	36000	23400	23317	3190.36	0.65
	Urt			37649		
	Peyrhorade			56059		
exp n°3	Lesseps	79200	25200	73167	3421.57	0.92
	Urt	79200	21600	57798	1568.66	0.73
	Peyrhorade	59400	7200	39389	191.88	0.66
exp n°4	Lesseps	68400	9000	60446	462.4	0.88
	Urt	70200	9000	45078	237.89	0.64
	Peyrhorade	57600	5400	26668	54.26	0.46

FIG. 40 – Résultats du calcul de la diffusion pour l'Adour.

On pourra ainsi déterminer les portées adéquates des fonctions gaussiennes de mes matrices \underline{B} et \underline{B} pour l'Adour

On remarque que pour les expériences 1 et 2 on ne peut déterminer les paramètres qui nous intéressent. En effet les courbes issues des différences entre les jeux de données A – AB ou AB – B ne sont pas assimilables à des fonctions erf.

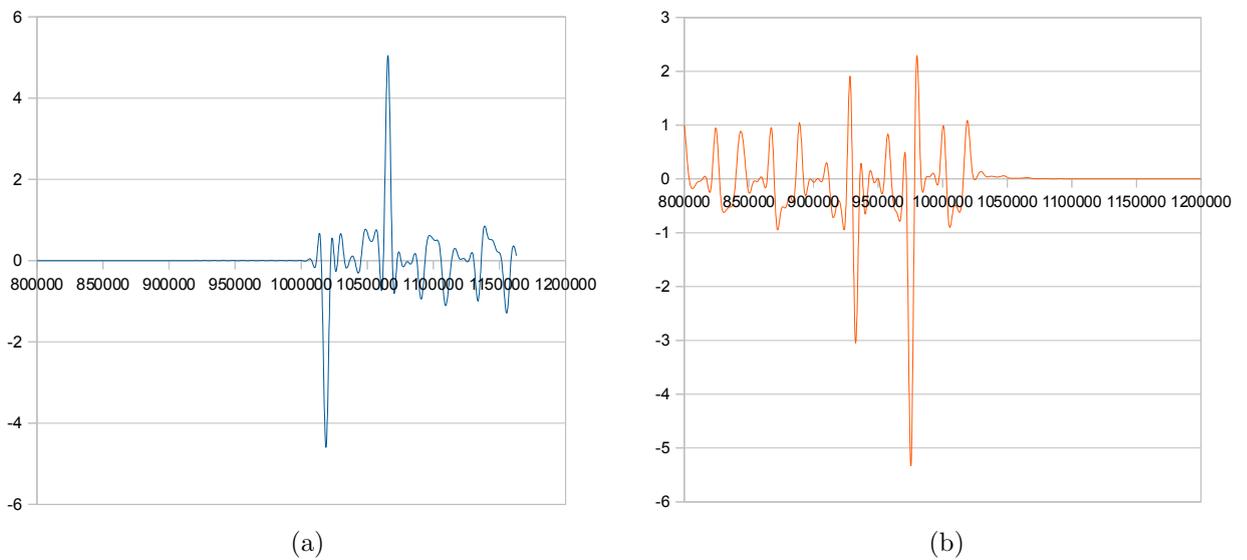


FIG. 41 – Différence entre les débits simulés à Urt pour l'exp 2 a) entre A - AB, b) entre B - AB.

Cela s'explique par le fait que pour cette expérience la station Urt se situe en amont de la confluence qui recevra le nouveau débit. Elle ne subira pas les effets directs de cette augmentation. En revanche on remarque que le débit simulé à ces stations subit quand même des variations. On assiste à un phénomène de déphasage de la marée.

6 Conclusion et perspectives

La démarche expérimentale exposée en deuxième partie de ce rapport a permis de mieux comprendre le fonctionnement d'une méthode d'Assimilation de Données, ce de son implémentation à son résultat. Bien que le modèle étudié soit simple, les résultats obtenus sont déjà très satisfaisants ce qui nous montre bien que l'assimilation de données est une méthode mathématique adaptée à l'optimisation des modèles géophysiques actuels. Cependant ces expériences nous montrent que la mise en place d'une méthode d'assimilation de données ne se fait pas sans une étude approfondie des paramètres utilisés pour sa construction. L'importance capitale de la connaissance du modèle physique souligne ici la nécessité croissante d'une collaboration entre deux disciplines scientifiques telles que les mathématiques appliquées et l'hydrologie.

On retrouve cette conclusion au travers l'établissement de la maquette Mascaret. En effet on a constaté au cours du paragraphe §5 la difficulté de déterminer avec précision les erreurs effectuées sur les observations ainsi que sur l'ébauche. Seule une connaissance poussée des cas d'études et du logiciel nous ont permis de mener à bien cette maquette. Cette dernière, bien que encore incomplète, nous permet tout de même de mettre en évidence les corrections que l'assimilation de données peut apporter sur un modèle plus complexe.

Les résultats obtenus jusqu'ici nous montrent que cette méthode permet d'améliorer significativement la prévision des pics de crue pour de nombreux épisodes pluvieux. Néanmoins, afin de les valider mathématiquement, il serait avantageux pour la suite d'implanter différents critères (biais, précision, efficacité ...) qui nous permettrait d'effectuer une comparaison avec le calage. De plus, il serait maintenant intéressant de caller les portées des matrices \underline{R} et \underline{B} en se servant de l'étude réalisée dans ce rapport, ainsi que le calcul des incréments.

D'autres perspectives sont aussi envisageables sur cette maquette. Il serait en effet intéressant par la suite d'affiner le travail effectué sur les matrices \underline{R} et \underline{B} en leur implantant une portée assymétrique, plus adaptée à nos cours d'eau étudiés. On pourrait aussi étendre cette approche d'AD à l'ajustement de nouveaux paramètres. Il serait avantageux pour la suite d'appliquer la démarche d'assimilation de données effectuée aux débits imposés en amont des cours d'eau. Avec l'implémentation de la maquette sous le coupleur PALM et l'utilisation de calculateurs de plus en plus puissant, on peut imaginer généraliser cette maquette sur tout les fleuves et rivières de France, ce qui permettrait ainsi une meilleure prévision des crues.

Outre les notions présentés dans ce rapport, ce stage a été très enrichissant tant sur la plan technique que scientifique. Au travers divers échanges riche et fructueux j'ai pu ainsi apprendre beaucoup de spécialistes et mieux saisir le travail attendu d'un ingénieur. De plus il m'a été aussi demandé de contribuer à "former" des novices, me faisant prendre conscience de l'importance des relations humaines, seules capables des faire avancer les choses.

Références

- **Bouvier C., Marchandise A., Brunet P., Crespy A.,** (). In modèle pluie-débit distribué événementiel parcimonieux pour la prédétermination et la prévision des crues éclair en zone méditerranéenne. Application au bassin du Gardon d'Anduze. 10 p.

