

# **Assimilation de données appliquée à un modèle hydrologique distribué : calage régional et assimilation des débits observés dans la méthode AIGA.**

La thèse sera co-encadrée par :

**Igor Gejadze** (Irstea, UMR G-EAU, Montpellier) et,

**Patrick Arnaud** (Irstea, UR RECOVER, Aix-en-Provence)

Accueil : Equipe RHAX de l'UR-RECOVER à Irstea Aix-en-Provence

Inscription à l'Ecole Doctorale GAIA à Montpellier

## **Résumé**

---

Le travail de thèse proposé vise à améliorer une approche d'anticipation des crues en milieu non-jaugé, appliquée sur l'ensemble du territoire national français (méthode AIGA implantée dans le service « Crues Soudaines » du Schapi : Vigicrue Flash – début 2017).

Cette approche, basée sur une modélisation hydrologique distribuée à la maille de 1 km<sup>2</sup>, alimentée par l'information des radars météorologiques, présente des marges d'amélioration sur deux points particuliers que l'on veut aborder dans ce travail de thèse :

- la pertinence en « descente d'échelle » ou « désagrégation spatiale », c'est-à-dire lorsque l'on applique la méthode sur des sous-bassins non-jaugés. Cela passe par un calage et une régionalisation des paramètres qui doivent être adaptés.
- la possibilité de prendre en compte les erreurs de modélisation en temps réel, lors de la réalisation de prévisions : pour une prévision en sites non jaugés, cela passe par l'assimilation des informations disponibles sur les sites voisins.

Ces deux points correspondent finalement à une même problématique d'assimilation de données par un modèle, pour le calage ou le recalage de ses paramètres. Ces points ont été abordés jusqu'à présent par des approches plus ou moins classiques en hydrologie. Les compétences en assimilation de données variationnelle de type 4D-Var portées par Irstea, nous conduisent à vouloir appliquer ces méthodes à nos approches. L'objectif de cette thèse est donc de mettre en œuvre des méthodes d'assimilation de données performantes pour le traitement d'approches de grandes dimensions (modèle distribué de prévision hydrologique à l'échelle du territoire français).

## **Contexte et enjeux**

---

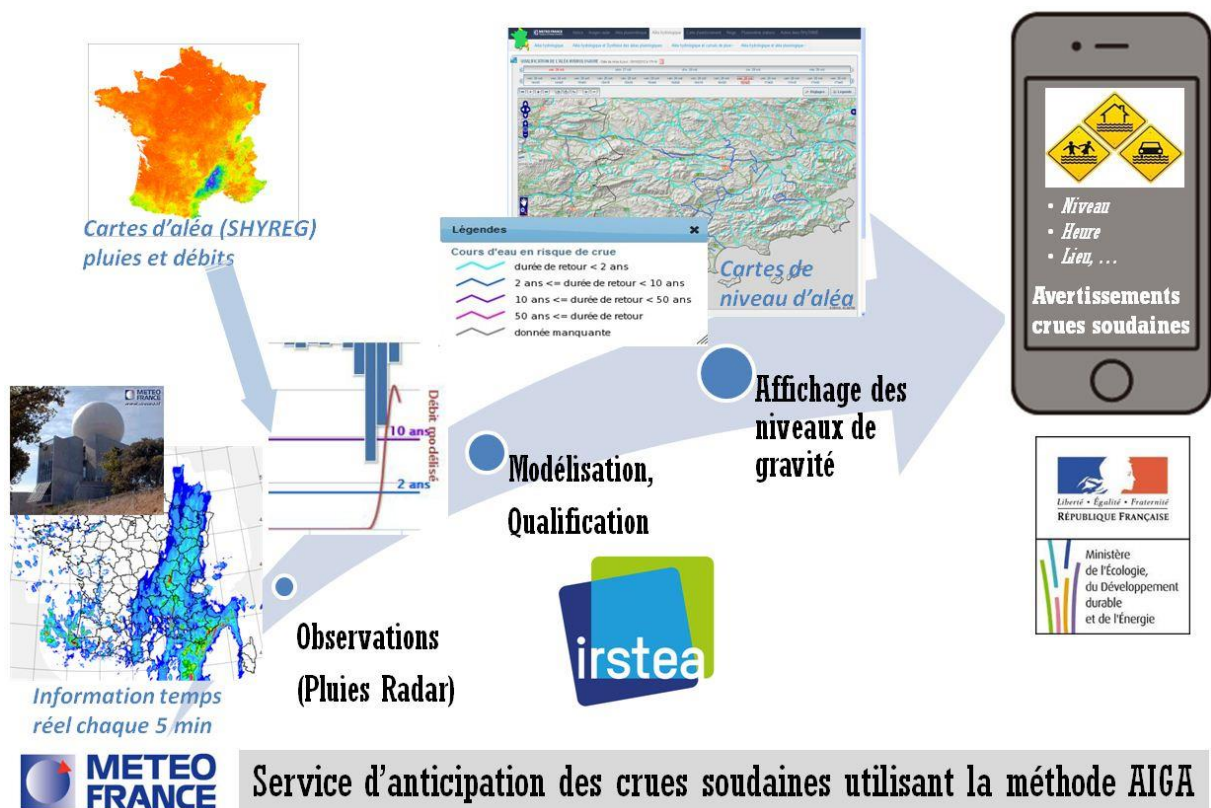
La méthode AIGA été développée depuis le début des années 2000 à Irstea et a pour objectif l'anticipation des crues rapides (Javelle et al., 2010; Lavabre and Gregoris, 2005). Elle trouve son intérêt sur les petits bassins non jaugés, pour lesquels il n'y a pas de suivi en temps réel des débits. Cette méthode est actuellement opérationnelle sur la plateforme RHYTMME, en région PACA. En janvier 2017, elle sera appliquée à l'ensemble du territoire national, dans le cadre du service d'avertissement « crues soudaines » du SCHAPI.

