

Université Sorbonne Université

Habilitation à Diriger des Recherches (HDR)

Qualité et valeur des prévisions hydrologiques d'ensemble

présentée par

Maria-Helena RAMOS

Chargée de Recherche à Irstea

Soutenue publiquement le 6 avril 2018, devant le jury composé de :

Sandrine Anquetin , Directrice de recherche, CNRS, IGE, Grenoble	Examinatrice
Christophe Cudennec , Professeur AGROCAMPUS OUEST, INRA, Rennes	Rapporteur
Luc Descroix , Directeur de Recherche, IRD, Muséum National d'Histoire Naturelle, Paris	Rapporteur
Agnès Ducharne , Directrice de recherche, CNRS, UPMC, UMR Metis, Paris	Examinatrice
Véronique Ducrocq , Ing. Général des Ponts, des Eaux et des Forêts, CNRM/Météo-France, Toulouse	Rapporteur
Eric Gaume , Ing. en Chef des Ponts, des Eaux et des Forêts, Ifsttar, Nantes	Examineur

Dossier scientifique

Préparé au sein de l'équipe Hydrologie des bassins versants de l'Unité de recherche HYCAR
(Hydrosystèmes Continentaux Anthropisés : Ressources, Risques, Restauration), Irstea (Centre d'Antony)

Institut National de Recherche en Sciences et Technologies pour l'Environnement et l'Agriculture

TABLE DES MATIERES

Résumé	5
Abstract.....	6
Liste des Figures	7
Liste des Tableaux	11
"Les Fleuves m'ont laissé descendre où je voulais" (A. Rimbaud).....	13
1. Introduction générale	15
1.1. Contexte thématique et enjeux.....	15
1.2. Démarche et cadres d'étude	18
1.3. Organisation du mémoire.....	21
2. La prévision intègre un raisonnement probabiliste	23
2.1. Données, modèles et prévisionnistes pour quantifier l'incertitude	23
2.2. La prévision hydrologique opérationnelle en cours de transformation majeure.....	25
2.3. Les prévisions probabilistes conduisent-elles à de meilleures décisions ?.....	30
2.4. Une approche multidisciplinaire inévitable	35
3. Qualité des prévisions	37
3.1. Attributs et critères de la qualité d'une prévision	37
3.2. Ce que les critères peuvent nous indiquer sur la qualité des prévisions	41
3.3. Quantifier les incertitudes prédictives : propagation de plusieurs sources d'erreurs ou post-traitement statistique ?	58
3.4. Le meilleur est peut-être devant nous	67
3.5. Que reste-t-il à faire une fois que les prévisions sont de bonne qualité ?.....	68
4. Liens entre qualité et valeur des prévisions en hydrologie.....	71
4.1. Pourquoi s'intéresser à l'interaction prévisionniste-utilisateur ?.....	71
4.2. Utilité en termes d'anticipation et d'alerte aux crues	73
4.3. Valeur économique relative d'un système de prévision et alerte aux crues.....	77
4.4. Apport des prévisions hydrologiques pour la gestion des réservoirs	79
5. Perspectives	91
Références bibliographiques	95

Résumé

Ce mémoire propose une synthèse des travaux de recherche que j'ai menés et dirigés en prévision hydrologique. Il s'agit majoritairement de travaux menés depuis mon arrivée au sein de l'équipe « Hydrologie de Bassins Versants » à Irstea, Centre d'Antony, en novembre 2007. Ils s'inscrivent pleinement dans le domaine des sciences hydrologiques et, plus particulièrement, dans la thématique de la prévision hydrologique pour l'anticipation des situations à risque (crues/inondations et étiages/sécheresses) et la gestion de la ressource en eau en avenir incertain.

De manière générale, la prévision hydrologique vise à traduire les observations et les prévisions météorologiques en estimations des débits futurs des rivières. Plusieurs techniques et modèles peuvent être employés afin de répondre à une variété d'objectifs d'un gestionnaire de l'eau ou d'un décideur sensible au risque d'un événement sévère, qu'il soit de crue ou d'étiage. Dans mes recherches, je me suis intéressée à la prévision hydrologique d'ensemble.

L'approche de prévision d'ensemble a été développée dans le but d'appréhender et quantifier l'incertitude inhérente à toute prévision. Elle ne cherche pas à fournir une unique prévision déterministe, mais un ensemble de scénarios représentatifs de l'incertitude liée à la prévision. Ces scénarios échantillonnent au mieux les différentes sources d'erreurs pouvant influencer la connaissance de l'évolution future des variables prévues. Ils représentent un système de prévision probabiliste, qui associe une probabilité (ou vraisemblance) à l'occurrence future d'un événement donné, lorsqu'ils sont statistiquement bien calibrés et sous l'hypothèse que le nombre de scénarios est suffisant pour bien représenter la distribution prédictive, c'est-à-dire, l'ensemble de probabilités d'occurrence de l'événement.

Dans ce mémoire, je me focalise sur deux aspects de la prévision hydrologique d'ensemble. Le premier concerne la qualité des prévisions, au vu des nombreux attributs qui qualifient une bonne prévision probabiliste. Le second aspect est la valeur (économique ou d'utilité sociale) des prévisions pour les utilisateurs et décideurs. J'ai cherché à mieux comprendre ces deux aspects afin de pouvoir apporter des éléments de compréhension dans l'usage des prévisions d'ensemble pour la prise de décision et dans la gestion de l'eau et des risques hydrologiques.

La qualité et la valeur d'une prévision hydrologique reposent sur au moins trois piliers essentiels d'une chaîne de prévision hydrométéorologique. Il s'agit (1) des données historiques, observées en temps réel ou prévues de forçage atmosphérique, (2) des modèles avec leurs capacités prédictives, et (3) des prévisionnistes avec leur perception du risque et des imperfections qui se cachent dans les données et les outils qu'ils manipulent. En s'appuyant sur ces trois piliers, la recherche du lien entre qualité et valeur des prévisions apparaît comme un fil conducteur au long de ce mémoire. L'idée est de faire dialoguer plusieurs points de vue qui définiront si une prévision est bonne et/ou utile pour ceux qui l'élaborent et pour ceux qui l'utilisent comme information de base pour prendre une décision.

Cet ensemble devrait nous fournir une approche intégrative de nos savoirs, nos ignorances et nos capacités à communiquer et agir lorsque nous sommes mobilisés par l'exercice de la prévision. Bien que le futur ne deviendra pas moins incertain, il sera, je l'espère, plus éclairé.

Abstract

This thesis presents a synthesis of the research work that I have conducted and directed in hydrological forecasting. Most of the work has been done since I joined the "Catchment Hydrology" team at Irstea, in Antony, in November 2007. It concerns the discipline of hydrological sciences and, in particular, the field of hydrological forecasting for the anticipation of risks (floods and droughts) and the management of water resources under uncertainty.

In general, hydrological forecasting aims to translate observations and weather forecasts into estimates of future river flows. Several techniques and models can be employed to meet a variety of objectives of a water manager or a decision maker concerned with the risk of a severe event, whether it is a flood or a drought. The research reported here focuses on hydrological ensemble forecasts.

The ensemble prediction approach has been developed to capture and quantify the uncertainty inherent in any forecast. It does not seek to provide a single deterministic forecast, but a set of scenarios representative of the uncertainty associated with forecasting. These scenarios would best sample the different sources of errors that may influence the knowledge of the future evolution of the predicted variables. They can represent a probabilistic forecasting system, which associates a probability (or likelihood) with the future occurrence of a given event, if they are statistically well calibrated and under the assumption that the number of scenarios is sufficient to represent the predictive distribution, i.e., the set of probabilities of occurrence of the event.

In this thesis, two aspects of ensemble hydrological forecasting are highlighted. These aspects are the quality of the forecasts, given the many attributes that are used to qualify a good probabilistic forecast, and the value (economic or societal utility) of the forecasts for users and decision-makers. They are issues I have sought to better understand in order to provide insights to the use of ensemble forecasts for decision making and in the management of water resources and hydrological risks.

These two aspects, forecast quality and value of hydrological forecasts, rely on at least three common and essential pillars of a hydrometeorological forecasting chain. These pillars are (1) the historical, the real-time and the atmospheric forcing data used in the forecasting process, (2) the models with their predictive capabilities, and (3) the forecasters with their perceptions of the risk and of the imperfections that hide in the data and the tools they handle. Building on these three pillars, the search for the links between forecast quality and forecast value appears as a common thread throughout this thesis. The idea is to address several viewpoints which define whether a forecast is good and/or useful as basic information to support a decision.

This ensemble should provide us with an integrative approach to our knowledge, our ignorance and our ability to communicate and act when engaged in the exercise of forecasting. Although the future will not become less uncertain, it will, I hope, be more enlightened.

Liste des Figures

FIGURE 1 : SCHEMA DE LA CHAINE DE PREVISION HYDROMETEOROLOGIQUE ET DES THEMATIQUES (CARRÉS EN VERT) QUI ORIENTENT LES DEMARCHES ET CADRES D'ÉTUDE DE MES RECHERCHES EN PREVISION HYDROLOGIQUE	18
FIGURE 2 : SCHEMA DES PRINCIPALES INTERCONNEXIONS DANS UNE CHAINE DE PREVISION HYDROMETEOROLOGIQUE (DANS RAMOS ET AL., 2010A).....	24
FIGURE 3 : UN ENSEMBLE DE PREVISIONS PRODUIT UNE GAMME DE SCENARIOS POSSIBLES PLUTOT QU'UNE SEULE VALEUR PREVUE. LA DISTRIBUTION DES MEMBRES DE L'ENSEMBLE DONNE UNE INDICATION DE LA PROBABILITE D'OCCURRENCE DE CES SCENARIOS (DISTRIBUTION PREDICTIVE) (D'APRES : HTTPS://WWW.ECMWF.INT).....	26
FIGURE 4 : VOULEZ-VOUS JOUER A UN JEU DE PREVISION HYDROLOGIQUE PROBABILISTE ?.....	33
FIGURE 5 : CADRE GENERAL POUR CARACTERISER LA QUALITE DES PREVISIONS SUR LA BASE DE LA RELATION STATISTIQUE ENTRE UNE PREVISION ET L'OBSERVATION QUI LUI EST ASSOCIEE.....	38
FIGURE 6 : PREVISION HYDROLOGIQUE D'ENSEMBLE BASEE SUR LA PREVISION D'ENSEMBLE METEOROLOGIQUE EPS DU CEPMMT (50 MEMBRES) ET LE MODELE GR3P (JOURNALIER) D'IRSTEA : CAS DE L'EVENEMENT DE MARS 2001 SUR LA SEINE A PARIS (DANS : RAMOS ET AL., 2008).....	42
FIGURE 7 : MEILLEURS DONNEURS DES PARAMETRES DU MODELE GR3P PARMIS LES VOISINS (LBN, L1N,...L20N) DU SITE NON JAUGE SELON QUATRE CRITERES DE LA QUALITE DES PREVISIONS HYDROLOGIQUES D'ENSEMBLE (NRMSE, CSI, BRIER 90 ET RPS). LES PARAMETRES SONT TRANSFERES DES SITES VOISINS MAIS LA MISE A JOUR A L'INSTANT DE PREVISION EST LOCALE (DANS : RANDRIANASOLO, 2012)	44
FIGURE 8 : APPORT DE LA SIMULATION CONDITIONNELLE GEOSTATISTIQUE SAMPO TBM POUR LA PREVISION IMMEDIATE PI ET L'ALERTE AUX CRUES RAPIDES (ADAPTE DE : CASERI, 2017)	45
FIGURE 9 : CRITERE CSI (INDICE DE SUCCES CRITIQUE, EN FRANÇAIS) D'ÉVALUATION DE LA QUALITE DES PREVISIONS POUR LES ECHEANCES DE 1 JOUR (GAUCHE), 5 JOURS (MILIEU) ET 9 JOURS (DROITE), APPLIQUE A LA PREVISION DETERMINISTE DES DEBITS (DET, EN VERT) ET LA PREVISION D'ENSEMBLE DE DEBITS DEPASSANT LES QUANTILES Q95% (A GAUCHE DE LA LIGNE POINTILLEE) ET Q75% (A DROITE DE LA LIGNE POINTILLEE) (AU MOINS 80% DES MEMBRES), BASEE SUR LA PREVISION DETERMINISTE DE PLUIE HABILLEE (HABILLAGES A PARTIR DE DEUX MODELES D'ERREUR TESTES, A ET B, EN BLEU ET ROUGE, RESPECTIVEMENT). STATISTIQUES SUR 74 BASSINS VERSANTS. LA MEILLEURE PERFORMANCE EST CELLE DONT LE CRITERE EST PLUS PROCHE DE 1.0	48
FIGURE 10 : CRITERES NRMSE ET CRPSS (SKILL SCORE PAR RAPPORT A UNE PREVISION NAÏVE CONSTITUEE DES VALEURS HISTORIQUES DE DEBITS OBSERVES AVANT LA PERIODE D'ÉVALUATION) POUR LES PREVISIONS HYDROLOGIQUES D'ENSEMBLE A L'ÉCHEANCE DE 2 JOURS DES MODELES GR3P ET SIM, AVEC ET SANS ASSIMILATION DE DONNEES, UTILISANT LA PEARP EN ENTREE. ÉVALUATION SUR 86 BASSINS VERSANTS ET LA PERIODE DU 10/03/2005 AU 31/07/2006. LES MEILLEURES PERFORMANCES SONT CELLES QUI ONT LA PLUS PETITE VALEUR DE NRMSE ET LA PLUS GRANDE VALEUR DE CRPSS (DANS : RANDRIANASOLO ET AL., 2014)	49
FIGURE 11 : PERFORMANCES DU SYSTEME DE PREVISION PEARP-GRP AVEC ET SANS ASSIMILATION DE DONNEES (DA ET NODA) ET AVEC OU SANS POST-TRAITEMENT (PP ET NOPP), SELON UN CRITERE DE MESURE DE LA DISPERSION (OU FINESSE) DE L'ENSEMBLE (DISTANCE INTERQUANTILE MOYENNE NORMALISEE NMIQR, EN HAUT) ET UN CRITERE DE MESURE DE LA FIABILITE DE L'ENSEMBLE (DISTANCE ALPHA DU DIAGRAMME DE PIT A SA DIAGONALE, EN BAS). POUR LE PREMIER CRITERE, LE MEILLEUR SCORE EST LA PLUS PETITE VALEUR ; POUR LE DEUXIEME, LA VALEUR PLUS PROCHE DE L'UNITE. ÉVALUATION DES DEBITS PREVUS SUR 212 BASSINS VERSANTS ET QUATRE ECHEANCES HORAIRES (6H, 12H, 24H, 48H) (DANS : BOURGIN ET AL., 2014).....	51
FIGURE 12 : QUELQUES CARACTERISTIQUES TECHNIQUES DE LA BASE TIGGE (D'APRES E. ZSOTER, « THE TIGGE ARCHIVE AND ITS APPLICATIONS », JUIN 2014).....	52
FIGURE 13 : PERFORMANCE DES PREVISIONS SAISONNIERES BRUTES (SANS POST-TRAITEMENT) DES PRECIPITATIONS DE BASSIN (SYS4 CEPMMT) ET DE DEBITS (SYS4 ET MODELE HYDROLOGIQUE GR6J), SELON LES CRITERES IQRSS (INTERQUANTILE RANGE), QUI MESURE LA DISPERSION DES MEMBRES, ET CRPSS, QUI MESURE LA PRECISION ET LA FIABILITE, ET POUR LES SAISONS DE DECEMBRE A FEVRIER (DJF ; SAISON HUMIDE) ET JUIN A AOUT (JJA ; SAISON SECHE). CHAQUE LIGNE REPRESENTA UN BASSIN VERSANT. LES « SKILL SCORES » ONT POUR REFERENCE L'HISTORIQUE DES PRECIPITATIONS OBSERVES (POUR L'ÉVALUATION DES PRECIPITATIONS ; A GAUCHE) ET LA METHODE ESP (POUR L'ÉVALUATION DES DEBITS ; A DROITE). UNE PREVISION EST MEILLEURE QUE SA REFERENCE SI LE « SKILL SCORE » EST SUPERIEURE A ZERO (DANS : CROCHEMORE ET AL., 2016B)	57
FIGURE 14 : BIAIS DES PREVISIONS SAISONNIERES D'ENSEMBLE DU CEPMMT (SYSTEM 4) POUR QUATRE BASSINS VERSANTS (#2, #4, #7 ET #14), ÉVALUES SUR LA PERIODE 1981-2010 ET POUR LES ECHEANCES ENTRE 31 ET 60 JOURS. LE BIAIS EST INDIQUE POUR TOUTE L'ANNÉE (PREMIERE LIGNE) ET POUR CHAQUE MOIS CIVIL (LIGNES SUIVANTES) (DANS : CROCHEMORE ET AL., 2016B)	57
FIGURE 15 : QUALITE DES PREVISIONS HYDROLOGIQUES D'ENSEMBLE POUR LES CAS OU (2) LES PREVISIONS METEOROLOGIQUES SONT PRETRAITEES, (3) LES PREVISIONS HYDROLOGIQUES SONT POST-TRAITEES, ET (4) LES PREVISIONS METEOROLOGIQUES SONT PRETRAITEES ET LES PREVISIONS HYDROLOGIQUES SONT POST-TRAITEES. LES CRITERES BS ET RPS SONT CALCULES PAR RAPPORT AUX PREVISIONS BRUTES (1). ÉVALUATION SUR 48 MOIS (2005-2008), 11 BASSINS VERSANTS ET POUR TROIS ECHEANCES : 3, 5 ET 7 JOURS (DANS : ZALACHORI ET AL., 2012)	60

FIGURE 16 : BIAIS DES PREVISIONS SAISONNIERES DE PRECIPITATIONS DU CEPMMT (SYSTEM 4) POUR QUATRE BASSINS VERSANTS (#2, #4, #7 ET #14), EVALUES SUR LA PERIODE 1981-2010 ET POUR LES ECHEANCES ENTRE 31 ET 60 JOURS, POUR HUIT METHODES DE CORRECTION DE BIAIS APPLIQUEES AUX PREVISIONS BRUTES (VOIR TABLEAU EN DESSOUS). LE BIAIS EST INDIQUE POUR TOUTE L'ANNEE (PREMIERE LIGNE) ET POUR CHAQUE MOIS CIVIL (LIGNES SUIVANTES). LE BIAIS ORIGINAL DES PREVISIONS BRUTES EST MONTRE DANS LA FIGURE 14 (DANS : CROCHEMORE ET AL., 2016B).....	62
FIGURE 17 : CRITERES DE QUALITE DES PREVISIONS DE PRECIPITATION (P) ET DE DEBITS (Q) APRES APPLICATION D'UNE CORRECTION DE BIAIS AUX PRECIPITATIONS (EDMD-M, FIG. 16). LES PREVISIONS BRUTES DE PRECIPITATIONS ET DE DEBITS SONT UTILISEES COMME REFERENCE. LES VALEURS DES CRITERES SONT MOYENNEES SUR LES ECHEANCES DE 10 A 90 JOURS. LES RESULTATS SONT PRESENTES POUR LES QUATRE SAISONS DE L'ANNEE ET QUATRE BASSINS VERSANTS. LES CRITERES MESURENT LA PERFORMANCE GLOBALE (CRPSS), LA FINESSE (IQRSS), ET LA FIABILITE (PITSS) DES PREVISIONS (DANS : CROCHEMORE ET AL., 2016B)	62
FIGURE 18 : PERFORMANCE DES PREVISIONS DE DEBIT (APPROCHE ESP CONDITIONNEE AU SPI3) PAR RAPPORT AUX PERFORMANCES DES PREVISIONS DE DEBIT BASEES SUR L'APPROCHE ESP TRADITIONNELLE (TRIANGLES EN GRIS) ET LES PREVISIONS SAISONNIERES DE PRECIPITATION CEPMMT SYSTEM 4 (RONDS EN ROUGE), SELON L'ECHEANCES DE 10, 30 ET 90 JOURS. LES CRITERES MESURENT LA PERFORMANCE EN FINESSE (IQRSS) ET FIABILITE (PITSS) (DANS : CROCHEMORE ET AL., 2017)	64
FIGURE 19 : GRAPHIQUES DU RISQUE D'AVOIR LE NOMBRE DE JOURS (Y-AXIS) ET LE VOLUME (X-AXIS) SOUS LE SEUIL DE BASSES EAUX Q80 (20% DES VALEURS HISTORIQUES SONT AU-DESSOUS DE CE SEUIL) POUR LES PREVISIONS SAISONNIERES DE DEBIT BASEES SUR LES PREVISIONS SAISONNIERES DE PRECIPITATION DU CEPMMT (SYST4), L'APPROCHE ESP CONDITIONNEE (ESP_SPI3) ET L'APPROCHE ESP TRADITIONNELLE (ESP). LE CAS ILLUSTRE EST CELUI DE LA SECHERESSE DE 2003 DANS LE BASSIN DE L'AZERGUES A LOZANNE ET D'UNE PREVISION EMISE LE 1ER JUILLET POUR LES 90 JOURS SUIVANTS (DANS : CROCHEMORE ET AL., 2017).....	65
FIGURE 20 : TESTS EFFECTUES POUR LA PRISE EN COMPTE EXPLICITE DES DIFFERENTES SOURCES D'INCERTITUDE ET LEUR PROPAGATION DANS LA CHAINE DE PREVISION DE DEBIT (DANS : VAN PHAM, 2011)	67
FIGURE 21 : DECOPAGE DE L'ECHELLE TEMPORELLE DE LA CHAINE DE PREVISION EN DELAI DE PREVISION ET TEMPS D'ATTENUATION DES DOMMAGES AVANT L'ARRIVEE D'UN EVENEMENT.	73
FIGURE 22 : REPRESENTATION SPATIALISEE DE LA PROBABILITE DE DEPASSEMENT DU DEBIT HORAIRE DE PERIODE DE RETOUR DE 10 ANS DANS LE DEPARTEMENT DU VAR POUR LE 05 DE NOVEMBRE 2011 A 20H, GENEREE A PARTIR DES DEBITS SIMULES AIGA : A GAUCHE, EN UTILISANT LES PREVISIONS D'ENSEMBLE SAMPO-TBM (METHODE PI-GAE DE PREVISION IMMEDIATE) POUR L'ECHEANCE DE 2 HEURES ; A DROITE, EN UTILISANT LA PLUIE OBSERVEE PAR LE RADAR ANTILOPE (DANS : CASERI, 2017)	74
FIGURE 23 : GAIN OU PERTE EN DELAI D'ANTICIPATION MOYEN DE LA PREVISION D'ENSEMBLE PAR RAPPORT A LA PREVISION DETERMINISTE (EN JOURS) POUR LE DEPASSEMENT DU QUANTILE 99% DES DEBITS JOURNALIERS OBSERVES POUR 208 BASSINS VERSANTS EN FRANCE LORSQU'UN SEUIL SPECIFIQUE AU BASSIN VERSANT EST APPLIQUE (DANS : WEEINK, 2010)	76
FIGURE 24 : VALEUR ECONOMIQUE POTENTIELLE DE PREVISIONS PROBABILISTES SYNTHETIQUES DE PRECIPITATIONS SUPERIEURES A 10 MM/J POUR DIFFERENTS SEUILS DE PROBABILITE (LIGNES GRISES). LES RESULTATS POUR LES SEUILS DE PROBABILITE DE 10% (LIGNE ROUGE) ET 60% (LIGNE VERTE) SONT SOULIGNES. LA VALEUR ECONOMIQUE POTENTIELLE DE L'ENSEMBLE EST DEFINIE PAR L'ENVELOPPE DE TOUTES LES COURBES (LIGNE BLEUE). A GAUCHE : ENSEMBLE FIABLE ; A DROITE : ENSEMBLE BIAISE (D'APRES : BUIZZA ET AL., 2007)	78
FIGURE 25 : RELATIONS ENTRE REV ET L'ERREUR ABSOLUE MOYENNE (MAE ; A GAUCHE) ET LA FIABILITE (RATIO ISSU DU DIAGRAMME DE RANG ; A DROITE), CALCULEE POUR UN CLR = 0,5. LES CERCLES INDIQUENT L'ECHEANCE 1 JOUR ET LES ASTERISQUES LES ECHEANCES JOURNALIERES SUIVANTES. LES SYSTEMES EVALUES COMPRENNENT DIFFERENTES SOURCES D'INCERTITUDE (TABLEAU EN BAS DE FIGURE) (DANS : THIBOULT ET AL., 2017)	79
FIGURE 26 : LORS DE PERIODES PROLONGEES DE FAIBLE DEBIT, DES CONFLITS D'USAGE PEUVENT ETRE OBSERVES DANS LA GESTION DU BARRAGE-RESERVOIR D'ARZAL, A L'EMBOUCHURE DU BASSIN DE LA VILAINE (SOURCE : HTTPS://WWW.OUEST-FRANCE.FR/)	81
FIGURE 27 : BARRAGE D'ARZAL - UN OUVRAGE AUX USAGES MULTIPLES (POSTER L. CROCHEMORE).....	81
FIGURE 28 : GRAPHIQUES DE RISQUE PROPOSES POUR REPRESENTER LA PROBABILITE D'ETRE EN DESSOUS DU NIVEAU DE RESERVOIR MINIMUM ACCEPTABLE DE 1,80 M NGF POUR DES DUREES DIFFERENTES (0 JOUR JUSQU'A 25-31 JOURS). LA LIGNE EN DESSOUS (« END OF TENSIONS ») CORRESPOND A LA PROBABILITE DE REVENIR A UN NIVEAU SUPERIEUR A 1,80 M NGF ET POUVOIR REVENIR A UNE GESTION HORS CONFLITS D'USAGE. LES GRAPHIQUES SONT PROPOSES POUR SIX ANNEES DIFFERENTES (2005 A 2010) ET DES DATES DE DEBUT DE PREVISION ALLANT DU 1ER MAI (A) AU 1ER AOUT (D) ; PREVISIONS VALABLES POUR LES HORIZONS DE PREVISION MENSUELS NOTES EN BAS DES GRAPHIQUES ET COMPTES A PARTIR DE LA DATE DE DEBUT DE PREVISION. CLASSEMENT DES ANNEES LES PLUS SECHES AUX ANNEES LES PLUS HUMIDES : 2005, 2010, 2009, 2006, 2007, 2008 (DANS : CROCHEMORE, 2016).....	83
FIGURE 29 : CLASSIFICATION GENERALE DES AMENAGEMENTS HYDROELECTRIQUES EN FONCTION DU COEFFICIENT DE TURBINAGE ET DE LA CAPACITE UTILE (EN JOURS D'APPORTS MOYEN) (SOURCE DES DONNEES : EDF ; DANS : ZALACHORI, 2013).....	85
FIGURE 30 : SCHEMA DE LA PROCEDURE DEVELOPPEE POUR L'OPTIMISATION HEURISTIQUE D'UNE RESERVE HYDROELECTRIQUE (DANS : POSTER PRESENTE A L'EGU EN 2017, M. CASSAGNOLE ET AL.)	86
FIGURE 31 : GAIN DE GESTION EN FONCTION DU COEFFICIENT DE CAPACITE DU RESERVOIR HYDROELECTRIQUE (AXE DES ABSCISSES, EXPRIME EN FACTEUR MULTIPLICATIF DU DEBIT MOYEN JOURNALIER) POUR DES PREVISIONS HYDROLOGIQUES D'ENSEMBLE BRUTES (PE), CLIMATOLOGIQUES (BASEES SUR L'HISTORIQUE DES DEBITS) (PC), PARFAITES (BASEES SUR L'OBSERVATION DES DEBITS) (PP), LA	

<i>MOYENNE D'ENSEMBLE (PME) ET LA PREVISION DETERMINISTE (PD). RESULTATS BASES SUR LES APPORTS EN DEBITS AU BASSIN VERSANT DE LA DURANCE A SERRE PONÇON ET UNE PERIODE DE PREVISION JOURNALIERE ALLANT DE 2005 A 2008. LE COEFFICIENT DE TURBINAGE DU MODELE DE GESTION EST FIXE A 4 FOIS LE DEBIT MOYEN JOURNALIER (DANS : ZALACHORI, 2013)</i>	<i>87</i>
<i>FIGURE 32 : EXEMPLE DE LA QUALITE DE QUATRE PREVISIONS SYNTHETIQUES D'ENSEMBLE DE DEBITS CREEES, TELLE QU'ÉVALUEE PAR LE DIAGRAMME DE PIT (ECHEANCE : 7 JOURS) (DANS : CASSAGNOLE ET AL., POSTER PRESENTE A L'EGU EN 2017)</i>	<i>88</i>
<i>FIGURE 33 : GAIN DES ENSEMBLES SYNTHETIQUES ILLUSTRÉS DANS LA FIGURE 32 (% PAR RAPPORT AU GAIN ECONOMIQUE OBTENU AVEC PREVISIONS DE DEBITS EGALES AUX OBSERVATIONS) (DANS : CASSAGNOLE ET AL., POSTER PRESENTE A L'EGU EN 2017)</i>	<i>89</i>

Liste des Tableaux

<i>TABLEAU 1 : SYNTHÈSE DES PRINCIPALES ÉTUDES QUE J'AI MENEES OU DIRIGÉES POUR ABORDER LES DIFFÉRENTES THÉMATIQUES DE LA CHAÎNE DE PRÉVISION HYDROMÉTÉOROLOGIQUE</i>	<i>19</i>
<i>TABLEAU 2 : QUELQUES CARACTÉRISTIQUES PRINCIPALES DES CRITÈRES D'ÉVALUATION COURAMMENT UTILISÉS POUR ÉVALUER LA QUALITÉ DES PRÉVISIONS HYDROMÉTÉOROLOGIQUES (DANS : ANCTIL ET RAMOS, 2018)</i>	<i>40</i>
<i>TABLEAU 3 : PHASES DANS LE DÉVELOPPEMENT DE LA GESTION (COMMUNICATION) DES RISQUES (TRADUIT D'APRÈS : FISCHHOFF, 1995)</i>	<i>72</i>
<i>TABLEAU 4 : TABLEAU DE CONTINGENCE AVEC LES COÛTS ASSOCIÉS À CHAQUE TYPE D'ÉVÈNEMENT (DANS : THIBOULT ET AL., 2017)</i>	<i>77</i>

"Les Fleuves m'ont laissé descendre où je voulais" (A. Rimbaud)¹

En remontant le cours du temps, je ne saurais pas trouver la raison pour laquelle j'ai voulu suivre la formation Génie Civil de l'école d'ingénieurs de l'Université Fédérale de Minas Gerais au Brésil en 1989. Pourtant, ce qui m'a poussée à ne pas la quitter avant son terme a été le premier cours de la discipline de Mécanique de Fluides, proposé par le professeur Marcus Vianna. Ce tout premier contact avec le domaine de l'eau a été marqué par une série de photos d'aménagements et ouvrages fluviaux, commentées avec beaucoup de richesse technique et d'enthousiasme professionnel. Par la suite, les cours d'Hydraulique (Pr. Márcio Baptista) et d'Hydrologie (Pr. Mario Cicareli) ont consolidé mon choix. A cette époque, j'avais été beaucoup aidée par la motivation et la patience de mes enseignants pour éclairer également ma curiosité concernant les différents champs d'application dans le domaine de l'eau, qu'ils soient académiques, scientifiques ou opérationnels, parfois même en dehors de leurs heures de travail. Ces premiers pas m'ont conduit à m'engager dans un master (dirigé par Márcio Baptista, UFMG, Brésil), suivi d'un doctorat (dirigé par Jean-Dominique Creutin, LTHE, Grenoble) et deux post-doctorats (accompagnés par Etienne Leblois à Irstea, Lyon, et Jutta Thielen au JRC à Ispra, Italie, respectivement). Ces expériences ont également été riches de rencontres, avec des passionnés et passionnées de l'hydrologie qui ont su me convaincre que je pouvais naviguer dans les mêmes eaux.

Mon parcours de recherche a été enrichi de réflexions et études portant sur les aspects urbanistiques, légaux et méthodologiques en hydrologie urbaine, la variabilité spatio-temporelle de l'aléa pluviométrique et sa sévérité, la prévision des débits de rivières et les interactions homme-environnement-eau dans la gouvernance des risques. J'ai pu confronter les résultats des modèles à l'évidence empirique reportée, au gré de la nature, dans les observations hydrométéorologiques et m'interroger sur la valeur sociétale et économique d'une meilleure connaissance des aléas et une meilleure gestion des risques hydrométéorologiques.

J'ai rencontré la prévision hydrologique pour la première fois lors de mon deuxième post-doctorat. Depuis, elle ne m'a jamais quittée. Aujourd'hui, elle est le point central de ce mémoire, qui a pour but de retracer quelques-unes des voies que j'ai empruntées pour relier mon enthousiasme et ma motivation aux défis scientifiques qu'il me semblait nécessaire de relever dans le domaine de la qualité et de la valeur des prévisions hydrologiques.

Ce mémoire ne prétend pas être exhaustif et, loin de vouloir reporter en détail des méthodes et résultats, j'ai surtout essayé de partager une expérience. J'ai essayé de projeter sur ce texte, avec tout le recul possible, un regard qui, avant tout, ne prétend pas s'imposer ou, encore moins, clôturer des sentiers. Au vu de l'importance que cette thématique a prise dans mon parcours professionnel, j'ai choisi de me laisser guider par une « démangeaison scientifique » qui me donne envie de mieux comprendre le lien entre qualité et valeur d'une prévision hydrologique. Il s'agit d'une question qui n'est pas encore exhaustivement abordée dans la littérature et encore moins parfaitement résolue. Le sujet prête ainsi à des questionnements scientifiques et opérationnels intéressants et d'actualité.

¹ *Le bateau ivre*, Arthur RIMBAUD (1854-1891)

1. Introduction générale

1.1. Contexte thématique et enjeux

Mes travaux de recherche s'intéressent aux situations hydrométéorologiques à risque (précipitations sévères/extrêmes, crues/inondations, étiages/sécheresses) dans le contexte de la prévision des aléas et de la réduction des risques liés à l'eau. Ils se développent en explorant plus particulièrement l'interface entre météorologie/climatologie et hydrologie, et la quantification et propagation des incertitudes, jusqu'à l'utilisateur final, au sein des chaînes de prévisions hydrométéorologiques.

Dans ce contexte, je m'intéresse aux différentes sources d'incertitude des débits prévus, ainsi qu'au développement de prévisions probabilistes fiables et utiles aux gestionnaires/décideurs exposés aux risques hydrométéorologiques à différentes échelles. Le but est de mieux comprendre le processus de prise de décision concernant des situations de mise en alerte de la population et de gestion des ressources hydriques, afin de pouvoir développer et exploiter au mieux les techniques de modélisation et de communication des prévisions les plus appropriées.

Tout au long de mon parcours, depuis ma thèse de doctorat (Ramos, 2002) jusqu'à aujourd'hui, mes thématiques de recherche se sont structurées progressivement autour de :

- l'analyse de la structure (formation et développement) des forçages climatiques des systèmes hydrologiques (notamment les précipitations) ;
- la caractérisation des échelles et structures spatio-temporelles en jeu dans les systèmes précipitants et leur impact dans l'hydrologie des bassins versants ;
- l'évaluation des prévisions de précipitations et de débits, leur prédictibilité et la gestion des risques associés ;
- l'amélioration de ces prévisions par le post-traitement statistique des sorties de modèles et la quantification des incertitudes associées ;
- les défis liés à l'usage opérationnel de scénarios futurs du risque hydrométéorologique et de la disponibilité de la ressource en eau, et à la valeur des prévisions.

D'une part, le choix de mener des recherches en hydrologie en intégrant des éléments, produits et concepts innovants du domaine de la météorologie a été motivé par l'envie d'apporter une vision plus intégrée, multidisciplinaire et interactive des systèmes hydrologiques, et d'aller au-delà de la démarche plus traditionnelle de simple utilisation des forçages météorologiques en entrée aux modèles pluie-débit. Il s'agit de développer des voies d'intégration plus efficaces de l'information météorologique, capables de mieux valoriser cette information dans l'analyse de situations à risque hydrologique. Apporter une vision « hydrologique » aux problématiques de prédictibilité et de caractérisation de phénomènes météorologiques sévères m'a paru essentiel pour un meilleur dialogue entre sciences atmosphériques et hydrologiques au service des sociétés.

D'autre part, je me suis progressivement intéressée aux aspects économiques et de gouvernance liés à la prévision et à la gestion de l'eau. Dans les deux cas, il s'agit de mieux expliciter les liens entre nos capacités à bien modéliser les phénomènes hydrologiques et prévoir les événements futurs, et nos outils/instruments de gestion de l'eau dans les bassins

versants. Le but est, d'une part, de mieux lier la qualité et la valeur (économique ou en termes d'utilité des services) des prévisions, et, d'autre part, de mieux comprendre comment les éléments de la gouvernance de l'eau (acteurs, échelles, instruments, responsabilités, etc.) peuvent jouer un rôle dans la mise en place de mesures adaptées aux situations extrêmes hydrométéorologiques, et ce face aux défis actuels et aux besoins d'adapter nos modes de gestion aux scénarios de changement climatique.

Ces douze dernières années, je me suis particulièrement intéressée à la qualité des prévisions hydrométéorologiques, au gain apporté par les prévisions en termes d'anticipation des situations à risque et à la valeur économique que les informations, produits et services de prévision peuvent apporter à un gestionnaire de l'eau (par ex., producteur d'hydroélectricité, gestionnaire de réservoirs d'approvisionnement en eau potable). Ces sujets présentent des enjeux sociaux et économiques forts.

D'une part, le risque d'inondation par débordement de cours d'eau concerne 17,1 millions de résidents permanents en France, dont 16,8 en métropole (évaluations réalisées en 2011 par les services de l'État² ; Lang *et al.*, 2013). Parmi les 113 événements naturels à l'origine de plus de 10 morts ou de plus de 30 millions d'euros de dommages matériels survenus en France entre 1950 et 2014, 62 % correspondent à des inondations, « dont environ la moitié s'est produite au cours de la dernière décennie. »³ Selon les données de l'EPRI (SOeS)², « le coût moyen annuel lié aux dommages causés par les inondations en France, pris en charge par le régime de solidarité nationale catastrophe naturelle, est estimé à environ 400 millions d'euros par an ». Ce coût n'intègre néanmoins pas des événements majeurs survenus pendant la période plus récente. Par exemple, les inondations engendrées par l'événement exceptionnel qui a touché les bassins de la Seine et de la Loire du 25 mai au 6 juin 2016 pourraient être les plus coûteuses en dégâts assurés depuis 1982, année de création du régime d'indemnisation des catastrophes naturelles, avec 1,2 milliard d'euros de dégâts indemnisés et plus de 2 000 communes touchées⁴. Sur le globe, on estime qu'environ un milliard de personnes vivent dans des zones inondables (Di Baldassarre *et al.*, 2013), et des recherches récentes montrent une corrélation positive entre le réchauffement de la planète et une intensification du cycle hydrologique au niveau mondial, avec impact attendu sur le risque d'inondation (Alfieri *et al.*, 2017).

D'autre part, les sécheresses météorologiques, vagues de chaleur et étiages sévères sont également des phénomènes qui peuvent gravement affecter les populations et plusieurs secteurs économiques. Les événements les plus meurtriers en France correspondent aux deux vagues de chaleur de 2003 et 2006 (avec 17 065 décès, elles totalisent 94 % de l'ensemble des décès provoqués par des événements naturels très graves depuis 1950, selon les *Chiffres clés de l'environnement*³). L'épisode de canicule de 2003 a aussi engendré un épisode de sécheresse hydrologique qui a impacté plusieurs secteurs économiques ainsi que l'environnement. Selon le

² Données : EPRI (SOeS) - 2011 / Source : *Première évaluation nationale des risques d'inondation*, Ministère chargé de l'environnement, 2012. Dernière consultation le 27/12/2017 : https://www.ecologie-solidaire.gouv.fr/sites/default/files/EPRI-Principaux-resultats_120712.pdf

³ *Chiffres clés de l'environnement* – Édition 2016, 72 p. Service de l'observation et des statistiques (SOeS) http://www.eaufrance.fr/IMG/pdf/chiffrescleenvironnement_ed2016_201702.pdf (consulté le 27/12/2017)

⁴ CCR, Caisse Centrale de Réassurance, *Inondations de mai-juin 2016 en France - Modélisation de l'aléa et des dommages*. Dernière consultation le 27/12/2017 : https://www.ccr.fr/documents/23509/29230/Inondations+de+Seine+et+Loire+mai+2016_version+13072016.pdf

bulletin de l'UNEP de mars 2004⁵, la chute de la production céréalière dans l'Union européenne a atteint plus de 23 millions de tonnes en 2003, par rapport à 2002. En France, 75% des départements ont été soumis à des restrictions d'eau. La hausse des températures des cours d'eau et les bas débits observés ont impacté le refroidissement adéquat des rejets des centrales nucléaires de production d'électricité, amenant EDF et ses filiales à une « gestion délicate de l'équilibre entre une consommation en hausse et une production limitée par la sécheresse et la chaleur »⁶, et à des arrêtés gouvernementaux autorisant EDF à rejeter des eaux plus chaudes qu'en temps normal dans les cours d'eau. Au total, l'INSEE a estimé que l'événement de canicule de 2003 « avait amputé la croissance de 0,1 à 0,2 point de PIB sur l'ensemble de l'année, ce qui représente de 15 à 30 milliards d'euros. »⁷

Probablement en raison de leurs durées et de l'extension spatiale de leurs impacts, les épisodes de sécheresse ont souvent un impact direct sur les politiques publiques liées à la gestion du risque. Ils mettent en lumière la complexité des relations entre décideurs et usagers de l'eau dans la gouvernance des sécheresses (La Jeunesse *et al.*, 2015). Dans le cadre du projet Interreg DROP, auquel j'ai participé en tant que coordinatrice au niveau d'Irstea (2013-2015), nous avons mis en lumière plusieurs cas illustratifs, où des aides financières ou plans nationaux sont intervenus à la suite de ces aléas hydro-climatiques extrêmes (par ex., la Circulaire du 30 mars 2004 établissant le « Plan d'Action Sécheresse », qui visait une meilleure coordination des restrictions d'usage de l'eau à l'échelle des bassins versants et une anticipation de la crise par l'élaboration d'arrêtés cadre de restriction des usages, suite à la sécheresse de 2003).

De même, les événements majeurs de crues peuvent déclencher des actions politiques, comme cela a été le cas lors de la réorganisation du système de prévision en France en 2003 avec la création du Service central d'hydrométéorologie et d'appui à la prévision des inondations (SCHAPI), à la suite d'une série d'inondations catastrophiques, dont celle des 8-9 septembre 2002 dans le Gard (Delrieu *et al.*, 2005) et celle du Rhône en 2003.

Dans les recherches s'intéressant aux événements hydrologiques sévères, il importe ainsi de comprendre également les relations entre acteurs, leurs implications et influences dans les processus de prise de décision, les stratégies (techniques ou institutionnelles) proposées pour répondre à d'éventuels problèmes engendrés par ces événements, la répartition des responsabilités et des ressources mobilisées, et les processus par lesquels les politiques et les actions proposées sont mises en œuvre. Ces aspects se trouvent, en effet, être révélateurs d'un mode de gestion de l'eau et de réponse aux situations de crise, et permettent de relier les événements hydrométéorologiques à des réponses techniques, sociales et politiques (La Jeunesse *et al.*, 2016 ; Ruin *et al.*, 2014).