- **Daget N.,** (2008). Estimation d'ensemble des paramètres des covariances d'erreur d'ébauche dans un système d'assimilation variationnelle de données océaniques. *Thèse de Doctorat*. Université Paul Sabatier - Toulouse III. 352 p.

- **Desroziers G.,** (2009). Assimilation de données. Cours de Master 2. Université Paul Sabatier - Toulouse III. 121 p.

- **Gouin P. et Le Pape E.,** (2008). MASCARET en temps réel. Rapport interne SCHAPI. 32 p.

- **Habert J. et Halgrain J.,** (2007). Rapport de callage, de modélisation et de test temps réel du modèle MASCARET "Marne Vallage". SPC SAMA. 46 p.

- **Le Pape E.,** (2008). Démarche pour la mise en oeuvre d'un modèle MASCARET en temps réel. Rapport interne SCHAPI. 12 p.

- **Marchandise A.,** (2007). Modélisation hydrologique distribuée sur le Gardon d'Anduze ; Étude comparative de différents modèles pluie-débit, extrapolation de la normale à l'extrême et test d'hypothèses sur les processus hydrologiques. *Thèse de Doctorat*. Université Montpellier II. 212 p.

- **Ricci S.,** (2004). Assimilation variationnelle océanique : modélisation multivariée de la matrice de covariance d'erreur d'ébauche. *Thèse de Doctorat*. Université Paul Sabatier - Toulouse III. 275 p.

- **Sabater J. M., Massart S., Ricci S. et Bouriquet B.,** (2008). Introduction à l'assimilation de données ; applications actuelles dans les projets du CERFACS. *Technical Report CMGC*. 41 p.

- **Sabater J. M., Bouriquet B., Ricci S., Argaud J.-Ph., Piacentini A., Massart S.,**

- Weaver A. et Thual O.**, (2008). Application du CERFACS en assimilation de données. *7^{èmes} Journées Scientifiques et Techniques du CETMEF*. 10 p.
- **Schoorens J. et Le Pape E.**, (2007). Rapport de calage du modèle MASCARET ‘Adour Maritime’ pour la prévision des crues. SCHAPI. 44 p.
 - **Thual O.**, (2002). L’assimilation de données pour la prévision : une application du concept de contrôle optimal. *EPI-N7, Échange de Projets et Idées à l’ENSEEIH, Séminaire CPGE*. 29 p.
 - **Thual O.**, (2009). Question sur la convergence d’une méthode d’assimilation de données itérative. *epI-ASSIM Échange de Projets et Idées “Assimilation de Données”*. 6 p.
 - **Thual O.**, (2009). Sciences de l’eau et de l’environnement. *Cours de l’École polytechnique*. 224 p.

Annexes

Annexe 1 : Démonstration de la stabilité du schéma explicite centré

Soit le schéma explicite centré suivant :

$$\frac{\tilde{h}_j^{n+1} - \tilde{h}_j^n}{\delta t} + \alpha \frac{\tilde{h}_{j+1}^n + \tilde{h}_{j-1}^n}{2 \delta x} - \beta \frac{\tilde{h}_{j+1}^n - 2 \tilde{h}_j^n + \tilde{h}_{j-1}^n}{\delta x^2} = 0.$$

On peut le réécrire sous la forme :

$$\tilde{h}_j^{n+1} = \tilde{h}_j^n - \alpha \frac{\delta t}{2 \delta x} (\tilde{h}_{j+1}^n + \tilde{h}_{j-1}^n) + \beta \frac{\delta t}{\delta x^2} (\tilde{h}_{j+1}^n - 2 \tilde{h}_j^n + \tilde{h}_{j-1}^n).$$

En regroupant les termes on a :

$$\tilde{h}_j^{n+1} = \left(1 - 2 \beta \frac{\delta t}{\delta x^2}\right) \tilde{h}_j^n + \left(\beta \frac{\delta t}{\delta x^2} - \alpha \frac{\delta t}{2 \delta x}\right) \tilde{h}_{j+1}^n + \left(\alpha \frac{\delta t}{2 \delta x} + \beta \frac{\delta t}{\delta x^2}\right) \tilde{h}_{j-1}^n.$$

Posons :

$$a = 1 - 2 \beta \frac{\delta t}{\delta x^2},$$

$$b = \beta \frac{\delta t}{\delta x^2} - \alpha \frac{\delta t}{2 \delta x},$$

$$\text{et } c = \alpha \frac{\delta t}{2 \delta x} + \beta \frac{\delta t}{\delta x^2}.$$

On va supposer que a et b sont ≥ 0 :

$$a \geq 0 \Leftrightarrow 1 - 2 \beta \frac{\delta t}{\delta x^2} \geq 0$$

$$\Leftrightarrow 2 \beta \frac{\delta t}{\delta x^2} \leq 1$$

$$\Leftrightarrow \beta \frac{\delta t}{\delta x^2} \leq \frac{1}{2}, \quad (1)$$

et

$$b \geq 0 \Leftrightarrow \alpha \frac{\delta t}{2 \delta x} \leq \beta \frac{\delta t}{\delta x^2}$$

$$\Leftrightarrow \alpha \frac{\delta t}{2 \delta x} \leq \frac{1}{2}$$

$$\Leftrightarrow \alpha \frac{\delta t}{\delta x} \leq 1. \quad (2)$$

Sous les conditions précédentes on va étudier la stabilité :

$$\begin{aligned} |\tilde{h}_j^{n+1}| &\leq a |\tilde{h}_j^n| + b |\tilde{h}_{j+1}^n| + c |\tilde{h}_{j-1}^n|, \\ &\leq a \sup_{j \in Z} |\tilde{h}_j^n| + b \sup_{j \in Z} |\tilde{h}_{j+1}^n| + c \sup_{j \in Z} |\tilde{h}_{j-1}^n|. \end{aligned}$$

$$|\tilde{h}_j^{n+1}| \leq a \sup_{j \in Z} |\tilde{h}_j^n| + b \sup_{j \in Z} |\tilde{h}_j^n| + c \sup_{j \in Z} |\tilde{h}_j^n|,$$

$$\Leftrightarrow \sup |\tilde{h}_j^{n+1}| \leq (a + b + c) \sup_{j \in Z} |\tilde{h}_j^n|.$$

Or $a + b + c = 1$,

d'où

$$\sup |\tilde{h}_j^{n+1}| \leq \sup_{j \in Z} |\tilde{h}_j^n| \Leftrightarrow \sup |\tilde{h}_j^n| \leq \sup_{j \in Z} |\tilde{h}_j^0|,$$

$$\Leftrightarrow \|\tilde{h}^n\|_{\infty, h} \leq \|\tilde{h}^0\|_{\infty, h}, \quad \forall n \in N^*.$$

Le schéma explicite centré est donc L^∞ stable sous les conditions (1) et (2).

cqfd

Annexe 2 : Situation géographique de la Marne.