Cette méthode met en œuvre un modèle distribué simplifié pluie-débit (de la famille GR) qui transforme en débit l'information pluviométrique fournie par les radars météorologiques (Fig

1). Ces débits sont ensuite comparés à la climatologie du modèle (débits simulés à partir de toute la réanalyse pluviométrique de Météo-France, COMEPHORE) afin de déterminer la période de retour de l'événement en cours et caractériser ainsi la gravité de l'événement.

Développée spécifiquement pour le non-jaugé, la méthode n'assimile pas de débit. Cependant, des résultats sont également calculés aux exutoires de bassins jaugés. Une piste d'amélioration est donc l'utilisation en temps réel de l'information disponible aux stations jaugées pour corriger les résultats sur l'ensemble des sorties du modèle, jaugées ou non.

Ce travail ne remet pas en cause l'intérêt ni les performances attendues et obtenues des approches en bassins jaugés, qui mieux ciblées sur un bassin particulier et avec peu de paramètres et peu d'états n'ont pas à propager des corrections sur plusieurs centaines ou milliers de mailles réparties sur tout ce bassin, ni à traiter des éventuelles corrélations inter-bassins. Et la thèse ne visera pas à comparer ces 2 approches.



*Fig 1. Principe de la méthode AIGA mise en œuvre dans le projet « Crues Soudaines »*

L'assimilation de données est une méthode couramment répandue dans le domaine de la prévision, notamment en prévision météorologique. Même si certains principes de base sont anciens (ex. : interpolation optimale, moindres carrés, optimisation, filtrage), de nombreux travaux récents ont mis au point de nouvelles méthodes et leurs implémentations et applications ont été rendues possibles grâce aux progrès tant théoriques que numériques et informatiques. Le Centre Européen pour les prévisions météorologiques à moyen terme de Reading a adopté la méthode dite 4D-Var en 1997 (soit 11 ans après l'article présentant cette méthode : Le Dimet, Talagrand 1986), avec une amélioration notable des performances de la prévision par rapport aux méthodes précédentes (Rabier et al., 2000). Les concepts peuvent vite s'avérer extrêmement compliqués à mettre en œuvre du fait du grand nombre de variables présentes dans le système, du nombre important d'observations tant spatialement que

temporellement, de la nécessaire mais difficile caractérisation des incertitudes sur les observations et sur les modèles, et éventuellement des dynamiques non-linéaires du système étudié. En hydrologie, ce type d'approche a déjà été testé pour le modèle à base physique Safran-Isba-Modcou (Thirel et al, 2010ab). Dans le cadre des modèles GR globaux développés à Irstea, les méthodes d'assimilation peuvent être plus simples. Ainsi, le modèle de prévision des crues GRP utilise une simple mise à jour de son réservoir de routage : une seule observation de débit modifie un seul état du modèle (Berthet, 2010). Cette méthode ne peut pas fonctionner dans le cadre de la méthode AIGA où différents débits sont à assimiler, et où les paramètres à modifier sont spatialisés. Des premiers tests d'assimilation de données ont été réalisés avec le modèle GRDc (GR Distribué continu, Fig 2) avec la méthode variationnelle 4D-Var en collaboration avec Irstea-Montpellier courant 2016 (Jay-Allemand et al., 2016). La mise en œuvre technique de la méthode 4D-Var sur ce modèle est donc fonctionnelle (calcul du modèle adjoint et premiers tests en calage Fig 3 et en prévision Fig 4) et les premiers résultats sont très encourageants sur des cas simples.

L'objectif de ce travail de thèse est de poursuivre les travaux initiés avec cette méthode 4D-Var, en appliquant la technique d'assimilation aux deux points importants posés par la mise en œuvre du modèle AIGA :

- Réaliser le « calage distribué » du modèle afin d'améliorer à la fois la régionalisation des paramètres et la pertinence de la désagrégation de l'information à l'échelle des mailles de calcul ( $1 \text{ km}^2$ ).
- Assimiler des données de débits observés afin d'améliorer la prévision des débits en sites non-jaugés.

Ces 2 points ne sont pas 2 thèses différentes, mais bien 2 volets complémentaires de la même thèse. En effet, les outils pour résoudre ces 2 problèmes dans un cadre d'AD Variationnelle sont exactement les mêmes : un algorithme de minimisation basé sur l'adjoint du modèle utilisé dans la formule du gradient. Les incertitudes sont traitées et présentes également dans les 2 points. Les techniques de régularisation sont également présentes dans les 2 points. Enfin il est possible, de par la puissance de ces méthodes, de séparer et réaliser en séquence les 2 points (calage puis assimilation des états initiaux pour la prévision) ou également de tout réaliser en même temps (assimilation des états initiaux pour la prévision et amélioration du calage simultanément, ce qui permettrait de faire évoluer les paramètres en cours d'évènement). Seul les variables et les paramètres mis dans le vecteur de contrôle sont différents.

Sur les aspects « assimilation de données » plusieurs questions scientifiques devront également être étudiées du fait des spécificités de cette application, en particulier sur l'estimation des incertitudes certainement non Gaussiennes, sur la fonction coût à minimiser et sur la présence de contraintes inégalités sur les paramètres (paramètres A et B > 0) et les variables (états  $0 < X_A < A$  et  $0 < X_B < B$ ).

Ces travaux seront réalisés en collaboration entre l'équipe Hydrologie d'Irstea-Aix pour les aspects liés à la méthode AIGA et l'équipe GHOSTE - G-EAU à Montpellier pour les aspects liés à l'utilisation de la méthode 4D-Var. La première a une reconnaissance nationale et internationale dans le domaine de l'hydrologie, avec en particulier des outils utilisés en opérationnel par le Schapi et les SPC. La seconde a une reconnaissance nationale et internationale dans le domaine de l'Assimilation de Donnée appliquée aux domaines hydrologiques et hydrauliques (Malaterre et al. 2016 conférence plénière au Colloque

National d'Assimilation de données, Oubanas et al. 2016 publication franco-américaine, Malaterre P.O. Membre de la Science Team de la Mission CNES-NASA SWOT, Oubanas et al. 2015).