⁵ Environment Alert Bulletin, United Nations Environment Programme (UNEP), *Impacts of summer 2003 heat wave in Europe*. Dernière consultation le 27/12/2017: http://www.unisdr.org/files/1145_ewheatwave.en.pdf

⁶ RTE, *Bilan 2003 de la sûreté du système électrique français*, juillet 2004. Dernière consultation le 27/12/2017 : http://clients.rte-france.com/htm/fr/mediatheque/telecharge/bilan_surete_2003_externe.pdf

⁷ Une chronologie intéressante des événements et impacts de la canicule de 2003 est présentée dans le Rapport d'information n° 195 (2003-2004) de Mme Valérie LÉTARD, MM. Hilaire FLANDRE et Serge LEPELTIER, fait au nom de la mission commune d'information, déposé le 3 février 2004 au Sénat, *La France et les Français face à la canicule : les leçons d'une crise* ; <https://www.senat.fr/rap/r03-195/r03-19511.html>. Pour les conséquences sur le secteur énergétique, le lien direct est ici : <https://www.senat.fr/rap/r03-195/r03-19511.html#toc409> (Dernière consultation le 27/12/2017).

1.2. Démarche et cadres d'étude

Depuis plus d'une dizaine d'années, la prévision hydrologique est au cœur de mes recherches. Plus particulièrement, je m'intéresse à examiner l'apport de la prévision météorologique d'ensemble (à scénarios multiples, physiquement cohérents dans l'espace et dans le temps) à l'hydrologie. Je m'intéresse également à la prise en compte de l'approche d'ensemble dans les systèmes opérationnels de prévision hydrologique, en considérant l'impact de cette démarche sur l'expertise des prévisionnistes et la prise de décision des utilisateurs de ces prévisions.

En parcourant la chaîne de prévision hydrométéorologique (**Figure 1**), nous nous retrouvons avec plusieurs étapes qui comportent en elles-mêmes plusieurs voies de recherche. De manière générale, elles regroupent les questions qui se posent autour de la description et la quantification des incertitudes dans la prévision hydrométéorologique, de l'évaluation de la qualité des prévisions, de la communication efficace des incertitudes prédictives, et, enfin, de l'usage, l'utilité et la valeur économique des prévisions dans un contexte décisionnel.

Ma démarche scientifique a suivi un processus itératif et interactif, consistant à (1) extraire des prévisions météorologiques l'information pertinente pour la prévision hydrologique, (2) évaluer la qualité des prévisions sur la base de retours d'expérience et de simulations d'événements passés, (3) comprendre le rôle des différents attributs qui qualifient une bonne prévision dans leur capacité à informer si un système de prévision répond à ces objectifs, (4) évaluer les techniques d'assimilation de données et de traitement statistique de correction de biais au regard de leur capacité à jouer un rôle dans l'amélioration de la qualité des prévisions, et (5) comprendre l'usage des prévisions d'ensemble ou probabilistes pour soutenir un processus de décision dans le domaine de la gestion de la ressource en eau et des risques liés à l'eau.

Le **Tableau 1** synthétise les principales études que j'ai menées ou dirigées, en les plaçant dans les différentes thématiques qui émergent de la chaîne de prévision de la **Figure 1**.

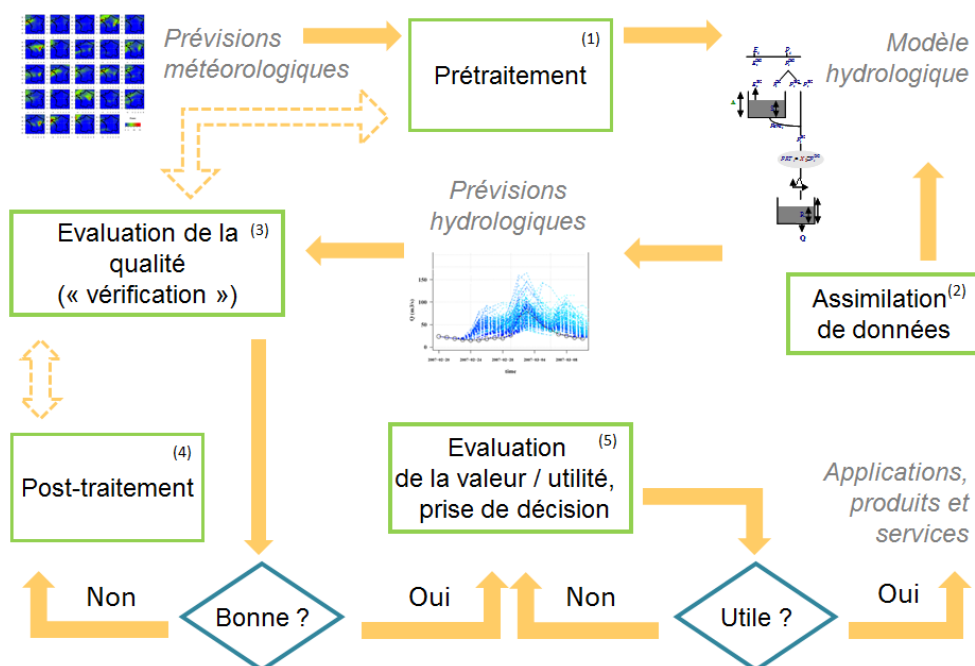


Figure 1 : Schéma de la chaîne de prévision hydrométéorologique et des thématiques (carrés en vert) qui orientent les démarches et cadres d'étude de mes recherches en prévision hydrologique

Tableau 1 : Synthèse des principales études que j'ai menées ou dirigées pour aborder les différentes thématiques de la chaîne de prévision hydrométéorologique

Etudes et articles *	Prévision météorologique	Modèle hydrologique	Cadre d'étude, application	Thématique principale dans la chaîne de prévision **				
				1	2	3	4	5
Thielen <i>et al.</i> , 2009 Bartholmes <i>et al.</i> , 2009 (Projet EFAS)	ECMWF DWD	LISFLOOD	Anticipation de la mise en vigilance (échéances > 3 jours)			x		x
Ramos <i>et al.</i> , 2008 Younis <i>et al.</i> , 2008 Kalas <i>et al.</i> , 2008 Ramos <i>et al.</i> , 2009 Ramos <i>et al.</i> , 2017 Master C. Vrignaud (Convention SCHAPI/SPC)	ECMWF DWD BP AROME PI (événements)	LISFLOOD GRP (opérationnel)	Evaluation d'événements de crues divers en France et en Europe (divers bassins versants)			x		
Master A. Randrianasolo Randrianasolo <i>et al.</i> , 2010, 2014 (Convention SCHAPI/SPC)	PEARP (2005 – 2009)	GR3P (J) SIM	Alerte aux crues à courte échéance (82 à 211 BVs)		x	x		
Master P. Nicolle Nicolle <i>et al.</i> , 2012	PEARP (2005 – 2009)	GR3P (J) GR5P (avec CemaNeige)	Crues nivales (176 BVs)			x		
Master W. Weeink (Convention SCHAPI/SPC)	DET et ENS ECMWF (mars 2005 – août 2006)	GR3P (J)	Etude de seuils de décision pour l'alerte aux crues (29 à 208 BVs)			x		x
Master T. van Pham (Projet FAST)	ENS ECMWF PEARP (2005 – 2009)	GR3P (J)	Propagation des incertitudes (3 BVs)		x	x	x	
Master H. Dhouioui (Convention SCHAPI/SPC)	DET et ENS ECMWF (mars 2005 – octobre 2006)	GR3P (J)	Habillage des prévisions déterministes pour la prévision de crues (74 BVs)	x				
Doctorat A. Randrianasolo Randrianasolo <i>et al.</i> , 2011 (Thèse ED GRNE)	PEARP (2005 – 2009)	GR3P (J) GR4J	Prévision de crues à courte échéance sur les bassins non-jaugés (211 à 756)		x	x		
Doctorat I. Zalachori Zalachori <i>et al.</i> , 2012 (Collab. EDF)	DET et ENS ECMWF TIGGE (2005 – 2008)	GR3P (J) MORDOR	Réservoirs hydroélectriques et lien qualité-valeur ; moyenne-échéance (11 à 74 BVs)	x		x	x	x

Tableau 1 – (cont.)

Doctorat L. Crochemore Crochemore <i>et al.</i> , 2016a,b, 2017 (Projet DROP ; Collab. EPTB IAV)	ECMWF System 4 (1981 – 2010)	GR6J	Prévisions saisonnières, apports au réservoir d'Arzal (eau potable) (17 BVs)	x		x			x
Doctorat A. Caseri Caseri <i>et al.</i> , 2016a,b) (Projet RHYTMME ; Collab. TR ARCEAU)	Simulations stochastiques basées sur radar et pluviomètres (17 événements 2009 – 2013)	Système AIGA	Prévision immédiate d'ensemble pour l'alerte aux crues soudaines dans les bassins non-jaugés (Dép. du Var)	x		x			x
Master de J. Norville (Projet IMPREX)	-	-	Développent d'un « scoreboard » (package R-Shiny)					x	
Master et doctorat (<i>en cours</i>) M. Cassagnole (Projet IMPREX ; collab. EDF)	ENS ECMWF, Harmonie, System 5 (<i>prévu</i>)	GR4H GR6J MORDOR	Hydroélectricité et lien qualité-valeur ; courte, moyenne et longue échéances (11 à 15 BVs)	x		x	x		x
Doctorat D. Peredo (<i>en cours</i>) (Thèse ED GRNE ; Projet PICS)	Simulations stochastiques, PEARP, AROME PE (<i>prévu</i>)	GRP (opérationnel) GRSD (<i>prévu</i>)	Prévision immédiate et à courte-échéance de crues soudaines par approches d'ensemble	x	x	x	x		x
Doctorat A. Assis dos Reis (Collab. UFMG, Brésil)	ECMWF System5, Extended, GEFS NCEP (variable)	HEC-HMS FEWS-Cemig	Prévisions saisonnières pour le secteur hydroélectrique (41 BVs)	x		x			x
Velazquez <i>et al.</i> , 2011 (Collab. Univ. Laval, Québec)	DET et ENS ECMWF (2005 – 2006)	16 modèles globaux	Propagation des incertitudes (29 BVs)				x	x	
Bourgin <i>et al.</i> , 2014	PEARP (2005 – 2006)	GRP horaire	Impact de l'assimilation et du post-traitement (212 BVs)		x	x	x		
Thiboult <i>et al.</i> (2017) (Collab. Univ. Laval, Québec)	ENS ECMWF (2008 – 2010)	20 modèles globaux	Propagation des incertitudes et lien qualité-valeur (crues) (20 BVs au Québec)		x	x	x		x
Zhao <i>et al.</i> (2017) (Collab. CSIRO, Australie/HEPEX)	POAMA GCM (BoM, Australie) (1981 – 2011)	-	Comparaison méthodes de prétraitement (pluies) (Grille sur l'Australie)	x		x			

* voir liste de publications et d'encadrements dans <http://www.irstea.fr/ramos>

** (1) Prétraitement ; (2) Assimilation des données ; (3) Evaluation de la qualité ; (4) Post-traitement ; (5) Evaluation de la valeur, utilité et prise de décision (voir Figure 1)

1.3. Organisation du mémoire

Loin d'être exhaustif, ce mémoire d'habilitation à diriger des recherches présente quelques éléments de réflexion autour de la qualité et de la valeur des prévisions. Il s'articule en cinq chapitres (incluant ce chapitre d'introduction) :

- Le **Chapitre 2** introduit des concepts, réflexions, points de vue et perceptions qui se sont développés au long de mes recherches autour de la prévision. Bien évidemment, je ne prétends pas avoir mené ces réflexions seule. Au contraire, j'ai bénéficié de plusieurs échanges lors des colloques et rencontres auxquels j'ai pu participer, des réflexions menées lors de mes séjours scientifiques au JRC (Italie, 2005-2007), à l'Université Laval (Canada, 2013), au CEPMMT (Angleterre, 2014), et au NCAR (USA, 2015), ainsi que des nombreux échanges dans le cadre du groupe HEPEX (*Hydrologic ensemble prediction experiment*⁸) et avec des services opérationnels en France (notamment, EDF DTG, CNR, réseau SCHAPI/SPC et EPTB IAV).
- Le **Chapitre 3** est consacré aux travaux centrés sur l'évaluation de la qualité des prévisions. Ce chapitre contient quelques résultats des travaux que j'ai dirigés en Master 2 ou en thèse, et qui m'ont permis d'avancer sur la compréhension de ce qui définit une bonne prévision hydrologique. Des questions clés liées aux impacts des techniques d'assimilation de données et des méthodes de post-traitement statistique sur les critères de qualité des prévisions y sont également abordées.
- Le **Chapitre 4** concerne les études qui ont examiné le lien qualité-valeur des prévisions. Il se base sur des résultats de travaux que j'ai dirigés en Master 2 ou en thèse, mais aussi sur des études auxquelles j'ai contribué, menées par des collègues à l'étranger. La valeur d'une prévision est ici illustrée par sa valeur économique ou décisionnelle (utilité que peut avoir l'information fournie par la prévision dans un contexte de décision).
- Enfin, le **Chapitre 5** présente les perspectives et pistes qui resteraient à explorer.

⁸ Depuis plus de 13 ans d'existence, HEPEX relie la communauté de chercheurs, prévisionnistes et utilisateurs des prévisions et facilite l'échange d'idées, données, méthodes et expériences. Aujourd'hui, cette initiative internationale compte plus de 450 membres dans le monde. J'ai participé à HEPEX lors de mon post-doctorat en Italie et j'ai poursuivi mon engagement une fois rentrée en France : j'ai été animatrice régionale pour l'Europe de 2012 à 2014 et membre du comité de présidence de 2014 à 2018. L'animation du réseau HEPEX consiste à contribuer aux orientations scientifiques du réseau, à l'organisation de sessions scientifiques et ateliers, et à l'administration et animation scientifique du Portail HEPEX, www.hepex.org (site hébergé par Irstea depuis 2013).

2. La prévision intègre un raisonnement probabiliste

2.1. Données, modèles et prévisionnistes pour quantifier l'incertitude

Les systèmes de prévision des débits des rivières peuvent servir à des applications diverses et cibler différents horizons de prévision. De manière générale, leur but est de pouvoir anticiper le plus possible les événements à venir et activer des réponses adéquates face à la sévérité de l'aléa prévu. Il s'agit notamment de :

- systèmes de prévision à court terme, dont l'objectif est de capturer des signaux d'alerte et fournir des avertissements à quelques jours d'avance pour les crues et inondations, voire quelques heures pour la prévision immédiate des phénomènes rapides ;
- systèmes de prévision à moyenne échéance (jusqu'à 10-15 jours) dans le but d'anticiper des actions de prévention face à des événements de crues prévus ou optimiser la gestion des apports en eau aux barrages-réservoirs ;
- systèmes dédiés à l'estimation de la ressource en eau disponible pour les mois à venir, en particulier, pour mieux planifier l'usage de l'eau et anticiper des situations de sécheresse, pendant lesquelles des conflits d'usage de l'eau peuvent avoir lieu.

Bien que les pratiques et les préoccupations des agences et services de prévision opérationnelle peuvent varier selon les pays et les services, notamment puisque certains aspects de la prévision opérationnelle dépendent de facteurs culturels, sociétaux, politiques et environnementaux (Pagano *et al.*, 2014), les systèmes de prévision hydrologique, indépendamment de l'horizon de prévision ciblé, ont au moins trois piliers en commun, qui servent de base à la production des prévisions :

1. les données, qu'elles soient historiques, observées en temps réel ou prévues (pour les entrées météorologiques),
2. les modèles, qu'ils soient des modèles hydrologiques distribués à base physique, conceptuels ou des modèles hydrauliques de propagation amont-aval, et
3. les prévisionnistes, confrontés à de nombreux défis liés à la gestion des données et la modélisation des bassins versants pour la prévision des débits, ainsi qu'aux contraintes administratifs et organisationnels des services de prévision pour lesquels ils travaillent.

Ces trois piliers sont essentiels dans une chaîne de prévision hydrométéorologique. Ils agissent dans la quantification des incertitudes et dans le flux d'expertise et de décisions qui traverse la chaîne, de la prévision et alerte météorologique et hydrologique à la cartographie des zones potentiellement inondables (**Figure 2** ; Ramos *et al.*, 2010a).

Tandis que les données et modèles aident à quantifier l'évolution future des variables d'intérêt à différentes échelles, une séquence d'actions à différents niveaux de complexité doit être entreprise par le prévisionniste. Cela comprend la gestion de bases de données et de leurs interconnexions en temps réel, le couplage ou chaînage des modèles mathématiques utilisés, la mise en place de routines diverses de pré- et post-traitement, la visualisation des sorties des modèles et l'évaluation de la qualité des prévisions, en appui aux produits et services qui seront communiqués aux utilisateurs finaux des prévisions (comme cela est illustré dans la **Figure 1**).

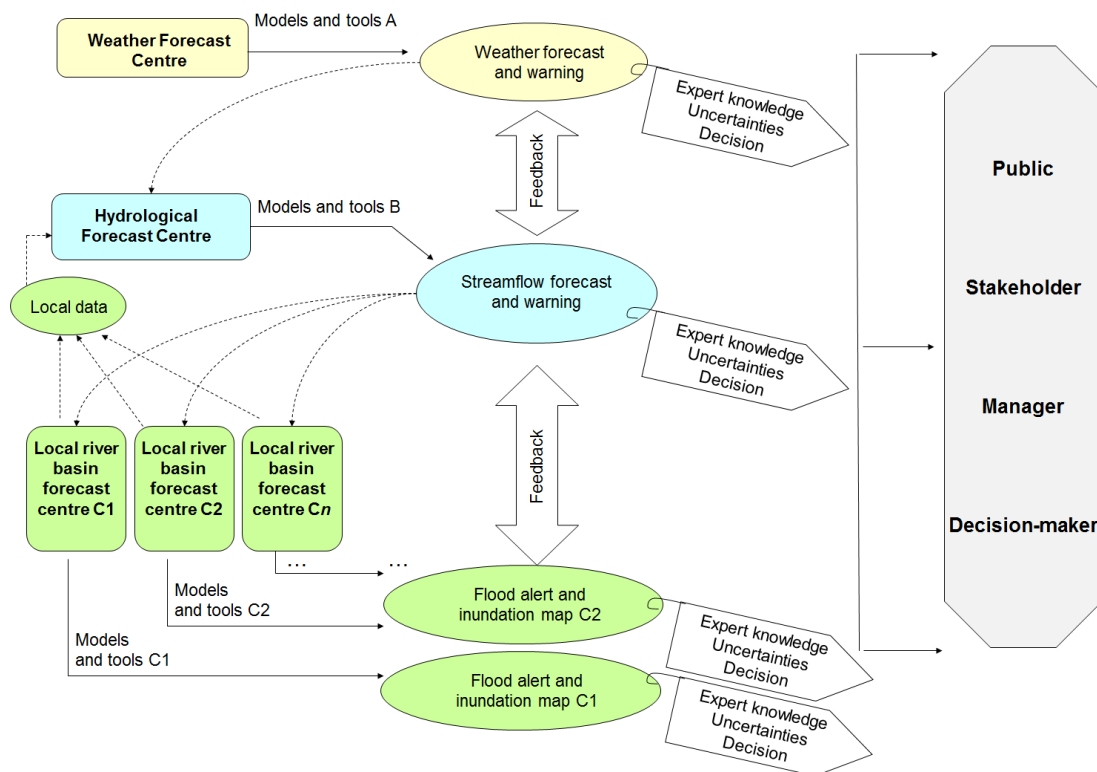


Figure 2 : Schéma des principales interconnexions dans une chaîne de prévision hydrométéorologique (dans Ramos et al., 2010a)

Au vu de sa complexité, un système de prévision efficace repose sur l'utilisation des bonnes données et des bons outils pour la bonne quantification de l'incertitude, mais aussi sur l'intériorisation (non seulement par les prévisionnistes, mais aussi par les développeurs des modèles, les utilisateurs, les décideurs et les instances institutionnelles) que la prévision intègre inévitablement un raisonnement probabiliste. Ce qui paraît évident n'est cependant pas immédiat dans la pratique ou lors de la mise en place de systèmes de prévision. En effet, plusieurs facteurs contribuent à imaginer, calculer, communiquer et vouloir recevoir des prévisions sous forme d'une évolution unique du futur le plus probable, qui sera observé ou très proche de l'observé (du moins, c'est ce que l'on souhaite au moment de la prévision).

Par exemple, l'exercice de modélisation hydrologique a été longtemps déterministe, dans la recherche du « meilleur jeu de paramètres » ou de « la meilleure simulation ». Bien qu'il ait été largement reconnu que les modèles hydrologiques n'étaient pas parfaits (ainsi que, par ailleurs, notre connaissance du système physique qu'ils représentent), et que des méthodes de prise en compte des incertitudes liées aux paramètres des modèles hydrologiques aient été proposées depuis maintenant plusieurs décennies (p.ex., la méthode GLUE a été proposée par Keith Beven et Andrew Binley en 1992 ; Beven et Binley, 1992, 2014), la prise en compte des incertitudes dans les simulations hydrologiques, en tant que partie intégrante du produit final d'un modèle, n'est aujourd'hui toujours pas une pratique adoptée par tous et partout. Ceci peut être expliqué, d'une part, par la difficulté à évaluer explicitement les erreurs de modélisation et les incertitudes des observations hydrométéorologiques, et, d'autre part, par le manque de pratique dans l'évaluation de simulations hydrologiques incertaines (par ex., Franz et Hogue, 2011, proposent un article intéressant sur la transposition des critères d'évaluation de prévisions probabilistes à l'évaluation d'un ensemble d'hydrogrammes obtenus à partir d'un modèle hydrologique conceptuel de simulation).

La dissociation du processus de production des prévisions du processus de prise de décision dans plusieurs systèmes opérationnels peut aussi contribuer à éclipser les incertitudes et probabilités des prévisions, ou au moins à les rendre moins visibles lors du passage de l'information du prévisionniste au décideur. Dans plusieurs cas, la prise de décision est vue comme un processus déterministe (du fait, par exemple, que la décision sera, de toute façon, binaire, du type oui/non). En extrapolation de raisonnement, on pourrait être tenté de faire l'hypothèse qu'un décideur n'a donc pas besoin de connaître (voire reconnaître), au moins quantitativement et explicitement, l'ensemble des incertitudes qui entourent les informations qui viennent en amont de sa décision pour agir. En d'autres termes, le décideur n'aurait besoin que du scénario le plus probable et non pas de toute la distribution prédictive.

Enfin, il pourrait y avoir aussi une question de base : et si le système de prévision ne peut pas quantifier toutes les sources d'incertitude et nous n'avons donc pas accès aux probabilités prédictives totales ? Heureusement, plusieurs décideurs peuvent se baser sur des scénarios pour prendre leurs décisions (ou, comme le note L. Smith dans son livre introductif sur le chaos⁹, « l'utilité n'exige pas d'extraire des estimations de probabilité utiles »). J'ajouterais également que, heureusement, nous pouvons aussi faire appel aux techniques de post-traitement statistique (mais je reviendrai plus tard sur ce sujet dans le **Chapitre 3**).

2.2. La prévision hydrologique opérationnelle en cours de transformation majeure

Dans un article en préparation, Wood *et al.* (2018)¹⁰ proposent d'examiner les transformations que les systèmes de prévision de crue ont subies (ou sont actuellement en train de subir) de manière générale. On y distingue quatre grands paradigmes (ou courants de pensées) de la prévision hydrologique opérationnelle, utilisés ou développés aujourd'hui : (1) le paradigme traditionnel, le plus souvent déterministe, centré sur la modélisation hydrologique conceptuelle, et marqué par une forte présence de l'intervention manuelle du prévisionniste ; (2) le paradigme de la modélisation à haute-résolution (souvent en référence à la résolution spatiale des modélisations), (3) le paradigme de la prévision hydrologique d'ensemble, et (4) le paradigme du système couplé atmosphère-surface, avec routage du ruissellement de surface des modèles numériques de prévision météorologique ou couplage d'un schéma de surface à un modèle hydrologique.

Ces paradigmes se distinguent dans leur philosophie et leurs objectifs, leurs séquences de travail (« workflow ») et choix de modélisation, et leurs potentiels à réduire et représenter l'incertitude des prévisions. Leur délimitation est, bien entendu, une simplification, dans la mesure où certains systèmes de prévision peuvent ne pas correspondre exactement à une seule vision ou cadre de travail, mais plutôt combiner des éléments de plusieurs paradigmes. Néanmoins, la caractérisation proposée peut être utile pour mieux comprendre comment les nouveaux systèmes de prévision se développent dans le monde, ainsi que leurs avantages et compromis (« tradeoffs ») potentiels, en lien avec l'approche sur laquelle ils se basent ou de

⁹ Leonard Smith (2007) *Chaos, A Very Short Introduction*, Oxford University Press, 180 p. La phrase originale est en page 149 : « *Luckily, utility does not require extracting useful probability estimates.* »

¹⁰ Wood, A., M.-H. Ramos, F. Pappenberger, M. Clark *et al.* (2018), *The Rise of New Paradigms in Flood Prediction*, en préparation.

laquelle ils s'inspirent le plus. L'objectif ici n'est pas de décrire en détail les paradigmes, mais plutôt de faire ressortir ce que l'approche de la prévision hydrologique d'ensemble, sur laquelle ce manuscrit est centré, apporte à (ou retire de) l'exercice de prévision, et ce notamment vis-à-vis du paradigme traditionnel de la prévision.

L'approche de prévision d'ensemble a été développée dans le but d'appréhender et quantifier l'incertitude inhérente à toute prévision. Elle ne cherche pas à fournir une unique prévision déterministe, mais un ensemble de scénarios représentatifs de l'incertitude liée à la prévision. Ces scénarios (aussi appelés « membres » ou « runs ») ont pour objectif d'échantillonner au mieux les différentes sources d'erreurs pouvant influencer la connaissance de l'évolution future des variables prévues.

Il est important de noter que la prévision d'ensemble n'a pas l'ambition d'échantillonner le « monde réel » que l'on observera dans le futur. En faisant des prévisions d'ensemble, nous souhaitons examiner quelles alternatives nous pouvons avancer sur la base de nos connaissances actuelles du système (conditions initiales) et des imperfections de nos modèles à représenter l'évolution de ce système (**Figure 3**). De plus, la prévision d'ensemble cherche à proposer un ensemble d'alternatives cohérentes dans le temps et dans l'espace, ce qui s'avère être un atout important lorsque des modèles s'enchaînent dans un système de prévision (par ex., des sorties des modèles hydrologiques sont utilisées en entrée des modèles hydrologiques dont les sorties sont, à leur tour, utilisées en entrée d'un modèle de gestion des niveaux d'eau et lâchers d'un réservoir).

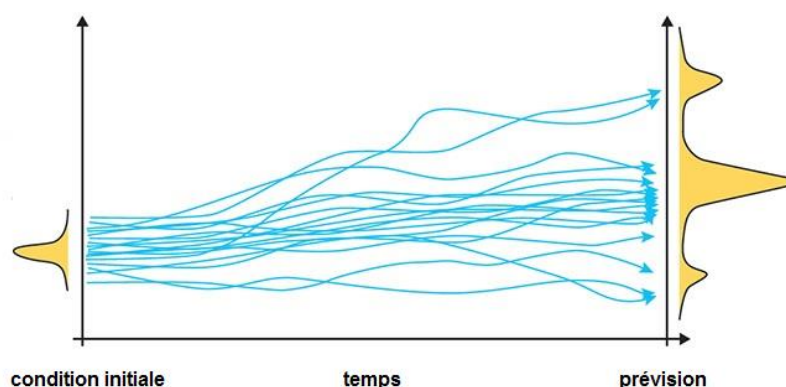


Figure 3 : Un ensemble de prévisions produit une gamme de scénarios possibles plutôt qu'une seule valeur prévue. La distribution des membres de l'ensemble donne une indication de la probabilité d'occurrence de ces scénarios (distribution prédictive) (d'après : <https://www.ecmwf.int>)

Il semblerait que cette notion de « quantification de l'incertitude » n'a pas été une controverse scientifique ou opérationnelle dans la prévision à longue échéance. L'utilisation de plusieurs scénarios futurs pour prévoir l'avenir aux échelles mensuelles ou saisonnières est une pratique répandue et bien acceptée, autant en météorologie qu'en hydrologie. En hydrologie, la méthode ESP (*Ensemble Streamflow Prediction*; Day, 1985), basée sur l'utilisation d'un historique de conditions météorologiques observées dans le passé en entrée d'un modèle

hydrologique continu de simulation en temps réel pour appréhender les conditions initiales des bassins versants à l'instant de la prévision, apparaît dès les années 1970¹¹. Un autre cas illustratif est celui des modèles statistiques (régressions statistiques avec intervalles de confiance ou recherche de situations météorologiques analogues associées aux précipitations locales) utilisés à EDF depuis également les années 1970 pour la prévision probabiliste à longue échéance (horizon de prévision de plusieurs mois) (Desaint *et al.*, 2009).

En météorologie, Tim Palmer rappelle un de ses articles co-écrit avec J. Murphy en 1986 dans sa présentation durant le colloque « *Ensemble prediction: past, present and future* » (11-14 septembre 2017, Reading, UK), et note¹² : « *The application of ensemble ideas to monthly (and seasonal) forecasting was not controversial. What was more controversial is that ensemble forecast methods should be developed within what was referred to as “the limit of deterministic predictability”. Medium-range ensemble forecasting was treated with scepticism. Instead of probabilistic prediction, the thought at the time was that we should be trying to “forecast the forecast skill” of the deterministic forecast. This turned out to be a forlorn hope, though it led to some very important scientific spin offs, such as how to perturb the initial conditions in a medium-range ensemble forecast.* » (diapositive 26, cachée à la fin de sa présentation).

Alors que la prévision d'ensemble à moyenne échéance a été développée dans les années 1990 en météorologie¹³ (Molteni *et al.*, 1996 ; Toth et Kalnay, 1997), ce type de prévision n'a débuté en hydrologie qu'au début des années 2000, notamment avec la création du groupe international HEPEX en 2004, lors d'un atelier tenu au CEPMMT (UK) et organisé conjointement avec le Service météorologique national des États-Unis (NWS) et la Commission européenne (Schaake *et al.*, 2007 ; Thielen *et al.*, 2008). Les premiers démonstrateurs commencent alors à voir le jour, dans le but de mettre en place des systèmes opérationnels de prévision hydrologique d'ensemble (Zappa *et al.*, 2008 ; Cloke et Pappenberger, 2009).

Si, dans un premier temps, les applications ont concerné la simple propagation des ensembles météorologiques dans un modèle hydrologique, des développements ont progressivement ciblé les défis inhérents aux étapes hydrologiques de la prévision de débits, telles que l'assimilation de la donnée de débit prise à l'exutoire du bassin versant (voir la synthèse proposée par Liu *et al.*, 2012), l'influence des erreurs de calage du modèle hydrologique (Olsson et Lindstrom, 2008 ; Engeland *et al.*, 2010), ou la prise en compte du lien probabiliste entre l'erreur de prévision et des variables explicatives liées au processus de prévision hydrologique (Montanari et Grossi, 2008 ; Coccia et Todini, 2011).

L'intérêt pour la prévision d'ensemble hydrométéorologique s'est également tourné vers d'autres échéances, avec la prise en compte de développements ciblant la prévision météorologique d'ensemble à courte échéance (par ex., systèmes SREF/NWS USA, PEARP/Météo-France, STEPS/UK MetOffice et COSMO LEPS/COSMO Consortium) et

¹¹ Le blog post HEPEX d'A. Wood *et al.*, *Tracing the Origins of ESP*, propose un historique intéressant sur la méthode ESP : <https://hepex.irstea.fr/tracing-the-origins-of-esp/> (dernière consultation le 28/12/2017)

¹² Tim Palmer (2017), *The ECMWF Approach to Ensemble Prediction* (dernière consultation le 28/12/2017) : <https://www.ecmwf.int/sites/default/files/elibrary/2017/17606-ecmwf-approach-ensemble-prediction.pdf>

¹³ 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 ».

l'incorporation de systèmes de prévisions météorologiques saisonnières en hydrologie (en complément ou en alternative à l'approche traditionnelle ESP) (plusieurs exemples sont présentés dans le récent numéro spécial de la revue HESS, *Sub-seasonal to seasonal hydrological forecasting*, 2016-2017)¹⁴.

Aujourd'hui, de nombreux systèmes opérationnels de prévision hydrologique d'ensemble existent ou sont en cours d'expérimentation (Pappenberger *et al.*, 2016). Par rapport au paradigme traditionnel, la prévision d'ensemble propose une progression logique à l'idée du « meilleur/pire scénario ». Ce qui est intéressant dans ce passage d'une vision qui s'appuie sur des prévisions à valeur unique (tout en reconnaissant, qualitativement, l'incertitude de l'exercice de prévision) à une prévision à plusieurs scénarios, qui cherche à quantifier l'incertitude prédictive, est le changement d'orientation qui s'opère dans le système de prévision. L'accent qui était auparavant mis sur la précision que l'on pourrait obtenir d'une simulation du modèle hydrologique se déplace vers l'objectif d'assurer que le système quantifie de manière statistiquement fiable les incertitudes prédictives.

De même, l'utilisation opérationnelle de la prévision d'ensemble nécessite un plus grand degré d'automatisation dans la chaîne de prévision par rapport à ce qui est souvent observé dans les approches plus traditionnelles de prévision. Cet aspect est à la base de l'approche appelée « *human over-the-loop* » (« prévisionniste au-dessus de la boucle »), qui contraste avec l'approche traditionnelle « *human in-the-loop* » (« prévisionniste dans la boucle »), où le prévisionniste intervient manuellement dans plusieurs étapes de la prévision (Wood *et al.*, 2018 ; en préparation).

Dans l'approche traditionnelle, les prévisionnistes passent par un mélange de processus semi-automatisés et manuels pour mettre en œuvre et caler des modèles hydrologiques, contrôler la qualité des données météorologiques et hydrologiques, initialiser les simulations des modèles, évaluer et modifier les forçages, les états et les résultats, parmi une myriade d'autres tâches. Effectuer toutes ces tâches sur 20 ou 50 membres d'une prévision d'ensemble, et parfois sur 10, 20 ou une centaine de bassins versants, devient, logiquement, compliqué dans les délais de la prévision opérationnelle. De plus, il y a le risque de générer des incohérences physiques dans les résultats finaux ou de générer des prévisions qui ne peuvent pas être reproduites dans le cas d'un retour d'expérience systématique, restant ainsi « invérifiables ».

Le passage d'un système de prévision déterministe peu automatisé à une approche de prévision d'ensemble peut être rapidement limité par sa capacité à fonctionner avec des techniques plus sophistiquées d'assimilation de données ou de post-traitement statistique. Une faible automatisation peut aussi compromettre la capacité du système à atteindre des résolutions spatiales élevées ou à faire tourner plusieurs modèles hydrologiques, ce qui peut contraindre les options de modélisation et déléguer l'amélioration de la qualité des prévisions à d'autres compartiments du système (forçage météorologique, collecte des données, produits de communication, etc.). De plus, l'évaluation des prévisions peut aussi devenir une tâche plus ardue puisqu'elle nécessite de longues séries de prévisions passées et que, si le système est très manuel, la qualité d'une prévision devient un mélange complexe de la qualité du système et du jugement du prévisionniste.

¹⁴ https://www.hydrol-earth-syst-sci.net/special_issue824.html (dernière consultation le 28/12/2017).

En contraste, l'approche « *human over-the-loop* » favoriserait l'automatisation de plusieurs étapes de la prévision (par ex., *runs* des simulations d'ensemble, collecte, critique et assimilation automatiques des données hydrométéorologiques, post-traitement automatique des sorties). Cela n'enlèverait pas l'expertise humaine de l'exercice de prévision, mais plutôt permettrait au prévisionniste de mieux se consacrer à des tâches pour lesquelles son expertise est primordiale et pourrait être plus valorisante (par ex., intégration d'informations qualitatives disponibles en temps réel, prise en compte de processus ou ouvrages non représentés dans les modèles, interprétation des résultats et communication de la prévision avec ses incertitudes en temps réel ; Pagano *et al.*, 2016).

En résumé, l'adoption d'un paradigme ou d'un autre révèle le point central auquel un service ou un développeur serait (explicitement ou inconsciemment) attaché (par ex., modèle hydrologique, résolution spatiale, part d'automatisme et de l'expertise humaine, quantification des incertitudes, couplages surface-atmosphère et bilans énergétiques/hydriques) et suppose une allocation de ressources (informatiques ou humaines) parfois très différente.

Bien que l'automatisation de certaines tâches devienne inévitable quand il s'agit de mieux quantifier les principales sources d'incertitudes de la prévision hydrologique, les prévisionnistes continueront certainement à jouer un rôle essentiel dans le processus de prévision. Cela tient en partie au fait que la génération de ruissellement dans un bassin versant est fortement compliquée par l'interaction complexe entre le paysage naturel et les interventions humaines sur ce paysage (occupation du sol, infrastructures, ouvrages de rétention et prélèvements, etc.). Cependant, ce rôle pourrait être modifié ou actualisé vers des nouvelles tâches une fois que certaines étapes du système de prévision en temps réel seront graduellement automatisées. Ces nouvelles tâches peuvent comprendre, par exemple, une évaluation plus détaillée des prévisions émises et la prise en compte de processus plus complexes et parfois singuliers lors de l'occurrence d'un événement rare. L'avantage est également d'avoir, hors temps réel, des outils plus automatisés pour pouvoir reproduire des événements passés dans un objectif de retour d'expérience sur la qualité du système de prévision et/ou se consacrer au développement de nouvelles composantes de modélisation statistique ou de processus physiques dans le système.

Pour les utilisateurs des prévisions et les décideurs, les changements qui aujourd'hui transforment la prévision hydrologique pourraient entraîner une plus grande disponibilité d'informations liées aux prévisions, menant à un plus large champ d'applications possibles et à des cadres quantitatifs de décisions fondées sur le risque plus élaborés. Enfin, ces changements ne seront pas sans conséquence non plus pour la science et l'éducation. L'évolution du paysage exige un large éventail de connaissances et de compétences qui, traditionnellement, ne sont pas offertes dans une seule discipline. Cela implique une science davantage multidisciplinaire et des offres de cours et formation qui puissent lier les différentes notions utilisées en prévision, issues de la météorologie, l'hydrologie, la théorie des probabilités et les statistiques appliquées, la programmation informatique et, de plus en plus, la théorie de la décision, l'économie de l'eau et les sciences sociales (qui éclairent la communication et l'utilisation des prévisions).

Pour finir, j'aimerais citer une autre phrase du livre de L. Smith : « *Interpréter une prévision d'ensemble du monde réel comme s'il s'agissait d'une prévision probabiliste d'un modèle parfait d'un système mathématique est une maladresse de prévision des plus naïves.* »¹⁵. Les systèmes de prévision d'ensemble peuvent représenter un système de prévision probabiliste, qui associe une probabilité (ou vraisemblance) à l'occurrence future d'un événement donné, s'ils sont statistiquement bien calibrés et sous l'hypothèse que le nombre de scénarios est suffisant pour bien représenter la distribution prédictive, c'est-à-dire, l'ensemble de probabilités d'occurrence de l'événement.

Comme à la fin de la **section 2.1**, nous revenons à la question autour de la quantification des incertitudes prédictives totales. On pourrait penser que nous avançons peu dans nos propos, mais, ne nous impatientons pas, le **Chapitre 3** arrive bientôt. Voyons d'abord ce que nous pouvons dire sur les prévisions probabilistes et leur contribution à la prise de décision.

2.3. Les prévisions probabilistes conduisent-elles à de meilleures décisions ?

Pendant mes recherches au sein de l'équipe qui développait le système EFAS (*European Flood Alert System*)¹⁶ au JRC à Ispra (Italie), j'ai eu l'opportunité de jouer un rôle de chercheur développeur, pour ce qui est de la mise en place du système de prévision d'ensemble, mais également de prévisionniste - décideur, car je participais activement aux prévisions (pré)opérationnelles faites en temps réel. J'ai eu la possibilité d'avoir été dans ce rôle lors des graves événements de crues de l'été 2005 et du printemps 2006 (Ramos, 2006) :

- Du 22 au 27 août 2005, le nord de la région alpine a été touché par de graves inondations, qui ont eu lieu sur les affluents du Rhin en Suisse et en Autriche et sur plusieurs affluents du Danube en Allemagne, Autriche et Hongrie. EFAS avait prévu ces événements et les avait signalés aux services hydrologiques des Etats membres européens concernés. L'analyse de l'historique des dépassements des niveaux critiques d'alerte EFAS a permis de montrer qu'une augmentation de l'anticipation d'au moins 1 à 2 jours aurait été possible en prenant en compte la persistance et la cohérence des prévisions hydrologiques d'ensemble (Ramos *et al.*, 2010b). Le vécu en temps réel a permis d'identifier que de meilleures règles de décision à appliquer en temps réel étaient nécessaires et a déclenché plusieurs développements autour de la mise en place d'outils de visualisation de la persistance des prévisions et des probabilités de dépassement des seuils (Ramos *et al.*, 2007).
- Du 29 mars au 09 avril 2006, les bassins de la Morava, dans le bassin du Danube, et de la Vltava, dans le bassin de l'Elbe, ont été frappés par de graves inondations causées par la fonte des neiges, combinée à des précipitations. Les inondations ont notamment affecté la Slovaquie et la République tchèque. Dans la Morava aval, les périodes de retour des débits de pointe observés ont été estimées à plus de 100 ans. Le système EFAS avait prévu l'événement avec plus de cinq jours d'anticipation. En effet, l'analyse *a posteriori*

¹⁵ Leonard Smith (2007) *Chaos, A Very Short Introduction*, Oxford University Press, 180 p. La phrase originale se retrouve en page 147-148 : « *Interpreting an ensemble forecast of the real world as if it were a perfect model probability forecast of a mathematical system is to make the most naïve forecasting blunders.* »

¹⁶ Plus tard, le système a été rebaptisé en « *European Flood Awareness System* ».

de l'événement avait montré un signal persistant à partir des prévisions de débit basées sur les prévisions météorologiques d'ensemble. Cependant, l'absence de cohérence entre les simulations basées sur les différentes prévisions météorologiques utilisées à l'époque (prévisions déterministes et d'ensemble) avait conduit à une diffusion tardive des prévisions aux partenaires des pays concernés, ce qui s'est répercuté sur la valeur des prévisions en tant qu'alerte précoce (Kalas *et al.*, 2008 ; Younis *et al.*, 2008).

Autant dire que ces deux événements vécus en temps réel m'ont totalement convaincue du potentiel de la prévision hydrologique d'ensemble pour anticiper les événements de crues. Il n'en restait pas moins que le manque de pratique sur les systèmes d'ensemble (ou probabilistes) laissait les prévisionnistes naturellement incertains quand il s'agissait de passer à l'action lorsqu'un événement extrême était prévu.

Bien que plusieurs acteurs de la prévision acceptent l'idée que l'information sur l'incertitude des prévisions est utile et doit être fournie, certains restent encore dubitatifs quant à la capacité d'une telle information à vraiment améliorer le processus décisionnel dans son ensemble et, par conséquent, sur le fait que ce soit une information qui doive être communiquée aux décideurs. Des questions sont souvent soulevées lors de discussions et échanges sur le sujet : si une décision doit être unique (oui/non ; ceci ou cela), alors comment décider face à plusieurs scénarios possibles ? Sur quel scénario les utilisateurs baseront-ils leur décision ? Quel niveau de probabilité (ou pourcentage de membres de l'ensemble) est crucial pour déclencher une alerte ? A quel point l'éventail des possibles fourni par le système de prévision est-il fiable ? Les prévisions probabilistes (ou d'ensemble) conduisent-elles vraiment à de meilleures décisions ?

