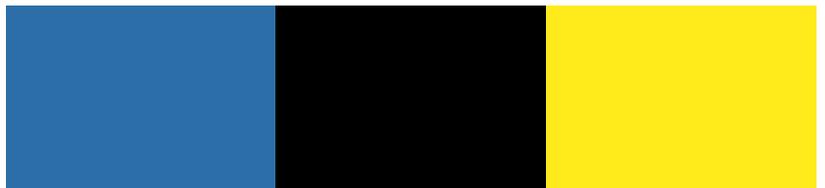




31 décembre 2013

Georges-Marie Saulnier, William
Castains, Anne Johannet



Tâche 05 Assimilation de données

Sommaire

1. INTRODUCTION.....	2
2. QUELLE APPROCHE DE L'ASSIMILATION DE DONNEES ?	4
2.1 L'APPROCHE SEQUENTIELLE	5
2.2 L'APPROCHE VARIATIONNELLE	6
3. EXPERIENCES D'ASSIMILATION DE DONNEES.	7
3.1 CONTEXTE OPERATIONNEL.....	7
3.2 CHOIX DES VARIABLES DE L'ASSIMILATION	8
3.3 SENSIBILITE AUX VARIABLES DE CONTROLE.....	10
3.4 SENSIBILITE A LA DUREE DE LA FENETRE D'ASSIMILATION	13
3.5 INFLUENCE DE L'HORIZON DE PREVISION	14
4. CONCLUSIONS - PERSPECTIVES.....	16

1. INTRODUCTION

Comme on le verra plus en détail par la suite, l'assimilation de données est un outil mathématique permettant de corriger les prévisions d'un modèle numérique en prenant en compte certaines observations de terrain, acquises dans un laps de temps plus ou moins court précédant l'instant auquel est élaborée la prévision numérique.

Cet outil est très largement utilisé dans les domaines de la météorologie et de l'océanographie mais encore rarement en hydrologie opérationnelle, du moins dans notre pays (l'assimilation de données est déjà utilisée en opérationnelle par exemple aux Etats-Unis).

Si elle ne compense pas la nécessité d'améliorer le réalisme et la performance des modèles représentant la relation pluie – débit (modèles physiques ou modèles statistiques), l'assimilation de données offre un potentiel, à explorer, pour réduire l'écart entre les prévisions effectuées par les modèles numériques et la réalité du terrain. Cette opportunité d'améliorer le réalisme de la prévision numérique prend évidemment tout son sens dans le cas de la prévision des crues pour une meilleure protection des populations.

Si ce potentiel demeure intact quelque soit la qualité intrinsèque des modèles numériques de prévision, c'est que, comme l'illustre la figure 1 suivante, **les modèles, nécessairement imparfaits, de prévision sont toujours contraints et forcés par des données incertaines.**

La problématique de la prévision hydrologique numérique consiste :

- à partir de de variables de forçages (rectangle bleu), par exemple une carte de pluies mesurée par radar ou un une évaluation de la pluie moyenne sur le bassin à un pas de temps donné, etc.
- à prédire la valeur de la réponse hydrologique (rectangle bleu), par exemple un débit d'une section donnée, le dépassement ou non d'une cote (hauteur d'eau) d'alerte à une section donnée, etc.
- au moyen d'un modèle hydrologique (rectangle gris foncé).

L'utilisation du modèle hydrologique introduit de fait deux types d'incertitudes :

- l'imperfection, incompressible, du modèle hydrologique (par rapport à la réaliste complexe des processus naturels) et des incertitudes : c'est l'**incertitude liée à la structure du modèle** (rectangle gris clair)
- l'**incertitude sur les conditions initiales et les paramètres du modèle**. En effet, même dans le cas d'un modèle physique les paramètres et les conditions initiales peuvent être plus ou moins difficilement fiables à une mesure de terrain. Cela peut être le cas d'un paramètre prescrit dans le modèle à une résolution spatiale différente de celle de l'appareil de mesure. Ce peut être aussi le cas si le paramètre prescrit dans le modèle est davantage un concept mathématique (par ex. l'épaisseur moyenne efficace d'un sol sur quelques

centaines de m2) qu'une réalité observable et quantifiable par un appareil de mesure existant.

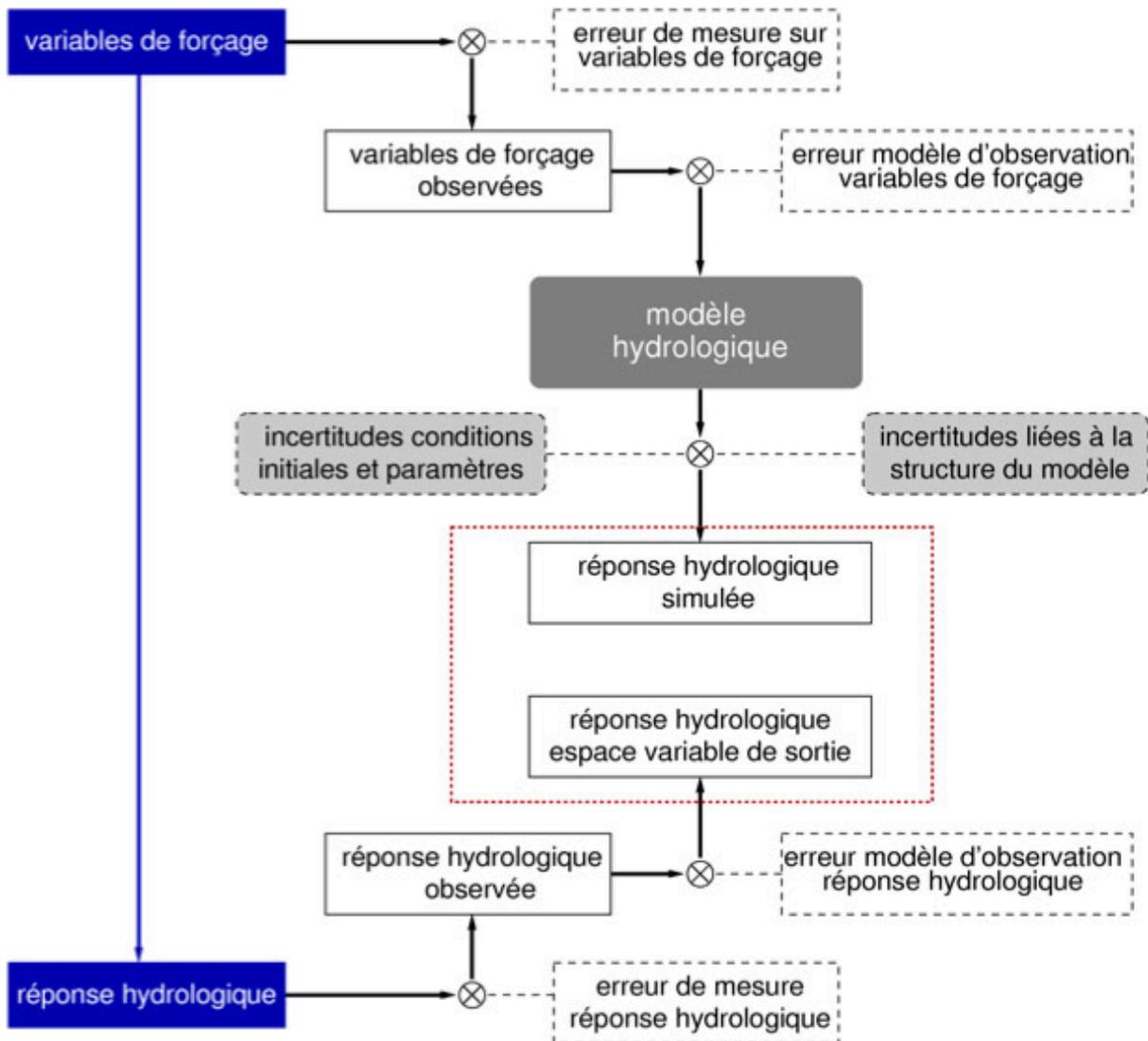


Figure 1 : Chaîne d'incertitudes et d'erreurs dans la prévision hydrologique

De plus, plusieurs sources d'erreurs sont également introduites dans cette chaîne de prévision hydrologique :

