

**Convention 2010-2011
CERFACS-SCHAPI-LNHE-CETMEF-SCP**

**Maquette d'assimilation de données
avec MASCARET**

Présentation maquette assimilation MASCARET

A. Piacentini, S. Ricci, O. Thual, G. Jonville, S. Massart, Ch. Ivanoff, E. Le Pape, N. Goutal, F. Zaoui

Document du 29 Avril 2011

Index

Introduction	3
Partie I : Rappel théorique	4
I.1 Assimilation de la ligne d'eau.....	4
Principes algorithmiques.....	4
Choix du contrôle.....	4
Vecteur d'observations.....	4
Opérateur d'observation.....	4
Matrice R des variances des erreurs d'observation.....	5
Matrice B des variances/covariances des erreurs d'ébauche.....	5
Formule d'interpolation optimale.....	5
Lissage et seuillage des observations.....	5
Modélisation de la forme des fonctions de covariances d'erreur d'ébauche.....	6
Détermination expérimentale des longueurs de corrélation des erreurs d'ébauche...	6
Détermination de la correction sur Q (débit).....	7
I.2 Assimilation pour le calage en temps réel des lois aux extrémités libres.....	7
Principes algorithmiques.....	7
Choix du contrôle.....	8
Vecteur d'observations.....	8
Opérateur d'observation.....	8
Matrice R des variances des erreurs d'observation.....	8
Matrice B des variances/covariances des erreurs d'ébauche.....	8
Formule d'interpolation optimale.....	8
Lissage et seuillage des observations.....	9
Détermination de G et GT par différences finies.....	9
Partie II : Exemple d'application	11
II.1 Cas d'étude :.....	11
Description du bassin	11
Description du cas dans le fichier Parametres.cas.....	12
II.2 Exécution du run libre et assimilation linéique.	18
Compilation de l'exécutable MASCARET sans PALM.....	18
Mise en place du cas test.....	19
Execution du run libre (dryrun).....	20
Run libre avec apport latéral arbitraire.....	22
Experience avec assimilation linéique.....	23
II. 3 Calage des lois par assimilation.....	26
Compilation et mise en place du cas test.....	26
Explication de la partie .cas qui concerne le calage.....	29
Cas de calage de l'apport supplémentaire.....	31
Cas de calage des lois amont.....	34
II.4 Combinaison avec l'assimilation linéique.....	36

Cas de calage apport supplémentaire + assimilation linéique.....	36
Bouclage des expériences.....	37

Annexes :

Annexe 1 : Récapitulatif Adour

Annexe 2 : Récapitulatif Marne

Annexe 3 : 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 (30 Novembre 2009)

Annexe 4 : Assimilation de données sur un modèle d'ondes de crues diffusantes – Propagation des fonctions de covariance (Rapport de stage ENM 2011).

Annexe 5 : Présentation CERFACS 29 Avril 2011

Annexe 6 : Guide d'utilisation de la machine virtuelle SCHAPI

Annexe 7 : Présentation SCHAPI 29 Avril 2011

Annexe 8 : Présentation SPC-SACN 29 Avril 2011

Introduction

Ce document décrit la chaîne d'assimilation de données pour le modèle MASCARET.

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 ». Dans ce document tous les exemples sont tirés d'une application sur le bassin « Adour Maritime ».

En se basant sur une approche d'assimilation de type BLUE (Best Linear Unbiased Estimate), c'est à dire un Filtre de Kalman sans évolution des matrices de covariance d'erreur d'ébauche, 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.

Les rappels algorithmiques de la première partie mettent en évidence les paramètres de l'assimilation qui doivent être fixés dans l'utilisation pratique illustrée en deuxième partie.

Partie I : Rappel théorique

I.1 Assimilation de la ligne d'eau

La première approche 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 dans le cas de l'Adour Maritime.

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

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

\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 rejette 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 à 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.

Pour comparer les observations aux valeurs simulées, 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}^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 R des variances des erreurs d'observation

Les informations statistiques sur les erreurs d'observation sont contenues dans la matrice R, de taille $p \times p$. Nous négligeons les corrélations d'erreur entre différentes stations. La matrice 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 `assim.cas`.

Matrice B des variances/covariances des erreurs d'ébauche

En théorie, la matrice 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 `assim.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{BH}^T , c'est à dire seulement la partie pour Z des p colonnes de 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 le biais de la courbe de tarage. La correction sur Q est ensuite propagée aux autres sections par multiplication par les colonnes de 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 $\underline{d} = \underline{y}^o - \underline{Hx}^b$.

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

Dans la pratique on construit directement les p colonnes de \underline{BH}^T et la matrice de taille p \underline{HBH}^T . p étant petit, la matrice $(\underline{HBH}^T + \underline{R})$ 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 par valeur constante les données brutes au delà du dernier temps observé. La moyenne peut ainsi être calculée jusqu'au dernier temps utile, mais sur la deuxième moitié de la dernière fenêtre on remplace la valeur calculée par une combinaison linéaire avec la valeur brute, de façon à faire coïncider la dernière valeur lissée avec la dernière valeur brute. 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 rejette 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 rejette 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.

Modélisation de la forme des fonctions de covariances d'erreur d'ébauche

Les fonctions de covariance d'erreur d'ébauche aux points d'observation sont modélisées par 2 demi gaussiennes de longueur de portée différente en amont et en aval du point d'observation. Ce choix repose sur les conclusions d'une étude de la propagation d'une fonction de covariance initialement Gaussienne par un Filtre de Kalman complet, implémenté sur un modèle d'ondes de crues diffusantes. Les détails de cette étude se trouvent en Annexe 4. Les longueurs de portée en amont et aval de chaque point d'obs, pour Z et Q sont précisées dans le fichier assim.cas.

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 la scansion récursive 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 expérimentale des longueurs de corrélation des erreurs d'ébauche

Les longueurs de corrélation aux point d'observation ont été déterminées de manière expérimentale.

Une brusque perturbation est ajoutée à chaque station amont du réseau. La propagation de cette perturbation au point d'observation permet de caractériser le coefficient de diffusion κ ainsi que la célérité des ondes c pour l'écoulement. Si on note s_r l'abscisse curviligne de la station d'obs par rapport à la station amont et t_r le temps de propagation de la perturbation au point d'obs, on peut écrire $c = s_r / t_r$. De plus la longueur de corrélation

du signal au point d'observation est donnée par $I^2(t_r) = I^2(t_0) + 2 \kappa t_r$. $I(t_r)$ peut être calculé approximativement en mesurant l'étalement de la perturbation en débit (en temps T) tel que $I(t_r) = c T$. En supposant $I^2(t_0) = 0$, on peut écrire $\kappa = s_r^2 T / 2 t_r^3$.

La longueur de portée de la fonction de covariance spatiale L_p du signal au point d'observation x_{obs} est donnée par $L_p^2(x_{obs}) = L_p^2(x_0) + 4 \kappa s_r/c$ avec $L_p^2(x_0)$ la longueur de portée de la fonction de covariance initiale propagée par MBM^T. En supposant que $4 \kappa s_r/c \gg L_p^2(x_0)$, on obtient une estimation de $L_p^2(x_{obs})$.

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. Pour cela on utilise la courbe de tarage au point d'observation (connue ou établie expérimentalement) de la forme $Q=f(Z)=a(Z)^r+b$. Les coefficients a,r,b sont spécifiés dans le fichier assim.cas. L'identification des coefficients a,b,r n'est pas toujours évidente, notamment en présence de la marée, la fonction f est non bijective. Dans ce cas, on fait l'hypothèse $f = Id$ (soit $a=1, r=1, b=0$).

La correction en débit au point d'observation s'écrit :

$$dQ = [f(Z+\delta Z)-f(Z^b)] * Q^b / Z^b$$

Cette correction est égale à $f(Z+\delta Z)-f(Z^b)$ si (Z^b, Q^b) est tel que $Q^b = f(Z^b)$.

Cette correction est égale à $\delta Z * Q^b / Z^b$ si $f = Id$.

Si l'on disposait d'une courbe de tarage valable en tout point et à tout instant on obtiendrait facilement cette relation.

1.2 Assimilation pour le calage en temps réel des lois aux extrémités libres

La deuxième approche 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.

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. Après une étude de sensibilité et de stabilité, nous avons décidé de contrôler les lois imposées aux extrémités libres. Après avoir fixé un contrôle représentatif de ces lois, il faut définir un opérateur d'observation généralisé \mathbf{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 $\underline{\mathbf{G}}$. En définissant les matrices des variances/covariances d'erreur sur les observations ($\underline{\mathbf{R}}$) et sur l'ébauche du contrôle ($\underline{\mathbf{B}}$), le contrôle issu de l'assimilation sera

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

où $\underline{\mathbf{d}}$ est le misfit sur l'intervalle choisie $\underline{\mathbf{d}} = \mathbf{y}^o - \mathbf{G}(\mathbf{x}^b)$

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

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

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^0 est un temps de mesure, nous faisons correspondre l'observation au pas de temps du modèle t_j le plus près de t^0 . Pour cette chaîne nous prenons en compte toutes les observations sur un intervalle temporel fixe.

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 un 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 G relie les éléments du vecteur de contrôle (donc un jeu de coefficients a_j, b_j, c_j) aux observations aux stations de mesure. Dans la pratique il peut être vu comme une intégration de Mascaret sur toute l'étendue de la simulation avec lois aux extrémités libres déterminées par le contrôle, accompagnée d'une sélection des valeurs de Z aux sections observées dans l'intervalle temporel d'assimilation.

Matrice R des variances des erreurs d'observation

Les informations statistiques sur les erreurs d'observation sont contenues dans la matrice R, de taille pxp. Nous négligeons les corrélations d'erreur entre différentes stations et entre différents pas d'observation. La matrice 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 B des variances/covariances des erreurs d'ébauche

Pour la matrice 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 \underline{d} est le misfit sur l'intervalle choisie $\underline{d} = \underline{y}^o - \mathbf{G}(\underline{x}^b)$.

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 \mathbf{G} et son adjoint, ou la matrice qui représente l'opérateur linéaire tangent et sa transposée. Puisque \mathbf{G} 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 fournit aussi le vecteur $\underline{y}^b = \mathbf{G}(\{a_j^b, b_j^b, c_j^b\})$ pour le calcul du misfit \underline{d} .

Une observation importante concerne le seuillage des valeurs des lois. Pour 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. De plus on peut négliger la relation entre le contrôle et certaines stations d'observation : par exemple sur le bassin de l'Adour les observations à Urt et Peyrehorade n'influencent pas le contrôle de la loi à Cambo-les-Bains (Nive Amont).

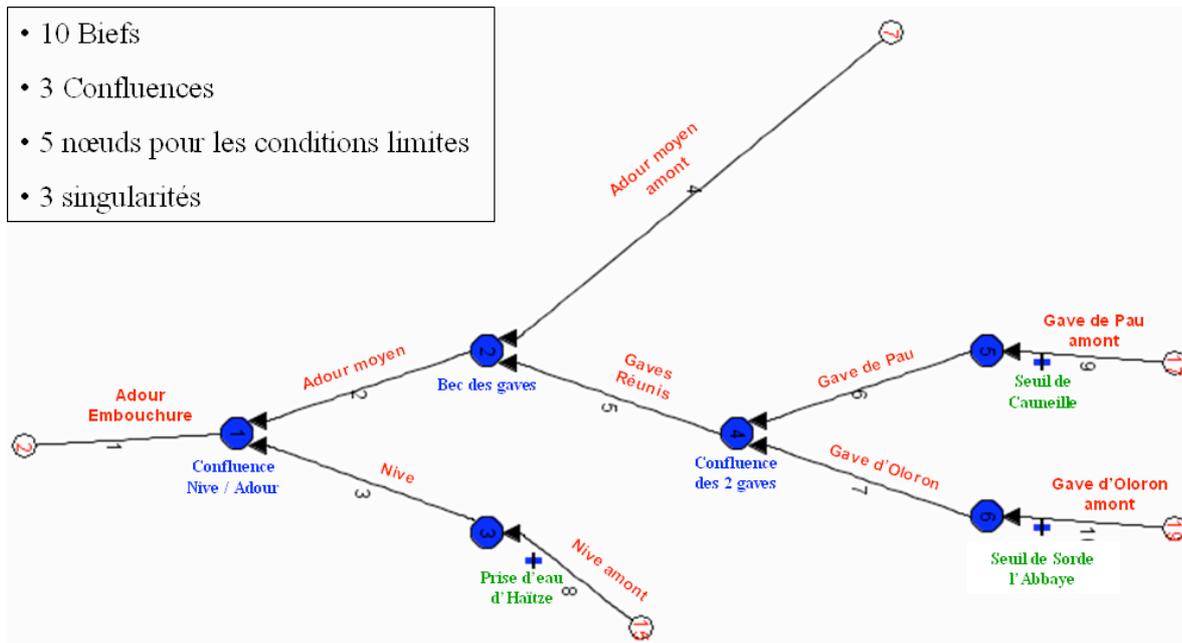
Partie II : Exemple d'application

II.1 Cas d'étude :

Description du bassin

Bassin : Adour Maritime

Schéma du réseau hydraulique :



Régime : Modélisation en régime non permanent fluvial

Bathymétrie/Topographie : Le modèle comprend des berges du lit mineur infranchissables : verticales et infinies (pas de prise en compte le lit majeur) sauf à Peyrehorade. Une surestimation à Urt à partir des cotes de débordement est donc probable.

Conditions limites : 5 extrémités pour imposer les conditions limites

- Conditions limites amont : Débit imposé $Q(t)$ à Cambo-les-Bains, Dax, Orthez, Escos
- Condition limite aval : Cote imposée $H(t)$ à Convergent

Points de sortie du modèle :

Points de sortie	
Code .pre	nom
11	Peyrehorade
41	Urt
52	Lesseps
151	Resplandy
150	Petit Bayonne
85	Pont Blanc

Calage : d'après SCHAPI/SPC. Simulation sur 12 évènements :

- 4 épisodes courants (faible débit, débit constant)
- 8 épisodes de crue atteignant les niveaux de vigilance jaune jusqu'à la limite du rouge

Validation des résultats avec les données des marégraphes du SPC (Lesseps pour l'Adour embouchure, Station provisoire du Pont Blanc pour la Nive, Urt pour l'Adour moyen, Dax pour l'Adour moyen amont, Peyrehorade pour l'Adour moyen)

Le modèle est calé pour donner les meilleurs résultats possibles pour des niveaux élevés.

Évènement : crue de novembre 2000 (épisode 12 du rapport de calage SCHAPI/SPC)
 Nous nous concentrons sur 3 jours et demi à partir du jour 20

Observations : hauteur d'eau aux stations de Peyrorade, Urt, Lesseps.

Modification pour cette démo : ajout d'un apport linéique d'un km, en amont de la station de Peyrehorade sur le bief n. 5 (Gaves Réunis) pour représenter un apport par des affluents non modélisés. Cette possibilité cherche à approcher une configuration pour laquelle les lois amonts sont connues mais où l'incertitude provient d'un apport ou d'un prélèvement dont on ne connaît ni la localisation précise, ni la valeur.

Description du cas dans le fichier Parametres.cas

```

/*****
/ FICHER DE PARAMETRES MASCARET ADOUR MARITIME
/ version 1.4 - MAJ 2007/06/13
/*****
/
/PARAMETRES GENERAUX (1)
/
VERSION DU CODE = 2
NOYAU DE CALCUL = 2
FICHER DES MOT-CLES = '././DonneesStat/ParametresMascaret.cas'
DICTIONNAIRE = 'dico.txt'
PROGRAMME PRINCIPAL = 'princi.f'
SAUVEGARDE DU MODELE = FAUX
FICHER SAUVEGARDE DU MODELE = 'mascaret1.tmp'
CALCUL POUR VALIDATION DU CODE = FAUX
TYPE DE CALCUL DE VALIDATION EFFECTUE = 1
PRESENCE DE CASIERS = FAUX
/
/BIBLIOTHEQUES (2)
/
BIBLIOTHEQUES = 'mascaretV5P1.a damoV3P0.a'
/
/MODELISATION PHYSIQUE (1)
/
PERTES DE CHARGE AUTOMATIQUE AUX CONFLUENTS = FAUX
COMPOSITION DES LITS = 1
CONSERVATION DU FROTTEMENT SUR LES PAROIS VERTICALES = VRAI
ELEVATION DE COTE ARRIVEE DU FRONT = 0.05
INTERPOLATION LINEAIRE DES STRICKLER = VRAI
/
/DEBORDEMENT PROGRESSIF (2)
/
DEBORDEMENT PROGRESSIF LIT MAJEUR = VRAI
DEBORDEMENT PROGRESSIF ZONES DE STOCKAGE = FAUX
/
/PARAMETRES NUMERIQUES (1)
/
CALCUL D'UNE ONDE DE SUBMERSION = FAUX
FROUDE LIMITE POUR LES CONDITIONS LIMITES = 1000.0

```

```

TRAITEMENT IMPLICITE DU FROTTEMENT = FAUX
HAUTEUR D'EAU MINIMALE = 0.0050
IMPLICATION DU NOYAU TRANSCRITIQUE = FAUX
OPTIMISATION DU NOYAU TRANSCRITIQUE = FAUX
PERTES DE CHARGE AUTOMATIQUE NOYAU TRANSCRITIQUE = VRAI
/
/PARAMETRES TEMPORELS (1)
/
PAS DE TEMPS = 300.0
TEMPS INITIAL = 1728000 → 20ème jour
CRITERE D'ARRET DU CALCUL = 1
NOMBRE DE PAS DE TEMPS = 2304
TEMPS MAXIMUM = 2030400 → Durée 3 jours et demi
PAS DE TEMPS VARIABLE SUIVANT NOMBRE DE COURANT = FAUX
NOMBRE DE COURANT SOUHAITE = 0.8
COTE MAXIMALE DE CONTROLE = 0.0
POINT DE CONTROLE ABSCISSE = 0.0
POINT DE CONTROLE BIEF ASSOCIE = 1
/
/GEOMETRIE-RESEAU (1)
/
/
/GEOMETRIE (2)
/
FICHER DE GEOMETRIE = '..../DonneesStat/Geometrie.geo'
FORMAT DU FICHER DE GEOMETRIE = 2
PROFILS EN ABSCISSE ABSOLUE = FAUX
/
/BRANCHES (2)
/
NOMBRE DE BRANCHES = 10
BRANCHE NUMERO =
1;2;3;4;5;6;7;8;9;10
ABSCISSE DEBUT =
0.0;0.0;0.0;0.0;0.0;0.0;0.0;0.0;0.0;0.0
ABSCISSE FIN =
5730.0;25190.0;11989.0;34496.0;9431.0;2168.0;2287.0;10810.0;35860.0;2302
0.0
NUM DE L'EXTREMITE DE DEBUT =
1;3;5;7;9;11;13;15;17;19
NUM DE L'EXTREMITE DE FIN =
2;4;6;8;10;12;14;16;18;20
/
/NOEUDS (2)
/
NOMBRE DE NOEUDS = 6
NOEUD 1 =
1;4;6;0;0
NOEUD 2 =
3;8;10;0;0
NOEUD 3 =
5;16;0;0;0
NOEUD 4 =
9;12;14;0;0
NOEUD 5 =
11;18;0;0;0
NOEUD 6 =
13;20;0;0;0
/
/EXTREMITES LIBRES (2)
/
NOMBRE D'EXTREMITES LIBRES = 5
EXTREMITE LIBRE NUMERO =
1;2;3;4;5
EXTREMITE NUMERO =
2;7;15;17;19
NOM EXTREMITE =
'limite2';'limite7';'limite15';'limite17';'limite19'
TYPE DE CONDITION =
2;1;1;1;1
NUMERO DE LA LOI =
5;2;1;3;4
/
/CONFLUENTS (1)

```

```

/
NOMBRE DE CONFLUENTS = 6
/
/CONFLUENT NUMERO 1 (2)
/
NOMBRE D'AFFLUENTS DU CONFLUENT 1 = 3
NOM DU CONFLUENT 1 = 'Noeud1'
ABSCISSE DE L'AFFLUENT DU CONFLUENT 1 =
0.0;0.0;0.0
ORDONNEE DE L'AFFLUENT DU CONFLUENT 1 =
0.0;0.0;0.0
ANGLE DE L'AFFLUENT DU CONFLUENT 1 =
0.0;0.0;0.0
/
/CONFLUENT NUMERO 2 (2)
/
NOMBRE D'AFFLUENTS DU CONFLUENT 2 = 3
NOM DU CONFLUENT 2 = 'Noeud2'
ABSCISSE DE L'AFFLUENT DU CONFLUENT 2 =
0.0;0.0;0.0
ORDONNEE DE L'AFFLUENT DU CONFLUENT 2 =
0.0;0.0;0.0
ANGLE DE L'AFFLUENT DU CONFLUENT 2 =
0.0;0.0;0.0
/
/CONFLUENT NUMERO 3 (2)
/
NOMBRE D'AFFLUENTS DU CONFLUENT 3 = 2
NOM DU CONFLUENT 3 = 'Noeud3'
ABSCISSE DE L'AFFLUENT DU CONFLUENT 3 =
0.0;0.0
ORDONNEE DE L'AFFLUENT DU CONFLUENT 3 =
0.0;0.0
ANGLE DE L'AFFLUENT DU CONFLUENT 3 =
0.0;0.0
/
/CONFLUENT NUMERO 4 (2)
/
NOMBRE D'AFFLUENTS DU CONFLUENT 4 = 3
NOM DU CONFLUENT 4 = 'Noeud4'
ABSCISSE DE L'AFFLUENT DU CONFLUENT 4 =
0.0;0.0;0.0
ORDONNEE DE L'AFFLUENT DU CONFLUENT 4 =
0.0;0.0;0.0
ANGLE DE L'AFFLUENT DU CONFLUENT 4 =
0.0;0.0;0.0
/
/CONFLUENT NUMERO 5 (2)
/
NOMBRE D'AFFLUENTS DU CONFLUENT 5 = 2
NOM DU CONFLUENT 5 = 'Noeud5'
ABSCISSE DE L'AFFLUENT DU CONFLUENT 5 =
0.0;0.0
ORDONNEE DE L'AFFLUENT DU CONFLUENT 5 =
0.0;0.0
ANGLE DE L'AFFLUENT DU CONFLUENT 5 =
0.0;0.0
/
/CONFLUENT NUMERO 6 (2)
/
NOMBRE D'AFFLUENTS DU CONFLUENT 6 = 2
NOM DU CONFLUENT 6 = 'Noeud6'
ABSCISSE DE L'AFFLUENT DU CONFLUENT 6 =
0.0;0.0
ORDONNEE DE L'AFFLUENT DU CONFLUENT 6 =
0.0;0.0
ANGLE DE L'AFFLUENT DU CONFLUENT 6 =
0.0;0.0
/
/PLANIMETRAGE ET MAILLAGE (1)
/
METHODE DE CALCUL DU MAILLAGE = 5
/
/PLANIMETRAGE (2)

```

```

/
NOMBRE DE PAS DE PLANIMETRAGE = 126
NOMBRE DE ZONES DE PLANIMETRAGE = 10
VALEUR DU PAS =
0.5;0.5;0.2;0.2;0.5;0.2;0.2;0.2;0.2;0.2
NUMERO DU PREMIER PROFIL =
1;24;123;332;427;477;484;492;500;539
NUMERO DU DERNIER PROFIL =
23;122;331;426;476;483;491;499;538;548
/
/MAILLAGE (2)
/
MODE DE SAISIE DU MAILLAGE = 2
SAUVEGARDE MAILLAGE = FAUX
/
/MAILLAGE PAR CLAVIER (3)
/
NOMBRE DE SECTIONS DE CALCUL = 0
NOMBRE DE PLAGES DE DISCRETISATION = 10
NUMERO DU PREMIER PROFIL DE LA SERIE =
1;24;123;332;427;477;484;492;500;539
NUMERO DU DERNIER PROFIL DE LA SERIE =
23;122;331;426;476;483;491;499;538;548
PAS D'ESPACE DE LA SERIE =
50.0;50.0;20.0;50.0;50.0;20.0;20.0;100.0;100.0;100.0
NOMBRE DE ZONES DE DISCRETISATION = 0
/
/SINGULARITE (BARRAGE-SEUIL) (1)
/
NOMBRE DE SEUILS = 3
/
/SEUIL 1 (2)
/
NOM SEUIL 1 = 'Seuil n°1'
TYPE SEUIL 1 = 4
NUM BRANCHE SEUIL 1 = 8
ABSCISSE SEUIL 1 = 10650.0
COTE CRETE SEUIL 1 = 4.24
COTE RUPTURE SEUIL 1 = 10000.0
COEFF DEBIT SEUIL 1 = 0.38
EPAISSEUR SEUIL 1 = 2
/
/SEUIL 2 (2)
/
NOM SEUIL 2 = 'Seuil n°2'
TYPE SEUIL 2 = 4
NUM BRANCHE SEUIL 2 = 9
ABSCISSE SEUIL 2 = 35700.0
COTE CRETE SEUIL 2 = 6.51
COTE RUPTURE SEUIL 2 = 10000.0
COEFF DEBIT SEUIL 2 = 0.38
EPAISSEUR SEUIL 2 = 1
/
/SEUIL 3 (2)
/
NOM SEUIL 3 = 'Seuil n°3'
TYPE SEUIL 3 = 4
NUM BRANCHE SEUIL 3 = 10
ABSCISSE SEUIL 3 = 22020.0
COTE CRETE SEUIL 3 = 8.81
COTE RUPTURE SEUIL 3 = 10000.0
COEFF DEBIT SEUIL 3 = 0.38
EPAISSEUR SEUIL 3 = 1
/
/APPORTS ET DEVERSOIRS (1)
/
NOMBRE DE DEBITS D'APPORTS = 1
NOM DE L'APPORT = 'Debit apport-Singularite n.1' → Apport linéique supplémentaire
NUMERO BRANCHE APPORT = 5
ABSCISSE APPORT = 300
LONGUEUR APPORT = 1000
NUMERO LOI APPORT = 6
/
/CALAGE (1)

```

```

/
/
/FROTTEMENT (2)
/
LOI DE FROTTEMENT = 1
NOMBRE DE ZONES DE FROTTEMENT = 21
NUMERO DE BRANCHE DE ZONE DE FROTTEMENT =
1;2;3;3;3;3;3;3;3;3;3;3;3;3;4;5;6;7;8;9;10
ABSCISSE DEBUT ZONE DE FROTTEMENT =
0.0;0.0;0.0;10652.0;10741.0;11278.0;11327.0;11512.0;11532.0;11666.0;1167
6.0;11812.0;11855.0;0.0;0.0;0.0;0.0;250.0;0.0;0.0;0.0
ABSCISSE FIN ZONE DE FROTTEMENT =
5730.0;25190.0;10652.0;10741.0;11278.0;11327.0;11512.0;11532.0;11666.0;1
1676.0;11812.0;11855.0;11989.0;34496.0;9431.0;2168.0;250.0;2287.0;10810.
0;35860.0;23020.0
VALEUR DU COEFFICIENT LIT MINEUR =
30.0;27.5;40.0;3.0;22.5;3.0;22.5;3.0;22.5;3.0;22.5;3.0;22.5;27.5;52.5;37
.5;20.0;37.5;15.0;30.0;30.0
VALEUR DU COEFFICIENT LIT MAJEUR =
30.0;27.5;40.0;3.0;22.5;3.0;22.5;3.0;22.5;3.0;22.5;3.0;22.5;30.0;40.0;37
.5;20.0;37.5;15.0;30.0;5.0
/
/ZONES DE STOCKAGE (2)
/
NOMBRE DE PROFILS COMPORTANT DES ZONES DE STOCKAGE = 0
/
/LOIS HYDRAULIQUES (1)
/
NOMBRE DE LOIS HYDRAULIQUES = 6 → une lois en plus pour l'apport linéique
/
/LOI NUMERO 1 (2)
/
LOI 1 NOM = 'bief8_amont'
LOI 1 TYPE = 1
/
/DONNEES LOI 1 (3)
/
LOI 1 MODE D'ENTREE = 1
LOI 1 FICHER = '../DonneesDyn/1023.loi'
/
/LOI NUMERO 2 (2)
/
LOI 2 NOM = 'bief4_amont'
LOI 2 TYPE = 1
/
/DONNEES LOI 2 (3)
/
LOI 2 MODE D'ENTREE = 1
LOI 2 FICHER = '../DonneesDyn/1029.loi'
/
/LOI NUMERO 3 (2)
/
LOI 3 NOM = 'bief9_amont'
LOI 3 TYPE = 1
/
/DONNEES LOI 3 (3)
/
LOI 3 MODE D'ENTREE = 1
LOI 3 FICHER = '../DonneesDyn/1049.loi'
/
/LOI NUMERO 4 (2)
/
LOI 4 NOM = 'bief10_amont'
LOI 4 TYPE = 1
/
/DONNEES LOI 4 (3)
/
LOI 4 MODE D'ENTREE = 1
LOI 4 FICHER = '../DonneesDyn/1030.loi'
/
/LOI NUMERO 5 (2)
/
LOI 5 NOM = 'bief1_aval'
LOI 5 TYPE = 2

```

```

/
/DONNEES LOI 5 (3)
/
LOI 5 MODE D'ENTREE = 1
LOI 5 FICHER = '../DonneesDyn/1467.loi'
/
/LOI NUMERO 6 (2)           → loi pour l'apport linéique supplémentaire
/
LOI 6 NOM = 'apport_bief5'
LOI 6 TYPE = 1
/
/DONNEES LOI 6 (3)
/
LOI 6 MODE D'ENTREE = 1
LOI 6 FICHER = '../DonneesDyn/Qappor5.loi'   → Initialement nulle
/
/CONDITIONS INITIALES (1)
/
/REPRISE ETUDE (2)
/
REPRISE DE CALCUL = FAUX
/
/LIGNE D'EAU (2)
/
PRESENCE LIGNE D'EAU INITIALE = VRAI
MODE D'ENTREE DE LA LIGNE D'EAU = 1
FICHER LIGNE D'EAU = '../DonneesStat/LigneEauInitiale_in.fig'
FORMAT LIGNE D'EAU = 2
/
/IMPRESSIONS - RESULTATS (1)
/
TITRE DU CALCUL = 'Chaine AdourMaritime Temps Reel'
/
/IMPRESSION (2)
/
IMPRESSION DE LA GEOMETRIE = FAUX
IMPRESSION DU PLANIMETRAGE = FAUX
IMPRESSION DU RESEAU = FAUX
IMPRESSION DES LOIS HYDRAULIQUES = FAUX
IMPRESSION DE LA LIGNE D'EAU INITIALE = FAUX
IMPRESSION CALCUL = FAUX
/
/PAS DE STOCKAGE ET D'IMPRESSION (2)
/
PREMIER PAS DE TEMPS A STOCKER = 1
PAS DE STOCKAGE = 12           → Sorties horaires (pas de temps = 300 secs)
PAS D'IMPRESSION = 144
/
/RESULTATS (2)
/
FICHER RESULTATS = '../Resultats/ResultatsOphyca.opt'
POST-PROCESSEUR = 2
/
/LISTING (2)
/
FICHER LISTING = '../Resultats/ResultatsListing.lis'
/
/FICHER DE REPRISE (2)
/
FICHER DE REPRISE EN ECRITURE = 'mascaret1.rep'
/
/RUBENS (2)
/
ECART ENTRE BRANCHES = 1.0
/
/STOCKAGE (2)
/
OPTION DE STOCKAGE = 2       → Stockage au format Ophyca aux stations d'observation
NOMBRE DE SITES = 3
BRANCHE DU SITE =
5;2;1
ABSCISSE DU SITE =
1361.0;10340.0;518.0

```

```

/
/VARIABLES CALCULEES (1)
/
VARIABLES CALCULEES =
FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FA
UX
/
/VARIABLES STOCKEES (1)
/
VARIABLES STOCKEES =
FAUX;FAUX;FAUX;FAUX;FAUX;VRAI;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FA
UX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX
;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX;FAUX
/

```

II.2 Exécution du run libre et assimilation linéique.

Un descriptif des étapes clés pour la préparation et l'exécution des simulations est donné en Annexe 1 pour le bassin de l'Adour et en Annexe 2 pour le bassin de la Marne. La partie II.2 donne des explications détaillées de ces étapes.

Compilation de l'exécutable MASCARET sans PALM

Les sources de la nouvelle version de Mascaret avec assimilation respectent le principe de la séparation IRF (Initialisation, Run, Finalisation) adopté au LNHE.

Pour faciliter la compatibilité avec la compilation de l'application piloté par PALM, l'arborescence des sources a été légèrement modifiée.

Le programme principal (appel.f90), le point d'entrée au système MASCARET (systeme_mascaret.f90), le module de partage des variables globales (partage.f90) et la procédure personnalisable d'arrêt du système (arret.f90) se trouvent dans le répertoire Super.

Le répertoire EXEC contiendra au final l'exécutable mascaret7 du code MASCARET. Au départ il contient le Makefile de plus haut niveau et les fichiers des dictionnaires (mots clés et position des fichiers .cas) qui serviront au moment de l'exécution.

Le Makefile du répertoire EXEC lance la compilation des programmes de haut niveau de Super et effectue l'édition de liens après avoir passé la main au Makefile qui se trouve dans le répertoire COMPLIB et qui s'occupe de la compilation de la totalité des sources MASCARET, de l'assemblage d'une librairie d'objets MASCARET et de la copie des fichiers de service issus de la compilation des modules dans le répertoire include.

Il suffit donc de se placer dans le répertoire EXEC et lancer la commande make

Les Makefile sont indépendants de la plateforme de compilation. Toutes les options dépendantes de la machine et du compilateur sont regroupées dans le fichier Make.commands.XXX qui se trouve dans les répertoire COMPLIB et qui est utilisé par les Makefile des répertoires COMPLIB et EXEC.

A titre d'exemple voyons le contenu du fichier COMPLIB/Make.commands.pgi pour un PC linux 64 bits avec environnement de compilation Portland Group.

```

FC=pgf90
LD=pgf90
#
SAVEFLAGS=-Msave
#
#LIBS
LIBS=-llapack -lblas
#
#OPTI
FFLAGS= -Mpreprocess -pc 64 -fastsse -O3
FFLAGSASSIM=$(FFLAGS)

```

LD_FLAGS=-pc 64 -fastsse -O3

Cet environnement marchera si la commande de compilation pgf90 fait partie du chemin d'accès aux exécutables (*path*)

```
NEW_MASCARET/SOURCES/COMPLIB:279>which pgf90  
/usr/local/pgi/linux86-64/2010/bin/pgf90
```

```
NEW_MASCARET/SOURCES/COMPLIB:285>pgf90 -V  
pgf90 10.2-0 64-bit target on x86-64 Linux -tp nehalem-64
```

La compilation s'effectue simplement dans le répertoire EXEC par la commande make.

Mise en place du cas test

Pour mettre en œuvre un cas test il faut préparer une arborescence de travail cohérente avec les fichiers descriptifs de l'application. Nous sommes restés cohérents avec l'arborescence opérationnelle du SCHAPI. Le choix étant purement arbitraire nous n'en discuterons pas les raisons ici. Certains noms de répertoires pourront pour cette raison apparaître inappropriés par la suite.

A chaque cas il faut attribuer un sous-répertoire de la racine MASCARET, au même niveau du répertoire SOURCES. Le nom est arbitraire, mais son niveau dans l'arborescence est fondamental, car toutes les références se font en relatif. A l'intérieur il faut créer les répertoires suivants : DonneesDyn, DonneesStat, Observations, Programmes et Resultats.

Le répertoire DonneesDyn contient les lois hydrauliques aux extrémités libres et pour les débits d'apport. De façon cohérente avec ce qui a été décrit dans le fichier .cas, dans notre cas il contiendra les fichiers 1023.loi, 1029.loi, 1030.loi, 1049.loi, 1467.loi Et Qappor5.loi.

Le répertoire DonneesStat contient la description de l'expérience sous la forme du fichier ParametresMascaret.cas, la description des paramètres d'assimilation, dont nous verrons plus tard le détail, ParametresMascaretAssim.cas, le fichier de description de la géométrie Geometrie.geo, et la ligne d'eau initiale dans le fichier LigneEauInitiale_in.lig.

Le répertoire Programmes contient les sous-répertoires Mascaret et PALM. Nous nous intéressons ici à Mascaret. Celui-ci contient des liens vers les fichiers qui se trouvent, après compilation dans SOURCES/EXEC et plus précisément :

```
dicoassim.txt -> ../../SOURCES/EXEC/dicoassim.txt  
dico.txt -> ../../SOURCES/EXEC/dico.txt  
FichierCasAssim.txt -> ../../SOURCES/EXEC/FichierCasAssim.txt  
FichierCas.txt -> ../../SOURCES/EXEC/FichierCas.txt  
systeme_mascaret -> ../../SOURCES/EXEC/mascaret7
```

Le répertoire Observations contient des fichiers ascii (extension .dat) avec deux colonnes pour les secondes et les hauteurs en mètres. Ils sont issus de fichiers excel au format csv par extraction et ajustement du zéro d'échelle (le répertoire contient aussi un petit script obs_split_AM.ksh basé sur awk pour le faire automatiquement).

Le répertoire Resultats est vide au départ sauf pour la présence de deux scripts gnuplot pour l'affichage plogp_ref.gp, pour l'affichage des résultats d'un run libre et superposition des valeurs observées, et plogp_ana.gp, pour l'affichage des résultats d'un run avec assimilation et superposition des valeurs observées et des résultats du run libre (qui doit avoir tourné au préalable et dont les résultats doivent avoir été stockés dans un sous-

répertoire appelé DRYRUN).

Execution du run libre (dryrun)

Le fichier .cas que nous avons commenté permet de simuler 3 jours et demi à partir du 20ème jour de l'évènement. Cet intervalle comprend le pic de crue. Pour lancer le modèle libre et obtenir les résultats de référence, il suffit de s'assurer que l'assimilation linéique n'est pas activée. Il faut contrôler qu'à la ligne 49 du fichier ParametresMascaretAssimilation.cas (qui se trouve dans le répertoire DonneesStat) la clé soit positionnée à *faux* :

```
/
/ASSIMILATION (3)
/
ACTIVATION DE L'ASSIMILATION = FAUX
/
```

N.B. Nous faisons ici et dans la suite référence à la numérotation des lignes des fichiers .cas de l'exemple utilisé tout le long de ce document, mais elle dépend, bien sûr, de chaque cas.

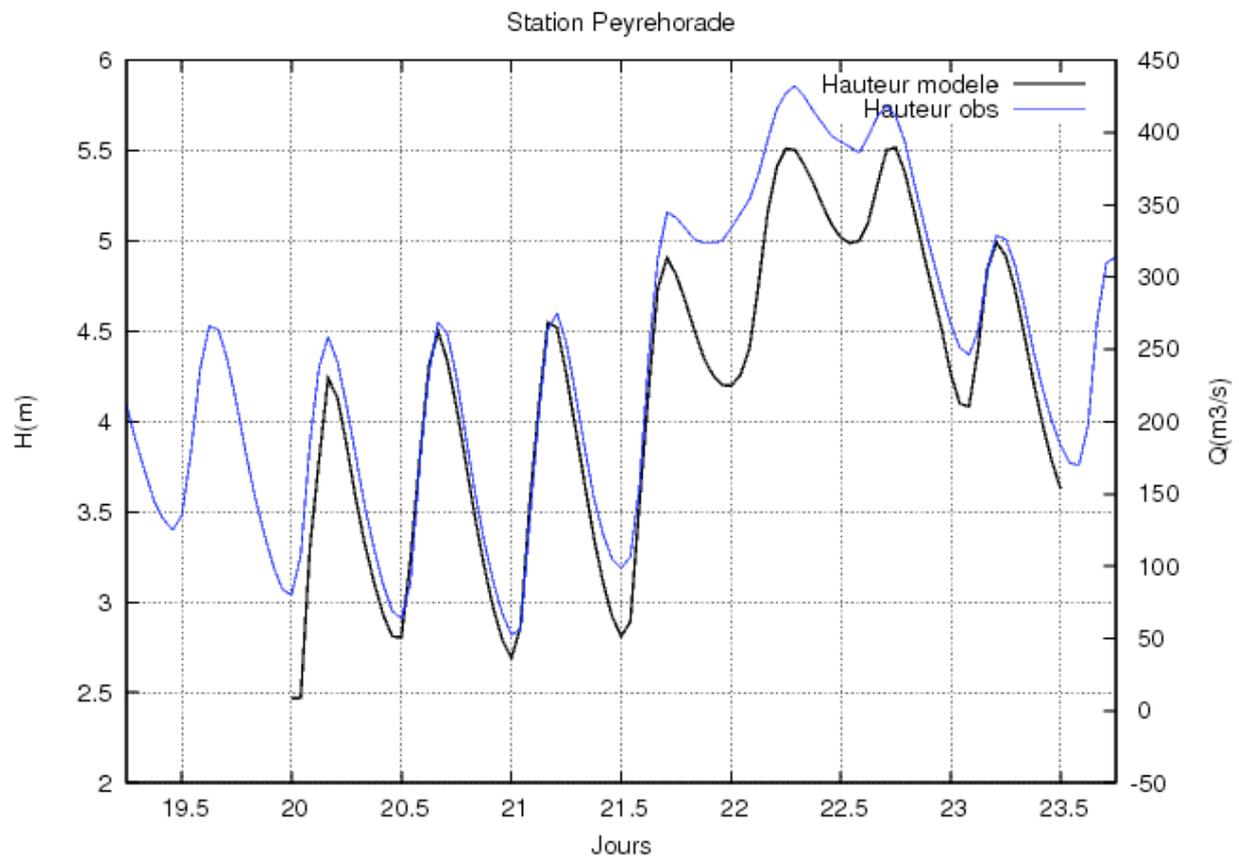
Pour lancer le run, dans le répertoire Programmes/Mascaret lancer la commande
./systeme_mascaret

Le run se termine en affichant OK.

Dans le répertoire Resultats on trouve les deux fichiers ResultatsListing.lis et ResultatsOphyca.opt. Puisque ces fichiers serviront comme référence pour la comparaison avec les résultats des expériences d'assimilation, nous allons les sauvegarder dans un répertoire qui doit nécessairement s'appeler DRYRUN. Il est conseillé de conserver aussi une copie des fichiers .cas qui ont été utilisés pour cette expérience, de façon à pouvoir la rejouer si nécessaire.

Pour afficher les résultats nous allons utiliser une procédure sous gnuplot (<http://www.gnuplot.info/>) qui va extraire les informations nécessaires du fichier Ophyca et générer un fichier au format postscript que l'on pourra ensuite visualiser avec ghostview sous linux ou GSview sous Windows (<http://pages.cs.wisc.edu/~ghost/>).

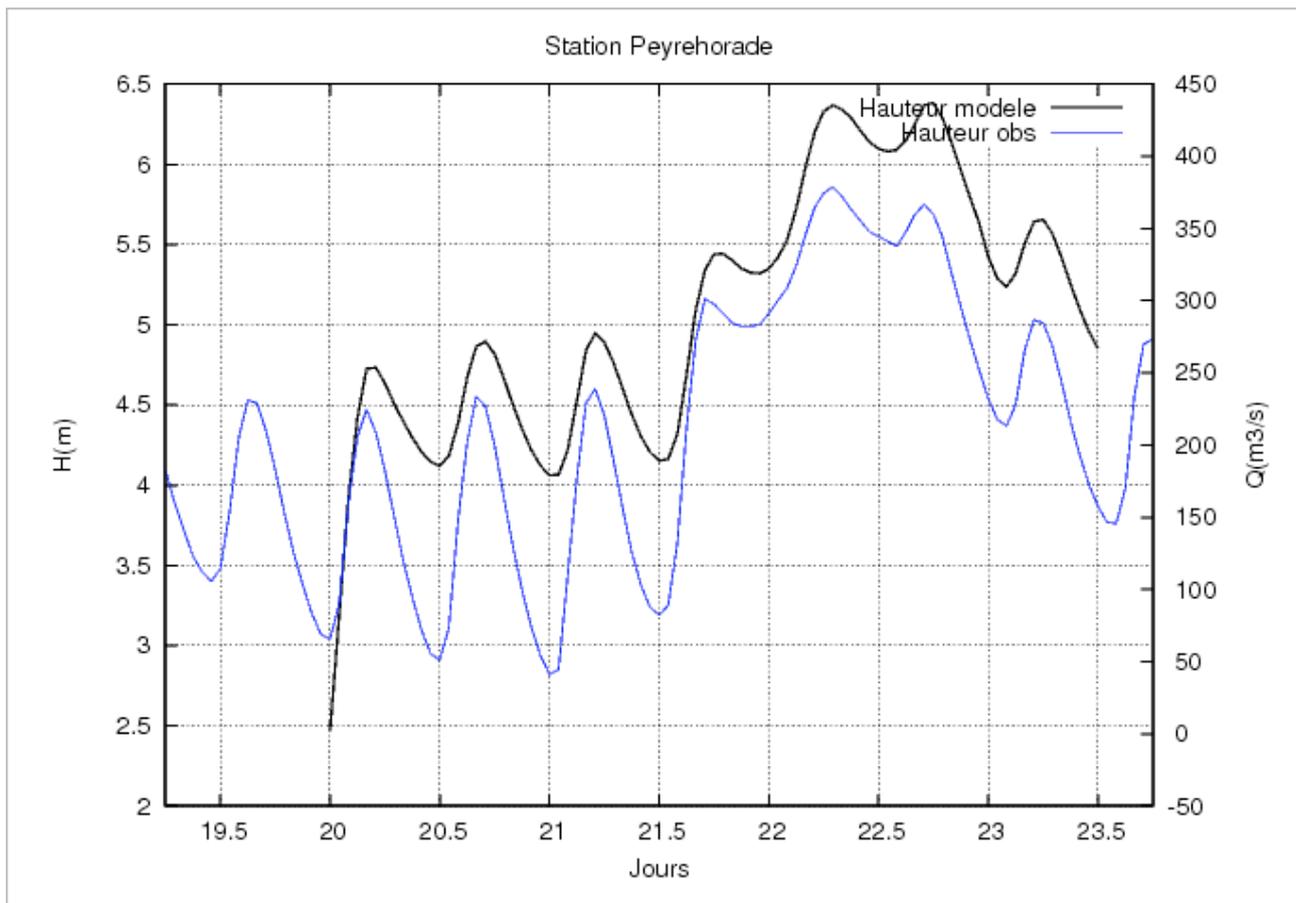
La procédure plogp_ref.gp doit être copiée (un lien suffit) dans DRYRUN et lancée avec gnuplot plogp_ref.gp. Elle génère le fichier plot_ps_2030400.ps qui sera visualisé avec gv plot_ps_2030400.ps



Run libre avec apport latéral arbitraire

Pour évaluer l'impact de l'apport supplémentaire nous allons relancer Mascaret après avoir modifié la loi `DonneesDyn/Qappor5.loi`. Nous allons indiquer une valeur constante en temps de $0.5 \text{ m}^3/\text{s}/\text{m}$ en modifiant la deuxième colonne des deux lignes.

En rejouant l'expérience de façon identique et en sauvegardant les résultats dans un autre sous-répertoire afin de ne pas écraser les résultats du run libre de référence, nous pouvons voir comment l'ajout d'un débit arbitraire en amont de Peyrehorade conduit à une surestimation de la hauteur d'eau



Expérience avec assimilation linéique

Nous allons procéder à une expérience au cours de laquelle nous allons utiliser les données d'observations de hauteur d'eau aux stations de Peyrehorade, Urt et Lesseps pour corriger périodiquement la ligne d'eau en Z et Q au cours de l'intégration.

Le code ne doit pas être modifié ni recompilé. Toute les informations nécessaires sont introduites dans le fichier ParametresMascaretAssimilation.cas.

Nous faisons référence à la présentation théorique de la méthode pour la compréhension des paramètres d'assimilation que nous allons d'abord énumérer et ensuite retrouver dans les parties du fichier .cas qui concernent l'assimilation linéique.

L'assimilation est activée (ligne 49)

```
/
ACTIVATION DE L'ASSIMILATION = VRAI
```

Puisque la ligne d'eau initiale est arbitraire, les résultats sur les premiers pas de temps ne sont pas significatifs et leur comparaison aux observations peut conduire à des écarts trop importants pour l'assimilation. Pour cette raison, la procédure d'assimilation n'est activée qu'après un jour simulé (mise à l'équilibre avec les forçages) et, dans notre cas, on la garde active jusqu'à la fin de la simulation (modalité sans prévision) (ligne 54)

```
/TEMPS DE DEBUT DE L'ASSIMILATION EN SEC PAR RAPPORT AU DEBUT DE LA SIMULATION
TEMPS DE DEBUT DE L'ASSIMILATION = 86400
```

```
/
/TEMPS DE FIN DE L'ASSIMILATION EN SEC PAR RAPPORT AU DEBUT DE LA SIMULATION
TEMPS DE FIN DE L'ASSIMILATION = 302400
```

Les observations sont regroupées aux heures pleines. Les pas de temps du modèle est de 5 minutes, mais la ligne d'eau ne sera corrigée que toutes les heures (ligne 60)

```
/FREQUENCE D'ASSIMILATION EN SEC  
FREQUENCE D'ASSIMILATION = 3600.0
```

La localisation des stations de mesure reprend la syntaxe des points de stockage. Elles sont aussi caractérisées par leur nom et par le décalage, si présent de l'origine de l'observation par rapport à la côte marine (ligne 72)

```
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'  
/DECALAGE DE L'ORIGINE DES OBSERVATIONS PAR RAPPORT A LA COTE MARINE EN METRES  
DECALAGE DE LA STATION =  
0.;0.;0.
```

Pour éviter que l'assimilation n'ait à traiter des écarts modèle-observations trop importants pour être réellement utiles ou significatifs, certaines observations sont rejetées. Il y a un premier critère de représentativité : puisque le modèle est calé pour les régimes moyens et forts, les observations sont rejetées si le niveau est trop bas (ligne 83)

```
/SEUIL DE REJET DES OBS POUR ASSIMILATION ET CALAGE PAR STATION (min en mt)  
SEUIL REJET OBSERVATIONS =  
1.;1.;1.
```

Le deuxième critère évite directement que l'on assimile des observations qui s'écartent excessivement des valeurs modélisées. L'unité de mesure de l'écart est l'écart type d'erreur d'ébauche. Mettre une très haute valeur revient à ne pas filtrer d'observations (ligne 87)

```
/SEUIL REJET MISFIT ASSIM EN NOMBRE D'ECARTS TYPE  
SEUIL REJET MISFIT ASSIM EN NOMBRE D'ECARTS TYPE = 20000.0
```

Pour la lecture des observations dans les fichiers, il faut en indiquer le chemin. On fait l'hypothèse que tous les fichiers aient le même nom indexé par le numéro de la maille qui contient la station (ligne 90)

```
/PREFIXE : LES FICHIERS SONT "$PREFIXE$NB_STAT.dat"  
PREFIXE FICHIERS OBSERVATIONS = '././Observations/Observations_stat'
```

La fréquence de stockage des observations ne coïncide pas nécessairement avec la fréquence d'assimilation. Nous indiquons la fréquences des observations et, si celle-ci est différente de la fréquence d'assimilation, nous pouvons indiquer aussi la taille d'une fenêtre glissante pour le lissage des oscillations à haute fréquence (ligne 93)

```
/FREQUENCE DES DONNEES DANS LE FICHIER D'OBSERVATIONS  
FREQUENCE DES DONNEES DANS LE FICHIER D'OBSERVATIONS = 3600.  
/  
/TAILLE DE LA FENETRE POUR LA MOYENNE GLISSANTE DU LISSAGE DES OBS  
TAILLE DE LA FENETRE POUR LE LISSAGE DES OBS = 3600.
```

Dans notre approche d'assimilation la matrice **R** est diagonale. Elle est donc entièrement déterminée par l'écart type de l'erreur d'observation aux stations (ligne 99)

```
/ECART TYPE ERREUR D'OBSERVATION EN METRES  
ECART TYPE ERREUR D'OBSERVATION A LA STATION =  
0.1;0.1;0.1
```

Nous avons vu dans la Partie I, que pour caractériser la matrice **B** nous avons besoin des

écarts type d'erreur d'ébauche sur la hauteur en correspondance des stations de mesure, des longueurs de corrélation (aussi dites de portée) en aval et en amont des stations d'observation pour les variables Z et Q et des coefficients (a,r,b) des courbes de tarage aux points d'observation qui sont de la forme

$$Q = a Z^r + b$$

/ECART TYPE ERREUR D'EBAUCHE AUX POINTS D'OBSERVATION EN METRES

ECART TYPE ERREUR D'EBAUCHE A LA STATION =

0.3;0.25;0.25

/

/LONGUEUR DE PORTEE DE L'ERREUR D'EBAUCHE POUR Z EN METRES (AMONT)

LONGUEUR DE PORTEE AMONT POUR Z =

20204.34;6613.27;34180.65

/

/LONGUEUR DE PORTEE DE L'ERREUR D'EBAUCHE POUR Z EN METRES (AVAL)

LONGUEUR DE PORTEE AVAL POUR Z =

2020.43;661.33;3418.06

/

/LONGUEUR DE PORTEE DE L'ERREUR D'EBAUCHE POUR Q EN METRES (AMONT)

LONGUEUR DE PORTEE AMONT POUR Q =

20204.34;6613.27;34180.65

/

/LONGUEUR DE PORTEE DE L'ERREUR D'EBAUCHE POUR Q EN METRES (AVAL)

LONGUEUR DE PORTEE AVAL POUR Q =

2020.43;661.33;3418.06

/

/COEFF A COURBE DE TARAGE $Q=A*Z^R+B$

COEFF A COURBE DE TARAGE =

1.;1.;1.

/

/COEFF R COURBE DE TARAGE $Q=A*Z^R+B$

COEFF R COURBE DE TARAGE =

1.;1.;1.

/

/COEFF B COURBE DE TARAGE $Q=A*Z^R+B$

COEFF B COURBE DE TARAGE =

0.;0.;0.

Pour lancer le run, dans le répertoire Programmes/Mascaret,lancer exactement comme avant, la commande

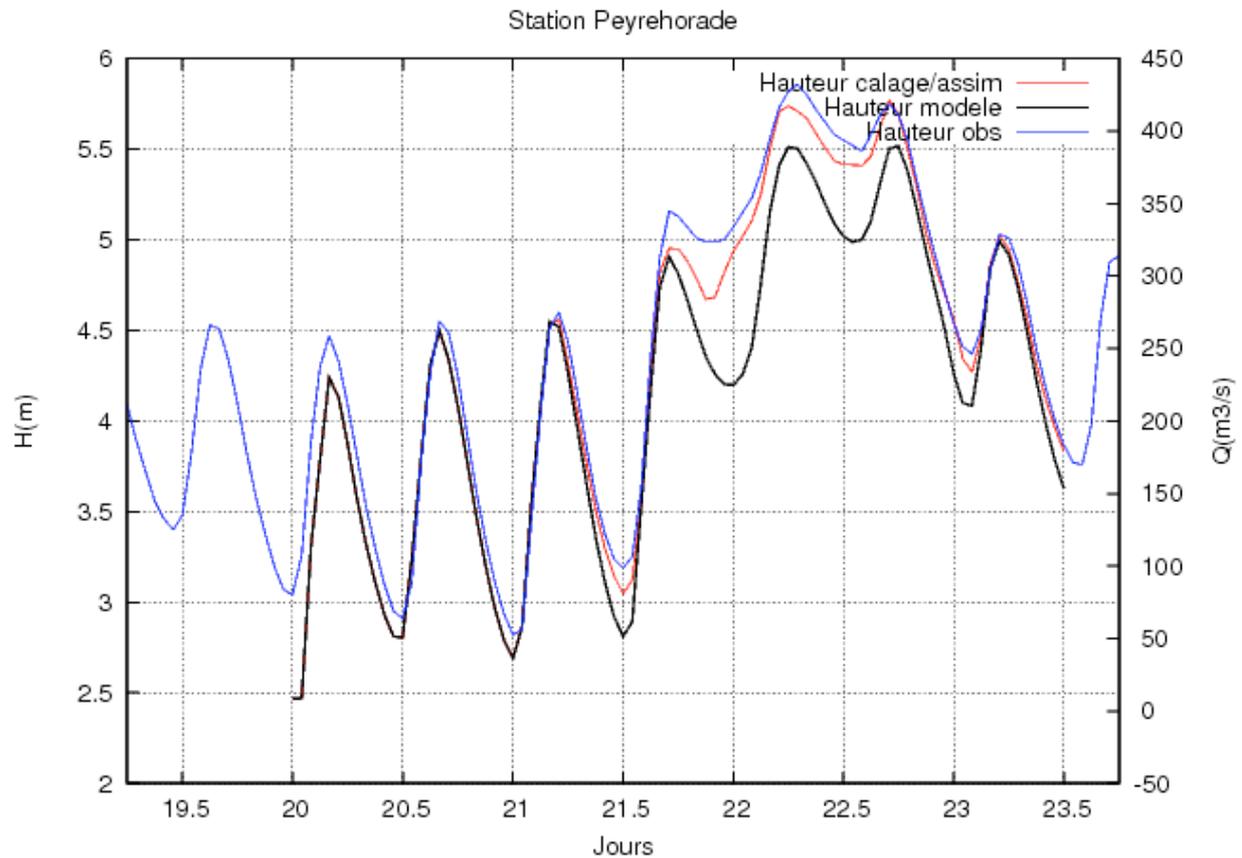
```
./systeme_mascaret
```

Le run affiche le nombre d'observations chargées dans les différents fichiers et se termine en affichant OK.

Dans le répertoire Resultats on trouve encore les deux fichiers ResultatsListing.lis et ResultatsOphyca.opt. Nous allons les sauvegarder dans un répertoire que l'on peut appeler ASSIM. Rappelons qu'il est conseillé de conserver aussi une copie des fichiers .cas qui ont été utilisés pour cette expérience, de façon à pouvoir la rejouer si nécessaire.

Pour afficher les résultats nous allons utiliser une autre procédure gnuplot qui va encore générer un fichier au format postscript.

La procédure plogp_ana.gp doit être copiée (un lien suffit) dans ASSIM et lancée avec gnuplot plogp_ana.gp. Elle génère le fichier plot_ps_ana_2030400.ps qui sera visualisé avec gv. On y voit en rouge la courbe simulée avec assimilation, superposée aux observations et à la courbe de référence du run libre.



II. 3 Calage des lois par assimilation

Cette approche d'assimilation prévoit un enchaînement de tâches complètement différent. Pour déterminer un nouveau jeu de lois, puisque on ne dispose pas de calcul explicite des jacobiniennes par rapport aux paramètres que l'on contrôle, il faut procéder par différences finies. Il faut donc effectuer un run libre de référence, qui sert aussi pour le calcul de l'écart aux observations, et un run perturbé pour chaque paramètre que l'on contrôle. Après une série de calculs algébriques d'assimilation, on dispose du jeu optimal de paramètres que l'on utilise pour un dernier run dit d'analyse.

Tous ces runs et les procédures d'algèbre sont pilotés par le coupleur PALM qui s'occupe aussi d'acheminer les données échangées. La gestion dynamique des tâches par PALM permet de compiler et installer l'application une seule fois et de l'utiliser pour les différentes expériences en modifiant tout simplement les fichier .cas

Compilation et mise en place du cas test

On considère que le coupleur PALM et son interface graphique PrePALM sont déjà installés sur la machine et que l'utilisateur qui s'occupe de la compilation de l'application a déjà été familiarisé avec l'utilisation du coupleur. N.B. l'utilisateur final peut utiliser l'application compilée en modalité *black box*.

L'organisation des fichiers et la modalité de compilation ressemblent à celles décrites dans la première partie du document. Cette fois la librairie MASCARET sera créée dans le répertoire COMPPALMLIB et les modules seront sauvegardés dans includepalm.

Dans le répertoire COMPPALMLIB, de façon tout à fait analogue à COMPLIB, il faut éditer un fichier de commandes de compilation. Il ne faut pas oublier, dans ce cas, d'activer la clé de preprocessing -DCALAGE_PALM

A titre d'exemple voyons le contenu du fichier COMPPALMLIB/Make.commands.pgi pour un PC linux 64 bits avec environnement de compilation Portland Group.

```
FC=pgf90
LD=pgf90
#
SAVEFLAGS=-Msave
#
PALMHOME = /home/palm/INSTALL/PALM_MP/linux64r4lam
INCPALM= -I$(PALMHOME)/include
#
#LIBS
LIBS= -llapack -lblas
#
#OPTI
FFLAGS=-Mpreprocess -pc 64 -fastsse -O3 -DCALAGE_PALM
FFLAGSASSIM=$(INCPALM) -Mpreprocess -pc 64 -fastsse -O3 -DCALAGE_PALM
LDFLAGS=-pc 64 -fastsse -O3
```

Nous avons mis en évidence les différences par rapport au fichier analogue pour la compilation sans PALM.

Le point de départ pour la compilation est le répertoire Assimilation/PALM. La première étape est la génération des fichiers de service avec PrePALM. On charge le fichier mascaret_calage.ppl et on génère tous les fichiers de service (sauf Make.include et Makefile).

Pour la compilation, on n'utilise pas le Makefile produit par PrePALM, mais le Makefile.AP qui est fourni, après avoir adapté le fichier Make.include à l'environnement de calcul. On lance enfin la compilation avec la commande make -f Makefile.AP.

A titre d'exemple voyons le contenu du fichier Assimilation/PALM/Make.include pour un PC linux 64 bits avec environnement de compilation Portland Group et bibliothèque de message passing LAM 7.1.4.

N.B. La bibliothèque MPI et les compilateurs doivent nécessairement être les mêmes qui ont été utilisés pour compiler la version du coupleur PALM indiquée dans le fichier de commandes dans COMPPALMLIB.

```
PALMHOME = /home/palm/INSTALL/PALM_MP/linux64r4lam
#
F90 = mpif77
F90FLAGS = -pc 64
F90PREPROC = -Mpreprocess
LF90FLAGS = -pc 64
F90EXTLIB = -L./../SOURCES/lib
#
F77 = mpif77
F77FLAGS = -pc 64 -fixedform
LF77FLAGS = -pc 64
F77EXLIB =
#
FPPFLAGS = -cpp
#
CC = mpicc
CCFLAGS = -pc 64
```

```

LCCFLAGS = -pc 64
CCEXTLIB =
#
C++ = mpiCC
C++FLAGS = -pc 64
LC++FLAGS = -pc 64
C++EXTLIB = -lpgf90 -lpgf90_rpm1 -lpgf902 -lpgf90rtl -lpgftrntl
#
OMPFLAGS = -mp
#
INCLUDES = -I/usr/local/include -I../..../SOURCES/includepalm
LIBS = -llapack -lblas -lnetcdf -L/home/palm/INSTALL/ALGEBRA/SCRIPPALM/linux64r4 -lscip
#
USERINCF =

```

En fonction de l'environnement de calcul il pourrait être nécessaire d'indiquer explicitement la position des bibliothèques de calcul et d'I/O lapack, blas et netcdf.

Il faut s'assurer que les *wrappers* des compilateurs (commandes d'enrobage fournies avec les bibliothèques MPI) soient cohérentes avec la version de MPI utilisée pour PALM et la version du compilateur utilisé pour les sources MASCARET.

```

NEW_MASCARET/SOURCES/Assimilation/PALM:282>which mpif77
/usr/local/pgi10/lam7.1.4_shared/bin/mpif77

```

```

NEW_MASCARET/SOURCES/Assimilation/PALM:283>mpif77 -V
pgf90 10.2-0 64-bit target on x86-64 Linux -tp nehalem-64

```

Pour installer l'application dans le répertoire de l'expérience il faudrait copier (ou du moins établir des liens symboliques avec) un grand nombre de fichiers. Pour cette raison dans le sous-répertoire Programmes/PALM du répertoire de l'expérience on trouve le script `inst_links` qui établit automatiquement tous les liens symboliques nécessaires.