FIG. 42 – Situation globale.

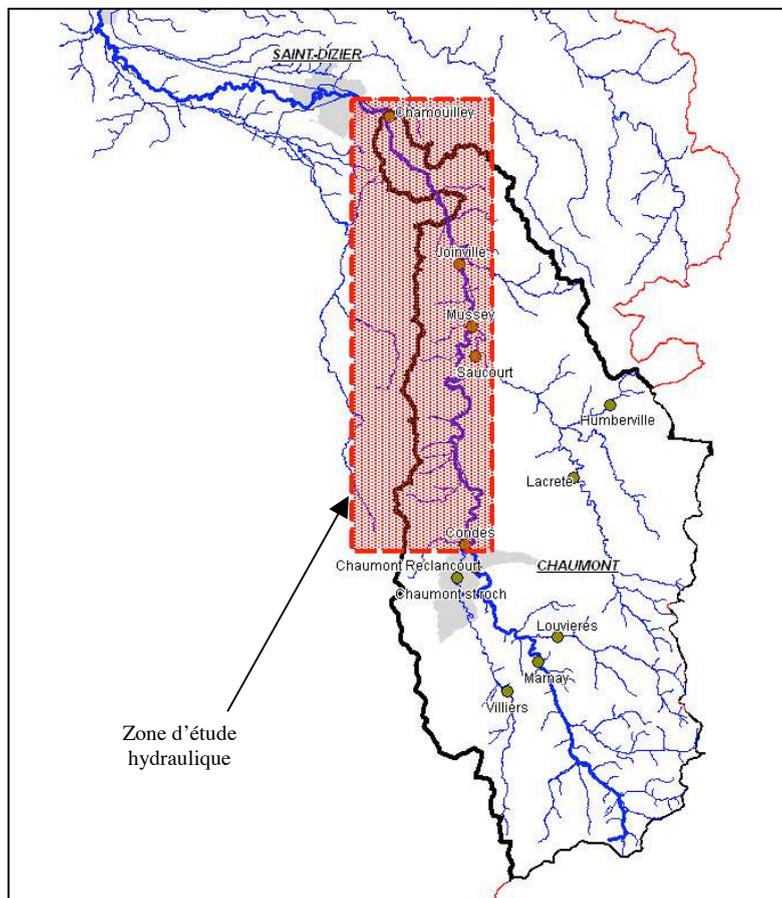


FIG. 43 – Limite de la zone d'étude pour la Marne.

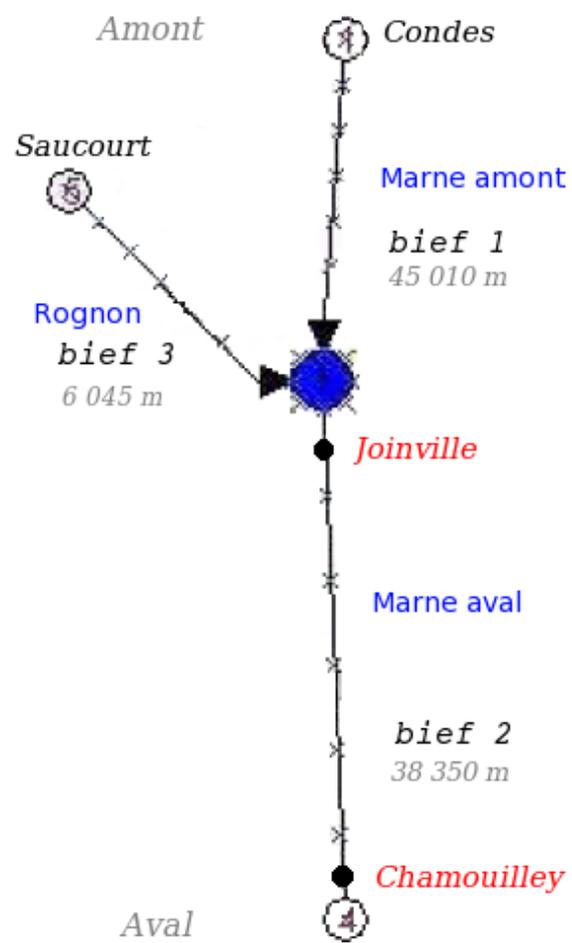


FIG. 44 – Schéma représentant la Marne sous Mascaret.

Annexe 3 : Situation géographique de l'Adour.

