
La prévision en hydrologie enjeux et incertitudes

Selon l'Agence Européenne de l'Environnement (EEA), entre les années 1998 et 2009, les crues observées en Europe ont causé 1126 morts, 500 000 déplacés et 52 milliards d'euros de pertes pour les assurances. De plus, l'Agence souligne que le risque d'observer des événements de crue plus fréquemment à l'avenir est élevé. Pour éviter les pertes de vies humaines et minimiser les dégâts économiques liés à ces événements extrêmes, la mise en place d'un système de prévision et alerte des crues solide et fiable est essentiel. L'objectif général est de pouvoir anticiper l'arrivée d'un événement potentiellement dangereux, ainsi que de bien capturer son intensité et localisation. Prévision et prévention des risques vont ainsi de paire. La Directive Européenne relative à l'évaluation et à la gestion des risques d'inondation de 2007 (DIRECTIVE 2007/60/CE) rappelle ainsi que " les plans de gestion des risques d'inondation englobent tous les aspects de la gestion des risques d'inondation, en mettant l'accent sur la prévention, la protection et la préparation, y compris la prévision des inondations et les systèmes d'alerte précoce (...) " (Chapitre IV, Article 7, §3).

Cependant, l'importance de la prévision hydrologique ne se restreint pas à l'anticipation des événements extrêmes, qu'ils soient caractérisés par des situations de crues ou d'étiages. Un système de prévision hydrologique efficace peut également contribuer à la gestion optimale de la ressource en eau pour les activités économiques qui en dépendent (production hydroélectrique, irrigation, tourisme, navigation, etc.). Il s'agit de mieux répartir l'utilisation de l'eau dans le temps et dans l'espace, à partir de l'information apportée par le système de prévision sur les apports en eau à venir. Au delà du gain en sécurité face aux risques des événements extrêmes, la valeur ajoutée de cette connaissance se traduira en gains économiques ou de gestion des ouvrages

Idéalement, pour que les services de protection civile ou les gestionnaires d'ouvrages aient le temps suffisant pour intervenir, une alerte doit être déclenchée au minimum dans les 12 à 24 heures précédant l'événement. De plus, les services opérationnels de prévision doivent pour cela avoir des procédures efficaces de prévision, interprétation et communication des alertes pour répondre aux besoins liés à ces temps de réaction.

Dans ces cas, il est évident qu'une prévision hydrologique à moyen terme (3 à 10 jours en avance) a toute sa place. Pour cela, la prévision météorologique du temps à moyen terme est indispensable comme source d'information sur les conditions météorologiques à venir qui seront utilisées en entrée des modèles hydrologiques.

Pour pouvoir faire une prévision du temps à moyen terme et aller au-delà des limites de prévisibilité classiques, la prise en compte des incertitudes est incontournable. La prévision déterministe, à scénario unique, n'est pas en mesure de répondre à ces objectifs. Pour cela, les services météorologiques font appel aux systèmes de prévision d'ensemble, générés sur la

base de variations imposées dans les conditions initiales des modèles numériques et de variations stochastiques de leur paramétrisation (Palmer *et al.*, 2005). La prévision météorologique d'ensemble propose des scénarios équiprobables de l'évolution de l'atmosphère pour des horizons de prévision pouvant aller jusqu'à 10-15 jours, ce qui serait en dehors de la place de prévisibilité des solutions traditionnelles déterministes. Le recours à la technique de la prévision d'ensemble permet ainsi, d'une part, d'étendre les échéances de prévision des variables à faible prédictibilité, telles les précipitations, et, d'autre part, de donner une mesure de l'incertitude des prévisions.

Par contraste à une solution unique déterministe, la prévision probabiliste est vue comme mieux adaptée à l'analyse de risque et la prise de décision (Rodwell, 2005). La valeur économique potentielle des prévisions probabilistes ressort notamment face aux utilisateurs exposés aux risques climatiques et hydrologiques à différentes échelles ; ainsi qu'aux gestionnaires dont les décisions dépendent de la disponibilité nécessaire en eau : pour la génération d'électricité, agriculture et irrigation, navigation, sécurité et santé publiques, etc. (Buizza, 2008, 2001 ; Houdant, 2004).

L'approche probabiliste ou d'ensemble est actuellement encore considérée comme faisant partie de l'état de l'art et plusieurs services météorologiques (France, Suisse, Belgique, Allemagne, Suède, etc.) développent ou envisagent de constituer un système de prévisions hydrologiques basé sur les prévisions d'ensemble (Cloke et Pappenberger, 2009b). La prévision hydrologique d'ensemble est au cœur des initiatives internationales regroupant plusieurs centres de recherche et services opérationnels (HEPEX *Hydrological Ensemble Prediction Experiment*, EFAS *European Flood Awareness System*), ainsi que dans les actions de recherche et développement de plusieurs services nationaux (Centre National de Recherche Météorologique, Météo-France, EDF, Institut Royal Météorologique de Belgique, etc.). Les résultats de prévision des crues à partir de la prévision d'ensemble de pluies sont encourageants.

En météorologie, les systèmes de prévision d'ensemble sont développés pour prendre en compte l'incertitude liée aux prévisions. L'application de ces prévisions dans un modèle pluie-débit pour la prévision hydrologique constitue une étape importante pour la prévision des débits, puisque les précipitations sont une des sources principales d'incertitude.

Néanmoins, l'incertitude se trouve dans toute la chaîne de prévision hydrométéorologique et se propage (cascade des incertitudes) de la donnée d'entrée jusqu'à la sortie du modèle. Les incertitudes de prévision sont des sources d'erreur, mais le manque de connaissances des incertitudes constitue également une source d'erreur. La prise en compte de l'ensemble de sources d'incertitude est nécessaire afin d'améliorer la qualité des prévisions et de fournir des informations fiables et utiles.