- Les variables de forçages et les variables pronostiques du modèle hydrologique ne sont pas nécessairement celles qui sont effectivement mesurées ni celles qui sont souhaitées pour la prévision. Par exemple le radar météorologique mesure une réflectivité électromagnétique alors que le modèle hydrologique a besoin d'une intensité pluvieuse. La relation convertissant la mesure du radar en information exploitable par le modèle hydrologique (par exemple une relation Z-R) possède ses propres sources d'erreurs. De même, les modèles hydrologiques prédisent souvent des valeurs de débits (du fait de la résolution de l'équation de conservation de la masse) alors qu'une information souvent plus pertinente d'un point de vue opérationnelle est la

hauteur d'eau de la rivière (dépassement ou non de cote d'alertes liées à des phénomènes d'inondations). La relation hauteur – débit (courbes de tarage) présente également des sources d'erreurs. La mise en relation (modèle d'observation) des variables mesurées ou comparées sur le terrain et celles utilisées ou prédies par le modèle présente donc des sources d'erreurs qualifiées d'**erreurs du modèle d'observation**.

- Enfin, les appareils de mesures qui in fine fournissent les mesures brutes à la base de la prévision (variables de forçage) ou à sa comparaison sur le terrain (réponse hydrologique) produisent eux mêmes des **erreurs de mesures** plus ou moins significatives.

Pour toutes ces raisons, quelque soit la qualité du modèle numérique de prévision hydrologique, l'ensemble de la chaîne de prévision est affectée par des sources d'incertitudes et d'erreurs. Ces erreurs ne sont pas toujours quantifiables et leurs interactions encore moins. L'assimilation de données propose un cadre mathématique pour tenter, non pas de quantifier ces incertitudes et erreurs, mais de réduire leurs impacts sur la précision de la prévision hydrologique.

Cette tâche s'intéresse à cette méthodologie. Du fait de l'impossibilité dans le cadre de ce projet de construire des modèles semi-physiques (cf. le rapport de la tâche 3), l'assimilation de donnée a été expérimentée exclusivement sur le modèle physique.

Le modèle physique n'est pas présenté dans cette tâche. Ces descriptions sont en effet présentes dans les rapports des autres tâches (notamment celui de la tâche 3). Nous concentrons la partie descriptive à l'assimilation de données telle qu'elle a été utilisée dans ce travail.

2. QUELLE APPROCHE DE L'ASSIMILATION DE DONNEES ?

Comme expliqué précédemment la propagation et l'interaction des différentes sources d'incertitudes (forçage, conditions initiales, paramètres et structure du modèle hydrologique) impliquent un écart entre les réponses hydrologiques simulée et observée, même après calibration du modèle.

Le principe de l'assimilation de données consiste :

- à mettre à jour de façon régulière (**intervalles de temps**)
- certains **facteurs d'entrée** de la modélisation (à déterminer)
- ou certaines **variables d'état** du modèle hydrologique
- afin de **réduire l'écart** aux **observations caractérisant la dynamique du système**.

La figure 2 suivante illustre ce principe.

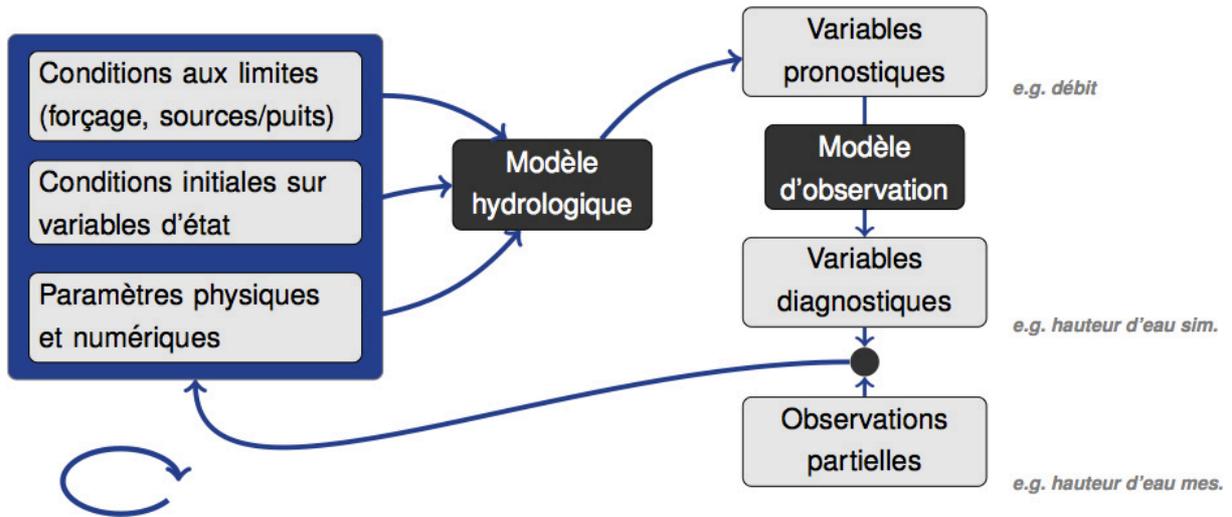


Figure 2 : principe de l'assimilation de données

Plusieurs techniques peuvent être utilisées pour mettre en œuvre ce principe de correction des états du modèle à partir des observations disponibles : une approche séquentielle ou une approche variationnelle.

2.1 L'APPROCHE SEQUENTIELLE

La figure 3 suivante illustre le principe de l'approche séquentielle.