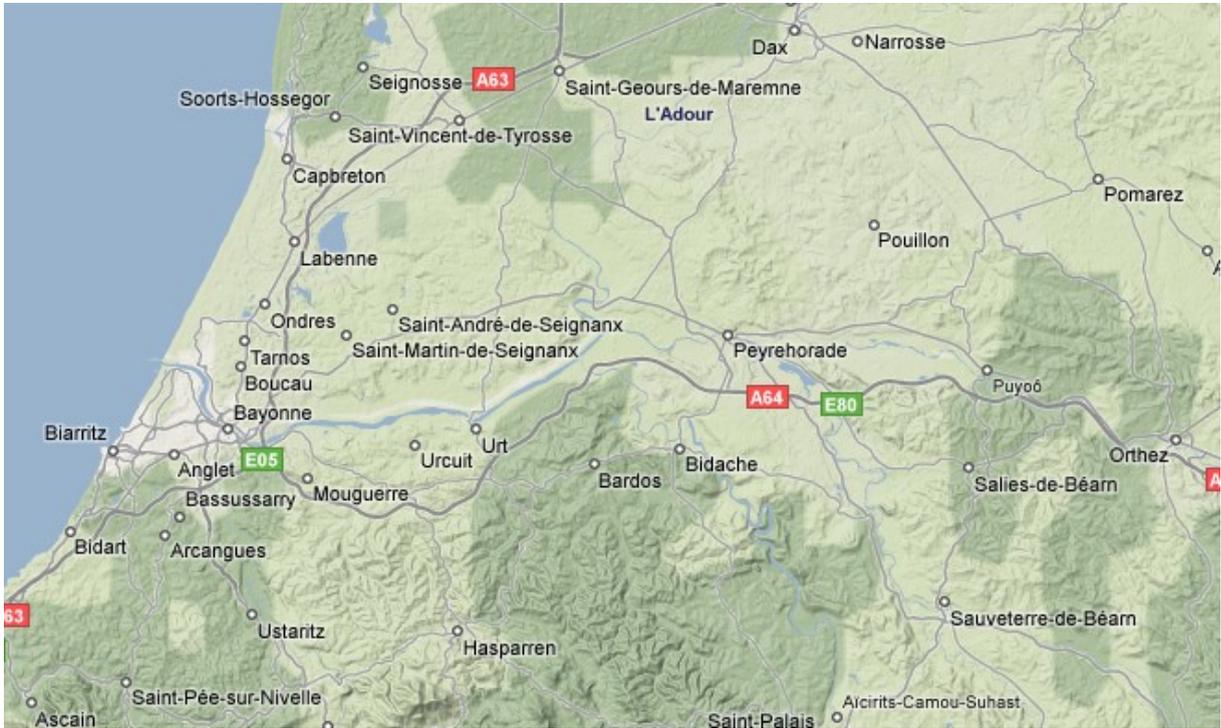


FIG. 45 – Localisation de l'Adour

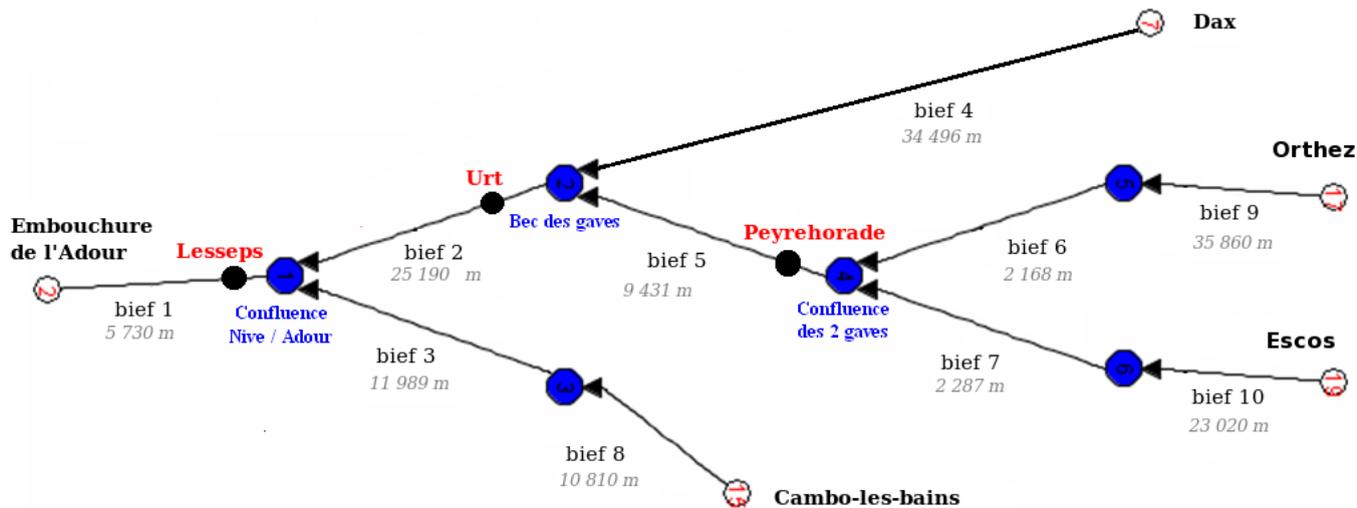


FIG. 46 – Schéma représentant l'Adour.

Expériences jumelles sur un modèle des ondes de crues

O. Thual, 18 janvier 2010

À l’aide d’expériences jumelles d’assimilation de données sur le modèle des ondes de crues forcé par des débits amonts aléatoires, on détermine empiriquement une portée asymétrique de la fonction de corrélations qui propage avec succès des mesures de débits en un point. La symétrie de la fonction de corrélation des erreurs d’ébauche est recouverte en considérant que ces corrélations sont spatio-temporelles.

1 Simulation du modèle des ondes de crues**1.1 Les équations du modèle**

Si $h(x, t)$ est la hauteur d’une rivière et $U = K_s h^{2/3} \sqrt{\sin \gamma}$ sa vitesse, où K_s est le nombre de Strickler et γ l’angle entre le fond plat et l’horizontale, les petites perturbations $(h, U) = (h_n + \tilde{h}, U_n + \tilde{U})$ d’un équilibre (h_n, U_n) obéissent au modèle

$$\frac{\partial \tilde{h}}{\partial t} + c \frac{\partial \tilde{h}}{\partial x} = K \frac{\partial^2 \tilde{h}}{\partial x^2}, \quad (1)$$

où c est la vitesse de propagation des ondes de crue et K est la constante de diffusion. On suppose $c > 0$. On montre ([1]), à partir des équations de Saint-Venant, que $c = 5U_n/3$ et $K = U_n h_n / (2 \tan \gamma)$.

1.2 Les conditions aux limites

On s’intéresse au domaine $x \in [0, L]$. En aval, on suppose que $\frac{\partial \tilde{h}}{\partial t}(L, t) + c \frac{\partial \tilde{h}}{\partial x}(L, t) = 0$ en $x = L$, ce qui correspond à des conditions aux limites ouvertes.

En amont, on impose un débit aléatoire, ce qui s’écrit $\tilde{h}(0, t) = q(t)$ en $x = 0$ dans la mesure où \tilde{U} dépend de \tilde{h} . On suppose que $q(t)$ est une fonction aléatoire qui possède une statistique gaussienne d’écart type q_m et une fonction d’autocorrélation est $R(\tau) = \langle q(t) q(t + \tau) \rangle = q_m^2 \exp\left(-\frac{\tau^2}{2\sigma_q^2}\right)$. La figure 1 montre un réalisation de $q(t)$ pour deux valeurs de σ_q .

1.3 Résultat de simulation

Nous intégrons le modèle avec un schéma Euler explicite en temps et différences finies centrées d’ordre 1 en espace. Le pas de temps est alors $dt = \min(dt_{cfl}, dt_{dif})$ avec $dt_{cfl} = dx/c$ et $dt_{dif} = dx^2/(2K)$ où $dx = L/N$ est le pas d’espace.

Un résultat de simulation avec $c = 1$, $K = 500$, $L = 500 \cdot 10^3$, $N = 1000$ sur un intervalle de temps $t \in [0, T]$ avec $T = 500 \cdot 10^3$ est présenté sur la figure 2. Une condition initiale

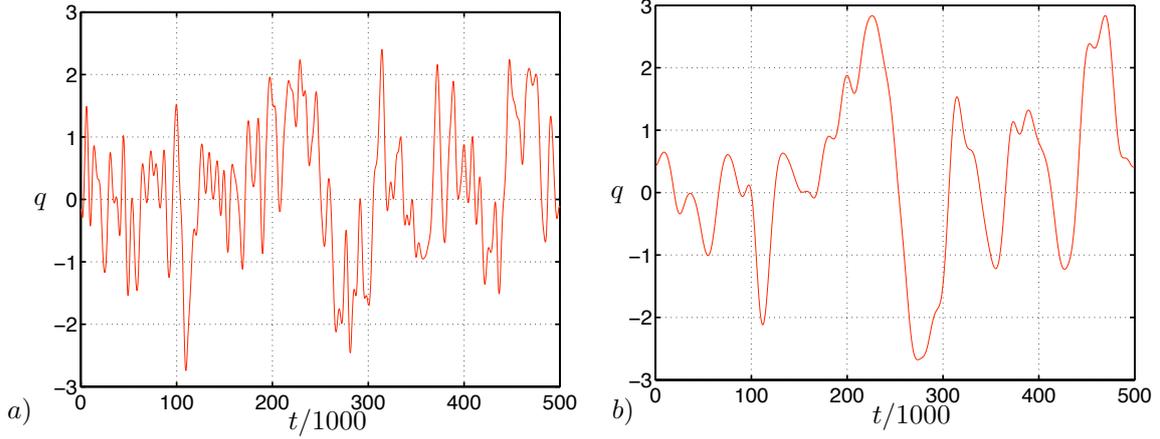


Figure 1: Débit amont aléatoire $q(t)$ pour $q_m = 1$. a) $\sigma_q = 3000$. b) $\sigma_q = 9000$.