Mes premières études en lien avec la perception du risque, la communication des incertitudes et la prise de décision en contexte incertain ont été menées lors d'un atelier que j'avais organisé en 2005 avec les utilisateurs des prévisions EFAS, en collaboration avec des hydrologues et chercheurs en sciences sociales de Lancaster University et King's College London. Cet atelier a eu pour but d'identifier plusieurs facteurs qui jouent un rôle sur la compréhension et communication des prévisions probabilistes quand des prévisionnistes sont confrontés à une situation de prévision de crue et doivent prendre des décisions.

Le concept de l'atelier était d'avoir un petit groupe de prévisionnistes de différents bassins fluviaux en Europe pour travailler sur un nombre de cas d'étude spécialement préparés pour les mettre en situation de prévision de crues en temps réel avec information de l'incertitude des prévisions. Les participants ont eu un temps limité pour analyser les situations proposées et décider s'ils enverraient ou non une alerte pour la procédure d'urgence. Des invitations ont été adressées à 11 prévisionnistes de huit pays européens, représentant neuf services hydrologiques nationaux différents.

Les participants ont été divisés en quatre groupes de travail pour analyser trois cas d'étude, l'un après l'autre. Chaque cas correspondait à la prévision d'une situation de crue potentielle sur un bassin hydrographique différent. Il s'agissait de vrais cas prévus par le système EFAS avec plus de 2-3 jours d'anticipation, qui avaient été anonymisés pour ne pas être reconnus par les prévisionnistes participant à l'atelier.

Pour chaque cas d'étude, nous avons préparé des informations sur trois jours consécutifs de prévision, qui pourraient être utilisées par les participants, à leur convenance, pour soutenir leurs décisions. Il s'agissait d'informations sur les derniers débits et précipitations journaliers observés, les précipitations prévues, les débits prévus par EFAS pour la moyenne échéance (synthèse et hydrogrammes pour les 7 à 10 jours à venir) et par un système supposé local pour la courte échéance (2 jours à venir), et des images d'un radar météorologique couvrant le bassin versant. Chaque jour de prévision était mis à disposition de manière séquentielle : le jour suivant n'était fourni que lorsqu'une décision avait été prise au jour qui le précédait. A chaque jour de prévision, chacun des groupes devait analyser la situation et décider si oui ou non ils contactaient les autorités de protection civile avec une alerte aux crues pour les jours suivants. Ils devaient aussi qualifier le niveau de gravité de la situation : faible, moyen ou élevé.

Afin d'évaluer l'influence des différentes informations sur les analyses et décisions des prévisionnistes, l'information disponible était distribuée de manière inégale (sans que les participants en soient au courant avant de démarrer leurs analyses) : à chaque cas d'étude, un groupe était affecté pour recevoir une quantité réduite d'information, qui comprenait toutes les informations mentionnées dans le paragraphe précédent, sauf les débits prévus par EFAS et par le système « local ». Ce groupe était ainsi sans aucune information sur l'incertitude hydrologique ou sur la prévision hydrologique d'ensemble. Cette situation « d'information réduite » pouvait arriver lors du premier cas pour certains groupes ou lors du deuxième ou du dernier cas d'étude pour les autres groupes. De plus, les cas d'étude ont été sélectionnés de manière à ce que, dans chaque cas, le rôle des prévisions d'ensemble était légèrement différent : parfois ces prévisions étaient en accord avec les prévisions déterministes, parfois elles indiquaient un signal contradictoire à celui donné par la prévision déterministe.

Avant de commencer l'atelier, un aperçu des aspects concernant le système de prévision EFAS et les systèmes de prévision d'ensemble en hydrométéorologie de manière générale avait été donné afin d'amener les participants à un même niveau de connaissance. Les détails de la méthodologie et des résultats de cet atelier ont été publiés dans deux articles, Demeritt *et al.* (2007) et Ramos *et al.* (2010a), et dans le rapport technique Thielen *et al.* (2005).

De manière générale, il est ressorti que les prévisions probabilistes étaient souvent utilisées pour soutenir une décision basée sur les prévisions déterministes. Il faut noter que, à l'époque, la prévision déterministe de précipitation avait une résolution bien plus fine que la prévision d'ensemble (même par rapport à la prévision du membre contrôle de l'ensemble, qui partait des mêmes conditions initiales de la prévision déterministe). Néanmoins, il a été également observé que l'information hydrologique d'ensemble manquait fortement aux prévisionnistes, notamment quand, après l'avoir eue sur un cas d'étude, elle n'était plus disponible pour les cas suivants.

Plus tard, c'est en cherchant à mieux comprendre comment quantifier l'importance de l'information sur les incertitudes de prévision pour la prise de décision que l'idée de mettre en place des jeux de rôle est venue collectivement au sein des discussions avec mes collègues du réseau international HEPEX. Jusqu'à présent, j'ai coordonné ou participé au développement de cinq jeux qui ont été mis à disposition de la communauté nationale et internationale, certains traduits en plusieurs langues et aujourd'hui largement utilisés en conférences, formation et enseignements.

La **Figure 4** illustre les trois premiers jeux de la liste ci-après¹⁷:

1. Flood control game: do probabilistic forecasts lead to better decisions? (2012) (Ramos *et al.*, 2013 ; Fan *et al.*, 2015b)¹⁸,
2. Water management game (2013) (Crochemore *et al.*, 2016a)¹⁹,
3. Pay for a forecast game (2015) (Arnal *et al.*, 2016)²⁰,
4. The Shopkeepers dilemma: a decision-making game using probabilistic forecasts (2016, M. Werner)²¹,
5. Pathways to running a flood forecasting centre: an adventure game (2017, L. Arnal)²².

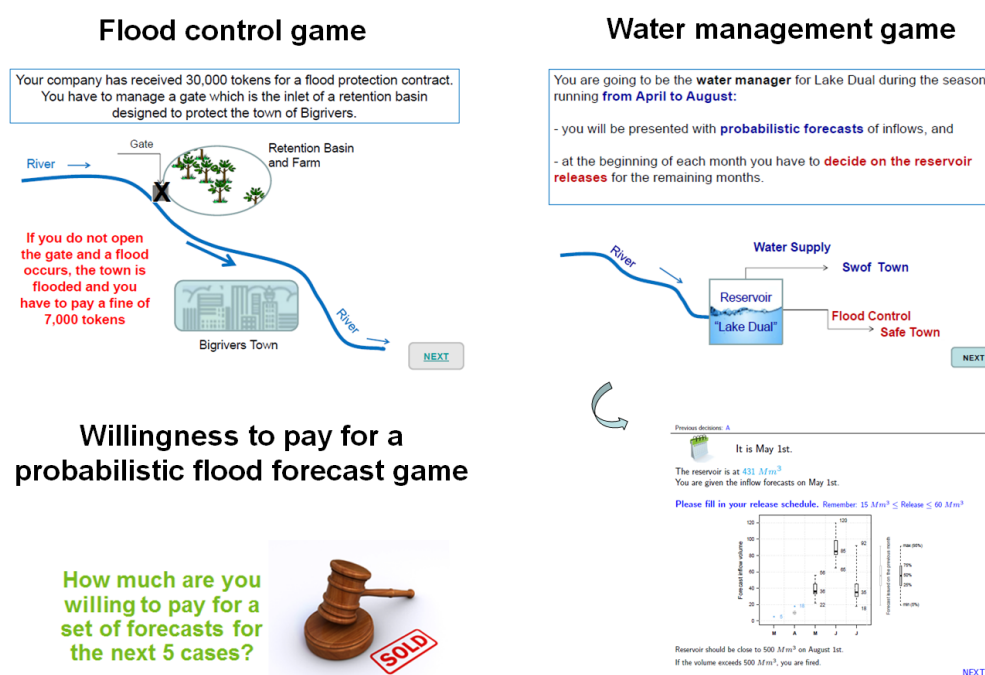


Figure 4 : Voulez-vous jouer à un jeu de prévision hydrologique probabiliste ?

¹⁷ Les jeux sont disponibles via le portail web d'HEPEX : <http://hepex.irstea.fr/resources/#gamestraining>.

¹⁸ Ce jeu a été repris et adapté par des prévisionnistes opérationnels en Ecosse et au Pays-Bas. Fan *et al.* (2015b) présentent une application faite lors d'un congrès au Brésil. La version française est utilisée dans la formation IFORE « Prévisions hydrologiques et incertitudes », offerte aux prévisionnistes du réseau national SCHAPI/SPC depuis 2016 en deux séances annuelles.

¹⁹ Ce jeu a été adapté par des gestionnaires du barrage Uersfeld de contrôle de crues, localisé à Aix-la-Chapelle et géré par WVER (Wasserverband Eifel Rur) en Allemagne dans le cadre du projet Interreg IVB NWE DROP. L'auteure principale, L. Crochemore, a participé à ce projet, et j'ai dirigé sa thèse de doctorat, dont le jeu fait partie.

²⁰ Ce jeu existe aussi en version en ligne. Il a été développé par L. Arnal, que j'avais co-encadrée en master (en partenariat avec Deltares) en 2014 et avec qui j'avais gardé contact (L. Arnal est actuellement en thèse à l'University of Reading et au CEPMMT, UK).

²¹ Les résultats du jeu ont été présentés dans un poster à l'EGU en 2017 et dans le blog post : <https://hepex.irstea.fr/the-elusive-optimal-decision-rule-and-the-impact-of-forecasting/>

²² Ce jeu a été développé dans le cadre du projet H2020 Imprex (<http://www.imprex.eu/stepping-boots-of-a-flood-manager>) et, grâce à son succès, sera incorporé au produit final du projet (« Risk Outlook »), qui est en cours de développement par l'UK MetOffice.

Sans vouloir gâcher l'expérience de la découverte des jeux (les lecteurs plus curieux sont fortement encouragés à passer directement à la **section 2.4** et à ne pas lire les prochaines lignes avant de jouer eux-mêmes aux jeux cités), plusieurs résultats intéressants ont été mis en évidence à l'aide des expérimentations menées :

- Les décisions sont souvent basées sur une combinaison de ce qui est affiché comme la valeur plus probable (de la prévision) et ce qui est donné comme information sur l'incertitude autour de cette valeur.
- Des prévisions avec une même valeur probable mais des incertitudes différentes ne conduiront pas à la même décision, ce qui montre que l'information sur l'incertitude d'une prévision joue un rôle dans le processus de prise de décision. La tendance observée est que plus l'information d'incertitude est élevée, plus les décisions seront marquées par une aversion au risque.
- La présence de l'information sur l'incertitude des prévisions semble conduire à des réponses (décisions) plus cohérentes parmi les différents décideurs.
- En absence d'informations sur l'incertitude de prévision, les décideurs sont plus contraints d'adopter une attitude d'aversion au risque, ce qui conduit à une moins bonne performance économique dans la prise de décision par rapport à une situation où cette information est disponible (vis-à-vis d'une situation où l'aversion au risque a un coût lié à l'activation d'une protection qui s'avèrera, par la suite, inutile).
- Une meilleure gestion des apports en eau prévisionnels à un réservoir est obtenue lorsque des prévisions incertaines et leurs impacts sont pleinement compris et pris en compte de manière anticipée dans la prise de décision concernant la gestion des niveaux d'eau et des lâchers.
- La qualité des prévisions utilisées pour la prise de décision, telle que perçue par l'utilisateur, semble être un des facteurs influençant le consentement à payer pour avoir de nouvelles prévisions.
- Les prévisions sont généralement perçues comme des éléments clés du processus de décision. Cependant, l'interprétation qu'un utilisateur peut faire d'une prévision peut avoir autant d'importance que l'information contenue dans la prévision elle-même. La tendance à oublier cet aspect et à se fier « aveuglément » à ce qui est affiché par une prévision pourrait potentiellement expliquer pourquoi, dans plusieurs situations, les décideurs auraient tendance à blâmer (remercier) les fournisseurs des prévisions pour leurs propres mauvaises (bonnes) décisions.
- Les jeux de rôles sont très efficaces pour soulever des discussions sur les défis de l'utilisation des prévisions probabilistes dans la prise de décision, où les différents choix ont des conséquences différées et les effets de la décision prise sont irréversibles, et ce que ce soit dans le cas de décisions séquentielles en arbre ou dans les cas où les décisions prises à chaque instant de prévision sont indépendantes entre elles.

Ces travaux sous forme de jeux mettent en évidence l'importance de la qualité des prévisions pour un utilisateur et la valeur d'une communication efficace des incertitudes.

Et, enfin, pour répondre à la question du titre de cette section, « oui, c'est sûr et certain ! » (Essayez le jeu « Flood Control » pour en avoir la preuve !).

2.4. Une approche multidisciplinaire inévitable

Une approche intégrée de la prévision implique la connaissance de plusieurs aspects qui intègrent les différentes étapes de production d'une prévision opérationnelle : l'apport des scénarios météorologiques futurs en hydrologie, la description et la propagation des incertitudes dans la prévision des débits, l'évaluation de la qualité des prévisions émises, la communication efficace des prévisions incertaines pour leur intégration dans le processus de décision, que ce soit pour l'alerte aux crues/inondations ou la gestion des ressources en eaux, et l'évaluation de l'utilité (sociale ou économique) apportée par une bonne prévision.

La mise en place d'outils adaptés pour modéliser le flux de l'information probabiliste au sein d'une chaîne de prévision hydrométéorologique est indispensable pour pouvoir évaluer l'impact de la propagation des incertitudes sur la qualité et l'utilité opérationnelle des prévisions hydrologiques. Un système de prévision doit également être suffisamment flexible pour pouvoir incorporer les nouvelles techniques qui se développent continuellement pour améliorer la performance des prévisions et s'approprier, sans difficulté, les évolutions des systèmes de modélisation numérique en hydrométéorologie (par ex., prévisions météorologiques à haute résolution, modélisations hydrologiques distribuées ou semi-distribuées).

Pour la recherche en sciences hydrologiques, la construction de partenariats pour évaluer les meilleures méthodes disponibles et identifier les principales questions scientifiques et opérationnelles à traiter est une étape cruciale. Elle doit s'appuyer, d'une part, sur les centres météorologiques à l'origine des forçages météorologiques et, d'autre part, sur les services hydrologiques opérationnels intéressés par l'utilisation des prévisions hydrologiques à différentes échéances et concernés par les situations où les incertitudes de prévision jouent un rôle prépondérant sur les décisions opérationnelles. Une approche multidisciplinaire et participative permet de mieux identifier les différents éléments de base d'une chaîne de prévision et ainsi mieux cibler les démarches à mener pour répondre à des besoins spécifiques, exprimés par les défis sociétaux et les enjeux économiques et de sûreté d'aujourd'hui.

Les chapitres qui suivent portent sur deux éléments de base d'une chaîne intégrée de prévision hydrométéorologique : l'évaluation de la qualité (**Chapitre 3**) et l'évaluation de la valeur/utilité des prévisions (**Chapitre 4**). Ces deux éléments n'existent pas sans connexion à d'autres éléments-clés de la prévision, à savoir : l'interfaçage météorologie/hydrologie, la prise en compte des principales sources d'incertitude prédictive, le post-traitement statistique ou correction de biais des prévisions, et la mise en place de produits de visualisation et communication des prévisions incertaines.

3. Qualité des prévisions

3.1. Attributs et critères de la qualité d'une prévision

La qualité d'une prévision peut être évaluée à l'aide d'un large éventail de critères, aussi appelés *scores*. En contexte probabiliste ou d'ensemble, les critères d'évaluation de la qualité des prévisions hydrologiques sont fortement basés sur les critères développés en météorologie²³ (Wilks, 2011 ; Jolliffe et Stephenson, 2011 ; Casati *et al.*, 2008). Les approches visant à mieux répondre aux besoins spécifiques en hydrologie ou à faire face aux contraintes typiques des systèmes hydrologiques (par ex., évaluer des événements rapides de crues ou les faibles débits saisonniers fortement corrélés dans le temps) sont de plus en plus nombreuses (Laio et Tamea, 2007 ; Pappenberger *et al.*, 2008b ; Demargne *et al.*, 2009 ; Brown *et al.*, 2010 ; Liu *et al.*, 2011 ; Zappa *et al.*, 2013 ; Trambauer *et al.*, 2015 ; Yuan *et al.*, 2015).

Plusieurs facteurs peuvent jouer un rôle dans la qualité des prévisions hydrologiques et intéresser particulièrement des utilisateurs spécifiques en hydrologie, comme, par exemple, la localisation des amas pluvieux et intensités extrêmes au sein d'un bassin versant, la concomitance entre débits de pointe de crue prévu et observé, ou encore l'intensité et la durée des débits en dessous de certains seuils critiques d'étiage. La qualité d'une prévision hydrologique dépendra donc de la qualité des prévisions météorologiques utilisées en entrée du modèle hydrologique, de la qualité de la simulation du modèle hydrologique (en termes de production de ruissellement et routage vers l'exutoire) ainsi que de la qualité du débit mesuré (ou de la courbe de tarage hauteur-débit utilisée).

Que ce soit en météorologie ou en hydrologie, l'utilisation de plusieurs critères pour mieux évaluer la variété des attributs d'une prévision est fortement recommandée. Ces attributs sont très bien présentés dans le chapitre de Bradley *et al.* (2016), qui fait partie de la section « *Verification of Hydrometeorological Ensemble Forecasts* », que j'ai coéditée avec J. Demargne (Pays Bas) et N. Voisin (États-Unis), appartenant au livre « *Handbook of Hydrometeorological Ensemble Forecasting* ».

En me basant sur ce chapitre, j'ai fait une présentation synthétique pour illustrer le cadre général sur lequel se fondent souvent les évaluations qui cherchent à caractériser la qualité des prévisions (**Figure 5**). Il s'agit d'un exercice qui s'intéresse à la relation statistique entre les prévisions et les observations. Le cadre d'évaluation peut être appliqué au cas élémentaire où la prévision est une valeur unique (déterministe ou probabiliste) et l'observation est une valeur unique (discrète ou continue). Nous pouvons aussi imaginer le cas plus complexe où la prévision est une distribution de probabilité à évaluer par rapport à une observation. L'observation peut être représentée par une fonction de Heaviside (prenant la valeur unique 0 pour tous les cas inférieurs à l'observation et la valeur 1 partout ailleurs) ou une distribution de probabilité, si l'incertitude d'observation peut être quantifiée.

²³ On notera que la « vérification » des prévisions est souvent appelée « évaluation » en hydrologie, tandis qu'en météorologie, « évaluation » fait spécifiquement référence à la « valeur » d'une prévision, donc à la manière dont les prévisions aident l'utilisateur à prendre de meilleures décisions.

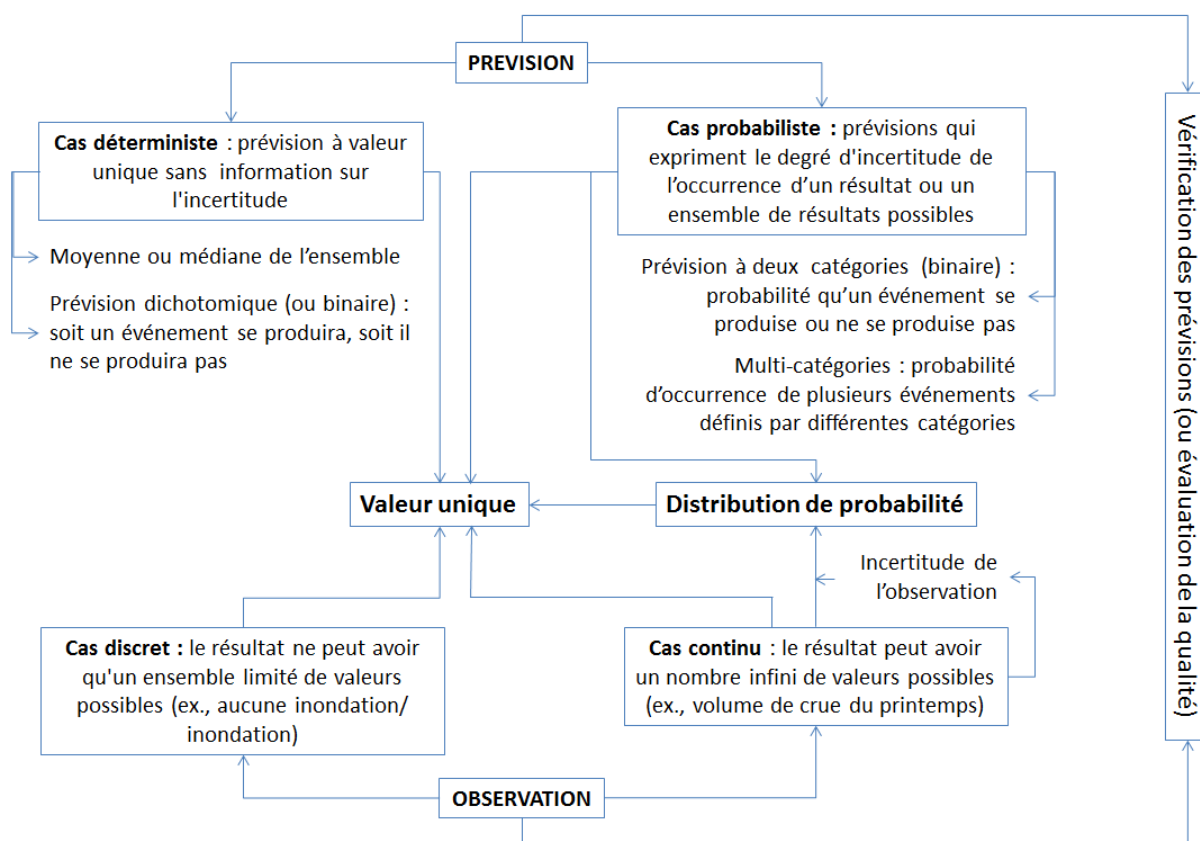


Figure 5 : Cadre général pour caractériser la qualité des prévisions sur la base de la relation statistique entre une prévision et l'observation qui lui est associée.

Le **Tableau 2** présente les principaux attributs qui décrivent les différents aspects de la qualité des prévisions, ainsi qu'une sélection de critères couramment appliqués aux prévisions hydrométéorologiques pour mesurer ces attributs de qualité. Ce tableau, publié dans Anctil et Ramos (2018), est basé sur l'article de blog de J. Brown et J. Demargne publié dans le portail HEPEX en 2013²⁴ et sur des définitions que l'on retrouve dans Bradley *et al.* (2016). On y distingue plusieurs attributs, tels que :

- la précision ou justesse (« *accuracy* »), qui correspond à la différence (ou distance) moyenne entre prévision et observation (dans une prévision probabiliste, cette distance peut être celle entre la distribution de probabilités prévue et celle de l'observation) ;
- la fiabilité (« *reliability* », parfois appelée « *calibration* »), qui indique dans quelle mesure un événement prévu avec une probabilité $p\%$ est effectivement observé $p\%$ du temps. Cet attribut permet de détecter des biais statistiques systématiques (par ex., des observations toujours ou très souvent inférieures aux valeurs prévues par les membres d'une prévision d'ensemble) ;

²⁴ <https://hepex.irstea.fr/hepex-science-and-challenges-verification-of-ensemble-forecasts-24/>

- la résolution (« *resolution* »), qui mesure la capacité de la prévision à trier l'ensemble des événements en sous-ensembles avec différentes distributions de probabilités. Cet attribut indique dans quelle mesure la prévision peut séparer différentes situations (par ex., quand on prévoit des débits forts, les débits observés sont effectivement forts, et quand on prévoit des débits faibles, les débits observés sont effectivement faibles). Cet attribut mesure à quel point les prévisions sont substantiellement différentes de la climatologie (historique des observations passées) ;
- la finesse (« *sharpness* »), qui mesure la dispersion de la prévision (ou l'étendue des valeurs prévues). Il s'agit d'un attribut indépendant de l'observation. La dispersion d'une prévision d'ensemble doit être représentative de l'incertitude prédictive ;
- la discrimination (« *discrimination* »), qui mesure la capacité de la prévision à associer une probabilité de prévision plus élevée à un événement quand celui-ci est observé.

On observe que plusieurs critères peuvent être associés à un même attribut. De même, plusieurs critères sont nécessaires pour pouvoir mesurer l'ensemble des attributs. On notera que, lors de l'évaluation de systèmes de prévision distincts, ceux-ci peuvent être classés comme meilleur, deuxième meilleur, etc., jusqu'au pire des systèmes selon un critère numérique donné, bien que ce classement ne soit pas nécessairement le même si un autre critère est choisi pour caractériser la performance de chaque système (Jolliffe et Stephenson, 2011).

La compréhension des liens entre critères numériques ou graphiques d'évaluation et attributs de la qualité des prévisions est nécessaire pour sélectionner des critères adaptés au système évalué et aux objectifs de la procédure d'évaluation (Anctil et Ramos, 2018). Cela permet d'éviter la « *misevaluation* » (tirer des conclusions erronées), la surévaluation (utilisation de critères redondants et efforts excessifs pour évaluer les mêmes qualités) ou la sous-évaluation (oublier l'évaluation d'un attribut qui pourrait mettre en évidence d'importants aspects liés aux forces ou faiblesses du système).

Indépendamment du fait que les métriques soient appliquées aux prévisions déterministes ou probabilistes des événements discrets (par ex., dépassements d'un seuil) ou continus, une évaluation robuste nécessite, au moins, des séries chronologiques longues de données homogènes sur la période d'évaluation. Ces séries, composées de paires « prévision-observation » doivent être représentatives de ce que l'on souhaite évaluer et, si un contexte décisionnel est en jeu, doivent renseigner les décideurs sur, au moins, les principaux attributs de qualité qui peuvent influencer le plus leurs décisions.

Tableau 2 : Quelques caractéristiques principales des critères d'évaluation couramment utilisés pour évaluer la qualité des prévisions hydrométéorologiques (dans : Anctil et Ramos, 2018)

Quality attribute	Metric name	Type of forecast	Discrete events?
Bias (difference between the mean forecast and the mean observation)	Relative Mean Error (or Relative Bias)	Single-valued	No
	Frequency Bias	Both	Yes
Accuracy (average difference between individual forecasts and observations)	Mean Absolute Error	Single-valued	No
	Mean Square Error	Single-valued	No
	Root Mean Square Error	Single-valued	No
	Mean Continuous Rank Probability Score (CRPS)	Probabilistic	No
	Brier Score	Probabilistic	Yes
Critical Success Index CSI (or Threat Score)	Both	Yes	
Correlation (linear relationship between forecasts and observations)	Pearson Correlation Coefficient	Single-valued	No
	Spearman Rank Correlation	Single-valued	No
Skill (accuracy of forecast relative to a reference forecast)	Mean Absolute Error Skill Score	Single-valued	No
	Mean Square Error Skill Score	Single-valued	No
	Mean Continuous Rank Probability Skill Score	Probabilistic	No
	Brier Skill Score	Probabilistic	Yes
	Equitable Threat Score (or Gilbert Skill Score)	Both	Yes
Reliability (agreement of forecast with observation conditioned on the forecast issued)	Mean Square Error Reliability	Single-valued	No
	Mean CRPS Reliability	Probabilistic	No
	Brier Score Reliability	Probabilistic	Yes
	Reliability Diagram	Probabilistic	Yes
	Rank Histogram	Probabilistic	Yes
	Success Ratio	Both	Yes
Resolution (differences in outcomes for different forecasts issued)	Mean Square Error Resolution	Single-valued	No
	Mean CRPS Resolution	Probabilistic	No
	Brier Score Resolution	Probabilistic	Yes
Type-2 conditional bias (agreement of observation with forecast conditioned on outcome)	Mean Square Error Type-2 Conditional Bias	Single-valued	No
	Brier Score Type-2 Conditional Bias	Probabilistic	Yes
Discrimination (differences in forecasts for different outcomes)	Mean Square Error Discrimination	Single-valued	No
	Brier Score Discrimination	Probabilistic	Yes
	Relative Operating Characteristic Score	Both	Yes
	Relative Operating Characteristic Diagram	Both	Yes
	Probability Of Detection (Hit Rate)	Both	Yes
	Probability Of False Detection (False Alarm Rate)	Both	Yes
Sharpness (degree of variability of the forecasts or concentration of the predictive distributions)	Forecast Frequency Histogram	Probabilistic	Yes
	Average width of the prediction intervals	Probabilistic	No

3.2. Ce que les critères peuvent nous indiquer sur la qualité des prévisions

Nous retrouvons de nombreuses applications des critères d'évaluation de la qualité des prévisions hydrologiques dans la littérature (Hashino *et al.*, 2007 ; Thirel *et al.*, 2008 ; Reggiani *et al.*, 2009 ; Bartholmes *et al.*, 2009 ; Wejis *et al.*, 2010 ; Randrianasolo *et al.*, 2010 ; Voisin *et al.*, 2011 ; Liechti *et al.*, 2013 ; Brown *et al.*, 2014 ; Engeland et Steinsland, 2014 ; Roulin et Vannitsem, 2015 ; Verkade *et al.*, 2017, pour ne citer que quelques exemples). Ces applications illustrent la variété d'objectifs, contextes et approches des études faites. Elles révèlent également une progression dans le temps de ce à quoi un exercice d'évaluation de la qualité des prévisions peut s'intéresser.

De manière générale, les premières applications des critères de qualité des prévisions en hydrologie ont principalement cherché à évaluer l'apport des prévisions d'ensemble (ou de la moyenne des membres de l'ensemble) par rapport à une prévision déterministe à scénario unique, ou à une prévision déterministe habillée par un modèle statistique des erreurs de calage du modèle hydrologique utilisé. Progressivement, les évaluations se sont intéressées à quantifier les améliorations que l'introduction d'une nouvelle technique, d'assimilation de données ou de post-traitement par exemple, pouvait apporter à la qualité des prévisions. Les études plus récentes semblent s'intéresser davantage à l'interprétation des critères par rapport aux attributs de qualité qu'ils mesurent ou à en créer des nouveaux, plus adaptés à l'évaluation des forces et faiblesses d'un système de prévision donné. Ces études s'intéressent également à quantifier les complémentarités entre systèmes de prévision et les compromis à faire quand on cherche à améliorer globalement la performance des prévisions sur un ensemble de bassins versants et pour un utilisateur bien ciblé.

Si, d'un côté, évaluer une prévision d'ensemble par rapport à une prévision déterministe permet de mieux se familiariser avec la prévision d'ensemble, une telle démarche n'est pourtant pas tout à fait satisfaisante intellectuellement. En effet, la richesse de la prévision d'ensemble, avec l'ensemble de ses membres et sa capacité à approcher la distribution prédictive par l'échantillonnage intelligente des scénarios futurs possibles, n'est pas exploitée à son maximum. Ceci dit, évaluer un système de prévision d'ensemble en soi, sans référence à laquelle se comparer, n'est pas une tâche entièrement satisfaisante non plus. Plusieurs critères (par ex., BS et CRPS²⁵) ont une valeur de score parfait (meilleur possible) qui vaut ce que l'on obtiendrait avec une « prévision déterministe parfaite » (si on pouvait l'avoir²⁶). Dans ce cas, évaluer si un système de prévision d'ensemble est « bon », dans l'absolu, nous laisse sans repères. Il est alors plus intéressant de pouvoir évaluer si un système est meilleur qu'un autre système de référence. Sauf à avoir deux (ou plusieurs) systèmes concurrents déjà disponibles à évaluer, le choix du système de référence à considérer devient lui aussi, un sujet d'étude en soi (Pappenberger *et al.*, 2015).

²⁵ Le BS (*Brier Score*) mesure l'erreur de probabilité quadratique moyenne de la prévision d'un événement donné et répond à la question « Quelle est l'ampleur des erreurs de probabilité de la prévision? ». Le CRPS (*Continuous Ranked Probability Score*) mesure la somme des différences quadratiques entre prévisions et observations dans l'espace de probabilité cumulée (version continue et multi-événement du BS) (Jolliffe et Stephenson, 2011).

²⁶ Wejis *et al.* (2010) notent qu'un utilisateur forcément affectera « *some probability to outcomes that the forecast did not speak about.* »

Dans mes recherches, j'ai parcouru quelques-uns de ces nombreux sentiers qui s'ouvrent quand on s'interroge sur la qualité des prévisions d'ensemble. A ma connaissance, avant mon arrivée à Irstea, les équipes de recherche en hydrologie n'avaient jamais utilisé les prévisions météorologiques (d'ensemble ou autres) dans leurs outils de modélisation hydrologique²⁷. Il s'agissait ainsi d'y apporter une nouvelle compétence.

La première application que j'ai faite en arrivant à Irstea (Centre d'Antony) s'est basée sur le modèle de prévision hydrologique GR3P (version journalière ; Tangara, 2005) et la prévision météorologique d'ensemble produite par le CEPMMT²⁸. Le but était d'évaluer la performance du système sur l'événement de crue de mars 2001 dans le bassin de la Seine (communication SHF : Ramos *et al.*, 2008 ; **Figure 6**). Ce fut une manière « douce » d'essayer la prévision d'ensemble avec l'un des modèles conceptuels globaux développés à Irstea pour le contexte opérationnel en France. Jusqu'alors, le modèle avait été testé avec seulement un scénario unique de pluie future observée (connue ou nulle). Le choix de l'événement de 2001 était fait pour pouvoir comparer, au moins visuellement, la performance du modèle conceptuel GR3P à celle du modèle à base physique SIM de Météo-France, qui avait également simulé cet événement et pour lequel les résultats avaient été publiés dans Rousset-Regimbeau *et al.* (2007).

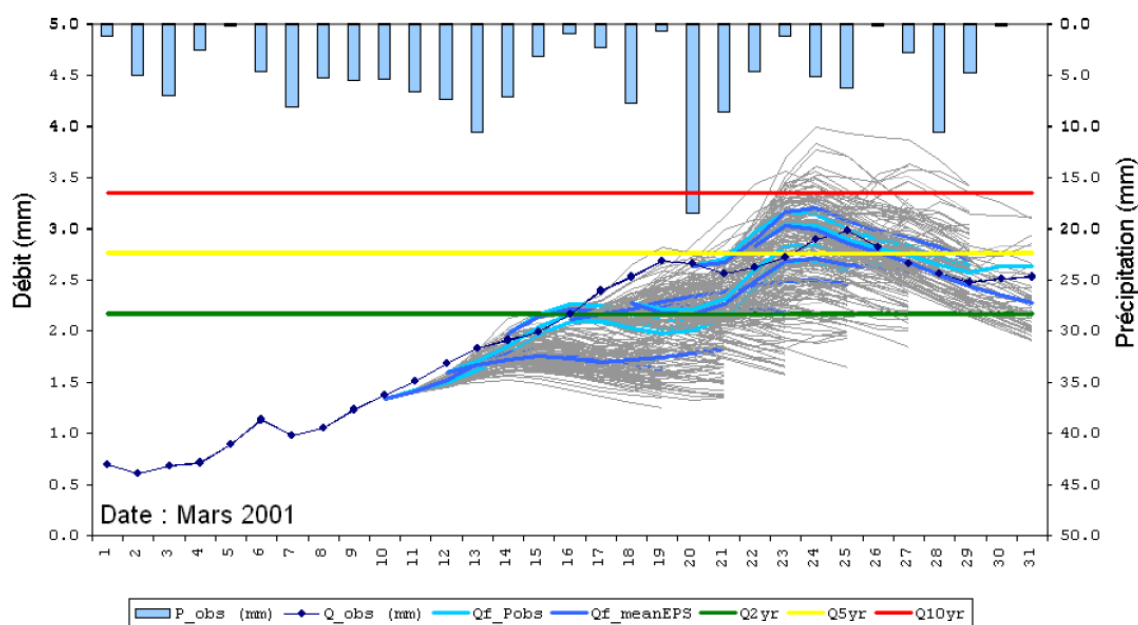


Figure 6 : Prévision hydrologique d'ensemble basée sur la prévision d'ensemble météorologique EPS du CEPMMT (50 membres) et le modèle GR3P (journalier) d'Irstea : cas de l'événement de mars 2001 sur la Seine à Paris (dans : Ramos *et al.*, 2008)

²⁷ La thèse d'Aurélien Ben Daoud (2010) avait démarré peu avant que je quitte le Centre d'Irstea à Lyon pour rejoindre le centre d'Antony. Elle s'est centrée sur la prévision de pluies par la méthode des analogues, avec une application pour la prévision de débits sur les bassins de la Saône et de la Seine. A. Ben Daoud avait alors utilisé la version « prévision d'ensemble » du modèle pluie-débit GR3P journalier que j'avais mis en place en arrivant à Irstea (Antony). Nous avons publié ensemble l'application sur le bassin de la Seine : Ben Daoud *et al.* (2011).

²⁸ Centre européen de prévision météorologique à moyen terme (ECMWF, en anglais) : <http://www.ecmwf.int/>

Avec l'automatisation de la chaîne de modélisation « GR3P-ensemble », j'ai pu ensuite passer de l'analyse événementielle à l'évaluation statistique de longues séries temporelles de prévisions hydrologiques d'ensemble. J'ai ainsi collecté des nouvelles bases de données de prévisions météorologiques d'ensemble et encadré plusieurs études pour pouvoir les exploiter (voir **Tableau 1, section 1.2**). Le but était d'examiner différents aspects de la prévision hydrologique d'ensemble basée sur différentes prévisions météorologiques d'ensemble (précipitations et températures), telles que celles du centre européen CEPMMT (Weeink, 2010 ; Velázquez *et al.*, 2011 ; Dhouioui, 2011 ; van Pham, 2011), de la base TIGGE (THORPEX WMO) (Zalachori, 2013 et séjour scientifique de T. Arandia-Martinez à Irstea en 2015), et de Météo-France (PEARP) (Randrianasolo *et al.*, 2010, 2011 ; Nicolle *et al.*, 2012 ; Berthet, 2010). Les évaluations menées ont pu également confronter différentes options de modélisation ou de techniques dans un système de prévisions d'ensemble (Zalachori *et al.*, 2012 ; Randrianasolo *et al.*, 2014 ; Bourgin *et al.*, 2014) et également cibler les échéances très courtes de la prévision immédiate (Caseri *et al.*, 2016a,b) ainsi que celles plus lointaines de la prévision saisonnière (Crochemore *et al.*, 2016b, 2017).

Ci-après, je présente quelques résultats sélectionnés, concernant particulièrement la qualité des prévisions d'ensemble et obtenus majoritairement à partir des recherches que j'ai dirigées (les thèses et Masters d'Irstea sont disponibles sur <https://webgr.irstea.fr/publications/>).

□ *Apport de la prévision d'ensemble à la prévision dans les bassins versants non jaugés : courte échéance et prévision immédiate*

La **thèse d'A. Randrianasolo (2009-2012)** s'est intéressée à des méthodes simples et robustes qui peuvent apporter de l'information pertinente pour quantifier les incertitudes de prévision dans les bassins versants non jaugés. Le but a été d'étudier la meilleure stratégie pour chercher l'information dans les bassins jaugés (« donneurs ») et pour la transférer vers un site cible non jaugé. Nous avons étudié cette recherche d'information sous deux angles : information nécessaire pour mettre en place un modèle de simulation pluie-débit (définition des paramètres) et information nécessaire pour effectuer une mise à jour des conditions initiales du modèle de prévision en temps réel (à partir de la dernière observation de débit disponible sur le site cible ou les sites voisins jaugés). Ces deux composantes de la prévision ont été ainsi découplées dans notre approche. L'étude a été menée sur 211 bassins versants, la prévision météorologique d'ensemble PEARP de Météo-France et les modèles hydrologiques conceptuels journaliers GR3P et GR4J. Les travaux ont été menés en collaboration avec Vazken Andréassian (Irstea, Antony).

Les principaux résultats ont montré que la fiabilité des prévisions peut être nettement améliorée sur un bassin versant non jaugé par l'utilisation de plusieurs jeux de paramètres issus du calage du modèle hydrologique sur des bassins versants jaugés voisins. La variabilité apportée par le voisinage géographique influe sur la dispersion des membres, et augmente ainsi la fiabilité des prévisions. Cependant, cette option n'améliore pas forcément la précision des prévisions. Une plus grande dispersion favoriserait ainsi la fiabilité, mais dégraderait la précision (ou justesse) des prévisions aux courtes échéances étudiées (et telle que mesurée par les critères évalués).

Les cartes de la **Figure 7** illustrent les résultats obtenus à l'aide du modèle GR3P journalier et selon quatre critères d'évaluation de la qualité des prévisions. Elles identifient les meilleurs donneurs de jeux de paramètres pour chaque bassin versant parmi : le voisin avec la meilleure performance en calage (LNB), et 1 à 20 voisins les plus proches géographiquement (L1N à L20N).

Les cartes des critères qui mesurent la précision des prévisions (NRMSE, pour la RMSE normalisée par l'écart-type des débits observés, et le critère CSI, « *critical success index* », issu d'un tableau de contingence établi pour le dépassement du quantile 90% des débits observés) associent la meilleure performance le plus souvent aux cas avec un seul jeu de paramètres, venant soit du bassin le plus proche (L1N), soit du meilleur voisin (LNB). Cela a été observé aussi bien dans le nord de la France, où la densité des bassins étudiés était plus importante, que dans le sud.

Au contraire, les cartes des critères probabilistes (Brier 90, BS calculé pour le dépassement du quantile 90%, et RPS, version discrétisée en multi-catégories du CRPS) indiquent, plus fréquemment, une meilleure performance associée à l'utilisation du transfert de plusieurs jeux de paramètres des bassins voisins (L10N à L20N). Ces critères semblent donc favoriser les gains de performance en termes de fiabilité. L'analyse d'autres critères qui mesurent la fiabilité seule (à savoir, le diagramme de PIT) a confirmé l'apport des membres de la prévision d'ensemble pour améliorer cet attribut important pour les systèmes de prévision probabilistes.

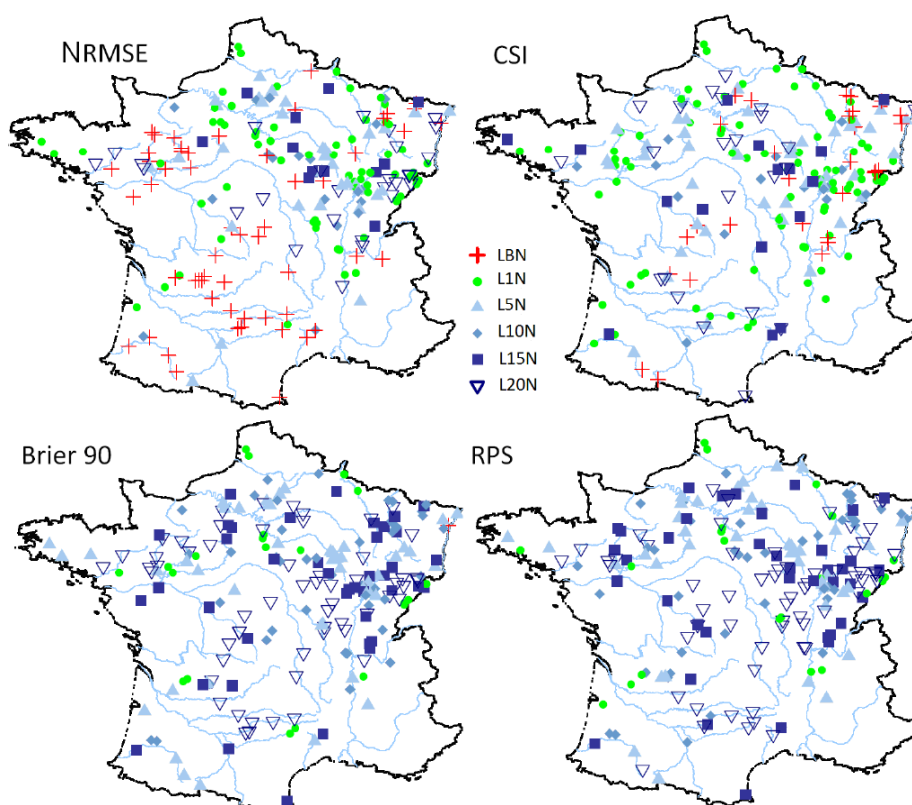


Figure 7 : Meilleurs donneurs des paramètres du modèle GR3P parmi les voisins (LBN, L1N,...L20N) du site non jaugé selon quatre critères de la qualité des prévisions hydrologiques d'ensemble (NRMSE, CSI, Brier 90 et RPS). Les paramètres sont transférés des sites voisins mais la mise à jour à l'instant de prévision est locale (dans : Randrianasolo, 2012)

L'étude a également montré que la précision des prévisions d'ensemble sur le site non jaugé peut être améliorée par l'intermédiaire du transfert (des bassins versants jaugés vers le site non jaugé) des corrections faites lors de la mise à jour des bassins voisins jaugés. Cela serait une manière d'améliorer la prise en compte des conditions initiales avant l'instant de prévision dans les sites non jaugés. La meilleure option pour améliorer la précision des prévisions serait de considérer les corrections effectuées sur le bassin le plus proche à chaque pas de temps de prévision (Randrianasolo *et al.*, 2011).

