

# Applications du CERFACS en assimilation de données

J. Munoz Sabater <sup>1</sup>, B. Bouriquet <sup>1</sup>, S. Ricci <sup>1</sup>, J.-Ph. Argaud <sup>2,1</sup>,  
A. Piacentini <sup>1</sup>, S. Massart <sup>1</sup>, A. Weaver <sup>1</sup> et O. Thual <sup>3,1</sup>

<sup>1</sup>Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique  
42, avenue Coriolis, 31057 Toulouse cedex – France  
[thual@cerfacs.fr](mailto:thual@cerfacs.fr)

<sup>2</sup>Electricité de France, Recherche & Développement,  
1, avenue du Général de Gaulle, 92141 Clamart cedex - France

<sup>3</sup>Institut National Polytechnique, Université de Toulouse  
IMFT, allée du Professeur Camille Soula, 31400 Toulouse - France

## Résumé

L'assimilation de données est une pratique qui devient indispensable pour la modélisation d'un nombre croissant d'applications. Nous présentons brièvement trois exemples de développement de chaînes d'assimilation réalisées au CERFACS dans des domaines différents : océanographie, chimie atmosphérique et neutronique.

Nous présentons ensuite l'application des concepts de l'assimilation de données à la prévision des crues rapides sur le bassin versant des Gardons d'Anduze à l'aide d'un modèle pluie-débit de la plateforme d'hydrologie ATHYS forcé par des pluies mesurées. Les variables contrôlées sont le paramètre  $S$  du modèle SCS, qui représente une capacité de rétention potentielle maximale des sols en eau, et le paramètre de transfert  $V_0$ , qui représente la vitesse de maille. Les observations sont constituées de mesures de débit à l'exutoire.

Les premiers tests effectués montrent que le choix des matrices des covariances d'erreurs d'ébauche  $\mathbf{B}$  et d'observation  $\mathbf{R}$  est déterminant pour l'ajustement de deux paramètres pertinents. Les scores de prévision des pics de crues obtenus à l'aide de cette approche sont améliorés par rapport à des approches plus classiques dans ce domaine.

## Mots clés

Assimilation de données, Hydrologie, Prévision des crues

## 1. Introduction

L'assimilation de données est devenue une composante intrinsèque de la démarche de modélisation, dès lors que l'on est en présence d'observations issues d'un système réel que l'on cherche à décrire ou prévoir. Nous présentons tout d'abord les notations qui seront utilisées pour décrire les méthodes choisies pour les différentes applications que nous allons décrire (Lagarde et al., 2001).

### 1.1 Notations génériques de l'assimilation de données

La plupart des méthodes d'assimilation de données utilisées pour des applications réalistes se laissent décrire par la minimisation de la fonction coût  $J$  suivante :

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}[\mathbf{y}^o - G(\mathbf{x})]^T \mathbf{R}^{-1}[\mathbf{y}^o - G(\mathbf{x})], \quad (1)$$

où  $\mathbf{x}$  est le vecteur de contrôle dont on cherche la valeur optimale,  $\mathbf{x}^b$  l'ébauche,  $G$  l'opérateur d'observation généralisé,  $\mathbf{y}^o$  le vecteur d'observation,  $\mathbf{B}$  la matrice des covariances d'erreur d'ébauche et  $\mathbf{R}$  la matrice des covariances d'erreur d'observation. On note  $\mathbf{x}^a$  l'analyse qui minimise la fonction coût.

Le vecteur de contrôle peut contenir l'état d'un modèle de simulation numérique, ses paramètres de calages ou les deux à la fois. Le vecteur d'observation contient toutes les mesures du système réel que l'on souhaite assimiler. L'opérateur d'observation permet de simuler ces mesures dans le modèle.

Les inverses  $\mathbf{B}^{-1}$  et  $\mathbf{R}^{-1}$  peuvent être vues comme des matrices de « certitude » qui pondèrent les différentes informations, ébauche ou observations, en fonction du degré de confiance qu'on leur attribue. Enfin, l'analyse est la meilleure estimation du vecteur de contrôle produite par la méthode d'assimilation choisie.

Dans certaines applications, la minimisation de  $J$  peut s'avérer fastidieuse et l'on approxime la méthode d'assimilation par sa version incrémentale qui consiste à minimiser

$$J_{\text{inc}}(\mathbf{x}) = J_{\text{inc}}(\mathbf{x}^b + \delta \mathbf{x}) = \frac{1}{2} \delta \mathbf{x}^T \mathbf{B}^{-1} \delta \mathbf{x} + \frac{1}{2}(\mathbf{d} - \mathbf{G} \delta \mathbf{x})^T \mathbf{R}^{-1}(\mathbf{d} - \mathbf{G} \delta \mathbf{x}), \quad (2)$$

où  $\mathbf{G}$  est l'opérateur d'observation  $G$  linéarisé autour de l'ébauche  $\mathbf{x}^b$  et  $\mathbf{d}$  le vecteur innovation défini par  $\mathbf{d} = \mathbf{y}^o - G \mathbf{x}^b$ .