$h_0(x) = h_m \exp\left(-\frac{x^2}{2l_0^2}\right)$ est choisie pour donner un aperçu de l'advection et de la diffusion dans la mesure où elle évolue en $h(x, t) = h_m \frac{l_0}{l(t)} \exp\left[-\frac{(x-ct)^2}{2l^2(t)}\right]$ avec $l^2(t) = l_0^2 + 2Kt$.

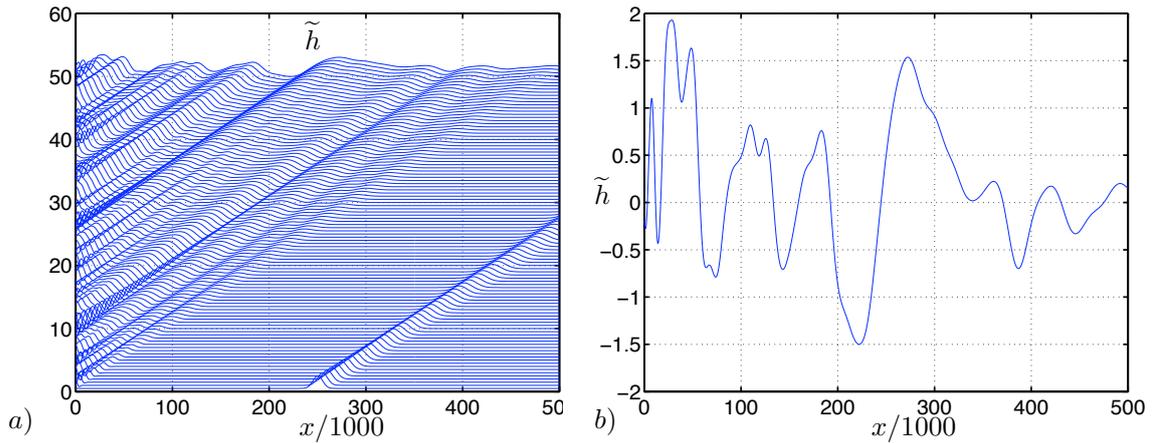


Figure 2: Solution $\tilde{h}(x, t)$ pour $c = 1$ et $K = 500$. a) Superposition des profils $\tilde{h}(x, t)$ décalés. b) Temps $T = 500$.

On voit sur la figure 2b que la longueur de corrélation de la solution $\tilde{h}(x, T)$ au temps final est plus petite en amont ($x = 0$) qu'en aval ($x = L$). On peut estimer qu'elle vaut $l_0 = c\sigma_q$ en amont et $l(x) = \sqrt{l_0^2 + 2Kx/c}$ en x , à l'aide d'un raisonnement simple. De même, le temps de corrélation $\sigma(x) = l(x)/c$ croît avec x comme on peut le voir sur la figure 3.

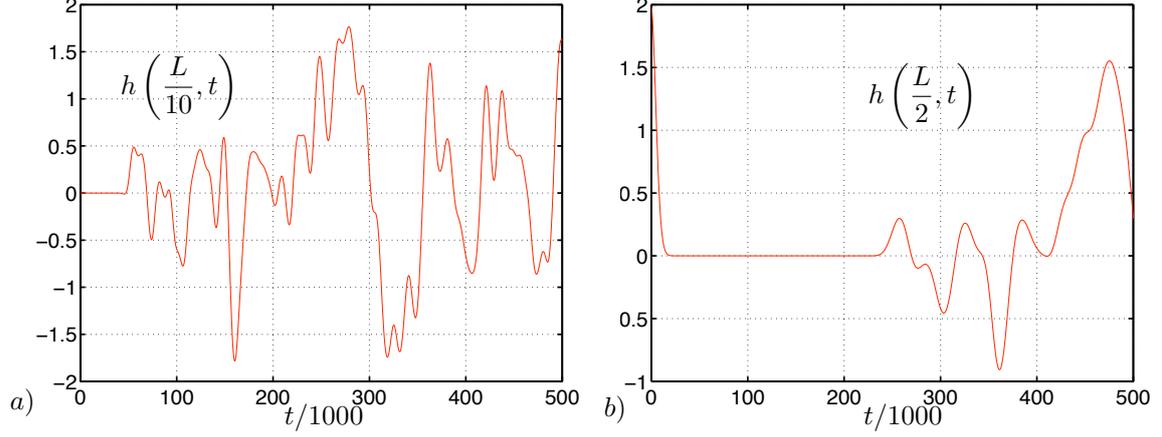


Figure 3: Solution $\tilde{h}(x, t)$ pour $c = 1$ et $K = 500$. Signal temporel en a) $x = L/10$, b) $x = L/2$.

2 Assimilation d'une donnée en un point

2.1 Expérience jumelle

On effectue une première expérience numérique avec $c = c_t$, $K = K_t$ forcé par les débits q_t généré avec les paramètre $(\sigma_q, q_m) = (\sigma_{qt}, q_{mt})$. On considère la solution $h_t(x, t)$ qui en découle comme étant l'état vrai.

On considère que le modèle a pour paramètres $c = c_s$, $K = K_s$ et qu'il est forcé par des débits q_s . On choisit $(c_t, c_s) = (1, .98)$ et $(K_t, K_s) = (500, 1000)$. On choisit $\sigma_{qt} = 30 \cdot 10^3$ et $q_{mt} = 1$ pour générer la fonction aléatoire q_t que l'on perturbe avec une fonction aléatoire δq avec $\sigma_{\delta q} = \sigma_{qt}$ et $\delta q_m = .5$ pour construire $q_s = q_t + \delta q$. La figure 4a compare q_t et q_s ainsi obtenu. La figure 4 compare les solutions $h_t(x, T)$ et $h_s(x, T)$ pour $T = 500 \cdot 10^3$.