Le fait de disposer d'une prévision sous forme probabiliste, et non pas de manière déterministe, est sans doute un avantage considérable pour les services opérationnels qui font face à la prise de décision en contexte incertain. Cependant, une telle approche soulève également de nouvelles questions : comment interpréter les sorties probabilistes d'un modèle de prévision ? Comment communiquer les prévisions avec leurs incertitudes ? Comment prendre en compte toutes les sources d'incertitude dans un système de prévision ? Comment évaluer la qualité et l'utilité d'un système de prévision ?

Certaines de ces questions seront abordées dans le cadre de cette thèse. Le cadre général dans lequel se développe cette étude est présenté ci-après (§1.2 et §1.3), suivi des défis actuels auxquels s'intéresse cette thèse (§1.4) et des objectifs spécifiques de nos travaux (§1.5).

1.2 "De l'incertitude des prévisions à la prévision des incertitudes"

Cette thèse s'intéresse aux prévisions hydrologiques probabilistes (ou multi-scénario) à moyenne échéance. Le système de prévision hydrologique est une chaîne qui a pour principaux éléments (Figure 1) : les données observées (précipitation, température et débits en temps réel observés jusqu'à l'instant de prévision), les prévisions météorologiques (scénarios futurs jusqu'à l'échéance cible) et le modèle hydrologique de transformation de la précipitation en débit. Le but est de fournir des prévisions hydrologiques pour une des utilisations cibles (génération d'électricité, agriculture et irrigation, navigation, sécurité et santé publiques, etc.).

Dans cette thèse, nous avons utilisé l'approche de la prévision d'ensemble pour prendre en compte des incertitudes de prévision et nous avons cherché à explorer l'usage de la prévision météorologique d'ensemble pour la prévision du temps à venir dans la prévision hydrologique des débits des rivières. Pour mieux comprendre nos objectifs de thèse, quelques notions introductives en prévision du temps et approches d'ensemble sont présentées par la suite.

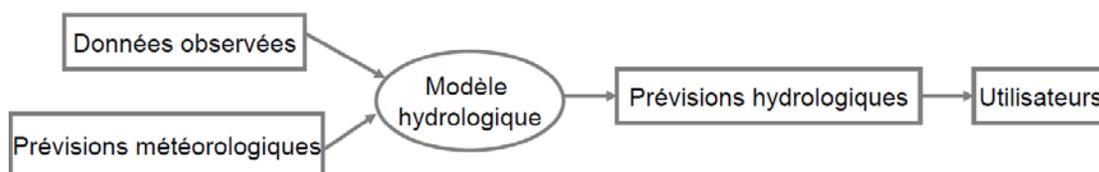


Figure 1 : Schéma général d'une chaîne de prévision hydrologique.

1.2.1 La prévision numérique du temps : un peu d'histoire

Les efforts pour prévoir le temps sont plus que centenaires. En 1904, le physicien V. Bjerknes était le premier à écrire sur la modélisation et la prévision du temps, en utilisant des équations de l'hydrodynamique basées sur des lois de Newton et de la thermodynamique. Le principe général de la prévision y était formulé : " connaissant l'état de l'atmosphère à un instant donné, et les équations qui régissent son évolution, il était possible de connaître l'état de l'atmosphère à tout instant ultérieur " (Rochas, 2000). Le mathématicien L.F. Richardson a tenté de résoudre ces équations par la méthode des différences finies et a été ainsi le précurseur de la modélisation numérique de prévision du temps. Après avoir consacré plusieurs années à la prévision du temps, il est arrivé à la conclusion que pour prévoir le temps plus vite que la vitesse avec laquelle le temps évolue il lui fallait 64 000 personnes travaillant simultanément ! En effet, de nombreux progrès étaient encore nécessaires pour que les méthodes de prévision du temps arrivent à être appliquées avec succès : il fallait des ordinateurs plus puissants, des moyens de mesure des paramètres atmosphériques en altitude couvrant une bonne partie de la planète, ainsi qu'un réseau de communication rapide de ces mesures vers les centres de prévision.

C'est dans les années 1950/60 que les premiers modèles opérationnels de prévision ont vu le jour. Depuis, les modèles ont considérablement évolué grâce à la meilleure représentation des processus physiques, à la meilleure connaissance de la microphysique des nuages, à la prise en compte de mesures nouvelles provenant des réseaux radar et des satellites, et à l'amélioration de la capacité de calcul des ordinateurs, qui a permis, notamment, d'augmenter la résolution spatiale des modèles.

Néanmoins les modèles numériques sont encore soumis à des erreurs liées aux conditions initiales, à leur paramétrage et structure (Murphy *et al.*, 2004 ; Bell *et al.*, 2004 ; Zupanski et Zupanski, 2006 ; Johnson et Swinbank, 2009 ; Goldstein et Rougier, 2009). Les observations qui sont assimilées dans le modèle afin de produire une bonne estimation de l'état actuel de l'atmosphère ont des incertitudes considérables (Leutbecher et Palmer, 2007). Ces erreurs se propagent et s'y ajoutent les erreurs du modèle, notamment sur les régions où la grille spatiale reste trop large pour modéliser les phénomènes locaux et où la représentation physique des phénomènes atmosphériques est imparfaite (Buizza et Palmer, 1995). Les principales sources des fléaux de la prévision se trouvent donc exactement dans le berceau de la prévision : les données, les modèles et le caractère chaotique de l'atmosphère.

En effet, c'est pour cette raison qu'il est impossible de prédire son état exact (Lorenz, 1969). A cause du caractère chaotique, non-linéaire et complexe de l'atmosphère, l'évolution temporelle de deux systèmes météorologiques ressemblant au départ à un système observé ne sera pas nécessairement la même. Ceci est expliqué par les petites différences dans leurs conditions initiales, lesquelles grandissent pendant l'évolution temporelle du système (Lorenz, 1982).