## Etat de l'art

---

- **Modèles distribués : le problème de l'équifinalité et de la descente d'échelle**

Depuis maintenant deux décennies, la modélisation hydrologique peut compter sur l'information spatiale de la pluie fournie par les radars météorologiques. En effet, que ce soit une information « réanalysée » (SAFRAN, COMEPHORE) ou une information en « temps réel » (lame d'eau PANTHERE), les chroniques de pluies discrétisées à la maille de 1 km<sup>2</sup> sont disponibles en France mais aussi dans de nombreux pays. Cette évolution a progressivement conduit les hydrologues à utiliser cette information spatiale dans leurs modèles, qu'ils soient globaux ou distribués.

De nombreuses études ont débattu sur l'intérêt de prendre en compte cette information spatiale par le biais d'approches globales ou distribuées (Andréassian et al., 2004; Arnaud et al., 2011; Chaubey et al., 1999; Zoccatelli et al., 2010). Cependant, les conclusions ne sont pas toujours tranchées et dépendent souvent du type de bassin étudié (nature, taille, etc.), de sa climatologie, de la disponibilité des données, des objectifs de la modélisation, etc.

Dans l'approche AIGA, la modélisation hydrologique du modèle GRD est dite « Distribuée » (Javelle et al., 2010), c'est-à-dire que l'on considère une transformation de la pluie en débit sur des mailles régulières, puis les débits générés sur chacune de ces mailles sont agglomérés / routés sur le bassin versant pour fournir une information à l'exutoire du bassin versant. Dans cette approche, l'information spatiale de la pluie peut donc être prise en compte. Cependant, le manque d'information de calage ne permet pas toujours de déterminer un jeu de paramètres hydrologiques différent sur chaque maille de modélisation.

On rejoint ici le problème de l'équifinalité des paramètres, le degré de liberté du modèle devenant trop grand par rapport aux points de calage (Beven, 1993). Dans ce cas, différentes approches sont possibles pour le calage des paramètres d'un modèle hydrologique distribué : (1) soit on fait l'hypothèse que les paramètres sont spatialement homogènes sur le bassin versant ou sur des sous-entités du bassin (Emery 2017), mais on perd alors la variabilité spatiale du fonctionnement hydrologique du bassin versant (mais pas de la pluie) ; (2) soit on pondère les valeurs de paramètres en fonction de descripteurs physiques pour prendre en compte une certaine variabilité spatiale du fonctionnement hydrologique (Roux et al., 2011), mais on part alors d'un a priori sur les causes du fonctionnement hydrologique ; (3) soit on utilise une procédure de calage numérique plus ou moins performante (Huot, 2014) mais on est confronté plus fortement au problème de l'équifinalité des paramètres évoqué plus haut.

Quelle que soit la méthodologie adoptée, la difficulté demeure dans la validation des hypothèses sur la variabilité spatiale des paramètres. Les procédures de validation croisée sont difficilement adaptées pour explorer à la « descente d'échelle » ou la « désagrégation spatiale des modèles » (c'est-à-dire la validation du modèle sur des petits bassins non jaugés). En effet la variabilité spatiale des paramètres à l'intérieur du bassin versant peut difficilement être vérifiée par l'observation lorsque l'on sait qu'en France, seulement 3 % des bassins versants

jaugés ont une surface de moins de 10 km<sup>2</sup>, alors qu'ils représentent 40 % des exutoires non jaugés (calcul effectué à partir de la BNBV – Base Nationale des Bassins Versants).

L'intérêt des méthodes distribuées, qui proposent de par leur nature une information sur la « descente d'échelle », pourrait alors être remis en cause par la difficulté à valider l'information fournie. Il est donc important de trouver des méthodes pertinentes face à la descente d'échelle. L'enjeu est d'autant plus important que c'est à travers cette descente d'échelle que l'on pourra fournir une information sur les bassins versants non-jaugés qui sont majoritairement des sous-entités de bassins jaugés.

- **Prévision des crues en sites non-jaugés : « assimilation » régionale**

La prévision des crues en sites non-jaugés présente plusieurs spécificités liées au fait que l'information débitométrique n'est pas disponible à l'exutoire où la prévision va être faite. Cette information va manquer lors du calage des modèles, mais aussi lors de leur mise en œuvre en temps réel puisqu'ils ne pourront pas s'appuyer sur une information locale pour corriger leurs erreurs, contrairement aux approches en milieu jaugé (Seo et al., 2003, Bessière, 2008, Berthet, 2010). Il faut alors utiliser l'information des bassins versants jaugés voisins pour caler les modèles et éventuellement corriger les erreurs (Randrianasolo et al., 2011).

Si l'on considère que pour un modèle, l'assimilation de données correspond à un calage de ses paramètres et/ou une estimation de ses états, dans le cas des modèles de prévision de crues ce calage peut être réalisé sur l'ensemble des données (on parlera du « calage » à proprement dit du modèle) ou sur une partie des données pour corriger des erreurs (on parlera alors de « recalage des états » ou de « mise à jour » du modèle). Dans le cas de la prévision en sites non-jaugés, la correction des erreurs devra se faire à partir de l'information des bassins voisins. Outre la problématique classique associée à la correction des erreurs (méthode à choisir, influence du critère de recalage, quelle information utiliser, etc.), se rajoute la problématique du choix des bassins versants voisins permettant le transfert de l'information.

En effet, les performances des procédures d'assimilation, que ce soit pour le calage des paramètres ou pour la mise à jour du modèle, devront être jugées en fonction de différentes hypothèses de transfert de l'information de bassins jaugés vers un bassin non-jaugé. Les procédures d'assimilation sont alors confrontées aussi aux problèmes de régionalisation. Ces procédures d'assimilation dites « régionales » sont plus complexes car il n'y a pas de consensus sur la meilleure information et/ou la meilleure méthode de régionalisation à utiliser pour estimer des paramètres (ou débits) d'un modèle en site non-jaugé.