Tandis que la thèse d'A. Randrianasolo s'est intéressé à la prévision de débits journaliers à courte échéance, la **thèse d'Angélica Caseri (2013-2017)** a eu pour but la mise en place d'une méthode de prévision d'ensemble pour la quantification des incertitudes de la prévision immédiate des crues soudaines (également dans les bassins non jaugés). Cette thèse a été co-encadrée avec Pierre Javelle de l'UR RECOVER d'Irstea à Aix-en-Provence (financement projet RHYTMME).

Le but principal de la thèse a été d'étudier le potentiel d'une méthode de simulation conditionnelle géostatistique pour générer un ensemble de scénarios de pluies qui peuvent être utilisés par un système d'alerte aux crues soudaines. Il s'agissait de générer un ensemble de champs de pluie fiable (au sens statistique), tout en exploitant au mieux les points forts des mesures souvent disponibles pour la prévision immédiate : les propriétés spatio-temporelles fournies par les données radar et les intensités pluviométriques mesurées par les pluviomètres (**Figure 8**). L'étude s'est basée sur les données radar et pluviométriques de 17 événements pluvieux observés dans le département du Var. Son volet hydrologique s'est appuyé sur la modélisation distribuée AIGA (Javelle *et al.*, 2010), adaptée pour la prévision d'ensemble, pour la cartographie de la sévérité des débits (dont nous reparlerons dans le **Chapitre 4, section 4.2**).

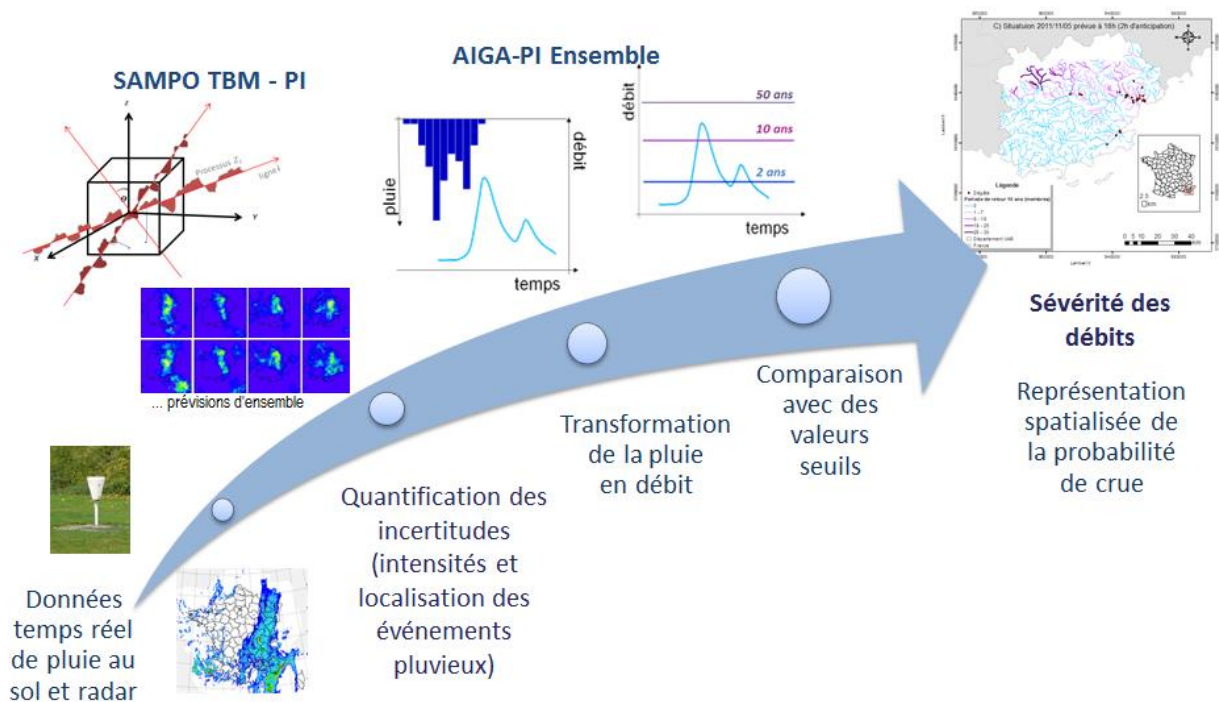


Figure 8 : Apport de la simulation conditionnelle géostatistique SAMPO TBM pour la prévision immédiate PI et l'alerte aux crues rapides (adapté de : Caseri, 2017)

Les travaux ont été menés en collaboration avec Etienne Leblois (Irstea, Lyon) pour ce qui est de l'adaptation de la méthode de simulation géostatistique SAMPO-TBM²⁹ pour considérer l'imagerie radar des champs de pluies et les propager par advection dans le temps pour la prévision immédiate. Les premiers efforts pour bâtir ce modèle de génération stochastique de champs de pluie ont été menés lors de ma thèse au LTHE à Grenoble (1998-2002, dirigée par J.-D. Creutin ; Ramos, 2002), quand nous avons adopté une approche basée sur la méthode géostatistique des bandes tournantes pour simuler des champs de pluies et mettre en évidence une approche unificatrice de différentes méthodologies d'analyse fréquentielle des pluies fortes. Ce travail impliquait la proposition d'une description générale de l'aléa pluviométrique selon son intensité, sa durée, sa fréquence et son support géométrique, c'est-à-dire, la surface sur laquelle la pluie est moyennée et le domaine de gestion opérationnelle où se trouve cette surface (Ramos *et al.*, 2006). Durant mes séjours à Irstea Lyon (2003-2005, 2007), j'avais repris les travaux avec E. Leblois, avec une implication mineure dans le suivi de la thèse de J.-M. Lepioufle (dirigée par E. Leblois et J.-D. Creutin). Après mon départ d'Irstea Lyon, d'importants développements ont été apportés au modèle, ce qui a conduit au simulateur SAMPO-TBM, utilisé aujourd'hui dans plusieurs projets et thèses à Irstea.

Les résultats de la thèse d'A. Caseri ont indiqué que l'approche développée était capable de simuler des scénarios de pluie fiables et représentatifs de l'organisation spatio-temporelle des champs de pluie observés par le radar et des intensités enregistrées par les pluviomètres. Notre approche géostatistique basée sur la simulation conditionnelle de champs de pluie représente une bonne option pour quantifier l'incertitude sur les intensités et la localisation des événements pluvieux observés (Caseri *et al.*, 2016b).

En prévision immédiate, l'approche est plus performante qu'une approche simple et naïve basée sur la persistance, dans les heures futures, du dernier champ de pluie enregistré par le radar météorologique. Les prévisions de pluie obtenues ont montré une bonne précision et performance globale, avec une performance légèrement meilleure et plus robuste pour une méthode de calage des paramètres du générateur par fenêtre glissante. Nous avons aussi proposé des méthodes d'ajustement des prévisions de pluie qui permettent d'exploiter davantage la dernière information de pluie disponible avant la prévision. Dans ce contexte, l'ajustement des paramètres du modèle de simulation géostatistique, suivi d'une correction par report d'erreur appliquée aux sorties du générateur de champs de pluie futurs, s'est avéré être l'approche la plus performante (Caseri *et al.*, 2016a et article en préparation³⁰).

□ *Evaluation de l'apport des prévisions déterministes habillées*

L'objectif des travaux de **Master 2 de Hajer Dhouioui (2011)** était de développer et évaluer une procédure permettant de produire des faisceaux de prévisions (prévisions d'ensemble) à partir d'une prévision déterministe à scénario unique et de la connaissance des statistiques des erreurs passées des prévisions. L'idée était d'examiner si une telle approche permettrait aux prévisionnistes qui ne disposaient pas d'une prévision d'ensemble d'en créer une à partir d'un seul scénario de prévision quantitative de pluie.

²⁹ Simulateur Amélioré, à Meso-échelle, des Précipitations et de leurs Occurrences (Irstea, Lyon)

³⁰ Article en préparation : « *A Conditional simulation based method for ensemble nowcasting using radar rainfall fields and raingauges* ».

Dans un premier temps, nous avons mis en place une méthode de perturbation (habillage ou « ensemble dressing », en anglais) des prévisions météorologiques de précipitation à scénario unique pour construire un ensemble de scénarios probables, en tenant compte des erreurs passées de la prévision déterministe. La méthode était basée sur un modèle Gaussien multivarié d'erreur, avec tirage aléatoire tenant compte de la corrélation entre les échéances pour respecter la cohérence temporelle des prévisions. Dans un deuxième temps, nous avons utilisé les différents scénarios construits en entrée au modèle hydrologique journalier GR3P, dans le but de générer une prévision de débits à plusieurs scénarios (construction d'une prévision hydrologique d'ensemble). Enfin, nous avons quantifié l'apport de la méthode développée à l'aide de critères d'évaluation de la qualité des prévisions.

Il s'agissait de mes premières recherches concernant le prétraitement statistique des entrées de pluie au modèle hydrologique. La méthode développée permettait d'inclure une information sur l'incertitude au scénario unique dont disposaient les prévisionnistes et de corriger les biais des prévisions déterministes à l'aide des scénarios générés. Dans cette étude, nous avons utilisé les prévisions déterministes et d'ensemble (50 membres) du CEPMMT, fournies, à l'origine, sur une grille de $0.5^\circ \times 0.5^\circ$ de résolution, émises à 12:00 UTC et s'étendant sur 240 heures d'horizon maximal de prévision. L'application a été faite sur 74 bassins versants français, allant d'environ 1 000 km² à 44 000 km², peu ou pas influencés par la neige.

Cette étude nous a éclairés sur les avantages et limites des approches d'habillage statistique pour construire une prévision d'ensemble. D'une part, nous avons réussi à générer des scénarios réalistes et cohérents en fonction de l'échéance de prévision, qui étaient comparables à ceux de la prévision météorologique d'ensemble du modèle numérique du CEPMMT. La performance de ces prévisions habillées était meilleure que celle de la prévision déterministe de départ (sur laquelle la méthode d'habillage avait été appliquée). La méthode mise en place nous a également permis d'améliorer la prévision de crues (dépassements des quantiles 75% et 95% des débits observés). De plus, l'évaluation de la qualité par échéance de prévision a permis d'illustrer l'effet de la mise à jour du modèle GR3P sur la première échéance journalière : les débits prévus étaient tellement rapprochés que les valeurs des critères d'évaluation différaient peu selon les méthodes (**Figure 9**).

D'autre part, l'étude nous a aussi appris que l'habillage était une méthode assez laborieuse et peu prête à la généralisation. La méthode requérait la vérification de plusieurs hypothèses sur la distribution des erreurs, dépendait de la disponibilité d'une période de calage pour le modèle d'erreur (donc d'une série de prévisions archivées issues du même modèle de prévision utilisé lors de l'application de la méthode) et pouvait laisser de côté plusieurs autres facteurs responsables d'erreurs importantes commises en prévision (par ex., erreurs de localisation des précipitations ou erreurs dans la répartition temporelle des quantités prévues ; Engeland *et al.*, 2010 ; Schaepli *et al.*, 2007 ; Vincendon *et al.*, 2011 ; Wu *et al.*, 2011).

Nous avons par ailleurs appliqué et transféré la méthode d'habillage au cours d'une action pour le SCHAPI en 2012 (exécutables construits pour pouvoir coupler la méthode au code GRP utilisé en temps réel). Dans cette étude de transfert, les prévisions utilisées ont été les prévisions de pluie de Météo-France au pas de temps de 3 heures (RR3). Les résultats obtenus avec l'habillage des RR3 n'ont cependant pas été très encourageants, avec beaucoup d'erreurs restant mal identifiées par le modèle d'erreur. A ma connaissance, la méthode n'a jamais été appliquée en temps réel.

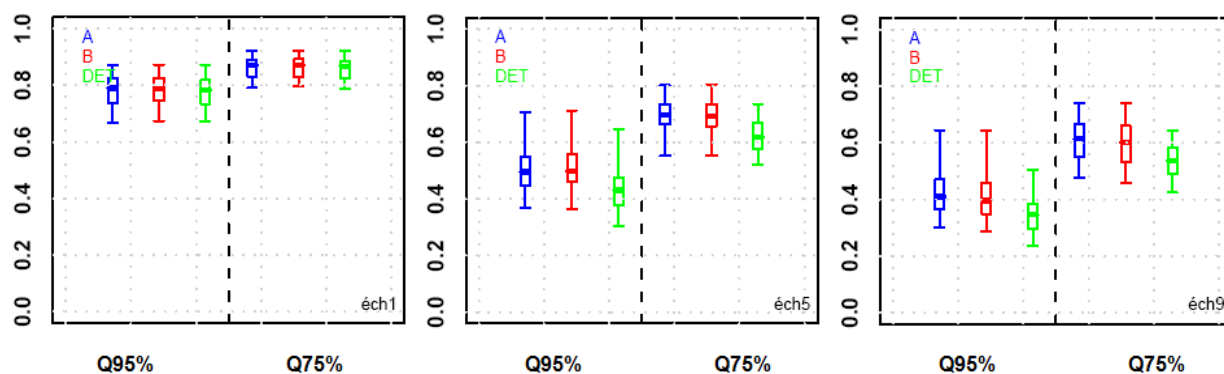


Figure 9 : Critère CSI (Indice de succès critique, en français) d'évaluation de la qualité des prévisions pour les échéances de 1 jour (gauche), 5 jours (milieu) et 9 jours (droite), appliqué à la prévision déterministe des débits (DET, en vert) et la prévision d'ensemble de débits dépassant les quantiles Q95% (à gauche de la ligne pointillée) et Q75% (à droite de la ligne pointillée) (au moins 80% des membres), basée sur la prévision déterministe de pluie habillée (habillage à partir de deux modèles d'erreur testés, A et B, en bleu et rouge, respectivement). Statistiques sur 74 bassins versants. La meilleure performance est celle dont le critère est plus proche de 1.0 (dans : Dhouioui, 2011)

□ Evaluation de la qualité de débits prévus à partir de modélisations hydrologiques différentes

Également dans le cadre d'une action pour le SCHAPI, j'ai encadré les travaux de **Master 2 d'Annie Randrianasolo (2009)**. Il s'agissait de la première étude que j'avais menée dans le but d'évaluer l'utilisation de la prévision d'ensemble PEARP de Météo-France en hydrologie. Pour avoir une référence de comparaison de la qualité des prévisions hydrologiques, nous nous sommes intéressés à évaluer les sorties de débits de deux modèles différents : le modèle hydrologique conceptuel, global GR3P (Irstea) et la chaîne de modélisation distribuée à base physique SIM (Météo-France). Les débits du modèle SIM avaient été générés par Guillaume Thirel pendant sa thèse de doctorat à Météo-France (Thirel, 2009), et les travaux ont été réalisés dans le cadre d'un protocole de coopération entre le SCHAPI, Irstea et le CNRM.

Les deux modélisations étaient ainsi alimentées par une même entrée de précipitation prévue (11 membres d'ensemble PEARP, jusqu'à deux jours d'échéance et période d'évaluation de 17 mois). Le but était d'étudier la performance (qualité) des prévisions hydrologiques d'ensemble sur un large échantillon de bassins versants en France (211 bassins), ainsi que de mieux comprendre les approches et critères utilisés couramment pour évaluer les prévisions d'ensemble en météorologie (alors peu utilisés en hydrologie).

Les résultats ont été publiés dans Randrianasolo *et al.* (2010). Ils ont montré des meilleures performances pour les prévisions hydrologiques du modèle GR3P. Il a été observé que cette meilleure performance était en grande partie due à la mise à jour effectuée par le modèle aux instants de prévision, à l'aide de la dernière observation de débits (à l'époque, le modèle SIM n'avait pas encore un module d'assimilation de données hydrologiques). Une fois la mise à jour du modèle GR3P désactivée, les deux modèles affichaient des performances plutôt équivalentes, surtout à l'échéance plus lointaine de 2 jours, avec des critères légèrement meilleurs et moins variés parmi les bassins versants pour le modèle GR3P.

Enfin, nous avons également montré que, pour la configuration PEARP-GR3P étudiée, l'évaluation des prévisions effectuée sur une période supérieure à 24 mois de débits journaliers était nécessaire pour fournir des valeurs moyennes des critères proches des valeurs finales qui seraient obtenues avec une période de vérification plus longue (Randrianasolo *et al.*, 2014).

De manière générale, les études utilisant la PEARP ont également montré une sous-dispersion des prévisions d'ensemble à la fois des pluies et des débits, mettant en évidence une sous-représentation des incertitudes de prévision et la nécessité d'appliquer des post-traitements statistique pour avoir des prévisions de débits fiables au final.

□ Influence de l'assimilation de données hydrologiques sur la qualité des prévisions

Une deuxième étude comparative des performances du modèle GR3P et SIM a été menée avec, cette fois-ci, des débits prévus par SIM avec assimilation de données. Elle a été publiée par Randrianasolo *et al.* (2014). Nous avons pu confirmer et mieux quantifier l'impact de l'assimilation de données sur la qualité des prévisions des deux modèles. En général, lorsque l'assimilation de données est effectuée, la qualité des prévisions augmente, avec des différences entre valeurs médianes de performance (sur un ensemble de 86 bassins versants étudiés), entre modélisation avec et sans assimilation de données, de l'ordre de 16%, variant selon le bassin versant, l'échéance et le critère utilisés pour évaluer les prévisions.

Nous avons également observé que, bien que le modèle GR3P présentait des performances globalement meilleures que le modèle SIM, il perdait relativement plus en performance quand son module de mise à jour était désactivé que le modèle SIM dans le même cadre d'évaluation (**Figure 10**). De plus, la méthode d'assimilation de données du modèle SIM, basée sur la modification de la teneur en eau du sol à partir du dernier débit observé, a montré un effet positif sur la qualité des prévisions plus durable avec l'échéance que la méthode utilisée dans GRP, qui modifie l'état du réservoir de routage sur la base d'une mise à jour directe de son niveau pour obtenir une valeur de débit modélisée égale à la celle observée à l'instant de la prévision. Bien qu'analysées, des différences importantes en performance entre les deux modèles en fonction des surfaces des bassins versants ou régimes hydrologiques n'ont pas été observées.

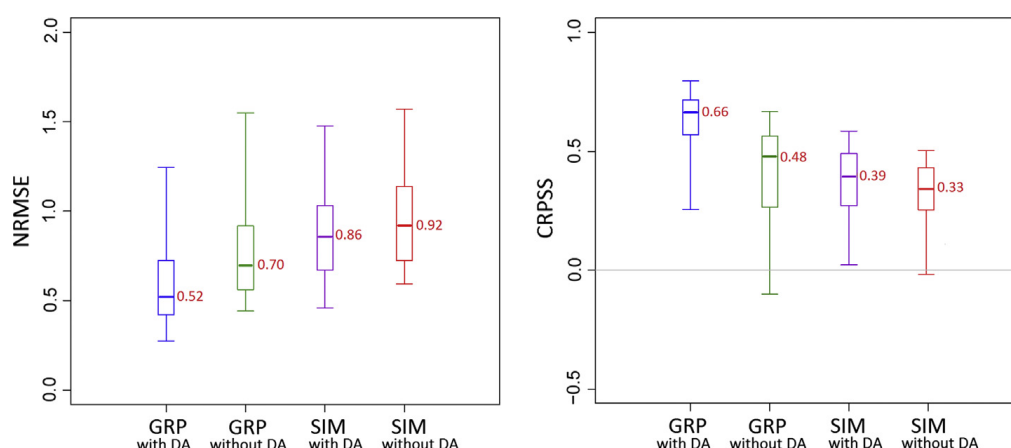


Figure 10 : Critères NRMSE et CRPSS (skill score par rapport à une prévision naïve constituée des valeurs historiques de débits observés avant la période d'évaluation) pour les prévisions hydrologiques d'ensemble à l'échéance de 2 jours des modèles GR3P et SIM, avec et sans assimilation de données, utilisant la PEARP en entrée. Evaluation sur 86 bassins versants et la période du 10/03/2005 au 31/07/2006. Les meilleures performances sont celles qui ont la plus petite valeur de NRMSE et la plus grande valeur de CRPSS (dans : Randrianasolo *et al.*, 2014)

J'ai également contribué à des études qui ont examiné l'influence de l'assimilation de données sur la qualité des prévisions hydrologiques au pas de temps horaire (contributions aux **thèses de Lionel Berthet, 2010 et François Bourgin, 2016**).

Dans Berthet *et al.* (2009), nous avons étudié l'impact de l'assimilation de données (mise à jour du réservoir de routage) et de la correction des sorties (régression simple et ARIMA), utilisées pour améliorer la qualité de la prévision hydrologique du modèle GRP (Berthet, 2010 ; modèle au pas de temps horaire et à 3 paramètres), sur un jeu de 178 bassins versants français. Dans cette étude, nous n'avons pas introduit de prévisions de précipitation, afin de regarder seulement la contribution de la modélisation hydrologique à la qualité des débits prévus. L'entrée au modèle GRP était constituée de « prévisions parfaites », c'est-à-dire que les précipitations observées à chaque échéance étaient considérées comme les précipitations prévues. Plus particulièrement, nous nous sommes intéressés à l'impact de la disponibilité de la donnée en temps réel pour effectuer la mise à jour. Nous avons ainsi évalué les pertes de performance du modèle GRP lorsque le délai entre la dernière observation de débit disponible et le moment où la prévision est émise augmente.

Les résultats ont montré une meilleure performance de la mise à jour de l'état du modèle (réservoir de routage) par rapport aux corrections des sorties, dont l'influence se dissipait après environ 12 heures de prévision dans la majorité des bassins versants étudiés. La mise à jour de l'état du réservoir de routage s'est révélée efficace même lorsque les données observées ne sont disponibles pour l'assimilation qu'une fois tous les cinq jours.

Dans Bourgin *et al.* (2014), nous nous sommes intéressés à étudier l'influence de l'assimilation de données sur les débits prévus, mais aussi les interactions avec un module de post-traitement statistique. L'assimilation de données suivait le même principe que celle utilisée dans Berthet *et al.* (2009) (mise à jour du réservoir de routage). Le post-traitement, appliqué aux débits prévus, était basé sur les erreurs de calage du modèle hydrologique. Ce module avait été implémenté au sein d'une chaîne de prévision hydrologique d'ensemble composée de la prévision météorologique d'ensemble PEARP et du modèle hydrologique GRP horaire. L'étude a porté sur un échantillon de 212 bassins versants en France.

Nos résultats ont clairement montré que les deux stratégies (assimilation de données et post-traitement) ont des effets complémentaires. L'assimilation de données a principalement un effet positif sur la précision des prévisions. Son impact diminue avec l'augmentation de l'échéance de prévision. Le post-traitement a un effet positif, et plus durable, surtout sur la fiabilité des prévisions (**Figure 11**). En conséquence, l'utilisation des deux techniques était fortement recommandée pour la prévision hydrologique d'ensemble.

Les révisions reçues sur l'article Bourgin *et al.* (2014) ont été très intéressantes et m'ont fait réfléchir à la distinction entre ce qu'on appelle « assimilation de données » et « post-traitement statistique ». J'ai ainsi publié un article dans le blog HEPEX à ce sujet : « *Interactions between data assimilation and post-processing* », publié le 4 août 2015³¹. La discussion a été courte, mais néanmoins intéressante, faisant ressortir, en effet, une « zone grise » en prévision hydrométéorologique où les méthodes peuvent se ressembler. Des différences ont été surtout mentionnées par rapport à la technique que l'on utiliserait : minimiser les erreurs d'analyse serait plus commun dans l'assimilation de données, tandis que les méthodes de post-

³¹ <https://hepex.irstea.fr/interactions-between-data-assimilation-and-post-processing/>

traitement seraient plutôt caractérisées par des étapes d'entraînement (ou calage), basées sur l'évaluation de la qualité de prévisions passées.

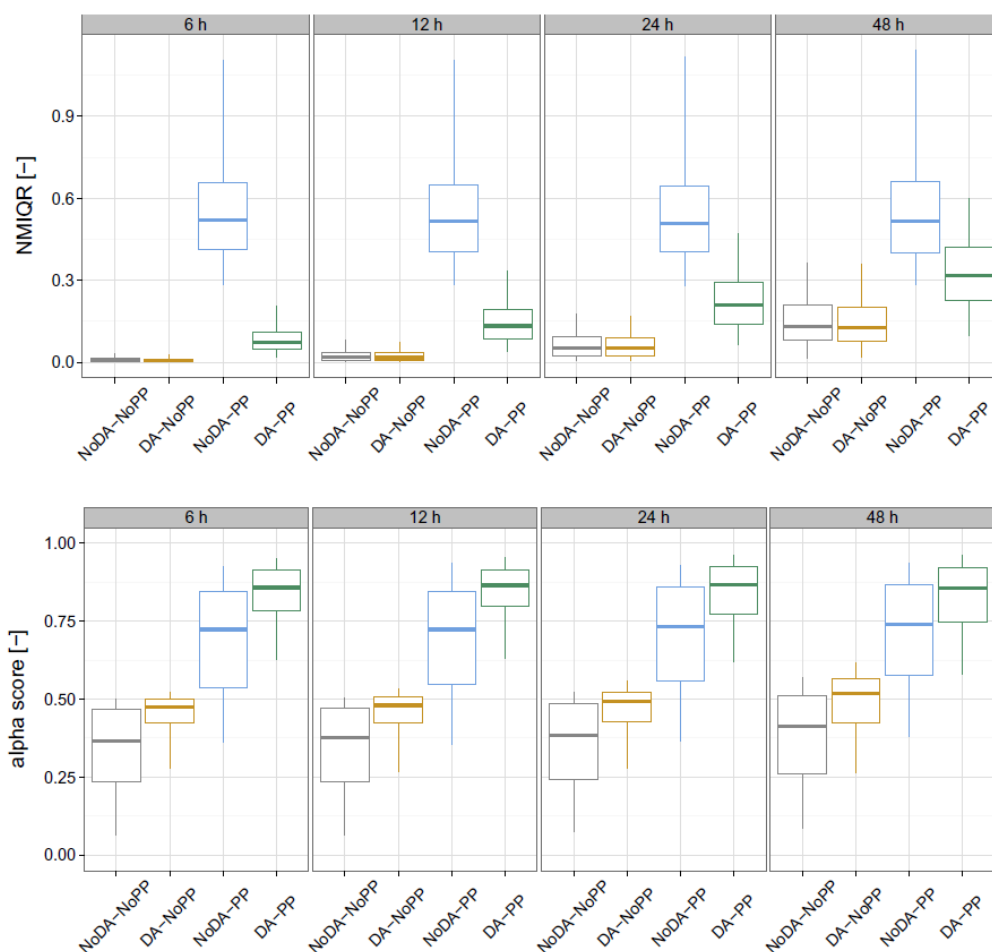


Figure 11 : Performances du système de prévision PEARP-GRP avec et sans assimilation de données (DA et NoDA) et avec ou sans post-traitement (PP et NoPP), selon un critère de mesure de la dispersion (ou finesse) de l'ensemble (distance interquartile moyenne normalisée NMIQR, en haut) et un critère de mesure de la fiabilité de l'ensemble (distance alpha du diagramme de PIT à sa diagonale, en bas). Pour le premier critère, le meilleur score est la plus petite valeur ; pour le deuxième, la valeur plus proche de l'unité. Evaluation des débits prévus sur 212 bassins versants et quatre échéances horaires (6h, 12h, 24h, 48h) (dans : Bourgin et al., 2014)

□ Evaluation de l'apport d'approches multi-modèles

Dans les approches multi-modèles, qu'il s'agisse de plusieurs modèles de prévision météorologique ou plusieurs modèles de prévision hydrologique, l'idée est de pouvoir prendre en compte les incertitudes dues à la structure/paramétrisation des modèles de prévision, et parfois aussi celles dues aux différentes méthodes d'assimilation des données (ou initialisation) qui sont particulières à chaque modèle.

En météorologie, il existe deux grandes approches multi-modèles (Myne, 2000). L'approche appelée prévision d'ensemble « du pauvre »³² regroupe plusieurs prévisions déterministes de différents modèles numériques de prévision du temps. Cette approche s'est montrée efficace pour la courte échéance (horizon de prévision inférieur à 2 jours ; Ebert, 2001 ; Arribas *et al.*, 2005), mais a montré aussi quelques inconvénients liés, notamment, au nombre souvent très limité de modèles à disposition pour une interprétation probabiliste, la difficulté (voire impossibilité) d'estimer de manière homogène les erreurs des différents modèles, ainsi que le fait que tous les modèles peuvent avoir des « mauvaises » prévisions en même temps s'ils partent de la même analyse globale de l'état initial de l'atmosphère.

La deuxième approche, appelée « super ensembles », regroupe plusieurs prévisions météorologiques d'ensemble de différents centres de prévision météorologique. L'exemple le plus connu est celui de la base TIGGE (pour *THORPEX Interactive Grand Global Ensemble*) de l'Organisation Mondiale de la Météorologie (OMM). Le projet TIGGE a été mené afin d'examiner l'efficacité d'une approche multi-modèle pour la prévision à moyenne-échéance. TIGGE propose plusieurs prévisions d'ensemble de modèles météorologiques indépendants à travers le monde, avec des caractéristiques différentes et des approches différentes de construction des ensembles (Bougeault *et al.*, 2010 ; **Figure 12**).

Centre	Ensemble members	Output data resolution	Forecast length	Forecasts per day	Fields (out of 73)	Start date
BoM* (AU)	33	TL119 (1.5°; 210km)	10 day	2	55	10 Mar 07
CMA (CHI)	15	T213 (0.56°; 70km)	10 day	2	60	15 May 07
CMC (CAN)	21	600x300 (0.6°; 75km)	16 day	2	56	10 Mar 07
CPTEC (BR)	15	T126 (0.94°; 120km)	15 day	2	55	1 Feb 08
ECMWF (EU)	51	TL639 (0.28°; 35km) TL319 (0.56°; 70km)	15 day	2	70	1 Oct 06
JMA (JAP)	51	TL479 (0.38°; 40km)	9 day	2	61	1 Oct 06
KMA (KOR)	24	N320 (0.56°; 70km)	10 day	2	46	28 Dec 07
Météo-France	35	TL358 (stretched 2.4)	4.5 day	2	62	25 Oct 07
NCEP (USA)	21	T254 (0.70°; 90km) T190 (0.95°; 110km)	16 day	4	69	5 Mar 07
UKMO (UK)	24	N216 (0.70°; 90km)	15 day	2	72	1 Oct 06

* Delivery of BoM data currently suspended

Figure 12 : Quelques caractéristiques techniques de la base TIGGE (d'après E. Zsótér, « The TIGGE archive and its applications », juin 2014)³³.

³² « Poor man's ensemble », en anglais.

³³ <https://www.ecmwf.int/sites/default/files/elibrary/2014/13760-tigge-archive-and-its-applications.pdf> ; dernière consultation le 28/12/2017.

Les premiers résultats de comparaison de la performance de chaque modèle de la base TIGGE en météorologie avaient indiqués que, contrairement à l'échelle saisonnière, où il peut être difficile de définir un meilleur modèle unique qui surpasse tous les autres modèles sur pratiquement tous les aspects, à la moyenne échéance, il était beaucoup plus facile d'identifier un modèle unique qui est clairement supérieur à tous les autres modèles (Hagedorn *et al.*, 2012). La question pour la moyenne-échéance était, d'ailleurs, bien posée par Hagedorn *et al.* (2012) : « *Does adding information from less skilful models to the best model lead to a more skilful multimodel ensemble?* » La combinaison multi-modèle est ainsi vue comme une méthode de post-traitement statistique des prévisions, qui serait une alternative au besoin de calibrer/post-traiter individuellement les prévisions d'ensemble de chaque centre météorologique.

Dans mes recherches, j'avais examiné deux approches pour la prévision « multi-modèle » de débits : (1) une approche basée sur la base mondiale TIGGE de prévisions d'ensemble de plusieurs centres météorologiques et (2) une approche basée sur l'utilisation de plusieurs modèles hydrologiques, forcés par une prévision d'ensemble d'un seul centre météorologique.

- *Approche basée sur le multi-modèle météorologique « super ensemble »*

Dans la **thèse de Ioanna Zalachori (2009-2013)**, nous nous sommes intéressés à déterminer si l'utilisation d'un grand ensemble de prévisions météorologiques d'ensemble peut corriger les faiblesses en performance liées à l'attribut de fiabilité des prévisions d'ensemble (météorologiques et hydrologiques) à la moyenne-échéance (jusqu'à 10 jours). Ces faiblesses se manifestent souvent sous forme de prévisions sous-dispersées (c'est-à-dire, avec faible dispersion des membres et un biais statistique), notamment sur les échéances les plus proches de l'initialisation de la prévision.

Au moment où ces travaux de thèse ont été lancés, l'utilisation de la base de prévisions météorologiques d'ensemble TIGGE en hydrologie était très récente et avait jusqu'alors concerné notamment quelques bassins versants et/ou périodes courtes d'évaluation (Pappenberger *et al.*, 2008a ; He *et al.*, 2010)³⁴. Les perspectives énoncées par les différents auteurs soulignaient l'importance d'appliquer la base de prévision TIGGE sur un échantillon plus large de bassins versants et d'examiner des combinaisons pondérées des prévisions tirées des différents centres météorologiques pour la construction du « super-ensemble ».

Les travaux d'I. Zalachori ont été menés avec un jeu de 74 bassins versants, le modèle GR3P journalier et les prévisions de précipitation de 8 centres météorologiques internationaux, avec un nombre de membres par centre variant de 14 à 50 (membres perturbés des prévisions d'ensemble utilisées), sur la période d'octobre 2006 jusqu'à septembre 2008. L'évaluation des prévisions hydrologiques a montré qu'aucun centre météorologique n'était le « plus performant de tous » pour l'ensemble des bassins versants étudiés. Les prévisions du centre européen CEPMMT ont néanmoins montré de meilleurs critères de qualité sur un plus large nombre de bassins versants, et ce notamment à partir de l'échéance de 3 jours.

³⁴ Voir aussi l'article de blog proposé par F. Pappenberger, « *TIGGE and the multi-model approach in hydrology (a brief review)* », publié le 5 novembre 2013: <https://hepex.irstea.fr/tigge-and-the-multi-model-approach-in-hydrology-a-brief-scientific-review/>

La combinaison des prévisions des différents centres météorologiques (avec ou sans pondération par rapport à la performance passée de chaque ensemble de prévision) pouvait apporter une amélioration à la qualité des prévisions en termes de fiabilité et discrimination des événements forts. Cette amélioration était plus importante pour les horizons plus courts de prévision. En effet, aux horizons de prévision plus lointains, la différence entre la performance du meilleur centre et la performance du grand ensemble était pratiquement négligeable.

Malgré l'amélioration de la performance des prévisions de débits aux horizons plus courts, les prévisions restaient néanmoins avec un important défaut de fiabilité. A nouveau, nous avons constaté le fait que la mise à jour de GR3P « écrase » la dispersion des prévisions d'ensemble au début de la prévision. Bien que cela favorise la précision des prévisions aux courtes échéances, le fait que l'observation assimilée soit considérée « parfaite » (sans incertitude) ne permet pas d'intégrer dans la prévision les erreurs associées aux conditions initiales hydrologiques. Cela impacte négativement la fiabilité des prévisions et laisse entrevoir le besoin de mieux prendre en compte l'incertitude des conditions initiales du modèle hydrologique ou d'introduire une démarche de post-traitement des débits prévus pour les rendre fiables.

En complément à ce travail, j'ai aussi apporté une petite contribution aux travaux de **master de F. T. Arandia Martinez** (« Conception d'un système de prévisions hydrologiques d'ensemble multimodèle en contexte opérationnel », soutenus en 2015 et encadrés par Marie-Amélie Boucher à l'Université du Québec à Chicoutimi). Pendant son séjour de mai à juillet 2014 à Irstea Antony, nous avons examiné l'apport d'une méthode bayésienne (BMA, *Bayesian Model Averaging*; Raftery *et al.*, 2005) pour combiner les prévisions de débits basées sur les prévisions météorologiques des centres TIGGE en une seule distribution prédictive.

La méthode BMA a été appliquée au même jeu de bassins versants de la thèse d'I. Zalachori³⁵. L'exploitation des résultats a néanmoins été bien décevante, car de nombreux problèmes sont apparus lors du traitement des séries de débits prévus, liés, par exemple, à la convergence des calculs, à l'utilisation d'une fenêtre mobile pour déterminer les paramètres du post-traitement, à la stabilisation des poids de la combinaison BMA et leur définition à travers les horizons de prévision, et à la reconstitution de la dépendance temporelle entre les distributions marginales du BMA pour différents horizons de prévision.

Depuis les travaux cités, je n'ai pas poursuivi dans cette voie d'une approche multi-modèle météorologique. On notera cependant que des études plus récentes confirment le potentiel de la base TIGGE à améliorer la fiabilité des prévisions et montrent qu'une combinaison de la base TIGGE avec des modèles à haute résolution peut fournir des ensembles précis (attribut qui viendrait des modèles à haute résolution), avec une dispersion adéquate (attribut qui viendrait de la diversité de la base TIGGE) (Wu *et al.*, 2016 ; Fan *et al.*, 2015a).

La mise en place récente de la base TIGGE-LAM³⁶, une extension des archives TIGGE pour inclure les prévisions météorologiques des ensembles de modèles à aire limitée (LAM), pourrait être une opportunité pour développer de nouvelles approches d'utilisation d'un grand ensemble en prévision hydrologique. Swinbank *et al.* (2016) notent que le succès de la base TIGGE a permis de prolonger le projet jusqu'à 2019, mais aussi que la grande majorité des centres météorologiques continueront à fournir des prévisions d'ensemble (archivées avec un

³⁵ Le package R « ensembleBMA » avait été utilisé.

³⁶ <https://www.ecmwf.int/en/research/projects/tigge> (dernière consultation le 28/12/2017)

délai de quelques jours) à des fins de recherche. Nous pouvons alors nous attendre à ce que des nouvelles études voient le jour et que le « TIGGE Museum », tenu par Dr. Mio Matsueda (University of Tsukuba et University of Oxford), maintienne ses portes virtuelles ouvertes au public³⁷.

- Approche basée sur le multi-modèle hydrologique

C'est avec une collaboration québécoise également que j'ai travaillé sur cette deuxième voie d'approche multi-modèle. Cette collaboration a été menée avec François Anctil (Université Laval) dans le cadre de la **thèse de Juan-Alberto Velázquez** (soutenue en 2010 et à laquelle j'ai contribué pendant son séjour de 6 mois à Irstea Antony en 2009). Dans Velázquez *et al.* (2011), nous présentons les résultats d'une étude menée sur 29 bassins versants en France, 16 modèles hydrologiques conceptuels et la prévision météorologique déterministe et d'ensemble (50 membres) du CEPMMT.

Les résultats ont montré que les prévisions hydrologiques d'ensemble produites par une combinaison de plusieurs structures de modèles hydrologiques et une prévision météorologique d'ensemble (dans notre cas, un système à $16 * 50 = 800$ membres) ont des performances plus élevées que les prévisions d'ensemble basées sur un unique modèle hydrologique alimenté par la prévision météorologique d'ensemble (50 membres) ou sur plusieurs modèles hydrologiques alimentés par une prévision météorologique déterministe (16 membres).

Le multi-modèle hydrologique à 16 membres a montré des bons résultats en termes de précision sur les courtes échéances (jusqu'à 3 jours). Le grand ensemble a été le seul système qui a fourni des prévisions de débit fiables. La fiabilité n'a cependant été atteinte qu'à partir de 4-5 jours d'échéance. Cela indique que, même avec une diversité de modèles hydrologiques et la quantification des incertitudes des prévisions météorologiques, le grand ensemble ne capture pas toutes les sources d'incertitude affectant la courte échéance. Cela pourrait être dû à l'ensemble météorologique du CEPMMT, qui est calé pour bien reproduire les incertitudes au-delà de 3 jours d'échéance, ou à l'absence de prise en compte des incertitudes des conditions initiales (états internes des modèles) dans les configurations étudiées.

Ce travail a pratiquement marqué le début d'une longue collaboration scientifique avec F. Anctil, qui perdure jusqu'à nos jours sur des sujets liés à la prévision d'ensemble et à la modélisation hydrologique. Entre 2010-2015, j'ai été examinatrice extérieure/rapporteur de plusieurs thèses dirigées par F. Anctil (thèses de M.A. Boucher, G. Seiller et A. Thibault). J'ai effectué un séjour scientifique de six mois à l'Université Laval en 2013. Nous avons également collaboré étroitement dans l'organisation de l'atelier HEPEX en 2016 à Québec et à l'écriture d'un chapitre de livre sur l'évaluation de la qualité des prévisions (Anctil et Ramos, 2018). En 2017, nous avons initié une collaboration sur le couplage d'un module d'évapotranspiration réelle calé à partir de données observées (tours à flux) et le modèle hydrologique horaire GR4H. Le stage de Master de Daniela Peredo (UPMC) que j'ai encadré en 2017 sur ce sujet a permis de prendre en main le modèle MEP (*Maximum Entropy Production*), mis en place à l'Université Laval en coopération avec des chercheurs américains, et d'examiner son potentiel sur des données du bassin versant expérimental de l'Orgeval (Irstea Antony, Equipe ORACLE).

³⁷ <http://gpvjma.ccs.hpcc.jp/TIGGE/> (dernière consultation le 28/12/2017)

Enfin, sur le sujet concernant les approches multi-modèles, j'ai aussi partagé mon expérience et questionnements scientifiques avec la communauté HEPEX, en contribuant à l'article de blog « *Multi-model approaches for river flow forecasting: blessing or burden?* », publié le 1^{er} novembre 2013 et co-écrit avec François Anctil et Florian Pappenberger³⁸. Dans cet article, nous soulevons quelques interrogations concernant la recherche et l'application des approches multi-modèles. Notamment, nous faisons ressortir les questions liées au fait que les approches multi-modèles peuvent facilement converger vers l'infini (nous avons atteint les 800 membres avec le travail de J. A. Velázquez, sans utiliser des approches multi-modèles météorologiques d'ensemble du type TIGGE). Les prévisionnistes opérationnels curieux de savoir si le modèle qu'ils utilisent est vraiment « le meilleur » et souhaitant tester une combinaison de modèles, peuvent vite se retrouver avec une importante charge de travail au moment de la prévision opérationnelle, notamment si leur système de prévision est peu automatisé. De plus, la question de la diversité des modèles se pose également : comment diffèrent vraiment les différents modèles ? Le commentaire posté par Tom Pagano dans ce post de blog résume bien la réflexion : la question est de savoir s'il y a plus de valeur dans un prévisionniste « connaissant » très bien un modèle, ou dans un prévisionniste ayant une diversité de modèles autour de lui.³⁹

□ *Qualité des prévisions saisonnières d'ensemble : précipitations et débits*

L'objectif de la **thèse de Louise Crochemore (2016)**, financée dans le cadre du projet Interreg DROP et que j'ai dirigée entre 2013 et 2016, a été de faire progresser les connaissances sur la prévision saisonnière pour la gestion de réservoirs. Il s'agissait, pour mes recherches, des premières expérimentations avec la prévision à longue échéance.

