Prévisions d'ensemble hydrologiques et hydrauliques pour la vigilance crues

Anne-Laure Tiberi

▶ **To cite this version:**

Anne-Laure Tiberi. Prévisions d'ensemble hydrologiques et hydrauliques pour la vigilance crues. Ingénierie de l'environnement. Université Paris-Est, 2021. Français. NNT : 2021PESC2031 . tel-03539466v1

HAL Id: tel-03539466

https://theses.hal.science/tel-03539466v1

Submitted on 21 Jan 2022 (v1), last revised 3 Feb 2022 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS-EST

École Doctorale SIE

Laboratoire d'Hydraulique Saint-Venant

Thèse

Présentée pour l'obtention du grade de DOCTEUR
DE L'UNIVERSITE PARIS-EST
par

Anne-Laure Tiberi-Wadier

Prévisions d'ensemble hydrologiques et hydrauliques pour la vigilance crues

---

Spécialité : Sciences et Techniques de l'Environnement

Soutenue le 27 septembre 2021 devant un jury composé de :

Mme Maria-Helena Ramos	INRAE	Présidente du jury
Mme Hélène Roux	IMFT	Rapporteur
M Eric Gaume	UGE	Examinateur
M Philippe Naveau	LSCE	Examinateur
Mme Nicole Goutal	EDF R&D/LHSV	Directrice de thèse
Mme Sophie Ricci	CERFACS	Co-directrice de thèse
M Philippe Sergent	CEREMA	Invité
M Bruno Janet	SCHAPI	Invité
M Rémy Garçon	EDF-DTG	Invité

Thèse effectuée au sein du **Laboratoire d'Hydraulique Saint-Venant**

de l'Université Paris-Est

6, quai Watier

BP 49

78401 Chatou Cedex

France

En collaboration avec le **CERFACS**

42 avenue Gaspard Coriolis, 31057 Toulouse Cedex 01

Résumé

Cette thèse s'inscrit dans le cadre d'un partenariat entre le CEREMA (Centre d'Études et d'expertise sur les Risques, l'Environnement, la Mobilité et l'Aménagement), EDF R&D, le CERFACS (Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique) et le SCHAPI (Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations). Afin de réaliser les cartes de vigilance bi-quotidiennes, le SCHAPI et les 19 SPC (Services de Prévision des Crues) répartis sur le territoire utilisent entre autres des résultats de modèles numériques généralement lancés de manière déterministe (prévisions météorologiques, modélisations hydrologique et hydraulique). L'objectif de la thèse est la mise en place et l'évaluation de prévisions d'ensemble hydrologiques et hydrauliques dans le cadre de la vigilance crue réalisée par les services de l'État afin de mieux appréhender et réduire les incertitudes dans un contexte de prévision à courte et moyenne échéance (24 heures). L'originalité de ce travail réside dans l'utilisation hybride de modèles à base physique et de modèles d'apprentissage sur un important volume de données. Dans cet objectif, les prévisions météorologiques forcent un modèle chaîné hydrologie-hydraulique afin de fournir des prévisions de débits et de hauteurs d'eau. Afin de prendre en compte les diverses sources d'incertitude liées aux modèles numériques, aux paramètres des modèles et aux données associées, l'approche déterministe est remplacée par une approche ensembliste. On fournit ainsi un ensemble de prévisions de débits et hauteurs d'eau.

Le bassin d'étude est le bassin versant de l'Odet situé dans le Finistère. La partie amont du bassin est modélisée par un modèle hydrologique (GRP ou MORDOR-TS). Il fournit une prévision de débit qui sert de forçage au modèle hydraulique 1D MASCARET, qui lui prévoit des hauteurs d'eau aux stations de vigilance en aval.

Dans un premier temps, une étude de sensibilité globale (GSA) est menée sur les modèles hydrologiques et hydrauliques. Ceci est un préalable à la génération des prévisions d'ensemble. La GSA permet d'identifier les sources principales d'incertitude et ainsi de perturber les paramètres incertains significatifs pour la représentation des débits et des hauteurs d'eau prévus. La propagation de ces incertitudes aboutit à la création d'un ensemble brut pour l'hydrologie et pour l'hydraulique, les ensembles hydrologiques étant utilisés pour forcer les ensembles hydrauliques. Deux méthodes de correction des ensembles sont alors investiguées dans la thèse : la calibration statistique via la méthode des forêts aléatoires « Quantile Regression Forests » et la calibration par assimilation de données via un filtre de Kalman d'ensemble (EnKF). On a montré que ces deux approches améliorent significativement les performances de l'ensemble en termes de fiabilité et réso-

lution. Enfin la comparaison des performances des prévisions d'ensemble est finalement réalisée pour l'hydrologie et l'hydraulique et des préconisations sont émises pour la génération opérationnelle de prévisions d'ensemble au sein des SPC.

Mots-clé:

Prévisions d'ensemble, modélisation hydrologique, modélisation hydraulique, GRP, MORDOR-TS, MASCARET, études de sensibilité, ANOVA, calibration statistique, forêts aléatoires, quantile regression forests, assimilation de données, filtre de Kalman d'ensemble, bassin de l'Odet.

Hydrologic and hydraulic ensemble forecasts for flood monitoring and forecasting

Abstract

This PhD originates from a cooperation between CEREMA (Centre for Studies and Expertise on Risks, Environment, Mobility and Urban Planning), EDF R&R, CERFACS (European Center for Advanced Research and Training in Computational Sciences) and SCHAPI (French National Service for Flood Forecasting). In order to realize the twice-daily vigilance maps, the SCHAPI and the 19 SPC (Services de Prévision des Crues) distributed on the territory, use, among others, the results of numerical models generally run with a deterministic approach (meteorological forecasts, hydrological and hydraulic modeling). The objective of the thesis is the implementation and evaluation of hydrological and hydraulic ensemble forecasts within the framework of the flood surveillance and forecast carried out by governmental agencies in order to identify and reduce uncertainties for short to mid-term forecasts (24 hours). The originality of this work lies in the hybrid use of physics-based models and learning models on a large volume of data. For this purpose, meteorological forecasts are used to force a hydrological-hydraulic chained model to provide discharge and water level forecasts. In order to take into account the various sources of uncertainty related to the numerical models, model parameters and associated data, the limits of the deterministic are overcome by an ensemble approach; an ensemble of flow and water level forecasts is thus generated.

The study basin is the Odet watershed located in Finistère (France, Brittany). The upstream part of the basin is represented by a hydrological model (GRP or MORDOR-TS). It provides a forecasted inflow for the 1D hydraulic model MASCARET that represents the dynamics of the river, and represents water heights at the downstream observing stations.

First, a global sensitivity study (GSA) is carried out for the hydrological and hydraulic models. This is a prerequisite for the generation of ensemble forecasts. The GSA allows to identify the main sources of uncertainty and thus to perturb the significant uncertain parameters for the generation of an ensemble of forecasted discharges and water levels. The propagation of these uncertainties results in the generation of a raw ensemble for both hydrology and hydraulics, with the hydrological ensembles used to force the hydraulic ensembles. Two methods of ensemble correction are then investigated in the PhD: statistical calibration via the Quantile Regression Forest method and data assimilation calibration via an ensemble Kalman Filter (EnKF). It was shown that both approaches significantly improve the performance of the ensemble in terms of reliability and resolution. Finally, the performance of ensemble forecasts is compared for hydrology and hydraulics and recommendations are made for the operational generation of ensemble flow forecasts within

SPCs.

Keywords:

Ensemble forecast, hydrological modeling, hydraulic modeling, GRP, MORDOR-TS, MASCARET, sensitivity analysis, ANOVA, statistical calibration, quantile regression forests, data assimilation, ensemble Kalman filter, Odet catchment.

Remerciements

Je tiens tout d'abord à remercier très chaleureusement ma directrice de thèse Nicole Goutal et ma co-directrice Sophie Ricci pour m'avoir accompagnée et guidée dans ce travail de recherche. Nicole, Sophie, j'ai énormément apprécié travailler à vos côtés dans une ambiance toujours positive, bienveillante et constructive. Votre binôme d'encadrement est parfait, aucun doctorant ne pourrait rêver mieux ! Votre complémentarité est exceptionnelle. J'ai apprécié vos grandes qualités scientifiques et humaines. Merci pour votre disponibilité et votre investissement dans mon travail de thèse pendant ces quatre années. Nos petites réunions de travail vont me manquer !

Je tiens également à remercier très chaleureusement Philippe Sergent qui a suivi ce projet de thèse pour le Cerema. Philippe, je te remercie pour avoir eu confiance en ce projet et pour tous les conseils que tu m'as donnés lors du montage de la thèse ou en cours de travail. Je te remercie également d'avoir accepté de faire partie de mon comité de suivi de thèse.

Je remercie grandement tous les membres de mon comité de suivi de thèse et plus généralement toutes les personnes qui ont guidé l'orientation de ces travaux et ont permis leur avancée. Je pense en particulier à Maxime Taillardat, Olivier Mestre, Michaël Zamo, François Bouttier, Denis Dartus, Bruno Janet, Etienne Le Pape, Anne Belleudy, Charles Perrin, Julie Viatgé, Céline Monteil, Fabrice Zaoui, Pamphile Roy, Isabelle Mirouze, Andrea Piacentini et Laurent Colonnette. Et puisqu'on n'est jamais à l'abri d'un oubli, je m'excuse par avance pour ceux que j'aurais pu oublier...

L'aboutissement de cette belle aventure se réalise grâce au jury de thèse qui évalue et valide les travaux de recherche menés. Je vous remercie sincèrement, Maria-Helena Ramos, Hélène Roux, Eric Gaume, Philippe Naveau, Bruno Janet, Rémi Garçon, pour m'avoir fait l'honneur d'évaluer mes travaux, pour votre lecture attentive du manuscrit et pour les échanges constructifs que nous avons eus pendant la soutenance. J'espère avoir l'occasion de vous recroiser au hasard des événements se déroulant autour du thème de la prévision hydrométéorologique.

Cette thèse n'aurait pas pu voir le jour sans un accueil positif du projet par le Cerema. Le Cerema m'a donné les moyens pour qu'elle se déroule dans d'excellentes conditions, en particulier grâce à une décharge des mes activités courantes pour libérer du temps pour la recherche. Je remercie la Direction Générale du Cerema, la Direction de la Recherche, de l'Innovation et de l'International (DRII) ainsi que la Direction du Cerema Risques, Eau et Mer de m'avoir permis de mener à bien ce projet. Je remercie mes supérieurs passés et actuels, Nathalie Metzler, Joël L'Her, Ronan Roué, Boris Leclerc et Philippe Joscht, pour m'avoir accordé leur confiance dans l'élaboration et la mise en œuvre de ce projet. Je souhaite également remercier le SCHAPI qui a soutenu la réalisation de ces travaux,

ainsi que le Service de Prévision des Crues Vilaine et Côtiers Bretons pour l'ensemble des échanges que nous avons eus sur la modélisation du bassin versant de l'Odet. La réussite de ce projet est étroitement liée à la participation active d'EDF R&D et du CERFACS pour l'encadrement. J'adresse mes remerciements à ces deux organismes. Merci également à Météo-France pour la mise à disposition des données hydrométéorologiques nécessaires, et à INRAE et EDF-DTG pour la mise à disposition des codes de modélisation hydrologique. Enfin, je tiens à remercier le Pôle de Calcul et de Données Marines (PCDM) pour l'accès aux ressources de calcul DATARMOR.

L'organisation pratique de la soutenance à Chatou depuis Brest n'aurait pas été aussi facile sans l'aide précieuse de Vanessya Laborie et Marissa Yates. Je vous remercie sincèrement toutes les deux pour votre aide et votre dynamisme. Je me souviendrai que vous avez fait partie des quelques spectateurs en présentiel de ma soutenance !

Pour finir, je souhaite remercier toutes celles et ceux qui m'ont soutenue au cours de ces quatre années, mes amis et ma famille. Merci à Pierre (et ses conseils python !), Pauline, Matthieu, Roman et Juliette.

Table des matières

Introduction **1**

I CONTEXTE ET MÉTHODOLOGIES DÉVELOPÉES **7**

1 Contexte de la prévision des crues et objectifs de la thèse **9**

1.1 La chaîne de prévision hydrométéorologique et incertitudes associées 11

1.2 La prévision opérationnelle des crues en France et la prise en compte des incertitudes 27

1.3 Objectifs de la thèse et méthodologie suivie 37

2 Le bassin versant de l'Odet, données et modèles **39**

2.1 Présentation générale du bassin de l'Odet 40

2.2 Les données d'observation et de prévision 47

2.3 La modélisation du bassin de l'Odet 51

3 Deux approches pour générer des prévisions d'ensemble **71**

3.1 La vérification des ensembles 73

3.2 L'étude d'incertitude : un préalable commun aux deux approches 87

3.3 La calibration statistique pour corriger les ensembles de prévision bruts . . 96

3.4 L'assimilation de données 104

II LES PRÉVISIONS D'ENSEMBLE HYDROLOGIQUES **129**

4 Études de sensibilité et calibration statistique QRF **133**

4.1 Article soumis à *La Houille Blanche* 134

4.2 Article accepté par *Journal of Hydrology* . . . 136

5 Filtre de Kalman d'ensemble (EnKF) dans le modèle MORDOR-TS 143

5.1 Structure de l'assimilation de données mise en place . . . 144

5.2 Réalisation des expériences jumelles . . . 146

5.3 Réalisation des expériences réelles . . . 170

5.4 Synthèse . . . 171

6 Comparaison des méthodes QRF et EnKF 173

6.1 Résultats comparés des ensembles *raw-HEF*, *QRF-hydro*, *QRF-hydro-TS* et *DA* en prévision . . . 175

6.2 Synthèse et perspectives . . . 181

III LES PRÉVISIONS D'ENSEMBLE HYDRAULIQUES 183

7 Étude de sensibilité du modèle MASCARET de l'Odet 187

7.1 Cadre d'étude . . . 189

7.2 Mise en œuvre des expériences . . . 193

7.3 Synthèse . . . 202

8 Calibration statistique QRF des ensembles bruts hydrauliques 205

8.1 Méthodologie . . . 207

8.2 Calibration QRF à Kervir et Moulin-Vert . . . 217

8.3 Calibration QRF à Justice . . . 236

8.4 Conclusion générale . . . 253

9 Filtre de Kalman d'ensemble (EnKF) dans le modèle MASCARET 255

9.1 Structure de l'assimilation de données mise en place . . . 256

9.2 Réalisation des expériences jumelles . . . 258

9.3 Réalisation des expériences réelles . . . 282

10 Comparaison des méthodes QRF et EnKF 291

10.1 Résultats à Kervir . . . 293

10.2 Résultats à Moulin-Vert . . . 296

10.3 Résultats à Justice 301
10.4 Synthèse 304

11 Conclusion générale 307

11.1 Synthèse des résultats obtenus 309
11.2 Perspectives en termes de recherche 313
11.3 Perspectives d'application opérationnelle 315

IV ANNEXES 317

L Compléments au chapitre 1 319

L.1 Bulletin de vigilance crues du SPC Vilaine et Côtiers Bretons du 23 décembre 2013 à 10h00 319

M Compléments au chapitre 2 325

M.1 Bulletin de Précipitations émis par Météo-France 326
M.2 Description du modèle hydrologique GRP 327
M.3 Graphiques des débits mesurés et simulés avec GRP 329
M.4 Description du modèle hydrologique MORDOR-TS 333
M.5 Graphiques des débits mesurés et simulés avec MORDOR-TS 338

N Compléments au chapitre 3 343

N.1 Méthodes d'analyse de sensibilité globale 344
N.2 Calibration statistique 348
N.3 Formulation de l'analyse avec deux estimateurs 353

O Compléments au chapitre 5 357

O.1 Expérience *corr1par* - Résultats à Tréodet 358
O.2 Expérience *corr1par* - Résultats à Kerjean 363
O.3 Expérience *corr1par* - Résultats à Ty-Planche 368
O.4 Expérience *corr8par* - Résultats à Kerjean 373
O.5 Expérience *corr8par* - Résultats à Ty-Planche 375

P Compléments au chapitre 6 377

P.1 Ensembles à Tréodet 378
P.2 Ensembles à Kerjean 380
P.3 Ensembles à Ty-Planche 385

Q Compléments au chapitre 8 **391**
Q.1 Calibration à Kervir et Moulin-Vert 392
Q.2 Calibration à Justice 400

R Compléments au chapitre 9 **407**
R.1 Expériences jumelles : correction simultanée des Ks et des conditions limites amont et aval 408
R.2 Expériences réelles 418

S Compléments au chapitre 10 **425**
S.1 Ensembles à Kervir 426
S.2 Ensembles à Moulin-Vert 432
S.3 Ensembles à Justice 439

Bibliographie **440**

Introduction

Les prévisions hydrométéorologiques diffusées au grand public et aux services régaliens de gestion de crise sont de plus en plus souvent publiées avec une estimation de leurs incertitudes. On connaît depuis longtemps maintenant la notion d'indice de confiance relatif aux prévisions météorologiques. Cette tendance se développe également depuis plusieurs années pour les prévisions hydrologiques ou hydrauliques opérationnelles. Le travail de thèse présenté dans ce manuscrit s'inscrit dans cette dynamique. Réalisé dans le cadre d'un partenariat entre le CEREMA[1], EDF R&D[2], le CERFACS[3] et le SCHAPI[4], il a pour objectifs la mise en place et l'évaluation de prévisions d'ensemble hydrologiques et hydrauliques dans le cadre de la vigilance crues-inondations réalisée par les services de l'État. Aujourd'hui, afin de réaliser les cartes de vigilance bi-quotidiennes, le SCHAPI et les 19 SPC[5] répartis sur le territoire utilisent entre autres des résultats de modèles numériques généralement lancés de manière déterministe (prévisions météorologiques, modélisations hydrologique et hydraulique). Étant donné le niveau d'incertitudes sur les trois modélisations en jeu, une prévision déterministe chaînée des trois modèles est peu fiable. Récemment, des outils de post-traitement des prévisions déterministes[6] ont été mis en place en opérationnel au SCHAPI et dans les SPC afin d'estimer l'incertitude prédictive. Cette incertitude est diffusée avec les prévisions. L'objectif de cette thèse est d'étudier plus généralement l'apport d'une prévision d'ensemble pour appréhender les incertitudes. L'objectif final est donc la mise au point de méthodologies et leur évaluation pour la génération de prévisions d'ensemble hydrologiques et hydrauliques.

Le travail est mené sur le bassin versant de l'Odet dans le Finistère. C'est un bassin versant d'environ 720 km^2 dont l'exutoire se trouve en mer. L'Odet traverse la ville de Quimper. Des débordements surviennent relativement régulièrement dans l'agglomération quimpéroise en raison de fortes pluies sur le bassin versant. Ils sont parfois amplifiés par des événements d'origine maritime qui génèrent des surcotes remontant dans l'estuaire. Sur

[1] Centre d'Études et d'expertise sur les Risques, l'Environnement, la Mobilité et l'Aménagement
[2] Électricité de France, Recherche et Développement
[3] Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique
[4] Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations
[5] Services de Prévision des Crues
[6] OTAMIN pour OuTil Automatique d'estiMation des INcertitudes prédictives. Cet outil sera présenté au chapitre 1.

ce bassin, la vigilance crues est réalisée de manière opérationnelle par le SPC Vilaine et Côtiers Bretons (SPC VCB). En opérationnel, les trois sous-bassins amont sont modélisés avec le modèle hydrologique GRP (modèle conceptuel global développé par INRAE[7]). Ces prévisions sont ensuite utilisées comme forçage en entrée d'un modèle hydraulique 1D MASCARET. Dans le cadre de cette thèse, la configuration opérationnelle est conservée et adaptée :

- Deux types de modèles hydrologiques sont utilisés sur les trois sous-bassins amont : le modèle GRP comme en opérationnel et le modèle MORDOR-TS (modèle conceptuel semi-distribué développé par EDF-DTG[8]). La mise en place de ces deux modélisations ainsi que les calages associés sont réalisés spécifiquement dans le cadre de la thèse.

- La modélisation hydraulique de la partie aval du bassin versant est réalisée à partir du modèle hydraulique MASCARET opérationnel. Dans le cadre de ce travail, la géométrie du modèle a été modifiée pour permettre une meilleure stabilité numérique du modèle. Il a ensuite fait l'objet d'un recalage.

Deux méthodologies ont été développées pour la génération de prévisions d'ensemble : (i) la calibration statistique d'ensembles bruts par la méthode *Quantile Regression Forests*, parfois appelée *méthode QRF* par la suite, et (ii) la mise en place d'assimilation de données dans les modèles de prévision avec un filtre de Kalman d'ensemble, appelée aussi *EnKF*. Ces deux méthodologies peuvent être développées après la réalisation d'études de sensibilité globale des modèles qui permettent de connaître les paramètres les plus significatifs pour les débits ou hauteurs d'eau simulés. Il s'agit de caractériser au mieux les incertitudes des modèles afin de les prendre en compte dans les prévisions d'ensemble. Une démarche commune a été suivie pour les prévisions d'ensemble générées, qu'elles soient issues de modélisations hydrologiques ou hydrauliques :

- Réalisation d'une étude de sensibilité globale du modèle.

- Génération d'une prévision d'ensemble brute suite aux résultats de l'étude de sensibilité : les paramètres incertains mis en évidence grâce à l'étude de sensibilité sont perturbés par tirage aléatoire.

- Calibration statistique de la prévision d'ensemble brute avec la méthode QRF sur une période de calage. On obtient alors une prévision d'ensemble calibrée, et on est capable de calibrer n'importe quelle nouvelle prévision brute dans un contexte opérationnel.

- Mise en place d'assimilation de données avec un filtre de Kalman d'ensemble au sein du modèle. Les variables corrigées sont celles pour lesquelles l'étude de sensibilité a montré une importance significative pour les débits ou hauteurs simulées. En prévision, les membres de l'EnKF forment une prévision d'ensemble.

[7]Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement, anciennement IRSTEA

[8]Direction Technique Générale

- Comparaison des deux méthodologies sur une période différente de la période de calage des forêts aléatoires, afin d'évaluer les performances respectives des ensembles dans un contexte opérationnel.

Au cours de ce travail, plusieurs partenariats ont été montés avec des organismes français qui interviennent dans la prévision hydrométéorologique. Tout d'abord, un travail a été mené en collaboration avec Météo-France et le CNRM[9]. Le CNRM nous a mis à disposition un historique de prévisions d'ensemble météorologiques PEAROME sur 112 jours en 2015-2016. Les extractions sur les points de grille des sous-bassins versants d'intérêt ont été réalisées dans le cadre du travail de thèse, puis ces données ont été retournées à Météo-France qui ont procédé à leur calibration statistique. Ces prévisions calibrées ont été utilisées pour la réalisation des études de sensibilité globale des modèles hydrologiques GRP et MORDOR-TS. Elles n'ont ensuite pas pu être utilisées comme forçage de la chaîne de prévision d'ensemble hydrologie-hydraulique du fait de leur couverture temporelle insuffisante. Par ailleurs, des partenariats ont été montés avec les organismes qui développent les deux modèles hydrologiques utilisés : INRAE pour le modèle GRP et EDF-DTG pour le modèle MORDOR-TS. En particulier, INRAE a adapté le modèle opérationnel GRP pour qu'il soit utilisable pour la réalisation d'une étude de sensibilité globale. La mise à disposition du modèle MORDOR-TS par EDF-DTG a été accompagnée d'une aide à sa prise en main.

Ce mémoire de thèse s'articule en trois grandes parties présentées ci-après.

Partie I : CONTEXTE ET MÉTHODOLOGIES DÉVELOPPÉES

La partie I regroupe trois chapitres. Elle a pour objectifs la présentation de la problématique, du bassin d'étude et des méthodes développées.

Le chapitre 1 présente le contexte de la prévision des crues en France, les incertitudes de prévision relatives aux modélisations météorologiques, hydrologiques et hydrauliques ainsi que les objectifs de la thèse.

Le chapitre 2 est consacré à la description physique du bassin versant de l'Odet et des modélisations hydrologiques et hydrauliques mises en œuvre dans ce travail.

Cette première partie se termine par le chapitre 3 dans lequel sont décrites les méthodologies mises en œuvre pour la génération des prévisions d'ensemble hydrologiques et hydrauliques. Nous y décrivons en premier lieu les caractéristiques recherchées pour une prévision d'ensemble. C'est à partir des scores présentés dans cette partie que seront analysées les prévisions d'ensemble construites par la suite. Ensuite, les études de sensibilité sont présentées. Puis, la suite du chapitre présente les deux grandes méthodes utilisées dans le travail de thèse pour générer des prévisions d'ensemble : la calibration statistique

[9] Centre National de Recherches Météorologiques

et l'assimilation de données ensembliste. Nous verrons comment ces deux méthodologies nécessitent de connaître les résultats de l'étude de sensibilité pour être développées.

Partie II : LES PRÉVISIONS D'ENSEMBLE HYDROLOGIQUES

La partie II regroupe trois chapitres également. Elle traite de la génération et de l'évaluation des prévisions d'ensemble hydrologiques.

Le chapitre 4 traite des études de sensibilité des modèles hydrologiques GRP et MORDOR-TS, de la génération d'un ensemble hydrologique brut avec MORDOR-TS et de sa calibration par la méthode des forêts aléatoires. Il est constitué du résumé de deux articles. Le premier article est intitulé *Étude de sensibilité globale comparée des modèles hydrologiques GRP et MORDOR-TS*. Il a été soumis au journal *La Houille Blanche*, des révisions mineures sont demandées. Le second article a été accepté au journal *Journal of Hydrology*. Il s'intitule *Strategies for hydrologic ensemble generation and calibration: on the merits of using model-based predictors*[10].

Le chapitre 5 présente l'assimilation de données ensembliste mise en place dans le modèle hydrologique MORDOR-TS. Il présente en particulier un ensemble d'expériences jumelles qui permettent de valider l'algorithme d'assimilation de données construit. Il présente également les résultats des expériences d'assimilation de données dans un cadre réel, qui pourraient être menées dans un contexte opérationnel. Les résultats sont alors uniquement présentés sur la fenêtre d'assimilation. Les résultats en prévision opérationnelle sont présentés dans le chapitre 6.

En effet, dans le chapitre 6 sont comparées les prévisions d'ensemble générées par calibration statistique et par assimilation de données. On se place alors sur une période différente de celle qui a servi à la construction des forêts aléatoires pour évaluer les performances des deux méthodologies dans un contexte opérationnel. Ce chapitre permet de conclure sur les stratégies à mettre en place pour une prévision d'ensemble hydrologique opérationnelle.

Partie III : LES PRÉVISIONS D'ENSEMBLE HYDRAULIQUES

Pour finir, la partie III présente la génération et l'évaluation des prévisions d'ensemble hydrauliques. Cette partie regroupe quatre chapitres.

Les chapitres 7 et 8 sont analogues au chapitre 4 mais pour la modélisation hydraulique. Le chapitre 7 présente l'étude de sensibilité globale menée sur le modèle hydraulique MASCARET. Il conclut sur les paramètres à prendre en compte pour générer les prévisions d'ensemble brutes et pour l'assimilation de données. Le chapitre 8 présente la génération des ensembles hydrauliques bruts et leur calibration par la méthode des forêts aléatoires.

[10] Tiberi-Wadier, A-L., Goutal, N., Ricci, S., Sergent, P., Taillardat, M., Bouttier, F.,Monteil, C., Strategies for hydrologic ensemble generation and calibration: on the merits of using model-based predictors, Journal of Hydrology (2021), doi: https://doi.org/10.1016/j.jhydrol.2021.126233

Dans le chapitre 9, analogue du chapitre 5 mais pour la modélisation hydraulique, on présente d'abord des expériences jumelles d'assimilation de données avant de mener des expériences dans un cadre réel. Les expériences réelles sont là aussi évaluées uniquement sur la fenêtre d'assimilation.

Pour finir, le chapitre 10, analogue du chapitre 6 mais pour l'hydraulique, compare les performances des ensembles générés par calibration statistique et par assimilation de données au cours de trois hivers indépendants de la période ayant servi à la calibration des forêts aléatoires. Ce chapitre permet de conclure sur les stratégies à mettre en place pour une prévision d'ensemble hydraulique opérationnelle.

Le mémoire se poursuit naturellement par des conclusions générales sur les méthodologies développées et oriente vers les meilleures stratégies à mettre en place en opérationnel. Des perspectives de poursuite et d'approfondissement du travail sont données. Enfin, pour une lecture plus fluide du manuscrit, de nombreux éléments ont été placés en annexe. Le lecteur intéressé pourra s'y reporter pour davantage de détails sur les méthodologies employées ou sur les résultats obtenus.

Partie I

CONTEXTE ET MÉTHODOLOGIES DÉVELOPÉES

Chapitre 1

Contexte de la prévision des crues et objectifs de la thèse

Sommaire

1.1 La chaîne de prévision hydrométéorologique et incertitudes associées . . . 11

1.1.1 La chaîne de prévision météorologie-hydrologie-hydraulique . . . 12

1.1.1.1 Terminologie relative à la prévision . . . 12

1.1.1.2 De la prévision météorologique à la prévision hydraulique 13

1.1.1.3 Comment tenir compte des incertitudes de prévision ? . 14

1.1.2 Les incertitudes dans la chaîne de prévision . . . 16

1.1.2.1 Les incertitudes météorologiques . . . 16

1.1.2.2 Les incertitudes hydrologiques . . . 17

1.1.2.3 Les incertitudes hydrauliques . . . 18

1.1.3 Produits opérationnels et de recherche pour les prévisions hydrométéorologiques . . . 20

1.1.3.1 Les prévisions d'ensemble météorologiques . . . 20

1.1.3.2 Les prévisions d'ensemble hydrologiques et hydrauliques 21

1.2 La prévision opérationnelle des crues en France et la prise en compte des incertitudes . . . 27

1.2.1 La prévision opérationnelle au sein du réseau *Vigicrues* . . . 27

1.2.1.1 Historique de la prévision des crues . . . 27

1.2.1.2 Le réseau *Vigicrues* . . . 29

1.2.1.3 La prise en compte des incertitudes de prévision . . . 33

1.2.1.4 Le système Vigicrues-Flash . . . 34

1.2.2 La prévision opérationnelle à EDF . . . 35

1.2.2.1 Description de la chaîne opérationnelle de prévision . . 35

1.2.2.2 La prise en compte des incertitudes de prévision 36
1.2.3 Conclusion 37
1.3 Objectifs de la thèse et méthodologie suivie 37

Les inondations représentent le premier risque naturel en France[1]. Les inondations par débordement de cours d'eau concernent potentiellement 17 millions d'habitants et plus de 9 millions d'emplois. Les submersions marines concernant quant à elles 1.4 millions d'habitants et plus de 850 000 emplois. En France, le Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations (SCHAPI) et les 19 Services de Prévision des Crues (SPC) répartis sur le territoire sont responsables de la veille hydrométéorologique et surveillent à ce titre environ 22 000 km de cours d'eau, formant le réseau *Vigicrues*. Ce réseau émet des bulletins de vigilance deux fois par jour pour prévenir le public et les acteurs de la gestion de crise en cas de risque de crue sur les cours d'eau surveillés par l'État.

Ce travail de thèse a pour objectif la réalisation de prévisions hydrométéorologiques pour la vigilance crues nationale qui tiennent compte des incertitudes. Pour cela, des prévisions d'ensemble hydrologiques et hydrauliques sont mises en place. Dans ce chapitre, nous décrivons au paragraphe 1.1 la chaîne de prévision hydrométéorologique opérationnelle et les incertitudes présentes le long de cette chaîne. Nous décrivons également les projets de recherche et produits opérationnels qui prennent en compte les incertitudes prédictives par la réalisation de prévisions probabilistes. Le paragraphe 1.2 présente quant à lui deux chaînes de prévision opérationnelles françaises : le réseau *Vigicrues* et les centres opérationnels de prévision d'EDF, qui gèrent la sûreté des grands barrages et l'optimisation de la ressource hydroélectrique. Pour finir, le paragraphe 1.3 présente les objectifs de la thèse et la méthodologie suivie.

1.1 La chaîne de prévision hydrométéorologique et incertitudes associées

Dans cette section, le paragraphe 1.1.1 présente la chaîne de prévision opérationnelle des services de prévision des crues. Celle-ci est constituée d'un chaînage de modèles hydrologiques et hydrauliques. Dans notre cas, le modèle hydrologique est alimenté par des prévisions météorologiques. Il peut également être envisagé de prendre en compte la pluie dans le modèle hydraulique mais cela n'a pas été considéré dans cette chaîne de prévision. Plusieurs méthodes permettant de tenir compte des incertitudes de prévision sont également décrites. En particulier, la réalisation de prévisions d'ensemble est une méthode permettant de les appréhender. Dans le paragraphe 1.1.2, les incertitudes associées aux prévisions météorologiques, hydrologiques et hydrauliques sont décrites en détail. Enfin, le paragraphe 1.1.3 présente des systèmes opérationnels et des études qui tiennent compte des incertitudes de prévision le long de la chaîne de prévision hydrométéorologique.

[1]Source Ministère de la Transition Écologique, Plaquette inondation

1.1.1 La chaîne de prévision météorologie-hydrologie-hydraulique

1.1.1.1 Terminologie relative à la prévision

Cette thèse a pour objectif la construction de prévisions opérationnelles hydrologiques et hydrauliques. Le terme *prévision opérationnelle* signifie que la prévision est réalisée de manière récurrente, à heure fixe, par des services dédiés. L'objectif est d'estimer l'évolution au cours du temps d'une variable temporelle. Dans notre cas, nous chercherons à estimer l'évolution des débits ou des hauteurs d'eau.

Avant de s'intéresser plus en détail à la chaîne de prévision hydrométéorologique investiguée dans ce travail, il est nécessaire de décrire la terminologie relative aux prévisions. Le moment auquel est réalisé la prévision, c'est-à-dire le moment auquel sont lancés les modèles de prévision météorologiques, hydrologiques ou hydrauliques, est appelé *temps de base*. Il est noté T_B par la suite. L'*échéance* d'une prévision correspond à la durée séparant le temps de base et l'instant futur pour lequel la prévision est faite. L'*horizon* d'un système de prévision correspond à l'échéance maximale de prévision disponible. Le schéma 1.1 illustre les différents termes utilisés : temps de base, échéance et horizon de prévision. Dans cette thèse, l'horizon des prévisions hydrologiques et hydrauliques construites est de 24 heures. Cela correspond à l'horizon de prévision classique pour la prévision opérationnelle en France, même si cet horizon de prévision peut être supérieur dans certains cas (voir paragraphe 1.2.1).

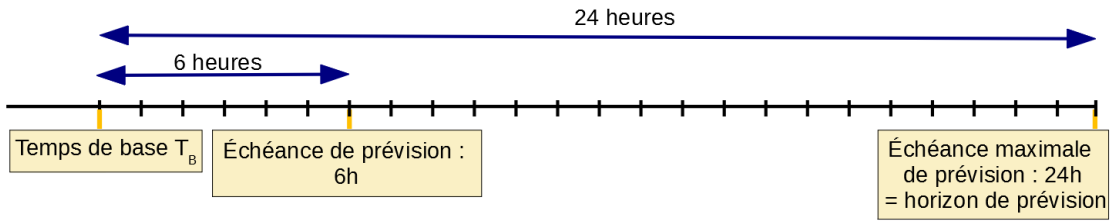


Figure 1.1: Illustration des termes relatifs à la réalisation d'une prévision : temps de base T_B, échéance d'une prévision et horizon de prévision.

Les usages des prévisions hydrométéorologiques sont multiples. On peut citer l'optimisation des ressources hydroélectriques, la prévision des étiages, et la vigilance pour les crues. C'est ce dernier aspect qui nous intéresse dans ce travail. Aussi, les prévisions construites seront analysées spécifiquement pendant des événements pluvieux. L'organisation de la vigilance crues au niveau national est décrite au paragraphe 1.2.1, mais nous nous plaçons dès à présent dans le cadre des prévisions opérationnelles réalisées par les Services de Prévision des Crues (SPC).

1.1.1.2 De la prévision météorologique à la prévision hydraulique

La chaîne de prévision opérationnelle au sein des Services de Prévision des Crues peut être représentée par le schéma de la figure 1.2. Deux types de modélisations sont effectuées de manière séquentielle : une modélisation hydrologique et une modélisation hydraulique. Ces modélisations s'appuient sur deux données essentielles :

(i) La connaissance de l'état passé et actuel : pluies tombées sur le bassin versant pour la prévision hydrologique, mesures de débits et de hauteurs d'eau pour la prévision hydraulique.

(ii) D'autres prévisions : prévisions météorologiques pour la prévision hydrologique, prévision de débits, prévision de cote le cas échéant pour une prévision hydraulique.

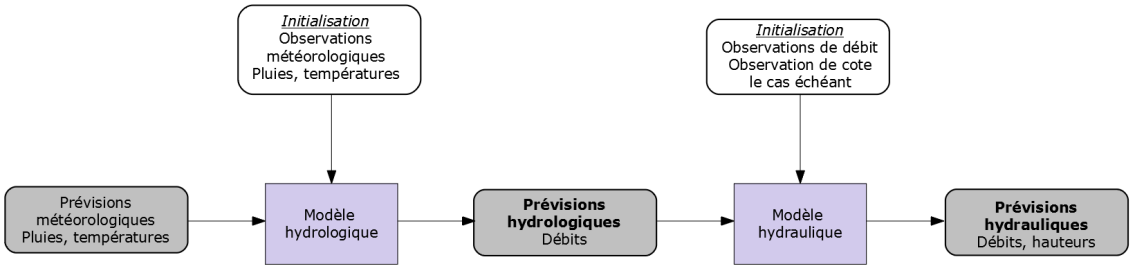


Figure 1.2: Chaîne de prévision opérationnelle dans les SPC.

Les prévisions météorologiques représentent le forçage initial de la chaîne de prévision. Tous les modèles hydrologiques ont besoin d'être alimentés par des précipitations. Certains utilisent également la température, ce qui leur permet de calculer l'évapotranspiration (cas du modèle hydrologique MORDOR-TS, voir chapitre 2), alors que d'autres utilisent une courbe d'évapotranspiration moyenne annuelle calculée au préalable (cas du modèle hydrologique GRP, voir chapitre 2). Les prévisions météorologiques sont issues de prévisions numériques du temps (Numerical Weather Prediction, NWP) et sont réalisées par des centres de prévision météorologique (Météo-France, CEPMMT[2], ...)

Le modèle hydrologique est initialisé avant le lancement de la prévision. Durant cette phase d'initialisation, fixée généralement à un an, il est alimenté par des mesures de pluie et de température le cas échéant. Cela permet au modèle de se trouver dans un état initial cohérent avec l'historique des précipitations et des températures sur le bassin versant au moment du lancement de la prévision. Après la phase d'initialisation, l'état du modèle n'est plus influencé par l'état initial. En prévision, le modèle hydrologique est alimenté par des prévisions météorologiques. Les résultats de la simulation sont une chronique de prévision de débits à l'exutoire de chaque bassin versant modélisé.

Le modèle hydraulique doit lui aussi être initialisé pour être dans un état cohérent avec

[2]Centre Européen pour les Prévisions Météorologiques à Moyen Terme. En anglais ECMWF, European Centre for Medium-range Weather Forecasts.

l'historique des débits et des hauteurs d'eau avant le temps de base. Pour un modèle hydraulique, la durée d'initialisation est nettement plus réduite que pour l'hydrologie, quelques jours sont suffisants. Il est possible d'initialiser le modèle avec une ligne d'eau cohérente avec les conditions aux limites imposées au premier pas de temps. Pour cela, le modèle doit être préalablement lancé en permanent avec les conditions limites égales à celles du premier pas de temps. En prévision, le modèle hydraulique est alimenté par les prévisions hydrologiques de débit, et le cas échéant par des prévisions de cotes à l'aval (cas d'une condition limite maritime). Les résultats de cette simulation sont une chronique de prévision de débits et de hauteurs d'eau aux points d'intérêt, qui sont les stations de vigilance des Services de Prévision des Crues.

1.1.1.3 Comment tenir compte des incertitudes de prévision ?

Il s'avère que des incertitudes sont présentes tout au long de la chaîne de prévision, au niveau de la prévision météorologique, et des modélisations hydrologiques et hydrauliques. Ces incertitudes seront décrites en détail au paragraphe 1.1.2, mais nous pouvons d'ores et déjà affirmer qu'une prévision déterministe chaînée des modèles est vouée à se tromper. Il est admis aujourd'hui que les prévisions doivent être accompagnées d'informations permettant d'estimer leur incertitude. La fourniture de prévisions probabilistes à la place d'une prévision déterministe est alors tout à fait adaptée, et c'est la tendance générale vers laquelle tendent les systèmes de prévision hydrométéorologiques opérationnels (voir paragraphe 1.1.2 pour des exemples). De nombreux projets de recherche ont également pour objectif la génération de prévisions probabilistes. Avec des prévisions probabilistes, l'estimation de la probabilité d'un événement est estimée par la fraction des membres qui prévoient l'événement. La dispersion des prévisions reflète le degré de confiance dans la prévision.

Il existe plusieurs manières d'obtenir des prévisions probabilistes [Jolliffe and Stephenson, 2003]. Tout d'abord, on peut utiliser des méthodes statistiques basées sur des données historiques, encore appelée méthode des analogues [Lorenz, 1969]. Cette méthode est largement utilisée pour la prise en compte des incertitudes de prévision météorologiques. Elle part du constat que la prévision des variables synoptiques (pression, vent, température) est de bien meilleure qualité que celle des précipitations, et repose sur le principe que deux situations météorologiques similaires à l'échelle synoptique produiront des situations météorologiques locales similaires. Une situation est caractérisée par des prédicteurs qui sont des variables décrivant la situation de l'atmosphère grande échelle. Une ou plusieurs situations analogues sont alors retenues selon un critère d'analogie préalablement défini. La prévision par analogie repose alors sur l'extraction du prédictand (la précipitation) dans une archive météorologique pour ces situations analogues.

Une autre méthode pour obtenir des prévisions probabilistes est d'appliquer à un modèle

déterministe des statistiques de performances passées. Cette méthodologie est par exemple utilisée pour perturber *a posteriori* des prévisions de précipitations afin de créer un ensemble caractérisant l'incertitude sur les prévisions (Vincendon et al. [2011], voir paragraphe 1.1.3.1). Elle est également à l'origine de la méthode QUOIQUE (QUantification Opérationnelle des Incertitudes par QUantiles de dEbit, Bourgin [2014], voir paragraphe 1.2.1.3).

Enfin, des prévisions probabilistes peuvent être produites par la réalisation d'un ensemble de prévisions déterministes, indépendantes et équiprobables au cours du temps. Cette dernière technique est nommée la prévision d'ensemble. Les différentes prévisions déterministes sont réalisées à partir de réalisations perturbées du modèle. L'ensemble des réalisations permet d'estimer la densité de probabilité de la prévision et de caractériser l'incertitude de prévision. Ce principe est illustré sur la figure 1.3.

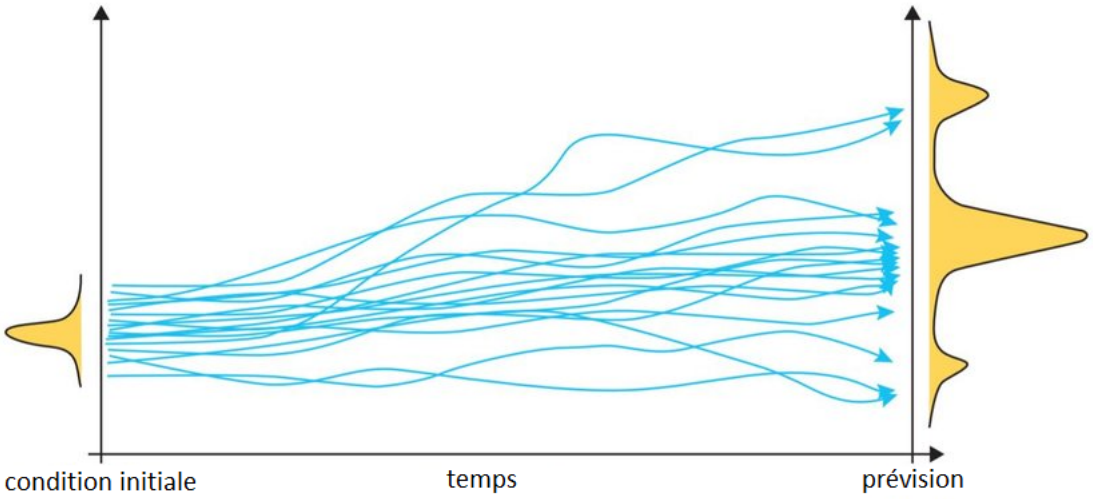


Figure 1.3: La prévision d'ensemble produit un ensemble de scénarios possibles, à partir desquels il est possible d'estimer la densité de probabilité prévue en fonction des différents membres de l'ensemble. D'après Grönquist et al. [2019].

La prévision d'ensemble permet aux prévisionnistes d'évaluer les différents scénarios et d'informer les usagers de leur vraisemblance [Bouttier and Buizza, 2018]. Elle permet d'estimer le scénario le plus probable et son incertitude. Par ailleurs, les prévisions d'ensemble fournissent des prévisions successives plus cohérentes entre elles que les prévisions déterministes. Dans cette thèse, nous nous intéressons uniquement aux méthodes de génération de prévisions d'ensemble. Les autres méthodes citées permettant de prendre en compte les incertitudes de prévision ne sont pas abordées.

1.1.2 Les incertitudes dans la chaîne de prévision

Les incertitudes sont couramment classées selon leur nature. Les incertitudes dites *épistémiques* sont celles liées à un manque de connaissance sur le système. Les incertitudes dites *aléatoires* sont en revanche liées à la nature stochastique du système. Cette dernière incertitude est irréductible. Ce paragraphe est dédié à la description des incertitudes de prévision météorologiques (section 1.1.2.1), hydrologiques (section 1.1.2.2) et hydrauliques (section 1.1.2.3).

1.1.2.1 Les incertitudes météorologiques

L'atmosphère étant par nature chaotique [Lorenz, 1963], des petites incertitudes sur la condition initiale peuvent avoir des conséquences importantes sur la précision de la prévision, et l'amplification de l'erreur initiale dépend de l'état de l'atmosphère. Les prévisions d'ensemble, dont les premières ont été lancées il y a 25 ans, fournissent des estimations précises des évolutions possibles du temps au lieu d'un scénario unique [Bouttier and Buizza, 2018]. Elles permettent au prévisionniste de connaître à la fois le scénario le plus probable et la probabilité que n'importe quelle autre évolution se produise.

Dans la prévision d'ensemble météorologique, le modèle numérique de prévision du temps est lancé plusieurs fois pour une même situation à prévoir. Les ensembles sont construits pour caractériser les sources d'erreur liées aux conditions initiales et aux faiblesses des modèles numériques de prévision [Bouttier and Buizza, 2018]. Les incertitudes sur les conditions initiales de l'atmosphère sont affectées par les erreurs d'observation, leur couverture limitée, et par des approximations dans les modèles statistiques et numériques utilisés. Elles sont prises en compte à partir de différentes conditions initiales différant entre elles par de petites quantités, appelées perturbations initiales, compatibles avec les incertitudes existant sur la connaissance de l'état initial de l'atmosphère [Palmer et al., 2002]. L'imperfection des modèles numériques est liée au fait qu'ils ne représentent que certains processus physiques sur une grille de résolution limitée. Les systèmes de prévisions d'ensemble peuvent aujourd'hui classiquement utiliser quatre approches différentes pour représenter les erreurs de modélisation : (i) l'approche multi-modèles dans laquelle différents modèles sont utilisés pour les différents membres de l'ensemble, (ii) la perturbation des paramètres du modèle [Houtekamer et al., 1996], (iii) la perturbation des tendances [Buizza et al., 1999] et (iv) la rétrodiffusion stochastique, dans laquelle un générateur stochastique d'énergie cinétique simule l'effet des processus que le modèle ne peut pas représenter. D'autres aspects sont essentiels pour la configuration d'un système de prévision d'ensemble : il s'agit de la résolution du modèle, de l'horizon de prévision et du nombre de membres. La plupart des ensembles opérationnels ont entre 20 et 50 membres. Pour la prévision des événements rares, 50 à 100 membres peuvent être nécessaires. Les prévisions d'ensemble météorologiques sont généralement sous-dispersives et biaisées

[Hamill and Colucci, 1997, Schaake et al., 2010]. Une étape de post-traitement, également appelée calibration statistique, est nécessaire avant leur utilisation. La description des méthodes de post-traitement est réalisée au chapitre 3.3.

1.1.2.2 Les incertitudes hydrologiques

Les incertitudes hydrologiques sont classées selon deux origines [Krzysztofowicz, 1999] : (i) les incertitudes sur le forçage et (ii) les incertitudes sur le modèle lui-même (conditions initiales, structure du modèle et ses paramètres). Les incertitudes sur le forçage, qui concerne les précipitations et éventuellement les températures, sont classiquement prises en compte en prévision avec l'utilisation de prévisions d'ensemble météorologiques.

L'étude des incertitudes sur le modèle a été largement investiguée dans la littérature. Dans la méthode GLUE (*Generalized Likehood Uncertainty Estimation*, voir paragraphe 3.2.4.1 pour la descritpion de la méthode), Beven and Binley [1992] partent du principe qu'il n'existe pas un unique jeu de paramètres qui permette de représenter le comportement du système. Ils proposent de conserver l'ensemble des jeux de paramètres dont les performances sont supérieures à un certain seuil. L'incertitude sur les débits est estimée grâce à l'ensemble des débits simulés avec les différents jeux de paramètres. Bourgin et al. [2014] a cependant montré que les approches ne s'appuyant que sur un ensemble différent de jeux de paramètres ne parviennent généralement pas à représenter correctement l'incertitude prédictive totale. L'utilisation d'un modèle d'erreur ou de modèles de post-traitement permet de mieux estimer les erreurs du modèle.

L'incertitude repose également sur la structure du modèle. Elle traduit le fait que le modèle est une représentation simplifiée de la réalité. Le modèle a donc une capacité limitée à représenter la dynamique réelle du système modélisé. Clark et al. [2008b] ont montré qu'il n'existe pas une structure de modèle hydrologique qui soit meilleure que les autres en toutes circonstances. Ainsi, cette incertitude peut être prise en compte par une approche multi-modèles [Bellier, 2018, Hopson and Webster, 2010, Thiboult et al., 2016, Velázquez et al., 2011].

Pour finir, l'état initial du bassin est également incertain. Cette information est néanmoins cruciale pour prévoir la réponse des cours d'eau [Le Lay and Saulnier, 2007]. Cette incertitude peut être reliée aux incertitudes sur les données, puisque ces dernières sont utilisées pour initialiser le modèle. Les données issues de l'observation servent au forçage (pluie, température ou ETP) et également à l'estimation de la variable de réponse (débit). Les erreurs peuvent être dues directement à la mesure mais également au traitement de la donnée. Les incertitudes concernant les précipitations sont en partie liées au sous-échantillonnage du champ de précipitations, puisqu'un nombre limité de pluviomètres est utilisé pour estimer la lame d'eau sur un bassin versant. Les débits sont estimés à partir de hauteurs d'eau mesurées par l'intermédiaire d'une courbe de tarage, ce qui introduit

également une incertitude due au nombre limité de jaugeages à partir desquels elle est construite, à l'interpolation réalisée dans la gamme de débits mesurés et l'extrapolation faite pour couvrir une gamme complète de débits. Thiboult et al. [2016] ont utilisé l'assimilation de données pour mettre à jour l'état initial du modèle.

1.1.2.3 Les incertitudes hydrauliques

De manière analogue à la modélisation hydrologique, les incertitudes en hydraulique ont pour origine les incertitudes sur le modèle (équations, résolution, hypothèses réalisées, paramètres du modèle et topographie/bathymétrie) et les incertitudes sur les forçages. Nous décrivons ci-après ces différentes sources d'incertitude.

Les incertitudes en hydraulique sont d'abord liées aux équations du modèle et aux hypothèses faites pour leur résolution. Des incertitudes sont liées à la simplification des équations de la physique, puisque les équations de Navier Stokes sont simplifiées par la formulation des équations de Saint-Venant. De même l'écoulement est simplifié. Les modèles hydrauliques sont classiquement en 1D ou 2D, alors qu'ils représentent des écoulements 3D. Les schémas numériques de résolution reposent sur une discrétisation spatiale et temporelle qui est également source d'incertitude.

Dans les modèles hydrauliques, des incertitudes sont également liées à la méconnaissance des caractéristiques hydrauliques de la rivière. En particulier, le frottement du fond de la rivière et des berges est caractérisé par un coefficient de frottement. La zone modélisée est découpée en sous-espaces sur lesquels les valeurs pour le frottement sont considérées comme homogènes. Dans le cas d'une modélisation 2D, le découpage pourra être assez souple et mettra en valeur en particulier le lit mineur et les plaines d'inondation. Plusieurs valeurs sont possibles dans la direction transverse à l'écoulement. Dans la plaine d'inondation, le frottement sera plus important dans les zones urbaines par exemple. Dans le cas d'une modélisation 1D, le découpage sera réalisé selon des tronçons de cours d'eau, pour lesquels sont imposés un coefficient pour le lit mineur et un pour le lit majeur, les lits majeurs droit et gauche pouvant être affectés d'une valeur différente. Selon la nature des surfaces pour les canaux artificiels ou selon les caractéristiques des cours d'eau naturels, des valeurs empiriques pour le frottement sont disponibles dans la littérature. La pratique est cependant de réaliser un calage de ces valeurs à partir de données observées. Le calage repose le plus souvent sur la minimisation d'une fonction d'écart entre les données observées et modélisées. Le problème d'équifinalité est alors d'autant plus important que le nombre de coefficients de frottement à caler est grand et que le nombre de points de mesure est faible.

La connaissance incomplète de la géométrie du cours d'eau et de la plaine d'inondation est également une source d'incertitude. Si les données topographiques peuvent être de bonne qualité lorsqu'elles sont réalisées par des relevés LIDAR à haute résolution (1 point par m^2,

précision verticale de 10 cm à 30 cm), la bathymétrie correspondant aux zones immergées n'est accessible que par la réalisation de profils bathymétriques. Cette information est donc toujours parcellaire. Lors de la construction d'un modèle 1D, une étape de planimétrage permet de traduire la hauteur d'eau en surface mouillée pour les sections transverses à l'écoulement. Plus le pas de planimétrage est fin, plus l'estimation de la surface mouillée est de bonne qualité. Le choix du pas de planimétrage peut introduire de l'incertitude si celui-ci n'est pas adapté à la hauteur d'eau dans la section. La construction d'un modèle 1D repose également sur le choix d'un pas pour le maillage. Plus le pas est fin, mieux on représente les variations spatiales du domaine, mais le temps d'exécution du modèle est d'autant plus long. Un compromis doit donc être fait. Un modèle hydraulique 2D repose quant à lui sur un maillage pour lequel chaque nœud est décrit par ses coordonnées (x,y,z). Une étape d'interpolation de la topographie/bathymétrie sur le maillage est réalisée et introduit des erreurs. Pour finir, les relevés topographiques/bathymétriques utilisés sont réalisés à une date donnée, qui n'est d'ailleurs pas forcément la même pour les différents relevés utilisés au sein d'un même modèle. En tout état de cause, les modifications de la bathymétrie et de la topographie avec le temps ne sont pas prises en compte, en particulier après le passage de crues.

Des incertitudes concernent enfin les conditions limites imposées dans le modèle. Comme pour les modèles hydrologiques, l'incertitude sur les données mesurées utilisées pendant la phase d'initialisation du modèle introduit une incertitude sur l'état initial du modèle. Cette incertitude se propage ensuite dans le réseau hydraulique. À l'amont, le forçage décrit les apports en eau. Il est donné sous la forme d'un débit. Des mesures de débit sont utilisées pour réaliser des simulations en mode rejeu, alors que les prévisions hydrauliques sont réalisées à partir de prévisions hydrologiques. Dans les deux cas, ces données sont entachées d'incertitudes. Dans le cas de prévisions, ces incertitudes sont celles liées à la modélisation hydrologique décrites précédemment, et augmentent avec l'échéance de prévision. À l'aval, deux types de conditions limites peuvent être imposées. Si des mesures sont disponibles, la cote d'eau est imposée. Dans le cas contraire, on peut utiliser une courbe de tarage. Cette dernière est également entachée d'erreurs comme nous l'avons vu précédemment. Dans le cas d'une condition aval maritime, les mêmes incertitudes que pour la condition limite amont dues aux erreurs de mesure se retrouvent en mode rejeu. En prévision, il s'agit d'imposer une chronique de cotes prévues à l'aval. Celle-ci est constituée de la somme de la marée astronomique, prévue par un modèle de marée, et d'une prévision de surcote. La prévision de surcote est entachée d'incertitudes, qui, comme dans le cas d'une prévision hydrologique, augmente avec l'échéance de prévision.

1.1.3 Produits opérationnels et de recherche pour les prévisions hydrométéorologiques

Dans cette partie nous décrivons section 1.1.3.1 les produits opérationnels et de recherche dont l'objectif est la prise en compte des incertitudes de prévision météorologiques. La prise en compte des incertitudes de prévision hydrologiques et hydrauliques est présentée section 1.1.3.2. La description des chaînes opérationnelles n'inclura pas les chaînes opérationnelles françaises qui seront décrites plus en détail au paragraphe 1.2.

1.1.3.1 Les prévisions d'ensemble météorologiques

Nous décrivons dans ce paragraphe les principaux produits opérationnels disponibles dans le monde, en Europe et en France en termes d'ensembles météorologiques. Le programme THORPEX[3] (2003-2014) de l'Organisation Météorologique Mondiale avait pour objet l'étude de l'atmosphère terrestre pour mieux prédire le temps jusqu'à l'horizon de 2 semaines afin en particulier d'amenuiser ses effets sur la population. Il est à l'origine de la création de la base de prévisions TIGGE[4] [Park et al., 2008]. Cette base de données, construite à des fins scientifiques, débute en 2006 et regroupe des prévisions d'ensemble d'une dizaine de centres de prévision météorologique à travers le monde. La base de données TIGGE sera prolongée *a minima* jusqu'à 2023. Au niveau européen, le CEPMMT[5] propose une prévision d'ensemble saisonnière (7 mois, résolution 80 km), mensuelle (46 jours, résolution 36 km) et de moyenne échéance (15 jours, résolution 18 km). Toutes ces prévisions comportent 51 membres. En France, Météo-France dispose de plusieurs produits de prévision d'ensemble [Bouttier and Buizza, 2018]. La prévision saisonnière a pour échéance maximale 7 mois, elle comporte 51 membres et a pour résolution 80 km. La prévision d'ensemble ARPEGE (PEARP) est une prévision globale construite à partir du modèle météorologique global ARPEGE[6] optimisé sur la France. Cette prévision d'ensemble possède 35 membres, a pour horizon de prévision 4.6 jours et est réalisée 2 fois par jour. La prévision d'ensemble AROME (PEAROME) est une prévision régionale. Elle a pour horizon de prévision 45 heures, a une résolution horizontale de 2.5 km et possède 12 membres. À partir de 2018, la PEAROME possède 16 membres. Ce système de prévision d'ensemble est celui utilisé dans ce travail de thèse et sera plus largement décrit au chapitre 2.

En parallèle au développement des systèmes de prévision d'ensemble à Météo-France, une autre méthodologie [Vincendon et al., 2011] a été développée au CNRM[7]. Cette méthode non opérationnelle est utilisée dans plusieurs projets de recherche concernant la réalisa-

[3] THe Observing system Research and Predictability EXperiment
[4] THORPEX Interactive Grand Global Ensemble
[5] Centre Européen pour les Prévisions Météorologiques à Moyen Terme
[6] Action de Recherche Petite Échelle Grande Échelle
[7] Centre National de Recherches Météorologiques

tion de prévisions d'ensemble dans un contexte de crues rapides méditerranéennes (voir paragraphe 1.1.3.2). Elle consiste à perturber *a posteriori* des prévisions de précipitations afin de créer un ensemble caractérisant l'incertitude sur les prévisions. Les perturbations sont réalisées à partir d'une climatologie des erreurs de prévision des pluies (localisation et intensité) lors d'événements pluvieux intenses dans la région méditerranéenne. Cette méthode est appelée *PERT* par la suite. Une étude comparant les prévisions de précipitations issues de la méthode *PERT* appliquée aux prévisions déterministes AROME avec celles issues de la PEAROME sur quatre événements pluvieux majeurs de 2014 montre que ces deux prévisions ont des caractéristiques semblables [de Saint-Aubin et al., 2018].

1.1.3.2 Les prévisions d'ensemble hydrologiques et hydrauliques

Dès les années 2000, Krzysztofowicz [2001] insiste sur l'importance de la réalisation de prévisions probabilistes en hydrologie. La prévision d'ensemble météorologique à moyenne échéance a commencé à être développée dans les années 1990. Les prévisions d'ensemble en hydrologie ont quant à elles débuté dans les années 2000. Elles se sont rapidement répandues, grâce notamment à des initiatives de coopération entre le monde de la recherche et celui de l'application opérationnelle. Deux principales initiatives méritent d'être citées : la création du groupe HEPEX (Hydrologic Ensemble Prediction EXperiment) en 2004 et le projet EFAS (European Flood Awarness System), initié en 2003. De nombreux projets de recherche passés ou en cours ont aussi pour objectif la prise en compte des incertitudes pour les crues rapides méditerranéennes et la réalisation de prévisions d'ensemble pour les appréhender. Nous décrivons également dans ce paragraphe la prise en compte des incertitudes de prévisions liées à la modélisation hydraulique.

Le groupe HEPEX La communauté HEPEX [Schaake et al., 2010, Thielen et al., 2008] rassemble des chercheurs, des producteurs opérationnels et des utilisateurs des prévisions. Cette initiative a pour objectifs l'amélioration des connaissances et des pratiques pour la prévision d'ensemble hydrologique, et la démonstration de leur utilité pour la prise de décision. Le groupe HEPEX organise ainsi de nombreuses activités : organisation de workshops, coordination d'expériences, organisation de publications scientifiques par l'intermédiaire de *special issues* dans des journaux scientifiques, mise en avant de systèmes de prévisions opérationnels ou expérimentaux pour aider les utilisateurs à comprendre comment les prévisions d'ensemble sont utilisées dans le monde pour des applications diverses, maintenance du lien entre les membres par le site web. Des premiers démonstrateurs sont construits dans le but de mettre en place dans un second temps des systèmes opérationnels de prévision. Cloke and Pappenberger [2009], Pappenberger et al. [2016] présentent des synthèses des systèmes de prévision d'ensemble hydrométéorologiques existants. Les systèmes de prévision d'ensemble ont d'abord été construits en propageant

les ensembles météorologiques dans un modèle hydrologique. Puis, ces systèmes se sont complexifiés, en intégrant les incertitudes liées à la modélisation hydrologique.

Le projet EFAS Le projet EFAS[8] est un système européen d'alerte aux crues initié en 2003 dans le but d'améliorer la préparation lors de crues concernant les grands bassins versants transnationaux. EFAS fournit des alertes très en amont (entre 3 et 10 jours en avance) aux services de prévision nationaux. Cela contribue à l'augmentation du temps de préparation des services opérationnels en cas de crues dangereuses [Thielen et al., 2009]. Plusieurs systèmes de prévision météorologique d'ensemble sont utilisés en entrée du modèle hydrologique distribué LISFLOOD [Van Der Knijff et al., 2010]. Ce modèle hydrologique se situe entre les modèles conceptuels et les modèles physiques. Il est spécialement adapté pour reproduire les principaux processus hydrologiques pour les moyens et grands bassins versants. Le système EFAS a été évalué dans de nombreuses études scientifiques [Demeritt et al., 2013, Pappenberger et al., 2011]. En particulier, Alfieri et al. [2014] comparent les prévisions de débits de EFAS à une simulation de référence utilisant des mesures en forçage pour chaque point de grille du réseau européen, pour des échéances de prévision allant jusqu'à 10 jours. Les meilleures prévisions sont réalisées pour les bassins moyens à grands pour les 10 jours d'horizon de prévision. La qualité de la prévision chute de manière significative pour les points de prévision pour lesquels le bassin amont est inférieur à 300 km^2. Cela est dû en particulier à la sous-estimation du débit au niveau des bassins de montagne.

La prise en compte des incertitudes dans les phénomènes cévenols Les crues rapides méditerranéennes sont engendrées par des événements dits méditerranéens ou cévenols, qui génèrent des précipitations considérables en un temps réduit. Ces événements surviennent essentiellement à l'automne. Nous décrivons ci-après plusieurs projets, dont le projet international HyMeX (Hydrological cycle in the Mediterranean eXperiment), son projet associé ANR PICS et le projet CHROME.

Le projet HyMeX est un programme de recherche international interdisciplinaire (météorologie, climat, océanographie, hydrologie, sciences humaines et sociales) dont l'objectif est d'améliorer la compréhension du cycle de l'eau en Méditerranée, et notamment d'améliorer les connaissances et la prévision des risques hydrométéorologiques. Il a permis des avancées en termes de réduction et de quantification des incertitudes de la prévision des pluies intenses et des crues rapides. Les recherches françaises au sein d'HyMeX dans cette thématique ont deux objectifs principaux [Ducrocq et al., 2018] : l'amélioration des prévisions de précipitations aux courtes échéances (0-48h) basées sur le modèle opérationnel AROME de Météo-France et l'exploitation de ces prévisions pour forcer des modèles hydrologiques

[8] EFAS fait partie du programme d'observation de la Terre de l'Union Européenne Copernicus - Gestion des urgences, https://www.copernicus.eu/fr/services/urgences.

adaptés aux crues rapides. Le programme HyMeX a ainsi offert un cadre pour le développement d'AROME déterministe, de la prévision d'ensemble AROME (12 membres en 2016, 16 membres à partir de 2018) et pour le développement de AROME-PI (Prévision Immédiate, Auger et al. [2015]). Divers systèmes de prévision d'ensemble météorologiques à résolution kilométrique ont ensuite été évalués du point de vue hydrologique. Des modèles hydrologiques ont ainsi été forcés par les ensembles, générant des prévisions d'ensemble de débits jusqu'à 30 heures d'échéances. Le modèle ISBA-TOP [Bouilloud et al., 2010] a été forcé d'une part avec les membres des ensembles de précipitations construits par la méthode de perturbation *PERT* décrite précédemment (voir paragraphe 1.1.3.1), et d'autre part avec les membres de la PEAROME. Sur les bassins cévenols à l'automne 2014, ces deux méthodes montrent globalement des performances très proches. Par ailleurs, la combinaison des deux méthodes conduit à des scores probabilistes meilleurs que ceux obtenus par les deux méthodes prises séparément. Le cadre multidisciplinaire du projet HyMeX permet d'aborder des questions transverses, telles que la prévision intégrée des pluies intenses aux impacts. C'est justement l'objet du projet ANR PICS associé (Prévision immédiate Intégrée et Impacts des Crues Soudaines, Payrastre et al. [2018]) qui s'étend sur la période 2018-2021.

La première tâche du projet ANR PICS concerne la prévision des pluies et des débits à courte échéance (0-6h). Cette tâche est divisée en trois parties. La première concerne la préparation et l'évaluation des champs de pluie estimés et prévus. L'objectif est de fournir des scénarios de pluies prévues comportant plusieurs dizaines de membres (50 à 100) optimisés pour les courtes échéances (0-6h). Les prévisions d'ensemble AROME (12 membres) et les prévisions immédiates AROME-PI sont combinées. Des perturbations spatiales sont également introduites. Une amélioration assez nette par rapport à la prévision d'ensemble AROME est mise en évidence. La seconde tâche concerne l'amélioration des modèles hydrologiques pour les bassins de petite taille. Dans la troisième partie est estimé le bénéfice lié à la combinaison des produits de pluie prévue à très courte échéance avec des modèles hydrologiques distribués à haute résolution pour former des chaînes de prévision hydrométéorologiques. Des efforts sont portés sur la quantification et la réduction des incertitudes qui affectent ces systèmes de prévision hydrométéorologique. Enfin, le bénéfice de l'assimilation de données de débits mesurés pour réduire les incertitudes de ces prévisions sera évalué. Les autres tâches du projet ANR PICS concernent l'estimation des zones inondées, pour convertir les débits en une estimation des zones inondables et des hauteurs/vitesses d'eau associées, et la modélisation des impacts socio-économiques. La dernière tâche concerne l'intégration et l'expérimentation, qui consistera en particulier à tester les chaînes de prévision intégrées sur différentes études de cas. La large représentation des utilisateurs finaux potentiels dans le projet (au sein du groupe utilisateurs et des partenaires) facilitera le transfert des résultats du projet vers des applications opérationnelles.

Le projet CHROME (Couplage Hydrométéorologique RégiOnal Multi-Ensembles, 2012-2017) est issu d'une collaboration entre Météo-France et le réseau *Vigicrues* [de Saint-Aubin et al., 2018]. Il a pour objectif d'estimer les apports de la réalisation d'une prévision d'ensemble hydrologique, tenant compte à la fois des incertitudes liées à la prévision des précipitations et de celles liées à la modélisation hydrologique. Le but final est l'amélioration de la prévision opérationnelle pour la vigilance crues dans les bassin méditerranéens jaugés. L'incertitude sur la localisation et l'intensité des prévisions de pluie est prise en compte par la génération de 50 scenarios de pluie créés par perturbation du champ de précipitation prévu par le modèle AROME [Seity et al., 2011] selon la méthode *PERT*. L'incertitude sur la modélisation hydrologique est prise en compte grâce à une approche multi-modèles, en utilisant quatre modèles hydrologiques distribués fonctionnant en événementiel et sans assimilation de données de débits. Cette chaîne de prévision hydrologique d'ensemble est mise en œuvre sur trois bassins cévenols pour 17 épisodes de crue marquants de 2008 à 2014. Ce projet a mis en évidence que l'utilisation de prévisions de pluie constituées de plusieurs scenarios était préférable à l'utilisation d'une prévision déterministe, même si cela ne permet pas de corriger complètement le biais négatif des prévisions de pluie. Par ailleurs, il a été montré que l'approche multi-modèles est pertinente pour appréhender les incertitudes hydrologiques. Pour finir, cette étude a montré que les prévisions hydrologiques d'ensemble ainsi construites permettent d'anticiper la plupart des passages des seuils de vigilance crues, bien qu'ayant une sous-dispersion marquée.

D'autres études ont pour objet la prise en compte des incertitudes dans un contexte de crues éclair méditerranéennes. Vincendon et al. [2019] ont travaillé sur les incertitudes dans un système de prévision d'ensemble des crues rapides méditerranéennes, en partant du constat que les prévisions hydrométéorologiques sont affectées par plusieurs types d'incertitude. La plus importante est celle qui provient des précipitations, mais les incertitudes concernant l'humidité initiale du sol et le modèle hydrologique lui-même ne sont pas à négliger. Différents systèmes de prévision d'ensembles sont comparés. Ils sont construits à partir du modèle de prévision AROME à l'échelle kilométrique de Météo-France et du modèle hydrologique ISBA-TOP, dédié aux bassins méditerranéens à réponse rapide. L'incertitude sur les prévisions de pluie est prise en compte de deux manières : soit avec la prévision d'ensemble AROME, soit avec l'introduction de perturbations selon la méthode *PERT*. L'incertitude sur la modélisation hydrologique est prise en compte par perturbation des paramètres les plus sensibles d'ISBA-TOP et de l'humidité initiale du sol. Cette étude montre qu'un petit nombre de membres est suffisant pour améliorer les performances par rapport à une prévision déterministe de débits. Par ailleurs, elle confirme l'intérêt de prendre en compte les incertitudes qui affectent la chaîne de prévision complète. En effet, le système de prévision d'ensemble des débits basé sur la PEAROME et la version ensembliste d'ISBA-TOP permet d'obtenir des ensembles de débits prévus englobant en général

les observations, à condition de corriger la sous-estimation des prévisions de pluies de la PEAROME. Enfin, Demargne et al. [2019] utilisent les prévisions AROME-PI de Météo France. Ces prévisions ont 6 heures d'échéance et sont réactualisées toutes les heures. Afin de prendre en compte les incertitudes des prévisions, les prévisions déterministes successives d'AROME-PI sont utilisées comme prévisions d'ensemble. Combinées avec différents jeux de paramètres hydrologiques régionalisés, ces prévisions d'ensemble définissent des avertissements probabilisés de risque de crue. L'évaluation sur 13 événements entre octobre 2015 et juin 2017 pour 750 bassins versants montre une amélioration significative en termes de détection et anticipation en comparaison au système sans prévision de pluie et l'intérêt des avertissements de crue de type probabiliste.

La prise en compte des incertitudes de prévision hydraulique Comme nous l'avons vu, les prévisions d'ensemble hydrologiques, tenant compte à la fois des incertitudes sur les forçages et les modèles, se sont largement développées depuis les années 2000, que ce soit du point de vue de la recherche ou des applications opérationnelles. En revanche, la cascade des incertitudes météorologiques et hydrologiques au sein d'un modèle de prévision hydraulique a été moins investiguée. La synthèse des systèmes de prévision d'ensemble opérationnels effectuée en 2009 dans Cloke and Pappenberger [2009] ne cite que des systèmes basés sur une modélisation hydrologique d'ensemble. Depuis, quelques systèmes opérationnels se sont développés. La synthèse la plus récente proposée par Pappenberger et al. [2016] en 2016 cite le centre de prévision italien FEWSOO/ER[9], qui réalise de manière opérationnelle un chaînage de modèles hydrologiques et hydrauliques, le tout étant alimenté par des prévisions d'ensemble météorologiques [Tonelli et al., 2011]. Georgas et al. [2016] présentent le système de prévision d'ensemble opérationnel des inondations des zones urbaines des régions de New York et du New Jersey, appelé SFAS (Stevens Flood Advisory System). La prévision est réalisée par le chaînage de modèles hydrologiques, hydrauliques et d'hydrodynamique maritime. Il émet des prévisions à un horizon de 81 heures. Les prévisions météorologiques d'ensemble utilisées sont constituées de 125 membres issues des modèles du NCEP[10], du CEPMMT et du CMC[11]. Ce système est opérationnel depuis 2015.

En parallèle aux systèmes opérationnels, plusieurs études ont également été réalisées concernant la prise en compte des incertitudes hydrauliques. Dans Pappenberger et al. [2005], des prévisions d'ensemble sont réalisées à échéance 10 jours, en chaînant les prévisions d'ensemble du CEPMMT (50 membres perturbés, un membre de contrôle et un membre déterministe), le modèle hydrologique distribué LISFLOOD [De Roo et al., 2001, 2000] et le modèle hydraulique LISFLOOD-FP [Bates and De Roo, 2000]. Les incertitudes sur les

[9] Flood Early Warning System for the Po River and the Emilia Romagna Region
[10] National Centers for Environmental Prediction
[11] Canadian Meteorological Centre's Ensemble

modèles hydrologiques et hydrauliques sont prises en compte par l'utilisation de respectivement 6 et 10 jeux de paramètres des modèles optimisés avec la méthode GLUE. Schumann et al. [2013] utilisent une approche analogue avec le modèle hydrologique VIC (Variable Infiltration Capacity, Liang et al. [1994]) et le modèle hydraulique LISFLOOD-FP sur le bassin du Zambezi (Mozambique). Di Baldassarre et al. [2010] adoptent eux aussi une approche similaire dans un contexte de simulation et non de prévision, en générant une carte d'inondation probabiliste de la crue centennale avec le modèle LISFLOOD-FP. Les différentes configurations du modèle sont construites en conservant un ensemble de jeux de paramètres selon la méthode GLUE. Mueller et al. [2016] proposent quant à eux la construction de cartes d'inondations probabilistes pour l'événement de juin 2013 en Europe Centrale (Elbe et Danube). Les prévisions d'ensemble météorologiques IFS (Integrated Forecasting System) du CEPMMT, composées de 50 membres perturbés et un membre de contrôle, sont utilisées pour forcer le modèle hydrologique LISFLOOD utilisé dans la configuration EFAS. Les prévisions hydrologiques d'ensemble sont alors utilisées comme forçage du modèle hydraulique LISFLOOD-FP [Bates and De Roo, 2000, Bates et al., 2010]). Cette étude montre que 80% de la zone inondée le long de l'Elbe aurait pu être prévue 7 jours en avance. Buahin et al. [2017] proposent de générer des courbes de tarage par simulation numérique avec le modèle hydraulique 1D HEC-RAS (Hydrologic Engineering Center's River Analysis System, Brunner [2016]) utilisé en mode permanent pour une gamme de débits. L'extraction de la relation hauteur/débit aux sections de calcul permet de construire les courbes de tarage. Les prévisions de débit d'un modèle hydrologique sont ensuite utilisées pour interpoler les hauteurs d'eau à chaque section, d'où sont calculés les cartes d'inondation. L'utilisation de prévisions d'ensemble hydrologiques permet de produire des cartes d'inondation probabilistes. Gomez et al. [2019] génèrent des cartes de prévision d'inondation en chaînant les prévisions météorologiques d'ensemble GEFSRv2 [Hamill et al., 2013], le modèle hydrologique distribué HL-RDHM [Koren et al., 2004] et le modèle hydraulique HEC-RAS. Le système chaîné génère des prévisions d'ensemble au pas de temps horaire pour des échéances comprises entre 1 et 7 jours entre 2008 et 2013. Les hauteurs d'eau brutes ont été post-traitées par la méthode Quantile Regression (voir annexe N.2.2.2 pour une description de la méthode). Les résultats montrent que l'utilisation des ensembles météorologiques ainsi que le post-traitement des hauteurs d'eau améliorent tous deux les performances des prévisions à courte/moyenne échéance. Les prévisions d'ensemble hydrauliques brutes présentent de meilleures performances que les prévisions déterministes. La calibration statistique des ensembles bruts améliore les performances surtout pour des échéances supérieures à 3 jours.

1.2 La prévision opérationnelle des crues en France et la prise en compte des incertitudes

En France, deux acteurs majeurs réalisent des prévisions de débit ou de cotes dans les cours d'eau de manière opérationnelle. D'une part le réseau *Vigicrues*, constitué de services de l'État, est responsable de la vigilance crues sur le territoire. Ses objectifs sont la gestion de crise institutionnelle et l'information au public. D'autre part, les centres de prévision d'EDF-DTG[12] sont en charge de prévisions hydrométéorologiques opérationnelles pour la sûreté des grands ouvrages hydrauliques et l'optimisation de l'usage de la ressource en eau. Quelles que soient les finalités recherchées, les chaînes opérationnelles visent un objectif commun [Berthet et al., 2019] : fournir une prévision fiable et la plus riche possible à des décideurs en différents points d'un réseau hydrographique. Pour réaliser cet objectif, les prévisions opérationnelles réalisées par ces deux organismes affichent une information quantitative sur l'incertitude associée aux prévisions.

La prévision opérationnelle au sein du réseau *Vigicrues* est présentée au paragraphe 1.2.1. Elle consiste généralement à chaîner un modèle de prévision hydrologique avec un modèle de prévision hydraulique, tel que présenté précédemment (voir figure 1.2). La prévision opérationnelle à EDF-DTG, présentée au paragraphe 1.2.2, est réalisée par chaînage de prévisions météorologiques d'ensemble avec un modèle hydrologique [Moulin et al., 2019]. Enfin, le paragraphe 1.2.3 présente la synthèse de ces deux systèmes de prévision opérationnels français.

1.2.1 La prévision opérationnelle au sein du réseau *Vigicrues*

1.2.1.1 Historique de la prévision des crues

Ce paragraphe concernant l'historique de la prévision des crues en France est en partie rédigé à partir du Schéma Directeur de Prévision des Crues du bassin Loire-Bretagne [DREALCentre, 2012]. C'est à la suite des crues exceptionnelles de 1856 et 1866 qu'a été initiée en France l'annonce des crues, avec la création des Services d'Annonces de Crues (SAC). Suite à des crues très importantes à la fin des années 1990 et au début des années 2000, l'organisation des services de l'État dans ce domaine a été modifiée par la circulaire du 30 octobre 2002 et par la loi du 30 juillet 2003 sur les risques naturels et technologiques. Cette loi affirme que l'organisation de la surveillance, de la prévision et de la transmission de l'information sur les crues est assurée par l'État, ses établissements publics et, le cas échéant, les collectivités territoriales ou leurs groupements.

Les 52 SAC sont désormais remplacés par 19 Services de Prévision des Crues (SPC) dont

[12] Électricité de France - Direction Technique Générale

le périmètre est élargi. Ils sont situés au sein des DREAL[13]. L'objectif n'est plus l'annonce des crues, mais leur prévision, ce qui implique une plus grande anticipation des événements. Au niveau national, le Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations (SCHAPI) est créé en 2003. Ce service central à compétence nationale du MTE[14]-DGPR[15] est situé dans le Météopôle de Toulouse. Il a pour objectif d'assurer la coordination opérationnelle, scientifique et technique de la prévision des crues et de l'hydrométrie. Nous reviendrons plus en détail par la suite sur les rôles respectifs des SPC et du SCHAPI.

L'organisation de la prévision des crues est désormais définie pour chacun des six bassins hydrographiques français (Artois-Picardie, Seine-Normandie, Rhin-Meuse, Loire-Bretagne, Adour-Garonne, Rhône-Méditerranée-Corse) par un Schéma Directeur de Prévision des Crues (SDPC). Celui-ci fixe les principes selon lesquels s'effectue la prévision des crues et l'organisation de la transmission de l'information. Il détermine également le territoire de compétence de chaque SPC situé dans le bassin. Les SPC représentent des unités hydrographiques homogènes, assemblées de manière à obtenir une taille critique nécessaire pour faire face aux exigences de la prévision des crues, sans dépasser une taille maximale pour garder le contact avec les réalités du territoire. La carte 1.4 représente les bassins hydrographiques français métropolitains et les territoires associés à chacun des 19 SPC. L'organisation de la surveillance, de la prévision et de la transmission de l'information sur les crues par l'État fait l'objet de règlements arrêtés par le préfet. Ces règlements sont appelés *RIC*, pour Règlement de surveillance, de prévision et de transmission d'Information sur les Crues. Les SPC sont responsables de la production opérationnelle de la vigilance crues sur leur territoire, selon leur périmètre d'intervention défini dans le RIC.