- **Assimilation de données de type 4D-Var**

L'assimilation de données variationnelle de type 4D-Var est basée sur la minimisation d'un critère J, appelé également « fonction coût », intégrant les mesures spatialisées et temporalisées. Pour que cette minimisation soit efficace il est nécessaire de disposer de l'adjoint du modèle. Ce travail a été réalisé récemment en collaboration entre les équipes de Montpellier et d'Aix (Jay-Allemand et al., 2016). La définition du critère J dépend du choix d'un vecteur d'ébauche (information a priori sur la solution) et de matrices de covariances d'erreur. Ces choix sont essentiels et devront être l'objet de réflexions intégrant des considérations physiques et algorithmiques. C'est en particulier par ces choix que l'on peut

indiquer comment répartir spatialement et/ou temporellement une information locale sur les états à corriger.

Même si les techniques de type 4D-Var sont appliquées avec succès depuis près de 20 ans sur des modèles océan-atmosphère pour les prévisions météorologiques, intégrant des états de dimension dépassant le milliard de variables, et utilisant plusieurs centaines de millions de mesures, le problème étudié ici est considéré comme de grande dimension. En effet, le modèle AIGA étant décliné à l'échelle nationale sur des mailles de  $1 \text{ km}^2$ , il y a environ 500 000 mailles, chacune possédant plusieurs états et paramètres. Les mesures, quant-à-elles, même si elles sont restreintes aux bassins versant jaugés représentent quand même plusieurs dizaines de milliers de mesures. Les matrices de covariances correspondantes seront donc délicates à mettre au point. Le modèle est cependant suffisamment simple et à peu de paramètres (pour chaque maille individuelle) pour qu'on puisse espérer des bonnes performances numériques permettant une utilisation opérationnelle de l'outil, qui est bien entendu un objectif du travail. Même si les Bassins Versants sont des entités pouvant être individualisées, il est probable que les corrélations sur les entrées (pluies), sur les états initiaux (remplissages des réservoirs) ainsi que sur les paramètres (caractéristiques physiques régionales) obligent ou au minimum incitent à considérer le problème comme global (tous les BV ensemble), ou au minimum semi-global (un BV et ses voisins), plutôt que chaque BV indépendamment, ce qui aurait l'avantage de diminuer la taille du problème à résoudre.

Le travail devant être effectué n'est pas une « simple application » de la méthode 4D-Var à un nouveau contexte, mais présente plusieurs questions scientifiques originales nécessitant des développements théoriques :

1 : le premier défi est de réaliser l'estimation des erreurs dans les entrées (pluies), les sorties (débits) et les modèles par réanalyse. Ces erreurs doivent aussi dépendre des bassins versants considérés, et ces erreurs sont certainement non Gaussiennes. Cet aspect oblige à utiliser un cadre bien moins classique et nécessitera des adaptations.

2 : le second défi est de choisir le meilleur estimateur a posteriori pour définir la fonction coût à minimiser et ses fonctions ou éventuellement matrices de pondération afin d'obtenir le meilleur mode (ce n'est pas le  $Q^{-1/2}$  classique). Ce point est difficile car les erreurs sont non Gaussiennes.

3 : le troisième défi est d'estimer les meilleures prévisions sur un horizon temporel futur, puisque le modèle avec assimilation de données sera utilisé dans ce cadre, en combinant ce modèle avec assimilation et des prévisions météo. Etudier par exemple des prédictions d'ensemble sur les prévisions de pluies.

4 : enfin il faudra tenir compte de contraintes inégalités lors du processus de minimisation (sur les débits  $Q$  et certains paramètres devant être positifs par exemple), ce qui est également une spécificité non classique (Bardsley J.M., Vogel 2003).

Une fois la méthodologie et les outils mis en place (une base a été réalisée en 2016 pour en tester la faisabilité, avec succès), l'adjoint peut également être utilisé pour analyser les sensibilités des sorties du modèle aux différents paramètres, conditions aux limites et états. Il peut également être utilisé pour optimiser un réseau de mesures afin d'améliorer au mieux la qualité du modèle et en particulier les prévisions (« experimental design » : si on peut ajouter une mesure, où la placer de manière optimale pour fournir l'information la plus riche).

## Descriptif des travaux

---

- **Les données**

L'équipe d'accueil dispose déjà d'un jeu de données important pour tester les différentes méthodes et options à mettre en œuvre dans ce travail de thèse. Les données d'entrée des modèles à tester sont des données hydro-climatiques disponibles sur l'ensemble du territoire français. Cette zone d'étude est très large et permet de tester les approches dans des contextes climatiques (tempérés, méditerranéens, alpins) et hydrologiques très différents (gamme de surface, occupation des sols, nature des sols...).

Les données pluviométriques disponibles sont des chroniques d'images radar (ré-analyse COMEPHORE et lames d'eau ANTILOPE) sur le territoire français de 1997 à 2015 au pas de temps de 15 minutes. Les données SAFRAN sont aussi disponibles pour fournir une information sur la partie solide des précipitations et pour estimer l'ETP quotidienne.

Les données hydrométriques sont disponibles sur les bassins de la banque HYDRO sélectionnés lors de multiples études précédentes.

Les informations en sites non jaugés sont aussi disponibles suite à une mutualisation des données fournies par des REX (retours d'expérience) suite aux derniers événements majeurs.

→ Détermination des jeux de données : le travail sur les données consistera surtout à faire un choix sur la période et la zone (ou bassins) d'étude. Dans un premier temps les méthodes seront testées sur un échantillon réduit. Une généralisation sera alors possible lorsque les méthodes seront bien définies.

- **Les outils**

Le modèle hydrologique utilisé sera le modèle GRD, modèle hydrologique simplifié de la famille des modèles conceptuels GR, dans sa version distribuée à la maille kilométrique (Fig 2). Les dernières évolutions de ce modèle permettent son utilisation au pas de temps de 15 minutes de façon continue.