Dans un premier volet des travaux, la qualité des prévisions saisonnières de pluies et de débits a été évaluée dans seize bassins versants français, à régime pluvial prépondérant. La prévision saisonnière de précipitation du CEPMMT (System 4) a été utilisée en entrée au modèle hydrologique GR6J. Les performances des systèmes de prévision évalués ont été comparées à celles de quelques systèmes de référence choisis. Un système de référence souvent utilisé pour les prévisions de précipitation est la climatologie. Dans le cas des débits, la référence la plus couramment utilisée est celle de l'approche ESP (*Ensemble Streamflow Prediction*, Day, 1985). Il s'agit d'utiliser la climatologie des pluies (donc l'historique des séries temporelles observées) en entrée au modèle hydrologique, avec celui-ci initialisé aux conditions hydrologiques observées au moment de la prévision. Dans le cas où un modèle hydrologique n'est pas disponible, l'historique des débits observés peut aussi être utilisé comme référence.

Les résultats ont montré une bonne performance des prévisions brutes de précipitation (prévision sans post-traitement) à l'échelle du bassin versant, et par rapport à la climatologie (ensemble de précipitations observées aux mêmes jours par le passé), jusqu'à 2-3 semaines

³⁸ <https://hepex.irstea.fr/multi-model-approaches-for-river-flow-forecasting-blessing-or-burden/> (les commentaires postés sur cet article de blog méritent aussi d'être lus).

³⁹ « *The question is if there is more value in a forecaster "knowing" one model very well, or having a diversity of models running around* », T. Pagano (2013), commentaire dans : <https://hepex.irstea.fr/multi-model-approaches-for-river-flow-forecasting-blessing-or-burden/>

dans les bassins versants étudiés (Crochemore *et al.*, 2016b). Les prévisions hydrologiques obtenues à partir de la prévision saisonnière de précipitation du CEPMMT (System 4) ont montré une tendance à être moins dispersées, mais aussi moins fiables que celles obtenues avec la méthode ESP (**Figure 13**).

De manière générale, les prévisions saisonnières de précipitation du CEPMMT (System 4) ont présenté des biais de sous ou surestimation. Ces biais varient selon le bassin versant et la saison, avec parfois des biais mensuels importants qui ont tendance à se compenser si on ne les évalue que sur l'année (**Figure 14**). Il a été mis en évidence qu'une correction de biais des prévisions saisonnières est nécessaire avant son usage dans le modèle hydrologique. Son impact sur la qualité des prévisions hydrologiques est présenté dans la prochaine section de ce chapitre (**section 3.3**).

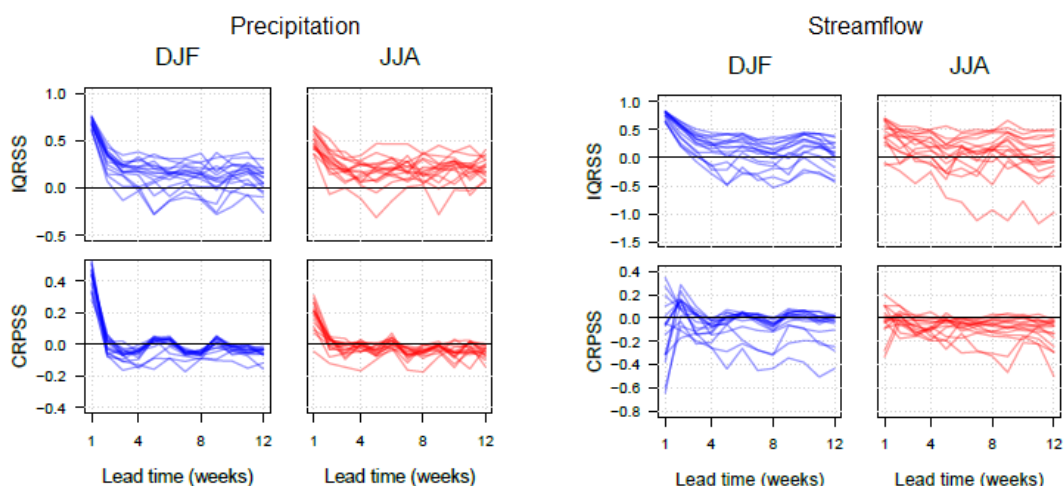


Figure 13 : Performance des prévisions saisonnières brutes (sans post-traitement) des précipitations de bassin (Sys4 CEPMMT) et de débits (Sys4 et modèle hydrologique GR6J), selon les critères IQRSS (Interquantile range, qui mesure la dispersion des membres, et CRPSS, qui mesure la précision et la fiabilité, et pour les saisons de décembre à février (DJF ; saison humide) et juin à août (JJA ; saison sèche). Chaque ligne représente un bassin versant. Les « skill scores » ont pour référence l'historique des précipitations observés (pour l'évaluation des précipitations ; à gauche) et la méthode ESP (pour l'évaluation des débits ; à droite). Une prévision est meilleure que sa référence si le « skill score » est supérieure à zéro (dans : Crochemore *et al.*, 2016b)

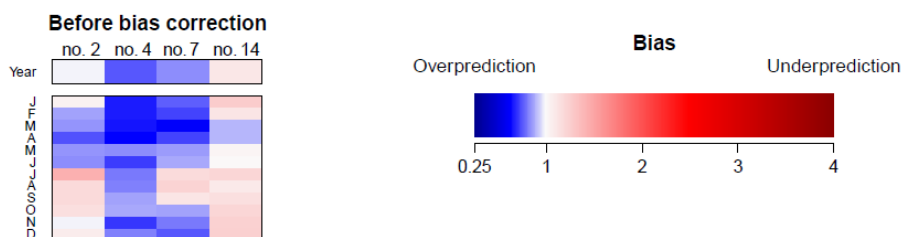


Figure 14 : Biais des prévisions saisonnières d'ensemble du CEPMMT (System 4) pour quatre bassins versants (#2, #4, #7 et #14), évalués sur la période 1981-2010 et pour les échéances entre 31 et 60 jours. Le biais est indiqué pour toute l'année (première ligne) et pour chaque mois civil (lignes suivantes) (dans : Crochemore *et al.*, 2016b)

3.3. Quantifier les incertitudes prédictives : propagation de plusieurs sources d'erreurs ou post-traitement statistique ?

Les prévisions hydrologiques d'ensemble sont caractérisées par des incertitudes provenant de différentes sources (données observées, prévisions météorologiques, paramétrage et structure du modèle, conditions initiales, etc.) et peuvent être biaisées dans leur moyenne (sous ou sur estimation), leur dispersion (sous ou sur dispersion), ou les deux. Les techniques de post-traitement statistique et de correction de biais cherchent à corriger ou réduire certains de ces biais. De manière générale, elles comportent des procédures qui caractérisent la distribution de probabilité conjointe des prévisions brutes et des observations, ou les statistiques des erreurs résiduelles de modélisation. Le but est de quantifier l'incertitude prédictive totale et, dans le cadre de la prévision d'ensemble, rendre les prévisions justes et fiables, c'est-à-dire, calibrées statistiquement, pour ainsi pouvoir convertir les scénarios d'une prévision d'ensemble en probabilités de prévision.

En 2013, j'ai contribué à trois articles de blog pour le portail HEPEx sur le post-traitement des prévisions, où nous avons proposé une brève revue de la littérature sur le sujet⁴⁰. Nous avons considéré les premiers articles proposant des approches basées sur l'inférence bayésienne pour quantifier l'incertitude hydrologique, combinée ou non à une approche de prévision météorologique d'ensemble pour quantifier l'incertitude des forçages (Krzysztofowicz, 1999, 2001 ; Reggiani et Weerts, 2008), ainsi que les développements et applications d'approches non paramétriques, à noyaux, basées sur des copules ou la transformation en ondelettes (par ex., Brown et Seo, 2010 ; Bogner et Pappenberger, 2011 ; Boucher *et al.*, 2012 ; Madadgar *et al.*, 2012). Une revue de la littérature très extensive et détaillée a été proposée récemment par Li *et al.* (2017).

De manière générale, les approches proposées dans la littérature diffèrent principalement par le contexte d'application (simulation ou prévision) et le formalisme mathématique utilisé pour, soit quantifier les différentes sources d'erreurs séparément et les propager jusqu'aux débits, soit quantifier l'incertitude prédictive des débits de manière globale, sans séparer les différentes sources d'erreurs. De même que pour les critères d'évaluation de la qualité des prévisions, nous constatons qu'il existe une grande variété de méthodes et plusieurs regards sont possibles. Ci-après, quelques actions de recherche auxquelles j'ai contribué sur ce sujet sont présentées.

□ *Post-traitement pour quantifier les incertitudes des simulations hydrologiques*

En 2012, une **expérimentation HEPEx** a été lancée dans le but de mettre en place une étude comparative de différentes méthodes de post-traitement (van Andel *et al.*, 2012). J'ai contribué à la mise en place du protocole expérimental et aux comparaisons faites sur des simulations hydrologiques basées sur des forçages météorologiques observés (c'est-à-dire que les incertitudes des prévisions météorologiques n'étaient pas prises en compte). Nous avons mis à disposition de la communauté HEPEx des données de 12 bassins versants et des simulations venant de 7 modèles hydrologiques utilisés dans l'expérimentation MOPEX, sur une période de 36 ans (1962 à 1997).

⁴⁰ Les trois sont accessibles à partir du premier article : <https://hepex.irstea.fr/hepex-sip-topic-post-processing-13/>

Les résultats obtenus à partir de cinq techniques différentes ont été publiés dans un article de blog et un poster⁴¹. Ils ont suggéré que le choix de la technique de post-traitement peut avoir peu d'importance quand on regarde la performance moyenne sur un grand nombre de cas. Cependant, des différences peuvent apparaître quand on s'intéresse à des sous-échantillons (par ex., en évaluant les débits par saison ou les hauts et bas débits séparément). L'étude n'a pas fait apparaître une méthode qui était la meilleure partout et dans toutes les conditions d'évaluation.

A Irstea, cette étude et les nombreux échanges qu'elle a suscités ont été mis à profit pour les travaux de **thèse de François Bourgin (2014)**, dirigés par mes collègues d'équipe V. Andréassian et C. Perrin. J'ai contribué à la thèse de F. Bourgin à travers deux chapitres (chapitres 4 et 9), dont l'un a été publié au *Journal of Hydrology* (Bourgin *et al.*, 2014 ; voir **section 3.2**). Enfin, on notera également que les travaux de thèse de F. Bourgin sont à l'origine de la méthode qui est à la base de l'outil OTAMIN, conçu dans l'objectif d'associer de manière automatique et systématique des intervalles prédictifs aux prévisions hydrologiques et hydrauliques obtenues avec différents modèles déterministes. L'outil est aujourd'hui utilisé par les services opérationnels du réseau SCHAPI/SPC⁴².

❑ *Prétraitement des prévisions de précipitation, post-traitement des débits prévus, ou les deux?*

Si l'on se met dans la perspective du modélisateur hydrologue, on peut se poser des questions sur la nécessité de faire un prétraitement des entrées au modèle hydrologique (ce qui serait un post-traitement pour la météorologie, mais ciblé vers les variables qui seront effectivement utilisées en entrée au modèle hydrologique), avant de faire un post-traitement des sorties du modèle hydrologique : quelle influence le prétraitement peut-il avoir sur la qualité des sorties du modèle hydrologique ? Est-ce que le prétraitement est suffisant ou faut-il ajouter ensuite un post-traitement hydrologique ? Dans quelle situation les efforts à faire pour le prétraitement des entrées météorologiques sont-ils vraiment nécessaires (notamment si un post-traitement des débits sera indispensable par la suite) ?

Dans le cadre de la **thèse d'I. Zalachori (2013)**, l'intérêt d'un post-traitement statistique à la sortie du modèle hydrologique par rapport à un prétraitement des forçages en amont a été clairement démontré. L'étude publiée dans Zalachori *et al.* (2012) a porté sur la performance d'un système de prévision composé des prévisions météorologiques d'ensemble du CEPMMT et du modèle hydrologique MORDOR (Garçon, 1999). Ce système a été évalué lorsque le traitement statistique est appliqué : (1) aux prévisions de précipitation et de température seulement (prétraitement), (2) aux prévisions de débits (post-traitement) et (3) aux deux. Le prétraitement combinait les prévisions d'ensemble et une approche de prévision par analogues, tandis que le post-traitement était basé sur les erreurs du modèle hydrologique observées lors de son calage. L'étude a été faite sur 11 bassins versants en France et une période d'évaluation allant de 2005 à 2008.

⁴¹ <https://hepex.irstea.fr/intercomparison-of-streamflow-post-processors-post-processing-hydrologic-model-simulations-phase-1/>

⁴² OuTil Automatique d'estiMation de l'INcertitude prédictive OTAMIN : <https://webgr.irstea.fr/modeles/otamin/>

La **Figure 15** illustre le fait que les critères de qualité s'améliorent avec les traitements statistiques, et proportionnellement de manière plus significative lorsque l'on ajoute le post-traitement des débits. La qualité des prévisions des hauts débits (supérieures au quantile 80% des débits observés ; critère BSSQ80) bénéficie plus du prétraitement des entrées, tandis que le critère global de qualité (critère RPSS), appliqué sur l'ensemble des débits, est moins impacté par le prétraitement des forçages météorologiques (et d'autant moins avec l'augmentation de l'échéance).

Ces résultats ont montré que les corrections statistiques apportées aux prévisions de précipitations peuvent perdre leur effet lorsqu'elles sont propagées à travers le modèle hydrologique. Cela implique que, même après avoir fait un effort de prétraitement des prévisions météorologiques pour les rendre précises et bien fiables, les prévisions hydrologiques risquent de manquer de qualité, et notamment de fiabilité, ce qui demanderait, de toute façon, un effort supplémentaire de post-traitement des débits aux prévisionnistes.

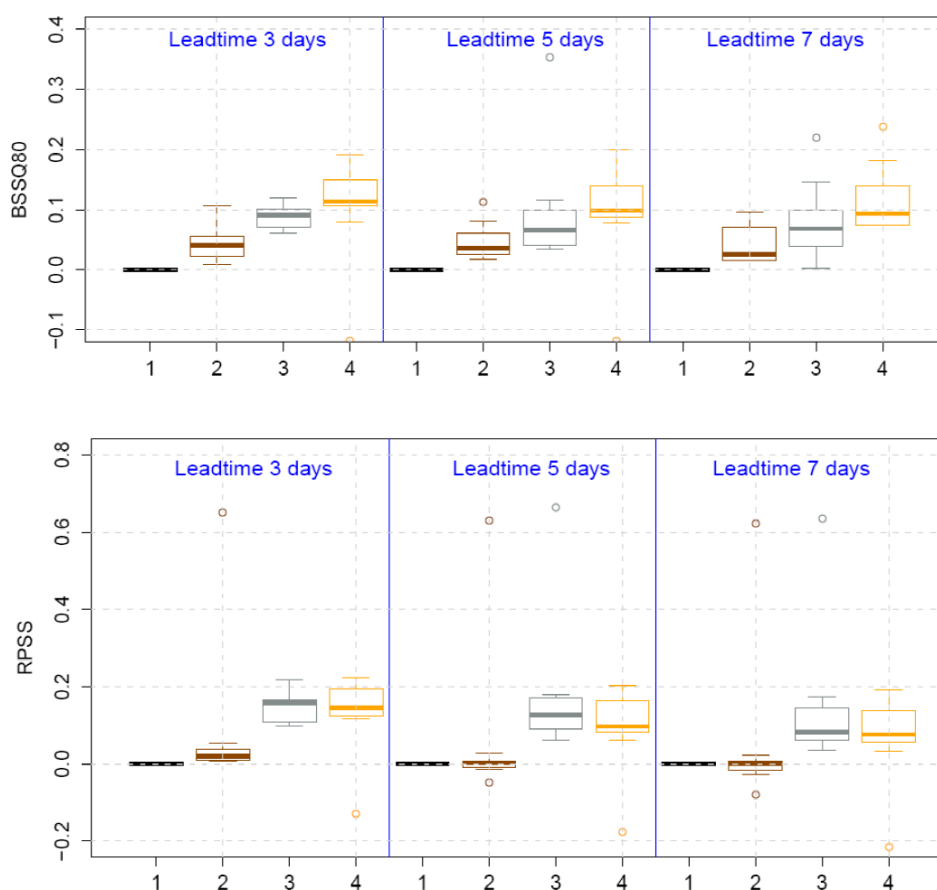


Figure 15 : Qualité des prévisions hydrologiques d'ensemble pour les cas où (2) les prévisions météorologiques sont prétraitées, (3) les prévisions hydrologiques sont post-traitées, et (4) les prévisions météorologiques sont prétraitées et les prévisions hydrologiques sont post-traitées. Les critères BS et RPS sont calculés par rapport aux prévisions brutes (1). Évaluation sur 48 mois (2005-2008), 11 bassins versants et pour trois échéances : 3, 5 et 7 jours (dans : Zalachori et al., 2012)

L'enjeu ici est lié à la question : est-ce qu'en hydrologie on peut accepter des prévisions météorologiques bonnes, sans des biais notables, mais sans être forcément parfaites en fiabilité (étant donné que, de toute façon, il faudra appliquer un post-traitement aux débits prévus en sortie de la chaîne) ? De manière générale, d'autres auteurs ont confirmé nos résultats, bien que sur des configurations différentes (Kang *et al.*, 2010 ; Verkade *et al.*, 2013 ; Roulin and Vannitsem, 2015).

Pour aller plus loin et répondre à la question « prétraitement, post-traitement ou les deux ? », je laisserai le lecteur se reporter à l'article de blog de Marie-Amélie Boucher, publié dans le portail HEPEX le 15 décembre 2015, ainsi qu'aux discussions qui ont suivi son texte⁴³.

□ *Quels impacts de la correction des biais des prévisions saisonnières de précipitation sur les prévisions de débits ?*

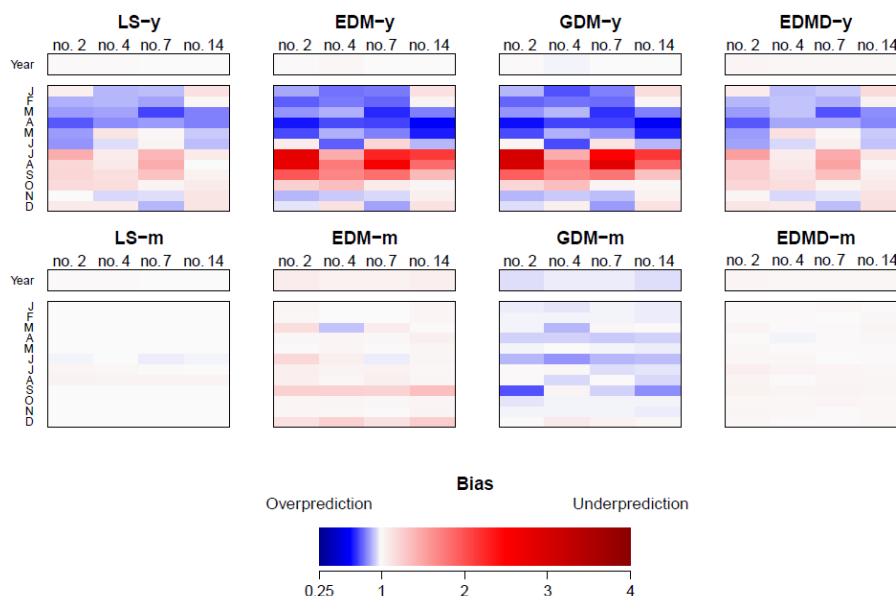
Dans le but de mieux comprendre le fait que certaines améliorations apportées aux précipitations par l'application de méthodes de correction de biais ne se traduisent pas forcément, et proportionnellement, en amélioration de la qualité des prévisions de débit, j'ai aussi examiné la question sous l'angle des prévisions saisonnières, pour lesquelles le prétraitement des entrées (prévisions de précipitations) en hydrologie est pratiquement inévitable. Ceci a été fait dans le cadre des travaux de **thèse de Louise Crochemore (2016)**.

L'évaluation de la qualité des prévisions saisonnières a été accompagnée de la mise en place de huit méthodes de correction de biais des prévisions saisonnières de précipitations, suivi de l'étude de l'impact des corrections faites sur la qualité des prévisions saisonnières de débits. L'étude a été publiée par Crochemore *et al.* (2016b). Les huit méthodes avaient des degrés croissants de complexité, allant d'une correction linéaire des biais calée sur l'ensemble de l'année, jusqu'à la correction des quantiles de la distribution journalière des valeurs prévues de précipitation, après ajustement d'une loi gamma et calage mensuel des paramètres de la correction.

La correction de biais la plus efficace, sur le jeu de données étudié, a été celle fournie par la méthode de correction quantile-quantile (avec distribution empirique) des valeurs journalières, calée pour chaque mois de l'année (EDMD-m dans la **Figure 16**). La correction des biais mensuels des prévisions saisonnières du CEPMMT (System 4) a permis, en effet, d'améliorer la fiabilité des prévisions, et d'harmoniser les performances obtenues dans les seize bassins étudiés.

En prenant en compte la méthode EDMD-m, nous avons évalué comment l'amélioration de la qualité des prévisions de précipitations se propage dans la qualité des prévisions de débits. Pour comparer les deux, nous avons pris en compte le gain en performance (scores relatifs) par rapport aux prévisions brutes non-corrigées (de précipitations ou de débits). La **Figure 17** illustre les résultats obtenus.

⁴³ <https://hepex.irstea.fr/pre-post-processing-or-both/>



Abbreviation	Calibration based on	Description
LS-y LS-m	the whole year calendar months	Linear scaling of monthly values
EDM-y EDM-m	the whole year calendar months	Empirical distribution mapping of monthly values
GDM-y GDM-m	the whole year calendar months	Gamma distribution mapping of monthly values
EDMD-y EDMD-m	the whole year calendar months	Empirical distribution mapping of daily values

Figure 16 : Biais des prévisions saisonnières de précipitations du CEPMMT (System 4) pour quatre bassins versants (#2, #4, #7 et #14), évalués sur la période 1981-2010 et pour les échéances entre 31 et 60 jours, pour huit méthodes de correction de biais appliquées aux prévisions brutes (voir tableau en dessous). Le biais est indiqué pour toute l'année (première ligne) et pour chaque mois civil (lignes suivantes). Le biais original des prévisions brutes est montré dans la Figure 14 (dans : Crochemore et al., 2016b)

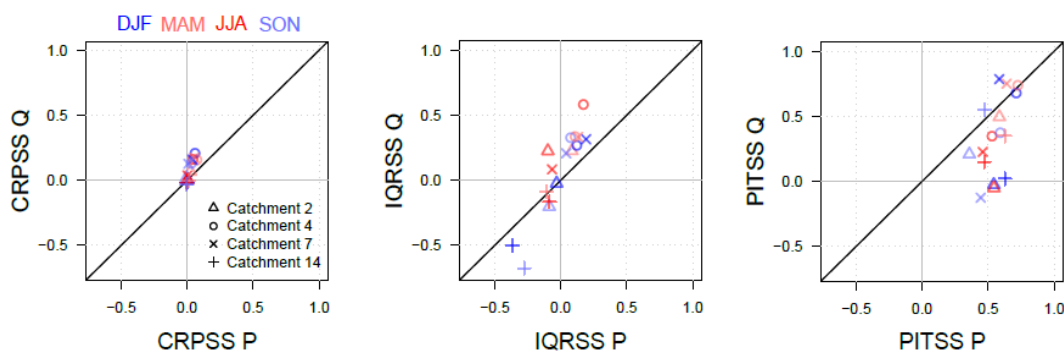


Figure 17 : Critères de qualité des prévisions de précipitation (P) et de débits (Q) après application d'une correction de biais aux précipitations (EDMD-m, Fig. 16). Les prévisions brutes de précipitations et de débits sont utilisées comme référence. Les valeurs des critères sont moyennées sur les échéances de 10 à 90 jours. Les résultats sont présentés pour les quatre saisons de l'année et quatre bassins versants. Les critères mesurent la performance globale (CRPSS), la finesse (IQRSS), et la fiabilité (PITSS) des prévisions (dans : Crochemore et al., 2016b)

Nous observons que :

- en termes de performance globale (critère CRPSS), la correction de biais des prévisions de précipitation conduit à un gain de performance (par rapport aux prévisions brutes) équivalent dans les prévisions de précipitation et de débit, ou supérieur dans les prévisions de débit (par rapport au gain des prévisions de précipitation) ;
- en termes de finesse (critère IQRSS), on observe plus généralement que, quand les précipitations gagnent en performance, les débits gagnent également et plus fortement ; quand les précipitations perdent en performance, les débits perdent également et aussi plus fortement (même si quelques cas de gain en débit ont été observés) ;
- en termes de fiabilité (critère PITSS), les gains en performance sont toujours positifs pour les précipitations. Les gains en débits sont majoritairement positifs, mais pas toujours ; certains gains en fiabilité des précipitations ne se transposent donc pas en gain en fiabilité des débits (encore une fois le besoin d'un post-traitement est mis en évidence pour rendre les prévisions hydrologiques fiables pour un utilisateur dans le secteur de l'eau).

Bien que la correction du type « quantile mapping » (QM) ait fourni des bons résultats sur les bassins étudiés en France, elle soulève des questions quant à sa capacité à aller au-delà d'une correction des biais et à effectuer un véritable post-traitement des prévisions de précipitations, c'est-à-dire les rendre fiables et statistiquement cohérentes vis-à-vis les fréquences d'observation des événements.

Dans une collaboration avec des collègues australiens (Zhao *et al.*, 2017), suite à une discussion que nous avons tenue lors du workshop HEPEX « Seasonal Hydrological Forecasting » à Norrköping, Suède, en 2015, nous avons illustré cet aspect en comparant la performance de la méthode QM à celle d'une méthode du type bayésienne (BJP). L'application a été faite sur des prévisions saisonnières de précipitation sur l'Australie. Les résultats ont montré les défauts de la méthode QM vis-à-vis de la correction de la dispersion (fiabilité) des prévisions, malgré sa capacité à corriger les biais. Nous avons ainsi suggéré l'utilisation de méthodes de post-traitement qui prennent explicitement en compte la relation entre prévisions et observations à travers leur distribution conjointe. Ces méthodes sont néanmoins plus coûteuses en termes de calcul et, dans certains cas, peuvent converger très rapidement à la climatologie, ce qui peut détériorer la finesse des prévisions aux échéances plus courtes.

Dans Crochemore *et al.* (2017), nous avons voulu examiner si les prévisions saisonnières de précipitation, corrigées de leurs biais, pouvaient apporter une information à la climatologie des précipitations passées, qui conditionnerait cette climatologie et apporterait une meilleure qualité aux prévisions de débit basées sur cette « climatologie conditionnée » des précipitations passées. Les résultats de l'étude ont montré que les sélections basées sur le SPI3 (*Standardized Precipitation Index*, avec cumul de précipitations sur 3 mois) conduisent à des meilleures performances. Une comparaison a été faite avec les prévisions de débit basées sur l'approche ESP traditionnelle (sans conditionnement) et les prévisions de débits basées sur les prévisions de précipitation du CEPMMT (System 4). La conclusion principale a été que les ensembles conditionnés ont généralement plus de fiabilité, mais moins de finesse que les ensembles basés directement sur les prévisions saisonnières du System 4. Ils ont aussi plus de qualité en finesse, mais moins en fiabilité par rapport aux prévisions basées sur l'approche traditionnelle de l'ESP (**Figure 18**).

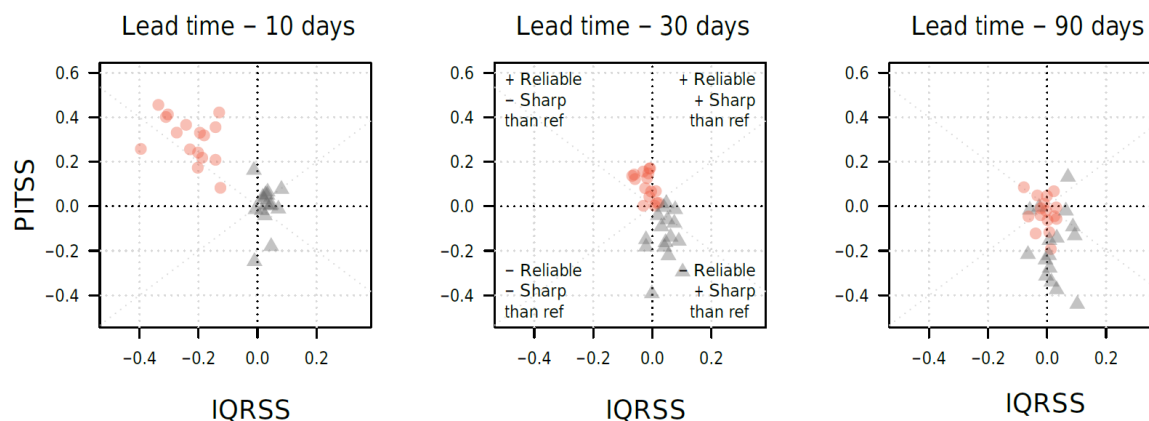


Figure 18 : Performance des prévisions de débit (approche ESP conditionnée au SPI3) par rapport aux performances des prévisions de débit basées sur l'approche ESP traditionnelle (triangles en gris) et les prévisions saisonnières de précipitation CEPMMT System 4 (ronds en rouge), selon l'échéance de 10, 30 et 90 jours. Les critères mesurent la performance en finesse (IQRSS) et fiabilité (PITSS) (dans : Crochemore et al., 2017)

Nos résultats ont montré que, dans certains cas, le conditionnement proposé pourrait améliorer la fiabilité et la finesse simultanément, en particulier pour les horizons de prévision inférieurs à un mois. Cependant, le plus souvent, lorsque l'on améliore un attribut, on finit par dégrader l'autre, ce qui nous laisse penser qu'un compromis entre fiabilité et finesse pourrait exister, ce qui a aussi été suggéré par d'autres auteurs (Hamlet et Lettenmaier, 1999 ; Yao et Georgakakos, 2001).

La question que l'on se pose est de savoir si, dans le cas où on n'arrive pas à avoir un système amélioré en fiabilité et finesse simultanément, un utilisateur serait plus intéressé à avoir des prévisions plus fiables ou, finalement, ce serait la finesse qui compte le plus dans son processus de décision ?

Nous avons cherché à voir comment le compromis « fiabilité / finesse » pourrait impacter l'évaluation d'un utilisateur qui s'intéresse à des critères plus hydrologiques dans la gestion d'un événement d'étiages. La **Figure 19** montre un outil de visualisation orienté vers des variables d'intérêt typiques en période d'étiages : le nombre de jours et le volume déficitaire sous un seuil de basses eaux. Le cas illustré est celui de la sécheresse de 2003 dans le bassin de l'Azergues à Lozanne (798 km²) et d'une prévision émise le 1^{er} juillet pour les 90 jours suivants.

Bien que ce ne soit pas possible d'évaluer les propriétés statistiques de la qualité des prévisions avec l'analyse d'un seul événement, nous pouvons observer quelques aspects intéressants. Par exemple, la prévision conditionnée (ESP_SPI3) donne plus de probabilités pour les déficits moins sévères que la prévision basée sur System 4 (Sys4), ce qui pourrait être une indication de sa moins bonne performance en termes de finesse. En revanche, quand on la compare à la prévision non conditionnée (ESP), on observe que le conditionnement a permis de détecter une probabilité à l'événement observé (rond noir en haut à droite de la figure), même si cette probabilité était faible. Sur ce bassin, l'événement de sécheresse de 2003 a été plus sévère (au regard de ces deux indices de déficit hydrologique) que l'événement historique de 1976. L'approche ESP, qui n'utilise pas une information météorologique actualisée aux conditions présentes (c'est-à-dire qui n'utilise pas des prévisions météorologiques à l'instant de la prévision), n'a pas été capable de détecter cet événement extrême de 2003 dans son faisceau d'ensemble.

Enfin, on notera que Gneiting *et al.* (2007) proposent de suivre une approche d'évaluation de la performance des prévisions selon laquelle nous devons chercher à avoir des prévisions qui sont d'abord fiables et ensuite à maximiser leur finesse, tout en garantissant qu'elles restent fiables. Les auteurs font l'hypothèse que les deux objectifs (prévision parfaite/idéale et maximisation de la finesse étant donnée la fiabilité) sont équivalents.⁴⁴

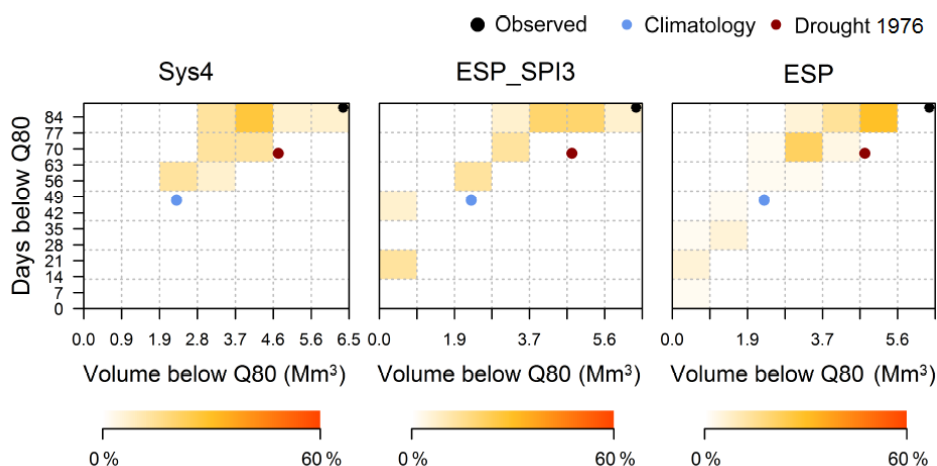


Figure 19 : Graphiques du risque d'avoir le nombre de jours (Y-axis) et le volume (X-axis) sous le seuil de basses eaux Q80 (20% des valeurs historiques sont au-dessous de ce seuil) pour les prévisions saisonnières de débit basées sur les prévisions saisonnières de précipitation du CEPMMT (Syst4), l'approche ESP conditionnée (ESP_SPI3) et l'approche ESP traditionnelle (ESP). Le cas illustré est celui de la sécheresse de 2003 dans le bassin de l'Azergues à Lozanne et d'une prévision émise le 1er juillet pour les 90 jours suivants (dans : Crochemore *et al.*, 2017)

❑ Est-ce qu'il ne suffit pas de quantifier toutes les sources d'incertitudes ? Avons-nous vraiment besoin de faire du post-traitement statistique des prévisions ?

Certaines études (Thibault, 2015 ; Thibault *et al.*, 2017) questionnent le besoin de faire du post-traitement statistique en hydrologie si on réussit à quantifier correctement (et de manière efficace) les principales sources d'incertitude qui jouent un rôle dans la qualité de la prévision hydrologique. Dans mes recherches, je n'ai pas adopté une telle approche, car je trouve cette vision très optimiste quant à notre capacité à élucider les différentes sources d'incertitude de la prévision hydrométéorologique et à les prendre en compte de façon complémentaire afin d'avoir des prévisions précises et fiables tout au long des horizons de prévision qui intéressent l'hydrologie. A mon avis, une telle approche peut aussi être peu pratique si on considère les contraintes de gestion de données et modélisation d'un service opérationnel (dans Thibault *et al.*, 2017, l'ensemble qui contient les trois sources d'incertitude étudiées – prévisions météorologiques, modèles hydrologiques et conditions initiales – contient 50 000 membres !).

⁴⁴ On parle du « (...) paradigm of maximizing the sharpness of the predictive distributions subject to calibration. » et les auteurs indiquent que leur « (...) conjectured sharpness principle contends that the two goals (ideal forecasts and the maximization of sharpness subject to calibration) are equivalent. » (Gneiting *et al.*, 2007)

Néanmoins, j'avais aussi fait quelques expérimentations de recherche sur la propagation des incertitudes par une approche explicite, basée sur l'identification et la quantification de différentes sources d'incertitude. J'ai ainsi encadré les travaux de **Master de Trang van Pham en 2011** (en collaboration avec Martijn Booij de l'Université de Twente au Pays-Bas et Benjamin Renard et Etienne Leblois d'Irstea Centre de Lyon). Cette étude a été d'un cran plus ambitieuse que celle menée avec J.-A. Velázquez (**section 3.2**). Nous avons considéré les sources d'incertitude suivantes :

- incertitudes des conditions initiales données par les observations de précipitation : pour cela, nous avons considéré un ensemble de pluies de bassin générées à l'aide de la méthode géostatistique de bandes tournantes, conditionnée aux pluviomètres au sol (simulations faites à l'aide du logiciel SAMPO TBM d'Irstea Lyon ; version mise en place lors de la thèse de J.M. Lepioufle, 2009, encadrée par E. Leblois) ;
- incertitudes des conditions initiales données par les observations des débits qui sont assimilées à l'instant de la prévision : pour cela nous avons utilisé la méthode de quantification des incertitudes des courbes de tarage proposée par Renard *et al.* (2010) ;
- incertitudes des prévisions de précipitation : nous avons utilisé les prévisions météorologiques d'ensemble du CEPMMT et la PEARP de Météo-France ;
- incertitudes de la modélisation hydrologique : nous avons considéré les incertitudes des paramètres du modèle hydrologique GR3P à l'aide de deux méthodes, à savoir, (1) des calages faites sur 10 périodes différentes, de 5 ans chacune, prise à l'intérieur de la période historique de données, et (2) la méthode GLUE (*Generalized Likelihood Uncertainty Estimator*), pour générer 125 000 jeux de paramètres dans les limites données par les paramètres du modèle GR3P calés sur 3 070 bassins versants en France.

L'étude a été faite sur un petit échantillon de trois bassins versants, pour lesquels nous avons les simulations géostatistiques (disponibles sur un seul bassin) et les courbes de tarage avec incertitudes (disponibles sur trois bassins). La période d'évaluation des prévisions allait de mars 2005 à juillet 2009. Au total, 11 tests ont été faits. La **Figure 20** illustre le schéma de trois d'entre eux et liste l'ensemble de tests effectués.

Les résultats ont montré que l'incertitude des observations de précipitation, telle qu'elle avait été quantifiée dans notre étude, n'avait pas un impact important sur la prévision de débit. Pour les autres sources d'incertitude, l'incertitude des conditions initiales hydrologiques a montré un impact plus important pour les échéances plus courtes (jusqu'à 2 jours). L'incertitude des précipitations prévues a également eu un impact important sur la qualité des prévisions de débit. Cet impact a été plus marqué pour les échéances plus longues. Selon le bassin versant, l'incertitude des paramètres peut avoir un impact non négligeable, selon si elle est évaluée à partir de différentes périodes de calage ou de la méthode GLUE. En résumé, les trois sources d'incertitude qui nous ont paru plus importantes pour avoir des prévisions de débit fiables ont été celles des précipitations prévues, des conditions initiales du modèle hydrologique et des jeux de paramètres du modèle hydrologique.

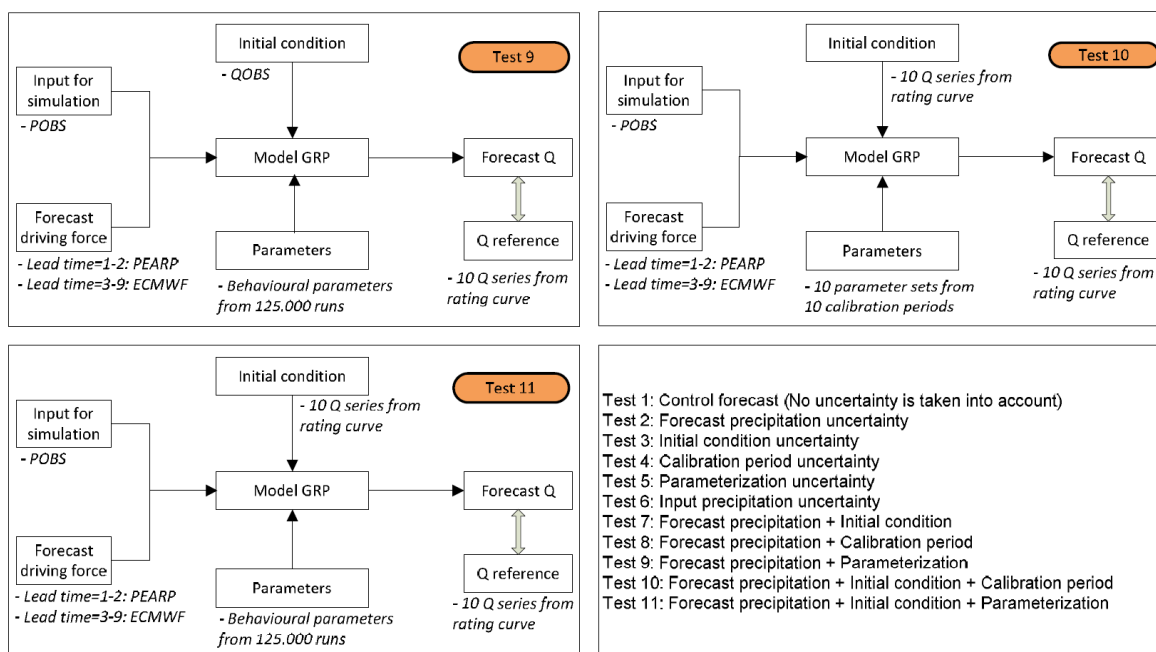


Figure 20 : Tests effectués pour la prise en compte explicite des différentes sources d'incertitude et leur propagation dans la chaîne de prévision de débit (dans : van Pham, 2011)

3.4. Le meilleur est peut-être devant nous

J'aimerais souligner deux aspects qui sont, à mon avis, fondamentaux pour pouvoir progresser dans le domaine de l'évaluation de la qualité des prévisions hydrologiques d'ensemble :

- l'importance de disposer de prévisions météorologiques d'ensemble pour pouvoir progresser dans la compréhension et l'utilisation opérationnelle des prévisions hydrologiques d'ensemble. *A priori*, le traitement de données (accès et quantité) peut sembler être un obstacle en hydrologie, mais, aujourd'hui, l'usage de codes et logiciels de lecture de données météorologiques est bien plus répandu et les moyens de stockage de données, plus puissants. Il est crucial, à mon avis, d'expérimenter largement avec ce type de prévision pour pouvoir se familiariser avec la façon dont l'ensemble capture les incertitudes de prévision et pour mieux comprendre les avantages de la prévision d'ensemble pour le secteur de l'eau ;
- une utilisation plus répandue des prévisions d'ensemble permettrait de disposer de résultats d'évaluation des prévisions de débit sur un plus large éventail de bassins versants aux caractéristiques variées et soumis à des conditions climatiques diverses, ainsi que sur une plus grande quantité d'événements marquants en hydrologie, à intérêt particulier dans la gestion des risques.

L'ensemble d'études présenté dans ce chapitre a montré la part d'importance des incertitudes des précipitations prévues, des modèles hydrologiques et des observations de débits sur la qualité des prévisions. Ceci est rassurant et enrichissant, mais soulève la question : « plus on en met, mieux c'est ? ». Cette question peut nous laisser un peu déconcertés, car la multiplication des scénarios possibles convient peu à un contexte opérationnel.