La minimisation de cette fonction quadratique conduit à l'analyse

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{K}[\mathbf{y}^o - G(\mathbf{x}^b)] \quad \text{avec} \quad \mathbf{K} = (\mathbf{B}^{-1} + \mathbf{G}^T \mathbf{R}^{-1} \mathbf{G})^{-1} \mathbf{G}^T \mathbf{R}^{-1} = \mathbf{B} \mathbf{G}^T (\mathbf{G} \mathbf{B} \mathbf{G}^T + \mathbf{R})^{-1} \quad (3)$$

où  $\mathbf{K}$  est la matrice de gain qui relie donc l'innovation à l'incrément d'analyse par la relation  $\delta \mathbf{x}^a = \mathbf{K} \mathbf{d}$ . Les deux expressions analytiques de  $\mathbf{K}$  en fonction des matrices  $\mathbf{B}$ ,  $\mathbf{R}$  et  $\mathbf{G}$  peuvent être utilisées indifféremment.

Cette matrice permet de déterminer la matrice des covariances d'erreur d'analyse  $\mathbf{A} = (\mathbf{I} - \mathbf{K} \mathbf{G}) \mathbf{B}$  qui peut être utilisée comme une nouvelle matrice des covariances d'erreur d'ébauche lorsque l'analyse  $\mathbf{x}^a$  est utilisée comme nouvelle ébauche pour améliorer l'approximation de  $J$  par  $J_{\text{inc}}$ .

## 1.2 Contrôle de l'état initial d'un modèle

Lorsque le vecteur de contrôle est l'état initial d'un modèle dynamique  $M_{0 \rightarrow i}$  qui associe à un état initial  $\mathbf{x}$  l'état  $M_{0 \rightarrow i}(\mathbf{x})$  du modèle au temps  $i$ , on écrit la fonction coût sous la forme

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2} \sum_{i=0}^N [\mathbf{y}_i^o - H_i M_{0 \rightarrow i}(\mathbf{x})]^T \mathbf{R}_i^{-1} [\mathbf{y}_i^o - H_i M_{0 \rightarrow i}(\mathbf{x})], \quad (4)$$

où  $H_i$  est l'opérateur d'observation au temps  $i$  et  $\mathbf{R}_i$  la matrice des covariances d'erreur d'observation au temps  $i$  (on suppose donc que les erreurs des mesures effectuées à des temps différents ne sont pas corrélées). Le gradient  $\nabla J$  de la fonction coût  $J$  de cette méthode d'assimilation, que l'on nomme « 4D-Var », s'écrit

$$\nabla J(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + \sum_{i=0}^N \mathbf{M}_{0 \rightarrow i}^T \mathbf{H}_i^T \mathbf{R}_i^{-1} [\mathbf{y}_i^o - H_i M_{0 \rightarrow i}(\mathbf{x})], \quad (5)$$

où les matrices  $\mathbf{M}_{0 \rightarrow i}$  et  $\mathbf{H}_i$  sont obtenues en linéarisant les opérateurs  $M_{0 \rightarrow i}$  et  $H_i$  autour de la trajectoire issue de  $\mathbf{x}$ . Lorsque qu'il n'est pas possible de stocker les matrices  $\mathbf{M}_{0 \rightarrow i}^T$ , on peut calculer leur action sur un vecteur en considérant ce dernier comme la condition initiale au temps  $i$  du modèle adjoint qui intègre en remontant le temps les opérateurs différentiels transposés du modèle tangent. De plus, on utilise souvent la version incrémentale de cette méthode.

Le développement de ce modèle adjoint est souvent un exercice fastidieux et de nombreux efforts informatiques sont effectués pour le générer automatiquement à partir du code du modèle direct. Lorsque les matrices  $\mathbf{M}_{0 \rightarrow i}^T$  ne peuvent ni être stockées ni être calculées, on les remplace par l'opérateur identité  $\mathbf{Id}$  dans l'expression du gradient  $\nabla J_{\text{inc}}$  et l'on baptise alors « FGAT » (First Guess at Appropriate Time) l'approximation ainsi obtenue de la méthode d'assimilation.

## 1.3 Plan du présent document

Nous allons maintenant utiliser ce formalisme pour décrire quatre exemples d'applications. Dans le paragraphe 2, nous décrivons les grandes lignes des méthodes d'assimilation de données développées pour des exemples issus de trois domaines. Dans le paragraphe 3, nous décrivons en détail une expérience d'assimilation de données réalisées dans le cadre de la prévision des crues.