Des travaux sur la régionalisation du modèle ont été effectués mais suivant des hypothèses simplificatrices d'homogénéité de certains paramètres sur les bassins versants. Ce travail de thèse doit permettre d'aller plus loin dans la régionalisation de ce modèle. Les premiers tests sont prometteurs, avec 2 jeux de paramètres différents sur le même bassin en expérience jumelle, avec vérification de la convergence de la méthode, ainsi qu'avec des données réelles (Fig 3) (Jay-Allemand *et al*, 2016).

Le modèle a déjà été préparé pour pouvoir fonctionner avec la méthode 4D-Var, en particulier le modèle « adjoint », nécessaire pour la mise en œuvre de la méthode d'assimilation a déjà été codé et testé en 2016 (Jay-Allemand *et al*, 2016).

→ Etude de sensibilité du modèle à ses paramètres : ce travail permettra l'appropriation du modèle par le doctorant et la détermination des paramètres importants.

→ Construction du modèle « adjoint » : le modèle « adjoint » a déjà été construit sur le modèle GDR pour pouvoir mettre en œuvre les procédures d'assimilation de données variationnelles. Si des améliorations du modèle GRD sont disponibles au cours de la thèse, il sera alors possible de les prendre en compte par la construction de modèle « adjoint » associé.

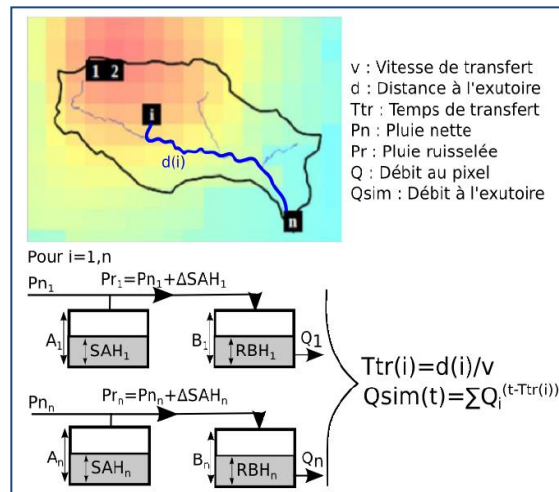


Fig 2. Structure du modèle GRDc

### • Cadre méthodologique

L'objectif est de proposer des améliorations sur les performances d'une méthode de prévision des crues en sites non jaugés. Il est alors important de mettre en œuvre la méthode jusque dans sa configuration d'utilisation finale pour pouvoir juger les différents tests effectués. Ainsi, il faudra réaliser les différentes étapes de mise en œuvre de la méthode.

→ Mise en place de la chaîne de calcul pour :

- le calage des paramètres du modèle sur un jeu de bassin « jaugés » et validation sur des bassins n'ayant pas participé au calage ;
- la prévision sur les bassins dits « de validation » avec assimilation des débits grâce aux observations sur les bassins « jaugés », *i.e.* ayant participé au calage.

→ Détermination des critères d'évaluation : un travail devra être réalisé pour définir les critères de performances de la méthode afin d'évaluer les différents tests effectués. Le panel des critères classiques pour l'évaluation des modèles de prévisions et d'alerte, largement utilisés par l'équipe d'accueil (critères d'écarts sur les débits, critères de contingence sur les alertes, critères sur l'anticipation...), pourra être sélectionné et/ou étoffé.

→ Mise en place de tests permettant de juger de la robustesse des résultats, basés sur différents échantillonnages des données de calage et de validation (ou contrôle). Par exemple, pour tester la descente d'échelle, le calage sur les grands bassins et la validation sur les petits.



- **Procédures de calage du modèle hydrologique distribué**

L'objectif est de proposer une méthode de calage du modèle GRD. Pour les raisons présentées plus haut, le calage d'un modèle distribué reste délicat et plusieurs options se présentent pour la prise en compte de la variabilité spatiale des paramètres. Plusieurs calages possibles seront testés :

- Calage classique de la méthode : les paramètres calés sont supposés spatialement homogènes sur le bassin versant. Une procédure de régionalisation est alors effectuée pour tester la pertinence de la paramétrisation sur les bassins de validation assimilés aux bassins non-jaugés. C'est la version classique de mise en œuvre de la méthode, pouvant servir de « version initiale » à améliorer.
- Calage par la méthode 4D-Var : la méthode 4D-Var est capable de caler les paramètres du modèle distribué en attribuant un jeu de paramètres à chaque maille du bassin. L'intérêt est de proposer une variabilité spatiale des paramètres à l'intérieur du bassin de calage. On peut aussi prendre en compte le calage simultané de plusieurs bassins versants emboîtés. Dans ce travail, il faudra étudier la robustesse des calages proposés (en lien avec l'équifinalité des calages possibles).
- Calage 4D-Var « contraint » : le calage des paramètres peut être contraint à suivre une variabilité spatiale plus proche d'un *a priori* physique du phénomène. Des tests seront effectués pour imposer une certaine variabilité aux paramètres optimisés, pour intégrer une expertise hydrologique aux calages (zones urbaines, occupation du sol homogènes,...). Ces contraintes peuvent par exemple limiter la variabilité des paramètres sur des zones homogènes imposées par des classes d'occupation du sol, d'hydrogéologie...