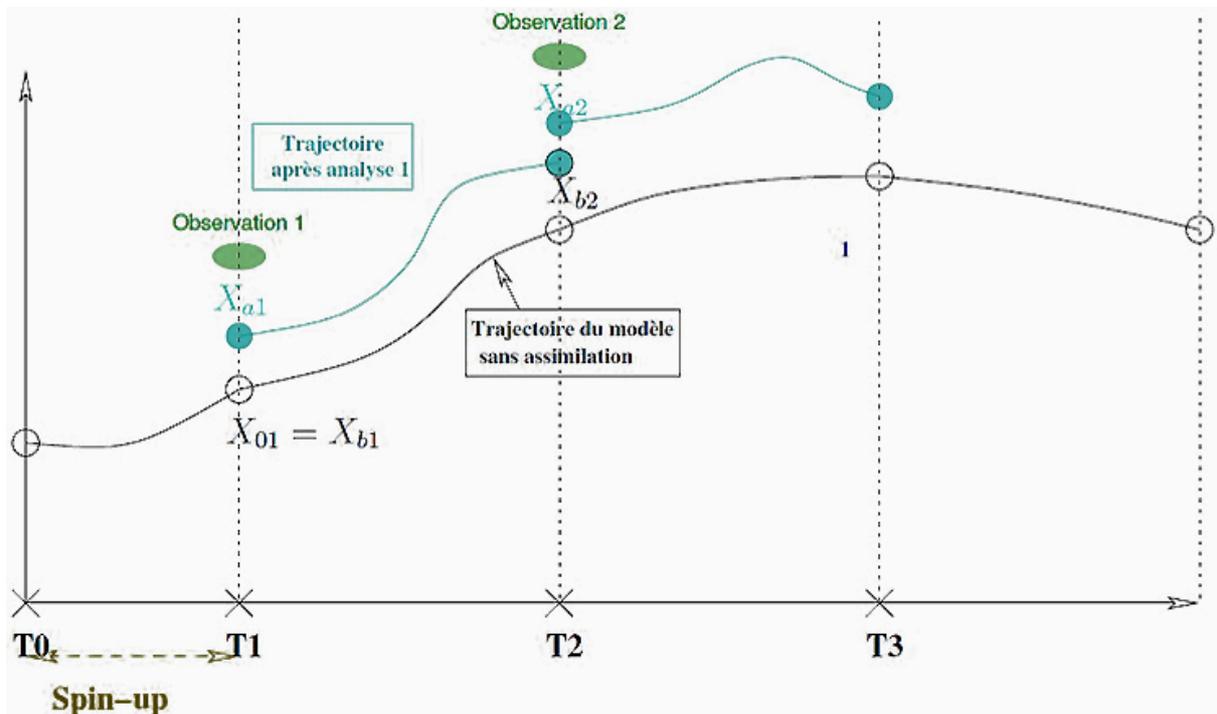


Figure 3 : approche séquentielle

Dans cette approche, certaines variables (à déterminer) du modèle sont régulièrement mises à jour à partir d'observations collectées et correspondant au pas de temps de l'assimilation de données. On voit bien sur cette figure que l'objectif de l'assimilation de données n'est pas de faire « coller » les prévisions du modèle aux observations mais de s'en approcher tout en respectant la cohérence interne du modèle. Autrement dit, un compromis est réalisé entre la diminution de l'écart entre la prévision aux observations et la modification des variables d'état du modèle.

Une des contraintes de cette approche séquentielle est la coïncidence entre le pas de temps de l'analyse de l'assimilation de données et celui des observations utilisées. Or ce pas de temps n'est pas toujours choisi en fonction des seules contraintes du meilleur bénéfice pour le modèle mais également en fonction du pas de temps des observations, du temps de calcul nécessaire pour réaliser l'analyse de l'assimilation de données, etc.

Pour ces raisons, entre autres, **nous avons préféré choisir l'analyse variationnelle** brièvement décrite ci-dessous.

2.2 L'APPROCHE VARIATIONNELLE

La figure 4 suivante illustre le principe de l'approche variationnelle.

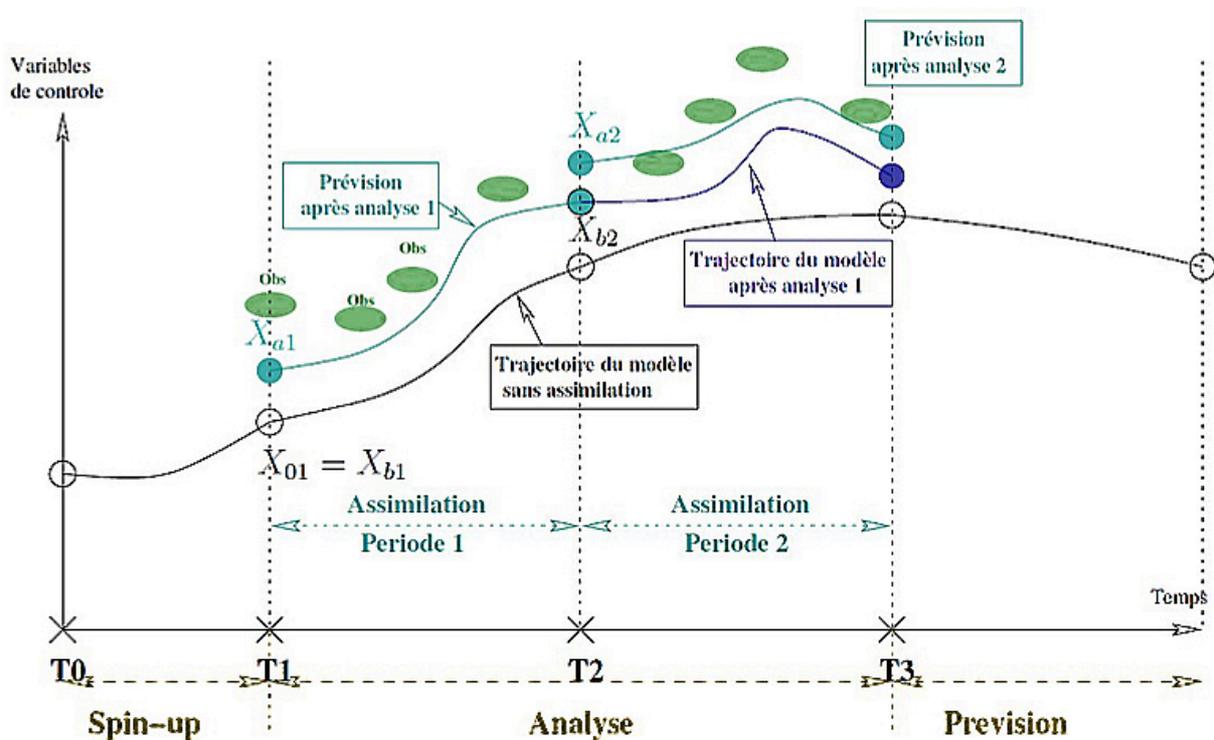


Figure 4 : approche variationnelle

Dans cette approche, le pas de temps de la période d'assimilation de données n'est pas nécessairement celui des observations. Cela permet en général d'utiliser

davantage d'observations pour corriger les variables internes du modèle de prévision. Si par exemple, le phénomène physique décrit par le modèle possède une certaine inertie par rapport aux variables d'observations, cela permet de densifier et de mieux renseigner la période d'analyse de l'assimilation par l'utilisation de davantage d'observations que dans l'approche séquentielle.

Plusieurs techniques d'utilisation du modèle de prévision sont également utilisables dans cette méthode variationnelle : par exemple utilisation des différences finies ou du modèle adjoint pour l'estimation de la Jacobienne. **Dans le cadre de ce projet nous avons choisi d'utiliser le modèle adjoint du modèle physique** pour sa rapidité d'exécution et la précision du calcul des dérivés.

3. EXPERIENCES D'ASSIMILATION DE DONNEES.

3.1 CONTEXTE OPERATIONNEL

Le cas d'étude de ce projet concerne la prévision des crues. Ce contexte opérationnel se caractérise par :

- des observations disponibles du forçage météorologique le plus souvent réduites à des précipitations moyennes sur le bassin. Ces valeurs de précipitations obtenues par agrégation de données de pluviomètres sont présentent des incertitudes en terme d'intensité et de phasage temporel du fait de l'impact de l'échantillonnage spatial et temporel de mesures ponctuelles d'un phénomène pouvant avoir une grande variabilité spatio-temporelle. Dans le cas de l'utilisation de données radar, la conversion des mesures de réflectivités en intensités de précipitations présentent encore de grandes incertitudes, notamment en fonction de la distance entre la localisation entre l'estimation de la précipitation radar et la localisation du radar. Même si des progrès significatifs ont été réalisés dans ce domaine (algorithmes de post-traitement) les erreurs résiduelles ont encore un grand impact hydrologique.
- des observations disponibles de la réponse hydrologique le plus souvent réduites aux hauteurs d'eau observées. Ces données sont intégratrices du fonctionnement de tout le bassin amont à la section de la rivière où sont mesurées ces hauteurs d'eau. Or les modèles hydrologiques (physiques et semi-physiques) utilisent plutôt la variable débit qui représente une variable de contrôle de la conservation de la masse. Or les courbes de tarage qui constituent le modèle d'observation (définition dans le paragraphe 1. Introduction) peuvent être très incertaines (interval de temps entre deux mises à jour, difficulté à jauger les débits extrêmes (très forts ou très faibles), etc.).