On considère une observation $h_{obs}(t_j) = h(x_{obs}, t) = h_t(x_{obs}, t) + \epsilon_r(t)$ où $\epsilon(t)$ est un bruit blanc gaussien d'écart type σ_r . On choisit $x_{obs} = L/2$ et $\sigma_r = 0.1$. La figure 4 montre l'observation obtenue pour $T = 500 \cdot 10^3$. On choisit d'assimiler cette observation tous les $\Delta t = 10 \cdot 10^3$.

2.2 Méthode 3D-Var

On choisit de mettre en place une assimilation de type BLUE pour chaque nouvelle observation et d'utiliser l'analyse $h_a(x, t_j)$ comme nouvelle condition initiale $h_b(x, t_j)$ pour le modèle (c_s, K_s) et q_s qui la transforme en la nouvelle ébauche $h_b(x, t_{j+1})$ pour l'observation suivante.

On a donc $\underline{R} = \sigma_r^2$, $\underline{H} = (0, \dots, 0, 1, 0, \dots, 0)^T$ avec le 1 à la position n_{obs} où n_{obs} est l'indice du point x_{obs} dans le maillage spatial.

Pour la matrice \underline{B} , on choisit une fonction de corrélation $B(x - x_{obs}) = \sigma_b^2 \exp\left[-\frac{(x - x_{obs})^2}{2l_b(x)^2}\right]$ avec $l_b(x) = l_g$ pour $x < x_{obs}$ et $l_b(x) = l_d$ pour $x > x_{obs}$.

La matrice de gain $\underline{\underline{K}} = \underline{\underline{B}} \underline{\underline{H}}^T (\underline{\underline{H}} \underline{\underline{B}} \underline{\underline{H}}^T + \underline{\underline{R}})^{-1}$ et la formule $\underline{x}^a = \underline{x}^b + \underline{\underline{K}} \underline{d}$ où l'innovation est ici $\underline{d} = h_{obs}(t_j) - h_b(x_{obs}, t_j)$ conduit à

$$\tilde{h}_a(x, t_j) = \tilde{h}_b(x, t_j) + [h_{obs}(t_j) - h_b(x_{obs}, t_j)] \frac{\sigma_b^2}{\sigma_b^2 + \sigma_r^2} \exp \left[-\frac{(x - x_{obs})^2}{2 l_b(x)^2} \right]. \quad (2)$$

On choisit $l_g = l(x_{obs})$ et $l_d = \frac{1}{10} l(x_{obs})$ avec $l(x_{obs}) = \sqrt{l_0^2 + 2 K_b x_{obs} / c_b}$ et $l_0 = c_b \sigma_{qt}$. Ce choix et cette asymétrie résultent d'un ajustement "ad hoc" issu de plusieurs expériences.

2.3 Résultat

La figure 4 montre que l'analyse permet de reconstituer de manière précise le profil $\tilde{h}(x, t)$ en aval du point de mesure et de lisser ce profil en amont.

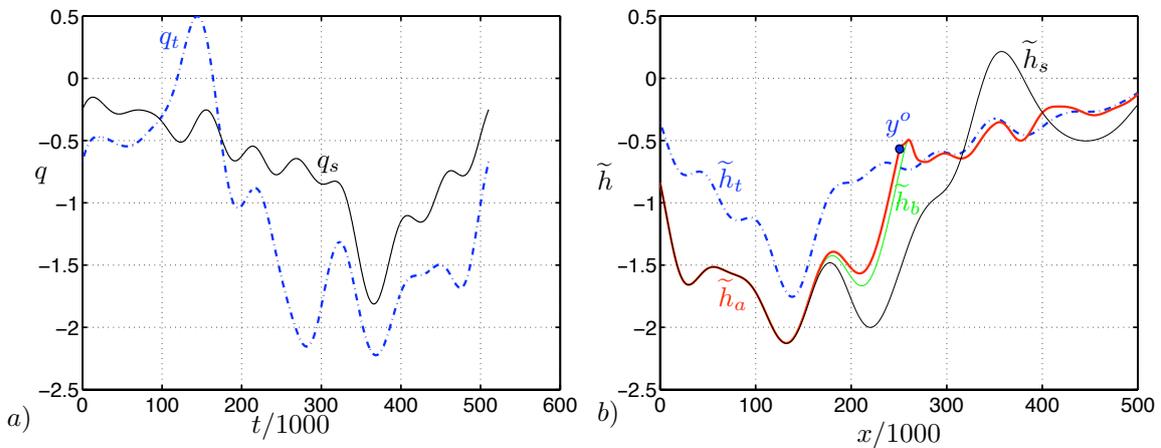


Figure 4: Assimilation BLUE tous les $\Delta t = 5 \cdot 10^3$. a) Débit amont vrai q_t (bleu pointillé) et modélisé q_s (noir plein). b) Profil assimilé à $t = T$ vrai h_t (bleu gras pointillé), sans assimilation h_s (noir plein), analyse h_a (rouge gras), ébauche h_b (vert plein).

3 Discussion

On a montré, sur un exemple concret, que l'introduction de portées asymétriques dans la fonction de corrélation spatiale d'une méthode de type "3D-Var" améliorerait la prévision. Cette constatation pragmatique a une justification théorique pour recouvrer la symétrie d'une fonction de corrélation d'erreurs d'ébauche. Une extension du domaine spatial au domaine spatio-temporel est très certainement le cadre théorique adéquat.

Remerciements

Je remercie S. RICCI et P. ERHARD pour leurs remarques concernant le problème de la symétrie de la fonction de corrélation d'erreur d'ébauche.

References

- [1] O. THUAL, Hydrodynamique de l'environnement, Éditions de l'École Polytechnique (2010).
<http://thual.perso.enseeiht.fr/see/index.htm>
- [2] J. M. SABATER, B. BOURIQUET, S. RICCI, J-Ph. ARGAUD, A. PIACENTINI, S. MASSART, A. WEAVER et O. THUAL, Applications du CERFACS en assimilation de données, *7èmes Journées Scientifiques et Techniques du CETMEF*, (2008).
- [3] O. THUAL, Conférences sur l'assimilation de données (2002-2010).
<http://thual.perso.enseeiht.fr/assimilation.htm>

logbook_prepa_MARNE.txt

```
=====
05/02/2009 - CERFACS - Gabriel Jonville, Andrea Piacentini
Guide d'utilisation pour le SCHAPI
Maquette MASCARET d'assimilation de données
Modèle MARNE-VALLAGE
=====
```

```
Installation dans /home/mascuser/MASCARET_ASSIM_GNU
avec les compilateurs gnu4.4.0
sous /usr/local/gfortran
```

..... 1/ COMPILATION

1.1 MASCARET assimilation linéique -----

```
cd SOURCES/SUPERVISEUR
make clean
cd ../SOURCES/COMPLIB
make clean
rm include/*
rm lib/libmascaret.a
```

```
[Si machine en 64 bits changer le lien symbolique
Make.commands.gnu dans SOURCES/COMPLIB/ ]
```

```
cd SUPERVISEUR
make
```