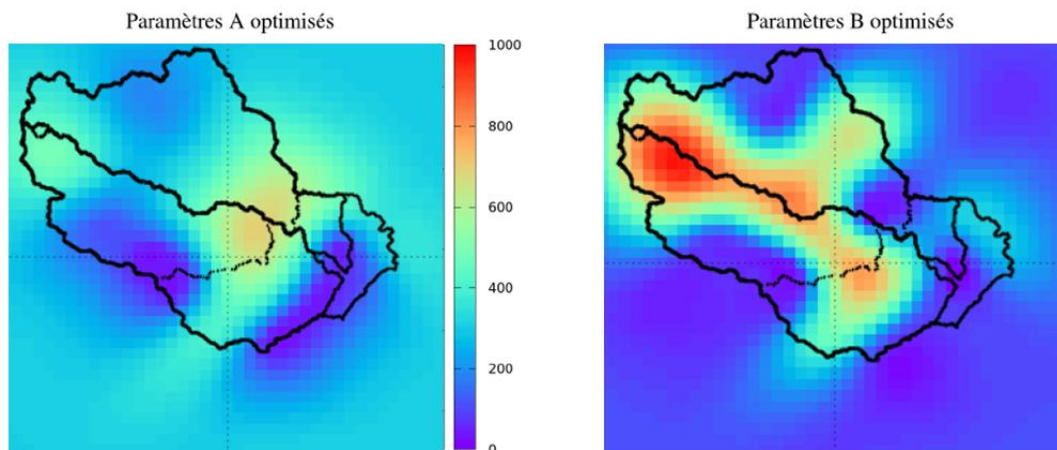


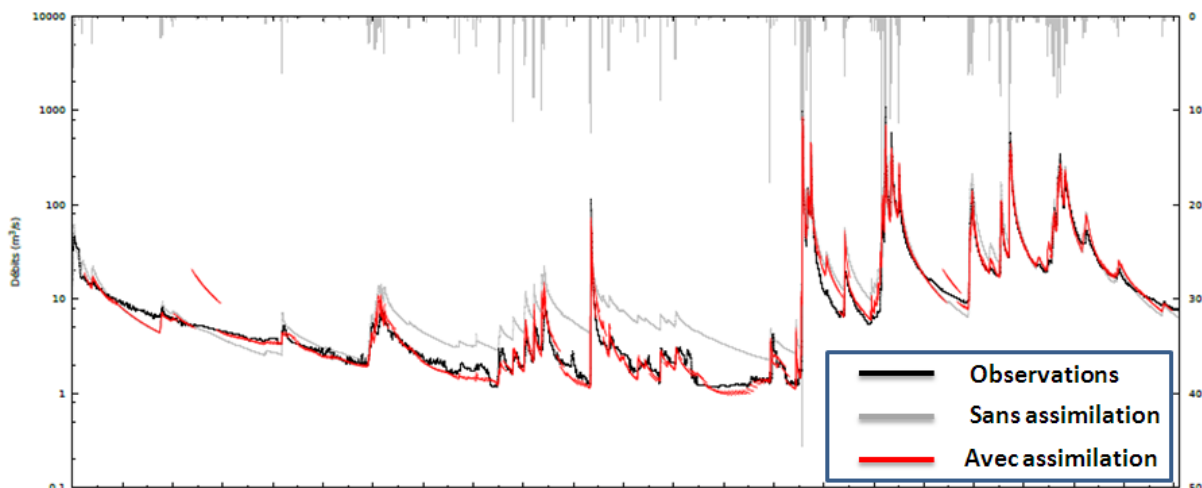
Fig 3. Essai de calage du modèle GRDc par 4D-Var sur les Gardons d'Anduze (Jay-Allemand et al., 2016)

La comparaison et l'appréciation des performances des différentes procédures de calage pourront être évaluées sur les jeux de bassins versants emboîtés où le calage sera effectué sur les exutoires aval et le contrôle effectué sur les exutoires amont. On récupèrera aussi des données de validation sur les petits bassins versants du réseau RBV (strengbach, Orgeval, Yzeron, Draix, OHM CV,...) non présents dans la banque Hydro.

- **Procédures d'assimilation pour la prévision**

L'objectif est d'améliorer la prévision des débits par la transposition des erreurs de modélisation observées sur les bassins versants jaugés. Ce travail consiste dans un premier temps à étudier la correction des erreurs des bassins jaugés, puis étudier le moyen de transférer cette correction sur les bassins jaugés. La faisabilité de ce travail a également été testée en 2016 (Fig 4, Jay-Allemand *et al*, 2016). Dans ces tests une assimilation des données est réalisée pour déterminer les paramètres initiaux, en utilisant les observations sur les dernières 75h, puis les prévisions sont faites avec le modèle direct pour les 48h heures futures. Ce calcul est réalisé périodiquement sur une fenêtre glissante toutes les 24h. La Fig 4 montre les améliorations réalisées avec Assimilation de Données (période 01/03/2014 à 01/01/2015. Le même travail a été réalisé pour la période 01/07/2013 à 01/03/2014 avec des résultats similaires). On observe des décrochages pour certaines périodes indiquant probablement un problème de convergence vers un minimum local, exemple des problèmes à résoudre.

- Correction des données sur les bassins jaugés : différentes méthodes de correction des erreurs seront testées, de la plus simple (décalage des débits) à la plus complexe (recalage des paramètres et/ou états au cours de la crue). Les outils d'assimilation étudiés dans la première partie seront testés. Les travaux effectués sur les modèles globaux peuvent aussi être testés.
- Transfert des corrections en sites non-jaugés : les méthodes de régionalisation testées lors du calage du modèle seront appliquées pour la transposition des erreurs. L'utilisation de l'outil 4D-Var pourrait permettre de fournir une correction des paramètres directement utilisable en site non-jaugés (dans le cas des sous-bassins).



*Fig 4. Essai d'assimilation des débits par 4D-Var : hydrogramme prévu à 48 heures avec pluies futures connues (Jay-Allemand *et al.*, 2016)*

## Résultats attendus

Les résultats attendus de ce travail de thèse sont importants et concernent à la fois l'amélioration des méthodes d'assimilation et leurs adaptations aux problématiques de modélisation hydrologique, et à la fois l'amélioration des performances des modèles de prévision hydrologiques en contexte non jaugés.