Ainsi, plus la prévision que l'on souhaite fournir est lointaine, plus l'état futur devient incertain. Cette incertitude est aussi dépendante du flux atmosphérique et, par conséquent, elle est considérablement plus importante pour les phénomènes extrêmes, qui sont alimentés par de grands flux d'énergie.

Le caractère chaotique de l'atmosphère contrôle l'imprédictibilité des événements naturels. Selon Koutsoyiannis (2010), dans les processus physiques, les composantes déterministes (qui sont équivalentes à la prédictibilité) et chaotiques (qui sont équivalentes à l'imprédictibilité) coexistent. Selon la composante dominante, la prédictibilité d'un phénomène change. Cela veut dire que les processus très chaotiques peuvent avoir une prédictibilité de quelques jours ou même de quelques heures, alors que, pour des conditions plus stables, la prédictibilité peut s'éteindre, par exemple, à plusieurs jours ou même des semaines.

Joly (2008) offre la définition suivante de la prévisibilité : "*l'échéance au-delà de laquelle une prévision s'écarte trop de l'évolution observée pour rester utile*". Nous notons que dans cette définition, les notions de qualité (proximité ou écart de l'observation) et utilité se côtoient. De plus, nous observons que la notion d'incertitude devient grandissante avec l'horizon de prévision. Le besoin de quantifier cette incertitude pour que les prévisions à horizon lointain restent utiles devient une évidence.

1.2.2 La prévision météorologique d'ensemble

Pendant plusieurs décades, les centres opérationnels de prévision du temps lançaient leurs modèles numériques de prévision à partir de la meilleure estimation des conditions initiales. La prévision était ainsi limitée à une seule indication de l'état futur possible de l'atmosphère, le "*best guess*". En 1992, des modèles de prévision ont commencé à fonctionner en considérant des conditions initiales légèrement perturbées. Le service national de prévision des Etats-Unis (NCEP) et le centre européen de prévision météorologique à moyen terme (CEPMMT) ont été pionniers dans cette démarche (Molteni et Palmer, 1993 ; Toth et Kalnay, 1993). La prévision d'ensemble est ainsi née.

Palmer *et al.* (2002) définissent la prévision d'ensemble comme "une technique de prévision météorologique dans laquelle on fait tourner un modèle numérique de prévision plusieurs fois, pour une même situation à prévoir, à partir de conditions initiales différant entre elles par de petites quantités [les perturbations initiales] compatibles avec les incertitudes existant sur la connaissance de l'état initial de l'atmosphère". Le problème fondamental de la prévision d'ensemble est d'effectuer un choix judicieux des conditions initiales de façon à pouvoir obtenir un maximum de solutions éloignées les unes des autres avec un minimum d'états initiaux. Il faut construire un ensemble initial qui sera un échantillon représentatif de la loi de probabilité qui décrit l'incertitude des conditions initiales.

L'idée est d'échantillonner la loi de probabilité des conditions initiales et de suivre la propagation des différentes perturbations dans l'évolution de l'état de l'atmosphère pour ainsi avoir différents scénarios ou trajectoires de prévision (Figure 2). La prévision d'ensemble est ainsi souvent faite d'une trajectoire provenant de la meilleure estimation de l'état de l'atmosphère au moment de la prévision (trajectoire ou membre dit "contrôle") et des autres trajectoires obtenues en perturbant les conditions initiales. Idéalement les perturbations de la meilleure estimation de l'état initial conduiront à un ensemble avec une dispersion similaire à la dispersion des erreurs de prévision au bout d'un certain intervalle de temps appelé "période d'optimisation" (Ehrendorfer et Tribbia, 1997).

Plusieurs méthodes permettant de créer les perturbations des conditions initiales existent. Parmi les méthodes opérationnelles les plus courantes se trouvent la méthode des Vecteurs Singuliers (Buizza et Palmer, 1995), la méthode des erreurs croissantes (en anglais "*Breeding of Growing Modes*"), (Toth et Kalnay, 1993, 1997) et la méthode du filtre Kalman d'ensemble (Bishop *et al.*, 2001 ; Kalnay, 2003). De manière générale, toutes ces méthodes visent à appréhender l'incertitude liée aux conditions initiales pour la production des prévisions d'ensemble correctement dispersées. Néanmoins, comme le remarque bien Smith (2007), l'incertitude des conditions présentes ne va se traduire en incertitudes bien quantifiées dans le futur que si le modèle est parfait. De ce fait, une approche complémentaire dans la création des ensembles a été introduite, celle de la perturbation des paramétrisations physiques des modèles (Buizza *et al.*, 1999 ; Houtekamer et Lefaire, 1997). Il s'agit d'introduire des perturbations à la meilleure estimation de la paramétrisation physique (celle de la trajectoire de contrôle), afin de prendre en compte également les incertitudes liées à la formulation du modèle. Souvent il s'agit de perturbations stochastiques des paramètres.

De plus, on notera l'existence également d'approches qui combinent les sorties de différents modèles numériques du temps pour créer des prévisions d'ensemble :

- l'ensemble du pauvre (en anglais, "*poor man's ensemble*") : combinaison de prévisions déterministes (uni-valeurs) provenant de différents centres météorologiques pour former un ensemble,
- le grand ensemble TIGGE (acronyme en anglais de *THORPEX Interactive Grand Global Ensemble*) : combinaison de prévisions d'ensemble provenant de différents centres météorologiques autour du globe. L'objectif derrière cette initiative est de capturer les incertitudes liées aux conditions initiales, à la paramétrisation, à la structure et à l'assimilation des données des différents modèles numériques du temps utilisés par les différents services météorologiques (Park *et al.*, 2008).