Pour l'exécution de l'application il est conseillé d'utiliser le shell *bash* ou, à défaut, *ksh*, car la commande *nice* a un comportement différent dans un environnement basé sur *csh*.

Se placer dans le répertoire Programmes/PALM et s'assurer de ne pas avoir gardé des fichiers de sortie d'expériences précédentes avec la commande `maketidy`.

Si on utilise un environnement MPI LAM, lancer le daemon LAM par `lamboot`

Lancer la simulation sous PALM avec en dernier argument le nombre de processeurs disponibles sur la machine

```
nice -n 19 ./palm_main -nbprocmax 8
```

A la fin de l'exécution, arrêter le daemon LAM par `lamhalt`

Les fichiers de sortie de PALM et des statistiques d'assimilation se trouveront dans Programmes/PALM, tandis que les résultats du run MASCARET d'analyse se trouveront, comme d'habitude dans Resultats et de là on pourra les archiver dans un sous-répertoire et les afficher avec la procédure `plogp_ana.gp`, exactement comme pour l'expérience d'assimilation linéique.

Explication de la partie .cas qui concerne le calage

Les expériences de calage par assimilation sont, tout comme l'assimilation linéique, entièrement décrites dans les fichiers ParametresMascaretAssimilation.cas. Nous allons en détailler le contenu spécifique pour le calage et nous verrons ensuite les valeurs à attribuer aux différentes clés pour réaliser une expérience où l'on cale seulement le débit d'apport supplémentaire ou une expérience où l'on cale aussi les débits amont.

La première clé que l'on trouve dans le fichier .cas (ligne 8) active l'impression dans le fichier de listing des valeurs numériques des lois après perturbation ou correction

```
IMPRESSION DES LOIS HYDRAULIQUES = FAUX
```

Ensuite, pour chaque loi hydraulique décrite dans ParametresMascaret.cas, on donne des informations complémentaires. Pour diminuer la taille du contrôle d'assimilation, chaque loi hydraulique est représentée sous la forme $a * f(t-c) + b$, où $f(t)$ est la loi hydraulique lue dans le fichier d'entrée. Le contrôle se réduit donc à un maximum de 3 paramètres par loi et l'ébauche correspond à $a=1$, $b=0$, $c=0$. Pour chaque loi, on indique quels paramètres (entre a , b et c) on veut contrôler et quelles stations d'observation on utilise pour les caler. On peut enfin imposer, pour éviter des instabilités numériques, un seuil min et un seuil max pour les valeurs des lois hydrauliques.

Pour la n -ème loi, on aura donc un bloc de la forme (ligne 10), où n est à remplacer par le numéro de la loi

```
/DONNEES LOI n (3)
/
LOI n CALAGE = 'abc'           → Quels paramètres on contrôle
LOI n STATIONS CALAGE = VRAI;VRAI;VRAI → Quelles stations contribuent
LOI n VALEUR MIN = 4.         → Seuil min
LOI n VALEUR MAX = 660.      → Seuil max
```

La ligne 52 permet d'activer le tracé automatique gnuplot et la visualisation des courbes avec ghostview pour le run non perturbé de référence et pour le run d'analyse. Afin que cette procédure marche de façon automatique il est nécessaire qu'une copie des fichiers des observations soit présente dans le répertoire Results.

```
/RENDU GRAPHIQUE DES RESULTATS
TRACER RESULTATS = FAUX
```

Ensuite on retrouve des parties communes ou tout à fait semblables aux clés de l'assimilation linéique. En particulier les limites de l'intervalle de prise en compte des observations pour le calage et la fréquence de comparaison aux valeurs modélisées (ligne 64)

```
/TEMPS DE DEBUT DU CALAGE DES DEBITS D'APPORT PAR ASSIMILATION
TEMPS DE DEBUT DU CALAGE PAR ASSIMILATION = 86400
/
/TEMPS DE FIN DU CALAGE DES DEBITS D'APPORT PAR ASSIMILATION
TEMPS DE FIN DU CALAGE PAR ASSIMILATION = 302400
/
/FREQUENCE DES OBS POUR LE CALAGE PAR ASSIMILATION
FREQUENCE DES OBS POUR LE CALAGE = 3600.0
```

La localisation des stations et le seuil min des observations prises en compte sont en commun avec l'assimilation linéique (ligne 72)

```
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'
/DECALAGE DE L'ORIGINE DES OBSERVATIONS PAR RAPPORT A LA COTE MARINE EN METRES
DECALAGE DE LA STATION =
0.;0.;0.
/
/SEUIL DE REJET DES OBS POUR ASSIMILATION ET CALAGE PAR STATION (min en mt)
SEUIL REJET OBSERVATIONS =
1.;1.;1.

```

de même que l'indication de l'emplacement des fichiers d'observations, de la fréquence de stockage, de la modalité éventuelle de lissage et de l'écart type d'erreur d'observation (ligne 91)

```

/PREFIXE : LES FICHIERS SONT "$PREFIXE$NB_STAT.dat"
PREFIXE FICHIERS OBSERVATIONS = '..../Observations/Observations_stat'
/
/FREQUENCE DES DONNEES DANS LE FICHIER D'OBSERVATIONS
FREQUENCE DES DONNEES DANS LE FICHIER D'OBSERVATIONS = 3600.
/
/TAILLE DE LA FENETRE POUR LA MOYENNE GLISSANTE DU LISSAGE DES OBS
TAILLE DE LA FENETRE POUR LE LISSAGE DES OBS = 3600.
/
/ECART TYPE ERREUR D'OBSERVATION EN METRES
ECART TYPE ERREUR D'OBSERVATION A LA STATION =
0.1;0.1;0.1

```

A noter, par contre, que comme le contrôle d'assimilation n'est plus directement l'état du modèle, nous ne pouvons pas utiliser un multiple de l'écart type sur le contrôle comme deuxième critère de rejet des observations. On indique plus loin (ligne 179) un critère basé sur la comparaison directe de l'écart obs-modèle à la valeur modélisée. Le seuil est indiqué en pourcentage

```

/SEUIL DE REJET DES OBS POUR LE CALAGE EN % DE LA VALEUR MODELE
SEUIL REJET MISFIT CALAGE = 50.

```

Pour la construction par différences finies de la matrice Jacobienne de sensibilité aux paramètres (a,b,c) contrôlés, il faut indiquer l'amplitude des perturbations. On ne le fera pas loi par loi, mais par type de loi : loi amont débit (ou débit latéral ponctuel), loi d'apport linéique, loi amont cote (ligne 136)

```

/PERTURBATION DU PARAMETRE A POUR LOI DEBIT (adim)
PERTURBATION PARAMETRE A DEBIT = 1.1
/PERTURBATION DU PARAMETRE B POUR LOI DEBIT (m3/sec)
PERTURBATION PARAMETRE B DEBIT = 6
/PERTURBATION DU PARAMETRE C POUR LOI DEBIT (sec)
PERTURBATION PARAMETRE C DEBIT = 7200
/
/PERTURBATION DU PARAMETRE A POUR LOI DEBIT LINEIQUE (adim)
PERTURBATION PARAMETRE A DEBIT LINEIQUE = 1.05
/PERTURBATION DU PARAMETRE B POUR LOI DEBIT LINEIQUE (m3/sec/m)
PERTURBATION PARAMETRE B DEBIT LINEIQUE = 0.0005
/PERTURBATION DU PARAMETRE C POUR LOI DEBIT LINEIQUE (sec)
PERTURBATION PARAMETRE C DEBIT LINEIQUE = 0
/
/PERTURBATION DU PARAMETRE A POUR LOI COTE (adim)
PERTURBATION PARAMETRE A COTE = 1.1
/PERTURBATION DU PARAMETRE B POUR LOI COTE (m)
PERTURBATION PARAMETRE B COTE = 0.5
/PERTURBATION DU PARAMETRE C POUR LOI COTE (sec)
PERTURBATION PARAMETRE C COTE = 7200

```

Avec le même type de regroupement on indique aussi les écarts type d'erreur d'ébauche (ligne 157)

```

/ECART TYPE PARAMETRE A POUR LOI DEBIT (adim)
ECART TYPE PARAMETRE A DEBIT = 0.2
/ECART TYPE PARAMETRE B POUR LOI DEBIT (m3/sec)
ECART TYPE PARAMETRE B DEBIT = 6

```

```

/ECART TYPE PARAMETRE C POUR LOI DEBIT (sec)
ECART TYPE PARAMETRE C DEBIT = 7200
/
/ECART TYPE PARAMETRE A POUR LOI DEBIT LINEIQUE (adim)
ECART TYPE PARAMETRE A DEBIT LINEIQUE = 1.11
/ECART TYPE PARAMETRE B POUR LOI DEBIT LINEIQUE (m3/sec/m)
ECART TYPE PARAMETRE B DEBIT LINEIQUE = 0.01
/ECART TYPE PARAMETRE C POUR LOI DEBIT LINEIQUE (sec)
ECART TYPE PARAMETRE C DEBIT LINEIQUE = 0
/
/ECART TYPE PARAMETRE A POUR LOI COTE (adim)
ECART TYPE PARAMETRE A COTE = 0.2
/ECART TYPE PARAMETRE B POUR LOI COTE (m)
ECART TYPE PARAMETRE B COTE = 0.5
/ECART TYPE PARAMETRE C POUR LOI COTE (sec)
ECART TYPE PARAMETRE C COTE = 7200

```

La dernière ligne qui concerne le calage, n'est pas destinée à un usage régulier. Elle sert pendant la phase de mise au point de la chaîne d'assimilation pour sortir, grâce à une méthode itérative, des indications statistiques sur la validité des estimations des erreurs d'ébauche et d'observation. Pour l'utilisation courant on ne demande qu'un seul passage d'assimilation (ligne 182)

```

/NOMBRE MAX ITERATIONS OPTIMISATION COVARIANCES CALAGE
NOMBRE MAX ITERATIONS OPTIMISATION COVARIANCES CALAGE = 1

```

Cas de calage de l'apport supplémentaire

Comme premier exemple d'application du calage par assimilation prenons le cas où l'on veut contrôler le débit linéique constant en temps du bief 5. La valeur d'ébauche est 0 (loi Qappor5.loi). Puisque le débit est constant en temps, il suffit d'estimer le paramètre b qui donnera la valeur du débit par mètre linéaire. N.B. si on considérait un apport supplémentaire qui varie dans le temps, une manière de corriger le point de localisation de cet apport serait de contrôler un décalage en temps de la loi, comme nous le faisons pour les lois amonts (le paramètre c).

Pour les autres lois on ne demandera pas de contrôle (lignes 12, 18, 24, 30, 37)

```

LOI 1 CALAGE = "
...
LOI 5 CALAGE = "

```

La loi qui représente le débit d'apport linéique est la 6 (on fait toujours référence au .cas de MASCARET), donc nous demandons le contrôle du paramètre b, avec utilisation des observations provenant de toutes les stations et avec la possibilité de calculer des valeurs négatives (ligne 43)

```

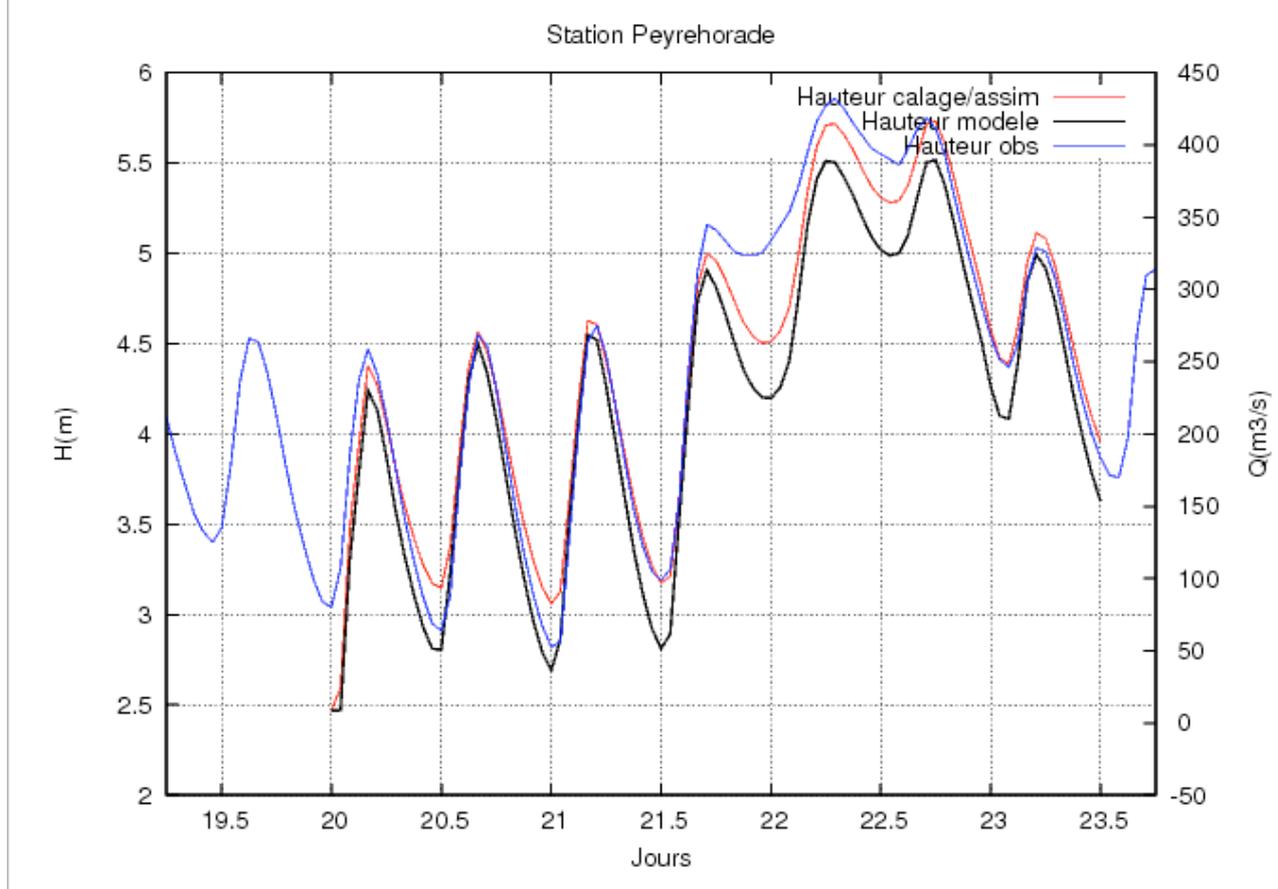
LOI 6 CALAGE = 'b'
LOI 6 STATIONS CALAGE = VRAI;VRAI;VRAI
LOI 6 VALEUR MIN = -100.

```

Pour toutes les autres valeurs (perturbations et écarts type on laisse les valeurs montrées dans la partie générale.

Après avoir exécuté la chaîne comme nous l'avons expliqué plus haut, dans le répertoire Resultats on trouve des fichiers de listing ResultatsListing_XXX.lis et deux fichiers de résultats au format Ophyca dont ResultatsOphyca.opt contient les résultats du run d'analyse. Nous allons tous les sauvegarder dans un répertoire que l'on peut appeler CALAGE_DEBIT. Rappelons qu'il est conseillé de conserver aussi une copie des fichiers .cas qui ont été utilisés pour cette expérience, de façon à pouvoir la rejouer si nécessaire.

Pour afficher les résultats nous allons utiliser la même procédure gnuplot que pour l'assimilation linéique. La procédure `plogp_ana.gp` doit être copiée (un lien suffit) dans `CALAGE_DEBIT` et lancée avec `gnuplot plogp_ana.gp`. Elle génère le fichier `plot_ps_ana_2030400.ps` qui sera visualisé avec `gv`. On y voit en rouge la courbe simulée avec assimilation, superposée aux observations et à la courbe de référence du run libre.



Pour connaître la valeur du débit latéral calculée par assimilation, il faut regarder le contenu du fichier `xa.txt` dans le répertoire d'exécution `Programmes/PALM`. On y trouve la valeur optimale de $0.129 \text{ m}^3/\text{m}$

Cas de calage des lois amont

Si on veut plutôt caler les lois des débits amont, sans toucher au débit latéral, on modifiera la première partie du fichier `.cas` de cette façon

```
/DONNEES LOI 1 (3)
/
LOI 1 CALAGE = 'abc'
LOI 1 STATIONS CALAGE = FAUX;FAUX;VRAI
LOI 1 VALEUR MIN = 6.
/
/DONNEES LOI 2 (3)
/
LOI 2 CALAGE = 'abc'
LOI 2 STATIONS CALAGE = VRAI;VRAI;VRAI
LOI 2 VALEUR MIN = 0.
/
/DONNEES LOI 3 (3)
```

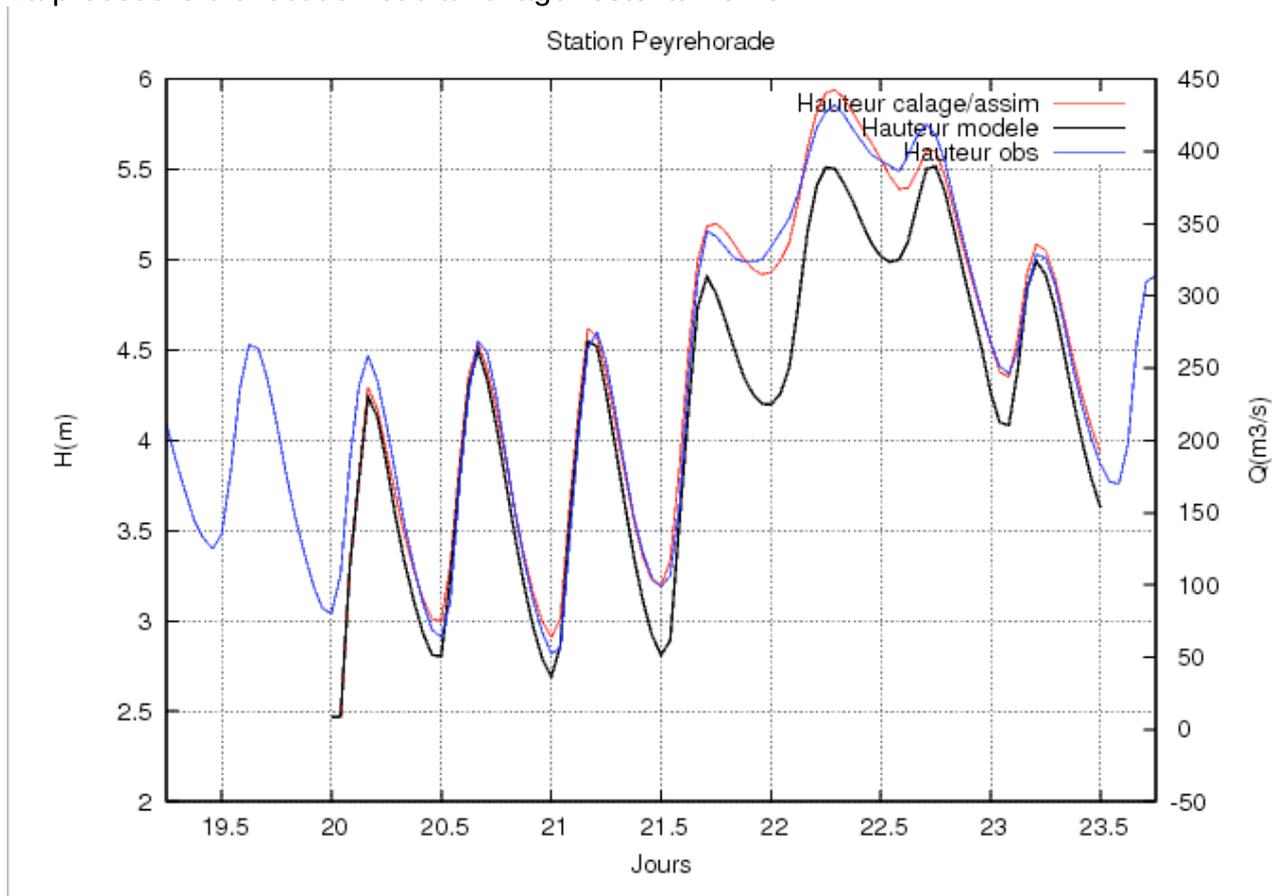
```

/
LOI 3 CALAGE = 'abc'
LOI 3 STATIONS CALAGE = VRAI;VRAI;VRAI
LOI 3 VALEUR MIN = 18.
/
/DONNEES LOI 4 (3)
/
LOI 4 CALAGE = 'abc'
LOI 4 STATIONS CALAGE = VRAI;VRAI;VRAI
LOI 4 VALEUR MIN = 4.
LOI 4 VALEUR MAX = 660.
/
/DONNEES LOI 5 (3)
/
LOI 5 CALAGE = ""
LOI 5 STATIONS CALAGE = FAUX;VRAI;VRAI
LOI 5 VALEUR MIN = 1.4
/
/DONNEES LOI 6 (3)
/
LOI 6 CALAGE = ""
LOI 6 STATIONS CALAGE = VRAI;VRAI;VRAI
LOI 6 VALEUR MIN = -100.
/

```

Il faut remarquer qu'après étude attentive du réseau, seulement une station d'observation (Lesseps, la troisième dans l'ordre des déclarations) contribue au calage des coefficients de la loi 1.

La procédure d'exécution et d'affichage reste la même.



Les coefficients optimaux sont dans le fichier xa.txt

a loi 1 0.9933860839305266

b loi 1 -0.0934008228071360
c loi 1 -325.6182973892540000
a loi 2 1.2276872262790380
b loi 2 0.9276446111528177
c loi 2 785.2746799174954000
a loi 3 1.3546149182880680
b loi 3 4.7182876112441590
c loi 3 17755.4813304947400000
a loi 4 1.1257117693669160
b loi 4 -1.4524870448816540
c loi 4 6110.3941953154790000

II.4 Combinaison avec l'assimilation linéique

Puisque l'assimilation linéique n'intervient qu'au cours du run d'analyse qui rejoue l'épisode avec les débits corrigés, la procédure de lancement est la même que pour le calage sous PALM.

La différence réside dans le .cas d'assimilation où l'on active la clé d'assimilation linéique. Par rapport à l'expérience illustrée précédemment, on utilise de plus petites variances d'erreur d'ébauche, puisque l'ébauche est calculée avec des débits corrigés.

Cas de calage apport supplémentaire + assimilation linéique

Dans le fichier .cas nous allons remettre la partie sur les lois comme pour le calage du seul débit latéral (sixième loi).

Nous allons par contre activer la clé d'assimilation linéique (ligne 49)

ACTIVATION DE L'ASSIMILATION = **VRAI**

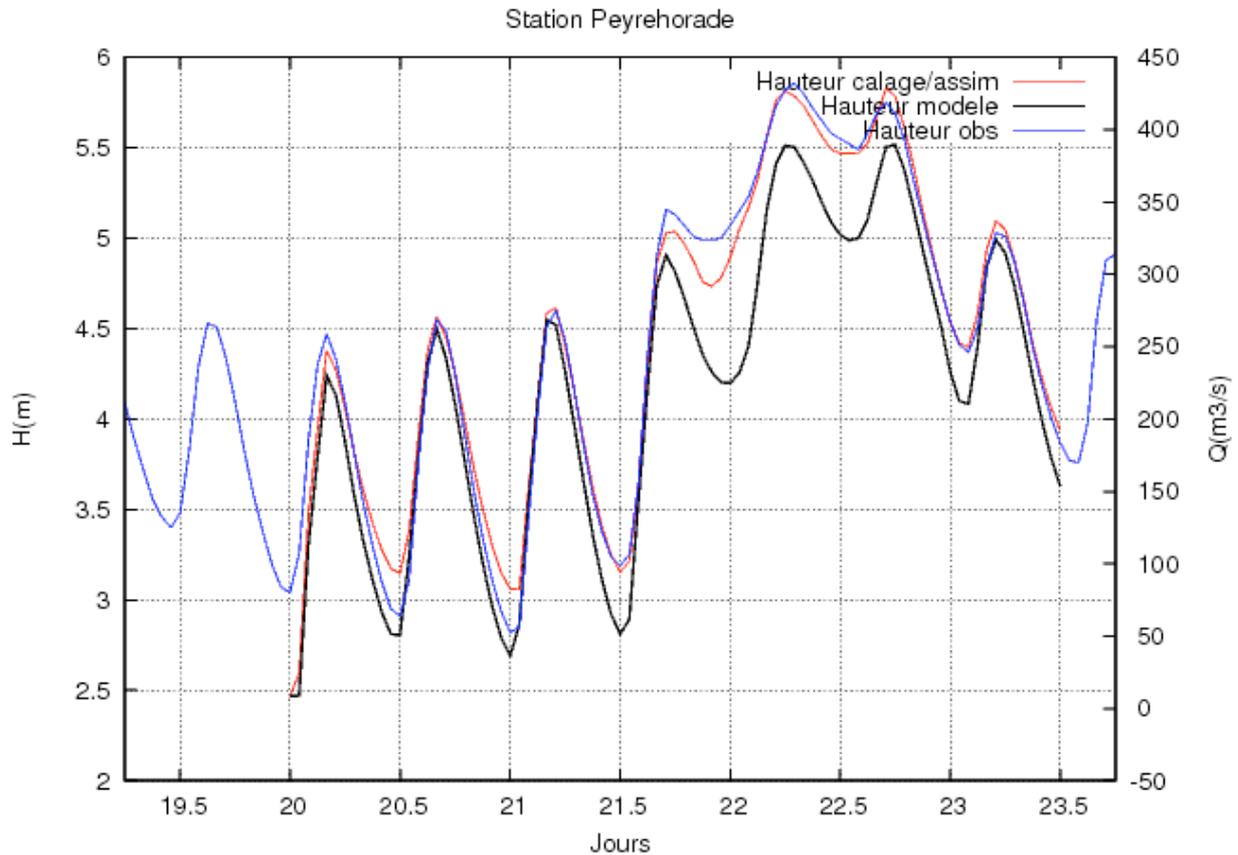
Cette fois les valeurs d'ébauche de l'assimilation linéique prennent déjà en compte la correction du débit latéral. Nous utilisons donc de plus petits écarts type d'erreur d'ébauche (ligne 104)

/ECART TYPE ERREUR D'EBAUCHE AUX POINTS D'OBSERVATION EN METRES

ECART TYPE ERREUR D'EBAUCHE A LA STATION =

0.15;0.12;0.12

L'exécution et l'affichage se font exactement comme pour une expérience de calage pilotée par PALM



Bouclage des expériences

Pour reproduire une situation proche de l'enchaînement de prévisions d'une chaîne temps réel, nous avons mis en place un système de relance automatique piloté par un script en korn shell.

Dans cette situation on imagine que chaque run de MASCARET prévoit une phase de mise à l'équilibre de l'état du modèle avec des débits amont observés prescrits et une phase de prévision pendant laquelle les débits amont restent constants en temps et égaux à la dernière valeur observée disponible. Si pendant la phase de mise à l'équilibre on dispose aussi d'observations de hauteur d'eau il est possible d'utiliser les procédures d'assimilation pour donner une meilleure condition initiale (assimilation linéique) et une meilleure estimation de la dernière valeur des débits amont (calage par assimilation) pour la prévision.

Nous sommes partis des épisodes de la base de calage, avec donc des lois hydrauliques connues sur tout l'évènement. Le script doit donc s'occuper pour une date de début de prévision donnée, de créer les fichiers d'entrée en coupant les lois au moment du début de la prévision et de remplacer les valeurs suivantes avec la dernière valeur observée.

On fait donc tourner l'épisode sur quelques jours avant la date de début de la prévision avec un temps initial d'équilibrage et ensuite avec calage des paramètres et/ou assimilation linéique jusqu'à la date de début de la prévision et on laisse ensuite continuer le modèle libre pendant quelques heures en mode prévision. On archive enfin les résultats.

A ce moment on décale le temps de début de prévision, typiquement d'une heure et on répète la procédure. Le script va donc aussi modifier les fichiers .cas (pour le modèle et pour l'assimilation) pour ajuster les repères en temps et les clés qui pourraient changer d'une exécution à l'autre.

En particulier, les dernières clé du fichier ParametresMascaretAssim.cas concernent la sauvegarde et la relecture d'une ligne d'eau complète de façon à pouvoir relancer le modèle à chaque fois d'une condition initiale cohérente avec les conditions aux limites au lieu de démarrer d'une ligne standard avec le risque d'introduire des instabilités non physiques (ligne 158)

```
/ACTIVATION DE LA SAUVEGARDE LIGNE D'EAU INITIALE
ACTIVATION DE LA SAUVEGARDE LIGNE D'EAU INITIALE = FAUX
/
/TEMPS DE SAUVEGARDE LIGNE D'EAU INITIALE (RELATIF AU DEBUT DE LA SIM. secs)
TEMPS DE SAUVEGARDE LIGNE D'EAU INITIALE = 302400
/
/CHEMIN SAUVEGARDE LIGNE D'EAU INITIALE
CHEMIN SAUVEGARDE LIGNE D'EAU INITIALE = '../DonneesStat'
```

Le script de pilotage n'a pas vocation à être un outil générique. Chaque vraie chaîne temps réel a son propre système de pilotage. Il s'agit simplement d'un outil de maquettage pour lequel la convivialité n'était pas le premier critère de conception. Sa syntaxe de lancement et quelques subtilités codées en dur dans le script méritent des explications.

En guise d'avant-propos, nous définissons quelques termes. *Temps de base* indique le temps de début de prévision. Le *spinup* est le temps de simulation avant le temps de base, donc avec débits amont connus et comprend le temps d'équilibrage et, éventuellement, le temps d'assimilation. Le *forecast* est la partie de prévision qui continue après le temps de base avec débits amont constants.

Le repérage en temps est fait par rapport au temps de base. Donc quand on demandera au pilote de balayer un évènement on lui indiquera le premier et le dernier temps de base et le décalage entre les temps de base de deux runs consécutifs. Remarquez donc, que le premier instant simulé sera (*Premier temps de base – spinup*) et le dernier sera (*Dernier temps de base + forecast*).

Le pilote utilise un mot clé pour identifier le type d'expérience. Ce même mot clé correspond aussi au nom du sous-répertoire de Resultats dans lequel seront stockés les résultats et les statistiques des critères d'évaluation des performances.

DRYRUN → Run libre
ASSIM → Run avec assimilation linéique
CALAGE → Run avec calage des lois hydrauliques
CALAGEASSIM → Run avec calage des lois hydrauliques et assimilation linéique

Si tous les runs de l'évènement on déjà été effectués et les résultats stockés dans le répertoire correspondant au type d'expérience et que l'on veut tout simplement recalculer les critères d'évaluation des performances, on indique un argument optionnel FAST pour indiquer au pilote de ne pas relancer le système mais de récupérer les résultats existants.

La syntaxe de la commande (telle que vous la verriez à l'écran si jamais le nombre d'arguments était incorrect) est

```
piloteAM.ksh first_time last_time frequency spinup_length forecast_length result_dir_prefix [FAST]
```

first_time est le premier temps de base traité

last_time	le dernier
frequency	est le décalage entre deux temps de base
spinup_lenght	est la durée du spinup avant le temps de base
forecast_lenght	est le durée de la prévision après le temps de base
result_dir_prefix	indique le type d'expérience [DRYRUN ASSIM CALAGE CALAGEASSIM]
FAST	est l'argument optionnel pour le recalcul des critères

Par exemple, pour lancer une séquence de runs avec temps de base qui se décale d'une heure à la fois du 10ème au 15ème jour d'un évènement avec 3 jours de spinup et 12 heures de prévision avec calage des lois hydrauliques mais sans assimilation linéique la commande sera

```
piloteAM.ksh 864000 1296000 3600 259200 43200 CALAGE
```

C'est le script qui va modifier les fichiers .cas pour chaque exécution. C'est donc dans le code du pilote lui même que l'on a codé en dur certains choix algorithmiques.

En particulier, en modalité DRYRUN ou ASSIM, on gère la politique de sauvegarde de la ligne d'eau pour le restart de façon différente par rapport aux cas CALAGE ou CALAGEASSIM. Dans le code on verra donc des tests pour attribuer les valeurs correctes aux variables `isave` et `tsave` qui servent pour modifier les lignes 158 et 161 de `ParametresMascaretAssim.cas` (Nous faisons ici et dans la suite référence à la numérotation des lignes des fichiers .cas de l'exemple utilisé tout le long de ce document, mais elle dépend, bien sûr, de chaque cas).

De façon analogue, c'est le pilote qui fixe le moment de début de l'assimilation ou du calage par assimilation. Pour l'Adour Maritime en général on effectue le calage de paramètres en prenant en compte les observations sur les deux jours précédant le temps de base. L'assimilation linéique se fait sur les 24 heures précédant le temps de base en cas de CALAGEASSIM. Dans le cas ASSIM elle se fait sur les 24 heures précédant le temps de base si le run a démarré d'un ligne d'eau standard ou sur la totalité du spinup si le run a redémarré d'une ligne d'eau calculée et sauvegardée. Dans le code on verra donc des tests pour attribuer les valeurs correctes aux variables `tiniassim` et `tinicalage` qui servent pour modifier les lignes 45 et 54 de `ParametresMascaretAssim.cas`.

Comme nous l'avons déjà souligné, on n'utilise pas les mêmes écarts type d'erreur d'ébauche pour l'assimilation linéique si celle ci est le seul traitement d'assimilation ou si elle suit le calage des lois hydrauliques. C'est donc dans le pilote que l'on attribue les bonnes valeurs à la variable `varstring` qui sert pour modifier la ligne 95 de `ParametresMascaretAssim.cas`.

Le pilote s'occupe aussi de lancer les exécutions. La syntaxe de lancement, dans le cas sous PALM, dépendra de l'environnement MPI installé sur votre machine. Il suffira d'adapter les lignes qui suivent le commentaire `Executing mascaret on $nbpr procs with initial time $tstart and final time $tstop`

Le pilote lance après chaque exécution une procédure d'extraction des données significatives des fichiers `Ophyca` et un traitement statistique pour le calcul des critères de performances.

Le calcul est effectué par le programme `Resultats/calcul_criteresAM` qui s'obtient par simple compilation F90 des sources `SOURCES/SCRIPTS/ADOUR/calcul_criteresAM.f90`

Dans sa version standard ce programme calcule 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'efficience 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.

Ces statistiques sont organisées dans des fichiers lus par des procédure gnuplot qui tracent les graphiques des statistiques par station d'observation. Le pilotage se conclut avec l'affichage des fichiers postscript des graphiques.

Dans les Annexes 1 et 2, vous trouverez des instructions spécifiques pour faire tourner l'enchaînement des simulations sur quelques cas tirés des rapports de calage des bassins Adour Maritime et Marne Vallage.

Les résultats de ces simulations sont regroupés et commentés dans l'Annexe 3 qui avait été rédigée à la fin de la validation de la maquette d'assimilation basée sur un jeu de sources MASCARET plus ancien que l'actuel. Cependant elle restent qualitativement comparables puisque les versions du modèle sont compatibles et les algorithmes d'assimilation n'ont pas changé.

Annexe 1 : Récapitulatif Adour

A- Compilation

```
Dans SOURCES/EXEC  
make clean  
make
```

```
Dans SOURCES/Assimilation/PALM  
make -f Makefile.AP allclean  
prepalm mascaret_calage.ppl  
Generation des fichiers de service des unites, des entites, des parametres  
PAS de Makefile  
PAS de Make.include  
make -f Makefile.AP
```

B- Mise en place d'un événement :

```
mkdir EXECAM_01  
cd EXECAM_01  
ln -s ../SOURCES/SCRIPTS/ADOUR/piloteAM.ksh .
```

```
mkdir Observations  
cd Observations  
cp (from somewhere)/obs_ep12.csv .  
cp ../SOURCES/SCRIPTS/ADOUR/obs_split_AM.ksh .  
./obs_split_AM.ksh
```

```
cd ..  
mkdir DonneesDyn  
cd DonneesDyn  
cp (fromwhere)/*.loi .  
(si nécessaire)  
## bief8_amont : mv crue*_ep*_0.loi 1023.loi  
## bief4_amont : mv crue*_ep*_1.loi 1029.loi  
## bief9_amont : mv crue*_ep*_2.loi 1049.loi  
## bief10_amont : mv crue*_ep*_3.loi 1030.loi  
## bief1_aval : mv crue*_ep*_4.loi 1467.loi
```

```
cd ..  
mkdir DonneesStat  
cd DonneesStat  
cp (fromwhere)/Geometrie.geo .  
cp (fromwhere)/LigneEauInitiale.lig .  
cp ../SOURCES/SCRIPTS/ADOUR/ParametresMascaret.cas .  
cp ../SOURCES/SCRIPTS/ADOUR/ParametresMascaretAssim.cas .
```

```
cd ..  
mkdir Resultats
```

```
cd Resultats
ln -s ../../SOURCES/SCRIPTS/ADOUR/calcul_criteresAM.f90 .
gfortran -o calcul_criteresAM calcul_criteresAM.f90
ln -s ../../SOURCES/SCRIPTS/ADOUR/plot_crit_mod1754.gp .
ln -s ../../SOURCES/SCRIPTS/ADOUR/plot_crit_mod296.gp .
ln -s ../../SOURCES/SCRIPTS/ADOUR/plot_crit_mod10.gp .
ln -s ../../SOURCES/SCRIPTS/ADOUR/plot_crit_assim1754.gp .
ln -s ../../SOURCES/SCRIPTS/ADOUR/plot_crit_assim296.gp .
ln -s ../../SOURCES/SCRIPTS/ADOUR/plot_crit_assim10.gp .
ln -s ../../SOURCES/SCRIPTS/ADOUR/updres.bash .
```

```
cd ..
mkdir Programmes
cd Programmes
mkdir Mascaret
cd Mascaret
ln -s ../../SOURCES/EXEC/Abaques.txt .
ln -s ../../SOURCES/EXEC/dico.txt .
ln -s ../../SOURCES/EXEC/dicoassim.txt .
ln -s ../../SOURCES/EXEC/FichierCas.txt .
ln -s ../../SOURCES/EXEC/FichierCasAssim.txt .
ln -s ../../SOURCES/EXEC/mascaret7 systeme_mascaret
```

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

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

C- Simulation d'un événement :

Usage du piloteAM pour lancer une série de runs Mascaret :

```
./piloteAM.ksh Tdeb(s) Tfin(s) pas(s) tequilibrage(s) tprevision(s) DRYRUN|ASSIM|CALAGE|
CALAGEASSIM [FAST]
```

Typiquement avec le modele AdourMaritime on utilise le pilote avec :

```
pas = 1h = 3600s
tequilibrage = 3j = 259200s
tprevision = 1j = 86400s
```

Exemples d'événements simulés sur AdourMaritime :

EXECAM_01 :

```
./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 CALAGEASSIM

(3j) (~14j) (1h) (3j) (0.5j)

EXECAM_NOV09_2CRUES :

./piloteAM.ksh 604800 1382400 3600 259200 43200 CALAGEASSIM

(7j) (16j) (1h) (3j) (0.5j)

Annexe 2 : Récapitulatif Marne

A- Compilation

```
Dans SOURCES/EXEC  
make clean  
make
```

```
Dans SOURCES/Assimilation/PALM  
make -f Makefile.AP allclean  
prepalmmmp mascaret_calage.ppl  
Generation des fichiers de service des unites, des entites, des parametres  
PAS de Makefile  
PAS de Make.include  
make -f Makefile.AP
```

B- Mise en place d'un événement :

```
mkdir EXECMV_05  
cd EXECMV_05  
ln -s ../SOURCES/SCRIPTS/MARNE/piloteMV.ksh .
```

```
mkdir Observations  
cd Observations  
cp (fromwhere)/2006_03_23.csv .  
cp ../SOURCES/SCRIPTS/MARNE/obs_split_MV.ksh .  
./obs_split_MV.ksh
```

```
cd ..  
mkdir DonneesDyn  
cd DonneesDyn  
cp (fromwhere)/*.loi .
```

```
cd ..  
mkdir DonneesStat  
cd DonneesStat  
cp (fromwhere)/Geometrie.geo .  
cp (fromwhere)/1489CT.loi .  
cp (fromwhere)/Qapport.loi .  
cp (fromwhere)/Qapport2.loi .  
cp (fromwhere)/LigneEauInitiale.lig .  
cp ../SOURCES/SCRIPTS/MARNE/ParametresMascaret.cas .  
cp ../SOURCES/SCRIPTS/MARNE/ParametresMascaretAssim.cas .
```

```
cd ..  
mkdir Resultats  
cd Resultats  
ln -s ../SOURCES/SCRIPTS/MARNE/calcul_criteresMV.f90 .  
gfortran -o calcul_criteresMV calcul_criteresMV.f90  
ln -s ../SOURCES/SCRIPTS/MARNE/plot_crit_mod381.gp .
```

```
ln -s ../../SOURCES/SCRIPTS/MARNE/plot_crit_mod258.gp .
ln -s ../../SOURCES/SCRIPTS/MARNE/plot_crit_assim381.gp .
ln -s ../../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/EXEC/Abaques.txt .
ln -s ../../SOURCES/EXEC/dico.txt .
ln -s ../../SOURCES/EXEC/dicoassim.txt .
ln -s ../../SOURCES/EXEC/FichierCas.txt .
ln -s ../../SOURCES/EXEC/FichierCasAssim.txt .
ln -s ../../SOURCES/EXEC/mascaret7 systeme_mascaret
```

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

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

C- Simulation d'un événement :

Usage du piloteMV pour lancer une série de runs Mascaret :

```
./piloteMV.ksh Tdeb(s) Tfin(s) pas(s) tequilibrage(s) tprevision(s) DRYRUN|ASSIM|CALAGE|
CALAGEASSIM [FAST]
```

Typiquement avec le modele MarneVallage on utilise le pilote avec :