# 2. **Océanographie, chimie atmosphérique et neutronique**

Le CERFACS est spécialisé dans le développement d'outils et d'algorithmes au service de plusieurs communautés scientifiques. L'océanographie opérationnelle est l'une des premières applications dans laquelle nous nous sommes lancés avec le développement du coupler PALM pour le projet MERCATOR Océan. Depuis, le CERFACS a développé son expertise en assimilation de données sur des applications diverses dont les trois plus importantes sont décrites dans ce paragraphe.

## 2.1 Développement de NEMOVAR

Le système NEMOVAR repose sur le développement de méthodes d'assimilation de données variationnelles incrémentales pour le modèle d'océan NEMO. Le vecteur de contrôle est constitué des champs de température, de salinité, de vitesses horizontales et de hauteurs de surface de la mer sur une grille 3D de plus d'un million de points. Les observations utilisées (plusieurs milliers) dans

ce système sont les profils de température et de salinité mesurées par des bouées plongantes, les anomalies de hauteur de surface de la mer mesurées ou des températures de surface de la mer mesurées par des satellites.

La version actuelle de cette chaîne d'assimilation est un 3D-VAR FGAT incrémental (Waeber et al., 2005 ; Daget et al., 2008) dont la fonction coût  $J_{inc}$  a été décrite dans l'introduction. La méthode 4D-VAR incrémentale, qui nécessite énormément de temps de calcul, est en cours de développement. Elle nécessite l'intégration du modèle adjoint  $\mathbf{M}_{0 \rightarrow i}^T$  qui est approximé par l'adjoint d'une version dégradée du modèle direct.

Un des points clés du développement de ce système repose sur la modélisation de la matrice des covariances d'erreur d'ébauche  $\mathbf{B}$ . Il faut tout d'abord prendre en compte des corrélations entre les différents champs de l'espace de contrôle. On parle alors d'assimilation de données multivariée. Par exemple, la température et la salinité ne varient pas indépendamment au sein d'une même masse d'eau. De même, l'équilibre géostrophique, qui résulte de la rotation rapide de la terre, induit une liaison entre les courants et la hauteur de la surface libre. La matrice  $\mathbf{B}$  doit prendre en compte les corrélations d'erreurs qui en résultent.

Comme il n'est pas possible de stocker la matrice  $\mathbf{B}$ , une méthode performante consiste à modéliser son action sur un vecteur par un opérateur de diffusion que l'on approxime par une méthode implicite. On peut alors considérer cette diffusion comme étant anisotrope pour privilégier les corrélations entre champs le long des surfaces isopycnales (iso-densité), comme c'est le cas pour les champs observés. Enfin, la minimisation de la fonction coût nécessite le choix d'algorithmes performants et surtout de pré-conditionneurs pour atteindre des coûts de calculs abordables.

La chaîne d'assimilation de données NEMOVAR sert à l'initialisation des modèles couplés pour la prévision saisonnière et décennale ainsi que pour la production de ré-analyses océaniques pour diverses études scientifiques.

## 2.2 Assimilation des espèces chimiques de l'atmosphère

L'assimilation variationnelle (3D-FGAT) des espèces chimiques de l'atmosphère mesurées par des instruments embarqués à bord de satellites de recherche, constitue la seconde application majeure où le CERFACS est présent en mode recherche et développement (Massart et al., 2005). Le système MOCAGE-PALM est basé sur le modèle de chimie transport MOCAGE du Centre National de Recherches Météorologiques (CNRM) de Météo-France. L'utilisation du coupleur PALM du CERFACS pour développer l'algorithme d'assimilation a été un choix primordial pour conserver la modularité du modèle (plusieurs maillages, différentes paramétrisations physiques, ...). Cette modularité permet d'avoir un outil commun au service de toute une communauté.

Les deux principaux thèmes de recherche de cette activité sont la valorisation des données spatiales et la modélisation de la matrice de covariance d'erreur d'ébauche  $\mathbf{B}$ . Dans le premier cas, il s'agit d'utiliser l'assimilation soit comme une référence à laquelle comparer de nouvelles données (Massart et al., 2007b) soit comme une référence pour étudier des phénomènes physiques particuliers (perte d'ozone dans le vortex polaire, échange entre les différentes couches de l'atmosphère, ...). Dans le second cas, les travaux se font en étroite collaboration avec ceux menés en océanographie au CERFACS en particulier autour de l'équation de diffusion (Pannekoucke et Massart, 2008).

### 2.3 Modélisation des coeurs de centrales nucléaires

L'assimilation de données n'est pas cantonnée aux applications géophysiques. En étroite collaboration avec EDF, nous avons montré ses potentialités pour la modélisation des coeurs de centrales nucléaires, qu'il s'agisse de l'estimation de champs neutroniques ou encore du calage de paramètres des modèles.