Plus récemment, la circulaire interministérielle [16] du 28 avril 2011, adressée aux préfets de département, poursuit sur les voies tracées par les réformes de 2005, et demande l'organisation au niveau départemental, au sein des Direction Départementale des Territoires (et de la Mer) - DDT(M) - d'une mission de référent départemental pour l'appui technique à la préparation et la gestion des risques d'inondation, complémentaire à l'action du SPC. C'est la mission RDI (Référent Départemental Inondation). L'institution d'un relais départemental doit permettre d'améliorer sensiblement le service rendu en la matière, notamment de connaissance sur les conséquences locales des crues prévues, connaissance particulièrement précieuse pour l'organisation et la gestion des secours.

[13]Direction Régionale de l'Environnement, de l'Aménagement et du Logement
[14]Ministère de la Transition Écologique
[15]Direction Générale de la Prévention des Risques
[16]Ministère de l'écologie, du développement durable, des transports et du logement et Ministère de l'intérieur, de l'outre-mer, des collectivités territoriales et de l'immigration

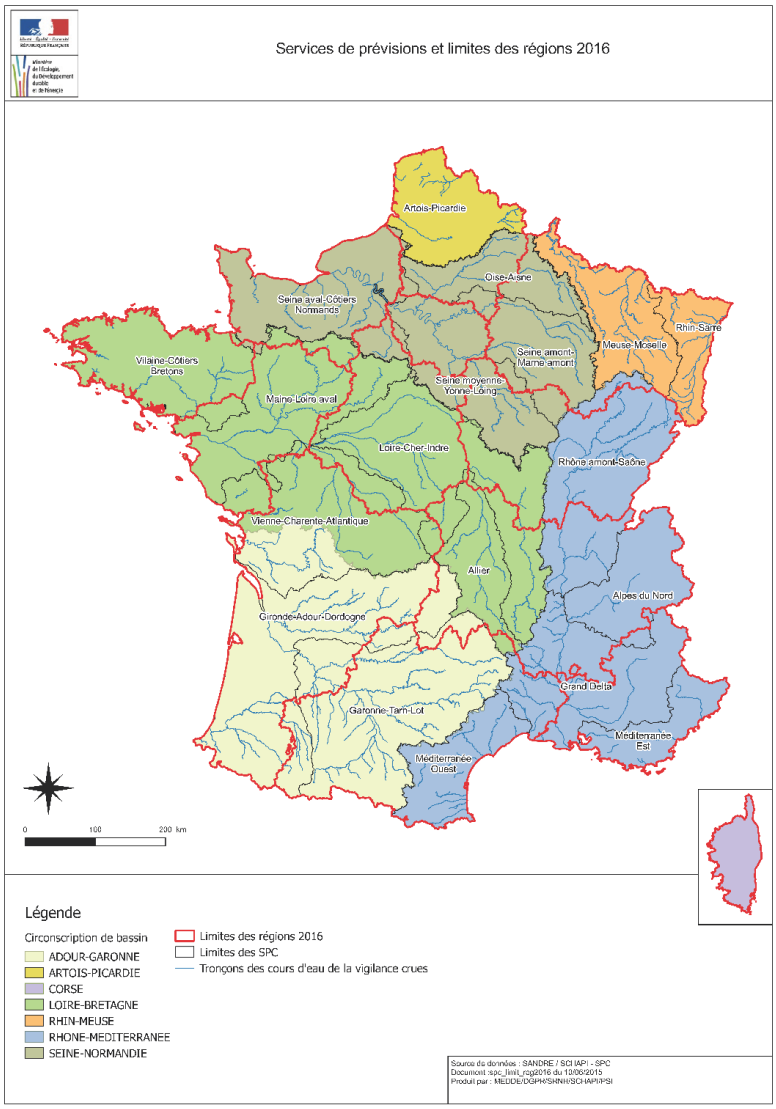


Figure 1.4: Bassins hydrographiques en France métropolitaine et territoires de compétence des 19 SPC - D'après DGPR-SRNH.

1.2.1.2 Le réseau *Vigicrues*

Depuis juillet 2006, la prévision des crues en France est donc désormais assurée par le réseau *Vigicrues*, composé des acteurs suivants (source Site du MTE, Prévision des inondations) :

- le SCHAPI ;
- les 19 services de prévision des crues,
- les 25 unités hydrométriques (UH) ;
- les cellules de veille hydrologique (CVH) dans les DOM.

Le réseau *Vigicrues* effectue en particulier la veille hydrométéorologique, et surveille à ce titre 22 800 km de cours d'eau sur le territoire français, aux abords desquels se situent plus de 75 % de la population qui vit ou travaille en zone inondable. Il est responsable de la transmission des niveaux d'eau et des débits en temps réel sur environ 1500 points de mesure, ainsi que leurs variations au cours des derniers jours. Il émet des prévisions sur les évolutions des niveaux et des débits à venir dans les 24 prochaines heures ou parfois davantage. Le réseau émet des bulletins de vigilance deux fois par jour (à 10h et à 16h) destinés aux particuliers et aux professionnels. Ils sont en particulier destinés aux acteurs de la gestion de crise en charge de la sécurité civile (préfets, maires, ...), qui déclenchent les alertes et mobilisent les moyens de secours en fonction du risque de débordement sur les cours d'eau surveillés par l'État. En période d'événement de crue, la mise à jour de la vigilance peut être plus fréquente. Les cours d'eau inclus dans le dispositif de vigilance apparaissent sur la carte de vigilance accessible sur le site *Vigicrues*. Ils sont découpés en tronçons, et à chaque tronçon est affectée une couleur (verte, jaune, orange, rouge) selon le niveau de vigilance adapté pour faire face au risque de crue susceptible de se produire dans les 24 heures à venir :

- vert : pas de vigilance particulière requise ;

- jaune : risque de crue ou de montée rapide des eaux n'entraînant pas de dommages significatifs, mais nécessitant une vigilance particulière dans le cas d'activités saisonnières et/ou exposées ;

- orange : risque de crue génératrice de débordements importants susceptibles d'avoir un impact significatif sur la vie collective et la sécurité des biens et des personnes ;

- rouge : risque de crue majeure. Menace directe et généralisée de la sécurité des personnes et des biens.

Des cartes locales sont également disponibles. On peut alors cliquer sur les stations de vigilance et accéder ainsi au graphique des dernières hauteurs d'eau ou débits mesurés à la station. Des niveaux de crue de référence peuvent également être affichés. Une illustration de ces graphiques sera donnée plus loin, dans le paragraphe relatif à la prise en compte et à l'affichage des incertitudes de prévision (paragraphe 1.2.1.3, figure 1.7).

La figure 1.5 présente à titre d'illustration la carte de vigilance crues émise le 24 décembre 2013 à 10h00. Les tronçons du bassin de l'Odet (Bretagne Sud, Finistère), qui correspond au bassin d'étude, sont alors en vigilance orange. On trouvera en annexe L.1 le bulletin de vigilance correspondant émis par le SCP Vilaine et Côtiers Bretons. La carte de vigilance météorologique mise en place par Météo-France en 2001 intègre la vigilance crues : en cas de passage en vigilance jaune, orange ou rouge, des pictogrammes 'Pluie-Inondation' ou 'Inondation' sont ajoutés sur les départements concernés, dont la couleur est déterminée par la vigilance météorologique. À titre d'illustration, la figure 1.6 présente la carte de vigilance météorologique émise le 24 décembre 2013 à 10h00. Le département du Finistère

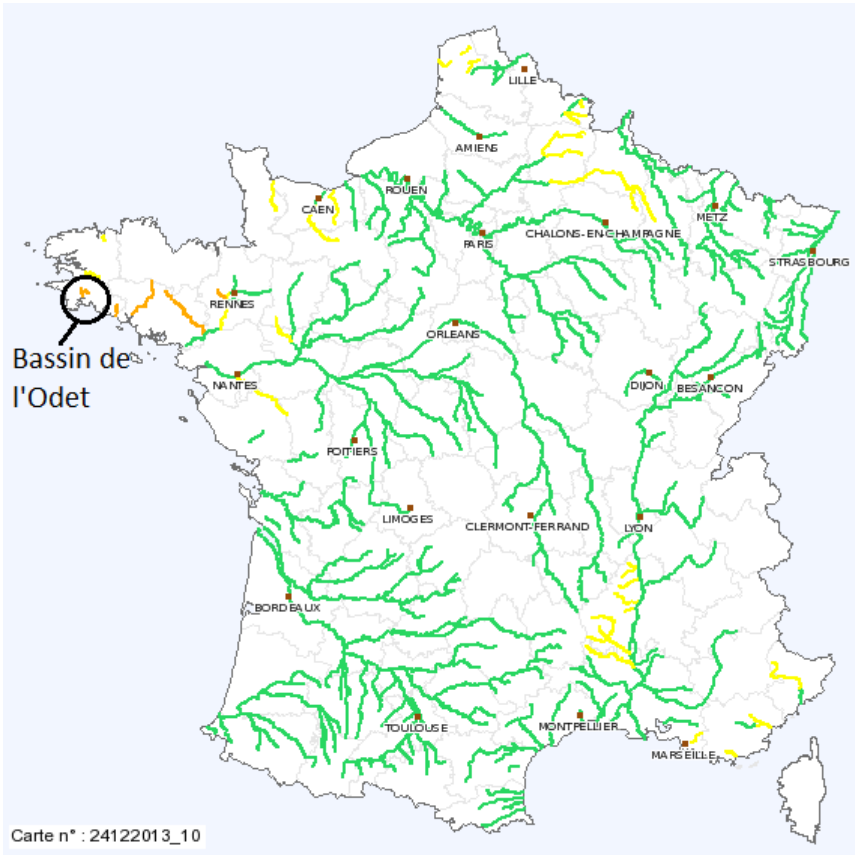


Figure 1.5: Carte de vigilance crues émise le 24 décembre 2013 à 10h sur le site *Vigicrues*. Les tronçons situés dans le bassin de l'Odet sont en vigilance orange.

est en vigilance pour le risque inondation.

Au sein du réseau *Vigicrues*, les SPC et le SCHAPI ont chacun des rôles propres clairement identifiés. Les SPC proposent une couleur de niveau de vigilance par tronçon spécifié au RIC, prévoient la cote maximale de crues à 24 heures aux stations de vigilance et rédigent des bulletins d'information qui accompagnent les cartes locales. En dehors de la surveillance hydrométéorologique opérationnelle, les SPC assurent également des tâches d'ingénierie comme la mise en place et l'amélioration des modèles de prévision hydrologiques ou hydrauliques. Ils ont également en charge la cartographie des Zones Inondées Potentielles (ZIP) : pour une hauteur donnée à une station de vigilance, la carte ZIP correspondante représente l'emprise de l'inondation sur la zone d'influence de la station. Ces cartes permettent ainsi de connaître les enjeux susceptibles d'être impactés pour un niveau de crue prévu. Les cartes ZIP sont produites pour différents scénarios, des premiers débordements à la crue de référence. Ces cartes sont destinée en premier lieu aux acteurs de la gestion de crise (COGIC[17], COZ[18] préfectures, DDT(M)[19], collectiv-

[17]Centre Opérationnel de Gestion Interministérielle des Crises
[18]Centre Opérationnel de Zone
[19]Direction Départementale des Territoires (et de la Mer)

ités). Elles sont également mises à disposition des services des ministères de l'écologie et de l'intérieur. Elles ont vocation à terme à être diffusées au grand public sur le site *Vigicrues*. Pour finir, les SPC assurent la maintenance et la mise à jour de la base nationale de données hydrométriques (Banque hydro).

Le SCHAPI quant à lui a principalement en charge deux domaines dans la chaîne de production de l'information de vigilance sur les crues : la publication des informations *Vigicrues* et l'appui technique aux SPC en matière d'outils de prévision. En ce qui concerne la publication des informations *Vigicrues*, le SCHAPI intègre et harmonise les informations produites par les différents SPC, l'objectif étant d'assurer une cohérence nationale de l'information. Cela l'amène à échanger fréquemment avec les SCP, en particulier lors de passages en vigilance jaune, orange ou rouge. Il est responsable de la production de la carte nationale de vigilance « inondation » et rédige le bulletin de vigilance joint à la carte, le tout étant diffusé sur le portail *Vigicrues*. Ce portail s'adresse d'une manière générale à tous les acteurs et au grand public. En parallèle, il communique spécifiquement des informations aux acteurs institutionnels et opérationnels de la gestion de crise au niveau national (COGIC[20], CNIR[21], DGPR, Météo-France), régional (CRICR[22], DIR[23]) ou départemental (préfectures et SDIS[24]). En dehors de la vigilance crues, le SCHAPI a pour mission d'assurer un appui technique aux SPC dans le champ de l'hydrologie et de la technique de prévision des crues. Il suit le développement et la maintenance d'outils de

[20] Centre Opérationnel de Gestion Interministérielle des Crises
[21] Centre National d'Information Routière
[22] Centre Régional d'Information et de Circulation Routières
[23] Direction Interrégionale des Routes
[24] Service Départemental d'Incendie et de Secours

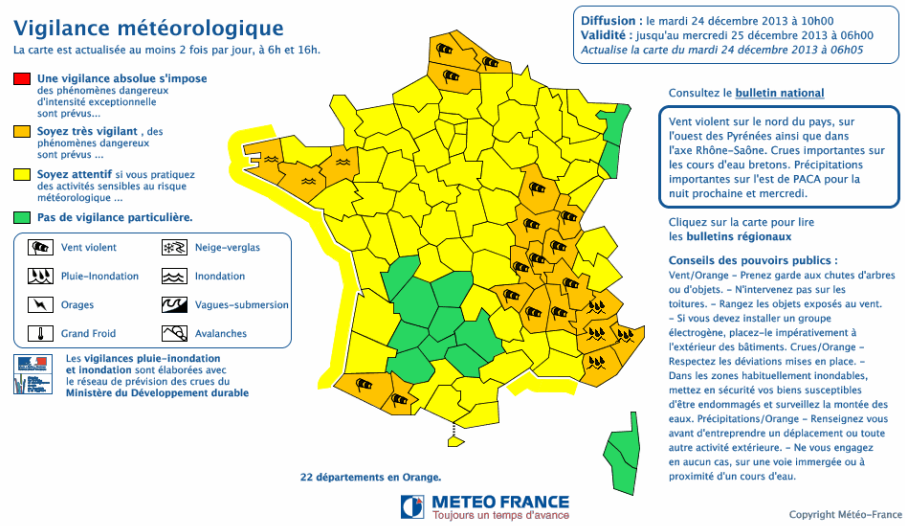


Figure 1.6: Carte de vigilance météorologique du 24 décembre 2013 à 10 heures. Le département du Finistère est en vigilance pour le risque inondation.

modélisation d'intérêt commun, et conseille les SPC dans leurs choix. Il est responsable de la veille scientifique et technique dans le domaine de la prévision des crues. Il anime le réseau et propose des formations pour les agents des SPC. Il est également responsable de la publication du BHYM (bulletin hydrométéorologique quotidien) diffusé à tous les SPC à 11h30.

1.2.1.3 La prise en compte des incertitudes de prévision

Dans les années 2012-2014, dans le cadre de l'amélioration du dispositif de prévision des crues en France, deux groupes de travail regroupant des membres du réseau *Vigicrues* et des experts scientifiques ont été mis en place pour mener une réflexion sur l'estimation et la communication des incertitudes associées aux prévisions hydrologiques ([Berthet, Lionel et al., 2016]). Cette réflexion avait pour objectif le développement d'outils à visée opérationnelle pour l'affichage sur le site *Vigicrues* des hauteurs/débits prévus ainsi que les incertitudes associées. La forme de ces graphes est fixée pour apporter une information accessible au grand public et utile pour la gestion de crise : la prévision est ainsi représentée par une tendance centrale, encadrée par deux tendances correspondant aux quantiles q_{10} et q_{90}. La production et l'affichage des prévisions graphiques sur *Vigicrues* accompagnées de la visualisation de l'incertitude est effective depuis octobre 2017, grâce en particulier à l'utilisation de deux outils :

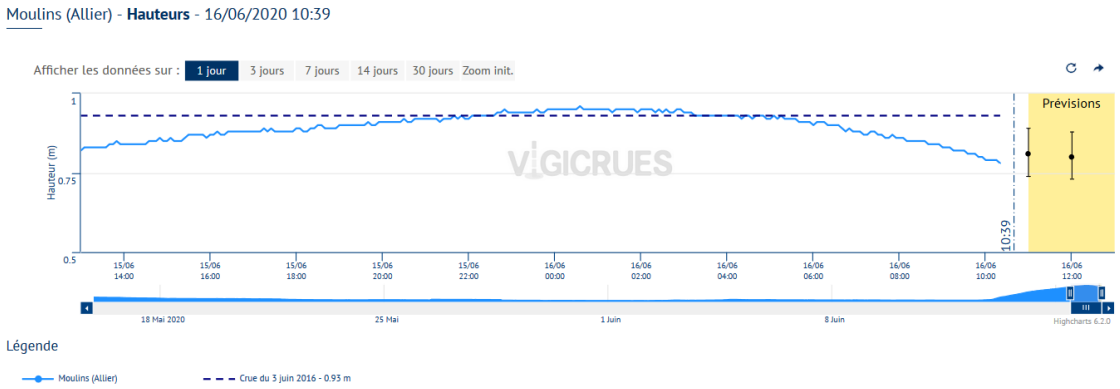
- OTAMIN (OuTil Automatique d'estiMation des INcertitudes prédictives) : outil d'estimation automatique et calibrée des incertitudes de modélisation, développé par l'INRAE[25] avec le soutien du SCHAPI, qui permet d'associer des intervalles de confiance à des prévisions déterministes. Cet outil peut être utilisé aussi bien pour des sorties de modèle hydrologique et hydraulique. Il repose sur la méthode QUOIQUE (QUantification Opérationnelle des Incertitudes par QUantiles de dEbit) développée dans [Bourgin, 2014, Viatgé, Julie et al., 2019a]. Elle se base sur l'analyse *a posteriori* des erreurs du modèle, calculées en fonction de l'échéance de prévision et du débit simulé, pour prendre en compte les différentes structures d'erreurs en fonction des gammes de débit. Ainsi, en opérationnel, en fonction du débit prévu par le modèle et de l'échéance de prévision, l'outil fournit les quantiles q_{10} et q_{90} de prévision.

- EAO (Expertise Assistée par Ordinateur) : les intervalles de confiance générés par l'outil OTAMIN sont ensuite expertisés par le prévisionniste, selon sa connaissance du bassin et de l'incertitude qu'il connaît sur les prévisions météorologiques. L'outil EAO lui permet de modifier graphiquement les trois tendances de prévision (basse, centrale et haute), correspondant aux quantiles de non-dépassement (respectivement

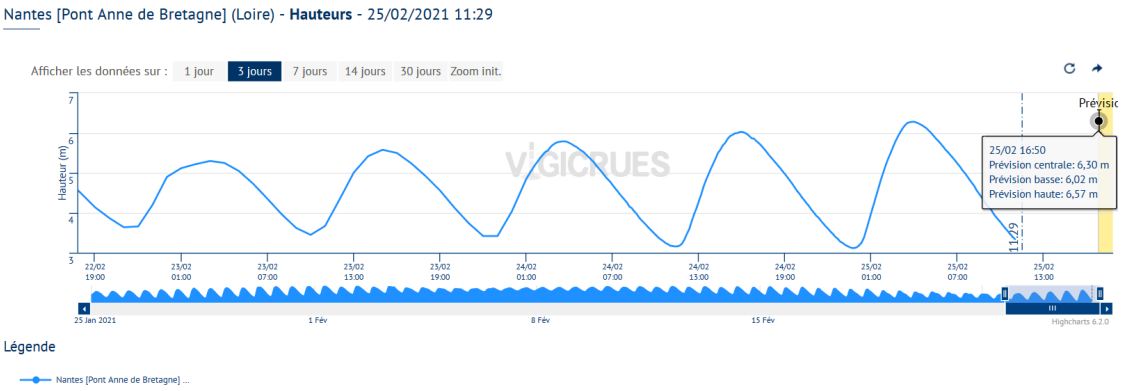
[25] Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement

q_{10}, q_{50} et q_{90}) de la variable prévue (débit ou hauteur) et de générer ainsi la prévision qui sera publiée.

L'affichage des prévisions sur *Vigicrues* est adaptée aux types de stations [Janet et al., 2018]. Pour les stations sous influence fluviale, les prévisions sont constituées de plusieurs points de prévision, comme illustré sur la figure 1.7a. En revanche, pour les stations sous influence maritime, seules les valeurs de pleine mer sont représentées. La figure 1.7b illustre ce cas pour des prévisions à Nantes.



(a) Moulins, prévision du 16 juin 2020. La hauteur de crue de référence du 3 juin 2016 est indiquée. Les prévisions sont constituées de plusieurs points de prévision.



(b) Nantes, prévision du 25 février 2021. La prévision concerne le pic de marée haute uniquement.

Figure 1.7: Visualisation de la chronique de hauteur mesurée et de la prévision pour deux stations : Moulins (régime fluvial) et Nantes (régime maritime). Les tendances de prévision basse, centrale et haute sont indiquées. D'après *Vigicrues*.

1.2.1.4 Le système Vigicrues-Flash

Depuis mars 2017, le système *Vigicrues Flash* permet de compléter le dispositif du système national de vigilance *Vigicrues* décrit précédemment. En effet, les cours d'eau à réaction rapide ne peuvent pas être gérés au sein du dispositif *Vigicrues*. Ce temps de réaction est incompatible avec une anticipation du risque de crues dans les 24 heures à venir. Ce

dispositif couvre environ 30 000 km de cours d'eau non surveillés dans le cadre de la vigilance. Étant donné le caractère automatique de la prévision, des critères d'éligibilité des bassins sont définis : bassins peu influencés par la neige, par les barrages, des temps de réponse supérieurs à 1h30 et la disponibilité des mesures radar de pluie de qualité suffisante. Il propose des avertissements automatiques sur le risque de crue. Il s'adresse aux communes, préfectures et acteurs de la gestion de crise qui peuvent s'y abonner. Ces acteurs sont avertis par SMS, téléphone et courriel en cas de risque de crue dans les heures à venir sur leur territoire. Ils peuvent alors consulter en ligne la cartographie des avertissements. Ce système repose sur une modélisation hydrologique semi-distribuée automatique réalisée avec le modèle GRD développé par IRSTEA [Javelle, 2014, Javelle et al., 2016], paramétré de manière régionale. Il est alimenté par la lame d'eau Panthère produite par Météo-France toutes les 5 minutes à partir de radars et permet d'estimer le débit en tous points d'intérêt des cours d'eau. Ces débits sont comparés à des quantiles de débits préalablement calculés à partir de chroniques simulées avec le même modèle. Les avertissements, mis à jour tous les quarts d'heure, indiquent un risque de crue forte ou de crue très forte sur chaque tronçon. Pour plus d'information sur ce dispositif, le lecteur pourra se reporter à Garandeau et al. [2018].

1.2.2 La prévision opérationnelle à EDF

1.2.2.1 Description de la chaîne opérationnelle de prévision

La prévision hydrologique à EDF a pour objectif d'optimiser la production électrique, tout en assurant la sûreté par une bonne gestion du passage des crues et la gestion de la ressource en eau par le respect de la réglementation en termes de débit réservé ou de variation de débit. Depuis les années 1970, EDF a ainsi développé en interne deux centres hydrométéorologiques de prévision, situés respectivement à Grenoble et à Toulouse. Ils fournissent des prévisions de manière régulière et assistent ponctuellement les exploitants pour l'anticipation et la gestion des crues. Les horizons de prévision sont adaptés aux enjeux : de quelques heures pour les phénomènes rapides comme les crues, quelques jours pour l'optimisation de la production hydroélectrique, à une prévision saisonnière pour la gestion des stocks des retenues et la prévision des étiages [Moulin et al., 2019].

La réalisation de la prévision opérationnelle à EDF est basée sur l'utilisation du modèle hydrologique MORDOR[26] déterministe semi-distribué (SD) [Garavaglia et al., 2017]. Le modèle MORDOR est un modèle pluie-débit conceptuel à réservoirs fonctionnant de manière continue [Garçon, 1996]. La structure de ce modèle sera décrite en détail au chapitre 2. Dans la prévision opérationnelle à EDF, l'expertise humaine est présente à chaque maillon de la chaîne de prévision ([Garçon et al., 2009], [Moulin et al., 2019]). Les

[26]Modèle à Réservoirs de Détermination Objective du Ruissellement

données d'entrée du modèle hydrologique MORDOR-SD sont tout d'abord expertisées, que ce soit pour la partie réanalyse ou la partie prévision. Les données observées (pluie, température, débits, hauteur de neige) permettent dans un premier temps de rejouer le modèle jusqu'à l'instant de prévision. Celles-ci sont préalablement expertisées à l'aide de tests statistiques de cohérence spatiale et temporelle, en croisant différentes sources disponibles : lame d'eau RADAR, pluviomètres, stations de mesures voisines. Les données manquantes ou aberrantes sont remplacées par des données plus réalistes. La suite du chapitre présente l'étape de prévision hydrologique et la prise en compte des incertitudes de prévision.

1.2.2.2 La prise en compte des incertitudes de prévision

Jusqu'en 2010, les prévisions opérationnelles étaient des prévisions de débits moyens journaliers déterministes, avec un horizon de prévision pouvant aller jusqu'à 6 jours. À partir de 2010, la chaîne opérationnelle a évolué vers une prévision probabiliste journalière pour prendre en compte l'incertitude sur les précipitations et températures futures. En complément des prévisions météorologiques des modèles AROME et ARPEGE de Météo-France sont utilisées les prévisions d'ensemble IFS-ENS du CEPMMT. Elles comportent 51 membres (50 membres et un *run* de contrôle, non utilisé pour les prévisions hydrologiques). Ce modèle propose des prévisions jusqu'à 14 jours d'échéance. Cet ensemble est corrigé avec des prévisions par analogues météorologiques [Courbariaux, 2017]. Cela permet de corriger la sous-dispersion des prévisions météorologiques. L'utilisation du modèle hydrologique MORDOR-SD est associée à un outil de recalage opérationnel qui utilise les dernières données de débit observées ainsi que la nature et la dynamique des erreurs passées (analyse par comparaison des séries simulées et mesurées sur une longue période). Il est alors lancé avec en entrée les 50 membres météorologiques. Les 50 prévisions de débit sont alors à leur tour post-traitées par gamme de débit et d'échéance, à partir d'une description empirique des erreurs. La connaissance approfondie qu'ont les prévisionnistes sur les outils utilisés leur permet de produire une prévision hydrologique expertisée qui est alors mise à disposition des utilisateurs finaux par l'intermédiaire d'une plate-forme intranet. Les informations sont communiquées de plusieurs manières : sous forme d'ensembles (trajectoires possibles), des synthèses ou encore des informations concernant les probabilités de dépassement de seuil [Berthet et al., 2019].

En parallèle au développement de la prévision d'ensemble journalière, la modélisation horaire se développe de plus en plus. Cette modélisation a l'intérêt de mieux prévoir les pics et gradients de crue. Elle permet également une meilleure estimation du stock de neige (meilleure estimation de la constitution du stock en cas d'un refroidissement en cours d'épisode et meilleure estimation de la fonte au printemps au pas de temps infra-journalier). Dans l'attente de la mise en place d'une véritable prévision d'ensemble au

pas de temps horaire, de manière analogue à ce qui existe au pas de temps journalier, EDF a élaboré une chaîne de prévision multi-scénarios (scénarios déterministes AROME, ARPEGE et IFS à haute résolution). L'objectif recherché à court terme est de converger vers une seule chaîne de prévision, horaire et probabiliste.

1.2.3 Conclusion

Les chaînes de prévision du réseau *Vigicrues* et d'EDF-DTG présentent plusieurs points communs [Berthet et al., 2019]. Elles sont toutes deux basées sur des modèles déterministes : hydrologique pour EDF-DTG, hydrologique et hydraulique pour *Vigicrues*. Les sorties sont ensuite post-traitées avec des méthodes similaires. De plus, l'expertise finale du prévisionniste est indispensable dans les deux chaînes de prévision. Elle permet en particulier d'assurer la cohérence des prévisions. Enfin, EDF et le réseau *Vigicrues* étudient actuellement des premières solutions visant à améliorer intrinsèquement la cohérence spatio-temporelle des prévisions.

Les deux chaînes de prévision ont pour principale différence l'utilisation ou non de prévision d'ensemble météorologique. EDF utilise des ensembles issus du centre européen dont chaque membre présente une structure temporelle respectant la dynamique des phénomènes hydrométéorologiques. Le réseau *Vigicrues* quant à lui décrit l'incertitude prédictive à partir d'une collection de quantiles de débits en fonction des différents horizons de prévision, qui ne constituent pas des trajectoires vraisemblables.

1.3 Objectifs de la thèse et méthodologie suivie

Au cours de ce premier chapitre, nous avons détaillé la chaîne de prévision hydrométéorologique opérationnelle et exploré les différentes sources d'incertitude qui affectent la prévision. Le travail de thèse présenté dans la suite de ce manuscrit a pour objectif d'explorer la cascade des incertitudes le long de la chaîne de prévision hydrométéorologique. L'objectif est de mettre en place et de démontrer les mérites d'une chaîne ensembliste hydrologie-hydraulique pour la vigilance réalisée au sein du réseau *Vigicrues*. Pour ce faire, nous allons mettre en place et étudier deux méthodologies pour créer un ensemble représentatif de l'incertitude prédictive :

- la première méthode consiste à générer des prévisions d'ensemble brutes en perturbant les paramètres incertains des modèles, puis à leur appliquer un post-traitement statistique pour les corriger ;

- la seconde méthode consiste à générer des prévisions d'ensemble en propageant en prévision des membres corrigés par assimilation de données ensembliste.

Ces deux méthodes sont appliquées aux deux maillons de la chaîne de prévision hy-

dométéorologique : pour réaliser des prévisions d'ensemble hydrologiques et hydrauliques. Les ensembles hydrologiques sont utilisés pour forcer les ensembles hydrauliques. Les ensembles hydrologiques et hydrauliques générés sont analysés de manière rigoureuse d'un point de vue statistique, et le gain par rapport à une prévision déterministe est évalué.

La démonstration est réalisée sur le bassin de l'Odet situé dans le Finistère. Deux modélisations hydrologiques sont mises en place, l'une avec le modèle hydrologique GRP [Perrin et al., 2003, Viatgé, Julie et al., 2019b], l'autre avec le modèle hydrologique MORDOR-TS [Garavaglia et al., 2017, Garçon, 1996, Rouhier et al., 2017]. Les prévisions hydrologiques alimentent un modèle hydraulique de prévision 1D, réalisé avec le code MASCARET [Goutal, 2012]. La description du bassin versant de l'Odet ainsi que les données et modèles utilisés sont présentés au chapitre 2.

Les deux méthodologies développées nécessitent de connaître les variables déterminantes pour les débits hydrologiques ou hauteurs d'eau modélisés. Dans ce but, une étude de sensibilité globale des modèles hydrologiques et hydrauliques est préalablement menée. Les différentes méthodologies mises en œuvre dans ce travail (étude de sensibilité, calibration statistique et assimilation de données) ainsi que les caractéristiques recherchées pour les ensembles construits sont présentées au chapitre 3.

Les méthodologies mises en œuvre pour générer des ensembles sont appliquées aux prévisions hydrologiques et hydrauliques. Suite aux résultats des études de sensibilité présentées respectivement aux chapitres 4 et 7 pour les modèles hydrologiques et hydrauliques, des prévisions d'ensemble brutes sont générées et calibrées par la méthode des forêts aléatoires. Cette étape est décrite dans le chapitre 4 pour les prévisions d'ensemble hydrologiques, et dans le chapitre 8 pour les prévisions d'ensemble hydrauliques. En parallèle, des ensembles sont construits par assimilation de données. La mise en place d'assimilation de données est d'abord analysée grâce à la réalisation d'expériences jumelles, décrites au chapitre 5 pour le modèle hydrologique MORDOR-TS et au chapitre 9 pour le modèle hydraulique MASCARET. Enfin, les résultats des prévisions d'ensemble réalisées avec les deux types de méthodes sont comparés en opérationnel. La comparaison pour les prévisions d'ensemble hydrologiques est réalisée au chapitre 6, tandis que la comparaison pour les prévisions d'ensemble hydrauliques est présentée au chapitre 10.

Chapitre 2

Le bassin versant de l'Odet, données et modèles

Sommaire

2.1 Présentation générale du bassin de l'Odet 40
2.1.1 Le bassin versant de l'Odet 40
2.1.2 La vigilance crues sur le bassin de l'Odet 43
2.1.3 La modélisation du bassin de l'Odet dans le cadre de la thèse . . 46
2.2 Les données d'observation et de prévision 47
2.2.1 Les données de pluie et de température 47
2.2.2 Les mesures hydrométriques 48
2.2.3 Les prévisions maritimes 50
2.3 La modélisation du bassin de l'Odet 51
2.3.1 La modélisation hydrologique 51
2.3.1.1 Le modèle hydrologique GRP 52
2.3.1.2 Le modèle hydrologique MORDOR-TS 54
2.3.2 La modélisation hydraulique 59
2.3.2.1 Le code MASCARET 59
2.3.2.2 Le modèle MASCARET du bassin de l'Odet aval 61

Les méthodes de génération de prévisions d'ensemble hydrologiques et hydrauliques développées dans cette thèse sont analysées dans le cadre de l'application au bassin de l'Odet situé dans le Finistère. Le paragraphe 2.1 présente le bassin de l'Odet et la vigilance crues opérationnelle effectuée sur ce bassin par le Service de Prévision des Crues *Vilaine et Côtiers Bretons*. Le paragraphe 2.2 présente les données observées nécessaires au calage des modèles de prévision ainsi que les historiques de prévisions météorologiques permettant de recréer des prévisions en temps différé. Enfin, le paragraphe 2.3 présente les modèles hydrologiques et hydrauliques utilisés ainsi que leur calage sur le bassin de l'Odet réalisé dans le cadre de ce travail.

2.1 Présentation générale du bassin de l'Odet

2.1.1 Le bassin versant de l'Odet

La figure 2.1 présente le bassin versant de l'Odet. C'est un bassin littoral situé en Bretagne dans le département du Finistère. Sa superficie est d'environ 720 km^2. La longueur de l'Odet jusqu'à son exutoire maritime est d'environ 60 km. Ses deux principaux affluents sont le Jet et le Steïr dont les confluences avec l'Odet sont situées dans l'agglomération de Quimper. La photo de la figure 2.2 illustre la confluence de l'Odet et du Steïr. Après avoir traversé Quimper, l'Odet s'écoule vers le sud en direction de la mer. Il se jette dans l'océan Atlantique au niveau de la cité balnéaire de Bénodet.

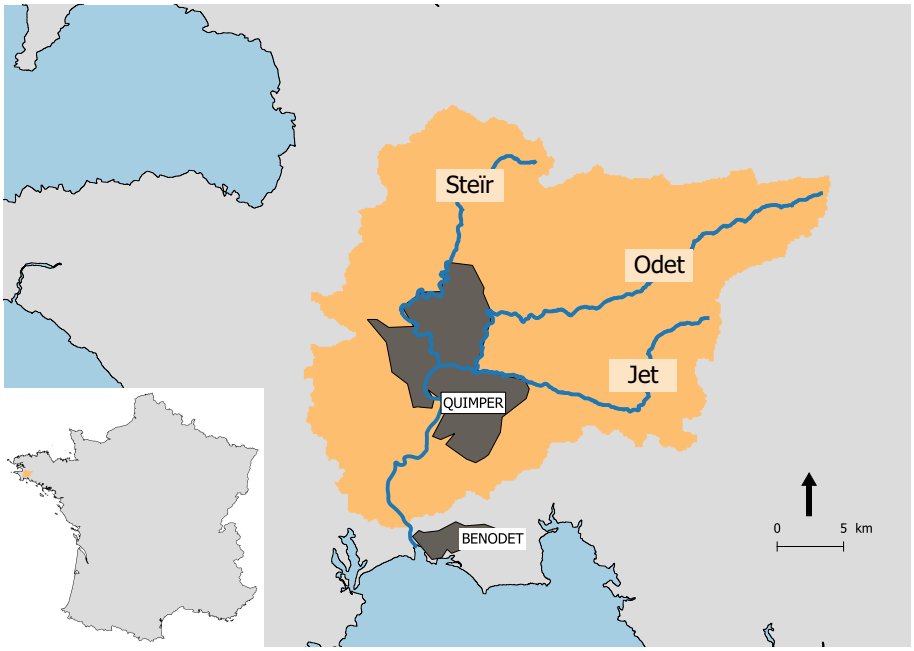


Figure 2.1: Présentation du bassin versant de l'Odet. L'Odet a deux principaux affluents : le Jet et le Steïr.

Figure 2.2: Confluence du Steïr et de l'Odet dans le centre-ville de Quimper.

Les caractéristiques hydrologiques des cours d'eau du bassin de l'Odet sont liées à la relative imperméabilité du sous-sol qui freine l'infiltration des pluies. Le débit des cours d'eau est ainsi directement influencé par les précipitations, avec des valeurs très variables d'une année à l'autre[1]. Dans un contexte de pluies hivernales quasi-continues, une pluie de 40 mm peut ainsi provoquer une crue débordante. Les pointes de crue peuvent être très élevées (plus de 100 m^3/s pour l'Odet, environ 80 m^3/s pour le Steïr et 50 m^3/s pour le Jet en décembre 2000). Mises en regard avec les capacités des lits mineurs dans l'agglomération de Quimper (environ 60 m^3/s sur l'Odet après sa confluence avec le Jet, et 50 m^3/s sur le Steïr), cela explique des crues relativement fréquentes dans Quimper. Les inondations de l'Odet ont lieu dans la période hivernale, généralement entre Octobre et Mars. Elles sont en premier lieu déclenchées par une forte pluviométrie sur le bassin, mais d'autres facteurs influencent la formation des crues sur ce bassin :

- saturation des sols suite à des précipitations durables ;
- faible capacité de stockage de l'eau du fait du substrat géologique imperméable ;
- forme des bassins versants responsable d'une vitesse élevée de propagation des crues ;
- marées hautes de fort coefficient à l'aval parfois amplifiées par des surcotes d'origine météorologiques ;
- direction des vents en hiver, majoritairement de secteur Ouest et Sud-Ouest, qui peut entraîner des gênes temporaires de l'écoulement de l'Odet de l'agglomération vers la zone littorale.

[1] d'après le Plan de Prévention des Risques d'Inondation (PPRI) - Communes de Quimper, Ergué-Gaberic et Guengat disponible ici.

Une importante crue de l'Odet et de ses affluents s'est produite en 1925. Elle a longtemps servi de crue de référence jusqu'au milieu du XX$^{\text{ième}}$ siècle. La période de retour de cette crue a été estimé de 50 à 100 ans. Dans la seconde moitié du XX$^{\text{ième}}$ siècle, une quinzaine de crues importantes sont répertoriées, dont les plus importantes sont les suivantes :

- Février 1974 : la période de retour est estimée entre 20 et 50 ans sur l'Odet, entre 10 et 20 ans sur le Steïr.

- Janvier 1995 : la période de retour est estimée entre 10 et 20 ans sur l'Odet et le Steïr.

- Décembre 2000 : cette crue est supérieure en moyenne de 20 cm à la crue historique de 1925. La période de retour est estimée entre 50 et 100 ans pour l'Odet et le Steïr. 175 hectares et 400 habitations ont été inondés. Le coût des dégâts occasionnés aux biens communaux est de 2.2 millions d'euros, et de 12 millions d'euros pour les particuliers[2].

- Hiver 2013-2014 : la tempête *Dirk* du 22 au 24 décembre 2013 et les dépressions *Gerhard* du 31 décembre 2013 au 7 janvier 2014 et *Qumeira* début février 2014 ont été à l'origine de précipitations exceptionnelles qui ont provoqué la crue de nombreux cours d'eau bretons et des inondations dans plusieurs sites, en particulier dans le Finistère. La période de retour de ces événements est d'environ 20 ans sur le bassin de l'Odet. Ces événements étant les premiers significatifs depuis la mise en place du Service de Prévision des Crues *Vilaine et Côtiers Bretons* (SPC-VCB) en 2006, ils ont été utilisés pour évaluer le fonctionnement global du système de prévision des crues à travers la mise en place d'une mission d'expertise[3].

À titre d'illustration, la figure 2.3 représente un repère de la crue de décembre 2000 sur le Steïr, environ 200 m à l'amont de sa confluence avec l'Odet. Les zones inondables du Steïr et de l'Odet définies par le Plan de Prévention des Risques d'Inondation couvrent environ 1000 locaux commerciaux ou industriels et 24 000 foyers.

[2] d'après le PPRI

[3] Mission d'expertise sur les crues de décembre 2013 à février 2014 en Bretagne. Disponible sur le site du ministère de l'intérieur.

(a) Vue d'ensemble de la position du repère de crue

(b) Zoom sur le repère de crue

Figure 2.3: Repère de la crue de décembre 2000 sur le Steïr, 200 m environ avant sa confluence avec l'Odet.

2.1.2 La vigilance crues sur le bassin de l'Odet

Le territoire de compétence du Service de Prévision des Crues *Vilaine et Côtiers Bretons* (SPC-VCB) a été fixé par le Schéma Directeur de Prévision des Crues Loire-Bretagne en 2012. Il couvre les quatre départements bretons à l'exclusion du bassin de l'Elorn (Finistère). Il est responsable en particulier de la surveillance du bassin de l'Odet. Pour la vigilance, celui-ci est subdivisé en quatre sous-bassins présentés sur la figure 2.4 :

- le sous-bassin amont correspondant à l'Odet dont l'exutoire est Tréodet ;
- le sous-bassin amont correspondant au Jet dont l'exutoire est Kerjean ;
- le sous-bassin amont correspondant au Steïr dont l'exutoire est Ty-Planche ;
- la partie aval de l'Odet dans l'agglomération de Quimper jusqu'à l'estuaire à Plaisance.

Les trois sous-bassins amont dont les exutoires sont Tréodet, Kerjean et Ty-Planche sont couverts par une modélisation hydrologique, tandis que le sous-bassin aval est couvert par une modélisation hydraulique 1D forcée par les prévisions hydrologiques. Cette dernière permet de prévoir les hauteurs d'eau aux stations de vigilance du SPC-VCB qui sont situées dans l'agglomération de Quimper et dont la localisation sera présentée à la figure 2.5.

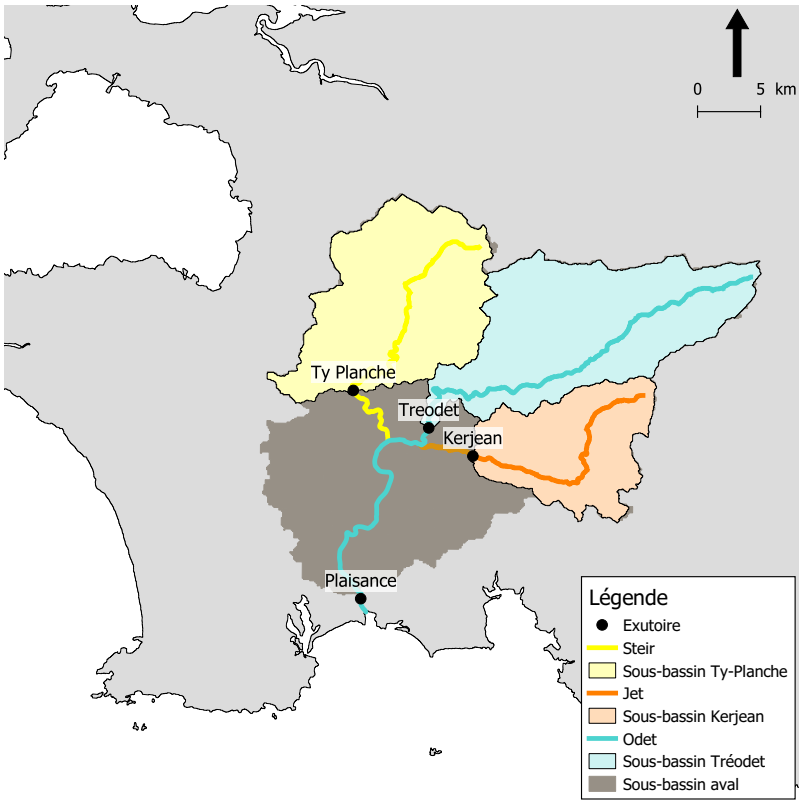


Figure 2.4: Bassin versant de l'Odet. Visualisation des trois sous-bassins amonts de l'Odet, du Jet et du Steïr, et de la partie aval couvrant l'agglomération de Quimper jusqu'à l'exutoire maritime.

Modélisation hydrologique Le tableau 2.1 présente les particularités physiques et les valeurs caractéristiques des débits des trois sous-bassins amonts.

Sous-bassin	Tréodet	Kerjean	Ty Planche
Altitude de la source	175 m	200 m	100 m
Longueur de la rivière dans le sous-bassin (km)	37	21	23
Surface du sous-bassin (km^2)	205	107	179
Débit moyen (m^3/s)	4.8	2.27	3.79
Débit journalier décennal (m^3/s)	55	19	39
Débit journalier cinquantenal (m^3/s)	75	25	53
Débit instantané crue 12-2000 (m^3/s)	110	46.6	81
Pluie moyenne annuelle (mm/yr)	743	672	671

Tableau 2.1: Caractéristiques des trois sous-bassins Tréodet, Kerjean et Ty-Planche. Données calculées sur la période 1969-2019 à Tréodet, 1966-2019 à Kerjean et 1976-2019 à Ty-Planche.

Pour la réalisation de la prévision opérationnelle, Météo-France met à disposition des services de prévision des crues des prévisions de pluie à l'échelle des bassins versants au pas de temps 24 heures jusqu'à échéance de 72 heures dans ce qui est appelé *BP* (Bulletin de Précipitations). Dans le BP, les prévisions sont donc identiques pour les trois sous-

bassins de l'Odet. Ces prévisions de pluie font l'objet d'une expertise par Météo-France avant leur diffusion, à partir de la connaissance des erreurs des modèles météorologiques et de la comparaison entre les observations et les dernières prévisions. Des scénarios de pluie minimale et maximale sont proposés. L'annexe M.1 reproduit à titre d'illustration un bulletin de précipitations émis par Météo-France en mars 2021. En parallèle, Météo-France produit des prévisions expertisées de pluie nommées *RR3* discrétisées au pas de temps tri-horaire (voir paragraphe 2.2). Ces prévisions sont réalisées sur des zones supposées homogène appelées *zones symposium*. Elles ne sont pas directement faites à l'échelle des bassins versants, une interpolation est nécessaire pour estimer la lame d'eau sur un bassin.

À partir des scénarios *min* et *max* du bulletin de précipitations, le SPC-VCB estime un scénario médian. Les cumuls quotidiens sont projetés sur le scénario infra-journalier *RR3* pour générer ces scénarios à un pas de temps plus fin, nécessaire pour l'alimentation des modèles hydrologiques. Les quatre scénarios *min*, *max*, *médian* et *RR3* sont alors utilisés en forçage des modèles hydrologiques GRP [Perrin et al., 2003, Viatgé, Julie et al., 2019b] utilisés au pas de temps horaire. L'incertitude de prévision de chaque scénario est estimée automatiquement grâce à l'utilitaire OTAMIN (voir paragraphe 1.2.1.3). Le prévisionniste choisit ensuite le scénario qu'il juge le plus pertinent. Il adapte la prévision à publier sur le site *Vigicrues* si nécessaire (pic de crue, incertitude, temporalité de l'événement).

Modélisation hydraulique Les prévisions hydrologiques issues du scénario médian du BP sont ensuite utilisées pour forcer un modèle hydraulique 1D MASCARET [Goutal, 2012] dont le linéaire est représenté sur la figure 2.5. À l'aval, ce modèle est alimenté par des prévisions de niveaux d'eau. On additionne pour cela les prévisions de marée astronomique réalisées par le SHOM[4] et les prévisions de surcote produites par Météo-France (modèle de surcote HYCOM du SHOM forcé avec ARPEGE). Cette modélisation chaînée des trois modèles hydrologiques et du modèle hydraulique fournit ainsi des prévisions de hauteur d'eau pour les trois stations de vigilance située dans l'agglomération de Quimper dont les localisations sont visibles sur la figure 2.5 :

- la station Kervir sur l'Odet ;
- la station Moulin-Vert sur le Steïr ;
- la station Palais de Justice sur l'Odet, située peu après sa confluence avec le Steïr, appelée dans la suite du travail *Justice*.

Le modèle hydraulique du bassin de l'Odet a été mis en service en 2017. Auparavant, le SPC-VCB utilisait uniquement des outils statistiques basés sur des régressions pour estimer les hauteurs aux trois stations de vigilance. Aujourd'hui, ces outils statistiques viennent en complément de la modélisation hydraulique. Les incertitudes de prévision sont évaluées à dire d'expert et communiquées sur le site *Vigicrues*.

[4]Service Hydrographique et Océanographique de la Marine

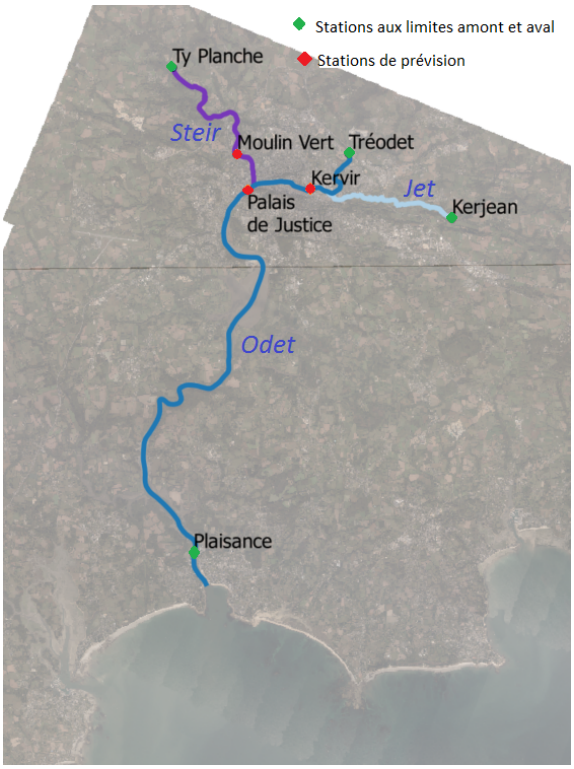


Figure 2.5: Bassin versant aval de l'Odet couvert par une modélisation hydraulique 1D avec le modèle MASCARET.

2.1.3 La modélisation du bassin de l'Odet dans le cadre de la thèse

Ce travail de thèse s'appuie sur la prévision opérationnelle en place au SPC. Une modélisation hydrologique est ainsi réalisée sur chacun des trois sous-bassins amont. Dans ce travail, deux modèles hydrologiques ont été utilisés : le modèle hydrologique GRP utilisé en opérationnel par le SPC-VCB comme mentionné ci-avant, et le modèle hydrologique MORDOR-TS [Garavaglia et al., 2017, Garçon, 1996, Rouhier et al., 2017] utilisé pour la prévision à EDF. Ces deux modèles ont été utilisés pour évaluer l'apport d'une modélisation semi-distribuée pour la prévision opérationnelle. La partie aval du réseau est modélisée avec le modèle 1D MASCARET. La géométrie du modèle opérationnel a été simplifiée pour une meilleure stabilité numérique. Suite à ces modifications, il a fait l'objet d'un recalage. Les modèles hydrologiques et le modèle hydraulique ainsi que leur calage sur le bassin de l'Odet seront présentés en détail au paragraphe 2.3, notamment après avoir présenté au paragraphe 2.2 les données d'observation nécessaire au calage des modèles.

2.2 Les données d'observation et de prévision

Dans ce paragraphe, nous présentons les données observées et les historiques de prévisions utilisés pour se placer dans un contexte de prévision opérationnelle. Les données observées servent d'abord au calage des modèles. Dans un contexte de prévision, elles sont également utilisées pour initialiser les modèles jusqu'au temps de base[5]. Les historiques de prévisions servent quant à eux à alimenter les modèles à partir du temps de base, afin de générer des prévisions dans un contexte opérationnel.

2.2.1 Les données de pluie et de température

Observations Météo-France met à disposition des services de prévision des crues la lame d'eau observée ANTILOPE (ANalyse par spaTIaLisation hOraire des PrEcipitations, Champeaux et al. [2009]). Ces données sont au pas de temps horaire, avec une résolution spatiale de 1 km^2. La lame d'eau ANTILOPE est obtenue par fusion des données radar et des pluviomètres. Les pluviomètres sont situés au sol et fournissent des observations ponctuelles des précipitations. Les radars météorologiques émettent des ondes électromagnétiques dans l'atmosphère qui sont réfléchies par les précipitations. Le signal reçu en retour est analysé pour estimer des cumuls de pluie. La calibration des radars est en particulier réalisée grâce aux observations des postes pluviométriques. La température de surface utilisée est issue de la réanalyse SAFRAN [Vidal et al., 2010]. Ces données sont au pas de temps horaire, avec une résolution spatiale de 8 km. Les données ANTILOPE et SAFRAN sont disponibles sur les trois sous-bassins amont de l'Odet à partir de 2007.

Prévisions En prévision, deux types de produits sont utilisés : les prévisions d'ensemble PEAROME et les prévisions RR3, que nous décrivons ci-après.

Dans ce travail de thèse, les prévisions d'ensemble PEAROME [Brousseau et al., 2011, Seity et al., 2011] sont utilisées pour décrire l'incertitude sur les pluies lors de la réalisation des études de sensibilité globale des modèles hydrologiques GRP et MORDOR-TS. Le maillage du modèle AROME couvre 1800 × 1700 km^2, englobant l'Irlande, l'Allemagne, le nord du Portugal et la Sicile. La prévision d'ensemble PEAROME est constituée de 12 simulations AROME équiprobables. Chaque membre est généré par une perturbation du modèle AROME, de manière à représenter les incertitudes sur les conditions initiales, les conditions aux limites, les conditions de surface et l'incertitude sur le modèle lui-même. Chaque membre est forcé à ses frontières par un membre de prévision d'ensemble ARPEGE (PEARP, Nicolau [2002]). Les 12 membres coupleurs de la PEARP sont sélectionnés automatiquement parmi les 34 membres disponibles. La PEAROME est opérationnelle

[5]Pour mémoire, le temps de base correspond au moment du lancement de la prévision. Pour plus d'informations, le lecteur pourra se reporter au paragraphe 1.1.1.1.

depuis fin 2016. Pour la réalisation d'une partie de ce travail (études de sensibilité globale des modèles GRP et MORDOR-TS), des données antérieures ont été reconstruites pour couvrir des événements sur le bassin de l'Odet (du 24 décembre 2015 au 15 mars 2016 et du 19 mai 2016 au 16 juin 2016). Chaque jour à 21 heures, une nouvelle prévision est disponible. Les prévisions sont au pas de temps horaire, sur une profondeur de 45 heures. La maille horizontale est de 2,5 km.

Les prévisions PEAROME disponibles au démarrage de la thèse et qui servent à la réalisation des études de sensibilité ne couvrent que 112 jours. Par ailleurs, depuis la production opérationnelle de la PEAROME, il n'y a pas eu d'événement significatif sur le bassin de l'Odet. Il n'a donc pas été possible d'utiliser la PEAROME pour la génération des prévisions d'ensemble hydrologiques car la période de disponibilité de la PEAROME pendant laquelle des événements significatifs sont survenus est trop petite. Ainsi, les données de prévision utilisées pour générer les prévisions d'ensemble hydrologiques sont les prévisions *RR3* disponibles de 2010 à 2019 inclus. Ce sont des prévisions expertisées par les prévisionnistes de Météo-France, déterministes au pas de temps tri-horaire. Comme nous l'avons vu, elles sont utilisées en opérationnel par les SPC.

2.2.2 Les mesures hydrométriques

L'hydrométrie a pour objet la mesure des hauteurs d'eau et des débits dans les cours d'eau. Le débit s'exprime en m^3/s et représente le volume d'eau s'écoulant en une seconde à travers une section de la rivière. La mesure de la hauteur d'eau est réalisée par la lecture d'une échelle limnimétrique. Elle s'exprime en mètres (m) par rapport à l'origine de l'échelle. En connaissant la cote NGF (Nivellement Général de la France) du zéro de l'échelle, on transforme facilement la hauteur d'eau en cote d'eau, exprimée en mètres également. Le suivi des débits et hauteurs d'un cours d'eau est réalisé à partir d'un réseau de stations hydrométriques. Si seule la hauteur d'eau est mesurée, on parle alors de station limnimétrique (observations discontinues) ou limnigraphiques (enregistrement continu).

La mesure de la hauteur d'eau Pour mesurer la hauteur d'eau, les stations sont composées d'un limnigraphe permettant l'enregistrement continu des hauteurs (sonde à pression, flotteur). Une échelle limnimétrique permet la lecture directe de la hauteur et le contrôle de l'appareil de mesure. Les données sont télétransmises en temps réel grâce à une carte SIM ou modem installé dans la centrale d'acquisition. Le schéma représentant le fonctionnement d'une station limnigraphique est représenté en figure 2.6. La figure 2.7 présente les échelles limnimétriques des stations Tréodet et Kerjean.

La mesure du débit Les observations de débit sont obtenues par conversion des hauteurs mesurées grâce à une courbe de tarage. Celle-ci associe un débit à chaque hau-

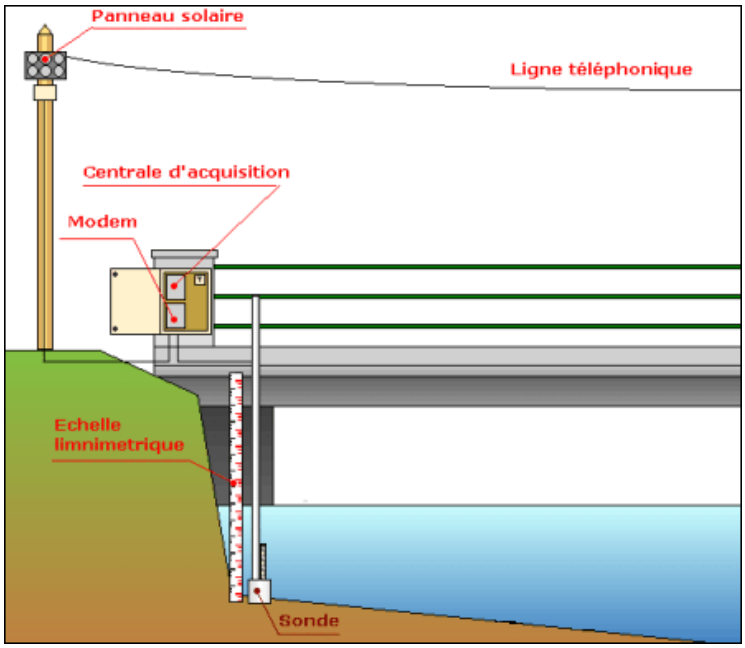


Figure 2.6: Schéma de fonctionnement d'une station limnigraphique. D'après DREAL Centre.

(a) Tréodet

(b) Kerjean, vue vers l'aval. L'échelle est en rive gauche, la centrale d'acquisition est en rive droite.

Figure 2.7: Stations hydrométriques de Tréodet et Kerjean.

teur d'eau mesurée ($Q = f(H)$). Pour construire une courbe de tarage, un ensemble de jaugeages doit être réalisé au droit de la station. Un jaugeage est une mesure du débit instantané pour une hauteur lue à l'échelle. La courbe de tarage est interpolée à partir du nuage de points de jaugeages. Elle est souvent extrapolée à ses extrémités. La courbe de tarage est naturellement propre à chaque station hydrométrique. Elle peut varier dans le temps, après le passage d'une crue par exemple. Des jaugeages réguliers sont donc nécessaires pour différentes gammes de débit pour détecter un éventuel détarage de la station.

Données hydrométriques disponibles pour le travail de thèse Pour ce travail, nous avons à disposition :

- les mesures des débits au pas de temps horaire aux trois stations amont Tréodet, Kerjean et Ty-Planche respectivement à partir des années 1970, 1966 et 1975. La station Ty-Planche a été déplacée en mai 2016 suite à la formation d'un renard au droit de la station. Au début du travail de thèse, nous ne disposions donc pas de mesures de débit de bonne qualité à Ty-Planche après 2016, puisqu'il faut du temps pour la construction d'une nouvelle courbe de tarage. Nous nous limitons donc aux données mesurées avant le déplacement de la station.
- les mesures de hauteur d'eau à la station aval Plaisance et aux trois stations de vigilance Kervir, Moulin-Vert et Justice au pas de temps 6 minutes. Les données sont disponibles à partir de 2006 pour toutes les stations sauf pour la station Justice qui a été mise en service fin août 2012.

2.2.3 Les prévisions maritimes

La vigilance crues réalisée sur les bassins littoraux doit prendre en compte les prévisions de marée et de surcote, car le niveau d'eau en mer conditionne les hauteurs d'eau dans les estuaires et sur la partie aval du cours d'eau. Comme nous le verrons plus tard (paragraphe 2.3.2.2), la station Justice est très influencée par le niveau de la mer. Les hauteurs d'eau mesurées à Justice correspondent à celles d'un marégramme.

La marée est un phénomène lié à l'attraction de la Lune (responsable pour $\frac{2}{3}$) et du Soleil (responsable pour $\frac{1}{3}$). C'est un phénomène déterministe mais non strictement périodique. Sur les côtes françaises, on observe environ deux cycles de marée par jour (période de 12 heures et 25 minutes par cycle). Selon la position relative de la Lune et du Soleil, l'amplitude des marées est plus ou moins importante. On parle alors de marées de *morte eau* ou de *vive eau*. On compte deux marées de morte eau et de vive eau par mois. En France, le SHOM est chargé de prévoir les marées astronomiques. Pour ce travail, les prévisions de marée astronomique sont disponibles sur la période 2010-2020.

Le niveau de la mer n'est pas défini uniquement par le niveau de la marée astronomique. En effet, les phénomènes météorologiques modifient les niveaux d'eau. On parle alors de *surcotes* ou *décotes météorologiques*. Les surcotes se produisent en présence de dépression. Elles sont dues à l'effet barométrique inverse, lié à la modification de la pression atmosphérique lors de la dépression (le niveau d'eau s'adapte au poids de l'atmosphère qu'il supporte), à un effet dynamique lié à la vitesse de déplacement de la dépression et à une surélévation des masses d'eau dues au vent. Lors d'événements maritimes importants, les surcotes peuvent atteindre des valeurs très importantes. Le marégraphe de La Pallice à La Rochelle a enregistré une surcote de pleine mer de 1.53 m lors du passage de la tempête Xynthia en février 2010.

La prévision opérationnelle sur le bassin de l'Odet intègre les prévisions de surcote réalisées par Météo-France. Pour ce travail cependant, nous n'avons pas eu à disposition un historique de prévisions de surcotes à Plaisance. En prévision, nous considérerons donc uniquement les prévisions de marée astronomique, à laquelle nous ajouterons la surcote mesurée au temps de base.

2.3 La modélisation du bassin de l'Odet

Ce paragraphe est dédié à la description de la modélisation mise en place sur le bassin de l'Odet. Le paragraphe 2.3.1 est consacré à la description des deux modèles hydrologiques GRP et MORDOR-TS, tandis que le paragraphe 2.3.2 décrit le modèle hydraulique MASCARET.

2.3.1 La modélisation hydrologique

Deux modèles hydrologiques sont utilisés pour modéliser les trois sous-bassins amont de l'Odet : le modèle GRP et le modèle MORDOR-TS. Ils sont tous les deux construits et calés spécifiquement dans le cadre de ce travail. Ce sont deux modèles hydrologiques conceptuels à réservoirs. Le modèle non distribué GRP est développé par INRAE[6] en collaboration avec le SCHAPI[7] pour la prévision des crues. Le modèle semi-distribué MORDOR-TS est développé par EDF-DTG[8] pour l'optimisation de la production électrique. Les deux modèles fonctionnent de manière continue et sont utilisés au pas de temps horaire. Leur module neige n'est pas activé du fait de la localisation côtière du bassin d'étude. Ils nécessitent tous les deux une période d'initialisation de 1 an. Les algorithmes de calage des deux modèles sont différents, mais dans les deux cas le calage est réalisé selon une procédure de calage-validation. Les paramètres sont d'abord optimisés

[6] Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement, anciennement IRSTEA

[7] Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations

[8] Électricité de France - Direction Technique Générale

sur une période de calage et ensuite évalués sur une période distincte de validation. Nous considérons les périodes de calage et de validation suivantes :

- période de calage P_{Cal} : entre le 1er janvier 2007 et le 31 décembre 2013 ;
- période de validation P_{Val} : entre le 1er janvier 2014 et le 31 mai 2017 (jusqu'au 31 mai 2016 pour Ty-Planche car la station a été déplacée à cette date).

La période de calage contient une dizaine d'événements d'importance moyenne à forte. La période de validation est plus courte et en dehors des événements de janvier à mars 2014, elle contient des événements plus modérés. Le paragraphe 2.3.1.1 présente le modèle hydrologique GRP et les résultats du calage de ce modèle sur les trois sous-bassins de l'Odet. Le paragraphe 2.3.1.2 est analogue pour le modèle MORDOR-TS.

2.3.1.1 Le modèle hydrologique GRP

Description du modèle Le modèle hydrologique GRP dérive du modèle GR4J [Perrin et al., 2003, Viatgé, Julie et al., 2019b]. Il est largement utilisé par les services de prévision des crues (SPC) au sein du réseau national de vigilance *Vigicrues*. La structure du modèle GRP est présentée en figure 2.8. Ce modèle est forcé par la pluie horaire cumulée et par l'évapotranspiration potentielle, estimée à partir de la température par la formule d'Oudin [Oudin, 2004, Oudin et al., 2005]). Il est composé d'un module de production et d'un module de routage. Il a trois paramètres libres à caler. Ils sont résumés dans le tableau 2.2. Dans le module de production, la quantité totale d'eau allant alimenter l'écoulement est corrigée par le facteur *Corr*. Le module de routage permet de transférer la pluie efficace par l'intermédiaire de l'hydrogramme unitaire de temps de base T_B[9] et d'un réservoir de routage de capacité *Rout*.

Le modèle opérationnel inclut une étape d'assimilation du dernier débit observé [Berthet, 2010] qui corrige le contenu en eau du réservoir de transfert, ainsi qu'une exploitation de l'erreur du modèle [Tangara, 2005]. Ces deux fonctionnalités ont été désactivées pour ce travail. En effet, l'étude de sensibilité a pour objectif de se focaliser sur le modèle seul, ses paramètres et son forçage. Par ailleurs, un des objectifs de la thèse étant l'introduction d'assimilation de données ensembliste dans la chaîne de prévision, il n'était pas possible de conserver l'assimilation déterministe présente dans le code opérationnel GRP. Une description complète du modèle est présentée en annexe M.2.

[9] Attention, le paramètre *temps de base* du modèle hydrologique GRP n'a rien à voir avec le temps de base d'une prévision opérationnelle, qui correspond au moment de lancement de la prévision (voir paragraphe 1.1.1.1).

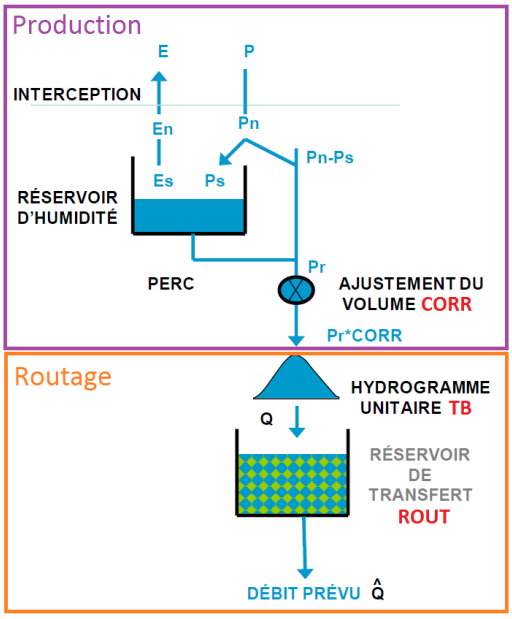


Figure 2.8: Schéma fonctionnel du modèle hydrologique GRP sans assimilation de données - D'après IRSTEA [2017].

Paramètre	Unité	Rôle
Corr	-	Coefficient correctif pour les précipitations
T_B	heures	Temps de base de l'hydrogramme unitaire
Rout	mm	Capacité du réservoir de routage

Tableau 2.2: Bilan des paramètres à caler dans le modèle GRP.

Calage du modèle Le calage du modèle GRP sur les trois sous-bassins amont de l'Odet est réalisé avec un utilitaire spécifiquement créé par INRAE pour travailler sans le module d'assimilation de données. La fonction objectif minimisée est l'erreur quadratique moyenne RMSE (*Root Mean Squarred Error*) :

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(Q_{obs}(t) - Q_{prev}(t))^2} \tag{2.1}$$

où n représente le nombre de pas de temps testés, $Q_{obs}(t)$ le débit observé au temps t et $Q_{sim}(t)$ le débit simulé par le modèle au temps t. L'optimisation des paramètres est réalisée pour les pas de temps pour lesquels les observations sont supérieures au quantile q_{60}. Le tableau 2.3 présente les valeurs des paramètres calés sur la période de calage P_{Cal} pour chacun des trois sous-bassins et tableau 2.4 synthétise les scores obtenus pour les débits simulés par le modèle sur les périodes P_{Cal} et P_{Val}. Nous notons que les scores sont du même ordre de grandeur sur les deux périodes à Tréodet et Kerjean. En revanche, l'écart entre les périodes de calage et de validation est plus important à Ty-Planche.

Bassin versant	Corr	T_B (heures)	Rout (mm)
Tréodet	1.01	7.26	8433.78
Kerjean	0.98	0.5	28001.12
Ty Planche	0.96	1.70	13629.62

Tableau 2.3: Paramètres calés pour le modèle GRP sans assimilation de données sur les sous-bassins de l'Odet

Bassin versant	Tréodet		Kerjean		Ty Planche	
Période	P_{Cal}	P_{VCal}	P_{Cal}	P_{Val}	P_{Cal}	P_{Val}
RMSE sup q_{60} (m3/s)	3.03	3.32	1.20	1.39	1.97	2.95

Tableau 2.4: Scores de calage du modèle GRP sans assimilation de données sur les sous-bassins de l'Odet.

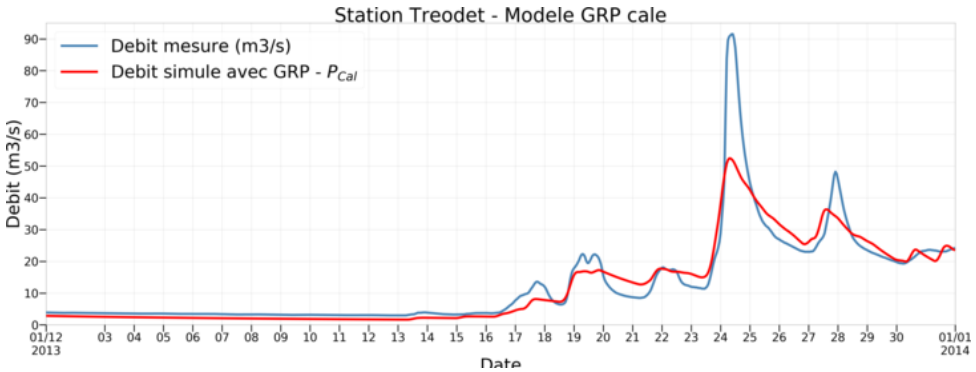
Graphiques des débits mesurés et simulés La figure 2.9 illustre les débits simulés par le modèle GRP à la station Tréodet pour le mois de décembre 2012 (période de calage, courbe rouge) et de janvier 2014 (période de validation, courbe orange). La courbe bleue représente la mesure. Le lecteur pourra se reporter à l'annexe M.3 pour visualiser les graphes pour les mois de décembre 2012, et de décembre 2013 à février 2014 pour les trois stations Tréodet, Kerjean et Ty-Planche. Ces graphiques mettent en évidence que le modèle GRP peine à reproduire les pics de crue. Nous rappelons qu'il est utilisé ici sans assimilation de données, donc en-dehors du cadre opérationnel.

2.3.1.2 Le modèle hydrologique MORDOR-TS

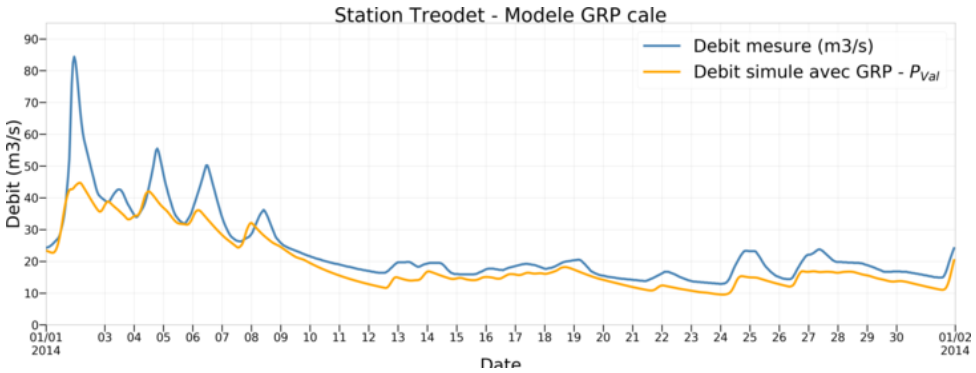
Description du modèle Le modèle hydrologique MORDOR-TS[10] [Garavaglia et al., 2017, Garçon, 1996, Rouhier et al., 2017] est un modèle spatialisé, dans lequel le bassin versant est découpé en mailles hydrologiques. Chacune de ces mailles peut être découpée en bandes d'altitudes comme dans le modèle hydrologique utilisé en opérationnel MORDOR-SD (Semi-Distribué). Les forçages sont alors étagés afin de mieux représenter leur variabilité spatiale en fonction de l'altitude, ce qui est surtout nécessaire pour les bassins de montagne. Sur le bassin versant de l'Odet, l'altitude varie peu et le modèle est utilisé avec une seule bande d'altitude. Les mailles sont reliées horizontalement entre elles selon le réseau hydrographique. La structure du modèle MORDOR-TS et le maillage réalisé dans le cadre de ce travail sur le bassin de l'Odet sont présentés en figure 2.10. Le bassin versant de l'Odet est découpé en 39 mailles dont la surface varie entre 12 km^2 et 22 km^2. Le sous-bassin de l'Odet est couvert par 11 mailles, celui du Jet par 6 mailles et celui du Steïr par 10 mailles. Les 12 autres mailles couvrent l'aval du bassin versant.

Le modèle est forcé par la pluie horaire cumulée (P) et la température (T), toutes deux spatialisées sur le maillage. Il est composé d'un module de production dans lequel la production est calculée pour chaque maille. Le bilan en eau du modèle est corrigé par les

[10]MOdèle à Réservoirs de Détermination Objective du Ruissellement - Totalement Spatialisé



(a) Décembre 2013 - Période de calage



(b) Janvier 2014 - Période de validation

Figure 2.9: Débits mesurés et simulés avec le modèle GRP à la station Tréodet.

paramètres *cp* et *kmin*, ce dernier paramètre étant utilisé pour le calcul de l'évapotranspiration ETR. Le module de production est ensuite composé de six réservoirs. Deux sont dédiés à la neige et la glace et sont désactivés ici. Les quatre autres représentent respectivement : (i) un réservoir superficiel U de capacité *Umax*, (ii) un réservoir intermédiaire L de capacité *Lmax* et dont le niveau de remplissage est contrôlé par le paramètre *evl*, (iii) un réservoir Z représentant le piégeage capillaire du sol de capacité *Zmax*, et (iv) un réservoir profond N dont le niveau de remplissage est contrôlé par le paramètre *lkn*. Trois flux sont transférés à partir du module de production vers le module de routage : le débit de surface Qs, le débit de versant Qv et le débit de base Qb. Le paramètre *kr* contrôle la part de l'eau alimentant d'une part le réservoir N et d'autre part directement le débit. Le module de routage permet ensuite de transférer la production de chaque maille jusqu'aux points de sortie du modèle. Il est composé du routage intra-maille, correspondant au transfert de l'écoulement produit sur une maille jusqu'à son exutoire, et du routage inter-mailles, qui correspond au transfert maille à maille jusqu'aux points de sortie. La fonction de transfert est issue du modèle de l'onde diffusante, avec la célérité *Cel* et la diffusion *Diff* indépendants du débit [Hayami, 1951]. Le modèle MORDOR-TS a ainsi 10 paramètres libres à caler. Ils sont synthétisés dans le tableau 2.5. Une description complète du modèle est présentée en annexe M.4.

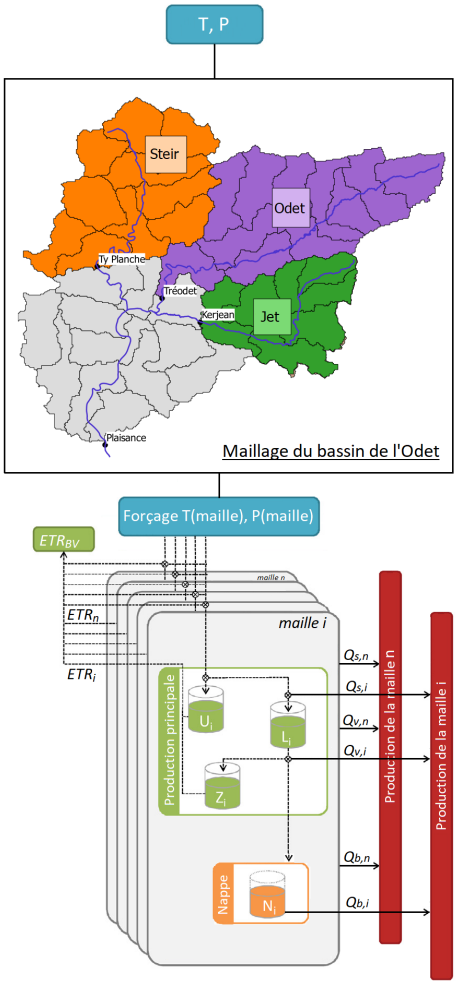


Figure 2.10: Schéma fonctionnel du module de production du modèle hydrologique MORDOR-TS sur le bassin de l'Odet avec une seule bande d'altitude - D'après Rouhier et al. [2017].

Paramètre	Unité	Rôle	Module
c_p	-	Coefficient correctif pour les précipitations	Bilan en eau
k_{min}	-	Paramètre pour le coefficient cultural	
U_{max}	mm	Capacité du réservoir de surface U	Production
L_{max}	mm	Capacité du réservoir de intermédiaire L	
ev_L	-	Exposant de vidage du stock intermédiaire L	
Z_{max}	mm	Capacité du réservoir évaporant Z	
k_r	-	Clé de répartition entre réservoir profond et ruissellement retardé	
lk_N	$mm.h^{-1}$	Logarithme du coefficient de tarissement du réservoir profond N	Écoulements lents
Cel	$m.s^{-1}$	Célérité de l'équation 1D de propagation (Hayami)	Transfert
$Diff$	$m^2.s^{-1}$	Diffusion de l'équation 1D de propagation (Hayami)	

Tableau 2.5: Bilan des 10 paramètres à caler dans le modèle MORDOR-TS.

La version distribuée MORDOR-TS comporte plusieurs avantages [Rouhier, 2018]. Tout d'abord, elle permet de mieux prendre en compte la variabilité spatiale des forçages météorologiques puisque les pluies et les températures sont moyennées à l'échelle de la maille hydrologique et non plus d'un bassin versant. De plus, un seul modèle peut permettre la modélisation de plusieurs points de sortie simultanément puisque des débits modélisés sont disponibles à chaque exutoire de maille. Cela évite la multiplication des modèles sur un grand bassin versant et garantit également la cohérence des débits simulés amont-aval. Enfin, les jeux de paramètres peuvent être différents d'une maille à l'autre, ce qui permet d'optimiser le calage du modèle par zones supposées homogènes.

Calage du modèle Le calage des paramètres est réalisé grâce à l'algorithme génétique multi-objectifs caRamel[11], disponible en ligne pour le logiciel R[12] [Le Moine et al., 2015, Monteil et al., 2019]. Cet algorithme permet l'optimisation simultanée de plusieurs fonctions objectif, en faisant évoluer une population de jeux de paramètres par itérations successives. À chaque itération, seuls les jeux les plus adaptés sont conservés. La population estime de mieux en mieux le front de Pareto optimum, correspondant à l'ensemble des jeux de paramètres pour lesquels l'amélioration d'une des fonctions objectif ne peut se faire qu'en en dégradant une autre. La fonction objectif utilisée pour le calage des sous-bassins de l'Odet regroupe trois scores :

- NSE (*Nash-Sutcliffe Efficiency*, Nash and Sutcliffe [1970]) sur les débits horaires ;
- NSE sur le régime hydrologique (débits moyens journaliers) ;
- NSE sur les débits classés.

Un jeu de paramètres est calé sur chacun des trois sous-bassins amont Tréodet, Kerjean et Ty-Planche. Le tableau 2.6 présente les valeurs des paramètres calés sur la période de calage P_{Cal} pour chacun des trois sous-bassins et tableau 2.7 synthétise les scores obtenus pour les débits simulés par le modèle sur les périodes P_{Cal} et P_{Cal}. Nous mettons en évidence le fait que les valeurs des NSE sur les débits horaires sont meilleurs à Tréodet et Kerjean sur la période de validation que sur la période de calage.