```
pas = 1h = 3600s
tequilibrage = 3j = 259200s
tprevision = 1j = 86400s
```

Exemples d'événements simulés sur MarneVallage :

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]
(4j2h) (15j) (1h) (4j) (1j)

Annexe 3 : 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 (30 Novembre 2009)

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

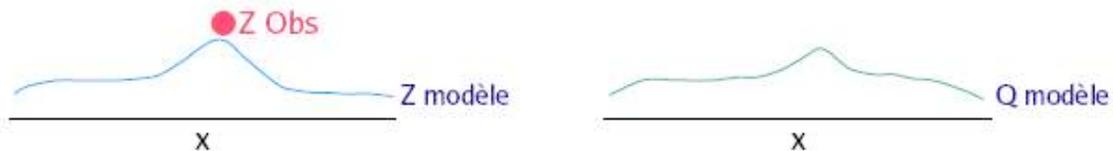
G. Jonville, A. Piacentini, O. Thual, C. Ivanoff, E. Le Pape

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

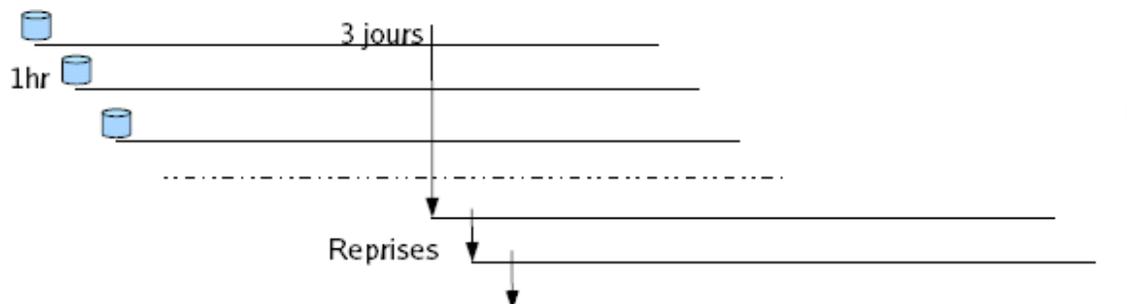
Des épisodes de crue pour les réseaux Marne Vallage et Adour Maritime ont été sélectionnés par Etienne Le Pape.

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.

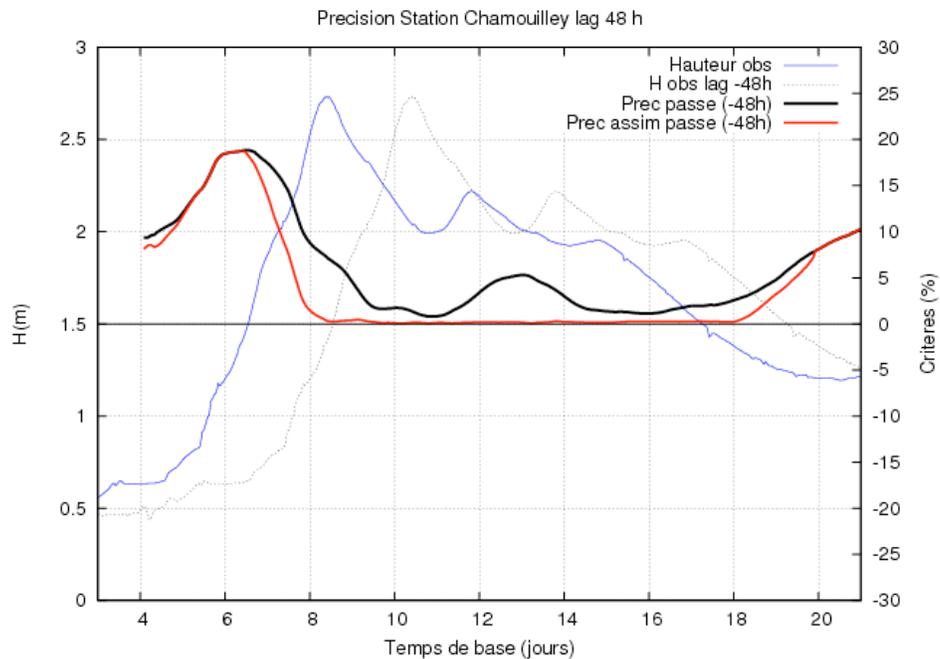
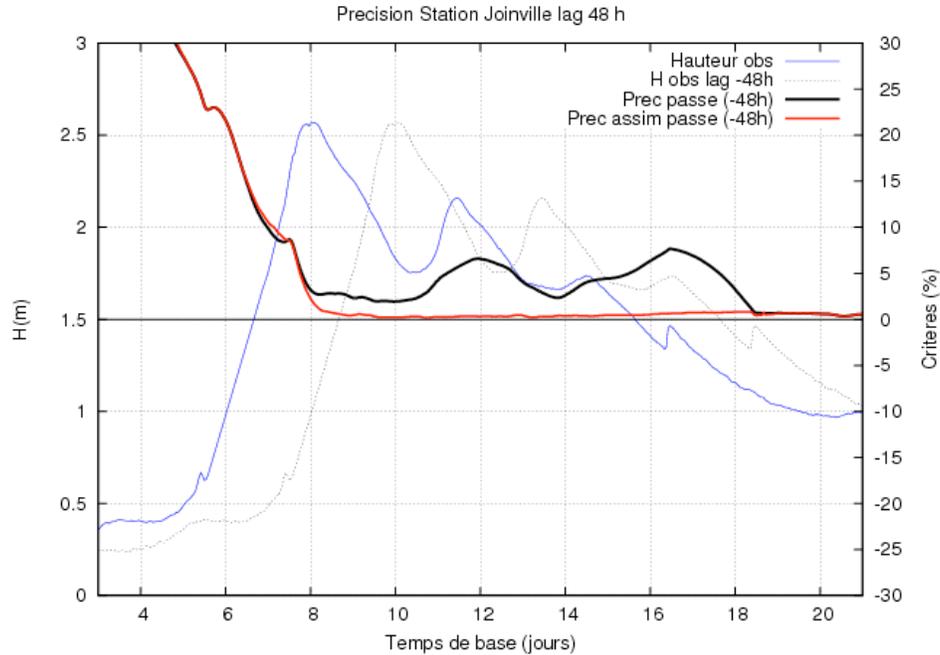
La courbe en rouge indique les précisions ou les hauteurs du modèle avec assimilation.

Episode MARNE VALLAGE 01 : janvier 2004

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

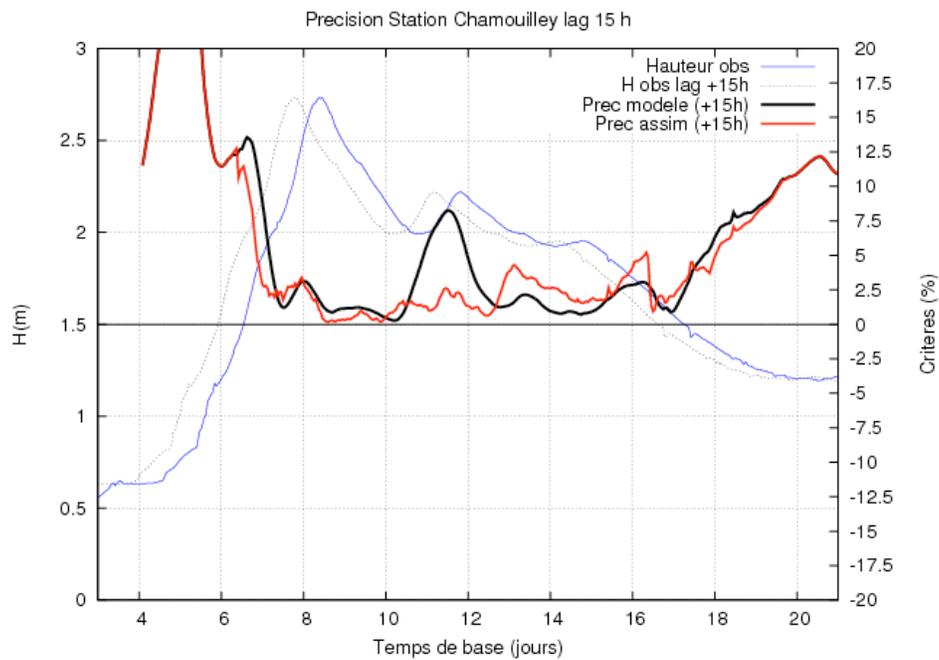
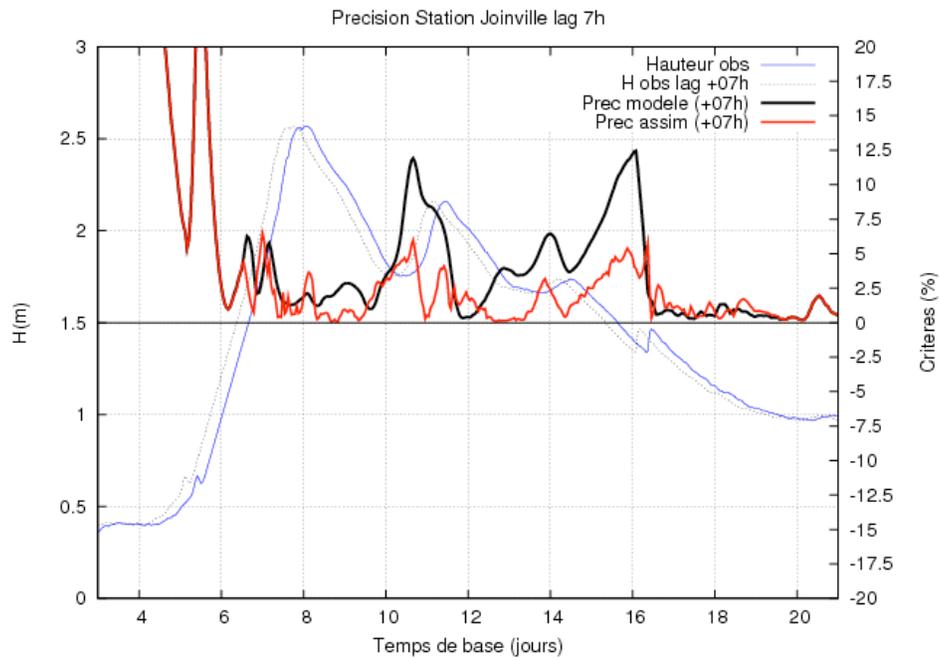
1- Précisions en mode simulation

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.



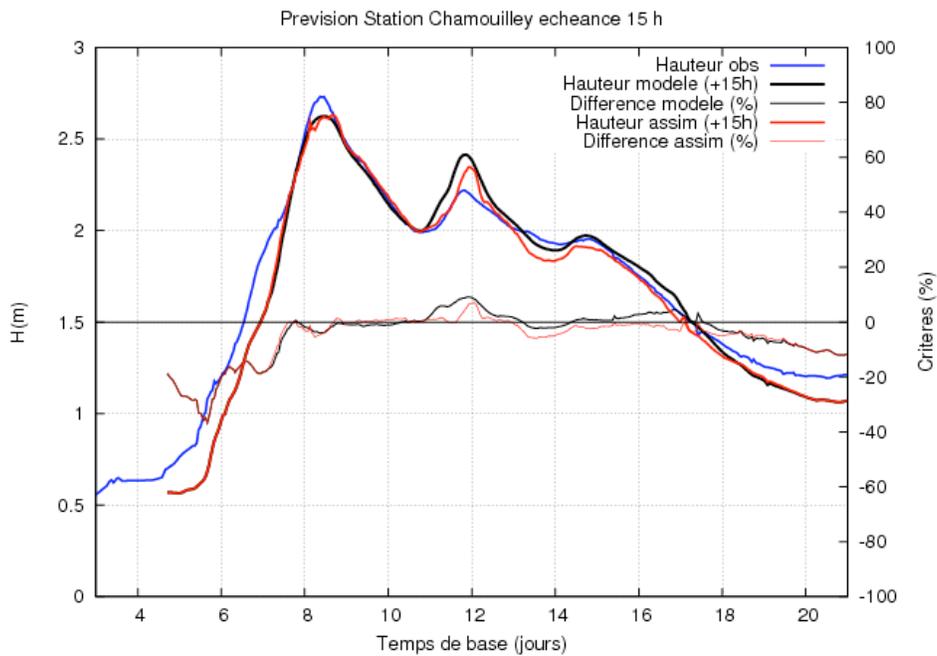
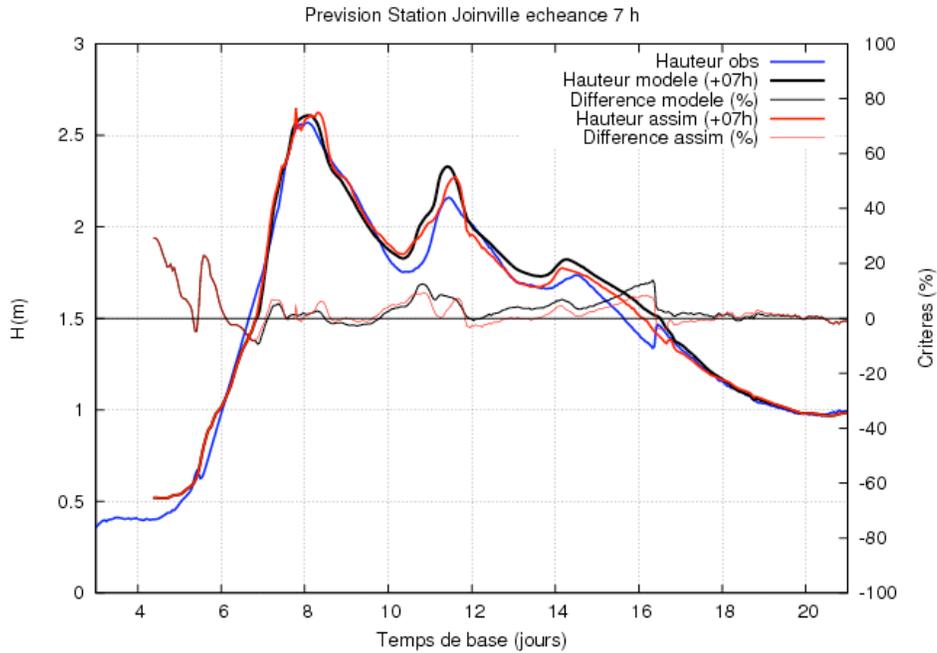
2- Précisions en mode prévision

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.



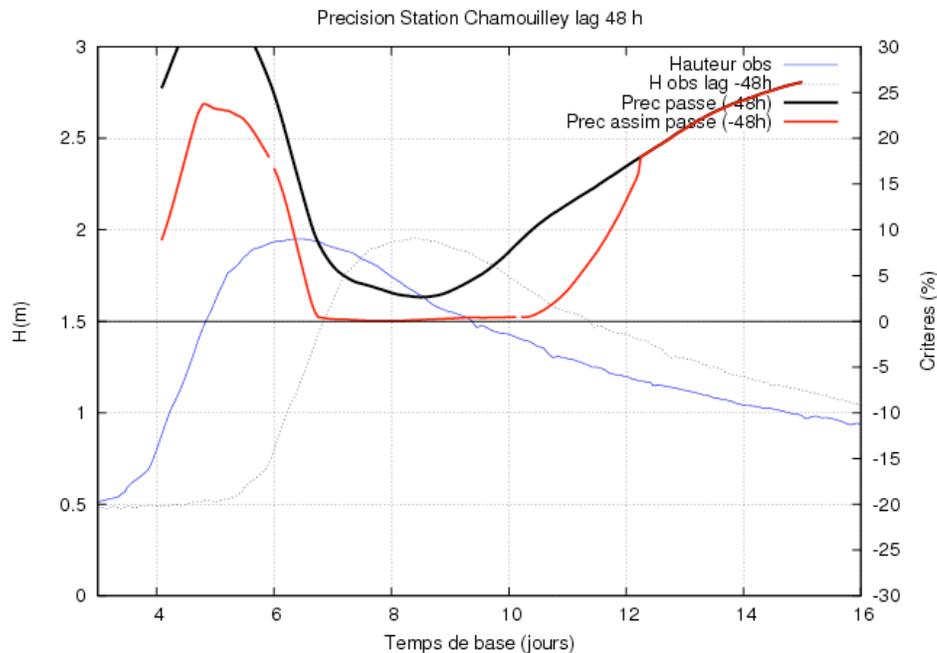
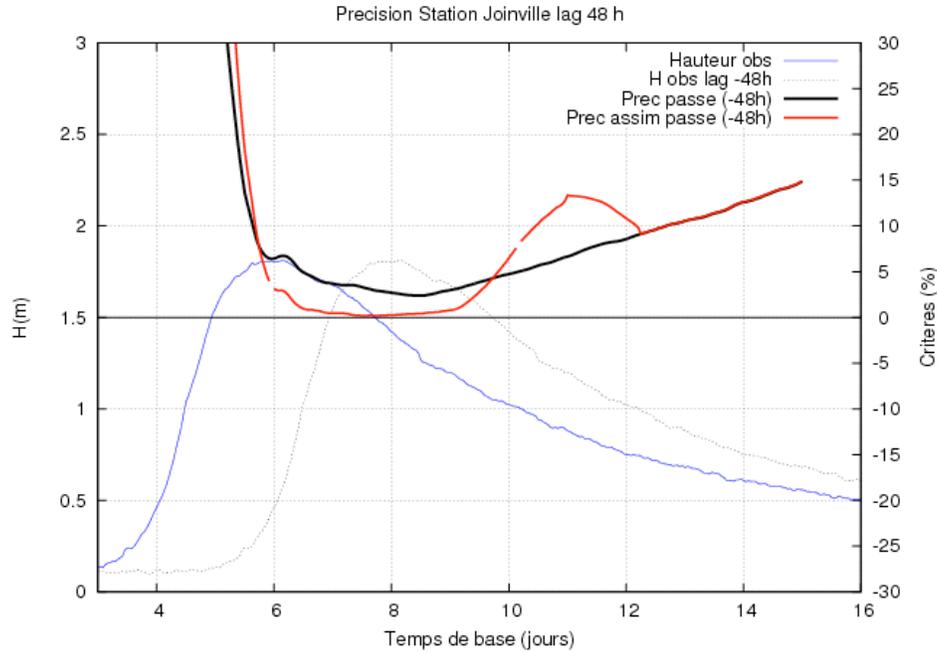
L'assimilation en mode CALAGEASSIM a apporté peu d'amélioration au modèle.

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).

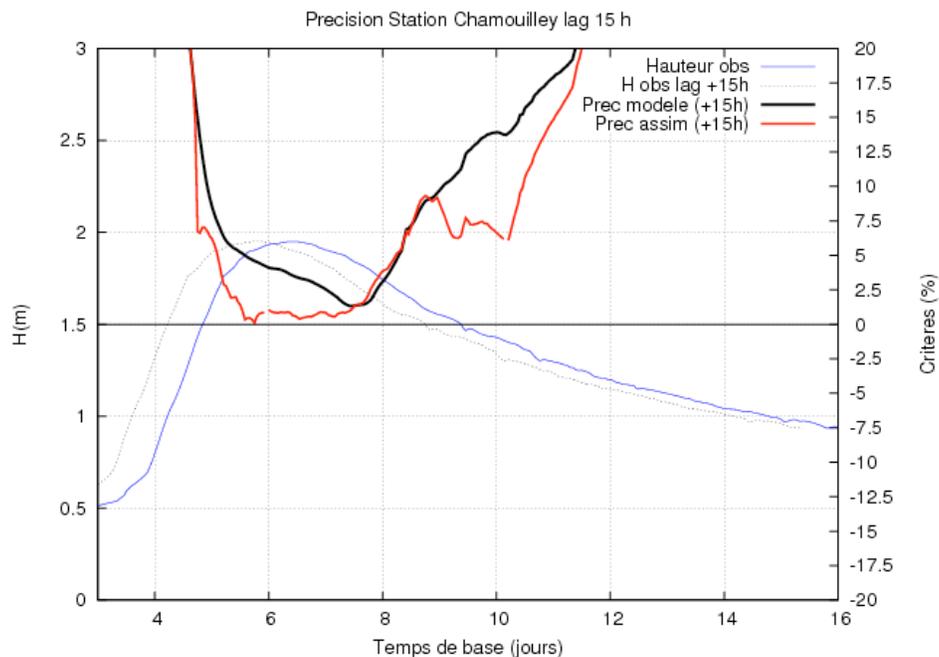
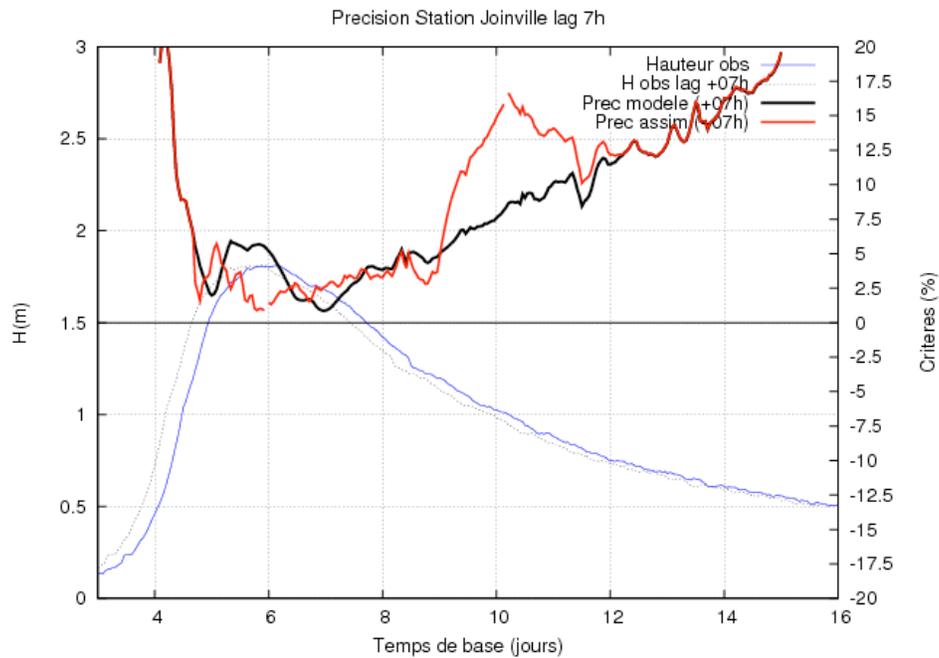
1- Précisions en mode simulation

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.



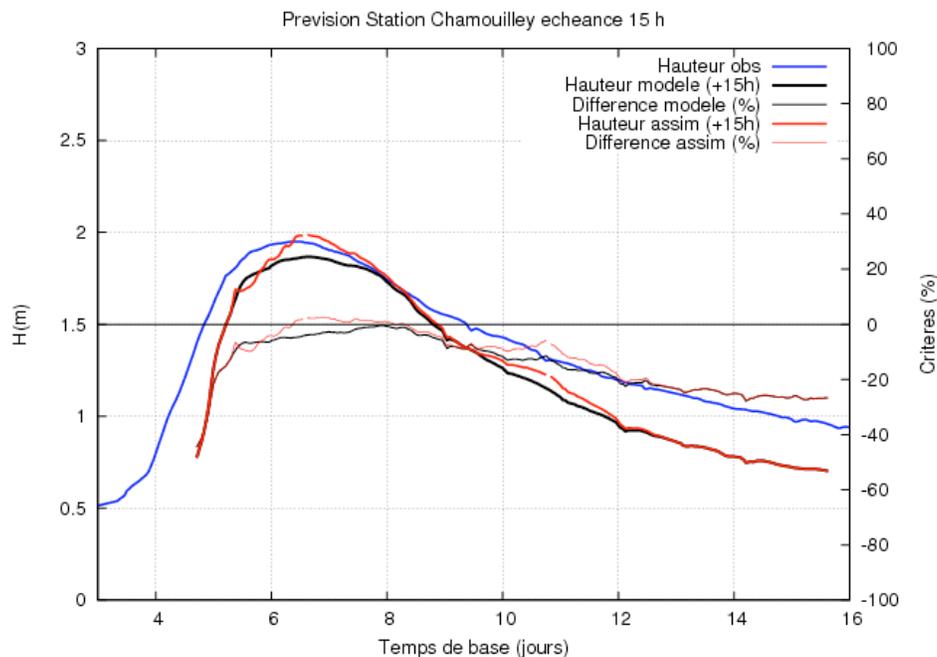
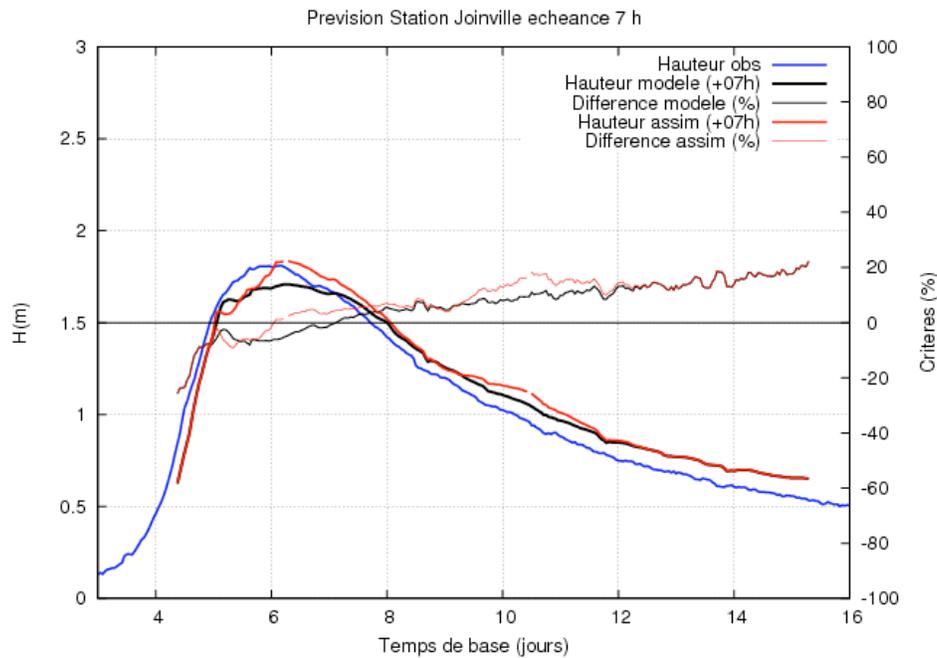
2- Précisions en mode prévision

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.



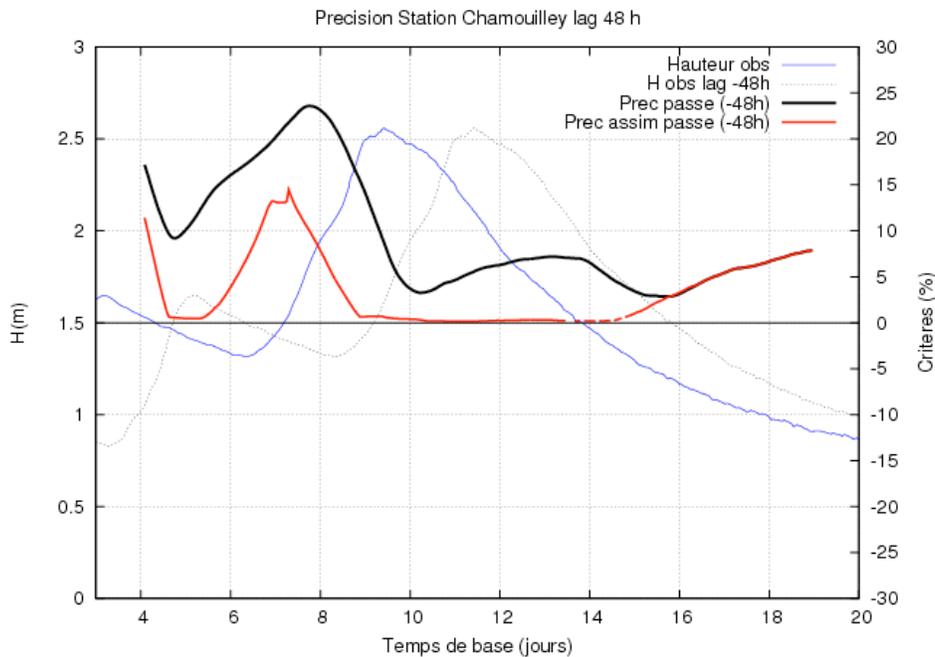
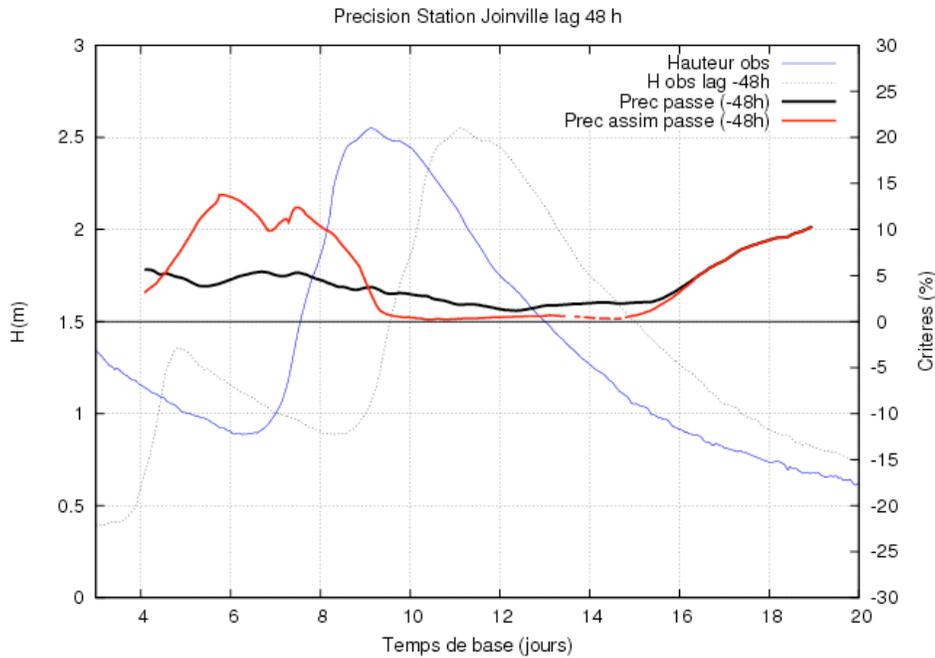
L'assimilation en mode CALAGEASSIM a amélioré la prévision de la crue aux deux stations.

Episode MARNE VALLAGE 03 : mars 2006

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

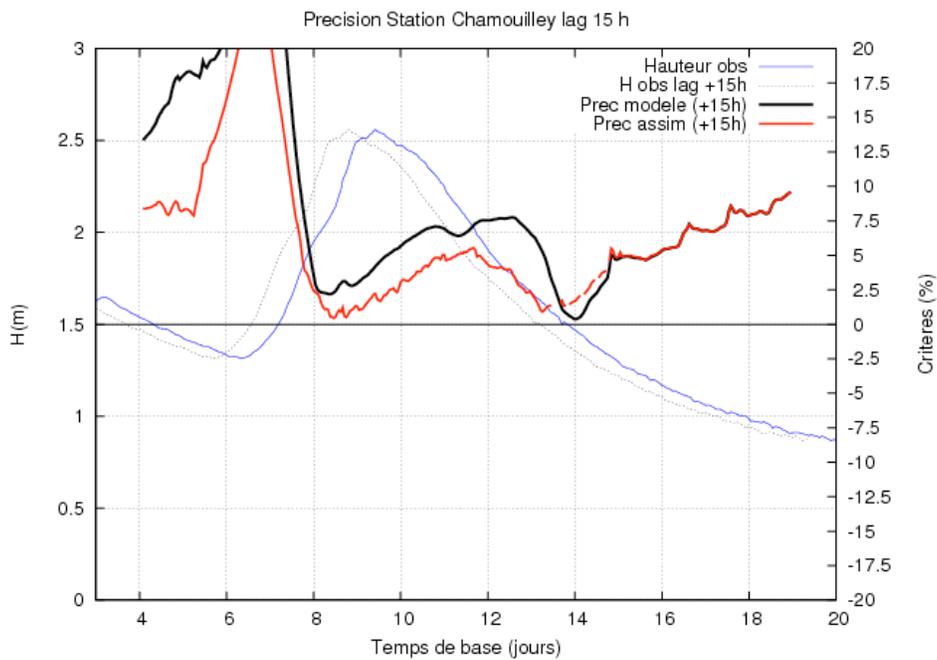
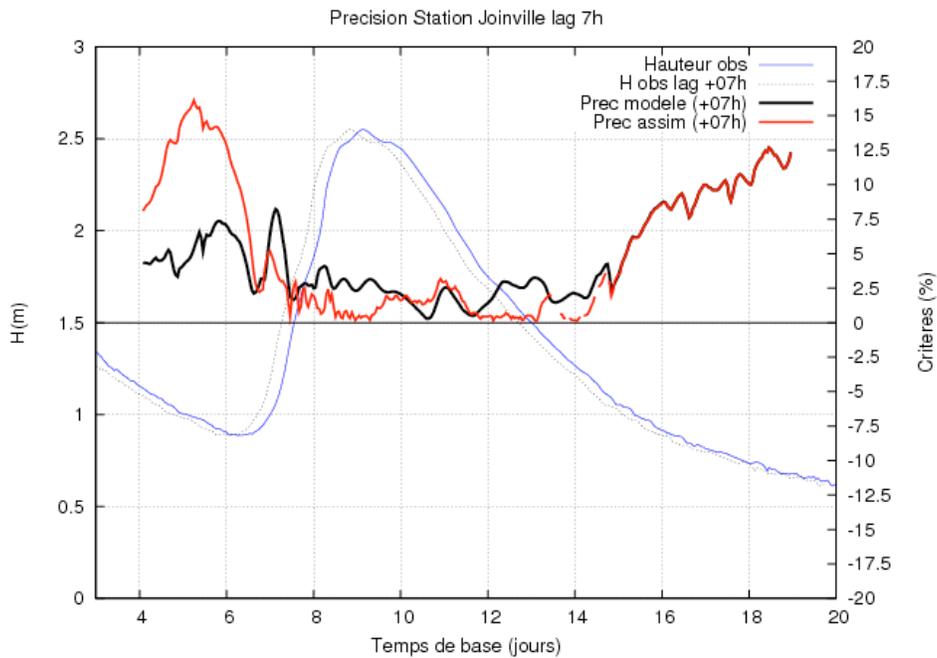
1- Précisions en mode simulation

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.



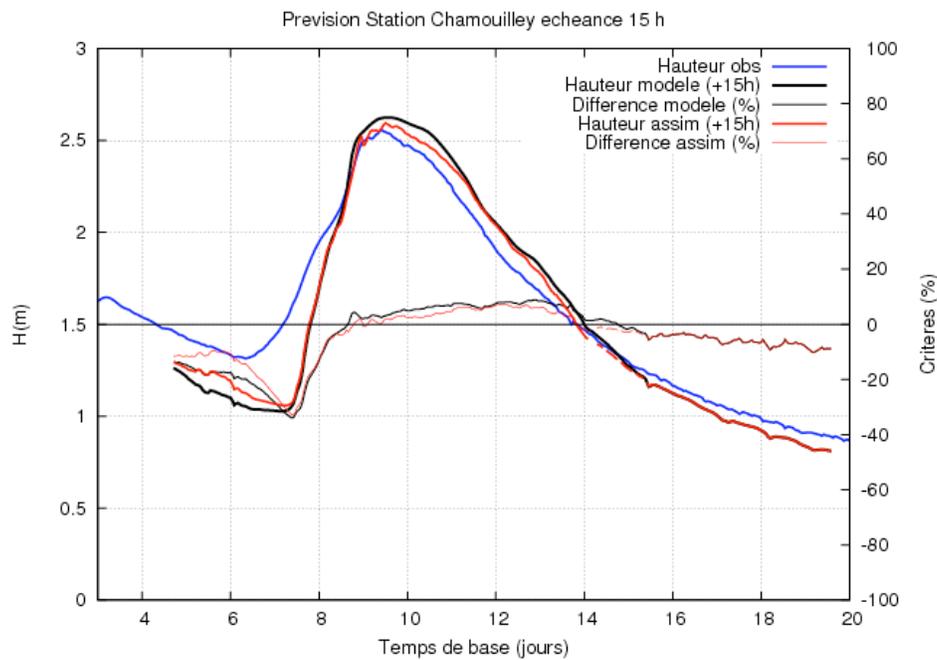
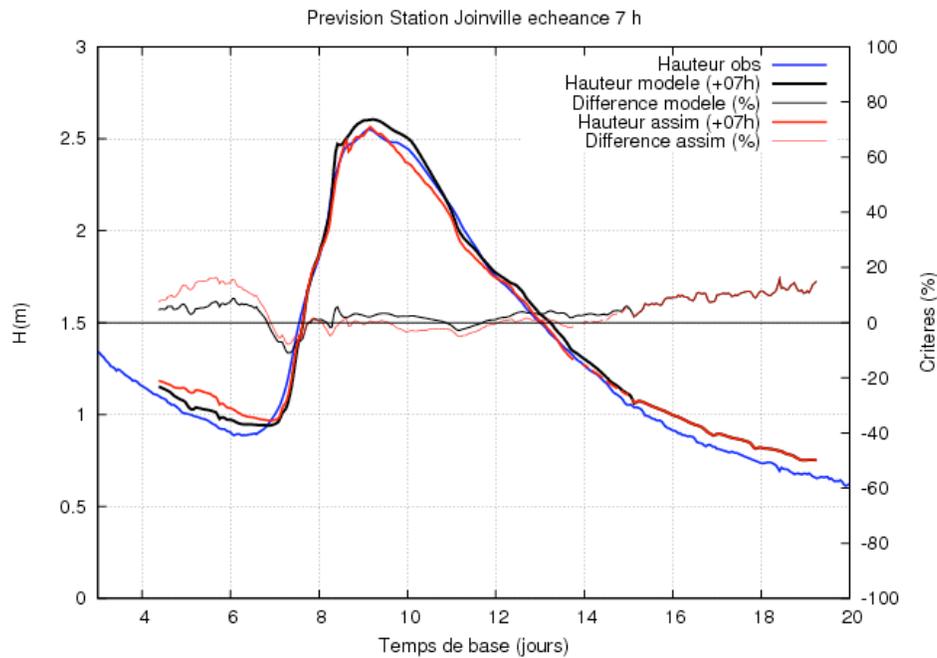
2- Précisions en mode prévision

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.



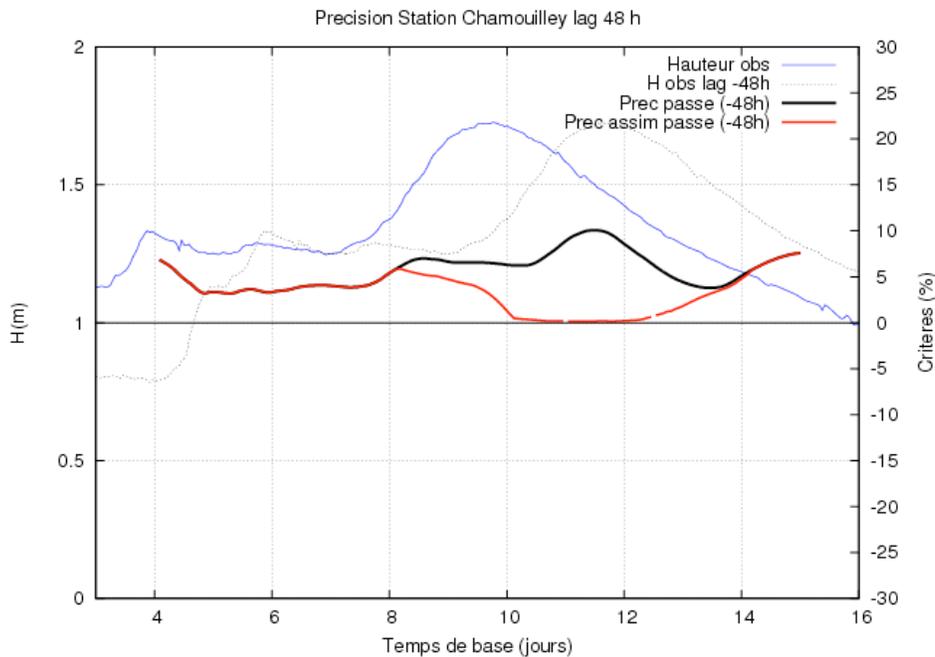
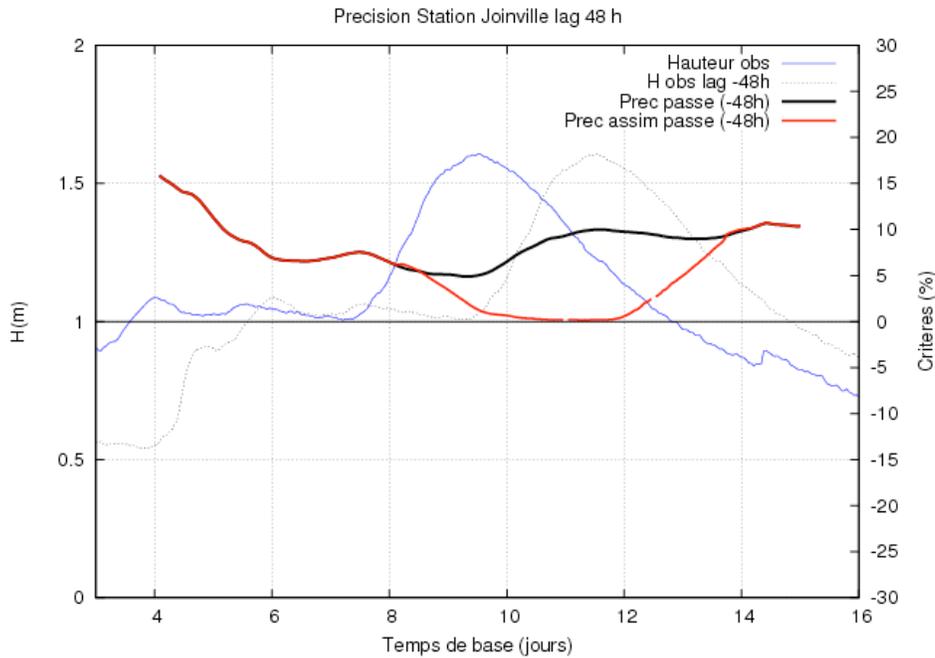
L'assimilation en mode CALAGEASSIM a amélioré la prévision de la crue aux deux stations.

Episode MARNE VALLAGE 05 : avril 2006

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

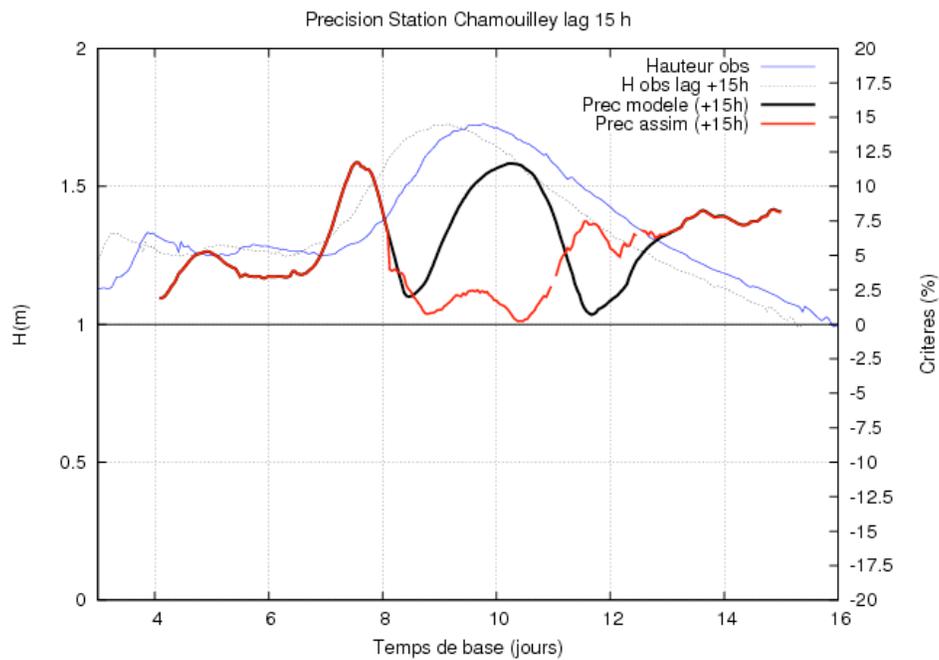
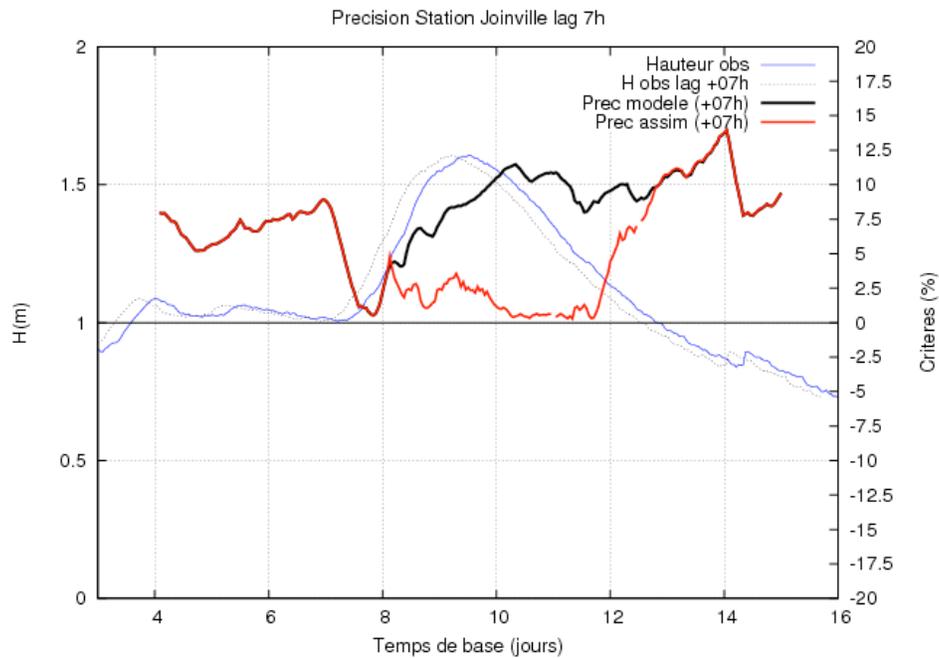
1- Précisions en mode simulation

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



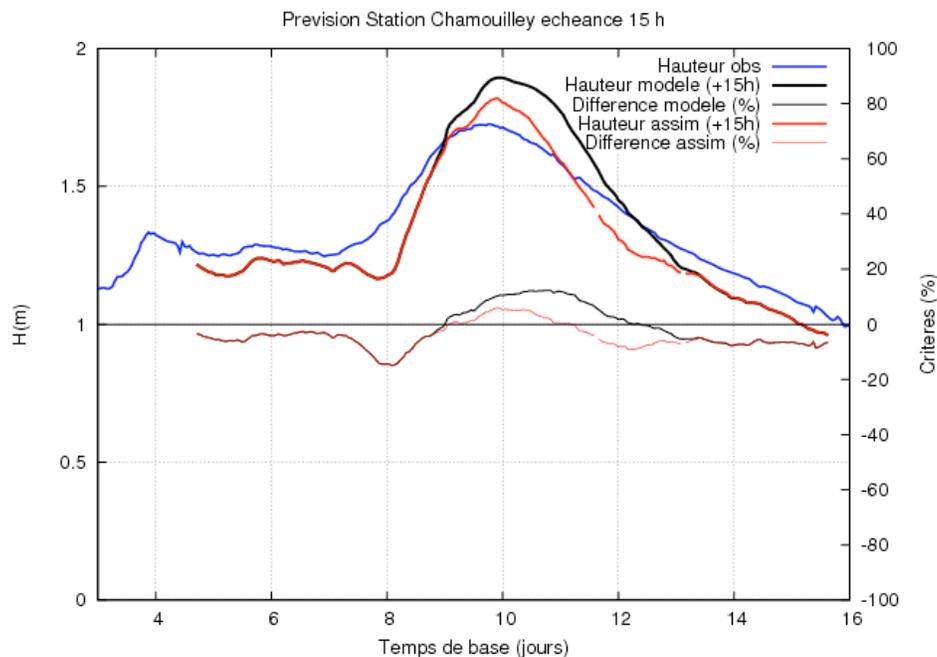
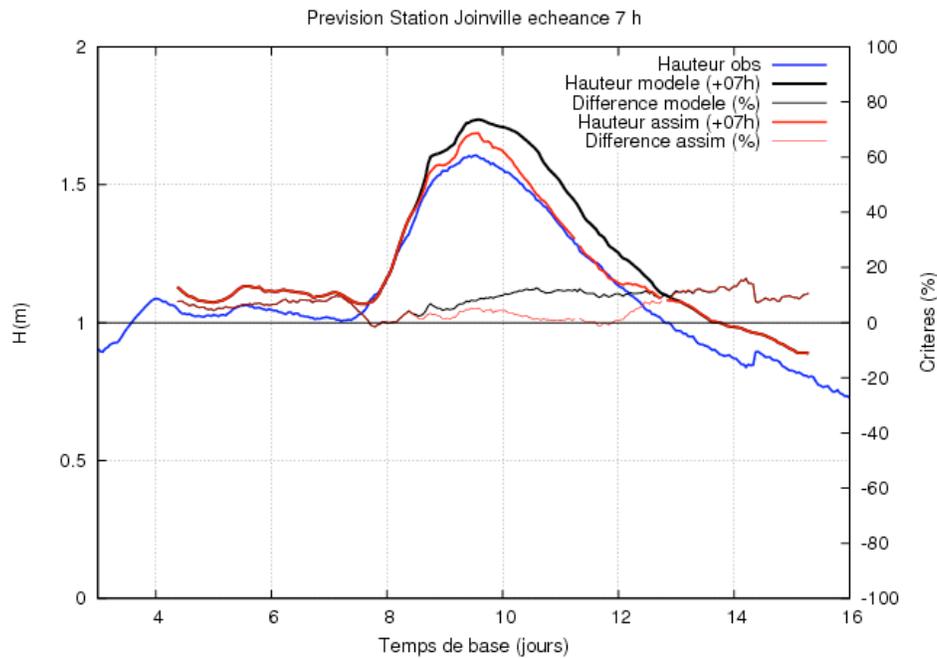
2- Précisions en mode prévision

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.



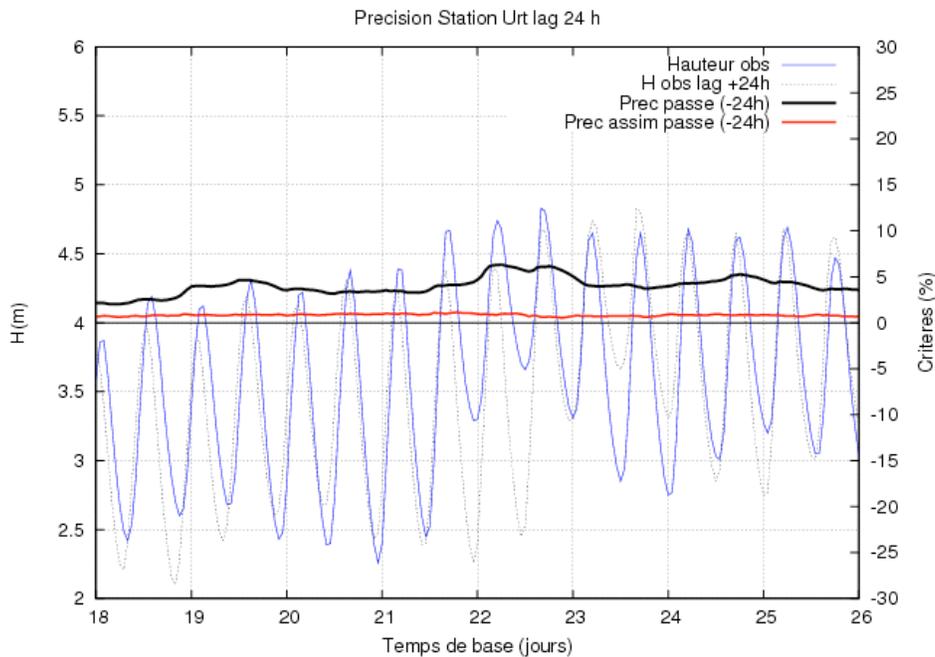
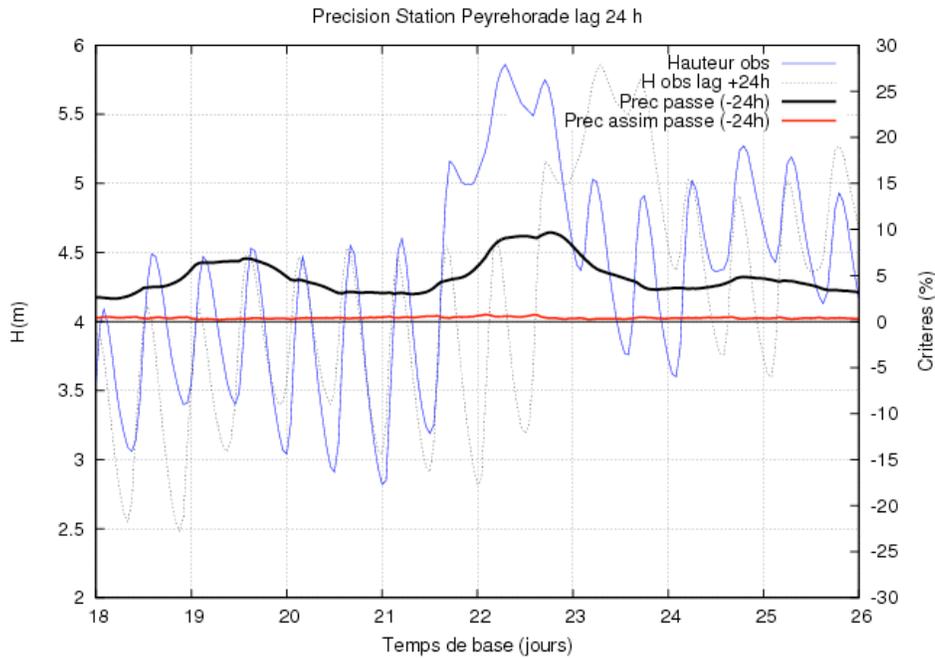
Le modèle direct surestime la crue. L'assimilation en mode CALAGEASSIM a amélioré nettement la prévision avec une légère surestimation du pic.

Episode ADOUR MARITIME 01

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

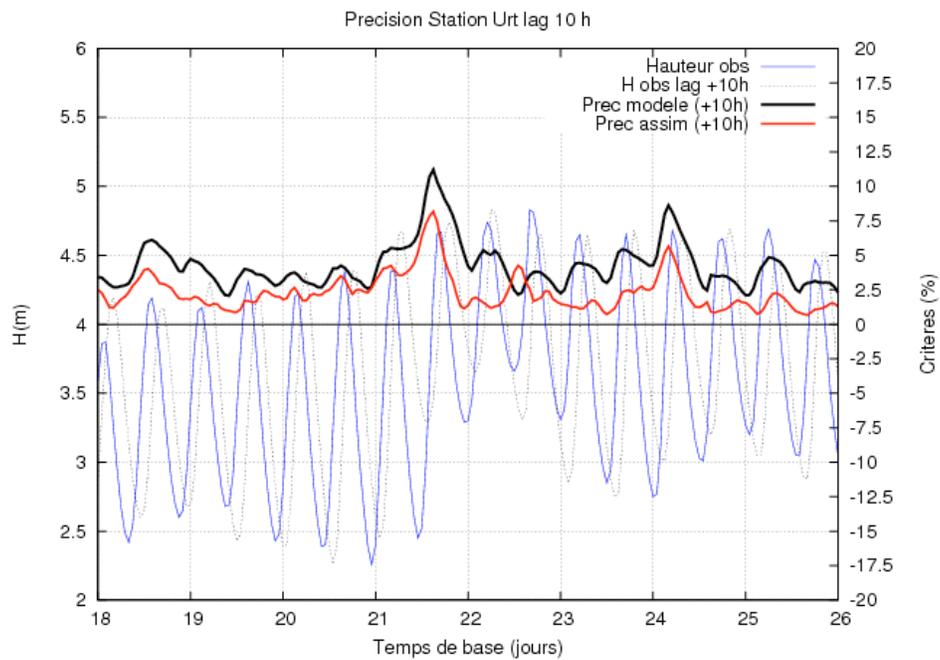
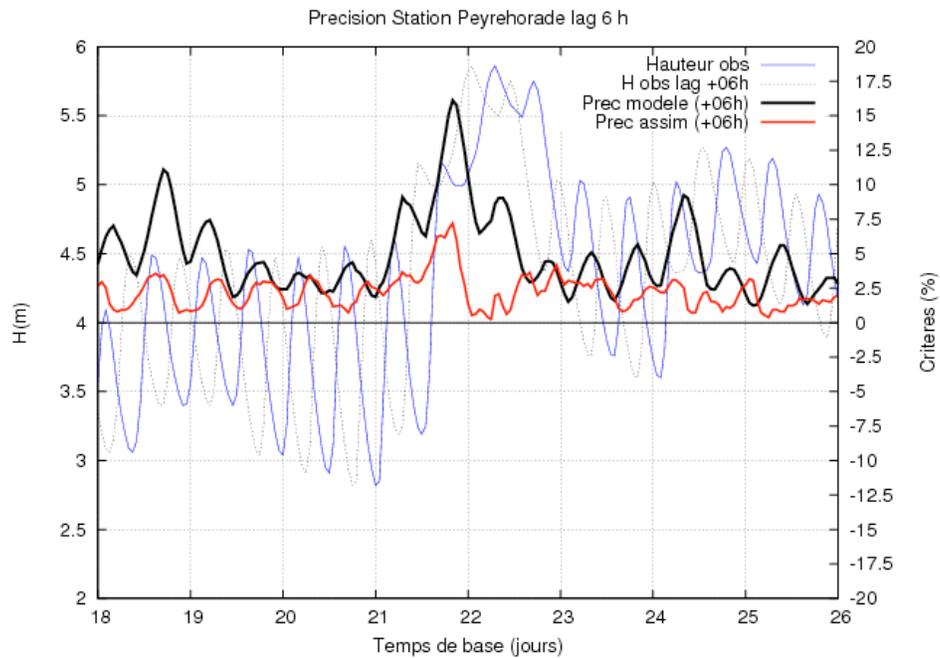
1- Précisions en mode simulation

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



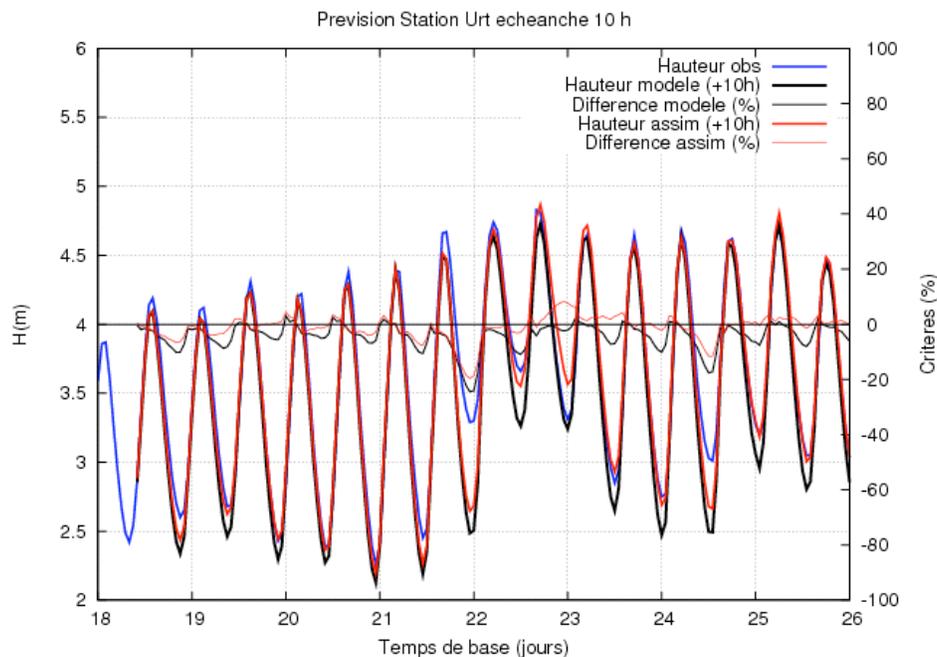
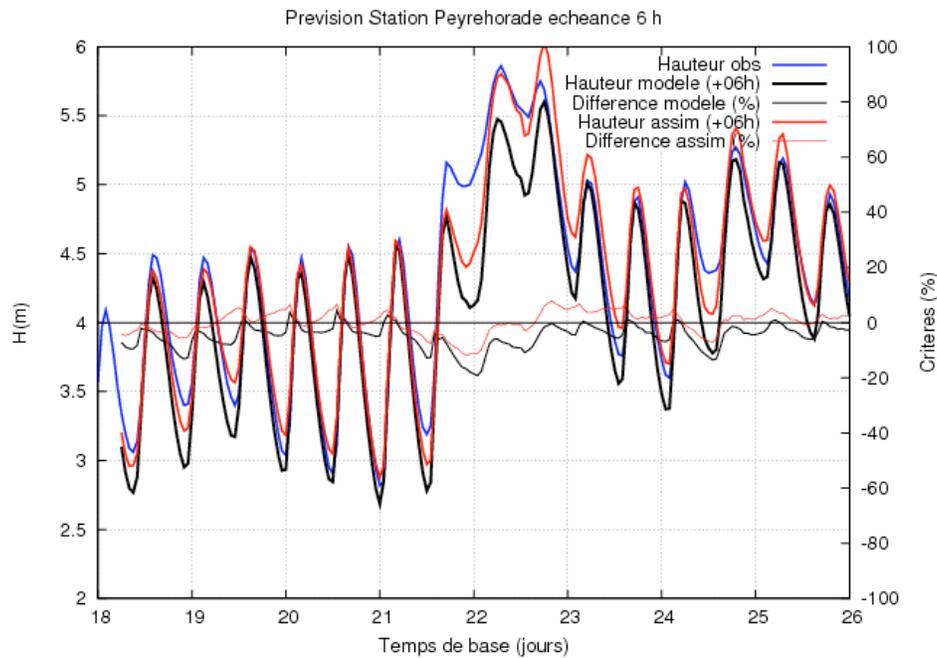
2- Précisions en mode prévision

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.



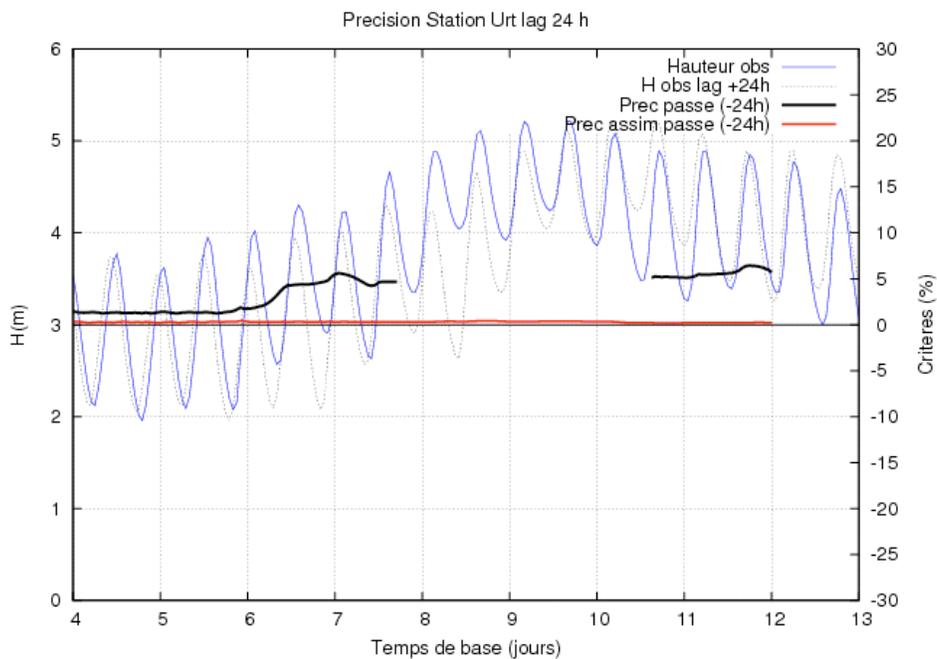
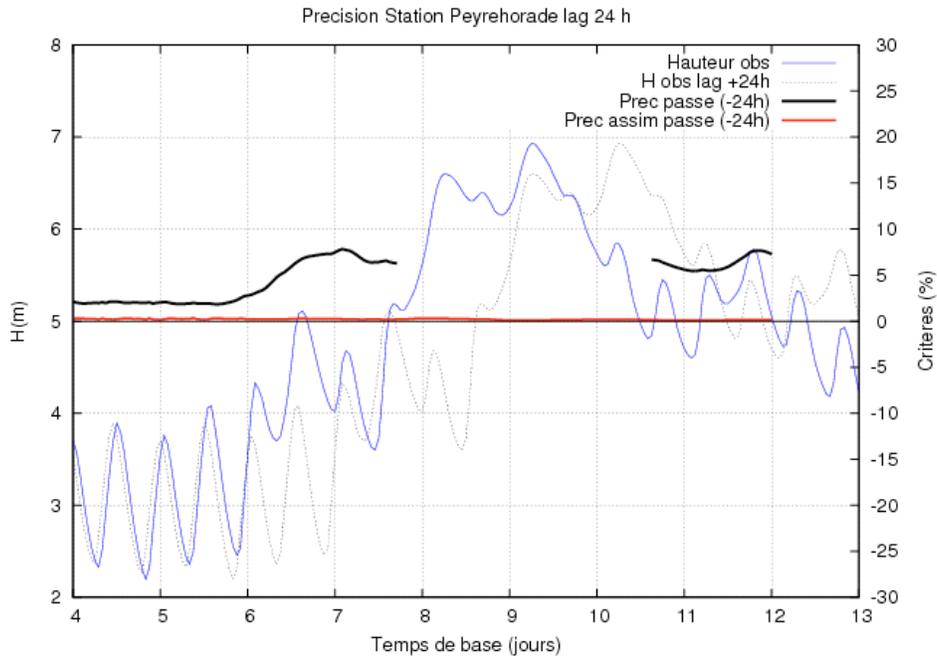
L'assimilation en mode CALAGEASSIM a amélioré nettement la prévision de crue aux deux stations.

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).

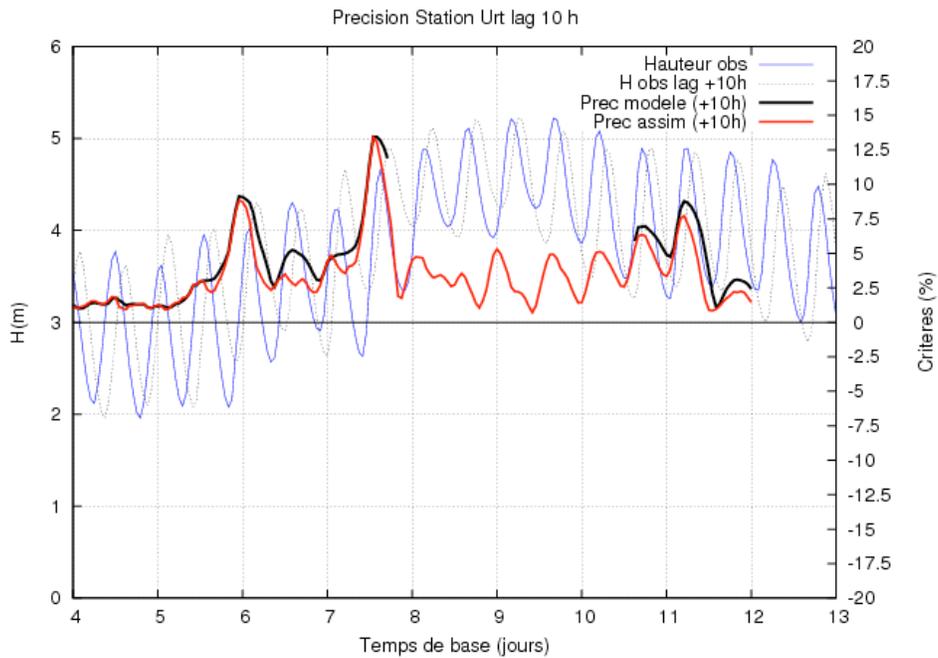
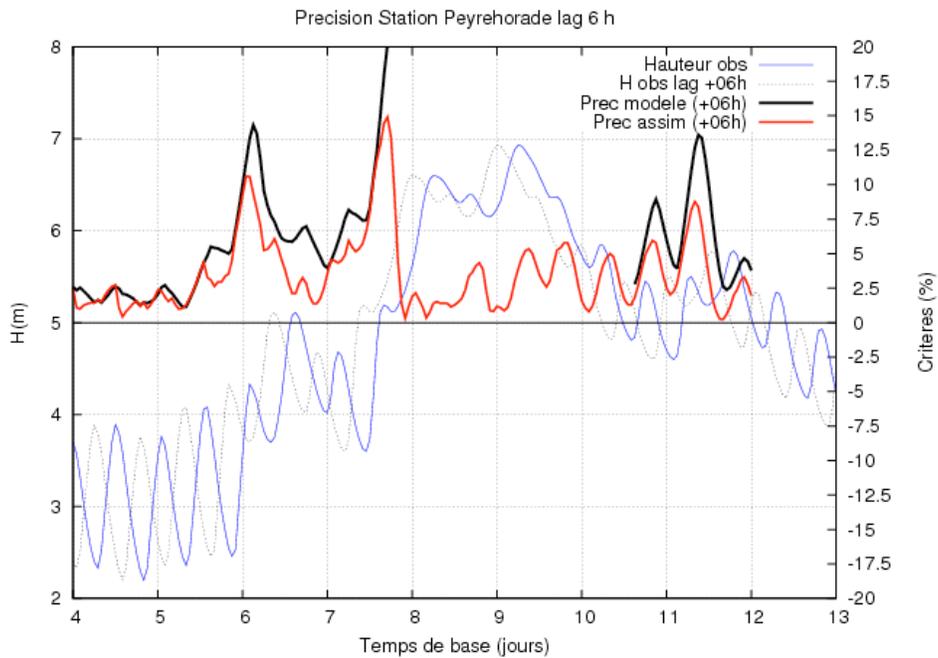
1- Précisions en mode simulation

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



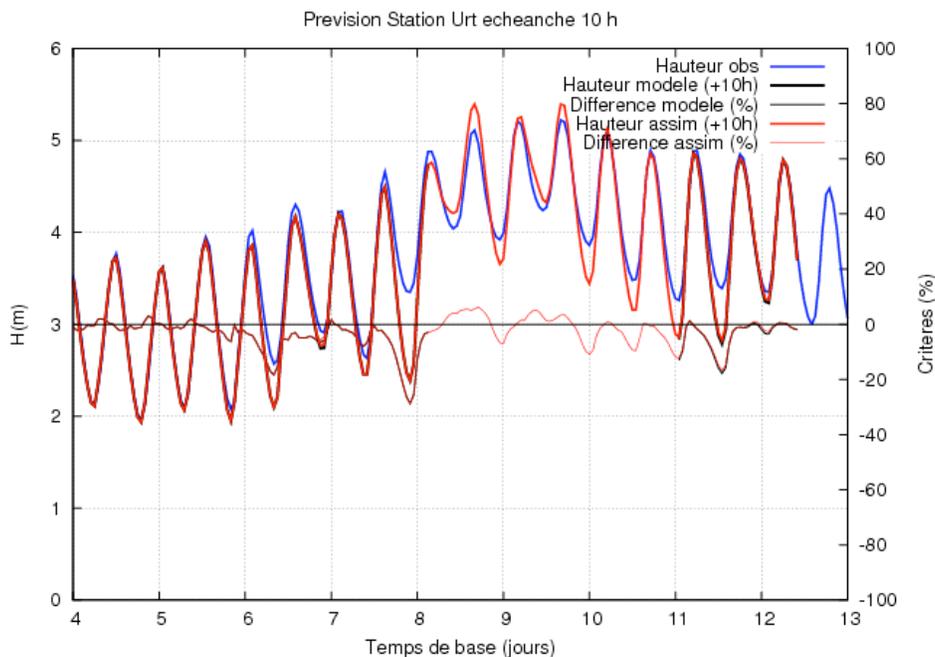
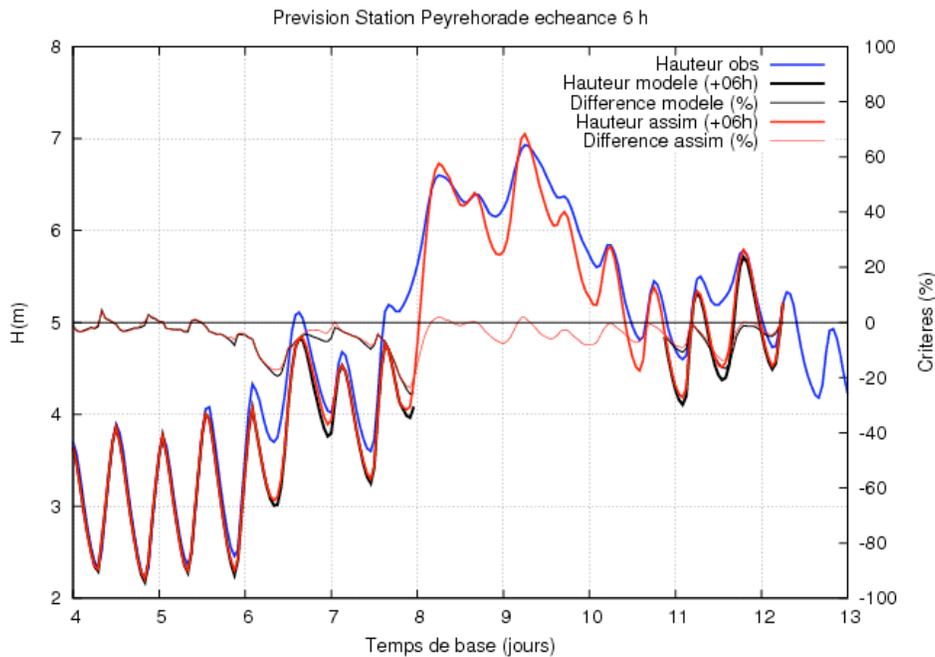
2- Précisions en mode prévision

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



3- Hauteurs prévues à l'échéance usuelle

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



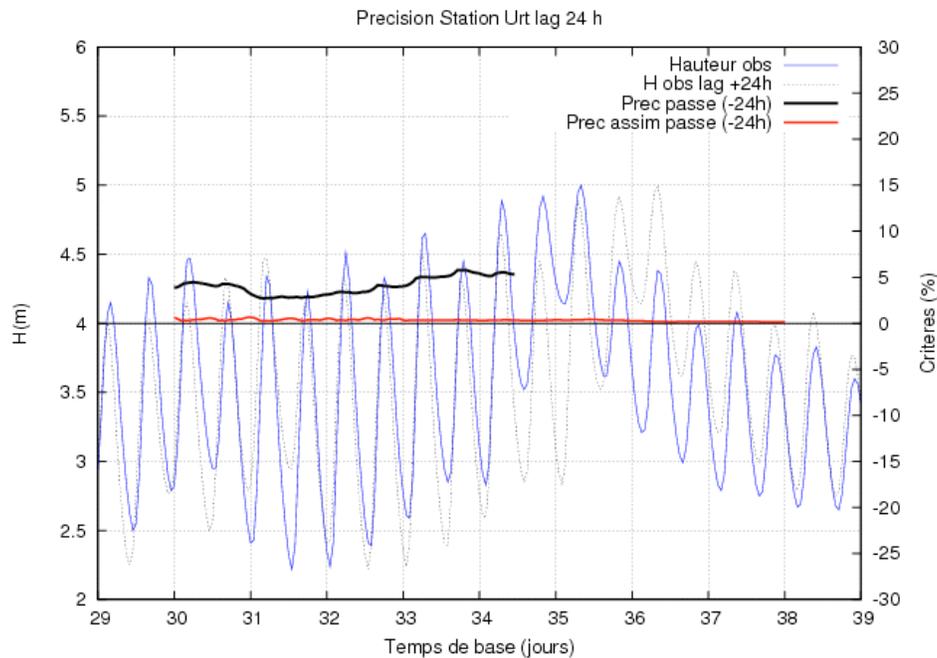
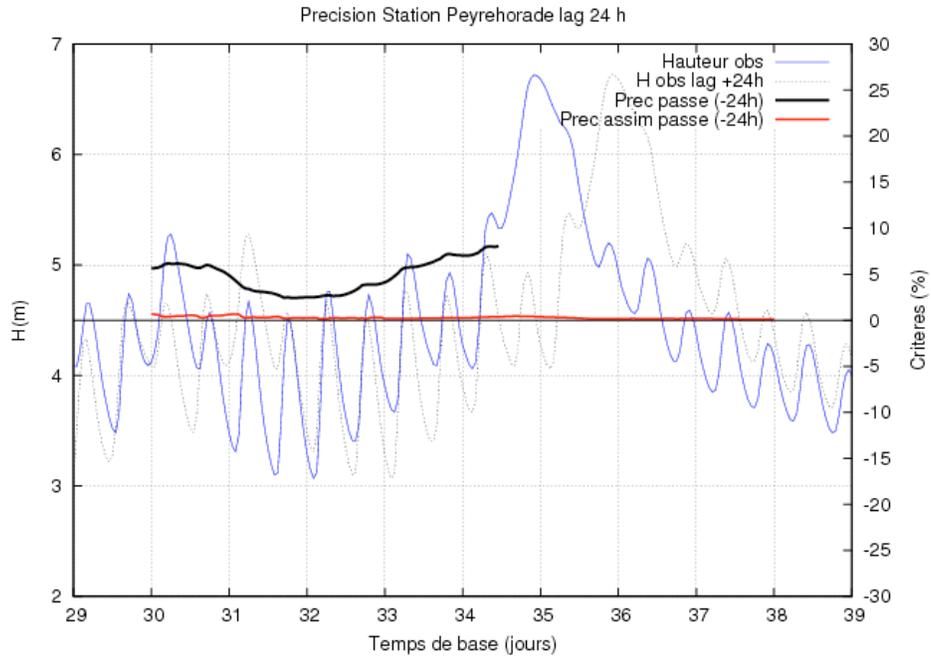
L'assimilation en mode ASSIM a permis de simuler l'épisode et donne de bons résultats de prévision de cette double crue.

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).

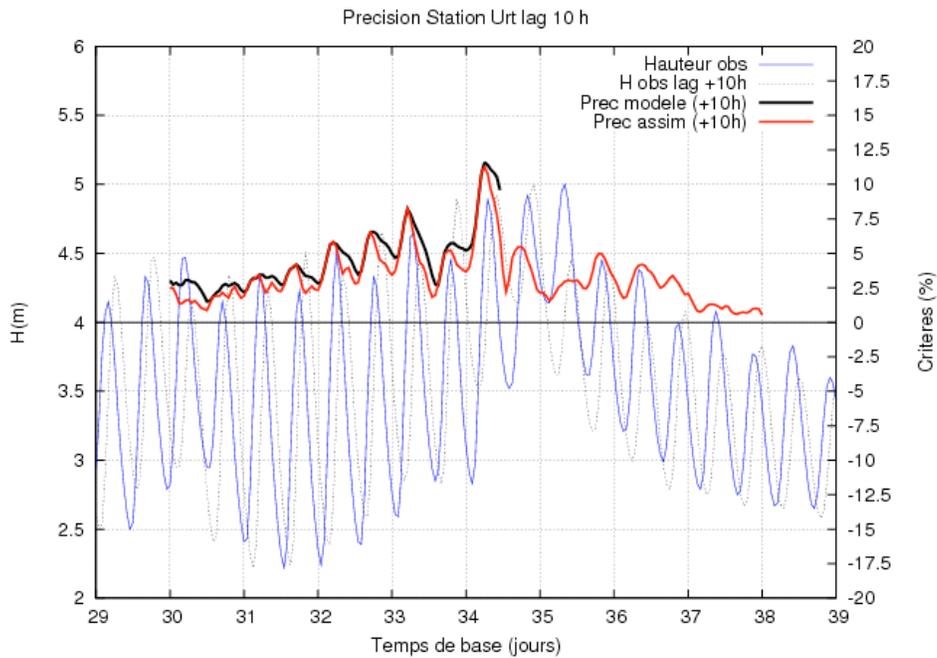
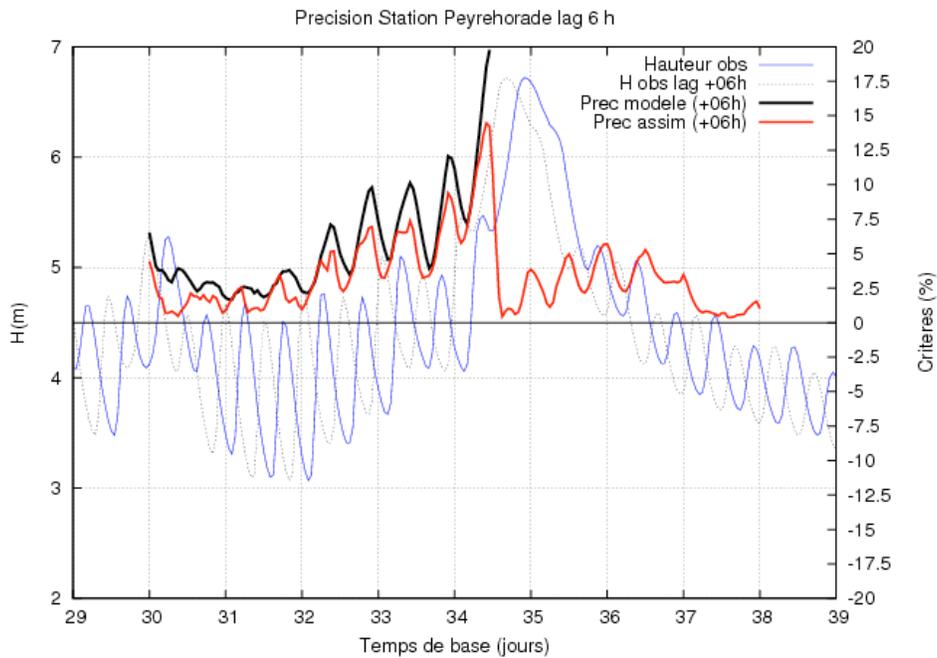
1- Précisions en mode simulation

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



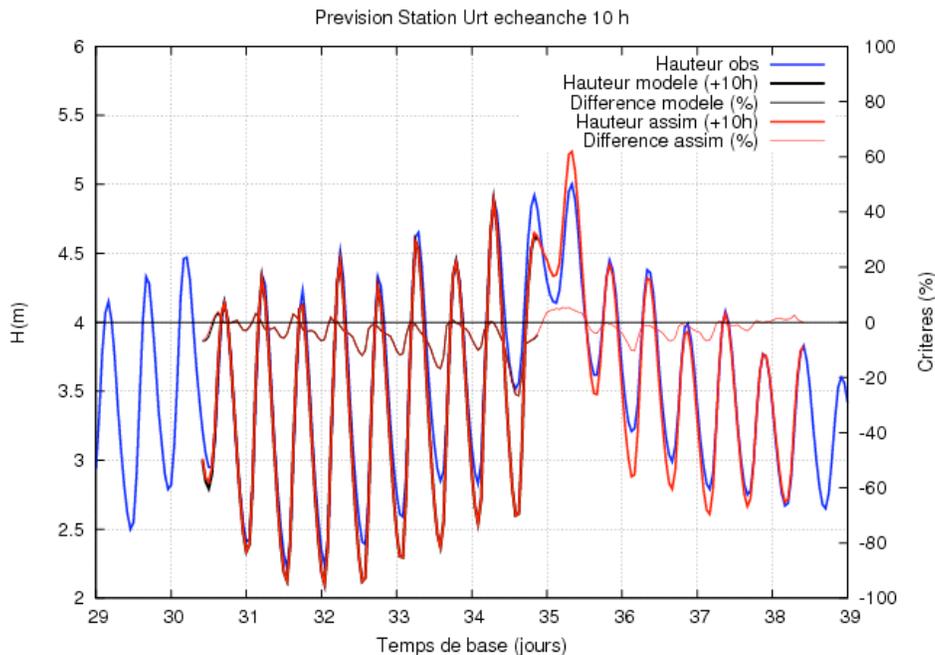
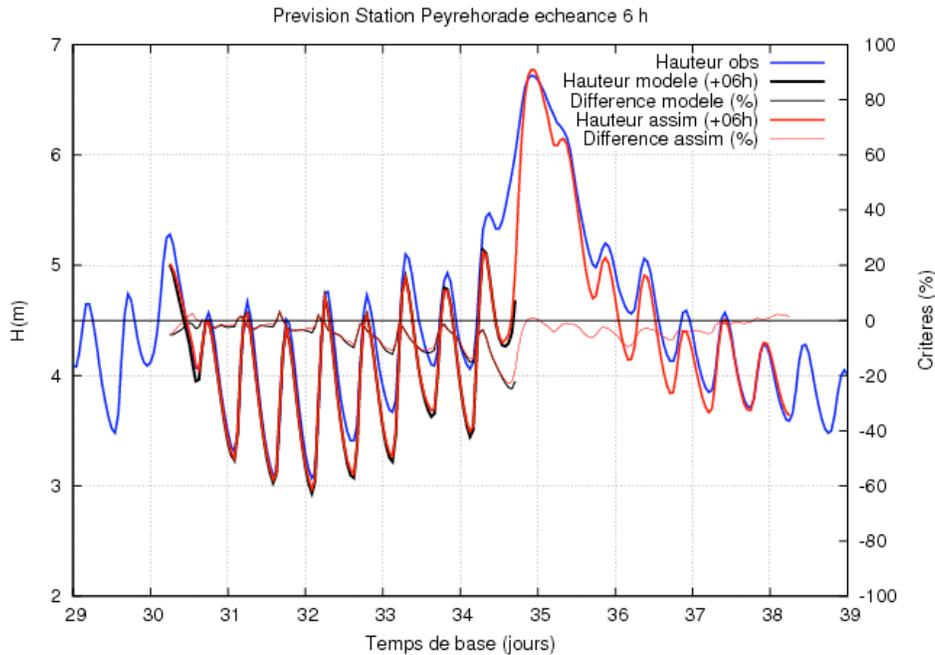
2- Précisions en mode prévision

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.



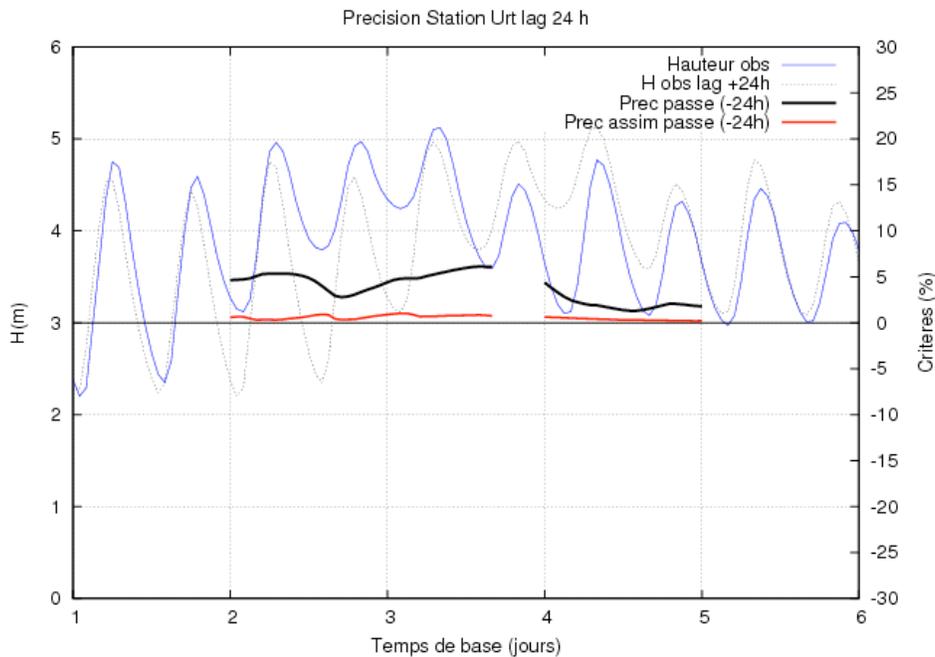
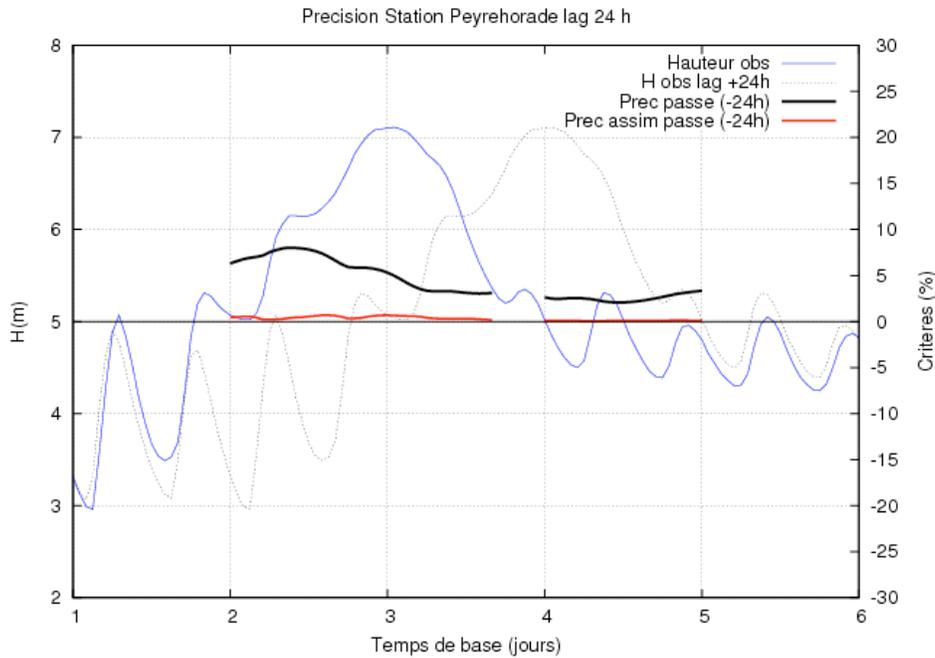
L'assimilation en mode ASSIM a permis de simuler l'épisode et a donné de bons résultats.

Episode ADOUR MARITIME 05

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

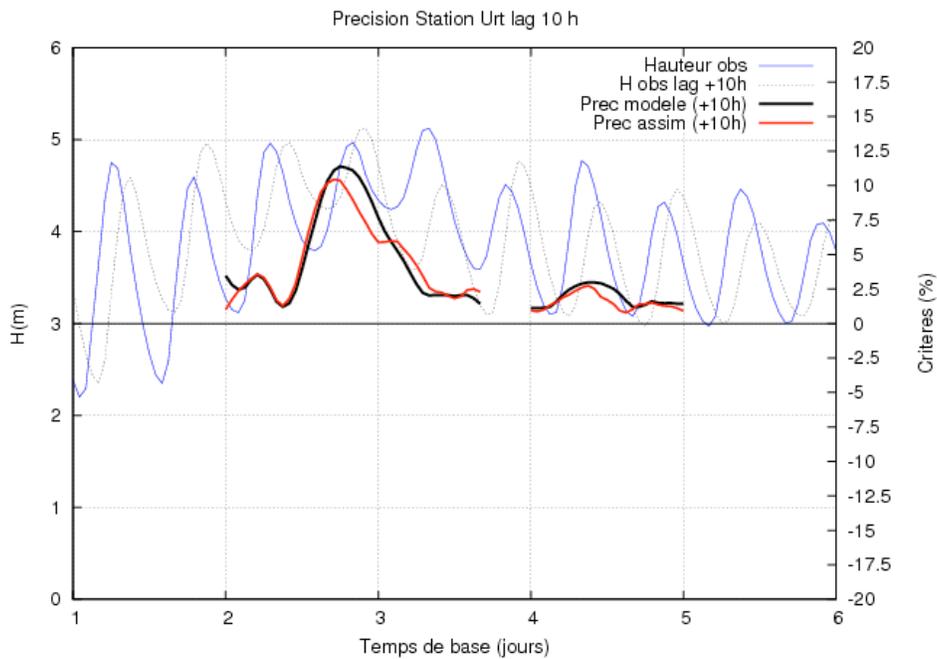
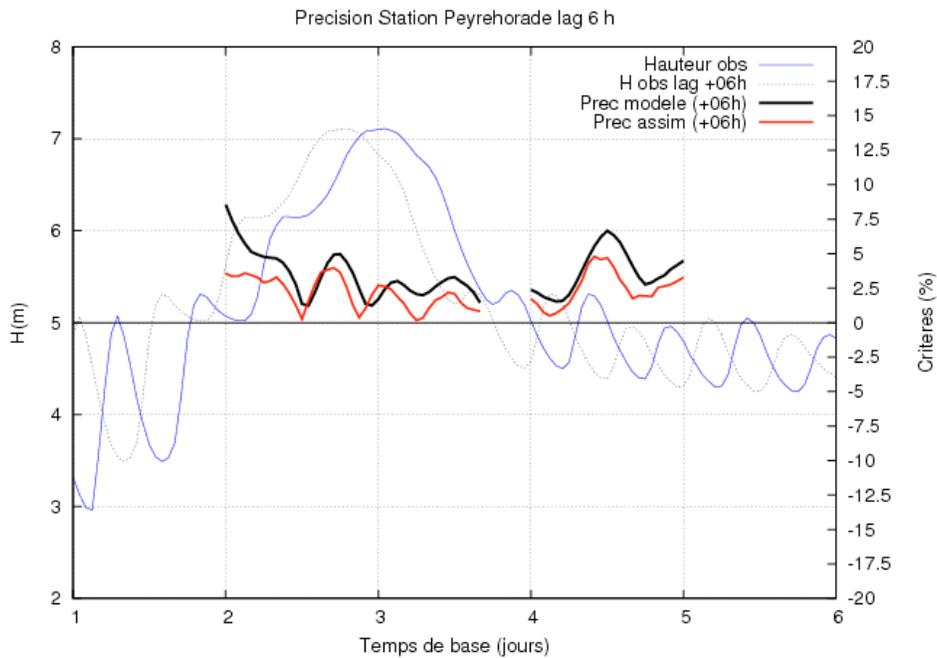
1- Précisions en mode simulation

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



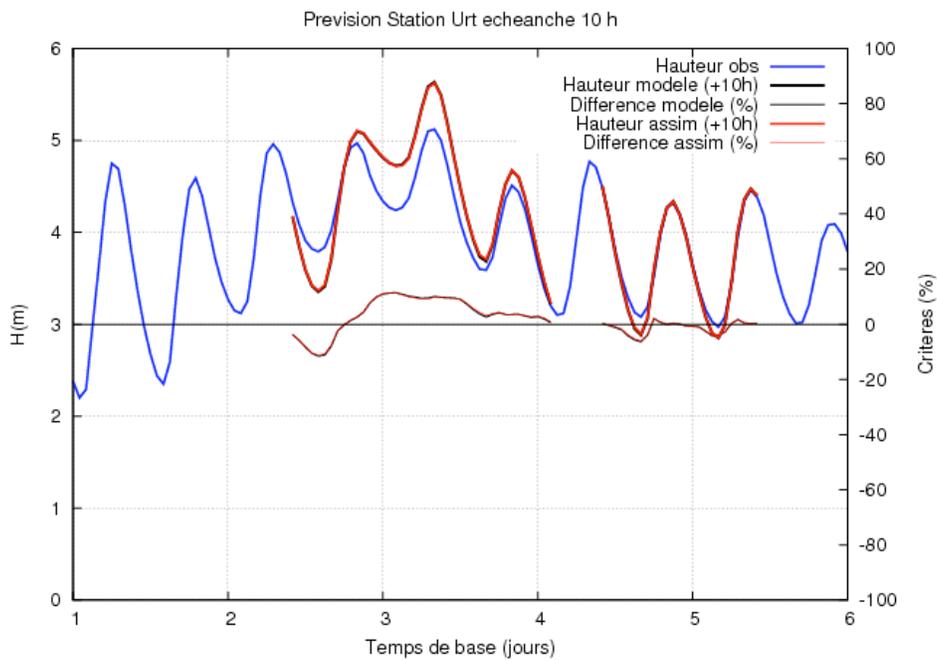
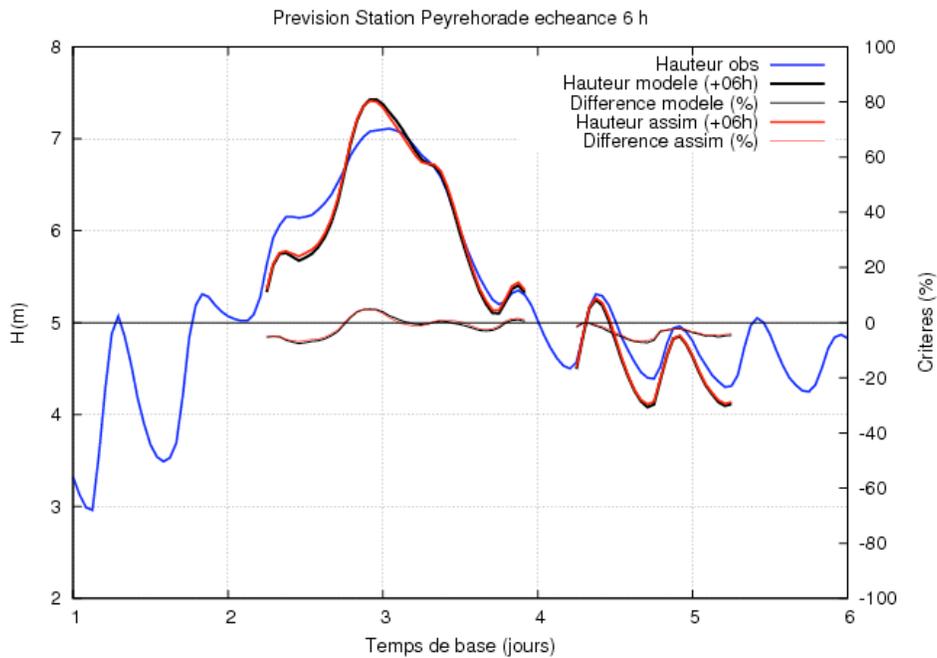
2- Précisions en mode prévision

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.



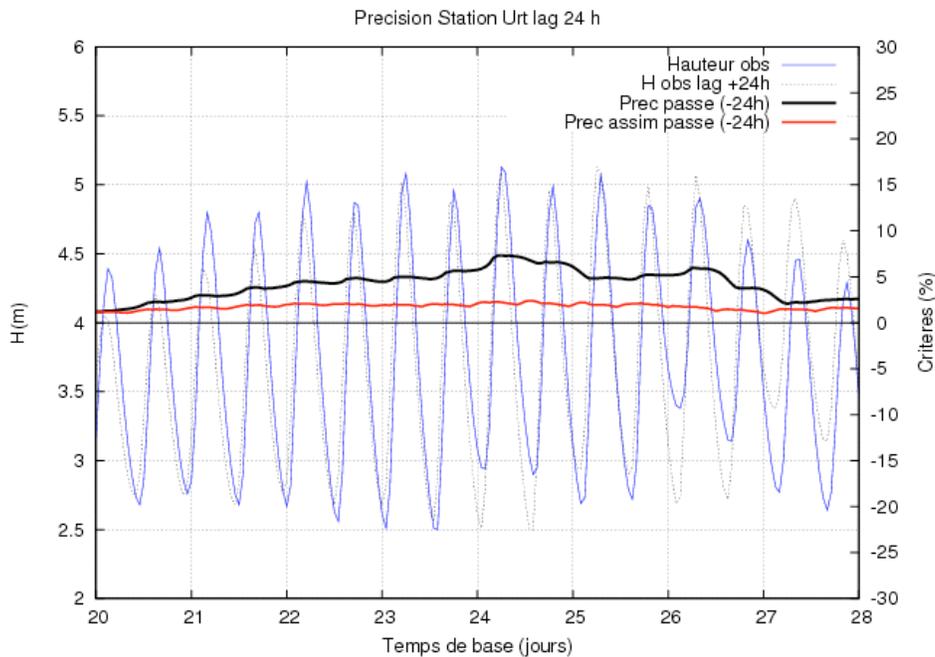
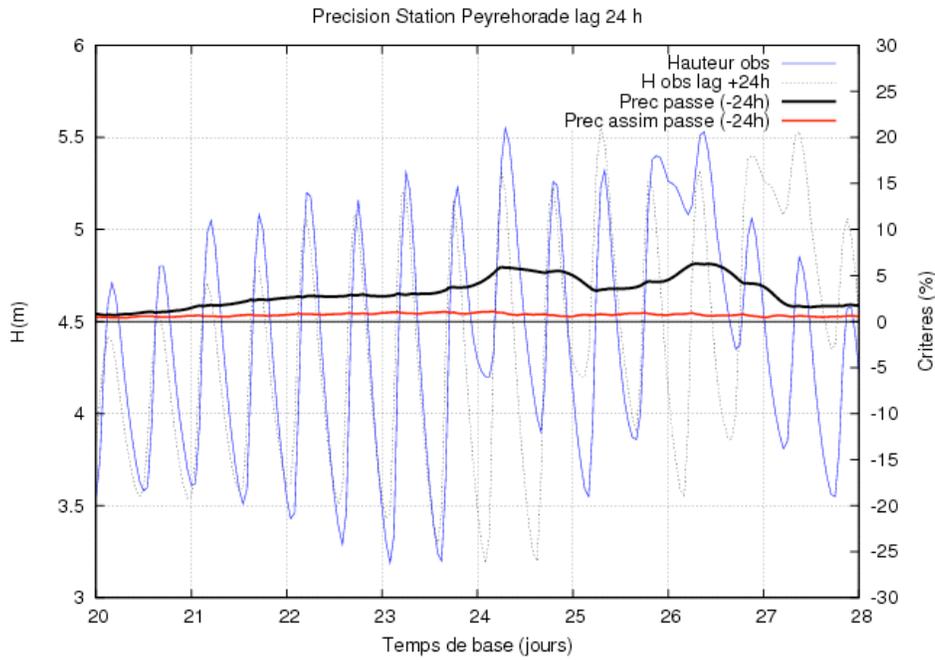
L'assimilation en mode ASSIM a apporté peu d'amélioration au modèle direct.

Episode ADOUR MARITIME 07

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

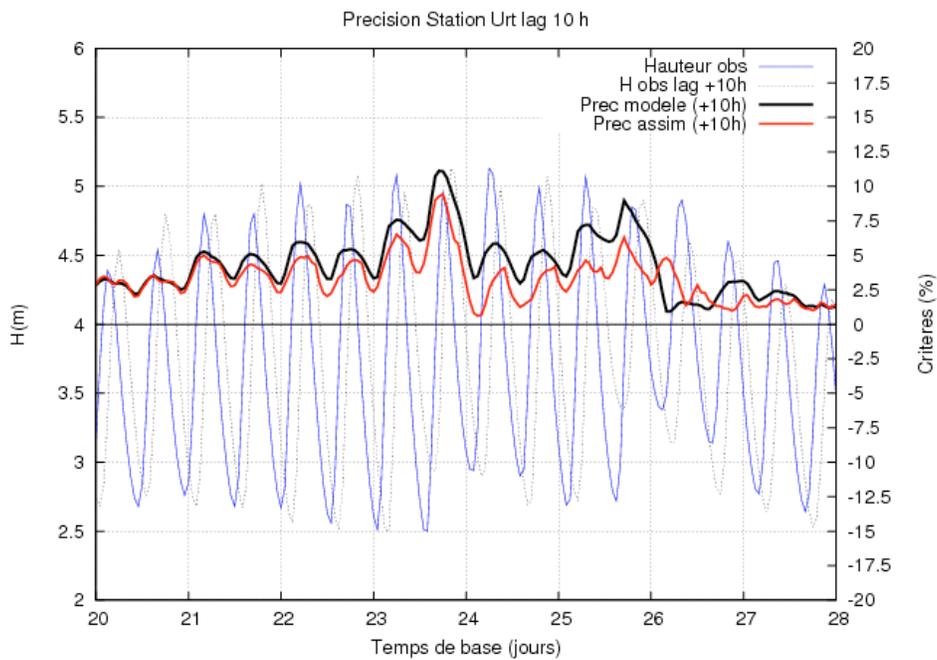
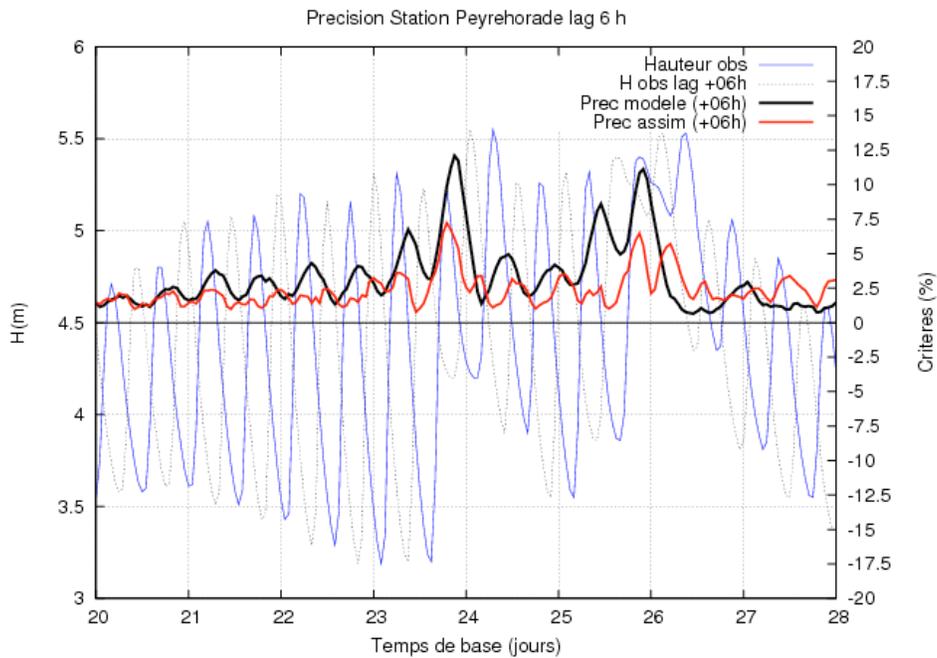
1- Précisions en mode simulation

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



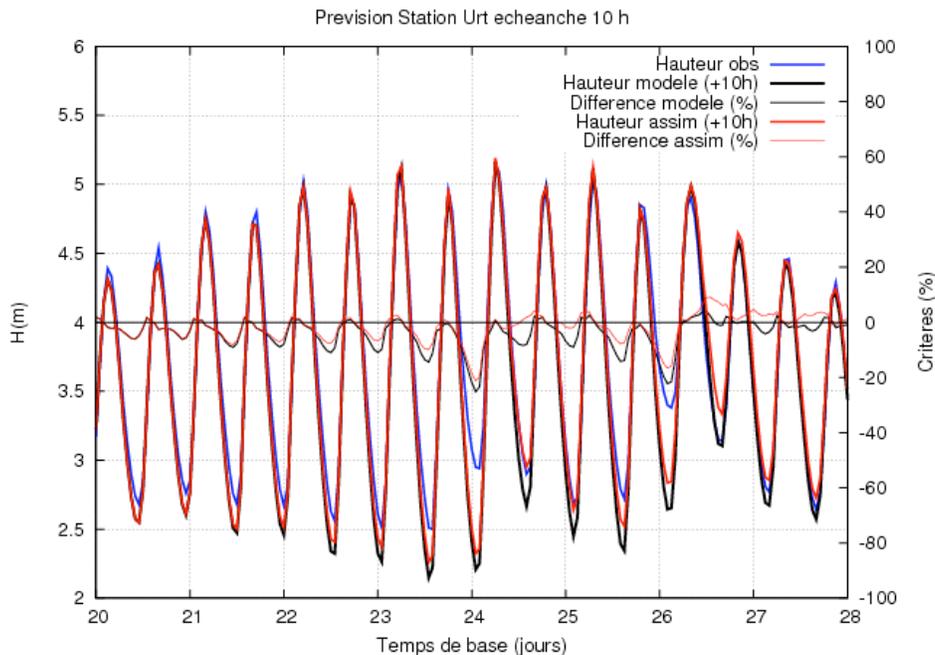
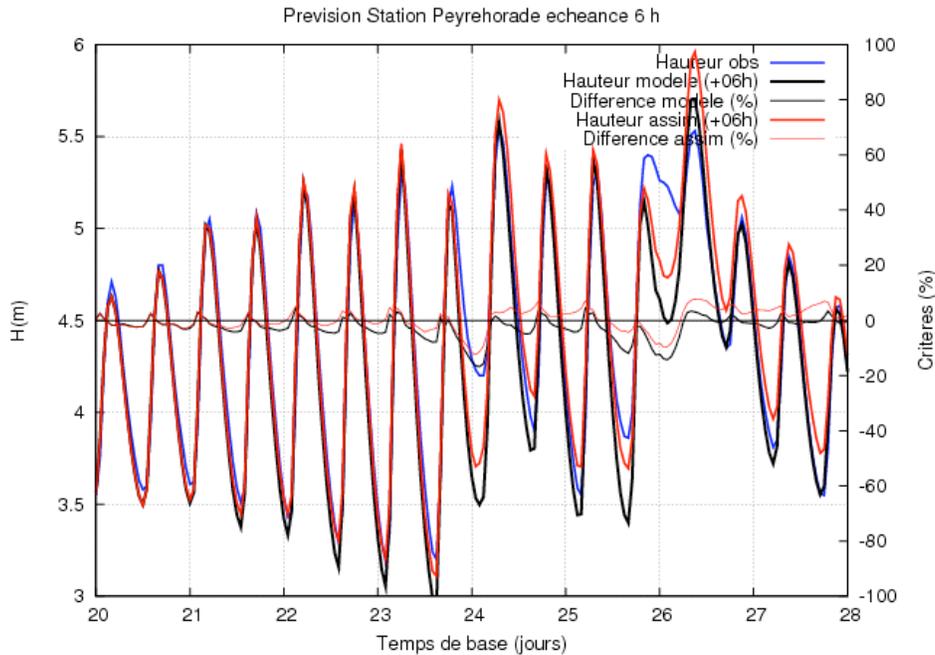
2- Précisions en mode prévision

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.



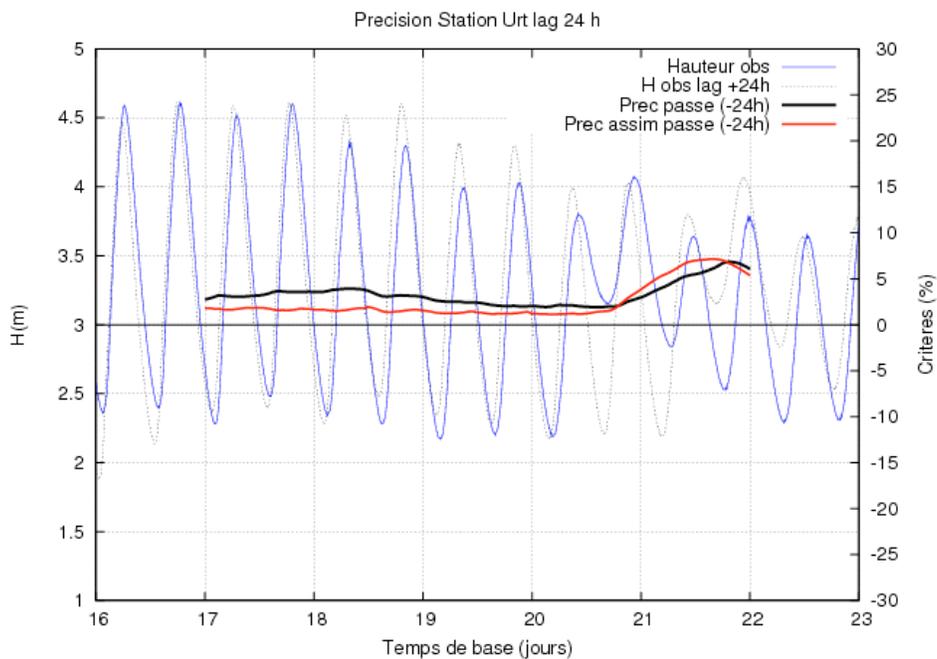
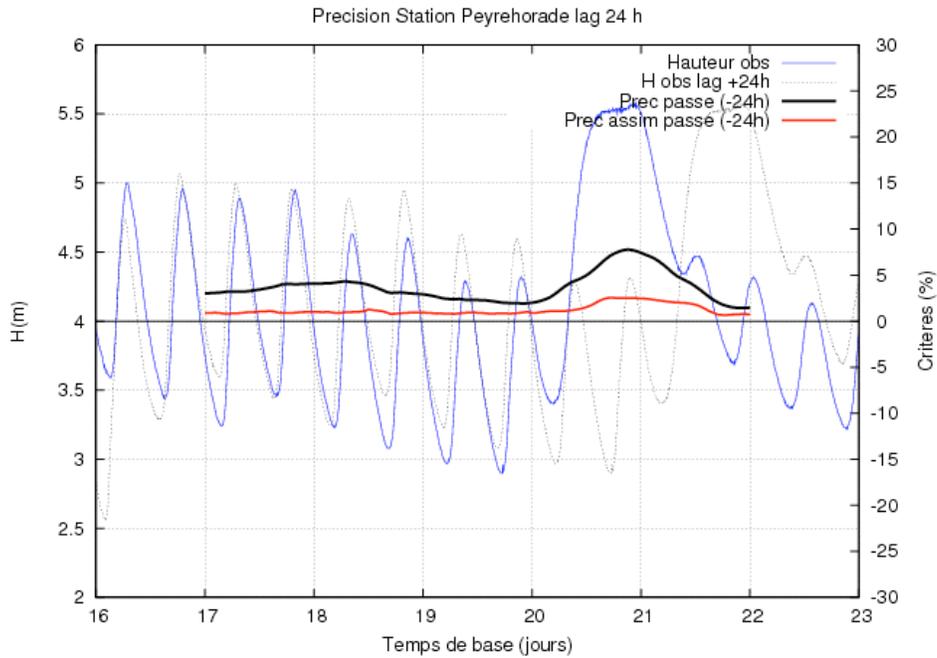
L'assimilation en mode CALAGEASSIM a globalement amélioré la prévision de l'épisode. Le dernier pic est néanmoins surestimé.

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.

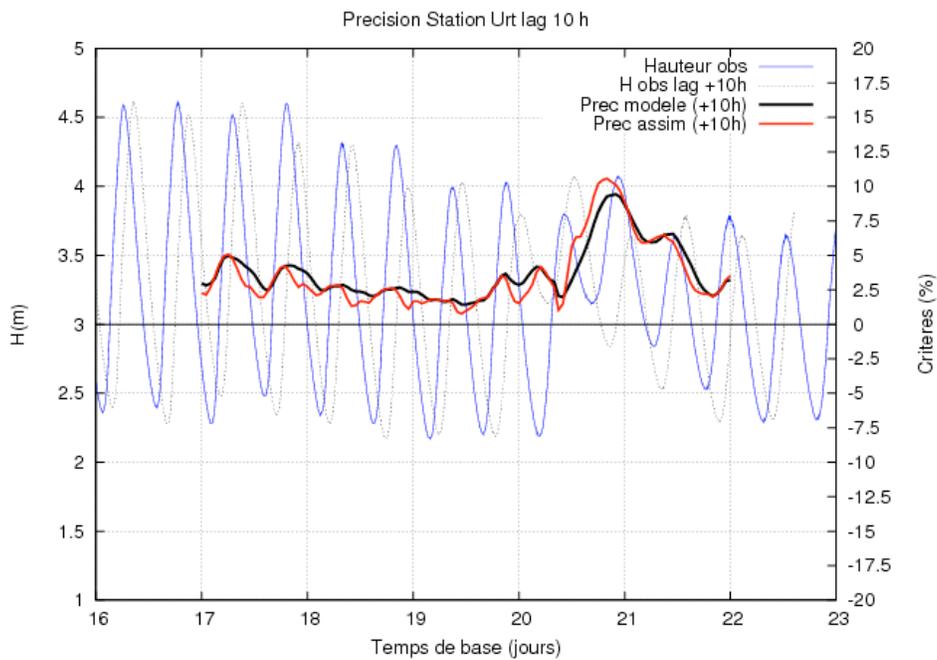
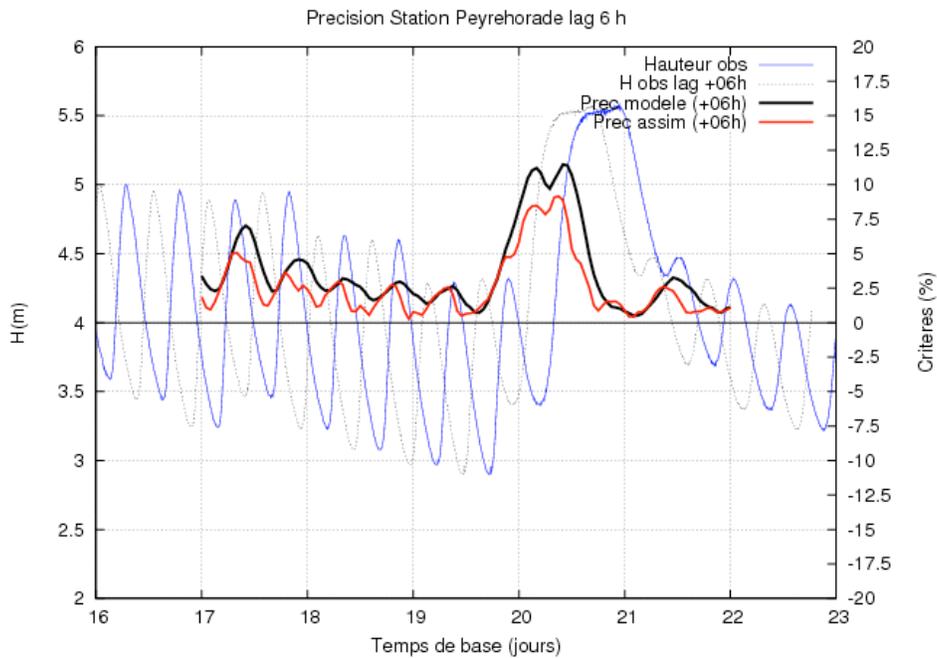
1- Précisions en mode simulation

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



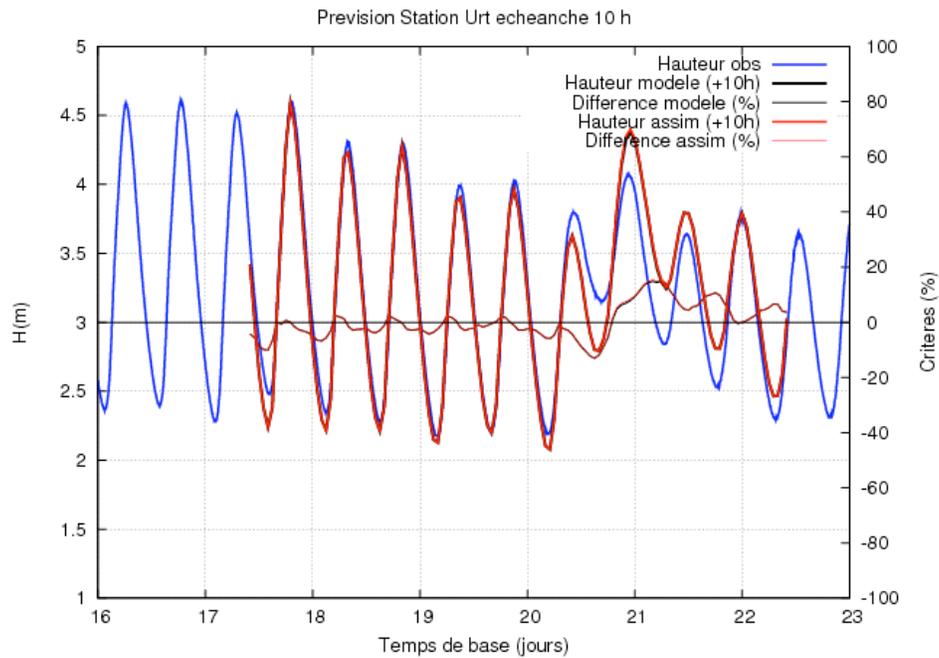
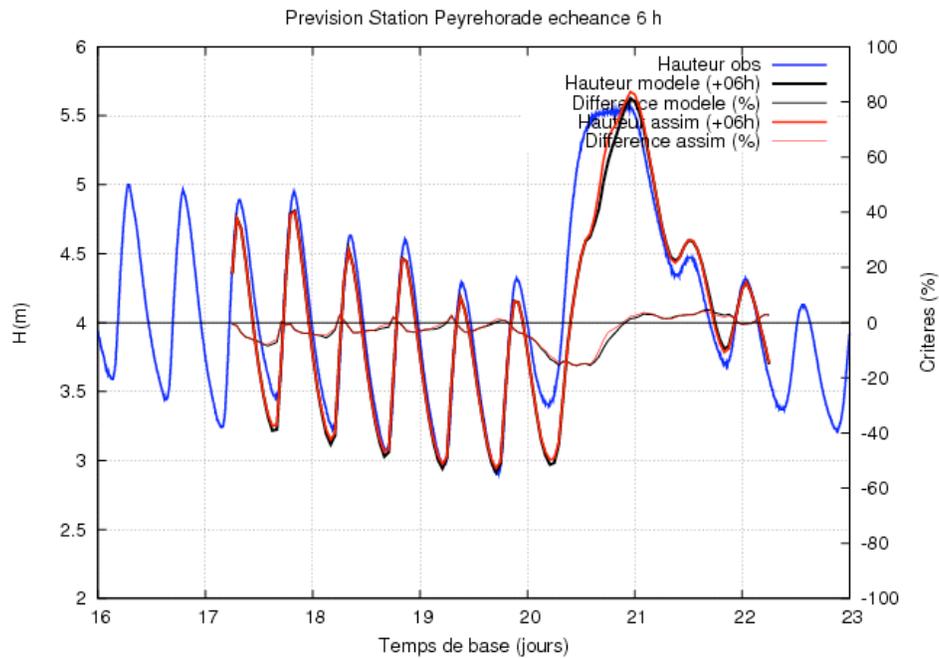
2- Précisions en mode prévision

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.



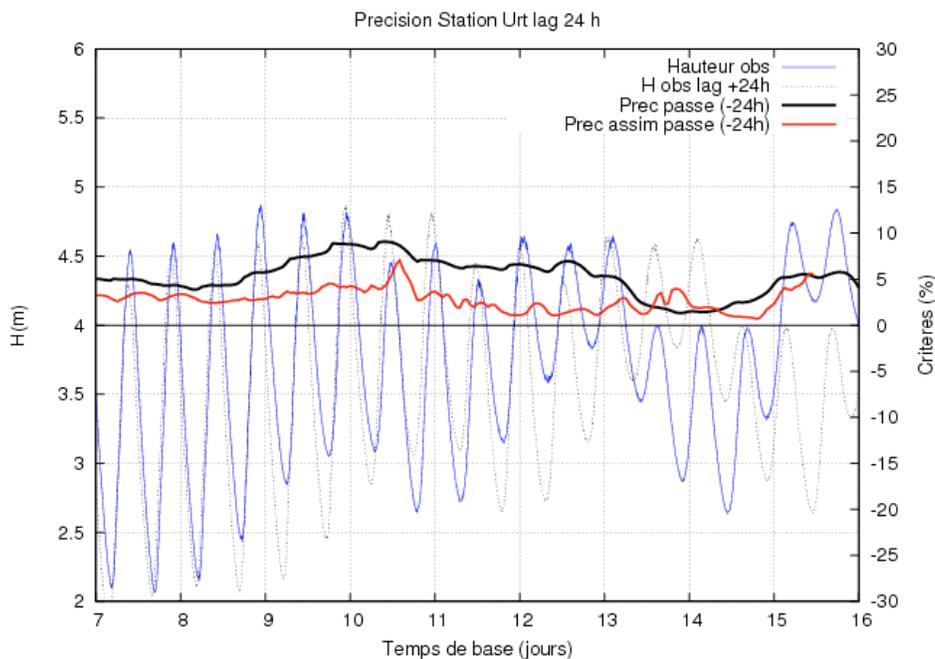
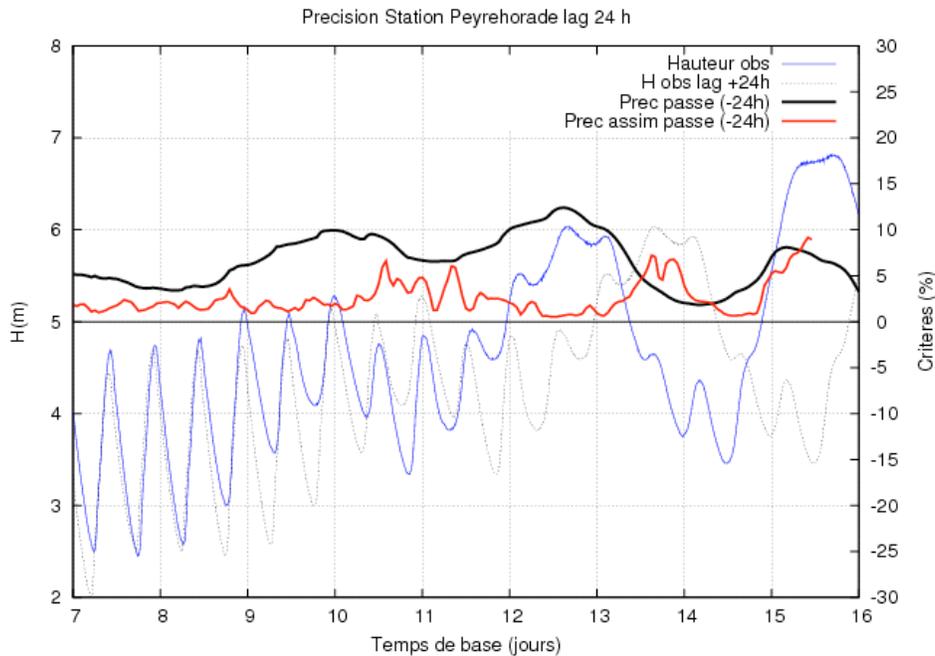
L'assimilation en mode ASSIM a apporté peu d'amélioration au modèle direct.

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 $660\text{m}^3/\text{s}$.

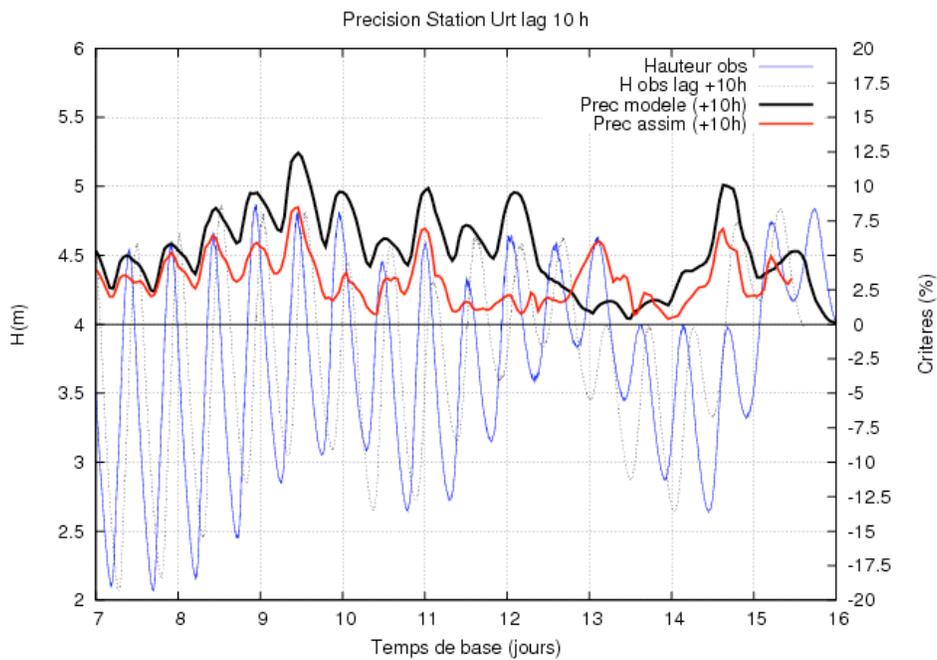
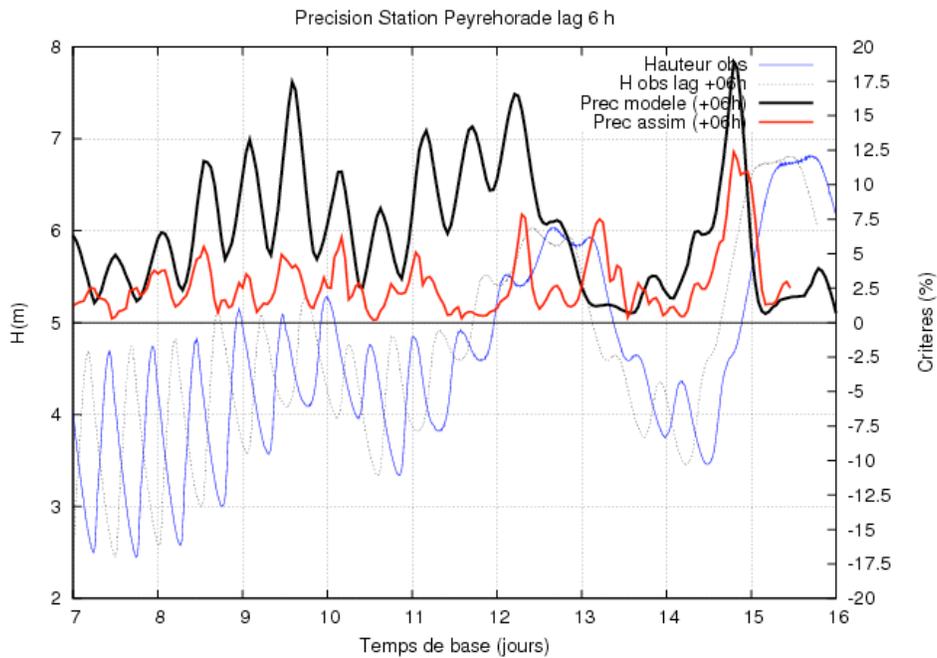
1- Précisions en mode simulation

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



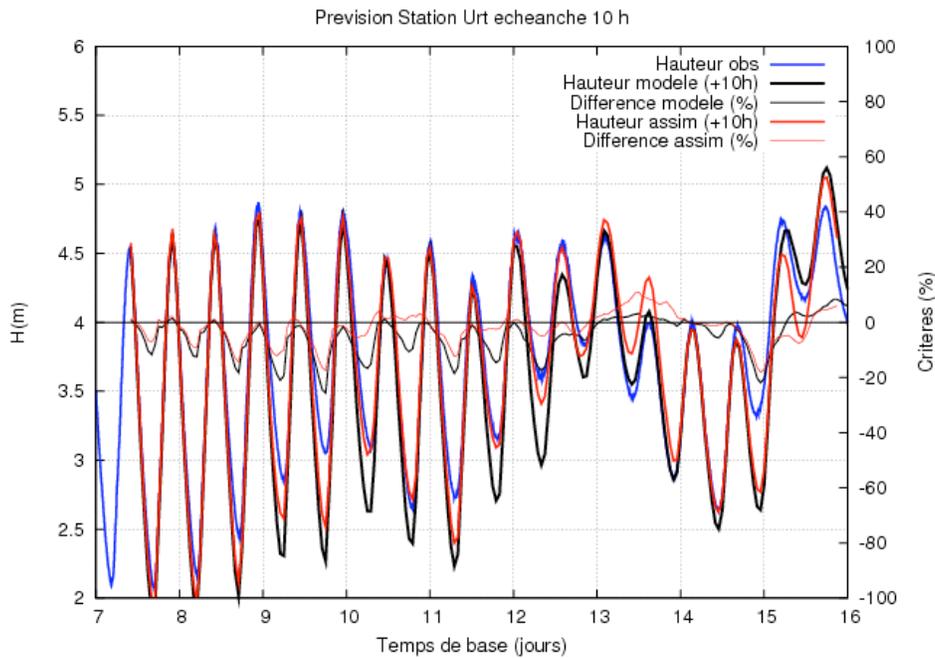
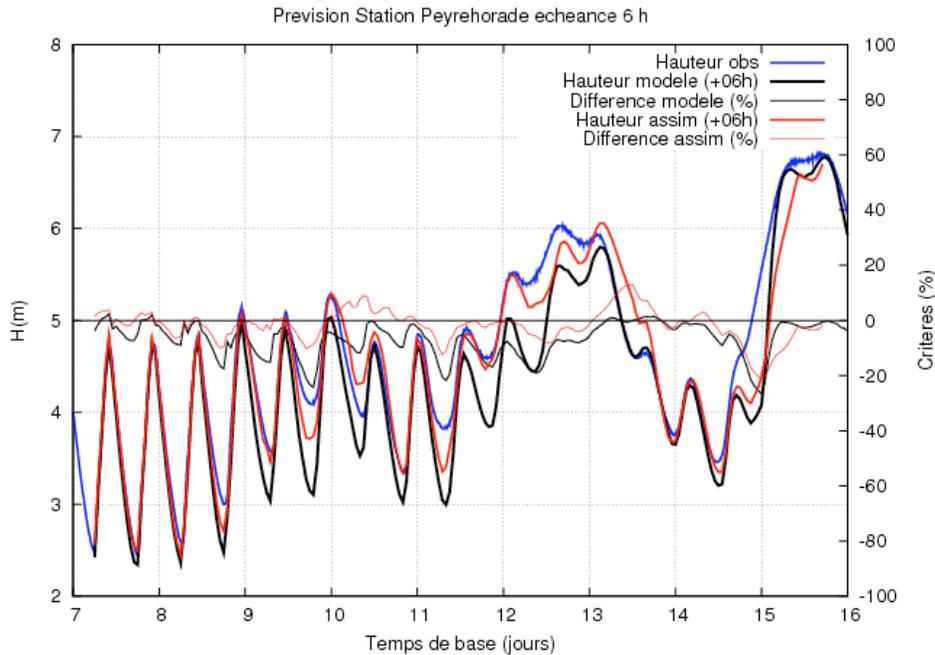
2- Précisions en mode prévision

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.

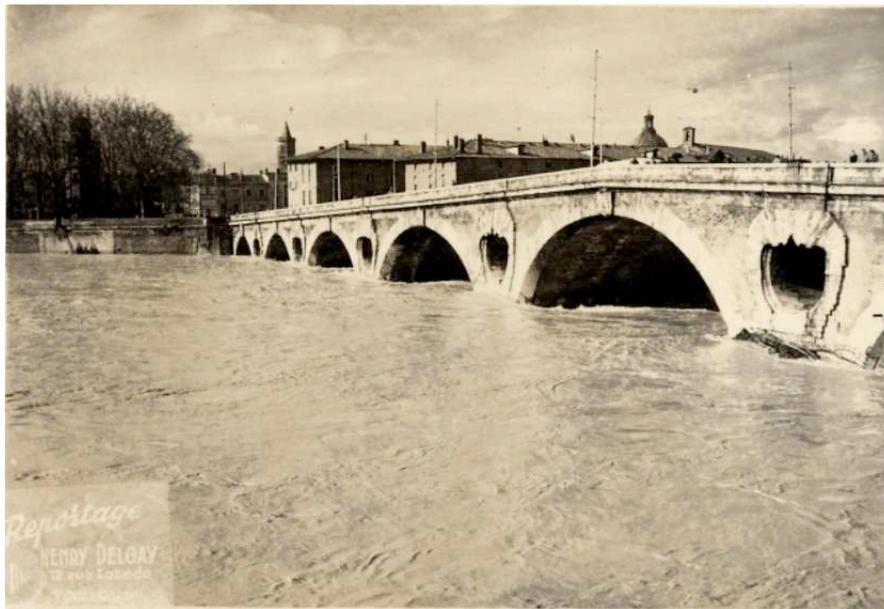
Annexe 4 : Assimilation de données sur un modèle d'ondes de crues diffusantes – Propagation des fonctions de covariance (Rapport de stage ENM 2011).



ÉCOLE NATIONALE
de la MÉTÉOROLOGIE
ENM

Projet modélisation janvier/février 2011 :

Assimilation de données sur un modèle d'ondes de crues



Crédit photo : Henry Delgay

Encadré par Sophie Ricci (CERFACS) et Olivier Pannekoucke (CNRM/GAME)

Etudiants IENM 2009/2012

Delplanque Alain

Guth Jonathan

Raynaud Damien

Résumé

La prévision des crues demeure un domaine incontournable lorsque l'on parle de sécurité des biens et des personnes. La multiplication des sites à risques ou à enjeux économiques et environnementaux (les centrales hydroélectriques ou nucléaires) ainsi que le développement de zones urbaines proches de zones inondables ne font que renforcer l'importance d'une bonne prévision. Des efforts dans le domaine de la modélisation et de l'observation permettent d'améliorer la prévision de la propagation d'une crue et d'en limiter les risques humains et technologiques. En France, c'est le Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations (SCHAPI), rattaché au ministère de l'écologie, du Développement Durable, des Transports et du Logement, en lien avec le Service de Prévision de crues (SPC) qui assure une surveillance des crues ¹. Les enjeux sont également économiques. En effet, une prévision fiable de la hauteur d'eau permet une gestion optimale des ressources en eau notamment pour la production électrique par Électricité de France (EDF). Ce projet s'inscrit dans le cadre de la mise en œuvre de la prévision hydraulique en temps réel. Il s'agit ici de travailler sur l'assimilation des observations de hauteurs d'eau et de débits afin d'améliorer les prévisions.

Lors de l'implémentation d'un algorithme d'assimilation de données, il est important de bien modéliser la matrice de covariance des erreurs de prévisions que l'on ne sait pas estimer exactement. En effet, ce sont les covariances d'erreurs d'ébauche qui déterminent l'amplitude et la forme de l'incrément d'analyse. A l'aide d'un modèle physique simple codé à partir de l'équation d'ondes de crues diffusantes on cherche à déterminer les caractéristiques de la matrice des erreurs de prévision.

¹<http://www.vigicrues.ecologie.gouv.fr/>

Table des matières

1	L'assimilation de données	5
1.1	Généralités	5
1.2	Méthode d'assimilation BLUE (Best Linear Unbiased Estimator)	6
1.3	Rôle de B	7
1.4	Diagnostic de la longueur de portée	9
1.5	Filtre de Kalman	10
1.6	Filtre de Kalman d'ensemble	10
2	Modèle d'onde de crue	13
2.1	Physique du problème	13
2.2	Schéma numérique de l'équation d'advection-diffusion	15
2.2.1	Schéma de discrétisation	15
2.2.2	Stabilité du schéma	16
2.2.3	Équation équivalente	16
2.2.4	Impact des termes diffusifs et dispersifs sur la propagation des covariances	17
2.3	Construction du forçage amont	18
3	Dynamique des covariances : évolution des longueurs de portée	21
3.1	Dynamique des covariances sans observations	21
3.1.1	Cas de l'advection pure	22
3.1.2	Cas de l'advection diffusion	23
3.1.3	Bilan de la dynamique des longueurs de portée sans assimilation	24
3.2	Dynamique des covariances avec observation	25
4	Expériences numériques	27
4.1	Diagnostic de B et des longueurs de portée sans assimilation	27
4.1.1	Calcul de la matrice de covariances d'erreurs d'ébauche	27
4.1.2	Calcul des longueurs de portée	28
4.2	Résultats des expériences sans assimilation	28
4.2.1	Modèle avec advection seule	28
4.2.2	Modèle avec advection et diffusion	29
4.3	Assimilation avec un BLUE	30
4.4	Implémentation du filtre de Kalman	33
4.4.1	Cadre de la simulation	33
4.4.2	Évolution des fonctions de corrélation	36
4.4.3	Méthode d'estimation des longueurs de portée	36
4.4.4	Diagnostic des longueurs de portée	38
4.4.5	Influence du quotient σ_b/σ_o	39
4.4.6	Influence du temps caractéristique du signal	40

TABLE DES MATIÈRES

4.5	Implémentation filtre de Kalman d'ensemble	40
A	Annexe	44
A.1	Condition de stabilité du schéma numérique	44
A.2	Influence du schéma numérique	45

Introduction

La prévision de phénomènes naturels complexes nécessite une modélisation mathématique précise. Les incertitudes sur les modèles numériques proviennent des conditions aux limites (état initial, forçages extérieurs) et des simplifications et paramétrisations de la physique (paramètres numériques, géométriques, physiques). En résolvant un problème inverse, l'algorithme d'assimilation de données combine les informations issues de l'observation et de la simulation pour estimer des paramètres du modèle ou corriger l'état initial. Ainsi l'assimilation de données permet de réduire l'incertitude sur la simulation et sur la prévision. Cette technique est communément utilisée en météorologie, océanographie, et tend à être utilisée dans d'autres domaines des géosciences ou même dans des domaines plus industriels tels que la neutronique.

Les phénomènes hydrauliques font partie des phénomènes naturels complexes à prévoir. Le Centre d'Etudes Techniques Maritimes et Fluviales (CETMEF) et le Laboratoire national d'Hydraulique et Environnement (LNHE, service R&D d'EDF) ont mis au point un code numérique visant à modéliser ces phénomènes hydrauliques : MASCARET. Ce code numérique 1D, basé sur les équations de Saint-Venant, est capable de représenter une large gamme de phénomènes hydrauliques (ondes de crues, ondes de marées, etc). Cette étude se place dans le cadre d'une collaboration entre le Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS), le SCHAPI, le CETMEF et le LNHE. Cette collaboration vise à construire un système d'assimilation pour le modèle MASCARET [Ricci et al., 2010].

Une difficulté rencontrée lors de la mise en place de la plupart des systèmes d'assimilation de données est de spécifier les statistiques des erreurs associées à une prévision : variance et covariances des erreurs. L'estimation directe des statistiques de l'erreur de prévision est impossible et la propagation explicite des covariances par la dynamique d'un modèle est généralement difficile à mettre en œuvre. C'est le cas pour le modèle MASCARET. Une manière efficace mais coûteuse en temps de calcul est d'utiliser un filtre de Kalman. Dans le cadre de cette étude, une hypothèse forte est alors émise. Les processus opérés par le modèle sont considérés comme identiques ou proches de ceux faits par un modèle plus simple. Ce modèle simplifié est basé sur l'équation de propagation des ondes de crues diffusantes. Le coût de calcul d'un filtre de Kalman devient alors plus raisonnable. Dans le travail proposé ici, on utilise donc cet algorithme d'assimilation sur le modèle d'onde de crues pour estimer une forme paramétrée des fonctions de covariance spatiale. Cette paramétrisation est alors utilisée pour modéliser la matrice de covariance d'erreur de prévision dans un algorithme de type BLUE.

Le premier chapitre présente un rappel sur l'assimilation de données. Le chapitre 2 présente le modèle numérique utilisé. Puis les chapitres 3 et 4 rapportent les résultats de l'étude. Le chapitre 3 correspond à l'étude théorique. Le chapitre 4 s'intéresse aux résultats numériques. Les conclusions et perspectives sont au chapitre 5.

Chapitre 1

L'assimilation de données

1.1 Généralités

L'assimilation de données est un processus mathématique qui permet d'obtenir la représentation la plus probable de l'état d'un système à partir de toutes les sources d'informations disponibles. En pratique, sous l'hypothèse de Gaussiannité, il s'agit de trouver la meilleure combinaison entre un état à priori du système que l'on appelle l'ébauche et les observations. L'utilisation de l'ébauche permet de palier le problème de sous-détermination dans le cas où le nombre d'observations est inférieur à la dimension du problème. Le critère d'optimalité pour obtenir le meilleur estimateur statistique de cette combinaison est la détermination du minimum de variance d'erreur d'estimation.

Les informations issues de l'ébauche et des observations sont entachées d'erreur. La combinaison des deux par le biais de divers algorithmes d'assimilation (BLUE, filtre de Kalman...) permet de se rapprocher de l'état réel de notre système. L'état vrai du système, généralement inconnu, est représenté par le vecteur \mathbf{X}^t . L'ébauche est elle notée \mathbf{X}^b . Ces deux vecteurs sont de dimension n . \mathbf{Y}^o désigne le vecteur d'observations, il est de dimension p . Le processus d'assimilation de données formule l'analyse \mathbf{X}^a de dimension n . L'erreur d'ébauche est définie par le vecteur ε^b telle que $\mathbf{X}^b = \mathbf{X}^t + \varepsilon^b$.

Les observations diffèrent également de l'état réel. L'erreur d'observation est définie par ε^o où $\mathbf{Y}^o = \mathcal{H}\mathbf{X}^t + \varepsilon^o$, avec \mathcal{H} est l'opérateur permettant de passer de l'espace du modèle à celui des observations. Par exemple si \mathbf{Y}^o contient des mesures de débit et que les observations ne sont pas faites en un point de la discrétisation de la rivière, \mathcal{H} correspond à une interpolation spatiale du vecteur d'état vers les points d'observations. Cet opérateur peut être non-linéaire. Par exemple en météorologie, l'assimilation des mesures de radiances satellitaires implique un opérateur \mathcal{H} non linéaire lié à l'équation de transfert radiatif.

Toutes les erreurs, d'observations ou d'ébauche, sont considérées comme non-biaisées. Les matrices de covariances d'erreurs sont notées $\mathbf{B} = \mathbb{E} [\varepsilon^b(\varepsilon^b)^T]$ pour l'ébauche et $\mathbf{R} = \mathbb{E} [\varepsilon^o(\varepsilon^o)^T]$ pour les observations. Par construction, ces matrices sont symétriques, définies et positives. Elles permettent de quantifier la confiance que l'on a en l'ébauche et les observations. Enfin on considère qu'il n'y a pas de corrélation entre erreurs d'ébauche et erreurs d'observation.

$$\mathbb{E} [\varepsilon^o(\varepsilon^b)^T] = \mathbf{0}_{n \times p}$$

Dans le cadre de cette étude, les vecteurs \mathbf{X}^t , \mathbf{X}^b et \mathbf{X}^a contiennent respectivement les hauteurs d'eau réelles (connues dans la cadre d'expériences jumelles ¹), simulées, et analysées. \mathbf{Y}^o

¹Simulations numériques où l'état vrai et l'ébauche sont connus et synthétiques.

1.2. MÉTHODE D'ASSIMILATION BLUE (BEST LINEAR UNBIASED ESTIMATOR)

représente les mesures faites le long de la rivière. Il s'agit également de hauteurs d'eau. Le lien entre vecteur d'état et observations est simple ; il ne s'agit que d'une sélection de points. L'opérateur \mathcal{H} est linéaire et sera à présent noté $\mathbf{H} = (0 \ \cdots \ 0 \ 1 \ 0 \ \cdots \ 0)$. Enfin dans des cas complexes faisant intervenir différentes mesures faites par un même instrument (satellites...), les erreurs d'observations sont corrélées en espace et en temps. Si, comme dans le cadre de ce projet, il y a indépendance entre erreurs d'observations, alors la matrice \mathbf{R} est diagonale.

1.2 Méthode d'assimilation BLUE (Best Linear Unbiased Estimator)

Dans le problème traité, on recherche le meilleur estimateur linéaire non biaisé de l'état vrai du système comme combinaison linéaire de l'ébauche et des observations. Pour construire un estimateur linéaire, deux matrices \mathbf{L} et \mathbf{K} de tailles respectives $n \times n$ et $n \times p$ doivent être déterminées. Elles sont telles que :

$$\mathbf{X}^a = \mathbf{L}\mathbf{X}^b + \mathbf{K}\mathbf{Y}^o.$$

Comme l'estimateur linéaire recherché est non biaisé, cela introduit des contraintes sur \mathbf{L} et \mathbf{K} :

$$\begin{aligned} \mathbb{E}[\varepsilon^a] &= \mathbb{E}[(\mathbf{X}^a - \mathbf{X}^t)] \\ &= \mathbb{E}[\mathbf{L}\mathbf{X}^t + \mathbf{L}\varepsilon^b + \mathbf{K}\mathbf{H}\mathbf{X}^t + \mathbf{K}\varepsilon^o - \mathbf{X}^t] \\ &= 0. \end{aligned} \tag{1.1}$$

De plus les erreurs d'observation et d'ébauche sont supposées non biaisées, l'équation 1.1 devient alors

$$\begin{aligned} (\mathbf{L} - \mathbf{I}_{n \times n})\mathbb{E}[\mathbf{X}^t] + \mathbf{K}\mathbf{H}\mathbb{E}[\mathbf{X}^t] &= (\mathbf{L} - \mathbf{I}_{n \times n} + \mathbf{K}\mathbf{H})\mathbb{E}[\mathbf{X}^t] \\ &= 0. \end{aligned} \tag{1.2}$$

La condition pour que l'estimateur \mathbf{X}^a soit non biaisé devient donc :

$$\mathbf{L} + \mathbf{K}\mathbf{H} = \mathbf{I}_{n \times n}.$$

L'estimateur optimal, minimise la trace de la matrice de covariance d'erreurs d'analyse $\mathbf{A} = E(\varepsilon^a \varepsilon^{aT})$. Elle s'exprime [Bouttier and Courtier, 1999] :

$$\mathbf{A} = (\mathbf{I}_{n \times n} - \mathbf{K}\mathbf{H})\mathbf{B}(\mathbf{I}_{n \times n} - \mathbf{K}\mathbf{H})^T + \mathbf{K}\mathbf{R}\mathbf{K}^T.$$

La minimisation permet d'établir une expression de la matrice \mathbf{K} :

$$\min(\text{Tr}(\mathbf{A})) \implies \mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}.$$

Grâce aux conditions trouvées à partir des deux hypothèses du BLUE, \mathbf{X}^a s'écrit :

1.3. RÔLE DE B

$$\begin{aligned}\mathbf{X}^a &= \mathbf{X}^b + \mathbf{K}(\mathbf{Y}^o - \mathbf{H}\mathbf{X}^b) \\ &= \mathbf{X}^b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}(\mathbf{Y}^o - \mathbf{H}\mathbf{X}^b).\end{aligned}\tag{1.3}$$

La matrice \mathbf{K} , appelée matrice de gain, permet de définir l'état assimilé comme la somme de l'ébauche et de l'incrément d'analyse

$$\delta\mathbf{X}^a = \mathbf{K}(\mathbf{Y}^o - \mathbf{H}\mathbf{X}^b) = \mathbf{K}d,\tag{1.4}$$

avec d le vecteur d'innovation.

1.3 Rôle de B

\mathbf{B} représente les statistiques sur l'incertitude sur l'état du système prévu à un instant donné. L'équation (1.4) de $\delta\mathbf{X}^a$ met en évidence le rôle primordial de la matrice de covariances d'erreurs d'ébauche :

$$\delta\mathbf{X}^a = \underbrace{\mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}}_{\delta\hat{\mathbf{x}}}(\mathbf{Y}^o - \mathbf{H}\mathbf{X}^b) = \mathbf{B}\delta\hat{\mathbf{X}}.$$

Le vecteur $\delta\hat{\mathbf{X}}$ contient des 0 sauf aux points où sont réalisées des observations. C'est la seule matrice \mathbf{B} qui répartit l'information contenue dans ce vecteur sur tout le domaine. Dans le cas multivarié ² elle permet également de corriger l'ensemble des variables du vecteur de contrôle. \mathbf{B} contient sur sa diagonale les variances d'erreurs d'ébauche à chaque point de grille. Les termes non-diagonaux représentent les covariances d'erreurs entre les différents points. En raisonnant sur la k^{ieme} ligne (ou colonne) de la matrice \mathbf{B} on observe la fonction de covariance du k^{ieme} point du domaine.

Plus concrètement, en considérant une assimilation faite au point d'observation, la variance détermine quel "poids" donner à l'observation face à la hauteur d'eau simulée par l'ébauche. Les covariances permettent elles de répartir spatialement l'information du point d'observation vers les autres points de grille.

La figure 1.1 montre, à partir d'une ébauche (a), et de deux fonctions de covariances (une "étroite" (b) et l'autre plus "large" (c)) la correction faite sur l'analyse. Les états analysés (d) et (e) montrent que plus la fonction de covariance est "étalée" plus la correction apportée au point d'observation se répercute sur des points éloignés. L'étalement des fonctions de covariance peut être caractérisé par une longueur appelée longueur de portée définie au paragraphe 1.4. La figure 1.2.(a) montre une fonction de covariance gaussienne. Celle-ci étant isotrope, l'information apportée par une assimilation au k^{ieme} point sera répartie de manière identique de part et d'autre du point. Il peut parfois être judicieux de rendre ces fonctions de covariance anisotropes. C'est ce qu'illustre la figure 1.2.(b). En confrontant ce profil de fonction de covariance à ce qui a été dit sur la figure 1.1, on s'attend à ce que la correction faite sur l'analyse se répercute plus avant le point d'observation qu'après. La caractérisation de l'anisotropie se fait à partir de la définition de deux longueurs de portée dites amont et aval.

²Le vecteur de contrôle comporte des variables de différentes natures.

1.3. RÔLE DE B