Nous avons développé une méthode de type 3D-VAR incrémentale, qui revient à minimiser la fonction coût  $J_{inc}$  de l'introduction dans le cas  $N=0$ , pour estimer la nappe de flux neutronique d'un coeur de centrale nucléaire à partir de mesures de capteurs neutroniques et thermohydrauliques disposés en plusieurs points de manière permanente ou épisodique. La construction de la matrice  $\mathbf{B}$  est primordiale pour propager correctement l'information des mesures vers les parties non instrumentées du coeur (Massart et al, 2007a).

Parallèlement, nous avons mis au point une méthode permettant le calage de certains paramètres du modèle de coeur COCCINELLE à partir de données issues d'un grand nombre de campagnes, c'est-à-dire de périodes d'exploitation d'environ 18 mois de centrales entre deux rechargement du combustible. Nos travaux ont porté principalement sur les paramètres relatifs au réflecteur du réacteur qui est la partie entourant la zone active à l'intérieur du coeur.

Enfin une méthode 4D-VAR a été développée sur un modèle simplifié unidimensionnel afin de contrôler les champs de concentration de xénon et d'iode qui sont responsables de la dynamique intrinsèque du coeur à l'échelle de l'heure ou de la journée (Ponçot, 2008). Ces travaux nous ont permis de proposer des améliorations pour les méthodes de pilotage des centrales.

L'ensemble de ces travaux permet de systématiser les comparaisons entre calculs et mesures, qui sont cruciales pour la qualité des simulations et du fonctionnement, d'améliorer les modèles et leur paramétrisation, et de prendre en compte de manière homogène l'ensemble des mesures disponibles.

## **3. L'assimilation de données pour la prévision des crues**

En collaboration avec le Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations (SCHAPI), le CERFACS a entamé une démarche visant à explorer les apports d'une approche de type assimilation de données pour la prévision des crues. Le premier exemple choisi pour cet exercice est la plateforme logicielle ATHYS développée à HydroSciences Montpellier (HSM). Nous présentons ici les premières conclusions de cet exercice.

### 3.1 La plateforme ATHYS et le jeu de données

Le travail présenté ici s'inscrit dans le prolongement de la thèse de doctorat de A. Marchandise (Marchandise, 2007) qui s'est plus particulièrement intéressé au Bassin du Gardon à Anduze (525 km<sup>2</sup>, voir figure 1) et au modèle MERCEDES de la plateforme logicielle ATHYS (Bouvier et al., 1996 ; Bouvier et al., 2007). Il s'agit d'un modèle pluie-débit distribué parcimonieux événementiel dans une de ses versions qui combine une fonction de production de type SCS et une fonction de transfert de type translation-stockage. Ce modèle et son application à ce bassin sont décrits dans (Bouvier et al., 2008).

La fonction de production SCS est paramétrée par trois paramètres dont le coefficient  $S$  qui modélise la capacité de stockage en eau des sols. Les deux autres coefficients sont l'intensité de vidange du stock  $ds$  et la fraction  $w$  de la vidange qui rejoint l'exutoire du bassin. Cette vidange permet de simuler le ressuyage des sols entre deux averses.

Le ruissellement produit par la maille est ensuite acheminé à l'exutoire du bassin à l'aide d'un modèle de translation-stockage caractérisé par une vitesse de transfert  $V_0$  constante pour toutes les mailles et un coefficient de diffusion  $K_0$ .

Le modèle est forcé par les pluies horaires enregistrées par 6 postes pluviométriques répartis sur le bassin. Les simulations du modèle sont réalisées au pas de temps horaire. L'initialisation de ce modèle pluie-débit événementiel doit être effectuée au début de chacun des épisodes pluvieux simulés. On compare alors le débit simulé au débit mesuré à l'exutoire du bassin (à Anduze).

Si les paramètres  $ds$ ,  $w$  et  $K_0$  peuvent être fixés à des valeurs constantes quelque soit l'événement, les valeurs de  $S$  et  $V_0$  sont soumises à une variabilité inter-événementielle dont les sources principales sont les incertitudes sur la connaissance de la pluie et la connaissance de l'état hydrique initial du bassin versant (Marchandise, 2007). Les stratégies pour déterminer les valeurs optimales de ces deux paramètres résident dans deux directions principales :

- une prédétermination empirique, dans la mesure où on arrive à relier ces paramètres à un indicateur d'humidité initiale du bassin (Bouvier et al., 2008 ; Marchandise et Viel, 2008),
- une réajustement par assimilation de données sur la base des observations de débits, soit sur la durée totale de l'épisode pluvieux (temps différé), soit sur le début de la crue (temps réel).