Bassin versant	c_p	k_{min}	ev_L	k_r	lk_N	U_{max}	L_{max}	Z_{max}	*Cel*	*Diff*
Tréodet	1.05	0.86	3.99	0.2	-5.45	108.2	104.5	79.5	0.52	422.5
Kerjean	1.03	0.72	1.50	0.15	-5.94	65.0	386.5	162.2	0.9	594.8
Ty Planche	0.97	1.22	2.85	0.10	-5.75	101.0	30.0	34.6	0.42	1645.4

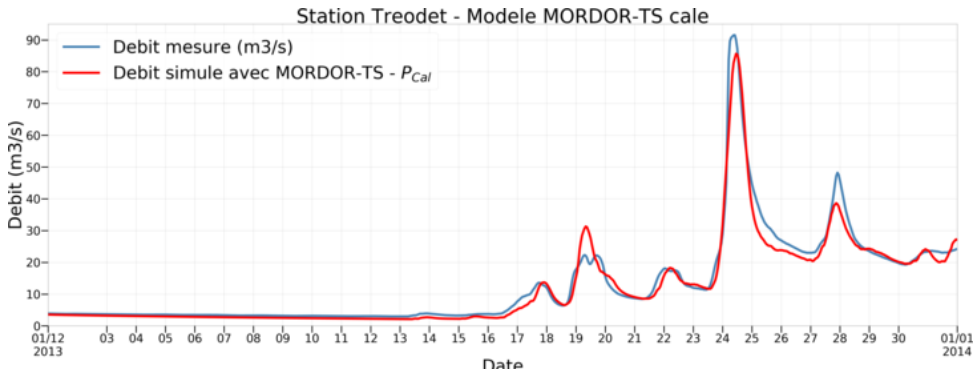
Tableau 2.6: Paramètres calés pour le modèle MORDOR-TS sur les sous-bassins de l'Odet.

[11] CAlage Rapide par Algorithme Multi-objectif EvoLutionnaire
[12] https://cran.r-project.org/web/packages/caRamel/index.html

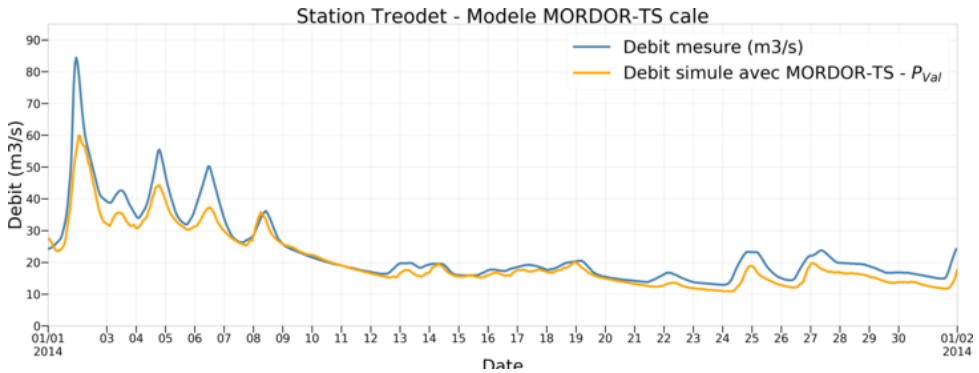
Bassin versant	Tréodet		Kerjean		Ty Planche	
Période	P_{Cal}	P_{Val}	P_{Cal}	P_{Val}	P_{Cal}	P_{Val}
Nash débits horaires	0.93	0.95	0.93	0.96	0.94	0.92
Nash régime journalier	0.99	0.96	0.99	0.97	0.99	0.94
Nash débits classés	0.99	0.97	0.99	0.98	0.99	0.95

Tableau 2.7: Scores de calage du modèle MORDOR-TS sur les sous-bassins de l'Odet.

Graphiques des débits mesurés et simulés La figure 2.11 illustre les débits simulés par le modèle GRP à la station Tréodet pour le mois de décembre 2012 (période de calage, courbe rouge) et de janvier 2014 (période de validation, courbe orange). La courbe bleue représente la mesure. Le lecteur pourra se reporter à l'annexe M.5 pour visualiser les graphes pour les mois de décembre 2012, et de décembre 2013 à février 2014 pour les trois stations Tréodet, Kerjean et Ty-Planche.



(a) Décembre 2013 - Période de calage



(b) Janvier 2014 - Période de validation

Figure 2.11: Débits mesurés et simulés avec le modèle MORDOR-TS à la station Tréodet.

2.3.2 La modélisation hydraulique

Développé depuis plus de 20 ans par EDF et le CEREMA[13], MASCARET est un logiciel de modélisation monodimensionnelle (1D) des écoulements à surface libre. Il résout les équations de Saint-Venant. Ces équations sont dérivées des équations de Navier-Stokes à surface libre qui sont les équations de base de la mécanique des fluides. Le code MASCARET dispose de trois modules qui permettent de faire des calculs en régime fluvial permanent, fluvial non-permanent et en torrentiel. Ces trois modules peuvent être couplés avec le module *casiers* qui permet de représenter le stockage dans les plaines d'inondation relativement isolées du lit majeur. Dans le cadre de ce travail, le module fluvial non permanent a été utilisé sans casiers. Le paragraphe 2.3.2.1 présente les équations de Saint-Venant et les données nécessaires pour réaliser un calcul. Le paragraphe 2.3.2.2 présente quant à lui le modèle construit sur le bassin aval de l'Odet ainsi que les résultats du calage.

2.3.2.1 Le code MASCARET

Les équations de Saint-Venant Les hypothèses faites pour la mise en œuvre d'une modélisation hydraulique monodimensionnelle sont les suivantes :

- les écoulements sont à faible courbure dans le plan horizontal ; les accélérations verticales sont négligeables et les pressions sont quasi hydrostatiques ;
- la pente moyenne de l'écoulement est faible ;
- les contraintes de viscosité sur le fond et les berges sont prises en compte à l'aide des lois empiriques de frottement (loi de Strickler) ;
- chaque bief a un axe préférencié d'écoulement, les vecteurs vitesse sont toujours considérés parallèles à l'axe du bief ;
- l'influence du vent est négligée ;
- les apports latéraux sont supposés être perpendiculaires à l'axe principal d'écoulement et n'apportent ainsi pas de quantité de mouvement.

Dans chaque plan perpendiculaire à l'axe du bief, l'écoulement est alors entièrement déterminé par la connaissance de la vitesse moyenne (ou du débit total) et de la cote de la surface libre. Les grandeurs calculées sont toujours relatives à une section d'écoulement perpendiculaire à l'axe du bief. Chaque section est identifiée par son abscisse le long de cet axe. Les sections d'écoulement sont considérées comme étant la réunion de trois sous-ensembles (figure 2.12) : (i) le lit mineur, correspondant au lit naturel de l'écoulement hors crue, (ii) le lit majeur, qui est activé lorsque la cote de l'eau passe par-dessus la crête des berges du lit mineur et (iii) les zones de stockage considérées comme des réservoirs qui se remplissent et se vident avec la dynamique de la crue. Contrairement au lit majeur, ces

[13] Centre d'Études et d'Expertise sur les Risques, l'Environnement, la Mobilité et l'Aménagement

zones ne participent pas à l'écoulement et les vitesses dans le sens de l'écoulement sont supposées nulles. Un bief composé uniquement d'un lit mineur est dit *lit unique*. Dans le cas contraire, on parle de *lit composé*.

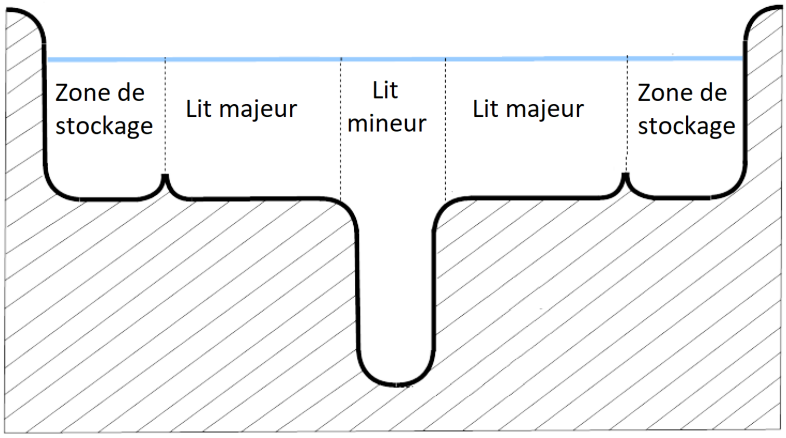


Figure 2.12: Visualisation du lit mineur, du lit majeur et des zones de stockage. D'après [Goutal and Zaoui, 2020]

Pour écrire les équations de Saint-Venant, on considère dans chacune des sections verticales de la rivière le débit Q(x,t) (m^3/s), la section mouillée S(x,t) (m^2), la cote de la surface de l'eau Z(x,t) (m) et le rayon hydraulique R(x,t) (m). Les équations de Saint-Venant en 1D sont composées de deux équations.

- La première est l'équation de continuité (2.2) et traduit la conservation des débits :

$$\frac{\partial S}{\partial t} + \frac{\partial Q}{\partial x} = q_l \tag{2.2}$$

  avec $q_l(x,t)$ les apports latéraux.

- La seconde correspond à l'équation fondamentale de la dynamique (2.3) :

$$\frac{\partial Q}{\partial t} + \frac{\partial}{\partial x}\left(\frac{\beta(x,S)Q^2}{S}\right) + gS\frac{\partial Z}{\partial x} = -gSJ. \tag{2.3}$$

  Le terme gSJ traduit l'effet des forces de frottement. J est sans dimension et traduit le taux moyen de dissipation de l'énergie. Il dépend du débit, des caractéristiques hydrauliques de la rivière et du coefficient de rugosité. J est évalué par la loi de Strickler en lit unique :

$$J = \frac{Q^2}{K_S^2 S^2 R^{4/3}} \tag{2.4}$$

  où K_S représente le coefficient de rugosité de Strickler.

  Le coefficient β est adimensionnel et résulte de la variation de la vitesse réelle de

l'écoulement dans une section. Sa définition est :

$$\beta = \frac{S}{Q^2} \int v^2 ds \tag{2.5}$$

En lit unique, compte-tenu des hypothèses formulées, il vient $\beta = 1$. Cela n'est plus vrai en lit composé.

Une confluence entre deux cours d'eau est modélisée par un nœud. Les équations résolues sur le noeud en régime fluvial non-permanent sont la continuité des débits ainsi que l'égalité des cotes.

Construction d'un modèle avec MASCARET La construction d'un modèle MASCARET repose tout d'abord sur la description de la géométrie du cours d'eau. Chaque tronçon de rivière appelé bief est décrit par une série de profils en travers décrivant les lits mineur et majeur du cours d'eau et les zones de stockage le cas échéant. Pour chaque profil on détermine son abscisse curviligne le long du bief auquel il appartient. Le calcul de la section mouillée au niveau de chaque section est réalisé à partir d'un planimétrage. Cette étape est réalisée une seule fois au lancement du calcul. Elle correspond à la discrétisation verticale des profils en travers. Le domaine doit également être discrétisé le long de l'axe des biefs. Cela forme le maillage. Les sections du maillage ne correspondant pas à une section fournie par l'utilisateur sont construites par interpolation entre les sections amont et aval disponibles. Une fois la géométrie définie, le modélisateur procède au calage du modèle en ajustant les coefficients de Strickler des lits mineur et majeur en minimisant une fonction coût préalablement définie.

2.3.2.2 Le modèle MASCARET du bassin de l'Odet aval

Le modèle hydraulique du bassin versant aval de l'Odet utilisé dans ce travail est issu du modèle opérationnel. Il est constitué de 12 zones de frottement telles qu'indiquées sur la figure 2.13. La géométrie du modèle a été simplifiée par rapport à celle du modèle opérationnel pour obtenir une meilleure stabilité numérique du code de calcul, indispensable pour réaliser des simulations perturbées. Le modèle reconstruit dans le cadre de la thèse est composé de 172 sections pour une modélisation d'environ 35 km de cours d'eau. La résolution du maillage est comprise entre 5 m et 10 m. Le pas de planimétrage est 10 cm.

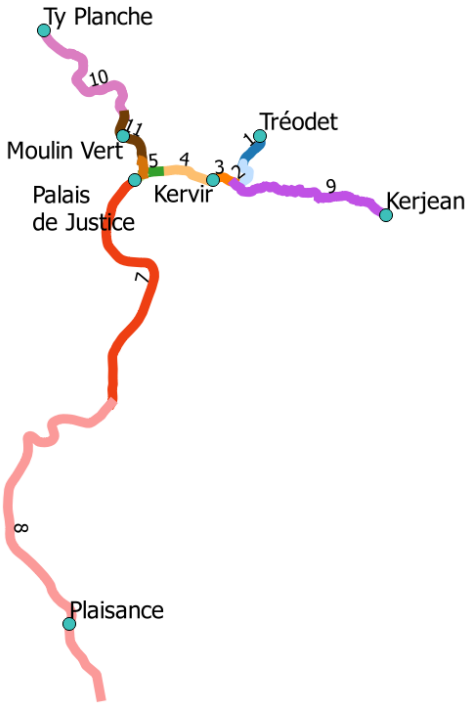


Figure 2.13: Visualisation des 12 zones de frottement du modèle MASCARET de l'Odet.

Calage du modèle Suite à la modification de la géométrie du modèle, celui-ci a été recalé pour obtenir une bonne modélisation des hauteurs d'eau aux trois stations de vigilance Kervir, Moulin-Vert et Justice. De même que pour le calage des modèles hydrologiques, cette étape est réalisée avec une procédure de calage-validation. Nous considérons les périodes de calage et de validation suivantes :

- période de calage P_{Cal} : hiver 2013-2014, entre le 1er novembre 2013 et le 28 février 2014. Cet hiver a été marqué par de nombreux événements de crue et de surcote maritime (voir paragraphe 2.1, tempête *Dirk*, dépressions *Gerhard* et *Qumeira*).
- période de validation P_{Val} : hiver 2012-2013, entre le 1er novembre 2012 et le 28 février 2013. Cet hiver a été moins touché par des événements de crue ou tempétueux.

La fonction coût utilisée pour l'optimisation des coefficients de Strickler est la suivante :

$$f_C = \frac{4}{5} * RMSE_{Chronique} + \frac{1}{5} * RMSE_{pics} \tag{2.6}$$

$RMSE_{Chronique}$ (*Root Mean Squarred Error*) représente le calcul de la RMSE sur toute la période P_{Cal} (du 1er novembre 2013 et le 28 février 2014). $RMSE_{pics}$ correspond au calcul de la RMSE sur les valeurs des pics de crue simulés (valeurs ponctuelles). L'introduction de la $RMSE_{pics}$ pondérée par le coefficient 0.2 permet d'améliorer la modélisation des pics

de crue sans pour autant dégrader de manière significative les débits simulés de manière courante.

La figure 2.14 montre les chroniques de hauteur d'eau observées aux trois stations de vigilance sur la période de calage. Cette figure met en évidence que la station Justice est dominée par l'influence de la marée même lorsque les niveaux sont hauts à l'amont. Pour les stations Kervir et Moulin-Vert, on compte 11 événements de crue mis en évidence en rouge sur lesquels la $RMSE_{pics}$ sera calculée. À Justice, étant donné le caractère maritime de cette station, la stratégie est légèrement différente. La RMSE sur la chronique est calculée uniquement pour les hauteurs d'eau supérieures au seuil 1.5 m. De cette manière, nous privilégions les hauteurs supérieures à la mi-marée dans une optique de vigilance crues. Par ailleurs, la RMSE sur les pics est calculée pour toutes les pleines mers dont la cote est supérieure à 2 m. Cela correspond environ à 200 pics de pleine mer.

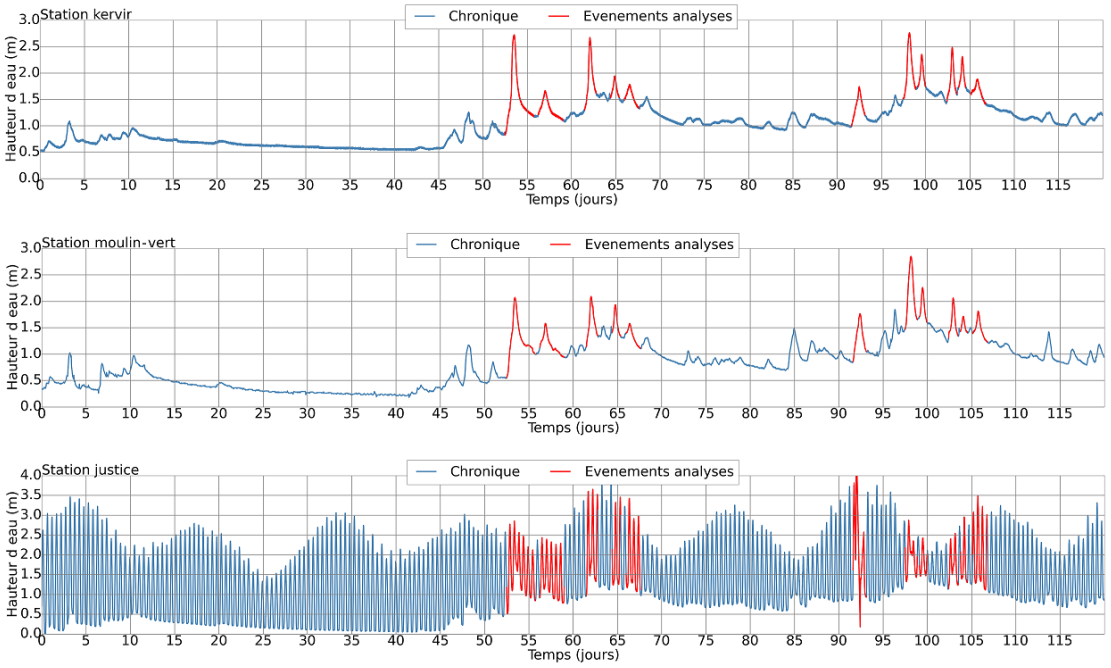


Figure 2.14: Chroniques aux trois stations de vigilance Kervir, Moulin-Vert et Justice entre le 1er novembre 2013 et le 28 février 2014. Mis en évidence des événements de crue en rouge.

Les trois stations sont calées de manière indépendante en faisant varier les valeurs de coefficients de Strickler des lits mineur et majeur à l'aval de chaque station et en imposant des valeurs standard sur les zones qui n'ont pas ou très peu d'impact sur les hauteurs simulées aux stations. Le tableau 2.8 présente les valeurs des coefficients de Strickler déterminés par l'étape de calage. La zone 11 à l'aval de la station Moulin-Vert possède un coefficient de Strickler de *1* en lit majeur ce qui est extrêmement faible. Nous consacrerons un paragraphe par la suite pour expliquer cette anomalie.

Zone	K_S mineur	K_S majeur
1	30	20
2	20	10
3	15	10
4	38	13.5
5	15	10
6	36.5	34
7	37	25
8	25.5	5
9	20	10
10	20	10
11	32.5	1
12	20	10

Tableau 2.8: Valeur des coefficients de Strickler calés sur le bassin de l'Odet

RMSE sur les chroniques Le tableau 2.9 présente la $RMSE_{chronique}$ calculée sur les chroniques simulées en périodes de calage et de validation pour les trois stations. La $RMSE_{chronique}$ est présentée pour toutes les hauteurs d'eau et par classe de hauteur. Les limites des classes correspondent aux seuils de basculement dans la vigilance jaune ou orange. Les valeurs de ces seuils étant confidentielles, elles ne sont pas mentionnées dans la thèse. Ces résultats montrent que les performances des simulations réalisées par le modèle sont similaires entre la période de calage et de validation. Le calage du modèle est relativement robuste.

Bassin versant	Kervir		Moulin-Vert		Justice	
Période	P_{Cal}	P_{VCal}	P_{Cal}	P_{Val}	P_{Cal}	P_{Val}
Toutes hauteurs	2.9	2.4	4.8	4.5	6.0	5.3
Vigilance verte	2.6	2.3	4.3	4.4	6.2	5.5
Vigilance jaune	6.0	6.3	7.2	7.2	5.7	5.0
Vigilance orange	3.8	–	15.46	–	–	–

Tableau 2.9: RMSE (cm) pour les périodes de calage et de validation du modèle MASCARET du bassin de l'Odet.

Modélisation des pics de crue à Kervir et Moulin-Vert Le tableau 2.10 présente les valeurs des 11 pics de l'hiver de calage et l'erreur de simulation de la hauteur du pic correspondante. Le tableau 2.11 présente les résultats analogues pour l'hiver de validation, pour lequel seulement deux pics d'hydrogrammes ont été recensés. Ces tableaux mettent en évidence que les pics de crue sont relativement bien représentés par le modèle sur les périodes P_{Cal} et P_{Val}.

Événement	Hauteur d'eau mesurée à Kervir (m)	Erreur (cm)	Hauteur d'eau mesurée à Moulin-Vert (m)	Erreur (cm)
Dec 2013 - 1	2.72	-4.8	2.07	-1.9
Dec 2013 - 2	1.66	6.7	1.58	8.9
Jan 2014 - 1	2.67	-3.1	2.09	-3.4
Jan 2014 - 2	1.94	3.3	1.93	4.5
Jan 2014 - 3	1.78	3.7	1.58	6.8
Jan 2014 - 4	1.74	-0.6	1.76	6.2
Feb 2014 - 1	2.76	-6.7	2.85	3.2
Feb 2014 - 2	2.35	-0.5	2.26	-5.7
Feb 2014 - 3	2.48	7.1	2.06	0.6
Feb 2014 - 4	2.31	4.6	1.71	5.6
Feb 2014 - 5	1.88	1.3	1.81	2.3

Tableau 2.10: Valeur observée pour les 11 pics sur P_{Cal} et erreur de modélisation correspondante - Stations Kervir et Moulin-Vert.

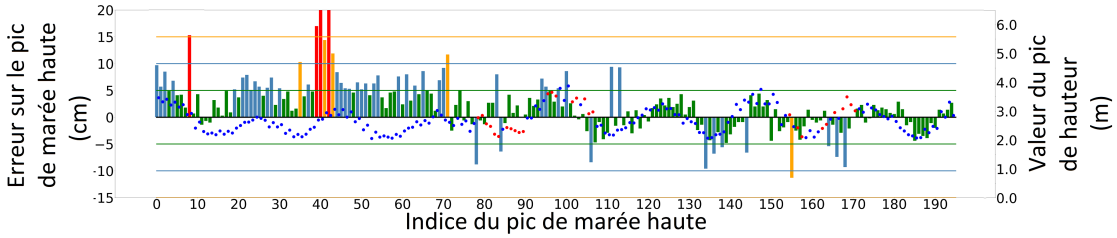
Événement	Hauteur d'eau mesurée à Kervir (m)	Erreur (cm)	Hauteur d'eau mesurée à Moulin-Vert (m)	Erreur (cm)
Dec 2012	2.22	3.0	2.09	-0.2
Jan 2013	1.38	-0.3	1.24	4.9

Tableau 2.11: Valeur observée pour les 2 pics sur P_{Val} et erreur de modélisation correspondante - Stations Kervir et Moulin-Vert.

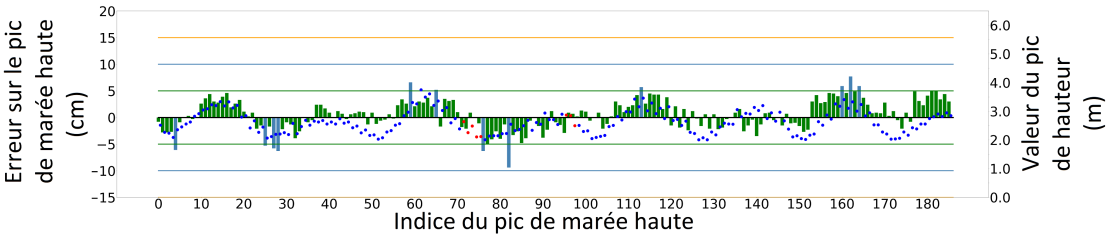
Modélisation des pics de pleine mer à Justice La figure 2.15 représente les différences entre les pics de marée haute simulés et mesurés pendant les hivers de calage et de validation pour toutes les pleines mers dont la hauteur à Justice dépasse 2.0 m. Pour chaque pleine mer, la valeur de l'erreur est représentée par une barre de couleur. Une barre verte correspond à une erreur absolue inférieure à 5 cm, une barre bleue correspond à une erreur absolue comprise entre 5 cm et 10 cm, une barre orange correspond à une erreur absolue comprise entre 10 cm et 15 cm et une barre rouge correspond à une erreur absolue supérieure à 15 cm. Une valeur positive de l'erreur au pic correspond à une surestimation du modèle par rapport aux mesures. Sur ces graphes sont superposés grâce à l'échelle de droite les hauteurs de pleine mer mesurées à l'échelle à Justice. Les points sont de couleur rouge lorsque le pic de pleine mer a lieu au cours d'un des 11 événements définis ci-avant.

Ces graphiques mettent en évidence que la plupart des pics sont simulés avec une erreur absolue inférieure à 5 cm. Seulement quatre marées hautes sont modélisées avec une erreur supérieure à 15 cm pendant la période de calage. On note une bonne robustesse du calage à Justice sur la période de validation pour l'estimation des pics de pleine mer.

Ces graphiques montrent également que les erreurs d'estimation des pics de marée haute ne



(a) Hiver 2013-2014 - Calage



(b) Hiver 2012-2013 - Validation

Figure 2.15: Erreurs sur les pics de marée haute à Justice sur les périodes P_{Cal} et P_{Val} représentées par les histogrammes de couleur. Les marées hautes correspondantes observées à Justice sont indiquée grâce à l'échelle de droite. Voir le texte pour la signification des couleurs.

sont pas plus importantes pendant les épisodes de crue. À une exception près pendant la phase de calage, les marées pour lesquelles l'erreur est importante (barres orange ou rouges) ne se produisent pas pendant un des 11 événements. En période de calage, l'erreur moyenne absolue d'estimation du pic de pleine mer est de 3.0 cm pendant les 11 événements. Elle est de 4.2 cm pour les pics en dehors des événements définis. La qualité de la simulation des pics de pleine mer ne semble donc pas significativement différente pendant les événements fluviaux. Il n'est pas pertinent d'étudier cet aspect en période de validation car seuls deux événements ont été retenus.

Par ailleurs, sur le graphique en période de calage, on remarque que jusqu'à environ 70 marées hautes, soit jusqu'au 20 décembre 2013, le modèle ne sous-estime jamais les hauteurs des pics de pleine mer. Après cette date, le modèle surestime et sous-estime les valeurs de pics. On calcule en effet les erreurs moyennes suivantes :

- moyenne des erreurs sur les 71 premiers pics de marée haute : 5.8 cm
- moyenne des erreurs sur les pics de marée haute suivants : -0.1 cm

Le graphique de la figure 2.14 montre qu'à partir du 20 décembre (ce qui correspond à 50 jours depuis le 1er novembre 2013), les débits du Steïr et de l'Odet deviennent significativement importants. Cette date marque le début d'une période de débits soutenus à l'amont. Ainsi, lorsque les débits fluviaux sont bas, le modèle a tendance à surestimer les pics à Justice. En revanche, lorsque ceux-ci sont plus soutenus, le modèle ne présente plus

cette tendance. Il surestime et sous-estime alternativement les niveaux à marée haute, dans des gammes qui restent raisonnables.

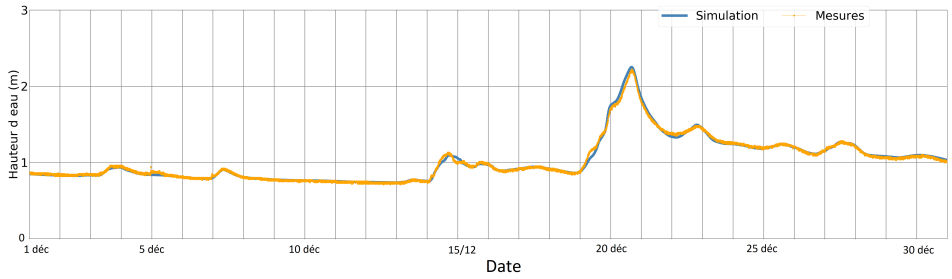
Pour finir, pour les périodes P_{Cal} et P_{Val}, les pics correspondants à des marées importantes ont plutôt tendance à être surestimés, alors que les pics pour les marées plus faibles ont plutôt tendance à être sous-estimés. En effet, la courbe présentant les niveaux à Justice semble suivre une variation analogue à la variation des surestimations/sous-estimations. Cela signifie que les marées de fort coefficient ont tendance à remonter davantage dans l'estuaire que les marées de faible coefficient.

Pour conclure, nous avons mis en évidence que la qualité de la modélisation des pics de pleine mer à Justice est dépendante :

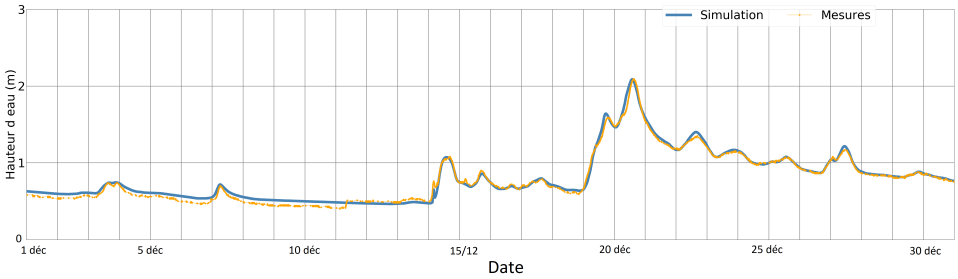
- des débits amont : les niveaux bas à l'amont sont responsables d'une tendance à la surestimation des pics de pleine mer à Justice ;
- de l'importance de la marée à l'aval : les niveaux bas à Plaisance ont du mal à remonter dans l'estuaire, et engendrent ainsi une sous-estimation du pic.

C'est la combinaison des deux tendances qui donnera l'erreur finale à Justice. Elle n'est pas spécifiquement liée à des pics de crue particuliers.

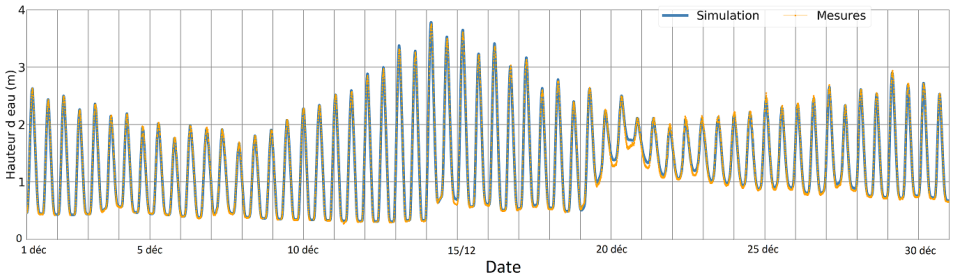
Graphiques des débits mesurés et simulés La figure 2.16 montre les hauteurs d'eau simulées avec le modèle MASCARET aux trois stations de vigilance du 1er au 31 décembre 2012 (période de validation). Ces graphiques mettent en évidence une bonne modélisation des chroniques de hauteurs d'eau, ce qui est cohérent avec les résultats précédemment présentés.



(a) Station Kervir



(b) Station Moulin-Vert



(c) Station Justice

Figure 2.16: Hauteurs d'eau mesurées et simulées aux trois stations de vigilance en décembre 2012 - Période P_{Val}.

Focus sur le calage de la station Moulin-Vert La station Moulin-Vert est située à environ 1 km à l'amont d'un seuil important placé perpendiculairement à l'écoulement. La carte figure 2.17 présente la localisation de cet ouvrage, et la figure 2.18 présente deux photographies de l'ouvrage (vue amont et vue aval). Dans cette zone, les hypothèses de modélisation monodimensionnelle ne sont pas respectées : les vecteurs vitesse ne sont pas ici parallèles à l'axe du bief. Cette configuration peut être à l'origine de la valeur *1* obtenue pour le coefficient de Strickler en lit majeur de la zone 11 (bief situé entre Moulin-Vert et cet ouvrage, voir figure 2.13 pour la localisation des zones de frottement et tableau 2.8 pour les valeurs de coefficients de Strickler calés). En effet, ce coefficient de frottement cumule en réalité une perte de charge linéaire et une perte de charge ponctuelle artificielle liée à la configuration du bief au niveau du seuil où les effets 2D sont très marqués. Cette valeur anormalement faible peut également être liée au fait que l'axe du bief présente un

coude à l'aval de la station. Cela crée une perte de charge, rattrapée artificiellement avec le coefficient de Strickler.

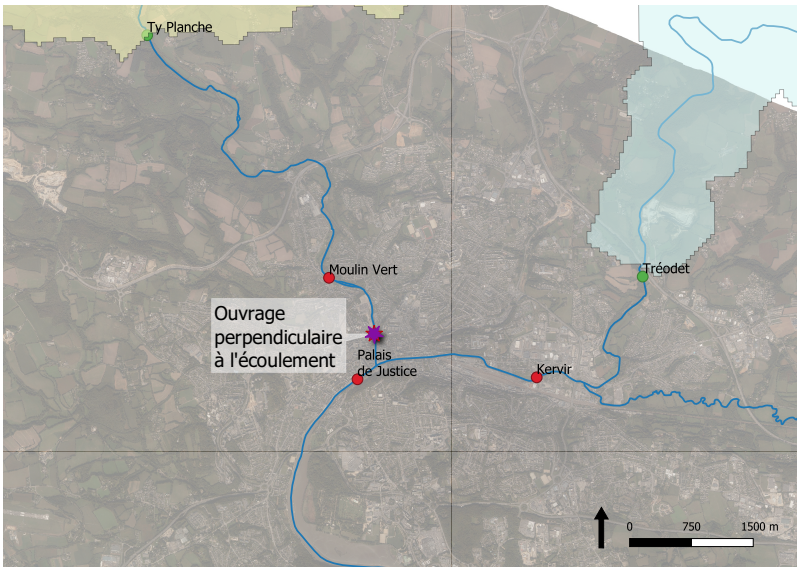


Figure 2.17: Positionnement de l'ouvrage perpendiculaire à l'écoulement par rapport à la station Moulin-Vert.

(a) Vue amont de l'ouvrage

(b) Vue aval de l'ouvrage

Figure 2.18: Vue de l'ouvrage placé perpendiculairement à l'écoulement.

Ce seuil est régulièrement manœuvré par les services de la ville de Quimper, mais le SPC-VCB n'est pas averti de ces manœuvres. Un relevé des dates des manœuvres passées avec les cotes correspondantes n'est pas disponible. À titre d'illustration, la figure 2.19 présente les hauteurs mesurées à Ty-Planche et à Moulin-Vert de janvier à avril 2012. La chronique présente des sauts à Moulin-Vert qui ne sont pas présents à la station Ty-Planche située environ 5 km à l'amont. Ces sauts correspondent à des manœuvres du seuil. Pour le calage du modèle, nous nous sommes placés en période de seuil bas correspondant à sa position hivernale. Nous verrons dans la suite de l'étude que le fait de ne pas connaître la position exacte du seuil peut néanmoins poser des problèmes pour une bonne mise en place des méthodes de génération des prévisions d'ensemble. En effet, le modèle est calé pour une seule position du seuil. Lorsque le seuil est remonté pour soutenir les niveaux à l'amont, le modèle aura tendance à sous-estimer les niveaux d'eau mesurés.

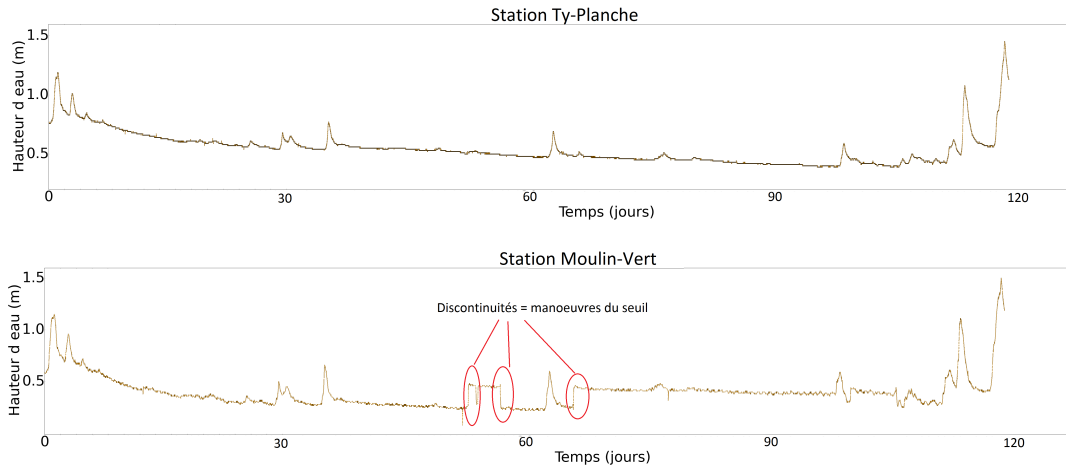


Figure 2.19: Chroniques des hauteurs d'eau aux stations Ty-Planche (amont) et Moulin-Vert (aval) pour les mois de janvier à avril 2012. Mise en évidence de la manœuvre du seuil.

Chapitre 3

Deux approches pour générer des prévisions d'ensemble

Sommaire

3.1 La vérification des ensembles 73
3.1.1 Contexte de la vérification 74
3.1.1.1 Quelle prévision vérifie-t'on ? 74
3.1.1.2 Fiabilité, résolution et finesse 75
3.1.2 Quels outils pour la vérification des ensembles ? 78
3.1.2.1 Vérification relative à un événement : diagramme de fiabilité et courbe ROC 79
3.1.2.2 Histogramme de rang, CRPS et RMSE 82
3.1.2.3 La stratification 86
3.1.3 Synthèse 86
3.2 L'étude d'incertitude : un préalable commun aux deux approches 87
3.2.1 Analyse locale ou globale 88
3.2.2 Différentes méthodes d'analyse de sensibilité globale 88
3.2.2.1 Une approche commune 88
3.2.2.2 Méthodes basées sur la décomposition de la variance . . 89
3.2.3 Mise en œuvre informatique 91
3.2.4 Études de sensibilité dans les domaines hydrologiques et hydrauliques 92
3.2.4.1 La méthode GLUE 92
3.2.4.2 Applications de la méthode GLUE et études ANOVA . 93
3.2.5 Synthèse 95
3.3 La calibration statistique pour corriger les ensembles de prévision bruts 96
3.3.1 Quantile Regression Forests (QRF) 97

3.3.2 Mise en œuvre informatique . . . 101
3.3.3 La reconstruction des membres post-traités . . . 101
3.3.4 Synthèse . . . 103
3.4 L'assimilation de données . . . 104
3.4.1 Notations et concepts . . . 105
3.4.1.1 Aide mémoire - Notations . . . 105
3.4.1.2 Vecteur d'état et vecteur de contrôle . . . 106
3.4.1.3 Notion de fenêtre d'assimilation . . . 107
3.4.1.4 Modèle d'évolution du système . . . 107
3.4.1.5 Modèle d'observation . . . 107
3.4.1.6 Différentes estimations du vecteur de contrôle . . . 109
3.4.2 Formulation bayésienne du problème d'assimilation de données . 110
3.4.3 Filtres de Kalman . . . 112
3.4.3.1 Le Filtre de Kalman . . . 112
3.4.3.2 Le Filtre de Kalman Étendu . . . 115
3.4.3.3 Le filtre de Kalman d'ensemble stochastique (EnKF) . 116
3.4.4 Mise en œuvre informatique . . . 120
3.4.5 Assimilation de données dans les modèles hydrologiques et hydrauliques . . . 120
3.4.5.1 L'assimilation de données pour la modélisation et la prévision hydrologiques . . . 121
3.4.5.2 L'assimilation de données pour la modélisation et la prévision en hydraulique fluviale . . . 124
3.4.6 Synthèse . . . 126

Ce chapitre présente les méthodes et outils utilisés dans la thèse pour la génération de prévisions d'ensemble. En premier lieu, le chapitre 3.1 présente les caractéristiques recherchées pour les prévisions d'ensemble et les outils utilisés pour les estimer. C'est grâce à ces outils que nous pouvons ensuite comparer les performances des prévisions d'ensemble obtenues avec deux familles de méthodes : la calibration statistique et l'assimilation de données ensembliste. La réalisation d'une étude de sensibilité telle que présentée au paragraphe 3.2 s'avère être un préalable commun aux deux méthodologies développées.

La première méthodologie développée, présentée au paragraphe 3.3, est la calibration statistique d'ensembles bruts par la méthode des forêts aléatoires QRF (*Quantile Regression Forests*, Meinshausen [2006]). Pour cela, il s'agit dans un premier temps de générer un ensemble brut de prévision en propageant un ensemble de modèles déterministes créés par échantillonnage des paramètres incertains. Les performances de l'ensemble brut sont liées à la manière dont les sources d'incertitude sont prises en compte pour sa génération. Une étude de sensibilité est donc nécessaire afin d'identifier et de hiérarchiser les sources d'incertitude du modèle de prévision. Les variables d'entrée qui n'influencent pas la sortie seront fixées pour la génération de l'ensemble brut. Une fois l'ensemble brut généré, celui-ci est calibré grâce à la construction d'une forêt aléatoire. En opérationnel, les prévisions brutes sont calibrées avec la forêt aléatoire construite pour fournir des prévisions d'ensemble calibrées.

La seconde méthodologie pour générer des prévisions d'ensemble est présentée au paragraphe 3.4. Elle consiste en l'introduction d'une assimilation de données ensembliste dans la chaîne de prévision. Les membres corrigés sur la fenêtre d'analyse sont ensuite propagés par le modèle en prévision, formant ainsi une prévision d'ensemble. L'étude de sensibilité est ici nécessaire pour connaître quels sont les paramètres d'entrée que l'assimilation de données devra corriger et pour valider la distribution des paramètres incertains.

3.1 La vérification des ensembles

Des prévisions sont réalisées dans un grand nombre de disciplines : en économie, dans le domaine médical, de l'environnement, en climatologie, en hydrométéorologie, ... Pour une utilisation éclairée des systèmes de prévision dans un contexte opérationnel, il est indispensable d'évaluer *a posteriori* les qualités des prévisions en les confrontant aux observations. Cette évaluation fait partie d'une discipline particulière appelée la *vérification*. Ce terme, initialement spécifique au domaine de la météorologie [Jolliffe and Stephenson, 2003], sera étendu dans ce manuscrit à l'évaluation des prévisions d'ensemble hydrologiques et hydrauliques. Les éléments décrits dans cette section permettront d'analyser et de comparer de manière objective les performances des prévisions d'ensemble hydrologiques et hydrauliques construites dans le cadre de ce travail.

Le paragraphe 3.1.1 présente tout d'abord les caractéristiques des ensembles que nous allons vérifier, ainsi que les propriétés recherchées pour les prévisions d'ensemble. En particulier, les concepts de *fiabilité* et *résolution* y sont définis. Le paragraphe 3.1.2 présente quant à lui les outils qui seront utilisés pour la vérification des ensembles : courbes de fiabilité, courbes ROC (*Relative Operating Characteristic*), histogrammes de rang, CRPS (*Continuous Ranked Probability Score*) et RMSE (*Root Mean Squared Error*). Le concept de stratification qui permet d'appliquer la vérification à des sous-échantillons homogènes y sera abordé. La synthèse de cette section est présentée au paragraphe 3.1.3.

3.1.1 Contexte de la vérification

3.1.1.1 Quelle prévision vérifie-t'on ?

Nous décrivons dans ce paragraphe la nature et la forme des prévisions sur lesquelles va porter le processus de vérification.

Qualité et valeur des prévisions d'ensemble Les prévisions ont souvent vocation à être utilisées dans un processus de décision. C'est là que se situent les interactions entre le modélisateur et le décideur. L'évaluation de la prévision peut donc se situer à deux niveaux, selon que l'on adopte le point de vue du modélisateur ou du décideur [Murphy, 1993] :

- lorsque l'évaluation consiste en la confrontation des valeurs prévues et des observations, on parle alors de *qualité* des prévisions ;
- lorsque l'évaluation porte plus généralement sur la prise de décision qui découle de la prévision, on parle alors d'*utilité* ou de *valeur* d'une prévision. Cela permet de mesurer l'intérêt du système de prévision pour un usage précis. La prise de décision, propre à chaque utilisateur, peut être guidée par l'application d'un modèle ou d'une fonction coût. La définition de la valeur d'une prévision d'ensemble permet en particulier d'évaluer ce qu'elle apporte par rapport à une prévision déterministe.

Dans cette thèse, nous nous intéressons uniquement aux qualités intrinsèques des prévisions. Les aspects relatifs à la prise de décision ne sont pas abordés.

Évaluer des prévisions probabilistes Les prévisions que nous cherchons à évaluer sont des prévisions probabilistes de grandeurs discrètes (pluie horaire) ou continues (débit hydrologique ou hauteur d'eau). Les prévisions d'ensemble étant constituées d'un ensemble de membres équiprobables $(x_1, ...x_{N_{ens}})$, il s'agit d'une prévision discrète. Elle est caractérisée par la fonction de répartition empirique en escalier :

$$F(x) = \frac{1}{N_{ens}} \sum_{i=1}^{N_{ens}} H_{x_i}(x)$$

où H_{x_i} représente la fonction échelon, ou *Heaviside*, en x_i :

$$\begin{cases} H_{x_i}(x) = 0 \ si \ x < x_i \\ H_{x_i}(x) = 1 \ si \ x \geq x_i. \end{cases} \tag{3.1}$$

La loi de probabilité définie par l'ensemble converge vers la loi de probabilité de la prévision de la variable d'intérêt appelée f_p quand le nombre de membres augmente. L'ensemble représente donc la discrétisation d'une loi de probabilité que l'on ne peut appréhender que de manière discrète. Dans le contexte de vérification, nous cherchons à évaluer la représentativité de l'ensemble par rapport aux observations en termes de loi de probabilité, c'est-à-dire que nous comparons la densité de probabilité f_p avec la densité de probabilité de l'observation f_o, ces deux densités de probabilité étant en pratique estimées de manière discrète.

3.1.1.2 Fiabilité, résolution et finesse

D'après Jolliffe and Stephenson [2003], les deux caractéristiques les plus importantes d'une prévision probabiliste sont la fiabilité et la résolution. Ce sont deux caractéristiques qui dépendent de la prévision probabiliste et des observations. La finesse quant à elle, appelée également acuité, ne dépend que de la forme des prévisions probabilistes. Ces différents concepts sont décrits ci-après.

Fiabilité La fiabilité (*calibration* ou *reliability* en anglais) représente la cohérence statistique entre les prévisions réalisées a priori et les fréquences observées *a posteriori*. Pour un système de prévision fiable, la densité de probabilité de l'observation f_o, compilée sur tous les cas pour lesquels la densité de probabilité de la prévision probabiliste est égale à f_p, est exactement égale à f_p. f_o est donc la densité de probabilité de l'observation quand f_p a été prévue. La figure 3.1 illustre les différences pouvant exister entre f_o et f_p en termes de moyenne et de variance et les conséquences sur les caractéristiques de la prévision. Lorsque l'ensemble est fiable, les membres de l'ensemble sont statistiquement indiscernables de la vraie distribution, donc $f_o = f_p$. Si $f_o \neq f_p$, la prévision n'est pas fiable. La prévision ne converge pas vers l'observation en termes de loi de probabilité. Si les prévisions présentent un biais systématique, la distribution de la prévision est décalée par rapport à celle de l'observation. Les prévisions sous-dispersives n'arrivent pas à représenter la gamme totale des incertitudes, tandis que les prévisions sur-dispersives surestiment la gamme d'incertitude.

La fiabilité devrait donc être évaluée pour l'ensemble des distributions prévues f_p. Cependant, une difficulté réside dans le fait que la densité de probabilité de la prévision f est différente à chaque nouvelle prévision. De plus, à chaque réalisation de la prévision telle

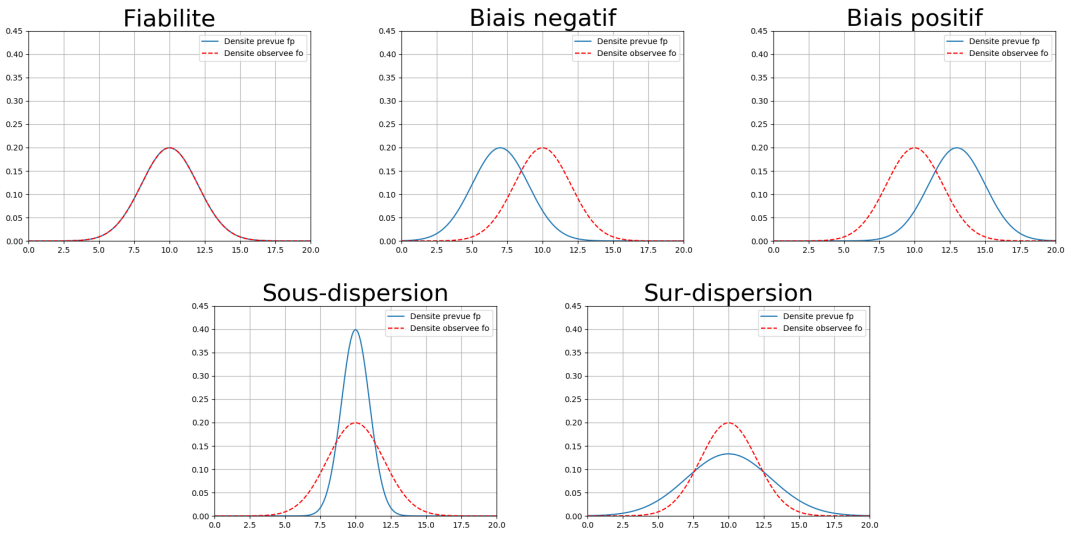


Figure 3.1: Différences caractéristiques entre les densités de probabilité prévues f_p et observées f_o. Schéma réalisé d'après Weigel [2012].

que $f = f_p$, l'observation associée ne correspond qu'à une seule réalisation de la loi f_o. Il faut donc attendre que $f = f_p$ un grand nombre de fois pour pouvoir construire f_o et évaluer la fiabilité. La fiabilité définie de cette manière est en pratique difficile à évaluer. Les principaux outils de vérification présentés par la suite se basent sur une définition moins stricte de la fiabilité : ils permettent d'évaluer le comportement moyen des réalisations pour une variété de densités prévues.

Résolution La fiabilité est une condition nécessaire pour l'utilité dans la pratique d'une prévision d'ensemble, mais elle n'est pas suffisante. Prenons l'exemple d'une prévision probabiliste qui soit égale à la fréquence issue de la climatologie établie à partir de mesures sur une période passée. Par construction, cette prévision probabiliste est fiable. Cependant, elle n'apporte aucune information supplémentaire en termes de prévision par rapport à la climatologie. Une prévision d'ensemble doit donc être capable d'apporter plus d'information qu'une prévision climatologique. La capacité d'une prévision à discriminer a priori les cas pour lesquels l'événement se produit plus ou moins fréquemment que la climatologie de référence est appelée résolution [Murphy, 1973]. Cet attribut indique dans quelle mesure la prévision peut séparer différentes situations, autrement dit si l'ensemble *prend des risques* [Bouttier, 2021]. La figure 3.2 illustre la notion de résolution en présentant des prévisions de résolution croissante. Le système climatologique (a) n'a pas de résolution, la densité de probabilité prévue f_p n'est pas discernable de la densité de probabilité issue de la climatologie. Les graphiques (b) et (c) caractérisent une résolution non nulle et plus importante pour (c). Le système déterministe (d) présente une résolution maximale.

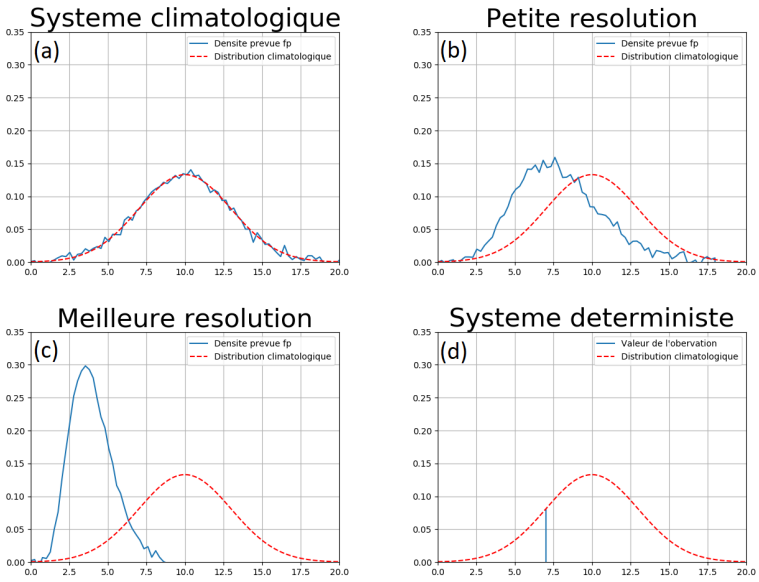


Figure 3.2: Différentes formes de la densité de probabilité de la prévision f_p par rapport à la distribution climatologique. Les prévisions (a) à (d) présentent une résolution croissante, avec une résolution nulle pour le système climatologique (a) et une résolution maximale pour le système déterministe (d). Schéma réalisé d'après Candille, 2006.

Finesse Une autre propriété, la finesse (*sharpness* en anglais), mesure la variabilité de la prévision probabiliste. Pour chaque prévision, plus la dispersion des membres est faible, meilleure est sa finesse, moins il y a d'incertitude. Cette propriété étant fonction uniquement des prévisions, son évaluation n'est pas un moyen de vérification des ensembles. La figure 3.3 présente les densités de probabilité prévues par deux ensembles de finesse différente. Sur cet exemple, la prévision caractérisée par la densité de probabilité f_{p_2} est plus fine que celle caractérisée par f_{p_1}.

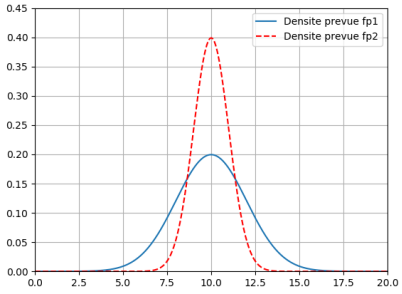


Figure 3.3: Différentes formes de la densité de probabilité prévue. La prévision caractérisée par la densité de probabilité f_{p_2} est plus fine que celle caractérisée par f_{p_1}.

Pour un système de prévision parfaitement fiable, la finesse est identique à la résolution.

Gneiting et al. [2007] proposent le paradigme de maximiser la finesse des prévisions probabilistes, sous réserve que celles-ci soient fiables.

3.1.2 Quels outils pour la vérification des ensembles ?

Les systèmes de prévision probabilistes ne peuvent être validés qu'avec une approche statistique [Jolliffe and Stephenson, 2003]. Les outils pour l'évaluation des performances des ensembles utilisés dans cette thèse ont été développés dans le domaine de la météorologie. Ils ont pour objectif de vérifier la fiabilité et/ou la résolution des prévisions. Certains outils se focalisent sur un événement particulier, par exemple le dépassement d'un seuil. C'est le cas des courbes de fiabilité et des courbes ROC, qui permettent d'évaluer respectivement la fiabilité et la résolution. Elles sont décrites dans le paragraphe 3.1.2.1. Le paragraphe 3.1.2.2 présente quant à lui les méthodes d'évaluation qui ne sont pas basées sur l'évaluation d'événements particuliers. Nous y décrirons l'histogramme de rang, le CRPS et la RMSE. Ce dernier score sera utilisé pour comparer les prévisions d'ensemble moyennées aux prévisions déterministes. Pour finir, le paragraphe 3.1.2.3 abordera le concept de stratification. L'application d'une stratification est nécessaire pour l'évaluation des performances des ensembles dans des conditions spécifiques. Dans le cadre de ce travail, une stratification est appliquée aux ensembles pour leur évaluation lors d'événements pluvieux importants.

On considère désormais une prévision d'ensemble constituée de N_{ens} membres. Cette prévision d'ensemble est lancée au temps de base T_B. On considère une échéance de prévision, par exemple 6 heures. Le temps de vérification est donc $T_V = T_B + 6h$. Dans la suite, on appelle *prévision* l'ensemble des N_{ens} valeurs prévues par l'ensemble à T_V. À cette prévision est associée une observation qui correspond à la mesure de la variable d'intérêt au même temps T_V. La figure 3.4 illustre ces concepts. La vérification des prévisions est réalisée sur un ensemble de N_{prev} réalisations de la prévision, donc pour N_{prev} temps de base différents.

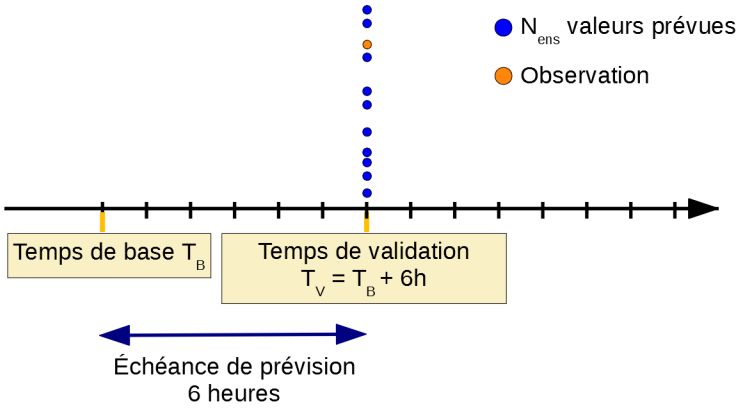


Figure 3.4: Notion de *prévision* utilisée pour la description des outils de vérification des ensembles. La prévision d'ensemble est lancée au temps de base T_B. Dans cet exemple, la vérification est réalisée pour l'échéance 6 heures à $T_V = T_B + 6h$. À T_V, nous avons N_{ens} valeurs prévues et une observation correspondante, cela forme un couple $(prévision, observation)$. L'évaluation des ensembles est réalisée sur N_{prev} couples $(prévision, observation)$.

3.1.2.1 Vérification relative à un événement : diagramme de fiabilité et courbe ROC

La vérification des ensembles peut être relative à un événement, par exemple le dépassement d'un seuil. Le diagramme de fiabilité permet de vérifier la fiabilité relative à un événement, tandis que le tracé des courbes ROC permet d'estimer la capacité de l'ensemble à discriminer les situations menant à l'occurrence ou non d'un événement. Ces deux courbes sont décrites ci-après.

Diagramme de fiabilité Le diagramme de fiabilité [Wilks, 2011] est un outil graphique qui résume la distribution jointe des prévisions et des observations pour des prévisions probabilistes d'événements binaires, par exemple la probabilité de dépassement d'un seuil $(X \geq x_{seuil})$. La courbe de fiabilité est obtenue en traçant les fréquences d'observation de l'événement en fonction de probabilités prévues. Le diagramme se construit point par point, chaque point correspondant à une gamme de valeurs de probabilité prédite. Pour cela, l'intervalle de probabilités [0, 1] est découpé en sous-intervalles. Dans le cas d'une prévision d'ensemble et d'un événement de dépassement de seuil, les sous-intervalles sont définis par les probabilités de dépassement atteintes par l'ensemble. Il s'agit donc de la suite d'intervalles $([0, \frac{1}{N_{Ens}}], [\frac{1}{N_{Ens}}, \frac{2}{N_{Ens}}], ..., [\frac{N_{Ens}-1}{N_{Ens}}, 1])$. Pour un intervalle de probabilités donné $[p_{i-1}, p_i]$, on considère l'ensemble des prévisions parmi les N_{prev} prévisions pour lesquelles la probabilité estimée de l'événement sur la base de l'ensemble est dans cet intervalle. Parmi ces prévisions, on calcule la fréquence f_i d'observation de l'événement. Le point $(\frac{p_{i-1}+p_i}{2}, f_i)$ est placé sur le graphique. La courbe de fiabilité d'un ensemble par-

faitement fiable est la première bissectrice. Un point en dessous de la première bissectrice correspond à une surestimation de la probabilité, et inversement. La figure 3.5 présente un exemple de courbe de fiabilité.

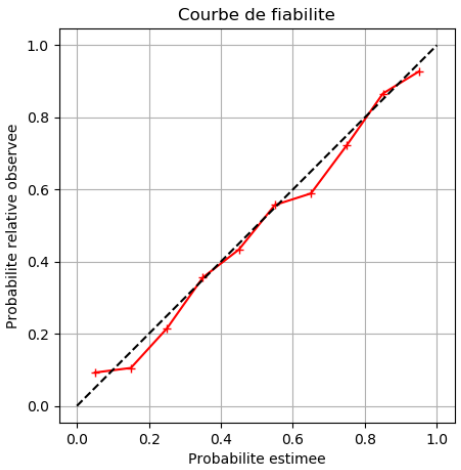


Figure 3.5: Exemple de courbe de fiabilité.

Les courbes ROC Une propriété intéressante des prévisions d'ensemble est leur capacité à discriminer les situations menant à l'occurrence d'un événement (par exemple, le dépassement d'un seuil) des situations menant à la non-occurrence de cet événement. Le diagramme ROC (*Relative Operating Characteristic*) est un graphique qui permet d'analyser la capacité d'un système à réaliser correctement cette discrimination. Il permet donc d'estimer la résolution pour un événement donné.

La courbe ROC est tracée point par point, en faisant varier des seuils de probabilité testés. En pratique, l'ensemble des seuils testés correspond aux probabilités de dépasser le seuil critique par 0, 1, 2, ... ou N_{ens} membres. L'ensemble des seuils testés est donc $(0, \frac{1}{N_{ens}}, \frac{2}{N_{ens}}, ..., 1)$. Pour chaque seuil de probabilité testé p, on construit la table de contingence du système relative à cet événement (Tableau 3.1). Elle permet de classer les prévisions (couples $(prévision, observation)_i$) selon la réalisation ou non de l'événement. Pour chaque prévision, si on appelle N_{EVT_i} le nombre de membres qui prévoient l'événement, la probabilité prévue de l'événement est $p_i = \frac{N_{EVT_i}}{N_{ens}}$. L'événement est prévu par cette prévision pour le seuil p si $p_i \geq p$.

À partir du tableau de contingence on peut calculer :

- le taux de réussite (*Hit Rate*) : $H = \frac{a}{a+c}$;
- le taux de faux positifs (*False Alarm Rate*) : $F = \frac{b}{b+d}$.

La courbe ROC représente le taux de réussite H en fonction du taux de faux positifs F, lorsqu'on fait varier le seuil de probabilité p de l'événement étudié. Le taux de faux positifs

est représenté sur l'axe des abscisses, le taux de réussite sur l'axe des ordonnées. Le point de coordonnées (F, H) est donc ajouté sur la courbe ROC.

	Dépassement observé	Dépassement non observé
Dépassement prévu	a	b
Dépassement non prévu	c	d

Tableau 3.1: Table de contingence pour le dépassement d'un seuil relative à un seuil de probabilité p.

Nous détaillons ci-après les caractéristiques des points formant une courbe ROC :

- Toutes les prévisions prévoient l'événement avec un seuil de probabilité $p = 0$. Pour ce seuil, $c = d = 0$, et $H = F = 1$. C'est le point en haut à droite du graphique.

- Si aucune prévision ne prévoit l'événement avec un seuil de probabilité donné, alors $a = b = 0$, et $H = F = 0$. C'est le point origine en (0,0). Dans tous les cas on fermera le graphique par ce point.

- Pour chaque seuil de probabilité testé autre que $p = 0$, une prévision parfaite a un taux de réussite égal à 1 (H=1) et un taux de faux positifs nul (F=0). Ainsi, plus le point de la courbe ROC est proche du point (0,1) (point en haut à gauche du graphique), meilleure est la capacité de discrimination du système.

- Pour une courbe ROC égale à la première bissectrice, le système n'est pas capable de discriminer entre les bonnes détections et les faux positifs. C'est le cas d'un système climatologique qui n'a pas de résolution.

À partir d'une courbe ROC, on peut définir le score AUC (*Area Under Curve*) correspondant à l'aire située sous la courbe. Plus AUC est proche de 1, meilleure est la discrimination du système de prévision pour l'événement testé. Un score AUC égal à 0.5 correspond à un système de prévision incapable de discriminer. La figure 3.6 représente des courbes ROC pour différents systèmes de prévision. Dans le cas (a), le système présente une discrimination parfaite pour l'événement testé. Dans le cas (b), le système ne sait pas discriminer. Le cas (c) est caractéristique d'un système de prévision réel, dont la discrimination est imparfaite.

Il est à noter que la courbe ROC ne permet pas d'évaluer la fiabilité de la prévision. En effet, la prévision climatologique est parfaitement fiable, et sa courbe ROC est la première bissectrice. L'utilisation des courbes ROC doit donc être réalisée en compléments d'autres outils de vérification présentés dans ce chapitre.

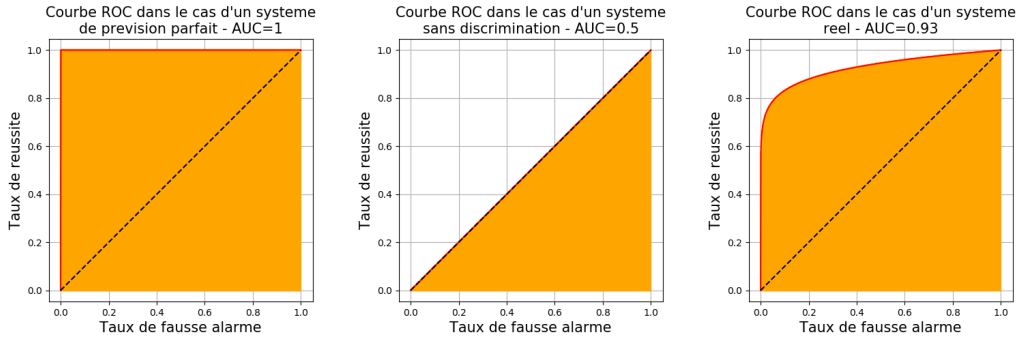


Figure 3.6: Courbes ROC pour différents systèmes de prévision : (a) discrimination parfaite, (b), aucune discrimination et (c) cas d'un système réaliste.

3.1.2.2 Histogramme de rang, CRPS et RMSE

Contrairement aux outils présentés au paragraphe précédent 3.1.2.1 qui visent à examiner les performances d'un ensemble vis-à-vis d'un événement binaire, l'histogramme de rang, le CRPS et la RMSE sont des outils qui évaluent les performances d'un ensemble (ou de sa moyenne dans le cas de la RMSE) sans le comparer à un événement.

L'histogramme de rang L'histogramme de rang, ou diagramme de Talagrand [Anderson, 1996, Hamill and Colucci, 1997, Talagrand et al., 1997], permet de connaître la fiabilité d'un système de prévision d'ensemble. Pour construire un histogramme de rang, on détermine pour chaque prévision (couple $(prévision, observation)$) le rang de l'observation parmi les membres de la prévision. En cas d'égalité de l'observation avec un ou plusieurs membres de l'ensemble, ou d'égalité de plusieurs membres de l'ensemble, le rang est affecté de manière aléatoire entre les différents rangs possibles. L'histogramme de rang tracé correspond à l'effectif du nombre d'observations dans chaque rang. Un ensemble fiable a un histogramme de rang plat, mais ce n'est pas une condition suffisante pour s'assurer de la fiabilité d'un ensemble [Hamill, 2001]. En effet, l'histogramme de rang mesure la fiabilité dans sa définition moins stricte (voir paragraphe 3.1.1.2) puisque la densité de probabilité des prévisions f_p est différente à chaque réalisation. [Hamill, 2001] montrent qu'un histogramme de rang plat peut être généré avec un ensemble non fiable, si différents biais conditionnels se compensent. Plus généralement, la forme de l'histogramme de rang nous donne une indication sur les caractéristiques et éventuels défauts de l'ensemble. La figure 3.7 illustre les formes caractéristiques que peut prendre un histogramme de rang. Les ensembles sous-dispersifs ont un histogramme en forme de U, les ensembles sur-dispersifs présentent une bosse, et les histogrammes non symétriques caractérisent un biais (les quantiles d'ordre élevé ou bas sont alors surreprésentés).

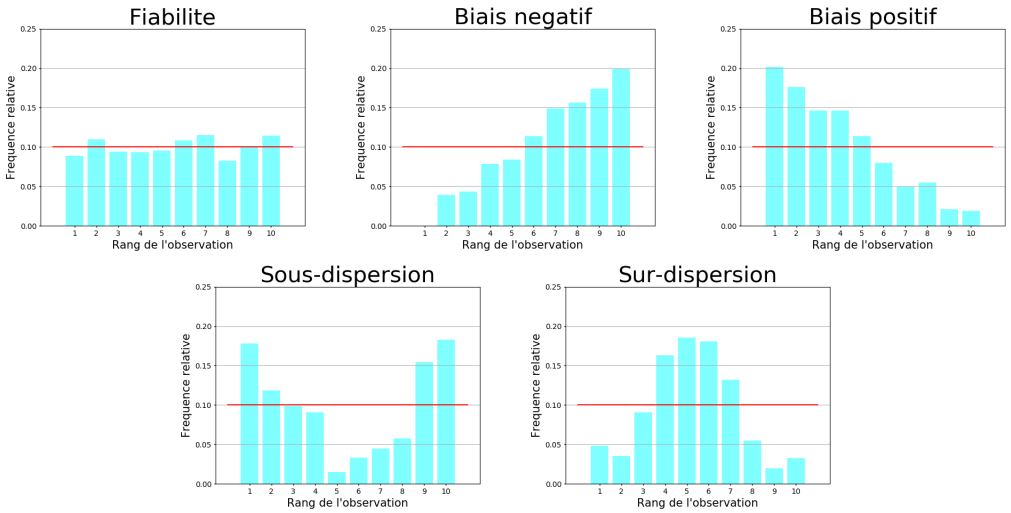


Figure 3.7: Formes caractéristiques des histogrammes de rang. Les caractéristiques des ensembles qui génèrent ces histogrammes de rang sont indiquées en titre de chaque graphique.

Candille and Talagrand [2005], puis Delle Monache et al. [2006] et Taillardat et al. [2016] ont introduit la notion de norme d'un histogramme de rang comportant $(N_{ens}+1)$ rangs. La norme $\|.\|_2$ est définie par :

$$\|.\|_2 = \sqrt{\sum_{i=1}^{N_{ens}+1} (f_i - \frac{1}{N_{ens}+1})^2} \tag{3.2}$$

où f_i représente la fréquence des observations dans le i^{me} rang. La figure 3.8 illustre la contribution de chaque rang au calcul de la norme. Pour un ensemble fiable, l'histogramme de rang est plat et la norme $\|.\|_2$ vaut 0, aux erreurs d'échantillonnage près. On définit les normes $\|.\|_1$ et $\|.\|_\infty$ de manière analogue.

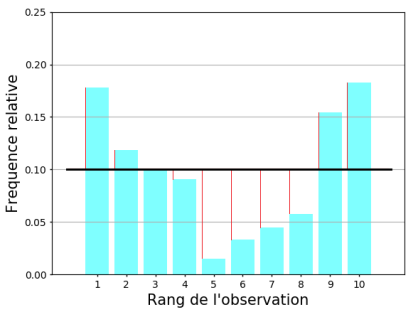


Figure 3.8: Visualisation de la contribution $|f_i - \frac{1}{N_{ens}+1}|$ (trait rouge) de chaque rang pour le calcul de la norme $\|.\|_2$.

Le Continuous Ranked Probability Score (CRPS) L'utilisation de scores propres permet d'attribuer à des prévisions probabilistes un résumé de leur performance en tenant compte à la fois de la fiabilité et de la résolution [Gneiting and Raftery, 2007]. Un score est dit strictement propre si et seulement si la valeur attendue pour le score est minimisée lorsque la distribution prédictive est issue de la même distribution que l'observation. Le *Continuous Ranked Probability Score* [Gneiting and Raftery, 2007, Hersbach, 2000, Matheson and Winkler, 1976] est un score propre, qui permet d'évaluer conjointement la fiabilité et la résolution. Il est orienté négativement, ce qui signifie que l'on cherche à obtenir les plus petites valeurs pour ce score. Pour une prévision i (couple $(prévision, observation)_i$), le $CRPS_i$ évalue la distance quadratique entre la fonction de répartition de la prévision F_i et la fonction échelon centrée sur l'observation notée H_i (voir définition 3.1) :

$$CRPS_i = \int_{-\infty}^{\infty} (F_i(x) - H_i(x))^2 dx. \tag{3.3}$$

Il s'exprime dans la même unité que l'observation. En pratique, le CRPS est calculé sur un large échantillon de prévisions. Si l'on appelle N_{prev} la taille de l'échantillon, il vient :

$$CRPS = \frac{1}{N_{prev}} \sum_{i=1}^{N_{prev}} CRPS_i. \tag{3.4}$$

La surface orange présentée sur l'image 3.9 illustre la contribution $CRPS_i$ d'un couple $(prévision, observation)_i$ au calcul du CRPS moyen. La contribution au CRPS d'une prévision dont la distribution est peu dispersée et centrée à proximité de l'observation sera nettement moins pénalisante que celle d'une distribution très dispersée ou très éloignée de l'observation.

Hersbach [2000] a montré que la CRPS peut être décomposé en la somme de trois termes caractéristiques de la fiabilité, de la résolution et d'une incertitude. Le terme d'incertitude est uniquement lié à la distribution des observations et est proportionnel à l'écart-type de la distribution. Il est donc constant pour un échantillon de N_{prev} prévisions donné. Il ne dépend pas de la prévision à évaluer mais des temps de validation considérés. La résolution et l'incertitude sont donc regroupées au sein d'un même terme $CRPS_{pot}$, qui correspond à la valeur du CRPS si l'ensemble était fiable. Le terme $CRPS_{pot}$ est d'autant plus important que la dispersion de l'ensemble est importante. Il vient :

$$CRPS = Reli + CRPS_{pot}. \tag{3.5}$$

Comme le CRPS, les deux termes $Reli$ et $CRPS_{pot}$ sont négativement orientés.

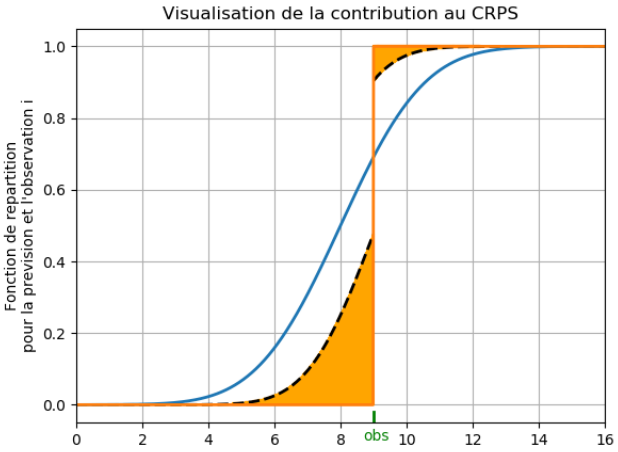


Figure 3.9: Visualisation graphique de la contribution d'un couple $(prévision, observation)_i$ au calcul du CRPS moyen. La courbe bleue représente la fonction de répartition de la prévision F_i. La courbe noire en pointillés représente F_i^2 ou $(1 - F_i)^2$ selon la position par rapport à l'observation. La contribution au CRPS moyen correspond à la surface orange.

Le CRPS peut être interprété comme la version probabiliste de l'erreur absolue moyenne (*Mean Absolute Error*, MAE) [Gneiting and Raftery, 2007], calculée par la formule :

$$MAE = \frac{1}{N_{prev}} \sum_{i=1}^{N_{prev}} |prev_i - obs_i| \tag{3.6}$$

où $prev_i$ représente la valeur de la prévision déterministe et obs_i l'observation associée. Le CRPS et la MAE sont deux scores qui peuvent être utilisés conjointement pour comparer les performances respectives d'un système déterministe et d'un système ensembliste.

RMSE Nous avons choisi dans ce travail d'utiliser la RMSE (*Root Mean Squared Error*) pour comparer les performances des systèmes de prévision d'ensemble construits et de la prévision déterministe actuellement opérationnelle. Le calcul de la RMSE est alors appliqué à la moyenne de la prévision d'ensemble. Si l'on appelle $prev_i$ la valeur de la prévision à évaluer, la RMSE est définie de la manière suivante :

$$RMSE = \sqrt{\frac{1}{N_{prev}} \sum_{i=1}^{N_{prev}} (prev_i - obs_i)^2}. \tag{3.7}$$

3.1.2.3 La stratification

Dans le domaine de la vérification, la stratification représente le découpage de l'échantillon des N_{prev} prévisions (couples $(prévision, observation)_i$) en un certain nombre de sous-échantillons les plus homogènes possibles. L'objectif est l'évaluation des performances des ensembles dans des conditions spécifiques. Dans le cadre de ce travail, la stratification est utilisée pour évaluer les prévisions d'ensemble réalisées en période de forts débits ou hautes eaux. Le concept de stratification s'applique uniquement aux outils utilisés dans un cadre global et non à ceux relatifs à la détection d'un événement. La stratification est ainsi appliquée aux histogrammes de rang et au CRPS selon les recommandations émises dans Bellier et al. [2017].

Pour réaliser une stratification, il est nécessaire de définir un critère de stratification. On peut dans un premier temps imaginer un critère basé sur l'observation, par exemple *« l'observation de débit dépasse la valeur 75 m^3/s »*, ce critère aboutissant à la création de deux sous-échantillons. Cette stratification est attractive puisqu'elle permet d'évaluer les prévisions uniquement lorsque les débits ont été significatifs. Cependant, ce type de stratification est proscrire. Bellier et al. [2017] ont montré que les histogrammes de rang d'ensembles fiables stratifiés de cette manière ne sont pas forcément plats. Par ailleurs, d'un point de vue strictement opérationnel, la mise en place d'un système de prévision d'ensemble dépendant de la valeur de l'observation est impossible puisque par définition l'observation n'est pas accessible au moment de la prévision. Il est donc recommandé d'utiliser un critère basé sur la prévision. Cette méthodologie permet de se rapprocher de la définition de la fiabilité dans sa version la plus stricte (paragraphe 3.1.1.2). Le cas extrême d'une stratification réalisée selon que la loi de probabilité prévue soit égale à f_p correspond exactement à cette définition. Ce type de stratification est également compatible avec l'évaluation du CRPS, qui permet de conserver sa propriété de *score propre* [Lerch et al., 2017].

3.1.3 Synthèse

La confrontation des prévisions aux observations couvre un champ thématique particulier appelé *vérification* dans le domaine hydrométéorologique.

Dans cette section, nous avons d'abord décrit les variables que nous allons vérifier dans la suite du travail. C'est la qualité des prévisions d'ensemble qui sera évaluée et non leur utilité, c'est-à-dire que l'on évaluera des sorties de modèles ensemblistes et non les décisions produites par les décideurs connaissant la prévision. Les prévisions à vérifier sont des prévisions d'ensemble discrètes. Nous avons également décrit les principales caractéristiques recherchées pour une prévision d'ensemble : nous nous attacherons à évaluer la fiabilité et la résolution des ensembles construits.

Dans un second temps, plusieurs outils permettant de vérifier les ensembles ont été présentés. Les diagrammes de fiabilité seront utilisés dans ce travail uniquement pour la vérification des pluies PEAROME (Tiberi-Wadier et al. [2018] et chapitre 4). Les histogrammes de rang, le CRPS et les courbes ROC seront utilisés pour évaluer les performances des prévisions d'ensemble météorologiques, hydrologiques et hydrauliques. Le score RMSE sera quant à lui uniquement utilisé pour la comparaison de la moyenne des prévisions d'ensemble hydrologiques et hydrauliques avec la prévision déterministe utilisée actuellement en opérationnel. Le concept de stratification, qui permet d'évaluer les performances des ensembles sur des sous-échantillons judicieusement choisis, a été présenté et sera utilisé dans la suite du travail pour se focaliser sur les périodes de forts débits ou de hautes eaux.

3.2 L'étude d'incertitude : un préalable commun aux deux approches

Ce paragraphe décrit le principe de la réalisation des études d'incertitudes. Il est en partie rédigé à partir de l'ouvrage Faivre et al. [2013]. Il existe une grande diversité de modèles mathématiques utilisés dans les sciences de l'environnement : météorologie, biologie, écologie, halieutique, agronomie, hydraulique, hydrologie... Tous ces modèles ont en commun d'être composés de quatre éléments [Faivre et al., 2013] :

(i) Des variables d'entrée : elles sont potentiellement issues de mesures.

(ii) Des variables de sortie : ce sont les variables d'intérêt qui sont le résultat de la simulation.

(iii) Des paramètres : leur valeur est à déterminer au préalable. Dans le cas des modèles hydrologiques et hydrauliques, ces valeurs sont déterminées par une procédure de calage-validation. Les paramètres à caler pour les modèles GRP, MORDOR-TS et MASCARET sont décrits dans le chapitre 2.

(iv) Des équations mathématiques : ces équations permettent d'estimer les variables de sortie, étant données les variables d'entrée et la valeur des paramètres. Les équations des modèles GRP, MORDOR-TS et MASCARET sont respectivement décrites en annexes M.2 et M.4 et au paragraphe 2.3.2.

Les erreurs de prédiction des modèles sont donc liées aux incertitudes sur les variables d'entrée, sur les paramètres des modèles et sur leur formulation mathématique, ce dernier point incluant également la résolution numérique des équations. L'analyse de sensibilité permet de connaître les principales sources d'incertitude parmi différents facteurs incertains considérés, et elle permet en outre de les hiérarchiser selon l'influence qu'ils ont sur les sorties du modèle. Elle permet de mesurer la contribution des différentes sources d'incertitude du modèle à l'incertitude des sorties. Pour mémoire, les principales incer-

titudes concernant les modélisations météorologiques, hydrologiques et hydrauliques ont été décrites au paragraphe 1.1.2.

3.2.1 Analyse locale ou globale

Les méthodes d'analyse de sensibilité sont classées dans deux grandes familles : les analyses de sensibilité *locales* et *globales*.

Historiquement, les analyses de sensibilité ont longtemps été menées de manière *locale*. On étudie alors la variabilité de la sortie $Y = f(X)$, induite par une variation de l'entrée X autour d'une valeur de référence X_0. Cette approche permet d'appréhender le comportement du modèle autour de certaines valeurs de paramètres et/ou variables d'entrée, mais elle ne permet pas de comprendre les effets de la variation des facteurs incertains dans une large gamme. Les méthodes locales sont basées sur des dérivées partielles du modèle en un point précis. Elles sont maintenant couramment utilisées grâce aux approches adjointes.