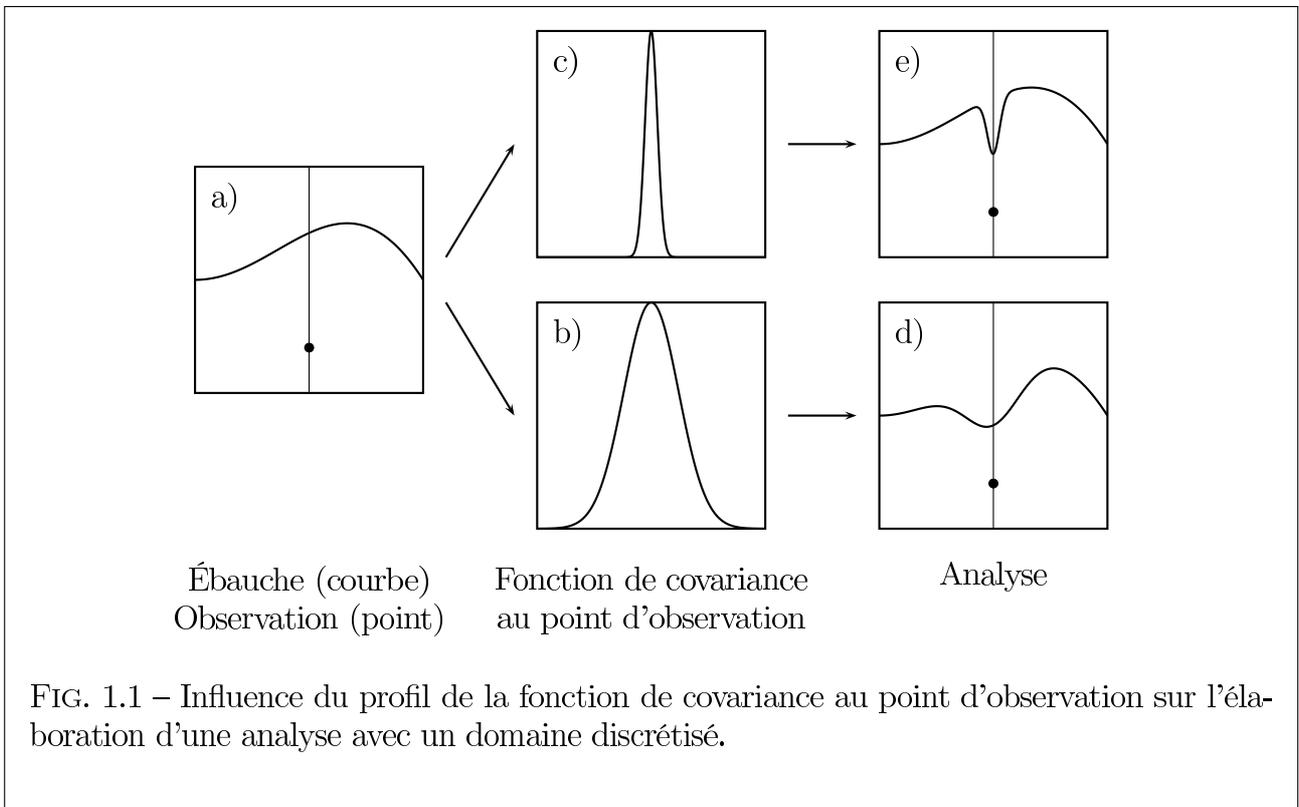


FIG. 1.1 – Influence du profil de la fonction de covariance au point d'observation sur l'élaboration d'une analyse avec un domaine discrétisé.

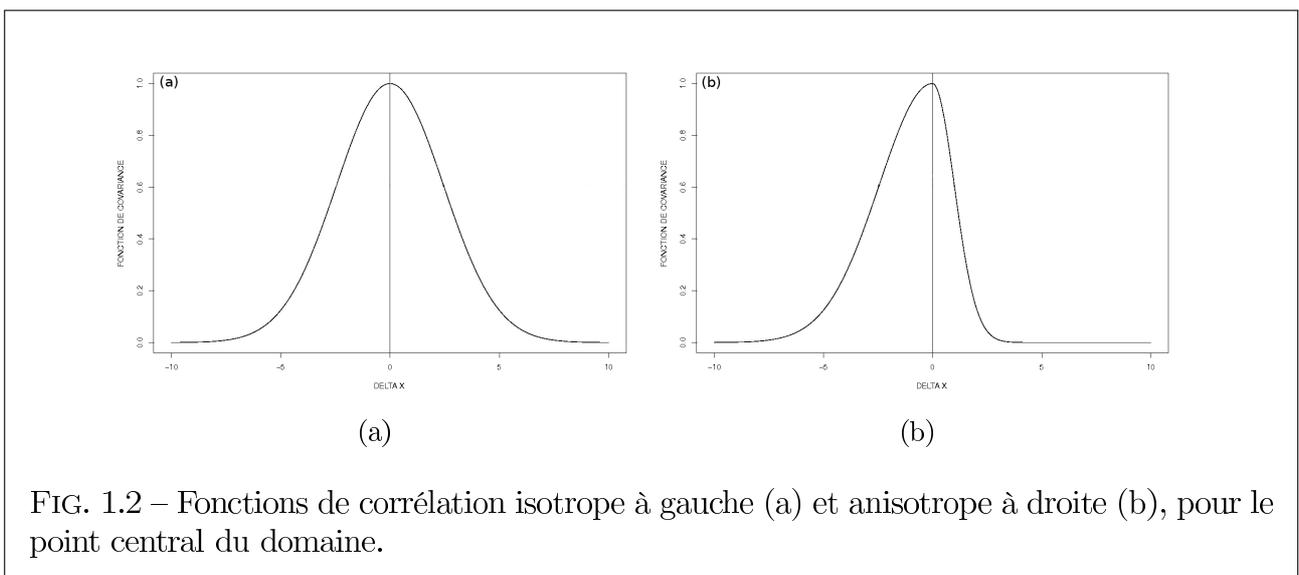


FIG. 1.2 – Fonctions de corrélation isotrope à gauche (a) et anisotrope à droite (b), pour le point central du domaine.

1.4 Diagnostic de la longueur de portée

Une façon d'estimer la distance de corrélation significative est d'introduire le diagnostic de la longueur de portée. Elle caractérise la corrélation d'une erreur en un point avec les erreurs aux points voisins. Daley [Daley, 1991] en donne une formule générale pour les problèmes à une dimension :

$$L_p = \sqrt{\frac{-1}{\nabla^2 \rho(0)}} \iff L_p = \sqrt{\frac{-1}{\frac{d^2 \rho}{dx^2}(0)}} \iff \frac{d^2 \rho}{dx^2}(0) = \frac{-1}{L_p^2}. \quad (1.5)$$

En effectuant un développement de Taylor à l'ordre 2 en δx d'une fonction de corrélation $\rho(x)$ on obtient

$$\rho(\delta x) \approx \rho(0) + \frac{\delta x^2}{2} \frac{d^2 \rho(0)}{dx^2}. \quad (1.6)$$

L'expression (1.5) reportée dans (1.6) donne alors, en notant que $\rho(0) = 1$, car ρ est une fonction de corrélation :

$$\rho(\delta x) \approx 1 - \frac{\delta x^2}{2L_p^2}.$$

D'où l'expression de L_p :

$$L_p = \frac{\delta x}{\sqrt{2(1 - \rho(\delta x))}}.$$

La figure 1.3 montre en trait plein une demi-fonction de corrélation gaussienne. La courbe en pointillés correspond au développement de Taylor à l'ordre 2 de cette fonction. Cette courbe, appelée parabole osculatrice à l'origine permet de définir la longueur de portée. Cette dernière correspond à la demi-largeur de la parabole lorsque sa hauteur est égale à 0,5.

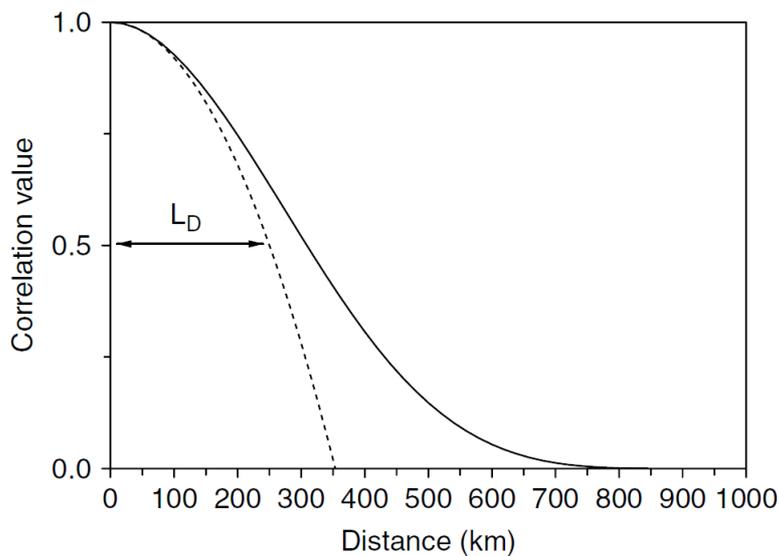


FIG. 1.3 – Illustration graphique de la longueur de portée d'une fonction de corrélation (courbe noire) pour laquelle la parabole osculatrice à l'origine (courbe en tireté) vaut 0.5 [Pannekoucke et al., 2007].

1.5 Filtre de Kalman

Dans le filtre de Kalman le calcul de l'analyse est similaire à celui fait dans l'algorithme du BLUE. Entre autre, il fait également intervenir la notion d'incrément d'analyse

$$\mathbf{d}_i = \mathbf{Y}_i^o - \mathbf{H}_i \mathbf{X}_i^b, \quad (1.7)$$

$$\mathbf{K}_i = \mathbf{B}_i \mathbf{H}_i^T (\mathbf{H}_i \mathbf{B}_i \mathbf{H}_i^T + \mathbf{R}_i)^{-1}, \quad (1.8)$$

$$\delta \mathbf{X}_i^a = \mathbf{K}_i (\mathbf{Y}_i^o - \mathbf{H}_i \mathbf{X}_i^b) = \mathbf{K}_i \mathbf{d}_i. \quad (1.9)$$

En revanche on remarque une différence importante dans la suite des algorithmes : le code du BLUE paramètre la matrice \mathbf{B} au début du processus. Celle ci est statique durant toute la phase d'assimilation (Par. 1.2). Le filtre de Kalman permet lui, l'évolution à chaque cycle d'assimilation de l'ensemble des matrices intervenant dans l'algorithme. Ceci est vrai en particulier pour \mathbf{B} :

$$\mathbf{A}_i = (\mathbf{I} - \mathbf{K}_i \mathbf{H}_i) \mathbf{B}_i \quad (1.10)$$

$$\mathbf{B}_{i+1} = \mathbf{M} \mathbf{A}_i \mathbf{M}^T + \mathbf{Q}_i, \quad (1.11)$$

où \mathbf{Q}_i est la matrice de covariance d'erreurs modèle, \mathbf{A}_i est la matrice de covariance d'erreurs de prévision et $\mathbf{M} = \mathbf{M}_{i,i+1}$ est l'opérateur modèle de prévision depuis l'instant i jusqu'à l'instant $i + 1$. Cette méthode permet en particulier aux différentes matrices entrant en jeu de "suivre" la dynamique du problème. La dernière étape de l'algorithme est à nouveau identique à celui du BLUE : l'analyse sert d'ébauche au cycle d'assimilation suivant.

$$\mathbf{X}_{i+1}^b = \mathbf{M} \mathbf{X}_i^a. \quad (1.12)$$

Dans le cadre de notre étude la matrice de covariance d'erreurs modèle \mathbf{Q} est considérée comme nulle. Cette hypothèse est forte et entraîne une diminution progressive des fonctions de covariances contenues dans \mathbf{B} . Concrètement l'erreur d'ébauche diminue peu à peu. On prend donc de moins en moins en compte les observations. Cependant, dans la mesure où cette étude s'intéresse aux longueurs de portée, un travail sur les fonctions de corrélation (qui elles sont normalisées) est satisfaisant.

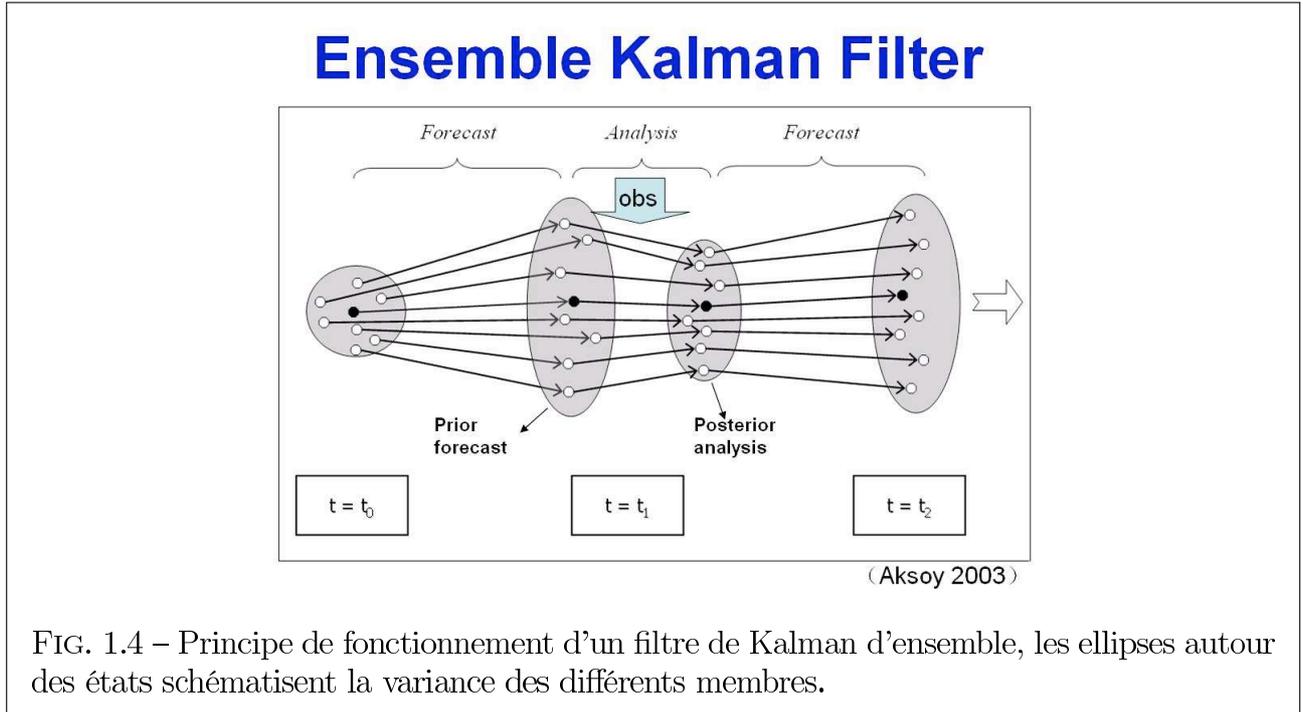
1.6 Filtre de Kalman d'ensemble

L'objectif du filtre de Kalman d'ensemble [Evensen, 1994] reste identique au filtre de Kalman vu précédemment, soit prendre en compte dans l'assimilation de données la dynamique d'évolution des erreurs associées à une prévision. Pour atteindre cet objectif, le filtre de Kalman applique les équations du modèle caractérisées par l'opérateur \mathbf{M} à la matrice \mathbf{B} , au travers de l'équation 1.11.

Le principe du filtre de Kalman d'ensemble est différent, il consiste à établir la matrice \mathbf{B} de covariance d'erreurs ébauche d'un point de vue statistique à partir d'un ensemble de couples ébauche–analyse représentant une même réalité et évoluant parallèlement.

Il opère sur un ensemble comportant N analyses, $\mathcal{E}^a = \{\mathbf{X}^{a,k}, 1 \leq k \leq N\}$, dont chacun des membres représente une même réalité perturbée par une erreur propre supposée non biaisée. Selon le même principe, un ensemble d'ébauches, $\mathcal{E}^b = \{\mathbf{X}^{b,k}, 1 \leq k \leq N\}$, est associé aux analyses.

1.6. FILTRE DE KALMAN D'ENSEMBLE



Un cycle d'assimilation au temps t_i se déroule tel qu'illustré par la figure 1.5. Tout d'abord, un ensemble d'ébauches est obtenu en appliquant le modèle à chacun des membres de l'ensemble d'analyses du pas de temps précédent. Ensuite la matrice de covariance d'erreurs d'ébauche, $\mathbf{B}_i = \mathbb{E}(\mathbf{X}_i^b - \overline{\mathbf{X}}_i^b)(\mathbf{X}_i^b - \overline{\mathbf{X}}_i^b)^T$, est estimée statistiquement ;

$$\mathbf{B}_i = \frac{1}{N} \sum_{k=1}^N (\mathbf{X}_i^{b,k} - \overline{\mathbf{X}}_i^b)(\mathbf{X}_i^{b,k} - \overline{\mathbf{X}}_i^b)^T; \quad \text{avec} \quad \overline{\mathbf{X}}_i^b = \frac{1}{N} \sum_{k=1}^N \mathbf{X}_i^{b,k}.$$

De même que dans le BLUE et le filtre de Kalman, cette matrice fournit la matrice de gain selon l'équation 1.8. Enfin, l'ensemble d'analyses au temps t_i est obtenu, en assimilant pour chaque membre de l'ensemble d'ébauches l'observation \mathbf{Y}_i^o perturbée par un bruit ε_i^k , selon l'équation (1.9). On observe alors une baisse de la variance, ces états analysés servent alors après propagation par le modèle d'ébauche pour l'analyse suivante (Fig. 1.4).

La convergence de la matrice \mathbf{B} s'effectue en $1/\sqrt{N}$ selon le théorème central limite. La taille de l'ensemble est déterminée par la précision souhaitée, ainsi que par la complexité des calculs d'ébauche et d'analyse de chacun des membres.

Le filtre de Kalman d'ensemble, en permettant le calcul séparé des différentes analyses et ébauches, est particulièrement bien adapté aux machines massivement parallèles.

1.6. FILTRE DE KALMAN D'ENSEMBLE

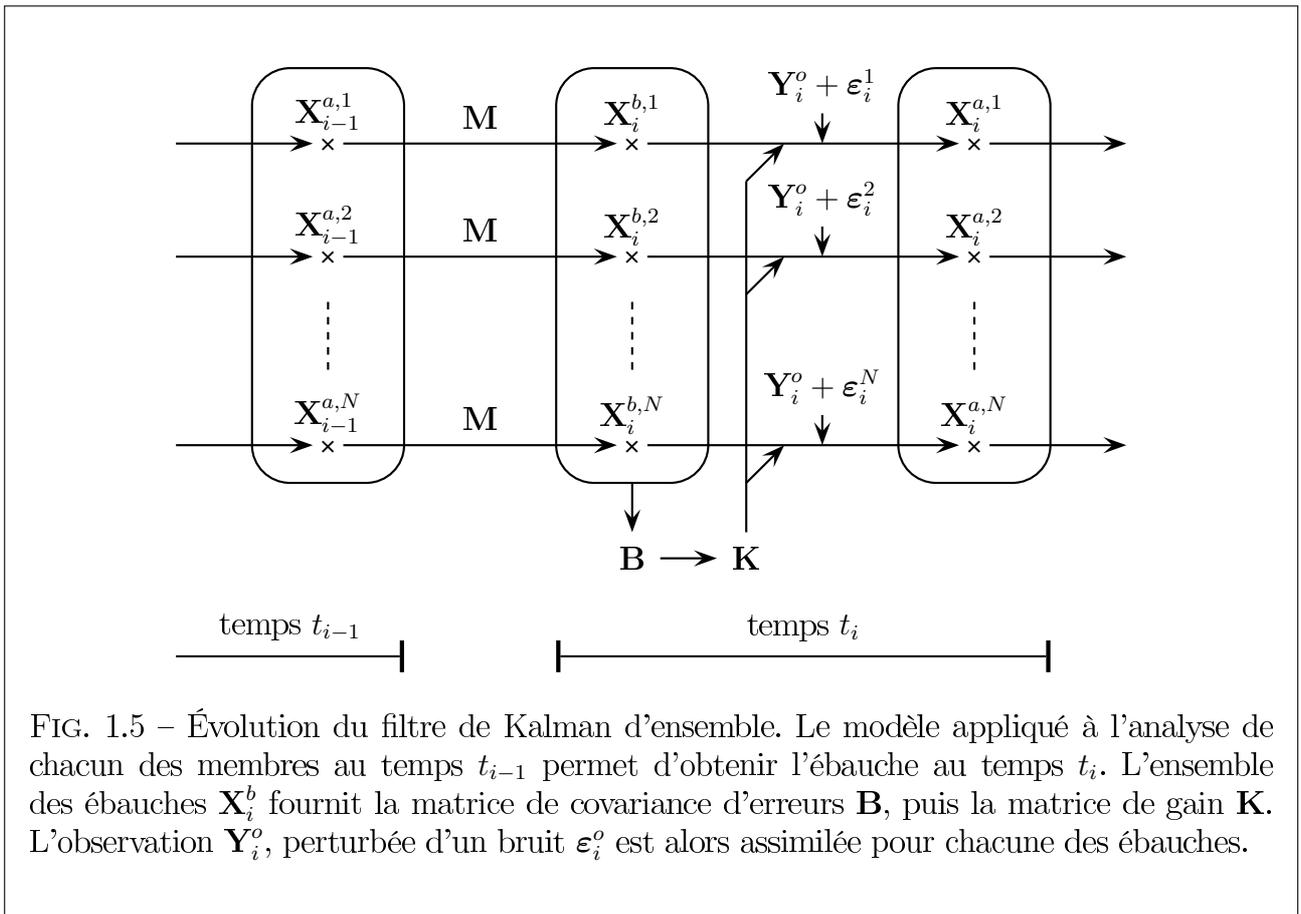


FIG. 1.5 – Évolution du filtre de Kalman d'ensemble. Le modèle appliqué à l'analyse de chacun des membres au temps t_{i-1} permet d'obtenir l'ébauche au temps t_i . L'ensemble des ébauches \mathbf{X}_i^b fournit la matrice de covariance d'erreurs \mathbf{B} , puis la matrice de gain \mathbf{K} . L'observation \mathbf{Y}_i^o , perturbée d'un bruit ϵ_i^o est alors assimilée pour chacune des ébauches.

Chapitre 2

Modèle d'onde de crue

Ce chapitre introduit dans un premier temps le modèle d'écoulement à surface libre dérivé des équations de Saint-Venant, utilisé dans cette étude [Thual, 2010], avant de présenter sa résolution numérique.

2.1 Physique du problème

Un bief est un canal à pente faible utilisant la gravité pour acheminer l'eau. On s'intéresse à la modélisation des ondes de crues sur un bief incliné d'un angle γ avec l'horizontale (Fig.2.1). La hauteur d'eau est notée h . On suppose de plus une invariance selon y , c'est à dire $\forall l, \frac{\partial l}{\partial y} = 0$. Pour un écoulement turbulent, les équations de Navier-Stokes s'écrivent sous la forme suivante, avec p_t la pression turbulente et ν_t la viscosité turbulente, de façon à traiter la turbulence d'une manière paramétrée

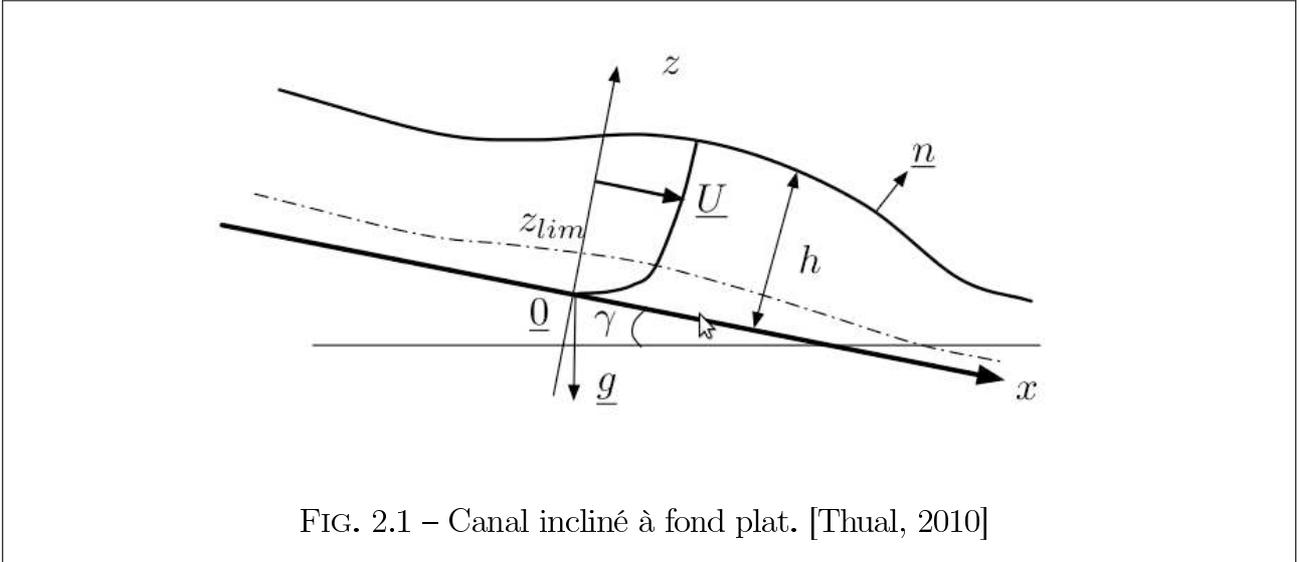
$$\left\{ \begin{array}{l} \frac{\partial u}{\partial x} + \frac{\partial z}{\partial w} = 0, \\ \frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + w \frac{\partial u}{\partial z} = -\frac{1}{\rho} \frac{\partial p_t}{\partial x} + g \sin \gamma + \nu_t \Delta u, \\ \frac{\partial w}{\partial t} + u \frac{\partial w}{\partial x} + w \frac{\partial w}{\partial z} = -\frac{1}{\rho} \frac{\partial p_t}{\partial z} - g \cos \gamma + \nu_t \Delta w. \end{array} \right. \quad (2.1)$$

A partir des équations 2.1 précédentes, le système d'équations de Saint-Venant se déduit à l'aide de différentes hypothèses :

- l'équation de la surface libre est de la forme $F(x, z, t) = z - h(x, t) = 0$, ce qui permet d'exclure les déformations de type déferlement.
- Au niveau de la couche limite de fond, z_{lim} est considéré comme très petit.
- La profondeur est supposée faible devant l'échelle de variation des phénomènes dans la direction de l'écoulement.
- La vitesse longitudinale moyenne U devient une variable du problème, en intégrant les équations sur la hauteur du canal.
- Les frottements sont paramétrés par la relation : $\tau_{frot} = \frac{1}{2} C_f \rho U |U|$, avec C_f un coefficient de frottement qui dépend de la structure de la rivière (rugosité du fond, hauteur, etc).¹

¹Il existe plusieurs manières de modéliser ce coefficient. Nous considérons ici l'approximation de Manning-Strickler.

2.1. PHYSIQUE DU PROBLÈME



Les équations de Saint-Venant qui est le système d'équations codées dans le modèle MAS-CARET [Ricci et al., 2010] sont alors,

$$\begin{cases} \frac{\partial h}{\partial t} + U \frac{\partial h}{\partial x} = -h \frac{\partial U}{\partial x}, \\ \frac{\partial U}{\partial t} + U \frac{\partial U}{\partial x} = -g \cos \gamma \frac{\partial h}{\partial x} + g \sin \gamma - \frac{CfU|U|}{2h}. \end{cases} \quad (2.2)$$

Dans le cadre de ce projet, ce n'est pas ce système d'équations qui est utilisé. En effet, des approximations supplémentaires sont réalisées afin de simplifier le problème. En particulier, ces simplifications permettent de passer d'un problème multivarié (h, U) à un problème univarié (h) .

- La pente est supposée faible, c'est à dire que le terme en $g \sin \gamma$ est négligeable dans le système de Saint-Venant.
- Le nombre de Froude est également supposé faible, ce qui signifie physiquement que la vitesse de l'écoulement est faible devant la vitesse des ondes de surface en milieu peu profond.

Pour de petites perturbations autour d'un état de référence (h_r, U_r) on écrit :

$$h = h_r + \tilde{h} \quad \text{et} \quad U = U_r + \tilde{U}.$$

En réalisant cette approximation et en négligeant les termes d'ordre 2 dans les équations 2.2, apparaît finalement l'équation des ondes de crues diffusantes, avec $\kappa_n = \frac{U_n h_n}{2 \tan \gamma}$, U_n la vitesse moyennée sur la verticale et h_n la hauteur d'eau moyenne :

$$\frac{\partial \tilde{h}}{\partial t} + \frac{5U_n}{3} \frac{\partial \tilde{h}}{\partial x} = \kappa_n \frac{\partial^2 \tilde{h}}{\partial x^2}.$$

L'équation régissant les ondes de crues diffusantes est en fait une équation d'advection/diffusion qui sera noté par la suite ² :

²Pour plus de simplicité, \tilde{h} est remplacé par h .

2.2. SCHÉMA NUMÉRIQUE DE L'ÉQUATION D'ADVECTION-DIFFUSION

$$\frac{\partial h}{\partial t} + c \frac{\partial h}{\partial x} = \kappa \frac{\partial^2 h}{\partial x^2}. \quad (2.3)$$

2.2 Schéma numérique de l'équation d'advection-diffusion

L'étude physique précédente permet de modéliser l'écoulement dans un cadre continu : les anomalies de hauteur sont régies par l'équation d'advection-diffusion (Eq. 2.3). Sa modélisation numérique nécessite une discrétisation spatiale et temporelle, puis une approximation des dérivées. La méthode retenue est celle des différences finies, en approximant les dérivées par des taux d'accroissement sur des intervalles de petites dimensions.

2.2.1 Schéma de discrétisation

L'écoulement est modélisé sur un domaine $[0; L]$ discrétisé au moyen d'une grille de pas Δx et comportant n points. Le temps est discrétisé selon un pas de temps Δt . L'écoulement est alors représenté numériquement par l'ensemble discret des valeurs $h_i^j = h(j\Delta x, i\Delta t)$. La modélisation de cet écoulement consiste à établir une relation $\mathbf{h}_{i+1} = \mathbf{M}\mathbf{h}_i$ où $\mathbf{h}_i = (h_0^i \cdots h_n^i)^T$ est le vecteur des valeurs discrétisées de h à l'instant $i\Delta t$. On utilise un schéma centré d'ordre 2 pour la dérivée spatiale $\frac{\partial h}{\partial x}$, centré d'ordre 1 pour la dérivée $\frac{\partial^2 h}{\partial x^2}$ et un schéma d'Euler explicite d'ordre 1 pour la dérivée temporelle $\frac{\partial h}{\partial t}$. En dehors des extrémités du domaine ($1 \leq j \leq n-1$), l'équation du modèle numérique s'écrit :

$$\frac{h_{i+1}^j - h_i^j}{\Delta t} + c \frac{h_{i+1}^{j+1} - h_i^{j-1}}{2\Delta x} = \kappa \frac{h_{i+1}^{j+1} - 2h_i^j + h_{i+1}^{j-1}}{\Delta x^2}. \quad (2.4)$$

En l'absence de forçage, la condition limite en $x = 0$ est modélisée par $h_i^0 = 0$. En $x = L$, elle est modélisée par une sortie libre, $\frac{\partial h}{\partial t}(L, t) + c \frac{\partial h}{\partial x}(L, t) = 0$, discrétisée par $\frac{h_{i+1}^j - h_i^j}{\Delta t} + c \frac{h_i^j - h_{i+1}^{j-1}}{\Delta x} = 0$. Les équations du modèle du temps i au temps $i+1$ s'écrivent matriciellement $\mathbf{h}_{i+1} = \mathbf{M}\mathbf{h}_i$ avec la matrice \mathbf{M} du modèle :

$$\mathbf{M} = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ a_0 & a_1 & a_2 & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & a_0 & a_1 & a_2 \\ 0 & \cdots & 0 & a_3 & a_4 \end{pmatrix}$$

Les coefficients a_0, \dots, a_4 dépendent des paramètres de la modélisation physique, c et κ , ainsi que des paramètres de la modélisation numérique, Δx et Δt :

$$a_0 = \frac{c\Delta t}{2\Delta x} + \frac{\kappa\Delta t}{\Delta x^2}; \quad a_1 = 1 - \frac{2\kappa\Delta t}{\Delta x^2}; \quad a_2 = -\frac{c\Delta t}{2\Delta x} + \frac{\kappa\Delta t}{\Delta x^2}; \quad a_3 = \frac{c\Delta t}{\Delta x}; \quad a_4 = 1 - \frac{c\Delta t}{\Delta x}.$$

La prise en compte d'un forçage en $x = 0$, $\mathbf{f}_i = (q(i\Delta t) \ 0 \ \cdots \ 0)^T$ entraîne l'ajout d'un terme non linéaire :

$$\mathbf{h}_{i+1} = \mathbf{M}\mathbf{h}_i + \mathbf{f}_{i+1}.$$

2.2. SCHÉMA NUMÉRIQUE DE L'ÉQUATION D'ADVECTION-DIFFUSION

2.2.2 Stabilité du schéma

L'étude de stabilité permet de vérifier que l'énergie, c'est-à-dire la norme carré du signal, reste bornée au cours du temps. Or la discrétisation de l'équation d'advection en utilisant un schéma centré pour la dérivée spatiale est connue pour être instable [Gicquel and Vermorel, 2010]. Il s'agit donc de voir sous quelles conditions la présence du terme de diffusion permet de rendre ce schéma stable.

L'analyse est effectuée selon la méthode de Von Neumann, dont le principe consiste à utiliser une transformation de Fourier afin d'étudier la stabilité de chaque mode séparément. Plus formellement cette méthode repose sur le théorème de Plancherel, selon lequel une fonction de carré intégrable possède la même norme carrée que sa transformée de Fourier : $\|h\|_2 = \|\widehat{h}\|_2$ [Revuz, 1997]. Les propriétés sur les transformés de Fourier permettent alors de transformer les translations en multiplications :

$$h^j(\widehat{x + \Delta x})(\zeta) = e^{-2i\pi\Delta x} \widehat{h^j(x)}(\zeta).$$

Ce qui est ensuite appliqué à l'équation du schéma numérique (Éq. 2.4) :

$$\frac{\widehat{h^{i+1}} - \widehat{h^i}}{\Delta t} + \frac{c}{2\Delta x} (e^{2i\pi\zeta\Delta x} - e^{-2i\pi\zeta\Delta x}) \widehat{h^i} - \frac{\kappa}{\Delta x^2} (e^{2i\pi\zeta\Delta x} - 2 + e^{-2i\pi\zeta\Delta x}) \widehat{h^i} = 0$$

Le gain g , quotient de la transformée de Fourier du signal entre deux instants successifs, est alors :

$$g(\zeta) = \frac{\widehat{h^{i+1}}}{\widehat{h^i}} = 1 - \frac{c\Delta t}{\Delta x} i \sin(2\pi\zeta\Delta x) + 2\frac{\kappa\Delta t}{\Delta x^2} (\cos(2\pi\zeta\Delta x) - 1). \quad (2.5)$$

La stabilité du schéma est alors acquise si pour chaque mode, soit $\forall \zeta \in \mathbb{R}$, le module du gain est inférieur à 1. Le détail des calculs figure en annexe A.1, la condition de stabilité prend la forme de trois inéquations :

$$\alpha = c \frac{\Delta t}{\Delta x} \leq 1; \quad \beta = 2\kappa \frac{\Delta t}{\Delta x^2} \leq 1; \quad \frac{\alpha^2}{\beta} = \frac{c^2 \Delta t}{2\kappa} \leq 1.$$

Ce qui conduit à la condition sur Δt , voir figure 2.2 :

$$\Delta t \leq \min \left\{ \frac{\Delta x}{c}; \frac{\Delta x^2}{2\kappa}; \frac{2\kappa}{c^2} \right\}.$$

Ces inégalités, nommées condition CFL³, font apparaître la CFL d'un schéma d'advection, $c \frac{\Delta t}{\Delta x} \leq 1$ et la CFL d'un schéma de diffusion, $2\kappa \frac{\Delta t}{\Delta x^2} \leq 1$. Le schéma d'advection seul n'étant pas stable, la dernière inéquation, $\frac{c^2 \Delta t}{2\kappa} \leq 1$ indique dans quelle mesure la diffusion permet de le stabiliser. Par exemple, avec les paramètres suivants, $c = 1 \text{ m}\cdot\text{s}^{-1}$, $\kappa = 500 \text{ m}^2\cdot\text{s}^{-1}$, un bief de longueur $L = 500 \text{ km}$, discrétisée en $n = 1000$ points, il vient un pas de temps maximum $\Delta t = 250 \text{ s}$.

2.2.3 Équation équivalente

Un développement limité en $(j\Delta x, i\Delta t)$ permet de vérifier que le schéma résout effectivement l'équation d'advection-diffusion (Eq. 2.3). La fonction ν solution du schéma numérique est telle

³Condition Courant-Friedrichs-Lewy, du nom de Richard Courant, Kurt Friedrichs, et Hans Lewy

2.2. SCHÉMA NUMÉRIQUE DE L'ÉQUATION D'ADVECTION-DIFFUSION

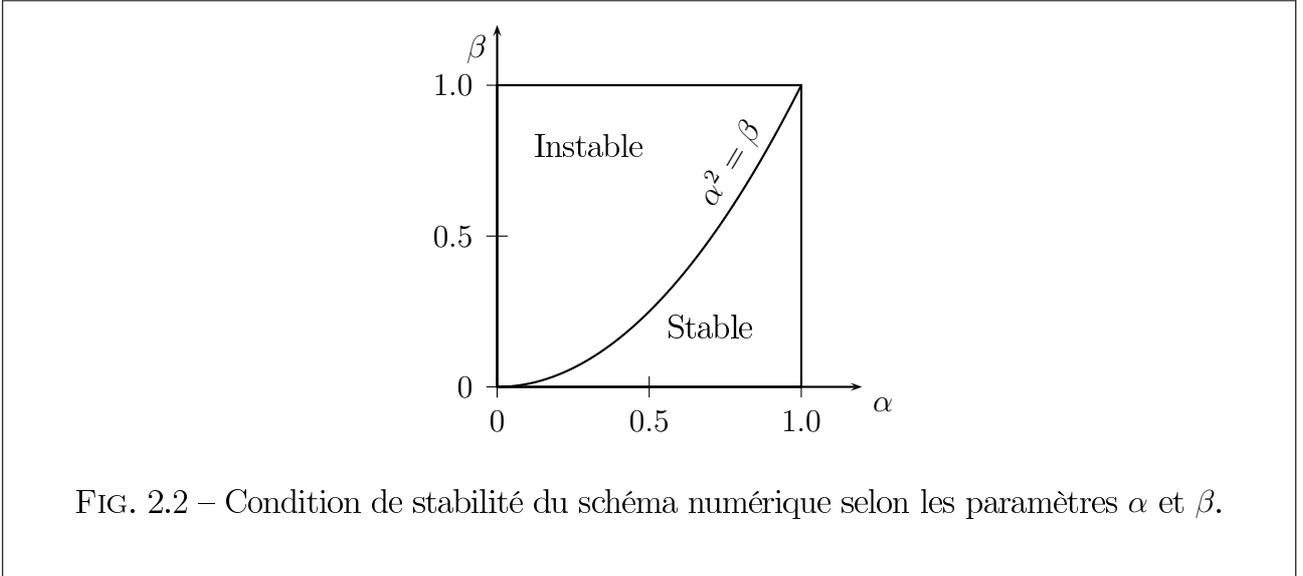


FIG. 2.2 – Condition de stabilité du schéma numérique selon les paramètres α et β .

que :

$$\frac{\partial \nu}{\partial t} + c \frac{\partial \nu}{\partial x} - \kappa \frac{\partial^2 \nu}{\partial x^2} + \underbrace{\frac{\Delta t}{2} \frac{\partial^2 \nu}{\partial t^2} + \Delta x^2 \left(\frac{c}{6} \frac{\partial^3 \nu}{\partial x^3} - \frac{\kappa}{12} \frac{\partial^4 \nu}{\partial x^4} \right)}_{\varepsilon} + O(\Delta t^2, \Delta x^3) = 0. \quad (2.6)$$

À l'ordre 1 en temps et 2 en espace, le schéma résout l'équation d'advection-diffusion, à l'erreur ε près. D'après l'équation 2.6, le schéma numérique introduit une erreur ε . Cette erreur engendre des comportements du modèle numérique non souhaités dans le modèle physique, en particulier en agissant sur le coefficient de diffusion, mais également en introduisant de la dispersion.

La dérivation de l'équation d'advection-diffusion (Eq. 2.3) permet d'exprimer les dérivées temporelles en fonction des dérivées spatiales.

$$\frac{\partial^2 \nu}{\partial t^2} = c^2 \frac{\partial^2 \nu}{\partial x^2} - 2c\kappa \frac{\partial^3 \nu}{\partial x^3} + \kappa^2 \frac{\partial^4 \nu}{\partial x^4}$$

Cela permet de réécrire l'équation résolue par le schéma numérique, et fait apparaître à l'ordre 2 des termes supplémentaires :

$$\frac{\partial \nu}{\partial t} + c \frac{\partial \nu}{\partial x} - \underbrace{\left(\frac{c\Delta x^2}{6} + c\kappa\Delta t \right)}_{\text{dispersion}} \frac{\partial^3 \nu}{\partial x^3} = \underbrace{\left(\kappa - \frac{c^2\Delta t}{2} \right)}_{\kappa'} \frac{\partial^2 \nu}{\partial x^2} - \left(\frac{\kappa^2\Delta t}{2} - \frac{\kappa\Delta x^2}{12} \right) \frac{\partial^4 \nu}{\partial x^4} + O(\Delta t^2, \Delta x^3) \quad (2.7)$$

2.2.4 Impact des termes diffusifs et dispersifs sur la propagation des covariances

L'équation (2.7), ou équation équivalente résolue par le schéma numérique, fait apparaître un terme de dispersion et deux termes de diffusion supplémentaires. Il faudra prendre en compte ces aspects lors des expériences numériques. En particulier, la diffusion apparente du schéma numérique $\kappa' = \kappa - \frac{c^2\Delta t}{2}$, doit être prise en compte lors de la comparaison des résultats. En effet le coefficient de diffusion intervient dans la loi d'évolution de la longueur de portée (Par. 3).

2.3. CONSTRUCTION DU FORÇAGE AMONT

En ce qui concerne le terme dispersif, ce dernier déforme les signaux propagés. En effet, la vitesse de phase est alors dépendante de la fréquence, et les différentes composantes d'un signal ne se déplacent pas à la même vitesse. D'autant plus que dans certains cas, une des trois conditions CFL impose $\kappa' = 0$. De plus, les deux autres termes, dispersif et diffusif d'ordre 4, affectent particulièrement les petites échelles ce qui peut être gênant lors des simulations. Pour plus de détails, voir l'annexe A.2.

2.3 Construction du forçage amont

Afin d'avoir un signal de covariance spatiale gaussienne en entrée du modèle, la solution choisie est de construire un signal de covariance temporelle gaussienne et de forcer le modèle à son bord à chaque pas de temps avec cette fonction. Il faut construire en amont une condition limite de covariance temporelle Gaussienne [Pannekoucke, 2010]. Pour cela, on peut considérer un bruit blanc défini par sa décomposition de Fourier

$$\zeta = \sum_n \zeta_n e_n(t),$$

où les ζ_n sont des variables aléatoires suivant une loi normale centrée réduite. En particulier, l'indépendance deux à deux des variables aléatoires ζ_n impose la propriété suivante :

$$\langle \zeta_p \zeta_q^* \rangle = \delta_{pq}. \quad (2.8)$$

On applique à ce bruit blanc un opérateur L permettant de passer du bruit blanc au bruit de covariance gaussienne. Cet opérateur est déterminé à partir de coefficients de Fourier. En utilisant le fait que les e_n sont vecteurs propres de L associés aux valeurs propres l_n , il vient que :

$$q = L\zeta = \sum_n \zeta_n l_n e_n. \quad (2.9)$$

Il suffit alors de déterminer les l_n pour construire le signal voulu. En effet les ζ_n assurent le caractère aléatoire, tandis que les l_n assurent la coloration Gaussienne (i.e la covariance temporelle Gaussienne) du signal. Une façon d'accéder aux coefficients l_n nécessaires à la fabrication du signal d'entrée q est d'écrire la décomposition de Fourier de la fonction d'auto-covariance :

$$\rho(\delta t) = \langle q(t)q(t + \delta t)^* \rangle.$$

Il vient alors en utilisant l'équation 2.9 et la propriété 2.8 :

$$\begin{aligned} \rho(\delta t) = \langle q(t)q(t + \delta t)^* \rangle &= \sum_{p,q} \langle (\zeta_p l_p e_p(t)) (\zeta_q l_q e_q(t + \delta t))^* \rangle \\ &= \sum_{p,q} \langle \zeta_p \zeta_q^* \rangle l_p l_q^* e_p(t) e_q(t + \delta t)^* \\ &= \sum_p l_p l_p^* e_{-p}(\delta t) \\ \rho(\delta t) &= \sum_p |l_p|^2 e_p(\delta t). \end{aligned} \quad (2.10)$$

2.3. CONSTRUCTION DU FORÇAGE AMONT

La phase permet simplement de décaler le signal dans le temps, ce qui n'est pas important ici. Le signal voulu est alors facilement déductible à partir d'une fonction d'auto-covariance ρ connue. En effet, la décomposition de Fourier selon la base des e_n donne :

$$\rho(\delta t) = \sum_n \rho_n e_n(\delta t).$$

Par identification avec l'équation (2.10), il vient alors que $|l_n|^2 = |\rho_n|$ et

$$q(t) = \sum_n \zeta_n l_n e_n = \sum_n \zeta_n \sqrt{|\rho_n|} e_n(t).$$

Dans le cadre de cette étude, il sera considéré une fonction de covariance gaussienne, avec τ le temps caractéristique d'auto-corrélation :

$$\rho(t) = e^{\frac{-t^2}{2\tau^2}}.$$

L'algorithme permettant de réaliser une fonction d'entrée telle que celle présentée sur la figure (2.3) est détaillé ci-après :