Nous explorons ici cette dernière stratégie et nous présentons le gain du développement d'une procédure d'assimilation de données en mettant l'accent sur la prévision en temps réel.

### 3.2 Mise en oeuvre de la méthode d'assimilation de données

Les crues du bassin du Gardon d'Anduze surviennent essentiellement en automne lors d'épisodes pluvieux pouvant atteindre plusieurs centaines de mm en 24 heures avec des temps de réponse de l'ordre de 3 à 5 heures. Pour modéliser et prévoir ces crues, l'ajustement des paramètres  $S$  et  $V_0$  doit être envisagé pour chaque événement pluvieux.

On développe ici une méthode d'assimilation de données dans laquelle le vecteur de contrôle est  $\mathbf{x}=(S, V_0)^T$  et le vecteur d'observations  $\mathbf{y}^o=(y_1^o, y_2^o, \dots, y_N^o)^T$  est constitué de mesures de débits  $y_i^o$  effectués à des temps  $i$  quelconques. On note  $G_i(\mathbf{x})$  le débit simulé par le modèle au temps  $i$ , l'opérateur d'observation généralisé  $G_i$  associant un débit à un jeu de paramètres.

La méthode d'ajustement des paramètres utilisée dans (Marchandise, 2007) et (Bouvier et al., 2008) repose sur la minimisation du critère de Nash

$$\text{Nash} = 1 - \frac{\sum_{i=0}^N [y_i^o - G_i(\mathbf{x})]^2}{\sum_{i=0}^N (y_i^o - \bar{y}^o)^2}, \quad \text{avec} \quad \bar{y}^o = \frac{1}{N} \sum_{i=0}^N y_i^o, \quad (6)$$

où seuls les  $N$  débits mesurés supérieurs à  $50 \text{ m}^3/\text{s}$  sont considérés. Ce calage peut être vu comme une méthode d'assimilation moindre carrés sans ébauche pour laquelle la matrice des covariances d'erreurs d'observation  $\mathbf{R}$  est l'identité.

Si cette méthode permet de bien estimer  $S$  et  $V_0$  pour un épisode pluvieux donné, ou de bâtir un atlas de valeurs mensuelles à partir d'un grand nombre d'épisodes, elle n'est pas appropriée pour ajuster ces paramètres en mode prévision.

C'est pourquoi nous avons choisi une méthode d'assimilation comportant un vecteur d'ébauche  $\mathbf{x}^b = (S^b, V_0^b)^T$ , une matrice de covariances d'erreur d'ébauche  $\mathbf{B}$  et une matrice de covariance d'erreur d'observation  $\mathbf{R}$  dont les poids respectifs vont être déterminants pour la qualité de la prévision. Pour accélérer les temps de calcul, notre méthode est incrémentale, et nécessite donc la mise en oeuvre de l'opérateur d'observation linéarisé

$$\mathbf{G}_i = \left[ \frac{\partial G_i}{\partial S}(\mathbf{x}^b), \frac{\partial G_i}{\partial V_0}(\mathbf{x}^b) \right], \quad (7)$$

que nous calculons par différences finies en effectuant deux simulations complémentaires avec des valeurs de paramètres très légèrement perturbées.

La méthode a été implémentée à l'aide du logiciel PALM (Lagarde et al., 2001) et validée par comparaison avec le calage par minimisation du critère de Nash qui utilise la méthode du SIMPLEX. Notre approche incrémentale s'avère être beaucoup plus rapide en temps de calcul, la convergence nécessitant très peu d'itérations.