Les méthodes dites *globales* ont été développées à partir de la fin des années 1980 pour pallier les limites des méthodes locales (hypothèse de linéarité, variations locales uniquement). Elles permettent au contraire d'analyser la variabilité de la sortie $Y = f(X)$ lorsque X varie sur l'ensemble de son domaine de validité ou dans un domaine caractéristique de son incertitude [Saltelli, 2008]. L'incertitude peut alors porter sur les paramètres des équations du modèle, sur les conditions initiales ou conditions limites du modèle, et également sur des variables numériques de résolution des équations du modèle. Ces méthodes impliquent l'intégration du modèle un grand nombre de fois, ce qui permet de propager les incertitudes de l'espace des entrées vers la sortie. Elles nécessitent le plus souvent des moyens informatiques importants. Il existe de nombreuses méthodes d'analyse de sensibilité globale. Iooss and Lemaître [2015] font la synthèse des principales méthodes existantes.

3.2.2 Différentes méthodes d'analyse de sensibilité globale

3.2.2.1 Une approche commune

Toutes les méthodes ont pour objectif le calcul d'indices de sensibilité qui caractérisent la contribution des différents facteurs d'entrée incertains à la variabilité de la sortie du modèle, également appelée la quantité d'intérêt. Le calcul de ces indices n'étant que rarement possible de manière analytique, leur estimation est généralement réalisée grâce à une estimation stochastique à partir d'un ensemble de réalisations du modèle numérique. L'analyse de sensibilité est alors menée selon les quatre étapes suivantes :

(i) Choix des k entrées incertaines dont l'on souhaite étudier l'influence et définition des distributions de probabilité caractérisant l'incertitude de chacun de ces paramètres d'entrée. Cette étape est essentielle car les résultats de l'étude de sensibilité en dépendent

directement.

(ii) Tirage d'un échantillon de N jeux d'entrée $X = (X_1, ..., X_k)$ selon les distributions de probabilité déterminées. Des contraintes peuvent être imposées pour le tirage selon la méthode d'analyse de sensibilité retenue (nombre N, algorithme pour générer les N valeurs). La taille de l'échantillon choisie impacte également les résultats si celle-ci est trop petite pour assurer la convergence des méthodes de calcul de la sensibilité.

(iii) Calcul par simulation numérique de la sortie du modèle $Y_i = f(X_i)$ pour chaque jeu d'entrée X_i.

(iv) Calcul des indices de sensibilité selon la méthode choisie. Ces indices peuvent mettre en évidence des interactions entre les variables d'entrée pour la valeur de la quantité d'intérêt simulée.

Les indices de sensibilité de la sortie par rapport aux entrées peuvent être obtenus par criblage ou régression linéaire dans le cas de modèles linéaires. Ces méthodes sont décrites en annexe N.1.1. Pour les modèles non-linéaires, ces méthodes ne sont plus applicables. On peut alors utiliser des méthodes basées sur la décomposition de la variance. Elles sont décrites au paragraphe 3.2.2.2. Enfin, la méthode basée sur les moments consiste à analyser des densités de probabilité non conditionnelles et conditionnelles en sortie de modèle. Elle permet de pallier les limites des méthodes ANOVA (hypothèse de normalité et de paramètres indépendants). Elle est décrite en annexe N.1.3.

3.2.2.2 Méthodes basées sur la décomposition de la variance

Pour les modèles non-linéaires, les méthodes basées sur la régression linéaire décrites en annexe N.1.1 ne sont pas applicables. Les méthodes basées sur l'analyse de la variance peuvent alors être utilisées (méthodes dites ANOVA, pour ANalysis Of VAriance, Efron and Stein [1981]). Elles sont valables pour les modèles non-linéaires et non monotones, avec l'hypothèse que la distribution des variables est normale puisque l'on se focalise sur le calcul des premier et second moments (moyenne et variance) [Saltelli and Bolado, 1998]. Ces méthodes sont parmi les plus utilisées [Ferretti et al., 2016]. Dans les méthodes ANOVA, la variance de la sortie du modèle est analysée lorsque les paramètres d'entrée varient simultanément sur leur domaine d'incertitude. Elles consistent en la décomposition de la variance de la variable de sortie en tant que somme de variances élémentaires associées aux différents paramètres d'entrée. Les indices de sensibilité sont les variances élémentaires normalisées.

Si X est considérée comme une variable aléatoire, alors la sortie du modèle $Y = f(X)$ est également une variable aléatoire. En supposant que l'espérance et la variance de Y sont finies et que les variables d'entrée sont indépendantes, la variance de Y peut être écrite

grâce à la décomposition de Hoeffing [Hoeffding, 1948] :

$$V(Y) = \sum_i V_i + \sum_i \sum_{j>i} V_{i,j} + ... + V_{1,2,3,...,k} \tag{3.8}$$

où V_i représente la contribution élémentaire de X_i à la variance de sortie $V(Y)$, et chaque V indicé par un sous-ensemble U représente la contribution due aux interactions entre les entrées $X_i, i \in U$ à la variance totale de sortie $V(Y)$. La division de l'équation 3.8 par V(Y) permet de définir les indices de sensibilité :

$$1 = \sum_i S_i + \sum_i \sum_{j>i} S_{i,j} + ... + S_{1,2,3,...,k} \tag{3.9}$$

où $S_i = \frac{V_i}{V(Y)}$ est l'indice du premier ordre de X_i et représente la contribution élémentaire normalisée de X_i à la variance totale $V(Y)$, $S_{i,j} = \frac{V_{i,j}}{V(Y)}$ est l'indice du second ordre de X_i and X_j et représente la contribution normalisée due aux interactions entre X_i et X_j à la variance totale $V(Y)$, et $S_{1,2,...,k} = \frac{V_{1,2,...,k}}{V(Y)}$ représente la contribution normalisée des interactions de toutes les entrées à la variance totale $V(Y)$. Ainsi, les indices de sensibilité sont compris entre 0 et 1 et la somme de tous les indices de sensibilité est égale à 1. La somme des indices de sensibilité du premier ordre est égale à 1 s'il n'y a pas d'interaction entre les paramètres d'entrée. Plus une variable d'entrée est influente, plus l'indice de sensibilité correspondant est élevé. Les indices totaux ST_i représentent toutes les contributions relatives à une variable d'entrée X_i :

$$S_{T_i} = S_i + \sum_{j \neq i} S_{i,j} + \sum_{j \neq i, k \neq i, j<k} S_{i,j,k} + ... + S_{1,2,3,...,k} = \sum_{l \in \sharp i} S_l \tag{3.10}$$

où $\sharp i$ représentent tous les sous-ensembles de $\{1, ..., k\}$ contenant i. La figure 3.10 représente l'ensemble des indices de sensibilité existants dans le cas simplifié d'une étude de sensibilité avec 3 facteurs incertains. Les termes situés dans la zone jaune correspondent aux termes pris en compte pour le calcul de l'indice total relatif à la variable X_1.

Dans la pratique, les indices de sensibilité ne sont pas tous calculés, et certains présentent plus d'intérêt que d'autres :

(i) Les indices du premier ordre sont liés aux effets principaux. Ils permettent le plus souvent d'expliquer une part majeure de la variabilité de sortie. Ils sont utilisés pour hiérarchiser les facteurs d'entrée (*Factor prioritization*, Saltelli [2008]). Ils permettent ainsi de connaître les paramètres qui sont davantage responsables de la variance de sortie du modèle. C'est donc sur ces paramètres qu'il faut réduire l'incertitude pour réduire l'incertitude sur Y.

(ii) Les indices totaux S_{T_i} fournissent une mesure de la variance moyenne de Y lorsque

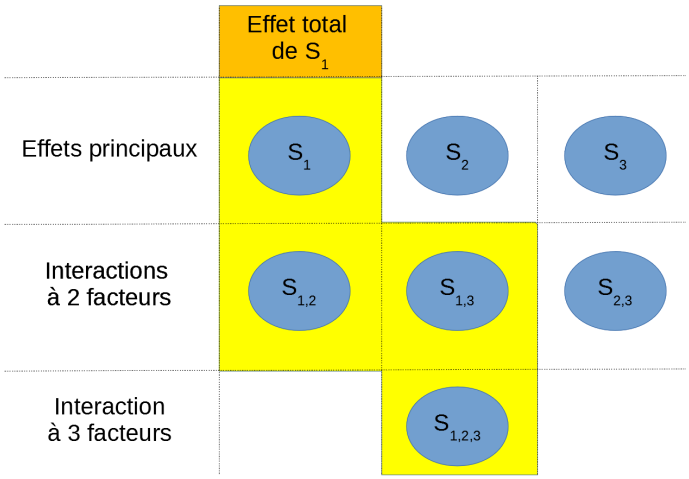


Figure 3.10: Indices de sensibilité dans le cas simplifié avec 3 facteurs incertains. Les indices situés dans la région jaune sont les termes pris en compte pour le calcul de l'indice total S_{T_1}. Schéma réalisé d'après Faivre et al. [2013].

toutes les entrées sauf X_i sont fixées. Ces indices permettent de déterminer les paramètres qui peuvent être fixés sans conséquence sur la sortie, quelles que soient les valeurs des autres entrées (*Factor Fixing*, Saltelli [2008]).

L'annexe N.1.2 présente deux méthodes de calcul des indices de sensibilité des méthodes ANOVA : la méthode FAST et la méthode de Sobol [Sobol, 2001], cette dernière étant la méthode utilisée dans cette thèse. La méthode de Sobol est basée sur une procédure stochastique. Elle nécessite un très grand nombre de simulations, classiquement plusieurs dizaines de milliers. Pour les modèles coûteux en temps de calcul, on peut appliquer cette méthode aux résultats d'un métamodèle. Le métamodèle, ou *surface de réponse*, a pour objectif la construction d'une fonction qui simule la réponse du modèle dans une certaine gamme de variation de ses variables d'entrée, à partir d'un certain nombre d'évaluations directes du modèle [Moçayd et al., 2017, Nabil, 2017, Roy et al., 2017, Sacks et al., 1989].

3.2.3 Mise en œuvre informatique

Dans cette thèse, les indices de sensibilité calculés sont les indices de Sobol. Les calculs sont calculés grâce aux modules python OpenTURNS[1] et OT-Batman[2,3] [T.Roy et al., 2018] sur le calculateur Datarmor hébergé et administré par l'IFREMER. Les étapes suivantes ont été suivies :

(i) Génération de l'échantillon $X_i = (X_1, ..., X_k)_i$ de taille N. N est un multiple de $k + 2$ et l'échantillon est généré selon les contraintes imposées par l'algorithme de Saltelli

[1] http://openturns.org/
[2] https://gitlab.com/cerfacs/batman
[3] https://batman.readthedocs.io/en/develop/

[Saltelli et al., 2010] décrit en annexe N.1.2. Le tirage est réalisé selon une suite de Sobol.

(ii) Réalisation des simulations pour chaque entrée X_i. Les simulations sont lancées en multiprocesseur sur un nœud du calculateur Datarmor qui possède 56 processeurs.

(iii) Relecture des résultats des simulations et post-traitement : calcul des indices de Sobol et calcul des densités de probabilité de l'ensemble.

La convergence des indices de Sobol a été assurée en réalisant des calculs pour des échantillons de taille croissante. La convergence est supposée atteinte lorsque les résultats sont très proches pour deux tailles d'échantillons croissantes.

3.2.4 Études de sensibilité dans les domaines hydrologiques et hydrauliques

Les études de sensibilité globales sont largement utilisées en hydrologie et en hydraulique pour appréhender et hiérarchiser les différentes sources d'incertitude dans la modélisation des débits et/ou hauteurs d'eau. Parmi les différentes méthodes, les méthodes de décomposition de variance ANOVA sont couramment utilisées. Les coûts de calculs étant significativement plus importants avec les modèles hydrauliques, le recours à la métamodélisation peut être réalisé. Avant de présenter une revue d'études de sensibilité globale réalisées pour la modélisation hydrologique ou hydraulique (paragraphe 3.2.4.2), nous présentons au préalable au paragraphe 3.2.4.1 la méthode GLUE qui connaît un fort succès en hydrologie.

3.2.4.1 La méthode GLUE

En parallèle à l'utilisation des méthodes d'analyse de sensibilité globales, la méthode GLUE (*Generalized Likehood Uncertainty Estimation*, Beven and Binley [1992]) est largement utilisée en hydrologie. Bien qu'elle ne s'inscrive pas dans les méthodes d'analyse de sensibilité classiques décrites au paragraphe 3.2.2, sa description est importante pour la compréhension de la bibliographie sur l'évaluation des incertitudes en hydrologie.

Cette méthode part du principe qu'il n'existe pas un unique jeu de paramètres capable de représenter le modèle de manière acceptable. GLUE permet d'identifier les entrées les plus vraisemblables connaissant les observations. Elle permet ainsi d'appréhender l'incertitude des modèles hydrologiques liée à l'estimation des paramètres. Un échantillonnage de tout l'espace des paramètres est réalisé de manière à explorer une large gamme de comportements du modèle. La performance des différents jeux de paramètres est évaluée selon un ou plusieurs critères. Une valeur seuil doit être choisie pour classer les jeux de paramètres acceptables (*behavioral*) ou non (*non-behavioral*), au regard des critères d'efficacité choisis. À chaque réalisation du modèle est attribué un poids traduisant sa vraisemblance, i.e. sa capacité à représenter l'échantillon de calibration. Les réalisations de modèles considérées non acceptables sont affectées d'un poids nul. Les débits simulés avec les différents jeux de

paramètres retenus sont utilisés pour caractériser l'incertitude prédictive, avec l'hypothèse que les erreurs de modélisation sont les mêmes en période de calage et en prévision. Cette technique permet en parallèle de construire la fonction de densité de probabilité des entrées. Cette méthode a été critiquée à cause de l'impact des choix subjectifs inhérents à son utilisation : choix de la fonction objectif, choix du seuil qui permet de sélectionner les jeux de paramètres, méthode d'échantillonnage et gamme de variation des paramètres. L'importance de ces choix a été discutée dans Beven [2006], Beven and Freer [2001], Freer et al. [1996]. La simplicité de sa mise en œuvre est vraisemblablement la raison de sa large utilisation.

3.2.4.2 Applications de la méthode GLUE et études ANOVA

Dans ce paragraphe, on décrit tout d'abord plusieurs études d'analyse de sensibilité réalisées pour des modélisations de complexité croissante, allant de la modélisation hydrologique petite échelle à la modélisation hydraulique 1D ou 2D. Par ailleurs, plusieurs publications ont pour objectif la comparaison des résultats associés à l'utilisation de différentes méthodes d'analyse de sensibilité, et nous nous attacherons à décrire les principaux résultats.

Garambois et al. [2013] calculent les indices de Sobol pour 5 paramètres du modèle hydrologique distribué MARINE [Roux et al., 2011] pendant 10 événements de crues éclair dans des bassins méditerranéens. Aux pics d'événements, la profondeur du sol explique plus de 80 % de la variance du débit modélisé. Dans Michon and Castaings [2017], une étude de sensibilité globale sur le score KGE (Kling-Gupta Efficiency, Gupta et al. [2009]) du modèle semi-distribué MORDOR-TS fonctionnant au pas de temps horaire ou journalier a été menée sur plusieurs bassin français. L'objectif est de guider les stratégies de spatialisation des paramètres. Pappenberger et al. [2010] réalisent une étude de sensibilité globale avec le calcul des indices de Sobol et une étude d'incertitude GLUE sur le modèle de routage global TRIP2 (Total Runoff Integrating Pathways-2, Ngo-Duc et al. [2007]) couplé avec le modèle de surface HTESSEL (Hydrology Tiled ECMWF Scheme of Surface Exchanges over Land, Balsamo et al. [2009]). Il s'avère que c'est le paramètre de latence pour les eaux souterraines qui est le paramètre de calage le plus important. Emery et al. [2016] réalisent également une étude ANOVA pour mettre en évidence les paramètres clés qui impactent le modèle de routage TRIP2 sur le bassin de l'Amazone, dans le but de corriger ensuite les paramètres les plus influents sur la modélisation des débits et hauteurs d'eau par assimilation de données. Les paramètres géomorphologiques expliquent environ 95 % de la variance sur les hauteurs d'eau modélisées : coefficient de Manning, pente du lit, largeur de la rivière. Les résultats montrent ici aussi que le débit est très dépendant de la valeur du paramètre de latence pour les eaux souterraines, qui explique plus de 90 % de la variance. Les paramètres géomorphologiques influencent les

valeurs de débit modélisées en montée et descente de crue. Romanowicz and Beven [1998] appliquent la méthode GLUE à la prévision distribuée dans le temps et l'espace des probabilités d'inondation en chaque point d'une plaine inondable. Dans Hall et al. [2005], les auteurs réalisent une étude de sensibilité globale pour évaluer l'influence des différentes zones de frottement pour la modélisation de la zone inondée et des hauteurs d'eau avec le modèle 2D LISFLOOD-FP [Bates and De Roo, 2000, Bates et al., 2010] sur un tronçon de la rivière Thames (United Kingdom). Dans Thomas Steven Savage et al. [2016], les auteurs ont pour objectif d'évaluer l'influence relative de la résolution spatiale du modèle hydraulique LISFLOOD-FP par rapport aux autres variables d'entrée incertaines que sont le coefficient de frottement et les débits amont pour la modélisation de la zone inondable, de la hauteur d'eau et de la durée d'inondation. Les résultats mettent en évidence que les facteurs les plus importants dépendent de la variable de sortie étudiée, et également de la localisation du point d'intérêt et du moment dans l'événement de crue : ce sont d'abord les hydrogrammes d'entrée pendant la phase de montée, puis les coefficients de frottement en lit mineur autour du pic de l'événement, puis le coefficient de frottement en lit majeur pendant la régression de la crue. L'influence de la résolution spatiale, prise en compte par ré-échantillonnage des données topographiques, est plus marquée pour la prévision des hauteurs d'eau, bien que celle-ci soit également sensible aux autres facteurs d'entrée. Laborie et al. [2020] réalisent une étude d'incertitude globale du modèle opérationnel TELEMAC2D utilisé pour la prévision des hauteurs d'eau dans l'estuaire de la Gironde. Les sources d'incertitudes sont hiérarchisées grâce au calcul des indices de Sobol lors d'un événement de tempête en 2003. Les variables incertaines considérées sont les coefficients de frottement et les chroniques pour les conditions limites amont (chroniques de débit) et aval (cote à la condition limite maritime). Les conditions limites, variables en temps, sont supposées être représentées par un Processus Gaussien, caractérisé par une longueur de corrélation temporelle calculée à partir des observations. Les auteurs mettent en évidence un comportement différencié par zone : (i) dans la zone estuarienne, la condition limite maritime et les coefficients de Strickler ont un rôle prépondérant, avec une influence cyclique liée à la phase de la marée, et (ii) dans les zones fluviales à l'amont, ce sont les coefficients de Strickler et la condition limite amont qui sont prédominants. Deux études [Moçayd et al., 2017, Roy et al., 2017] mettent en évidence une bonne efficacité de l'utilisation de métamodèles pour représenter les équations SWE (*Shallow Water Equations*) avec le modèle 1D MASCARET. Dans Moçayd et al. [2017], les auteurs comparent l'utilisation d'un modèle réduit de type polynômes du chaos généralisés par rapport à une approche de Monte Carlo classique pour l'estimation de la densité de probabilité de la hauteur d'eau modélisée. Il en résulte que l'utilisation du métamodèle, nécessitant seulement 49 évaluations directes du modèle MASCARET, permet de représenter correctement la distribution des hauteurs d'eau ainsi que la matrice de covariance des erreurs sur la hauteur d'eau modélisée. Dans Roy et al. [2017], les auteurs comparent la qualité de deux modèles

réduits : décomposition en polynômes du chaos et approximation par processus gaussiens (appelé également *kriegeage*). Les deux méthodes présentent des résultats intéressants pour l'estimation des fonctions de densité de probabilité, des indices de sensibilité et des matrices de covariance des hauteurs d'eau modélisées, à moindre coût par rapport à un calcul de référence de type Monte Carlo. En effet, un nombre réduit de simulations (10 à 100) du code direct MASCARET est suffisant pour construire les modèles réduits.

Plusieurs auteurs analysent les résultats comparés obtenus avec différentes méthodes d'analyse de sensibilité. Dans Pappenberger et al. [2008], les auteurs comparent plusieurs méthodes d'analyse de sensibilité globale (Sobol, Morris, étude régionalisée, régression) sur le modèle hydraulique 1D HEC-RAS. Ils concluent que les différentes méthodes hiérarchisent les incertitudes de manière différente. Dans Hall et al. [2009], les auteurs font la revue de plusieurs méthodes pour la réalisation d'études de sensibilité (méthodes locales, ANOVA). Ils mettent en évidence que les méthodes locales fournissent une compréhension incomplète de la réponse du modèle sur une large gamme de variation des entrées. Les méthodes ANOVA sont plus performantes pour représenter les effets non-linéaires et les interactions entre les variables. Goutal et al. [2017] comparent différentes méthodes d'analyse de sensibilité locales ou globales pour la modélisation d'un tronçon de la Garonne avec différents solveurs 1D ou 2D. Khorashadi Zadeh et al. [2017] comparent le calcul des indices de Sobol et la méthode basée sur les moments PAWN pour 26 paramètres du modèle SWAT (Soil and Water Assessment Tool, Arnold et al. [1998]). Les deux méthodes mettent en évidence les 12 mêmes paramètres influents, mais la hiérarchisation des paramètres est différente selon les méthodes. Les auteurs formulent l'hypothèse que cette différence est probablement dûe aux limites de la méthode ANOVA quand les distributions de sortie ne sont pas symétriques.

3.2.5 Synthèse

Dans cette section ont été exposés les objectifs poursuivis lorsqu'on met en œuvre une étude de sensibilité, et la méthode de décomposition de la variance ANOVA a été décrite. Dans la suite du travail, les analyses de sensibilité des modèles hydrologiques GRP et MORDOR-TS et du modèle hydraulique MASCARET sur le bassin de l'Odet seront réalisées de manière globale selon la méthodologie de décomposition de la variance, avec le calcul des indices de Sobol. Ces travaux font respectivement l'objet des chapitres 4 et 7. Étant donné que ces modèles ne sont pas très coûteux en temps de calcul, l'évaluation des modèles sera réalisée de manière directe, sans construction de métamodèle.

3.3 La calibration statistique pour corriger les ensembles de prévision bruts

Les approches ensemblistes ont pour objectif de représenter les incertitudes de prévision. Cependant, celles-ci sont rarement complètement couvertes. Les prévisions d'ensemble météorologiques sont réputées sous-dispersives et biaisées pour les variables de surface [Hamill and Colucci, 1997, Schaake et al., 2010]. Les prévisions d'ensemble hydrologiques et hydrauliques brutes, c'est-à-dire sans post-traitement, n'ont pas non plus les caractéristiques recherchées (fiabilité et résolution, voir paragraphe 3.1.1.2). Elles sont le plus souvent caractérisées par des biais concernant leur moyenne et/ou leur dispersion [Ramos, 2018]. Le post-traitement des ensembles bruts, appelé encore *calibration statistique*, est donc indispensable avant une utilisation opérationnelle des ensembles.

La calibration statistique des ensembles est un domaine de recherche actif en météorologie et hydrologie. C'est l'exploitation de la relation entre les prévisions passées et les observations correspondantes qui permet de corriger les prévisions [Wilks, 2018b]. Nous considérons toujours un ensemble de N_{prev} réalisations d'une prévision d'ensemble à une échéance donnée comportant N_{ens} membres ainsi que l'observation correspondante tel que présenté à la figure 3.4. Cet ensemble de N_{prev} couples $(prévision, observation)$ constitue un échantillon d'apprentissage. Pour mettre en place une calibration statistique permettant de corriger des prévisions à une échéance donnée, on va investiguer de manière statistique la relation entre les valeurs des prévisions et la valeur des observations correspondantes. Wilks [2018b] présente l'état de l'art concernant les méthodes de post-traitement des ensembles météorologiques, tandis que Li et al. [2017] fait une revue des méthodes de post-traitement statistique des ensembles hydrologiques. Un grand nombre de ces méthodes sont bien sûr communes. Toutes ces méthodes ont pour objectif la construction d'un modèle statistique qui relie les prévisions aux observations. L'étape préalable de calage permet d'apprendre cette relation à partir d'un échantillon historique de couples $(prévision, observation)$. Une fois calé, le modèle peut être utilisé en opérationnel. Il permet alors de fournir pour chaque nouvelle prévision brute (i.e. avant post-traitement) une estimation de la loi de la densité de probabilité de l'observation sachant la prévision. Cette densité conditionnelle remplace alors la prévision brute.

On peut classer les méthodes selon qu'une hypothèse est formulée ou non sur la forme de la variable à prévoir. Dans les méthodes dites *paramétriques*, la densité de probabilité de la variable à prévoir est prescrite initialement. Les principales méthodes paramétriques utilisées en météorologie et en hydrologie sont présentées en annexe N.2.1 : il s'agit la méthode EMOS (*Ensemble Model Output Statistics*, Gneiting et al. [2005]) et la méthode BMA (*Bayesian Model Averaging*, Raftery et al. [2005]). Au contraire, dans les méthodes dites *non paramétriques*, aucune hypothèse n'est formulée sur la forme de la variable à prévoir. Si ces méthodes sont largement utilisées dans le domaine météorologique, elles sont à notre

connaissance appliquées de manière plus confidentielle dans le domaine de l'hydrologie et de l'hydraulique. Les méthodes non paramétriques sont entièrement basées sur les données, contrairement aux méthodes paramétriques qui se basent sur une densité de probabilité paramétrique, donc sur une forme mathématique pré-imposée. L'annexe N.2.2 présente deux méthodes non paramétriques que sont la recalibration d'histogrammes de rang et la régression quantile. Le paragraphe 3.3.1 ci-après présente quant à lui la méthode non paramétrique *Quantile Regression Forests* [Meinshausen, 2006] et sa mise en œuvre dans le cadre de ce travail de thèse. Cette méthode est utilisée pour la calibration statistique à chaque maillon de la chaîne de prévision : pour la calibration des prévisions d'ensemble de pluie PEAROME, des prévisions d'ensemble de débit hydrologiques et des prévisions d'ensemble de hauteur d'eau avec le modèle hydraulique.

Toutes les méthodes décrites dans cette section (corps principal et annexes) permettent de réaliser un post-traitement univarié, c'est-à-dire qu'elles sont appliquées à une seule variable, en un point, et pour une échéance de prévision. Elles ne sont donc pas appropriées pour la prise en compte des structures de dépendance spatiales et temporelles [Wilks, 2018b]. Il est cependant indispensable d'en tenir compte pour de nombreuses applications. Dans notre cas particulier, l'utilisation des prévisions d'ensemble au sein de la chaîne de prévision doit impérativement tenir compte de la dépendance temporelle. Nous avons besoins de membres, c'est-à-dire de chroniques, en entrée des modèles hydrologiques et hydrauliques. Une étape de reconstruction des membres à l'issue de la calibration statistique est donc nécessaire. Cette étape sera décrite au chapitre 3.3.3.

3.3.1 Quantile Regression Forests (QRF)

Les méthodes d'apprentissage statistique [Hastie et al., 2009] représentent une grande famille de méthodes non paramétriques. Ce sont des méthodes coûteuses en temps de calcul, et leur application au post-traitement de prévisions d'ensemble est relativement récent. Elles nécessitent un important volume de données d'apprentissage. En contrepartie, elles sont extrêmement flexibles, s'adaptent aux données d'entrée, et sont adaptées aux relations non-linéaires entre les entrées et sorties [Wilks, 2018b]. La méthode QRF (*Quantile Regression Forests*, Meinshausen [2006]) fait partie de la famille des méthodes d'apprentissage statistique. Elle est détaillée dans Taillardat et al. [2019], Zamo et al. [2014]. C'est à partir de ces publications que la description de la méthode est réalisée dans ce paragraphe.

Cette méthode d'apprentissage statistique est une généralisation de la méthode des forêts aléatoires (*random forests*, Breiman [2001]), qui, au lieu d'estimer la moyenne conditionnelle, permet d'estimer des quantiles conditionnels de la variable de réponse, donc une approximation de la distribution conditionnelle. Le principe de la méthode QRF est l'agrégation des observations selon leur prévision, cette dernière étant résumée par dif-

férents prédicteurs. L'hypothèse sous-jacente est que des prédicteurs proches mènent à des observations proches. En ce sens, cette méthode se rapproche de la méthode des analogues [Delle Monache et al., 2013, Hamill and Whitaker, 2006, Zalachori et al., 2012]. Elle consiste en la construction d'une forêt aléatoire constituée d'un ensemble d'arbres de décisions binaires, chacun de ces arbres étant construit à partir d'un ensemble de prédicteurs. Nous décrivons ci-après la méthodologie de construction des forêts aléatoires, en abordant les points suivants : choix des prédicteurs, construction d'un arbre de décision binaire, construction d'une forêt. L'application de la méthode QRF pour calibrer une nouvelle prévision à partir de la forêt aléatoire est ensuite décrite. On suppose que l'on dispose d'un échantillon d'apprentissage constitué de N_{prev} couples $(prévision, observation)_i$, où chaque prévision est constituée de N_{ens} membres d'une prévision d'ensemble.

Choix des prédicteurs Plusieurs types de prédicteurs peuvent être utilisés pour appliquer la méthode QRF. Ils peuvent naturellement être des statistiques de la variable à calibrer calculées à partir de l'ensemble : moyenne, variance, quantiles. Ils peuvent aussi être des statistiques relatives à la prévision d'autres variables. Par exemple, des statistiques relatives à la prévision de température ou d'humidité peuvent être utilisées pour calibrer une précipitation. Enfin, les prédicteurs peuvent ne pas être liés du tout à la prévision émise. Des mesures avant ou au moment de la prévision peuvent par exemple être pertinentes, comme le débit mesuré au temps de base pour la calibration des ensembles hydrologiques. Lorsque le prédicteur prend sa valeur dans $\mathbf{R}$, on dit que le prédicteur est *quantitatif*. Le prédicteur peut aussi être *qualitatif*, ou *à modalités* lorsqu'il est caractérisé par une classe. Par exemple, le prédicteur *moment de la prévision* dont les valeurs possibles sont *matin, après-midi, soir* ou *nuit* est un prédicteur à modalités. L'échantillon d'apprentissage est à présent représenté par N_{prev} n-uplets $(pred_1, pred_2, .., pred_k, observation)_i$, $pred_i$ représentant l'un des k prédicteurs définis.

Construction d'un arbre de décision binaire Un arbre de décision binaire se construit de manière itérative. À chaque nœud de l'arbre, l'échantillon d'apprentissage est partagé en deux selon la valeur d'un des prédicteurs. Pour un prédicteur quantitatif, le partage est réalisé selon la valeur d'un seuil, alors que pour un prédicteur qualitatif, il est réalisé selon une de ses modalités. Le choix du prédicteur et du seuil/modalité est réalisé de manière à minimiser la variabilité des observations dans les deux sous-groupes résultants. Chaque sous-groupe est alors à son tour divisé en deux, et ainsi de suite, jusqu'à l'obtention d'un critère d'arrêt (par exemple un nombre de données minimal dans les sous-groupes, ou une diminution de la variance non suffisante). Chaque groupe final constitue une feuille de l'arbre et contient un ensemble d'observations. L'algorithme permettant la construction d'un arbre est illustré en figure 3.11, avec 2 prédicteurs p_1 et p_2 variant entre 0 et 1. Dans cet exemple, le partage de l'échantillon d'apprentissage est d'abord réalisé

selon le prédicteur p_1 avec le seuil 0.25. Les membres de l'échantillon d'apprentissage pour lesquels $p_1 \geq 0.25$ sont dans un second temps divisés en deux sous-groupes, selon la valeur du prédicteur p_2 par rapport au seuil 0.65. L'arbre comporte trois feuilles finales et a permis de classer les observations en trois sous-groupes.

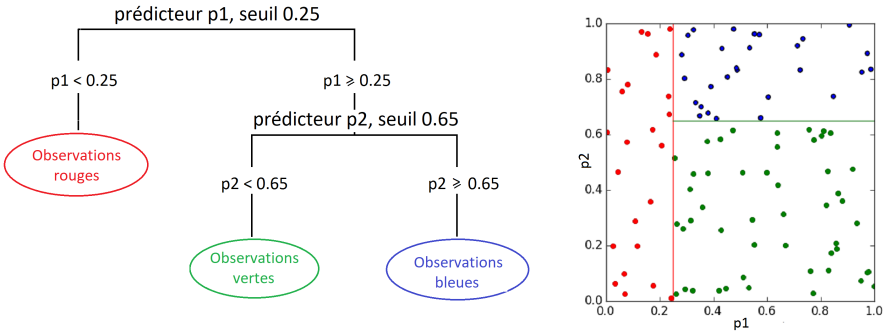
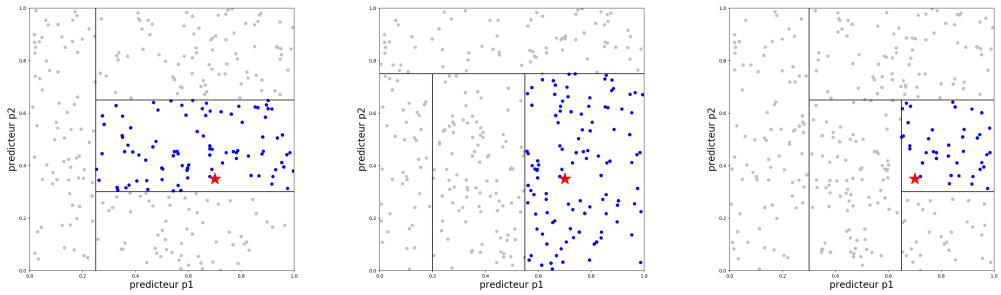


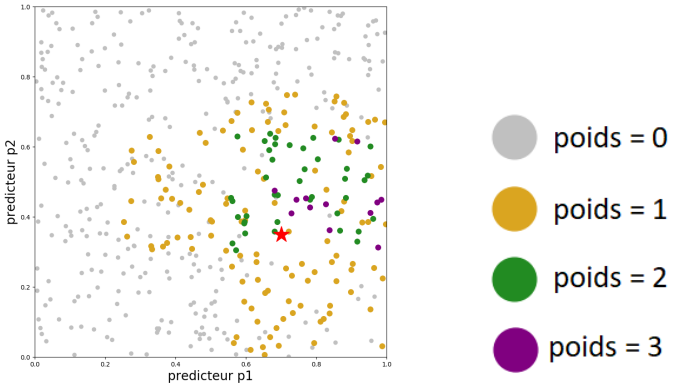
Figure 3.11: Illustration d'un arbre de décision binaire construit avec 2 prédicteurs p_1 et p_2.

Construction d'une forêt Un arbre construit de cette manière est très dépendant de l'échantillon d'apprentissage utilisé. Pour améliorer la robustesse de la prévision, Breiman [1996] a proposé de construire plusieurs arbres à partir de différents échantillons d'apprentissage. Cet ensemble d'arbres forme ainsi une *forêt*. En pratique, la quantité de données d'apprentissage nécessaire à la construction d'une forêt n'est pas disponible. Des méthodes d'échantillonnage par bootstrap (tirage avec remise à partir de l'échantillon initial) sont alors couramment utilisées. Cependant, puisque les différents arbres de la forêt sont construits avec les mêmes données, ils ne sont pas indépendants statistiquement. Breiman [2001] a alors proposé de restreindre l'ensemble des prédicteurs disponibles à chaque nœud de l'arbre de manière aléatoire. C'est ce qui vaut à la méthode son nom de *random* forest.

Utilisation de la forêt en opérationnel : Quantile Regression Forests En opérationnel, à chaque nouvelle prévision, les prédicteurs associés à cette prévision sont d'abord calculés. Puis, chaque arbre de la forêt est parcouru, sélectionnant ainsi un ensemble feuilles. Les valeurs des observations de toutes les feuilles retenues sont agrégées et permettent d'estimer la fonction de répartition empirique et les quantiles de la prévision calibrée. Ce principe est illustré sur la figure 3.12 pour une forêt aléatoire constituée de 3 arbres. Les prédicteurs p_1 et p_2 sont calculées pour la nouvelle prévision à calibrer représentée par l'étoile rouge sur chaque schéma. La figure 3.12a met en évidence les feuilles retenues pour chaque arbre pour cette prévision. L'ensemble des feuilles retenues sont aggrégées. Chaque observation est alors affectée d'un poids représentant le nombre de fois qu'elle a été sélectionnée par l'algorithme, comme illustré sur la figure 3.12b.



(a) Forêt aléatoire constituée de 3 arbres de décision. Les données pour construire ces arbres sont des sous-échantillons construits par *bootstrap*. La nouvelle prévision à calibrer est représentée par l'étoile rouge. Les deux prédicteurs associés sont $p_1 = 0.7$ et $p_2 = 0.35$. Pour chaque arbre, les points bleus correspondent aux observations conservées pour la construction de la PDF finale.



(b) Attribution d'un poids a chaque observation de l'échantillon d'apprentissage.

Figure 3.12: Illustration de la reconstruction d'une PDF pour une nouvelle prévision à partir d'une forêt aléatoire. Ici, le *bootstrap* est réalisé sans remise puisque le poids maximal affecté à une observation est égal 3.

Avantages et limites de la méthode Un avantage de la méthode est qu'aucune hypothèse n'est réalisée sur la forme *a priori* de la distribution de la variable à calibrer. Ainsi, la densité de probabilité prédictive peut prendre une grande variété de formes, qui peuvent d'ailleurs être différentes d'un pas de temps de prévision à l'autre. De plus, puisque la distribution prédictive est construite à partir des observations situées dans les feuilles des arbres, des prévisions incohérentes, par exemple une prévision de précipitation négative, ne peuvent pas être réalisées. La contrepartie est que la forêt aléatoire n'est pas capable de prévoir en-dehors des valeurs contenues dans l'échantillon d'apprentissage. Le choix de l'échantillon d'apprentissage est donc déterminant pour la performance de la méthode en opérationnel : il doit être représentatif des événements à prévoir. Il n'est pas toujours aisé de satisfaire cette contrainte dans la pratique, et ce point peut être problématique pour la prévision des extrêmes. Pour pallier cette limite, il est possible de calibrer l'anomalie

(écart par rapport à la moyenne) plutôt que la valeur prévue. L'hypothèse sous-jacente est alors que les erreurs sont homoscédastiques, c'est-à-dire qu'elles ne dépendent pas de la valeur de la variable. Il est également possible d'ajuster une fonction paramétrique sur la densité de probabilité prédictive obtenue avec QRF [Taillardat et al., 2019].

3.3.2 Mise en œuvre informatique

Aspects informatiques Dans la pratique, la construction d'une forêt aléatoire nécessite la prescription de trois hyper-paramètres :

- *mtry* : nombre de prédicteurs tirés aléatoirement parmi l'ensemble des prédicteurs à chaque nœud de chaque arbre ;
- *nodesize* : nombre d'éléments dans les feuilles ;
- *ntree* : nombre d'arbres dans la forêt.

L'optimisation de la forêt aléatoire sur la période de calage est réalisée par optimisation du triplet *(mtry, nodesize, ntree)*. Les critères d'optimisation utilisés seront développés par la suite lors de la mise en œuvre de la méthode QRF (chapitres 4 et 8). Nous pouvons indiquer à ce stade qu'un compromis est réalisé entre la minimisation du CRPS et celle de la valeur de la norme des histogrammes de rang des ensembles. En effet, la minimisation du CRPS ne va donc pas forcément de pair avec une bonne fiabilité. Nous avons vu au paragraphe 3.1.2.2 que le CRPS est composé d'un terme relatif à la fiabilité *Reli* et d'un terme relatif à la résolution/incertitude $CRPS_{pot}$ (équation 3.5). Une diminution du CRPS peut ainsi résulter d'une diminution du terme relatif à la résolution $CRPS_{pot}$ et d'une augmentation plus faible du terme relatif à la fiabilité *Reli*. En ce sens, nous rejoignons les préconisations de Wilks [2018a], qui proposent d'estimer les paramètres de post-traitement en ajoutant une pénalité au CRPS en cas de manquement à la fiabilité.

Les calculs d'optimisation des forêts aléatoires sont réalisés grâce au package *quantregForest* du logiciel R. Ils sont réalisés sur le calculateur Datarmor hébergé et administré par l'IFREMER.

3.3.3 La reconstruction des membres post-traités

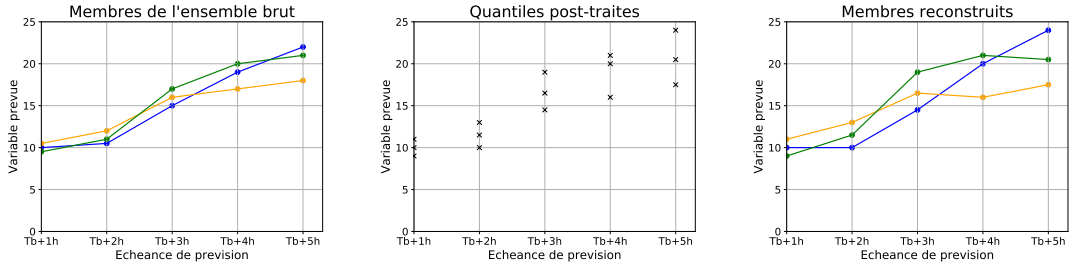
L'ensemble brut avant calibration est constitué de membres, c'est-à-dire de chroniques de prévision sur un certain nombre d'échéances de prévision. Les méthodes de calibration décrites s'appliquent à une échéance de prévision donnée, en un point donné. Elles utilisent des informations issues des membres bruts pour fournir des prévisions calibrées uniquement pour ce point et cette échéance : les méthodes paramétriques fournissent une loi de distribution de la variable à prévoir (annexe N.2.1) et les méthodes non paramétriques fournissent des quantiles (annexe N.2.2 et paragraphe 3.3.1). En sortie de ces méthodes,

la notion de membre a donc disparu. Elles ne sont donc pas adaptées à la représentation des structures de dépendance spatiales et temporelles [Vannitsem et al., 2018]. Cependant, l'utilisation des prévisions d'ensemble dans une chaîne de prévision ne peut se faire qu'avec des membres. L'utilisation d'ensembles météorologiques en entrée de modèles hydrologiques nécessitent que la répartition spatio-temporelle des membres météorologiques soit réaliste [Cloke and Pappenberger, 2009]. De même, il est nécessaire de reconstruire des chroniques de prévisions d'ensemble hydrologiques pour les utiliser en entrée d'un modèle hydraulique. Ainsi, lorsque le post-traitement est réalisé de manière univariée pour chaque variable, position et échéance, il faut reconstruire les dépendances : c'est l'étape de *reconstruction*.

Deux méthodes non paramétriques de reconstruction sont couramment employées dans la littérature : le *Schaake Shuffle* (SS, Clark et al. [2004]), et l'*Ensemble Copula Coupling* (ECC, Schefzik et al. [2013]). Le principe de ces deux méthodes est le réarrangement des prévisions univariées. Elles diffèrent par le choix du modèle qui sert à la reconstruction. Dans le *Schaake Shuffle*, c'est la structure de rang de scénarios passés qui est transposée. Ces scénarios sont sélectionnés pour leur ressemblance à la situation de la prévision actuelle [Schefzik, 2016]. L'hypothèse sous-jacente est que la structure de dépendance ne varie pas avec le temps. Dans cette méthode, on fait donc confiance aux structures spatio-temporelles observées par le passé. Les membres calibrés n'ont donc plus de lien direct avec l'ensemble brut. La méthode ECC est attractive dans le cas de prévisions d'ensemble, car la structure de rang de l'ensemble reconstruit est calquée sur celle de l'ensemble brut. Cette méthode génère un ensemble post-traité de même taille et avec la même structure que l'ensemble brut. La limite de cette méthode est que l'on fait confiance aux structures spatio-temporelles de l'ensemble qui peuvent être erronnées.

La figure 3.13 illustre la méthode ECC dans le cas d'une prévision d'ensemble constituée de 3 membres, pour des échéances de prévision allant de $T_b + 1$ heure à $T_b + 5$ heures, où T_b représente le temps de base. La figure 3.13a présente les chroniques de prévision de la prévision brute. La figure 3.13b présente quant à elle les valeurs des quantiles après le post-traitement pour chaque échéance de prévision. Les membres sont reconstruits figure 3.13c en conservant la même structure de rang que celle de l'ensemble brut, et les valeurs des quantiles post-traités remplacent désormais les valeurs des ensembles bruts. Dans notre exemple, le membre orange de l'ensemble brut occupe respectivement les rangs (1, 1, 2, 3, 3) selon les échéances. Dans l'ensemble reconstruit, il occupe les mêmes rangs, et la valeur pour chaque échéance est remplacée par celle des quantiles calibrés.

Cependant, les techniques de reconstruction atteignent des limites pour la reconstruction des prévisions d'ensemble hydrologiques au pas de temps horaire. En effet, elles génèrent des sauts de débits entre deux échéances successives non réalistes. Une procédure de lissage des trajectoires, décrite dans Bellier [2018], a donc été appliquée pour préserver la corrélation temporelle entre les pas de temps. Contrairement à l'application de la méthode



(a) Chroniques de prévision de l'ensemble brut

(b) Valeurs des quantiles post-traités pour chaque échéance de prévision

(c) Chroniques après reconstruction par ECC

Figure 3.13: Illustration de la méthode ECC. La structure de rang de l'ensemble brut est appliquée aux quantiles post-traitées pour générer des membres post-traités.

ECC, la procédure de lissage modifie les valeurs donc potentiellement les performances de l'ensemble calibré.

3.3.4 Synthèse

Dans ce chapitre, nous avons présenté en détails la méthode *Quantile Regression Forests*, puis nous avons décrit des méthodes de reconstruction des membres après calibration : le *Schaake Shuffle* et l'*Ensemble Copula Coupling*.

Dans ce travail de thèse, les prévisions PEAROME sont calibrées à l'échelle des sous-bassins versants amont avec les méthodes *Quantile Regression Forests* (QRF) et *Ensemble Model Output Statistics* (EMOS) [Tiberi-Wadier et al., 2018]. La reconstruction spatio-temporelle des membres est réalisée avec la méthode *Ensemble Copula Coupling* (ECC). Ce travail concernant la calibration des prévisions d'ensemble de précipitations est issu d'une collaboration avec Météo-France et la calibration à proprement parler a été effectuée par Maxime Taillardat (Météo-France & CNRM, Toulouse).

Après avoir déterminé les paramètres incertains significatifs pour les débits hydrologiques simulés avec le modèle MORDOR-TS grâce à une étude de sensibilité globale (voir paragraphe 3.2 pour la méthodologie), des prévisions d'ensemble hydrologiques brutes sont générées puis calibrées par la méthode QRF. La reconstruction des membres est réalisée grâce à la méthode ECC. Un lissage des trajectoires est appliqué pour pouvoir ensuite utiliser ces débits hydrologiques d'ensemble en entrée du modèle hydraulique MASCARET. Ce travail fait l'objet du chapitre 4.

De manière analogue, après avoir déterminé les paramètres incertains significatifs pour les hauteurs d'eau simulées par le modèle MASCARET aux trois stations de vigilance (travail décrit dans le chapitre 7), des prévisions d'ensemble hydrauliques brutes sont générées puis calibrées par la méthode QRF. Aucune reconstruction de membres n'est

effectuée puisqu'on est alors au bout de la chaîne de prévision. Ce travail fait l'objet du chapitre 8.

3.4 L'assimilation de données

Cette section est écrite à partir de Barthelemy [2015], Bonan [2013], Emery [2017], du cours de master SOAC de l'université Paul Sabatier à Toulouse[4] et d'une revue des méthodes d'assimilation de données du CERFACS[5].

Les méthodes d'assimilation de données ont pour objectif la meilleure estimation (en un sens qui sera défini ultérieurement) de l'état d'un système physique, en combinant deux sources d'information : d'une part une modélisation numérique de l'évolution de l'état du système qui, comme nous l'avons vu au paragraphe 1.1.2, est entachée d'incertitudes, et d'autre part des observations ponctuelles dans le temps et l'espace qui elles non plus ne sont pas parfaites. Le meilleur état estimé est appelé l'état *analysé*. Il résulte ainsi d'un compromis entre ces deux sources d'informations. On le construit en tenant compte des erreurs associées à chacune d'elles. On cite classiquement trois types d'applications de l'assimilation de données [Carrassi et al., 2018] :

- Le **filtrage** : l'assimilation utilise les observations passées et présentes pour donner la meilleure estimation de l'état du système étudié au temps présent. C'est ce qui est réalisé dans les systèmes opérationnels de prévision.

- La **prévision** : c'est la suite logique de l'étape de filtrage. L'assimilation utilise les observations passées et présentes pour donner la meilleure estimation de l'état futur du système étudié.

- Le **lissage** : l'assimilation utilise les observations passées, présentes et futures pour réaliser une analyse rétrospective de l'état du système. L'objectif est la meilleure estimation de la trajectoire du système étudié au cours du temps.

Les méthodes d'assimilation de données sont généralement classées en deux groupes : les méthodes variationnelles, pour lesquelles la meilleure estimation de l'état du système est obtenue en minimisant une fonctionnelle, et méthodes séquentielles également appelées filtres, dont fait partie le filtre de Kalman d'ensemble utilisé dans ce travail. Toutes ces méthodes sont fondées sur l'hypothèse d'approximation gaussienne des erreurs du modèle et d'observation.

Dans cette section, nous introduisons d'abord au paragraphe 3.4.1 les notations utilisées pour décrire les algorithmes de l'assimilation de données, ainsi que les concepts néces-

[4]Master Sciences de l'Océan, de l'Atmosphère et du Climat, cours d'assimilation de données, novembre 2004 par François Bouttier, https://assim.chocolate.fr/tcfr.pdf

[5]N. Daget, CERFACS, Revue des méthodes d'assimilation, 14 novembre 2007, http://www.cerfacs.fr/globc/publication/technicalreport/2007/revue_methodes_assimilation.pdf

saires à leur mise en place. Le paragraphe 3.4.2 présente la formulation bayésienne du problème d'assimilation de données. Le paragraphe 3.4.3 présente le filtre de Kalman et ses variantes, le filtre de Kalman étendu et enfin le filtre de Kalman d'ensemble utilisé dans ce travail. Le paragraphe 3.4.4 présente la mise en œuvre informatique de ce travail. Le paragraphe 3.4.5 présente une bibliographie des études utilisant l'assimilation de données pour les modélisations hydrologiques et hydrauliques. Pour finir, le paragraphe 3.4.6 réalise la synthèse de cette section et décrit la place de l'assimilation de données dans le contexte qui nous intéresse, à savoir la réalisation de prévisions d'ensemble hydrologiques et hydrauliques.

3.4.1 Notations et concepts

3.4.1.1 Aide mémoire - Notations

Avant d'entrer dans le vif du sujet, nous présentons un tableau récapitulatif des notations utilisées dans cette section. L'objet n'est pas les expliquer dès à présent ; le lecteur pourra s'y reporter lors la lecture de cette section.

Symbole	Définition
$\mathbf{s}^{\mathbf{c}}$	état réel continu d'un système physique
	Vecteurs
$\mathbf{x}^{\mathbf{t}}$	vecteur de contrôle vrai (dimension n_X)
$\mathbf{x}^{\mathbf{b}}$	vecteur de contrôle d'ébauche (dimension n_X)
$\mathbf{x}^{\mathbf{a}}$	vecteur de contrôle analysé (dimension n_X)
$\mathbf{s}^{\mathbf{t}}$	état vrai
$\mathbf{s}^{\mathbf{b}}$	état de l'ébauche (état du système suite à la propagation par le modèle du vecteur de contrôle d'ébauche)
$\mathbf{s}^{\mathbf{a}}$	état analysé (état du système suite à la propagation par le modèle du vecteur de contrôle analysé)
$\mathbf{y}^{\mathbf{o}}$	vecteur d'observation (dimension n_Y)
	Opérateurs non linéaires
$\mathcal{H}$	Opérateur d'observation entre le vecteur de contrôle et les observations (dimension (n_Y, n_X))
$\mathcal{M}$	Modèle d'évolution de l'état
	Matrices et opérateurs linéaires
$\mathbf{B}$	Matrice de covariances d'erreurs d'ébauche $(x^b - x^t)$ - dimension (n_X, n_X)
$\mathbf{R}$	Matrice de covariances d'erreurs d'observation $(y^o - \mathcal{H}x^t)$ - dimension (n_Y, n_Y)
$\mathbf{A}$	Matrice de covariances des erreurs d'analyse $(x^a - x^t)$ - dimension (n_X, n_X)
$\mathbf{Q}$	Matrice de covariances des erreurs modèle
$\mathbf{H}$	Opérateur d'observation généralisé linéarisé
$\mathbf{M}$	Modèle d'évolution de l'état linéarisé
Π	Projection de l'espace continu vers l'espace discret

Tableau 3.2: Aide-mémoire - Notations utilisées pour l'assimilation de données.

3.4.1.2 Vecteur d'état et vecteur de contrôle

On note $\mathbf{s^c}$ l'état réel continu d'un système physique à un instant donné. Cet état réel peut être multivarié. Dans le cas de modèles de prévision numérique du temps, il peut représenter par exemple la température et l'humidité de l'air. Les modèles numériques qui résolvent les équations représentant la physique du modèle travaillent dans un espace discrétisé avec un nombre fini de points. Avec ces modèles, on ne peut donc accéder qu'à une représentation discrète du champ physique modélisé. On appelle alors l'état vrai noté $\mathbf{s^t}$ (t pour *true*) la projection de $\mathbf{s^c}$ dans l'espace discrétisé du modèle grâce à l'opérateur de projection Π tel que :

$$\mathbf{s^t} = \Pi(\mathbf{s^c}).$$

Dans un contexte d'assimilation de données, on cherche à estimer au mieux l'état $\mathbf{s^t}$. Il doit être le plus cohérent possible avec la sortie d'un modèle numérique et des observations, selon un critère à définir. Dans le contexte de cette thèse, le vecteur d'état contient la valeur du débit modélisé à l'exutoire du bassin versant pour les modélisations hydrologiques avec MORDOR-TS. Dans le cas de la modélisation hydraulique avec MASCARET, le vecteur d'état contient la discrétisation des hauteurs d'eau et des débits sur chaque section de calcul.

L'assimilation de données a pour objectif d'estimer au mieux le vecteur d'état. Cependant, l'analyse n'est pas systématiquement effectuée sur le vecteur d'état. On peut corriger des variables et/ou des paramètres pour lesquels il n'existe pas d'observation correspondante. C'est par exemple le cas lorsque l'on souhaite corriger des paramètres des modèles numériques. Par ailleurs, on peut également avoir besoin de réduire la taille du problème pour des questions de coût de calcul informatique. On définit ainsi le vecteur de contrôle correspondant à l'ensemble des variables d'intérêt que l'on cherche à corriger avec l'assimilation de données. Il peut être constitué de variables d'état, de paramètres du modèle, de conditions aux limites... Dans le cadre de cette thèse, le vecteur de contrôle pour l'assimilation de données dans le modèle MORDOR-TS est constitué uniquement des paramètres du modèle. L'assimilation de données mise en place dans le modèle MASCARET utilise comme variables de contrôle les coefficients de Strickler et des paramètres additifs ou multiplicatifs pour corriger les forçages. Dans notre cas applicatif, les variables d'état ne font donc pas partie du vecteur de contrôle. Dans la suite, on considère donc que le vecteur de contrôle noté $\mathbf{x}$ est constitué de paramètres du modèle et de paramètres permettant la correction des forçages.

Lorsque l'on tient compte de l'aspect temporel en temps discret, on note respectivement $\mathbf{x_k}$ et $\mathbf{s_k}$ le vecteur de contrôle et l'état du modèle au temps t_k.

3.4.1.3 Notion de fenêtre d'assimilation

On considère une *fenêtre d'assimilation* correspondant à la période sur laquelle est réalisée un cycle d'assimilation et sur laquelle sont disponibles des observations. Cette fenêtre est délimitée par deux pas de temps du modèle consécutifs ou non. Dans un souci de simplification, nous considérons ici que la fenêtre d'assimilation est constituée de deux pas de temps consécutifs $[t_{k-1}, t_k]$ et que des observations sont disponibles sur cette fenêtre. Cette fenêtre correspond au cycle d'assimilation c_k. La fin de cette fenêtre correspond au temps de base, c'est-à-dire le moment auquel une prévision peut être lancée en opérationnel après assimilation de données.

3.4.1.4 Modèle d'évolution du système

La simulation de l'évolution du système physique est réalisée grâce à un modèle mathématique noté $\mathcal{M}$ qui résout les équations de la physique et permet de décrire l'évolution temporelle du système. Sur le cycle d'assimilation $c_k = [t_{k-1}, t_k]$, ce modèle qui a pour paramètres le vecteur de contrôle $\mathbf{x_k^t}$ fait évoluer l'état du système $\mathbf{s}$ de t_{k-1} à t_k en partant de l'état $\mathbf{s_{k-1}^t}$. Nous faisons l'hypothèse que le modèle d'évolution du vecteur de contrôle est l'identité, ce qui signifie que les paramètres sont constants sur la fenêtre d'assimilation. Si le modèle $\mathcal{M}$ est parfait, alors on a :

$$\mathbf{s_k^t} = \mathcal{M}_{k-1,k}(\mathbf{s_{k-1}^t}, \mathbf{x_k^t}). \tag{3.11}$$

Dans la pratique cependant, une erreur modèle est présente et la relation devient :

$$\mathbf{s_k^t} = \mathcal{M}_{k-1,k}(\mathbf{s_{k-1}^t}, \mathbf{x_k^t}) + \eta_\mathbf{k}.$$

avec $\eta_\mathbf{k}$ représentant l'erreur due au modèle. Toutes ces erreurs sont liées entre elles et sont indissociables. L'erreur modèle est considérée comme étant une variable aléatoire de moyenne $\overline{\eta_\mathbf{k}}$ et de matrice de covariance $\mathbf{Q_k}$ telle que

$$\mathbf{Q_k} = \overline{(\eta_\mathbf{k} - \overline{\eta_\mathbf{k}})(\eta_\mathbf{k} - \overline{\eta_\mathbf{k}})^T}.$$

Avec l'hypothèse gaussienne et l'hypothèse d'erreur modèle non biaisée, on a $\eta_\mathbf{k} \sim \mathcal{N}(0, \mathbf{Q_k})$. Pour mémoire, le paragraphe 1.1.2 traite des incertitudes des modèles météorologiques, hydrologiques et hydrauliques et explique les multiples causes des erreurs modèle.

3.4.1.5 Modèle d'observation

Sur la fenêtre d'assimilation du cycle c_k, nous disposons d'un ensemble d'observations de l'état du système. Nous appelons $\mathbf{y_k^o}$ le vecteur de taille n_Y qui contient les observations sur

le cycle c_k. Il est à noter que la taille n_Y peut changer au fil des cycles d'analyse selon la disponibilité des observations sur chaque fenêtre d'assimilation. On note également que le nombre d'observations n_Y est généralement inférieur à la taille du vecteur de contrôle notée n_X, ce qui implique que le problème d'analyse est sous-déterminé en l'absence d'autres informations. Comme nous l'avons vu, les variables de l'espace de contrôle ne sont pas forcément observées. C'est le cas lorsque l'on corrige des paramètres du modèle. Ainsi, la manière la plus commode pour faire la comparaison entre les observations et le modèle est de définir une application de l'espace de contrôle vers l'espace des observations : c'est l'*opérateur d'observation* noté $\mathcal{H}_k$. Cet opérateur simule des valeurs d'observations $\mathcal{H}_k(\mathbf{x_k^t})$ correspondant au vecteur de contrôle $\mathbf{x_k^t}$ que l'on observerait en l'absence d'erreur modèle. On a :

$$\mathbf{y_k^o} = \mathcal{H}_k(\mathbf{x_k^t}) + \epsilon_\mathbf{k}^\mathbf{o}. \tag{3.12}$$

avec $\epsilon_\mathbf{k}^\mathbf{o}$ l'erreur d'observation à l'instant t_k. Cette erreur représente les erreurs instrumentales dues aux imprécisions des appareils de mesure, des erreurs dans la modélisation de l'opérateur $\mathcal{H}$ et des erreurs de représentativité, ce qui correspond notamment aux erreurs de discrétisation qui empêchent $\mathbf{s_k^t}$ d'être une représentation parfaite de la réalité. L'erreur d'observation est considérée comme étant une variable aléatoire de moyenne $\overline{\epsilon_\mathbf{k}^\mathbf{o}}$ et de matrice de covariance $\mathbf{R_k}$ telle que :

$$\mathbf{R_k} = \overline{(\epsilon_\mathbf{k}^\mathbf{o} - \overline{\epsilon_\mathbf{k}^\mathbf{o}})(\epsilon_\mathbf{k}^\mathbf{o} - \overline{\epsilon_\mathbf{k}^\mathbf{o}})^T}.$$

Avec l'hypothèse gaussienne et l'hypothèse d'erreur d'observation non biaisée, on a $\epsilon_\mathbf{k}^\mathbf{o} \sim \mathcal{N}(0, \mathbf{R_k})$.

La définition de l'opérateur d'observation $\mathcal{H}$ dépend des variables de contrôle choisies :

- Dans le cas où le vecteur de contrôle n'est constitué que de variables d'état du modèle, l'opérateur d'observation peut être l'identité si toutes les variables d'état sont observées aux mêmes points de grille et aux mêmes instants. En pratique, l'opérateur d'observation représentera une sélection si toutes les variables d'état ne sont pas observées, et/ou un opérateur d'interpolation si les variables ne sont pas observées sur la même grille que le modèle ou aux mêmes temps. On note alors :

$$\mathcal{H}_k = \mathcal{S}_k.$$

- Dans le cas où le vecteur de contrôle est constitué de paramètres du modèle, l'opérateur d'observation est la composition de l'opérateur de sélection et d'interpolation précédent $\mathcal{S}_k$ et du modèle d'évolution $\mathcal{M}_{k-1,k}$:

$$\mathcal{H}_k = \mathcal{S}_k \circ \mathcal{M}_{k-1,k}.$$

En effet, le modèle transforme d'abord les paramètres en variables d'état, et ces variables

d'état sont projetées dans l'espace des observations. Cet opérateur est appelé opérateur d'observation généralisé, il est souvent noté $\mathcal{G}$ dans la littérature. Dans un objectif de simplification, on conservera dans la suite la notation $\mathcal{H}_k$, que l'opérateur d'observation soit généralisé ou non.

3.4.1.6 Différentes estimations du vecteur de contrôle

Deux terminologies sont employées selon l'étape à laquelle est estimé le vecteur de contrôle : ébauche ou analyse. Par extension, dans ce manuscrit, ces termes s'appliquent également au vecteur d'état $\mathbf{s_k}$ estimé au temps t_k par simulation par le modèle numérique sur la fenêtre d'assimilation du cycle c_k avec les paramètres d'ébauche ou d'analyse. On parle alors de trajectoire d'ébauche ou d'analyse.

Ébauche Si $\mathbf{x_k}$ est une estimation *a priori* du vecteur de contrôle avant assimilation, $\mathbf{x_k}$ est appelé *ébauche* et est noté $\mathbf{x_k^b}$, le b signifiant *background*. L'ébauche peut être issue d'une climatologie dans le cas du contrôle de l'état du modèle, d'une prévision, d'un processus d'assimilation antérieur... L'erreur d'ébauche $\boldsymbol{\epsilon_k^b}$ vaut

$$\boldsymbol{\epsilon_k^b} = \mathbf{x_k^b} - \mathbf{x_k^t}.$$

Elle est supposée être une variable aléatoire de moyenne $\overline{\boldsymbol{\epsilon_k^b}}$ et de matrice de covariance

$$\mathbf{B_k} = \overline{(\boldsymbol{\epsilon_k^b} - \overline{\boldsymbol{\epsilon_k^b}})(\boldsymbol{\epsilon_k^b} - \overline{\boldsymbol{\epsilon_k^b}})^T}.$$

Avec l'hypothèse gaussienne et l'hypothèse d'erreur d'ébauche non biaisée, on a $\boldsymbol{\epsilon_k^b} \sim \mathcal{N}(0, \mathbf{B_k})$.

Analyse Si $\mathbf{x_k}$ est une estimation du vecteur de contrôle suite à l'assimilation d'une observation, alors $\mathbf{x_k}$ est appelé *analyse* et est noté $\mathbf{x_k^a}$. L'erreur d'analyse $\boldsymbol{\epsilon_k^a}$ vaut

$$\boldsymbol{\epsilon_k^a} = \mathbf{x_k^a} - \mathbf{x_k^t}.$$

Elle est supposée être une variable aléatoire de moyenne $\overline{\boldsymbol{\epsilon_k^a}}$ et de matrice de covariance

$$\mathbf{A_k} = \overline{(\boldsymbol{\epsilon_k^a} - \overline{\boldsymbol{\epsilon_k^a}})(\boldsymbol{\epsilon_k^a} - \overline{\boldsymbol{\epsilon_k^a}})^T}.$$

Ce sont les erreurs d'estimation de l'analyse que l'on cherche à minimiser dans le processus d'assimilation. Une mesure de l'amplitude $\overline{||\boldsymbol{\epsilon_k^a} - \overline{\boldsymbol{\epsilon_k^a}}||}$ est fournie par la trace de la matrice des covariances d'erreur $\mathbf{A_k}$. L'assimilation de données cherchera donc à minimiser

$$Tr(\mathbf{A_k}) = \overline{||\boldsymbol{\epsilon_k^a} - \overline{\boldsymbol{\epsilon_k^a}}||^2}.$$

3.4.2 Formulation bayésienne du problème d'assimilation de données

Comme nous l'avons expliqué, nous possédons deux sources d'information complémentaires pour améliorer l'état du système : le modèle numérique décrit au paragraphe 3.4.1.4, et les observations décrites au paragraphe 3.4.1.5. Il s'agit à présent d'estimer le vecteur de contrôle du système à partir de ces deux sources d'information en tenant compte de leurs incertitudes respectives. Pour construire des algorithmes d'assimilation de données, on représente mathématiquement l'incertitude de ces deux sources d'information grâce à l'utilisation de fonctions de densité de probabilités (PDFs). Parmi toutes les fonctions de densité de probabilité, la fonction Gaussienne a des caractéristiques intéressantes comme nous allons le voir. L'approche bayésienne offre alors un cadre pour la formalisation du problème d'assimilation de données.

On considère des pas de temps discrets de l'intervalle $[t_0; t_k]$ notés $t_{0:k} = (t_0, t_1, ..., t_k)$. Sur ces pas de temps, les états du vecteur de contrôle sont notés $\mathbf{x_{0:k}} = (\mathbf{x_0}, \mathbf{x_1}, ..., \mathbf{x_k})$, et les observations correspondantes $\mathbf{y^o_{0:k}} = (\mathbf{y^o_0}, \mathbf{y^o_1}, ..., \mathbf{y^o_k})$. Le but de l'assimilation de données est d'améliorer l'estimation du vecteur de contrôle $\mathbf{x_k}$ étant donné ses états passés $\mathbf{x_{0:k}}$ et les observations $\mathbf{y_{0:k}}$.

On considère que le vecteur de contrôle d'ébauche $\mathbf{x^b_k}$ est une variable aléatoire de densité de probabilité $p^b(\mathbf{x_k})$. Elle représente la probabilité que le vecteur d'ébauche représente le vecteur de contrôle vrai $\mathbf{x^t_k}$. On considère également le vecteur d'observation $\mathbf{y^o_k}$ comme une variable aléatoire de densité de probabilité $p(\mathbf{y^o_k})$. On suppose à présent que la suite de variables aléatoires est une chaîne de Markov discrète telle que :

- la PDF du vecteur de contrôle $\mathbf{x_k}$ au temps t_k ne dépend que de sa valeur à l'instant précédant t_{k-1}, c'est-à-dire :

$$p(\mathbf{x_k}|\mathbf{x_{0:k-1}}) = p(\mathbf{x_k}|\mathbf{x_{k-1}}).$$

La PDF $p(\mathbf{x_k}|\mathbf{x_{k-1}})$ représente l'évolution temporelle du vecteur de contrôle entre t_{k-1} et t_k. Elle correspond à la PDF d'ébauche $p^b(\mathbf{x_k})$, qui représente la probabilité *a priori* du vecteur de contrôle $\mathbf{x_k}$ avant l'assimilation de nouvelles observations :

$$p(\mathbf{x_k}|\mathbf{x_{k-1}}) = p^b(\mathbf{x_k}).$$

$p^b(\mathbf{x_k})$ est appelé *prior*.

- la PDF des observations $\mathbf{y^o_k}$ ne dépend que du contrôle au même pas de temps :

$$p(\mathbf{y^o_k}|\mathbf{x_{0:k}}) = p(\mathbf{y^o_k}|\mathbf{x_k}).$$

$p(\mathbf{y^o_k}|\mathbf{x_k})$ est la fonction de vraisemblance de $\mathbf{y^o_k}$ connaissant $\mathbf{x_k}$.

Sous ces hypothèses, le but de l'assimilation de données devient l'amélioration de l'estimation du vecteur de contrôle $\mathbf{x_k}$ à partir des observations $\mathbf{y_k^o}$. Il s'agit donc d'estimer la probabilité conditionnelle de $\mathbf{x_k}$ sachant $\mathbf{y_k^o}$ notée $p(\mathbf{x_k}|\mathbf{y_k^o})$. Elle correspond à la PDF de l'analyse $p^a(\mathbf{x_k})$. Le théorème de Bayes nous permet de l'estimer :

$$p^a(\mathbf{x_k}) = p(\mathbf{x_k}|\mathbf{y_k^o}) = \frac{p^b(\mathbf{x_k})\ p(\mathbf{y_k^o}|\mathbf{x_k})}{p(\mathbf{y_k^o})}.$$

Notons que le dénominateur n'a pas besoin d'être calculé. Il est constant puisque la PDF est évaluée en $\mathbf{y_k^o}$. Il représente ainsi un facteur de normalisation. En pratique, on retiendra la relation de proportionnalité suivante :

$$p^a(\mathbf{x_k}) = p(\mathbf{x_k}|\mathbf{y_k^o})\ \alpha\ p^b(\mathbf{x_k})\ p(\mathbf{y_k^o}|\mathbf{x_k}). \tag{3.13}$$

Ainsi, il suffit donc de connaître le prior $p^b(\mathbf{x_k})$ et la vraisemblance $p(\mathbf{y_k^o}|\mathbf{x_k})$ pour obtenir la PDF de l'analyse $p^a(\mathbf{x_k})$. Il s'agit à présent de choisir un estimateur de cette PDF afin d'estimer de manière optimale le vecteur de contrôle connaissant la réalisation de $\mathbf{y_k^o}$. L'estimation optimale dépend du critère d'optimalité choisi et plusieurs estimateurs existent. La moyenne et le mode représentent deux approches classiques. Les estimateurs de la moyenne minimisent l'erreur quadratique moyenne. Le vecteur de contrôle analysé doit être aussi proche que possible du contrôle vrai dans le sens de minimum de variance. Ces estimateurs sont à l'origine des méthodes séquentielles par filtrage, en particulier le filtre de Kalman et ses dérivés, qui seront exposées à partir du paragraphe 3.4.3. Le mode représente le pic de la distribution du vecteur de contrôle, c'est son état le plus probable. Les estimateurs utilisés sont des estimateurs de maximum *a posteriori* et sont à l'origine des méthodes variationnelles. Elles ne seront pas abordées en détail dans ce chapitre mais le lecteur intéressé pourra se reporter à Bonan [2013], Carrassi et al. [2018] ou Mirouze [2010] pour un approfondissement.

Comme nous venons de le voir, le problème d'assimilation de données a pour objectif la recherche de la PDF $p^a(\mathbf{x_k})$. Cependant, en pratique, cette densité est difficile à estimer. C'est pourquoi la plupart des méthodes d'assimilation de données ne s'intéressent qu'aux deux premiers moments des variables aléatoires, c'est-à-dire la moyenne et la variance. On se place désormais dans le cas linéaire gaussien, c'est-à-dire que les distributions des lois de probabilités $p^b(\mathbf{x_k})$ et $p(\mathbf{y_k^o}|\mathbf{x_k})$ sont des lois gaussiennes et que l'opérateur d'observation $\mathcal{H}$ est linéaire. Nous faisons les hypothèses suivantes :

- La densité de probabilité $p^b(\mathbf{x})$ est normale de moyenne $\mathbf{x^b}$ et de matrice de covariance $\mathbf{B}$. L'erreur d'ébauche $\epsilon^\mathbf{b}$ est sans biais ($\overline{\epsilon^\mathbf{b}} = 0$), non triviales donc $\mathbf{B}$ est définie positive.
- La densité de probabilité $p(\mathbf{y_k^o}|\mathbf{x_k})$ est normale de moyenne nulle et de matrice de covariance $\mathbf{R}$. L'erreur d'observation $\epsilon^\mathbf{o}$ est sans biais ($\overline{\epsilon^\mathbf{o}} = 0$), non triviale donc $\mathbf{R}$ est

définie positive.

- Les erreurs d'ébauche et d'observation ont des corrélations mutuelles nulles : $\overline{\epsilon^{\mathbf{b}}(\epsilon^{\mathbf{o}})^T} = 0$.
- L'opérateur d'observation $\mathcal{H}_k$ est linéaire au voisinage de l'ébauche : pour tout $\mathbf{x_k}$ assez proche de $\mathbf{x_k^b}$, $\mathcal{H}_k(\mathbf{x_k}) - \mathcal{H}_k(\mathbf{x_k^b}) = \mathbf{H}_k(\mathbf{x_k} - \mathbf{x_k^b})$ avec $\mathbf{H}_k$ la linéarisation de l'opérateur d'observation autour de $\mathbf{x_k^b}$.

L'annexe N.3 présente la formulation de l'analyse sous ces hypothèses en utilisant respectivement les estimateurs de maximum *a posteriori* et de minimum de variance dans le cas linéaire gaussien. Dans ce cas, les deux estimateurs sont équivalents. Il faut garder à l'esprit que leur équivalence disparaît dès qu'une des hypothèses n'est plus vérifiée. C'est à partir de cette formulation que sont construits les algorithmes des filtres de Kalman présentés au paragraphe 3.4.3.

3.4.3 Filtres de Kalman

Les filtres de Kalman [Kalman, 1960, Kalman and Bucy, 1961] sont une généralisation des équations d'analyse présentées en annexe N.3 aux problèmes d'évolution temporelle sous certaines hypothèses. Il nous faut donc à présent utiliser la notion de modèle d'évolution du système présenté au paragraphe 3.4.1.5. Comme nous l'avons vu, l'utilisation du modèle d'évolution est à l'origine d'un nouveau type d'erreur appelé *erreur modèle*. Nous faisons cependant l'hypothèse que celle-ci est nulle.

Nous décrivons à présent la réalisation d'un cycle d'assimilation de données dans le cas de la correction des paramètres d'un modèle. Au début du cycle d'assimilation, le vecteur de contrôle d'ébauche est défini à partir du vecteur de contrôle analysé au cycle précédent. L'état d'ébauche sur la fenêtre d'assimilation est fourni par une simulation du modèle lancé avec les paramètres d'ébauche. La comparaison des observations avec les états d'ébauche simulés par le modèle permet de calculer le vecteur de contrôle analysé. Le modèle est ensuite relancé sur la même fenêtre d'assimilation avec les paramètres analysés, ce qui permet d'obtenir les états du modèle en équilibre avec les valeurs du vecteur de contrôle analysé sur la fenêtre d'assimilation. On peut alors éventuellement fournir une prévision opérationnelle en alimentant le modèle avec des prévisions.

3.4.3.1 Le Filtre de Kalman

On considère une fenêtre d'assimilation $[t_{k-1}, t_k]$ correspondant au cycle c_k d'assimilation. Des observations sont disponibles sur cette fenêtre. Nous décrivons ci-après les hypothèses émises pour la mise en place d'un filtre de Kalman puis l'algorithme du filtre de Kalman.

Hypothèses pour le filtre de Kalman

- Les hypothèses émises à la fin du paragraphe 3.4.2 sont valables pour chaque temps t_k.
- Il n'y a pas de modèle d'évolution des paramètres entre les cycles d'assimilation.
- Le vecteur de contrôle $\mathbf{x_k}$ est constant sur le cycle c_k.
- On considère que les erreurs de modélisation η_k sont nulles.
- Le modèle de prévision est linéaire autour de $\mathbf{x_k^a}$: pour tout $\mathbf{x_k}$ assez proche de $\mathbf{x_k^a}$, $\mathcal{M}_{k-1,k}(\mathbf{s_{k-1}}, \mathbf{x_k}) - \mathcal{M}_{k-1,k}(\mathbf{s_{k-1}}, \mathbf{x_k^a}) = \mathbf{M_{k-1,k}}(\mathbf{s_{k-1}}, \mathbf{x_k} - \mathbf{x_k^a})$, avec $\mathbf{M}_{k-1,k}$ la linéarisation du modèle autour de $\mathbf{x_k^a}$.

Algorithme du filtre de Kalman

Phase d'analyse au cycle $c_k = [t_{k-1}, t_k]$

- Calcul de la trajectoire d'ébauche $\mathbf{s_k^b}$ du cycle c_k selon l'équation 3.11 à partir de la phase prévision du cycle précédent c_{k-1} : l'état du modèle à t_{k-1} est $\mathbf{s_{k-1}^a}$ et le vecteur d'ébauche pour le cycle c_k est $\mathbf{x_k^b}$:

$$\mathbf{s_k^b} = \mathbf{M_{k-1,k}}(\mathbf{s_{k-1}^a}, \mathbf{x_k^b}) \tag{3.14}$$

- Calcul de la matrice de gain de Kalman :

$$\mathbf{K_k} = \mathbf{B_k H_k^T}(\mathbf{H_k B_k H_k^T} + \mathbf{R_k})^{-1} \tag{3.15}$$

- Calcul du vecteur de contrôle analysé :

$$\mathbf{x_k^a} = \mathbf{x_k^b} + \mathbf{K_k}(\mathbf{y_k^o} - \mathbf{H_k x_k^b}) \tag{3.16}$$

- Calcul de la matrice de covariance d'erreur d'analyse :

$$\mathbf{A_k} = (\mathbf{I} - \mathbf{K_k H_k})\mathbf{B_k} \tag{3.17}$$

Phase de prévision pour le cycle suivant $c_{k+1} = [t_k, t_{k+1}]$

- Calcul de l'état analysé $\mathbf{s_k^a}$ au temps t_k selon l'équation 3.11 : l'état du modèle à t_{k-1} est $\mathbf{s_{k-1}^a}$ et le vecteur de contrôle analysé sur le cycle c_k est $\mathbf{x_k^a}$:

$$\mathbf{s_k^a} = \mathbf{M_{k-1,k}}(\mathbf{s_{k-1}^a}, \mathbf{x_k^a}) \tag{3.18}$$

- Propagation du vecteur de contrôle pour le cycle suivant :

$$\mathbf{x_{k+1}^b} = \mathbf{x_k^a} + \nu_\mathbf{k} \tag{3.19}$$

avec ν_k représentant un bruit de moyenne nulle.

La figure 3.14 illustre le fonctionnement du filtre dans le cas simplifié où :

- Il n'y a qu'une seule observation par fenêtre d'assimilation. Pour le cycle c_k, celle-ci est disponible à t_k.
- L'espace de contrôle est l'espace des observations.
- On ne décrit pas l'étape de propagation de la trajectoire d'analyse.
- Le vecteur de contrôle d'ébauche au cycle c_{k+1} est le vecteur de contrôle analysé du cycle c_k.

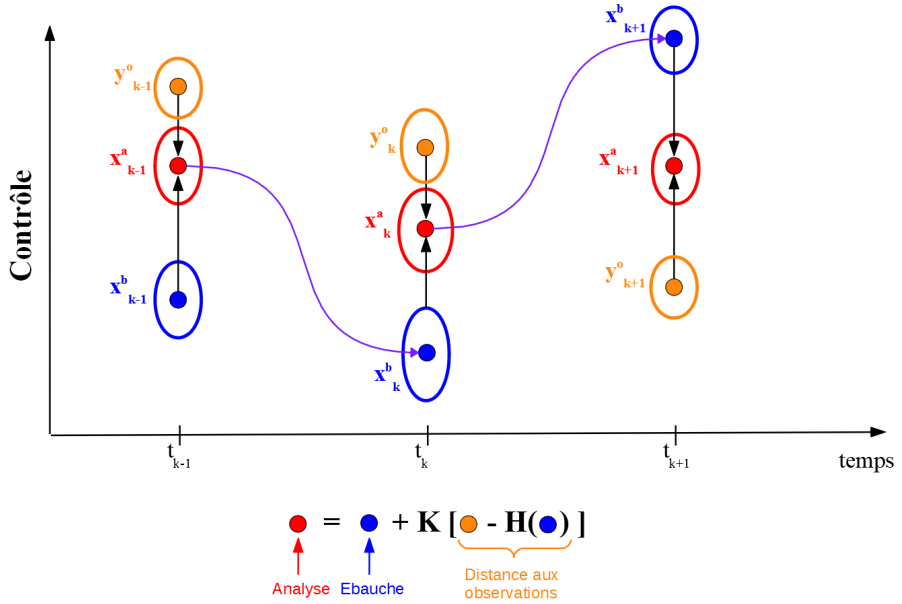


Figure 3.14: Fonctionnement séquentiel du filtre de Kalman - Cas simplifié (voir texte). À chaque instant t_k, nous avons à disposition : (i) un vecteur de contrôle d'ébauche $\mathbf{x_k^b}$ (point bleu), obtenu à partir de l'analyse effectuée à l'instant précédent $\mathbf{x_{k-1}^a}$ et sa matrice de covariance d'erreur associée $\mathbf{B_k}$ (symbolisée par l'ellipse bleue), et (ii) une observation $\mathbf{y_k^o}$ (point orange) et sa matrice de covariance d'erreur associée $\mathbf{R_k}$ (symbolisée par l'ellipse orange). La comparaison entre le vecteur de contrôle et les observations passe par l'intégration du modèle sur la fenêtre d'assimilation $[t_{k-1}, t_k]$. Les équations du filtre de Kalman permettent d'estimer le vecteur de contrôle analysé $\mathbf{x_k^a}$ (point rouge) comme une correction de l'ébauche $\mathbf{x_k^b}$ étant donnée l'observation $\mathbf{y_k^o}$, ainsi que sa matrice de covariance d'erreur associée $\mathbf{A_k}$ (ellipse rouge).