Si d'une part le travail de propagation des incertitudes par des approches ensemblistes est tout à fait intéressant pour mieux évaluer la part contributive des différentes sources d'incertitude, il devient vite très contraignant quand il y a plusieurs sources à prendre en compte, avec plusieurs scénarios à la fois. Cet aspect est d'autant plus limitant si les sorties des modèles hydrologiques sont utilisées en entrée d'autres modélisations, telles que les modélisations hydrauliques pour les inondations ou les modèles d'optimisation d'apports pour la gestion des réservoirs.

Le rôle d'un post-traitement devient primordial pour pouvoir fiabiliser les prévisions d'ensemble. Par contre, son utilisation soulève d'autres questions comme, par exemple, la disponibilité de prévisions archivées (émises par le même système de prévision opérationnel) pour le calage des paramètres, la vérification des hypothèses gaussiennes de certaines techniques, et le maintien des corrélations spatiales et temporelles (au long des échéances), et entre variables météorologiques et hydrologiques après l'application d'un post-traitement statistique. Ce ne sont pas des défis insurmontables, mais il ne faut pas les sous-estimer.

3.5. Que reste-t-il à faire une fois que les prévisions sont de bonne qualité ?

La qualité des prévisions hydrologiques d'ensemble a été un sujet qui a beaucoup capté l'attention de mes travaux de recherche. Il s'agissait de démontrer quantitativement et avec des critères compréhensibles, le potentiel des approches d'ensemble pour anticiper les risques d'événements sévères, favoriser l'action pour la protection des personnes et des biens, et rendre l'utilisation de la ressource en eau plus optimale, ou, au moins, la décision la concernant plus informée. Cette connaissance permet, à mon avis, de mieux cibler les efforts de développement dans un système de prévision et peut appuyer l'automatisation de certaines procédures, permettant au prévisionniste de mieux utiliser ses compétences dans l'interprétation de la prévision vis-à-vis de son objectif opérationnel ou de l'utilisation qui en sera faite (Pagano *et al.*, 2016).

Comme l'avait déjà bien indiqué Murphy en 1993 : « ... *forecasts possess no intrinsic value. They acquire value through their ability to influence the decisions made by users of the forecasts* » (Murphy, 1993). La question est de savoir comment mesurer cette « *ability* », ou la capacité des prévisions à jouer un rôle dans la décision. Dans mes recherches, je me suis aussi intéressée à ce lien entre qualité et valeur d'une prévision hydrologique. Hélas pour moi, qui pensait avoir juste affaire à un modèle de plus dans la chaîne de prévision, cela signifiait essentiellement qu'il fallait d'abord être en mesure de comprendre, modéliser et évaluer les décisions.

Une fois, j'ai demandé à un décideur (responsable des manœuvres d'une vanne de contrôle de crues d'un barrage-réservoir) comment il savait, à la suite d'un événement, si la décision qu'il avait prise était la « meilleure décision ». La réponse qu'il m'a donnée était (de mémoire, et non pas *ipsis verbis*) : « *C'est la meilleure décision : nous prenons la meilleure décision compte tenu des prévisions que nous recevons et d'autres informations complémentaires que nous avons sur la situation. Si le résultat n'est pas bon, le problème n'est pas dans la décision, mais dans les prévisions, qui n'étaient pas bonnes* »⁴⁵.

⁴⁵ Ce dialogue a eu lieu en 2013 pendant mon séjour de six mois à Québec. Il a été repris dans l'un de mes articles de blog HEPEX, publié en 2014 : <https://hepex.irstea.fr/economic-value-of-hydrological-ensemble-forecasts/>.

J'avais encore plus de questions après cet échange : est-ce que la mise à disposition des utilisateurs d'un système de prévision hydrologique efficace et de bonne qualité, qu'il soit plus ou moins complexe selon les objectifs auquel il doit répondre, est suffisante pour garantir un impact final positif d'un tel système ? Est-ce que les améliorations de qualité se traduisent par des améliorations de valeur ou, plus généralement, d'utilité des prévisions ? Comment mesurer la valeur d'une prévision en hydrologie ? Comment lier la qualité des prévisions à leur valeur ? A minima, cela m'intéressait pour pouvoir justifier de continuer à faire des recherches vers l'amélioration de la qualité des prévisions hydrométéorologiques.

4. Liens entre qualité et valeur des prévisions en hydrologie

4.1. Pourquoi s'intéresser à l'interaction prévisionniste-utilisateur ?

Dans mes recherches, l'intérêt porté sur l'étude des liens qualité-valeur des prévisions hydrologiques s'est traduit par l'émergence d'un nouveau regard, cette fois-ci sur l'interaction « prévisionniste (modélisateur) ⇔ utilisateur (décideur) ». Dans la majorité des services ou entreprises, la prévision et la décision constituent des actions séparées, étant donné la complexité de chacune prise isolément (même si le prévisionniste porte une part des décisions dans la chaîne complète, cf. **Figure 2**, et que, probablement, les décideurs « prévoient » eux-aussi un peu, au moins les conséquences de leurs décisions, lors de leur prise de décision).

Dans son article, Raftery (2016) suggère que le fait que les prévisionnistes et les décideurs sont des personnes différentes, comme c'est souvent le cas dans le contexte des décisions politiques, pourrait être une explication possible du fait que communiquer des prévisions incertaines (probabilistes) est un vrai défi. Selon l'auteur, cela amènerait à ce que les responsabilités soient très souvent transférées de l'un à l'autre, ce qui serait, cependant, plus compliqué à faire quand les incertitudes sont quantifiées et explicitement communiquées, par rapport aux situations où un seul chiffre est donné (cas de la prévision déterministe).

L'auteur liste cinq types d'utilisateurs de prévisions⁴⁶, allant du plus simple utilisateur, avec des enjeux faibles, aux plus complexes, qui quantifient leurs fonctions de coûts/pertes et prennent appui sur la théorie de la décision pour définir leurs règles de gestion. En faisant ressortir leurs différences en termes de besoin d'information probabiliste, Raftery (2016) conclut sur l'importance pour les prévisionnistes d'interagir avec les utilisateurs et de prendre en compte leurs objectifs⁴⁷. Il suggère qu'il s'agit ici d'un problème cognitif ainsi que d'un problème statistique, où des post-traitements statistiques pour aboutir à des prévisions fiables amèneraient les utilisateurs à avoir une confiance accrue dans les prévisions qu'ils reçoivent : « *It is important for trust in the forecast that the probabilistic statements be at least approximately calibrated (...)* ».

Raftery (2016) suggère également que l'interaction prévisionniste/utilisateur peut prendre la forme d'un contact direct entre les développeurs et les utilisateurs. Cela peut se faire dans le cadre d'un comité consultatif qui se réunit régulièrement, des réunions de groupes d'experts, des ateliers de plusieurs jours ou des sites internet interactifs, intégrant des enquêtes publiques, par exemple. Cependant, pour qu'une action (prévision) soit en harmonie avec l'autre (décision), il est essentiel d'aller au-delà de la communication en tant que « transmission d'informations justes », ce qui, finalement, n'est qu'une partie du processus de communication entre les acteurs (Ramos *et al.*, 2010a ; Morss *et al.*, 2015).

⁴⁶ A savoir, « *Low Stakes Users, General Assessors, Change Assessors, Risk Avoiders, and Decision Theorists* »

⁴⁷ A noter que l'auteur ajoute également ceci : « *While this may seem obvious, it is often not done* »

Fischhoff (1995) identifie huit phases dans le développement de la gestion (communication) des risques (**Tableau 3**). Un système de prévision qui livre des bonnes prévisions (justes et fiables) ne sera ainsi que dans la première phase de développement d'une communication efficace du risque. Communiquer ces prévisions ne l'avancera que jusqu'à la quatrième phase. Il lui restera à montrer que l'information fournie a de la valeur et transformer l'interaction envisagée en un véritable partenariat « prévisionniste-utilisateur ».

Tableau 3 : Phases dans le développement de la gestion (communication) des risques
(traduit d'après : Fischhoff, 1995)

<i>"All we have to do is get the numbers right"</i>	Tout ce que nous avons à faire est d'avoir les bons chiffres
<i>"All we have to do is tell them the numbers"</i>	Tout ce que nous avons à faire est de leur dire les chiffres
<i>"All we have to do is explain what we mean by the numbers"</i>	Tout ce que nous avons à faire est d'expliquer ce que nous voulons dire par les chiffres
<i>"All we have to do is show them that they've accepted similar risks in the past"</i>	Tout ce que nous avons à faire est de leur montrer qu'ils ont accepté des risques similaires dans le passé
<i>"All we have to do is show them that it's a good deal for them"</i>	Tout ce que nous avons à faire est de leur montrer que c'est une bonne affaire pour eux
<i>"All we have to do is treat them nice"</i>	Tout ce que nous avons à faire est de les traiter gentiment
<i>"All we have to do is make them partners"</i>	Tout ce que nous avons à faire est de faire d'eux nos partenaires
<i>"All of the above"</i>	Tout ce qui précède

Dans leur article sur l'avenir de la prévision d'ensemble, Palmer *et al.*, (2002) soulignaient déjà que « (...) si l'on développait l'usage du couplage entre les sorties de la prévision d'ensemble et l'entrée de modèles d'application, on se rendrait compte de la véritable valeur de la prévision d'ensemble en tant qu'outil de gestion du risque ». Plusieurs modèles d'application requièrent de comprendre comment les décisions sont (ou peuvent être) prises quand les décideurs disposent de prévisions.

Au long de mes recherches, je me suis intéressée à la question de la valeur d'une prévision. Par « valeur », je me réfère ici à une valeur économique, avec coûts et bénéfices explicitement quantifiés, mais également à une valeur décisionnelle, ou, en d'autres termes, l'utilité qu'une prévision peut avoir dans un contexte décisionnel où les pertes ne sont pas forcément quantifiées ou quantifiables.

Plus particulièrement, je me suis intéressée à examiner le lien entre qualité et valeur des prévisions. N'étant ni prévisionniste, ni décideur (de profession, car nous sommes tous des décideurs quand on regarde le temps qu'il fera avant de partir en weekend à la montagne), j'aborde la question de l'extérieur, mais avec un regard positionné plus près de celui d'une prévisionniste intéressée à mesurer l'utilité des informations qu'elle fournit, une fois que celles-ci sont le reflet de la connaissance qu'elle porte sur la situation à venir.

Quelques cas particuliers sont discutés ci-après :

1. l'utilisation des prévisions d'ensemble pour l'anticipation et l'alerte aux crues ;
2. le lien entre la quantification des incertitudes pour une meilleure qualité des prévisions hydrologiques et la valeur économique relative d'un système de prévision et alerte aux crues ;
3. l'apport des prévisions hydrologiques pour la gestion de réservoirs, avec deux sujets traités :
 - a. l'utilité des prévisions saisonnières (d'ensemble) pour la gestion d'un réservoir d'eau potable, avec potentiels conflits en période d'étiage en contexte multi-usager ;
 - b. la valeur économique des apports en eau dans la gestion optimale d'une réserve d'eau superficielle destinée à la production hydroélectrique.

4.2. Utilité en termes d'anticipation et d'alerte aux crues

Le but des prévisions d'ensemble est, d'une part, de pouvoir étendre l'échéance de prévision et, d'autre part, de fournir une quantification des incertitudes de prévision à partir des différents scénarios qui en sont issus. Une question qui se pose est de savoir si les prévisionnistes seraient en mesure de prévoir un événement de crue avec un délai plus long si une meilleure prévision de l'événement était émise par la chaîne de modèles hydrométéorologiques. En autres termes, il s'agit de savoir si les prévisionnistes seraient mieux guidés par le champ de scénarios possibles (dans le temps et dans l'espace) proposé par une prévision d'ensemble.

Du point de vue de l'utilisateur, la chaîne de prévision de crues peut être découpée en prévision, alerte et décision, et réponse (action), avant l'observation d'un événement (Parker et Fordham, 1996 ; Carsell *et al.*, 2004 ; Verkade, 2015) (**Figure 21**). Le délai de prévision est défini comme étant le temps écoulé entre la production d'une prévision et le début d'un événement de crue/inondation. Le temps qui reste après l'alerte et la prise de décision est appelé temps d'atténuation. C'est le temps qui peut être utilisé pour atténuer les dommages réels que l'événement peut engendrer. En supposant que le temps requis pour l'alerte et la prise de décision reste inchangé, un gain en délai de prévision (anticipation) se traduira automatiquement davantage de temps disponible pour l'action face à l'événement.

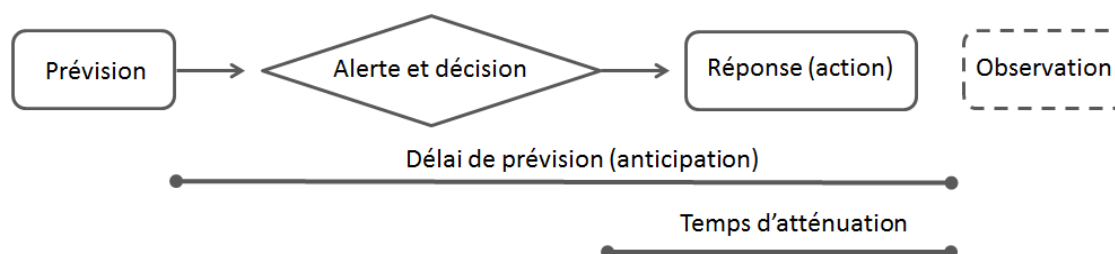


Figure 21 : Découpage de l'échelle temporelle de la chaîne de prévision en délai de prévision et temps d'atténuation des dommages avant l'arrivée d'un événement.

L'un des défis de la prévision immédiate des crues rapides est de fournir des avertissements sur des petits cours d'eau de temps de réponse inférieurs à six heures et d'informer les autorités et la population suffisamment en avance pour permettre la mise en sécurité des biens et des personnes (Ruin *et al.*, 2014 ; Gaume et Borga, 2013 ; Liguori et Rico-Ramirez, 2012). Dans cette configuration, on se heurte à plusieurs difficultés : par exemple, les phénomènes orageux à l'origine des crues soudaines sont difficiles à prévoir par les modèles numériques de prévision, autant en intensité qu'en localisation à l'intérieur du bassin versant ; les crues sur ces bassins ne sont qu'exceptionnellement mesurées par les réseaux classiques de l'hydrométrie ; le temps de réponse des petits bassins est très rapide, laissant très peu de temps pour avertir les services de secours.

L'utilité de la prévision d'ensemble pour la prévision immédiate a été abordée, de manière préliminaire, à la fin de la **thèse d'A. Caseri (2013-2016)**. Dans sa thèse, en plus des aspects liés à la qualité des prévisions, présentés dans la **section 3.2**, nous nous sommes également intéressés à évaluer le potentiel de l'ensemble pluviométrique généré à fournir une information utile pour la construction d'une représentation spatialisée de la probabilité de crue.

Nous avons montré que l'approche d'ensemble proposée dans la thèse peut fournir une indication anticipée du risque de crue, avec des alertes plus justes et moins dispersées sur le territoire de surveillance au fur et à mesure que le pic de l'événement s'approche. Le système AIGA d'avertissement aux crues rapides, développé à Irstea à Aix-en-Provence, avait été utilisé et nous l'avons adapté pour un fonctionnement en situation de prévision immédiate d'ensemble. Les prévisions émises par le système mis en place semblent bien se prêter à la détection du dépassement d'un seuil critique, ce qui a été confirmé en comparant les zones à risque indiquées par la prévision avec les dégâts enregistrés lors de l'événement (**Figure 22**).

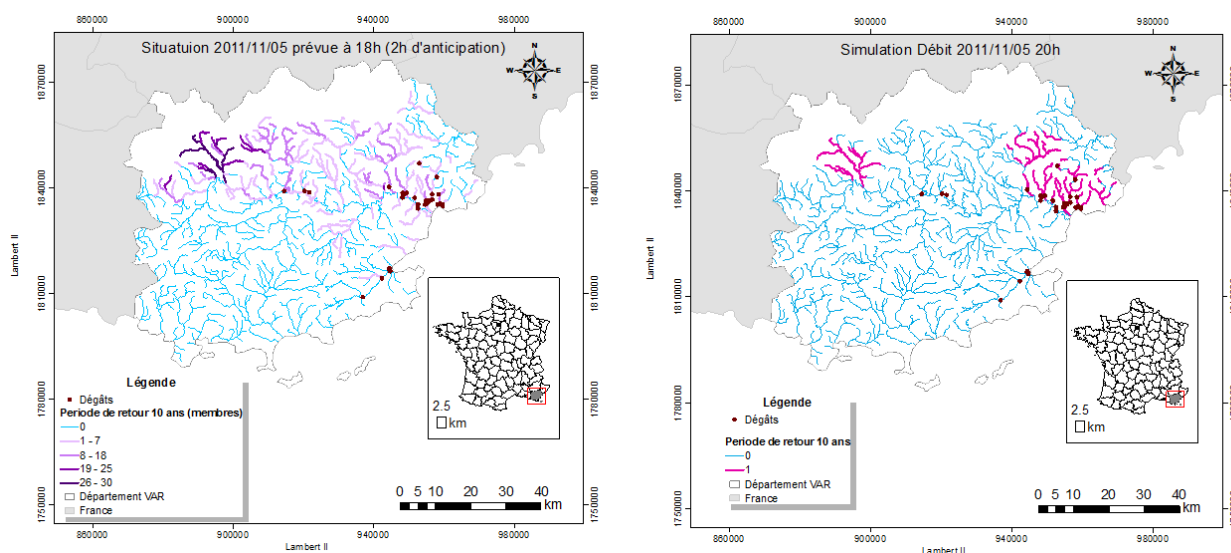


Figure 22 : Représentation spatialisée de la probabilité de dépassement du débit horaire de période de retour de 10 ans dans le Département du Var pour le 05 de novembre 2011 à 20h, générée à partir des débits simulés AIGA : à gauche, en utilisant les prévisions d'ensemble SAMPO-TBM (méthode PI-GAE de prévision immédiate) pour l'échéance de 2 heures ; à droite, en utilisant la pluie observée par le radar ANTILOPE (dans : Caseri, 2017)

Des pistes pour une poursuite de ce travail restent ouvertes et incluent une meilleure évaluation quantitative de la méthode de simulation (sur un plus grand nombre d'événements et d'autres domaines géographiques), un meilleur conditionnement de la simulation géostatistique à l'information de l'intermittence spatiale de la pluie (zones de pluie nulle) et la mise en place d'une étude où les méthodes géostatistique et physique (prévisions numériques issues de modèles météorologiques, comme les prévisions AROME PE de Météo-France) peuvent être combinées dans un système de prévision immédiate d'ensemble. Certains de ces aspects seront poursuivis dans le cadre de la thèse de Daniela Peredo (démarrée en octobre 2017, que je dirige avec V. Andréassian et L. Oudin, UPMC) et du projet ANR PICS (2018-2022, coordonné par Olivier Payrastre, Ifsttar), auquel je participe.

Sur le sujet de la prévision immédiate, des collaborations internationales ont été également établies avec le Brésil dans le cadre des travaux de Master de Luiz Bacelar (INPE/Cemaden, São Paulo ; accueilli à Irstea en février 2017) et du post-doctorat d'Angélica Caseri (Cemaden, São Paulo ; depuis juin 2017) pour la simulation géostatistique d'événements de pluie intense observés à Rio de Janeiro et à São Paulo. Ces collaborations sont complétées par celles établies dans le cadre de la thèse de Fernando Neves Lima (doctorant de l'Université Fédérale de Minas Gerais accueilli à Irstea à Antony durant l'année 2017), pour la simulation d'événements pluvieux à Belo Horizonte. Ce sont des opportunités ouvertes pour examiner l'utilité de l'approche développée dans la thèse d'A. Caseri en contexte d'orages convectifs en milieu fortement urbanisé.

Enfin, on notera que, bien que la simulation géostatistique ait donné des scénarios fiables de champs de pluie, notre étude portant sur son utilisation pour la prévision immédiate d'ensemble a montré un gain d'anticipation relativement faible, de 1 à 2 heures d'anticipation par rapport à une prévision naïve par persistance, où le dernier champ de pluie observé par le radar est considéré à l'identique au long des prochaines heures de prévision. Ceci dit, l'étude de Golian *et al.* (2014) suggère que l'augmentation du délai de prévision peut se traduire par une réduction des pertes économiques, et ce notamment dans la plage des courts délais horaires : une augmentation du temps d'alerte de 1 à 6 heures, amène à une réduction des dégâts de 18 à 30%. Leur étude considère le délai d'anticipation comme un facteur important pour déterminer l'efficacité d'un système de prévision, d'alerte et de réponse aux crues et inondations. Cela souligne l'importance de toute augmentation du délai de prévision, même si celle-ci peut paraître faible, comme dans le cas des résultats que nous avons obtenus sur les crues rapides du département du Var.

Les travaux de **Master de W. Weeink (2010)**, que j'ai encadrés en collaboration avec l'Université de Twente aux Pays-Bas, ont aussi porté sur la valeur ajoutée des prévisions d'ensemble, par rapport à une prévision déterministe classique, pour l'anticipation aux événements de crue. Différemment des travaux d'A. Caseri, où la composante spatiale des champs pluviométriques et le cas des bassins non-jaugés étaient dominants, cette étude s'est plutôt focalisée sur des prévisions à moyenne échéance (horizon de prévision à plusieurs jours).

Reprenant des études préliminaires que j'avais faites pendant mon post-doctorat au JRC sur les prévisions hydrologiques du système EFAS (Ramos *et al.*, 2009), nous nous sommes penchés sur la question du choix d'un seuil optimal de prévision d'ensemble. Ce seuil est lié à la probabilité de prévision d'un événement donné, ou à la probabilité de dépasser un certain seuil de débit requis pour émettre un avertissement. Dans le cas de la prévision d'ensemble, cette probabilité est estimée par la proportion de membres de la prévision d'ensemble dépassant le seuil qui

définit l'événement critique auquel on s'intéresse. Un seuil optimal de décision est ainsi un « seuil de décision » pour le prévisionniste quant à alerter ou non un décideur de la situation à venir. Il se traduit par la question : quel est le nombre minimum de membres de la prévision d'ensemble dépassant un seuil critique de débit nécessaire pour lancer un avertissement ? Ce choix de seuil cherche à minimiser le nombre de fausses alertes et alertes manquantes, tandis que le nombre de crues correctement prévues est optimisé. Dans l'étude de W. Weeink, le seuil de prévision d'ensemble optimal considérait aussi l'anticipation aux crues, traduite en gain en délai de prévision comparé aux prévisions déterministes.

L'étude a été basée sur 208 bassins versants en France et le modèle GR3P journalier que j'avais adapté pour la prévision d'ensemble. Le modèle a été alimenté par les prévisions du CEPMMT. Il s'agissait de prévisions déterministes et d'ensemble (51 membres) de précipitations pour un horizon maximal de prévision de 10 jours. Les données étaient disponibles sur une période de 18 mois.

L'étude a montré qu'un seuil unique, qui pourrait être appliqué à tous les bassins versants, et qui entraînerait une plus grande valeur de CSI et un gain en anticipation comparé à une prévision déterministe pour le dépassement de débits forts, ne pouvait pas être déterminé. Cependant, il a été montré que les critères de qualité considérés peuvent être optimisés quand il s'agit de rechercher un seuil de prévision d'ensemble spécifique à chaque bassin versant. Dans ce cas de figure, les prévisions d'ensemble ont montré un gain moyen en délai d'anticipation d'environ 2-3 jours pour les prévisions de dépassement de forts débits (**Figure 23**).

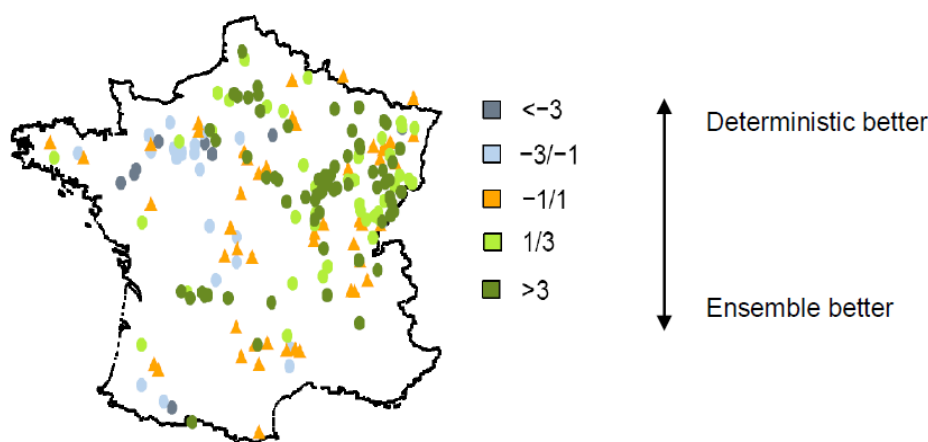


Figure 23 : Gain ou perte en délai d'anticipation moyen de la prévision d'ensemble par rapport à la prévision déterministe (en jours) pour le dépassement du quantile 99% des débits journaliers observés pour 208 bassins versants en France lorsqu'un seuil spécifique au bassin versant est appliqué (dans : Weeink, 2010)

Des règles de décisions quantitatives et généralisables sont difficiles à définir. Les « seuils probabilistes de décision » dépendent de la configuration et des propriétés du système de prévision considéré, des attributs de la qualité espérée du système et des tolérances des utilisateurs vis-à-vis des fausses alertes ou des alertes manqués. De plus, un seuil de dépassement ne tient pas nécessairement compte de la gravité des conséquences des crues et inondations. Dans ce cas, l'analyse « coût-bénéfice » peut fournir un cadre plus rationnel pour l'évaluation des systèmes de prévision, alerte et réponse aux crues.

4.3. Valeur économique relative d'un système de prévision et alerte aux crues

Toujours dans l'idée d'examiner les liens entre qualité et valeur des prévisions, j'ai contribué à une étude menée par A. Thiboult à l'Université Laval à Québec dans le cadre de sa thèse, dirigée par F. Anctil. Dans Thiboult *et al.* (2017), nous présentons les résultats d'une analyse menée sur six systèmes d'alerte construits pour étudier comment la quantification des incertitudes affecte la qualité et la valeur d'une décision.

Ces systèmes diffèrent par la façon dont ils prennent en compte les principales sources d'incertitude qui jouent un rôle dans la prévision hydrométéorologique et, par conséquent, par la quantité de l'incertitude totale qu'ils traitent. Ils sont évalués à l'aide de critères de qualité des prévisions d'ensemble et de la Valeur économique relative (REV), qui est issue de l'analyse du rapport coût/perte (« cost/loss ratio »).

La valeur économique relative suppose un modèle coût/perte de prise de décision dans lequel le coût de l'action est C (que l'événement se produise ou non) et la perte encourue si l'événement survient mais qu'aucune mesure n'est prise est L , donnant ainsi le rapport C/L . Ce rapport coût-perte (CLR) représente donc le ratio entre les coûts d'atténuation (mesures de protection) et les pertes (dégâts) évitables dues à un événement indésirable. L'analyse se base sur un tableau de contingence (**Tableau 4**).

Tableau 4 : Tableau de contingence avec les coûts associés à chaque type d'événement (dans : Thiboult *et al.*, 2017)

		Warning issued	
		Yes	No
Event observed	Yes	Hit (h) Mitigated Loss ($C + L_u$)	Miss (m) Loss ($L_a + L_u$)
	No	False Alarm (f) Warning Cost (C)	Correct negative (c) No Cost ($-$)

La valeur économique relative (REV) est ensuite évaluée. Il s'agit d'un facteur adimensionnel qui évolue entre le cas où aucune prévision n'est émise (donc aucun avertissement) et le cas de « prévision parfaite », où un avertissement est émis pour chaque événement indésirable. Un REV égal à 1 indique le meilleur système de décision possible, alors qu'un REV égal à 0 indique que le système est équivalent à un cas sans prévision.

Pour que les résultats puissent être utiles à des décideurs sensibles à des rapports coût/perte différents, les REV obtenues sont souvent représentées en fonction du rapport C/L . En outre, lorsque l'analyse est appliquée à des prévisions probabilistes, les courbes de valeurs relatives sont calculées pour plusieurs seuils de probabilité. L'enveloppe des courbes des valeurs de REV représente ainsi la valeur économique potentielle du système évalué. Un décideur peut maximiser la valeur des prévisions probabilistes en choisissant le seuil de probabilité le plus approprié pour le rapport C/L associé à son contexte décisionnel.

La **Figure 24** montre un exemple tiré de Buizza *et al.* (2007), où on observe les différences entre les courbes de REV obtenues pour un système fiable (graphique à gauche) et un système biaisé (graphique à droite). Le lien qualité-valeur économique potentielle est établi pour des prévisions synthétiques de précipitation. Ces prévisions sont tirées d'une loi Gaussienne : 51 prévisions synthétiques diffèrent légèrement les unes des autres, et le champ d'observation synthétique soit (a) coïncide avec l'un des champs de prévision (cas de la prévision non biaisée), soit (b) diffère du champ de prévision dans sa forme et positionnement (cas de la prévision biaisée, où les prévisions sous-estiment la quantité de précipitation et ne parviennent pas à placer les précipitations à l'emplacement correct).

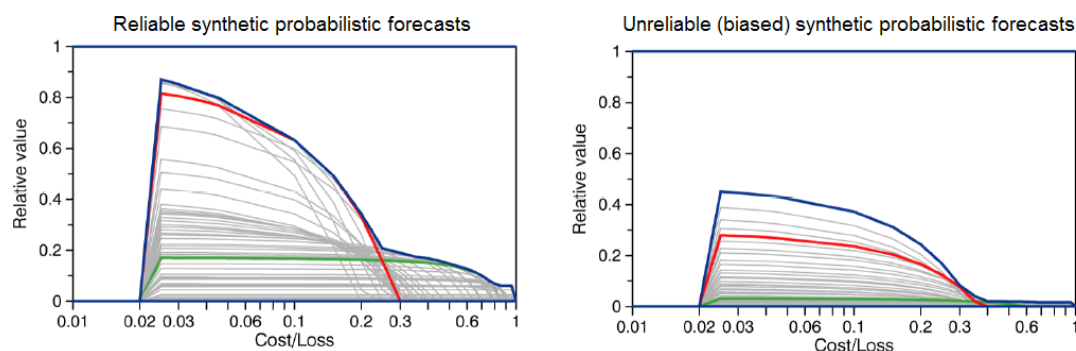
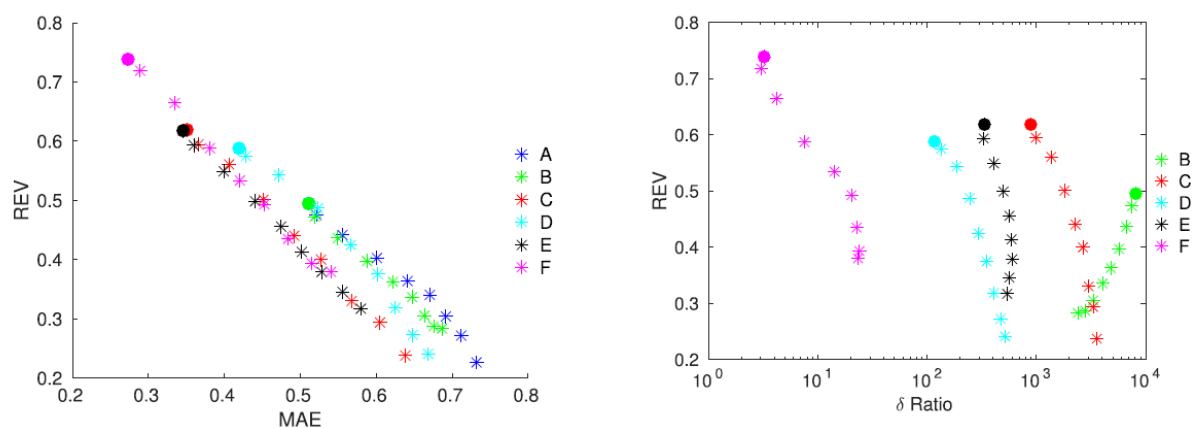


Figure 24 : Valeur économique potentielle de prévisions probabilistes synthétiques de précipitations supérieures à 10 mm/j pour différents seuils de probabilité (lignes grises). Les résultats pour les seuils de probabilité de 10% (ligne rouge) et 60% (ligne verte) sont soulignés. La valeur économique potentielle de l'ensemble est définie par l'enveloppe de toutes les courbes (ligne bleue). A gauche : ensemble fiable ; à droite : ensemble biaisé (d'après : Buizza *et al.*, 2007)

Plusieurs applications de l'approche REV sont proposées dans la littérature pour les prévisions hydrométéorologiques (Verkade et Werner, 2011 ; Roulin, 2007 ; Atger, 2001 ; Richardson, 2000). Dans l'étude que nous avons proposée dans Thiboult *et al.* (2017), nous avons voulu aller un peu plus loin, vers la quantification du lien entre critères de qualité des prévisions émises et valeurs de REV, en considérant des « vraies » prévisions d'ensemble.

L'étude est basée sur 20 bassins versants au Québec (Canada) et trois approches de quantification d'incertitude, utilisées seules ou de manière combinée : (1) quantification de l'incertitude des forçages météorologiques à l'aide de la prévision d'ensemble du CEPMMT, (2) quantification de l'incertitude des conditions initiales à partir d'une assimilation de données basée sur le filtre de Kalman d'ensemble (EnKF), et (3) quantification de l'incertitude de la modélisation hydrologique, basée sur la construction d'un ensemble multi-modèles constitué de 20 modèles hydrologiques globaux (Perrin, 2000 ; Seiller *et al.*, 2012).

Les résultats ont montré que la quantification des sources d'incertitude est toujours bénéfique : tous les systèmes fournissent des valeurs de REV positives jusqu'à 9 jours d'échéance. Des systèmes plus complexes (qui prennent en compte plusieurs sources d'incertitude) fournissent une valeur économique plus élevée. De manière générale, les systèmes de prévisions avec une meilleure précision et fiabilité ont aussi donné de meilleures valeurs de REV. Cependant, alors que le lien était plus clairement défini pour l'attribut de précision, il n'en était pas de même pour le lien entre REV et critère de fiabilité (**Figure 25**). J'avais déjà été confrontée à cela dans une autre étude, menée dans le cadre des travaux de thèse de Ioanna Zalachori (2009-2013), comme on le verra dans la **section 4.4** qui suit.



Systems	A	B	C	D	E	F
Multimodel	Off	Off	Off	On	Off	On
EnKF	Off	Off	On	Off	On	On
Met. ensemble	Off	On	Off	Off	On	On
Nb of members	(20×)1	(20×)50	(20×)50	20	(20×)2500	50,000

Figure 25 : Relations entre REV et l'erreur absolue moyenne (MAE ; à gauche) et la fiabilité (ratio issu du diagramme de rang ; à droite), calculée pour un CLR = 0,5. Les cercles indiquent l'échéance 1 jour et les astérisques les échéances journalières suivantes. Les systèmes évalués comprennent différentes sources d'incertitude (tableau en bas de figure) (dans : Thiboult et al., 2017)

4.4. Apport des prévisions hydrologiques pour la gestion des réservoirs

Autant l'importance d'une alerte précoce et d'une bonne anticipation des événements hydrologiques extrêmes est indéniable pour les pouvoirs publics, notamment s'agissant de la protection de vies humaines et de biens situés en zones vulnérables, autant l'enjeu économique des prévisions hydrologiques ressort davantage face aux utilisateurs industriels exposés aux risques climatiques et hydrologiques à différentes échelles (génération d'électricité, approvisionnement en eau potable, agriculture, navigation, etc.).

Dans mes recherches, je me suis intéressée à deux cas d'évaluation de l'apport des prévisions d'ensemble pour la gestion de réservoirs :

1. l'apport des prévisions saisonnières d'ensemble pour l'évaluation des risques liés aux périodes d'étiage pour le cas du réservoir d'Arzal (bassin versant de la Vilaine, Bretagne) (en lien avec le projet Interreg IVB DROP et en partenariat opérationnel avec l'EPTB IAV - Institution d'Aménagement de la Vilaine),
2. l'apport des prévisions à moyenne échéance, déterministes et d'ensemble, brutes ou post-traitées, dans la gestion hebdomadaire d'une réserve en eau superficielle destinée à la production hydroélectrique (en partenariat opérationnel avec EDF DTG Grenoble et en lien avec le projet H2020 IMPREX).

□ *Prévisions hydrologiques saisonnières pour la gestion d'un réservoir d'eau potable*

L'implémentation des prévisions saisonnières dans les processus de prise de décision des services hydrologiques opérationnels rencontre encore de nombreux obstacles. Ceux-ci sont notamment liés à la qualité actuelle des prévisions numériques de précipitation à l'échelle saisonnière ou à la difficulté de traduire l'information disponible en risque hydrologique face à un événement critique qui doit être géré en temps réel par les acteurs opérationnels.

Les travaux de **thèse de Louise Crochemore (2016)** ont été motivés par les besoins des gestionnaires du barrage d'Arzal en outil de prévision des étiages pour venir en appui à la gestion de leur réserve en eau potable pendant la saison estivale de basses eaux (période de gestion allant de mai à octobre). Le site était un des sites pilotes du projet DROP (Interreg IVB) (2013-2015), qui avait deux volets de développement : (1) un volet dédié à concevoir un outil d'évaluation de la gouvernance pour identifier les modalités d'action locale qui soutiennent ou desservent les mesures de lutte contre la sécheresse et la pénurie d'eau, et (2) un volet de développement et implémentation de mesures concrètes sur des sites pilotes choisis dans six régions du nord-ouest de l'Europe. C'est à ce deuxième volet que se réfèrent les travaux ici reportés.

Un des objectifs de la thèse de L. Crochemore a été d'étudier comment les prévisions saisonnières pourraient contribuer à la gestion actuelle du réservoir d'Arzal. Les apports en eau au réservoir d'Arzal proviennent du bassin versant de la Vilaine (10 000 km²). A son embouchure, la Vilaine se jette dans le réservoir d'Arzal puis dans l'Océan Atlantique. Le barrage d'Arzal est géré par l'EPTB IAV (Institution d'Aménagement de la Vilaine). Initialement construit pour arrêter les marées qui contribuaient aux inondations du bassin aval de la Vilaine, le barrage crée un réservoir d'eau douce de 50 Mm³, permettant d'approvisionner la région en eau potable.

Aujourd'hui le barrage d'Arzal est un ouvrage aux usages multiples : évacuation des crues, production d'eau potable, navigation, agriculture, gestion environnementale des marais et suivi des poissons migrateurs. Au printemps et à l'automne, les usages peuvent être satisfaits sans difficulté particulière. Cependant, pendant les crues de la Vilaine ou lors de périodes prolongées de faible débit, des conflits peuvent être observés. En période de crue, l'objectif d'évacuation des débits est prioritaire, ce qui peut gêner ou entraîner l'interdiction de la navigation. En période d'étiage, l'objectif d'alimentation en eau potable devient prioritaire et toute intrusion d'eau de mer dans le réservoir doit être évitée. Dans ce cas, des restrictions d'éclusages peuvent être décidées, ce qui est pénalisant pour la navigation⁴⁸. Cela a été le cas à la fin de l'été 2016 (**Figure 26**).

Grâce aux données des différentes entrées et sorties du réservoir d'Arzal, mises à disposition par l'IAV sur la période de 2005 à 2011, nous avons mis en place un modèle simple de bilan d'eau du réservoir (Crochemore, 2016). Les entrées du réservoir sont les apports de la Vilaine et les précipitations. Les sorties comprennent l'eau évacuée via les vannes et volets pour la gestion du réservoir, via la passe à poissons, l'écluse et les siphons, ainsi que l'eau prélevée pour l'eau potable et l'eau évaporée (**Figure 27**).

⁴⁸ <http://www.eptb-vilaine.fr/index.php/f-barrage/32-un-ouvrage-multi-usages>

Sécheresse. La fermeture de l'écluse du barrage d'Arzal prolongée



Modifié le 21/09/2016 à 16:46 | Publié le 21/09/2016 à 16:46



En raison de la sécheresse, l'Institution d'aménagement de la Vilaine a décidé de prolonger pour le mois d'octobre les restrictions d'éclusage au barrage d'Arzal. | OF

En raison de la sécheresse, l'Institution d'aménagement de la Vilaine a décidé de prolonger pour le mois d'octobre les restrictions d'éclusage au barrage d'Arzal.

Les pluies de la mi-septembre n'ont eu que très peu d'impact sur les conditions hydrologiques de la Vilaine. Les niveaux de chlorures restent préoccupants pour la production d'eau potable et les prévisions météorologiques ne font pas état de pluies significatives d'ici la fin du mois de septembre.

Cette situation oblige l'Institution d'aménagement de la Vilaine (IAV) à prolonger pour le mois d'octobre les restrictions d'éclusage.

Figure 26 : Lors de périodes prolongées de faible débit, des conflits d'usage peuvent être observés dans la gestion du barrage-réservoir d'Arzal, à l'embouchure du bassin de la Vilaine (Source : <https://www.ouest-france.fr/>)

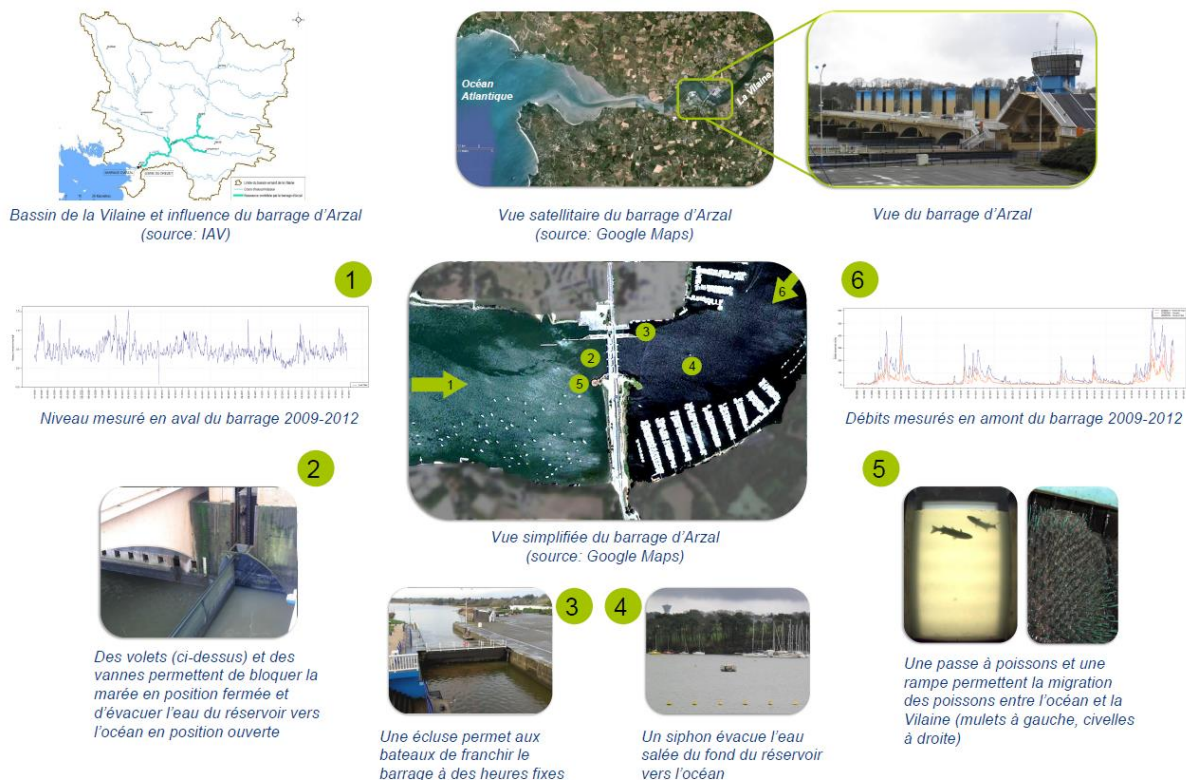


Figure 27 : Barrage d'Arzal - un ouvrage aux usages multiples (poster L. Crochemore)

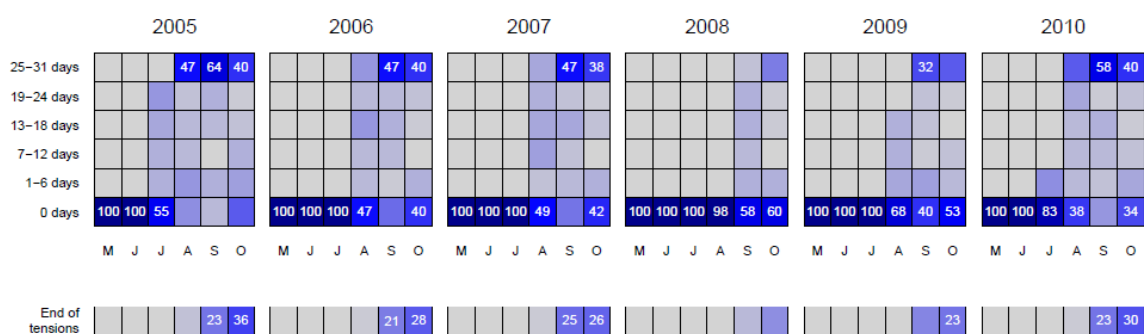
La **Figure 28** montre l'outil proposé, après plusieurs échanges avec l'IAV, visant à évaluer le risque de ne pas pouvoir maintenir un niveau minimum dans le réservoir d'Arzal entre mai et octobre (saison de gestion). La conception de l'outil était guidée par les objectifs et les besoins de gestion du barrage, et basée sur les niveaux d'eau et les restrictions d'usage, formulées relativement à des niveaux seuils du réservoir.