### 3.3 Apport de l'assimilation de données en mode prévision

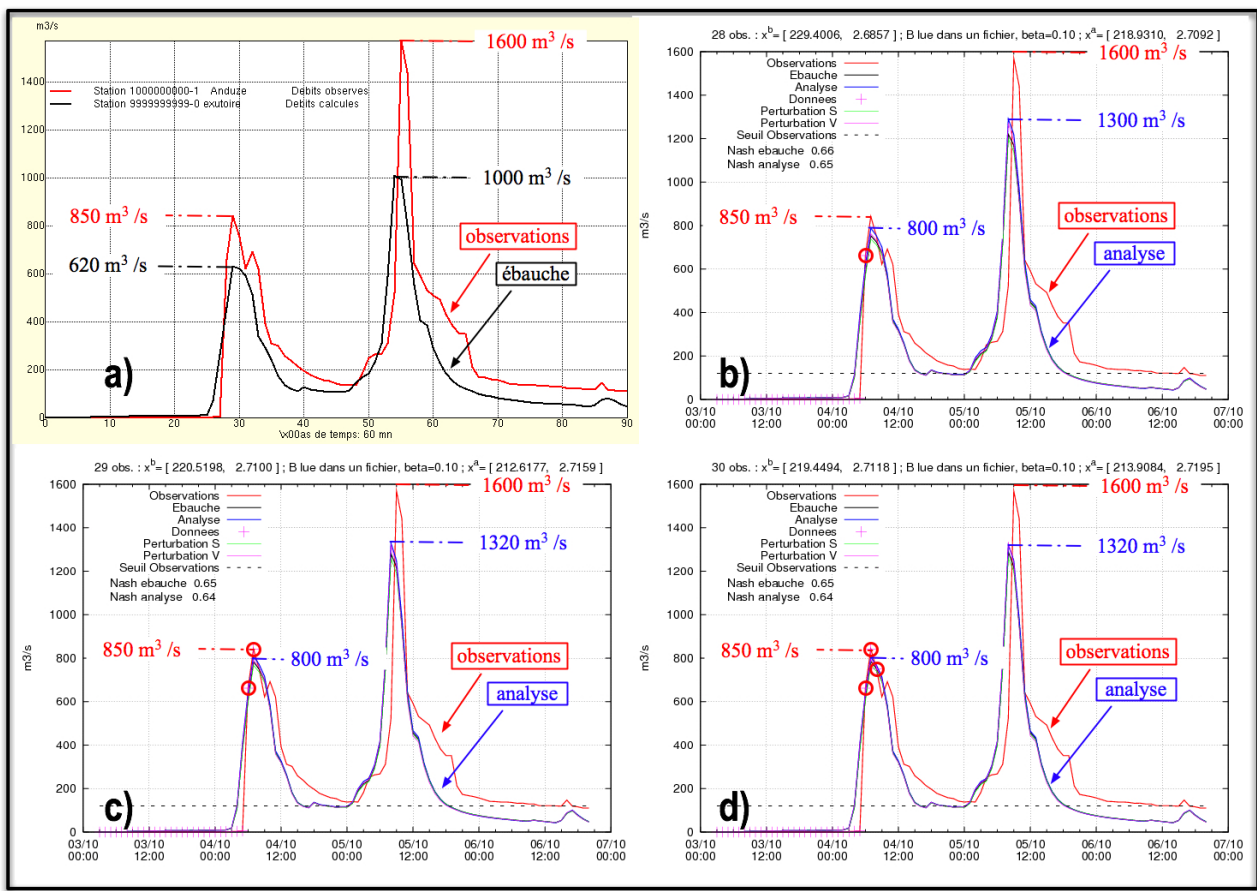
Nous avons réalisé des expériences de prévision a posteriori (temps différé) sur une dizaine d'épisodes pluvieux représentatifs des quinze dernières années. Nous avons tout d'abord choisi de ne retenir que les mesures de débits supérieurs à 120 m<sup>3</sup>/s et d'effectuer une assimilation de données suivie d'une prévision pour chaque nouvelle observation répondant à ce critère. Pour cet exercice, l'assimilation de  $N = 1, 2, 3$  observations est suffisante pour prévoir les pics de crues.

Pour un nombre d'observations  $N$  fixé, nous avons effectué plusieurs itérations de la méthode d'assimilation incrémentale en choisissant, respectivement en guise d'ébauche  $\mathbf{x}^b$  et de matrice  $\mathbf{B}$ , l'analyse  $\mathbf{x}^a$  et sa matrice de covariances d'erreur  $\mathbf{A}$  obtenues à l'étape précédente. Néanmoins, nous avons choisi de ne pas changer la matrice  $\mathbf{B}$  lorsque les incréments sont supérieurs à 33% sur  $S$  ou à 20% sur  $V_0$ .

Pour la première itération, la valeur utilisée pour  $V_0^b$  (en m/s) est de 2,6 et les valeurs pour  $S^b$  (en mm) varient selon les mois comme suit :

305 en septembre, 268 en octobre et 154 en novembre.

Ce choix correspond aux valeurs climatologiques calées par le SCHAPI sur de nombreux épisodes pluvieux. La matrice  $\mathbf{B}$  est diagonale et les variances d'erreurs d'ébauche sont prises égales à 10% de  $S^b$  et 7% de  $V_0^b$ . La matrice  $\mathbf{R}$  est diagonale et les variances d'erreurs d'observation sont prises égales à 10% de  $y_i^0$ .