Le filtre de Kalman se décompose en deux étapes qui s'enchaînent de manière itérative, la phase d'analyse et la phase de prévision :

- La phase d'analyse consiste à mettre à jour le vecteur de contrôle sur le cycle c_k grâce aux observations disponibles sur la fenêtre d'assimilation selon les équations d'analyse (voir annexe N.3). La trajectoire d'ébauche $\mathbf{s_k^b}$ est estimée par le modèle numérique lancé sur $[t_{k-1}, t_k]$ avec les paramètres d'ébauche $\mathbf{x_k^b}$ à partir de l'état analysé du cycle précédent

$\mathbf{x^a_{k-1}}$ (équation 3.14). Les équations d'analyse permettent d'obtenir le vecteur de contrôle analysé $\mathbf{x^a_k}$ (équation 3.16).

- La phase de prévision est réalisée en deux temps. L'objectif est préparer le cycle d'assimilation suivant c_{k+1}. Dans un premier temps, elle consiste à propager l'état du modèle sur $[t_{k-1}, t_k]$ à partir de l'état analysé à t_{k-1} avec les paramètres analysés sur le cycle c_k. L'état du modèle à t_k alors obtenu sert d'état initial pour le cycle d'assimilation suivant. Ensuite, il s'agit également de propager le vecteur de contrôle pour le cycle d'assimilation suivant. Le vecteur de contrôle issu de cette propagation sert ensuite de vecteur de contrôle d'ébauche au cycle suivant. Puisqu'on fait l'hypothèse qu'il n'y a pas d'évolution des paramètres entre les cycles d'assimilation, il est propagé par la fonction identité à laquelle une perturbation peut être ajoutée pour éviter un effondrement du filtre au fil des cycles d'assimilation. On parle alors d'inflation. Ce principe sera détaillé au paragraphe 3.4.3.3.

3.4.3.2 Le Filtre de Kalman Étendu

Dans l'algorithme du filtre de Kalman, le modèle d'évolution et l'opérateur d'observation sont supposés linéaires par rapport à l'état. Cependant, cette hypothèse n'est souvent pas valide dans la pratique. L'algorithme du filtre de Kalman peut être généralisé lorsque les opérateurs $\mathcal{M}$ et $\mathcal{H}$ sont non linéaires. Ce filtre est alors appelé filtre de Kalman étendu (*Extended Kalman Filter*, ou EKF). On utilise alors des formes linéarisées localement de l'opérateur d'observation et du modèle d'évolution dans les équations 3.15 et 3.17 du filtre de Kalman et les formes non linéaires dans les équations 3.14, 3.16 et 3.18. Les formes linéarisées sont définies par :

$$\mathbf{H_k} = (\frac{\partial \mathcal{H}_k}{\partial x_k})_{x\,=\,x^b_{k-1}}$$
$$\mathbf{M_{k-1,k}} = (\frac{\partial \mathcal{M}_{k-1,k}}{\partial x_k})_{x\,=\,x^b_{k-1}}$$

Ils représentent la linéarisation de $\mathcal{H}_k$ et de $\mathcal{M}_{k-1,k}$ au voisinage de x^b_{k-1}. Dans ce cas, contrairement au filtre de Kalman, l'analyse n'est plus optimale, c'est-à-dire que le filtre de Kalman étendu ne fournit pas la solution de variance minimale. Dans le cas faiblement non linéaire, l'utilisation de l'EKF permet d'obtenir de bonnes analyses. Il est à noter que la qualité de l'analyse est significativement dégradée dans le cas de modèles fortement non linéaires.

Un inconvénient majeur de l'EKF est qu'il est en pratique inutilisable lorsque le vecteur de contrôle est de très grande taille (taille supérieure à 100), comme cela peut être le cas des modèles météorologiques ou océanographiques où le vecteur de contrôle est constitué du vecteur d'état. En effet, l'algorithme nécessite le stockage et la manipulation de matri-

ces de covariance d'erreur de dimension (n_Y, n_X). Des alternatives à l'EKF ont alors été développées pour réduire le nombre d'intégrations du modèle et en ne propageant la matrice de covariance d'erreur que dans certaines directions. On peut citer le filtre de Kalman de rang réduit (RRSQRT pour *Reduced Rank SQuare RooT*, Verlaan and Heemink [1997]), le filtre SEEK (*Singular Evolutive Extended Kalman filter*, Tuan Pham et al. [1998]) ou sa variante SEIK (*Singular Evolutive Interpolation Kalman filter*, Pham [2001]).

3.4.3.3 Le filtre de Kalman d'ensemble stochastique (EnKF)

Le filtre de Kalman d'ensemble, ou EnKF (*Ensemble Kalman Filter*), a été introduit par Evensen [1994] puis amélioré quatre ans plus tard par Burgers [1998]. Une version détaillée de l'EnKF est disponible dans Asch et al. [2016], Evensen [2003]. L'EnKF représente la version stochastique du filtre de Kalman. L'utilisation d'une méthode stochastique permet de pallier les principales difficultés relatives à l'utilisation du filtre de Kalman étendu, en particulier son coût de calcul dans le cas de systèmes de grande taille. Etant basé sur l'utilisation d'un ensemble pour estimer l'état du modèle et les matrices de covariance d'erreur, il permet également de contourner le problème de la définition de la matrice de covariance d'erreur d'ébauche $\mathbf{B}$ dans le cas du contrôle de l'état et des matrices $\mathbf{BH^T}$ et $\mathbf{HBH^T}$ dans le cas du contrôle des paramètres. Tout comme le filtre de Kalman et le filtre de Kalman étendu, l'EnKF est un filtre gaussien. L'analyse est réalisée uniquement à partir de l'estimation des moyennes et des covariances. Cependant, contrairement au filtre de Kalman, l'EnKF ne fait pas l'hypothèse explicite de la linéarité du modèle. Il est à noter que l'EnKF ne converge avec l'augmentation du nombre de membres vers le filtrage bayésien que dans le cas linéaire gaussien. Un nombre limité de membre (jusqu'à 100 membres) est souvent suffisant dans de nombreuses applications. Les simulations de l'ensemble sont facilement réalisables sur plusieurs processeurs limitant ainsi le temps de calcul de l'algorithme.

L'idée à l'origine de l'EnKF est d'approcher l'estimation du vecteur de contrôle analysé $\mathbf{x^a}$ (respectivement du vecteur de contrôle d'ébauche $\mathbf{x^b}$) et sa matrice de covariance associée $\mathbf{A}$ (respectivement $\mathbf{B}$) par un ensemble de N vecteurs de contrôle du système $(\mathbf{x_k^1}, \mathbf{x_k^2}, ..., \mathbf{x_k^N})$. La figure 3.15 illustre le fonctionnement du filtre de Kalman d'ensemble appliqué dans le même cas simplifié que précédemment (voir paragraphe 3.4.3.1 pour les hypothèses sous lesquelles ce schéma est valable.)

Le vecteur de contrôle du modèle au temps t_k est estimé par la moyenne des membres :

$$\overline{\mathbf{x_k}} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{x_k^i}.$$

L'erreur d'ébauche ou d'analyse, appelée également anomalie, est estimée au temps t_k

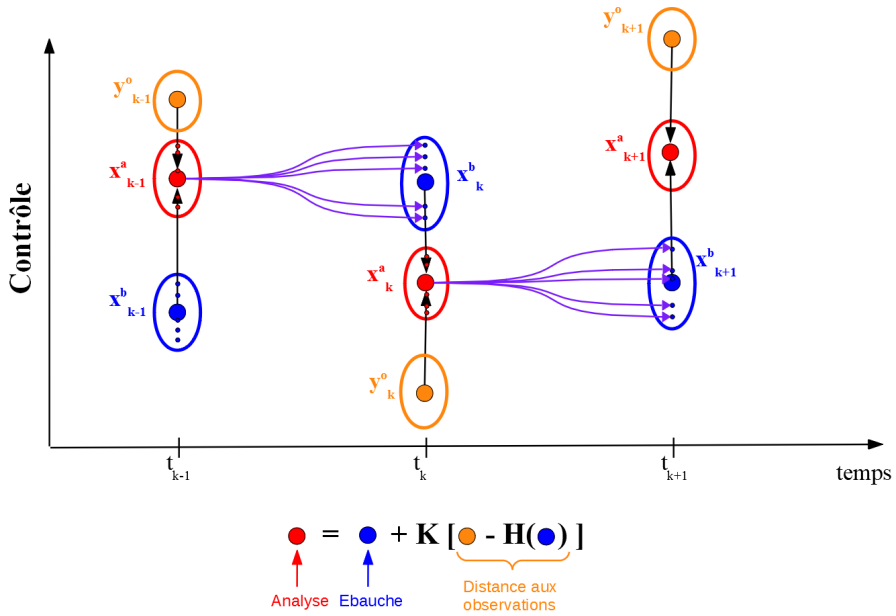


Figure 3.15: Fonctionnement séquentiel du filtre de Kalman d'ensemble - Cas simplifié (voir texte). À chaque instant t_k, nous avons à disposition : (i) un ensemble de vecteurs de contrôle d'ébauche $\mathbf{x_k^{b,i}}$ (petits points bleus) de moyenne $\overline{\mathbf{x_k^b}}$ (point bleu), obtenus à partir de l'ensemble analysé moyen à l'instant précédent $\overline{\mathbf{x_{k-1}^a}}$ auquel est ajouté une pertubation (tirage dans une loi centrée en 0 et d'écart-type choisi initialement), et sa matrice de covariance d'erreur associée $\mathbf{B_k}$ (symbolisée par l'ellipse bleue), et (ii) une observation $\mathbf{y_k^o}$ (point orange) et sa matrice de covariance d'erreur associée $\mathbf{R_k}$ (symbolisée par l'ellipse orange). La comparaison entre le vecteur de contrôle et les observations passe par l'intégration du modèle sur la fenêtre d'assimilation $[t_{k-1}, t_k]$ qui fournit les états d'ébauche correspondants. Les équations du filtre de Kalman permettent d'estimer le vecteur de contrôle analysé $\mathbf{x_k^{a,i}}$ pour chacun des membres de l'ébauche (petits points rouges). Ils ont pour moyenne $\overline{\mathbf{x_k^a}}$ (point rouge) et pour matrice de covariance d'erreur $\mathbf{A_k}$ (ellipse rouge).

pour un membre i par

$$\epsilon_{\mathbf{k}}^{\mathbf{i}} = \mathbf{x_k^i} - \overline{\mathbf{x_k}}.$$

On définit alors la matrice $\mathbf{X_k}$ dont les colonnes sont les vecteurs des anomalies de l'ensemble du vecteur de contrôle : $\mathbf{X_k} = [\epsilon_{\mathbf{k}}^{\mathbf{1}}, \epsilon_{\mathbf{k}}^{\mathbf{2}}, ..., \epsilon_{\mathbf{k}}^{\mathbf{N}}]$. De même, on définit la matrice $\mathbf{Y_k}$ dont les colonnes sont les vecteurs des anomalies de l'ensemble des états simulés par le modèle. Dans l'EnKF, les matrices des covariances d'erreur d'ébauche $\mathbf{B_k}$ ou d'analyse $\mathbf{A_k}$ peuvent théoriquement être estimées par la formule :

$$\overline{\epsilon_{\mathbf{k}}(\epsilon_{\mathbf{k}})^T} = \frac{1}{N-1}\sum_{i=1}^{N}(\mathbf{x_k^i} - \overline{\mathbf{x_k}}^b)(\mathbf{x_k^i} - \overline{\mathbf{x_k}})^T = \frac{1}{N-1}\mathbf{X_k}(\mathbf{X_k})^{\mathbf{T}}$$

donc

$$\mathbf{A_k} = \frac{1}{N-1}\mathbf{X_k^a}(\mathbf{X_k^a})^\mathbf{T}$$
$$\mathbf{B_k} = \frac{1}{N-1}\mathbf{X_k^b}(\mathbf{X_k^b})^\mathbf{T}.$$

En pratique, ces matrices ne sont pas estimées et on calcule directement les matrices du filtre de Kalman $\mathbf{BH^T}$ et $\mathbf{HBH^T}$ de manière stochastique :

$$\mathbf{B_k H_k^T} = \overline{\mathbf{X_k}(\mathbf{Y_k})^T} \tag{3.20}$$
$$\mathbf{H_k B_k H_k^T} = \overline{\mathbf{Y_k}(\mathbf{Y_k})^T}. \tag{3.21}$$

Comme pour le filtre de Kalman, le filtre de Kalman d'ensemble est constitué des phases d'analyse et de prévision.

Phase d'analyse au cycle $c_k = [t_{k-1}, t_k]$ Comme pour le filtre de Kalman, la trajectoire d'ébauche est calculée pour chaque membre selon l'équation 3.14 à partir de la phase prévision du cycle précédent :

$$\mathbf{s_k^{b,i}} = \mathcal{M}_{k-1,k}(\mathbf{s_{k-1}^{a,i}}, \mathbf{x_k^{b,i}}). \tag{3.22}$$

L'analyse est réalisée en appliquant les équations du filtre de Kalman (voir équations 3.15 et 3.16) à chaque membre de l'ensemble grâce à l'estimation de $\mathbf{BH_k^T}$ et $\mathbf{HBH_k^T}$ données précédemment (équations 3.20 et 3.21). La matrice de gain, commune à tous les membres de l'ensemble, est d'abord évaluée par par :

$$\mathbf{K_k} = \mathbf{B_k H_k^T}(\mathbf{H_k B_k H_k^T} + \mathbf{R_k})^{-\mathbf{1}}.$$

Le vecteur de contrôle analysé est ensuite calculé pour chaque membre de l'ensemble :

$$\mathbf{x_k^{a,i}} = \mathbf{x_k^{b,i}} + \mathbf{K_k}(\mathbf{y_k^{o,i}} - \mathbf{H_k x_k^{b,i}})$$

où $\mathbf{y_k^{o,i}} = \mathbf{y_k^o} + \epsilon_k^i$, les ϵ_k^i suivent une loi normale de moyenne nulle et de matrice de covariance $\mathbf{R_k}$ prescrite. Dans la version initiale du filtre de Kalman d'ensemble [Evensen, 1994], l'observation assimilée par chacun des membres est l'observation $\mathbf{y_k^o}$. C'est donc la même pour chacun des membres. Les conséquences sont alors une sous-estimation de la variance de l'ensemble analysé. Burgers [1998] a proposé d'assimiler des observations perturbées pour conserver la variance de l'ensemble. Dans ce cas, si l'opérateur d'observation est linéaire, la phase d'analyse tend avec l'augmentation du nombre de membres vers les estimations données par le BLUE dans le cas linéaire. L'EnKF approche alors le filtre de

Kalman. Cette version du filtre de Kalman d'ensemble est appelée *version stochastique*. D'autres versions appelées *versions déterministes* sont proposées pour éviter le problème de la sous-estimation de la variance de l'ensemble analysé. Elles ne seront pas évoquées dans ce manuscrit. [Carrassi et al., 2018] décrit de manière approfondie les deux familles de méthodes.

Phase de prévision pour le cycle suivant $c_{k+1} = [t_k, t_{k+1}]$ Comme pour le filtre de Kalman, la phase de prévision est réalisée en deux temps : il faut propager l'état du modèle et le vecteur de contrôle jusqu'au cycle d'assimilation suivant.

Chacun des membres analysés est d'abord propagé à nouveau sur la fenêtre d'assimilation du cycle $c_k = [t_{k-1}, t_k]$ pour obtenir les états analysés du modèle en équilibre avec le vecteur de contrôle analysé qui serviront d'état initial au cycle suivant[6]. Pour chaque membre de l'ensemble, la propagation est réalisée comme pour le filtre de Kalman (équation 3.18) :

$$\mathbf{s_k^{a,i}} = \mathcal{M}_{k-1,k}(\mathbf{s_{k-1}^{a,i}}, \mathbf{x_k^{a,i}}).$$

Cette étape permet en outre d'étudier *a posteriori* la qualité des simulations analysées. La seconde étape de la phase de prévision concerne la propagation du vecteur de contrôle. Un modèle d'évolution du vecteur de contrôle de cycle à cycle pourrait exister, mais nous faisons le choix de prendre l'identité. Pour chaque membre, le vecteur de contrôle utilisé au début du cycle d'assimilation suivant est donc :

$$\mathbf{x_{k+1}^{b,i}} = \mathbf{x_k^{a,i}} + \boldsymbol{\nu_k^i}.$$

ν_k^i représente un bruit additif permettant en particulier de prendre en compte l'erreur modèle considérée comme nulle et empêchant le filtre de s'effondrer. Dans ce travail, il a été choisi un bruit suivant une loi de probabilité uniforme centrée en 0 et dont la largeur est égale à celle de l'intervalle initial prescrit. Cela correspond à une inflation *additive* (voir ci-après).

Problèmes liés à l'échantillonnage et solutions Dans le filtre de Kalman d'ensemble, on approxime les matrices de covariance d'erreur d'ébauche et d'analyse à partir d'un nombre fini de membres. Cette méthode introduit des erreurs d'échantillonnage. En particulier, lorsque le nombre de membres utilisés pour estimer la matrice de covariance d'erreur d'ébauche $\mathbf{B}$ dans le cas de la correction de l'état et les matrices $\mathbf{BH^T}$ et $\mathbf{HBH^T}$ dans le cas de la correction des paramètres est trop faible, les variances peuvent être sous-estimées. Une solution peut être la mise en place du processus d'inflation.

[6] Cette étape n'est pas représentée sur la figure 3.15

Inflation L'application de l'inflation permet de compenser les problèmes dûs à l'échantillonnage et éventuellement à la non prise en compte de l'erreur modèle dans la phase de propagation. Anderson and Anderson [1999], Hamill et al. [2001] présentent la méthode d'inflation, qui consiste à réinjecter de la dispersion au sein de l'ensemble afin d'éviter son effondrement. On augmente ainsi les anomalies de l'ensemble grâce à l'utilisation d'un coefficient ρ supérieur à 1. Le vecteur de contrôle de chaque membre de l'ensemble est corrigé par la formule suivante :

$$\tilde{\mathbf{x}^{\mathbf{i}}} = \rho(\mathbf{x}^{\mathbf{i}} - \overline{\mathbf{x}}) + \overline{\mathbf{x}}$$

Cette inflation est dite *multiplicative*. La moyenne de l'ensemble n'est pas modifiée par ce processus, mais sa variance est augmentée. Ce processus d'inflation est généralement appliqué juste avant la phase d'analyse et permet d'augmenter l'incertitude sur l'estimation de l'ébauche. Une autre méthode peut être l'ajout de perturbations stochastiques aux membres de l'ensemble Mitchell and Houtekamer [2000]. L'inflation est alors dite *additive*. Comme nous l'avons décrit ci-avant, dans le cadre de ce travail de thèse, une inflation additive a été appliquée.

3.4.4 Mise en œuvre informatique

L'implémentation d'un filtre de Kalman d'ensemble dans les modèles MORDOR-TS et MASCARET a été implémentée à partir du module python SMURF (*System for Modelling with Uncertainty Reduction and Forecasting*, Mirouze and Ricci [2021])[7].

Les interfaces avec le module SMURF ont été complètement développées dans le cadre de cette thèse pour la mise en place d'assimilation de données dans le modèle MORDOR-TS. Dans le cadre de la mise en place d'assimilation de données dans le modèle MASCARET, la version distribuée propose déjà une interface avec le code MASCARET. Elle a été adaptée et améliorée dans ce travail pour travailler dans le contexte du bassin versant de l'Odet.

Les expériences d'assimilation de données sont réalisées sur le calculateur Datarmor hébergé et administré par l'IFREMER.

3.4.5 Assimilation de données dans les modèles hydrologiques et hydrauliques

Les méthodes d'assimilation de données, largement utilisées dans le domaine météorologique, sont désormais également employées de manière courante pour la modélisation hydrologique et hydrodynamique. Le vecteur de contrôle peut être constitué de variables d'état du système et/ou des paramètres des modèles et/ou des conditions aux limites des modèles.

[7] Logiciel open source, licence CecillB, disponible sur https://gitlab.com/cerfacs/Smurf.git

Naturellement, de nombreuses études assimilent des mesures de débit dans les modèles hydrologiques ou hydrauliques. Les observations de débit peuvent provenir de plusieurs sources [Noh et al., 2018]. La méthode la plus classique est la mesure de hauteurs d'eau associée à l'utilisation de courbes de tarage (voir description des mesures hydrométriques au paragraphe 2.2.2). Les mesures satellitaires ont par ailleurs permis l'extension de la couverture spatiale des observations de débits sur la majeure partie du globe. La future mission SWOT (*Surface Water and Icean Topography*) qui devrait démarrer en 2022 fournira les cotes de surface libre, les pentes et les largeur des rivières pour les cours d'eau de largeur supérieure à 100 m. Afin de préparer l'utilisation future de ces observations, de nombreux projets de recherche ont pour objet la réalisation d'expériences jumelles en construisant des observations de type SWOT. L'incertitude sur les mesures de débits est fondamentalement liée aux techniques utilisées pour les mesurer [McMillan et al., 2012].

Les deux paragraphes suivants s'attachent à décrire l'état de l'art concernant la mise en place d'assimilation de données pour la modélisation hydrologique (paragraphe 3.4.5.1) et hydraulique fluviale (paragraphe 3.4.5.2).

3.4.5.1 L'assimilation de données pour la modélisation et la prévision hydrologiques

Les filtres de Kalman sont les techniques d'assimilation les plus couramment utilisées en hydrologie, et particulièrement le filtre de Kalman d'ensemble [Moradkhani et al., 2019]. Les méthodes variationnelles peuvent en effet être difficiles à mettre en œuvre à cause de la complexité potentielle du développement du modèle adjoint pour les modèles fortement non linéaires. Bien que les méthodes utilisées pour l'assimilation de données en hydrologie ont pour origine les méthodes développées pour la météorologie et l'océanographie, la nature de l'assimilation est différente. En effet, les processus hydrologiques ont lieu sur une échelle de temps beaucoup plus longue. Par ailleurs, les modèles météorologiques sont chaotiques, ce qui signifie qu'une erreur dans la condition initiale peut être à l'origine d'erreurs qui s'amplifient avec le temps. L'assimilation de données en météorologie a donc surtout vocation à corriger la condition initiale. Au contraire, les modèles hydrologiques sont de nature dissipative, et les erreurs dans les conditions initiales vont s'effacer progressivement. L'objectif de correction de l'état initial est alors fréquemment associé à un objectif de correction des paramètres du modèle et/ou des forçages [Lü et al., 2011, Leisenring and Moradkhani, 2011, Moradkhani et al., 2005, Nie et al., 2011, Samuel et al., 2014]. Certains auteurs ont étudié l'impact de l'assimilation de données dans un contexte de prévision opérationnelle [Dechant and Moradkhani, 2011, DeChant and Moradkhani, 2011, Seo et al., 2003, 2009, Thirel et al., 2010a,b, Weerts et al., 2010]. Liu et al. [2012] font une synthèse des méthodes d'assimilation de données mises en œuvre pour la prévision hydrologique, que ce soit du point de vue de la recherche ou en opérationnel. Cet

article fait suite à un workshop qui a eu lieu de novembre 2010 à Delft (Pays-Bas) sur les méthode d'assimilation de données pour la prévision hydrologique opérationnelle et pour la gestion de la ressource en eau.

Dans le domaine de la modélisation hydrologique, les données assimilées sont variées. De nombreuses études ont pour objet l'assimilation de données d'humidité du sol. Kumar et al. [2014] assimilent des observations satellitaires d'humidité du sol et de hauteur de neige dans le modèle de surface continentale[8] Noah dans sa configuration NLDAS (North American Land Data Assimilation System) [Ek et al., 2003] avec un filtre de Kalman d'ensemble. Ils mettent en évidence que l'assimilation de l'humidité du sol améliore la modélisation des débits pour des échelles de temps courtes, et que l'impact de l'assimilation de données neigeuses est négligeable. Reichle et al. [2002] assimilent eux aussi des observations satellitaires d'humidité du sol dans un modèle de surface continental (décrit dans [Reichle et al., 2001]) avec un filtre de Kalman d'ensemble. Des expériences jumelles permettent d'évaluer l'effet de la taille de l'ensemble et d'erreurs modèle non gaussiennes sur la qualité de l'algorithme d'assimilation. Un ensemble de taille 30 (respectivement 100 et 500) permet ainsi de réduire les erreurs sur l'humidité du sol de 55% (respectivement 70% et 80%) par rapport aux valeurs obtenues sans assimilation. Dans De Lannoy et al. [2007], les auteurs corrigent le biais dans la prévision de l'humidité du sol avec un filtre de Kalman d'ensemble dans le modèle CLM (*Community Land Model*) grâce à des prélèvements *in situ*.

L'assimilation de données de couverture du manteau neigeux[9] et du volume d'eau équivalent[10] fait également l'objet de plusieurs études. Andreadis and Lettenmaier [2006] assimilent des observations satellitaires de couverture du manteau neigeux et de volume d'eau équivalent dans le modèle hydrologique grande échelle VIC (*Variable Infiltration Capacity*, Liang et al. [1994]) sur le bassin du Snake (affluent du Columbia, à la frontière entre les États-Unis et le Canada) grâce à un filtre de Kalman d'ensemble. Ils mettent en évidence que cette méthode améliore les prévisions de couverture neigeuse, mais l'amélioration est plus modeste concernant l'estimation du volume d'eau équivalent. Slater and Clark [2006] mettent à jour le volume d'eau équivalent à la neige dans le modèle conceptuel SNOW-17 [Anderson, 1973] du bassin amont du Colorado grâce à un filtre de Kalman d'ensemble. Ils montrent que les résultats sont meilleurs au début de la saison d'accumulation neigeuse. De Lannoy et al. [2012] assimilent des observations satellitaires de couverture de manteau neigeux et de volume d'eau équivalent dans le modèle de surface continental Noah appliqué dans le nord du Colorado grâce à un filtre de Kalman d'ensemble. Leisenring and Moradkhani [2011] comparent différentes méthodes d'assimilation de données pour améliorer la prévision saisonnière de volume d'eau équivalent dans le modèle SNOW-17 sur un bassin

[8] LSM pour *Land Surface Model* en anglais. Ce sont des modèles qui simulent les interactions entre le sol, la végétation et l'atmosphère

[9] SCE pour *Snow Cover Extent*

[10] SWE pour *Snow Water Equivalent*

californien. Ils concluent que le filtre à particules fournit de meilleurs résultats que le filtre de Kalman d'ensemble pour la prévision des états du modèle et de ses paramètres.

De nombreuses études assimilent des observations de débit pour mettre à jour l'état du modèle hydrologique et/ou ses paramètres. Clark et al. [2008a] assimilent des observations de débit pour mettre à jour les états dans le modèle hydrologique distribué Top-Net [Bandaragoda et al., 2004] avec un filtre de Kalman d'ensemble. Ils montrent que l'implémentation standard est inappropriée à cause de la relation non-linéaire entre les états du modèle et les observations. Une transformation logarithmique appliquée sur les débits avant de calculer les covariances d'erreur améliore les performances du filtre. Moradkhani et al. [2005] cherchent à améliorer la prévision hydrologique réalisée avec le modèle conceptuel HyMOD [Wagener et al., 2001] en corrigeant ses états et ses paramètres avec un filtre de Kalman d'ensemble. Ils montrent que l'état et les paramètres peuvent être corrigés de manière simultanée. Samuel et al. [2014] utilisent eux aussi un filtre de Kalman d'ensemble pour corriger à la fois l'état du modèle SAC-SMA (*Sacramento Soil Moisture Accounting*) et ses paramètres. Les données assimilées sont l'humidité du sol et des débits. Elles sont assimilées de manière séparée ou combinée. Lorsque seuls les débits sont assimilés, l'humidité du sol est notablement déformée, surtout pour les couches les plus basses. Lorsque seule l'humidité du sol est assimilée, les débits modélisés sont de mauvaise qualité. L'assimilation combinée des débits et de l'humidité du sol fournit de meilleures prévisions d'humidité du sol et de débits, surtout pour les courtes échéances. Emery et al. [2018] mettent en place un filtre de Kalman d'ensemble dans le modèle de surface continentale ISBA-CTRIP et assimilent des données de débits construites à partir de données satellitaires de hauteur d'eau et de courbes de tarage sur le bassin de l'Amazone. Le système d'assimilation corrige soit le volume d'eau stockée dans le modèle soit les débits. Les auteurs montrent que l'assimilation mise en place permet de réduire les erreurs de modélisation. Dans Emery et al. [2019], les auteurs poursuivent leurs investigations. Ils assimilent des observations satellitaires de hauteur d'eau de type SWOT dans des expériences jumelles avec un filtre de Kalman d'ensemble asynchrone pour corriger les coefficients de Manning du module de routage du modèle ISBA-CTRIP. L'algorithme permet le plus souvent de retrouver les vraies valeurs des coefficients au bout d'un cycle d'assimilation. Häfliger et al. [2019] réalisent eux aussi des expériences jumelles en assimilant des observations de hauteur d'eau de type SWOT dans un filtre de Kalman étendu dans le modèle hydrométéorologique régional ISBA-MODCOU de la Garonne pour corriger les coefficients de frottement.

En ce qui concerne le modèle hydrologique MORDOR, [Farouk et al., 2020] utilisent un filtre à particules dans le modèle MORDOR-SD[11] [Garavaglia et al., 2017] en assimilant des données de débit et de stock de neige. L'étude montre les avantages liés à l'assimilation de données *in situ* sur la qualité des simulations et des prévisions. Pour finir, Cheng et al.

[11] Semi-Distribué

[2021] comparent différentes stratégies d'assimilation de données variationnelle dans le modèle MORDOR-TS du Tarn (sud de la France) au pas de temps journalier. Ils corrigent les remplissages des réservoirs au début de la fenêtre d'assimilation et les précipitations journalières sur la fenêtre en assimilant les débits journaliers. Ils s'intéressent particulièrement à la détermination des matrices de covariances **B** et **R**, dont la définition impacte directement les performances de l'algorithme.

3.4.5.2 L'assimilation de données pour la modélisation et la prévision en hydraulique fluviale

Les méthodes d'assimilation de données appliquées aux modèles hydrauliques ont généralement pour objectif l'amélioration de la simulation et/ou de la prévision des niveaux d'eau et/ou des vitesses des systèmes modélisés, à partir de mesures *in situ* ou de mesures satellitaires. Afin de réduire l'incertitude du système modélisé, différentes approches sont employées dans la littérature, utilisées seules ou de manière combinée : correction de l'état hydraulique seul, correction des forçages et correction des paramètres des modèles au sens large (coefficients de frottement, bathymétrie).

De nombreuses études utilisent un filtre de Kalman pour assimiler des données dans les modèles hydrauliques. Ricci et al. [2011] utilisent un filtre de Kalman étendu pour corriger en deux étapes les forçages à l'amont et l'état hydraulique du modèle 1D MASCARET en assimilant des mesures de hauteur d'eau. Cette méthodologie est appliquée à deux bassins versants, celui de l'Adour (sud-ouest de la France) et celui de la Marne (à l'est de Paris). La matrice de covariance d'erreur d'ébauche est considérée invariante au cours du temps afin de réduire le coût de calcul de l'algorithme dans un objectif opérationnel. Habert et al. [2016] poursuivent cette étude sur le bassin de la Marne amont, en cherchant à corriger les forçages amont et les coefficients de frottement dans un cadre opérationnel. L'assimilation de données permet alors d'améliorer significativement les hauteurs prévues jusqu'à l'échéance 12 heures. Madsen and Skotner [2005] utilisent aussi un filtre de Kalman avec matrice de covariance invariante pour corriger un modèle opérationnel de prévision des crues aux Philippines. L'assimilation de données est associée à un système de prévision de l'erreur, basé sur la différence entre les prévisions passées et les observations correspondantes, qui permet d'améliorer les résultats en prévision sur une échéance plus longue. Barthélémy et al. [2017] mettent en place un système d'assimilation de données dans un cadre opérationnel sur le bassin de l'Adour fondé sur le code 1D MASCARET avec un filtre de Kalman d'ensemble. Le vecteur de contrôle est composé de l'état du système hydraulique (hauteurs et débits) et les données assimilées sont des observations in situ. L'estimation stochastique de l'ébauche est réalisée par des simulations du modèle dont les conditions aux limites hydrologiques sont perturbées. En prévision, les forçages hydrologiques sont supposés constants et égaux aux derniers forçages disponibles avant la

phase de prévision. Les performances de l'algorithme sont évaluées sur une série de 8 événements entre 2009 et 2014. Ils montrent que l'ensemble possède de bonnes caractéristiques tant que l'échéance de prévision est plus courte que le temps de transfert du réseau. Les variables simulées sont améliorées sur tout le réseau donc même pour des points non observés avec un ensemble de taille relativement réduit (40 membres), ce qui est compatible avec des contraintes opérationnelles. L'amélioration de la RMSE pour la variable hauteur d'eau est de 88% pour l'analyse et de 40% pour la prévision à échéance 4 heures par rapport au modèle sans assimilation. Neal et al. [2007] appliquent un filtre de Kalman d'ensemble pour corriger les états d'un modèle hydraulique 1D dans un cadre de prévision opérationnelle. Ils montrent que l'utilisation de l'EnKF permet d'améliorer les précisions des prévisions entre 50% et 70% selon les zones. Biancamaria et al. [2011] appliquent un filtre de Kalman d'ensemble à un modèle couplé hydrologie-hydraulique dans un cadre synthétique avec des observations virtuelles type SWOT (*Surface Water and Icean Topography*). Le modèle ISBA (Interactions between the Soil-Biosphere-Atmosphere, Noilhan and Mahfouf [1996]) fournit des débits qui servent de condition aux limites au modèle LISFLOOD-FP (Flood Plain, modèle couplé 1D-2D, Bates and De Roo [2000]). Yoon et al. [2016] appliquent eux aussi un filtre de Kalman pour corriger des débits à partir d'observations de type SWOT au modèle 1D HEC-RAS de la Garonne ou de la rivière de Sacramento. Andreadis et al. [2007] effectuent eux aussi des expériences jumelles dans le cadre d'un filtre de Kalman d'ensemble avec le modèle LISFLOOD-FP. Ils assimilent des hauteurs d'eau pour améliorer la simulation des débits. Le filtre permet de retrouver les hauteurs et débits synthétiques, en réduisant les RMSE de 10.0% à 23.2% par rapport à une simulation sans assimilation. Ils mettent en place un système de prévision des erreurs comme dans Madsen and Skotner [2005]. Durand et al. [2008] et Yoon et al. [2012] assimilent des données altimétriques avec un filtre de Kalman d'ensemble pour corriger la bathymétrie d'un modèle hydraulique respectivement sur les fleuves de l'Amazone et de l'Ohio dans le cadre d'expériences jumelles avec des données de type SWOT. Ils ont montré la capacité de l'assimilation de données à corriger les bathymétries, ce qui a pour conséquence une meilleure estimation des débits et des hauteurs d'eau. La correction de la bathymétrie peut être pertinente dans les zones pour lesquelles on possède peu de mesures, ou après le passage d'une forte crue.

Des méthodes variationnelles sont également appliquées pour l'assimilation de données dans les modèles hydrauliques. Hostache et al. [2010] appliquent une méthode variationnelle (4D-VAR) pour assimiler des observations satellitaires de hauteurs d'eau SAR dans un modèle 2D. Ils montrent une amélioration de la calibration du modèle en estimant des coefficients de Manning du modèle. Brisset et al. [2018] mettent en place une assimilation de données variationnelle dans le cadre d'expériences jumelles avec des observations altimétriques de type SWOT dans un modèle hydraulique 1D Saint-Venant. Ils corrigent le débit et les coefficients de frottement en fonction du temps. Oubanas [2018] réalisent des

expériences jumelles dans le cadre d'une assimilation de données variationnelle de type 4D-Var dans le modèle SIC^2 (*Simulation and Integration of Control for Canals*, INRAE) de la Garonne sur une période de 5 mois en 2010 pendant laquelle ont eu lieu plusieurs événements de crue. L'objectif est l'évaluation simultanée du débit, de la bathymétrie et de la rugosité en assimilant des observations satellitaires virtuelles de type SWOT et des mesures *in situ*. Ils montrent que les techniques d'assimilation mises en place fonctionnent pour la correction de toutes ces variables.

3.4.6 Synthèse

Dans ce paragraphe, nous avons décrit les principes généraux de l'assimilation de données dans les modèles numériques. Nous avons décrit en détails les algorithmes des filtres de Kalman (filtre de Kalman, filtre de Kalman étendu et filtre de Kalman d'ensemble) dans le cas d'un vecteur de contrôle constitué de paramètres du modèle et de paramètres pour la correction des conditions aux limites. La description de l'état de l'art s'est focalisée sur le domaine des géosciences, en particulier sur l'assimilation de données dans les modélisations hydrologiques et hydrauliques.

Nous présentons à présent l'assimilation de données mise en place au sein des modèles MORDOR-TS et MASCARET dans le cadre de ce travail. Pour les deux modèles, on considère une fenêtre d'assimilation de 3 heures et une fenêtre de prévision opérationnelle de 24 heures, comme présenté sur la figure 3.16.

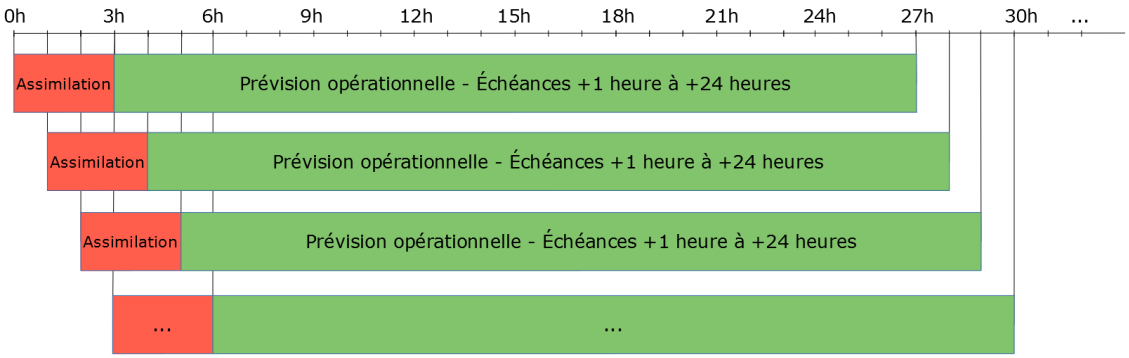


Figure 3.16: Fenêtres d'assimilation et de prévision opérationnelle mises en place pour l'assimilation de données dans les modèles MORDOR-TS et MASCARET.

Ici, la fenêtre d'assimilation est $[t_{k-3}, t_k]$, et des observations sont assimilées à t_{k-2}, t_{k-1} et t_k. Sur chaque fenêtre d'assimilation il y a donc trois pas de temps d'observation. La phase d'analyse de l'EnKF est réalisée sur la fenêtre d'assimilation. Nous attirons l'attention sur le fait que la phase de prévision opérationnelle n'est pas la même que la phase de prévision de l'algorithme EnKF présentée au paragraphe 3.4.3.3. La prévision dans l'algorithme EnKF permet de passer d'un cycle d'assimilation au suivant. La prévision opérationnelle

indiquée dans les fenêtres vertes de la figure 3.16 correspond à la réalisation des prévisions d'ensemble de manière opérationnelle.

Les fenêtres *Assimilation + Prévision opérationnelle* sont cyclées de manière horaire, c'est-à-dire que le début de la fenêtre d'assimilation est décalé d'une heure à chaque cycle d'assimilation. Le passage d'un cycle au suivant est réalisé à partir de la phase *prévision* de l'algorithme EnKF.

Phase d'analyse de l'EnKF Pour chaque cycle *Assimilation + Prévision opérationnelle*, la phase d'analyse est réalisée sur la fenêtre d'assimilation selon l'algorithme suivant :

- Pour chaque membre de l'ensemble, le vecteur de contrôle d'ébauche a été déterminé à la fin de la phase prévision de l'EnKF du cycle d'assimilation précédent. L'ensemble est propagé par le modèle ce qui constitue la trajectoire d'ébauche sur la fenêtre d'assimilation.
- À partir de la trajectoire d'ébauche et des mesures disponibles sur la fenêtre d'assimilation, l'équation du filtre de Kalman est appliquée pour corriger le vecteur de contrôle pour chacun des membres.
- Les membres sont à nouveau propagés par le modèle sur la même fenêtre d'assimilation avec pour paramètres le vecteur de contrôle analysé : c'est la trajectoire d'analyse qui a été simplifiée dans les schémas 3.14 et 3.15. L'objectif est d'obtenir pour chacun des membres un état sur la fenêtre d'assimilation qui soit en en équilibre avec les paramètres analysés. La façon de faire est alors différente selon que l'on travaille avec le modèle MASCARET ou le modèle MORDOR-TS. Avec le modèle MASCARET, la simulation est relancée depuis le début de la fenêtre d'assimilation. On rejoue donc 3 heures. L'état du modèle au bout d'une heure est sauvegardé pour chacun des membres pour servir d'état initial au cycle d'assimilation suivant puisqu'on décale la fenêtre d'assimilation d'une heure à chaque nouveau cycle. Cela correspond à la première étape de la phase de prévision de l'algorithme EnKF (voir paragraphe 3.4.3.3). Avec le modèle MORDOR-TS, il est nécessaire de relancer la simulation un an auparavant pour que l'état du modèle soit cohérent avec l'historique des précipitations et les nouveaux paramètres du modèle. Il n'est donc pas nécessaire de sauvegarder l'état au bout d'une heure pour le cycle suivant.

Après ces étapes, nous nous trouvons au temps de base (fin de la fenêtre d'assimilation) avec des membres dont l'état est en équilibre avec la valeur du vecteur de contrôle analysé. Nous sommes alors en mesure de réaliser une prévision d'ensemble sur la fenêtre de prévision opérationnelle. Les membres de la prévision d'ensemble sont les membres de l'algorithme EnKF. Les forçages alors utilisés sont des forçages disponibles en phase de prévision : prévision de pluie pour la prévision avec MORDOR-TS, prévision de débits et de cote aval pour la prévision avec MASCARET. L'ensemble des membres prévus est considéré comme une prévision d'ensemble dont les performances seront évaluées selon les critères définis au paragraphe 3.1.

Phase de prévision de l'EnKF La phase de prévision de l'EnKF permet de passer au cycle d'assimilation suivant. Elle est réalisée en prescrivant pour chacun des membres la valeur du vecteur de contrôle d'ébauche utilisé au cycle suivant. Il est tiré aléatoirement dans des distributions uniformes centrées sur la valeur moyenne analysée au cours du cycle courant et de largeur égale à l'intervalle initial. Pour le modèle MASCARET, l'état initial pour le cycle suivant est l'état sauvegardé au bout d'une heure au cours de la simulation de la trajectoire d'analyse. En revanche, le modèle MORDOR-TS est relancé un an auparavant avec les nouveaux paramètres d'ébauche.

La mise en place d'assimilation de données avec l'algorithme EnKF sera d'abord évaluée dans le cadre d'expériences jumelles puis d'expériences réelles. Ce travail fait l'objet du chapitre 5 pour la prévision hydrologique d'ensemble avec le modèle MORDOR-TS et du chapitre 9 pour la prévision hydraulique d'ensemble avec le modèle MASCARET.

Partie II

LES PRÉVISIONS D'ENSEMBLE HYDROLOGIQUES

Aide-mémoire sur les ensembles hydrologiques

Cette partie a pour objectif l'application des méthodes développées au chapitre 3 pour la construction de prévisions d'ensemble hydrologiques. Nous résumons dans les encarts ci-dessous tous les ensembles hydrologiques construits dans cette partie. Le lecteur pourra se reporter à cet aide-mémoire au cours de la lecture de la suite du manuscrit.

raw-HEF

Ensemble brut généré par perturbation des paramètres de calage significatifs pour les débits simulés avec le modèle MORDOR-TS.

QRF-hydro

Ensemble de quantiles générés par la méthode de calibration statistique *Quantile Regression Forests*.
⇒ Parmi les prédicteurs sont utilisés des prédicteurs issus de statistiques de l'ensemble brut *raw-HEF*.

QRF-nothydro

Ensemble de quantiles générés par la méthode de calibration statistique *Quantile Regression Forests*.
⇒ Aucun prédicteur n'est issu de statistiques de l'ensemble brut *raw-HEF*.

QRF-hydro-TS

Ensemble issu de l'ensemble de quantiles *QRF-hydro* :

- la reconstruction est réalisée par la méthode *Ensemble Copula Coupling* (ECC)
- un lissage des trajectoires est réalisé. TS signifie *Trajectory Smoothing*.

Data Assimilation (DA)

Ensemble obtenu après assimilation de données réalisée avec un filtre de Kalman d'ensemble.

La construction des ensembles *raw-HEF*, *QRF-hydro*, *QRF-nothydro* et *QRF-hydro-TS* est présentée dans le chapitre 4 qui traite des études de sensibilité globale des modèles hydrologiques, de la construction de l'ensemble brut *raw-HEF* et de la calibration statistique mise en œuvre. Les études de sensibilité globale ont été réalisées avec les prévisions d'ensemble PEAROME calibrées en forçage. En revanche, l'ensemble brut est ensuite généré avec les prévisions déterministes RR3 pour avoir une couverture temporelle plus importante. La construction de l'ensemble par assimilation de données *Data Assimilation* est décrite dans le chapitre 5. Les performances comparées des ensembles *raw-HEF*, *QRF-hydro*, *QRF-hydro-TS* et *Data Assimilation* sont présentées au chapitre 6.

Chapitre 4

Études de sensibilité et calibration statistique QRF

Sommaire

4.1 Article soumis à *La Houille Blanche* 134

4.2 Article accepté par *Journal of Hydrology* 136

Ce chapitre présente les études de sensibilité globale des modèles hydrologiques et la génération de prévisions d'ensemble hydrologiques brutes et calibrées. Il est construit à partir du résumé de deux articles.

Le premier article est intitulé *Étude de sensibilité globale comparée des modèles hydrologiques GRP et MORDOR-TS*. Il a été soumis au journal *La Houille Blanche*, des révisions mineures sont demandées. Il a pour objet la réalisation des études de sensibilité globale des deux modèles GRP et MORDOR-TS et la comparaison des résultats obtenus pour chacun des modèles. Le paragraphe 4.1 résume la méthodologie développée ainsi que les principaux résultats obtenus.

Les résultats des études de sensibilité globale permettent de définir la méthodologie de génération d'un ensemble de prévision brut appelé *raw-HEF*. Celui-ci est généré uniquement avec le modèle MORDOR-TS. Cet ensemble est ensuite calibré par la méthode *Quantile Regression Forests* (QRF). Cette étude fait l'objet d'un article accepté au journal *Journal of Hydrology*, dont le titre est *Strategies for hydrologic ensemble generation and calibration: on the merits of using model-based predictors*[1] [Tiberi-Wadier et al., 2021]. Le paragraphe 4.2 présente un résumé de cet article. Pour davantage de détails, le lecteur pourra se reporter à la version finale de l'article[2].

4.1 Article soumis à *La Houille Blanche*

Cet article est intitulé *Étude de sensibilité globale comparée des modèles hydrologiques GRP et MORDOR-TS*. Il présente les études de sensibilité globale des modèles GRP et MORDOR-TS sur les trois sous-bassins amont de l'Odet.

Résumé de l'étude La quantité d'intérêt sur laquelle porte l'étude de sensibilité est le débit prévu à l'exutoire de chaque sous-bassin versant hydrologique à un instant donné. L'étude de sensibilité est donc menée de manière temporelle pour les trois stations Tréodet, Kerjean et Ty-Planche. Les paramètres incertains considérés sont : (i) la pluie utilisée à l'entrée des modèles, et (ii) les paramètres de calage de chaque modèle. L'incertitude sur les variables d'entrée est qualifiée de la manière suivante :

- Les prévisions PEAROME étant composées de 12 membres équiprobables, le scénario de pluie est tiré aléatoirement dans une loi uniforme discrète entre 1 et 12. Les prévisions d'ensemble de pluie PEAROME ont été préalablement calibrées par la méthode QRF. Cette étude a fait l'objet d'un poster au colloque *De la prévision des crues à la gestion de*

[1] Stratégies pour la génération et la calibration d'ensembles hydrologiques : mérites de l'utilisation de prédicteurs issus du modèle.

[2] Tiberi-Wadier, A-L., Goutal, N., Ricci, S., Sergent, P., Taillardat, M., Bouttier, F.,Monteil, C., Strategies for hydrologic ensemble generation and calibration: on the merits of using model-based predictors, Journal of Hydrology (2021), doi: https://doi.org/10.1016/j.jhydrol.2021.126233.

crise de la SHF en novembre 2018 et un article a été rédigé[3] [Tiberi-Wadier et al., 2018].

- La plage de variation des paramètres des modèles est déterminée par une série de calages des modèles sur des périodes de 2 ans consécutives sur la période 2007-2017. Pour chaque paramètre, un tirage dans une loi uniforme continue sur la plage ainsi déterminée est réalisé.

L'étude de sensibilité est réalisée en se plaçant dans des conditions proches du contexte opérationnel : les modèles hydrologiques sont alimentés avec les mesures de pluie (ANTILOPE) et de température (SAFRAN) jusqu'au lancement de la prévision, puis par les prévisions PEAROME. En prévision, les données SAFRAN sont toujours utilisées pour la température. On n'est donc pas à proprement parler dans un contexte opérationnel, néanmoins, l'étude de sensibilité se focalise sur les pluies. Dans le mode rejeu appliqué, après une année d'initialisation, les conditions initiales sont successivement reprises des pas de temps antérieurs et les simulations sont réalisées de manière cyclée. Pour chaque membre perturbé, on détermine d'abord ses caractéristiques (valeurs des paramètres du modèle, choix de la pluie PEAROME). Puis, pour tous les pas de temps dans la fenêtre définie pour l'analyse de sensibilité, on lance une prévision de 45 heures alimentée par la PEAROME. Le calcul des indices de Sobol peut donc être effectué par échéance de prévision.

Le calcul de la chronique temporelle des indices de Sobol pour chacun des modèles est réalisé au cours de l'événement du 26 au 30 janvier 2016. Les résultats mettent en évidence une grande similarité dans la réponse des deux modèles bien que les deux modèles hydrologiques soient différents dans leur complexité : modèle spatialisé ou non, un seul ou plusieurs réservoirs de production. En effet, pour les deux modèles, la source d'incertitude prépondérante concerne le bilan en eau : il est contrôlé par le coefficient correctif pour les précipitations et par la chronique de prévision de pluie. À l'échéance 6 heures, nous mettons en évidence que la valeur du débit simulé en début d'événement dépend en partie de la capacité du réservoir qui alimente directement le débit (paramètre *Rout* pour GRP et L_{max} pour MORDOR-TS). À longue échéance (21 heures), la pluie devient une source d'incertitude significative et la courbe représentant l'indice de Sobol de la pluie a la même forme pour les deux modèles. Enfin, les deux modèles ont tendance à sous-estimer le pic de débit de l'événement étudié, que ce soit pour les modèles déterministes (modèles calés) ou ensemblistes (pdf). Nous mettons par ailleurs en évidence une divergence des deux modèles en ce qui concerne le temps de réponse du sous-bassin versant. Celui modélisé avec le modèle GRP semble plus court que celui modélisé avec le modèle MORDOR-TS.

Cette étude permet de valider la méthodologie utilisée pour générer un ensemble hydrologique brut (choix des variables incertaines, densités de probabilité et hyper-paramètres

[3] *Vers une approche ensembliste de la prévision des crues* disponible sur https://hal.archives-ouvertes.fr/hal-03250002.

associés). Cela permettra de créer dans un second temps des prévisions d'ensemble hydrologiques brutes qui seront calibrées par la méthode *Quantile Regression Forests* (voir paragraphe 4.2).

Remarques suite aux commentaires de la première soumission Un relecteur s'interroge sur le choix d'avoir réalisé des tirages dans des lois uniformes après avoir déterminé les valeurs minimales et maximales pour les paramètres. En effet, il est possible que la distribution ne puisse pas être considérée comme uniforme. Cependant, nous n'avons aucun argument pour justifier le choix d'une distribution plutôt qu'une autre. Pour pouvoir faire cela, il faudrait mener une étude plus en amont sur le processus physique que le paramètre est censé représenter. En l'absence d'autre information, nous avons préféré imposer des lois uniformes pour ne pas ajouter un degré de liberté supplémentaire qu'on ne saurait pas contraindre.

Il est également proposé de tirer des vecteurs de paramètres optimisés plutôt que des paramètres indépendants. En effet, notre méthodologie ne tient pas compte des corrélations entre les paramètres induites par la procédure de calage. On pourrait alors utiliser l'ensemble issu du front de Pareto déterminé au moment du calage dans le cas du modèle MORDOR-TS. Nous n'avons pas privilégié cette méthodologie. Tout d'abord, sous cette hypothèse, du fait de la dépendance entre les entrées, il n'est pas possible de mener l'étude de sensibilité globale avec le calcul des indices de Sobol. Il faudrait alors s'orienter vers une méthode pour laquelle les paramètres peuvent ne pas être choisis de manière indépendante, par exemple la méthode des moments décrite en annexe N.1.3. Par ailleurs, la construction d'un ensemble brut avec les vecteurs de paramètres issus du front de Pareto a été testée sur la période de disponibilité des PEAROME, en considérant les pluies et les paramètres comme variables incertaines. Nous avons alors obtenu des ensembles très sous-dispersifs, avec des histogrammes de rang en forme de U. Ainsi, cette méthodologie ne suffit pas à créer un ensemble brut calibré. Dans tous les cas, une étape de post-traitement, telle que décrite au paragraphe 4.2, est nécessaire.

4.2 Article accepté par *Journal of Hydrology*

Cet article est intitulé *Strategies for hydrologic ensemble generation and calibration: on the merits of using model-based predictors*. L'étude présentée dans cet article a pour objectifs la génération de prévisions d'ensemble hydrologiques brutes puis leur calibration avec la méthode *Quantile Regression Forests*. Cette méthode est traditionnellement employée pour la calibration statistique d'ensembles de prévision météorologique [Taillardat et al., 2016, 2017, Zamo et al., 2014].

La première partie de l'article est dédiée à l'analyse et à la hiérarchisation des sources

d'incertitude de la modélisation avec MORDOR-TS grâce à une étude de sensibilité globale. Elle permet d'identifier les sources d'incertitude les plus significatives à prendre en compte pour la génération de l'ensemble de prévision brut. La méthodologie utilisée est celle présentée dans l'article *La Houille Blanche* (voir paragraphe précédent). Une étude de sensibilité globale comparée entre des configurations tenant compte ou non de l'incertitude sur les pluies est également menée pour valider la méthodologie de construction des ensembles bruts forcés avec une prévision déterministe. L'ensemble hydrologique brut sera en effet construit sans prendre en compte les prévisions d'ensemble PEAROME faute de leur disponibilité sur une période suffisante. Ces études de sensibilité ont mis en évidence que les deux paramètres de MORDOR-TS k_{min} et $Diff$ n'ont pas d'influence significative sur les débits prévus aux trois stations.

La seconde partie de l'article est dédiée à la génération de prévisions d'ensemble hydrologiques brutes et à leur calibration statistique avec la méthode QRF. Les calibrations statistiques sont menées pour chacun des trois sous-bassins et pour chaque échéance de prévision de +1 heure à +24 heures. Plusieurs stratégies sont investiguées et résumées par la figure 4.1 également présentée dans l'article :

- L'ensemble brut *raw-HEF* est généré par perturbation des 8 paramètres de MORDOR-TS significatifs pour les débits simulés à l'exutoire de chaque bassin. Les paramètres k_{min} et $Diff$, non significatifs, sont fixés à leur valeur de calage. Cet ensemble est composé de 99 membres. Le tableau 4.1 présente les plages de variation des paramètres retenues pour la génération de l'ensemble brut. Le modèle est alimenté par les prévisions déterministes *RR3* afin de couvrir une période temporelle plus importante que celle de la disponibilité des PEAROME. Cet ensemble est issu d'une approche que l'on peut qualifier de *modèle uniquement*.

Paramètre	Tréodet min	Tréodet max	Kerjean min	Kerjean max	Ty-Planche min	Ty-Planche max
c_p (-)	0.98	1.21	0.97	1.13	0.99	1.18
U_{max} (mm)	75	196	30	111	39	116
L_{max} (mm)	64	153	30	493	74	298
ev_L (-)	3.95	4.00	1.50	3.98	3.38	4.00
Z_max (mm)	47	138	77	256	72	482
k_r (-)	0.10	0.28	0.10	0.41	0.10	0.30
lk_N (mm.h^{-1})	-5.8	-5.2	-6.2	-5.5	-5.7	-5.4
Cel (km.h^{-1})	0.44	0.56	0.45	1.10	0.50	0.55

Tableau 4.1: Plage de variation des paramètres pour la génération de l'ensemble hydraulique brut *raw-HEF* avec le modèle hydrologique MORDOR-TS.

- L'ensemble *QRF-hydro* est obtenu par calibration statistique QRF de l'ensemble brut *raw-HEF* : parmi les prédicteurs sont utilisés des prédicteurs issus de l'ensemble brut *raw-HEF* (moyenne et écart-type). Cet ensemble est issu d'une approche qualifiée de *modèle et données*. 99 quantiles calibrés sont extraits et analysés.

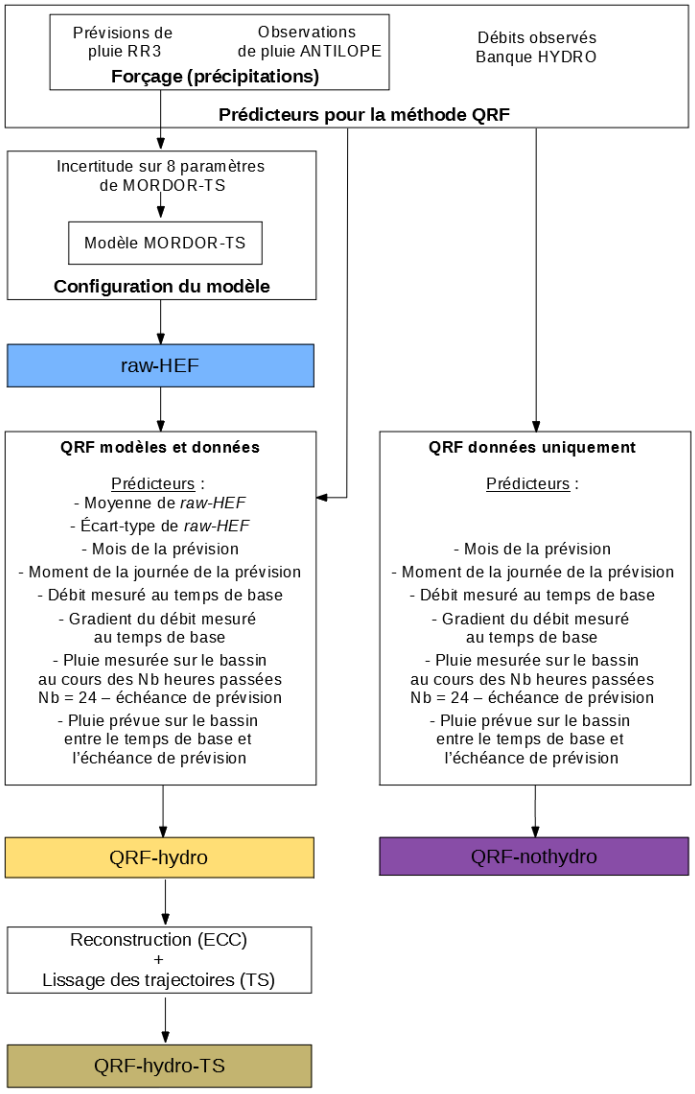


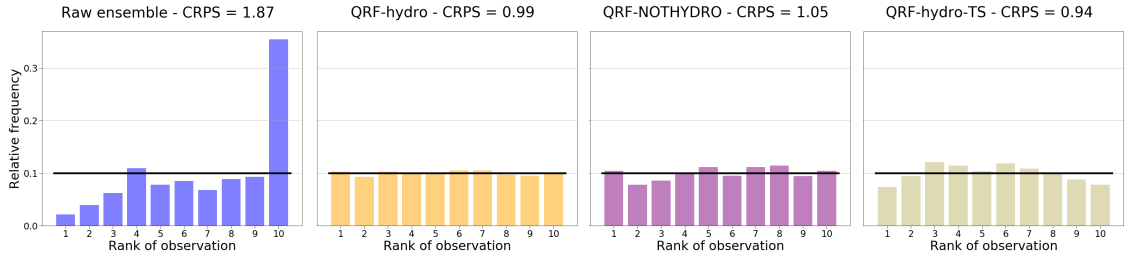
Figure 4.1: Stratégies mises en place pour la génération et calibration des prévisions d'ensemble hydrologiques. Les données utilisées pour les simulations du modèle hydrologique MORDOR-TS ainsi que les prédicteurs utilisés sont présentées. Les quatre ensembles résultants sont dans les boites de couleur.

• L'ensemble *QRF-nothydro* est obtenu par application de la méthode QRF, mais sans utiliser de prédicteur issu de l'ensemble brut *raw-HEF*. Les prédicteurs sont des mesures disponibles avant le temps de base, le moment de la journée... Cet ensemble est alors issu d'une approche appelée *données uniquement*. 99 quantiles calibrés sont extraits et analysés.

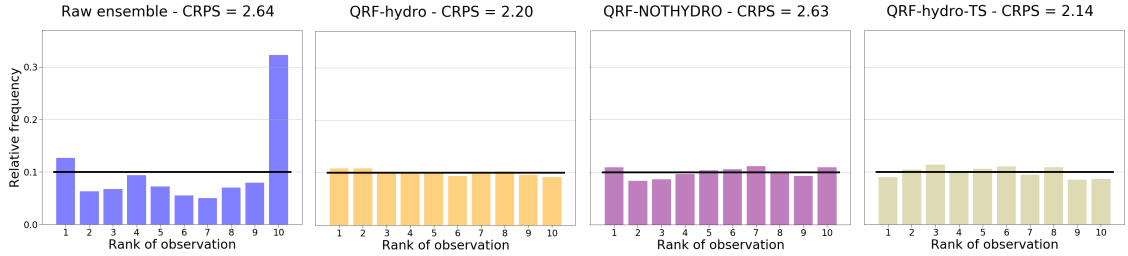
• Dans l'objectif d'utiliser les prévisions d'ensemble hydrologiques comme forçage du modèle hydraulique MASCARET, une étape de reconstruction est appliquée à l'ensemble *QRF-hydro* : reconstruction temporelle par la méthode *Ensemble Copula Coupling* et

lissage des trajectoires appelée *Trajectory Smoothing* (TS) (voir paragraphe 3.3.3). On obtient alors l'ensemble appelé *QRF-hydro-TS*. Il est composé de 99 membres reconstruits.

Quelques graphiques illustrant des résultats présentés dans l'article Nous présentons dans ce résumé des graphiques correspondant à quelques uns des résultats présentés dans l'article. La figure 4.2 présente les histogrammes de rang des quatre ensembles *raw-HEF*, *QRF-hydro*, *QRF-nothydro* et *QRF-hydro-TS* à la station Tréodet pour les échéances de prévision 6 heures et 21 heures. La tableau 4.2 présente le résultat du calcul de la moyenne des normes des histogrammes de rang pour toutes les échéances de prévision pour chacune des stations. La figure 4.3 présente les valeurs des CRPS obtenus pour les quatre ensembles aux trois stations Tréodet, Kerjean et Ty-Planche pour les échéances de prévision allant de +1 heure à +24 heures.



(a) Echéance 6 heures



(b) Echéance 21 heures

Figure 4.2: Histogrammes de rang et CRPS associé pour les ensembles hydrologiques *raw-HEF*, *QRF-hydro*, *QRF-nothydro* et *QRF-hydro-TS* aux échéances de prévision 6 heures et 21 heures à Tréodet.

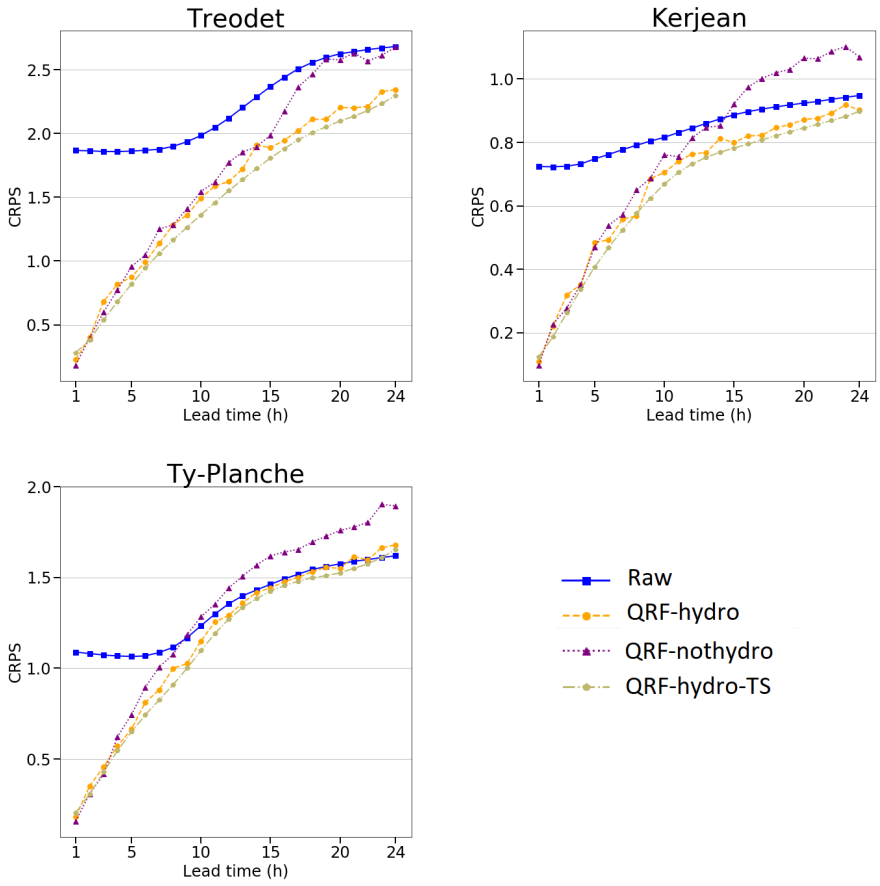


Figure 4.3: CRPS selon les échéances de prévision pour les ensembles *raw-HEF*, *QRF-hydro*, *QRF-nothydro* et *QRF-hydro-TS* aux trois stations Tréodet, Kerjean et Ty-Planche.

Catchment	$10^{-2}\|.\|_2$ *raw-HEF*	$10^{-2}\|.\|_2$ *QRF-hydro*	$10^{-2}\|.\|_2$ *QRF-nothydro*	$10^{-2}\|.\|_2$ *QRF-hydro-TS*
Tréodet	25.8	1.8	3.3	4.5
Kerjean	15.8	1.8	2.8	4.0
Ty-Planche	9.3	1.9	3.6	3.1

Tableau 4.2: Moyenne temporelle des normes $\|.\|_2$ des histogrammes de rang pour les échéances de prévision allant de +1 heure à +24 heures pour les 4 ensembles *raw-HEF*, *QRF-hydro*, *QRF-nothydro* et *QRF-hydro-TS* aux trois stations Tréodet, Kerjean et Ty-Planche.

Synthèse de l'ensemble des résultats présentés dans l'article Tout d'abord, il est intéressant de noter que les résultats obtenus pour des échéances de prévision inférieures au temps de transfert des bassins sont ceux que l'on obtiendrait en forçant le modèle MORDOR-TS avec des prévisions d'ensemble de pluie PEAROME. Nous avons en effet mis

en évidence que l'indice de Sobol de la pluie est négligeable pour le sous-bassin de l'Odet (exutoire Tréodet) pour des échéances de prévision inférieures à 6 heures. Par ailleurs, nous mettons en évidence que les prévisions d'ensemble brutes *raw-HEF* sont sous-dispersives et sous-estiment les débits observés. Cela montre qu'une approche *modèle uniquement* ne permet pas d'obtenir des prévisions d'ensemble bien calibrées. Les avantages de la calibration QRF sont évalués par l'étude de la fiabilité et de la résolution des ensembles par échéance de prévision, des ensembles des prévisions cumulées sur 24 heures et des ensembles du maximum de la prévision sur toutes les échéances. Ces deux derniers ensembles se rapportent respectivement aux notions de volume de crue et de pic de crue [Hemri et al., 2015]. Nous mettons en évidence que sur le bassin de l'Odet, la méthode QRF permet d'obtenir des ensembles avec une meilleure fiabilité et des meilleurs CRPS sur les 3 sous-bassins. Par ailleurs, les meilleures performances sont obtenues lorsque les prédicteurs issus de l'ensemble brut sont utilisés. L'approche *modèles et données* qui combine les résultats du modèle hydrologique avec les mesures disponibles avant le temps de base apparaît comme la plus pertinente. L'application de l'étape de lissage des trajectoires impacte légèrement les performances de l'ensemble *QRF-hydro* : l'ensemble *QRF-hydro-TS* présente une fiabilité légèrement moins bonne, mais des CRPS un peu meilleurs.

Perspectives Plusieurs perspectives peuvent être envisagées après la réalisation de cette étude. Tout d'abord, il serait pertinent d'appliquer cette méthodologie en utilisant une prévision d'ensemble météorologique comme forçage du modèle MORDOR-TS pour la génération de l'ensemble brut *raw-HEF*. Les incertitudes de la chaîne de prévision hydrométéorologique seraient alors plus largement couvertes pour les échéances de prévision supérieures au temps de transfert du bassin versant. Dans tous les cas, un post-traitement en sortie de modèle hydrologique restera *a priori* indispensable [Zalachori, 2013]. Il est également nécessaire de développer cette méthodologie sur d'autres bassins versants afin d'apprécier notamment sa robustesse. Les prédicteurs utilisés pour la calibration pourront alors être adaptés. Par exemple, le prédicteur représentant la pluie mesurée ou prévue sur le bassin versant pourrait être moyenné sur des périodes plus courtes dans le cas de bassins versant soumis à des crues éclair (dans notre cas, nous avons choisi 24 heures au maximum). Le prédicteur représentant le mois de la prévision aurait alors potentiellement plus d'importance, puisqu'il permettrait de différencier les crues éclair se produisant essentiellement à l'automne des autres phénomènes. Pour finir, un travail considérant l'ensemble issu du front de Pareto pour créer les membres de l'ensemble hydrologique brut est une perspective également intéressante. Comme nous l'avons vu au paragraphe précédent, cette méthodologie n'est pas suffisante pour générer des ensembles ayant de bonnes propriétés. Il s'agirait alors de développer la même méthodologie de calibration QRF à partir de ce nouvel ensemble brut. Il serait pertinent de comparer les performances des nouveaux ensembles calibrés avec celles des ensembles construits dans cette étude.

Chapitre 5

Filtre de Kalman d'ensemble (EnKF) dans le modèle MORDOR-TS

Sommaire

5.1 Structure de l'assimilation de données mise en place 144

5.2 Réalisation des expériences jumelles 146

5.2.1 Description des expériences jumelles 147

5.2.1.1 Principes généraux 147

5.2.1.2 Expériences réalisées 148

5.2.2 Expériences de correction d'un paramètre : *corr1par* 151

5.2.3 Expériences de correction simultanée des 8 paramètres : *corr8par* 155

5.2.3.1 Choix de l'écart-type de l'erreur d'observation et du nombre de membres 156

5.2.3.2 Analyse de l'impact de l'assimilation de données sur des mailles à l'amont de Tréodet 163

5.2.3.3 Résultats à Kerjean et Ty-Planche 168

5.2.4 Synthèse 168

5.3 Réalisation des expériences réelles 170

5.3.1 Caractéristiques des expériences 170

5.3.2 Résultats dans la phase d'analyse 170

5.4 Synthèse 171

Ce chapitre présente la mise en place d'assimilation de données au sein du modèle hydrologique MORDOR-TS. Celle-ci est réalisée de manière cyclée en appliquant l'algorithme du filtre de Kalman d'ensemble. Le paragraphe 5.1 présente la structure de l'assimilation de données mise en place dans ce travail. Elle s'applique aussi bien à la réalisation des expériences jumelles et réelles que nous décrivons ci-après.

Tout d'abord, des expériences dites *jumelles* ou *synthétiques* sont réalisées. Pour ces expériences, des observations fictives sont fabriquées grâce à l'intégration du modèle avec un autre jeu de paramètres que celui calé, et avec des forçages perturbés par rapport aux mesures. Ces observations sont alors considérées *vraies*. L'assimilation de données va chercher à s'en approcher en partant d'un autre jeu, qui dans notre cas est le modèle calé et les forçages réels. La réalisation d'expériences jumelles va permettre de qualifier la performance de l'algorithme d'assimilation de données sur le modèle. En revanche, la structure de l'erreur modèle se retrouve dans celle des observations puisqu'elles sont issues d'une modélisation avec le même modèle. On ne tient donc pas compte des erreurs de représentativité, c'est-à-dire des processus et de leur résolution non pris en compte par le modèle. Les résultats des expériences jumelles faites avec le modèle MORDOR-TS sont présentés au paragraphe 5.2.

Une fois que l'algorithme d'assimilation de données est validé grâce aux expériences jumelles, il est alors possible de l'utiliser en conditions réelles. Cela représente l'objectif final de notre travail. On assimile alors des mesures temps réel au sein du modèle calé. Cette expérience ainsi que les résultats sur la phase d'analyse sont décrits au paragraphe 5.3. En revanche, les résultats en phase de prévision seront analysés dans le chapitre 6. Ce chapitre aura en effet pour objet la comparaison des performances des ensembles obtenus par calibration statistique QRF et par assimilation de données ensembliste.

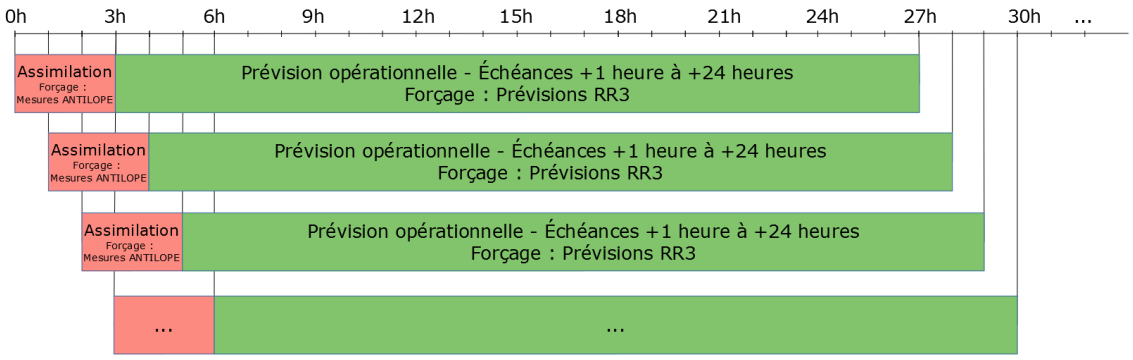
5.1 Structure de l'assimilation de données mise en place

Vecteur de contrôle L'assimilation de données mise en place pour la prévision hydrologique a pour objectif uniquement la correction des paramètres du modèle MORDOR-TS. On ne cherche à corriger ni l'état du modèle, ni ses forçages. Ce choix a été fait pour les raisons suivantes. Tout d'abord, contrairement à la correction des paramètres, les bénéfices de la correction de l'état s'estompent avec les échéances. Il n'y a pas de persistance de la correction. L'objectif étant de réaliser des prévisions opérationnelles jusqu'à 24 heures d'échéance, la correction des paramètres du modèle est privilégiée. Par ailleurs, nous ne corrigeons pas les forçages car dans un contexte opérationnel, les SPC[1] auront à terme à disposition des prévisions ensemblistes post-traitées. Par ailleurs, la correction des précipitations par l'algorithme EnKF supposerait que l'on puisse corriger l'intensité

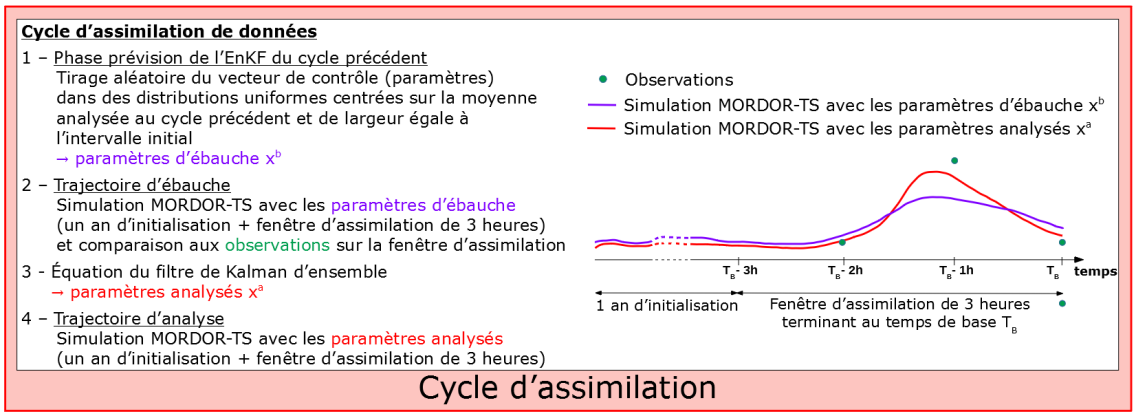
[1] Services de Prévision des Crues.

et la spatialisation des pluies avec une seule donnée de débit, ce qui n'est pas pertinent.

Cyclage La figure 5.1 présente l'assimilation de données mise en place pour le modèle hydrologique MORDOR-TS. La figure 5.1a présente le principe du cyclage des analyses et des prévisions opérationnelles. Cette figure a déjà été présentée au paragraphe 3.4 (figure 3.16). Elle est adaptée ici au cas du modèle MORDOR-TS. Le modèle est alimenté par les pluies mesurées ANTILOPE sur la fenêtre d'assimilation et par les prévisions de pluie RR3 sur la phase de prévision opérationnelle. La fin de la fenêtre d'assimilation correspond au temps de base T_B, c'est le moment où une prévision opérationnelle peut être lancée. Nous rappelons que nous avons fait le choix de cycler les fenêtres d'assimilation de manière horaire et que la fenêtre d'assimilation dure 3 heures. La fenêtre de prévision opérationnelle dure 24 heures. Le passage d'un cycle d'assimilation au suivant est réalisé à partir de la phase *Prévision* de l'EnKF (voir paragraphe 3.4.3.3).



(a) Cycles d'analyses et de prévisions mis en place pour l'assimilation de données dans MORDOR-TS.



(b) Description détaillée d'un cycle d'assimilation de données. À la fin de cette phase, l'état du modèle est en équilibre avec la valeur des paramètres analysés. Une prévision opérationnelle peut être lancée.

Figure 5.1: Description de la structure de l'assimilation de données mise en place pour MORDOR-TS. La fenêtre d'assimilation dure 3 heures. Les prévisions opérationnelles sont émises jusqu'à l'échéance 24 heures.

La figure 5.1b présente la réalisation d'un cycle d'assimilation de données en détail corre-

spondant aux phases *Prévision* et *Analyse* de l'EnKF. Pour plus de détail, le lecteur pourra se reporter au paragraphe 3.4.6 qui décrit l'EnKF appliqué dans le cadre de ce travail. L'intervalle initial dont la largeur est utilisée à chaque cycle pour le tirage d'un nouveau vecteur de contrôle d'ébauche (étape 1 du schéma 5.1b) est issu de la plage de variation des paramètres déterminée lors de l'étude d'incertitude. Le tableau 5.1 rappelle la largeur des intervalles utilisés dans ce travail (voir chapitre 4, tableau 4.1 pour les bornes de l'intervalle du premier cycle d'assimilation). Il est à noter qu'ici, les trajectoires d'ébauche et d'analyse sont propagées par le modèle MORDOR-TS sur la fenêtre d'assimilation après une initialisation du modèle pendant 1 an (étapes 2 et 4). Cette durée est nécessaire pour que l'état du modèle MORDOR-TS soit en équilibre avec l'historique des précipitations et avec les paramètres imposés. Pour ces simulations, le modèle MORDOR-TS est alimenté par les pluies mesurées ANTILOPE. À la fin du cycle d'assimilation (après l'étape 4), on se trouve au temps de base et la prévision opérationnelle peut être lancée. Elle correspond à l'intégration déterministe des membres analysés sur la fenêtre de prévision opérationnelle. Ils sont alors alimentés par les prévisions de pluie RR3. Les paramètres de MORDOR-TS analysés sur la fenêtre d'assimilation sont persistés sur la fenêtre de prévision opérationnelle.

Bassin versant	c_p	U_{max}	L_{max}	Z_{max}	ev_L	k_r	lk_N	Cel
Tréodet	0.23	121	89	91	0.1	0.2	0.5	0.11
Kerjean	0.16	81	463	179	2.5	0.3	0.7	0.65
Ty-Planche	0.18	77	224	410	0.6	0.2	0.3	0.05

Tableau 5.1: Largeur des plages de variation des paramètres de MORDOR-TS.

Nombre de membres La taille de l'ensemble est fixée à 49 membres. Lors de la réalisation des expériences jumelles, un test de sensibilité des résultats avec un ensemble plus grand à 279 membres est réalisé (voir paragraphe 5.2.3.1).

5.2 Réalisation des expériences jumelles

Ce paragraphe présente les expériences jumelles réalisées pour évaluer l'algorithme d'assimilation de données appliqué au modèle MORDOR-TS sur les trois sous-bassins de l'Odet dont les exutoires sont Tréodet, Kerjean et Ty-Planche. La partie 5.2.1 présente le principe et l'implémentation des expériences jumelles. Ensuite, deux types d'expériences sont réalisées : les expériences *corr1par* dans lesquelles un seul paramètre du modèle est corrigé sont présentées en partie 5.2.2, et les expériences *corr8par* dans laquelle les 8 paramètres de MORDOR-TS sont corrigés simultanément sont présentées en partie 5.2.3. La synthèse des résultats des expériences jumelles est présentée en paragraphe 5.2.4.