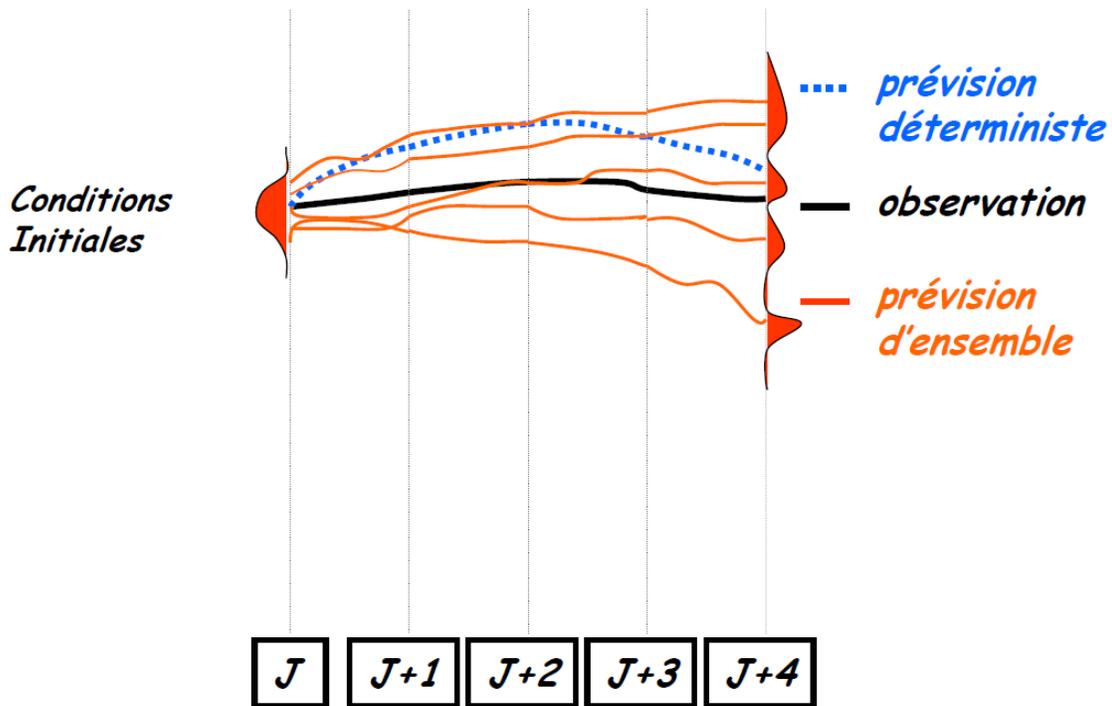


Figure 2: Schéma des prévisions de l'évolution de l'atmosphère pour le cas d'une prévision déterministe et des prévisions d'ensemble et des échéances allant jusque $J+4$.

Dans cette thèse, nous nous intéressons plus particulièrement à l'utilisation des prévisions météorologiques d'ensemble pour la prévision hydrologique d'ensemble. La description plus détaillée de ces techniques de modélisation d'ensemble n'est pas l'objectif de cette thèse. Nous retenons seulement qu'il existe plusieurs méthodes de génération d'ensembles actuellement utilisées par les centres météorologiques, sans qu'aucune ne soit universellement considérée comme la meilleure.

1.2.3 La prévision hydrologique d'ensemble

Même si les prévisions météorologiques d'ensemble à moyenne échéance sont produites par les services opérationnels depuis 20 ans, ce n'est que pendant la dernière décennie que les premiers systèmes de prévision hydrologique les utilisant ont vu le jour (De Roo et al., 2003).

En effet, l'utilisation des ensembles en hydrologie était plutôt répandue dans les cas de la prévision hydrologique à longue échéance (> 15 jours). La méthode ESP ("Ensemble Streamflow Prediction") a été développée par le service météorologique national des Etats-Unis ("National Weather Service") dans les années 1970-1980 (Day, 1985) dans le but de repousser l'horizon des prévisions hydrologiques. Il s'agit d'utiliser des données météorologiques historiques afin de générer un ensemble des scénarios possibles des débits futurs, conditionnellement aux états initiaux du bassin versant.

L'utilisation de la prévision météorologique d'ensemble pour la prévision hydrologique à des horizons plus courts est plus récente. Elle a été néanmoins rapidement répandue, notamment grâce à des initiatives significatives qui ont été proposées dans un large cadre de coopération entre instituts de recherche et services opérationnels. Dans ce contexte, il ressort notamment:

- l'EFAS (acronyme en anglais de *European Flood Alert System*, récemment rebaptisé *European Flood Awareness System*) : il s'agit d'un système européen d'alerte aux crues dont le développement a été initié en 2003 par le centre commun de recherche de la commission européenne (JRC) afin d'améliorer la prévision dans les grands bassins versants transnationaux et d'encourager la coopération entre Etats-Membres. L'un des objectifs de cette initiative est de fournir des pré-alertes (échéances supérieures à 3 jours) aux services nationaux de prévision à partir de l'utilisation de la prévision météorologique d'ensemble en entrée à un modèle hydrologique distribué, mis en place sur des bassins européens transnationaux, et ainsi contribuer à augmenter le temps de préparation des services opérationnels en cas de crues potentiellement dangereuses (Thielen *et al.*, 2009 ; Bartholmes *et al.*, 2009) ;
- l'HEPEX (acronyme en anglais de *Hydrological Ensemble Prediction Experiment*) : il s'agit d'une initiative internationale démarrée en 2003, suscitant fomentation de la coopération entre chercheurs en météorologie et hydrologie et services opérationnels dans le domaine de la prévision hydrologique d'ensemble (Schaake *et al.*, 2005, 2007). HEPEX a pour but de réunir les communautés de météorologues, hydrologues et gestionnaires de l'eau pour promouvoir la mise en place des systèmes de prévision hydrologique d'ensemble fiables, ainsi que la communication et l'utilisation des prévisions dans la prise de décision.