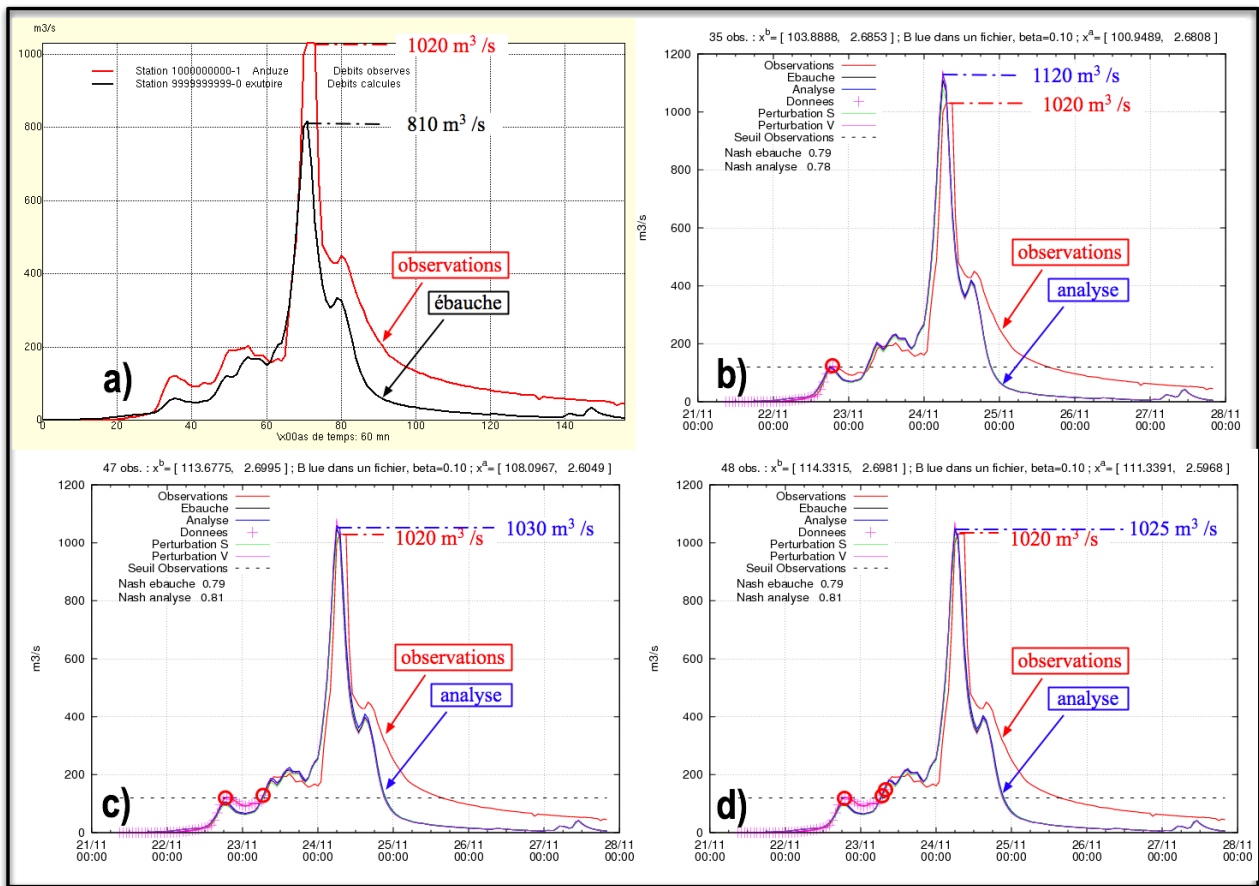
**Figure 1** : Evénement du 03/10/1995. a) Comparaison entre les débits de l'ébauche et des observations. Assimilation avec : b) une, c) deux et d) trois observations.

L'analyse de ces expériences de prévisions montrent que les valeurs climatologiques de  $S$  et  $V_0$  donnent souvent de bons résultats sans qu'il soit nécessaire d'assimiler des observations de débits. Ce résultat doit être relativisé par le fait que le modèle est forcé par des pluies mesurées alors qu'elles ne seraient que prévues dans un exercice de prévision réel.

Cependant, nous avons observé que l'assimilation de données améliorerait la prévision des pics de crues dans au moins un tiers des cas considérés. La figure 1 présente l'exemple de l'épisode pluvieux du 3 octobre 1995 avec un premier pic de crue de  $850 \text{ m}^3/\text{s}$  et un second, un jour plus tard, de  $1600 \text{ m}^3/\text{s}$  sous-estimé de 35% en l'absence d'assimilation (figure 1a). Les observations assimilées lors de la première crue permettent de mieux prévoir le second pic de crue (figure 1d) qui reste néanmoins sous-estimé de 20%.

Un deuxième exemple est illustré par la figure 2 où l'on voit que l'assimilation de données améliore significativement la prévision dès la première observation (figure 2b), soit un jour et demi avant le pic de crue. Les observations suivantes, assimilées un jour avant le pic, permettent de prévoir la crue avec une excellente précision alors que le modèle sans assimilation la sous-estime de 20%.