1.2 MASCARET-PALM calage par assimilation -----

```
cd SOURCES/Assimilation/PALM
make -f Makefile.AP allclean
cd ../SOURCES/COMPPALMLIB
make clean
rm includepalm/*
rm lib/libpalmmascaret.a
```

```
[Si machine en 64 bits changer les liens symboliques
Make.commands.gnu dans SOURCES/COMPPALMLIB
et
Make.include dans SOURCES/Assimilation/PALM ]
```

```
cd SOURCES/Assimilation/PALM
prepalm mascaret_calage.pp1
Generation des fichiers de service des unites, des entites, et des parametres
Ne pas generer le Makefile, ni le Make.include
make -f Makefile.AP
```

..... 2/ MISE EN PLACE D'UN EPISODE MARNE-VALLAGE

```
mkdir EXECMV_01
cd EXECMV_01
cp ../SOURCES/SCRIPTS/MARNE/piloteMV.ksh .

mkdir Observations
cd Observations
cp (base-de-donnees)/2004_01_07.csv .
cp ../../SOURCES/SCRIPTS/MARNE/obs_split_MV.ksh .
./obs_split_MV.ksh

cd ..
```

logbook_prepa_MARNE.txt

```
mkdir DonneesDyn
cd DonneesDyn
cp (base-de-donnees)/1332.loi .
cp (base-de-donnees)/1683.loi .

cd ..
mkdir DonneesStat
cd DonneesStat
cp ../../SOURCES/SCRIPTS/MARNE/Geometrie.geo .
cp ../../SOURCES/SCRIPTS/MARNE/1489CT.loi .
cp ../../SOURCES/SCRIPTS/MARNE/Qapport.loi .
cp ../../SOURCES/SCRIPTS/MARNE/LigneEauInitiale.lig .
cp ../../SOURCES/SCRIPTS/MARNE/ParametresMascaret.cas .

cd ..
mkdir Resultats
cd Resultats
cp ../../SOURCES/SCRIPTS/MARNE/calcul_criteresMV.f90 .
gfortran -o calcul_criteresMV calcul_criteresMV.f90
ln -s ../Observations/Observations_stat258.dat .
ln -s ../Observations/Observations_stat381.dat .
cp ../../SOURCES/SCRIPTS/MARNE/plot_crit_mod381.gp .
cp ../../SOURCES/SCRIPTS/MARNE/plot_crit_mod258.gp .
cp ../../SOURCES/SCRIPTS/MARNE/plot_crit_assim381.gp .
cp ../../SOURCES/SCRIPTS/MARNE/plot_crit_assim258.gp .
ln -s ../../SOURCES/SCRIPTS/MARNE/updres.bash .

cd ..
mkdir Programmes
cd Programmes
mkdir Mascaret
cd Mascaret
ln -s ../../../../SOURCES/SUPERVISEUR/Abaques.txt .
ln -s ../../../../SOURCES/SUPERVISEUR/dico.txt .
ln -s ../../../../SOURCES/SUPERVISEUR/FichierCas.txt .
ln -s ../../../../SOURCES/SUPERVISEUR/systeme_mascaret .

cd ..
mkdir PALM
cd PALM
cp ../../../../SOURCES/SCRIPTS/MARNE/inst_links .
cp ../../../../SOURCES/SCRIPTS/MARNE/uninst_links .
./inst_links

cd ../../..
```

.....

3/ SIMULATION D'UN EPISODE MARNE-VALLAGE

.....

3.1 Usage du piloteMV

Un épisode est simulé avec le script piloteMV.ksh qui enchaîne plusieurs simulations MASCARET par glissement d'une fenêtre temporelle.

Usage du piloteMV pour lancer une série de simulations MASCARET :
./piloteAM.ksh Tdeb(sec) Tfin(sec) pas(sec) tequilibrage(sec) tprevision(sec)
DRYRUN|ASSIM|CALAGE|CALAGEASSIM [FAST]

Typiquement avec le modèle MarneVallage on utilise le pilote avec :

pas	= 1h	= 3600s
tequilibrage	= 4j	= 345600s
tprevision	= 1j	= 86400s

3.2 Exemples d'utilisation du piloteMV

Exemples d'utilisation du pilote :

EXECMV_01 :
 ./piloteMV.ksh 352800 1900800 3600 345600 86400 DRYRUN|ASSIM|CALAGE|CALAGEASSIM
 [FAST]
 (4j,2h) (22j) (1h) (4j) (1j)

EXECMV_02 :
 ./piloteMV.ksh 352800 1296000 3600 345600 86400 DRYRUN|ASSIM|CALAGE|CALAGEASSIM
 [FAST]
 (4j,2h) (15j) (1h) (4j) (1j)

EXECMV_03 :
 ./piloteMV.ksh 352800 1638000 3600 345600 86400 DRYRUN|ASSIM|CALAGE|CALAGEASSIM
 [FAST]
 (4j,2h)(18j,23h)(1h) (4j) (1j)

EXECMV_05 :
 ./piloteMV.ksh 352800 1296000 3600 345600 86400 DRYRUN|ASSIM|CALAGE|CALAGEASSIM
 [FAST]
 (4j,2h) (15j) (1h) (4j) (1j)

logbook_prepa_ADOUR.txt

```
=====
05/02/2009 - CERFACS - Gabriel Jonville, Andrea Piacentini
Guide d'utilisation pour le SCHAPI
Maquette MASCARET d'assimilation de données
Modèle ADOUR-MARITIME
=====
```

```
Installation dans /home/mascuser/MASCARET_ASSIM_GNU
avec les compilateurs gnu4.4.0
sous /usr/local/gfortran
```

..... 1/ COMPILATION

1.1 MASCARET assimilation linéique

```
-----
cd SOURCES/SUPERVISEUR
make clean
cd ../SOURCES/COMPLIB
make clean
rm include/*
rm lib/libmascaret.a
```

```
[Si machine en 64 bits changer le lien symbolique
Make.commands.gnu dans SOURCES/COMPLIB/ ]
```

```
cd SUPERVISEUR
make
```

1.2 MASCARET-PALM calage par assimilation

```
-----
cd SOURCES/Assimilation/PALM
make -f Makefile.AP allclean
cd ../SOURCES/COMPPALMLIB
make clean
rm includepalm/*
rm lib/libpalmmascaret.a
```

```
[Si machine en 64 bits changer les liens symboliques
Make.commands.gnu dans SOURCES/COMPPALMLIB
et
Make.include dans SOURCES/Assimilation/PALM ]
```

```
cd SOURCES/Assimilation/PALM
prepalm mascaret_calage.pp1
Generation des fichiers de service des unites, des entites, et des parametres
Ne pas generer le Makefile, ni le Make.include
make -f Makefile.AP
```