Ces dernières années, des nombreuses études ont ainsi été dédiées à l'évaluation de l'intérêt de la prévision d'ensemble pour la prévision hydrologique, qu'il s'agit d'étudier :

- la viabilité de l'approche dans un contexte opérationnel de prévision à moyenne échéance (Rousset-Regimbeau *et al.*, 2007 ; Zappa *et al.*, 2008 ; Addor *et al.*, 2011) ou pour la prévision des crues rapides (Younis *et al.*, 2008a ; Alfieri et Thielen, 2012 ; Marty *et al.*, 2012),
- la qualité des prévisions émises soit sur la base de longues séries prévision-observation (Roulin et Vannitsem, 2005 ; Rousset-Regimbeau *et al.*, 2008 ; Olsson et Lindstrom, 2008 ; Jaun et Ahrens, 2009 ; Demargne *et al.*, 2009 ; Bartholmes *et al.*, 2009 ; Randrianasolo *et al.*, 2010), soit sur des événements particuliers (Bartholmes et Todini, 2005 ; Pappenberger *et al.*, 2005 ; Komma *et al.*, 2007 ; Jaun *et al.*, 2008 ; Kalas *et al.*, 2008 ; Younis *et al.*, 2008b),
- l'impact de la prévision d'ensemble pour la courte échéance et des petites échelles spatiales (Thirel *et al.*, 2008 ; Marsigli *et al.*, 2008),

- l'apport des approches multi-modèles hydrologiques (Velazquez *et al.*, 2009) ou multi-modèles météorologiques (Pappenberger *et al.*, 2008 ; Jasper *et al.*, 2002 ; Davoglio *et al.*, 2008),
- le développement de méthodes de correction de biais (Krzysztofowicz et Kelly, 2000 ; Reggiani et Weerts, 2008 ; Brown et Seo, 2010 ; Coccia et Todini, 2011),
- la valeur économique espérée des prévisions (Buizza, 2008 ; Verkade et Werner, 2011 ; Boucher *et al.*, 2012), ou
- la communication des incertitudes et l'utilisation des prévisions incertaines pour la prise de décision (Houdant, 2004 ; Norbert *et al.*, 2010 ; Ramos *et al.*, 2010 ; 2012).

Toutes ces études ont derrière elles au moins deux points en commun : elles s'intéressent à l'application opérationnelle des prévisions hydrologiques d'ensemble et confirment la valeur ajoutée apportée par la prévision d'ensemble à un système de prévision et prise de décision.

Actuellement, plusieurs services opérationnels en France (SCHAPI - Météo-France/CNRM, EDF) et ailleurs (Cloke et Pappenberger, 2009b) ont développé ou envisagent de constituer un système de prévision hydrologique basé sur les prévisions d'ensemble. Pour les services opérationnels, le fait que la prévision des débits (en crues ou périodes normales) à partir de la prévision météorologique d'ensemble soit présentée sous forme probabiliste ou multi-scénarios, et non pas de manière déterministe (scénario futur unique), est un avantage considérable, notamment pour la prise de décision en contexte incertain (Ramos *et al.*, 2010). En prenant en compte l'incertitude de la prévision météorologique, les prévisions d'ensemble peuvent couvrir des échéances de prévision plus lointaines et augmenter ainsi le temps de préparation (Cloke et Pappenberger, 2009a). De plus, la mise à disposition d'une information probabiliste permet aux décisionnaires de prendre leur décision eux-mêmes, au regard des risques qu'ils souhaitent prendre, au lieu de laisser cette décision dans les mains des modélisateurs ou des prévisionnistes, parfois éloignés des objectifs de l'utilisateur des prévisions (Weerts *et al.*, 2010).

Enfin, on notera que toutes les prévisions d'ensemble ne constituent pas de "bons" ensembles. Par exemple, souvent les ensembles n'arrivent pas à bien décrire toutes les évolutions possibles de l'atmosphère et, par conséquent, l'observation n'est pas capturée par le faisceau des prévisions défini par les trajectoires de l'ensemble. Des traitements statistiques sont alors nécessaires pour corriger les éventuels biais. De plus, afin de juger si une prévision d'ensemble est "bonne" ou pas, il y a deux points de vue différents, même si pas forcément disjoints : celui du prévisionniste et celui de l'utilisateur. Le prévisionniste est souvent intéressé par la qualité des prévisions en termes de proximité par rapport à l'observation, tandis que l'utilisateur focalise l'utilité de l'information pour la décision qu'il doit prendre. Par conséquent, l'évaluation des prévisions d'ensemble concernera souvent deux axes principaux : i) trouver où se situe l'observation par rapport à la gamme de valeurs prévues des différents membres de l'ensemble (*évaluation de la qualité*) et ii) déterminer

l'utilité de la prévision d'ensemble pour un utilisateur-décideur (valeur décisionnelle ou économique) par rapport aux prévisions déterministes et/ou à une prévision dite naïve, telle que la climatologie, c'est-à-dire comparer chaque observation à une prévision égale à la moyenne climatologique, ou la persistance, c'est-à-dire comparer chaque observation à une prévision égale au dernier débit observé (*évaluation de l'utilité*) (Jolliffe et Stephenson, 2003 ; Buizza *et al.*, 2007b).

1.3 Aspects clés de la chaîne de prévision hydrologique d'ensemble

Même si l'incertitude des prévisions météorologiques est souvent considérée comme la principale source d'erreur pour les prévisions hydrologiques, notamment pour les échéances supérieures à 2-3 jours, d'autres facteurs influent également sur la qualité des prévisions hydrologiques. Les principaux d'entre eux sont les erreurs liées à la modélisation hydrologique (structure et paramétrage du modèle) et aux observations utilisées en temps réel pour la mise à jour des conditions initiales du modèle hydrologique avant l'instant de prévision (Pierce *et al.*, 2005 ; Wood et Lettenmaier, 2008 ; Bogner et Kalas, 2008 ; Zhao *et al.*, 2011 ; Pagano *et al.*, 2012 ; Zalachori *et al.*, 2012). La compatibilité entre les échelles spatio-temporelles des sorties des modèles météorologiques et celles requises par les modèles hydrologiques, souvent plus fines, peut également devenir une source d'incertitude et mériter un traitement préalable de désagrégation ou descente d'échelle (Marty *et al.*, 2012).