5.2.1 Description des expériences jumelles

5.2.1.1 Principes généraux

Les expériences jumelles mises en place sont étudiées sur la fenêtre d'assimilation et sur la fenêtre de prévision opérationnelle. Nous décrivons ci-après les simulations nécessaires à leur mise en place.

Fenêtre d'assimilation La réalisation des expériences jumelles nécessite plusieurs intégrations du modèle sur la période de travail notée P_{jum}, comme l'illustre la figure 5.2 :

- La simulation de contrôle[2] correspond à une simulation continue sur la période P_{jum}. Les paramètres du modèle MORDOR-TS sont les valeurs calées. Cette simulation représente la connaissance *a priori* du système issue du calage, pas forcément représentative de la réalité. C'est à partir de cette connaissance *a priori* que l'on va chercher à s'approcher de la *réalité*. Les paramètres de la simulation de contrôle sont appelés les paramètres de contrôle.

- La simulation de référence est une simulation continue sur la période P_{jum}. Les paramètres du modèle MORDOR-TS sont différents des valeurs de calage. On construit à partir de cette simulation des observations fictives, dites *observations vraies* dans le cadre des expériences jumelles, en perturbant les résultats de cette simulation pour tenir compte des erreurs d'observation. Les paramètres de la simulation de référence sont appelés les paramètres de référence.

- L'assimilation de données correspond à un ensemble de simulations cyclées du modèle décrites au paragraphe 5.1. À chaque cycle d'assimilation de données, la trajectoire d'analyse est récupérée. La comparaison avec les valeurs *vraies* sera un des moyens d'évaluation de l'algorithme d'assimilation de données.

Il est à noter que le code couleur employé sur la figure 5.2 est cohérent avec les couleurs utilisées par la suite pour la présentation des résultats d'assimilation de données : la simulation de contrôle ou les paramètres de cette simulation sont en bleu, la simulation de référence ou les paramètres de cette simulation sont en noir, les observations issues de cette simulation sont en vert, et les simulations après la réalisation de l'analyse des paramètres est en rouge.

Prévision opérationnelle Les expériences jumelles servent également à étudier la pertinence de l'assimilation de données sur la phase de prévision opérationnelle. Pour cela, les trois simulations (de référence, de contrôle et d'assimilation) sont alors réalisées de manière cyclée sur les fenêtres d'assimilation et de prévision opérationnelle. La figure 5.1a présente

[2] Attention à ne pas confondre la simulation de contrôle avec le vecteur de contrôle, qui correspond aux paramètres de MORDOR-TS qui sont corrigés par l'assimilation de données.

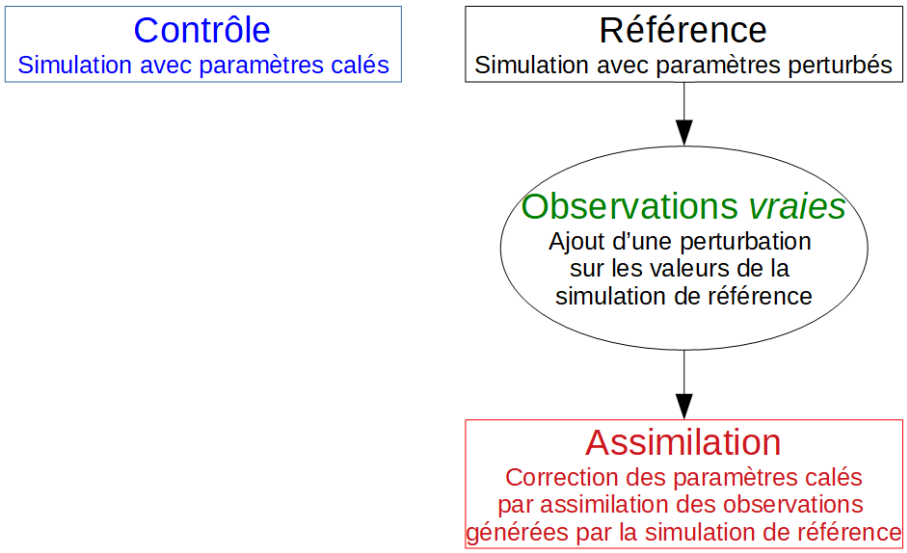


Figure 5.2: Principe de réalisation des expériences jumelles.

ce principe pour la simulation d'assimilation de données. Sur la fenêtre d'assimilation, les trois simulations sont alimentées par les pluies observées ANTIPOLE. Sur la fenêtre de prévision, les trois simulations sont alimentées par les prévisions RR3.

5.2.1.2 Expériences réalisées

Les expériences jumelles sont réalisées sur la période du 15 décembre 2013 au 30 décembre 2013 qui contient un événement de crue important (voir paragraphe 2.1.1, tempête Dirk).

Création des observations *vraies* Pour chacun des trois sous-bassins, la valeur choisie pour la simulation de référence pour chaque paramètre de MORDOR-TS correspond à sa valeur calée à laquelle on ajoute 1.2 fois la largeur de la plage de variation du paramètre retenue pour l'étude de sensibilité (voir tableau 5.1). Si la valeur ainsi obtenue est en-dehors des bornes valides des paramètres définies pour le modèle MORDOR-TS, cette valeur est réduite. Ceci est le cas pour L_{max} à Kerjean, Z_{max} à Ty-Planche et ev_L à Kerjean. Le tableau 5.2 présente les valeurs utilisées pour générer les simulations de contrôle et de référence. Il est à noter que la valeur des paramètres pour la simulation de contrôle correspond aux valeurs de calage sur la période 2007-2017. Cela explique les différences avec les valeurs des paramètres calés sur 2007-2013 indiquées au chapitre 2 dans le tableau 2.6. Ce point n'a aucune influence sur les résultats des expériences jumelles présentés dans cette section.

Pour la génération de la chronique d'observation, une perturbation est appliquée sur les résultats de la simulation de référence pour prendre en compte une erreur de mesure des débits observés. Elle est ici fixée à 5% de la valeur simulée. Des tests sur l'impact de la

Bassin versant	Simulation	c_p	U_{max}	L_{max}	Z_{max}	ev_L	k_r	lk_N	Cel
Tréodet	Contrôle	1.11	113	123	122	4.00	0.20	-5.49	0.53
	Référence	1.39	257	230	232	3.93	0.42	-4.83	0.67
Kerjean	Contrôle	1.07	70	30	165	1.50	0.10	-6.05	0.54
	Référence	1.26	168	430	380	3.50	0.47	-5.19	1.32
Ty-Planche	Contrôle	1.02	84	153	111	4.00	0.25	-5.34	0.54
	Référence	1.24	176	422	411	3.24	0.49	-4.93	0.60

Tableau 5.2: Valeur des paramètres pour les simulations de contrôle et de référence pour les expériences jumelles.

valeur de l'écart-type de l'erreur seront présentés en partie 5.2.3.1. Cette valeur de perturbation de 5% est également retenue pour la perturbation stochastique des observations dans l'algorithme EnKF (voir paragraphe 3.4.3.3). Les observations sont générées au pas de temps horaire. On assimilera donc 3 observations sur chaque fenêtre d'assimilation de 3 heures.

Travail à l'échelle des sous-bassins Les expériences jumelles d'assimilation de données sont menées sur un seul sous-bassin à la fois : sous-bassins de l'Odet (exutoire Tréodet), du Jet (exutoire Kerjean) ou du Steïr (exutoire Ty-Planche). Comme nous l'avons mentionné en introduction, deux types d'expériences sont réalisées :

- *corr1par* : dans ces expériences, le vecteur de contrôle est constitué d'un seul paramètre du modèle MORDOR-TS. Ce paramètre est modifié pour la réalisation de la simulation de référence selon les valeurs indiquées dans le tableau 5.2. L'assimilation de données va chercher à corriger uniquement ce paramètre. Une expérience est donc réalisée pour chacun des 8 paramètres de MORDOR-TS et pour chacun des 3 sous-bassins, ce qui fait 24 expériences différentes.

- *corr8par* : dans ces expériences, le vecteur de contrôle est constitué des 8 paramètres du modèle MORDOR-TS. Ils sont simultanément modifiés pour la réalisation de la simulation de référence selon les valeurs indiquées dans le tableau 5.2. L'assimilation de données va chercher à corriger simultanément les 8 paramètres. Une expérience est réalisée sur chacun des 3 sous-bassins, ce qui fait 3 expériences différentes.

Dans tous les cas, les paramètres de référence utilisés pour générer la simulation de référence sont appliqués à toutes les mailles du sous-bassin versant concerné. La correction apportée par le filtre de Kalman d'ensemble est la même pour toutes les mailles.

Application de l'algorithme EnKF et étude de ses performances Une fois les observations de référence générées, l'algorithme d'assimilation de données est lancé sur les 15 jours de l'événement de décembre 2013. L'ensemble comporte 49 membres. Un test de sensibilité des résultats avec un ensemble plus grand à 279 membres a été réalisé et présenté en partie 5.2.3.1. Nous rappelons que le vecteur de contrôle est constitué uniquement des

paramètres de MORDOR-TS. La performance de l'algorithme d'assimilation de données est analysé selon plusieurs critères :

- Étude temporelle des valeurs analysées pour les éléments du vecteur de contrôle (valeurs analysées pour tous les membres et moyenne de l'analyse) et comparaison avec la valeur imposée pour la simulation de référence.

- Calcul de l'erreur sur le débit simulé en fonction du temps à l'exutoire du sous-bassin traité pour la simulation de contrôle et la simulation avec assimilation de données par rapport à la simulation de référence. Pour la simulation avec assimilation de données, on calcule pour chaque pas de temps l'erreur ϵ_i entre la moyenne des membres et la simulation de référence.

- Visualisation des chroniques de débits des simulations de référence, de contrôle et d'analyse avec mise en évidence des observations réellement assimilées.

- Calcul de la RMSE intégrée sur l'événement calculée pour chaque maille du sous-bassin versant considéré. C'est un scalaire exprimé à chaque maille. En effet, le modèle MORDOR-TS étant spatialisé, il est possible d'obtenir le débit modélisé à chaque maille. Pour la simulation avec assimilation de données, pour chaque maille, on calcule d'abord pour chaque pas de temps l'erreur ϵ_i entre la moyenne des membres et la simulation de référence, puis on calcule la norme 2 de ces erreurs sur tous les pas de temps considérés : $\sqrt{\frac{1}{N_{pas}} \sum_{i=1}^{N_{pas}} \epsilon_i^2}$.

- Calcul de la RMSE des erreurs de prévision en fonction de l'échéance pour les simulations de contrôle et l'assimilation de données. C'est un scalaire exprimé pour chaque échéance. Pour une échéance donnée, on calcule d'abord pour chaque pas de temps l'erreur ϵ_i entre la moyenne des membres prévus et la simulation de référence, puis on calcule la norme 2 de ces erreurs sur tous les pas de temps considérés : $\sqrt{\frac{1}{N_{pas}} \sum_{i=1}^{N_{pas}} \epsilon_i^2}$.

Pour tous les graphiques présentés, les valeurs des simulations analysées sont prises en fin de fenêtre d'assimilation (temps de base).

Sur tous les graphiques présentés :

- la légende *asm* en rouge est relative à l'analyse ;
- la légende *ctl* en bleu est relative à la simulation de contrôle ;
- la légende *ref* en noir est relative à la simulation de référence.

Selon les variables représentées sur les graphiques, ces termes seront relatifs à l'espace de contrôle (les paramètres) ou à l'espace des observations (les valeurs de débit à l'exutoire). Sur les graphiques relatifs aux résultats dans la phase de prévision opérationnelle, deux nouveaux éléments de légende apparaissent :

- la légende *fcst* (pour *forecast*) en vert est relative à la prévision opérationnelle effectuée suite à la simulation de la trajectoire d'analyse ;

- la légende *bkg* en rose est relative à la simulation sans assimilation, c'est le *background*.

Pour l'étude des prévisions pseudo-opérationnelles réalisées après le cycle d'assimilation, on représentera également les histogrammes de rang (voir paragraphe 3.1.2.2). Dans le cadre de la réalisation des expériences jumelles, ceux-ci sont calculés sans stratification. L'observation prise en compte est construite de la même manière que pour la génération des *observations vraies* : la même perturbation est appliquée aux résultats de la simulation de référence.

5.2.2 Expériences de correction d'un paramètre : *corr1par*

Les expériences jumelles *corr1par* correspondent à la modification et la correction d'un seul paramètre de MORDOR-TS sur un seul sous-bassin versant. Le vecteur de contrôle est donc constitué d'un seul paramètre de MORDOR-TS. Nous rappelons qu'il y a 24 expériences *corr1par* différentes (3 sous-bassins * 8 paramètres). Dans ces expériences, l'ensemble est composé de 49 membres, et l'écart-type des erreurs d'observation est fixé à 5% de la valeur du débit de la référence. Nous présentons ici quelques résultats des expériences *corr1par*. L'intégralité des résultats de ces expériences est disponible à l'annexe O.1 pour Tréodet, O.2 pour Kerjean et O.3 pour Ty-Planche.

Résultats dans la phase d'analyse Nous présentons ici les résultats de deux expériences *corr1par* menées à Tréodet, dans le cas où le vecteur de contrôle est constitué du paramètre c_p ou du paramètre ev_L. La figure 5.3 présente pour ces deux expériences les valeurs analysées du vecteur de contrôle au cours de l'événement : les courbes orange représentent les valeurs pour chacun des membres, et la courbe rouge la valeur analysée moyenne. La figure 5.4 présente pour ces deux expériences l'erreur au cours du temps de la valeur du débit à Tréodet issu de la simulation de contrôle et de l'analyse par rapport à la simulation de référence. Pour finir, la figure 5.5 présente les chroniques de la simulation de contrôle, de la simulation de référence et de l'analyse pour la journée du 24 décembre 2013 pendant laquelle a eu lieu le pic d'événement. Sur cette figure sont représentées en vert les valeurs réellement assimilées, avec la prise en compte des erreurs d'observations ayant pour écart-type de 5% de la valeur du débit.

Les graphes de la figure 5.3 mettent en évidence une bonne capacité du modèle d'assimilation de données à retrouver la valeur du paramètre imposé en référence à Tréodet. Pour tous les paramètres sauf ev_L, la référence est atteinte dès les premiers cycles d'analyse. Pour ev_L, l'assimilation de données met environ une semaine à atteindre la valeur de référence. Nous notons qu'avec la gamme de variation retenue (voir tableau 5.1), ce paramètre a montré très peu d'influence dans l'étude d'incertitude à Tréodet. Ceci est également confirmé sur la figure 5.5 : nous notons très peu d'écart entre les simulations de contrôle de référence, et l'ensemble analysé est très peu dispersif. Les observations assimilées ont une variabilité

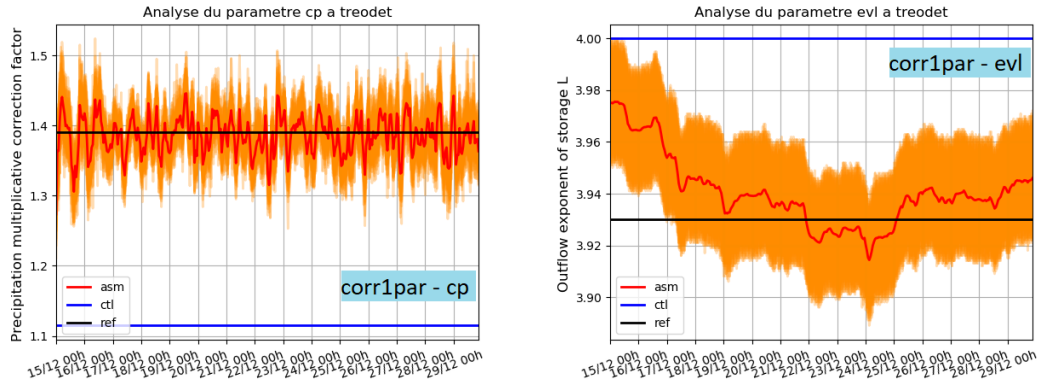


Figure 5.3: Valeurs analysées pour les paramètres c_p et ev_L en fonction du temps pour les expériences *corr1par* à Tréodet - Décembre 2013. Les courbes orange représentent les valeurs pour chacun des 49 membres et la courbe rouge la valeur analysée moyenne.

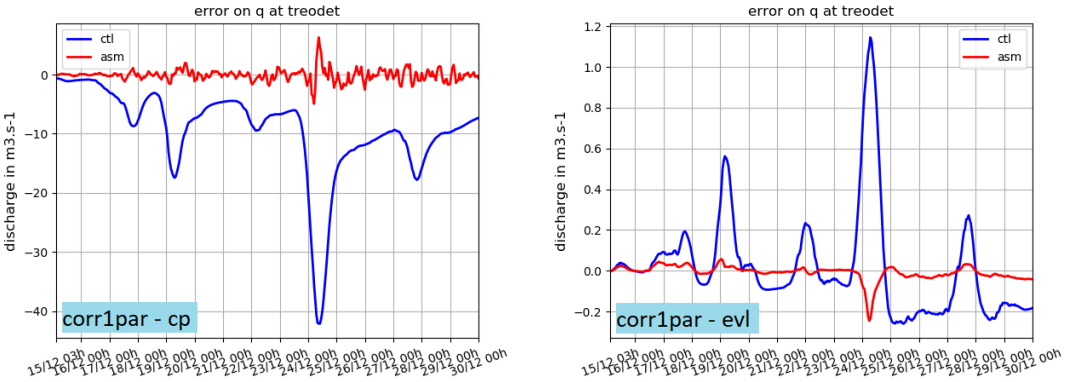


Figure 5.4: Chroniques d'erreurs de la simulation de contrôle et de l'analyse par rapport à la simulation de référence pour deux expériences *corr1par* à Tréodet (correction de c_p et ev_L) - Décembre 2013. La valeur de l'analyse est prise à la fin de la fenêtre d'assimilation.

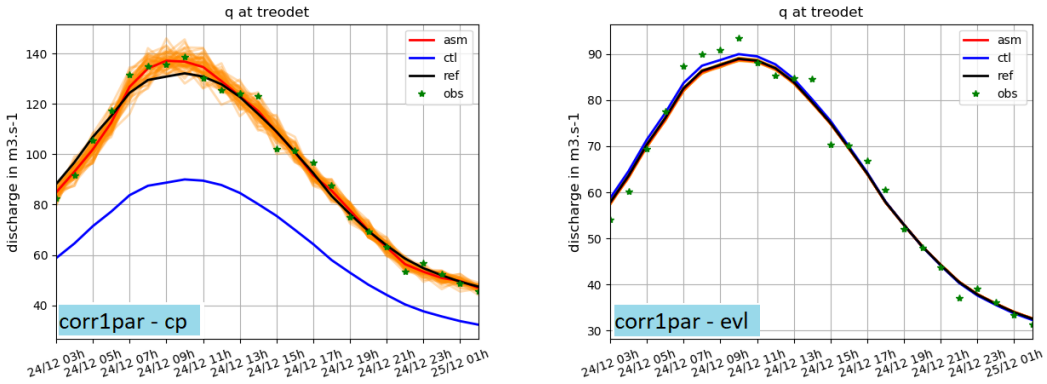


Figure 5.5: Simulation de contrôle, simulation de référence, analyses et observations assimilées au moment du pic d'événement le 24 décembre 2013 pour deux expériences *corr1par* à Tréodet (correction de c_p et ev_L).

significativement supérieure à la différence entre les simulations de référence et de contrôle. Ceci explique la dynamique assez lente observée pour la correction de ce paramètre. Nous notons que pour les stations Kerjean et Ty-Planche, la plage de variation de ce paramètre est significativement plus large et l'ensemble analysé possède une plus grande dispersion. L'algorithme d'assimilation de données réussit à corriger ce paramètre au bout de quelques cycles. Ceci confirme que la lente correction du paramètre ev_L est liée au fait que sa plage de variation est petite.

Les graphes de la figure 5.4 montrent une très bonne correction du débit en fin de fenêtre d'assimilation par l'algorithme d'assimilation de données pour toutes les expériences présentées. L'erreur de simulation est très nettement réduite après l'analyse.

Puisque les valeurs analysées des paramètres sont proches de celles imposées pour la simulation de référence, les chroniques assimilées présentées figure 5.5 sont naturellement proches de la simulation de référence. La chronique assimilée oscille autour de la chronique de référence, et les variations sont dues à l'erreur d'observation imposée, c'est-à-dire au fait que les observations assimilées (étoiles vertes) ne sont pas sur la courbe noire représentant la simulation de référence. Des tests ont été réalisés avec une erreur d'observation très faible ayant pour écart-type 10^{-5}. Dans ce cas, la chronique analysée se superpose parfaitement à la simulation de référence.

L'analyse des résultats des expériences *corr1par* sur les deux autres sous-bassin montre des résultats similaires (voir annexes O.2 et O.3). Chacune des expériences permet de retrouver une valeur analysée proche de la référence (figures O.6 et O.11). Les chroniques d'erreur aux deux stations sont très nettement réduites par l'assimilation de données (figures O.7 et O.12), ce qui se traduit par des chroniques assimilées proches de la référence (figures O.8 et O.13). Ainsi, la correction d'un seul paramètre dans le modèle MORDOR-TS par l'algorithme d'assimilation de données est de bonne qualité sur les trois sous-bassins étudiés et pour tous les paramètres.

Résultats dans la phase de prévision Nous cherchons à présent à analyser les performances de l'algorithme d'assimilation de données pour l'amélioration de la prévision. En prévision, toutes les simulations sont alimentées à partir du temps de base par les prévisions de pluie RR3.

Nous présentons ici les deux expériences *corr1par* consistant à corriger les paramètres c_p et k_r sur le sous-bassin de l'Odet (exutoire Tréodet). La figure 5.6 présente pour ces deux expériences les résultats du cycle d'assimilation pour lequel la fenêtre d'assimilation de 3 heures débute le 23 décembre 2013 à 20 heures. Sur chaque graphique sont représentés deux types d'ensembles : (i) l'ensemble d'ébauche sur les fenêtres d'assimilation et de prévision opérationnelle (*bkg* en rose), et (ii) l'ensemble analysé sur la fenêtre d'assimilation (*asm* en rouge) puis lancé en prévision à partir du temps de base (*fcst* en vert). Les

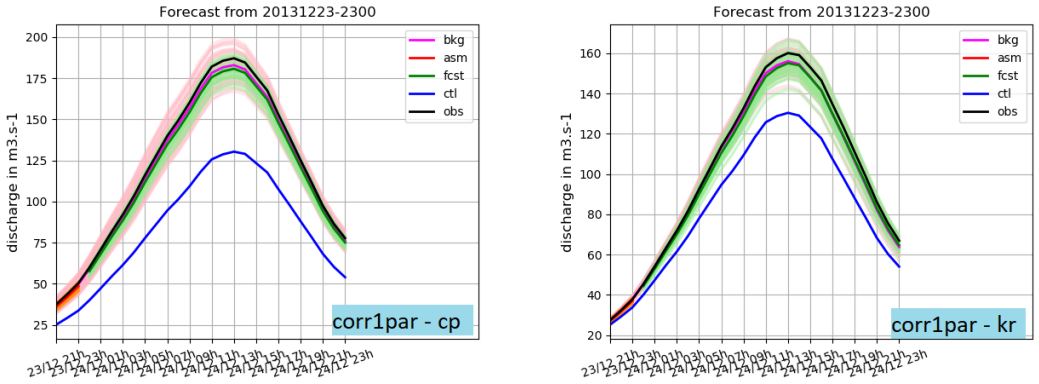


Figure 5.6: Visualisation pour 2 expériences *corr1par* à Tréodet (correction de c_p et k_r) du cycle d'assimilation de données débutant le 23 décembre 2013 à 20 heures : ensemble d'ébauche (bkg), ensemble analysé (asm) et ensemble en prévision suite à l'analyse (fcst), simulation de contrôle (ctl) et simulation de référence (ref). À partir du temps de base (23 décembre à 23 heures), toutes les simulations sont alimentées par les prévisions de pluie RR3.

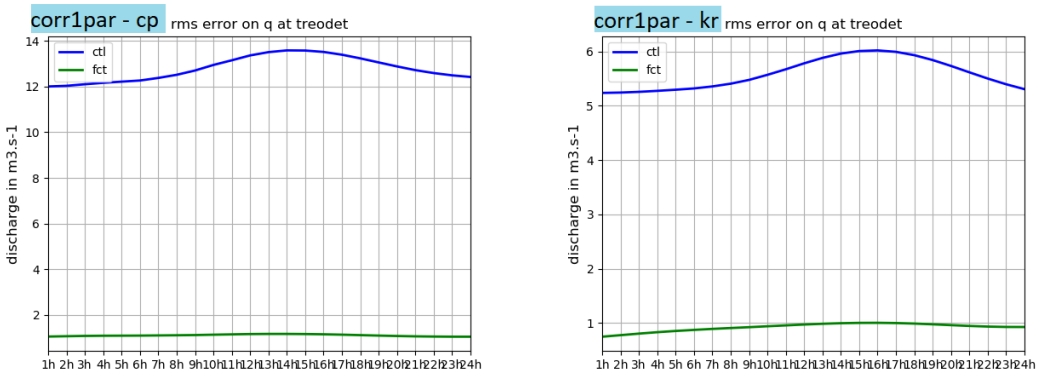


Figure 5.7: RMSE des erreurs de prévision en fonction de l'échéance pour la simulation de contrôle (*ctl* en bleu) et le modèle analysé (*fcst* en vert) par rapport à la simulation de référence pour 2 expériences *corr1par* à Tréodet (correction de c_p et k_r) - Décembre 2013. Toutes les simulations (de référence, de contrôle et le modèle corrigé) sont alimentées par les prévisions de pluie RR3 en prévision.

simulations de référence et de contrôle sont également tracées. Les résultats pour les 8 expériences à Tréodet sont présentées à la figure O.4 en annexe. Ces graphes mettent en évidence que les prévisions sont proches de la simulation de référence pour les 8 expériences menées. Nous notons que la dispersion des membres de l'ébauche n'est pas toujours réduite avec l'analyse, par exemple ici pour les paramètres k_r et Cel. Si l'on analyse ces graphes à d'autres moments caractéristiques de la crue, on s'aperçoit que pour le paramètre k_r, cette remarque n'est vraie que lorsque l'analyse a lieu en phase de montée (graphes non montrés ici). Cela correspond à une période pendant laquelle la valeur de l'indice de Sobol est faible. L'analyse n'est alors pas en mesure de réduire la dispersion de l'ébauche. Pour le paramètre Cel, cette tendance est générale quelle que soit la tendance de l'hydrogramme. Ces tendances se retrouvent également sur les stations Kerjean et Ty-Planche (voir figures O.9 et O.14 en annexe).

La figure 5.7 présente la RMSE par échéance de prévision pour la simulation de contrôle et pour la simulation issue du modèle analysé par rapport à la simulation de référence pour les deux expériences à Tréodet. Les figures O.5, O.10 et O.15 présentent les résultats pour les 24 expériences différentes menées. Ces graphes mettent en évidence le fait que la prévision est nettement améliorée par l'assimilation de données pour chaque paramètre pour les trois sous-bassins. Ceci est cohérent avec les résultats montrés précédemment, en particulier avec le fait que les valeurs analysées des paramètres sont proches de celles utilisées pour la simulation de référence pour chaque expérience. En effet, cela implique qu'au temps de base, le modèle analysé se trouve dans un état proche de l'état du modèle de référence qui a permis de générer les observations, et donc la prévision suit une dynamique similaire.

5.2.3 Expériences de correction simultanée des 8 paramètres : *corr8par*

Les expériences jumelles *corr1par* dans lesquelles un seul paramètre du modèle MORDOR-TS est perturbé ont montré que l'algorithme EnKF implémenté permettait de corriger de manière satisfaisante le modèle. Nous allons à présent nous intéresser à des expériences dans lesquelles les 8 paramètres de MORDOR-TS sont simultanément corrigés sur un sous-bassin. Il y a donc 3 expériences *corr8par* différentes, une sur chacun des trois sous-bassins. Les valeurs imposées pour les paramètres pour les simulations de contrôle et de référence sont les mêmes que pour les expériences *corr1par* (voir tableau 5.2).

Dans un premier temps, nous avons voulu appréhender l'impact de deux choix effectués pour la réalisation des expériences jumelles : (i) le choix de l'écart-type des erreurs utilisé pour la génération des observations à partir de la simulation de référence et également utilisé dans l'algorithme EnKF pour la perturbation des observations assimilées, et (ii) le nombre de membres de l'ensemble utilisés pour l'assimilation de données. Les résultats de ces tests préliminaires sont présentés en partie 5.2.3.1. Une fois fixés l'écart-type de l'erreur d'observation et le nombre de membres de l'expérience, plusieurs expériences ont

été réalisées sur le sous-bassin dont l'exutoire est Tréodet, en assimilant les débits uniquement à l'exutoire ou également des débits à l'exutoire des mailles à l'amont de Tréodet. Notons que cette dernière expérience n'est possible qu'en expériences jumelles puisqu'il n'existe pas de mesure de débit à l'amont de Tréodet. Les résultats de ces expériences sont présentés en partie 5.2.3.2. Pour finir, la partie 5.2.3.3 présente les résultats aux stations Kerjean et Ty-Planche.

5.2.3.1 Choix de l'écart-type de l'erreur d'observation et du nombre de membres

Trois expériences ont été menées dont les caractéristiques en termes de nombre de membres et d'écart-type des erreurs d'observation sont résumées dans la tableau 5.3. Cette partie présente les résultats comparés de ces trois expériences et permettra de choisir le nombre de membres et l'écart-type des erreurs d'observation pour les expérience réelles d'assimilation de données.

Nom	Nombre de membres	Ecart-type d'erreur d'observation
Expe 1	49	5.0% * observation
Expe 2	49	0.5% * observation
Expe 3	279	5.0% * observation

Tableau 5.3: Caractéristiques des trois expériences menées.

Résultats dans la phase d'analyse La figure 5.8 présente les chroniques des paramètres analysés pour les trois expériences pour les 8 paramètres de MORDOR-TS corrigés simultanément pour le sous-bassin dont l'exutoire est Tréodet. Ces graphiques montrent que la valeur de l'écart-type des erreurs d'observation peut influencer les valeurs des paramètres analysés. En effet, la comparaison des expériences *Expe 1* et *Expe 2*, qui ne diffèrent que par la valeur de l'écart-type de l'erreur d'observation, montre que pour certains paramètres comme c_p, U_{max} et Z_{max}, les chroniques des paramètres analysés sont quasiment identiques. En revanche, pour L_{max}, ev_L, k_r, lk_N et Cel, les chroniques montrent davantage de différences et l'expérience *Expe 2* aboutit à des valeurs analysées plus proches de la référence. Une erreur d'observation plus petite permet de mieux approcher les valeurs imposées en référence. La comparaison des graphiques pour les expériences *Expe 1* et *Expe 3*, qui ne diffèrent que par le nombre de membres de l'ensemble, ne met pas en évidence de différence notable lorsque l'on passe de 49 à 279 membres. Seule la variable lk_N présente une analyse légèrement différente avec les deux expériences : sa valeur tend à s'accroître dans l'expérience avec 49 membres, alors qu'elle n'évolue quasiment pas avec 279 membres. Nous pouvons donc conclure que l'augmentation du nombre de membres de 49 à 279 n'a pas d'impact notable sur les valeurs moyennes analysées des éléments du vecteur de contrôle.

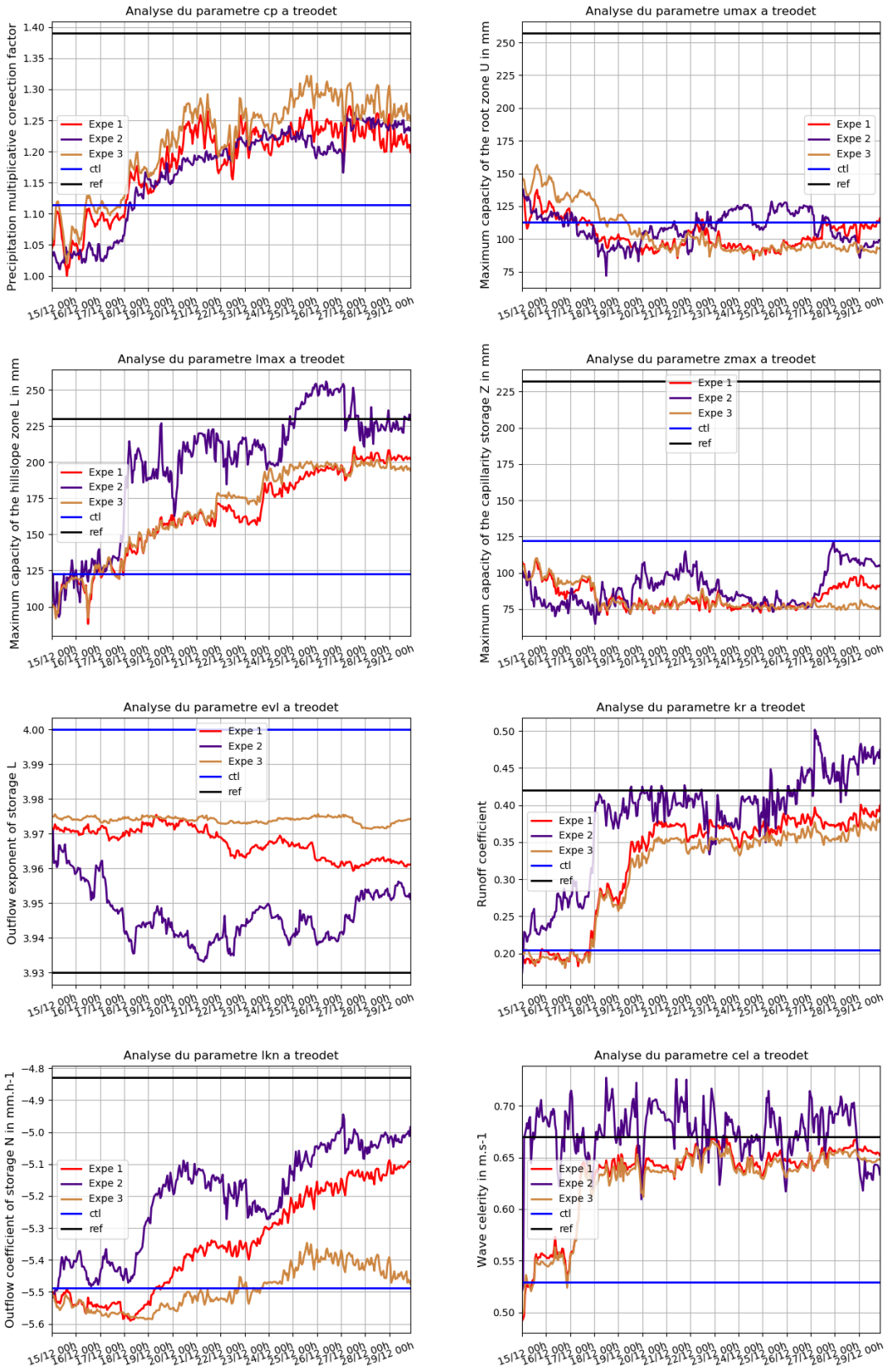


Figure 5.8: Valeurs analysées pour le vecteur de contrôle (paramètres) en fonction du temps pour les trois expériences *Expe 1*, *Expe 2* et *Expe 3* menées à Tréodet - Décembre 2013. Les expériences diffèrent par le nombre de membres et l'erreur considérée sur les observations. Leurs caractéristiques sont résumées dans le tableau 5.3.

Nous notons ici que contrairement aux expériences *corr1par*, les valeurs analysées dans le cadre de l'expérience *corr8par* convergent rarement vers la valeur imposée en référence. Les paramètres les mieux corrigés sont les paramètres L_{max}, k_r et Cel. Ce sont les paramètres pour lesquels les indices de Sobol sont les plus importants au cours d'un événement [Tiberi-Wadier et al., 2021].

La figure 5.9 montre pour les trois expériences les résultats du calcul de l'erreur sur le débit simulé avec la simulation de contrôle et avec l'assimilation de données par rapport à la simulation de référence à l'exutoire Tréodet. Nous notons que l'erreur sur le débit analysé à Tréodet est très nettement réduite par l'assimilation de données. Il possède une moins grande variabilité lorsque l'écart-type des erreurs est moindre (expérience *Expe 2*). Le nombre de membres utilisés dans l'algorithme d'assimilation de données n'a pas d'impact sur l'erreur du débit corrigé à Tréodet (expériences *Expe 1* et *Expe 3*).

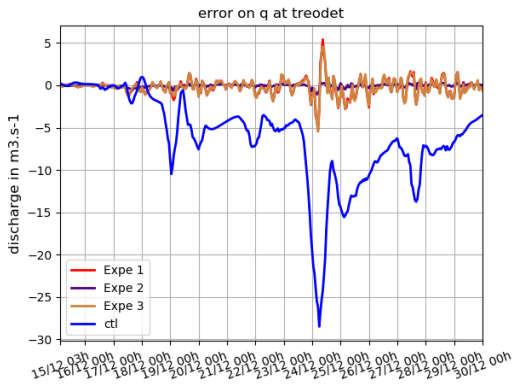


Figure 5.9: Chronique temporelle des erreurs à Tréodet pour la simulation de contrôle (bleu) et pour les analyses réalisées lors des trois expériences *Expe 1*, *Expe 2* et *Expe 3* par rapport à la simulation de référence - Décembre 2013.

Le code hydrologique MORDOR-TS étant spatialisé, un débit simulé est disponible à l'exutoire de chaque maille composant le bassin. Dans le cadre des expériences jumelles, il est alors facile et pertinent de calculer des erreurs sur les débits simulés à l'exutoire de chaque maille. Une nomenclature, présentée sur la figure 5.10, a été définie afin de pouvoir repérer les mailles du bassin de l'Odet :

- Les mailles sont indicées par *o* lorsqu'elles appartiennent au sous-bassin de l'Odet. Elles sont de couleur violette. L'exutoire du sous-bassin à Tréodet correspond à la maille *o21*.

- Les mailles sont indicées par *j* lorsqu'elles appartiennent au sous-bassin du Jet. Elles sont de couleur verte. L'exutoire du sous-bassin à Kerjean correspond à la maille *j28*.

- Les mailles sont indicées par *s* lorsqu'elles appartiennent au sous-bassin du Steïr. Elles sont de couleur orange. L'exutoire du sous-bassin à Ty-Planche correspond à la maille *s19*.

- Les mailles sont indicées par *oa* pour *Odet aval* lorsqu'elles sont situées à l'aval de

Tréodet. Lorsqu'elles sont à l'aval des trois stations, elles sont de couleur bleue. La maille *oa25* est située uniquement à l'aval de Tréodet et Kerjean. Elle est de couleur violet foncé.

- La maille *sa23*, pour Steïr aval, est située uniquement à l'aval de Ty-planche. Elle est de couleur jaune.
- Les mailles grises ne sont pas situées à l'aval des trois stations amont : une perturbation des paramètres sur les sous-bassins amont n'influence pas le débit simulé à leur exutoire.

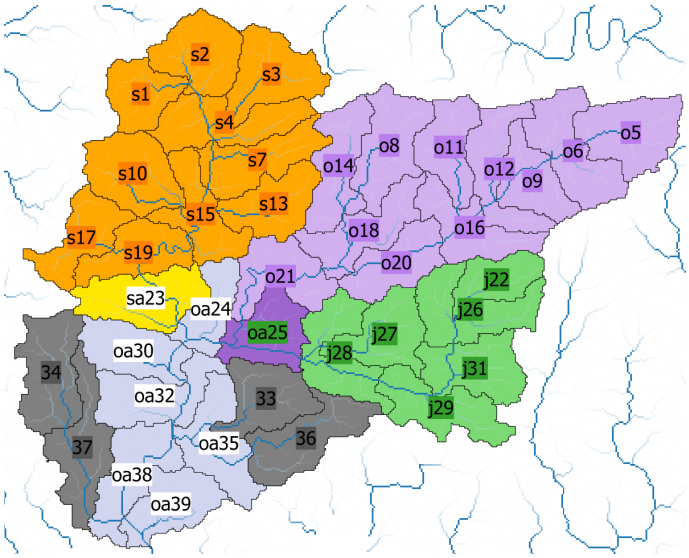


Figure 5.10: Nom attribué à chacune des 39 mailles du bassin de l'Odet. Tréodet est situé à l'exutoire de la maille *o21*, Kerjean à l'exutoire de la maille *j28* et Ty-Planche à l'exutoire de la maille *s19*.

La figure 5.11 présente les résultats du calcul de la RMSE sur chaque maille, calculée sur les débits obtenus avec la simulation de contrôle et avec la simulation issue de l'analyse par rapport à ceux obtenus avec la simulation de référence. Cette RMSE est calculée pour la simulation d'analyse en considérant le dernier pas de temps de la fenêtre d'assimilation. Les mailles sont présentées de l'amont vers l'aval, et celles correspondant à Tréodet et Plaisance sont spécifiquement mentionnées. Ce graphique met en évidence que la RMSE calculée sur tout l'événement à l'exutoire de chaque maille est nettement améliorée par l'assimilation de données. Nous notons que l'erreur résiduelle est plus importante lorsque l'écart-type sur les erreurs d'observation est plus important. Ici encore, nous ne notons pas de différence de performance entre les expériences à 49 et 279 membres.

À cette étape de l'étude, nous pouvons conclure que bien que les valeurs analysées ne convergent pas vers la valeur de référence pour quasiment tous les paramètres, cela n'empêche pas l'assimilation de données de fournir de bons résultats dans la phase d'analyse. Le modèle MORDOR-TS possède ici 8 paramètres libres, et il est possible que le problème soit sous-contraint pour qu'il existe une solution unique. Cela signifie qu'il existe

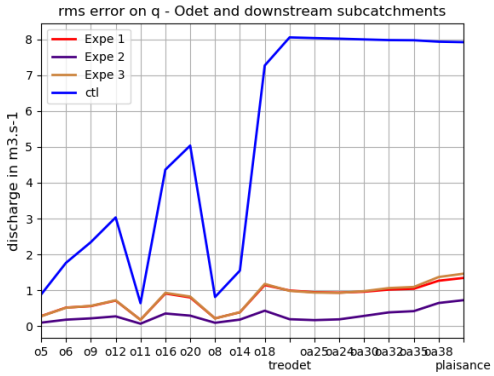


Figure 5.11: RMSE sur les mailles de l'Odet de l'amont vers l'aval pour la simulation de contrôle (bleu) et issue de l'analyse (rouge, violet et marron) par rapport à la simulation de référence. Événement de décembre 2013. La nomenclature des mailles est présentée figure 5.10.

plusieurs jeux de paramètres qui permettent d'avoisiner les observations sur chaque fenêtre d'assimilation. On parle alors d'équifinalité. Une densification du réseau d'observation doit pouvoir pallier ce problème. Ce sera l'objet du paragraphe 5.2.3.2. Il se peut alors que la phase de prévision suite à l'analyse ne soit pas optimale, puisque l'état dans lequel se trouve le modèle analysé n'est pas le même que celui de la simulation qui a généré la prévision de référence. La suite du paragraphe s'intéresse aux résultats en phase prévision.

Résultats dans la phase de prévision La figure 5.12 montre les résultats du calcul de la RMSE à Tréodet pour les prévisions pour les échéances variant de 1 heure à 24 heures. Nous notons tout d'abord que l'erreur de la simulation de contrôle n'est pas constante en fonction de l'échéance. Ceci aurait été le cas si les prévisions avaient été alimentées par ANTILOPE : en effet, les valeurs simulées auraient alors été identiques pour toutes les échéances de prévision. Ce point a fait l'objet d'une vérification. Ainsi, la forme de la courbe d'erreur est dûe à la chronique de pluie RR3 imposée en prévision : la nature des pluies est telle que pour les échéances entre 5 heures et 18 heures, l'erreur entre la simulation de contrôle et la référence est augmentée par rapport à celle des autres échéances. Après assimilation de données, les prévisions sont améliorées jusqu'à une échéance lointaine. L'erreur résiduelle est croissante en fonction des échéances pour les trois expériences. Nous notons que celle-ci est plus importante que dans les expériences *corr1par*. Ainsi, l'équifinalité mise en évidence précédemment n'empêche pas le modèle de rester relativement prédictif, mais elle dégrade un peu les performances en prévision. Nous avons vu que les paramètres les mieux corrigés sont ceux dont l'indice de Sobol est le plus important. Cela peut expliquer pourquoi le modèle reste prédictif. Par ailleurs, nous notons comme lors de la phase d'analyse que les erreurs résiduelles sur le débit simulé avec assimilation de données sont plus importantes lorsque l'écart-type des erreurs d'observations est plus important. La comparaison des expériences *Expe 1* et *Expe 3* met

en évidence le fait que pour les courtes échéances, le nombre de membres n'a aucune influence sur l'erreur résiduelle après assimilation de données à Tréodet. En revanche, à partir de l'échéance 8 heures environ, l'augmentation du nombre de membres de l'ensemble permet de diminuer l'erreur résiduelle. La réduction de l'erreur n'est cependant pas très importante.

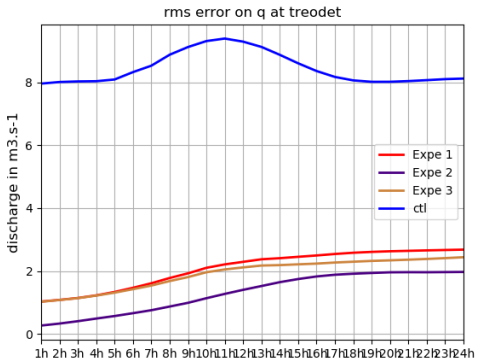


Figure 5.12: RMSE des erreurs de prévision à Tréodet en fonction de l'échéance pour la simulation de contrôle (bleu) et le modèle analysé pour les trois expériences par rapport à la simulation de référence. Événement de décembre 2013.

Pour finir cette analyse comparative, nous présentons sur la figure 5.13 les histogrammes de rang de l'analyse et des prévisions à échéance 6 heures pour les trois expériences. Les histogrammes de rang sont légèrement sur-dispersifs pour l'analyse et à courte échéance lorsque l'écart-type des erreurs d'observation vaut 0.5% (*Expe 2*). En revanche, avec un écart-type d'erreur d'observations de 5%, la forme des histogrammes de rang est représentative d'un ensemble sous-dispersif, quel que soit le nombre de membres utilisés (*Expe 1* et *Expe 3*). Les graphes de la figure 5.14 permettent d'avoir une meilleure compréhension de l'origine des formes des histogrammes de rang. Ils représentent pour les trois expériences la simulation de contrôle (bleu), la simulation de référence (noir) qui a servi à construire les observations perturbées prises en compte pour l'assimilation de données (étoiles vertes), ainsi que les membres assimilés (spaghettis orange) et leur moyenne (rouge). Nous remarquons que les membres présentent davantage de dispersion lorsque l'erreur prise en compte est plus importante. Cependant, les observations sont encore plus dispersées, d'où la forme sous-dispersive de l'histogramme de rang. Une piste d'amélioration de ce travail réside dans le choix de la stratégie de regénération de l'ensemble à chaque nouveau cycle. Il serait intéressant de voir s'il est possible d'améliorer la fiabilité des ensembles en testant plusieurs alternatives de regénération de l'ensemble.

Le tableau 5.6 résume les CRPS obtenus pour l'analyse et les échéances de prévision 1h, 6h, 12h et 18h pour les trois expériences. Il met en évidence le fait que les CRPS sont d'autant plus faibles que lorsque l'écart-type de l'erreur prise en compte est faible. Là encore, nous ne notons pas de différence notable du choix du nombre de membres de

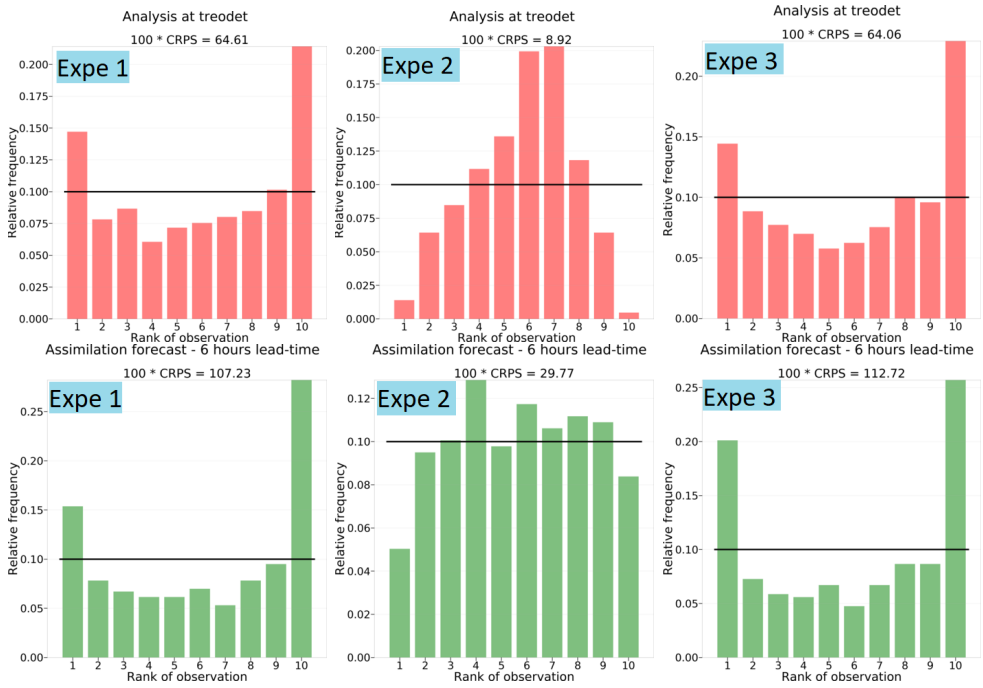


Figure 5.13: Histogrammes de rang de l'analyse et des prévisions à l'échéance 6 heures à Tréodet. Événement de décembre 2013.

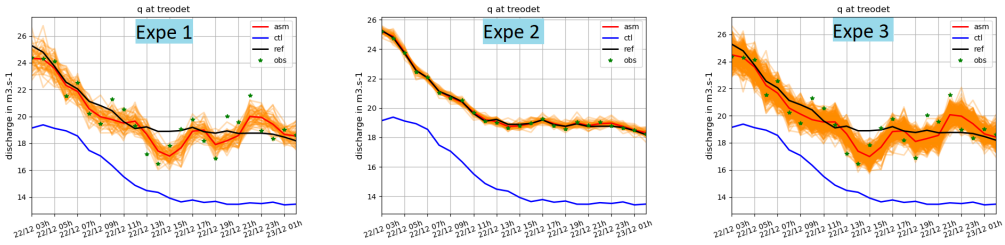


Figure 5.14: Simulation de contrôle (bleu), assimilation de données (rouge), membres (spaghettis orange), référence (noir) et observations (points verts). Décembre 2013.

l'ensemble. Dans tous les cas, les CRPS augmentent avec l'échéance de prévision.

Tableau 5.4: Bilan du CRPS pour les trois expériences.

Expérience	Analyse	Ech 1 h	Ech 6 h	Ech 12 h	Ech 18 h
Expe 1	64.6	95.0	107.2	147.3	147.1
Expe 2	8.9	13.8	29.8	57.2	84.0
Expe 3	64.0	97.4	112.7	131.0	142.5

Conclusion Dans ce paragraphe, nous avons étudié l'impact du choix du nombre de membres et de l'écart-type de l'erreur d'observation considérés pour la réalisation des expériences jumelles. Nous avons mis en évidence que dans tous les cas, des problèmes

d'équifinalité empêchent probablement le vecteur de contrôle d'être très proche de celui choisi en référence. Nous avons vu que les trois paramètres pour lesquels la valeur de contrôle est relativement bien approchée par l'analyse correspondent à ceux qui ont un indice de Sobol significatif au cours des événements de crue. L'erreur résiduelle après assimilation de données est plus faible en sur la fenêtre d'assimilation et en prévision lorsque l'écart-type des erreurs d'observations est plus faible. La forme des histogrammes de rang pour l'analyse et la prévision est directement liée au choix de cette erreur. Pour la suite, nous conservons une erreur d'observation égale à 5% de la valeur mesurée car c'est un ordre de grandeur cohérent avec celui obtenu avec des mesures réelles. Nous avons également vu qu'il n'y a pas d'impact notable du passage de 49 membres à 279 membres. Pour la suite des expériences, nous considérerons donc désormais 49 membres. C'est donc la configuration de l'expérience *Expe 1* qui est retenue.

5.2.3.2 Analyse de l'impact de l'assimilation de données sur des mailles à l'amont de Tréodet

Les expériences *corr8par* présentées précédemment étaient réalisées en assimilant uniquement le débit à Tréodet et en corrigeant les paramètres du modèle MORDOR-TS sur toutes les mailles de ce sous-bassin. Il n'existe pas de mesure de débit sur les mailles situées à l'amont de Tréodet, mais un des intérêts des expériences jumelles est que l'on peut tester l'apport que pourrait fournir une telle mesure. Dans cette partie, nous conservons l'expérience retenue à l'issue des tests précédents, à savoir l'expérience *Expe 1*. Dans ce paragraphe, elle sera désormais nommée *assim@treo* puisque l'on y assimile uniquement les débits à Tréodet. Elle sera comparée avec deux autres expériences nommées *assim@treo&20* et *assim@tous_sbv* pour lesquelles des débits sur des mailles à l'amont de Tréodet sont également assimilés. Les caractéristiques des trois expériences comparées dans ce paragraphe sont décrites dans le tableau 5.5. Les expériences possèdent désormais 49 membres et l'écart-type d'erreur d'observation est fixé à 5% de la valeur. L'événement étudié est toujours celui de décembre 2013.

Nom de l'expérience	Mailles pour lesquelles le débit est assimilé
assim@treo = Expe 1	Tréodet (maille o21)
assim@treo&20	Tréodet (maille o21) et maille o20
assim@tous_sbv	Les 11 mailles à l'amont de Tréodet (Tréodet, o5, o6, o9, o12, o11, o16, o20, o8, o14 et o18)

Tableau 5.5: Caractéristiques des trois expériences comparées dans le paragraphe 5.2.3.2. Pour la localisation des mailles, se référer à la figure 5.10. Toutes les expériences possèdent 49 membres et un écart-type pour les erreurs d'observation égal à 5% de la valeur.

Dans l'expérience *assim@treo&20*, les débits sont assimilés à l'exutoire Tréodet (maille o21) ainsi que à l'exutoire de la maille 20 (voir figure 5.10). La maille o20 draine 126

km² à l'amont, alors que l'exutoire à Tréodet draine 205 km². La surface du sous-bassin amont associé à la maille o20 représente ainsi environ 60% de la surface du sous-bassin dont l'exutoire est à Tréodet. Dans l'expérience *assim@tous_sbv*, les débits sont assimilés sur chacune des mailles situées à l'amont de Tréodet, soit sur 11 mailles au total. Comme pour toutes les expériences précédentes, les paramètres du modèle MORDOR-TS sont corrigés de manière uniforme sur tout le sous-bassin amont de Tréodet, c'est-à-dire que les 11 mailles à l'amont de Tréodet possèdent les mêmes paramètres.

Résultats dans la phase d'analyse La figure 5.15 présente les chroniques des paramètres de MORDOR-TS analysés pour les trois expériences. Ces graphiques mettent tout d'abord en évidence que les problèmes d'équifinalité présents dans l'expérience *assim@treo* évoqués en partie 5.2.3.1 sont levés dans l'expérience *assim@tous_ssbv* lorsque l'on assimile les débits à l'exutoire de chaque maille. Nous notons en effet que pour cette expérience, les analyses convergent toutes vers la valeur de référence. Le fait d'assimiler un seul débit supplémentaire dans l'expérience *assi@treo&20* améliore les analyses pour certains paramètres seulement. L'amélioration est très visible pour le paramètre lk_N (comparaison des courbes rouge et violette). Pour les paramètres k_r et Cel, la valeur de référence est atteinte beaucoup plus rapidement dans l'expérience *assi@treo&20*. Pour les autres paramètres, les chroniques des paramètres analysés sont analogues pour les deux expériences *assi@treo* et *assi@treo&20*. Ainsi, plus on assimile des débits différents à l'amont, meilleurs sont les valeurs analysées des paramètres que l'on cherche à corriger. Le problème d'équifinalité est dû à un problème sous-contraint. Le fait d'assimiler des débits sur plus de mailles à l'amont permet de contraindre davantage le problème et de lever ce problème.

La figure 5.16 montre pour les trois expériences à Tréodet les résultats du calcul de l'erreur sur le débit obtenu avec la simulation de contrôle et par l'assimilation de données par rapport à la simulation de référence. Nous notons que l'erreur sur le débit analysé à Tréodet possède une moins grande variabilité lorsque des débits sont assimilés en plus de celui à Tréodet.

Pour finir, la figure 5.17 présente la RMSE sur le débit mesuré à l'exutoire de toutes les mailles situées de l'amont du sous-bassin jusqu'à l'aval à Plaisance (voir figure 5.10 pour la position des mailles). Les erreurs résiduelles après assimilation de données dépendent de l'expérience et de la position de la maille. Ainsi, pour les mailles situées à l'amont de Tréodet (Tréodet inclus), l'erreur sur le débit à l'exutoire de chaque maille est réduite lorsque le nombre de débits assimilés sur ces mailles augmente. À partir de la maille oa30, soit trois mailles à l'aval de Tréodet, l'erreur la plus faible est obtenue avec l'expérience *assim@treo*, puis c'est l'expérience *assim@tous_sbv*, et c'est enfin l'expérience *assim@treo&20* qui présente l'erreur la plus importante.

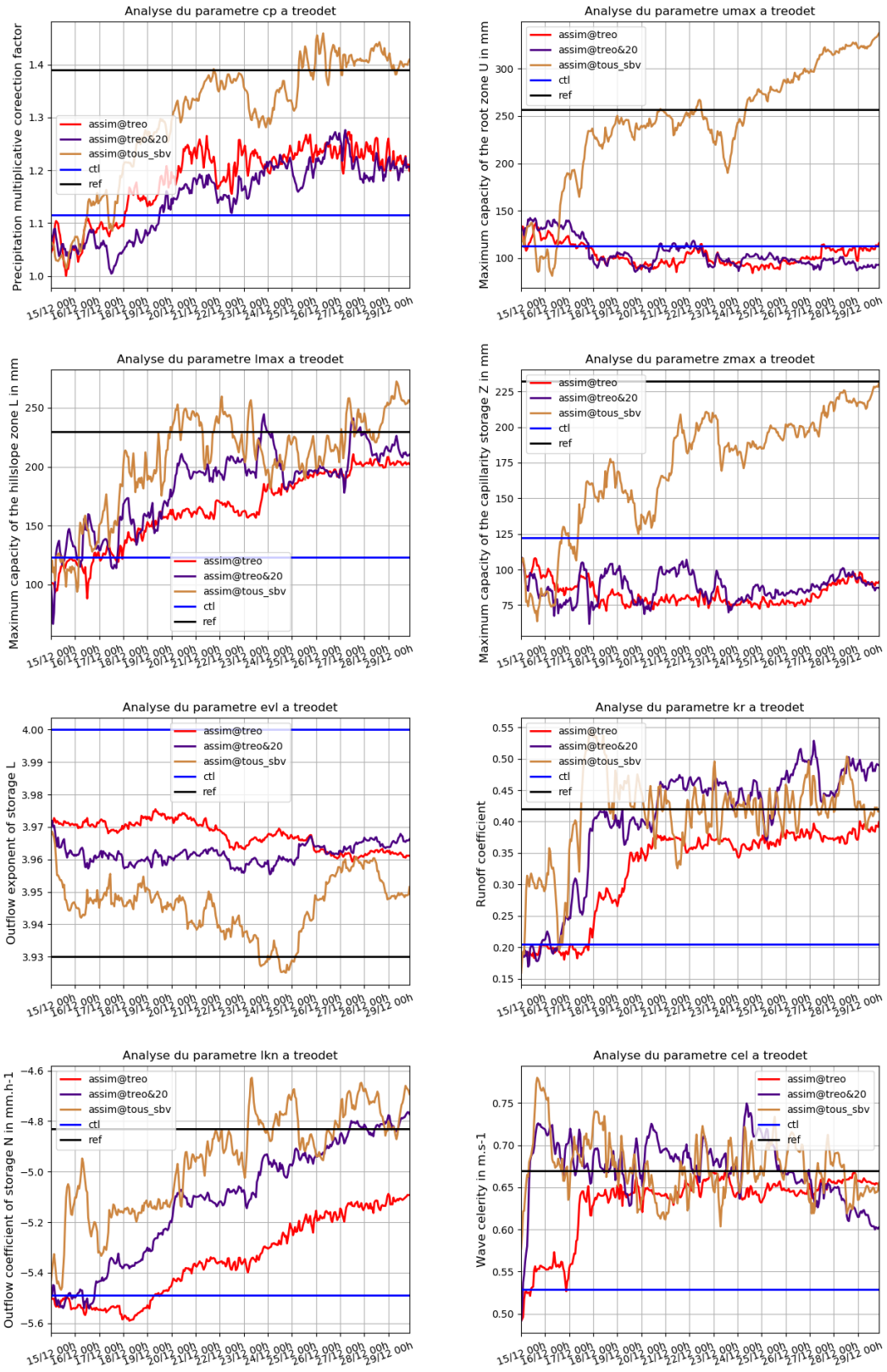


Figure 5.15: Valeurs analysées pour le vecteur de contrôle (paramètres) en fonction du temps pour les trois expériences *assim@treo*, *assim@treo&20* et *assim@tous_sbv* menées à Tréodet (décembre 2013). Leurs caractéristiques sont résumées dans le tableau 5.5.

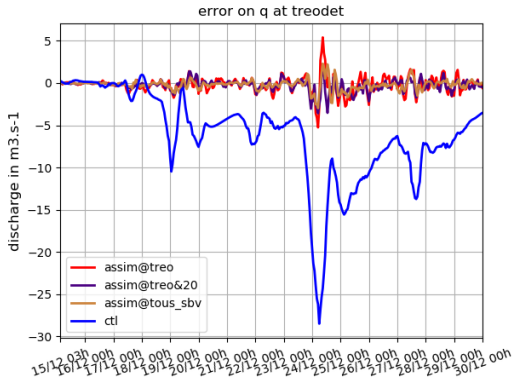


Figure 5.16: Chronique temporelle des erreurs à Tréodet pour la simulation de contrôle (bleu) et pour les analyses réalisées lors des trois expériences *assim@treo*, *assim@treo&20* et *assim@tous_sbv* par rapport à la simulation de référence. Evénément de décembre 2013.

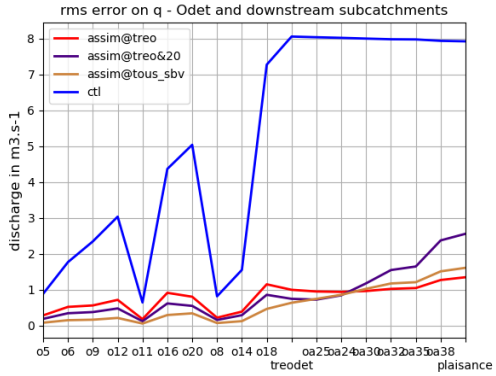


Figure 5.17: RMSE sur les mailles de l'Odet de l'amont vers l'aval pour le contrôle (bleu) et de l'analyse pour les trois expériences par rapport à la simulation de référence. Événement de décembre 2013.

Résultats dans la phase de prévision La figure 5.18 montre les résultats du calcul de la RMSE à Tréodet pour les prévisions pour les échéances variant de 1 heure à 24 heures. D'une manière générale, plus le nombre de débits assimilés à l'amont de Tréodet est important, moins l'erreur résiduelle en prévision est importante. Le fait d'assimiler des débits sur les mailles à l'amont de Tréodet permet au modèle d'être un peu plus prédictif. En effet, comme nous l'avons vu, cela permet de lever en partie les problèmes d'équifinalité. Dans ce cas, l'algorithme corrige les paramètres du bassin pour améliorer la valeur des débits simulés sur les mailles situées à l'amont, qui arriveront à l'exutoire Tréodet pendant la phase de prévision. Une correction des débits amont permet donc naturellement d'être davantage prédictif. À partir de l'échéance de prévision de 20 heures, on note cependant que l'erreur de l'expérience *assim@tous_sbv* dépasse celle de l'expérience *assim@treo&20*, pour rejoindre celle de l'expérience *assim@treo* à l'échéance 24 heures.

Pour compléter cette analyse comparative, nous présentons sur la figure 5.19 les his-

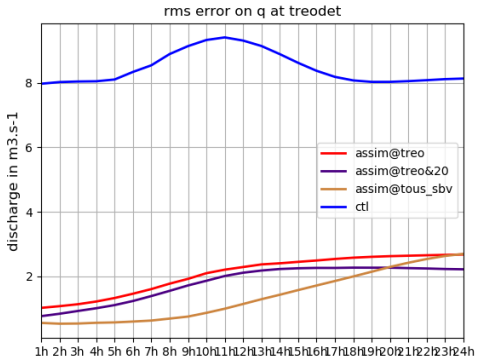


Figure 5.18: RMSE des erreurs de prévision à Tréodet en fonction de l'échéance pour la simulation de contrôle (bleu) et le modèle analysé pour les trois expériences par rapport à la simulation de référence. Événement de décembre 2013.

togrammes de rang de l'analyse et des prévisions à échéance 6 heures pour les trois expériences. La forme des histogrammes de rang est caractéristique d'ensembles sous-dispersifs. Cette tendance générale est retrouvée pour les trois expériences, cependant, on note que plus le nombre de débits assimilés est important, plus cette tendance semble marquée. Comme nous l'avons mentionné précédemment, il serait pertinent d'investiguer la méthode de regénération de l'ensemble à chaque cycle d'assimilation.

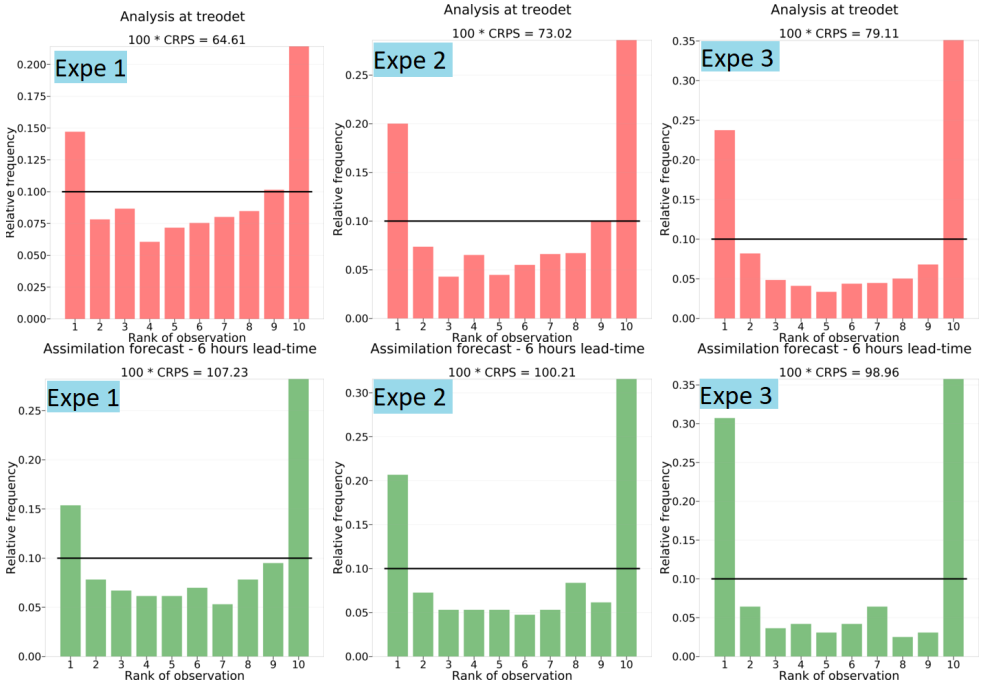


Figure 5.19: Histogrammes de rang de l'analyse et des prévisions pour l'échéance 6 heures à Tréodet. Événement de décembre 2013.

Enfin, le tableau 5.6 résume les CRPS obtenus pour l'analyse et les échéances de prévision

1h, 6h, 12h et 18h pour les trois expériences. Les cases jaunes (respectivement orange) correspondent au CRPS minimum (respectivement maximum) obtenu parmi les trois expériences. En phase d'analyse, le CRPS est plus faible dans l'expérience *assim@treo* dans laquelle seul le débit à Tréodet est assimilé, et il est maximal pour l'expérience *assim@tous_sbv* dans laquelle on assimile les débits de chaque maille. Ainsi, l'ensemble analysé à Tréodet est de meilleure qualité lorsque seuls les débits à Tréodet sont assimilés. En effet, le fait d'assimiler des débits à l'amont implique que l'analyse est optimisée pour que le débit simulé soit bon sur chaque maille en moyenne, donc il est moins bon sur un point particulier par construction. En revanche, les CRPS en prévision sont d'autant plus faibles que le nombre de mailles assimilées à l'amont est important. En effet, le fait d'assimiler des débits à l'amont permet de prendre en compte des débits qui passeront à l'exutoire Tréodet plus tard, c'est pour cela que la prévision est améliorée avec l'assimilation des débits amont.

Expérience	Analyse	Ech 1 h	Ech 6 h	Ech 12 h	Ech 18 h
assim@treo	64.6	95.1	107.2	147.3	147.1
assim@treo&20	73.0	97.4	100.2	119.0	136.4
assim@tous_sbv	79.1	95.4	99.0	106.1	123.5

Tableau 5.6: Bilan du CRPS pour les trois expériences. Événement de décembre 2013.

5.2.3.3 Résultats à Kerjean et Ty-Planche

Les résultats des expériences *corr8par* menées respectivement sur les sous-bassins du Jet (exutoire Kerjean) et du Steïr (exutoire Ty-Planche) sont disponibles en annexes O.4 et O.5. Ces expériences ont été menées avec 49 membres, une erreur d'observation égale à 5% de la valeur, et en assimilant les débits uniquement à l'exutoire du sous-bassin. Les résultats sont consistants avec ceux obtenus sur le sous-bassin de l'Odet dont l'exutoire est à Tréodet. Sur les trois sous-bassins, les valeurs analysées pour les paramètres ne convergent pas tous vers les valeurs imposées pour la simulation de référence. Ceci est *a priori* dû au problème d'équifinalité, car nous avons un système sous-contraint. Ceci n'est cependant pas un frein à l'obtention des résultats de simulation satisfaisants, que ce soit sur la fenêtre d'assimilation ou en prévision opérationnelle.

5.2.4 Synthèse

Dans ce paragraphe, nous avons décrit les expériences jumelles réalisées pour valider l'algorithme d'assimilation de données mis en place pour la modélisation hydrologique avec MORDOR-TS. Nous avons d'abord décrit le cadre des expériences jumelles. Toutes les expériences sont menées à l'échelle d'un sous-bassin. La fenêtre d'assimilation dure

3 heures sur laquelle 3 observations sont assimilées. La fenêtre de prévision opérationnelle dure 24 heures.

Des tests ont été réalisés sur le nombre de membres considérés dans l'algorithme EnKF. Nous avons mis en évidence que le passage de 49 membres à 279 membres n'avait pas de réel impact sur les résultats, que ce soit concernant les valeurs des paramètres analysés ou les erreurs sur les débits analysés ou prévus. Nous conservons donc un ensemble constitué de 49 membres. Des tests ont également été réalisés sur la valeur de l'écart-type de l'erreur d'observation utilisé pour générer les observations des expériences jumelles et au cœur de l'algorithme EnKF stochastique. Nous avons mis en évidence que les erreurs résiduelles sont moins importantes avec une petite erreur d'observation. Si cela est vrai dans le cadre d'expériences jumelles, il s'agit tout de même de conserver une valeur réaliste pour les expériences réelles. Nous conservons donc une erreur égale à 5% de la valeur observée.

Dans ce paragraphe, nous nous sommes intéressés à deux types d'expériences. Dans les expériences *corr1par*, un seul paramètre du modèle est corrigé. Ces expériences ont montré que l'algorithme d'assimilation de données permet de retrouver la valeur de référence pour tous les paramètres. Les résultats dans l'espace des observations (débits analysés, débits prévus) sont alors de bonne qualité puisque l'assimilation de données a permis de remettre le modèle dans un état proche de celui qui a construit les observations de référence. Dans les expériences *corr8par*, les 8 paramètres du modèle MORDOR-TS sont simultanément corrigés. Nous avons mis en évidence qu'alors, les paramètres analysés ne convergent pas vers la valeur de référence lorsqu'on assimile uniquement les débits à l'exutoire du sous-bassin. Le fait d'assimiler des débits sur les mailles à l'amont permet alors d'améliorer les valeurs des paramètres analysés, et tous les paramètres convergent vers la valeur de référence si l'on assimile les débits sur toutes les mailles amont. Dans l'espace des observations, ce problème d'équifinalité n'a pas d'impact sur la qualité des débits analysés sur la fenêtre d'assimilation. Ils restent de très bonne qualité. En prévision, l'erreur résiduelle est plus importante que dans les expériences *corr1par*. Cependant, l'assimilation de données permet là-encore de réduire de manière importante les erreurs de prévision.

Les expériences jumelles réalisées sur un événement type de décembre 2013 et décrites dans ce paragraphe permettent de valider l'algorithme d'assimilation de données mis en place pour la modélisation hydrologique, et nous sommes à présent en mesure de réaliser des expériences réelles sur une période plus longue dans un contexte opérationnel.

5.3 Réalisation des expériences réelles

5.3.1 Caractéristiques des expériences

Les expériences réelles d'assimilation de données nous permettent de construire des prévisions d'ensemble. Elles sont générées sur la période du 1er janvier 2010 au 31 décembre 2019. Cette période couvre exactement la disponibilité des prévisions RR3. Comme pour la calibration QRF, les trois mois d'été (juillet, août et septembre) ne sont pas traités. Ainsi, pour chaque année entre 2010 et 2019, l'assimilation est lancée de manière cyclée pendant 9 mois de manière continue. Le vecteur de contrôle d'ébauche initial est égal aux paramètres calés.

Pour la réalisation de l'assimilation de données dans un cadre réel, nous avons à disposition uniquement des mesures de débit aux trois exutoires Tréodet, Kerjean et Ty-Planche. Sur chaque sous-bassin, on assimile donc uniquement les débits mesurés à son exutoire. Suite à l'étude des expériences jumelles, nous retenons un ensemble de 49 membres, une fenêtre d'assimilation de 3 heures, et un écart-type sur les erreurs d'observation égal à 5% de la valeur de l'observation.

Nous analysons dans ce chapitre au paragraphe 5.3.2 les performances de l'assimilation de données sur la fenêtre d'assimilation. L'analyse des résultats pendant la phase de prévision sera l'objet du chapitre suivant (chapitre 6), qui comparera les méthodes de calibration statistique et d'assimilation de données pour la génération de prévisions d'ensemble dans un contexte opérationnel.

5.3.2 Résultats dans la phase d'analyse

Les résultats présentés dans cette partie sont calculés sur la période du 1er janvier 2010 au 31 décembre 2019, sauf les trois mois d'été qui sont toujours retirés. Le tableau 5.7 montre le résultat du calcul de la RMSE par rapport aux mesures sur toute la période d'assimilation de données pour la simulation issue du modèle calé et celle avec assimilation de données. Le calcul est réalisé : (i) pour tous les pas de temps de la période et (ii) pour les pas de temps pour lesquels le débit mesuré est supérieur au quantile q_{90} à la station.

	Tréodet		Kerjean		Ty-Planche	
	modèle calé	assimilation	modèle calé	assimilation	modèle calé	assimilation
Toute la période	1.65	0.29	0.64	0.16	1.23	0.21
Qobs > Q90	3.67	0.75	1.33	0.36	2.49	0.47

Tableau 5.7: RMSE calculée sur la période janvier 2010 - décembre 2019 (sauf mois d'été) pour le modèle MORDOR-TS issu du calage et avec assimilation de données pour les trois sous-bassins dont les exutoires sont à Tréodet, Kerjean et Ty-Planche. Calcul réalisé pour tous les débits (ligne 1) et pour les débits mesurés supérieurs au quantile q_{90} à la station (ligne 2).

Ce tableau met en évidence une nette diminution de l'erreur sur la fenêtre d'assimilation

avec l'assimilation de données, que ce soit pour tous les débits ou les débits supérieurs au quantile q_{90}. Les bons résultats en phase d'analyse sont illustrés sur les figures 5.20 et 5.21 qui représentent les chroniques des débits simulés par le modèle calé, par le modèle avec assimilation de données et la chronique observée pour les trois stations pour deux événements significatifs : événement du 28 février 2010 (tempête Xynthia) et période du 1er au 10 janvier 2014. Ces graphes mettent en évidence une très bonne correction du débit simulé dans la phase d'analyse. Les pics de débits ne sont pas toujours atteints parfaitement mais sont tout de même très proches de la mesure.

Pour finir, le tableau 5.8 présente les valeurs analysées moyennées sur toute la période et rappelle les valeurs calées par l'algorithme caRamel sur la période 2007-2013 (voir chapitre 2, paragraphe 2.3.1.2). Nous remarquons que pour la moitié des paramètres environ, l'intervalle centré sur la valeur analysée et de largeur 2 fois la valeur de l'écart-type de l'analyse contient la valeur calée. Pour l'autre moitié, la valeur calée est significativement différente de la valeur analysée. Cela justifie la correction par assimilation de données qui permet l'adaptation en fonction du temps des paramètres du modèle.

Bassin versant		c_p	U_{max}	L_{max}	Z_{max}	ev_L	k_r	lk_N	Cel
Tréodet	calage	1.05	108.2	104.5	79.5	3.99	0.2	-5.45	0.52
	analyse moyenne	1.10	211	270	102	3.98	0.27	-6.11	0.51
	écart-type	0.21	74	69	24	0.01	0.08	0.42	0.11
Kerjean	calage	1.03	65.0	386.5	162.2	1.50	0.15	-5.94	0.9
	analyse moyenne	1.14	93	455	326	3.41	0.36	-5.80	0.66
	écart-type	0.11	19	123	131	0.33	0.11	0.61	0.13
Ty-Planche	calage	0.97	101.0	30.0	34.6	2.85	0.10	-5.75	0.42
	analyse moyenne	1.13	108	386	603	3.66	0.29	5.68	0.60
	écart-type	0.20	54	123	160	0.45	0.10	0.21	0.06

Tableau 5.8: Valeurs analysées moyennes pour les 8 paramètres de MORDOR-TS et valeurs issues du calage hydrologique.

5.4 Synthèse

Ce chapitre a présenté les expériences jumelles réalisées dans le cadre de l'assimilation de données dans le modèle MORDOR-TS. Elles ont permis de valider l'algorithme mis en place et de choisir le nombre de membres pour la réalisation des expériences réelles.

Sur la fenêtre d'assimilation, les expériences réelles présentent de bons résultats. Les erreurs sur les débits simulés sont très nettement réduites, que ce soit pour les débits courants ou ceux supérieurs au seuil q_{90}. Les valeurs des paramètres analysés peuvent être proches ou éloignées des valeurs de calage. Il s'agit à présent d'étudier les performances de l'assimilation de données en prévision. C'est l'objet du chapitre 6, dans lequel nous allons comparer les résultats des expériences d'assimilation de données en prévision avec les résultats des ensembles construits par calibration statistique.

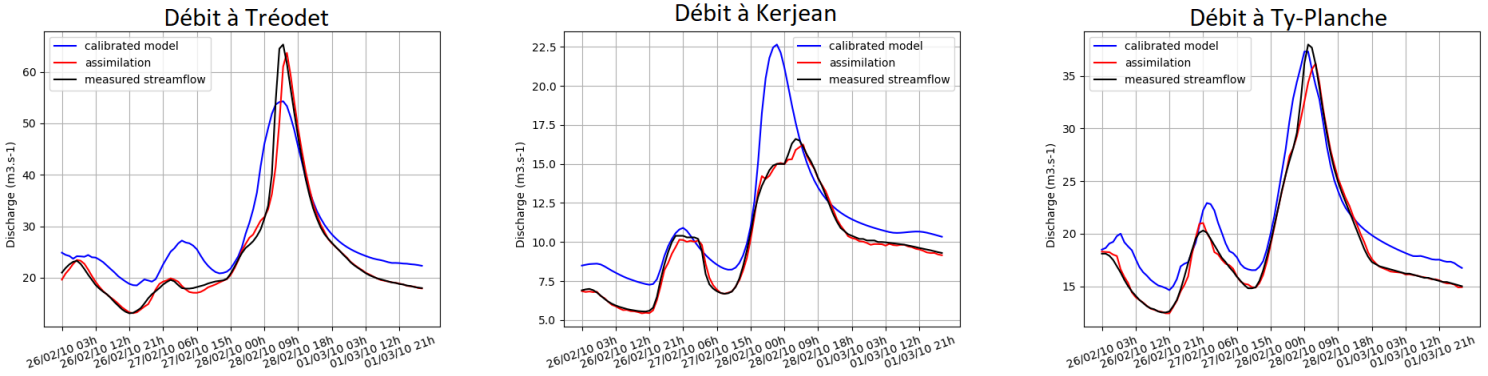


Figure 5.20: Chroniques de débits mesurés, simulés par le modèle calé et avec assimilation de données pour l'événement du 28 février 2010 (tempête Xynthia). Stations Tréodet, Kerjean et Ty-Planche.