L'assimilation de données va ainsi permettre de surtout corriger des erreurs de phase et d'amplitude entre les réponses hydrologiques simulée et observée. Pour mettre en œuvre cette technique, une phase préliminaire de choix des variables utilisées pour l'assimilation est nécessaire.

3.2 CHOIX DES VARIABLES DE L'ASSIMILATION

Le choix des variables mises à jour par l'assimilation de données est une étape cruciale. Ce choix doit s'opérer en fonction des objectifs finaux de l'utilisation du modèle hydrologique. En l'occurrence, dans ce projet nous visons à améliorer la prévision des débits tout au long de la crue pour à la fois correctement prédire les débits de pointe des crues, les moments de dépassement de certaines côtes d'alerte, le moment de fin de la crue, etc.

Une démarche intuitive est possible. Par exemple, comme cela est couramment effectué, les conditions initiales sont souvent utilisées comme variable de contrôle. Cela est pertinent. Par exemple, les figures suivantes comparent les débits observés (trait continu épais), les débits initialement simulés (trait pointillé) et les débits corrigés par assimilation de données (trait continu fin). Deux expériences ont été réalisées : l'une en sous-estimant volontairement la condition initiale en humidité moyenne du bassin (figure 5) et l'autre en sur-estimant volontairement cette même condition initiale (figure 6).

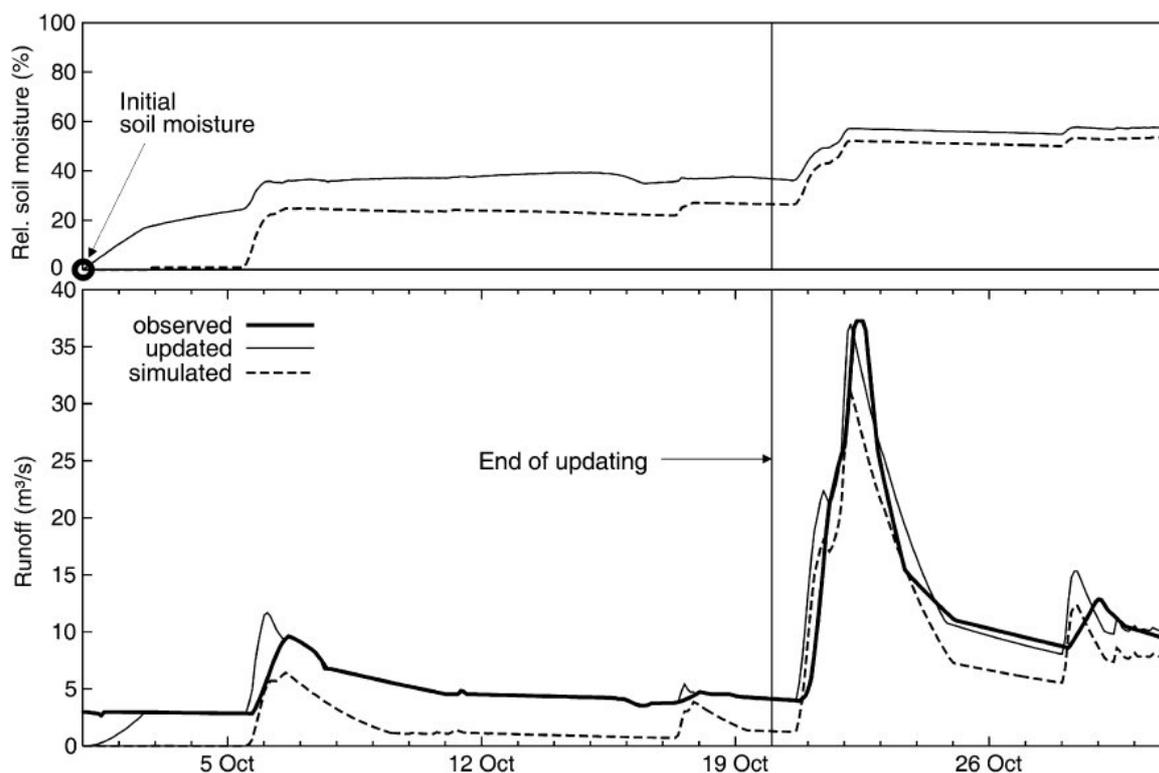


Figure 5 : utilisation de la condition initiale en humidité du bassin (sous-estimation) pour la correction des débits simulés par assimilation de données variationnelle.

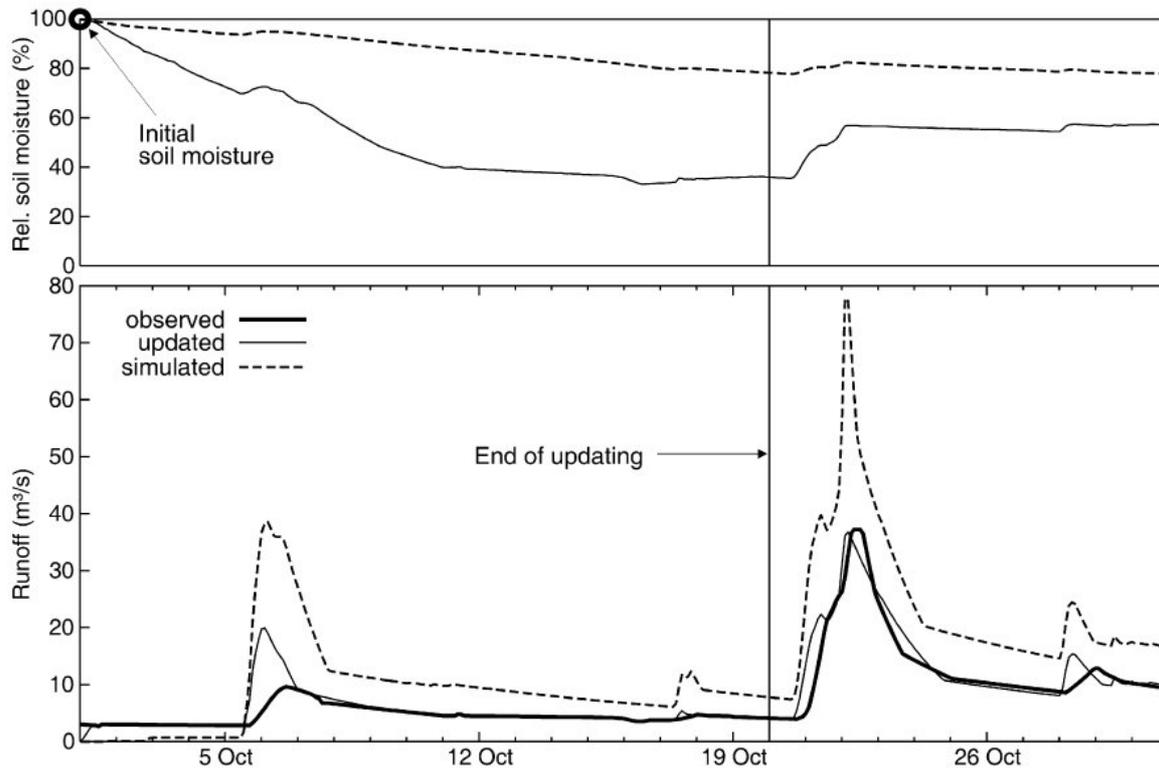


Figure 6 : utilisation de la condition initiale en humidité du bassin (sur-estimation) pour la correction des débits simulés par assimilation de données variationnelle.