Dans la réalité opérationnelle, le prévisionniste se trouve souvent face à une combinaison de plusieurs de ces sources d'incertitude (Pappenberger *et al.*, 2005). Etant donné qu'il est devenu évident qu'aucune amélioration de modèle n'éliminera totalement toutes les incertitudes de prévision, il est indispensable de prendre en compte cette incertitude dans la procédure de prévision, même si on doit le faire de manière globale et/ou implicite. Une chaîne de prévision hydrologique doit en effet être capable de gérer les différentes sources d'incertitudes et fournir des prévisions précises et fiables à des utilisateurs-décideurs exposés aux risques hydrologiques à différentes échelles. Elle doit en plus permettre d'évaluer l'impact de la prise en compte des incertitudes sur la qualité et l'utilité opérationnelle des prévisions hydrologiques.

Pour cela, plusieurs aspects doivent être considérés au sein d'une chaîne de prévision hydrologique d'ensemble, dont (au moins) les quatre points-clés suivants (Figure 3):

- I. *l'évaluation de la qualité* des prévisions hydrométéorologiques émises : la qualité d'une prévision est en général jugée par rapport à l'observation. Pour cela il est essentiel de comparer une longue série de prévisions avec la série correspondante des observations. Selon Jolliffe et Stephenson (2003) cette évaluation de la qualité d'une prévision d'ensemble consiste à trouver où se situe l'observation par rapport à la gamme des valeurs prévues des différents membres de l'ensemble ; et estimer ainsi

- l'écart entre les valeurs prévues et celles réellement observées ("biais"). Cette étape précède toute utilisation des prévisions hydrométéorologiques d'ensemble,
- II. le *prétraitement* des prévisions utilisées en entrée du modèle hydrologique : il comprend toutes les approches de correction statistique de biais mises en place en amont du modèle hydrologique afin de réussir au mieux à prendre en compte l'incertitude liée aux prévisions météorologiques,
 - III. le *post-traitement* statistique des prévisions : il s'agit de toutes les approches de correction statistiques employées à la sortie du modèle hydrologique pour corriger les biais éventuels des prévisions des débits,
 - IV. *l'évaluation de l'utilité* des prévisions : il s'agit d'examiner si à l'aide de l'information probabiliste fournie par les prévisions d'ensemble, les utilisateurs peuvent mieux décider des actions à mener pour maximiser les gains ou limiter les pertes liées à l'impact d'un événement à venir sur leurs activités. L'évaluation objective de l'utilité d'une prévision sur le long terme fait appel à des techniques économiques, qui cherchent à proposer une mesure de l'utilité, "valeur économique", d'une prévision pour aider à la prise de décision.

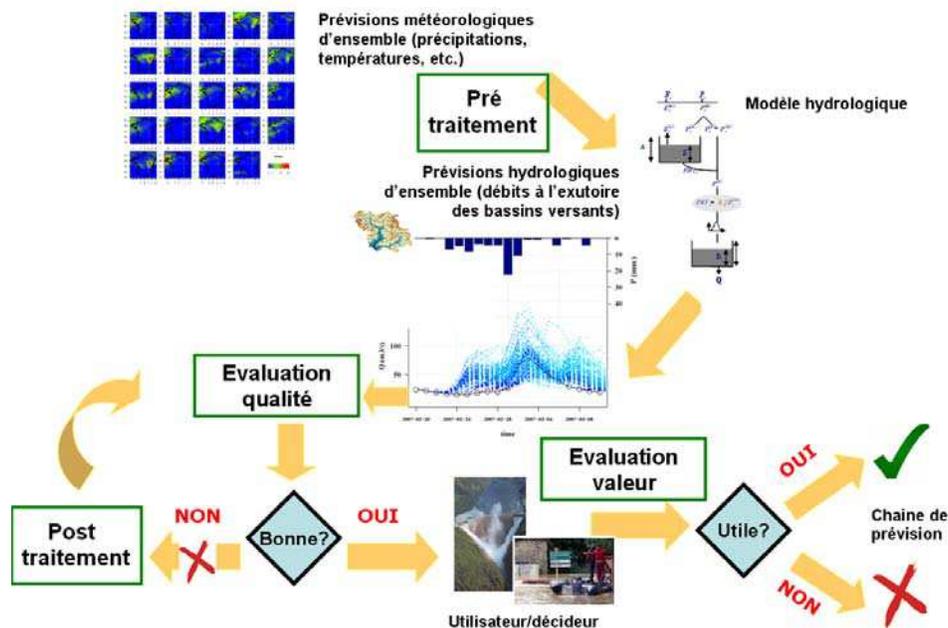


Figure 3: Aspects clés de la chaîne de prévision hydrologique d'ensemble

Ces quatre aspects constituent le cœur des travaux de cette thèse. Ils comportent de nombreux défis à relever pour que le système de prévision puisse émettre des prévisions de bonne qualité et utiles. Ils sont listés dans le §1.4 ci-dessous.

1.4 Défis actuels dans la recherche de la qualité et de l'utilité des prévisions

Dans l'objectif de fournir des prévisions hydrologiques *bonnes* et *utiles* aux utilisateurs, il est essentiel de répondre aux différents défis de la chaîne de prévision hydrologique. Pour cette raison, la première étape consiste à les identifier.

1.4.1 La qualité des prévisions

L'évaluation des prévisions d'ensemble montre souvent l'existence de différents biais dans les prévisions. Une fois identifiés, une correction statistique appropriée doit se mettre en place afin de corriger ces biais et minimiser les différences entre la série de données observées et les prévisions d'ensemble émises.