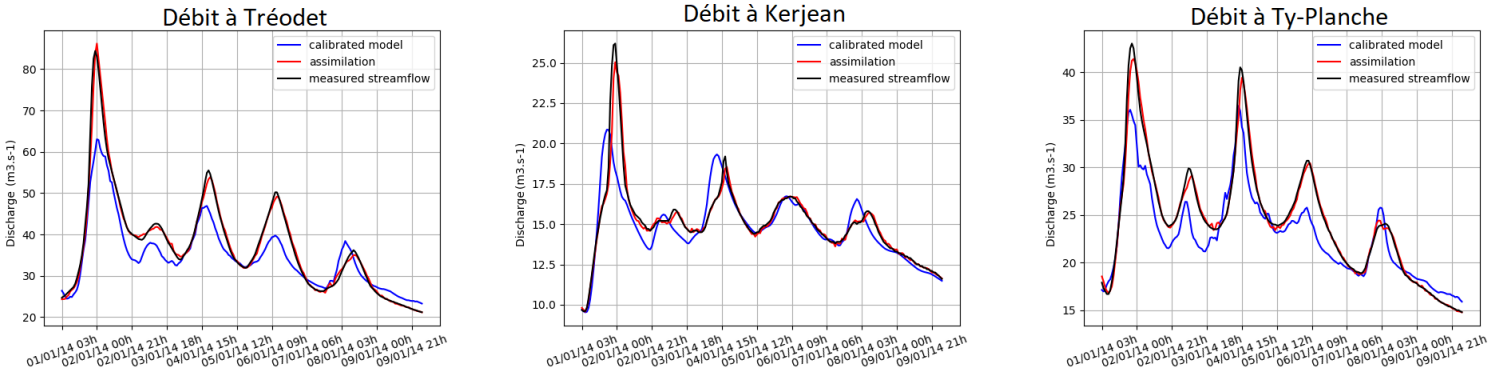


Figure 5.21: Chroniques de débits mesurés, simulés par le modèle calé et avec assimilation de données pour la période du 1er janvier 2014 au 10 janvier 2014. Stations Tréodet, Kerjean et Ty-Planche.

Chapitre 6

Comparaison des méthodes QRF et EnKF

Sommaire

6.1 Résultats comparés des ensembles *raw-HEF*, *QRF-hydro*, *QRF-hydro-TS* et *DA* en prévision 175

6.1.1 Analyse par échéance 175

6.1.2 Prévisions intégrées sur 24 heures : somme et maximum 179

6.2 Synthèse et perspectives 181

Ce chapitre a pour objectif de faire le bilan et de comparer les performances des prévisions d'ensemble hydrologiques générées dans le cadre de ce travail. Nous rappelons ci-après les deux méthodologies mises en œuvre et les ensembles correspondants :

- La calibration statistique des ensembles bruts hydrologique générés avec le modèle MORDOR-TS a fait l'objet du chapitre 4. Nous avons alors généré un ensemble brut appelé *raw-HEF*, et deux ensembles calibrés appelés *QRF-hydro* et *QRF-hydro-TS*. Ce dernier est construit à partir de *QRF-hydro* sur lequel est appliqué un lissage des trajectoires pour obtenir des chroniques de débits ayant un sens physique.

- L'introduction d'assimilation de données ensembliste dans le modèle hydrologique MORDOR-TS a fait l'objet du chapitre 5. L'ensemble en prévision est appelé *Data Assimilation* ou *DA* dans la suite du chapitre.

Afin de comparer les performances des ensembles dans un contexte opérationnel, tous les ensembles sont analysés en-dehors de la période d'apprentissage qui a servi à construire les forêts aléatoires. Nous travaillons désormais sur la période du 1er janvier 2010 au 30 juin 2011 et d'octobre 2014 à décembre 2019. Les mois d'été (juillet, août et septembre) sont toujours retirés. Cette période sera à présent nommée période de validation. Le tableau 6.1 synthétise la période utilisée pour le calage des forêts aléatoires sur les stations Tréodet, Kerjean et Ty-Planche, ainsi que la période de validation considérée. À Tréodet et Kerjean, la période de validation est homogène, puisqu'elle contient exactement 7 saisons d'octobre à juin. À Ty-Planche en revanche, nous avons vu que les données s'arrêtent fin avril 2016. La période de validation est donc moins homogène.

Année	Mois											
2010	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
2011	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
2012	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
2013	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
2014	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
2015	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
2016	**1**	**2**	**3**	**4**	5	6	7	8	9	10	11	12
2017	1	2	3	4	5	6	7	8	9	10	11	12
2018	1	2	3	4	5	6	7	8	9	10	11	12
2019	1	2	3	4	5	6	7	8	9	10	11	12

Période de calage des forêts aléatoires
Période de validation utilisée dans ce chapitre

Tableau 6.1: Périodes de calage des forêts aléatoires et de validation pour les stations Tréodet, Kerjean et Ty-Planche. Pour Ty-Planche, les mois disponibles sont uniquement ceux écrits en gras.

6.1 Résultats comparés des ensembles *raw-HEF*, *QRF-hydro*, *QRF-hydro-TS* et *DA* en prévision

Les résultats en prévision sont d'abord analysés par échéance de prévision (paragraphe 6.1.1). De même que pour l'étude des performances des forêts aléatoires sur la période de calage présentée au chapitre 4, les prévisions intégrées sur 24 heures seront également étudiées au (paragraphe 6.1.2) : somme des prévisions et maximum de la prévision.

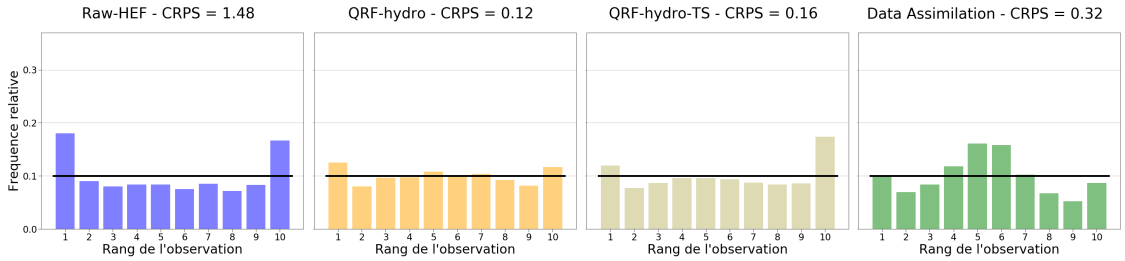
6.1.1 Analyse par échéance

La figure 6.1 présente les histogrammes de rang des prévisions pour les échéances 1h, 6h, 12h et 18h pour l'ensemble brut *raw-HEF*, les ensembles calibrés *QRF-hydro* et *QRF-hydro-TS*, et l'ensemble crée par assimilation de données *Data Assimilation* à station Tréodet. La stratification est réalisée selon les valeurs de l'ensemble brut. La figure 6.2 représente pour les échéances de prévision allant de 1 heure à 24 heures le CRPS, la RMSE, la norme $\|.\|_2$, et l'aire sous la courbe ROC pour les événements de dépassement des quantiles q_{90} et q_{98} pour les 4 prévisions d'ensemble à Tréodet. La RMSE est calculée à partir de la moyenne des ensembles. Elle est également calculée pour la prévision déterministe. Ce graphique permet ainsi d'évaluer l'apport de la prévision ensembliste par rapport à la prévision déterministe réalisée avec le modèle calé. Les résultats analogues pour les stations Kerjean et Ty-Planche sont respectivement présentés en annexes P.2 et P.3. Pour finir, le tableau 6.2 présente la valeur moyennée sur toutes les échéances de prévision de la norme $\|.\|_2$ des histogrammes de rang pour les trois stations.

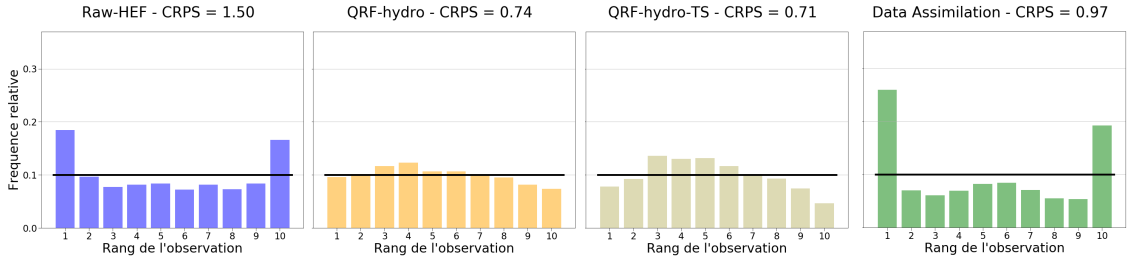
Sous-bassin	$10^{-2}\|.\|_2$ *raw-HEF*	$10^{-2}\|.\|_2$ *QRF-hydro*	$10^{-2}\|.\|_2$ *QRF-hydro-TS*	$10^{-2}\|.\|_2$ *DA*
Tréodet	18.2	5.6	7.8	28.1
Kerjean	13.6	6.7	9.2	16.7
Ty-Planche	23.7	11.3	11.4	23.3

Tableau 6.2: $\|.\|_2$ des histogrammes de rang moyennées sur toutes les échéances entre 1 heure et 24 heures pour les prévisions d'ensemble *raw-HEF*, *QRF-hydro*, *QRF-hydro-TS* et *Data assimilation* pour les trois stations Tréodet, Kerjean et Ty-Planche.

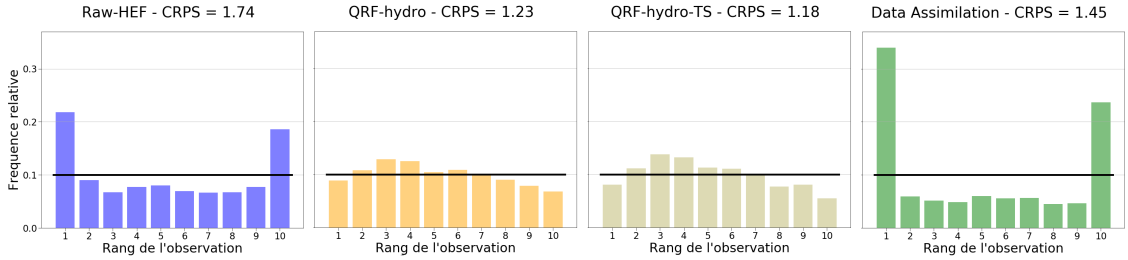
Comparaison des ensembles *QRF-hydro* et *QRF-hydro-TS* obtenus sur les périodes de calage et de validation des forêts aléatoires Avant d'effectuer la comparaison de tous les ensembles construits en opérationnel sur la période de validation (voir tableau 6.1), il est pertinent de comparer les performances de la calibration QRF sur la période de calage 2011-2014 et celle de validation. En effet, dans l'article *Strategies for hydrologic ensemble generation and calibration: on the merits of using model-based pre-*



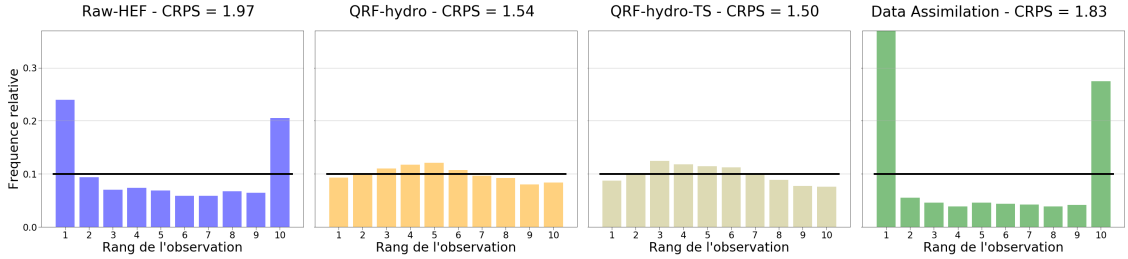
(a) Echéance 1 heure.



(b) Echéance 6 heures.



(c) Echéance 12 heures.



(d) Echéance 18 heures.

Figure 6.1: Histogrammes de rang des prévisions pour les ensembles *raw-HEF*, *QRF-hydro*, *QRF-hydro-TS* et *Data Assimilation*, pour les échéances 1 heure, 6 heures, 12 heures et 18 heures à Tréodet.

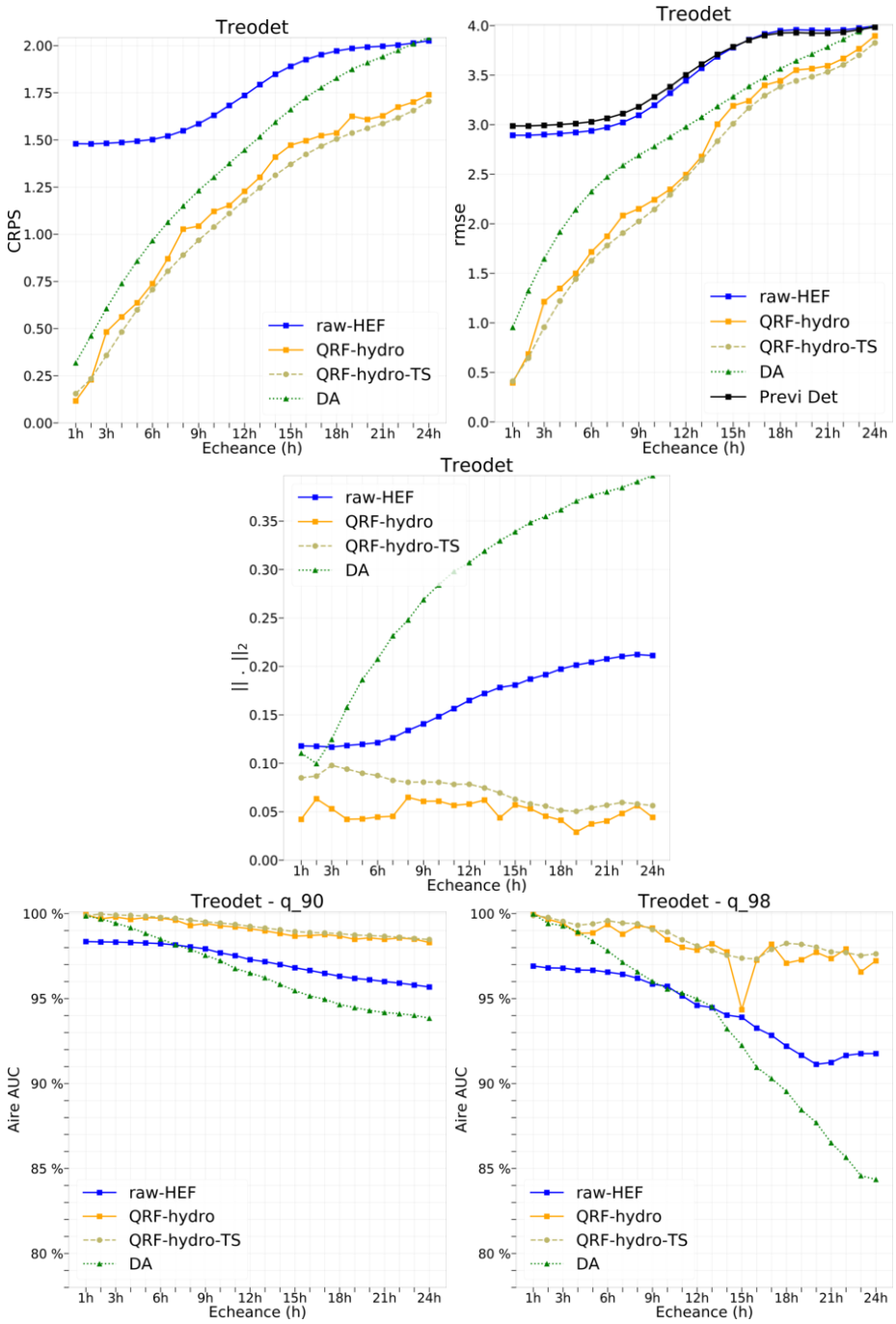


Figure 6.2: CRPS, RMSE, $||.||_2$ et aire ROC *AUC* pour les événements de dépassement des quantiles q_{90} et q_{98} pour les prévisions d'ensemble *raw-HEF*, *QRF-hydro*, *QRF-hydro-TS* et *Data assimilation (DA)* pour les échéances de 1 heure à 24 heures à Tréodet. La RMSE est calculée à partir de la moyenne des ensembles. Elle est également calculée pour la prévision déterministe avec la modèle calé (courbe noire de la figure en haut à droite).

dictors[1], l'étude des performances des forêts aléatoires construites n'a été réalisée que sur la période où elles ont été optimisées. Nous comparons donc les résultats présentés au chapitre 4 avec ceux présentés dans ce chapitre.

Nous notons que la valeur des CRPS par échéance présente une forme analogue pour les ensembles *raw-HEF*, *QRF-hydro* et *QRF-hydro-TS* pendant les période de calage et de validation des forêts aléatoires (comparer les CRPS figures 6.2, P.8 et P.19 avec ceux de la figure 4.3). L'amélioration du CRPS est donc consistante sur ces deux périodes. En ce qui concerne la fiabilité, l'étude de la forme des histogrammes de rang (comparaison des figures 6.1 et 4.2) et de leur norme $\|.\|_2$ (comparaison des tableaux 6.2 et 4.2) met en évidence une moins bonne correction de la fiabilité sur la période de validation, bien que celle-ci soit toujours bien meilleure que celle de l'ensemble brut.

Analyse comparative : CRPS et RMSE La courbe présentant les CRPS en prévision à Tréodet figure 6.2 met en évidence que les deux méthodologies développées, la calibration statistique et l'assimilation de données ensembliste, permettent d'améliorer les CRPS des ensembles à toutes les échéances. L'amélioration est plus prononcée aux courtes échéances. Par ailleurs, la calibration statistique permet d'obtenir de meilleurs CRPS que l'assimilation de données. L'analyse de la RMSE montre que l'ensemble brut *raw-HEF* présente une erreur semblable à la prévision déterministe. Les deux méthodologies ensemblistes mises en place (QRF et EnKF) permettent d'améliorer la RMSE par rapport à la prévision déterministe. Là aussi, la calibration statistique fournit les meilleurs résultats.

À Kerjean et Ty-Planche, les résultats montrent une tendance similaire (figures P.8 et P.19), mais les différences entre les deux méthodologies sont moins prononcées. La RMSE est d'ailleurs meilleure à Ty-Planche pour les échéances de prévision supérieures à 15 heures pour l'ensemble construit par assimilation de données. A Ty-Planche également, on note que l'ensemble *raw-HEF* présente une RMSE inférieure à celle de la prévision déterministe.

Analyse comparative : fiabilité L'analyse de la norme des histogrammes de rang met en évidence que les ensembles créés par assimilation de données *DA* présentent une très mauvaise fiabilité pour les trois stations. Ce point avait déjà été constaté au cours de la réalisation des expériences jumelles d'assimilation de données (voir chapitre 5). Ici, les normes des histogrammes de rang sont dégradées à partir de l'échéance 3 heures à Tréodet par rapport à celle de l'ensemble brut *raw-HEF*. Ces résultats se retrouvent également à Kerjean et Ty-Planche, même si l'échéance à partir de laquelle l'ensemble *DA* dégrade la fiabilité par rapport à l'ensemble brut *raw-HEF* est plus tardive. Cette caractéristique est

[1]Tiberi-Wadier, A-L., Goutal, N., Ricci, S., Sergent, P., Taillardat, M., Bouttier, F.,Monteil, C., Strategies for hydrologic ensemble generation and calibration: on the merits of using model-based predictors, Journal of Hydrology (2021), doi: https://doi.org/10.1016/j.jhydrol.2021.126233

peut-être liée à la stratégie de regénération de l'ensemble au fil des cycles d'assimilation. D'autres stratégies que celle employée pourraient être testées et il serait alors intéressant de voir s'il est possible d'améliorer la fiabilité de l'ensemble généré. Nous notons que la fiabilité des ensembles créés par calibration statistique est toujours améliorée par rapport à celle de l'ensemble brut *raw-HEF*.

Analyse comparative : discrimination des événements de dépassement de seuil
En ce qui concerne la capacité des ensembles à discriminer les événements de dépassement de seuil, l'analyse du score *AUC* met en évidence que là encore, l'ensemble *DA* présente des caractéristiques globalement plus mauvaises que celles de l'ensemble brut *raw-HEF*. L'assimilation de données ne permet d'améliorer la discrimination par rapport à l'ensemble brut que pour les très courtes échéances. Ce résultat est valable pour les trois stations et les deux seuils analysés. Les performances des ensembles *QRF-hydro* et *QRF-hydro-TS* sont analogues pour ce critère d'évaluation. Ainsi, le lissage des prévisions n'a pas d'impact significatif. Les deux ensembles *QRF-hydro* et *QRF-hydro-TS* améliorent nettement la capacité à discriminer les seuils testés par rapport à l'ensemble brut *raw-HEF*. Pour finir, nous notons que les performances pour la discrimination des événements a tendance à décroître avec les échéances.

À titre d'illustration, la figure 6.3 présente les prévisions cyclées à échéance 12 heures à Tréodet pour les 4 ensembles hydrologiques construits pour l'événement du 13 au 17 novembre 2010. Les graphiques analogues aux échéances 6, 18 et 24 heures ainsi que les résultats sur l'événement du 2 au 15 janvier 2016 sont présentés en annexe P.1.1 pour Tréodet. Les graphiques pour Kerjean et Ty-Planche sont respectivement placés en annexes P.2.3 et P.3.3.

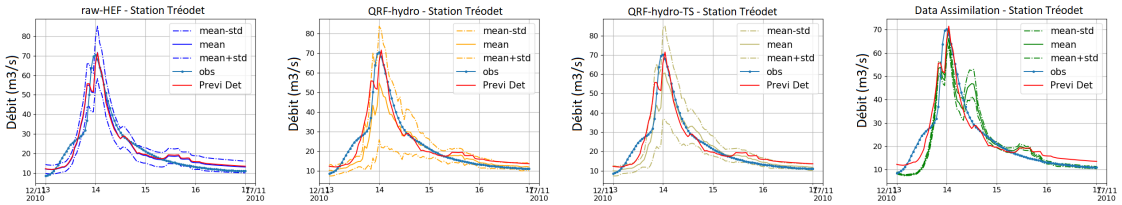


Figure 6.3: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 12 heures à Tréodet. Événement du 13 au 17 novembre 2010.

6.1.2 Prévisions intégrées sur 24 heures : somme et maximum

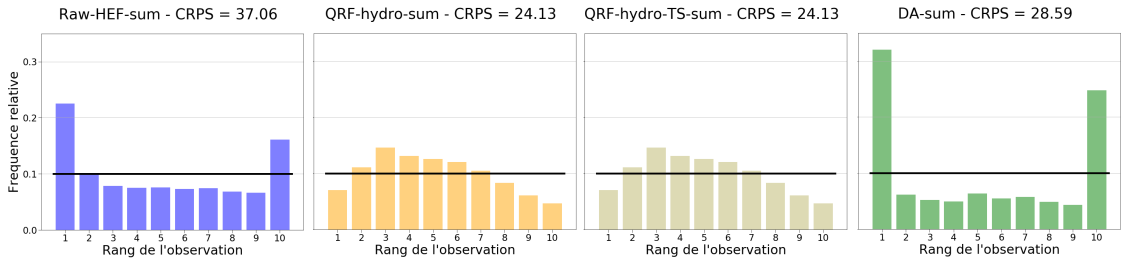
De même que lors de la mise en place de la calibration QRF présentée au chapitre 4, nous étudions à présent les performances de variables globales calculées à partir des prévisions d'ensemble par échéance :

- l'ensemble *sum* représente la somme des prévisions sur les 24 heures d'échéance ;

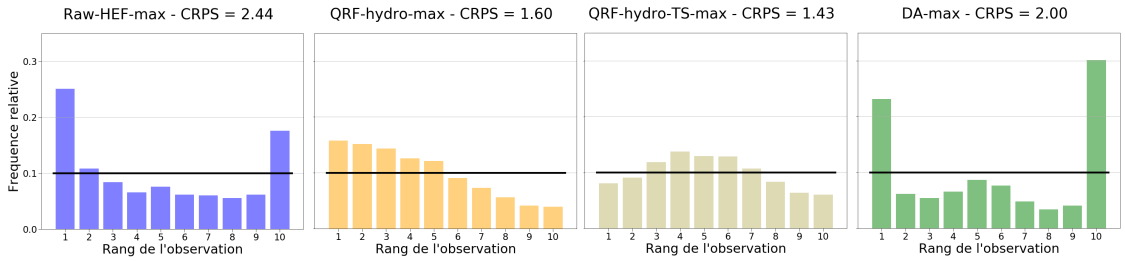
- l'ensemble *max* représente le maximum des prévisions sur les 24 heures d'échéances.

Ces deux variables sont directement liées aux notions de volume et de maximum d'hydrogramme [Hemri et al., 2015].

La figure 6.4 présente les histogrammes de rang pour la somme et le maximum des prévisions pour l'ensemble brut *raw-HEF*, les ensembles calibrés *QRF-hydro* et *QRF-hydro-TS*, et l'ensemble crée par assimilation de données *DA* à station Tréodet. La stratification est réalisée selon les valeurs de l'ensemble brut. Les résultats analogues pour les stations Kerjean et Ty-Planche sont respectivement présentés en annexes P.2 et P.3.



(a) Somme de la prévision



(b) Maximum de la prévision

Figure 6.4: Histogrammes de rang pour la somme et le maximum des prévisions pour les ensembles *raw*, *QRF-hydro*, *QRF-hydro-TS* et *Data Assimilation* à Tréodet.

Somme de la prévision Nous notons tout d'abord que les ensembles de la somme pour *QRF-hydro* et *QRF-hydro-TS* présentent les mêmes histogrammes de rang et des CRPS identiques. La procédure de lissage n'a donc pas d'impact sur le volume total prévu. Le lissage permet juste de répartir ce volume en évitant des marches non physiques entre les échéances de prévision. La calibration statistique améliore la forme des histogrammes de rang et le CRPS pour la somme des prévisions. En revanche, même si l'assimilation de données améliore le CRPS de la somme, l'histogramme de rang est dégradé. Il est davantage sous-dispersif. Ces caractéristiques se retrouvent pour les trois sous-bassins étudiés.

Maximum de la prévision Pour cette variable, les meilleurs CRPS sont obtenus avec l'ensemble *QRF-hydro-TS*. L'assimilation de données améliore le CRPS mais moins que la

calibration statistique. Les histogrammes de rang sont améliorés par rapport à l'ensemble brut avec la calibration statistique, mais ils ne sont pas plats. L'assimilation de données n'améliore pas la fiabilité de l'ensemble construit sur le maximum de la prévision.

6.2 Synthèse et perspectives

Dans ce chapitre, nous avons comparé les performances de l'ensemble brut *raw-HEF* avec celles des prévisions d'ensemble hydrologiques réalisées par calibration statistique et par assimilation de données dans un contexte opérationnel. Il ressort de l'analyse par échéance de prévision que les deux stratégies de construction de prévisions d'ensemble hydrologiques permettent d'améliorer l'erreur RMSE par rapport à une prévision déterministe. Cependant, la calibration statistique des ensembles hydrologiques paraît plus prometteuse que la mise en place d'une chaîne d'assimilation de données ensembliste telle qu'elle a été implémentée dans ce travail. Elle permet en effet d'obtenir des CRPS meilleurs, de manière plus ou moins prononcée selon les stations, mais surtout elle permet d'obtenir une meilleure fiabilité et une meilleure discrimination des événements de dépassement de seuil. L'analyse des ensembles *sum* et *max* aboutit aux mêmes conclusions. L'assimilation de données paraît ainsi moins performante pour la génération d'ensembles hydrologiques. Ceci est probablement en partie dû à l'équifinalité mise en évidence dans les expériences jumelles. Le modèle analysé simule des débits proches des mesures sur la fenêtre d'assimilation (voir paragraphe 5.3), cependant, il ne se trouve pas dans un état lui permettant de faire de bonnes prévisions. Il ne peut alors pas reproduire la dynamique hydrologique nécessaire pour générer de bonnes prévisions. Une densification du réseau d'observation permettrait de lever en partie ce problème d'équifinalité, comme nous l'avons vu lors de la réalisation des expériences jumelles.

D'un point de vue opérationnel, les deux méthodes de génération des prévisions d'ensemble sont comparables en termes de ressources nécessaires. La calibration statistique demande une forte puissance de calcul pour l'optimisation des forêts aléatoires, mais le calcul d'une calibration une fois la forêt construite ne nécessite pas des ressources informatiques importantes. Il est possible de réaliser cette étape sur un ordinateur personnel. L'assimilation de données est a priori un peu plus gourmande, mais il faut tout de même noter que le lancement de 49 membres avec le code MORDOR-TS est relativement rapide.

Dans la suite du travail, pour continuer à investiguer la chaîne de prévision hydrologique-hydraulique, nous conservons uniquement les prévisions d'ensemble brutes *raw-HEF* et les prévisions d'ensemble *QRF-hydro-TS* réalisées par calibration statistique sur lesquelles un lissage des trajectoires est appliqué pour obtenir des chroniques de débits ayant un sens physique. La prévision d'ensemble obtenue par assimilation de données *DA* est mise de côté car ses performances sont significativement inférieures à celles de l'ensemble *QRF-*

hydro-TS avec les choix d'assimilation mis en œuvre ici. Des pistes d'amélioration de ce travail résident dans l'optimisation de l'algorithme EnKF, en particulier les stratégies de rééchantillonnage, pour évaluer dans quelle mesure on peut améliorer la fiabilité des ensembles en prévision.

Des perspectives à ce travail émergent suite à la lecture de l'article Bourgin et al. [2014]. Les auteurs alimentent le modèle GRP au pas de temps horaire par les prévisions d'ensemble PEARP sur 202 bassins versants français. Ils étudient les apports de deux méthodes sur les performances des prévisions d'ensemble : (i) l'assimilation du débit mesuré au temps de base et la correction des prévisions à partir de l'erreur modèle au temps de base selon la méthodologie décrite dans Berthet [2010], Tangara [2005] (voir paragraphe 2.3.1.1) et (ii) un post-traitement basé sur les erreurs de calage du modèle hydrologique. L'assimilation de données consiste en la correction de l'état du réservoir de routage de GRP de telle sorte que le débit simulé soit égal au débit observé. Ils mettent en évidence que l'impact le plus important sur la fiabilité est dû au post-traitement. L'assimilation de données aide à améliorer la fiabilité dans une moindre mesure, surtout aux courtes échéances. Dans leur cas, ils montrent que les deux stratégies, assimilation de données et post-traitement, ont des effets bénéfiques complémentaires et recommandent l'utilisation simultanée de ces deux techniques. Les méthodes d'assimilation de données et de post-traitement qu'ils ont mises en place sont significativement différentes de celles employées dans ce travail. En particulier, nous ne corrigeons par les états du modèle mais uniquement ses paramètres. Le post-traitement n'est pas construit de la même manière. Nous notons des différences entre nos résultats et ceux de [Bourgin et al., 2014]. Dans notre travail, on note aussi une meilleure fiabilité aux courtes échéances avec l'assimilation de données. Cependant, aux plus longues échéances, la fiabilité est dégradée par rapport à l'ensemble sans assimilation. Le travail présenté dans cette thèse pourrait être complété en appliquant une calibration statistique à l'ensemble *DA* généré par assimilation de données. Il serait intéressant de voir si les performances finales seraient améliorées par rapport à l'ensemble *QRF-hydro-TS*. Cependant, nous émettons des réserves quant à la fiabilité de l'ensemble obtenu, puisque l'ensemble *DA* présente une nettement moins bonne fiabilité que celle de l'ensemble brut *raw-HEF*.

Partie III

LES PRÉVISIONS D'ENSEMBLE HYDRAULIQUES

Aide-mémoire sur les ensembles hydrauliques

Cette partie a pour objectif l'application des méthodes développées au chapitre 3 pour la construction de prévisions d'ensemble hydrauliques.

Dans la partie II (chapitres 4 à 6), nous avons mis en œuvre et comparé des méthodes pour la génération de prévisions d'ensemble hydrologiques. Nous avons décidé de mettre de côté l'ensemble généré par assimilation de données car il ne présente pas une bonne fiabilité dans le cadre des expériences menées ici. Pour la suite de l'exploration de la chaîne de prévision, nous conservons deux ensembles :

- l'ensemble brut *raw-HEF* généré par perturbation des paramètres de calage significatifs pour les débits simulés avec le modèle MORDOR-TS ;
- l'ensemble *QRF-hydro-TS* généré par calibration statistique QRF puis par l'étape de reconstruction et de lissage des trajectoires.

Ces deux ensembles sont constitués de 99 membres.

La génération et l'étude des performances des prévisions d'ensemble hydrauliques est réalisée selon la même méthodologie que pour les prévisions d'ensemble hydrologiques. Le chapitre 7 présente l'étude de sensibilité globale menée sur le modèle MASCARET du bassin de l'Odet. Les résultats de cette étude permettent ensuite de générer des ensembles bruts sur lesquels des calibrations statistiques *Quantile Regression Forests* sont appliquées. C'est l'objet du chapitre 8. Le chapitre 9 est consacré à la mise en place d'assimilation de données ensembliste dans le modèle hydraulique. On y décrira en particulier les expériences jumelles qui permettent de comprendre et de valider l'algorithme d'assimilation développé, puis les résultats des expériences réelles sur la fenêtre d'assimilation. Enfin, le chapitre 10 est consacré à la comparaison des deux approches développées pour générer des prévisions d'ensemble dans un contexte opérationnel : la calibration statistique des ensembles bruts et l'assimilation de données ensembliste.

L'intégralité des ensembles construits dans cette partie III est résumée dans la partie introductive du chapitre 10.

Chapitre 7

Étude d'incertitude du modèle MASCARET de l'Odet

Sommaire

7.1 Cadre d'étude 189

7.1.1 Espace incertain pour l'étude de sensibilité globale 189

7.1.2 Quatre événements de typologie différente 192

7.2 Mise en œuvre des expériences 193

7.2.1 Expérience S_{min} : perturbation des Stricklers en lits mineurs des 12 zones du modèle 195

7.2.2 Expérience $S_{min}S_{maj}$: perturbation des Stricklers en lits mineurs et majeurs des 6 zones influentes du modèle 197

7.2.3 $S_{min}S_{maj}CL$: perturbation des Strickler en lits mineurs et majeurs des 6 zones influentes du modèle et des conditions aux limites 200

7.3 Synthèse 202

Le modèle hydraulique MASCARET de l'Odet a pour objectif la modélisation des hauteurs d'eau en particulier aux stations de vigilance Kervir, Moulin-Vert et Justice situées dans le centre-ville de Quimper. La figure 7.1 rappelle la configuration des biefs modélisés par le modèle MASCARET et la localisation des stations de vigilance. Elle présente également les 12 zones de frottement du modèle.

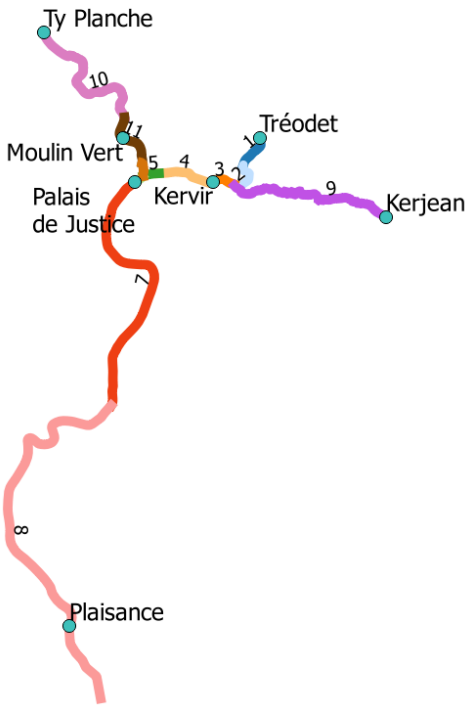


Figure 7.1: Bassin versant aval de l'Odet modélisé avec le modèle MASCARET. Les trois stations de vigilance sont Kervir, Moulin-Vert et Justice.

Ce chapitre décrit l'étude de sensibilité globale du modèle MASCARET du bassin de l'Odet. Elle a fait l'objet d'une communication au colloque *Telemac User Club* en 2019 et un article a été rédigé[1] [Tiberi-Wadier et al., 2019]. La partie 7.1 décrit le cadre général de l'étude menée. Dans un premier temps, la caractérisation des variables incertaines pour la réalisation de l'étude d'incertitude est réalisée. Puis, on y décrit quatre événements de typologie différente pour lesquels on analysera les résultats de l'étude d'incertitude : fluvial, fluvio-maritime, maritime et pas d'événement particulier. La partie 7.2 présente ensuite trois expériences différentes réalisées pour l'étude d'incertitude et les résultats associés. Enfin, les conclusions sont exposées en partie 7.3.

[1] *Sensitivity Analysis of the MASCARET model on the Odet River* disponible sur https://hal.archives-ouvertes.fr/hal-03250128.

7.1 Cadre d'étude

7.1.1 Espace incertain pour l'étude de sensibilité globale

La quantité d'intérêt sur laquelle porte l'étude d'incertitude est la hauteur d'eau simulée par le modèle MASCARET aux stations de vigilance (Kervir, Moulin-Vert ou Justice) à un instant donné. L'étude de sensibilité globale est ainsi réalisée temporellement aux trois stations d'intérêt. Les variables d'entrées incertaines considérées, qui résultent de l'analyse hydraulique du modèle, sont les suivantes :

- la valeur des coefficients de Strickler des lits mineur et majeur des 12 zones de frottement du modèle,
- les débits hydrologiques imposés à l'amont,
- la condition limite aval.

Les conditions initiales et la bathymétrie sont considérées comme déterministes.

Coefficients de Strickler Chaque coefficient de Strickler mineur ou majeur est une variable scalaire dont l'incertitude est décrite par une densité de probabilité. Celle-ci est considérée uniforme. La distribution est centrée sur la valeur calée, et les bornes sont à +/- 5 de la valeur calée. La plage de variation des coefficients de Strickler n'est pas très large par rapport aux valeurs calées car les débits qui ont servi au calage sont représentatifs de ceux que l'on souhaite simuler (débits relativement soutenus pendant l'hiver). Le tableau 7.1 présente les bornes utilisées pour chaque zone. Il est à noter que les bornes ont été réduites pour le tirage de coefficient de Strickler en lit mineur de la zone 5 pour assurer une meilleure stabilité numérique du code de calcul.

Zone	Ks_{min} mineur	Ks_{max} mineur	Ks_{min} majeur	Ks_{max} majeur
1	25	35	15	25
2	15	25	5	15
3	10	20	5	15
4	33	43	8.5	18.5
5	13	17	5	15
6	31.5	41.5	29	39
7	32	42	20	30
8	20.5	30.5	1	10
9	15	25	5	15
10	15	25	5	15
11	27.5	37.5	1	10
12	17.5	22.5	5	15

Tableau 7.1: Bornes minimales et maximales pour les coefficients de Strickler en lits mineur et majeur.

Condition limite amont La prise en compte de l'incertitude sur la condition limite amont est réalisée grâce aux prévisions d'ensemble hydrologiques brutes *raw-HEF* dont la génération est présentée au chapitre 4. Ces prévisions d'ensemble, qui comportent 99 membres, ont été générées par perturbation des 8 paramètres de MORDOR-TS pour lesquels l'étude d'incertitude a montré qu'ils avaient de l'influence sur les débits hydrologiques simulés. Elles sont générées dans un contexte opérationnel de manière cyclée et sont alimentées en prévision par les pluies RR3. Pour prendre en compte l'incertitude sur les débits hydrologiques en entrée du modèle, un tirage dans une loi uniforme discrète de bornes [1, 99] est ainsi réalisé.

Condition limite aval Nous n'avons pas à disposition de prévisions d'ensemble de niveaux d'eau à Plaisance qui permettraient de tenir compte l'incertitude sur la condition limite aval de la même manière qu'à l'amont. Nous avons alors créé des chroniques perturbées de la mesure à Plaisance en générant des chroniques de perturbation de la surcote mesurée s. Cette méthodologie est décrite dans Laborie [2020]. Considérer que les erreurs sur la surcote à Plaisance pour deux pas de temps proches sont des variables aléatoires indépendantes ne paraît pas légitime. La perturbation doit donc permettre de conserver la corrélation temporelle des erreurs. Pour cela, la surcote est supposée être la réalisation d'un processus gaussien.

Un processus gaussien est un processus aléatoire[2] tel que chaque combinaison linéaire finie de ses lois suit une loi gaussienne. Dans notre cas, les variables aléatoires sont indexées par le temps. Un processus gaussien est entièrement déterminé par sa moyenne et sa fonction de covariance. Puisque nous appliquons une perturbation sur la valeur de la surcote mesurée, nous supposons que le processus gaussien est de moyenne nulle. La fonction de covariance est définie par sa variance σ^2 et sa fonction de corrélation r telle que $cov(s(t), s(t')) = \sigma^2 r(t, t')$. Nous choisissons la fonction de corrélation exponentielle pour r. Les valeurs de σ et des hyper-paramètres de la fonction r sont estimés à dire d'expert, à partir des considérations suivantes :

(i) La longueur de corrélation est arbitrairement fixée à 6 heures, ce qui correspond approximativement à la durée du flot ou du jusant.

(ii) L'amplitude de la fonction de covariance est choisie de manière à ce que l'écart-type des perturbations soit du même ordre de grandeur que l'écart-type des surcotes observées. Les simulations perturbées sont réalisées sur trois grandes périodes P_{2012}, P_{2013} et P_{2014}. Le tableau 7.2 caractérise ces périodes, ainsi que les écart-types de la mesure de hauteur d'eau à Plaisance et les écart-types médians de l'ensemble généré par perturbation de la mesure pour chaque période.

[2] collection de variables aléatoires indexées par le temps ou l'espace

Période	P_{2012}	P_{2013}	P_{2014}
Dates	du 10/12/2012 au 10/01/2013	du 10/12/2013 au 10/01/2014	du 26/01/2014 au 18/02/2014
Écart-type de la mesure	16.1 cm	18.0 cm	18.5 cm
Écart-type médian de l'ensemble	16.0 cm	16.1 cm	17.8 cm

Tableau 7.2: Caractéristiques de l'ensemble issu de la perturbation de la surcote mesurée pour les 3 périodes simulées.

Une fois que le processus gaussien est correctement défini, il s'agit à présent d'en générer plusieurs réalisations. Pour cela, la surcote s est écrite selon la décomposition de Karhunen-Loève comme la forme tronquée de n_p fonctions orthogonales ϕ_i , pour lesquelles les coefficients de chaque mode ϵ_i sont des variables normales indépendantes :

$$s(t) = \sum_{i=0}^{n_p} \epsilon_i \sqrt{\lambda_i} \phi_i(t). \tag{7.1}$$

Nous faisons le choix de conserver un nombre de modes permettant d'expliquer 99% de la variance du signal. Plus de 100 modes ont ainsi été conservés. Il est à noter que l'objectif n'est pas ici de réduire la dimension de ce nouvel espace. λ_i and ϕ_i sont respectivement les valeurs propres et les fonctions propres de la matrice de corrélation R associée à la fonction de corrélation r du processus gaussien.

Un ensemble de 99 chroniques de surcotes perturbées est généré avec un échantillonnage de ϵ_i. Pour prendre en compte l'incertitude sur la condition limite aval, un tirage dans une loi uniforme discrète de bornes [1, 99] est ainsi réalisé. La figure 7.2 illustre la perturbation de la condition limite aval pour 10 membres sur la journée du 4 février 2014.

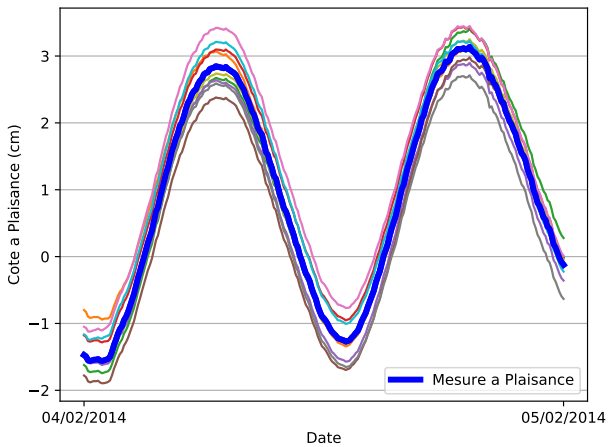


Figure 7.2: Illustration de la perturbation de la condition limite aval par rapport à la cote mesurée. Les spaghettis de couleur représentent 10 membres.

Comparaison de l'écart-type des perturbations des conditions limite amont et aval La figure 7.3 présente l'écart-type des perturbations des conditions limites amont et aval sur une sous-période de P_{2013} (du 15 décembre 2013 au 10 janvier 2014). Sur cette figure, les débits amont ont été convertis en hauteurs d'eau grâce aux courbes de tarage aux stations. Cette figure confirme que la condition limite amont est davantage dispersée lors des événements pluvieux. En effet, les pics visibles pour les trois stations amont sur ce graphique correspondent à des pics de débit. Ceci était déjà mis en évidence lors du tracé des chroniques de densité de probabilité des ensembles hydrologiques [Tiberi-Wadier et al., 2021]. Par ailleurs, ce graphique montre que l'amplitude en hauteur d'eau des perturbations imposées à l'amont et à l'aval sont du même ordre de grandeur.

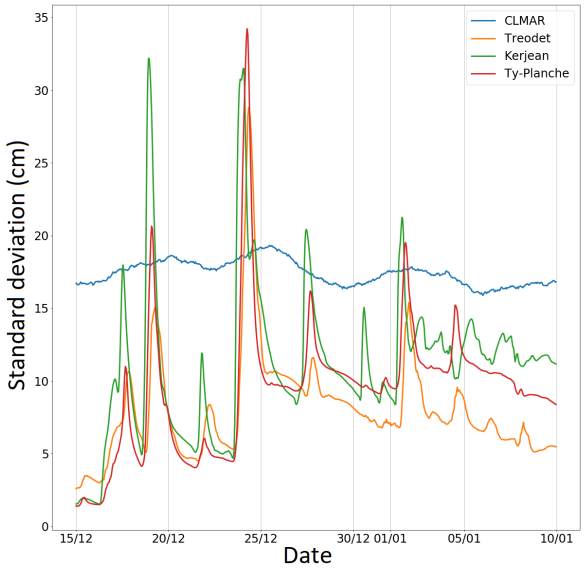


Figure 7.3: Écart-type de la perturbation des conditions limite amont et aval en cm sur une sous-période de P_{2013}. CLMAR, pour *Condition Limite MARitime*, correspond à l'écart-type de la condition limite aval.

7.1.2 Quatre événements de typologie différente

L'analyse de sensibilité est menée aux trois stations Kervir, Moulin-Vert et Justice au cours de quatre événements de typologie différentes : événement fluvial, fluvio-maritime, maritime et pas d'événement du tout. Cela permet de mieux appréhender les influences de chacun des paramètres incertains selon les types d'événements. Les quatre événements sont décrits dans le tableau 7.3 : dates, valeur des pics de débit à l'amont et des pics de surcote/décote à Plaisance. La dernière ligne du tableau indique pour le modèle calé le pourcentage de débit passant dans le lit majeur au niveau des biefs contenant chacune

des stations Kervir et Moulin-Vert sur une fenêtre de 2 heures autour du pic de débit mesuré. Pour chaque station, la portion de bief servant au calcul est constituée des profils dont le pk varie entre -200 m et +350 m par rapport à la station. Il est à noter qu'un pourcentage de débit nul ou très faible n'implique pas forcément que l'indice de Sobol de la zone concernée sera nul, puisque la perturbation des paramètres peut activer le lit majeur là où le modèle calé ne le fait pas.

Au cours de l'événement EVT_2 ont été mesurés les deuxièmes plus forts débits aux trois stations amont, avec une période de retour supérieure à 50 ans à Tréodet, et de l'ordre de 10 ans à Ty-Planche et Kerjean. C'est là qu'on note les débordements les plus importants à Kervir dans le modèle. L'événement EVT_3 est fortement maritime, avec un pic de surcote à Plaisance de 89 cm. Les débordements sont négligeables aux deux stations. L'événement EVT_4, pour lequel les débits sont faibles, ne génère naturellement aucun débordement dans le modèle.

Nom de l'événement	EVT_1	EVT_2	EVT_3	EVT_4
Dates	19-22/12/2012	23-26/12/2013	4-6/02/2014	10/12/2013
Pic de surcote/décote à Plaisance	23 cm	67 cm	89 cm	-18 cm
Pic de débit à Tréodet	69.7 m^3 s^{-1}	91.5 m^3 s^{-1}	40.5 m^3 s^{-1}	3.17 m^3 s^{-1}
Pic de débit à Kerjean	18.2 m^3 s^{-1}	17.6 m^3 s^{-1}	14.8 m^3 s^{-1}	1.22 m^3 s^{-1}
Pic de débit à Ty-Planche	44.0 m^3 s^{-1}	42.7 m^3 s^{-1}	37.4 m^3 s^{-1}	2.59 m^3 s^{-1}
Type d'événement	Fluvial	Fluvio-maritime	Maritime	Pas d'événement
% Q majeur Kervir	2.5%	5.8%	0.38%	0.0%
% Q majeur Moulin-Vert	0.09%	0.07%	0.01%	0.0%

Tableau 7.3: Synthèse des événements étudiés lors de l'étude d'incertitude globale.

7.2 Mise en œuvre des expériences

Nous rappelons ici que le nombre de simulations a été déterminé pour assurer la convergence des indices de Sobol calculés (voir chapitre 3.2.3). Dans cette partie sont analysées trois expériences, de complexité croissante :

- expérience S_{min} : cette expérience consiste à perturber uniquement les 12 coefficients de Strickler en lit mineur du modèle. Elle permet de mettre en évidence les zones de frottement qui contrôlent le niveau simulé aux trois stations de prévision. L'expérience possède ainsi 12 paramètres incertains, et 20800 simulations perturbées ont été lancées.

- expérience $S_{min}S_{maj}$: cette expérience consiste à considérer dorénavant comme incertaines uniquement les zones qui contrôlent le niveau simulé aux stations déter-

minées grâce à l'expérience précédente S_{min}. Dans ces zones, les coefficients de Strickler mineurs et majeurs sont maintenant perturbés. Les conditions aux limites amont et aval ne sont pas perturbées. L'expérience possède 12 paramètres incertains (6 coefficients de Strickler en lit mineur et 6 en lit majeur) et 20800 simulations perturbées ont été lancées.

- expérience $S_{min}S_{maj}CL$: cette expérience consiste à reprendre la configuration de l'expérience précédente $S_{min}S_{maj}$, et à y ajouter la perturbation des conditions aux limites amont et aval. L'expérience possède alors 14 paramètres incertains : les 12 de l'expérience précédente, l'indice du membre hydrologique et l'indice du membre pour la condition limite aval perturbée. 48000 simulations perturbées ont été lancées.

Le schéma figure 7.4 présente la mise en œuvre des simulations pour l'étude d'incertitude. Un membre perturbé correspond au tirage des coefficients de Strickler, et, le cas échéant pour l'expérience $S_{min}S_{maj}CL$, des indices des membres utilisés pour les conditions limites amont et aval. Pour chaque membre, le modèle dont les coefficients de frottement sont perturbés est lancé sur une période d'initialisation de 3 jours : il est alimenté par les mesures à l'amont et à l'aval. Après la période d'initialisation, la mise en œuvre est alors différente selon les expériences : (i) pour les expériences S_{min} et $S_{min}S_{maj}$ pour lesquelles les conditions limites ne sont pas perturbées, le modèle est toujours alimenté par les mesures à l'amont et à l'aval, et (ii) pour l'expérience $S_{min}S_{maj}CL$, le modèle est alimenté à l'amont par les prévisions hydrologiques cyclées à 1 heure, et à l'aval par la chronique de hauteurs d'eau perturbée à Plaisance. Nous n'avons pas mis en place ici des simulations cyclées comme cela a été fait pour l'étude d'incertitude hydrologique pour des considérations de temps de calcul.

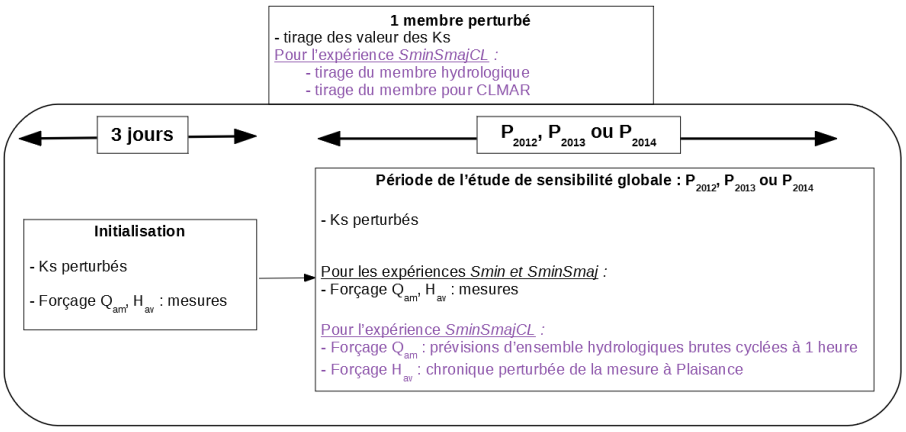


Figure 7.4: Schéma de principe des simulations non cyclées mises en place pour l'étude de sensibilité globale du modèle MASCARET de l'Odet. Le forçage sur la période P_{2012}, P_{2013} ou P_{2014} est différent selon que l'on perturbe les conditions aux limites ou pas.

Pour les trois expériences, nous présentons les résultats du calcul des indices de sensibilité du premier ordre. Il est à noter que ceux-ci sont très proches des indices de sensibilité totaux, ce qui permet de conclure qu'il n'y a pas d'interactions notable entre les différents facteurs d'entrée. L'étude de la hiérarchisation des indices de Sobol du premier ordre permettra donc de fixer les paramètres incertains n'ayant pas d'influence sur les sorties étudiées (*Factor Fixing*, voir paragraphe 3.2.2.2).

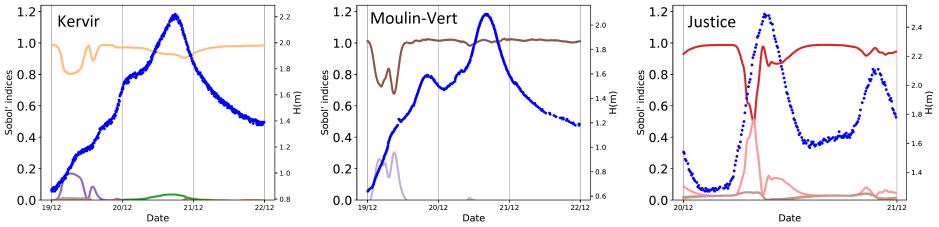
7.2.1 Expérience S_{min} : perturbation des Stricklers en lits mineurs des 12 zones du modèle

La figure 7.5 représente les résultats du calcul des indices de Sobol au cours des quatre événements étudiés pour les trois stations de vigilance. La figure 7.5e représente les 12 zones perturbées. La couleur de chaque zone est la même que celle de la chronique temporelle pour l'indice de Sobol correspondant représentée sur les graphiques 7.5a à 7.5d (axe des ordonnées de gauche). La hauteur d'eau mesurée à la station est également tracée sur ces graphiques (axe des ordonnées de droite).

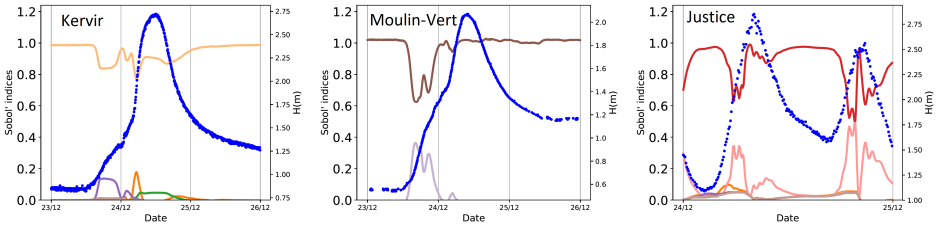
À Kervir, l'analyse des Sobols met en évidence que la hauteur d'eau simulée est contrôlée majoritairement par la valeur du coefficient de Strickler de la zone 4 située immédiatement à l'aval pour tous les types d'événements. En l'absence d'événement (EVT_4) ou pour un événement essentiellement maritime (EVT_3), c'est alors uniquement la valeur du Strickler de la zone 4 qui est importante. En présence d'un événement de type fluvial (EVT_1) ou fluvio-maritime (EVT_2), les coefficients de Strickler des zones 3 et 9 situées à l'amont de la station voient leur influence légèrement augmenter en montée de crue. La valeur de ces coefficients contrôle en partie la dynamique de la crue et sa vitesse de propagation vers l'aval de la station. Nous pouvons noter que le coefficient de Strickler de la zone 5 a une influence très faible au moment du pic de crue, mais cette zone a été moins perturbée pour des raisons de stabilité du modèle numérique.

À Moulin-Vert, les chroniques de coefficients de Sobols montrent que la hauteur d'eau simulée est majoritairement contrôlée par la valeur du coefficient de Strickler de la zone 11 située immédiatement à l'aval de la station. En l'absence d'événement (EVT_4), c'est uniquement la valeur du Strickler de cette zone qui est importante. En présence d'un événement pluvieux (EVT_1, EVT_2 et EVT_3), la valeur du coefficient de Strickler de la zone 10 située immédiatement à l'amont prend un peu d'importance lors de la montée de l'hydrogramme.

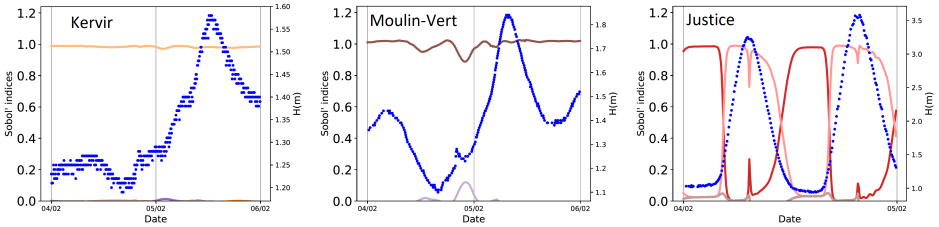
L'analyse des résultats à Kervir et Moulin-Vert met donc en évidence des comportements analogues : (i) la hauteur d'eau simulée est majoritairement contrôlée par le coefficient de Strickler de la zone située immédiatement à l'aval de la station, et (ii) les coefficients de Strickler des zones situées à l'amont des stations influencent légèrement la hauteur d'eau simulée lors des montées d'hydrogrammes.



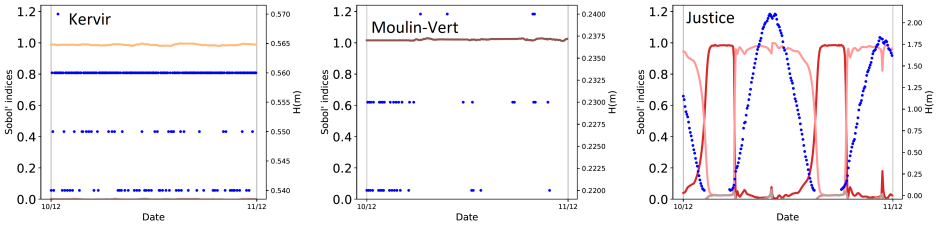
(a) EVT_1 : événement fluvial



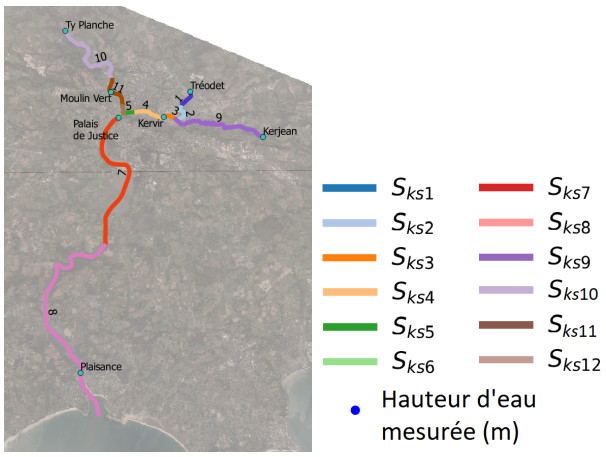
(b) EVT_2 : événement fluvio-maritime



(c) EVT_3 : événement fortement maritime



(d) EVT_4 : pas d'événement



(e) Zones perturbées pour l'expérience S_{min}

Figure 7.5: Chroniques temporelles des indices de Sobol pour les stations Kervir, Moulin-Vert et Justice, pour les 4 événements étudiés - Expérience S_{min}.

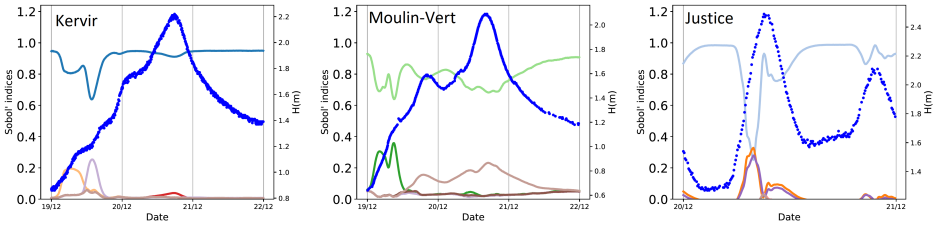
Dans cette expérience pour laquelle les conditions aux limites ne sont pas perturbées, le niveau simulé à Justice est quasiment uniquement contrôlé par le choix de la valeur des coefficients de Strickler des zones 7 et 8 (il est de manière marginale influencé par le coefficient de Strickler de la zone 12). Leur influence est cyclique et leur importance respective dépend largement du type d'événement considéré. En l'absence d'événement fluvial important (EVT_3 et EVT_4), la hauteur d'eau à marée haute est uniquement contrôlée par la valeur du coefficient de Strickler de la zone 8, qui est la zone située le plus à l'aval, au niveau de l'estuaire. À marée basse en revanche, c'est le coefficient de Strickler de la zone 7, zone immédiatement à l'aval de la station, qui contrôle la hauteur d'eau simulée à Justice. En revanche, lorsqu'il y a un événement fluvial (EVT_1 et EVT_2), les niveaux sont globalement plus hauts. La nature cyclique des chroniques des indices de Sobols est la même que précédemment, mais la zone 8 perd de son importance au profit de la zone 7. Si l'on se focalise sur les pics de marée haute, nous avons ainsi mis en évidence que l'écoulement du débit amont lors d'un événement fluvial est lié à la valeur du coefficient de Strickler de la zone 7. De manière symétrique, en l'absence d'événement fluvial, la remontée du débit aval lié au flot est contrôlée par le coefficient de Strickler de la zone 8.

Grâce à cette première expérience, nous avons mis en évidence les zones pour lesquelles la valeur du coefficient de Strickler influence de manière significative les hauteurs d'eau simulées aux trois stations de vigilance. Il s'agit des zones 4 et 9 pour Kervir, 10 et 11 pour Moulin-Vert et 7 et 8 pour Justice. Les zones 3 et 5 ont une très petite influence à Kervir : l'indice de Sobol est inférieur à 0.1, l'influence est mise en évidence uniquement pour les événements 1 et 2, et l'influence est très limitée dans le temps. Elles ne seront plus perturbées par la suite pour des raisons de stabilité du modèle numérique. Désormais, seules les 6 zones définies ici seront perturbées dans la suite de l'étude.

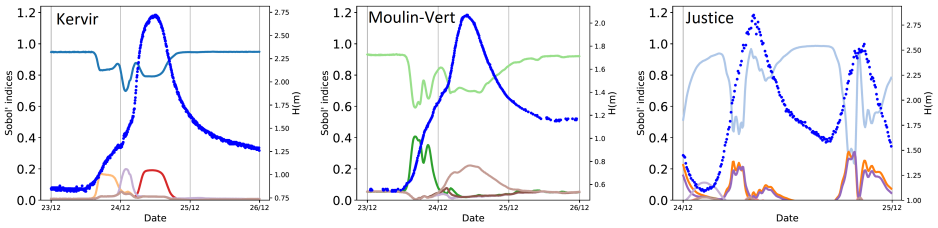
7.2.2 Expérience $S_{min}S_{maj}$: perturbation des Stricklers en lits mineurs et majeurs des 6 zones influentes du modèle

La figure 7.6 est analogue à la figure 7.5 pour l'expérience $S_{min}S_{maj}$ dans laquelle les paramètres incertains sont les coefficients de Strickler en lits mineurs et majeurs des zones 4, 7, 8, 9, 10, et 11. Par rapport à l'expérience précédente, cette expérience permet de mettre en évidence l'influence éventuelle des coefficients de Strickler en lit majeur dans les zones qui influencent les hauteurs d'eau simulées aux trois stations.

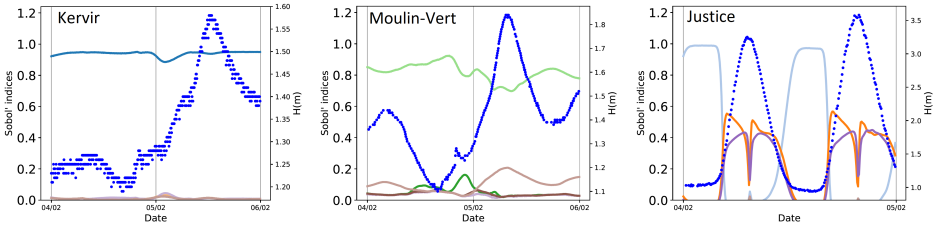
À Kervir, les niveaux sont naturellement toujours contrôlés par la valeur des coefficients de Strickler des zones 4 et 9. La valeur du coefficient de Strickler en lit mineur de la zone 4 est toujours prépondérant. En l'absence d'événement fluvial majeur (EVT_3 et EVT_4), c'est uniquement la valeur de ce coefficient qui a de l'importance. Le coefficient de Strickler en lit majeur de la zone 4 est activé uniquement au moment du pic de crue pour l'événement EVT_2 qui a créé des débordements. La zone 9 située à l'amont présente toujours une faible



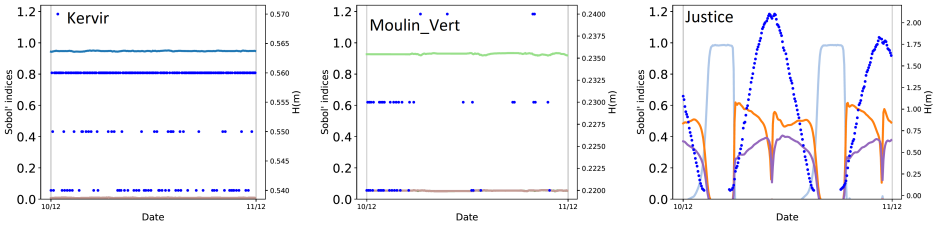
(a) EVT_1 : événement fluvial



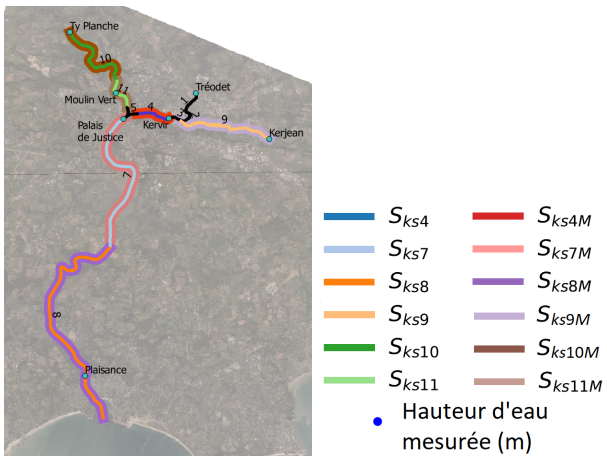
(b) EVT_2 : événement fluvio-maritime



(c) EVT_3 : événement fortement maritime



(d) EVT_4 : pas d'événement



(e) Zones perturbées pour l'expérience $S_{min}S_{maj}$

Figure 7.6: Chroniques temporelles des indices de Sobol pour les stations Kervir, Moulin-Vert et Justice, pour les 4 événements étudiés - Expérience $S_{min}S_{maj}$.

influence pour les événements ayant une composante fluviale importante : les indices de Sobol sont relativement faibles et l'influence est très limitée dans le temps (début de montée de crue uniquement). L'influence, limitée également, du coefficient de Strickler du lit majeur de la zone 9 arrive juste après celle du lit mineur.

À Moulin-Vert, on note toujours un contrôle majoritaire de la zone 11 située immédiatement à l'aval de la station. Le coefficient de Strickler en lit majeur correspondant est mobilisé au moment des pics dans les hydrogrammes pour les événement EVT_1, EVT_2 et EVT_3. Nous avons vu précédemment dans le tableau 7.3 que ces événements n'entraînent pas un débordement significatif pour le modèle calé. La perturbation des paramètres entraîne ici une activation du lit majeur qui n'était que peu actif avec le modèle calé. La zone 10, à l'amont de Moulin-Vert, présente toujours une petite influence en montée de crue, puisque le coefficient de Strickler mineur de cette zone est mobilisé lors de la montée de crue. Cela est moins marqué pour l'événement d'origine maritime. L'influence du coefficient en lit majeur en zone 10 est en revanche négligeable. En l'absence d'événement, c'est quasiment uniquement la valeur du coefficient de Strickler en lit mineur de la zone 11 qui influence la hauteur d'eau simulée à Moulin-Vert. L'indice de Sobol du coefficient de Strickler en lit majeur de cette zone présente une influence négligeable. Il est surprenant que son influence ne soit pas nulle. Ceci met probablement en évidence un défaut dans la définition des lits mineurs et majeurs au niveau de certaines sections du modèle dans la zone 11. Pour autant, cela ne remet pas en cause la suite du travail.

Les résultats de l'expérience $S_{min}S_{maj}$ mettent toujours en évidence des résultats analogues aux stations Kervir et Moulin-Vert, qui sont toutes les deux des stations à dominance fluviale : (i) c'est la zone de frottement située immédiatement à l'aval des stations qui contrôle majoritairement la hauteur d'eau à la station, (ii) le lit majeur dans cette zone est activé au moment des pics d'événements, (iii) le lit mineur de la zone située immédiatement à l'amont a une petite influence très limitée temporellement et (iv) le lit majeur de la zone amont peut être légèrement activé ou pas selon l'événement.

À Justice, on retrouve comme dans l'expérience précédente S_{min} un comportement cyclique de l'influence des coefficients de Strickler des deux zones aval lié à la phase de la marée. La zone 7 reste prépondérante en cas d'événement fluvial (EVT_1 et EVT_2) pendant quasiment tout le cycle de marée. Le lit majeur de la zone 7 étant quasiment inexistant dans le modèle, il est normal qu'il ne soit pas activé ici. Les lits mineur et majeur de la zone 8 (zone estuarienne) sont activés en même temps, lorsque le gradient de hauteur d'eau à Justice est important (phases de flot et de jusant). Dans le modèle, le lit majeur zone 8 correspond à la zone de marnage. Il est activé à chaque cycle de marée. Les résultats sur la sensibilité aux valeurs des coefficients de Strickler obtenus ici sont donc cohérents avec la géométrie du modèle.

7.2.3 $S_{min}S_{maj}CL$: perturbation des Strickler en lits mineurs et majeurs des 6 zones influentes du modèle et des conditions aux limites

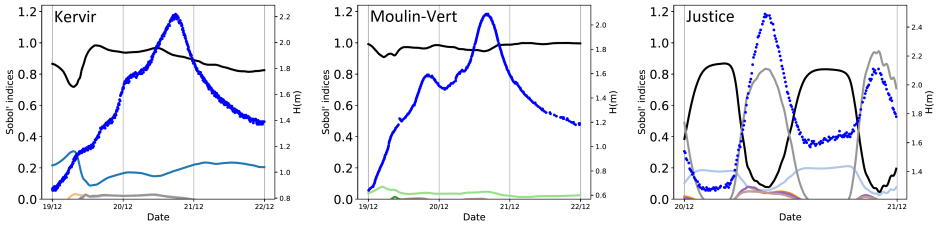
La figure 7.7 est analogue aux figures 7.5 et 7.6 pour l'expérience $S_{min}S_{maj}CL$. Dans cette expérience, les paramètres incertains sont toujours les coefficients de Strickler en lits mineurs et majeurs des zones 4, 7, 8, 9, 10, et 11. On y ajoute les conditions aux limites amont et aval.

À la station Kervir, on note une prépondérance de l'influence du choix du débit amont pour les quatre événements étudiés. La valeur du coefficient de Strickler mineur de la zone 4 conserve cependant une influence significative pour les quatre événements. Pour l'événement EVT_3 majoritairement maritime, l'importance du choix du débit amont est moins marquée que pour les autre événements (fluviaux ou pas d'événement). Le choix du coefficient de frottement de la zone 4 a alors plus d'influence. Nous notons que les coefficients de Strickler en lit majeur n'ont plus d'importance dans cette expérience puisque les coefficients de Sobol sont quasiment nuls. De manière analogue à Moulin-Vert, les graphiques mettent en évidence que la hauteur d'eau simulée est majoritairement déterminée par le choix de la condition limite amont. Cette influence est d'autant plus importante que l'événement a une forte composante fluviale. Les coefficients de Strickler en lit majeur n'ont là aussi plus d'importance pour le niveau simulé, et après le choix du débit amont, seul le coefficient de Strickler en lit mineur de la zone 11 influence le niveau d'eau simulé. À Kervir et à Moulin-Vert, le choix de la condition limite aval est négligeable quel que soit l'événement. Ceci n'est pas complètement fidèle à la réalité puisque lors des forts événements maritimes, les pics de marée haute peuvent générer un léger pic à Moulin-Vert. Ce léger pic est visible sur les graphiques 7.7a et 7.7c (événements EVT_1 et EVT_3) à Moulin-Vert en montée de crue. La modélisation n'est donc pas complètement représentative de la réalité, mais nous ne pouvons pas estimer ici si c'est le modèle lui-même qui est en cause ou seulement l'amplitude de la perturbation de la condition limite aval.

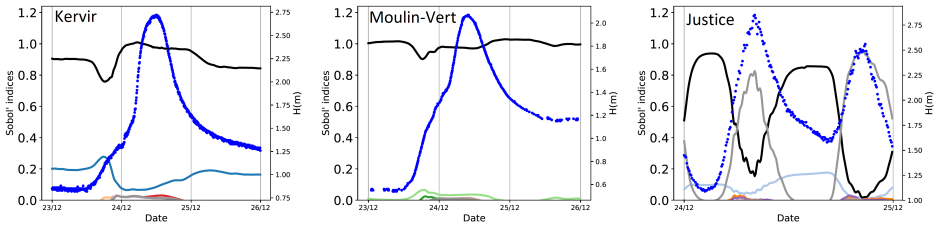
À Justice, nous notons là aussi que l'influence des conditions aux limites est prépondérante par rapport aux valeurs des coefficients de Strickler. Le comportement cyclique est toujours présent. Les événements qui n'ont pas de composante fluviale importante (EVT_3 et EVT_4) présentent un indice de Sobol pour la condition limite aval à marée haute plus important. On note encore une importance non négligeable de la valeur du coefficient de Strickler en lit mineur de la zone 7 pour les quatre événements, ainsi qu'une importante moindre mais pour autant non négligeable de la valeur des coefficients de Strickler en lits mineur et majeur de la zone 8 lorsque le gradient de hauteur à Justice est important (flot et jusant).

Pour conclure, les paramètres déterminants pour la modélisation des hauteurs d'eau aux trois stations de vigilance Kervir, Moulin-Vert et Justice pour l'expérience $S_{min}S_{maj}CL$ sont :

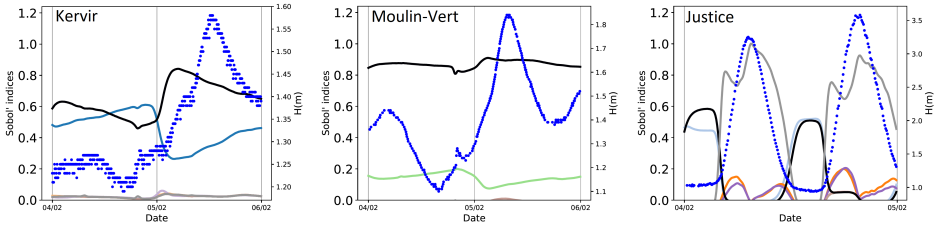
- les conditions aux limites amont (pour Kervir et Moulin-Vert) et aval (pour Justice)



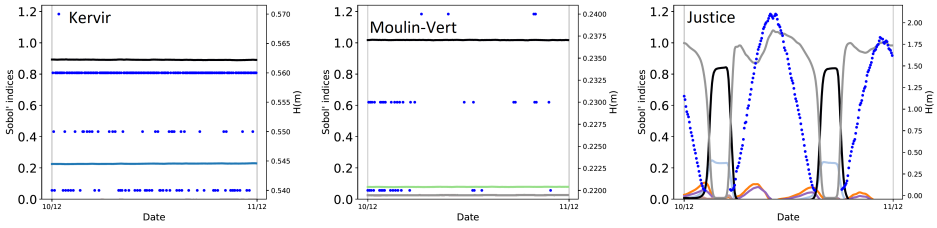
(a) EVT_1 : événement fluvial



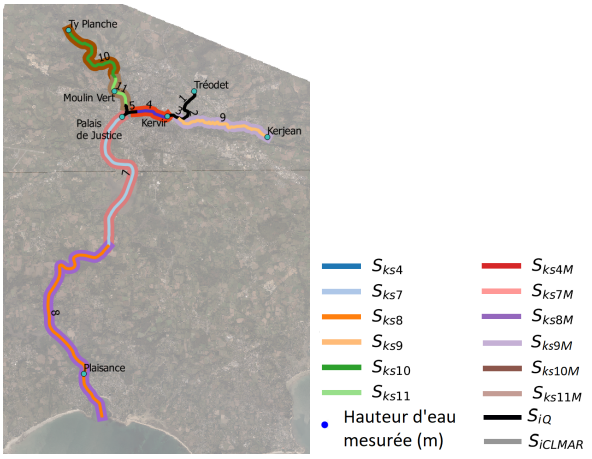
(b) EVT_2 : événement fluvio-maritime



(c) EVT_3 : événement maritime



(d) EVT_4 : pas d'événement



(e) Zones perturbées pour l'expérience $S_{min}S_{maj}CL$

Figure 7.7: Chroniques temporelles des indices de Sobol pour les stations Kervir, Moulin-Vert et Justice, pour les 4 événements étudiés - Expérience $S_{min}S_{maj}CL$.

imposées, quels que soient les types d'événements ;
- la valeur des coefficients de Strickler des mineurs des zones 4, 11, 7 et 8. Les zones 4 et 11 sont plus influentes lors d'un événement essentiellement maritime ;
- la valeur du coefficient de Strickler du lit majeur de la zone 8.

7.3 Synthèse

Dans ce chapitre, nous avons mené une étude de sensibilité globale du modèle MASCARET de l'Odet. Les indices de sensibilité ont été calculés de manière temporelle sur quatre types d'événements : fluvial, fluvio-maritime, maritime et pas d'événement. Il a été mis en évidence que la sensibilité du modèle aux différents paramètres d'entrée dépend du type d'événement considéré. L'étude de quatre types d'événements permet d'être le plus exhaustif possible pour l'évaluation et la hiérarchisation des paramètres influents. Trois expériences de complexité croissante ont été menées.

Grâce à l'expérience S_{min} dans laquelle les 12 zones de frottement mineur du modèle ont été considérées comme incertaines, nous avons pu mettre en évidence 6 zones qui ont effectivement une influence sur les hauteurs d'eau modélisées aux trois stations de vigilance. Pour les stations fluviales Kervir et Moulin-Vert, ce sont les zones de frottement situées immédiatement à l'aval de la station qui ont la plus grande importance. Les zones de frottement immédiatement à l'amont présentent une petite importance, limitée dans le temps au moment de la montée de l'hydrogramme. À Justice, ce sont les deux zones de frottement situées entre la station et l'exutoire maritime qui contrôlent la hauteur d'eau simulée, avec un comportement cyclique en fonction de la phase de la marée. Dans l'expérience $S_{min}S_{maj}$, les coefficients de Strickler des 6 zones retenues précédemment ont été perturbés en lit mineur et en lit majeur. Les conclusions en termes d'importance sur les zones de frottement sont identiques à celles émises avec l'expérience S_{min}. Cette expérience permet de mettre en outre en évidence les périodes pendant lesquelles le lit majeur est activé au niveau des différentes zones de frottement. À Kervir et Moulin-Vert, le coefficient de Strickler en lit majeur prend de l'importance au cours des événements, de manière plus prononcée au moment du pic. L'importance du coefficient de Strickler majeur de la zone amont est relativement faible pour ces deux stations. À Justice, le lit majeur de la zone 8 est activé en même temps que le lit mineur, et cela correspond aux moments où le gradient de hauteur d'eau à Justice est important (flot et jusant). Pour finir, la perturbation des conditions limites amont et aval a été ajoutée dans l'expérience $S_{min}S_{maj}CL$. Cette expérience montre une importance prépondérante des conditions aux limites par rapport aux valeurs de frottement pour les hauteurs d'eau simulées aux trois stations de vigilance. À Justice, les influences respectives de la condition limite maritime et des coefficients de frottement sont toujours cycliques et liées à la phase de la marée. Cette dernière expérience permet de conclure sur les variables d'entrée qui sont influentes

pour la simulation des hauteurs d'eau aux trois stations. Il s'agit :

- des conditions aux limites amont et aval imposées ;
- de la valeur des coefficients de Strickler des mineurs des zones 4, 11, 7 et 8 ;
- de la valeur du coefficient de Strickler du lit majeur de la zone 8.

Dans une étude de sensibilité globale portant sur un modèle de prévision TELEMAC2D sur l'estuaire de la Gironde, Laborie et al. [2020] obtiennent des résultats similaires : les variables prédominantes sont liées au caractère estuarien ou fluvial de la zone d'étude. Dans la zone estuarienne, les auteurs relèvent un comportement cyclique lié à la phase de la marée pour l'influence de la condition limite maritime et des coefficients de Strickler. Dans les zones fluviales à l'amont, les coefficients de Strickler et la condition limite amont sont prédominants.