Sur ces figures on peut constater que l'assimilation de données parvient à corriger les débits simulés volontairement sur-estimés et sous-estimés par modification arbitraire de la condition initiale en humidité moyenne du bassin. L'utilisation de cette variable de contrôle comme cela est classiquement fait est donc possible. On constate toutefois qu'à partir d'une certaine durée (ligne « End of updating » sur les deux figures), les débits simulés corrigés par l'assimilation de données ne sont plus influencés par cette correction : les deux simulations corrigées se confondent presque. **Corriger la condition initiale n'est pas suffisant pour améliorer les performances du modèle hydrologique sur l'ensemble de la chronique de crue.**

Une étude plus approfondie a donc été entreprise. Pour cela, par utilisation du modèle adjoint au modèle physique, **nous avons réalisé une décomposition temporelle des sensibilités du critère de Nash** (calculé entre les débits observés et les débits simulés après assimilation de données) **sur l'ensemble des épisodes de crues disponibles**. Un exemple est fourni dans la figure 7 suivante.

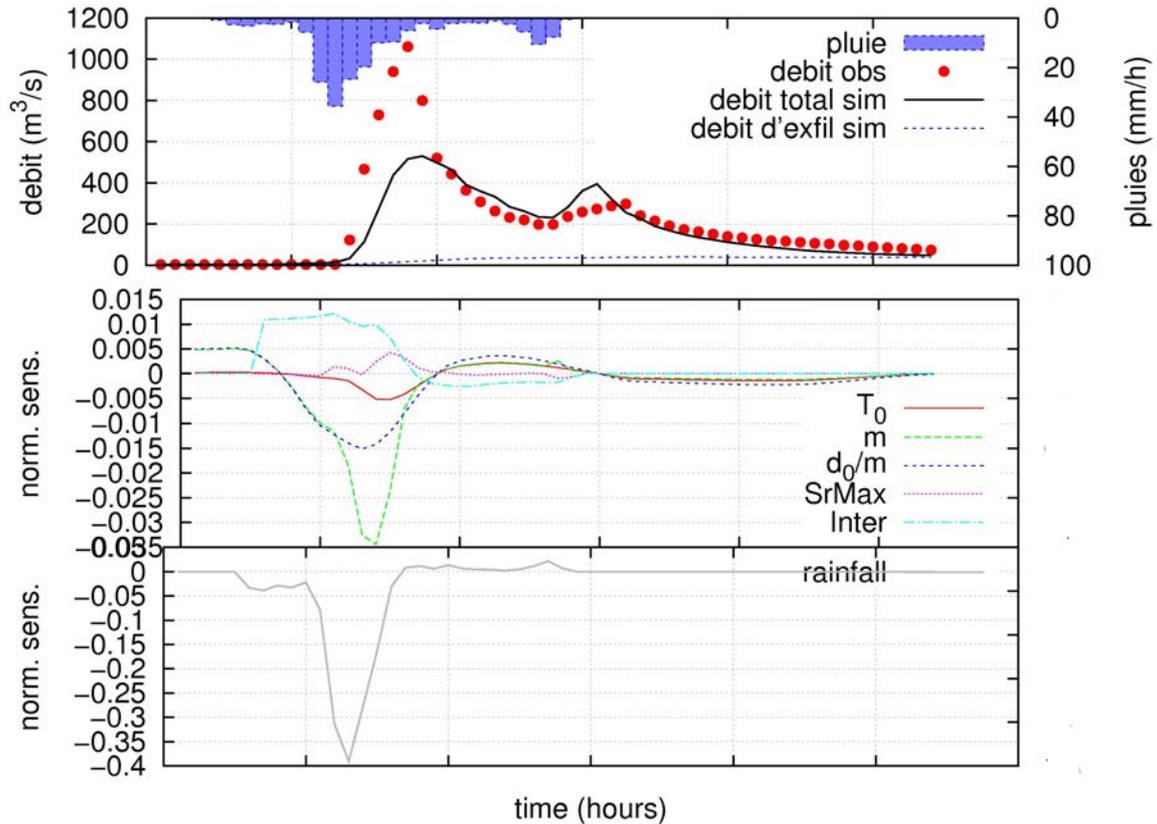


Figure 7 : décomposition temporelle de la sensibilité du critère de Nash aux paramètres du modèle et à la variable de forçage (ici la précipitation moyenne du bassin versant).

On constate sur cette figure 7 que les paramètres ont également une influence certaine sur la performance du modèle. Cette influence est variable selon les phases de la crue (montée, pic, récession). On retrouve ici les conséquences de l'incertitude de structure du modèle hydrologique : ses imperfections sont différentes selon les phases de la crue. Dans le cas du modèle physique cela s'explique par la pertinence variable des différents choix scientifiques et algorithmiques réalisés pour représenter les différents processus hydrologique (infiltration, ruissellement, transfert hydraulique, etc.) qui ont une importance relative variable selon la phase de la crue.

Cette étude, résumée ici, a permis *in fine* d'affiner le choix des variables de contrôles. Les paramètres du modèle et le contenu en eau moyen du bassin tout au long de l'épisode de crue (i.e. pas uniquement sa valeur initiale comme illustré précédemment) ont été identifiés comme les variables de contrôles optimales pour l'assimilation de données variationnelle du modèle physique.

3.3 SENSIBILITE AUX VARIABLES DE CONTROLE

Les figures suivantes illustrent la sensibilité de l'assimilation de données variationnelles en utilisant comme variable de contrôle (i) les paramètres du modèle, (ii) le contenu en eau moyen du bassin et (iii) les deux à la fois.