L'outil a été testé en utilisant les débits prévus à la station de Rieux au Pont de Cran, à l'aide de l'approche ESP de prévision saisonnière et du modèle hydrologique GR6J. Ces débits ont ensuite été utilisés dans un modèle de bilan d'eau du réservoir mis en place pendant la thèse de L. Crochemore. Les graphiques proposent de représenter la probabilité d'être en dessous du niveau de réservoir minimum acceptable de 1,80 m NGF pour différentes durées. Ils informent également (graphique en bas) la probabilité de revenir à un niveau supérieur au niveau minimum au cours du mois représenté.

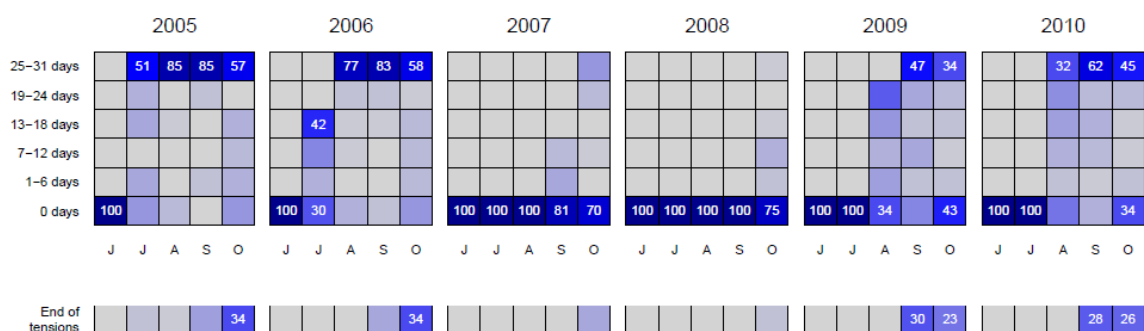
Les sorties des vannes et volets sont optimisées pour maintenir un niveau objectif dans le réservoir, tandis que les sorties liées aux autres usages sont supposées constantes et maximales. Les prévisions de niveaux de réservoir ainsi obtenues nous permettent d'identifier le risque de franchir le seuil minimal dans le réservoir et de déterminer le nombre de jours et le nombre de membres de la prévision d'ensemble atteignant ce seuil. Ces variables d'intérêt pour la gestion du risque sont ainsi réunies dans le graphique d'évaluation du risque. Les graphiques ont été produits rétrospectivement pour les périodes de basses eaux de 2005 à 2010 afin d'évaluer leur potentiel en conditions opérationnelles (Crochemore, 2016).

On notera qu'il n'a pas été possible d'évaluer l'outil proposé par rapport aux observations, car l'outil n'utilise pas les lâchers réels du réservoir, mais les lâchers optimisés pour maintenir un niveau constant et qui sont choisis pour toujours fournir de l'eau aux usages prescrits. Les lâchers réels sont décidés en fonction d'un processus complexe qui prend également en compte la qualité de l'eau, facteur qui n'a pas été introduit dans notre modélisation. Néanmoins, les graphiques permettent de visualiser le risque d'être en dessous d'un niveau minimum qui garantit les usages, et de distinguer les années les plus sèches (2005 et 2010), les années les plus humides (2007, 2008) et les années plutôt intermédiaires à ces deux situations (2009 et 2006).

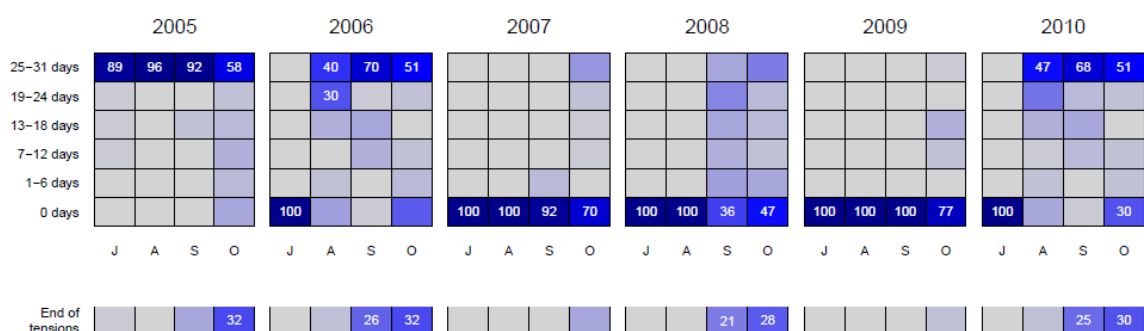
a) Prévisions d'apport au 1^{er} mai



b) Prévisions d'apport au 1^{er} juin



c) Prévisions d'apport au 1^{er} juillet



d) Prévisions d'apport au 1^{er} août

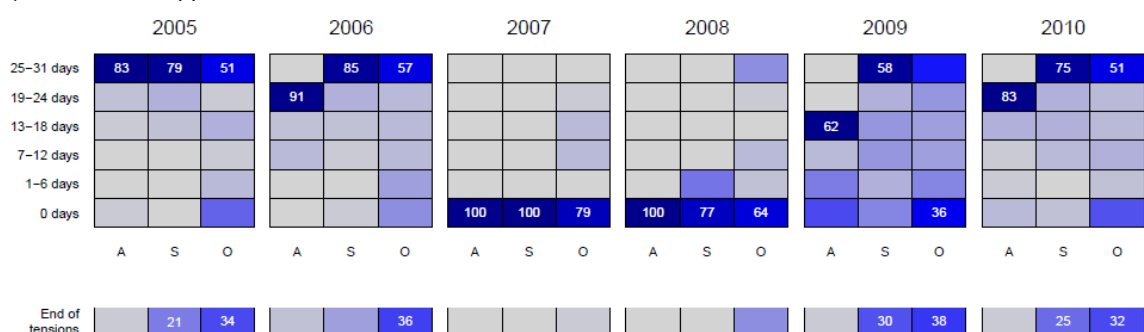


Figure 28 : Graphiques de risque proposés pour représenter la probabilité d'être en dessous du niveau de réservoir minimum acceptable de 1,80 m NGF pour des durées différentes (0 jour jusqu'à 25-31 jours). La ligne en dessous (« end of tensions ») correspond à la probabilité de revenir à un niveau supérieur à 1,80 m NGF et pouvoir revenir à une gestion hors conflits d'usage. Les graphiques sont proposés pour six années différentes (2005 à 2010) et des dates de début de prévision allant du 1er mai (a) au 1er août (d) ; prévisions valables pour les horizons de prévision mensuels notés en bas des graphiques et comptés à partir de la date de début de prévision. Classement des années les plus sèches aux années les plus humides : 2005, 2010, 2009, 2006, 2007, 2008 (dans : Crochemore, 2016)

□ *Valeur économique des apports en eau dans la gestion optimale d'une réserve en eau destinée à la production hydroélectrique*

Dans mes recherches, le cas de la valorisation des prévisions hydrologiques d'ensemble dans le secteur hydroélectrique a été d'abord abordé dans la **thèse de doctorat de Ioanna Zalachori** (2009-2013 ; co-financement EDF DTG et Irstea) et, ensuite, dans les **travaux de Master (2016) et thèse (en cours) de Manon Cassagnole** (financement projet H2020 IMPREX). Ces premiers pas dans l'évaluation économique des prévisions pour le secteur de l'hydroélectricité ont marqué le début d'une série de travaux liés au secteur hydroélectrique⁴⁹.

Dans les études citées, nous avons suivi une approche prescriptive pour modéliser la « meilleure décision » qu'un gestionnaire peut prendre. Cela veut dire que nous nous intéressons à évaluer le gain apporté par les prévisions étant donné l'action optimale d'un décideur, sans prendre en compte son comportement réel. Une étude prescriptive sert à identifier la meilleure décision sous l'hypothèse d'un décideur idéal (c'est-à-dire, juste et rationnel) et parfaitement informé (Wilks, 1997). L'approche décrit ainsi, théoriquement, comment les décideurs doivent prendre leurs décisions. Les outils opérationnels de gestion sont souvent développés sur ce principe⁵⁰.

En outre, en contraste avec les algorithmes d'optimisation qui assurent une convergence mathématique vers la solution optimale, nous avons adopté une approche heuristique d'optimisation de la gestion d'une réserve en eau destinée à la production hydroélectrique. Basée sur les connaissances empiriques et l'expérience des gestionnaires, l'approche heuristique ne garantit pas une convergence vers la décision optimale, globale ou locale, mais propose une solution approchée, qui a l'avantage d'être plus simple à mettre en œuvre.

Dans le cadre de la thèse d'I. Zalachori (2009-2013), cofinancée par Irstea et EDF DTG Grenoble, un modèle d'optimisation heuristique de la gestion d'une retenue du type éclusée hebdomadaire pour la production hydroélectrique a ainsi été développé. Gérer un aménagement hydroélectrique revient à « décortiquer le système contrôlable, ses entrées et les usages, puis à les modéliser » (Roche *et al.*, 2012). Il s'agit de simuler le système (entrées, variables, contraintes) et rechercher les meilleures règles de décision de façon à se rapprocher au mieux des objectifs de gestion.

⁴⁹ J'ai été invitée à participer au projet FP7 Complex (2012-2016, « *Knowledge based climate mitigation systems for a low carbon economy* ») sur les liens entre variabilité climatique et énergies renouvelables (éolien, solaire et hydroélectricité) (François *et al.*, 2014 ; Engeland *et al.*, 2017 ; Creutin *et al.*, 2016a,b), et à coordonner le WP « *Hydropower* » du projet H2020 Imprex (2015-2019) (Hurk *et al.*, 2016 ; Ramos *et al.*, 2016). A cela, s'ajoute ma participation au projet récemment initié ERA4CS / JPI Climate AQUACLEW (2017-2020, « *Advancing QUALity of CLimate services for European Water* »), financé par l'ANR pour les partenaires français, dans le cadre duquel j'encadre la thèse de A. Lemoine sur l'utilisation d'indicateurs climatiques pour la gestion de l'hydroélectricité.

⁵⁰ En pratique, deux voies sont souvent indiquées pour identifier la meilleure décision : les études prescriptives et les études descriptives (Wilks, 1997 ; Stewart, 1997). Contrairement à la démarche que j'ai adoptée dans mes recherches, les études descriptives cherchent à décrire le comportement réel des décideurs et sont basées sur le fait que, en réalité, les décideurs s'éloignent, pour des raisons diverses, de l'optimalité (Houdant, 2004 ; Ramos *et al.*, 2010a). Les deux approches, prescriptives et descriptives, sont, bien évidemment, liées entre elles : des hypothèses sur le comportement des décideurs peuvent être énoncées et testées face au comportement réel des décideurs devant le processus de prise de décision (Zalachori, 2013).

Typiquement, les aménagements d'écluse permettent de stocker l'eau arrivant durant les heures creuses de la semaine (nuit et week-end, par exemple) pour la turbiner pendant les heures pleines. Ces aménagements ont une capacité utile de l'ordre de 1 à 15 jours et une capacité de turbinage entre 2 à 4 fois le débit moyen (**Figure 29**). Dans notre démarche, le modèle de gestion est basé sur ces deux paramètres principaux que l'on peut faire varier pour mieux s'adapter à une situation donnée : le coefficient de turbinage de l'usine et la capacité de la retenue. Ces deux paramètres sont exprimés en fonction du débit journalier moyen au point du réseau hydrographique où se situe le réservoir modélisé.

Nous n'avons donc pas eu l'ambition de modéliser dans le détail toutes les variables et contraintes d'un aménagement réel, mais avons choisi de travailler avec des aménagements « fictifs », simplifiés, pour lesquels nous faisons varier la capacité utile et de turbinage dans les limites indiquées par les données de la **Figure 29**. Ceci dit, on notera que les apports en débits prévus et observés au droit de ces aménagements fictifs seront ceux des bassins versants réels. Cette configuration garantit la facilité d'application du modèle à n'importe quelle dimension d'aménagement hydroélectrique et à plusieurs séries d'entrée d'apports en eau (prévisions hydrologiques).

Dans nos études, nous avons appliqué le modèle de gestion aux apports en débit sur 7 jours observés sur un échantillon de 10 à 15 bassins versants de fort intérêt économique pour la production hydroélectrique en France. Ces bassins versants sont répartis dans les principaux massifs français (Alpes, Massif Central, Cévennes et Jura) et situés en amont des retenues d'eau, dont les apports alimentent les lacs-réservoirs des barrages gérés par EDF, ou font figure de témoins importants pour la génération d'électricité (Zalachori, 2013 ; Cassagnole, 2016).

Le gain économique des prévisions en gestion de production électrique a été évalué sur la base du prix de marché de l'énergie EPEX-SPOT (en €/MWh), supposé traduire la valeur de l'énergie. Nous travaillons donc dans le cas de figure des prix connus à l'avance.

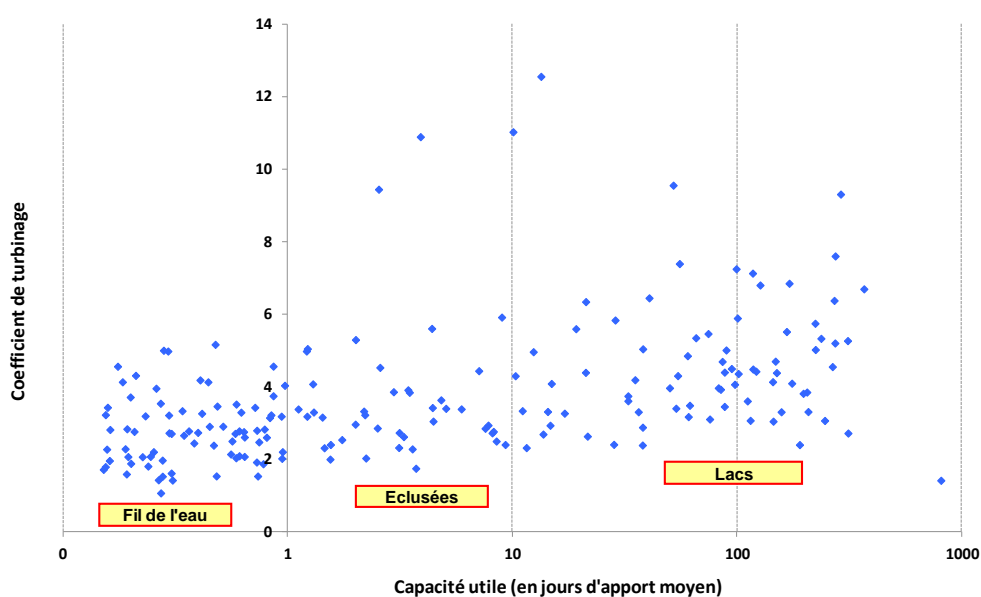


Figure 29 : Classification générale des aménagements hydroélectriques en fonction du coefficient de turbinage et de la capacité utile (en jours d'apport moyen) (source des données : EDF ; dans : Zalachori, 2013).

Dans l'outil de gestion développé, une règle de gestion se traduit par une programmation, pour la semaine à venir, des heures de turbinage et d'arrêt, dans le but d'optimiser la production électrique. Dans une logique d'optimisation des gains, il s'agit de choisir les heures de turbinage en fonction du prix horaire de l'électricité. Ce prix est variable dans le temps et suit les fluctuations de consommation, qu'elles soient journalières (heures de pointe du matin et du soir) ou hebdomadaires (différences entre jours ouvrables et week-end). Une fois la règle de gestion établie, on procède à son application. Pour cela, la gestion doit être simulée avec les débits observés. Cela permet d'évaluer, à chaque instant de prévision/observation (jour J de la période d'étude), le gain lorsque l'on turbine les débits observés aux heures programmées *a priori* par la gestion établie en fonction des apports prévus.

Le modèle de gestion est expliqué en détail dans Zalachori (2003) et Cassagnole (2016). La **Figure 30** illustre son fonctionnement. Dans son ensemble, l'approche développée pour l'estimation de la valeur d'une prévision comprend trois étapes de calcul :

- A chaque jour J de la période d'évaluation, le modèle de gestion est appliqué, sur la base d'une prévision d'apports sur 7 jours à venir, pour établir la règle de gestion à considérer : étape de gestion.
- Ensuite, les gains de gestion sont calculés en appliquant la règle de gestion aux débits observés à la fin du jour J : étape de simulation.
- A la fin de la période d'évaluation T , le gain total, sur l'ensemble des jours J , est évalué : étape d'évaluation.

Dans ce mémoire, je focalise sur la présentation et discussion des principaux résultats obtenus de l'application du modèle de gestion à différentes configurations d'étude. Les gains obtenus à la fin de la période d'évaluation pour différents apports prévus sont comparés entre eux.

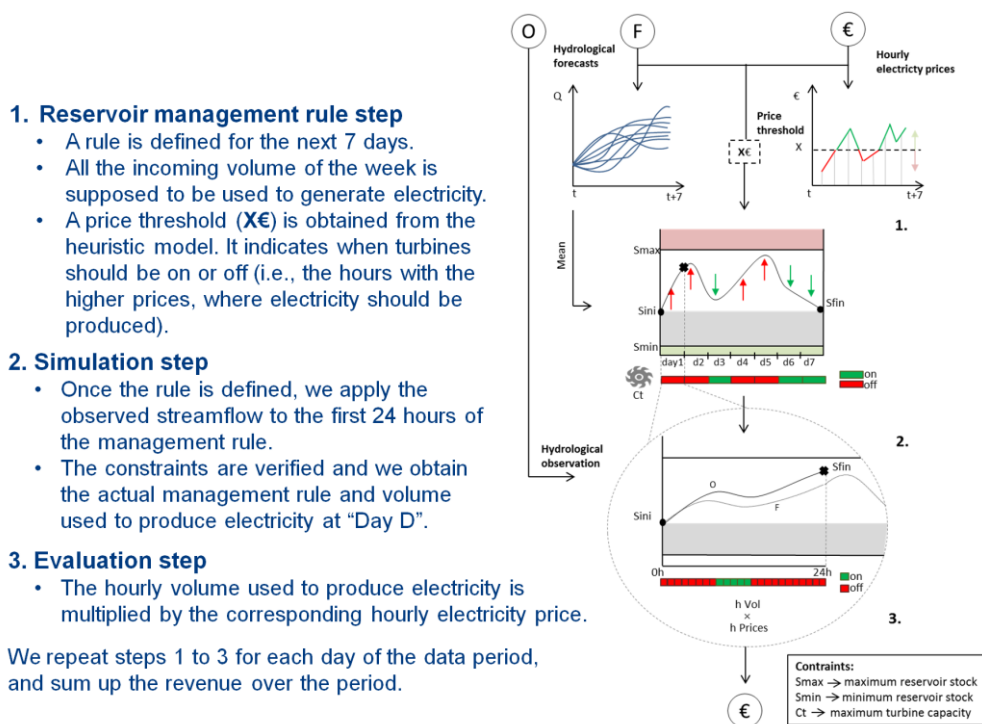


Figure 30 : Schéma de la procédure développée pour l'optimisation heuristique d'une réserve hydroélectrique (dans : poster présenté à l'EGU en 2017, M. Cassagnole et al.).

Les résultats de la thèse d'I. Zalachori ont montré que les prévisions d'ensemble apportent un gain économique supplémentaire par rapport à une prévision naïve (historique des débits passés) ou une prévision uni-scénario (déterministe ou moyenne d'ensemble) (**Figure 31**). A ma connaissance, c'était la première fois qu'une telle étude avait été menée sur une longue série de prévisions non-synthétiques (2005 à 2008)⁵¹.

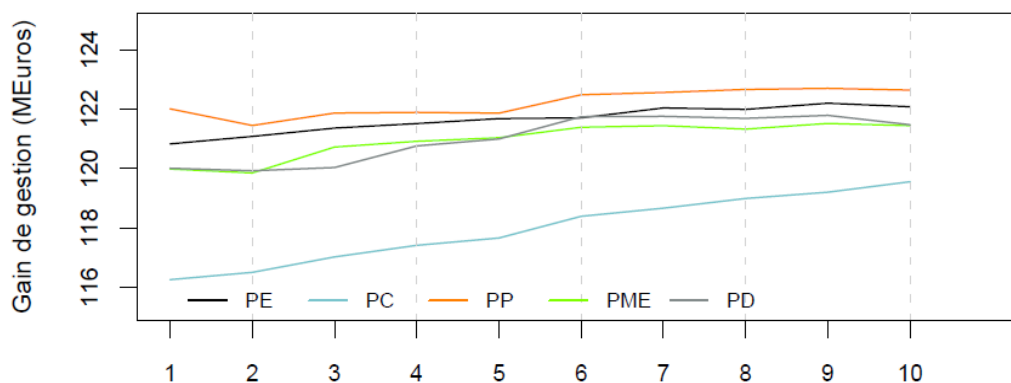


Figure 31 : Gain de gestion en fonction du coefficient de capacité du réservoir hydroélectrique (axe des abscisses, exprimé en facteur multiplicatif du débit moyen journalier) pour des prévisions hydrologiques d'ensemble brutes (PE), climatologiques (basées sur l'historique des débits) (PC), parfaites (basées sur l'observation des débits) (PP), la moyenne d'ensemble (PME) et la prévision déterministe (PD). Résultats basés sur les apports en débits au bassin versant de la Durance à Serre Ponçon et une période de prévision journalière allant de 2005 à 2008. Le coefficient de turbinage du modèle de gestion est fixé à 4 fois le débit moyen journalier (dans : Zalachori, 2013)

Dans la thèse d'I. Zalachori, nous avons aussi observé que les gains économiques de gestion reflètent, en général, la qualité de la prévision telle qu'elle peut être évaluée par un critère numérique de qualité : les prévisions brutes (de moindre qualité et sans aucun traitement statistique de correction de biais) ont tendance à fournir des gains économiques moins élevés que les prévisions post-traitées. Nous avons constaté que les approches statistiques de post-traitement qui arrivent à améliorer la qualité globale des prévisions, mais aussi la fiabilité et la discrimination, conduisent à un gain économique plus important. Néanmoins, nous n'avons pas pu dégager une « meilleure » méthode à appliquer pour avoir le meilleur gain partout. Le lien « qualité-valeur » n'était pas généralisable sur l'ensemble des bassins versants étudiés, ni valable pour l'ensemble des critères d'évaluation de la qualité considérés.

L'idée est alors venue de regarder la sensibilité du modèle de gestion heuristique à la qualité des prévisions utilisées en entrée, et ce sur un échantillon synthétique de prévisions hydrologiques pour lesquelles on pourrait contrôler la qualité des prévisions dans ces principaux aspects (fiabilité, dispersion, sur- et sous-estimation). Ceci a été le sujet du travail de **Master de M. Cassagnole (2016)** et d'une partie de sa première année de thèse (en cours).

⁵¹ Boucher *et al.* (2011) avaient utilisé un outil d'aide à la décision (SOHO), basé sur la programmation stochastique, sur un unique bassin versant au Québec et quatre mois de validation. Les auteurs ont conclu que les prévisions d'ensemble brutes (non post-traitées) conduisent à une production d'énergie moins importante (à cause de déversements plus élevés) par rapport aux prévisions déterministes. Ils ont également montré que les prévisions post-traitées sont plus performantes que les prévisions brutes et les prévisions déterministes.

Dans son travail, des prévisions synthétiques ont été construites à partir de perturbations faites aux séries de débits observés, afin de créer des prévisions qui s'éloignent de l'observation, en sur- ou sous-estimation, c'est-à-dire, qui ont une qualité variable. La qualité de ces prévisions est ainsi dégradée selon des critères numériques (l'exemple du diagramme de PIT est donné dans la **Figure 32**). Dans cette démarche, les prévisions fiables (« no bias ») sont d'abord créées à l'aide d'une méthode que j'ai co-développée avec B. Renard (Irstea Lyon) lors du projet FAST France-Australie (2010-2012). Cette méthode avait déjà été utilisée dans la thèse d'A. Randrianasolo que j'avais dirigée en 2009-2012 et y est expliquée en détail.

De manière générale, il s'agit de choisir aléatoirement dans une distribution uniforme la position de l'observation dans la distribution prédictive de l'ensemble à chaque jour J de prévision. Ensuite, à l'aide d'une loi log-normale, dont les paramètres sont définis en fonction de la valeur de l'observation et de sa position, nous tirons 50 membres aléatoirement, qui vont constituer la prévision d'ensemble du jour J . La procédure est répétée pour chaque jour de prévision.

Dans son travail, M. Cassagnole a dû rajouter une étape essentielle. Puisque la procédure ci-dessous était répétée, de manière indépendante, pour chaque échéance d'intérêt, il a fallu ajouter une dépendance temporelle aux valeurs tirées pour chaque échéance. Pour cela, une approche ECC (« *Ensemble Copula Coupling* »), similaire à la technique du Schaake Shuffle (Clark *et al.*, 2004) a été utilisée. Le modèle de dépendance de rang est hérité d'une prévision d'ensemble qui, dans notre cas, est celle donnée par le système de prévision utilisé à EDF DTG de Grenoble, constitué d'une chaîne formée par les 50 membres de la prévision d'ensemble du CEPMMT et le modèle hydrologique MORDOR (Cassagnole, 2016).

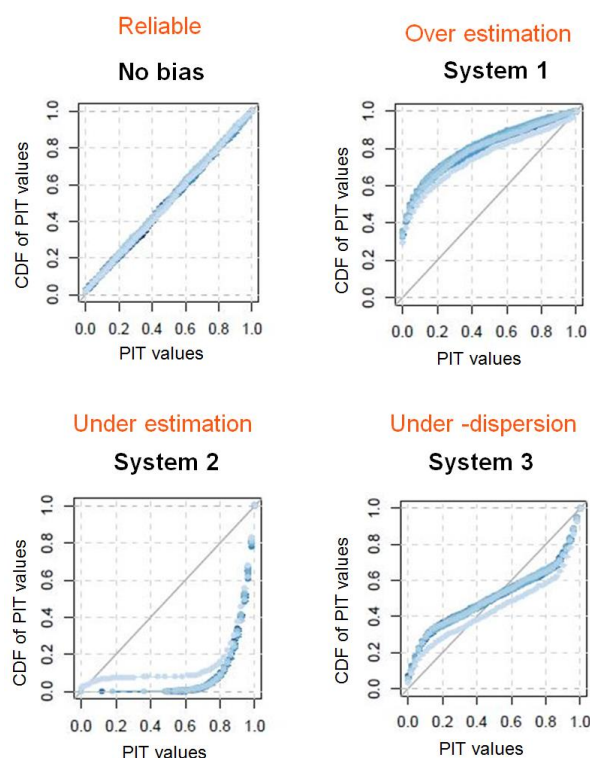


Figure 32 : Exemple de la qualité de quatre prévisions synthétiques d'ensemble de débits créées, telle qu'évaluée par le diagramme de PIT (échéance : 7 jours) (dans : Cassagnole *et al.*, poster présenté à l'EGU en 2017)

Les résultats ont permis de montrer clairement que le modèle de gestion est sensible à la qualité des prévisions hydrologiques utilisées en entrée. Nous avons observé qu'une meilleure prévision induit une meilleure gestion de la réserve hydraulique, avec un gain économique lié à la gestion plus forte au final. Cela est représenté dans la **Figure 33** en termes de pertes dans le gain économique de gestion (total sur la période évaluée, 2005-2008, et en valeur médiane sur les 10 bassins versants étudiés) par rapport à une « prévision parfaite » (où la prévision est identique aux observations utilisées pour évaluer le gain) : un système de prévision qui a tendance à surestimer les observations amène à une gestion mal programmée et à la pire performance économique vis-à-vis des autres systèmes de prévision étudiés (jusqu'à environ 3% de perte économique pour le système le plus dispersé de l'étude).

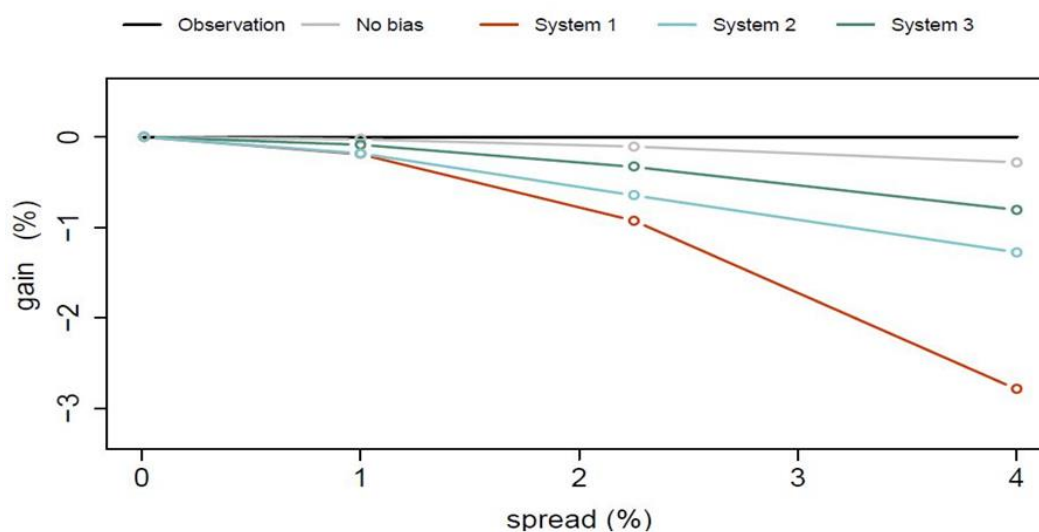


Figure 33 : Gain des ensembles synthétiques illustrés dans la Figure 32 (% par rapport au gain économique obtenu avec prévisions de débits égales aux observations) (dans : Cassagnole et al., poster présenté à l'EGU en 2017)

Ces travaux, en cours de publication, nous confortent dans l'idée que le modèle de gestion que nous avons développé dans le cadre de nos collaborations IRSTEA-EDF DTG est sensible à la qualité des apports en eau prévus. Ils ont été également essentiels pour mieux comprendre les forces et faiblesses du modèle de gestion. Aujourd'hui, la suite des travaux de thèse de M. Cassagnole concerne justement quelques limites que nous souhaitons examiner pour pousser le modèle (et les conclusions que l'on peut en tirer) plus loin.

Par exemple, il serait intéressant de savoir si la sortie « gain économique » du modèle peut être un critère en plus pour évaluer une méthode de post-traitement ou appuyer un choix de développement au sein d'un système de prévision existant. Dans un autre volet, nous pourrions nous intéresser à examiner comment la gestion se comporte face à un événement de crue extrême prévu, qui introduira des contraintes liées à des inondations en aval des réservoirs à éviter, ou encore face à des apports mensuels ou saisonniers, où des contraintes de niveau du réservoir, liées à d'autres usages de l'eau en période de basses eaux, par exemple, pourront venir jouer un rôle dans l'optimisation du revenu de la production hydroélectrique qui est faite par le modèle de gestion.

5. Perspectives

Comment utiliser notre temps et nos efforts de manière plus efficace ? Voici une question qui est au cœur du livre de Christian et Griffiths (2017), qui commence par un premier chapitre intitulé « *Optimal Stopping* », où on peut lire au dernier paragraphe : « Intuitivement, nous pensons que la prise de décision rationnelle signifie énumérer exhaustivement nos options, en pondérant soigneusement chacune d'entre elles, puis en sélectionnant la meilleure. Mais en pratique, quand l'horloge tourne – ou le cœur bat – peu d'aspects de la prise de décision (ou de la façon de penser, en général) sont aussi importants qu'un seul : savoir quand s'arrêter. »⁵².

On arrive ainsi à la fin de ce mémoire, où quelques perspectives qui s'ouvrent aujourd'hui sont brièvement présentées.

□ *Prévisions hydrologiques d'ensemble pour les crues et étiages*

Dans le contexte national, les problématiques liées aux crues et aux étiages sont au cœur des préoccupations de différents acteurs du domaine de l'eau, et ce à différentes échelles. Par exemple, les développements pilotés par le SCHAPI incorporent aujourd'hui plusieurs acquis de la recherche en prévision hydrologique développée à Irstea. L'implémentation des premiers tests en opérationnel des prévisions d'ensemble pourrait tout à fait avoir sa place dans une perspective à court terme. Il serait intéressant de construire une chaîne de prévision avec des méthodes adaptées à la prévision d'ensemble (post-traitement, évaluation de la qualité, visualisation des probabilités/scénarios) afin que cela puisse tourner en temps réel et apporter du savoir-faire aux prévisionnistes en hydrologie. Je pourrais contribuer à cet effort, ainsi qu'à de formations à l'emploi des systèmes de prévision d'ensemble dans la pratique.

Dans les années à venir, je poursuivrai également mes recherches concernant la prévision hydrologique immédiate des crues rapides avec des approches ensemblistes, faisant suite aux travaux de thèse d'A. Caseri (2017). Je souhaite d'abord focaliser mes travaux autour du contexte construit dans le cadre du projet ANR PICS, qui a démarré en janvier 2018 pour 4 ans. Ce projet nous permettra d'investiguer une approche basée sur scénarios pour aller jusqu'à la modélisation des inondations. Pour cela, on incorporera les acquis de la thèse en cours de C. Rebolho, dirigée par mon collègue V. Andréassian et qui devra être soutenue fin 2018, où une approche simplifiée basée sur les volumes de crues a été développée. Les travaux sur cette thématique seront conduits dans le cadre de la thèse de D. Peredo, que je dirige et qui a démarré en octobre 2017.

De même, il serait intéressant de poursuivre des actions pour la valorisation des prévisions météorologiques à long terme ou saisonnières en hydrologie. Cet aspect sera développé dans les années à venir dans le cadre de la thèse en cours de M. Cassagnole (2016-2019) pour la gestion des réserves en eau pour l'hydroélectricité, dans la continuité du projet H2020 Imprex (2015-2019), et dans le cadre de la thèse d'A. Assis dos Reis (co-encadrement avec l'UFMG au Brésil), également focalisée sur le secteur hydroélectrique. En lien avec ces travaux, un stage de

⁵² Dans Christian, B., Griffiths, T. (2017), *Algorithms to live by – The computer science of human decisions*, page 30 : "Intuitively, we think that rational decision-making means exhaustively enumerating our options, weighting each one carefully, and then selecting the best. But in practice, when the clock – or the ticker – is ticking, few aspects of decision-making (or of thinking more generally) are as important as one: when to stop".

Master démarrera prochainement, afin de reprendre les travaux menés par L. Crochemore (2016) lors de sa thèse et l'étendre à d'autres bassins versants, avec des régimes hydrologiques plus variés que ceux étudiés jusqu'à présent.

Les acquis en prévision saisonnière pourront également alimenter l'outil de prévision opérationnel PREMHYCE (Prévision des Étiages par des Modèles Hydrologiques, Comparaison et Évaluation), développé par mes collègues d'équipe P. Nicolle et C. Perrin dans le cadre d'un projet lancé en 2010 par l'ONEMA (devenu AFB) et la Direction de l'Eau et de la Biodiversité du Ministère de l'Environnement, de l'Énergie et de la Mer (devenu Ministère de la transition écologique et solidaire). Il serait intéressant d'explorer comment l'information météorologique actuelle peut améliorer la performance des prévisions hydrologiques basées sur un historique d'observations météorologiques passées et un ensemble multi-modèles hydrologique. Ce sujet pourrait, par exemple, faire l'objet d'une thèse en collaboration avec Météo-France et l'AFB. Il serait intéressant de pouvoir tester l'efficacité d'une telle approche utilisée en couplage à d'autres outils de modélisation de systèmes (ex., qualité des eaux ou gestion d'ouvrages).

A la fin de son article, Raftery (2016) note que : « (...) *once probabilistic forecasts become available in a domain, they will be used: "Build it and they will come."* »⁵³. Bien évidemment, il n'y a aucune garantie que le monde opérationnel sera rempli de systèmes de prévision hydrologique d'ensemble efficaces et utilisés par tous demain, si on se dépêche de les mettre en place aujourd'hui, avec les connaissances et techniques dont nous disposons. Néanmoins, j'aime bien penser qu'un certain état d'esprit peut y contribuer : si nous agissons, nous augmenterons la probabilité que cela se produise ; si nous y consacrons du temps et des efforts, nous avons des chances de réussir.

□ *Services climatiques et hydrologiques pour la prise de décision*

A travers le projet européen AQUACLEW, qui a démarré en octobre 2017 pour trois ans, je m'intéresse à la poursuite de mes recherches sur les performances des systèmes « eau-énergie », cette fois-ci en intégrant des indicateurs de services climatiques. Il s'agit ainsi de cibler la longue échéance et les projections hydrologiques en climat futur pour la prise de décision. Ce sujet sera développé en lien avec la thèse d'A. Lemoine, que je dirige et qui a démarré en novembre 2017. La thèse s'intéresse à l'amélioration des performances des systèmes eau-énergie à travers l'intégration efficace d'indicateurs issus des services climatiques à large échelle dans des modèles locaux d'impact pour le secteur hydroélectrique. Il s'agit de développer des méthodologies d'intégration de l'information climatique et hydrologique dans les modèles et schémas informatifs pour la prise de décision concernant la production hydroélectrique. Nous nous intéresserons ici au cas de la gestion stratégique (à l'horizon supérieur à 1 an) des réservoirs de production hydroélectrique et de leur intégration dans un système électrique multi-sources.

Les travaux seront conduits en étroite collaboration avec la thèse de P. Royer-Gaspard, également développée dans le cadre du projet AQUACLEW, dirigée par V. Andréassian, qui abordera plus particulièrement les défis de la modélisation hydrologique en contexte non-

⁵³ Il semblerait que la phrase, un cliché déformé de l'original, vienne du film « *Field of Dreams* » (1989), où le personnage principal entend une voix en marchant dans son champ de maïs qui lui dit « si tu le construis, il viendra » (« *If you build it, he will come* ») et, plus tard, il voit une vision d'un terrain de baseball. « Il » semblerait être Shoeless Joe Jackson, un joueur de baseball professionnel mort depuis longtemps (mais je n'ai pas vu le film).

stationnaire. De plus, les travaux bénéficieront des connaissances acquises lors du projet FP7 COMPLEX (2012-2016). Dans ce projet, ma participation a notamment consisté à contribuer au WP2 (« *Modelling Climate Related Energies* »), coordonné par J.-D. Creutin, IGE/LTHE (Grenoble, France). Il s'agissait, principalement, de développer une méthodologie permettant de mieux comprendre les interdépendances spatio-temporelles entre les différentes sources d'énergie renouvelable portées par le climat (éolienne, solaire et hydroélectrique), afin de pouvoir élaborer des options de politiques d'atténuation basées sur des connaissances acquises. Plus particulièrement, j'ai contribué à quatre produits de ce projet : (1) l'état de l'art de la modélisation des énergies renouvelables en lien avec les échelles typiques du climat et de l'hydrologie (Engeland *et al.*, 2017), (2) l'appel à ce que l'hydrologie se fasse plus présente dans le domaine, notamment via la capacité de l'hydroélectricité à stocker de l'eau (et donc de l'énergie) dans le temps et dans l'espace pour mieux gérer les sources d'énergie intermittentes (François *et al.*, 2014), (3) une réflexion sur le rôle des parties prenantes et leurs engagements dans la recherche dans ce domaine (Creutin *et al.*, 2016b), et (4) la rédaction d'une « Policy-Brief », en collaboration avec des industriels du domaine (EDF, CNR, Statkraft, Sintef) (Creutin *et al.*, 2016a). Ces connaissances renforcent et contribuent à l'ensemble de mes recherches sur les liens « eau-énergie ».

□ *Vers des outils cohérents de gouvernance des extrêmes hydrologiques de crue et sécheresse*

En 2013, j'ai pris la coordination au niveau Irstea du projet DROP (*Benefit of governance in DRought adaPtation*, 2013-2015), un projet Interreg IVB NWE. Mené en collaboration entre six agences régionales de l'eau et cinq institutions scientifiques européennes, DROP a eu pour but de mieux préparer les régions du Nord-Ouest de l'Europe à faire face aux périodes de sécheresse et de manque d'eau. Le projet a fait appel à des connaissances scientifiques, politiques et pratiques. Entre autres, il a permis de concevoir un outil d'évaluation de la gouvernance qui a été appliqué à des sites pilotes dans les six régions participantes. Cette évaluation a permis d'identifier les modalités d'action de gouvernance locale qui soutiennent ou desservent les mesures de lutte contre la sécheresse et la pénurie d'eau. En France, nous avons conduit des études sur le site pilote du barrage d'Arzal, géré par l'EPTB IAV.

Au-delà de la participation d'Irstea avec la thèse de L. Crochemore (2016), nous avons contribué à la mise en place de l'outil de gouvernance, en étroite collaboration avec des géographes et sociologues français et européens (Bressers *et al.*, 2015 ; La Jeunesse *et al.*, 2015, 2016). Les résultats obtenus ont été encourageants et nous avons contribué au montage d'un projet ANR en 2016-2017, malheureusement non retenu, dans le but d'étendre l'analyse de la gouvernance à d'autres sites en France et aussi au cas de la « co-gouvernance crues/sécheresses ». Il serait intéressant d'analyser les risques hydrométéorologiques dans quelques sites pilotes et les confronter à la cohérence (ou à l'absence d'une cohérence) des outils de gouvernance locale/régionale de la gestion des impacts des extrêmes de crue et sécheresse.

□ *Vers un meilleur bilan énergétique des modèles hydrologiques*

En 2017, j'ai initié un travail de recherche avec une étudiante en Master (Daniela Peredo, UPMC), qui s'est intéressé à l'évaluation des bilans en eau et énergie de surface pour mieux quantifier la part des précipitations qui contribuera aux écoulements de surface et des rivières. Ce travail a été mené en collaboration avec des collègues à Irstea Antony, V. Andréassian et C.