Il y a plusieurs raisons qui expliquent ces différences :

- les erreurs d'échantillonnage liées à la trop petite taille de l'ensemble ;
- les biais ou erreurs systématiques liés à l'observation ou au(x) modèle(s) hydrométéorologiques (paramètres, structures, conditions initiales, etc.) ;
- la dispersion inappropriée des membres, venant soit d'une mauvaise prise en compte des incertitudes pendant la génération des perturbations qui vont produire les ensembles météorologiques, soit des modifications induites lors du passage par le modèle hydrologique de prévision.

Afin d'améliorer ces défauts des prévisions, la solution idéale est de s'attaquer à la racine de ces problèmes. Une telle approche se concentrera alors sur l'élargissement des ensembles (augmentation de leur dispersion) et/ou sur l'amélioration des modèles hydro-météorologiques et leur façon de quantifier et propager les différentes sources d'incertitude qui affectent leurs simulations. Néanmoins, ces démarches sont considérablement chronophages : il faut bien identifier les sources d'incertitude, les quantifier de façon appropriée et s'assurer de la robustesse de la démarche de propagation des incertitudes. Une alternative plus rapide et efficace est de faire appel aux techniques de correction statistique des biais avant (pré-) ou après (post-) le modèle hydrologique, avec pour but d'améliorer la qualité des prévisions en termes de volume, d'évaluation temporelle et de distribution statistique. L'objectif final se résume à obtenir des prévisions fiables et précises en même temps.

Un nombre considérable d'approches de correction statistique de biais existent dans la littérature. Nous pouvons séparer ces approches en deux principales familles :

- les approches qui visent à corriger l'erreur totale de la prévision (Krzysztofowicz, 2002 ; Roulston et Smith, 2003 ; Gneiting *et al.*, 2005 ; Raftery *et al.*, 2005 ; Wang et Bishop, 2005 ; Fortin *et al.*, 2006 ; Seo *et al.*, 2006 ; Slaughter *et al.*, 2007 ; Duan *et al.*, 2007 ; Hamill *et al.*, 2008 ; Vrugt *et al.*, 2008 ; Regianni *et al.*, 2009 ; Wilks, 2009 ; Bao *et al.*, 2010 ; Brown et Seo, 2010, 2012 ; Weerts *et al.*, 2011 ; Roulin et Vannitsem, 2012) ;
- les approches qui traitent chaque source d'erreur séparément : conditions initiales, structure du modèle, paramétrage, etc. (Beven et Binley, 1992 ; Kuczera et Parent, 1998 ; Seo *et al.*, 2003 ; Vrugt *et al.*, 2003 ; Goswami *et al.*, 2005 ; Pappenberger *et al.*, 2005 ; Andreatis et Lettenmaier, 2006 ; Bogner et Kalas, 2008 ; Zhao *et al.*, 2011 ; Pagano *et al.*, 2012 ; Zalachori *et al.*, 2012).

Dans cette thèse nous nous sommes principalement intéressés par les deux familles, mais en ciblant particulièrement les approches de correction de biais de caractère empirique. Ce choix se justifie par le contexte général de la thèse qui se situe dans le cadre des applications opérationnelles de la prévision d'ensemble. Grâce à leur simplicité, les approches empiriques sont considérées comme des approches simplifiées d'agrégation des erreurs et elles sont souvent préférées par les services opérationnels. Dans la plupart des cas, il s'agit d'ajustements des quantiles ou d'habillage des membres (Hashino *et al.*, 2006 ; Seo *et al.*, 2006 ; Wood et Schaake, 2008 ; Olsson et Lindström, 2008 ; Kang *et al.*, 2010 ; Weerts *et al.*, 2011 ; Pagano *et al.*, 2012 ; Boucher *et al.*, 2012 ; Zalachori *et al.*, 2012).

Etant donné que les approches de traitement statistique constituent actuellement un sujet clé de la recherche sur les prévisions hydrologiques d'ensemble, la communauté scientifique se trouve à présent en plein débat sur les questions soulevées. Lors des derniers colloques HEPEX (Delft, 2011 et Beijing, 2012), quelques défis ont été identifiés :

- I. Comment évaluer les différentes approches de traitement statistique existantes sur une base comparable de calage/ validation ?
- II. Comment évaluer ces approches en utilisant des critères multiples d'évaluation de façon à faire ressortir leurs points forts et leurs points faibles ?
- III. Comment combiner les techniques d'assimilation de données et les approches de correction de biais appliquées en amont du modèle hydrologique (prétraitement) de manière efficace ?
- IV. Comment les approches de traitement statistique peuvent-elles contribuer à l'amélioration de prévisions probabilistes des événements extrêmes rares ?
- V. Comment évaluer ces approches en vue du gain qu'elles apportent à leur utilisateur final dans un contexte de prise de décision ?

- VI. Comment guider les utilisateurs vers des approches de post-traitement les plus adaptées à leurs besoins ?
- VII. Comment éviter que les approches de traitement statistique ne deviennent des boîtes-noires pour les utilisateurs finaux ? Comment impliquer ces derniers dans la conception de ces approches ?

Parmi ces nombreux défis, cette thèse s'intéressera aux questions suivantes :

- L'utilisation d'une approche multi-modèle peut-elle améliorer la qualité des prévisions ?
- L'utilisation des approches empiriques de correction de biais améliore-t-elle la qualité des prévisions hydrologiques d'ensemble ? Existe-t-il dans ce domaine une méthode plus performante que les autres ?
- Un prétraitement est-il essentiel ou un post-traitement est-il préférable ?
- La longueur de calage influence fortement la performance des approches de correction statistique de biais ?

1.4.2 L'utilité des prévisions

L'intérêt de maximiser l'utilité d'une prévision est lié à la recherche d'une prise de décision optimale, à moindre coût et un maximum de bénéfice pour l'utilisateur des prévisions. L'utilité d'une prévision est étroitement liée au concept de la valeur économique de cette prévision. Afin de mesurer et optimiser la valeur économique d'une prévision, il est primordial de recourir à des modèles de décision et des techniques d'optimisation, permettant de lier les événements prévus à ceux observés, les actions des décideurs à leurs conséquences. Ces modèles peuvent être basés sur des objectifs divers : par exemple, minimiser les fausses alertes pour la prévision des crues, maximiser la production d'énergie d'une réserve en eau, minimiser les coûts d'une mesure de protection, etc.