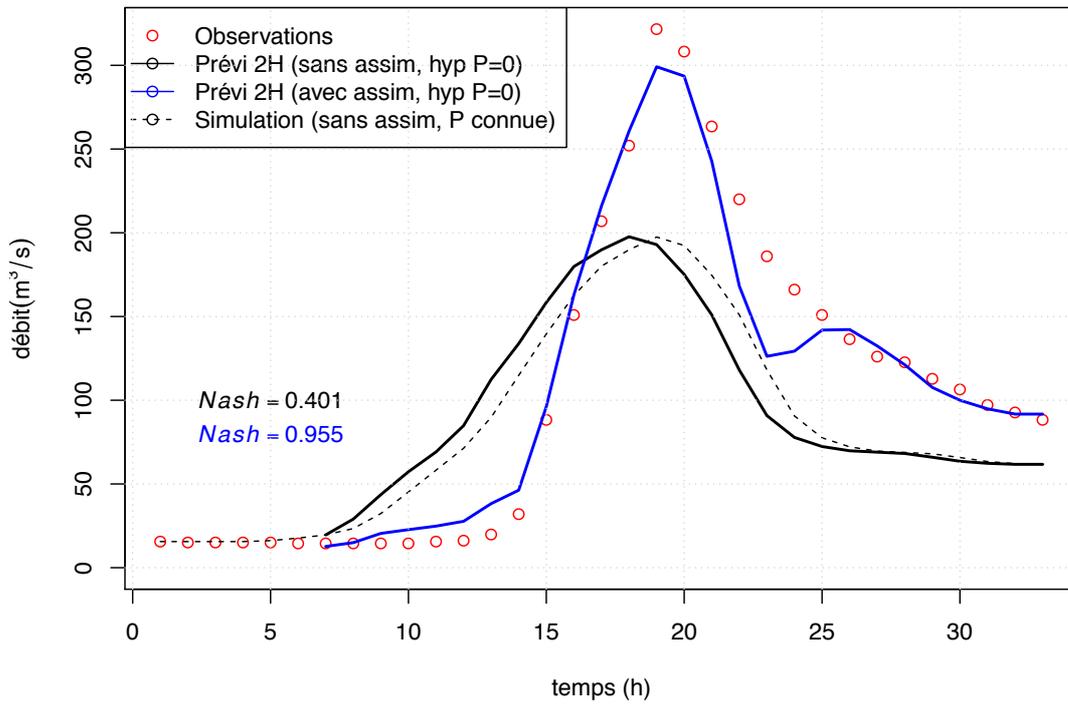


Figure 8 : Assimilation de données variationnelle avec utilisation des paramètres du modèle hydrologique comme variable de contrôle.

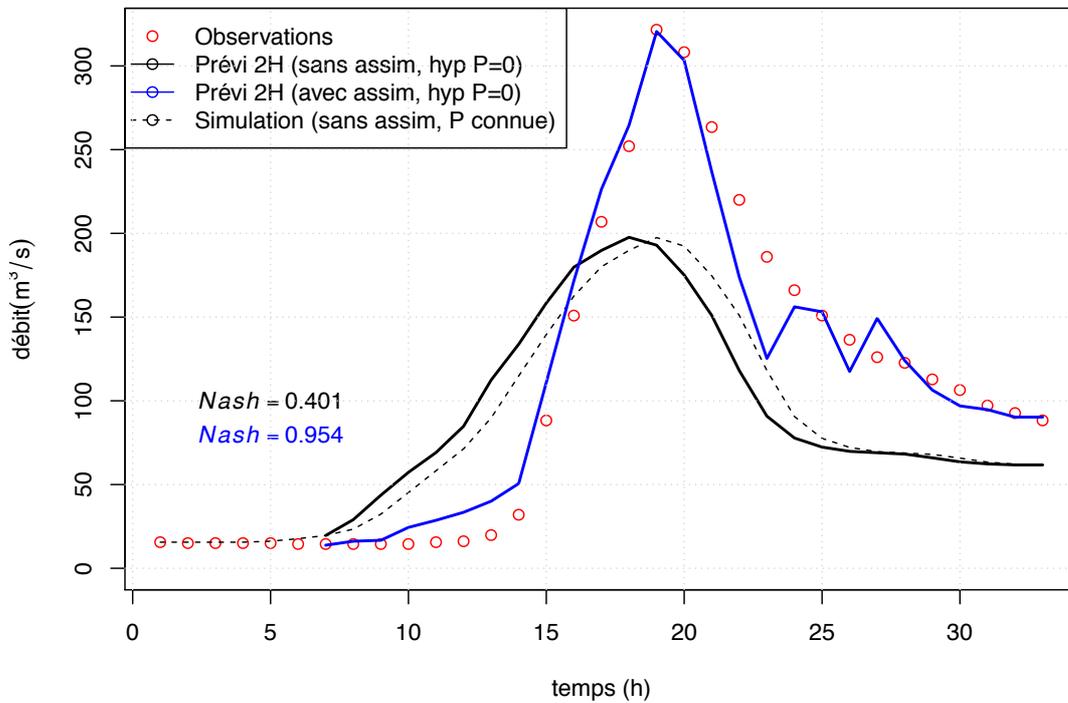


Figure 9 : Assimilation de données variationnelle avec utilisation du contenu en eau moyen du bassin comme variable de contrôle.

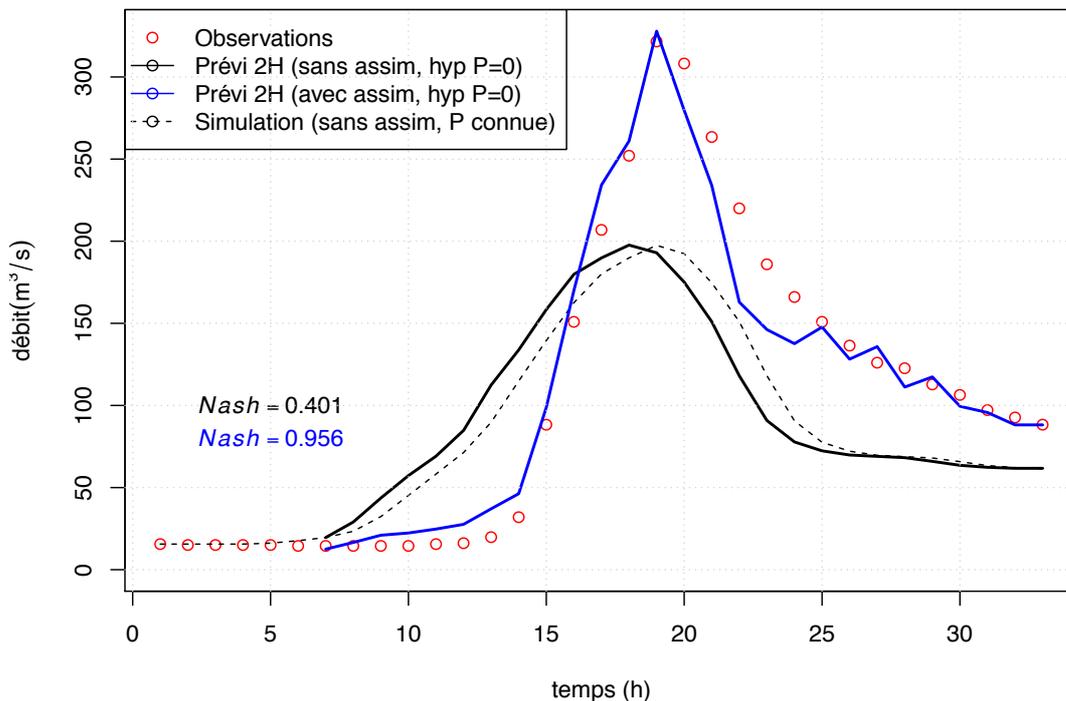


Figure 10 : Assimilation de données variationnelle avec utilisation du contenu en eau moyen du bassin et des paramètres du modèle hydrologique comme variable de contrôle.

Même si, comme on pouvait s'y attendre, l'assimilation de données utilisant comme variable de contrôle à la fois les paramètres du modèle hydrologique et le contenu en eau moyen du bassin donne les meilleurs résultats, on peut constater que l'assimilation n'utilisant que le seul contenu en eau donne de meilleur résultat que l'assimilation n'utilisant que les paramètres du modèle hydrologique (malgré l'avantage, non significatif, du critère de Nash). L'allure est mieux représentée et le pic de crue mieux simulé. C'est un constat général sur l'ensemble des épisodes de crues où a été appliqué cette assimilation de données :

- ajuster uniquement le contenu en eau moyen du bassin permet d'avoir déjà et toujours de bons résultats.
- Modifier uniquement les paramètres du modèle hydrologique n'est souvent pas suffisant.