Perrin (experts en modélisation hydrologique de l'équipe HYDRO) et A. Guérin (expert en micro-météorologie de l'équipe ORACLE), et aussi avec F. Anctil de l'Université Laval à Québec.

Le but du travail a été d'explorer une formulation d'évapotranspiration basée sur le modèle de production maximale d'entropie (MEP), qui assure la conservation de l'énergie à l'interface atmosphère-surface terrestre, dans un modèle hydrologique conceptuel à l'échelle du bassin versant. Le modèle propose une solution analytique du taux d'évaporation (flux de chaleur latente), du flux de chaleur sensible et celui du sol, à partir de la température de la surface, l'humidité de la surface et le rayonnement net. Il a été développé aux Etats-Unis et repris par l'équipe de F. Anctil.

J'avais déjà commencé à m'intéresser à l'utilisation de ce modèle dans la modélisation hydrologique des modèles GR développés à Irstea en 2013, lors de mon séjour de six mois à l'Université Laval. A cette époque, j'avais fait quelques tentatives avec A. Morais, professionnelle de recherche à l'université, et nous avons conclu au côté prometteur de la méthode, avec dans un premier temps la nécessité de mieux valider le modèle MEP en soi. Ceci a été fait par l'équipe canadienne sur des données américaines du réseau FLUXNET. Les résultats positifs et les échanges avec F. Anctil ont déclenché le Master de D. Peredo (2017).

Le stage de Master s'est intéressé à : (1) la paramétrisation du modèle MEP sur des données du bassin expérimental de l'Orgeval, équipé avec une tour à flux et un scintillomètre dans le cadre du projet Critex - équipements innovants pour la zone critique, (2) son évaluation en termes de quantification de la chaleur latente et du flux d'évapotranspiration par rapport à des données expérimentales issues de la tour à flux et du scintillomètre, et (3) à son chaînage avec le modèle hydrologique GR4H développé à Irstea et utilisé dans des nombreuses applications de quantification de la ressource en eau de surface.

Les travaux ont montré que le modèle MEP reproduit bien la dynamique des flux évaporatoires et les ordres de grandeur mesurés, notamment après comparaison avec les données issues de la mesure intégrée par scintillométrie sur le bassin des Avenelles (sous-bassin du bassin expérimental de l'Orgeval). Le chaînage avec le modèle hydrologique d'Irstea GR4H a montré une légère amélioration de la simulation des bas débits sur la période de mai à décembre 2016, surtout lorsque le calage du modèle GR4H est réalisé sur une période de plusieurs années.

Des pistes d'amélioration du chaînage MEP-GR4H ont été soulevées et un deuxième stage de Master a été proposé (démarrage en avril 2018). D. Peredo (qui poursuit sa formation en thèse à Irstea, sous ma direction) contribuera à son encadrement. Le but est d'investiguer : (1) la validation du chaînage MEP-GR4H sur une année additionnelle de données observées, (2) la prise en compte des travaux passés autour du modèle GR-HUM (Loumagne *et al.*, 1996), qui introduisait une couche de surface dans le réservoir de production des modèles GR afin de mieux prendre en compte l'humidité de surface du bassin versant, et/ou d'un réservoir d'interception mis en place par Ficchi (2017) (thèse dirigée par C. Perrin), et (3) l'évaluation de l'impact du modèle couplé sur les projections hydrologiques, en s'appuyant sur deux scénarios très contrastés du Portail Drias (en collaboration avec G. Thirel).

Références bibliographiques

- Alfieri, L., B. Bisselink, F. Dottori, G. Naumann, A. de Roo, P. Salamon, K. Wyser, L. Feyen (2017), Global projections of river flood risk in a warmer world. *Earth's Future*, 5: 171–182.
- Anctil, F., M.H. Ramos, (2018), Verification metrics for hydrological ensemble forecasts. In: Q. Duan, F. Pappenberger, J. Thielen-del Pozo, A. Wood, H. L. Cloke, J. Schaake (Eds.), *Handbook of Hydrometeorological Ensemble Forecasting*. Springer-Verlag Berlin Heidelberg 2016, DOI: 10.1007/978-3-642-40457-3_3-1
- Arandia-Martinez, F. T. (2015), *Conception d'un système de prévisions hydrologiques d'ensemble multimodèle en contexte opérationnel*. Mémoire de Maîtrise, Université de Québec à Chicoutini, Avril 2015, 124 pp.
- Arnal, L., M.-H. Ramos, E. Coughlan, H. L. Cloke, E. Stephens, F. Wetterhall, S. J. van Andel, F. Pappenberger (2016), Willingness-to-pay for a probabilistic flood forecast: a risk-based decision-making game. *Hydrol. Earth Syst. Sci.*, 20: 3109–3128.
- Arribas, A., K. B. Robertson, K. R. Mylne (2005), Test of a Poor Man's Ensemble Prediction System for Short-Range Probability Forecasting. *Monthly Weather Review*, 133: 1825–1839.
- Atger, F. (2001), Verification of intense precipitation forecasts from single models and ensemble prediction systems. *Nonlin. Proc. Geophys.*, 8: 401–417.
- Bartholmes, J. C., J. Thielen, M.H. Ramos, S. Gentilini (2009), The European Flood Alert System EFAS – Part 2: Statistical skill assessment of probabilistic and deterministic operational forecasts. *Hydrol. Earth Syst. Sci.*, 13 (2): 141–153.
- Ben Daoud, A. (2010), *Améliorations et développements d'une méthode de prévision probabiliste des pluies par analogie. Application à la prévision hydrologique sur les grands bassins fluviaux de la Saône et de la Seine*. Université de Grenoble, Irstea Lyon, 340 pp.
- Ben Daoud, A., E. Sauquet, M. Lang, M.H. Ramos (2011), Peut-on étendre l'échéance de prévision des crues en optimisant la prévision de pluies par recherche d'analogues ? Application au bassin de la Seine à Paris. *La Houille Blanche*, 1: 37–43.
- Berthet, L. (2010), *Prévision des crues au pas de temps horaire : pour une meilleure assimilation de l'information de débit dans un modèle hydrologique*. Thèse de Doctorat, Cemagref (Antony), AgroParisTech, Paris, 603 pp.
- Berthet, L., M.H. Ramos, C. Perrin, A. Andréassian, C. Loumagne (2009), Can discharge assimilation methods be used to improve flood forecasting when few data are available? *IAHS Red Book: New Approaches to Hydrological Prediction in Data-sparse Regions*. IAHS Publ. 333: 94–100.
- Beven, K., A. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6 (3): 279–298.
- Beven, K., A. Binley (2014), GLUE: 20 years on. *Hydrological Processes*, 28 (24): 5897–5918.
- Bogner, K., F. Pappenberger (2011), Multiscale error analysis, correction, and predictive uncertainty estimation in a flood forecasting system. *Water Resour. Res.*, 47 (7), W07524.
- Boucher, M.-A., D. Tremblay, L. Delorme, L. Perreault, F. Anctil (2012), Hydroeconomic assessment of hydrological forecasting systems. *Journal of Hydrology*, 416–417: 133–144.
- Boucher, M.-A., F. Anctil, L. Perreault, D. Tremblay (2011), A comparison between ensemble and deterministic hydrological forecasts in an operational context. *Advances in Geosciences*, 29: 85–94.
- Bougeault, P., Z. Toth, C. Bishop, B. Brown, C. Burridge, D.H. Chen, B. Ebert, M. Fuentes, T. M. Hamill, K. Mylne, J. Nicolau, T. Paccagnella, Y.-Y. Park, D. Parsons, B. Raoult, D. Schuster, P. Silva Dias, R. Swinbank, Y. Takeuchi, W. Tennant, L. Wilson, S. Worley (2010), The THORPEX interactive grand global ensemble. *Bull. Am. Meteorol. Soc.*, 91: 1059–1072.
- Bourgin, F., M.H. Ramos, G. Thirel, A. Andréassian (2014), Investigating the interactions between data assimilation and post-processing in hydrological ensemble forecasting. *Journal of Hydrology*, 519, Part D: 2775–2784.
- Bradley, A. A., J. Demargne, J. J. Franz, (2016), Attributes of Forecast Quality. In: Duan, Q., Pappenberger, F., Thielen, J., Wood, A., Cloke, H., Schaake, J. (Eds.) *Handbook of Hydrometeorological Ensemble Forecasting*, Springer-Verlag Berlin Heidelberg 2016, 46p. 10.1007/978-3-642-40457-3_2-1

- Bressers, H., N. Bressers, A. Browne, C. Furusho, I. La Jeunesse, C. Larrue, G. Özerol, M.-H. Ramos, U. Stein, J. Tröltzsch, R. Vidaurre (2015), *Benefit of Governance in Drought Adaptation – Governance Assessment Guide*. DROP Interreg IVB Project, The Netherlands, March 2015, 24 pp.
- Brown J. D., D. J. Seo (2010), A nonparametric post-processor for bias-correction of hydrometeorological and hydrologic ensemble forecasts. *Journal of Hydrometeorology*, 11(3): 642–665.
- Brown, J. D., J. Demargne, D. J. Seo, Y. Liu (2010), The Ensemble Verification System (EVS): a software tool for verifying ensemble forecasts of hydrometeorological and hydrologic variables at discrete locations. *Environ. Modell. Softw.*, 25(7): 854–872.
- Brown, J. D., M. He, S. Regonda, L. Wu, H. Lee, D. J. Seo (2014), Verification of temperature, precipitation, and streamflow forecasts from the NOAA/NWS Hydrologic Ensemble Forecast Service (HEFS): 2. Streamflow verification. *Journal of Hydrology*, 519: 2847–2868.
- Buizza, R., H. Asensio, G. Balint, J. Bartholmes, J. Bliedernicht, K. Bogner, F. Chavaux, A. de Roo, J. Donnadille, V. Ducrocq, C. Edlund, V. Kotroni, P. Krahe, M. Kunz, K. Lacire, *et al.* (2007), *EURORISK/PREVIEW report on the technical quality, functional quality and forecast value of meteorological and hydrological forecasts*. ECMWF Tech. Memorandum 516, Reading, UK, February 2007.
- Carsell, K. M., N. D. Pingel, D. T. Ford (2004), Quantifying the benefit of a flood warning system. *Natural Hazards Review*, 5: 131–140.
- Casati, B., L. J. Wilson, D. B. Stephenson (2008), Forecast verification: current status and future directions, *Meteorological Applications*, 15 (1): 3–18.
- Casari, A. (2017), *Apport de la simulation conditionnelle géostatistique pour la prévision immédiate d'ensemble de pluies et l'alerte aux crues rapides*. Thèse de doctorat, Irstea (Aix-en-Provence & Antony), AgroParisTech (Paris), 201 pp.
- Casari, A., M.-H. Ramos, P. Javelle, E. Leblois (2016a), A space-time geostatistical approach for ensemble rainfall nowcasting. *Proceed. FLOODrisk 2016 – 3rd European Conference on Flood Risk Management*, E3S Web of Conferences, 7 18001 (2016), 5 pp.
- Casari, A., P. Javelle, M.-H. Ramos, E. Leblois (2016b), Generating precipitation ensembles for flood alert and risk management, *Journal of Flood Risk Management*, 9 (4): 402–415.
- Cassagnole, M. (2016), *Etude du lien entre la qualité des prévisions hydrologiques et leur valeur économique : cas du secteur hydroélectrique*. Mémoire de Master 2 SDUEE, UPMC, Irstea, Antony, 51 pp.
- Christian, B., T. Griffiths (2017), *Algorithms to Live by: The Computer Science of Human Decisions*. William Collins Edition, London, 361 pp. ISBN 978-0-00-754799-9
- Clark, M., S. Gangopadhyay, L. Hay, B. Rajagopalan, R. Wilby (2004), The Schaake shuffle: A method for reconstructing space-time variability in forecasted precipitation and temperature fields. *Journal of Hydrometeorology*, 5(1): 243–262.
- Cloke, H.L., F. Pappenberger (2009), Ensemble flood forecasting: a review. *Journal of Hydrology*, 375: 613–626.
- Coccia, G., E. Todini (2011), Recent developments in predictive uncertainty assessment based on the model conditional processor approach. *Hydrol. Earth Syst. Sci.*, 15 (10): 3253–3274.
- Creutin, J.-D., M.-H. Ramos, B. Hingray, B. François, M. Borga, L. Tøfte L. (2016a), *Science-Policy Brief: a climate variability viewpoint on renewable energy and electricity systems*. Policy brief FP7 COMPLEX project, 4 pp.
- Creutin, J.-D., R. Hewitt, L. Tøft, M.-H. Ramos, M. Borga (2016b), Energy companies as stakeholders. In Winder, N., H. Liljenström and R. Seaton (Eds) *The Quest for a Model-Stakeholder Fusion*. COMPLEX Final Scientific Report, Vol. 1, 220 p. Uppsala, Sweden. ISBN: 978-91-976048-1-9
- Crochemore, L. (2016), *Seasonal streamflow forecasting for reservoir management*. Thèse de doctorat, Irstea (Antony), AgroParisTech (Paris), 213 pp.
- Crochemore, L., M.-H. Ramos, F. Pappenberger (2016b), Bias correcting precipitation forecasts to improve the skill of seasonal streamflow forecasts. *Hydrol. Earth Syst. Sci.*, 20: 3601–3618.
- Crochemore, L., M.-H. Ramos, F. Pappenberger, C. Perrin (2017), Seasonal streamflow forecasting by conditioning climatology with precipitation indices. *Hydrol. Earth Syst. Sci.*, 21: 1573–1591.
- Crochemore, L., M.-H. Ramos, F. Pappenberger, S. J. van Andel, A. Wood (2016a), An experiment on risk-based decision-making in water management using monthly probabilistic forecasts. *Bull. Amer. Meteor. Soc.*, 97:4, 541–551.

- Day, G. N. (1985), Extended Streamflow Forecasting Using NWSRFS. *Journal of Water Resources Planning and Management*, 111(2): 157–170.
- Delrieu, G., V. Ducrocq, E. Gaume, J. Nicol, O. Payrastre, E. Yates, P. E. Kirstetter, H. Andrieu, P. A. Ayrat, C. Bouvier, J.-D. Creutin, M. Livet, S. Anquetin, M. Lang, L. Neppel, C. Obled, J. Parent-du-Chatelet, G.-M. Saulnier, A. Walpersdorf, W. Wobrock (2005), The catastrophic flash-flood event of 8-9 September 2002 in the Gard region, France: A first case study for the Cevennes-Vivarais Mediterranean Hydrometeorological Observatory. *Journal of Hydrometeorology*, 6: 34–52.
- Demargne, J., M. Mullusky, K. Werner, T. Adams, S. Lindsey, N. Schwein, W. Marosi, E. Welles (2009), Application of forecast verification science to operational river forecasting in the US national weather service. *Bull. Amer. Meteor. Soc.*, 90: 779–784.
- Demeritt, D., H. Cloke, F. Pappenberger, J. Thielen, J. Bartholmes, M.H. Ramos (2007), Ensemble predictions and perceptions of risk, uncertainty, and error in flood forecasting. *Environmental Hazards*, 7 : 115-127.
- Desaint, B., P. Nogues, C. Perret, R. Garçon (2009), La prévision hydrométéorologique opérationnelle : l'expérience d'Electricité de France. *La Houille Blanche*, 5: 39–46.
- Dhouioui, H. (2011), *Construction d'un faisceau de prévisions hydrométéorologiques à partir d'une prévision à scénario-unique de précipitation*. Mémoire de Master 2, Cemagref Antony/UPMC, 64p.
- Di Baldassarre, G., A. Viglione, G. Carr, L. Kuil, J. L. Salinas, G. Blöschl (2013), Socio-hydrology: Conceptualising human-flood interactions. *Hydrol. Earth Syst. Sci.*, 17(8), 3295–3303.
- Ebert, E. E. (2001), Ability of a Poor Man's Ensemble to Predict the Probability and Distribution of Precipitation. *Monthly Weather Review*, 129: 2461–2480.
- Engeland, K., B. Renard, I. Steinsland, S. Kolberg (2010), Evaluation of statistical models for forecast errors from the HBV model. *Journal of Hydrology*, 384: 142–155.
- Engeland, K., I. Steinsland (2014), Probabilistic postprocessing models for flow forecasts for a system of catchments and several lead times. *Water Resour. Res.*, 50: 182–197.
- Engeland, K., M. Borga, J.-D. Creutin, B. François, M.H. Ramos, J.-P. Vidal (2017), Space-time variability of climate variables and intermittent renewable electricity production - a review. *Renewable & Sustainable Energy Reviews*, 79, 600–617.
- Fan, F. M., D. Schwanenber, W. Collischonn, A. Weerts (2015a), Verification of inflow into hydropower reservoirs using ensemble forecasts of the TIGGE database for large scale basins in Brazil. *Journal of Hydrology: Regional Studies*, 4(Part B): 196–227.
- Fan, F. M., M-H Ramos, W. Collischonn (2015b), Sobre o uso de previsões hidrológicas probabilísticas para tomada de decisão [On the use of probabilistic hydrological forecasts for decision making]. *Revista Brasileira de Recursos Hídricos*, RBRH, 20 (4): 914–926.
- Ficchi, A. (2017), *An adaptive hydrological model for multiple time-steps: Diagnostics and improvements based on fluxes consistency*. Thèse de doctorat, Irstea (Antony), GRNE (Paris), 281 pp.
- Fischhoff, B. (1995), Risk perception and communication unplugged: Twenty years of process. *Risk Analysis*, 15: 37–45.
- François, B., M. Borga, S. Anquetin, J.-D. Creutin, K. Engeland, A.-C. Fave, B. Hingray, M.-H. Ramos, D. Raymond, B. Renard, E. Sauquet, J.F. Sauterleute, J.P. Vidal, G. Warland (2014), Integrating hydropower and intermittent climate-related renewable energies: a call for hydrology. *Hydrological Processes*, 28 (21): 5465–5468.
- Franz, K. J., T.S. Hogue (2011), Evaluating uncertainty estimates in hydrologic models: borrowing measures from the forecast verification community. *Hydrol. Earth Syst. Sci.*, 15: 3367–3382.
- Garçon, R. (1999), Modèle global Pluie-Débit pour la prévision et la prédétermination des crues. *La Houille Blanche*, 7/8: 88–95.
- Gaume, E., M. Borga (2013), Etat des connaissances récentes acquises sur les crues éclair en Europe : bilan du projet de recherches européen hydrate (2006-2010). *La Houille Blanche* (2): 24–30.
- Gneiting, T., F. Balabdaoui, A. E. Raftery (2007), Probabilistic forecasts, calibration and sharpness, *J. R. Stat. Soc. B*, 69: 243–268.
- Golian, S., J. Yazdi, M. L. V. Martina, S. Sheshangosht (2014), A deterministic framework for selecting a flood forecasting and warning system at watershed scale. *Journal of Flood Risk Management*, 8 (4): 356–367.

- Hagedorn, R., R. Buizza, T. M. Hamill, M. Leutbecher, T. Palmer (2012), Comparing TIGGE multimodel forecasts with reforecast-calibrated ECMWF ensemble forecasts. *Q. J. Roy. Meteorol. Soc.*, 138: 1814–1827.
- Hamlet, A. F., D. P. Lettenmaier (1999), Columbia River Streamflow Forecasting Based on ENSO and PDO Climate Signals, *Journal of Water Resources Planning and Management*, 125 (6): 333–341.
- Hashino, T., A. A. Bradley, S. S. Schwartz (2007), Evaluation of bias-correction methods for ensemble streamflow volume forecasts. *Hydrol. Earth Syst. Sci.*, 11: 939–950.
- He, Y., F. Wetterhall, H. Bao, H. L. Cloke, Z. Li, F. Pappenberger., Y. Hu, D. Manful, Y. Huang (2010), Ensemble forecasting using TIGGE for the July–September 2008 floods in the Upper Huai catchment: a case study, *Atmos. Sci. Lett.*, 11 (2): 132–138.
- Houdant, B. (2004), *Contribution à l'amélioration de la prévision hydrométéorologique opérationnelle. Pour l'usage des probabilités dans la communication entre acteurs*. Thèse doctorat, ENGREF, EDF, Grenoble, 209 pp.
- Hurk, B. van den, L. Bouwer, C. Buontempo, R. Döscher, E. Ercin, C. Hananel, J. Hunink, E. Kjellström, B. Klein, M. Manez, F. Pappenberger, L. Pouget, M.-H. Ramos, P. Ward, A. Weerts, J. Wijngaard (2016), Improving Predictions and Management of Hydrological Extremes through Climate Services: www.imprex.eu. *Climate Services*, 1: 6–11.
- Javelle, P., C. Fouchier, P. Arnaud, J. Lavabre (2010), Flash flood warning at ungauged locations using radar rainfall and antecedent soil moisture estimations. *Journal of Hydrology*, 394 (1-2): 267-274.
- Jolliffe, I. T., D. B. Stephenson (2011), *Forecast Verification: A practitioner's guide in Atmospheric Science*, 2nd Edition. Wiley & Sons, Ltd, 292 pp. ISBN: 978-0-470-66071-3
- Kalas, M., M.H. Ramos, J. Thielen, G. Babiakova (2008), Evaluation of the medium-range European flood forecasts for the March-April 2006 flood in the Morava River. *Journal of Hydrology and Hydromechanics*, 56 (2): 116–132.
- Kang, T.-H., Y.-O. Kim, I.-P. Hong (2010), Comparison of pre- and post-processors for ensemble streamflow prediction. *Atmos. Sci. Lett.*, 11: 153–159.
- Krzysztofowicz, R. (1999), Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resour. Res.*, 35(9): 2739–2750.
- Krzysztofowicz, R. (2001), The case for probabilistic forecasting in hydrology. *Journal of Hydrology*, 249(1-4): 2–9.
- La Jeunesse, I., C. Larrue, C. Furusho, M.-H. Ramos, A. Browne, C. de Boer, R. Vidaurre, L. Crochemore, J.-P. Arrondeau, A. Penasso (2016), Chapter 6: The Governance Context of Drought Policy and Pilot Measures for the Arzal Dam and Reservoir, Vilaine Catchment, Brittany, France, p. 109-128. In: H. Bressers, N. Bressers, C. Larrue (eds.), *Governance for Drought Resilience – Land and Water Drought Management in Europe*. Springer eBook, Switzerland, 266 pp.
- La Jeunesse, I., C. Larrue, C. Furusho, M.-H. Ramos, A. Opeicle, A. Browne, C. de Boer, R. Vidaurre (2015), Gouvernance de la sécheresse : le cas du bassin versant de la Vilaine aval (Bretagne, France), *Revue SET*, 6 p.
- Laio, F., S. Tamea (2007), Verification tools for probabilistic forecasts of continuous hydrological variables. *Hydrol. Earth Syst. Sci.*, 11: 1267–1277.
- Lang, M., C. Coeur, A. Bard, B. Bacq, T. Becker, *et al.* (2013), Les inondations remarquables en France : premiers éléments issus de l'enquête EPRI 2011. *La Houille Blanche*: 37–47.
- Li, W., Q. Duan, C. Miao, A. Ye, W. Gong, Z. Di (2017), A review on statistical postprocessing methods for hydrometeorological ensemble forecasting. *WIREs Water* (4): e1246. doi: 10.1002/wat2.1246
- Liechti, K., M. Zappa, F. Fundel, U. Germann (2013), The potential of radar-based ensemble forecasts for flash-flood early warning in the southern Swiss Alps. *Hydrol. Earth Syst. Sci.*, 17: 3853–3869.
- Liguori, S., M. A. Rico-Ramirez (2012), Quantitative assessment of short-term rainfall forecasts from radar nowcasts and MM5 forecasts. *Hydrological Processes*, 26 (25): 3842–3857.
- Liu, Y., A. H. Weerts, M. Clark, H.-J. Hendricks Franssen, S. Kumar, H. Moradkhani, D.-J. Seo, D. Schwanenberg, P.J. Smith, A. I. J. M van Dijk, N. van Velzen, M. He, H. Lee, S. J. Noh, O. Rakovec, P. Restrepo (2012), Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. *Hydrol. Earth Syst. Sci.*, 16: 3863–3887.
- Liu, Y., J. D. Brown, J. Demargne, D.-J. Seo (2011), A wavelet-based approach to assessing timing errors in hydrologic predictions. *J. Hydrol.*, 397 (3-4): 210–224.

- Loumagne, C., N. Chkir, M. Normand, C. Ottlé, D. Vidal-Madjar (1996), Introduction of soil/vegetation/atmosphere continuum in a conceptual rainfall-runoff model. *Hydrological Sciences Journal*, 41(6): 889-902.
- Madadgar, S., H. Moradkhani, D. Garen (2012), Towards improved post-processing of hydrologic forecast ensembles. *Hydrological Processes*, 28, 10.1002/hyp.9562.
- Molteni, F., R. Buizza, T. Palmer, T. Petroliagis (1996), The CEPMMT Ensemble Prediction System: Methodology and validation. *Q. J. Roy. Meteor. Soc.*, 122, 73–119.
- Montanari, A., G. Grossi (2008), Estimating the uncertainty of hydrological forecasts: A statistical approach. *Water Resour. Res.*, 44, W00B08.
- Morss, R.E., J. L. Demuth, J. K. Lazo, A. Bostrom, H. Lazrus (2015), Flash Flood Risks and Warning Decisions: A Mental Models Study of Forecasters, Public Officials, and Media Broadcasters in Boulder, Colorado. *Risk Analysis*, Vol. 35 (11): 2009–2028.
- Murphy, A. H. (1993), What Is a Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting. *Weather and Forecasting*, 8: 281–293.
- Mylne, K. (2000), *Multi-model Ensembles, Poor Man's Ensembles and Super-Ensembles - a brief review*. Prepared for the WMO Workshop on Applications of Ensemble Forecasting, 16-20 October 2000, Beijing, China, 4 pp.
- Nicolle, P., A. Valéry, M.H. Ramos, C. Perrin, A. Andréassian (2012), Mieux prévoir les crues nivales : évaluation de prévisions probabilistes de débit sur des bassins versants de montagne français. *La Houille Blanche*, 2: 26–33.
- Olsson, J., G. Lindstrom (2008), Evaluation and calibration of operational hydrological ensemble forecasts in Sweden. *Journal of Hydrology*, 350: 14–24.
- Pagano, T. C., A. W. Wood, M. H. Ramos, H. L. Cloke, F. Pappenberger, M. P. Clark, M. Cranston, D. Kavetski, T. Mathevet, S. Sorooshian, J. S. Verkade (2014), Challenges of Operational River Forecasting. *Journal of Hydrometeorology*, 15 : 1692–1707.
- Pagano, T. C., F. Pappenberger, A. W. Wood, M. H. Ramos, A. Persson, B. Anderson (2016), Automation and human expertise in operational river forecasting. *WIRES Water*, 3, 5: 692–705.
- Palmer, T., J. Barkmeijer, R. Buizza, E. Klinker, D. Richardson (2002), L'avenir de la prévision d'ensemble. *La Météorologie*, 36: 22–30.
- Pappenberger, F., J. Bartholmes, J. Thielen, E. Anghel (2008a), TIGGE: Medium range multi model weather forecast ensembles in flood forecasting (a case study). *ECMWF Technical Memorandum*, 557, January 2008, 28 pp.
- Pappenberger, F., K. Scipal, R. Buizza (2008b), Hydrological aspects of meteorological verification. *Atmos. Sci. Lett.*, 9: 43–52.
- Pappenberger, F., M. H. Ramos, H. L. Cloke, F. Wetterhall, L. Alfieri, K. Bogner, A. Mueller, P. Salamon (2015), How do I know if my forecasts are better? Using benchmarks in Hydrological Ensemble Prediction. *Journal of Hydrology*, 522: 697–713.
- Pappenberger, F., T. C. Pagano, J. D. Brown, L. Alfieri, D. A. Lavers *et al.*, (2016), Hydrological Ensemble Prediction Systems around the Globe. In: Q. Duan, F. Pappenberger, J. Thielen-del Pozo, A. Wood, H. L. Cloke, J. Schaake (Eds), *Handbook of Hydrometeorological Ensemble Forecasting*. Springer-Verlag Berlin Heidelberg 2016, DOI: 10.1007/978-3-642-40457-3_47-1.
- Parker, P., M. Fordham (1996), An evaluation of flood forecasting, warning and response systems in the European Union. *Water Resources Management*, 10 (4): 279–302.
- Peredo, D. (2017), *Impact d'une meilleure prise en compte de l'évapotranspiration dans la modélisation hydrologique*. Mémoire de Master 2 SDUEE, UPMC, Irstea, Antony, 40 pp.
- Perrin, C. (2000), *Vers une amélioration d'un modèle global pluie-débit au travers d'une approche comparative*. Thèse de doctorat, Institut National Polytechnique de Grenoble - INPG, Cemagref Antony, 20/10/2000, 287 pp.
- Renard, B., Kavetski, D., Kuczera, G., Thyer, M., Franks SW. (2010), Understanding predictive uncertainty in hydrologic modelling: the challenge of identifying input and structural errors. *Water Resources Research* 46: W05521.
- Raftery, A. (2016), Use and Communication of Probabilistic Forecasts. *Stat Anal Data Min.*, 9 (6): 397–410.
- Raftery, A. E., T. Gneiting, F. Balabdaoui, M. Polakowski (2005), Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, 133 (5): 1155–1174.

- Ramos M.H., T. Mathevet, J. Thielen, F. Pappenberger (2010a), Communicating uncertainty in hydro-meteorological forecasts: mission impossible? *Meteorological Applications*, 17: 223–235.
- Ramos, M.H. (2002), *Analyse de la pluviométrie sous des systèmes nuageux convectifs – Etude de cas sur les données de la ville de Marseille et de la méthode ISIS de Météo-France*. Thèse de Doctorat, Université Joseph Fourier - UJF, LTHE, Grenoble, 27/06/2002, 169 pp.
- Ramos, M.H. (2006), *EFAS-EPS forecasts: in-depth case study analyses and statistical evaluation of summer 2005 and spring 2006 flood forecasts in the Danube catchment*. In: Thielen, J. (Ed.), *The benefit of probabilistic flood forecasting on European scale - Results of the European Flood Alert System for 2005/2006*. European Report EUR 22560 EN, European Communities, Italy, Chap. 4, pp. 36–58.
- Ramos, M.H., A. Castelletti, M. Pulido-Velazquez, D. Gustafsson (2016), Weather and climate services for hydropower management. *Proceed. Workshop Environnement & Hydroélectricité: HydroES*, SHF, 16-17 March 2016, Grenoble, France, 5 pp.
- Ramos, M.H., J. Bartholmes, J. Thielen (2007), Development of decision support products based on ensemble weather forecasts in the European Flood Alert System. *Atmos. Sci. Let.*, 8 (4): 113–119.
- Ramos, M.H., J. Thielen, A. de Roo (2009), Prévision hydrologique d'ensemble et alerte avec le système européen d'alerte aux crues (EFAS) : cas des crues du bassin du Danube en août 2005. In: Tanguy, J.-M. (Dir.), *Traité d'hydraulique environnementale - Volume 7 : applications des modèles numériques en ingénierie 1*, Chapitre 5, Ed. Hermès Lavoisier, Octobre 2009, Paris, France, 190 pp.
- Ramos, M.H., J. Thielen, A. de Roo (2010b), Ensemble Hydrological Forecasting and Alert with the European Flood Alert System (EFAS): Case of the Danube Basin Floods in August 2005. In: Tanguy, J.-M. (Ed.), *Environmental Hydraulics: Practical Applications in Engineering*, Vol. 4, Chapter 5, p.47-61. Wiley-ISTE, October 2010, Hoboken, USA, 448 pp.
- Ramos, M.H., J. Thielen, F. Pappenberger (2008), Utilisation de la prévision météorologique d'ensemble pour la prévision hydrologique opérationnelle et l'alerte aux crues. *Proceed. Colloque SHF Prévisions hydro-météorologiques*, Lyon, 18-19 November 2008: 191–199.
- Ramos, M.H., Leblois, E., Creutin, J.-D. (2006) From point to areal rainfall: linking the different approaches for the frequency characterization of rainfalls in urban areas. *Water Science and Technology*, 54 (6-7): 33-40.
- Ramos, M.H., S. J. van Andel, F. Pappenberger (2013), Do probabilistic forecasts lead to better decisions? *Hydrol. Earth Syst. Sci.*, 17: 2219–2232.
- Randrianasolo, A. (2012), *Généralisation de l'approche d'ensemble à la prévision hydrologique dans les bassins non jaugés*. Thèse de Doctorat, Irstea (Antony), AgroParisTech (Paris), 314 pp.
- Randrianasolo, A., G. Thirel, M.-H. Ramos, E. Martin (2014), Impact of streamflow data assimilation and length of the verification period on the quality of short-term ensemble hydrologic forecasts, *Journal of Hydrology*, 519, Part D: 2676–2691.
- Randrianasolo, A., M.H. Ramos, G. Thirel, V. Andréassian, E. Martin (2010), Comparing the scores of hydrological ensemble forecasts issued by two different hydrological models. *Atmos. Sci. Let.*, 11(2): 100–107.
- Randrianasolo, A., M.H. Ramos, V. Andréassian (2011), Hydrological ensemble forecasting at ungauged basins: using neighbour catchments for model setup and updating. *Advances in Geosciences*, 29: 1–11.
- Reggiani, P., A. H. Weerts (2008), A Bayesian approach to decision-making under uncertainty: An application to real-time forecasting in the river Rhine. *Journal of Hydrology*, 356 (1-2): 56–69.
- Reggiani, P., M. Renner, A. Weerts, P. van Gelder (2009), Uncertainty assessment via bayesian revision of ensemble streamflow predictions in the operational river Rhine forecasting system. *Water Resour. Res.*, 45, W02428.
- Richardson, D. S. (2000), Skill and relative economic value of the ECMWF ensemble prediction system. *Q. J. R. Meteorolog. Soc.*, 126: 649–667.
- Roche, P.-A., J. Miquel, E. Gaume (2012), *Hydrologie Quantitative : Processus, modèles et aide à la décision*. Springer, 620 pp.
- Roulin, E. (2007), Skill and relative economic value of medium-range hydrological ensemble predictions. *Hydrol. Earth Syst. Sci.*, 11: 725–737.
- Roulin, E., S. Vannitsem (2015), Post-processing of medium-range probabilistic hydrological forecasting: impact of forcing, initial conditions and model errors. *Hydrological Processes*, 29: 1434–1449.

- Rousset Regimbeau, F., F. Habets, E. Martin, J. Noilhan (2007), Ensemble streamflow forecasts over France. *ECMWF Newsletter* 111: 21–27.
- Ruin, I., C. Lutoff, B. Boudevillain, J.-D. Creutin, S. Anquetin, M. Bertran Rojo, L. Boissier, L. Bonnifait, M. Borga, L. Colbeau-Justin, L. Creton-Cazanave, G. Delrieu, J. Douvinet, E. Gaume, E. Grunfest, J.-P. Naulin, O. Payrastre, O. Vannier (2014), Social and Hydrological Responses to Extreme Precipitations: An Interdisciplinary Strategy for Postflood Investigation. *Weather, Climate & Society*, 6(1): 135–153.
- Schaake, J.C., T. H. Hamill, R. Buizza, M. Clark (2007), HEPEX – the hydrological ensemble prediction experiment. *Bulletin of the American Meteorological Society*, 88(10): 1541–1547.
- Schaefli, B., D. B. Talamba, A. Musy (2007), Quantifying hydrological modeling errors through a mixture of normal distributions. *Journal of Hydrology*, 332: 303–315.
- Seiller, G., F. Anctil, F., C. Perrin (2012), Multimodel evaluation of twenty lumped hydrological models under contrasted climate conditions. *Hydrol. Earth Syst. Sci.*, 16: 1171–1189.
- Stewart, T. R. (1997), Forecast value: descriptive decision studies. In: Katz, R.W. and Murphy, A. H. (Eds), *Economic value of weather and climate forecasts*, Chapter 5, Cambridge University Press: 147–181.
- Swinbank, R., M. Kyouda, P. Buchanan, L. Froude, T. M. Hamill, T. D. Hewson, J. H. Keller, M. Matsueda, J. Methven, F. Pappenberger, M. Scheuerer, H. Tittley, L. Wilson, M. Yamaguchi (2016), The TIGGE Project and Its Achievements. *Bull. Amer. Meteor. Soc.*, 97 (1): 49–67.
- Tangara, M. (2005), *Nouvelle méthode de prévision de crue utilisant un modèle pluie-débit global*. Thèse de doctorat, EPHE, Cemagref, Paris, 374 pp.
- Thiboult, A. (2015), *Toward coherent accounting of uncertainty in hydrometeorological modeling*. Thèse de doctorat, Université Laval, Québec, Canada, 180 pp.
- Thiboult, A., F. Anctil, M.H. Ramos (2017), How does the quantification of uncertainties affect the quality and value of flood early warning systems? *Journal of Hydrology*, 551: 365–373.
- Thielen, J., J. Schaake, R. Hartman, R. Buizza (2008), Aims, challenges and progress of the Hydrological Ensemble Prediction Experiment (HEPEX) following the third HEPEX workshop held in Stresa 27 to 29 June 2007. *Atm. Sci. Letters*: 29–35.
- Thielen, J., M.-H. Ramos, J. Bartholmes, A. de Roo, F. Pappenberger, H. Cloke, D. Demeritt (2005), *Summary report- 1st workshop on the use of ensemble prediction system in flood forecasting (Ispra, 21-22 November 2005)*. European Report EUR 22118 EN, European Communities, Italy, 23 pp.
- Thirel, G. (2009), *Amélioration des prévisions d'ensemble des débits sur la France de SAFRAN-ISBA-MODCOU*. Thèse de doctorat, Université Toulouse III - Paul Sabatier, 252 pp.
- Thirel, G., F. Rousset-Regimbeau, E. Martin, F. Habets (2008), On the impact of short-range meteorological forecasts for ensemble stream flow predictions. *Journal of Hydrometeorology*, 9(6): 1301–1317.
- Toth, Z., E. Kalnay (1997), Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, 125: 3297–3319.
- Trambauer, P., M. Werner, H. C. Winsemius, S. Maskey, E. Dutra, S. Uhlenbrook (2015), Hydrological drought forecasting and skill assessment for the Limpopo River basin, southern Africa. *Hydrol. Earth Syst. Sci.*, 19: 1695–1711.
- van Andel, S.J., A. Weerts, J. Schaake, K. Bogner (2012), Post-processing hydrological ensemble predictions intercomparison experiment. *Hydrological Processes*, 27: 158–161.
- van Pham, T. (2011), *Tracking the uncertainty in streamflow prediction through a hydrological forecasting system*. Master Civil Engineering & Management, Cemagref Antony/University of Twente, The Netherlands, 74 p.
- Velázquez, J. A., F. Anctil, M. H. Ramos, C. Perrin (2011), Can a multi-model approach improve hydrological ensemble forecasting? A study on 29 French catchments using 16 hydrological model structures. *Advances in Geosciences*, 29: 33–42.
- Verkade, J. S. (2015), *Estimating real-time predictive hydrological uncertainty*. Doctoral thesis, TU Delft, The Netherlands, 193 pp.
- Verkade, J. S., J. D. Brown, F. David, P. Reggiani, A. H. Weerts (2017), Estimating predictive hydrological uncertainty by dressing deterministic and ensemble forecasts; a comparison, with application to Meuse and Rhine. *Journal of Hydrology*, 555: 257–277.

- Verkade, J. S., J. D. Brown, P. Reggiani, A. H. Weerts (2013), Post-processing ECMWF precipitation and temperature ensemble reforecasts for operational hydrologic forecasting at various spatial scales. *Journal of Hydrology*, 501: 73–91.
- Verkade, J. S., M. G. F. Werner (2011), Estimating the benefits of single value and probability forecasting for flood warning. *Hydrol. Earth Syst. Sci.*, 15: 3751–3765.
- Vincendon, B., V. Ducrocq, O. Nuissier, B. Vié (2011), Perturbation of convection-permitting NWP forecasts for flash-flood ensemble forecasting. *Natural Hazards Earth Syst. Sci.*, 11: 1529–1544.
- Voisin, N., F. Pappenberger, D. P. Lettenmaier, R. Buizza, J. C. Schaake (2011), Application of a Medium-Range Global Hydrologic Probabilistic Forecast Scheme to the Ohio River Basin. *Weather and Forecasting*, 26: 425–446.
- Weeink, W. (2010), *Thresholds for flood forecasting and warning – Evaluation of streamflow and ensemble thresholds*. MSc thesis, Cemagref Antony/University of Twente, Enschede, June 2010, 88 pp.
- Weijs, S.V., G. Schoups, N. van De Giesen (2010), Why hydrological predictions should be evaluated using information theory. *Hydrol. Earth Syst. Sci.*, 14: 2545–2558.
- Wilks, D. S. (1997), Forecast value: prescriptive decision studies. In: Katz, R.W. and Murphy, A. H. (Eds), *Economic value of weather and climate forecasts*, Chapter 4. Cambridge University Press: 109–145.
- Wilks, D. S. (2011), *Statistical Methods in the Atmospheric Sciences: An Introduction*, Academic Press, 3rd Edition, 704 pp. ISBN: 9780123850225
- Wu, L., D.-J. Seo, J. Demargne, J. D. Brown, S. Cong, J. Schaake (2011), Generating of ensemble precipitation forecast from single-valued quantitative precipitation for hydrological ensemble prediction. *Journal of Hydrology*, 399: 281–298.
- Wu, Z., J. Wu, G. Lu (2016), A one-way coupled atmospheric-hydrological modeling system with combination of high-resolution and ensemble precipitation forecasting. *Front. Earth Sci.*, 10 (3): 432–443.
- Yao, H., A. Georgakakos (2001), Assessment of Folsom Lake response to historical and potential future climate scenarios: 2. Reservoir management. *Journal of Hydrology*, 249: 176–196.
- Younis, J., M. H. Ramos, J. Thielen (2008), EFAS forecasts for the March-April 2006 flood in the Czech part of the Elbe River Basin - a case study. *Atmos. Sci. Let.*, 9(2): 88–94.
- Yuan, X., J. Roundy, E. Wood, J. Sheffield (2015), Seasonal forecasting of global hydrologic extremes: system development and evaluation over GEWEX basins. *Bull. Amer. Meteor. Soc.*: 1895–1912.
- Zalachori, I. (2013), *Prévisions hydrologiques d'ensemble : développements pour améliorer la qualité des prévisions et estimer leur utilité*. Thèse de Doctorat, Irstea (Antony), AgroParisTech (Paris), 398 pp.
- Zalachori, I., M. H. Ramos, R. Garçon, T. Mathevet, J. Gailhard (2012), Statistical processing of forecasts for hydrological ensemble prediction: a comparative study of different bias correction strategies. *Advances in Science and Research*, 8: 135–141.
- Zappa, M., F. Fundel, S. Jaun (2013), A 'Peak-Box' approach for supporting interpretation and verification of operational ensemble peak-flow forecasts. *Hydrological Processes*, 27 (1): 117–131.
- Zappa, M., M. W. Rotach, M. Arpagaus, M. Dorninger, C. Hegg, A. Montani, P. R. Ranzi, F. Ament, U. Germann, G. Grossi, S. Jaun, A. Rossa, S. Vogt, A. Walser, J. Wehrhan, C. Wunram, C. (2008), MAP D-PHASE: real-time demonstration of hydrological ensemble prediction systems. *Atmos. Sci. Let.*, 9: 80–87.
- Zhao, T., J. Bennett, Q.J. Wang, A. Schepen, A. Wood, D. Robertson, M.H. Ramos (2017), How suitable is quantile mapping for post-processing GCM precipitation forecasts? *Journal of Climate*, 30: 3185–3196.