```
>  $\rho(t) = e^{\frac{-t^2}{2\tau^2}}$ 
% On effectue une discrétisation spatiale.
>  $\rho$ 
% vecteur contenant la discretisation spatiale de  $\rho$ ; ( $\rho_n$ )
>  $\hat{\rho} = \text{fft}(\rho)$ 
% Vecteur contenant les coefficients spectraux de  $\rho$ .
>  $\zeta$  = vecteur de la même taille que  $\rho$  rempli avec des tirages aléatoires selon une loi normale centrée réduite.
>  $\hat{q} = \sqrt{|\hat{\rho}|} * \zeta$ 
% Opération terme à terme.
% Vecteur contenant les coefficients spectraux de  $q$ ; ( $\zeta_n \sqrt{|\rho_n|}$ )
>  $q = \text{fft}^{-1}(\hat{q})$ 
% Vecteur de sortie, forçage amont.
```

2.3. CONSTRUCTION DU FORÇAGE AMONT

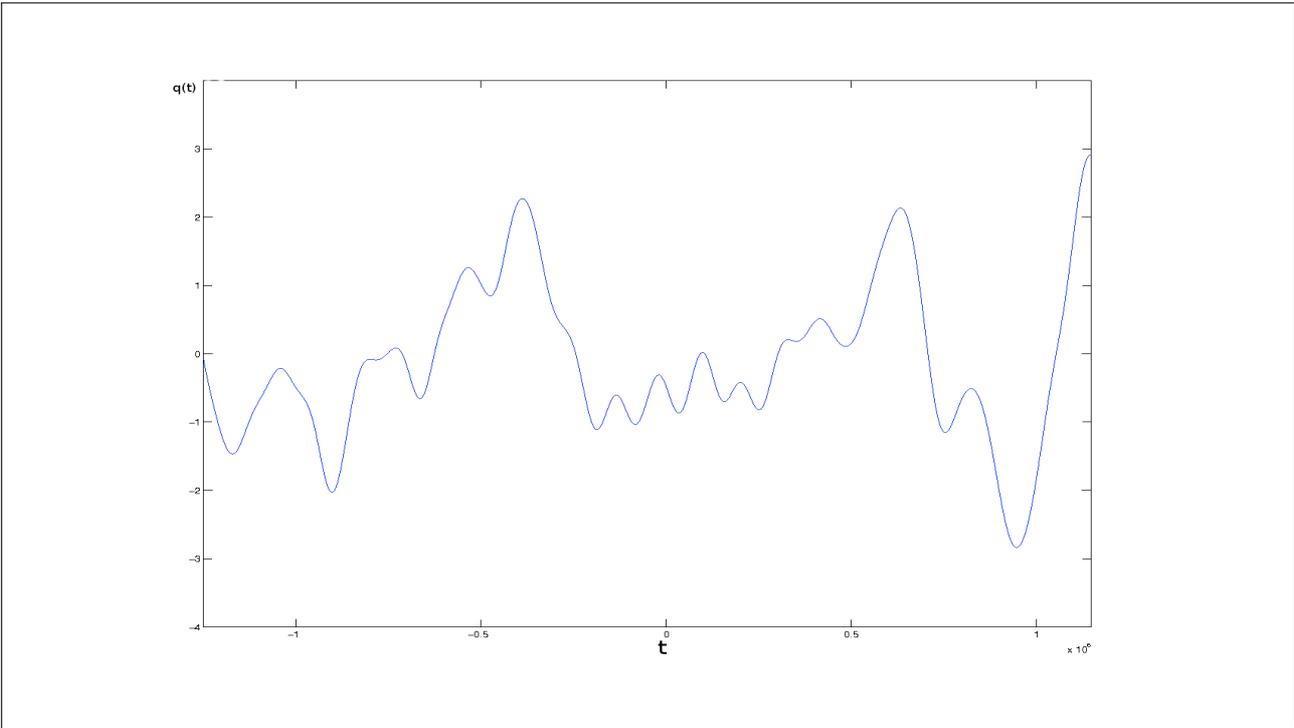


FIG. 2.3 – Exemple de réalisation de forçage amont.

Chapitre 3

Dynamique des covariances : évolution des longueurs de portée

L'étude se tourne désormais vers l'étude de la propagation des fonctions de covariances, et plus particulièrement vers l'évolution des longueurs de portée. Cependant cette étude n'est pas aisée. En effet, alors qu'une fonction de covariance est propagée par le modèle, elle n'est pas directement accessible comme la valeur de la hauteur d'eau. Dans cette partie sera traitée l'aspect théorique de la propagation de fonctions de covariance par la dynamique de l'équation 2.3.

3.1 Dynamique des covariances sans observations

Le but de cette partie est de trouver analytiquement la forme de la covariance spatiale d'un signal propagé par l'équation 2.3. La méthode proposée est de résoudre analytiquement la propagation du signal $h(x, t)$, avant d'écrire la fonction d'auto-covariance spatiale de la solution. Le système à résoudre est le suivant, avec c la vitesse d'advection, κ le coefficient de diffusion, $q(t)$ la condition limite en $x = 0$, et h_0 la condition initiale.

$$\begin{cases} \frac{\partial h}{\partial t} + c \frac{\partial h}{\partial x} = \kappa \frac{\partial^2 h}{\partial x^2} \\ h(0, t) = q(t) \\ h(x, 0) = h_0(x), \quad x > 0 \end{cases} \quad (3.1)$$

Le forçage $q(t)$ peut être vu comme une superposition de signaux harmoniques sinusoidaux :

$$q(t) = \int q_\omega e_\omega(t) d\omega.$$

Le fait que $q(t)$ soit construit aléatoirement, c'est à dire que les q_ω sont des réalisations d'une variable aléatoire (voir 2.3) implique la propriété suivante :

$$\langle q_{\omega_1} q_{\omega_2}^* \rangle = |q_{\omega_1}|^2 \delta_{12}, \quad (3.2)$$

avec δ_{12} le symbole de Kronecker. De plus, $q(t)$ réelle, impose la condition : $|q_\omega|^2 = |q_{|\omega|}|^2$. La fonction d'auto-covariance, en utilisant la propriété 3.2 est alors donnée par :

$$\rho(\delta t) = \langle q(t) q(t + \delta t)^* \rangle = \int \rho_\omega e^{i\omega \delta t} d\omega = \int |q_\omega|^2 e^{i\omega \delta t} d\omega \quad (3.3)$$

3.1. DYNAMIQUE DES COVARIANCES SANS OBSERVATIONS

Le signal $q(t)$ possède un bruit Gaussien (i.e. coloration Gaussienne d'un bruit blanc). Ce qui se traduit par :

$$\begin{cases} \rho(\delta t) = e^{-\delta t^2/2\tau^2} \\ |q_\omega|^2 = \rho_\omega = e^{\omega^2\tau^2/2} \end{cases}$$

La linéarité du problème permet de ne considérer qu'un forçage sinusoïdal $q(t) = q_\omega e^{i\omega t}$ pour la résolution. La solution est reconstruite à l'aide du théorème de superposition. Cette vision revient à chercher une solution de la forme $h(x, t) = q_\omega e^{-i\omega t} h_\omega(x)$. Dans la suite on cherchera donc des solutions sous cette forme, en posant $h_\omega(0) = 1$. La solution générale est dès lors reconstruite avec la relation suivante :

$$h(x, t) = \int q_\omega e^{-i\omega t} h_\omega(x) d\omega. \quad (3.4)$$

La fonction d'auto-covariance de $h(x, t)$ spatiale est alors calculée en reportant l'équation 3.4 dans :

$$f(x, x + \delta x) = \langle h(x, t) h(x + \delta x, t)^* \rangle, \quad (3.5)$$

ce qui permet d'en déduire des propriétés sur les longueurs de portée.

3.1.1 Cas de l'advection pure

Dans ce cas, le système à résoudre est le suivant :

$$\begin{cases} \frac{\partial h}{\partial t} + c \frac{\partial h}{\partial x} = 0 \\ h(x, t) = q_\omega e^{-i\omega t} h_\omega(x) \\ h_\omega(0) = 1 \end{cases}$$

En remplaçant la forme de la solution dans l'équation pronostique, il apparaît une équation différentielle ordinaire en h_ω :

$$\frac{\partial h}{\partial t} + c \frac{\partial h}{\partial x} = 0 \Leftrightarrow -i\omega h_\omega + c h'_\omega = 0 \Leftrightarrow h'_\omega = \frac{i\omega}{c} h_\omega$$

Les solutions de cette équation différentielle ordinaire, sont de la forme : $h(x, t) = A e^{i\omega x/c}$, avec A une constante. En utilisant la condition initiale, $h_\omega(0) = 1$, il vient en notant $k = \frac{\omega}{c}$

$$h(x, t) = q_\omega e^{ik(x-ct)}.$$

On utilise alors la linéarité du problème pour en déduire la solution complète :

$$\begin{cases} q(t) = \int q_\omega e^{-i\omega t} d\omega \\ h(x, t) = \int q_\omega e^{ik(x-ct)} d\omega. \end{cases}$$

Il ne reste alors plus qu'à déterminer le tenseur des covariances spatiales. Il s'agit donc d'évaluer en un instant donné, la quantité : $f(x, x + \delta x) = \langle h(x, t) h(x + \delta x, t)^* \rangle$. Il vient ainsi que :

$$\begin{aligned} f(x, x + \delta x) &= \langle h(x, t) h(x + \delta x, t)^* \rangle \\ &= \left\langle \int_\omega q_\omega h_\omega(x) d\omega \int_\beta q_\beta^* h_\beta(x + \delta x)^* d\beta \right\rangle \\ &= \int |q_\omega|^2 h_\omega(x) h_\omega(x + \delta x)^* d\omega \\ &= \int |q_\omega|^2 e^{i\omega\delta x/c} d\omega \end{aligned} \quad (3.6)$$

3.1. DYNAMIQUE DES COVARIANCES SANS OBSERVATIONS

On reconnaît alors la forme de la fonction d'auto-covariance de la fonction $q(t)$ en posant $\delta t = \delta x/c$. C'est à dire que l'on a :

$$f(x, x + \delta x) = \rho(\delta x/c)$$

avec $\rho(\delta x/c) = e^{-(\delta x/c)^2/2\tau^2} = e^{-\delta x^2/2Lp^2}$ en posant $Lp = c\tau$. Cette fonction ne dépend ni du temps t , ni de la position x , mais uniquement de la distance entre deux points du domaine. La solution analytique de Lp dans le cadre du forçage considéré, est

$$Lp = \text{Constante} = c\tau, \quad (3.7)$$

avec τ le temps caractéristique d'auto-corrélation de la condition limite à gauche.

3.1.2 Cas de l'advection diffusion

Dans ce cas, le système à résoudre est le suivant :

$$\begin{cases} \frac{\partial h}{\partial t} + c \frac{\partial h}{\partial x} = k \frac{\partial^2 h}{\partial x^2} \\ h(x, t) = q_\omega e^{-i\omega t} h_\omega(x) \\ h_\omega(0) = 1 \end{cases}$$

En remplaçant la forme de la solution dans l'équation pronostique, il apparaît une équation différentielle en h_ω :

$$\frac{\partial h}{\partial t} + c \frac{\partial h}{\partial x} = k \frac{\partial^2 h}{\partial x^2} \Leftrightarrow -i\omega h_\omega + c h'_\omega = k h''_\omega \Leftrightarrow k h''_\omega + \frac{i\omega}{c} h_\omega - h'_\omega = 0$$

Le polynôme caractéristique associé à cette équation différentielle est :

$$kZ^2 - cZ + i\omega = 0$$

Les racines de ce polynôme sont : $z^{+/-} = \frac{c}{2k} \pm \frac{\sqrt{c^2 - 4i\omega k}}{2k}$ ce qui permet d'en déduire la forme des solutions de l'équation différentielle :

$$h_\omega(x) = Ae^{z^+x} + Be^{z^-x}$$

On déduit ensuite les constantes A et B des données du problème. En particulier, si $A \neq 0$, alors $\lim_{x \rightarrow \infty} |h_\omega(x)| = \infty$. Cette solution étant physiquement absurde, on en déduit que $A = 0$. De plus, on a $h_\omega(0) = 1$, ce qui donne $B = 1$. La solution de l'équation différentielle ordinaire est donc pour un signal harmonique :

$$h_\omega(x) = e^{(\frac{c}{2k} - \frac{\sqrt{c^2 - 4i\omega k}}{2k})x}$$

On utilise alors la linéarité du problème pour en déduire la solution complète :

$$\begin{cases} q(t) = \int q_\omega e^{-i\omega t} d\omega \\ h(x, t) = \int q_\omega e^{(\frac{c}{2k} - \frac{\sqrt{c^2 - 4i\omega k}}{2k})x - i\omega t} d\omega \end{cases}$$

On note qu'en considérant un cas avec une diffusion peu intense, $k \ll 1$, la racine du polynôme caractéristique se simplifie : $z \approx i\frac{\omega}{c} - \frac{\omega^2 k}{c^3}$. L'ordre zéro, permet de retrouver la solution de l'advection pure, et l'ordre un permet de corriger cette solution par un terme d'atténuation en $e^{-\frac{\omega^2 k}{c^3}x}$.

3.1. DYNAMIQUE DES COVARIANCES SANS OBSERVATIONS

Il ne reste alors plus qu'à déterminer le tenseur des covariances spatiales. Il s'agit donc d'évaluer en un instant donné, la quantité : $f(x, x + \delta x) = \langle h(x, t)h(x + \delta x, t)^* \rangle$. Il vient ainsi que :

$$\begin{aligned}
 f(x, x + \delta x) &= \langle h(x, t)h(x + \delta x, t)^* \rangle \\
 &= \left\langle \int_{\omega} q_{\omega} h_{\omega}(x) d\omega \int_{\beta} q_{\beta}^* h_{\beta}(x + \delta x)^* d\beta \right\rangle \\
 &= \int |q_{\omega}|^2 h_{\omega}(x) h_{\omega}(x + \delta x)^* d\omega
 \end{aligned} \tag{3.8}$$

En considérant alors l'approximation à l'ordre 1 de la diffusion, il vient alors que :

$$f(x, x + \delta x) = \int \left(|q_{\omega}|^2 e^{-\omega^2 k(2x + \delta x)/c^3} \right) e^{i\omega \delta x/c} d\omega$$

La première remarque à faire est que le tenseur des covariances spatiales, ne dépend toujours pas du temps, mais dépend cependant cette fois-ci de la position x . De plus, en remplaçant $|q_{\omega}|^2$ par sa valeur.

$$\begin{aligned}
 f(x, x + \delta x) &= \int \left(e^{-\omega^2/2(\tau^2 + 4kx/c^3 + 2k\delta x/c^3)} \right) e^{i\omega \delta x/c} d\omega \\
 &\underset{\delta x \ll 1}{\approx} \int \left(e^{-\omega^2/2(\tau^2 + 4kx/c^3)} \right) e^{i\omega \delta x/c} d\omega \\
 &\underset{\delta x \ll 1}{\approx} \int e^{-\omega^2 Lp_x^2/2} e^{i\omega \delta x/c} d\omega \\
 &\underset{\delta x \ll 1}{\approx} e^{-\delta x^2/2Lp_x^2}
 \end{aligned} \tag{3.9}$$

Ce calcul nous donne donc une estimation de Lp donnée par la relation :

$$Lp_x^2 = c^2(\tau^2 + 4kx/c^3) = Lp_0^2 + 4kx/c. \tag{3.10}$$

3.1.3 Bilan de la dynamique des longueurs de portée sans assimilation

Dans le cas d'une advection pure, la solution analytique de Lp dans le cadre du forçage considéré, est (Fig. 3.1) :

$$Lp = \text{Constante} = c\tau, \tag{3.11}$$

avec τ le temps caractéristique d'auto-corrélation de la condition limite à gauche.

Dans le cas d'une advection diffusion, une estimation de Lp en supposant que la fonction de covariance est localement gaussienne (on se place dans le cas où $\delta x \ll 1$). Cette estimation est donnée par la relation (Fig. 3.1) :

$$Lp_x^2 = c^2(\tau^2 + 4kx/c^3) = Lp_0^2 + 4kx/c. \tag{3.12}$$

3.2. DYNAMIQUE DES COVARIANCES AVEC OBSERVATION

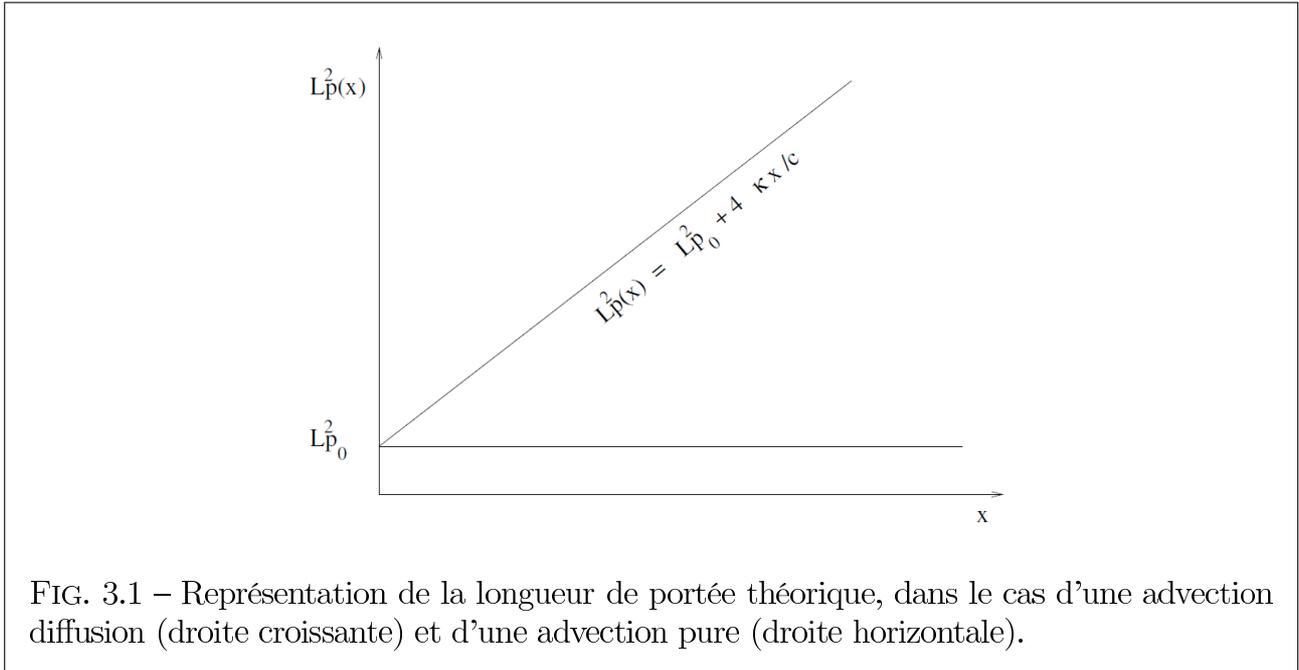


FIG. 3.1 – Représentation de la longueur de portée théorique, dans le cas d’une advection diffusion (droite croissante) et d’une advection pure (droite horizontale).

3.2 Dynamique des covariances avec observation

Le cas avec observation est beaucoup plus compliqué à mettre en œuvre analytiquement. L’assimilation de données implique une diminution de la variance au point d’observation. La propagation des covariances vers l’aval entraîne alors une diminution de la longueur de portée en aval du point d’observation (Fig. 3.2). La diminution de cette variance étant dépendante du rapport¹ σ_o^2/σ_b^2 , la diminution de la longueur de portée l’est également .

On se propose alors d’établir un abaque du rapport des longueurs de portée entre l’aval et l’amont au point d’observation en fonction du rapport des variances. Cet abaque a pour but de tenter de paramétrer la matrice de covariance d’erreurs d’ébauche de l’assimilation dans un BLUE pour le modèle MASCARET.

¹ σ_o^2 , la variance des erreurs d’observations et σ_b^2 la variance des erreurs sur l’ébauche

3.2. DYNAMIQUE DES COVARIANCES AVEC OBSERVATION

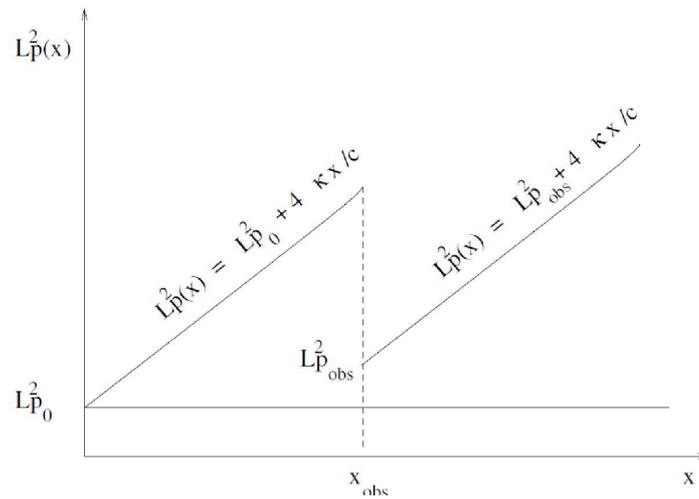


FIG. 3.2 – Longueurs de portée théoriques avec une observation.

Chapitre 4

Expériences numériques

Le paragraphe précédent résout mathématiquement la propagation des fonctions de covariance pour le schéma numérique utilisé. Cette partie traite de la vérification expérimentale de ces résultats mathématiques. Pour cela on diagnostique les longueurs de portée dans le cas d'une advection seule et dans celui d'une advection diffusion sans assimilation. Ensuite en se concentrant sur le modèle d'advection/diffusion, les mêmes diagnostics sont effectués dans le cas d'une assimilation à l'aide d'un algorithme de type BLUE ou d'un filtre de Kalman.

4.1 Diagnostic de \mathbf{B} et des longueurs de portée sans assimilation

La réalisation de ce diagnostic nécessite la détermination de la matrice \mathbf{B} . Une méthode ensembliste permet d'en obtenir une estimation. Désormais, les matrices de covariances d'erreurs déterminées statistiquement seront indicé par e .

4.1.1 Calcul de la matrice de covariances d'erreurs d'ébauche

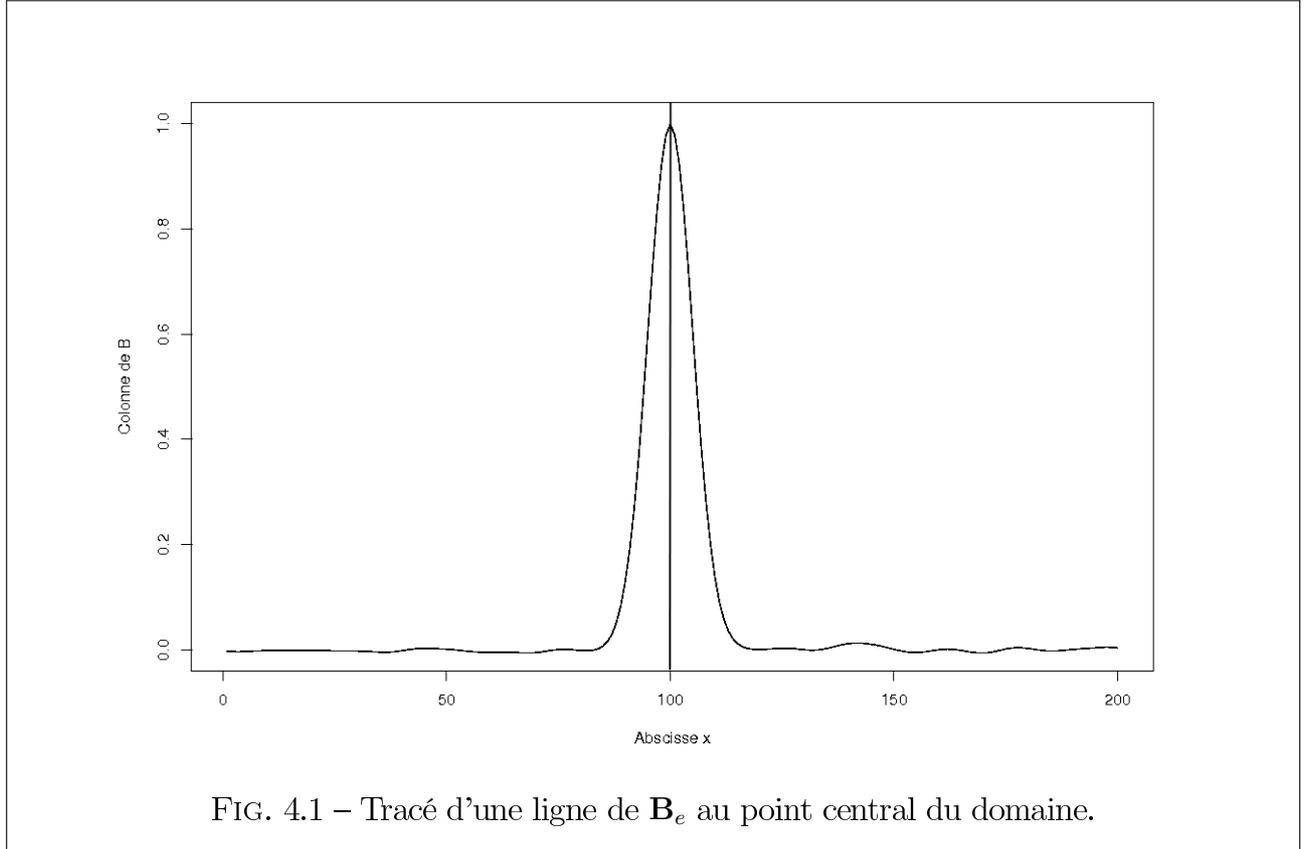
L'estimation de la matrice \mathbf{B}_e est faite statistiquement à partir de la réalisation d'un ensemble de propagations de signaux d'entrée $q(t)$ différents. Afin d'obtenir une matrice sans bruit d'échantillonnage, ce travail est réalisé sur un échantillon de 10^5 membres. Chaque forçage est propagé le long de la rivière. Le temps de propagation est choisi de façon à ce que le signal q ait pu atteindre la limite aval du domaine. Les différents vecteurs (\mathbf{X}_k) , $k \in [1, 10^5]$ sont gardés en mémoire (vecteur de taille n contenant la perturbation de la hauteur d'eau en chaque point de discrétisation spatiale). La matrice de covariances d'erreurs spatiales \mathbf{B}_e est alors directement déterminée à partir des vecteurs (\mathbf{X}_k) en calculant les covariances entre les différents vecteurs :

$$\mathbf{B}_e = \frac{1}{N} \sum_{k=1}^N ((\mathbf{X}_k - \bar{\mathbf{X}})(\mathbf{X}_k - \bar{\mathbf{X}})^T),$$

en notant $\bar{\mathbf{X}}$ = Vecteur des moyennes en chaque point sur l'ensemble des membres.

Par construction \mathbf{B}_e est carrée, de dimension n^2 et symétrique positive, mais pas nécessairement définie. En effet, son rang est au plus N . La figure 4.1 montre le tracé d'un ligne de \mathbf{B}_e quelconque. On y voit l'allure de la fonction de covariance au point central du domaine.

4.2. RÉSULTATS DES EXPÉRIENCES SANS ASSIMILATION



4.1.2 Calcul des longueurs de portée

A partir de cette matrice \mathbf{B}_e qui contient une estimation statistique des fonctions de covariances, on peut calculer les longueurs de portée amont et aval au point courant. La formule utilisée pour la longueur de portée est celle décrite dans le paragraphe 1.4. Le choix a également été fait de travailler sur les fonctions de corrélation plutôt que sur les fonctions de covariance étant donné que la définition de longueur de portée est valable pour les fonctions de corrélations. Finalement,

$$\rho(0) = 1, \quad \rho(\delta x) = \frac{\mathbf{B}_e(i, i+1)}{\sqrt{\mathbf{B}_e(i, i)\mathbf{B}_e(i+1, i+1)}} \quad \text{et} \quad \rho(-\delta x) = \frac{\mathbf{B}_e(i, i-1)}{\sqrt{\mathbf{B}_e(i, i)\mathbf{B}_e(i-1, i-1)}}.$$

4.2 Résultats des expériences sans assimilation

4.2.1 Modèle avec advection seule

Dans le cadre de l’advection seule, le signal d’entrée q est simplement propagé au cours du temps vers l’aval sans qu’il ne subisse aucune modification. D’après les calculs réalisés en 3 on s’attend à obtenir une longueur de portée constante par rapport à x sur l’ensemble du domaine. De plus $L_p^- = L_p^+$ implique que la fonction de covariance spatiale est identique (par rapport au point considéré) d’une limite à l’autre du domaine. Les résultats obtenus sont présentés dans les figures 4.2 et 4.3. Sur la figure 4.2 sont superposées les trois fonctions de covariances des points d’abscisses $x = \frac{L_{max}}{6}$, $x = \frac{L_{max}}{2}$ et $x = \frac{5L_{max}}{6}$. La fonction de covariance est donc, au même titre que le signal lui-même, simplement translatée par la schéma numérique sans être déformée. La figure 4.3 montre l’évolution du carré de la longueur de portée (leurs courbes

4.2. RÉSULTATS DES EXPÉRIENCES SANS ASSIMILATION

étant identiques, aucune distinction $L_p^+ - L_p^-$ n'est représentée) avec l'abscisse. Une régression linéaire ¹ (trait plein) permet de confronter l'expérience et la théorie. Avec une ordonnée à l'origine de $2494 km^2$ on est proche des $2500 km^2$ attendus. Le coefficient directeur estimé est lui d'un ordre de grandeur négligeable. On est là aussi en accord avec avec la théorie.

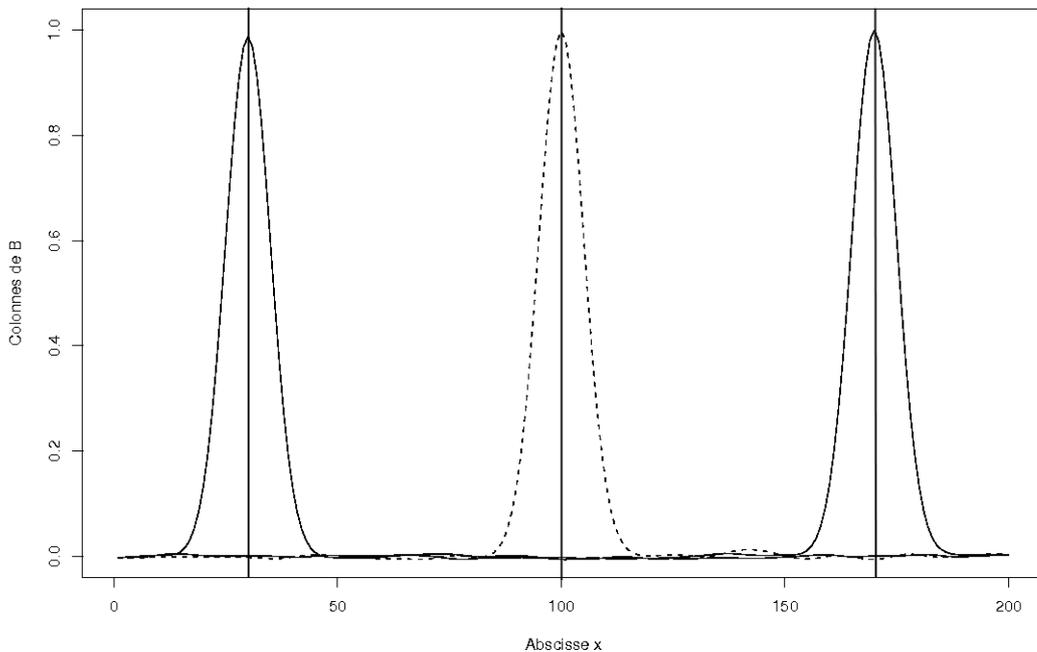


FIG. 4.2 – Tracé de trois colonnes de B_e dans le cadre d'une advection pure.

4.2.2 Modèle avec advection et diffusion

D'après la théorie (voir chapitre 3), la diffusion doit ici jouer sur "l'étalement" des fonctions de corrélation. En effet si les longueurs de portée augmentent bien linéairement le long de la rivière, on s'attend à un étalement progressif des fonctions de covariance. Expérimentalement la figure 4.4 illustre l'évolution du carré de la longueur de portée en fonction de x . Une régression linéaire a également été faite. Là encore, les coefficients directeurs et ordonnées à l'origine théoriques et expérimentaux sont en accord.

La figure 4.5 superpose plusieurs fonctions de corrélation aux mêmes points qu'en figure 4.2. L'influence de la diffusion qui "écrase et étale" les fonctions de covariance au fur et à mesure de leur propagation est bien visible.

Enfin la figure 4.6 montre un problème numérique auquel on peut être confronté. Le signal injecté en entrée possède un temps caractéristique faible. Autrement dit, ce signal contient beaucoup de "petites échelles". Au début du domaine, ces petites échelles ont pour effet de rendre non négligeable un terme de diffusion en $\frac{\partial^4 h}{\partial x^4}$ (Eq. 2.7) qu'on avait laissé de côté en raison de son ordre de grandeur dans le cadre général. La régression linéaire (en trait plein) réalisée sur les derniers points montre qu'on tend peu à peu vers une solution convenable. Il convient

¹réalisée avec le logiciel R : <http://www.r-project.org/>

4.3. ASSIMILATION AVEC UN BLUE

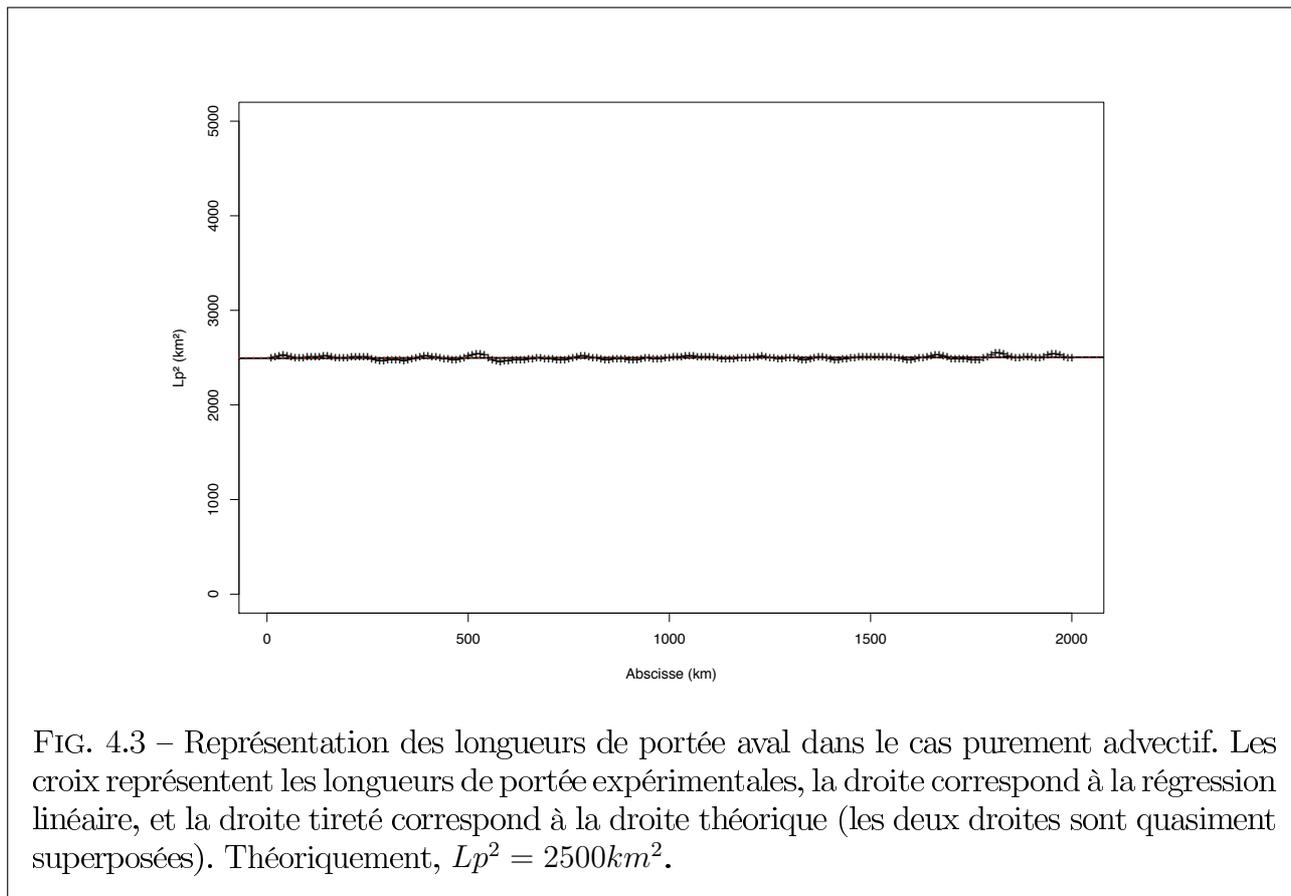


FIG. 4.3 – Représentation des longueurs de portée aval dans le cas purement advectif. Les croix représentent les longueurs de portée expérimentales, la droite correspond à la régression linéaire, et la droite tireté correspond à la droite théorique (les deux droites sont quasiment superposées). Théoriquement, $Lp^2 = 2500km^2$.

donc d'introduire un signal amont possédant un temps caractéristique d'auto-corrélation assez grand pour éviter cette difficulté.

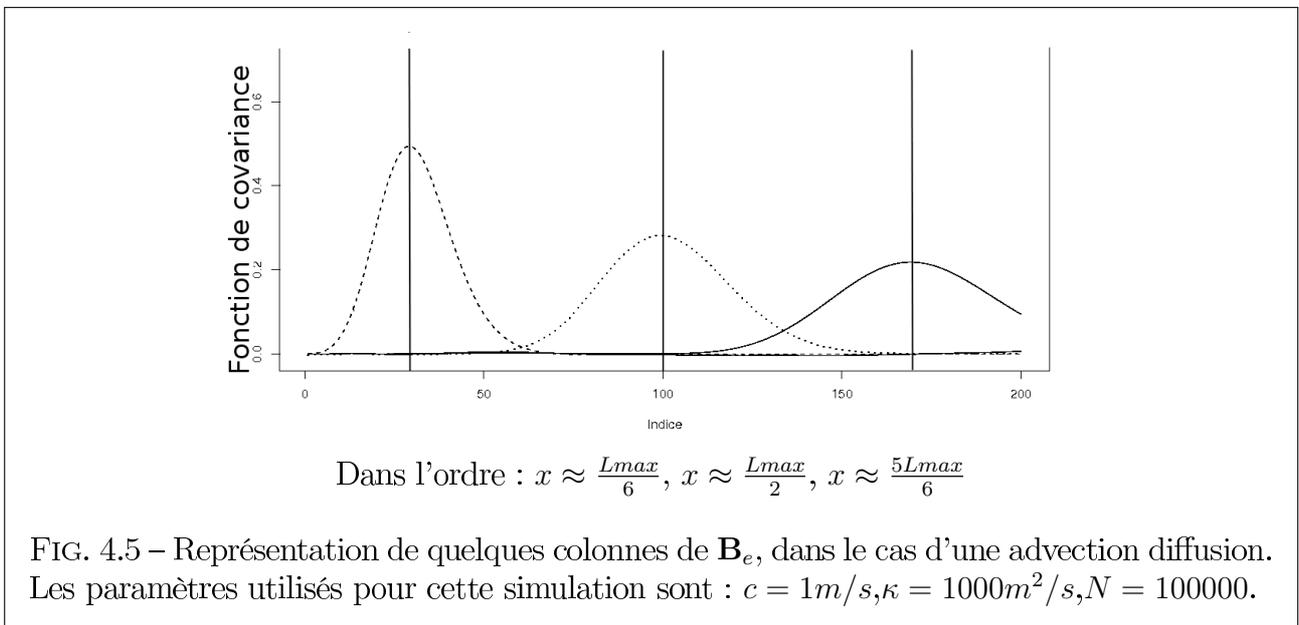
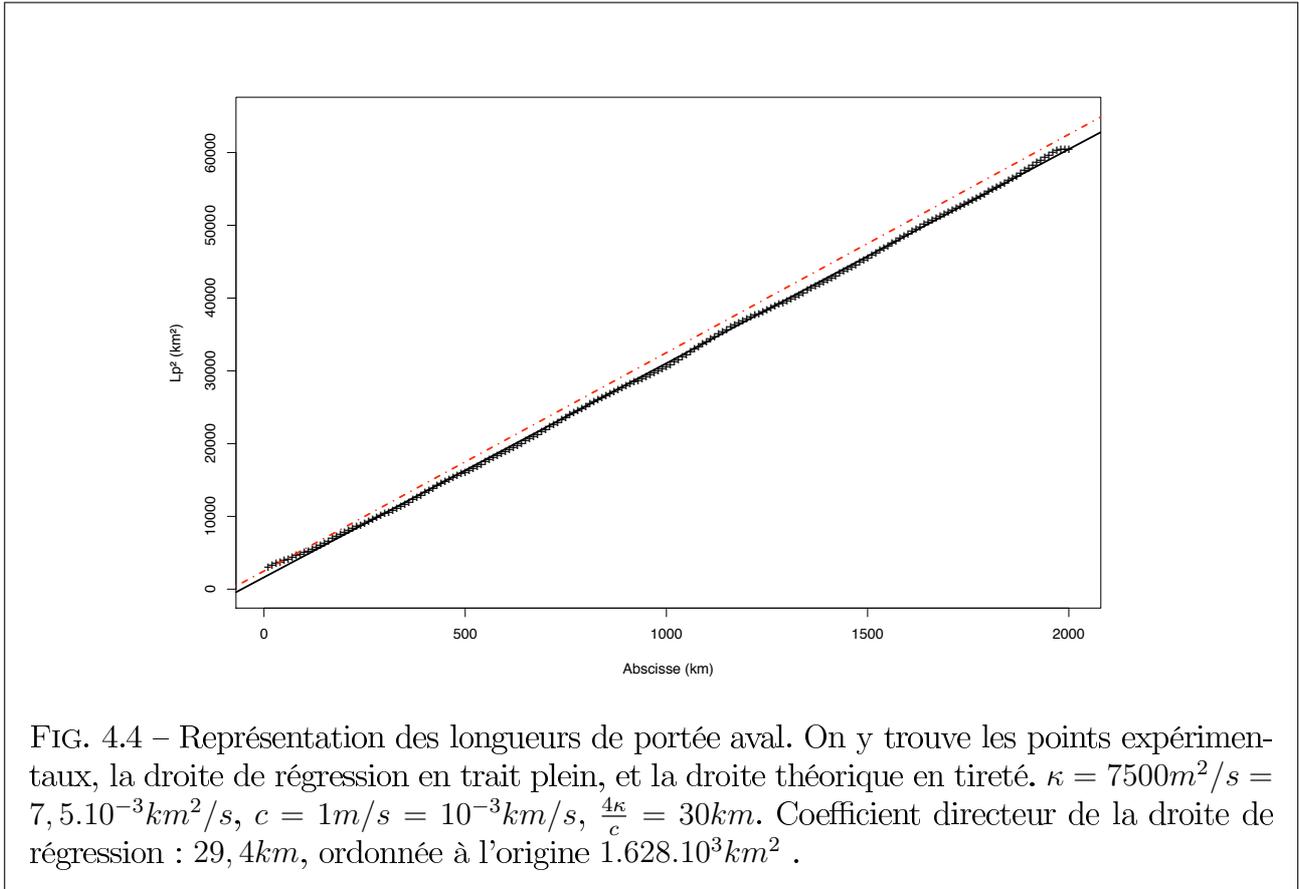
4.3 Assimilation avec un BLUE

On réalise ici le même travail que dans la partie 4.2 mais en effectuant une assimilation de données avec un algorithme de BLUE. La matrice de covariance d'erreur d'ébauche \mathbf{B} utilisée est celle déterminée dans la partie 4.1. La matrice de covariance d'erreurs de prévision obtenue est désormais notée \mathbf{A}_e . Le code numérique implémenté transcrit la formulation du BLUE présentée dans le chapitre 1. Comme précédemment le travail est fait sur 10^5 membres. On attend tout d'abord que le signal ait le temps de se propager sur tout le domaine. Puis la phase d'assimilation démarre. Le point d'observation est placé en milieu de domaine et une observation est assimilée tous les 5 pas de temps. Le temps d'assimilation est assez important pour que l'ajout d'information apporté par les observations ait pu atteindre les limites du domaine.

On considère dans un premier temps la figure ?? représentant les valeurs de la variance de l'erreur d'analyse (en trait plein) le long du cours d'eau. La courbe en tiretés correspond aux valeurs de la variance de l'erreur de prévision dans le cas sans observations. L'assimilation a donc permis une réduction de variance au point d'observation.

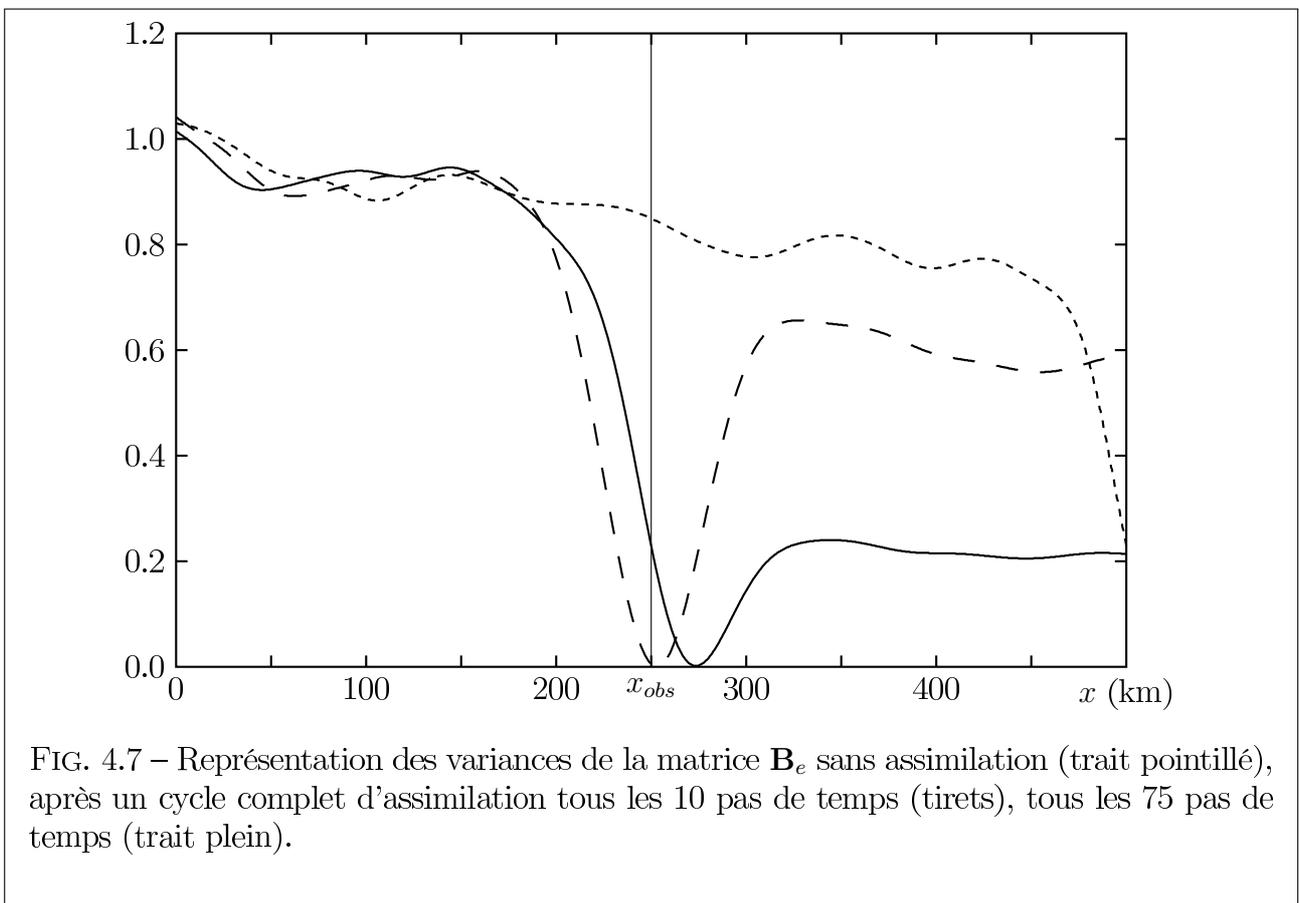
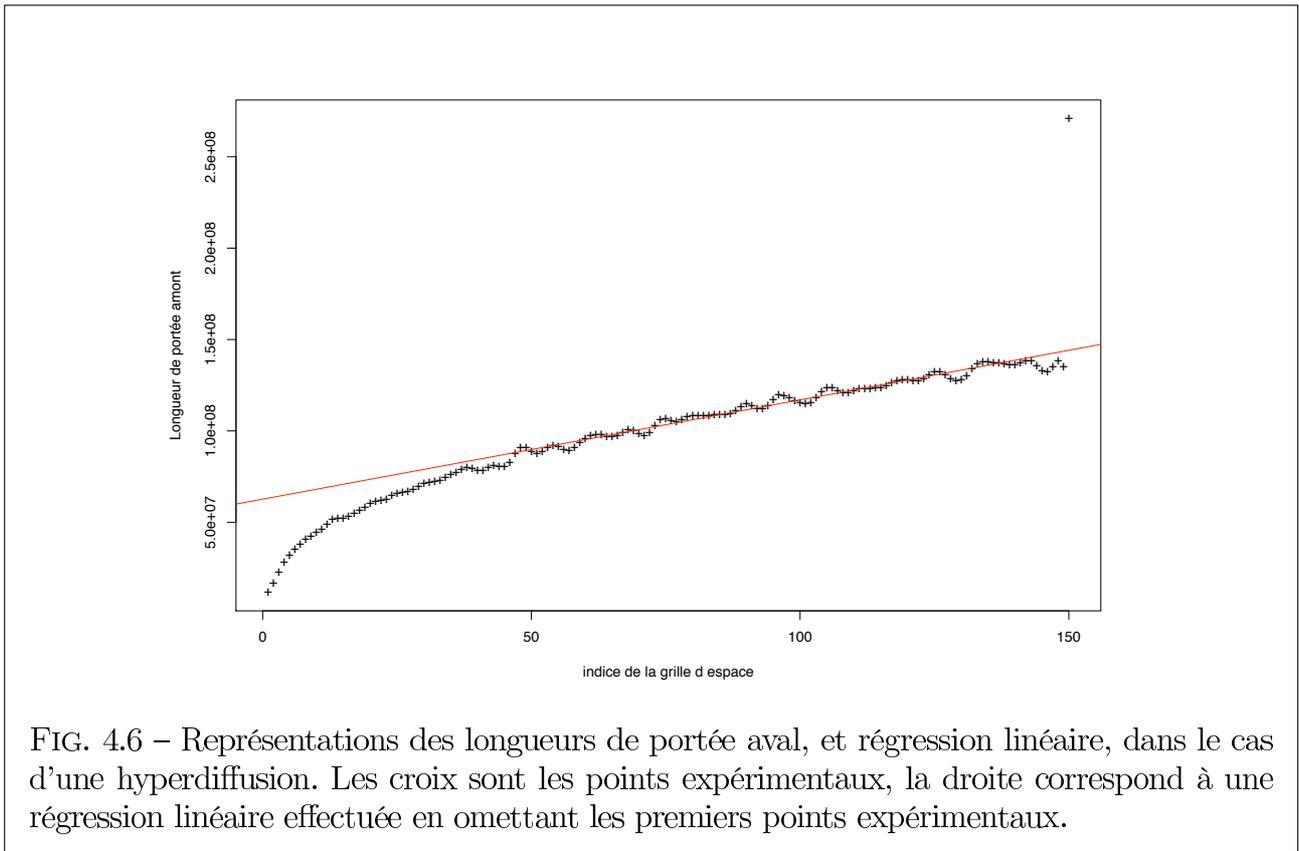
Une méthode pour la détermination de la structure de la matrice \mathbf{B}_e serait de boucler le travail précédent ; on remplace dans le code du BLUE la matrice \mathbf{B} par la matrice \mathbf{A}_e calculée à partir de l'ensemble des réalisations puis on recommence les calculs pour obtenir une nouvelle

4.3. ASSIMILATION AVEC UN BLUE



matrice \mathbf{A}_e qui servira à l'itération suivante. Cependant son coup de calcul est énorme. Le filtre de Kalman est à présent la méthode retenue pour construire une matrice \mathbf{B} convenable.

4.3. ASSIMILATION AVEC UN BLUE



4.4 Implémentation du filtre de Kalman

Dans le BLUE, la matrice \mathbf{B} de covariance des erreurs d'ébauche est statique. Elle ne tient pas compte de la réduction de l'incertitude engendrée par l'assimilation d'observations. Après quelques itérations, les covariances d'erreurs diffèrent sensiblement des conditions initiales, l'analyse n'est plus optimale.

Le principe du filtre de Kalman a été exposé dans la partie 1.5. Sa mise en œuvre va permettre de déterminer l'effet d'une série d'observations sur les erreurs d'ébauche, en tenant compte de la diffusion et de l'advection. Cela doit permettre un diagnostic des longueurs de portée des fonctions de corrélation au point d'observation.

4.4.1 Cadre de la simulation

Il s'agit de déterminer la matrice de covariance d'erreurs d'ébauche dans un cadre stationnaire (dans le sens où une observation est effectuée à chaque pas de temps) à laquelle est associée une erreur d'écart-type constant. Dans l'état initial, c'est-à-dire avant la période d'observation, les covariances d'erreurs ont été déterminées par méthode ensembliste tel que vu précédemment dans la partie 4.1, avec un écart-type $\sigma_q = 1$ m, et un temps caractéristique, $\tau = 30 \cdot 10^3$ s. La simulation est effectuée sur un domaine de longueur $L = 500$ km, discrétisé sur une grille de $n = 1000$ points. La condition CFL impose un pas de temps de $\Delta t = 250$ s. L'observation est affectée d'une erreur d'écart-type 0.5 m. En outre la vitesse d'advection est de $c = 1$ m·s⁻¹, la diffusion de $\kappa = 500$ m²·s⁻¹. En vertu des résultats de la partie 2.2.4, la diffusion réellement simulée est alors de $\kappa' = 375$ m²·s⁻¹.

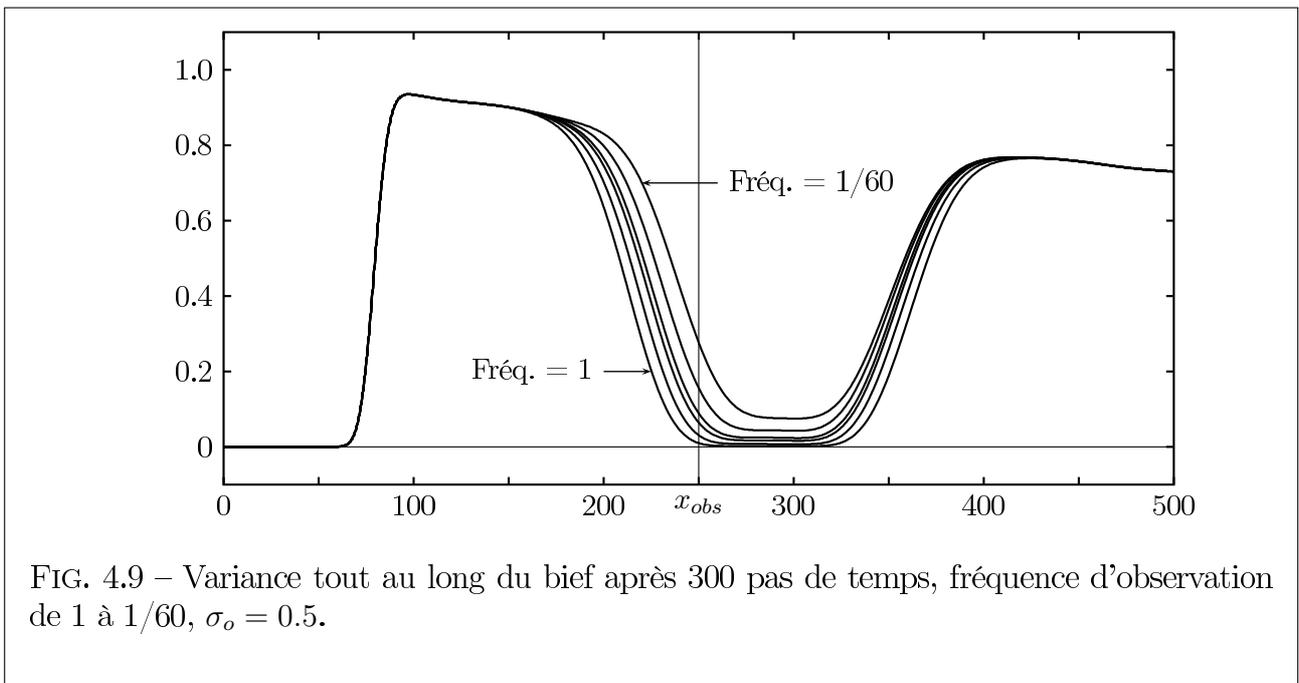
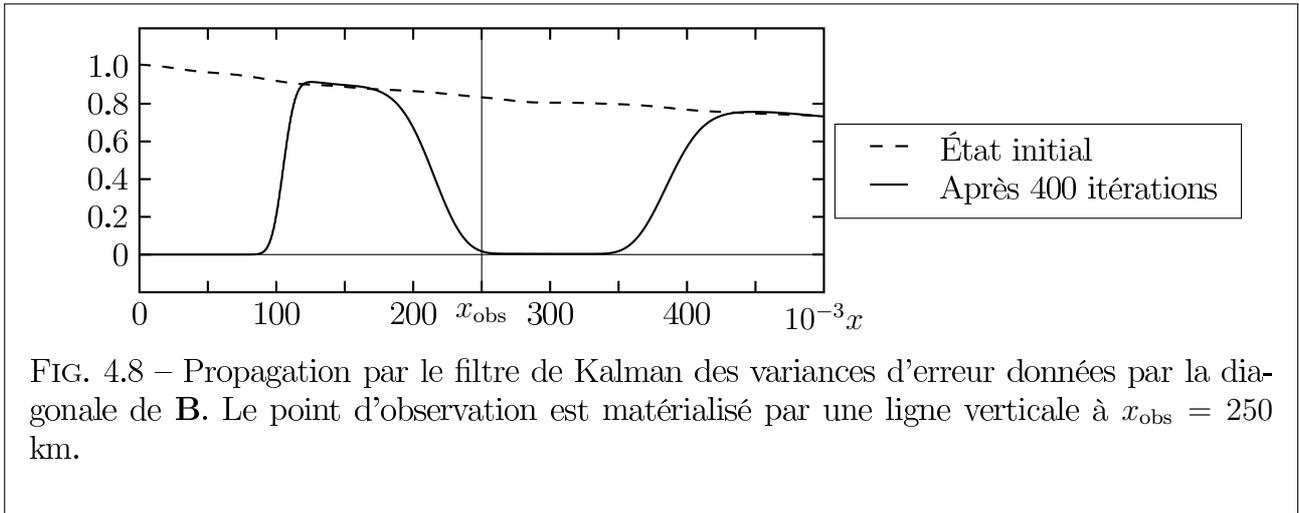
La figure 4.8 montre les variances d'erreurs d'ébauche tout au long du bief, c'est-à-dire la diagonale principale de la matrice \mathbf{B} . Les deux courbes correspondent à l'état initial, et à l'état après 400 itérations, soit une advection de 100 km. Dans l'état initial, les variances sont décroissantes. Sur la courbe représentant les variances après 400 itérations, une zone de très faibles variances s'est formée au delà du point d'observation. L'information apportée par l'observation a fortement réduit l'incertitude, au point d'observation, l'advection a ensuite étendue cette zone vers l'aval.

De plus, une zone dans laquelle les variances sont nulles ou quasi-nulles s'est artificiellement propagée à partir de l'amont, sur une distance correspondante à l'advection. En effet, cela provient de l'impossibilité de modéliser l'incertitude entrant en amont dans le système. Celle-ci est négligée et se propage ainsi ensuite. Afin que cette zone n'atteigne pas le point d'observation, la simulation est nécessairement limitée en temps, et les résultats suivants seront donnés sur un domaine restreint par rapport au domaine de calcul, avec un temps d'assimilation limité. Seul l'éloignement du point d'observation permet de limiter ces contraintes. Il est ainsi impossible de parvenir à un état rigoureusement stationnaire, mais les résultats suivants montrent une convergence assez rapide dans le domaine d'intérêt, c'est-à-dire autour du point d'observation. Seul un filtre de Kalman d'ensemble permet de dépasser ces limites.

En dehors de ces deux zones, les variances ne sont que très peu affectées.

La figure 4.9 montre, après 300 pas de temps, l'influence de la fréquence d'observation sur la variance d'erreur d'ébauche. L'augmentation de la fréquence d'observation réduit la variance dans la zone située au delà du point d'observation, jusqu'à la rendre quasiment nulle lorsque les observations sont effectuées à chaque pas de temps.

4.4. IMPLÉMENTATION DU FILTRE DE KALMAN



4.4. IMPLÉMENTATION DU FILTRE DE KALMAN

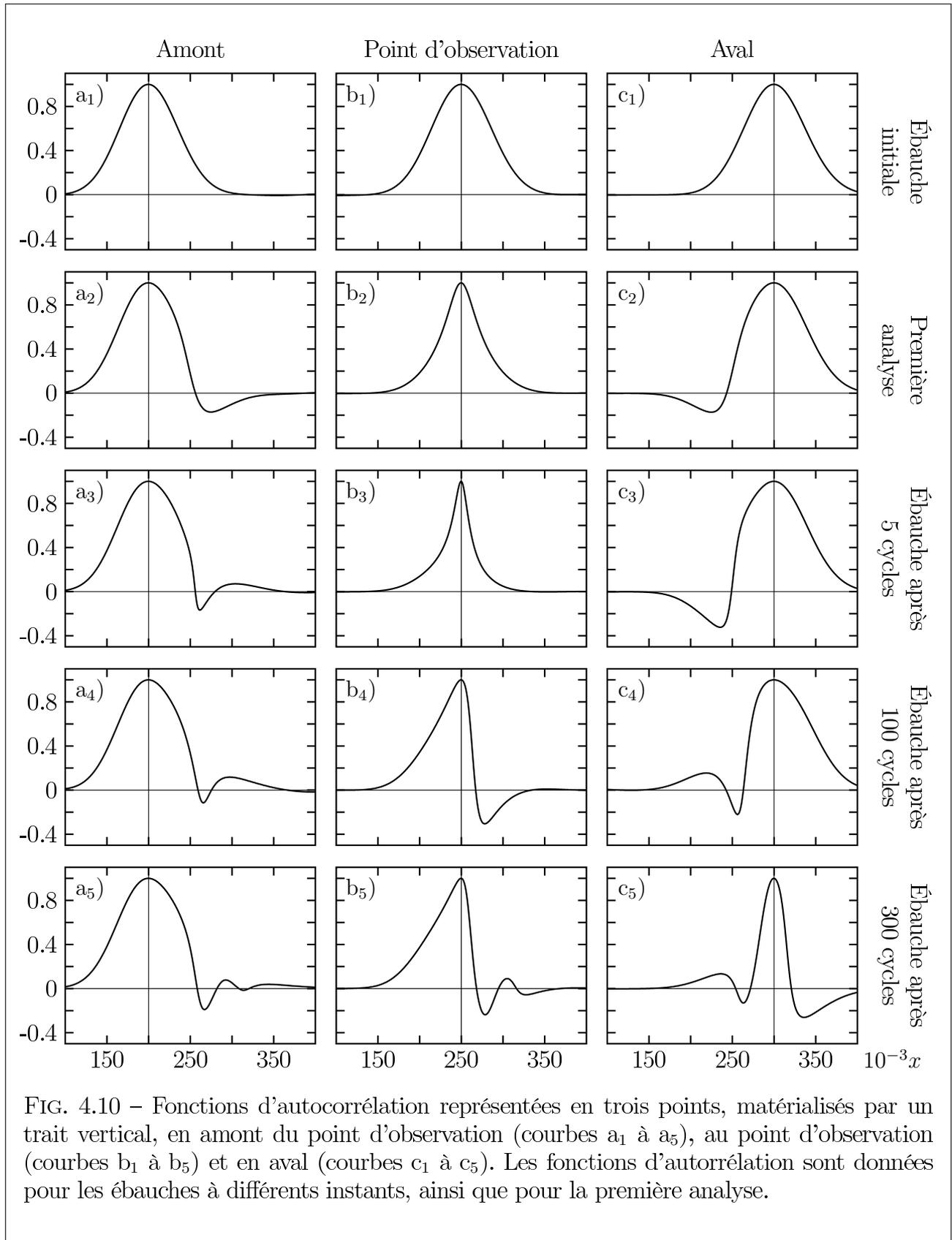


FIG. 4.10 – Fonctions d'autocorrélation représentées en trois points, matérialisés par un trait vertical, en amont du point d'observation (courbes a_1 à a_5), au point d'observation (courbes b_1 à b_5) et en aval (courbes c_1 à c_5). Les fonctions d'autocorrélation sont données pour les ébauches à différents instants, ainsi que pour la première analyse.

4.4. IMPLÉMENTATION DU FILTRE DE KALMAN

4.4.2 Évolution des fonctions de corrélation

Le graphique 4.10 montre les fonctions d'autocorrélation en trois points du bief, au point d'observation et en deux points amont et aval, chacun distant de 50 km du point d'observation. Comparée aux longueurs de portée initiales, cette distance est relativement faible, et ces deux points sont influencés par la présence des observations.

Dans l'état initial les fonctions d'autocorrélation sont celles de la partie 4.2 caractérisées par des longueurs de portée croissantes en allant vers l'aval. La première assimilation impacte de manière isotrope la fonction d'autocorrélation au point d'observation (courbe b_2) et réduit sensiblement les longueurs de portée. Les points amont et aval (courbes a_2 et c_2) sont impactés de manière symétriques. L'advection n'a pas pu, à ce stade, entraîner un déplacement de l'information vers l'aval. Cependant, dès les premiers cycles d'assimilation (courbes a_3 à c_3), une tendance à l'anisotropie semble se dégager.

Au cours des cycles suivants, les fonctions d'autocorrélation en amont, au point d'observation, et en aval évoluent de manières différentes. Au point amont (courbes a_4 et a_5), l'assimilation de données influence la fonction d'autocorrélation, mais celle-ci conserve des longueurs de portée relativement isotropes et constantes au cours du temps.

Au point d'observation, la fonction d'autocorrélation prend très rapidement un caractère anisotrope qu'elle conserve ensuite. Les longueurs de portée amont et aval diffèrent fortement, mais leur quotient semble converger assez rapidement. Ce point est étudié plus particulièrement dans la suite.

Au point aval, la fonction d'autocorrélation passe par trois étapes : isotrope avec des longueurs de portée importantes, anisotrope, puis isotrope avec des longueurs de portée réduites. Cela résulte de l'advection. Tout d'abord ce point n'est pas concerné par l'information provenant de l'observation, ensuite l'information parvient à ce point entraînant le caractère anisotrope, finalement ce point est traversé par un flot continu d'information. La fonction de corrélation retrouve ainsi un caractère isotrope mais avec des longueurs de portée réduites.

4.4.3 Méthode d'estimation des longueurs de portée

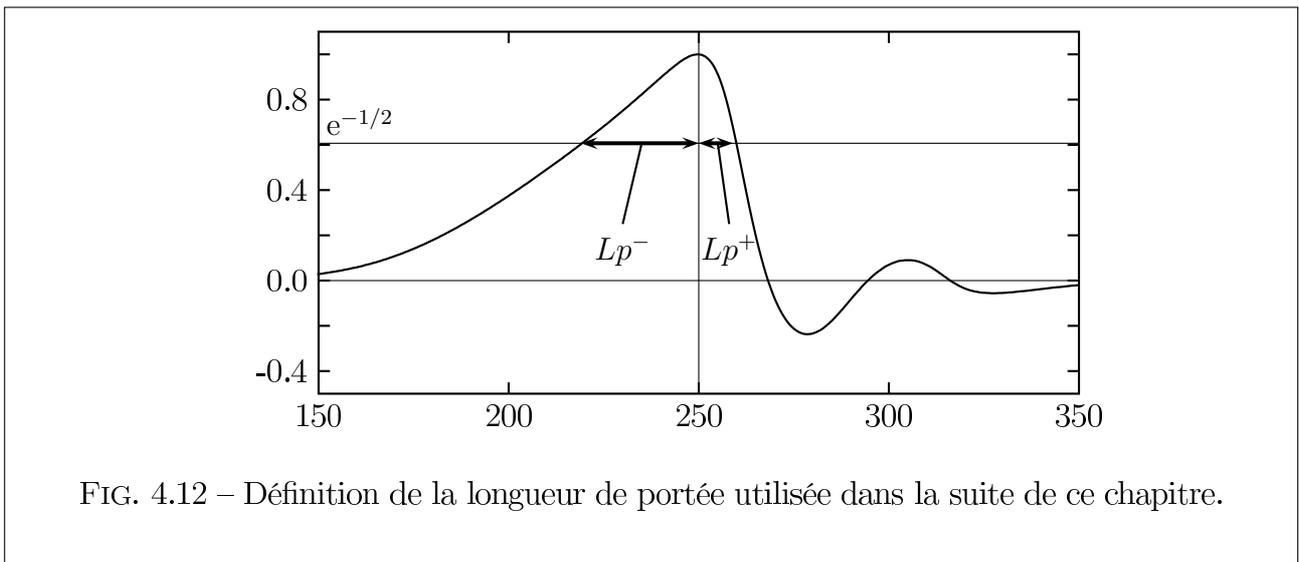
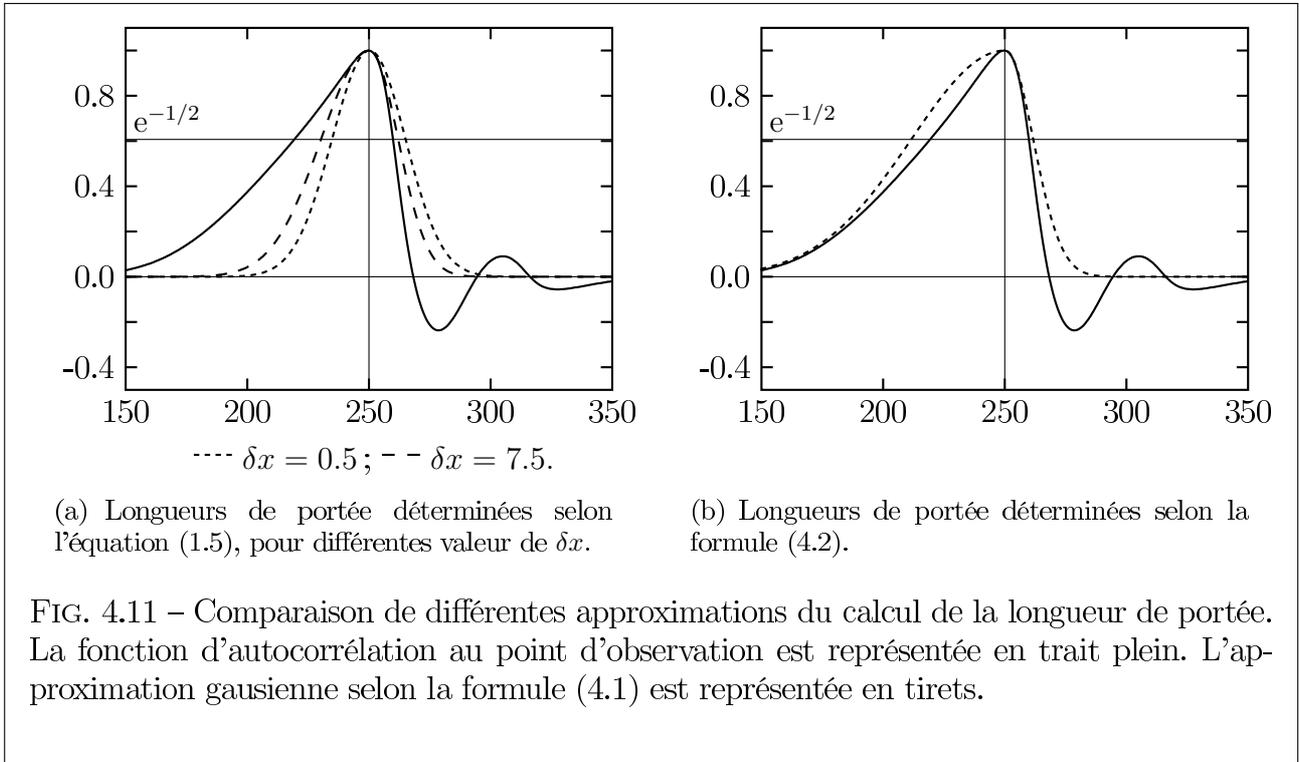
Les graphiques précédents (figure 4.10) ont montré que les fonctions d'autocorrélation au point d'observation ne sont plus gaussiennes. L'estimation des longueurs de portée selon la formule (Eq. 4.1), basée sur l'approximation de fonctions gaussiennes, est fortement dépendante de l'écart au point d'observation, δx , utilisé. La figure 4.11 illustre les approximations gaussiennes obtenues pour différentes valeurs de δx au moyen des fonctions :

$$\rho(dx) = \begin{cases} \exp\left(-\frac{dx^2}{2Lp^-}\right) & \text{pour } dx \leq 0; \\ \exp\left(-\frac{dx^2}{2Lp^+}\right) & \text{pour } dx \geq 0. \end{cases} \quad (4.1)$$

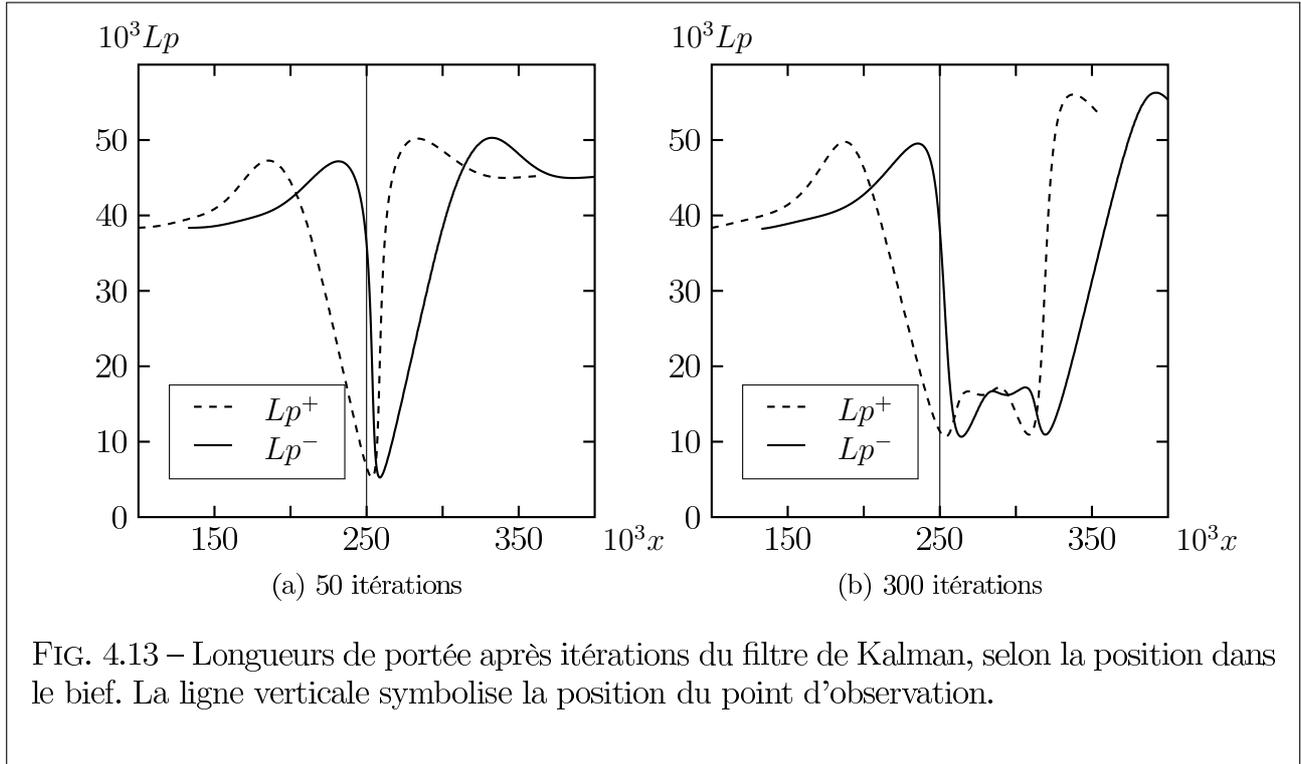
Il en ressort que cette méthode d'approximation traduit mal le caractère anisotropique de la fonction d'autocorrélation pour des valeurs trop petites de δx , et oblige à un choix arbitraire de δx . L'estimation de l'anisotropie par le quotient Lp^+/Lp^- devient alors dépendante non seulement des paramètres du modèle physique, mais également de ceux du modèle numérique.

La méthode proposée, illustrée par la figure 4.12, consiste à définir Lp^+ et Lp^- comme étant

4.4. IMPLÉMENTATION DU FILTRE DE KALMAN



4.4. IMPLÉMENTATION DU FILTRE DE KALMAN



les longueurs vérifiant :

$$\rho(Lp^+) = \rho(-Lp^-) = \exp\left(-\frac{1}{2}\right). \quad (4.2)$$

Cette définition, aussi arbitraire qu’elle soit, est plus robuste en cas de changement des paramètres du modèle numérique.

4.4.4 Diagnostic des longueurs de portée

Dans l’espace : La figure 4.13 donne les longueurs de portée en fonction de la position dans le bief. Au début de la simulation (figure 4.13(a)), les seuls points impactés se situent autour du point d’observation.

Cette zone s’élargit ensuite vers l’aval, figure 4.13(b). Deux zones présentent alors d’importantes différences entre les longueurs de portée à gauche et à droite; au point d’observation ($x = 250$ km), et en aval ($x = 325$ km). Cette deuxième zone correspond à la frontière jusqu’à laquelle l’advection a transporté l’information.

Conformément au graphique 4.10, chaque point situé en aval passe par trois étapes. Dans un premier temps l’information n’est pas encore parvenue et les longueurs de portée sont isotropes et importantes. Quand l’information arrive, les longueurs de portée deviennent anisotropes, après quoi elles reprennent un caractère isotrope, mais deviennent beaucoup plus réduites.

Après 300 itérations, le graphique 4.13(b) fait apparaître un intervalle dans lequel les longueurs de portée sont réduites, et semblent présenter une légère pente positive due à la diffusion. L’étendue de cette zone correspond à la longueur d’advection obtenue après 300 itérations, soit 75 km. Cela est conforme à la figure 3.2, mais le nombre limité d’itérations possibles ne permet pas de voir si cette tendance se poursuit.

4.4. IMPLÉMENTATION DU FILTRE DE KALMAN

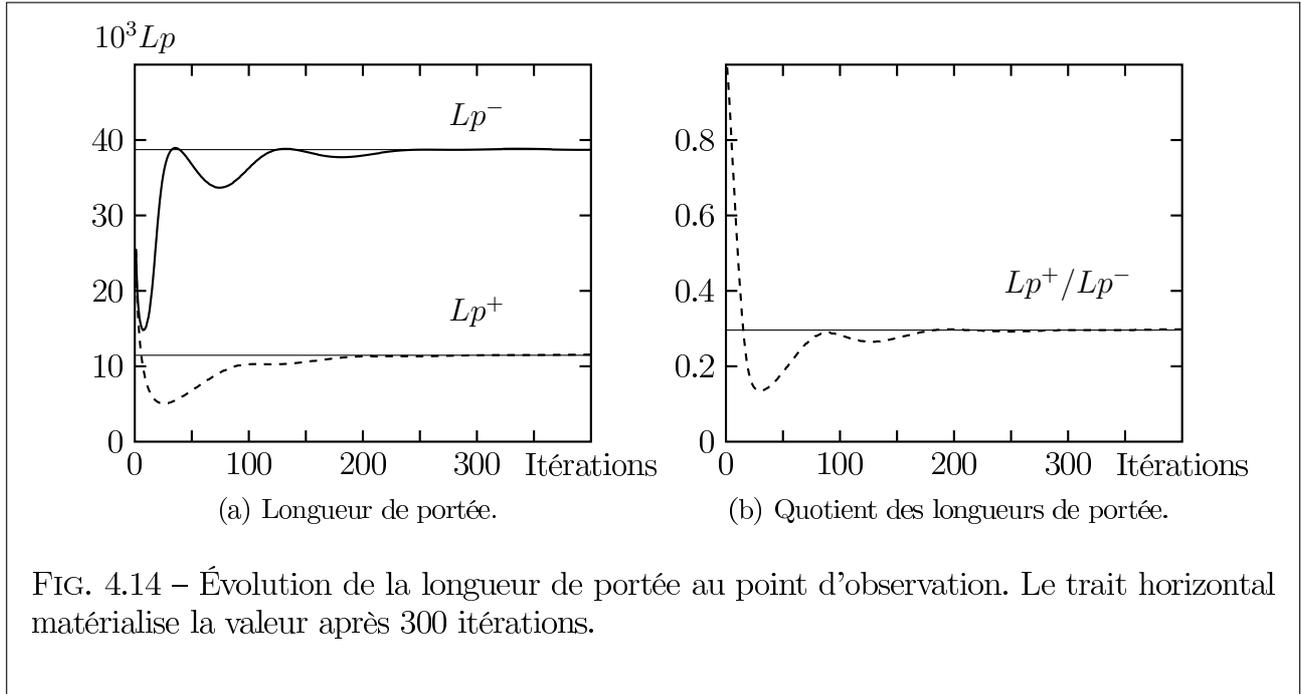


FIG. 4.14 – Évolution de la longueur de portée au point d’observation. Le trait horizontal matérialise la valeur après 300 itérations.

Dans le temps : La figure 4.14 montre l’évolution des longueurs de portée au point d’observation, selon le nombre d’itérations. Compte tenu de l’advection de l’information, les longueurs de portée devraient se stabiliser après un temps de l’ordre de grandeur de la longueur de portée divisée par la vitesse, soit 80 itérations pour une longueur de portée de 40 km. Quelques oscillations sont visibles au delà de 80 itérations, et les longueurs de portée se stabilisent à partir de 150 itérations. Dans la suite, les longueurs de portée au point d’observation sont définies comme celles obtenues après 300 itérations.

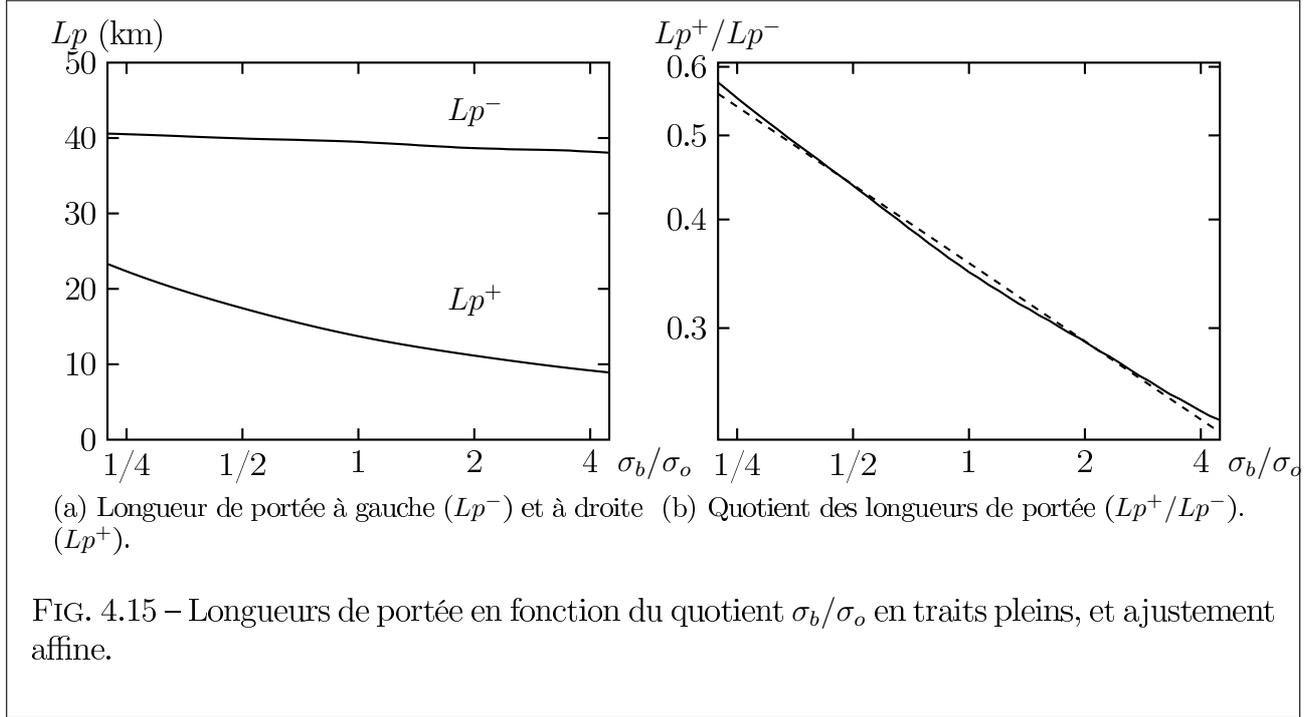
4.4.5 Influence du quotient σ_b/σ_o

Dans toutes les simulations précédentes, la précision de l’observation est restée constante. Cette précision est caractérisée par le quotient σ_b/σ_o , avec σ_b l’écart-type initial d’erreur d’ébauche au point d’observation, et σ_o celui de l’observation.