Ceci est un résultat intéressant pour les perspectives de ce projet : au contraire de ce qui est généralement effectué (assimilation sur la condition initiale et les paramètres du modèle), **l'assimilation sur la variable d'état interne du modèle hydrologique contrôlant la résolution de l'équation de conservation de la masse est la plus intéressante vis à vis de la prévision des crues.**

3.4 SENSIBILITE A LA DUREE DE LA FENETRE D'ASSIMILATION

Comme expliqué dans le paragraphe 2, les techniques d'assimilation utilisent une durée de la fenêtre d'assimilation. Nous avons testé la sensibilité de la performance de l'assimilation de données à la durée de cette fenêtre d'assimilation. Une illustration sur un seul épisode mal simulé par le modèle physique est présentée ci-dessous.

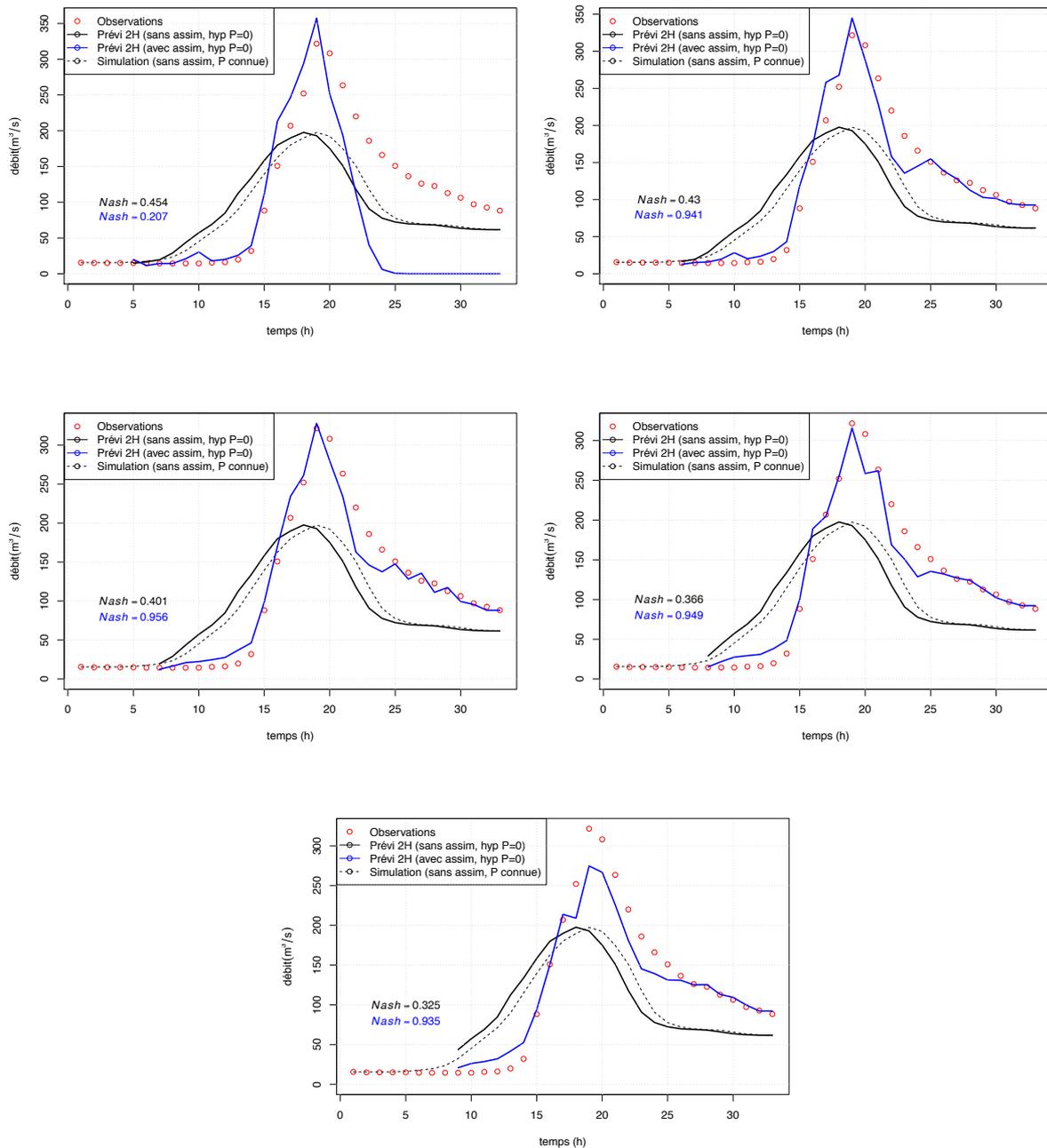


Figure 11 : Sensibilité de l'assimilation de données variationnelle à la durée de la fenêtre d'assimilation (de haut en bas , de gauche à droite : 3h, 4h, 5h, 6h et 7h)

Le résultat intéressant de cette étude de sensibilité est que **la durée optimale de la fenêtre d'assimilation est de 5h, c'est à dire égale au temps de réponse du bassin versant**. C'est un résultat que l'on peut estimer pertinent vis à vis de la physique, il est intéressant pour la généralisation de ce type de méthodologie de constater que ce résultat correspond également à l'optimum numérique. **Ce résultat facilitera l'application des résultats de ce projet à d'autres bassins versants** sans qu'il soit nécessaire de procéder à ces analyses numériques tout en étant rassuré sur la pertinence physique de cette approche.

3.5 INFLUENCE DE L'HORIZON DE PREVISION

Ces études de sensibilité ayant permis au sein du projet de dessiner un certain guide de pratiques à l'utilisation de l'assimilation de données variationnelle appliquée à la prévision des crues, il était intéressant de tester cette méthodologie non plus pour la correction des débits simulés en temps réel mais également pour la prévision des débits à des horizons temporels où les pluies sont encore inconnues.

Les figures suivantes illustrent la qualité de la prévision des débits simulés par le modèle hydrologique physique et corrigé par l'assimilation de données variationnelles pour des horizons de 2h, 4h et 6h sous hypothèse de pluies futures nulles (l'hypothèse couramment admise dans connaissance de pluies futures : cette hypothèse correspond également à la « prévision optimiste », c'est à dire celle correspondant à l'arrêt des pluies).