- **Calage d'un modèle hydrologique distribué, sous contrainte « physique »**
- **Modélisation performante sur les petits bassins versants**
- **Estimation des erreurs sur les entrées (pluies), les sorties (débits) et le modèle lui-même, par des techniques de réanalyse.**
- **Méthodologie pour le choix des matrices de covariance des erreurs de modèle et d'observation adaptées à ce contexte hydrologique en bassins versants non-jaugés, et de manière plus générale le choix de l'estimateur a posteriori en présence d'erreurs non Gaussiennes**
- **Amélioration des prévisions : justesse et anticipation**
- **Application pour l'optimisation du réseau de mesure.**

Deux articles sont prévus, axés chacun sur une application des méthodes d'assimilation adaptées aux deux problématiques posées : l'un sur l'apport des méthodes d'assimilation variationnelle pour un « calage régional » du modèle hydrologique distribué, l'autre sur l'assimilation de données pour la mise à jour de modèle de prévision en sites non-jaugés. Ces articles pourront être soumis à des journaux internationaux comme Journal of Hydrology ou Hydrological processes.

## **Organisation de la thèse**

---

- **Collaborations**

La thèse sera co-encadrée par **Igor Gejadze** (UMR G-EAU, HDR prévue courant 2017, Directeur de thèse) pour les aspects assimilation de données et **Patrick Arnaud** (Irstea, UR RECOVER, HDR, co-Directeur de thèse) pour les aspects régionalisation et descente d'échelle, avec le soutien de certains scientifiques de leur équipe (**Pierre Javelle** pour les aspects modélisation distribuée de la méthode AIGA, **Pierre-Olivier Malaterre** pour des aspects hydrauliques (routage) et assimilation de données). Une collaboration étroite avec le SCHAPI, METEO-France et le bureau d'étude HYDRIS est déjà active en lien avec les développements et la mise en œuvre opérationnelle de la méthode AIGA (fourniture et actualisation des données, transferts des résultats...).

Le comité de pilotage doit être pluridisciplinaire pour ce sujet à l'intersection des mathématiques appliquées et de l'hydrologie.

Proposition de composition : **Bruno Janet** (Schapi, besoins opérationnels) – **Isabelle Braud** (Irstea Lyon, approches distribuées) – **Ludovic Oudin** (UMR Sisyphe, Univ P & M Curie, régionalisation et descente d'échelle) – **Guillaume Thirel** (Irstea Antony, assimilation de données et hydrologie) – **Sophie Ricci** (Cerfacs, Toulouse, incertitudes et assimilation de données).

- **Planning**

Un planning prévisionnel peut être avancé en fonction des travaux envisagés. La structuration de ce planning s'appuie sur les compétences et connaissances disponibles au sein du collectif d'encadrement, sur les travaux antérieurs déjà réalisés et sur les données déjà disponibles.

	2017 S2	2018 S1	2018 S2	2019 S1	2019 S2	2020 S1
Analyse bibliographique et préparation des jeux de données	x	x				
« Calage régional »						
- Mise en place	x	x				
- Tests / variantes		x	x			
- Validation / apports			x			
« Mise à jour des débits »						
- Mise en place			x	x		
- Tests / variantes				x	x	
- Validation / apports					x	
Réaction d'articles			x		x	
Réaction de la thèse						x

*Table 1. Planning*

- **Accueil**

Le doctorant sera accueilli par l'équipe RHAX de l'UR-RECOVER à Irstea Centre d'Aix-en-Provence, pour orienter le doctorant sur les questions liées à la problématique de la modélisation hydrologique distribuée. Des séjours longs au centre de Montpellier seront prévus lors de la prise en main des logiciels d'assimilation de données (4D-Var) avec un encadrement direct par Igor Gejadze et Pierre-Olivier Malaterre. Le candidat sera inscrit à l'ED GAIA à Montpellier. Le fait que le doctorant soit basé à Aix-en-Provence permettra à cette équipe de réellement prendre en main ces techniques d'Assimilation de Données Variationnelle et les outils afférents pour continuer à faire évoluer ces méthodes au-delà de la thèse et de potentiellement les appliquer à d'autres modèles. L'encadrement par Igor Gejadze et l'inscription à l'ED GAIA à Montpellier favorisera une collaboration très étroite entre les 2 équipes.

Le candidat devra être a priori de formation hydrologique – hydraulique (Master, Ecole d'Ingénieur), familiarisé avec les méthodes numériques d'assimilation de données si possible, ou au minimum une bonne formation de base en mathématique et en programmation. Les techniques d'Assimilation de Données sont de plus en plus enseignées et connues dans le domaine de l'hydrologie. Certains prévisionnistes des SPC réalisent même des thèses parfois sur ce sujet (Habert 2016). Par ailleurs, les encadrants et Igor Gejadze en particulier, maîtrisent très bien ces techniques et une bonne base du travail de codage a déjà été réalisée en 2016. Le travail du doctorant sur ce point sera de les faire évoluer en fonction des réflexions autour des incertitudes, des descentes d'échelle, des critères à minimiser, des choix de régularisation, etc.