La figure 4.15 donne le quotient L_p^+/L_p^- de la longueur de portée à droite par la longueur de portée à gauche. Ce quotient dépend essentiellement de la longueur de portée aval. La figure 4.15(a) montre une division de la longueur de portée à droite par un facteur supérieur à deux, alors que la longueur de portée à gauche ne diminue que peu. L’information progressant vers l’aval, la longueur de portée amont n’est que peu influencée par l’observation. La figure 4.15(b), en coordonnées logarithmiques, permet de vérifier que lorsque la précision de l’observation augmente (fort quotient σ_b/σ_o), la longueur de portée aval diminue (faible quotient L_p^+/L_p^-) : Une augmentation de l’erreur d’observation réduit son importance et de ce fait réduit l’anisotropie.

L’objectif est maintenant d’obtenir une relation empirique du quotient L_p^+/L_p^- au point d’observation en fonction du quotient σ_b/σ_o . L’utilisation des coordonnées logarithmiques sur

4.5. IMPLÉMENTATION FILTRE DE KALMAN D'ENSEMBLE



la figure 4.15(b) montre qu'un ajustement affine est justifié.

$$\ln \left(\frac{Lp^+}{Lp^-} \right) = -1.03 \ln \left(\frac{\sigma_b}{\sigma_o} \right) - 0.30.$$

La paramétrisation des longueurs de portée semble donc envisageable.

4.4.6 Influence du temps caractéristique du signal

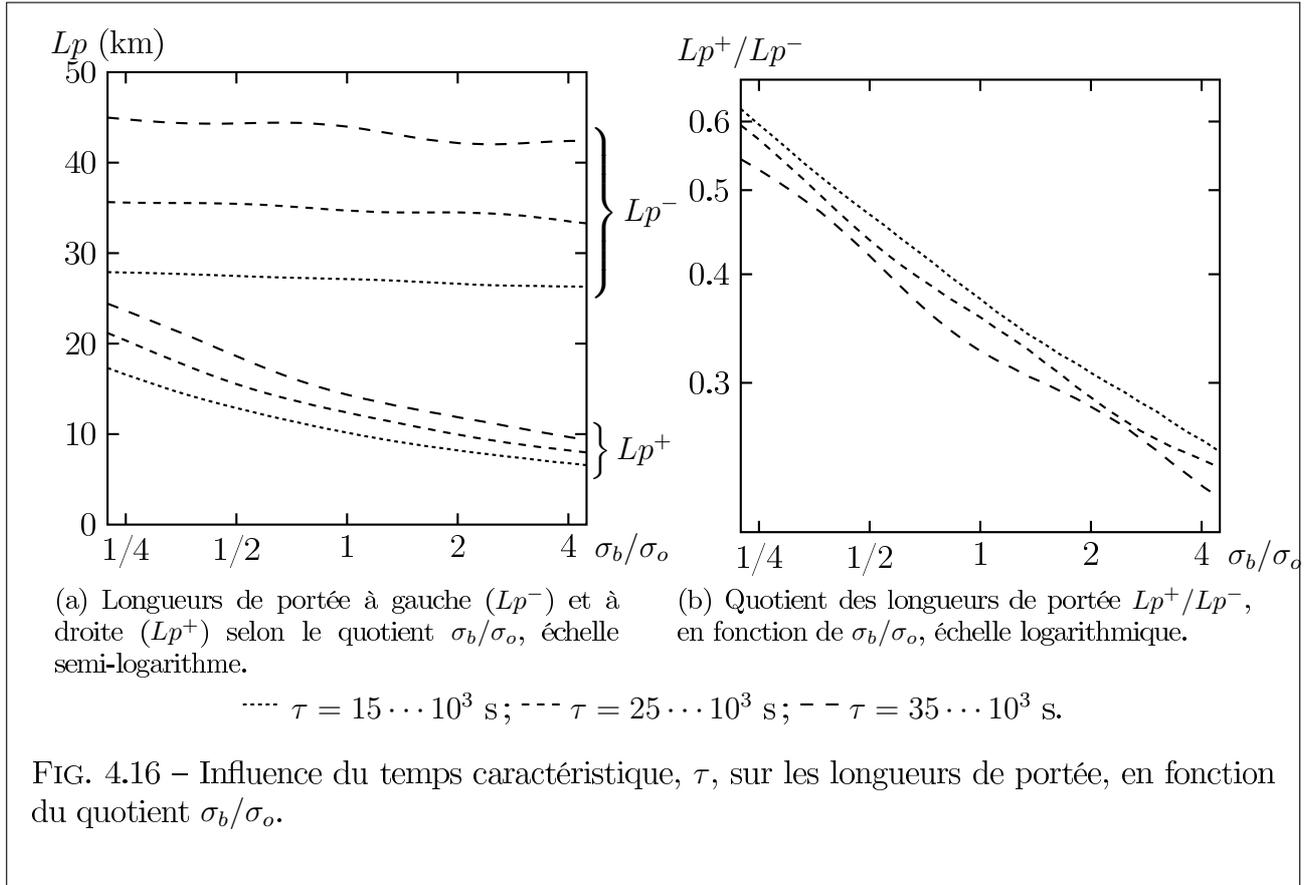
Le temps caractéristique du forçage amont détermine les longueurs de portée initiales, cela se répercute ensuite sur le quotient Lp^+/Lp^- . La figure 4.16 reprend les graphiques de la figure 4.15(b), pour différentes valeurs du temps caractéristique, τ . Le graphique 4.16(a), donnant les longueurs de portée, montre une forte dépendance de la longueur de portée amont. Celle-ci est essentiellement déterminée par les conditions initiales. En revanche, la longueur de portée aval est plus faiblement affectée, la précision de l'observation étant là le paramètre déterminant.

Une augmentation du temps caractéristique a alors comme principale conséquence une augmentation des longueurs de portée amont, puis une réduction du quotient Lp^+/Lp^- , tel qu'illustré sur la figure 4.16(b). Cette figure représente les courbes du quotient Lp^+/Lp^- en fonction de σ_b/σ_o pour différentes valeurs du temps caractéristique.

4.5 Implémentation filtre de Kalman d'ensemble

Le principe du filtre de Kalman d'ensemble a été étudié dans la partie 1.6. Sa mise en œuvre va permettre de lever les limitations vues en 4.4.1. Les paramètres de la simulation restent ceux de la partie 4.4.1, les résultats sont obtenus sur un ensemble de $N = 10000$ membres.

4.5. IMPLÉMENTATION FILTRE DE KALMAN D'ENSEMBLE

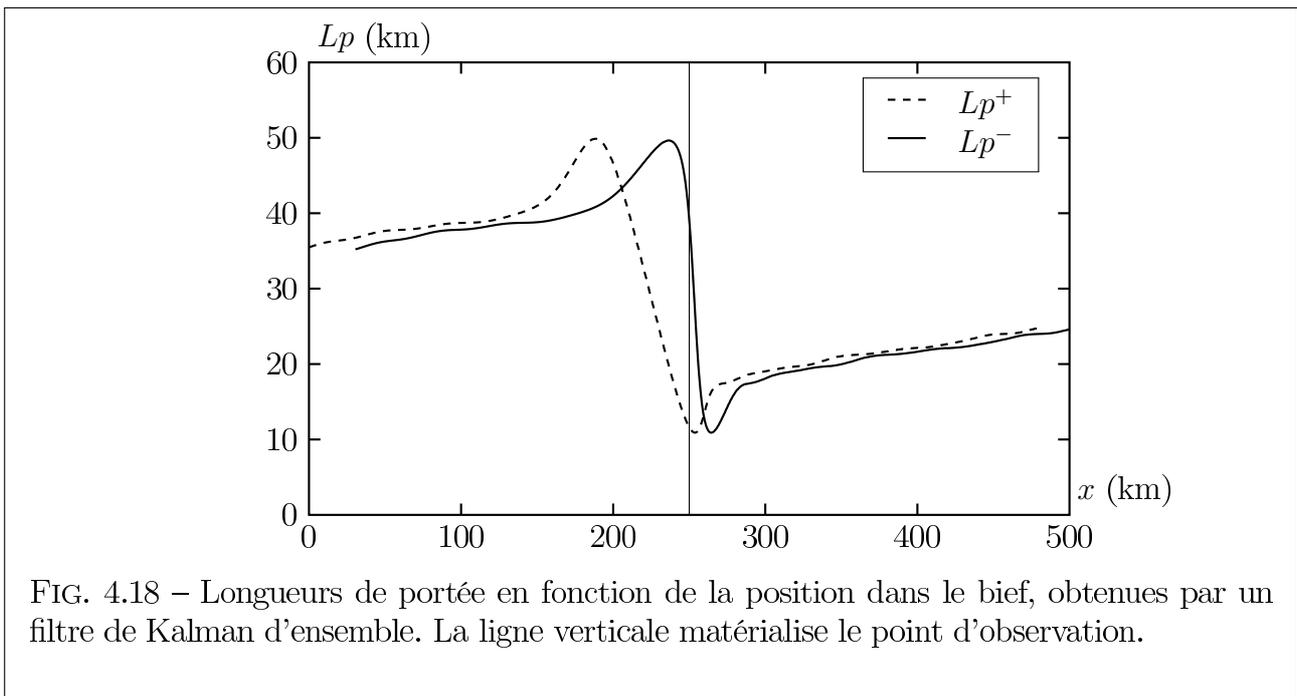
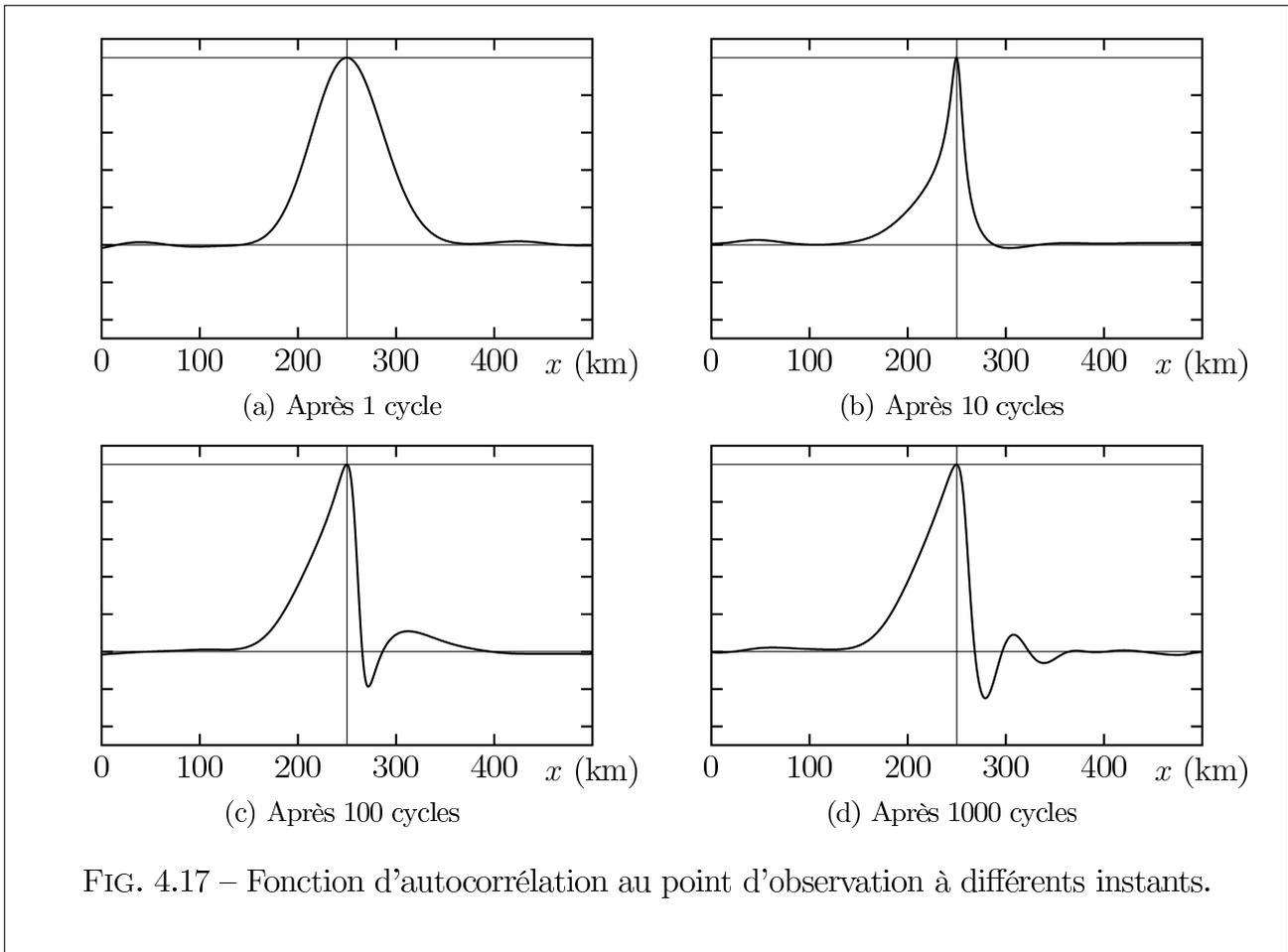


La figure 4.17 donnant la fonction d'autocorrélation au point d'observation, permet de constater qu'une centaine de cycles suffisent à déterminer les longueurs de portée. Au delà, seules quelques petites oscillations se propagent. Ces résultats sont similaires à ceux du filtre de Kalman (Fig. 4.10). L'utilisation d'un filtre de Kalman est donc suffisante.

La figure 4.18 donne les longueurs de portée, tout au long du bief, selon l'équation (4.2). Les calculs sont effectués après 1997 cycles, soit le temps nécessaire pour qu'un écoulement traverse complètement le bief. Au niveau du point d'observation il n'y a pas de différences notables avec les résultats du filtre de Kalman (Fig. 4.13). De plus la courbe obtenue est conforme à l'étude réalisée dans la partie 3.2.

Le filtre de Kalman d'ensemble n'améliore pas la qualité des résultats au point d'observation, il permet néanmoins une vision plus globale des longueurs de portée. Cependant son utilisation pourrait s'avérer nécessaire pour atteindre la convergence des longueurs de portée dans le cas d'une réduction de la fréquence d'observation.

4.5. IMPLÉMENTATION FILTRE DE KALMAN D'ENSEMBLE



Conclusion

Cette étude a permis, par le diagnostic des longueurs de portée, de quantifier le caractère anisotrope des fonctions de covariance d'erreurs d'ébauche en fonction du rapport $\frac{\sigma_b}{\sigma_o}$ au point d'observation. Une partie de l'information sur la dynamique de ces fonctions est représentée par leur caractère anisotrope. Les perspectives immédiates de ces travaux portent sur la modélisation d'une matrice B en but de l'utiliser dans un algorithme de BLUE. Bien que celle-ci reste statique, on peut s'attendre à une amélioration de l'analyse grâce à l'introduction de fonctions de covariance anisotropes correspondantes à la dynamique de l'équation d'onde de crues. Il s'agira donc de vérifier cette hypothèse et de quantifier l'amélioration obtenue dans le modèle d'ondes de crues, avant de tenter une modélisation similaire dans MASCARET.

D'autres pistes auraient également pu être approfondies. En effet l'influence du coefficient de diffusion sur le rapport $\frac{Lp^+}{Lp^-}$ n'a pas été étudiée. De plus il a été décidé d'assimiler une observation à chaque pas de temps. On aurait pu s'intéresser à l'effet d'une fréquence d'assimilation plus faible. Les conséquences peuvent être importantes notamment sur le nombre d'itérations nécessaires pour atteindre un état stationnaire des fonctions de covariance.

Annexe A

Annexe

A.1 Condition de stabilité du schéma numérique

En introduisant les coefficients positifs $\alpha = c \frac{\Delta t}{\Delta x}$ et $\beta = 2\kappa \frac{\Delta t}{\Delta x^2}$ et le réel $\theta = 2\pi\zeta\Delta x$, le gain à chaque cycle du schéma numérique (Eq. 2.5) s'écrit :

$$|g|^2 = \underbrace{\alpha^2 \sin^2 \theta}_A + \underbrace{(1 - \beta + \beta \cos \theta)^2}_B \quad (\text{A.1})$$

Dans l'équation (A.1) il est nécessaire d'avoir $A \leq 1$, d'où la condition $\alpha \leq 1$. Il est également nécessaire d'avoir $B \leq 1$, ce qui est vérifié dès que $\beta \leq 1$.

La dernière condition, $\frac{\alpha^2}{\beta} \leq 1$, lie les paramètres α et β , comme illustré par la figure 2.2. Cette condition peut s'obtenir en étudiant le gain comme un polynôme en $X = \cos \theta$:

$$\begin{aligned} |g|^2 &= ((1 - \beta) + \beta X)^2 + \alpha^2(1 - X^2) \\ &= (\beta^2 - \alpha^2) X^2 + 2(1 - \beta)\beta X + (1 - \beta)^2 + \alpha^2 \end{aligned}$$

- Cas où $\beta^2 - \alpha^2 \geq 0$:

$|g|^2$ prend ses valeurs maximales pour $X = -1$ et $X = 1$.

$|g|_{X=-1}^2 = \beta^2 - \alpha^2 - 2(1 - \beta)\beta + (1 - \beta)^2 + \alpha^2 = 4\beta^2 - 4\beta + 1$. La condition $\beta \leq 1$ entraîne bien $|g|_{X=-1}^2 \leq 1$.

$|g|_{X=1}^2 = \beta^2 - \alpha^2 + 2(1 - \beta)\beta + (1 - \beta)^2 + \alpha^2 = 1$

Dans ce cas aucune condition supplémentaire n'est nécessaire pour obtenir $|g| \leq 1$.

- Cas où $\beta^2 - \alpha^2 \leq 0$:

$|g|^2$ est représenté par une parabole d'axe $X = \frac{(1 - \beta)\beta}{\alpha^2 - \beta^2}$.

Or $\frac{(1 - \beta)\beta}{\alpha^2 - \beta^2} \geq 1 \Leftrightarrow \beta \geq \alpha^2$.

Donc en prenant $\beta \geq \alpha^2$, $|g|^2$ atteint son maximum en $X = 1$ et le calcul précédent fournit $|g|_{X=1}^2 = 1$.

En prenant $\beta < \alpha^2$, $|g|^2$ atteint son maximum en $X = \frac{(1 - \beta)\beta}{\alpha^2 - \beta^2}$.

$|g|_{X=\frac{(1-\beta)\beta}{\alpha^2-\beta^2}}^2 = \frac{\alpha^2(1-\beta)^2}{\alpha^2 - \beta^2} + \alpha^2$. On a alors $|g|_{X=\frac{(1-\beta)\beta}{\alpha^2-\beta^2}}^2 \leq 1 \Leftrightarrow \beta^2 - 2\alpha^2\beta + \alpha^4 \leq 0$. L'étude

A.2. INFLUENCE DU SCHEMA NUMERIQUE

du polynôme $\beta \mapsto \beta^2 - 2\alpha^2\beta + \alpha^4$ montre qu'aucune valeur de β ne permet d'obtenir un schéma stable.

– Cas où $\beta^2 - \alpha^2 = 0$:

$|g|^2 = 2(1 - \beta)\beta X + (1 - \beta)^2 + \alpha^2$ atteint son maximum en $X = 1$.

$|g|_{X=1}^2 = 2(1 - \beta)\beta + (1 - \beta)^2 + \alpha^2 = 1 + \alpha^2 - \beta^2 = 1$.

Dans ce cas aucune condition supplémentaire n'est nécessaire pour obtenir $|g| \leq 1$.

La stabilité du schéma numérique est définie par trois inéquations :

$$\alpha \leq 1; \quad \beta \leq 1; \quad \frac{\alpha^2}{\beta} \leq 1$$

Soit :

$$c \frac{\Delta t}{\Delta x} \leq 1; \quad 2\kappa \frac{\Delta t}{\Delta x^2} \leq 1; \quad \frac{c^2 \Delta t}{2\kappa} \leq 1$$

A.2 Influence du schéma numérique

L'équation 2.7, ou équation équivalente résolue par le schéma numérique fait apparaître un terme de dispersion. Ce terme, en $\frac{\partial^3 h}{\partial x^3}$, affecte tout particulièrement les petites échelles.

L'équation 2.7 présente également un coefficient de diffusion $\kappa' = \kappa - \frac{c^2 \Delta t}{2}$ cohérent avec les conditions de stabilité établies précédemment et dans lesquelles figure l'inéquation $\kappa - \frac{c^2 \Delta t}{2} \geq 0$. Cette inéquation s'interprète alors comme la nécessité d'avoir un coefficient de diffusion positif.

Le schéma numérique tend donc à sous-estimer la diffusion. Le graphique A.1 montre l'évolution d'une fonction gaussienne $h(x, t)$, en prenant comme condition initiale une fonction de solution analytique connue.

$$h_\kappa(x, t) = q_m \frac{l_0}{l^2(t)} \exp \left[-\frac{x - x_0 - ct}{2l^2(t)} \right], \quad l^2(t) = l_0^2 + 2\kappa t.$$

La fonction simulée est moins diffusée qu'attendu par la solution analytique, cela étant d'autant plus vrai que la valeur de κ est petite.

La fonction h simulée reste une gaussienne, advectée à la même vitesse que la solution analytique. À chaque instant, elle coïncide avec une solution analytique de coefficient de diffusion $\kappa_n(t)$, tel que $h(x, t) = h_{\kappa_n(t)}(x, t)$.

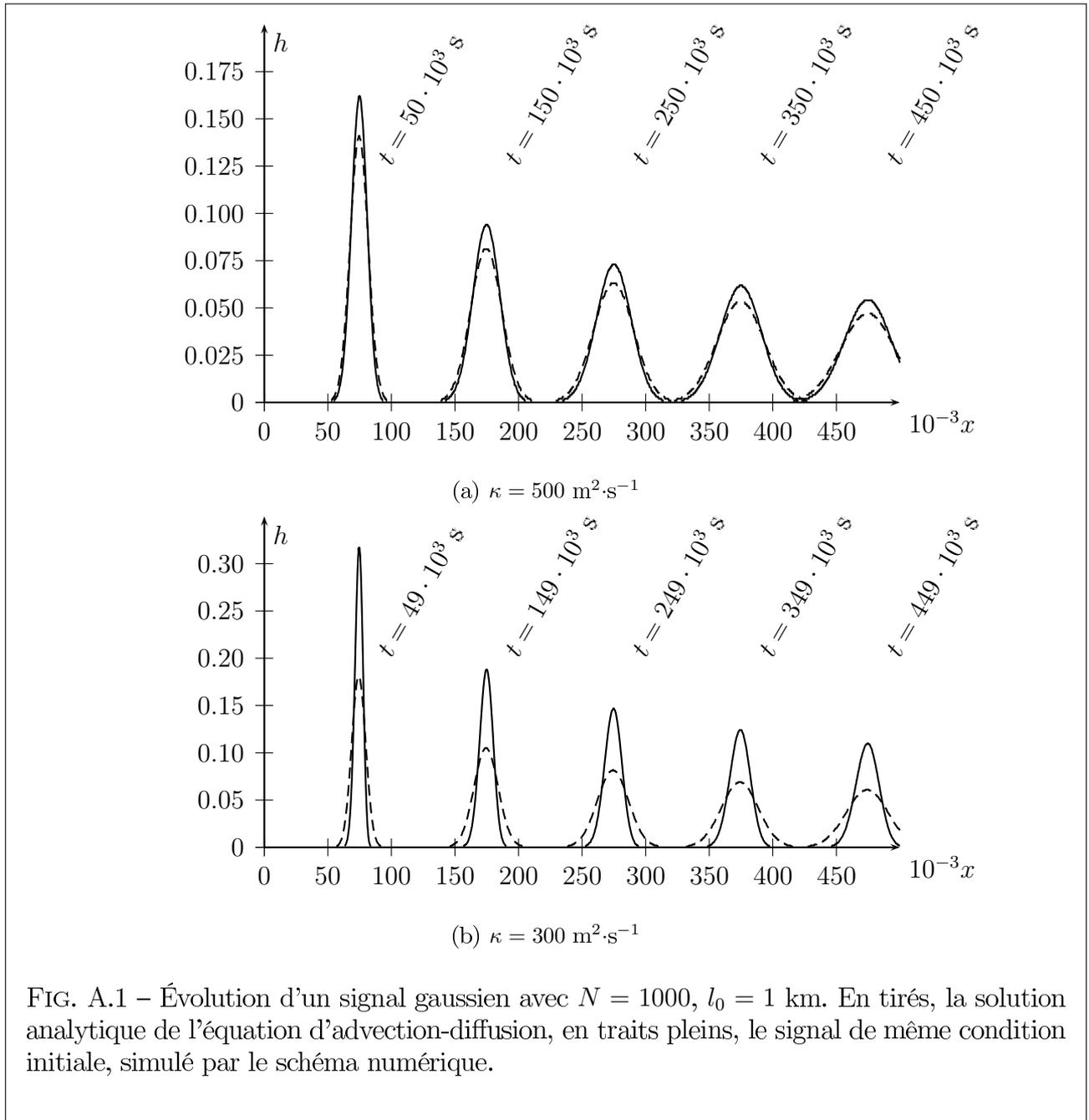
$\kappa_n(t)$, peut-être considéré comme le coefficient de diffusion effectif du schéma numérique. Son estimation permet de vérifier que le coefficient κ' , introduit dans l'équation 2.7, est bien représentatif de la diffusion réelle du schéma.

Il est déterminé en considérant le maximum $h_m(t)$ de $h(x, t) = h_{\kappa_n(t)}(x, t)$, et vérifie ainsi $h_m^2(t) = q_m^2 \frac{l_0^2}{l_0^2 + 2\kappa_n(t)t}$. D'où une estimation de $\kappa_n(t)$ par :

$$\kappa_n(t) = \left[\left(\frac{q_m}{h_m(t)} \right)^2 - 1 \right] \frac{l_0}{2t}.$$

Le graphique A.2, représentant l'évolution de $\kappa_n(t)$ à chaque pas de temps, montre une convergence de $\kappa_n(t)$. Cette valeur limite, κ_n , peut-être considérée comme le coefficient de diffusion du schéma numérique. Les fluctuations observées lors des premiers pas de temps sont liées

A.2. INFLUENCE DU SCHEMA NUMERIQUE



A.2. INFLUENCE DU SCHEMA NUMERIQUE

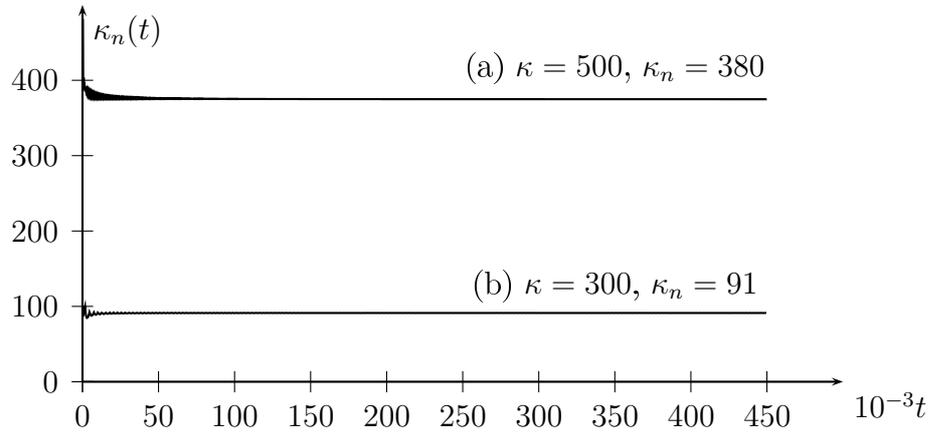


FIG. A.2 – Coefficient κ_n permettant à chaque itération d’obtenir une courbe théorique identique à la courbe simulée. Les simulations sont effectuées avec les paramètres $\Delta x = 500$ m et $c = 1 \text{ m}\cdot\text{s}^{-1}$.

(a) Courbe réalisée avec $\kappa = 500 \text{ m}^2\cdot\text{s}^{-1}$. La condition de stabilité donne $\Delta t = 250$ s, donc

$$\kappa' = \kappa - \frac{c^2 \Delta t}{2} = 375 \text{ m}^2\cdot\text{s}^{-1}, \text{ alors que } \kappa_n = 380 \text{ m}^2\cdot\text{s}^{-1}.$$

(b) Courbe réalisée avec $\kappa = 300 \text{ m}^2\cdot\text{s}^{-1}$, La condition de stabilité donne $\Delta t = 416.7$ s, donc $\kappa' = 91.7 \text{ m}^2\cdot\text{s}^{-1}$, alors que $\kappa_n = 91$.

aux petites échelles, particulièrement sensibles aux termes en $\frac{\partial^3 h}{\partial x^3}$ et $\frac{\partial^4 h}{\partial x^4}$. Ces petites échelles sont rapidement diffusées, d’où la convergence de $\kappa_n(t)$.

Dans chacun des cas le coefficient κ' permet d’expliquer la majeure partie de l’écart entre la diffusion souhaitée, κ , et la diffusion effective du schéma numérique. La prise en compte d’une correction du coefficient de diffusion κ devra donc être effectuée lorsque le but d’une simulation est de tester la dépendance d’un résultat par rapport au coefficient de diffusion.

Bibliographie

- F. Bouttier and P. Courtier. Data assimilation concepts and methods, 1999.
- R. Daley. *Atmospheric data analysis*. Cambridge atmospheric and space science series. Cambridge University Press, 1991.
- G. Desroziers. *Assimilation de données, notes de cours*. Ecole Nationale de la Météorologie, 2010.
- Geir Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *J. Geophys. Res.*, 99(C5) :10143–10162, 1994.
- L.Y.M. Gicquel and O. Vermorel. Computational fluid dynamics (cfd) : Discretization, stability, dispersion and dissipation. 2010.
- O. Pannekoucke. Note interne : Dynamique des covariances, 2010.
- O. Pannekoucke, L. Berre, and G. Desroziers. Background-error correlation length-scale estimates and their sampling statistics. *Quarterly Journal of the Royal Meteorological Society*, 2007.
- D. Revuz. *Mesure et intégration*. Collection Méthodes. Hermann, 1997.
- S. Ricci, A. Piacentini, O. Thual, E. Le Pape, and G. Jonville. Correction of upstream flow and hydraulic state with data assimilation in the context of flood forecasting. 2010.
- A. Tarantola. *Inverse problem theory : Methods for data fitting and model parameter estimation*. Elviesier, 1987.
- O. Thual. *Hydrodynamique de l'environnement*. Ecole Polytechnique, 2010.

Annexe 5 : Slides présentation CERFACS 29 Avril 2011



Correction de l'état hydraulique et des lois d'apport par assimilation de données.

S. Ricci ¹, A. Piacentini ¹, O. Thual ³, G. Jonville ¹,
E. Le Pape ², B. Janet ², J. Habert ², N. Goutal ⁴, F. Zaoui ⁴, R. Ata ⁴

¹ CERFACS, ² SCHAPI, ³ IMFT, ⁴ LNHE

CERFACS, 29 Avril 2011

Outline

- 1 Assimilation de données : Méthodologie
- 2 Modélisation de l'opérateur d'observation
- 3 Modélisation de **B**
- 4 Resultats des simulations de crues et prévision

Sources des incertitudes : choix du vecteur de contrôle

Objectif

Améliorer la qualité de la simulation de l'état hydraulique 1D, dans le contexte de la prévision des crues, avec le code MASCARET.

Réduction des incertitudes

La mise en place d'une méthode d'assimilation de données a pour but de réduire les incertitudes sur les entrées et/ou les sorties des simulations MASCARET en utilisant des observations de hauteurs d'eau.

⇒ On appelle *vecteur de contrôle*, le vecteur qui rassemble les variables à corriger : paramètres, variables d'état, forçages ...

Le vecteur de contrôle pour la procédure d'assimilation de données en 2 étapes.

Une procédure en 2 étapes :

Step 1 : Correction des lois amonts

Correction paramétrisée des débits aux stations amonts, notées i , sur une période d'intégration : $\tilde{q}_i(t) = a_i q_i(t - c_i) + b_i$

⇒ Vecteur de contrôle $\mathbf{x} = (\dots, a_i, b_i, c_i \dots)$

Step 2 : Correction de l'état hydraulique

Correction instantanée de la hauteur d'eau et du débit, aux temps d'observation, sur l'ensemble du réseau hydraulique

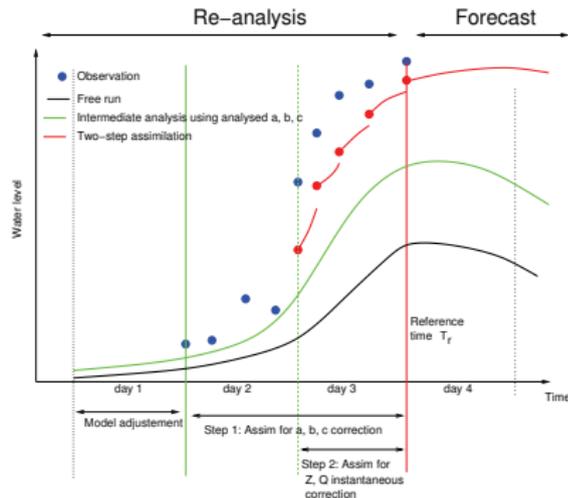
⇒ Vecteur de contrôle $\mathbf{x} = (Z_{x_1}, \dots, Z_{x_m}, Q_{x_1}, \dots, Q_{x_m})$

- Le vecteur d'observations rassemble les mesures de hauteur d'eau
 $\mathbf{y} = (Z_{obs_1}, \dots, Z_{obs_N})$
- Les 2 corrections sont appliquées séquentiellement.

Cyclage de l'algorithme

L'analyse est faite sur une fenêtre glissante de 4 jours, 3 jours en ré-analyse, 1 jour en prévision (pour l'exemple).

- Analyse à la fin du jour 3
- La hauteur d'eau est observée à plusieurs temps et positions entre les jours 1 et 3.
- Ajustement libre du modèle sur le jour 1.
- Correction des débits en amont avec les obs. des jours 2 et 3.
- Correction de l'état hydraulique avec les obs. du jour 3.



Algorithme d'assimilation de données.

Définition des matrices de covariance d'erreur

Erreur d'ébauche

$$\mathbf{x}^t - \mathbf{x}^b = \varepsilon^b$$

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

Erreur d'observation

$$H(\mathbf{x}^t) - \mathbf{y}^o = \varepsilon^o$$

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

Erreur d'analyse

$$\mathbf{x}^t - \mathbf{x}^a = \varepsilon^a$$

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

Filtre de Kalman Invariant

L'estimateur optimal \mathbf{x}^a du vecteur de contrôle est formulé comme la combinaison linéaire de l'ébauche et des observations

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

- La matrice de gain matrix \mathbf{K} est invariante car la matrice de covariance d'erreur d'ébauche \mathbf{B} n'est pas propagée par la dynamique du modèle.
- H est l'opérateur d'observation qui relie l'espace de contrôle à l'espace des observations. \mathbf{H} est la matrice Jacobienne de H .

Step 1 : Correction des débits amonts

Vecteur de contrôle

$$\mathbf{x} = (\dots, a_i, b_i, c_i \dots) \text{ tel que } \tilde{q}_i(t) = a_i q_i(t - c_i) + b_i$$

$$\mathbf{y} = (Z_{obs_1}, \dots, Z_{obs_N})$$

La relation entre le contrôle \mathbf{x} et l'espace des observations Z_{obs} est non linéaire puisqu'elle implique l'intégration du modèle numérique M .

Calcul approximatif du linéaire tangent de MASCARET

- Calcul local de \mathbf{H} and \mathbf{H}^T au voisinage de l'ébauche.
- \mathbf{H} est approximé avec un schéma aux différences finies, décentré.

Développement de Taylor

$$\mathbf{H} = \left. \frac{\partial H(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}^b} \text{ est approximé par } \mathbf{H} \approx \frac{H(\mathbf{x}^b + \delta \mathbf{x}) - H(\mathbf{x}^b)}{\delta \mathbf{x}}$$

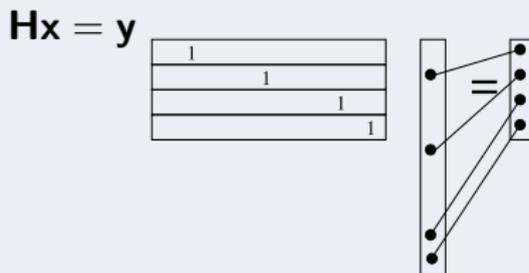
Step 2 : Correction de l'état hydraulique

Vecteur de contrôle

$$\mathbf{x} = (Z_{x_1}, \dots, Z_{x_m}, Q_{x_1}, \dots, Q_{x_m})$$

$$\mathbf{y} = (Z_{obs_1}, \dots, Z_{obs_N})$$

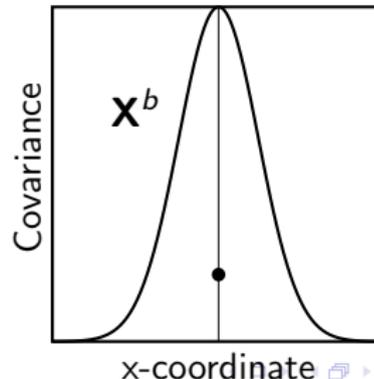
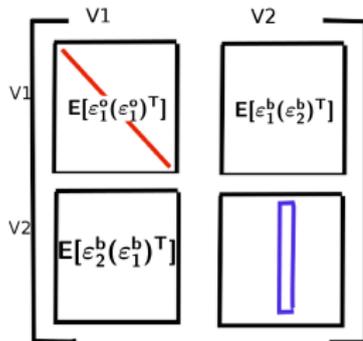
L'opérateur d'observation se résume à une matrice de sélection dont les valeurs non nulles correspondent aux positions des observations sur le vecteur \mathbf{x} .



Modélisation de la matrice de covariances d'erreur d'ébauche (**B**) pour le Step 2.

La matrice de covariances d'erreur d'ébauche traduit la correction en hauteur d'eau aux points d'obs

- en une correction en hauteur d'eau sur tout le domaine,
- en une correction en débit sur tout le domaine.



Les 4 étapes clés pour modéliser B

Variances

Quelles variances pour l'erreur d'ébauche et l'erreur d'observation, à chaque point d'observation ? Sens physique ?

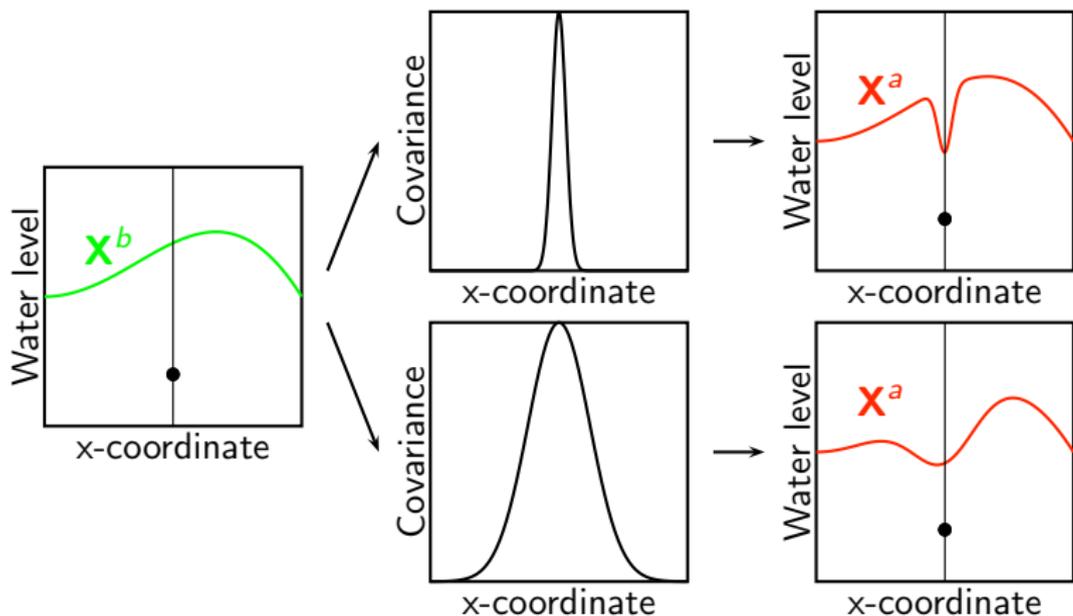
Covariances univariées en Z

- Quelle forme fonction de covariance spatiale en Z, à chaque point d'observation ? Gaussienne ?
- Quelle longueur de portée pour cette fonction ?

Covariances multivariées Z,Q

Quelle fonction de covariance multivariée entre Z et Q à chaque point d'observation ? Courbe de tarage ?

Influence de la fonction de covariance spatiale



Background
Observation

Covariance function
at observation point

Analysis

La fonction de covariance spatiale dans **B** pour **Z**.

Comment spécifier la fonction de covariance spatiale ?

- Une solution simple : Gaussienne. Est-ce représentatif de la corrélation spatiale des erreurs par le modèle, en présence d'observations ?
- Une solution correcte serait donnée par la Filtre de Kalman qui propage explicitement la matrice de covariance **A** : $\mathbf{B} = \mathbf{MAM}^T$. Mais ceci est coûteux et requiert la formulation du tangent linéaire du modèle **M** et son adjoint \mathbf{M}^T .
- Une solution correcte serait donnée par une approche ensembliste $\mathbf{B}_e = \frac{1}{m} \sum_{k=1}^m ((\mathbf{X}_k - \bar{\mathbf{X}})(\mathbf{X}_k - \bar{\mathbf{X}})^T)$. Mais ceci est coûteux dans le cadre de la prévi opérationnelle. (A explorer néanmoins)

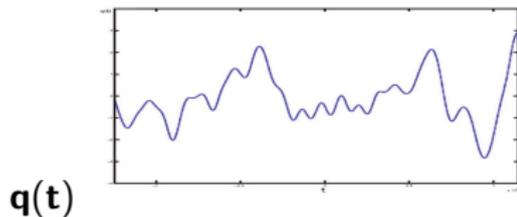
⇒ La forme de la fonction de covariance d'erreur d'ébauche au point d'obs est paramétrée pour émuler la propagation des covariances par la dynamique du modèle et l'analyse.

Approximation des ondes de crue diffusantes

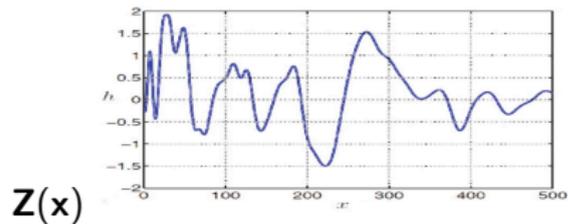
Propagation d'une fonction de covariance

- Implémentation du FK sur le modèle d'onde de crue diffusante.
- Propagation d'une fonction de covariance Gaussienne jusqu'à convergence du KF.
- Caractérisation de la fonction de covariance résultante.

Condition limite $q(t)$
avec une fonction de covariance
temporelle Gaussienne

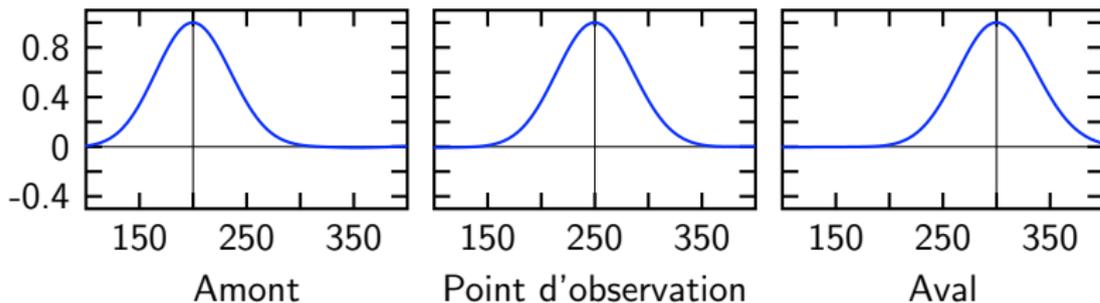


Hauteur d'eau simulée $Z(x)$
avec une fonction de covariance
spatiale Gaussienne

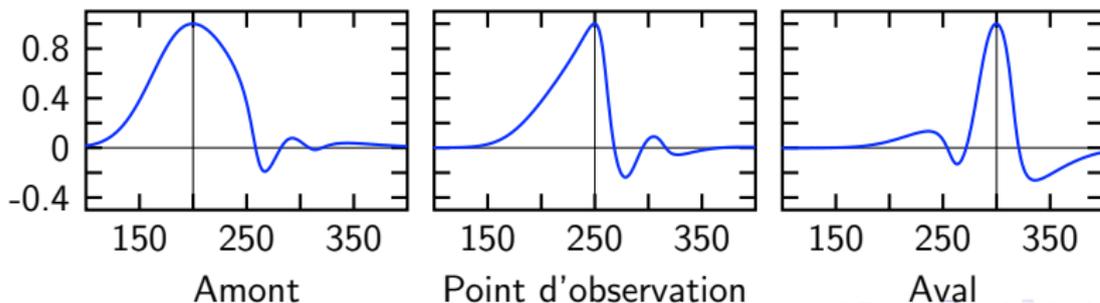


Kalman Filter covariance functions

Sans observation



Avec observations



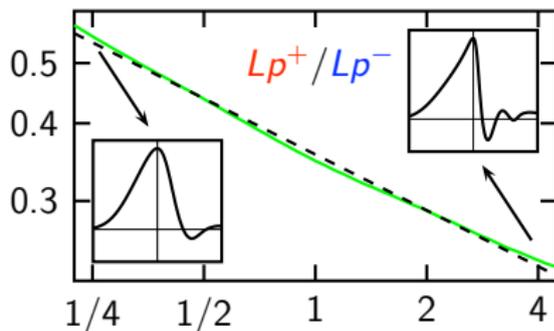
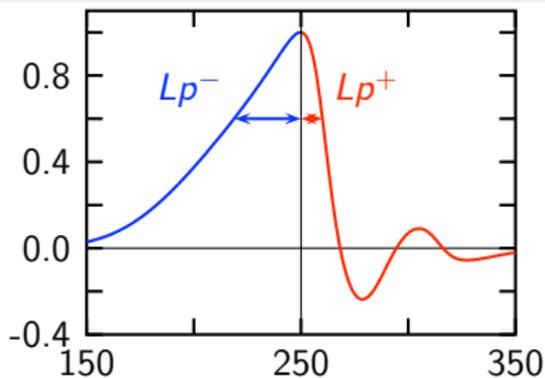
Caractérisation de la fonction de covariance propagée

Forme de la fonction de covariance

- La fonction de covariance au point d'observation est anisotrope selon le sens de l'écoulement.
- La longueur de corrélation est réduite en aval du point d'observation.

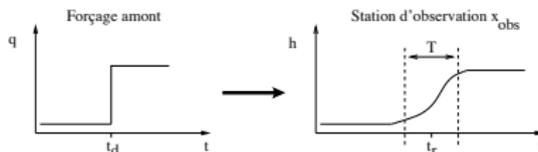
Sensibilité aux variances d'erreur

- Qd les obs sont incertaines ($\sigma_o^2 \gg \sigma_b^2$), l'anisotropie est faible $Lp^+ / Lp^- \approx 1$.
- Qd les observations sont certaines ($\sigma_o^2 \ll \sigma_b^2$), l'anisotropie est forte $Lp^+ / Lp^- < 1$.



Propagation d'une perturbation du forçage amont

On estime le coefficient de diffusion κ , la célérité des ondes c , puis la longueur de portée de la fonction de covariance L_p



- t_r : durée entre le temps de perturbation t_d à l'amont et l'arrivée de la perturbation à la station d'obs
- s_r : distance curviligne entre la station amont et la station d'observation
- $l^2(t_0)$: longueur de corrélation temporelle de la perturbation en amont (≈ 0).

$$c = \frac{s_r}{t_r} \quad | \quad l^2(t_r) = cT \quad | \quad l^2(t_r) = l^2(t_0) + 2\kappa t_r \quad \Rightarrow \quad \kappa = \frac{T^2 l_r^2}{2t_r^3}$$

Longueur de portée de la fonction de covariance :

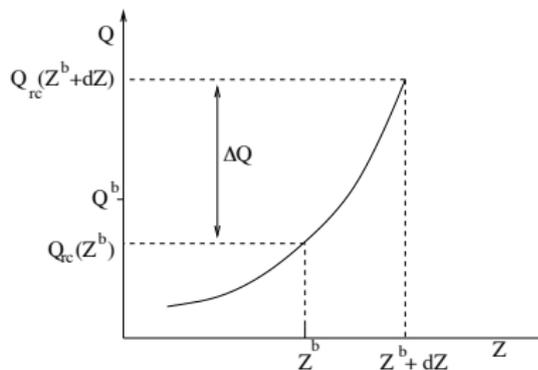
$$L_p(x)^2 = L_p(x_0)^2 + 4\kappa \frac{x_{obs}}{c}$$

Estimation locale de la courbe de tarage

On cherche α, γ, β tels que $Q_{rc} = \alpha Z^\gamma + \beta$

L'incrément en débit dQ s'écrit

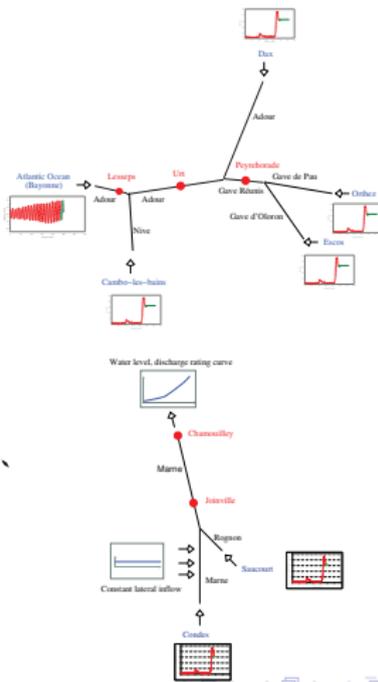
$$dQ = Q^b \frac{Q_{rc}(Z^b + dZ) - Q_{rc}(Z^b)}{Q_{rc}(Z^b)}$$



- si (Z^b, Q^b) suit la courbe de tarage : $dQ = \Delta Q$
- si $Q_{rc} = I$: $dQ = \frac{Q^b}{Z^b} dZ$

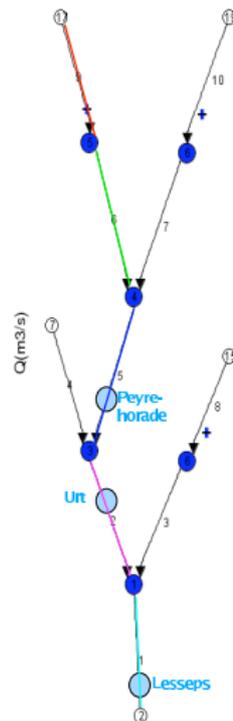
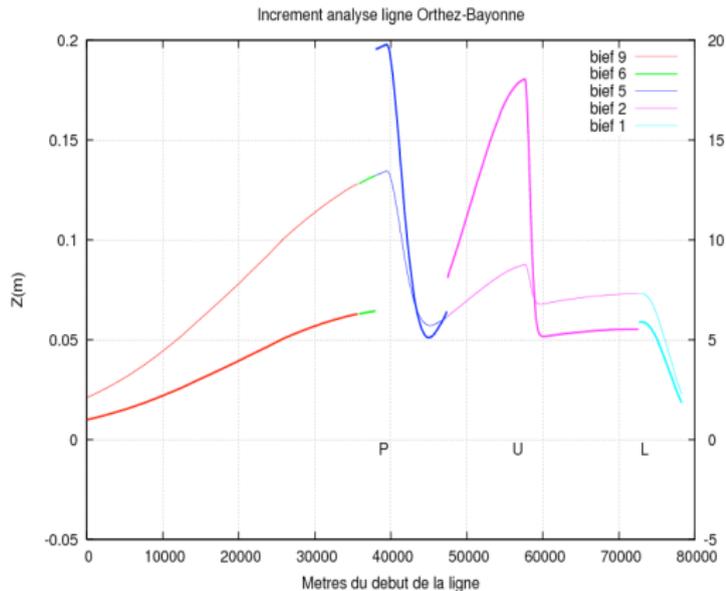
Application sur les bassins Adour Maritime et Marne Vallage

- 3 points d'obs : Lesseps, Urt, Peyrehorade
 - 4 débits amonts (12 paramètres)
 - 1 station maritime (pas contrôlée)
-
- 2 observation points : Joinville, Chamouilley
 - 2 débits amonts + 1 apport latéral (7 parameters)

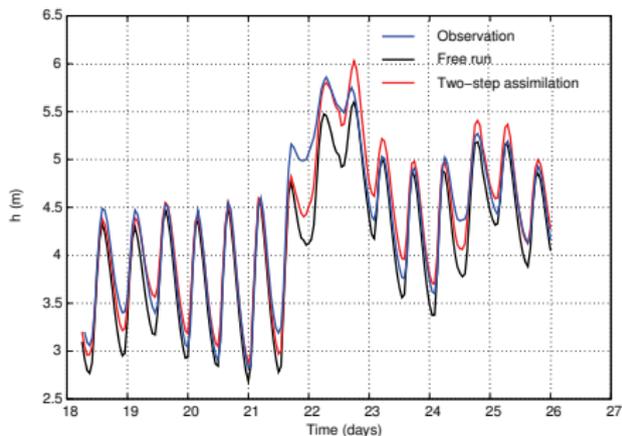
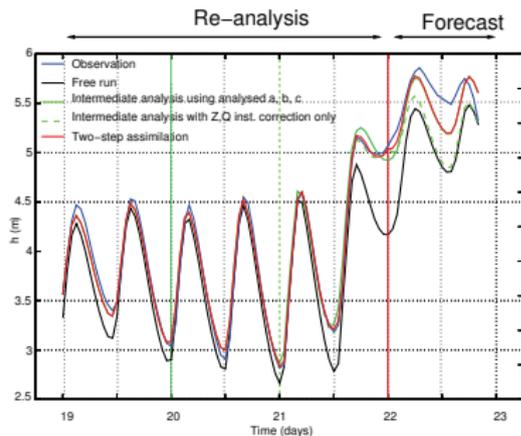


Correction spatiale en hauteur d'eau et débit

- Continuité de Z aux points de jonction.
- Correction en Q déduite de la correction en Z par le biais de la courbe de tarage locale.
- Additivité de Q aux jonctions.



Crue Novembre 2002, à Peyrehorade



- Bonne ré-analyse entre les jours 19 et 22.
- Faible prédictibilité en prévision avec la correction instantanée. La correction des débits amonts en revanche, permet d'être prédictif.

Critères pour l'interprétation statistique

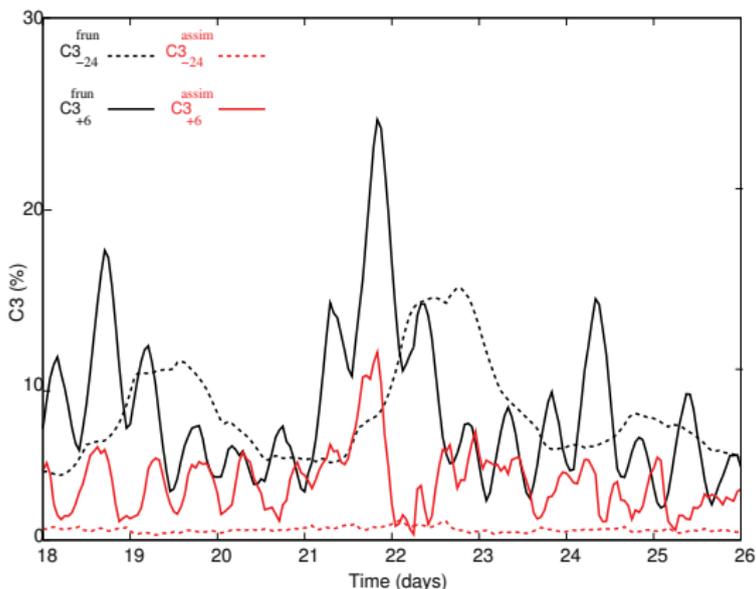
$$MmO = Z_{mod} - Z_{obs}$$

L'écart entre la simulation et l'observation MmO est défini pour les runs libre, assimilation linéique, calage et calage + assim, à t donné, aux points et temps d'observation.

$$C1 = \text{mean}(MmO) \text{ à } t \text{ donné} \quad | \quad C2 = \text{std}(MmO) \text{ à } t \text{ donné}$$

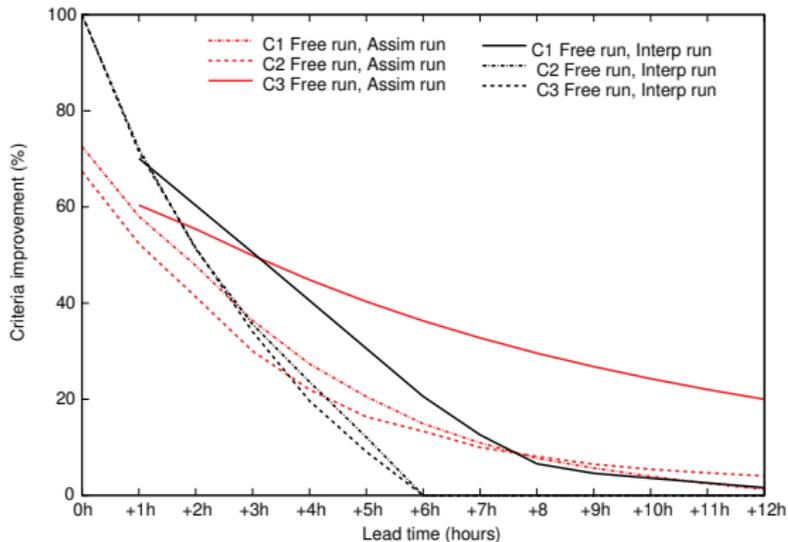
$$C3 = \frac{100}{N_{obs}} \sum_1^{N_{obs}} \frac{Z^{mod} - Z^{obs}}{Z^{obs}} \text{ entre } T_r \text{ et } t$$

Critère C3, Crue Novembre 2002, à Peyrehorade



Réduction de l'écart aux observations en ré-analyse et à +6h sur tout l'épisode de crue.

Amélioration des critères à Peyrehorade



Amélioration des 3 critères entre + 0h et 12h. Le post-traitement permet d'être très bon pour les très courtes échéances. Au delà de +3h, l'amélioration liée à l'assimilation est plus forte.

Amélioration de C1, C2 et C3 entre le Free run et Assim+Calage

-24h re-analysis (Adour)	Peyrehorade	Urt	Lesseps
C1 improvement (%)	72	60	54
C2 improvement (%)	67	58	50
C3 improvement (%)	80	65	54

+6h forecast (Adour)	Peyrehorade	Urt	Lesseps
C1 improvement (%)	15	15	0.5
C2 improvement (%)	13	11	0
C3 improvement (%)	36	25	3

C3 improvement (%)	Chamouilley	Joinville
-24h re-analysis	41	24
+6h forecast	24	10

Conclusions et perspectives

- Amélioration de la qualité de la simulation et de la prévision.
 - Rôle prépondérant du contrôle des débits amonts pour la prévision à longues échéances
 - Rôle prépondérant de la correction de l'état hydraulique pour la ré-analyse la prévision à courte échéance.
 - Hypothèses fortes sur la paramétrisation de la correction de $q(t)$ et de la modélisation de la matrice B .
-
- Extension du vecteur de contrôle : bathymétrie, coefficients de Strickler ...
 - Méthodes avancées : approche ensembliste, minimisation sans gradient.

Annexe 6 : Guide d'utilisation de la machine virtuelle SCHAPI



Direction Générale de la Prévention des Risques
SCHAPI (Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations)

M A S C A R E T A S S I M

Utilisation de la machine virtuelle
et de la maquette Cerfacs

Liste des révisions de ce document:

Révisi on	Date	Rédaction	Contenu
1a	22/03/11	Philippe Guoin	First shot
1b	14/04/11	Johan Habert	Ajout du paragraphe 4

Table des matières

<u>1.INSTALLATION DE VIRTUALBOX.....</u>	<u>3</u>
1.1.Introduction.....	3
1.2.Version à utiliser.....	3
1.3.Installation.....	3
<u>2.MISE EN PLACE DE LA MACHINE VIRTUELLE « MASCARET ASSIM ».....</u>	<u>4</u>
2.1.Récupération de l'archive.....	4
2.2.Importation de la machine dans le virtualiseur.....	4
2.3.Paramétrage final de la machine Mascaret.....	4
<u>3.UUTILISATION DE LA MACHINE VIRTUELLE « MASCARET ASSIM ».....</u>	<u>5</u>
<u>4.UUTILISATION DE LA MAQUETTE CERFACS « MASCARET ASSIM 2 ».....</u>	<u>5</u>
4.1.Présentation générale.....	5
4.2.Initialisation des répertoires.....	6
<u>ANNEXE 1 - CARACTÉRISTIQUES DE LA MACHINE VIRTUELLE</u>	
<u>« MASCARET ASSIMILATION ».....</u>	<u>8</u>

1. Installation de VirtualBox

1.1. Introduction

VirtualBox is a general-purpose full virtualizer for x86 hardware, targeted at server, desktop and embedded use. For a thorough introduction to virtualization and [VirtualBox](#) ^[1], please refer to the [online version of the VirtualBox User Manual's first chapter](#) ^[2].

[1] <http://www.virtualbox.org/wiki/VirtualBox> - [2] <http://www.virtualbox.org/manual/ch01.html>

1.2. Version à utiliser

Page de téléchargement: <http://www.virtualbox.org/wiki/Downloads>

Dans la cadre de Mascaret et pour bénéficier de la meilleur puissance de calcul disponible, privilégier une version 64 bits de VirtualBox à une version 32 bits, si votre machine et système d'exploitation le permettent:

ATTENTION, la machine virtuelle « Mascaret assim » possède un système d'exploitation 64 bits. Pour la faire tourner sur un système hôte est en 32 bits, il est indispensable que le(s) processeur(s) de la machine hôte supporte(nt) la virtualisation matérielle et que cette option soit activée dans le bios.

Windows hosts:

- Windows XP, all service packs (32-bit)
- Windows Server 2003 (32-bit)
- Windows Vista (32-bit and 64-bit)
- Windows Server 2008 (32-bit and 64-bit)
- Windows 7 (32-bit and 64-bit)

Linux hosts (32-bit and 64-bit)

1.3. Installation

suivre les instruction de la documentation officielle:

<http://www.virtualbox.org/manual/ch02.html>

2. Mise en place de la machine virtuelle « Mascaret assim »

2.1. Récupération de l'archive

Fournie sur un DVD disponible sur demande auprès du Schapi, l'archive se compose de 3 fichiers, dont le disque dur virtuel qui pèse environ 2Go (compressé):

Nom	Taille	Date
Mascaret assim.vmdk	1,8 Gio	22/03/2011 18:26
mascaret-assim.mf	145 o	22/03/2011 18:29
mascaret-assim.ovf	13,2 Kio	22/03/2011 18:26

2.2. Importation de la machine dans le virtualiseur

Lancer l'interface graphique de VirtualBox.

Sélectionner « Fichier », « Importer une application virtuelle ».

Sélectionner le fichier 'ovf' puis cliquer sur « Suivant ».

Vérifier ou modifier les paramètres de la machine puis cliquer sur « Terminer ».

Nb: en cas d'échec de cette procédure, il est envisageable de « créer » une nouvelle machine virtuelle en utilisant comme « disque virtuel » le fichier vdi ou vmdk complet (5 à 6 Go).

2.3. Paramétrage final de la machine Mascaret

Au même titre que l'on peut ajouter ou modifier des composants physiques d'un ordinateur, (presque) tous les réglages des machines virtuelles sont réversibles et modifiables. Ils ne sont possibles que sur une machine arrêtée.

Un réglage très important pour le calcul parallèle est celui du nombre de processeurs à affecter à chaque machine virtuelle, en fonction du matériel disponible et de l'utilisation qui en est faite. Plus ce nombre est important, plus rapides seront les calculs avec assimilation (calcul parallèle) mais bien sûr au détriment des ressources disponibles^[1]:

Nombre de processeurs (configuration >> système >> processeur)

Un autre paramètre à adapter, plus secondaire, est celui de la mémoire affectée à la machine. La valeur par défaut de 512 Mo peut être étendue à 1Go pour de meilleurs performances (2 Go ? benchmark TODO):

mémoire (configuration >> système >> carte mère >> mémoire vive)

Le partage de répertoire permet de faciliter les échanges de fichiers entre l'hôte et l'invité en rendant accessible directement depuis la machine virtuelle des répertoire de l'hôte (la machine principale).

Les autres réglages, assez nombreux, n'ont pas ou peu d'importance dans les configurations ou la machine invitée est installée directement sur le poste de l'utilisateur. Dans les autres cas il faudra soit régler les paramètres réseau soit utiliser le partage de bureau pour accéder à la machine. Se référer à la documentation de VirtualBox si besoin.

[1] si votre configuration matérielle et votre bios le permettent, il est conseillé d'activer au préalable l'hyperthreading (ou équivalent)

3. Utilisation de la machine virtuelle « Mascaret assim »

Le plus simple, si la machine virtuelle est installée sur le poste de l'utilisateur, est de l'utiliser directement dans une fenêtre VirtualBox. Une utilisation distante, via un partage de bureau ou bien une connexion ssh est également possible mais elle nécessite d'adapter au préalable la configuration réseau.

Pour lancer une machine, la sélectionner dans l'interface principale de VirtualBox puis cliquer sur « démarrer ».

On peut arrêter une machine virtuelle mais il est plus pratique de la mettre en pause, les temps d'arrêt et de reprise sont raccourcis.

Vous trouverez en annexe les caractéristiques de la machine virtuelle Mascaret.

Et pensez à sauvegarder vos données de temps à autre !

Nb: si le support de la souris ou de l'écran n'est pas parfait, il peut être nécessaire de ré-installer les additions clientes. Se référer à la documentation officielle.

4. Utilisation de la maquette Cerfacs « Mascaret assim 2 »

4.1. Présentation générale

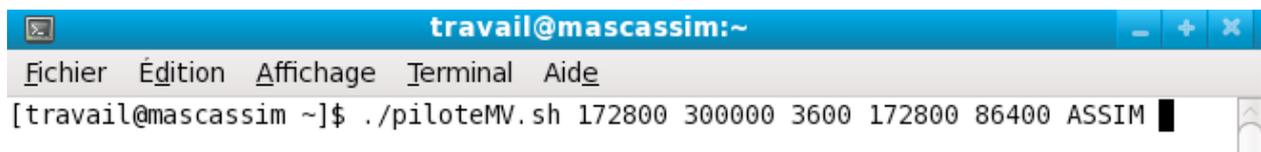
A partir du répertoire « modèle » /home/(utilisateur)/cerfacs, se créer un ou des répertoires de travail via la commande 'cp -r repertoire_source repertoire_destination' puis s'y déplacer.

Contenu d'un répertoire de travail:

```
drwxr-xr-x 2 marne users 4096 mars 22 10:02 DonneesDyn => fichiers .loi dynamiques
(hydrogrammes des stations d'entrée du modèle).
drwxr-xr-x 2 marne users 20480 mars 22 10:02 DonneesStat => configuration hydraulique et
fichiers .loi statiques (courbe de tarage, apport constants).
drwxr-xr-x 2 marne users 4096 mars 22 09:56 Observations => données observées afin que la
maquette puisse comparer les résultats obtenus à ceux réellement observés.
-rwxr-xr-x 1 marne users 12811 mars 22 09:56 piloteMV.sh
drwxr-xr-x 4 marne users 4096 mars 22 09:56 Programmes
drwxr-xr-x 10 marne users 4096 mars 22 10:02 Resultats => résultats
```

Les simulations sont gérées par le script piloteMV.sh, dont quelques exemples d'utilisation sont indiqués dans le fichier texte logbook.txt.

Pour lancer une simulation, ouvrir un terminal LINUX et taper la commande suivante:



./piloteMV.sh = lancer le pilote

Tdeb(s) = temps à partir duquel la maquette va produire des résultats

Tfin(s) = temps à partir duquel la maquette va stopper les calculs

pas(s) = pas de temps, fréquence des calculs

tequilibrage(s) = période à partir du premier temps (0 normalement) sur laquelle le modèle s'équilibre. Il doit être inférieur ou égal à Tdeb.?

tprevision(s) = échéance maximum des prévisions ?

DRYRUN|ASSIM|CALAGE|CALAGEASSIM [FAST] = il faut choisir parmi l'un de ses quatre termes et y ajouter facultativement FAST (séparé par un espace). Le DRYRUN réalise des prévisions brutes pour chaque temps de base (Tdeb) à différentes échéances (échéance max = tprevision ?). ASSIM réalise des prévisions avec l'assimilation de la ligne d'eau, la ligne d'eau est alors recalée en fonction des erreurs observées. CALAGE réalise des prévisions en recalant en temps réel des lois d'entrées du modèle (forçage amont). Enfin, CALAGEASSIM mixe les deux techniques d'assimilation de données. On ajoute FAST derrière l'un de ses termes lorsque l'on veut seulement effectuer le pst-traitement des données (cela implique donc que l'on a déjà des résultats).

Les résultats de la simulation sont disponibles dans le répertoire du même nom.

4.2. Initialisation des répertoires

Il est nécessaire d'adapter la maquette au modèle Mascaret dont vous voulez tester les procédures d'assimilation de données. Seul le dossier Programmes du répertoire de travail ne doit pas être modifié.

Dans un premier créer une copie du répertoire Cerfacs afin d'avoir toujours des fichiers propres au cas où.

Sous Windows, ouvrir Mascaret, faite un calcul avec votre modèle .masc que vous souhaitez installer dans la maquette.

Allez sous C:\Program Files\Fudaa-Mascaret\serveurs\mascaret\mascaret_7_0 et copiez les fichiers dynamiques (.loi hors courbe de tarage ou apport constants) dans le répertoire DonneesDyn, au préalable supprimez le contenu de ce dossier. Copiez les fichiers .cas .lig .geo et .loi statiques (courbe de targe, apports constants) dans DonneesStat, au préalable supprimez le contenu de ce dossier sauf les fichiers ParametresMascaret.cas et ParametreMascaretAssim.cas. Déposez le fichier des données observées au format .csv dans le dossier Observations afin de pouvoir comparer les données du modèle à l'observation.

Ouvrez le fichier .opt et notez les points kilométriques (PK) ainsi que les numéros des sections de calcul de vos points de sortie du modèle.

- **dossier Observations:** adapter le script obs_split_MV.sh au fichier .csv en modifiant les numéros des colonnes (\$4 = colonne 4 par exemple) que vous souhaitez exporter ainsi que les numéros de section de calcul. Renseignez autant de ligne (en copiant/collant) que de stations d'observations. Lancez le script (en double cliquant dessus, puis 'lancer dans un terminal'). Vous obtenez un fichier .dat par stations d'observation. Vérifiez les résultats et supprimez les anciens .dat.

- **dossier DonneesStat:** ouvrir les fichiers mascaret0.cas et ParametresMascaret.cas, remplacez dans ce premier les chemins des dossiers suivant (en les copiant/collant depuis le fichier d'origine ParametresMascaret.cas par exemple):

- fichier des mots clés: '.././DonneesStat/ParametresMasacret.cas
- fichier sauvegarde du modèle: 'Mascaret1.tmp'
- fichier de géométrie: '.././DonneesStat/Geometrie.geo'
- fichier ligne d'eau: '.././DonneesStat/LigneEeuIntiale_in.lig'
- fichier résultats: '.././Resultats/ResultatsOphyca.opt'
- fichier listing: '.././Resultats/ResultatsListing.lis'
- fichier de reprise en écriture: 'masacret1.rep'

Renommer le fichier .geo que vous avez importé en Geometrie.geo et le fichier .lig en LigneEaulInitiale.lig dans le dossier DonneesStat.

Procéder de la même façon pour les fichiers .loi: une loi statique au nom X (courbe de tarage, apport constant) aura logiquement pour chemin '.././DonneesStat/X.loi'. Une loi dynamique au nom Y aura pour chemin '.././DonneesDyn/Ycut.loi'. A partir des fichier .loi d'origine, la maquette va construire automatiquement pour chaque pas de temps et chaque loi dynamique un hydrogramme coupé et prolongé avec un débit constant.

Supprimer l'ancien ParametresMascaret.cas et renommer votre fichier mascaret0.cas .

Attention: le fichier ParametreMascaret.cas est limité à 73 colonnes sur une même ligne, le fichier .cas constitué avec l'interface Fudaa pour comporter plus de colonne donc n'hésiter pas à passer à la ligne pour éviter des plantages.

Seul ce fichier est nécessaire pour effectuer un DRYRUN. L'ASSIM, le CALAGE et L'ASSIMCALAGE utilise le fichier ParametreMascaretAssim.cas pour effectuer les calculs. Pour

adapter ce fichier à votre nouveau modèle, se référer au rapport 'Principe de la maquette d'assimilation de données pour le modèle Mascaret'

- **pilote:** ouvrir PiloteMV.sh et modifier les référence ds stations de sortie:

```
#Extrait temps, hauteur($6), debit lit mineur+majeur ($12) dans fichier Opthyca ($7=debir lit mineur)
awk -F ':' '{ /^ 870*/{print $1,$6-251.1,$12}' ResultatsOpthyca.opt > Resultats_stat870.dat
awk -F ':' '{ /^ 969*/{print $1,$6-245.94,$12}' ResultatsOpthyca.opt > Resultats_stat969.dat
awk -F ':' '{ /^ 1865*/{print $1,$6-193.16,$12}' ResultatsOpthyca.opt > Resultats_stat1865.dat
awk -F ':' '{ /^ 2037*/{print $1,$6-182.68,$12}' ResultatsOpthyca.opt > Resultats_stat2037.dat
awk -F ':' '{ /^ 2549*/{print $1,$6-149.95,$12}' ResultatsOpthyca.opt > Resultats_stat2549.dat
awk -F ':' '{ /^ 3667*/{print $1,$6-205.74,$12}' ResultatsOpthyca.opt > Resultats_stat3667.dat
```

Le premier numéro (878 par exemple) est celui de la section de calcul, le quatrième (251,1) est la cote du zéro de l'échelle de la station à retrancher à la cote de l'eau de la sortie Mascaret. Si la section de calcul comporte 3 chiffres, conserver un espace avant le numéro de la section. Les variances pour le calage et l'assimilation sont à rentrer en dur dans le pilote ainsi que plusieurs données temporelles (à préciser c'est encore flou).

Enregistrer, vous pouvez maintenant effectuer des calculs, les fichiers résultats sont disponibles sous Resultats/DRYRUN-ASSIM-CALAGE-CALAGEASSIM (selon votre choix), ce sont des fichiers .opt. Il est possible de visualiser ces résultats avec les graphiques Gnuplot mais cela implique des modifications plus lourdes du PiloteMV.sh et du fichier calcul_criteresMV.f90 à faire si nécessaire.

ANNEXE 1 - Caractéristiques de la machine virtuelle « Mascaret assimilation »

Système:

Système::Linux - Distribution::Fedora 11 (x86_64) - Bureau::Gnome (2.26) - Shell::bash

Compte utilisateur: travail - Mot de passe : travail (Mot de passe root: mascaret)

Maquette assimilation: Mascaret 7 modifié - palm 4.0.0 - Mpich2 1.2.1p1

Quelques outils graphiques du gestionnaire de bureau Gnome:

navigateur de fichier::Nautilus - éditeur de texte::gedit / terminal::gnome-terminal

Quelques commandes de base pour Linux:

(source <http://www.epons.org/commandes-base-linux.php>)

pwd - Affiche le chemin absolu du répertoire courant (*Print Working Directory*).

cd [répertoire] - Change de répertoire (*Change Directory*). Va dans répertoire ou dans le répertoire de l'utilisateur s'il n'y a pas d'argument. Si "-" est indiqué en argument, déplace dans le répertoire précédent.

ls [fichier ...] - Liste le contenu des répertoires ou le nom des fichiers passés en arguments (liste le répertoire courant si pas d'argument).

-l : affichage détaillé (long)

-a : affichage aussi des fichiers cachés dont le nom commence par un point (all)

-i : affichage des numéros d'inodes (inode)

-d : affichage du nom du répertoire et non de son contenu (directory)

-t : trie l'affichage suivant la date de modification des fichiers (time)

-r : inverse le tri d'affichage (reverse)

-R : affichage du contenu de tous les sous répertoires (recursive)

cp <source ...> <destination> - Copie (*CoPy*) les fichiers source vers destination.

-i : demande confirmation avant écrasement (interactive)

-f : écrase sans demander confirmation (force)

-R ou -r : copie aussi les répertoires (recursive)

mv <source ...> <destination> - Renomme/déplace (*MoVe*) les fichiers source vers destination.

-i : demande confirmation avant écrasement (interactive)

-f : écrase sans demander confirmation (force)

rm <fichier ...> - Supprime (*ReMove*) les fichiers passés en arguments.

-i : demande confirmation avant suppression (interactive)

-f : supprime sans demander confirmation (force)

-R : supprime aussi les répertoires (recursive)

cat <fichier ...> - Affiche le contenu des fichiers texte passés en arguments.

more <fichier ...> - Affiche page par page le contenu des fichiers texte passés en arguments.

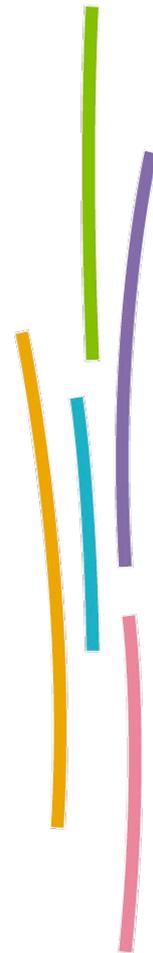
man [section] <argument> - Affiche la page de manuel électronique dont le nom est argument (se trouvant dans la section du manuel éventuellement spécifié).

-k : retourne le nom des pages de manuel contenant argument

exit [code_retour] - Quitte le shell en cours avec le code retour passé en argument ; par défaut exit retourne le code retour 0.

Annexe 7 : Présentation SCHAPI 29 Avril 2011

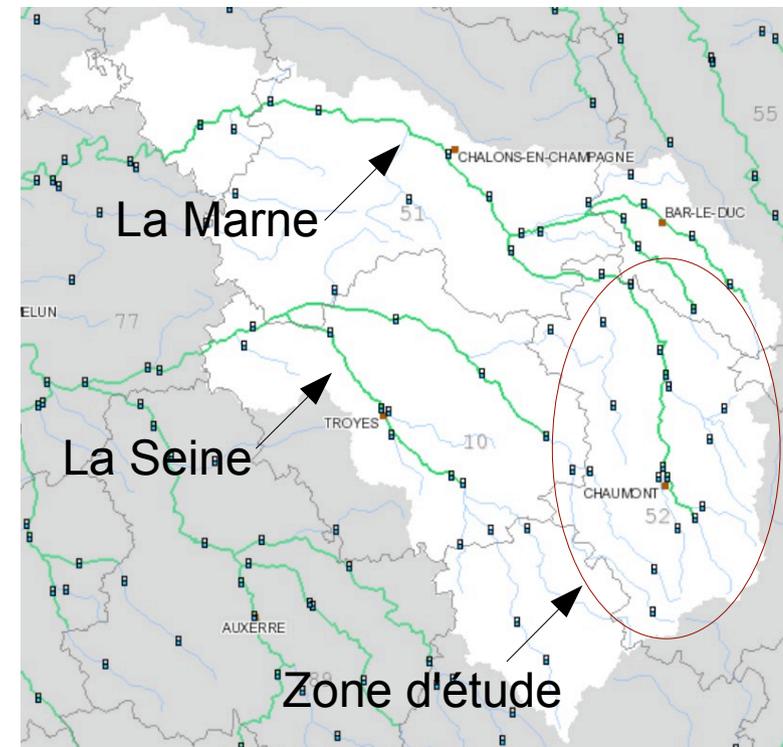
Mise en place de la maquette d'assimilation de données pour le modèle hydraulique Mascaret 'Marne amont global'



Le modèle Mascaret 'Marne amont global'

2

- La Marne est la plus longue rivière de France, 520 Km de long, elle est un affluent majeur de la seine.
- La prévision des crues sur la partie amont est assuré par le Service de Prévision des Crue Seine Amont Marne Amont (SPC SAMA) basé à Châlons-en-Champagne.



Le modèle Mascaret 'Marne amont global'

3

Le modèle résulte d'une fusion de deux modèles préexistants, le modèle marne valage, opérationnel depuis 2007 et le modèle marne amont opérationnel au SPC depuis 2008, afin d'accroître le temps de prévision.



Cette fusion a pu être possible par l'ajout de nouveaux profils, par une meilleure maîtrise de la modélisation hydraulique et par... du temps.

Conséquences:

- les apports intermédiaires doivent être modélisés
- la station de Saucourt est devenu limitante, il a fallu agrandir le modèle en amont de cette station en intégrant le bassin versant karstique du Rognon
- Le temps de prévision à la station la plus à l'aval (Chamouilley) est passé de 15 à 21h.

Le modèle Mascaret 'Marne amont global'

4



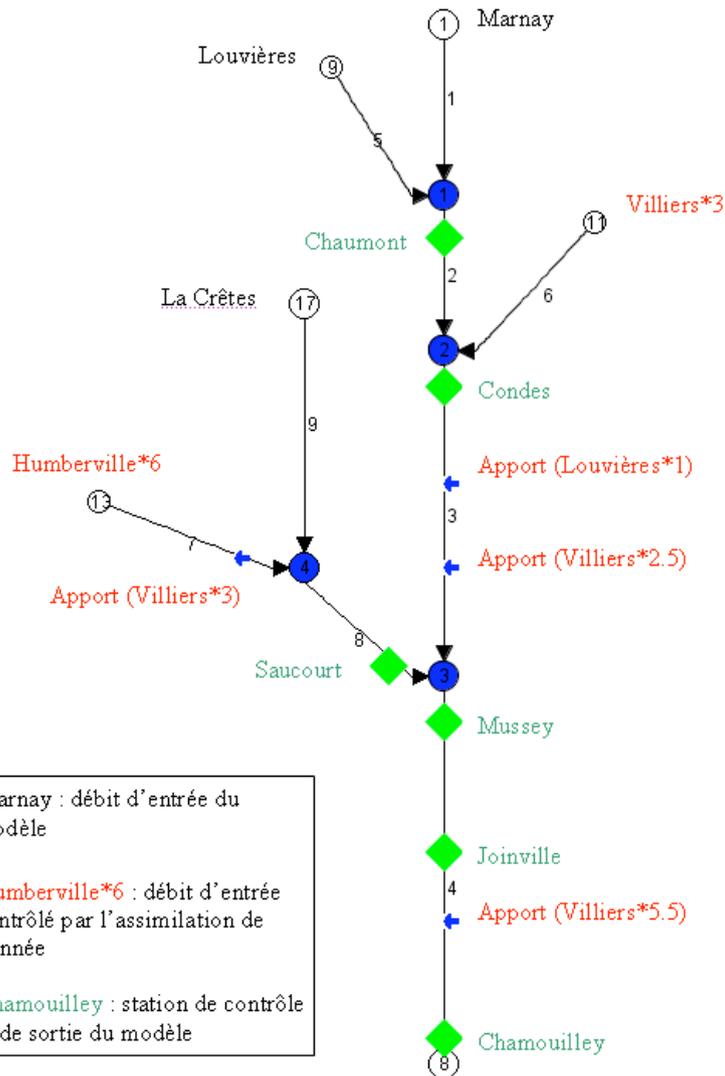
Le bassin versant du Rognon est karstique, le service d'hydrométrie gère 4 stations, Saucourt, à l'aval et trois autres stations très en amont du bassin. Une est à sec lorsqu'il ne pleut pas, donc inutilisable pour la modélisation hydraulique. La somme des deux autres ne représente pas la moitié du débit à l'aval à saucourt. De plus, on ne dispose que de peu de profil. Le Rognon peut apporter en crue autant de débit que la marne, sa modélisation est essentielle.

Resurgence de la Manoise, affluent du Rognon, dite du 'Cul du cerf'

Le modèle Mascaret 'Marne amont global'

5

Modèle hydraulique Mascaret Marne Amont Global



Marnay : débit d'entrée du modèle

Humberville*6 : débit d'entrée contrôlé par l'assimilation de donnée

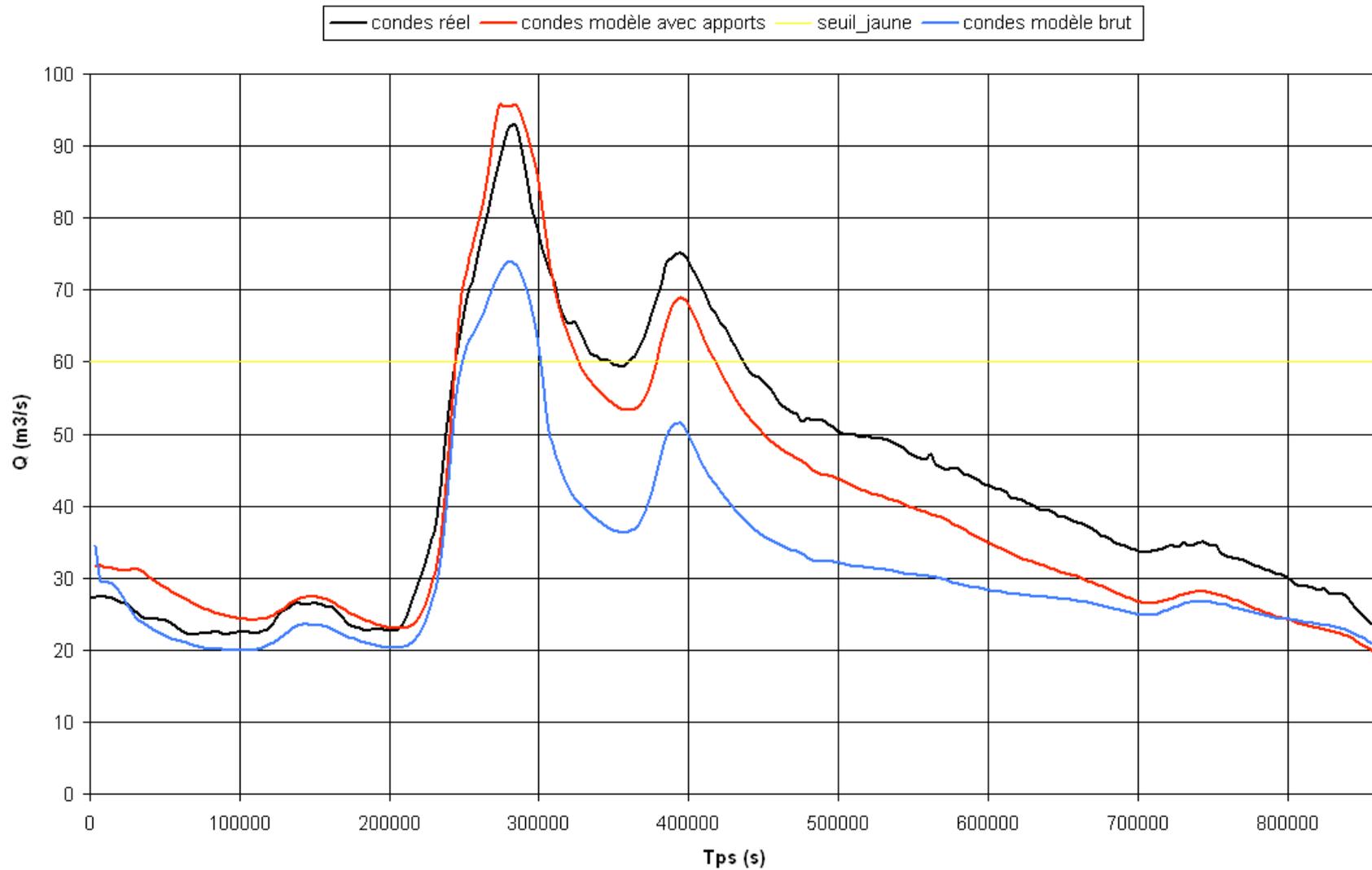
Chamouilley : station de contrôle et de sortie du modèle

- 170 km de cours d'eau modélisé, 2250 Km² de bassin versant modélisé
- 112 profils, 1 profil tous les 1,5 Km en moyenne
- 5 stations d'entrée, 4 apports intermédiaires, 6 stations de sortie.
- prévisions de 5 à 21h selon les stations.

Les apports intermédiaires ont été modélisés par transposition d'hydrogramme de stations d'entrée en différents points du modèle, multiplié par un coefficient moyen A déterminé après analyse hydrologique de dix crues.

Le modèle Mascaret 'Marne amont global'

6

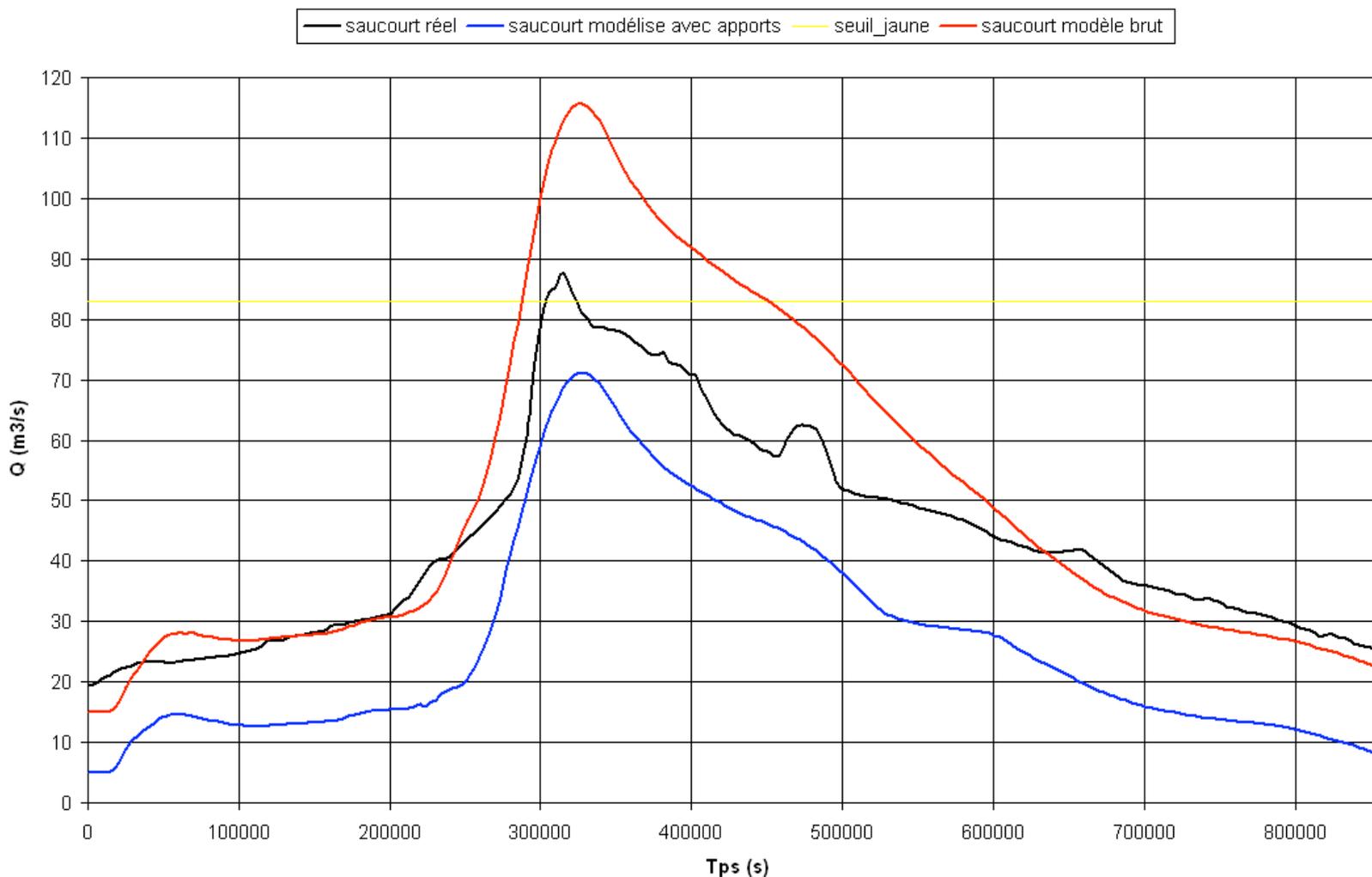


La multiplication, la transposition des hydrogrammes sont bénéfiques la plupart du temps mais...



Le modèle Mascaret 'Marne amont global'

7



... parfois ça explose! La réaction brutale des karst sur le bassin du Rognon multiplié par un coefficient A moyen peut provoquer des surestimation des débits à l'aval.

L'assimilation de données

8

Ces coefficients A peuvent être contrôlés par assimilation de données. Le CERFACS a développé une plate-forme d'assimilation de données compatible avec les modèles hydrauliques Mascaret.

En temps réel, A va être modifié en fonction des écarts observés aux stations de sortie entre les hauteurs d'eau observées et les hauteurs d'eau modélisées.

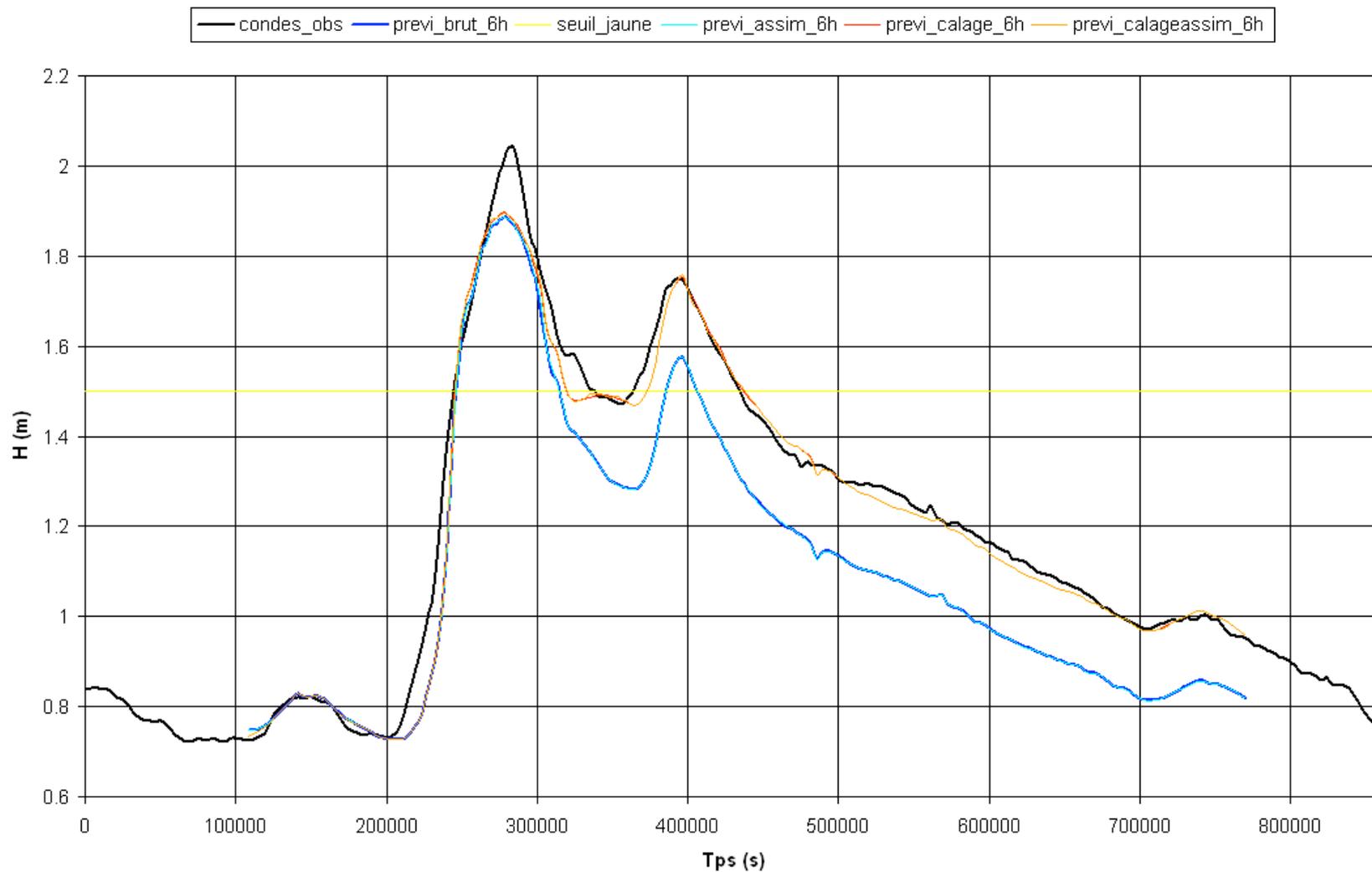
L'assimilation de données peut aussi jouer sur le décalage temporel des hydrogrammes d'entrée et sur la ligne d'eau initiale (cf CERFACS pour les équations, les matrices de covariance et les différences finies).

Les premiers résultats: Seuls les débits des stations d'entrées de Villiers et de Humberville sont contrôlés en assimilant les écarts entre hauteurs observés et modélisées sur les 24 dernières heures respectivement aux stations de Condes et de Saucourt.



L'assimilation de données

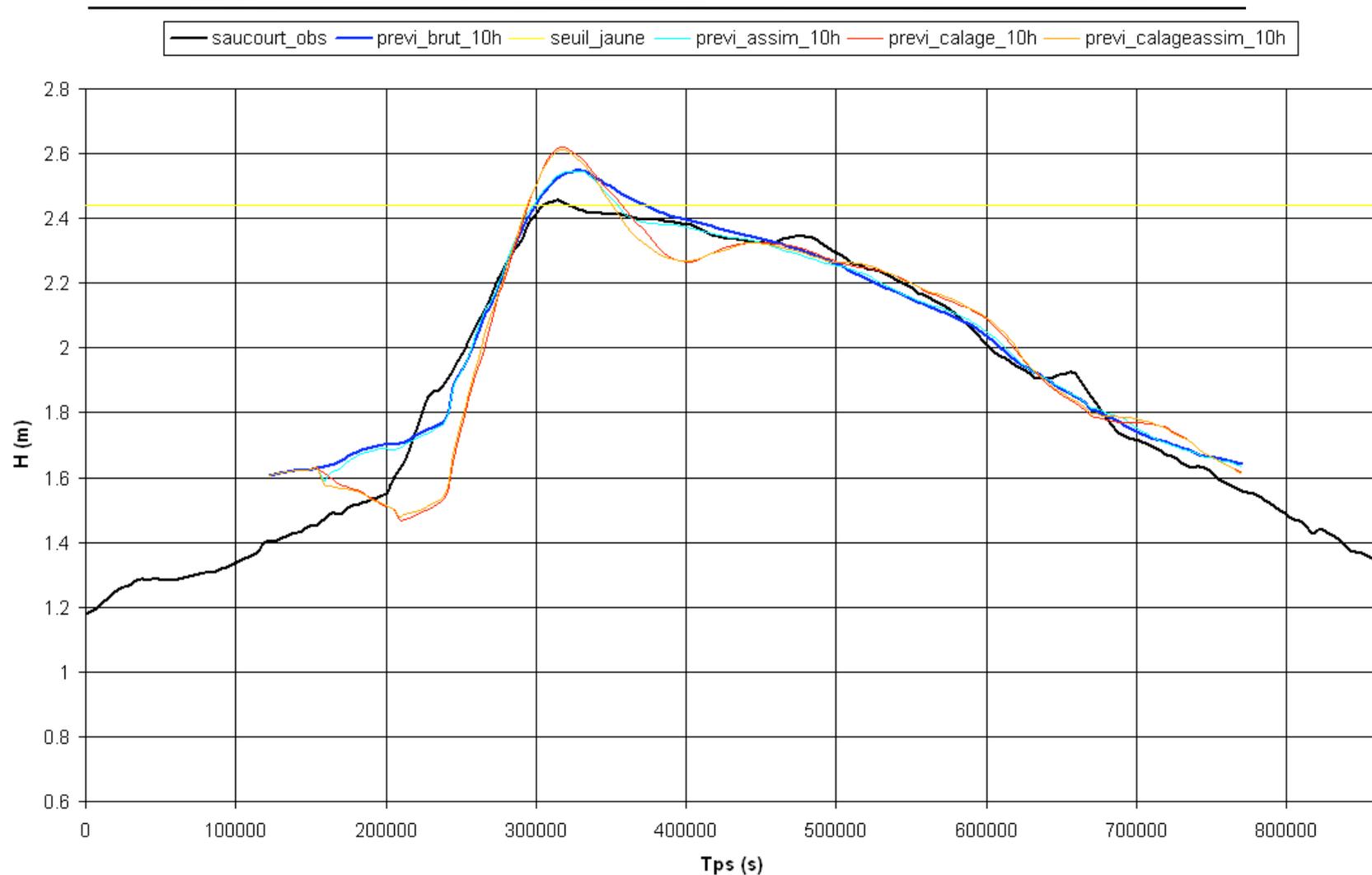
9



Prévision à 6h à Condes: les résultats sont encourageants. Le recalage de la ligne d'eau (assim) n'a plus d'effet à cette échéance à Condes

L'assimilation de données

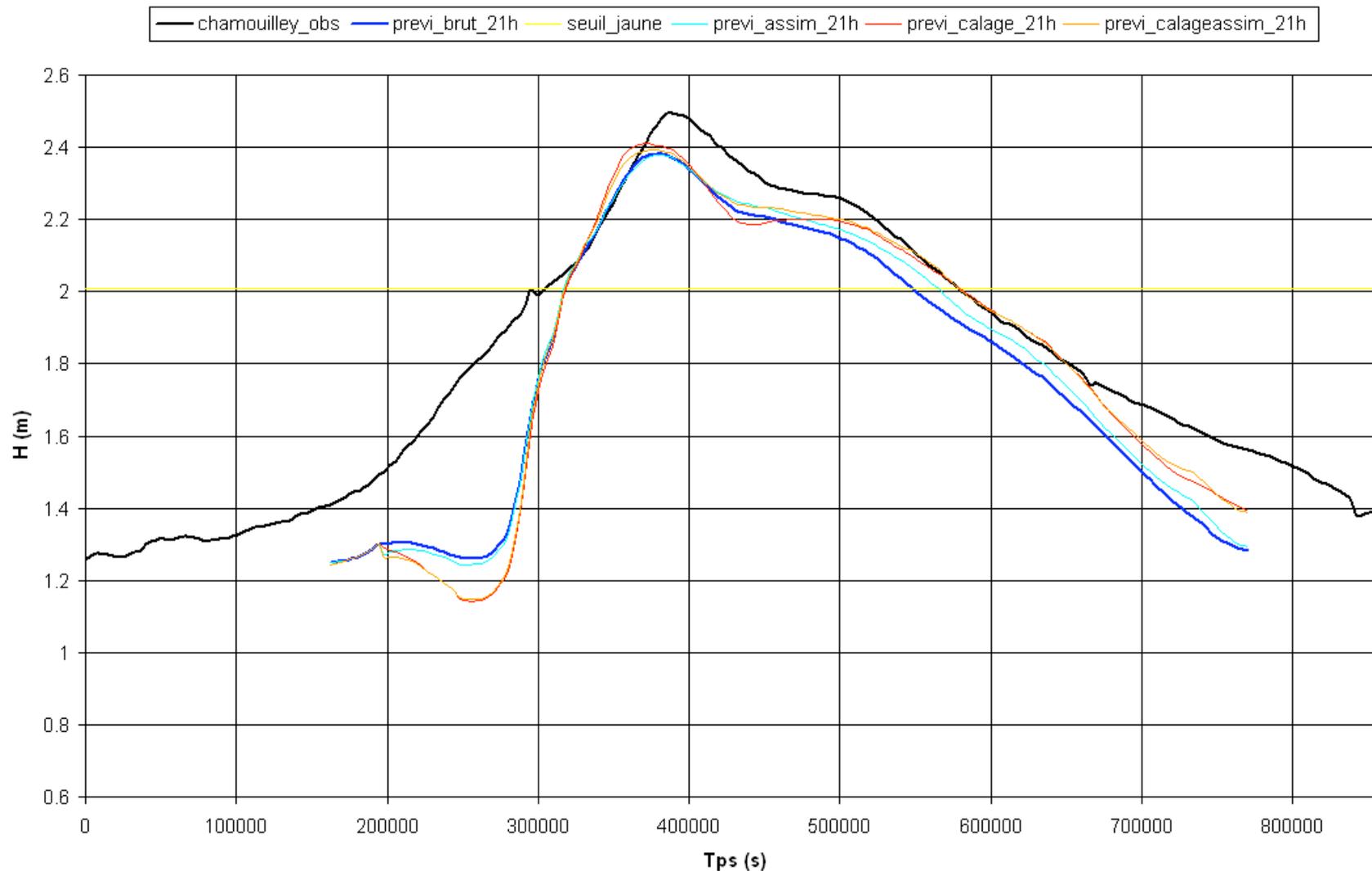
10



Prévision à 10h à Saucourt: les hauteurs d'eau assimilées dégradent la prévision par rapport à l'ébauche (previ brut), qui n'est pas mauvaise ceci-dit

L'assimilation de données

11



Prévision à 21h à Chamouilley: l'assimilation améliore les prévisions, notamment celle de la ligne d'eau initiale (assim)

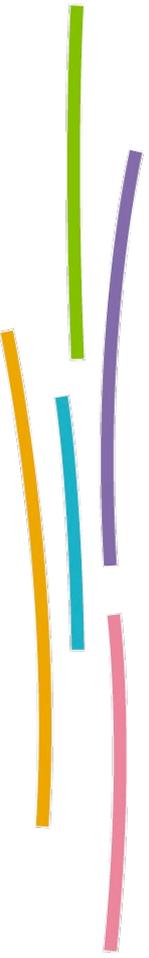
L'assimilation de données

12

Les résultats obtenus après assimilation de données sont encourageants, ils peuvent encore être améliorés. La plate-forme développée par le CERFACS est un outil puissant.

Mettre en place un couple assimilation de données-modèle hydraulique Mascaret pour effectuer des prévisions de hauteurs d'eau ou de débits réclame une bonne connaissance du modèle, du temps ...

et surtout des erreurs.



Annexe 8 : Présentation SPC-SACN 29 Avril 2011

Maquette MASCARET avec Assimilation de données : application au modèle de la Seine Aval à Rouen

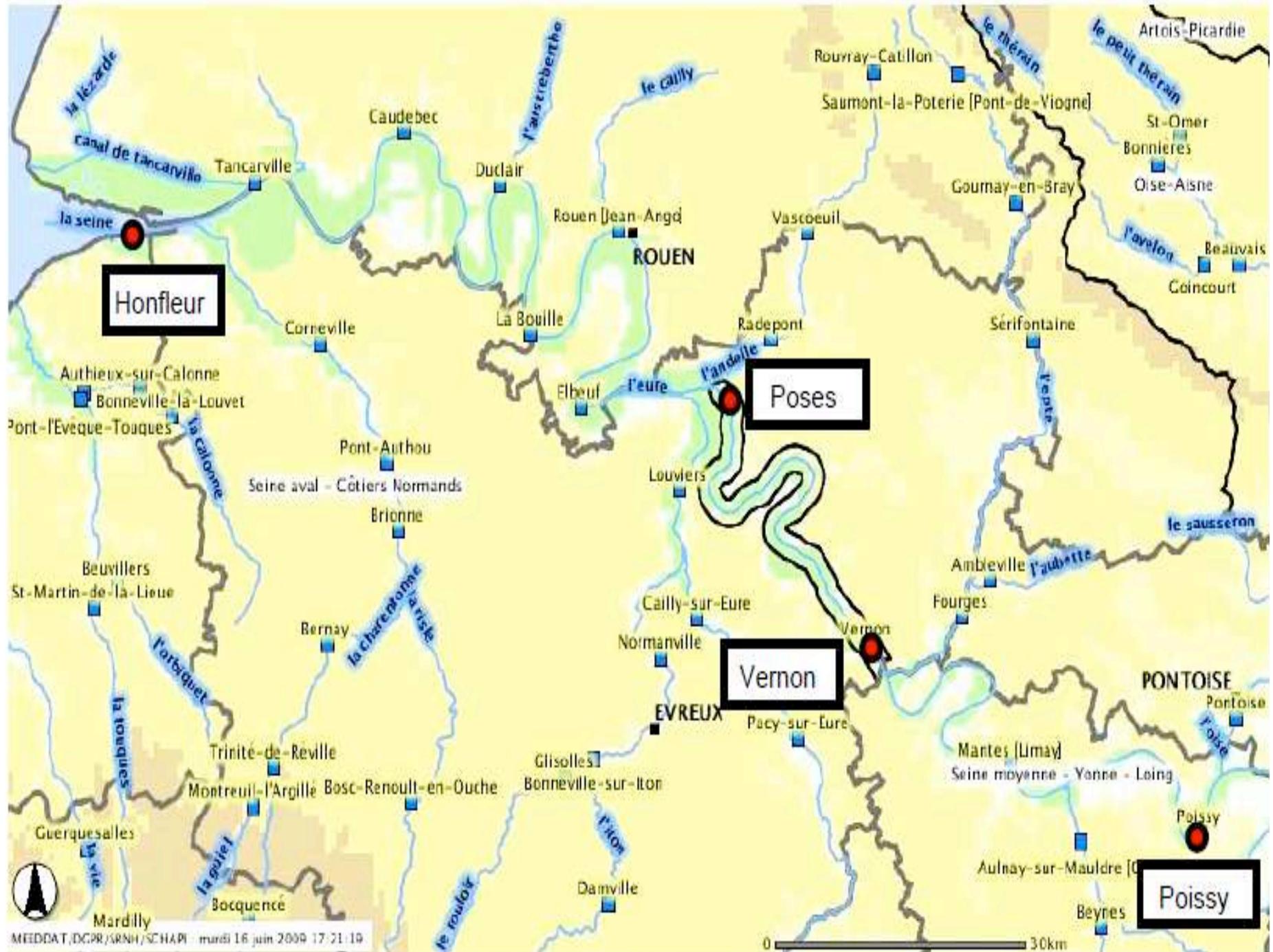
nt
ement durable
ts et mer

Marie-Pierre CRIBELLIER, SPC Seine Aval
et fleuves Côtiers Normands

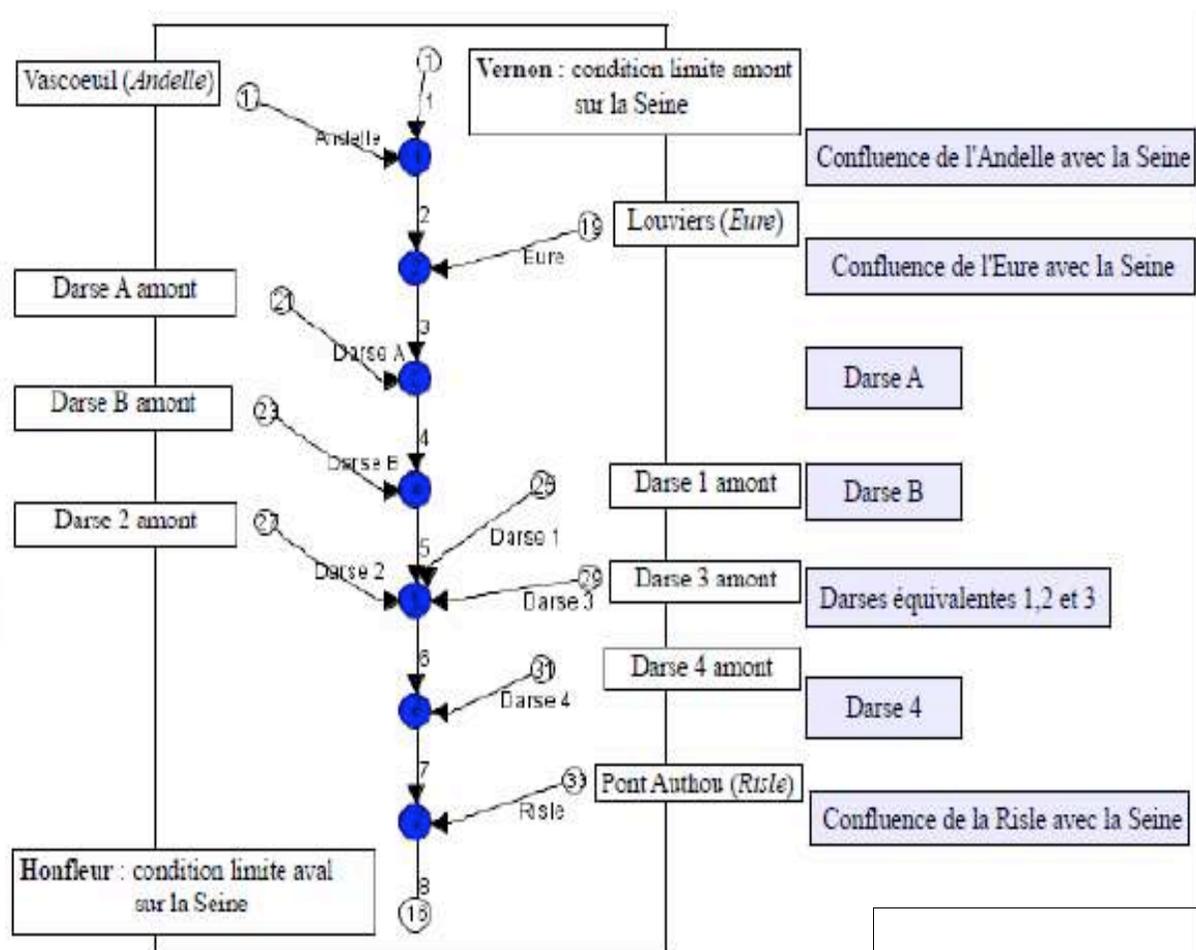
Vanessya LABORIE, CETMEF, Laboratoire
d'Hydraulique SAINT-VENANT



La zone d'étude



Le réseau hydraulique



STRUCTURE DU RESEAU:

- 17 biefs
- 7 noeuds
- 10 conditions limites amont (Seine, Andelle, Eure, Risle, 6 darses)
- 1 condition limite aval (Seine)

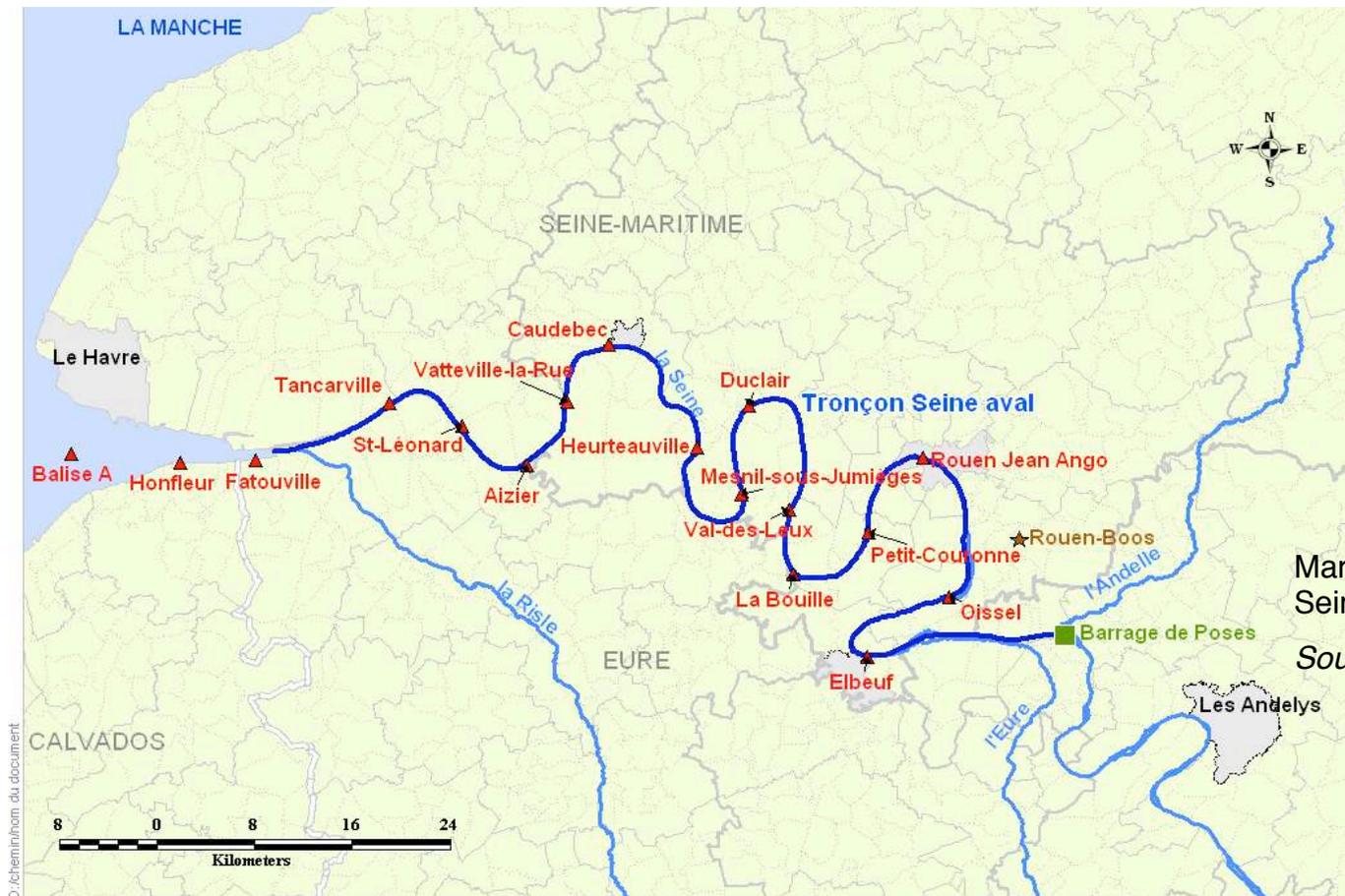
Fonctionnement estuarien

- Pas de temps de calcul : 5 mn
- Conditions fluviales en amont
 - Apports hydrologiques amont du fleuve
 - Apports hydrologiques des affluents:
Principalement l'Andelle, l'Eure et la Risle
 - Apports hydrologiques souterrains via la nappe de la craie
- Conditions maritimes en aval
 - Cote de surface libre (marée + surcote) à Honfleur



Données d'observation

- À l'amont : station débitmétrique de Vernon en place depuis août 2006 : débits horaires.
- Marégraphes de Rouen et Honfleur (limite aval du modèle) : pas de temps de 5 mn



Marégraphes sur le tronçon Seine aval.

Source: SPC SACN (2009).

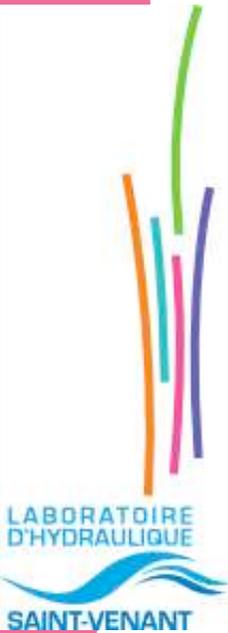
D:\chemin\nom du document

Sources : DDE 76 | IGN BdCarto©2006 |
© Direction Départementale de l'Équipement de Seine-Maritime - SSER | conception : SSER - BPC - C. Lanon - Juin 2008



Tests avec assimilation de données

- Contrôle des conditions fluviales en amont
 - Débits amont: station de Vernon
 - Apports latéraux des affluents : stations de Vascoeuil, Louviers et Pont-Authou
 - Evaluation des apports hydrologiques souterrains via la nappe de la craie
introduction d'un débit d'apport en amont de Rouen dans le modèle
- Contrôle de la condition limite aval (en mode prévision particulièrement) :
Surcote marine prévue par Meteo France à 48h



Merci de votre attention...