En météorologie, le concept de valeur économique (utilité) des prévisions a été très étudié dans les années 1970-80, suivi par une réapparition de l'intérêt porté à ce sujet à la fin des années 1990. En hydrologie, des études sur l'évaluation économique des prévisions de débits sont plus rares (Roulin, 2007 ; Kim *et al.*, 2007 ; McCollor et Stull, 2008 ; Van de Bergh et Roulin, 2010 ; Muluye, 2011 ; Verkade et Werner, 2011 ; Weijs, 2011 ; Boucher *et al.*, 2012), même si l'utilisation des prévisions d'ensemble pour la prévision des crues et la gestion des ressources en eau est déjà présente depuis une dizaine d'années (Cloke et Pappenberger, 2009). En effet, les études portant sur la qualité des prévisions hydrologiques sont beaucoup plus courantes que celles dédiées à l'utilité (valeur) des prévisions.

Les études déjà réalisées laissent percevoir la valeur économique potentielle des prévisions probabilistes et encouragent les recherches futures sur ce sujet. Néanmoins, elles se limitent à des événements particuliers ou des séries de prévisions-observation courtes, ou encore des aménagements de gestion à caractéristiques fixes. Les forts intérêts économiques liés à la gestion des ressources en eau imposent la réalisation d'études plus approfondies, ainsi que de nouveaux développements pour mieux comprendre comment l'utilisateur des prévisions peut en faire le meilleur usage dans la prise de décision.

Dans ce contexte, nous proposons dans cette thèse notre propre approche d'optimisation de la gestion d'un réservoir. Plus particulièrement, notre objectif est de construire un modèle de gestion qui permette d'examiner l'apport des prévisions, et notamment des prévisions probabilistes d'ensemble, à la gestion optimale d'une réserve en eau superficielle destinée à la production hydroélectrique.

Plus particulièrement, cette thèse s'intéressera aux questions suivantes :

- Quel est le gain apporté en valeur économique par les prévisions d'ensemble par rapport aux prévisions déterministes et climatologiques à la production hydroélectrique ?
- Quelle est la sensibilité du gain économique aux caractéristiques du réservoir ?
- Quel est l'impact de l'amélioration de la qualité des prévisions par des approches de traitement statistiques de biais sur le gain économique ?

1.5 Objectifs de la thèse et structure du mémoire

Cette thèse a deux grands objectifs généraux :

- évaluer des approches empiriques sélectionnées pour garantir une bonne qualité des prévisions pour les applications opérationnelles de la prévision hydrologique d'ensemble,
- développer une méthode simple pour pouvoir estimer la valeur d'une prévision d'ensemble pour le cas de la gestion des apports en eau aux retenues de production hydroélectrique.

Ce mémoire s'organise autour de sept chapitres, décrits sommairement ci-après.

Le *Chapitre 2* présente les données et modèles hydrologiques utilisés : les prévisions météorologiques, les jeux de bassins versants qui ont servi de base à cette étude et leurs séries historiques de données observées. Ces données sont utilisées en entrée de la modélisation hydrologique, effectuée par deux modèles pluie-débit globaux : le modèle *GRP* et le modèle *MORDOR*.

Le *Chapitre 3* examine la qualité des prévisions d'ensemble et présente ses principaux attributs. Nous présentons les différentes mesures de qualité utilisées lors de cette thèse et les principales sources de biais de la prévision d'ensemble. Les différentes approches de pré/post-traitement statistique existantes et celles exploitées lors de cette thèse pour répondre à des questions liées à l'amélioration de la qualité des prévisions sont aussi présentées.

Le *Chapitre 4* présente une autre caractéristique des prévisions d'ensemble : l'utilité, qui est souvent mesurée en termes monétaires (gain ou pertes). Un résumé des techniques les plus courantes est présenté ici, accompagné d'une synthèse des études menées en hydrologie pour évaluer la valeur économique des prévisions.

Dans le *Chapitre 5* nous évaluons la qualité des prévisions tirées de la base TIGGE en utilisant divers critères existants dans la littérature. Nous nous sommes concentrés sur la qualité globale des prévisions, ainsi que sur les attributs de fiabilité et discrimination. Nous évaluons les prévisions d'ensemble de pluie et de débit séparément en examinant leurs performances en fonction de l'horizon de prévision, la surface et la localisation du bassin versant. Différentes approches de combinaison des prévisions de la base TIGGE sont également examinées dans ce chapitre.

Dans le *Chapitre 6* nous utilisons les prévisions météorologiques d'ensemble du centre européen (CEPMMT) et les deux modèles hydrologiques à notre disposition (*MORDOR* et *GRP*) pour examiner différentes approches de traitement statistique appliquées soit sur les prévisions de pluie (pré-), soit sur les prévisions de débit (post-traitement).

Dans le *Chapitre 7* nous nous intéressons à évaluer le gain en termes d'impact économique, apporté par les prévisions, y compris par les prévisions d'ensemble, sur la gestion de la production d'énergie hydroélectrique d'une retenue. Le but est de quantifier la valeur d'une prévision pour la gestion des apports en eau aux barrages-réservoirs. Nous rappelons d'abord les concepts autour de la gestion des apports aux aménagements hydroélectriques. Ensuite, nous décrivons les étapes de construction d'un modèle de gestion heuristique des apports en débit. Enfin, les résultats de l'application du modèle de gestion aux prévisions hydrologiques du modèle *MORDOR* sont présentés et discutés.

Le *Chapitre 8* présente les principales conclusions et les perspectives qui découlent de ce travail de thèse.