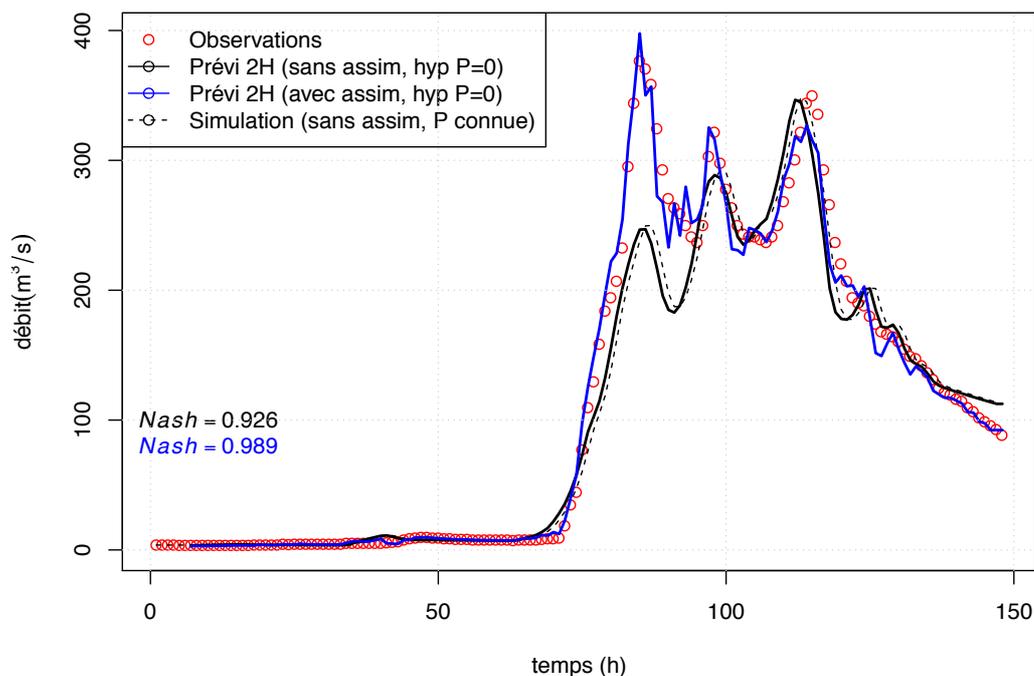


Figure 12 : Prévision à 2 heures avec assimilation

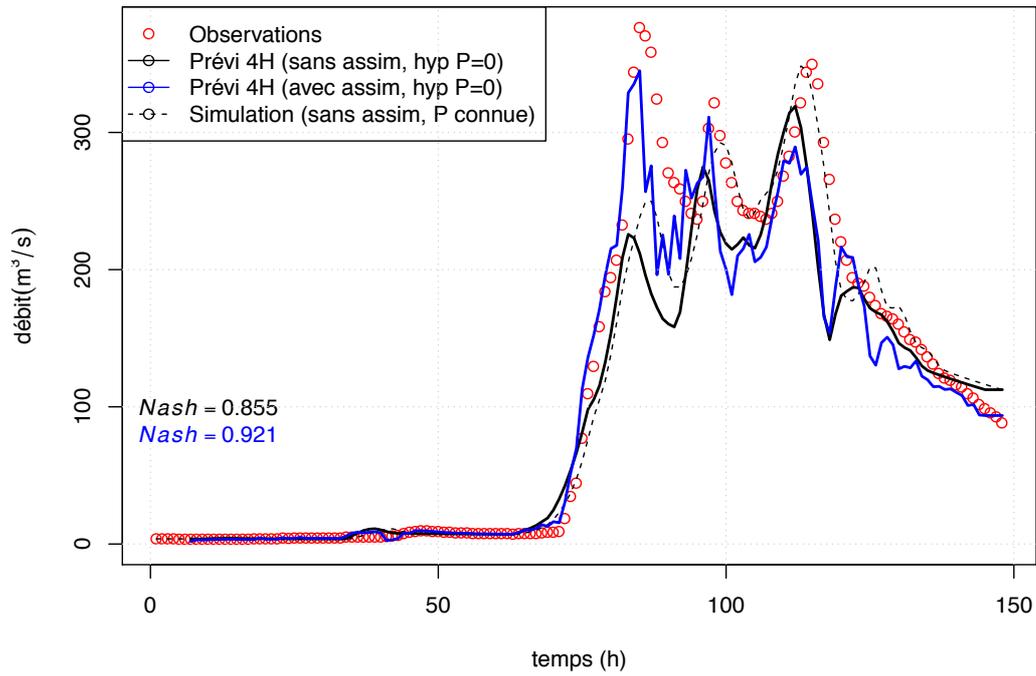


Figure 13 : Prévision à 4 heures avec assimilation

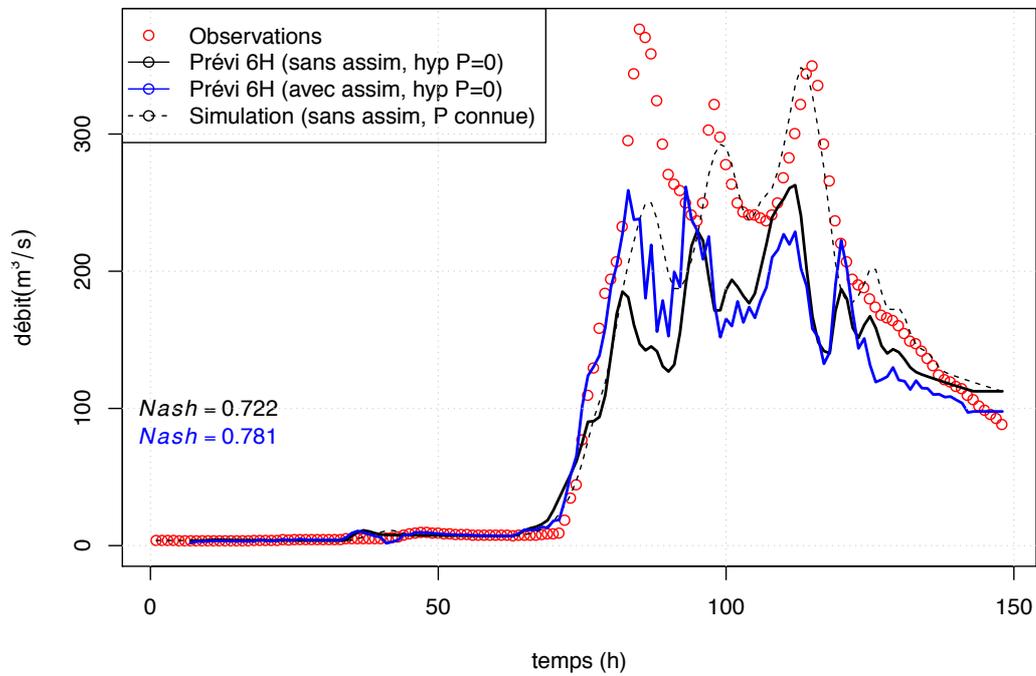


Figure 14 : Prévision à 6 heures avec assimilation

On constate que ces figures que **jusqu'à des horizons temporels inférieurs ou égaux au temps de réponse du bassin (ici 5h), la prévision des débits avec assimilation et sans connaissance des précipitations (hypothèse de pluies futures nulles), reste excellente** (par ex critère de Nash supérieur à 0.92 à 4h). Au delà, la performance des prévisions diminue sensiblement. Ceci est un résultat logique au sens ou au delà de l'inertie significative du bassin, les processus météorologiques conditionnent de façon prépondérante le devenir de la réponse hydrologique.

4. CONCLUSIONS - PERSPECTIVES

Ce projet a permis d'aborder l'utilité de l'assimilation de données pour la prévision des crues éclair. Les principaux résultats sont :

- L'assimilation de données variationnelle est la plus pertinente pour le contexte hydrologique
- La mise à jour des variables de contrôle en cours d'événement permet une prévision de très bonne qualité.
- Les résultats plaident pour une plus grande utilisation de modèle continu et non événementiel pour la prévision hydrologique
- Il est possible de fixer les paramètres de l'assimilation de donnée (durée de la fenêtre d'assimilation de donnée et horizon de prévision des débits sans connaissance des précipitations égales au temps de réponse du bassin).
- La principale limitation reste la prévision des pluies pour des horizons de prévision au delà du temps de réponse du bassin.
- La chaîne modèle + assimilation de données est très rapide quand l'assimilation utilise le modèle adjoint pour estimer la jacobienne.
- La prévision d'ensemble est facilitée. Il est possible d'assimiler différentes réalisations des courbes de tarages, des pluies prévues etc.).

Les résultats de ce projet sont donc transférables à d'autres bassins versants aux crues rapides. L'utilisation de l'assimilation de données, déjà habituelle en météorologie et océanographie, est donc possible et à encourager en hydrologie.

<http://blog.espci.fr/flash/>