..... 2/ MISE EN PLACE D'UN EPISODE ADOUR-MARITIME

```
mkdir EXECAM_01
cd EXECAM_01
cp ../SOURCES/SCRIPTS/ADOUR/piloteAM_pdtobs1h.ksh piloteAM.ksh (si pdt_obs = 1h)
[ou cp ../SOURCES/SCRIPTS/ADOUR/piloteAM_pdtobs5mn.ksh piloteAM.ksh (si pdt_obs
= 5mn) ]
```

```
mkdir Observations
cd Observations
cp (base-de-donnees)/obs_ep12.csv .
cp ../../SOURCES/SCRIPTS/ADOUR/obs_split_AM.ksh .
./obs_split_AM.ksh
```

logbook_prepa_ADOUR.txt

```
cd ..
mkdir DonneesDyn
cd DonneesDyn
cp (base-de-donnees)/1023.loi .
cp (base-de-donnees)/1029.loi .
cp (base-de-donnees)/1030.loi .
cp (base-de-donnees)/1049.loi .
cp (base-de-donnees)/1467.loi .

cd ..
mkdir DonneesStat
cd DonneesStat
cp ../../SOURCES/SCRIPTS/ADOUR/Geometrie.geo .
cp ../../SOURCES/SCRIPTS/ADOUR/LigneEauInitiale.lig .
cp ../../SOURCES/SCRIPTS/ADOUR/ParametresMascaret_pdtobs1h.cas
ParametresMascaret.cas (si pdt_obs = 1h)
[ou cp ../../SOURCES/SCRIPTS/ADOUR/ParametresMascaret_pdtobs5mn.cas
ParametresMascaret.cas (si pdt_obs = 5mn) ]

cd ..
mkdir Resultats
cd Resultats
cp ../../SOURCES/SCRIPTS/ADOUR/calcul_criteresAM.f90 .
gfortran -o calcul_criteresAM calcul_criteresAM.f90
ln -s ../Observations/Observations_stat1754.dat .
ln -s ../Observations/Observations_stat296.dat .
ln -s ../Observations/Observations_stat10.dat .
cp ../../SOURCES/SCRIPTS/ADOUR/plot_crit_mod1754.gp .
cp ../../SOURCES/SCRIPTS/ADOUR/plot_crit_mod296.gp .
cp ../../SOURCES/SCRIPTS/ADOUR/plot_crit_mod10.gp .
cp ../../SOURCES/SCRIPTS/ADOUR/plot_crit_assim1754.gp .
cp ../../SOURCES/SCRIPTS/ADOUR/plot_crit_assim296.gp .
cp ../../SOURCES/SCRIPTS/ADOUR/plot_crit_assim10.gp .
ln -s ../../SOURCES/SCRIPTS/ADOUR/updres.bash .
cp ../Observations/obs_ep12.csv .

cd ..
mkdir Programmes
cd Programmes
mkdir Mascaret
cd Mascaret
ln -s ../../SOURCES/SUPERVISEUR/Abaques.txt .
ln -s ../../SOURCES/SUPERVISEUR/dico.txt .
ln -s ../../SOURCES/SUPERVISEUR/FichierCas.txt .
ln -s ../../SOURCES/SUPERVISEUR/systeme_mascaret .

cd ..
mkdir PALM
cd PALM
cp ../../SOURCES/SCRIPTS/ADOUR/inst_links .
cp ../../SOURCES/SCRIPTS/ADOUR/uninst_links .
./inst_links

cd ../../

.....
3/ SIMULATION D'UN EPISODE ADOUR-MARITIME
.....

3.1 Usage du piloteAM
-----
Un épisode est simulé avec le script piloteAM.ksh qui enchaîne
plusieurs simulations MASCARET par glissement d'une fenêtre temporelle.

Usage du piloteAM pour lancer une série de simulations MASCARET :
./piloteAM.ksh Tdeb(sec) Tfin(sec) pas(sec) tequilibrage(sec) tprevision(sec)
```

DRYRUN|ASSIM|CALAGE|CALAGEASSIM [FAST]

Typiquement avec le modele AdourMaritime on utilise le pilote avec :
 pas = 1h = 3600s
 tequilibrage = 3j = 259200s
 tprevision = 1j = 86400s

3.2 Exemples d'utilisation du piloteAM

```
-----
EXECAM_01 : ep12_crue1
./piloteAM.ksh 1555200 2246400 3600 259200 86400
DRYRUN|ASSIM|CALAGE|CALAGEASSIM [FAST]
                (18j)  (26j)  (1h)  (3j)  (1j)

EXECAM_02 : ep13_crue2
./piloteAM.ksh 518400 1209600 3600 259200 86400 DRYRUN|ASSIM|CALAGE|CALAGEASSIM
[FAST]
                (6j)  (14j)  (1h)  (3j)  (1j)

EXECAM_03 : ep15_crue3
./piloteAM.ksh 259200 1036800 3600 259200 86400 DRYRUN|ASSIM|CALAGE|CALAGEASSIM
[FAST]
                (3j)  (12j)  (1h)  (3j)  (1j)

EXECAM_04 : ep16_crue4
./piloteAM.ksh 2592000 3283200 3600 259200 86400
DRYRUN|ASSIM|CALAGE|CALAGEASSIM [FAST]
                (30j)  (38j)  (1h)  (3j)  (1j)

EXECAM_05 : ep17_crue5
./piloteAM.ksh 172800 432000 3600 172800 86400 DRYRUN|ASSIM|CALAGE|CALAGEASSIM
[FAST]
                (2j)  (5j)  (1h)  (2j)  (1j)

EXECAM_07 :
./piloteAM.ksh 1555200 2592000 3600 259200 86400 DRYRUN|ASSIM|CALAGE|CALAGEASSIM
[FAST]
                (18j)  (30j)  (1h)  (3j)  (1j)

EXECAM_AVR09 :
./piloteAM.ksh 1468800 1900800 3600 259200 86400 DRYRUN|ASSIM|CALAGE|CALAGEASSIM
[FAST]
                (17j)  (22j)  (1h)  (3j)  (1j)

EXECAM_MAR09 :
./piloteAM.ksh 1123200 1728000 3600 259200 86400 DRYRUN|ASSIM|CALAGE|CALAGEASSIM
[FAST]
                (13j)  (20j)  (1h)  (3j)  (1j)

EXECAM_NOV09_1CRUE :
./piloteAM.ksh 259200 1208400 3600 259200 43200 DRYRUN|ASSIM|CALAGE|CALAGEASSIM
[FAST]
                (3j)  (~14j)  (1h)  (3j)  (0.5j)

EXECAM_NOV09_2CRUES :
./piloteAM.ksh 604800 1382400 3600 259200 43200 DRYRUN|ASSIM|CALAGE|CALAGEASSIM
[FAST]
                (7j)  (16j)  (1h)  (3j)  (0.5j)
```