**Figure 2** : Evénement du 21/11/2003. Même légende que pour la figure 1.

## Conclusions

Nous avons présenté une version synthétique du formalisme de l'assimilation de données afin de décrire plusieurs exemples d'applications issus de domaines différents : océanographie, chimie atmosphérique et neutronique. Nous avons ensuite utilisé ce formalisme pour décrire la méthode d'assimilation de données que nous avons conçue pour prévoir les crues de plusieurs épisodes pluvieux en s'appuyant sur un modèle pluie-débit pertinent pour un petit bassin versant sujet à des crues éclairs fréquentes. Les résultats montrent que cette méthode permet d'améliorer significativement la prévision des pics de crue pour de nombreux épisodes pluvieux. L'extension de cette approche à l'ajustement de nouveaux paramètres ou à d'autres modèles constitue une perspective intéressante à explorer.

## Remerciements

Nous remercions A. Marchandise pour son expertise et ses suggestions lors de la réalisation de ce travail, allant jusqu'à la participation à la rédaction de cet article. Nous remercions J.-M. Tanguy et B. Janet pour leurs conseils et leur soutien dans le cadre de la collaboration entre le CERFACS et le SCHAPI.

## Bibliographie

- Bouvier Ch., Delclaux F. et Crespy A.**, (1996), *ATHYS : Atelier Hydrologique Spatialisé*, L'hydrologie tropicale : géoscience et outil pour le développement, Publication AISH 238, pp 425-435 <http://www.athys-soft.org/>
- Bouvier C., Marchandise A., Brunet P. et Crespy A.**, (2008), *Un modèle pluie-débit distribué parcimonieux pour la prédétermination et la prévision des crues éclair. Application au bassin du Gardon d'Anduze (France)*, 13<sup>th</sup> IWRA World Water Congress 2008, 10 pages
- Bouvier C., Ayral P.A., Brunet P., Crespy A., Marchandise A., Martin C.**, (2007), *Recent advances in rainfall-runoff modelling: extrapolation to extreme floods in southern France.*, Proceedings of the 1<sup>st</sup> International Workshop on Hydrological Extremes, FRIEND-AMHY, Consenza; Italy, 3-4 may 2006, 229-238
- Daget N., Weaver A.T. et Balmaseda M.A.**, (2008), *Ensemble estimation of background-error variances in a three-dimensional variational data assimilation system for global ocean analysis*, in revision for Quarterly Journal of the Royal Meteorological Society
- Lagarde T., Piacentini A. et Thual O.**, (2001), *A new representation of data-assimilation methods ; the PALM flow-charting approach*, Quarterly Journal of the Royal Meteorological Society, 127, 189-207 [http://www.cerfacs.fr/globc/PALM\\_WEB/](http://www.cerfacs.fr/globc/PALM_WEB/)
- Marchandise A.**, (2007), *Modélisation hydrologique distribuée sur le Gardon d'Anduze ; étude comparative de différents modèles pluie-débit, extrapolation de la normale à l'extrême et test d'hypothèses sur les processus hydrologiques*, Thèse de l'Université Montpellier II, 214 pages
- Marchandise A. et Viel C.**, (2008), *Utilisation de l'indicateur d'humidité de la chaîne ISBA de Météo-France pour la prévision opérationnelle des crues*, 191<sup>ème</sup> Colloque de la Société Hydrotechnique de France, Lyon
- Massart S., Buis S., Erhard P. et Gacon G.**, (2007a), *Use of 3DVAR and Kalman Filter Approaches for Neutronic State and Parameter Estimation in Nuclear Reactors*, Nuclear Science and Engineering, 155(3), pp 409-424
- Massart S., Cariolle D. et Peuch V.-H.**, (2005), *Vers une meilleure représentation de la distribution et de la variabilité de l'ozone atmosphérique par l'assimilation des données satellitaires*, Comptes Rendus Geosciences 337, pp. 1305-1310
- Massart S., Piacentini A., Cariolle D., El Amraoui L. et Semane N.**, (2007b), *Assessment of the quality of the ozone measurements from the Odin/SMR instrument using data assimilation*, Canadian Journal of Physics 85, pp. 1209-1223
- Pannekoucke O. et Massart S.**, (2008), *Estimation of the local diffusion tensor and normalization for heterogeneous correlation modelling using a diffusion equation*, Quarterly Journal of the Royal Meteorological Society 134 (635), pp. 1425-1438
- Ponçot A.**, (2008), *Assimilation de données pour la dynamique du xénon dans les coeurs de centrale nucléaire*, Thèse de doctorat de l'Université de Toulouse, 266 pages
- Weaver A.T., Deltel C., Machu E., Ricci S. et Daget N.**, (2005), *A multivariate balance operator for variational ocean data assimilation*, Quarterly Journal of the Royal Meteorological Society 131, pp 3605-3625