## Références

---

- Andréassian, V. et al., 2004. Impact of spatial aggregation of inputs and parameters on the efficiency of rainfall-runoff models: A theoretical study using chimera watersheds. *Water Resour. Res.*, 40(5): W052091-W052099.
- Arnaud, P., Lavabre, J., Fouchier, C., Diss, S., Javelle, P., 2011. Sensitivity of hydrological models to uncertainty in rainfall input. *Hydrolog Sci J*, 56(3): 397-410. DOI:10.1080/02626667.2011.563742
- Bessière H. 2008, PhD. Assimilation de données variationnelle pour la modélisation hydrologique distribuée des crues à cinétique rapide. Univ Paul Sabatier, Toulouse, INP.
- Bardsley J.M., Vogel C.R., 2003. A nonnegatively constrained convex programming method for image reconstruction. *SIAM J. Sci. Comput.* Vol 25, No. 4, pp. 1326-1343
- Berthet, L., 2010. Préviation des crues au pas de temps horaire : pour une meilleure assimilation de l'information de débit dans un modèle hydrologique. Cemagref Thesis, AgroParisTech, 603 pp.
- Beven, K., 1993. Prophecy, reality and uncertainty in distributed hydrological modelling. *Advanced Water Resources*, 16: 41-51.
- Emery, C. Contribution de la future mission altimétrique large fauchée SWOT pour la modéliation hydrologique à grande échelle. PhD. ED SDU2E, Université Paul Sabatier Toulouse 3. 3 Février 2017.
- Chaubey, I., Haan, C.T., Grunwald, S., Salisbury, J.M., 1999. Uncertainty in the model parameters due to spatial variability of rainfall. *Journal of Hydrology*, 220(1-2): 48-61.
- Gejadze I., Malaterre P.-O., 2016. Discharge estimation under uncertainty using variational methods with application to the full Saint-Venant hydraulic network model. *International Journal for Numerical Methods in Fluids*. Version of Record online: 28 JUL 2016 | DOI: 10.1002/flid.4273.
- Gejadze I., Malaterre P.-O., 2016. Design of the control set in the framework of variational data assimilation. *Journal of Computational Physics*. August 2016. DOI: 10.1016/j.jcp.2016.08.029.
- Gejadze I., Shutyaev V., Vidard A., Le Dimet F.-X.. Optimal solution error quantification in variational data assimilation involving imperfect models: Optimal solution error quantification in variational DA. *International Journal for Numerical Methods in Fluids* · January 2016
- Habert J., 2016. Préviation des crues en temps réel sur le bassin de la Marne : assimilation in situ pour la correction du modèle hydraulique mono-dimensionnel Mascaret. PhD, Hydrologie, Hydrochimie, Sols, Environnement, Institut National Polytechnique de Toulouse, 2016
- Huot, P.-L., 2014. Évaluation de méthodes d'optimisation pour le calage efficace de modèles hydrologiques coûteux en temps de calcul. Mémoire de maîtrise électronique, Montréal, École de technologie supérieure.
- Javelle, P., Fouchier, C., Arnaud, P., Lavabre, J., 2010. Flash flood warning at ungauged locations using radar rainfall and antecedent soil moisture estimations. *Journal of Hydrology*, 394(1-2): 267-274. DOI:DOI 10.1016/j.jhydrol.2010.03.032
- Jay-Allemand, M., Organde, D., Fine, J.-A., 2016. Implémentation de l'algorithme 4DVAR dans un modèle hydrologique conceptuel, continu et distribué (GRDc). Application à la préviation des crues soudaines. Rapport d'étude Hydris: 60 pages.
- Lavabre, J., Gregoris, Y., 2005. AIGA : un dispositif d'alerte des crues sur l'ensemble du réseau hydrographique. In *Ingénieries*(n° 44): 3-12.

- Malaterre P.-O., H. Oubanas, I. Gejadze. 2016. Congrès français d'assimilation de données. Conférence plénière sur la quantification d'incertitude et l'assimilation de données appliquées à l'hydraulique à surface libre. Grenoble du 30 novembre au 2 décembre 2016.
- Malaterre P.-O., P.-A. Garambois, J. Monnier, S. Ricci, H. Roux. 2016. Modèles 1D/2D avec algorithmes d'assimilation : avancées, inter-comparaisons. Journées Nationales SWOT. 22-24 November 2016. Toulouse, France.
- Monnier J., P.-O. Malaterre, P.-A. Garambois, S. Biancamaria, S. Ricci, H. Roux, M. Gosset. 2016. SWOT river discharge algorithms. 2016 Surface Water and Ocean Topography (SWOT) Science Team Meeting. 13-16 June 2016. Pasadena, California, USA.
- Oubanas H., I. Gejadze, Malaterre P.-O., M. Durand (Univ Ohio). 2016. AGU San Francisco. "Simultaneous estimation of inflow discharge, river bathymetry and friction from synthetic SWOT data using variational data assimilation"
- Oubanas H., I. Gejadze, P.-O. Malaterre, F. Mercier. 2015. Estimation of river discharge from in-situ and remote sensing data, using variational data assimilation and a full Saint-Venant hydraulic model. 3rd Space for Hydrology Workshop, Frascati (Rome), Italy, September 15~17, 2015
- Rabier, F., Mahfouf, J.-F., Klinker, E., 2000. Une nouvelle technique d'assimilation des données d'observation au CEPMMT : l'assimilation variationnelle quadridimensionnelle. *La Météorologie*, 8e serie(30): 87-101.
- Randrianasolo, A., Ramos, M.H., Andréassian, V., 2011. Hydrological ensemble forecasting at ungauged basins: Using neighbour catchments for model setup and updating. *Advances in Geosciences*, 29: 1-11. DOI:10.5194/adgeo-29-1-2011
- Roux, H. et al., 2011. A physically-based parsimonious hydrological model for flash floods in Mediterranean catchments. *Nat. Hazards Earth Syst. Sci.*, 11(9): 2567-2582. DOI:10.5194/nhess-11-2567-2011
- Seo, D., V. Koren, N. Cajina. 2003. Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting. *Journal of Hydrometeorology* 4:627-641
- Thirel G., E. Martin, J.-F. Mahfouf, S. Massart, S. Ricci, F. Habets: A past discharges assimilation system for ensemble streamflow forecasts over France – Part 1: Description and validation of the assimilation system, *Hydrol. Earth Syst. Sci.*, 14, 1623-1637, doi:10.5194/hess-14-1623-2010, 2010.
- Thirel G., E. Martin, J.-F. Mahfouf, S. Massart, S. Ricci, F. Regimbeau, F. Habets: A past discharge assimilation system for ensemble streamflow forecasts over France – Part 2: Impact on the ensemble streamflow forecasts, *Hydrol. Earth Syst. Sci.*, 14, 1639-1653, doi:10.5194/hess-14-1639-2010, 2010.
- Zoccatelli, D., Borga, M., Zanon, F., Antonescu, B., Stancalie, G., 2010. Which rainfall spatial information for flash flood response modelling? A numerical investigation based on data from the Carpathian range, Romania. *Journal of Hydrology*, 394(1-2): 148-161. DOI:<http://dx.doi.org/10.1016/j.jhydrol.2010.07.019>