L'objectif de l'étude de sensibilité est d'appréhender les paramètres incertains du modèle pour pouvoir en particulier générer des prévisions d'ensemble brutes. Puisque nous n'avons pas à disposition de prévision d'ensemble de surcotes à l'aval, il ne sera pas possible en pratique de perturber la condition limite aval pour générer l'ensemble brut. On peut se demander si les paramètres d'entrée significatifs seront les mêmes si la condition limite aval n'est plus perturbée. L'analyse comparée des résultats des expériences $S_{min}S_{maj}$ et $S_{min}S_{maj}CL$ permet de conclure que ce sont effectivement les mêmes. En effet, l'expérience $S_{min}S_{maj}CL$ montre que seule la station Justice est impactée par le choix de la condition limite maritime. En plus de la condition limite maritime, ce sont les coefficients de Strickler mineurs des zones 7 et 8 et le coefficient de Strickler majeur de la zone 8 qui sont influents pour les hauteurs simulées à Justice. L'expérience $S_{min}S_{maj}$, pour laquelle la condition limite aval n'est pas perturbée, n'a pas mis en évidence d'autres zones de frottement déterminantes pour les hauteurs d'eau simulées à Justice. Ainsi, nous pouvons conclure que même en l'absence de perturbation de la condition limite maritime, les paramètres influents pour les hauteurs d'eau simulées par le modèle aux trois stations sont ceux définis grâce à l'expérience $S_{min}S_{maj}CL$.

Ces conclusions, qui sont un préalable à la suite de l'étude, sont naturellement liées aux hypothèses faites concernant la plage de variation des coefficients de Strickler, le choix de l'amplitude de la perturbation de la condition limite aval, et les caractéristiques des prévisions d'ensemble hydrologique. Elles sont aussi étroitement liées aux types d'événements étudiés, bien que l'étude de la sensibilité pour quatre types d'événements permette de limiter cet impact. Elles vont néanmoins permettre de déterminer la stratégie de génération d'un ensemble brut de prévision hydraulique qui sera dans un deuxième temps calibré statistiquement grâce à la méthode QRF. Cela permettra également de mettre en place une assimilation de données ensembliste dans le modèle MASCARET de l'Odet, et les paramètres mis en évidence ici seront corrigés.

Chapitre 8

Calibration statistique QRF des ensembles bruts hydrauliques

Sommaire

8.1 Méthodologie 207
8.1.1 Génération de l'ensemble brut 207
8.1.1.1 Variables incertaines pour la génération de l'ensemble brut 207
8.1.1.2 Réalisation des prévisions dans un cadre opérationnel . 210
8.1.2 Méthodologie de calibration QRF 211
8.1.2.1 Périodes de calibration et validation 211
8.1.2.2 Optimisation des forêts aléatoires sur la période de calage P_{Cal} 212
8.1.3 Synthèse des ensembles générés 214
8.2 Calibration QRF à Kervir et Moulin-Vert 217
8.2.1 Prédicteurs 217
8.2.2 Comparaison des deux critères pour le choix de la meilleure forêt 219
8.2.3 Résultats à Kervir et Moulin-Vert 219
8.2.3.1 Comparaison des ensembles bruts 222
8.2.3.2 Performances des ensembles calibrés par rapport à l'ensemble brut correspondant 224
8.2.3.3 Comparaison des quantiles calibrés avec et sans prédicteur issus d'un ensemble hydraulique 226
8.2.3.4 Meilleure stratégie à retenir 228
8.2.3.5 Résultats avec une prévision de pluie parfaite 229
8.2.3.6 Importance des prédicteurs à Kervir et Moulin-Vert . . 231
8.2.4 Synthèse 234
8.3 Calibration QRF à Justice 236

8.3.1 Prédicteurs . . . 236
8.3.2 Comparaison des deux critères pour le choix de la meilleure forêt 237
8.3.3 Résultats comparés des 3 expériences E_{PA}, E_{PASC} et E_{MP} . . . 238
8.3.3.1 Analyse des ensembles bruts . . . 243
8.3.3.2 Caractéristiques des ensembles calibrés par rapport aux ensembles bruts correspondants . . . 245
8.3.3.3 Meilleure stratégie à retenir concernant la condition limite aval en prévision . . . 248
8.3.3.4 Importance des prédicteurs . . . 249
8.3.4 Synthèse . . . 252
8.4 Conclusion générale . . . 253

Ce chapitre présente la génération de prévisions d'ensemble hydrauliques brutes, leur calibration statistique QRF et les résultats associés. Le paragraphe 8.1 décrit dans un premier temps la méthodologie employée pour la génération des ensembles bruts avec le modèle MASCARET et pour leur calibration avec la méthode QRF. La calibration à Kervir et Moulin-Vert est effectuée sur les hauteurs d'eau prévues au pas de temps horaire. Les résultats pour ces deux stations sont présentés au paragraphe 8.2. En revanche, la station Justice étant sous influence maritime, ce sont les valeurs des pics de pleine mer qui sont calibrés à cette station. Les résultats sont présentés au paragraphe 8.3. Pour finir, le paragraphe 8.4 propose une synthèse des résultats obtenus dans ce chapitre.

8.1 Méthodologie

Dans ce paragraphe, nous décrivons la méthodologie suivie pour la génération des ensembles bruts et leur calibration statistique. Différents ensembles bruts sont construits selon le choix effectué pour les conditions limites amont et aval. Ce point fait l'objet du paragraphe 8.1.1. La méthodologie suivie pour la calibration des ensembles bruts est présentée au paragraphe 8.1.2. Nous suivons la même approche que pour la calibration des prévisions d'ensemble hydrologiques, c'est-à-dire que les forêts aléatoires sont construites pour chaque station et par échéance de prévision. Les périodes de calage et de validation des forêts aléatoires sont présentées, ainsi que les critères testés pour l'optimisation des forêts aléatoires. Le paragraphe 8.1.3 fait la synthèse des ensembles bruts générés et des méthodes de calibration appliquées pour chacune des stations.

8.1.1 Génération de l'ensemble brut

8.1.1.1 Variables incertaines pour la génération de l'ensemble brut

Le chapitre 7 a présenté l'étude de sensibilité globale du modèle MASCARET pour les hauteurs d'eau prévues aux trois stations de vigilance Kervir, Moulin-Vert et Justice. Cette étude a permis d'identifier les paramètres d'entrée significatifs pour les hauteurs d'eau prévues par le modèle. Nous rappelons que pour des considérations de temps de calcul, l'étude de sensibilité n'a pas été réalisée dans un contexte de prévision. Il n'y a donc pas de notion d'échéance de prévision dans les conclusions émises. Les paramètres d'entrée significatifs mis en évidence sont les suivants :

- les débits hydrologiques imposés aux trois stations amont Tréodet, Kerjean et Ty-Planche ;
- la condition limite imposée à la station aval Plaisance pour la station Justice uniquement ;
- la valeur des coefficients de Strickler en lit mineur des zones 4, 7, 8 et 11 ;

- la valeur du coefficient de Strickler en lit majeur de la zone 8.

L'idéal est de générer des prévisions d'ensemble brutes qui tiennent compte de toutes ces sources d'incertitude. Nous décrivons ci-après de quelle manière elles sont prises en compte.

Débits hydrologiques amont Comme nous l'avons rappelé dans l'aide-mémoire de la partie III, les deux prévisions d'ensemble hydrologiques désormais considérées sont :

- l'ensemble brut *raw-HEF* généré par perturbation des paramètres de calage significatifs pour les débits simulés avec le modèle MORDOR-TS ;
- l'ensemble *QRF-hydro-TS* généré par calibration statistique QRF, suivie de l'étape de reconstruction et de lissage des trajectoires.

Des ensembles bruts hydrauliques sont donc construits pour chacune de ces deux conditions limites amont. Ces ensembles ont été construits avec 99 membres, un tirage dans une loi aléatoire discrète uniforme [1,99] est donc réalisé pour le choix du membre hydrologique.

Coefficients de Strickler Tous les ensembles hydrauliques bruts sont générés en tenant compte de l'incertitude sur les coefficients de Strickler pour lesquels l'étude de sensibilité a montré qu'ils ont un impact sur les hauteurs prévues. Ils sont perturbés de la même manière, avec un tirage aléatoire dans une loi uniforme dont les bornes sont rappelées dans le tableau 8.1. Les coefficients de Strickler pour lesquels la valeur n'est pas significative sont fixés à leur valeur de calage (voir tableau 2.8).

Zone	Ks_{min} minor	Ks_{max} minor	Ks_{min} flood plain	Ks_{max} flood plain
4	33	43	-	-
7	32	42	-	-
8	20.5	30.5	1	10
11	27.5	37.5	-	-

Tableau 8.1: Valeurs minimales et maximales des coefficients de Strickler perturbés pour la génération des ensembles hydrauliques bruts.

Condition limite aval À Plaisance, le niveau d'eau est la somme de la marée astronomique et d'éventuelles surcotes ou décotes météorologiques. Pour ce travail, nous avons à disposition des prévisions de marée astronomique réalisées par le SHOM[1]. Celles-ci sont supposées parfaites, et l'incertitude sur le niveau est reportée sur la surcote. En effet, l'ordre de grandeur des surcotes/décotes est significativement supérieur aux erreurs de prévision de la marée astronomique. Cependant, nous n'avons pas de prévision déterministe ou d'ensemble de surcote. Nous ne sommes donc pas en mesure de tenir compte de

[1]Service Hydrographique et Océanographique de la Marine

l'incertitude sur la condition limite aval. C'est pour cette raison que l'étude de sensibilité globale a été réalisée en construisant des chroniques de surcotes perturbées à partir des mesures disponibles *a posteriori*. Cette méthodologie n'est évidemment pas possible en opérationnel. En opérationnel, le Service de Prévision des Crues possède des prévisions déterministes de surcote, mais l'historique des prévisions n'est pas conservé au sein du réseau.

Dans le paragraphe 7.3, nous avons montré que même en l'absence de perturbation de la condition limite aval, les autres paramètres incertains restent les mêmes. Le fait de ne pas perturber la condition limite aval ne remet donc pas en cause la prise en compte des autres paramètres incertains décrits précédemment. Différentes conditions limites en prévision sont utilisées pour générer les ensembles hydrauliques bruts :

- Dans l'expérience E_{PA} (*Prévision Astronomique*), c'est la prévision de marée astronomique qui est utilisée en prévision. On ne tient pas compte des surcotes/décotes liées aux phénomènes météorologiques.

- Dans l'expérience E_{PASC} (*Prévision Astronomique Surcote Constante*), on ajoute à la chronique de prévision de marée astronomique la surcote mesurée au temps de base. Nous rappelons que le temps de base correspond au moment de lancement de la simulation de prévision.

- Dans l'expérience E_{MP} (*Mesure Perturbée*), la mesure de surcote réalisée *a posteriori* est perturbée de la même manière que lors de la réalisation de l'étude de sensibilité globale (voir paragraphe 7.1.1). Même si cette expérience n'est pas réalisable en opérationnel, elle est néanmoins testée ici car elle représente le cas d'une pseudo-prévision d'ensemble à l'aval. C'est l'objectif vers lequel il faut tendre pour la prise en compte des incertitudes de prévision au niveau de la condition aval. L'ensemble est généré à Plaisance avec 99 membres. Un tirage dans une loi aléatoire discrète uniforme [1,99] est donc réalisé. Avec cette expérience, on se place dans un cadre plutôt favorable puisque les membres perturbés sont créés à partir de mesures de surcote. Par construction, la moyenne de l'ensemble est proche de la valeur observée. En revanche, nous pouvons noter qu'il n'est pas fiable car très sur-dispersif.

Synthèse des expériences et terminologie Il y a deux possibilités pour le choix de la condition limite amont (*raw-HEF* et *QRF-hydro-TS*), et il y a trois possibilités pour le choix de la condition limite aval (*PA*, *PASC* et *MP*). La combinaison de ces différentes possibilités engendre six expériences différentes pour lesquelles des ensembles bruts sont construits. Elles sont synthétisées dans le tableau 8.2. Le nom des expériences $E_{el1/el2}$ est ainsi indicé par deux éléments *el1* et *el2* :

- *el1* correspond à l'ensemble hydrologique pris en compte à la condition limite amont pour la génération de l'ensemble brut : R pour l'ensemble brut *raw-HEF*, et Q pour

l'ensemble calibré *QRF-hydro-TS* ;

- *el2* correspond à la condition limite maritime utilisée : *PA*, *PASC* ou *MP*.

		Condition limite amont	
		raw-HEF	*QRF-hydro-TS*
Condition limite aval	*PA*	$E_{R/PA}$	$E_{Q/PA}$
	PASC	$E_{R/PASC}$	$E_{Q/PASC}$
	MP	$E_{R/MP}$	$E_{Q/MP}$

Tableau 8.2: Hypothèses pour la génération des 6 ensembles bruts et nom des expériences résultantes.

Pour chacune de ces expériences $E_{el1/el2}$, les ensembles bruts seront nommés $R_{el1/el2}$ (R pour *raw*), et les ensembles calibrés $Q_{el1/el2}$ (Q pour *QRF*). Lorsqu'il n'y aura pas de doute sur le choix de la condition limite aval utilisée (cas de la calibration à Kervir et Moulin-Vert, voir paragraphe 8.1.3), le second élément *el2* sera omis pour plus de lisibilité. On parlera alors des ensembles bruts R_R et R_Q et des ensembles calibrés Q_R et Q_Q.

8.1.1.2 Réalisation des prévisions dans un cadre opérationnel

Les prévisions brutes sont réalisées dans un cadre opérationnel de manière cyclée selon le schéma présenté figure 8.1. Cela signifie qu'en-dehors de l'expérience E_{MP} qui utilise une condition limite issue de mesures postérieures au temps de base, les autres expériences fournissent des prévisions de hauteurs d'eau qu'il serait possible d'obtenir en opérationnel.

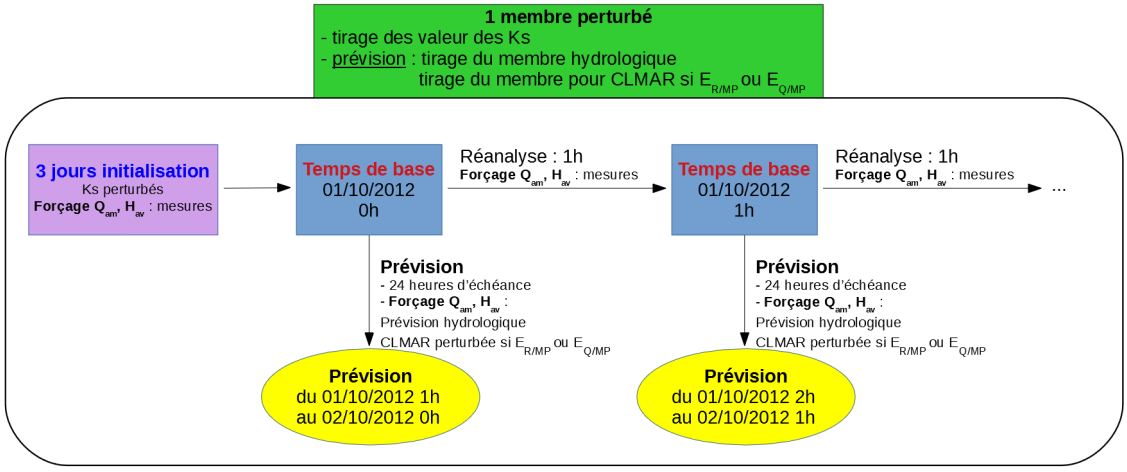


Figure 8.1: Réalisation des prévisions d'ensemble de manière cyclée.

Les prévisions d'ensemble hydrauliques sont constituées de 19 membres[2]. Chacun des

[2]Les ensembles hydrauliques bruts n'ont pas été construits avec 99 membres comme les ensembles hydrologiques à cause d'un problème de fuite mémoire dans les sources des API MASCARET au moment du lancement des simulations ensemblistes. Ce problème a été résolu depuis. Pour la calibration statistique, comme nous le verrons plus loin, nous utilisons uniquement comme prédicteurs la moyenne et l'écart-type de l'ensemble. Nous avons vérifié sur 30 jours de simulation pour les événements de crue de décembre 2013

membres est généré par le tirage d'une valeur pour chacun des coefficients de Strickler retenus et le tirage d'un membre hydrologique. Pour l'expérience E_{MP}, le tirage d'un membre perturbé pour la condition limite aval est également effectué. La simulation est d'abord lancée sur une période d'initialisation de 3 jours durant laquelle aucune prévision n'est réalisée : le modèle dont les coefficients de Strickler sont perturbés est alimenté avec les mesures à l'amont et à l'aval. Après cette période d'initialisation, une première prévision est réalisée en alimentant le modèle MASCARET par les prévisions hydrologiques choisies selon l'expérience, et par la condition limite aval perturbée le cas échéant (expérience E_{MP}). Pour se placer au temps de base suivant, une réanalyse d'une heure est lancée en repartant de l'état du modèle au temps de base courant. Une prévision est alors à nouveau lancée, et ainsi de suite, pendant toute la durée de l'expérience, c'est-à-dire pendant les périodes de calibration ou de validation. Ces périodes sont décrites dans le paragraphe 8.1.2.1.

8.1.2 Méthodologie de calibration QRF

Une fois que les ensembles hydrauliques bruts sont générés, une calibration statistique QRF leur est appliquée par échéance de prévision pour chacune des trois stations de vigilance. À partir des forêts alétaoires construites, 99 quantiles calibrés équirépartis sont extraits et forment ainsi l'ensemble calibré. Ce paragraphe décrit la mise en œuvre de la calibration statistique QRF des ensembles bruts.

8.1.2.1 Périodes de calibration et validation

Nous suivons ici la même méthodologie que pour la calibration statistique des ensembles hydrologiques et l'évaluation opérationnelle de cette calibration. Pour chaque ensemble hydraulique brut construit, la forêt aléatoire qui va permettre de le calibrer est optimisée sur une période dite de *calibration* ou de *calage* par le choix du triplet d'hyperparamètres *(mtry*[3]*, nodesize*[4]*, ntree*[5]*)* (voir paragraphe 3.3.1). Ces deux termes, *calibration* ou *calage*, pourront être employés indistinctement. Cette période est appelée P_{Cal} par la suite. Dans un second temps, la forêt aléatoire optimisée sur P_{Cal} est évaluée sur période de validation distincte appelée P_{Val}. Cela permet d'appréhender ses performances en opérationnel.

Le tableau 8.3 résume les périodes de calibration et de validation utilisées pour la calibration statistique QRF et son évaluation aux trois stations de vigilance. La période de calage P_{Cal} est la même que celle qui a servi à la calibration de l'hydrologie, à savoir d'octobre 2011 à juin 2014, les mois d'été de juillet à septembre étant retirés. Les mesures

que l'estimation de la moyenne n'est pas influencée par le nombre réduit de membres. Pour l'estimation de l'écart-type, l'écart est cependant plus significatif.

[3]nombre de prédicteurs tirés aléatoirement parmi l'ensemble de prédicteurs à chaque nœud de l'arbre

[4]nombre d'éléments dans les feuilles

[5]nombre d'arbres dans la forêt

de hauteur d'eau à la station Justice ne débutant qu'en septembre 2012, la calibration est effectuée sur une période de 2 ans à Justice, à partir d'octobre 2012. Dans le tableau 8.3, les mois écrit en gras correspondent à ceux utilisés à Justice. La période de validation P_{Val} s'étend de janvier 2010 à juin 2011 (Kervir et Moulin-Vert uniquement), puis d'octobre 2014 à avril 2016, les mois d'été étant toujours retirés. Il n'est pas possible d'étendre la période P_{Val} au-delà d'avril 2016 car cette date correspond au déplacement de la station Ty-Planche. Nous n'avions donc pas au moment du travail des chroniques de débits cohérentes sur cette station à partir de 2016 permettant d'alimenter le modèle MASCARET. Pour chaque station, la période de calage représente une période homogène d'un point de vue climatologique puisqu'elle contient chaque mois le même nombre de fois. En revanche, ceci n'est pas le cas pour la période de validation. Nous avons tout de même souhaité conserver un échantillon de validation le plus important possible.

Année	Mois											
2010	1	2	3	4	5	6	7	8	9	10	11	12
2011	1	2	3	4	5	6	7	8	9	10	11	12
2012	1	2	3	4	5	6	7	8	9	**10**	**11**	**12**
2013	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
2014	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
2015	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
2016	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**

Période de calage
Période de validation

Tableau 8.3: Périodes de calage et de validation pour les stations Kervir, Moulin-Vert et Justice. Pour Justice, les mois disponibles sont uniquement ceux écrits en gras.

8.1.2.2 Optimisation des forêts aléatoires sur la période de calage P_{Cal}

Pour optimiser le triplet d'hyperparamètres des forêts aléatoires, un ensemble de forêts aléatoires est créé en le faisant varier. Cela génère autant d'ensembles calibrés sur lesquels des scores sont calculés. La meilleure forêt au sens d'une optimisation d'un score est retenue. Nous décrivons ci-après trois points essentiels pour la mise en œuvre de la méthode : la méthode de validation croisée, la stratification réalisée et le critère d'optimisation choisi.

Méthode de validation croisée Une fois que le triplet d'hyperparamètres est choisi, la construction de la forêt aléatoire et son évaluation sont réalisées grâce à une stratégie de validation croisée (en anglais *cross validation*). Cette stratégie est employée sur les 27 mois que comporte la période (18 mois pour Justice). Chaque mois est alternativement retiré de

l'échantillon d'apprentissage. Les 26 mois restants sont utilisés pour l'apprentissage et le mois retiré est utilisé pour la validation de la forêt. Une période de 10 jours avant et après le mois utilisé pour la validation est retirée de l'échantillon d'apprentissage pour éviter les problèmes d'auto-corrélation entre les données d'apprentissage et de validation. Les 27 calibrations indépendantes effectuées sur chacun des 27 mois sont agrégées et forment ainsi l'échantillon calibré. Des scores peuvent alors être calculés. Attention à ne pas confondre la validation dont nous parlons ici, qui est un processus inhérent à l'optimisation des forêts aléatoires, et la validation de la forêt aléatoire finalement retenue sur une période P_{Val} indépendante de P_{Cal} dont nous avons parlé précédemment, et qui est effectuée dans un second temps.

Stratification de l'ensemble Le principe de la stratification a été présenté au paragraphe 3.1.2.3. La stratification des ensembles permet d'évaluer leurs performances sur des sous-échantillons. Dans notre cas, nous utilisons la stratification pour optimiser et évaluer les ensembles sur des périodes de hautes eaux. Justice étant sous influence maritime, le choix a été fait de calibrer les pics de marée haute, puisque ce travail a pour visée opérationnelle la prévision des hautes eaux. La calibration des pics de marée haute remplace la stratification qui n'est ainsi pas nécessaire. À Kervir et Moulin-Vert en revanche, la variable calibrée est la hauteur d'eau mesurée au pas de temps horaire. Le critère de stratification est de retenir pour la vérification des ensembles les prévisions pour lesquelles le maximum de l'ensemble a dépassé le quantile q_{90} des hauteur observées à la station. Nous notons que c'est le même critère qui a été utilisé pour la calibration statistique des ensembles hydrologiques. Ce critère est appelé M_{90} dans la suite. Dans la pratique, ce critère permet de retenir environ 25% des pas de temps de prévision sur les périodes de calage P_{Cal} et de validation P_{Val}.

Critère d'optimisation des forêts Pour chaque ensemble calibré relatif à un triplet d'hyperparamètres, les CRPS et les normes $\|.\|_1$, $\|.\|_2$ et $\|.\|_\infty$ des histogrammes de rang sont calculés. Comme nous l'avons expliqué au paragraphe 3.3.2 concernant les aspects informatiques, la minimisation du CRPS ne va pas forcément de pair avec une bonne fiabilité. Dans la suite, nous allons comparer deux stratégies d'optimisation des forêts aléatoires :

- Optimisation de la fiabilité : on conserve dans un premier temps les 3 forêts qui minimisent les 3 normes des histogrammes de rang. Ensuite, on conserve parmi celles-ci la forêt qui minimise le CRPS. Ce critère d'optimisation est appelé Opt_{Rel} par la suite.
- Optimisation du CRPS : on conserve la forêt pour laquelle le CRPS est minimal. Ce critère d'optimisation est appelé Opt_{CRPS} par la suite.

Il est à noter que la valeur de AUC (aire sous la courbe ROC) n'a pas été utilisée comme

critère pour optimiser les forêts aléatoires.

Le triplet d'hyperparamètres finalement retenu est celui qui optimise le critère d'optimisation choisi. Une seule forêt est ensuite reconstruite avec ce triplet sur toute la période P_{Cal} (rappel, la validation croisée avait créé 27 forêts) pour une utilisation en opérationnel, ou sur la période P_{Val} dans le cadre de ce travail.

8.1.3 Synthèse des ensembles générés

Considérations générales La figure 8.2 synthétise la construction des ensembles bruts et calibrés aux trois stations. Cette configuration est la même pour les trois expériences *PA (Prévision Astronomique), PASC (Prévision Astronomique Surcote Constante)* et *MP (Mesure Perturbée)*. Dans un souci de simplification, ce graphique concerne uniquement l'expérience *PA*. Il est facilement transposable aux deux autres expériences. Il est à noter que le code couleur employé sur ce graphique sera également appliqué dans le texte et sur tous les graphiques pour une meilleure lisibilité du manuscrit.

Dans un premier temps, deux types d'ensembles bruts sont générés en tenant compte des deux types de condition limite amont concernant l'hydrologie brute ou calibrée. Les ensembles bruts $R_{R/PA}$ (forçage *raw-HEF*) et $R_{Q/PA}$ (forçage *QRF-hydro-TS*) sont construits selon la méthodologie présentée dans le paragraphe 8.1.1.2. Notons que les simulations réalisées avec le modèle MASCARET permettent d'obtenir les hauteurs d'eau en tout point du réseau hydraulique. Les prévisions brutes $R_{R/PA}$ ou $R_{Q/PA}$ sont donc construites de manière simultanée pour les trois stations de vigilance.

Pour chaque station de vigilance, chacune de ces deux prévisions d'ensemble brutes est ensuite calibrée par échéance de prévision avec la méthode QRF, ce qui produit respectivement les ensembles de quantiles calibrés $Q_{R/PA}$ et $Q_{Q/PA}$. Pour ces deux calibrations, on utilise en particulier comme prédicteurs des statistiques issues des ensembles bruts : la moyenne et la variance de l'ensemble. Un prédicteur nommé Q_{am} est également utilisé. Il représente le débit mesuré ou prévu aux stations amont à un pas de temps particulier. Sa description détaillée sera effectuée respectivement au paragraphe 8.2.1 pour Kervir et Moulin-Vert et au paragraphe 8.3.1 pour Justice. Pour le moment, il est juste utile de préciser que pour la calibration de l'ensemble brut $R_{R/PA}$, celui-ci est calculé à partir de la modélisation de MORDOR-TS déterministe alimenté par RR3 en prévision. Pour la calibration de l'ensemble brut $R_{Q/PA}$, ce prédicteur est la moyenne de l'ensemble hydrologique calibré *QRF-hydro-TS*.

La méthode QRF est également appliquée sans utiliser de prédicteur issu de statistiques d'un ensemble brut. Cela correspond à la construction d'un modèle d'apprentissage basé uniquement sur les données. Sa construction ne nécessite pas de modèle hydraulique. Une calibration statistique analogue à $Q_{R/PA}$ est ainsi réalisée sans utiliser les prédicteurs issus

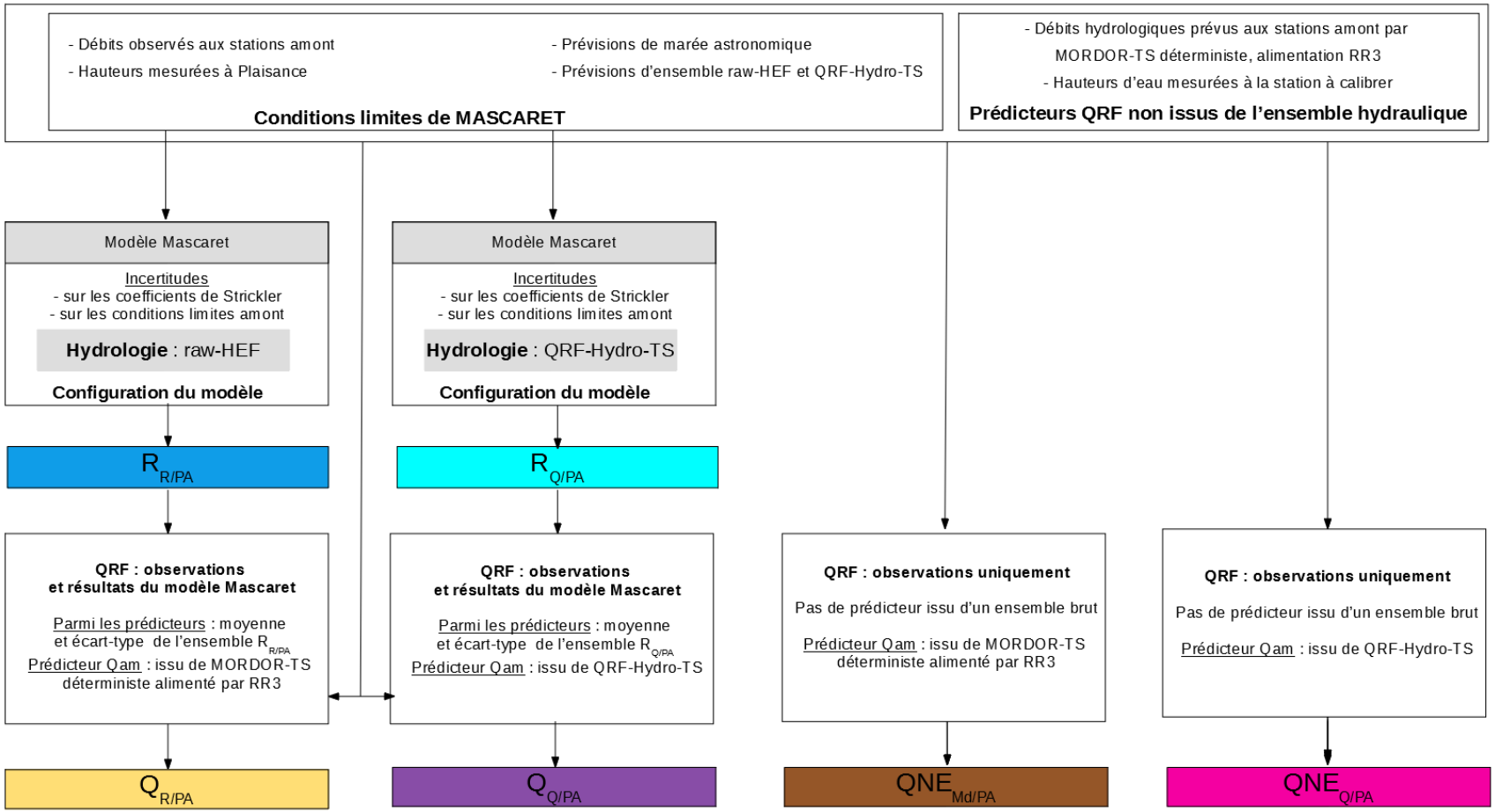


Figure 8.2: Génération des 6 ensembles de prévision à Kervir, Moulin-Vert ou Justice pour l'expérience E_{PA}. Les ensembles bruts $R_{R/PA}$ et $R_{Q/PA}$ (*«R» pour «Raw»*) diffèrent par le choix de l'ensemble hydrologique imposé à la condition limite amont (*raw-HEF* ou *QRF-hydro-TS*). Leur calibration produit respectivement les ensembles de quantiles calibrés $Q_{R/PA}$ et $Q_{Q/PA}$ (*«Q» pour «QRF»*). Les ensembles de quantiles calibrés $QNE_{Md/PA}$ et $QNE_{Q/PA}$ sont produits par des calibrations QRF sans utilisation de statistiques issues d'un ensemble brut.

de statistiques de la prévision d'ensemble brute. Cela signifie que tous les prédicteurs sont les mêmes, sauf que les prédicteurs représentant la moyenne et la variance de l'ensemble brut sont omis. Cette calibration fournit l'ensemble de quantiles $QNE_{Md/PA}$, *Md* signifiant *MORDOR déterministe*. Nous avons fait le choix de prendre MORDOR déterministe à la place de la moyenne de l'ensemble brut car ce prédicteur est beaucoup plus facile à obtenir : pour la prévision $QNE_{Md/PA}$, il n'est pas nécessaire de construire un ensemble hydrologique, mais seulement un modèle hydrologique déterministe. De même, l'ensemble de quantiles $QNE_{Q/PA}$ est obtenu avec une calibration QRF réalisée à partir des mêmes prédicteurs que pour la calibration $Q_{Q/PA}$, sauf ceux relatifs aux statistiques de l'ensemble brut.

Scores pour l'évaluation des ensembles Pour évaluer les performances des ensembles, les scores suivants seront présentés pour les ensembles bruts et calibrés pour les échéances de prévision allant de +1 heure à +24 heures :

- CRPS,
- norme 2 des histogrammes de rang,
- aire AUC sous les courbes ROC pour des événements de dépassement de seuil.

Spécificités à Kervir et Moulin-Vert À Kervir et Moulin-Vert, les hauteurs d'eau ne sont quasiment pas influencées par la condition limite aval maritime (voir les résultats de l'étude de sensibilité au chapitre 7). Pour ces deux stations, nous présenterons uniquement les résultats de la calibration pour l'expérience E_{PA} pour laquelle la prévision de marée astronomique est utilisée (voir tableau 8.2). Les expériences E_{PASC} (marée astronomique avec persistance de la surcote mesurée au temps de base) et E_{MP} (mesure perturbée) donneraient des résultats analogues sur ces deux stations. Dans la rédaction, pour une meilleure lisibilité, nous omettrons l'indice relatif à la condition limite aval pour le nom des ensembles bruts et calibrés. Pour ces deux stations, la variable calibrée est la hauteur d'eau horaire simulée par l'ensemble par échéance (entre +1h et +24h). La stratification décrite au paragraphe 8.1.2.2 est effectuée pour optimiser la calibration sur la prévision des hautes eaux. Les deux événements étudiés pour le tracé des courbes ROC sont les événements *Dépassement du quantile* q_{90} et *Dépassement du quantile* q_{99}.

Spécificités à Justice À Justice, la méthodologie est différente puisqu'on y calibre les pics de marée haute : les prédicteurs sont différents, aucune stratification n'est effectuée. Pour cette station, les trois expériences E_{PA}, E_{PASC} et E_{MP} seront comparées. L'événement considéré pour le tracé des courbes ROC est l'événement *Dépassement du seuil de vigilance vert-orange*. Les valeurs des seuils de vigilance utilisés en opérationnel sont confidentiels. Elles ne sont donc pas mentionnées dans ce manuscrit.

La suite du chapitre présente les calibrations réalisées aux trois stations de vigilance Kervir, Moulin-Vert et Justice. Les résultats de la calibration sur les périodes de calage et de validation sont présentés en partie 8.2 pour les stations Kervir et Moulin-Vert, et en partie 8.3 pour la station Justice.

8.2 Calibration QRF à Kervir et Moulin-Vert

Nous présentons dans cette partie les résultats concernant la génération des ensembles bruts avec l'expérience E_{PA} (prévision astronomique) et leur calibration statistique aux stations Kervir et Moulin-Vert. Nous travaillons exactement sur les six ensembles décrits en figure 8.2. L'indice PA est omis dans cette section. Les prédicteurs utilisés pour les calibrations sont présentés en partie 8.2.1. Dans la partie 8.2.2, les deux stratégies d'optimisation des forêts Opt_{Rel} et Opt_{CRPS} sont comparées. Les résultats en périodes de calage et de validation sont présentés en partie 8.2.3. La synthèse de la calibration à Kervir et Moulin-Vert est présentée en partie 8.2.4.

8.2.1 Prédicteurs

Les prédicteurs utilisés pour la calibration QRF à Kervir et Moulin-Vert sont présentés dans le tableau 8.4. Nous rappelons que des forêts aléatoires sont construites spécifiquement pour chacune des deux stations et pour chaque échéance de prévision. Les spécificités liées à l'une ou l'autre des stations sont indiquées en orange pour Kervir et en vert pour Moulin-Vert. Les ensembles de quantiles Q_R et Q_Q sont calibrés avec tous les prédicteurs présentés. Les prédicteurs représentant la moyenne et l'écart-type de l'ensemble (*Mean* et *Sigma*) sont différents pour chacune de ces deux calibrations puisque les ensembles bruts sont différents. L'ensemble de quantiles QNE_{Md} (respectivement QNE_Q) est généré avec la méthodologie QRF appliquée avec les mêmes prédicteurs que pour générer Q_R (respectivement Q_Q), sauf la moyenne et l'écart-type qui ne sont pas utilisés.

Les prédicteurs Q_{am1} et Q_{am2} correspondent aux débits mesurés ou prévus aux stations amont du réseau un certain nombre d'heures avant l'échéance calibrée ; le choix entre les deux dépend de la disponibilité ou non du débit mesuré à cet instant. Le temps de propagation moyen T_p entre chaque station amont et la station à calibrer est évalué à partir des mesures. Il dépend du débit et a été évalué pour des propagations de pic de crues à environ 2 heures pour Tréodet, 3 heures pour Ty-Planche et 4 heures pour Kerjean. Si l'échéance calibrée à la station de vigilance est inférieure au temps de propagation moyen T_p, alors le prédicteur est le débit mesuré T_p heures avant, puisque celui-ci est disponible. En revanche, si l'échéance calibrée est supérieure à T_p, le débit mesuré T_p heures avant n'est plus disponible. Le prédicteur utilisé est alors le débit prévu. Quand le modèle hydraulique est alimenté par l'ensemble hydrologique brut, le prédicteur Q_{am} est la prévision est réalisée

Nom	Unité	Description
Mean	m	Moyenne de l'ensemble brut
Sigma	m	Ecart-type de l'ensemble brut
Month		Mois de la prévision
Period		Période de la journée de la prévision (0h-6h, 6h-12h, 12h-18h, 18h-24h)
$GradH_0$	m/s	Gradient de la hauteur d'eau au temps de base
$PrecH_0$	m	Hauteur d'eau mesurée au temps de base
Q_{am1}	m^3/s	Débit mesuré ou prévu à une station amont du réseau (selon l'échéance calibrée) Kervir : débit à Tréodet 2 heures plus tôt Moulin-Vert : débit à Ty-Planche 3 heures plus tôt
Q_{am2}	m^3/s	Débit mesuré ou prévu à une station amont du réseau (selon l'échéance calibrée) Kervir : débit à Kerjean 4 heures plus tôt Moulin-Vert : pas de prédicteur Q_{am2}

Tableau 8.4: Prédicteurs utilisés pour la calibration QRF des stations sans influence maritime Kervir et Moulin-Vert.

par le modèle déterministe MORDOR-TS alimenté en prévision par les prévisions RR3. En revanche, quand le modèle hydraulique utilise les prévisions hydrologiques calibrées, le prédicteur Q_{am} est alors la prévision moyenne de l'ensemble hydrologique calibré *QRF-hydro-TS*. L'illustration des deux cas possibles selon la disponibilité du débit mesuré est donnée sur la figure 8.3 pour Tréodet ($T_p = 2$ heures) pour la calibration des échéances 1 heure et 6 heures.

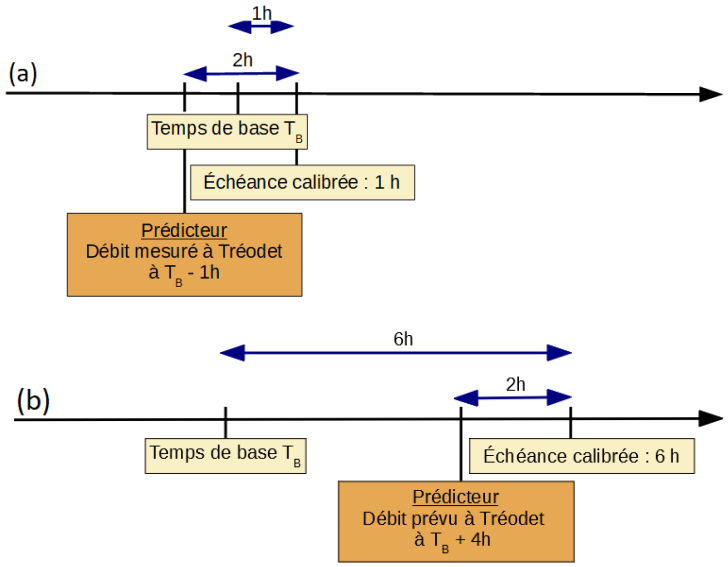


Figure 8.3: Illustration du choix du prédicteur Q_{am} selon la disponibilité de la mesure T_p heures précédant l'échéance calibrée. $T_p = 2$ heures à Tréodet. (a) : calibration de l'échéance +1h, le débit mesuré 2 heures avant est disponible. (b) : calibration de l'échéance +6h, le débit mesuré 2 heures avant n'est pas disponible, une prévision de débit hydrologique lancée au temps de base est utilisée pour le prédicteur.

La station Kervir étant située à l'aval de la confluence de l'Odet et du Jet, les deux prédicteurs Q_{am1} et Q_{am2}, représentant respectivement les débits mesurés ou prévus à Tréodet et à Kerjean, sont utilisés pour la calibration. À Moulin-Vert, seul le prédicteur Q_{am1} est utilisé, il représente le débit mesuré ou prévu à Ty-Planche.

8.2.2 Comparaison des deux critères pour le choix de la meilleure forêt

L'étude de la comparaison des deux critères d'optimisation des forêts aléatoires est présentée en annexe Q.1.1. Ces deux critères sont : optimisation du CRPS uniquement (Opt_{CRPS}) et optimisation de la fiabilité au préalable (Opt_{Rel}). Cette étude met en évidence les conclusions suivantes :

- Sur la période de calage, la stratégie Opt_{Rel} entraîne une courbe de CRPS moins lisse selon les échéances. La stratégie Opt_{CRPS} permet d'obtenir des performances en termes de CRPS légèrement meilleures que la stratégie Opt_{Rel}. En revanche, la stratégie Opt_{Rel} améliore plus nettement les normes des histogrammes de rang. Ainsi, en calage, si l'on accepte de dégrader légèrement le CRPS, on améliore significativement la fiabilité des forêts aléatoires.

- Sur la période de validation, la tendance des résultats est un peu différente. Les valeurs des CRPS pour les deux types d'optimisation sont quasiment équivalentes. En revanche, la fiabilité reste toujours un peu meilleure avec les forêts construites avec la stratégie $OptimRel$.

- Les courbes ROC ne sont pas significativement impactées par le choix de l'optimisation des forêts sur les deux périodes, même si l'aire AUC est très légèrement meilleure avec la stratégie Opt_{Rel}.

Pour la suite de l'étude, nous faisons le choix de conserver le critère Opt_{Rel} pour l'optimisation des forêts aléatoires à Kervir et Moulin-Vert.

8.2.3 Résultats à Kervir et Moulin-Vert

La figure 8.4 représente les valeurs des CRPS, la norme $\|.\|_2$ des histogrammes de rang et l'aire AUC de la courbe ROC pour les seuils q_{90} et q_{99}, pour les six ensembles R_R, R_Q, Q_R, Q_Q, QNE_{Md}, QNE_Q pour les périodes de calage et de validation à la station Kervir. Le tableau 8.5 présente la moyenne sur toutes les échéances des scores présentés sur les graphiques de la figure 8.4. La figure 8.5 et le tableau 8.6 sont analogues pour les résultats à Moulin-Vert. Dans l'annexe Q.1.2, les histogrammes de rang pour les échéances 6 heures, 12 heures, 18 heures et 24 heures sont présentés pour les stations Kervir et Moulin-Vert pour les périodes de calage et de validation. Les valeurs des CRPS sont indiquées au-dessus de chaque histogramme de rang. Nous rappelons que les résultats pour les CRPS et les histogrammes de rang sont présentés pour la stratification M_{90} (voir paragraphe 8.1.2.2).

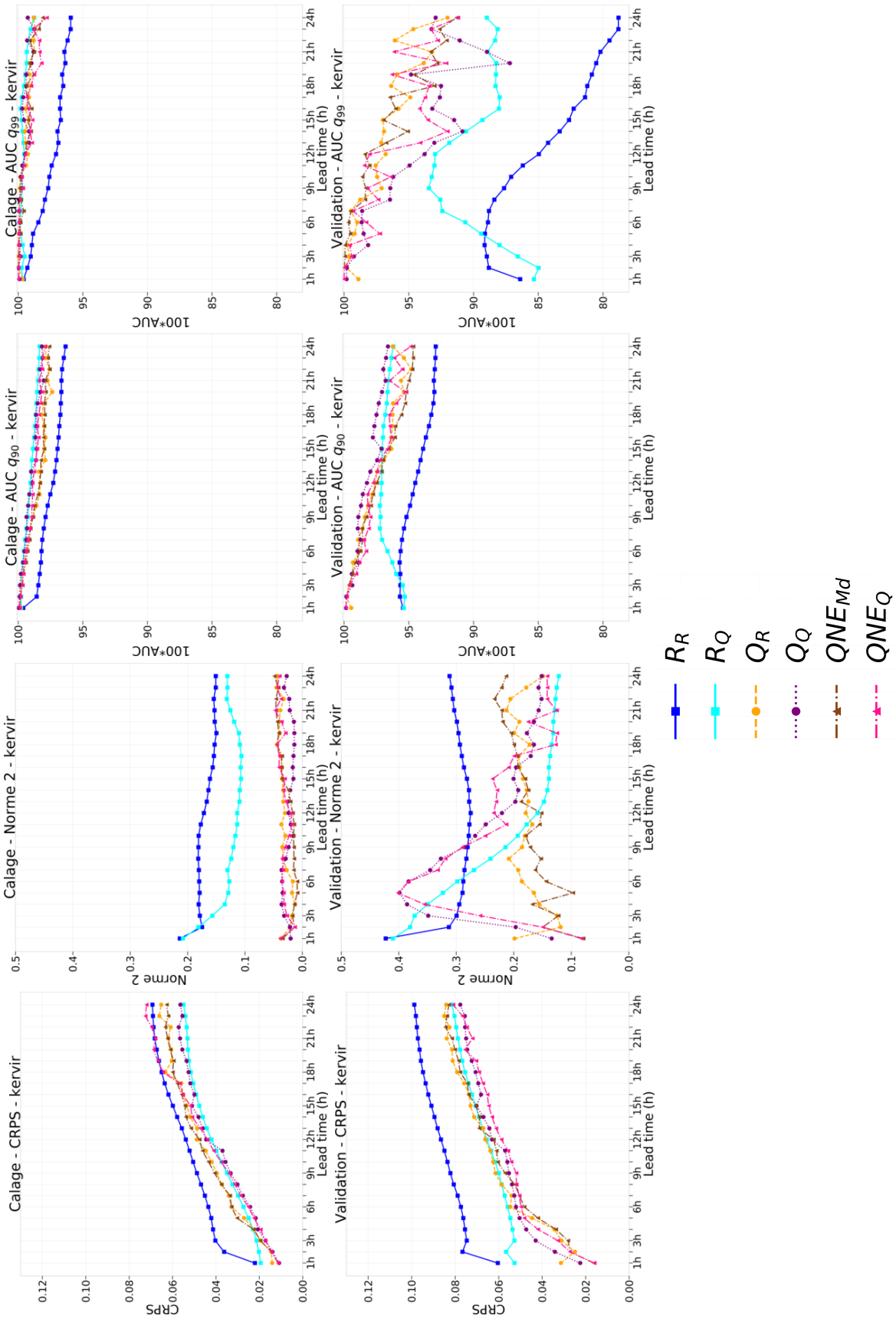


Figure 8.4: CRPS (col 1), $||.||_2$ (col 2), aire roc AUC pour les seuils q_{90} et q_{99} (col 3 et 4) pour la calibration à Kervir pour les périodes de calage (ligne 1) et de validation (ligne 2).

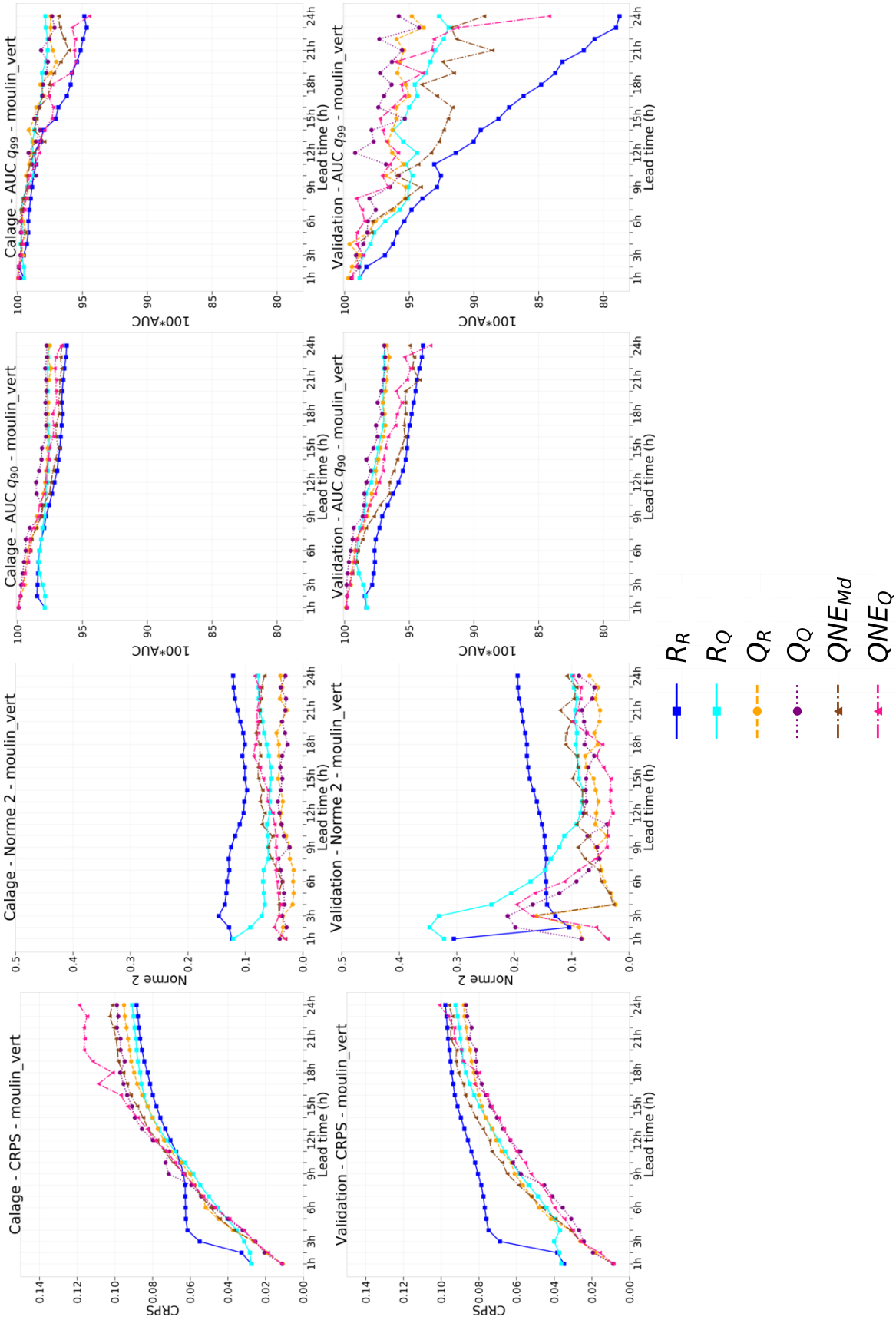


Figure 8.5: CRPS (col 1), $||\cdot||_2$ (col 2), aire roc AUC pour les seuils q_{90} et q_{99} (col 3 et 4) pour la calibration à Moulin-Vert pour les périodes de calage (ligne 1) et de validation (ligne 2).

Score	$10^{-2}\|.\|_2$		100*CRPS		100 * AUC - q_{90}		100 * AUC - q_{99}	
Période	Cal	Val	Cal	Val	Cal	Val	Cal	Val
R_R	16.9	29.7	5.4	8.6	97.5	94.4	97.4	84.5
R_Q	12.7	20.8	4.0	6.6	99.1	96.6	99.6	89.6
Q_R	3.2	17.9	4.5	6.4	98.6	97.5	99.4	97.0
Q_Q	2.4	23.4	4.0	6.0	99.0	98.1	99.5	94.7
QNE_{Md}	2.7	17.4	4.5	6.1	98.6	97.2	99.4	96.7
QNE_Q	3.3	22.3	4.5	5.7	98.8	97.4	99.3	96.0

Tableau 8.5: $\|.\|_2$, CRPS et aire ROC moyennés sur les 24 heures d'échéances à Kervir. Expérience E_{PA}.

Score	$10^{-2}\|.\|_2$		100*CRPS		100 * AUC - q_{90}		100 * AUC - q_{99}	
Période	Cal	Val	Cal	Val	Cal	Val	Cal	Val
R_R	11.7	16.8	7.0	8.3	97.3	96.0	97.6	89.7
R_Q	6.9	14.2	6.7	6.7	97.8	97.9	98.8	95.3
Q_R	3.3	6.1	6.8	6.4	98.3	97.9	98.7	96.5
Q_Q	3.5	8.8	7.2	5.9	98.5	98.3	98.7	97.3
QNE_{Md}	6.2	8.4	7.2	6.8	97.8	96.9	98.4	94.2
QNE_Q	6.0	7.5	7.6	6.2	98.1	97.3	98.0	96.2

Tableau 8.6: $\|.\|_2$, CRPS et aire ROC moyennés sur les 24 heures d'échéances à Moulin-Vert. Expérience E_{PA}.

8.2.3.1 Comparaison des ensembles bruts

Ensembles évalués dans ce paragraphe

⇒ les 2 ensembles bruts pour les deux stations Kervir et Moulin-Vert.

R_R : ensemble hydraulique brut, condition limite amont *raw-HEF*.

$\Longleftrightarrow$

R_Q : ensemble hydraulique brut, condition limite amont *QRF-hydro-TS*.

Dans ce paragraphe, nous comparons les performances des ensembles hydrauliques bruts R_R (alimenté par l'ensemble hydrologique brut) et R_Q (alimenté par l'ensemble hydrologique calibré) pour les deux stations Kervir et Moulin-Vert.

Nous notons tout d'abord que les performances des ensembles bruts sont moins bonnes sur la période de validation. Il est possible que cela soit lié au fait que les événements sont moins forts sur la période de validation et que les performances des ensembles bruts soient différentes selon les types d'événements.

D'une manière générale, les deux ensembles R_R et R_Q présentent des caractéristiques similaires aux très courtes échéances (1 heure voire 2 heures). Ainsi, aux très courtes échéances, la calibration de l'hydrologie au préalable n'apporte rien de significatif : ceci est lié au temps de propagation du débit, calibré ou non, vers l'aval où se trouvent les stations de vigilance.

À Kervir, le CRPS de l'ensemble brut est meilleur pour toutes les échéances si l'on a calibré l'ensemble hydrologique au préalable. Ceci est vrai sur les périodes P_{Cal} et P_{Val}. Cette conclusion est aussi valable à Moulin-Vert sur P_{Val}. Sur la période de calage en revanche, le comportement est légèrement différent. On note qu'à partir de l'échéance +10h, les deux ensembles R_R et R_Q présentent des CRPS similaires, légèrement meilleurs pour l'ensemble construit avec l'hydrologie non calibrée R_R. Pour cette gamme d'échéances en période de calage, il n'y a donc pas de gain à calibrer l'hydrologie pour générer les ensembles bruts, mais le fait de le faire ne dégrade pas les résultats. D'une manière générale, la calibration de l'ensemble hydrologique permet donc d'obtenir des ensembles hydrauliques bruts avec des meilleurs CRPS.

En ce qui concerne la norme des histogrammes de rang, l'intérêt à calibrer l'hydrologie au préalable est nettement visible sur la période P_{Cal}. Ceci n'est pas complètement vrai sur P_{Val} aux courtes échéances. En effet, entre les échéances 2 heures et 5 heures à Kervir et 2 heures et 6 heures à Moulin-Vert, le fait de calibrer l'hydrologie au préalable dégrade la fiabilité de l'ensemble hydraulique. On peut en partie imputer ce défaut de fiabilité à de moins bonnes performances de l'ensemble hydrologique calibré *QRF-hydro-TS* sur la période de validation. Cet ensemble présente un défaut de fiabilité aux courtes échéances par rapport à l'ensemble brut hydrologique à Kerjean (voir chapitre 6). Cela peut en partie expliquer le défaut de fiabilité aux courtes échéances à Kervir, mais ça ne permet pas d'expliquer cette tendance à Moulin-Vert (rappel : Kerjean est à l'amont de Kervir mais pas de Moulin-Vert).

L'ensemble R_R présente une très mauvaise discrimination des événements *dépassement du seuil q90* et *dépassement du seuil q99* à toutes les échéances. La discrimination des événements est nettement améliorée pour les deux stations avec l'ensemble R_Q qui utilise un ensemble hydrologique calibré.

Synthèse

Nous pouvons conclure à ce stade de l'analyse qu'il y a un réel intérêt à calibrer l'ensemble hydrologique avant son utilisation pour générer des prévisions d'ensemble hydrauliques brutes avec le modèle MASCARET.

8.2.3.2 Performances des ensembles calibrés par rapport à l'ensemble brut correspondant

Ensembles évalués dans ce paragraphe

⇒ Comparaison des ensembles hydrauliques calibrés par rapport à l'ensemble brut correspondant.

R_R : ensemble hydraulique brut, condition limite amont *raw-HEF*.	$\Longleftrightarrow$	Q_R : quantiles calibrés à partir de l'ensemble hydraulique brut R_R.
R_Q : ensemble hydraulique brut, condition limite amont *QRF-hydro-TS*.	$\Longleftrightarrow$	Q_Q : quantiles calibrés à partir de l'ensemble hydraulique brut R_Q.

Dans ce paragraphe, nous comparons les performances des ensembles hydrauliques calibrés par rapport à l'ensemble brut correspondant, c'est-à-dire que nous comparons l'ensemble calibré Q_R par rapport à l'ensemble brut R_R, et l'ensemble calibré Q_Q par rapport à l'ensemble brut R_Q.

Les graphiques 8.4 et 8.5 mettent en évidence que les CPRS des ensembles calibrés sont généralement meilleurs que ceux de l'ensemble brut correspondant : la courbe Q_R est sous la courbe R_R et la courbe Q_Q est sous la courbe R_Q, même si ces dernières sont parfois tangentes. Ceci n'est pas complètement vrai à Moulin-Vert en période de calage : le CRPS n'est pas amélioré par la calibration QRF à partir de l'échéance +10h. C'est alors l'ensemble R_R, construit sans aucune calibration dans la chaîne de prévision, qui fournit les meilleurs CRPS, suivi de près par le second ensemble brut R_Q. On pourrait penser qu'il n'y a alors aucun intérêt à calibrer l'hydrologie et/ou l'hydraulique, mais ces résultats ne se retrouvent pas en période de validation, pour laquelle l'intérêt de la calibration statistique reste présent. Nous pouvons conclure que la calibration de l'hydraulique en bout de chaîne permet d'améliorer les CRPS par rapport à l'ensemble brut.

Dans la période de calage, les normes des ensembles calibrés sont significativement meilleures que celles des deux ensembles bruts : les courbes Q_R et Q_Q sont respectivement très en-dessous des courbes R_R et R_Q. En période de validation, les résultats ne sont pas aussi clairs. L'ensemble R_R présente une fiabilité assez homogène selon les échéances (sauf pour l'échéance 1 heure). L'ensemble calibré correspondant Q_R présente aussi une fiabilité assez homogène selon les échéances, avec des normes d'histogrammes de rang plus petites que pour l'ensemble brut, sauf à Moulin-Vert à l'échéance isolée de 3 heures. En ce qui concerne les ensembles créés avec l'hydrologie calibrée, nous avons vu au paragraphe 8.2.3.1 que l'ensemble R_Q présente une mauvaise fiabilité jusqu'à environ 10 heures aux deux

stations. Cette tendance se retrouve sur les quantiles calibrés à partir de cet ensemble brut : les quantiles Q_Q présentent aussi un gros défaut de fiabilité aux courtes échéances. À Kervir, sur la période de validation, l'ensemble calibré Q_Q présente même pour toutes les échéances de prévision supérieures à 4 heures une fiabilité moins bonne que l'ensemble brut correspondant R_Q. Les normes des histogrammes de rang des ensembles calibrés Q_R et Q_Q sont nettement supérieures à celles obtenues sur P_{Cal} pour la station Kervir. À Moulin-Vert, cette caractéristique est également présente mais de manière plus mesurée, sachant que les ensembles bruts présentent également une moins bonne fiabilité sur P_{Val}. Ces éléments permettent tout d'abord de conclure que les résultats ne sont globalement pas consistants entre les périodes de calibration et de validation pour la fiabilité des ensembles calibrés. Seul l'ensemble calibré Q_R à Moulin-Vert présente une fiabilité relativement proche de celle obtenue sur P_{Cal}. La calibration statistique mise en place ne permet pas de corriger la fiabilité en validation de manière homogène sur toutes les échéances pour les ensembles issus de l'hydrologie calibrée. Nous avons également mis en évidence que la fiabilité des ensembles calibrés est en partie liée à la fiabilité des ensembles hydrologiques utilisés.

Les graphiques représentant les aires AUC des courbes ROC mettent en évidence le fait qu'en période de calage, c'est l'ensemble R_Q qui est le plus performant pour prévoir les événements *dépassement du seuil q90* et *dépassement du seuil q99* à Kervir. À Moulin-Vert, cet ensemble est également très performant par rapport aux autres ensembles pour l'événement *dépassement du seuil q99*, un peu moins pour *dépassement du seuil q90*. En période de validation, la discrimination de l'ensemble R_Q est nettement moins bonne. Sur la période de validation, les calibrations statistiques Q_R et Q_Q permettent d'améliorer la discrimination des événements de dépassement de seuil par rapport aux performances des ensembles bruts.

Synthèse

La conclusion de ce paragraphe est que d'une manière générale, la calibration statistique des ensembles bruts permet d'améliorer leurs performances en termes de CRPS et discrimination. Nous avons vu que la fiabilité est plus difficile à corriger en période de validation, surtout à Kervir.

8.2.3.3 Comparaison des quantiles calibrés avec et sans prédicteur issus d'un ensemble hydraulique

Ensembles évalués dans ce paragraphe

⇒ Comparaison des quantiles calibrés avec et sans prédicteur issu de statistiques de l'ensemble brut.

Q_R : quantiles calibrés à partir de l'ensemble hydraulique brut R_R. $\Longleftrightarrow$ QNE_{Md} : quantiles calibrés, mêmes prédicteurs que pour Q_R sauf les prédicteurs issus de statistiques de l'ensemble brut R_R.

Q_Q : quantiles calibrés à partir de l'ensemble hydraulique brut R_Q. $\Longleftrightarrow$ QNE_Q : quantiles calibrés, mêmes prédicteurs que pour Q_Q sauf les prédicteurs issus de statistiques de l'ensemble brut R_Q.

Dans ce paragraphe, nous cherchons à comparer les performances des ensembles calibrés Q_R et QNE_{Md} (respectivement Q_Q et QNE_Q). Les prédicteurs utilisés pour ces deux calibrations sont les mêmes, sauf que la calibration QNE_{Md} (respectivement QNE_Q) n'utilise pas comme prédicteur la moyenne et l'écart-type de l'ensemble brut Q_R (respectivement Q_Q).

À Kervir, les ensembles calibrés Q_R et QNE_{Md}, qui utilisent pour le prédicteur Q_{am} les prévisions émises par le modèle MORDOR-TS déterministe alimenté par RR3, présentent globalement les mêmes CRPS. De la même manière, les CRPS des ensembles calibrés Q_Q et QNE_Q, qui utilisent les prévisions hydrologiques calibrées, sont aussi du même ordre de grandeur, sauf pour QNE_Q à longue échéance en période de calage uniquement. Ainsi, à Kervir, l'utilisation du modèle hydraulique MASCARET qui fournit les prédicteurs *Moyenne* et *Ecart-type* n'est pas nécessaire pour améliorer les CRPS des ensembles calibrés. Une calibration réalisée uniquement à partir des historiques de données mesurées et des débits amont prévus apparaît aussi intéressante pour le CRPS. À Moulin-Vert en revanche, l'ensemble Q_R présente de meilleurs CRPS à longue échéance que l'ensemble QNE_{Md}, pour les périodes de calage et de validation. De même, l'ensemble Q_Q présente des meilleurs CRPS que QNE_Q à partir d'une certaine échéance (17 heures en calage et 19 heures en validation). Pour cette station, l'information fournie par les prédicteurs issus du modèle MASCARET permet d'améliorer la calibration pour les CRPS.

À Kervir, la fiabilité est améliorée de manière analogue sur la période P_{Cal} que l'on

tienne compte ou non de prédicteurs issus d'un ensemble hydraulique. À Moulin-Vert en revanche, les normes des histogrammes de rang des ensembles calibrés avec prédicteur issus du modèle hydraulique Q_R et Q_Q sont significativement plus faibles que celles des ensembles construits sans prédicteur issu du modèle hydraulique QNE_{Md} et QNE_Q sur P_{Cal}. Ces deux derniers ensembles présentent d'ailleurs une moins bonne fiabilité que l'ensemble brut R_Q pour les échéances à partir de +12h environ sur P_{Cal}. En période de validation, on note que les calibrations vont par paire : les quantiles Q_R et QNE_{Md} présentent les mêmes tendances, de même que les quantiles Q_Q et QNE_Q. Cela montre que le choix de l'hydrologie pour le prédicteur Q_{am} influence la fiabilité des ensembles calibrés, comme nous l'avons déjà évoqué en comparant les ensembles bruts. À Kervir, l'ajout de prédicteurs issus de statistiques de l'ensemble brut n'améliore pas la fiabilité des ensembles en validation. À Moulin-Vert, même si les courbes présentent les mêmes tendances, les calibrations avec prédicteurs issus de l'hydraulique présentent des normes pour les histogrammes de rang un peu plus faibles.

La tendance générale est une meilleure discrimination des événements pour les calibrations utilisant les prédicteurs issus de l'hydraulique à Kervir et Moulin-Vert. À Kervir, la calibration Q_Q permet d'obtenir les meilleures discriminations sauf sur P_{Val} pour le dépassement du quantile q_{99}. C'est alors la calibration Q_R qui est la meilleure.

Synthèse

À Kervir, l'utilisation de prédicteurs issus d'un ensemble hydraulique permet d'améliorer la discrimination, mais n'est pas indispensable pour les autres scores étudiés. À la station Moulin-Vert, l'utilisation de prédicteurs issus d'un ensemble hydraulique apporte une réelle plus-value pour tous les scores étudiés. Cette différence de comportement pour ces deux stations n'est pas expliquée à l'heure actuelle.

8.2.3.4 Meilleure stratégie à retenir

Ensembles évalués dans ce paragraphe

⇒ les 6 ensembles en période de validation

- R_R : ensemble hydraulique brut, condition limite amont *raw-HEF*.
- Q_R : quantiles calibrés à partir de l'ensemble hydraulique brut R_R.
- QNE_{Md} : quantiles calibrés, mêmes prédicteurs que pour Q_R sauf les prédicteurs issus de statistiques de l'ensemble brut R_R.
- R_Q : ensemble hydraulique brut, condition limite amont *QRF-hydro-TS*.
- Q_Q : quantiles calibrés à partir de l'ensemble hydraulique brut R_Q.
- QNE_Q : quantiles calibrés, mêmes prédicteurs que pour Q_Q sauf les prédicteurs issus de statistiques de l'ensemble brut R_Q.

Pour estimer la meilleure stratégie, nous évaluons ici les résultats des différents ensembles calibrés sur la période de validation uniquement.

À Kervir et Moulin-Vert, les ensembles qui présentent les meilleurs CRPS sont les quantiles calibrés Q_Q et QNE_Q (jusqu'à 19 heures pour ce dernier à Moulin-Vert). Ainsi, cela met en évidence qu'il est préférable d'utiliser de l'hydrologie calibrée pour obtenir des meilleurs CRPS. La meilleure stratégie est une donc calibration de l'hydrologie et de l'hydraulique. Les graphes mettent cependant en évidence le fait que c'est la calibration de l'hydrologie qui apporte le plus grand gain en termes de CRPS. La calibration de l'hydraulique améliore encore le CRPS, mais de manière moins importante.

En ce qui concerne la fiabilité, les ensembles Q_R et QNE_{Md} présentent les meilleures fiabilités à courte échéance pour les deux stations. Ceci signifie donc que l'ajout de prédicteurs relatifs à l'hydrologie calibrée dégrade la calibration statistique pour la fiabilité. Comme nous l'avons déjà mentionné, pour Kervir, cela peut être est en partie lié à un défaut de fiabilité de l'ensemble hydrologique calibré à Kerjean, puisque l'ensemble brut R_Q présente le même défaut de fiabilité. À longue échéance, ce sont les ensembles R_Q et dans une moindre mesure Q_Q et QNE_Q qui présentent les meilleures fiabilités à Kervir. À Moulin-Vert, les ensembles QNE_Q et Q_R présentent alternativement les meilleures normes d'histogrammes de rang. Il n'y a donc pas un unique ensemble qui minimise les normes

des histogrammes de rang pour toutes les échéances en validation à Kervir et Moulin-Vert. Il ne se dégage donc pas de stratégie de calibration en ce qui concerne l'optimisation de la fiabilité.

Pour les deux stations, l'ensemble Q_Q présente presque toujours les meilleures discriminations. Il présente de moins bonnes performances à Kervir pour l'événement *dépassement du seuil q99*. C'est alors la calibration Q_R qui est la meilleure. À Moulin-Vert, le second ensemble qui discrimine le mieux les événements après Q_Q est l'ensemble Q_R. Cela signifie que la stratégie de calibration à privilégier pour améliorer la discrimination des événements est une calibration QRF d'un ensemble hydraulique en utilisant des prédicteurs issus d'un ensemble brut généré avec le modèle MASCARET. Il est préférable que l'hydrologie ait été préalablement calibrée.

Synthèse

La meilleure stratégie pour la génération d'ensembles hydrauliques à Kervir et Moulin-Vert est une calibration à chaque maillon de la chaîne de prévision, c'est-à-dire calibrer l'hydrologie et l'hydraulique. Comme nous l'avons vu au paragraphe précédent, il est pertinent d'utiliser des prédicteurs issus de statistiques d'un modèle hydraulique.

8.2.3.5 Résultats avec une prévision de pluie parfaite

Ensemble évalué dans ce paragraphe

QNE_{Md} : quantiles calibrés, mêmes prédicteurs que pour Q_R sauf les prédicteurs issus de statistiques de l'ensemble brut R_R.

$\Rightarrow$ la calibration est réalisée avec le prédicteur Q_{am} valant le débit déterministe prévu par MORDOR-TS alimenté par les pluies parfaites ANTILOPE.

Nous avons souhaité étudier l'impact de la qualité du prédicteur Q_{am} dans la performance des ensembles générés. Pour cela, nous avons généré la calibration QNE_{Md} en modifiant le prédicteur Q_{am} : sa valeur est désormais égale au débit déterministe prévu par le modèle MORDOR-TS alimenté avec les mesures de pluie ANTILOPE à la place des prévisions RR3. Le prédicteur Q_{am} est supposé être de meilleure qualité. La figure 8.6 présente les résultats en période de validation pour les deux stations. Le tableau 8.7 présente le calcul des scores moyennés.

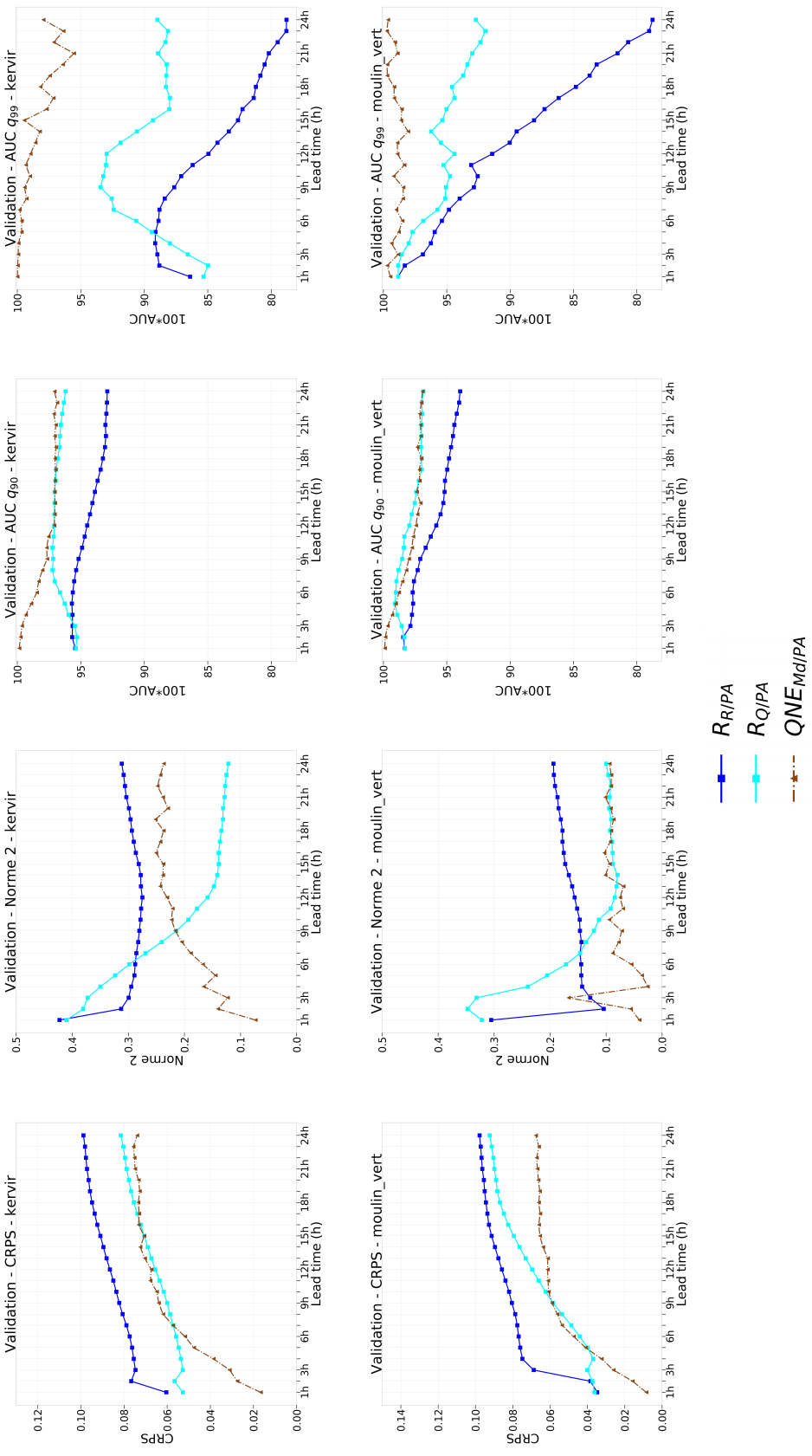


Figure 8.6: CRPS (col 1), $||.||_2$ (col 2), aire roc AUC pour les seuils q_{90} et q_{99} (col 3 et 4) pour la période de validation à Kervir (ligne 1) et à Moulin-Vert (ligne 2). Le prédicteur Q_{am} est construit à partir de prévisions MORDOR-TS en pluies parfaites ANTILOPE.

Score	$10^{-2}\|.\|_2$		100*CRPS		100 * AUC - q_{90}		100 * AUC - q_{99}	
Période	K	MV	K	MV	K	MV	K	MV
R_R	29.7	16.8	8.6	8.3	94.4	96.0	84.5	89.7
R_Q	20.8	14.2	6.6	6.7	96.6	97.9	89.6	95.3
QNE_{Md} - Qam RR3	17.4	8.4	6.1	6.8	97.2	96.9	96.7	94.2
QNE_{Md} - Qam ANTILOPE	20.8	8.2	6.1	5.5	97.8	97.9	98.5	99.0

Tableau 8.7: $\|.\|_2$, CRPS et aire ROC moyennés sur les 24 heures d'échéances à Kervir (K) et Moulin-Vert (MV). Période P_{Val}.

Ces résultats montrent que l'amélioration des prévisions de pluie permet d'améliorer les CRPS à Moulin-Vert de manière notable, et uniquement pour les plus longues échéances à Kervir par rapport à la calibration avec le prédicteur Q_{am} construit à partir des prévisions RR3. La fiabilité de l'ensemble calibré est analogue. La discrimination des événements de dépassement de seuil est améliorée aux deux stations, surtout à longue échéance. L'amélioration des prévisions de pluie permet donc d'améliorer la résolution des ensembles, mais la fiabilité n'est pas améliorée.

8.2.3.6 Importance des prédicteurs à Kervir et Moulin-Vert

Calibrations étudiées dans ce paragraphe

$\Rightarrow$ importance des prédicteurs pour les 4 ensembles calibrés

Q_R : quantiles calibrés à partir de l'ensemble hydraulique brut R_R.

QNE_{Md} : quantiles calibrés, mêmes prédicteurs que pour Q_R sauf les prédicteurs issus de statistiques de l'ensemble brut R_R.

Q_Q : quantiles calibrés à partir de l'ensemble hydraulique brut R_Q.

QNE_Q : quantiles calibrés, mêmes prédicteurs que pour Q_Q sauf les prédicteurs issus de statistiques de l'ensemble brut R_Q.

La méthode de construction des forêts aléatoires QRF permet d'évaluer *a posteriori* l'importance des prédicteurs utilisés pour la calibration. L'importance d'un prédicteur est caractérisée par l'augmentation de la valeur de la RMSE de la forêt lorsque ce prédicteur est permuté aléatoirement avec remplacement : les valeurs du prédicteur sont alors un réarrangement aléatoire des valeurs de départ. Les figures 8.7 et 8.8 présentent l'importance des prédicteurs pour les échéances 6 et 18 heures pour les ensembles calibrés Q_R, Q_Q, QNE_{Md} et QNE_Q, respectivement à Kervir et Moulin-Vert.

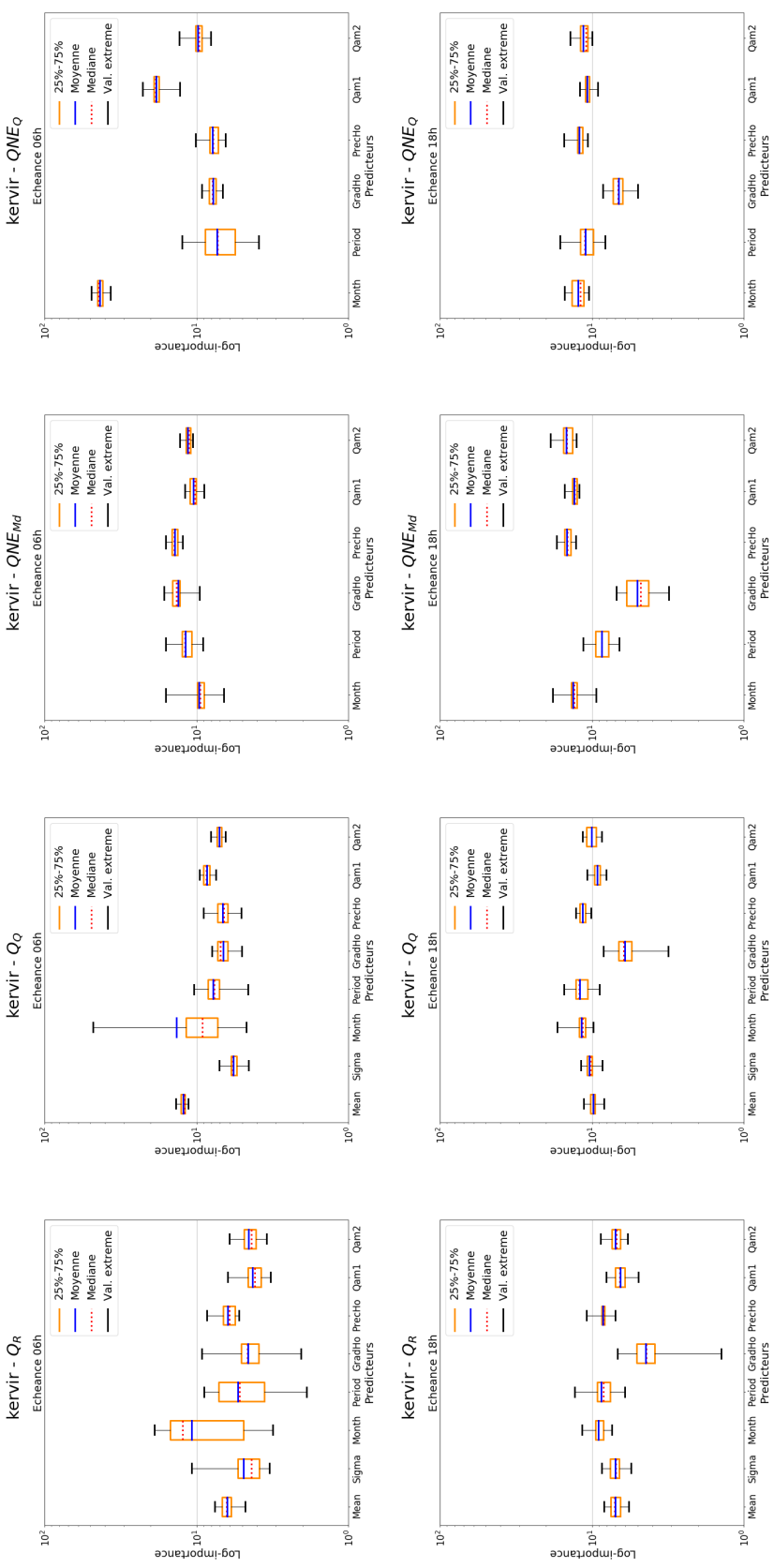


Figure 8.7: Importance des prédicteurs à Kervir pour les échéances 6 heures (ligne 1) et 18 heures (ligne 2) : Q_R (col 1), Q_Q (col 2), QNE_{Md} (col 3) et QNE_Q (col 4).

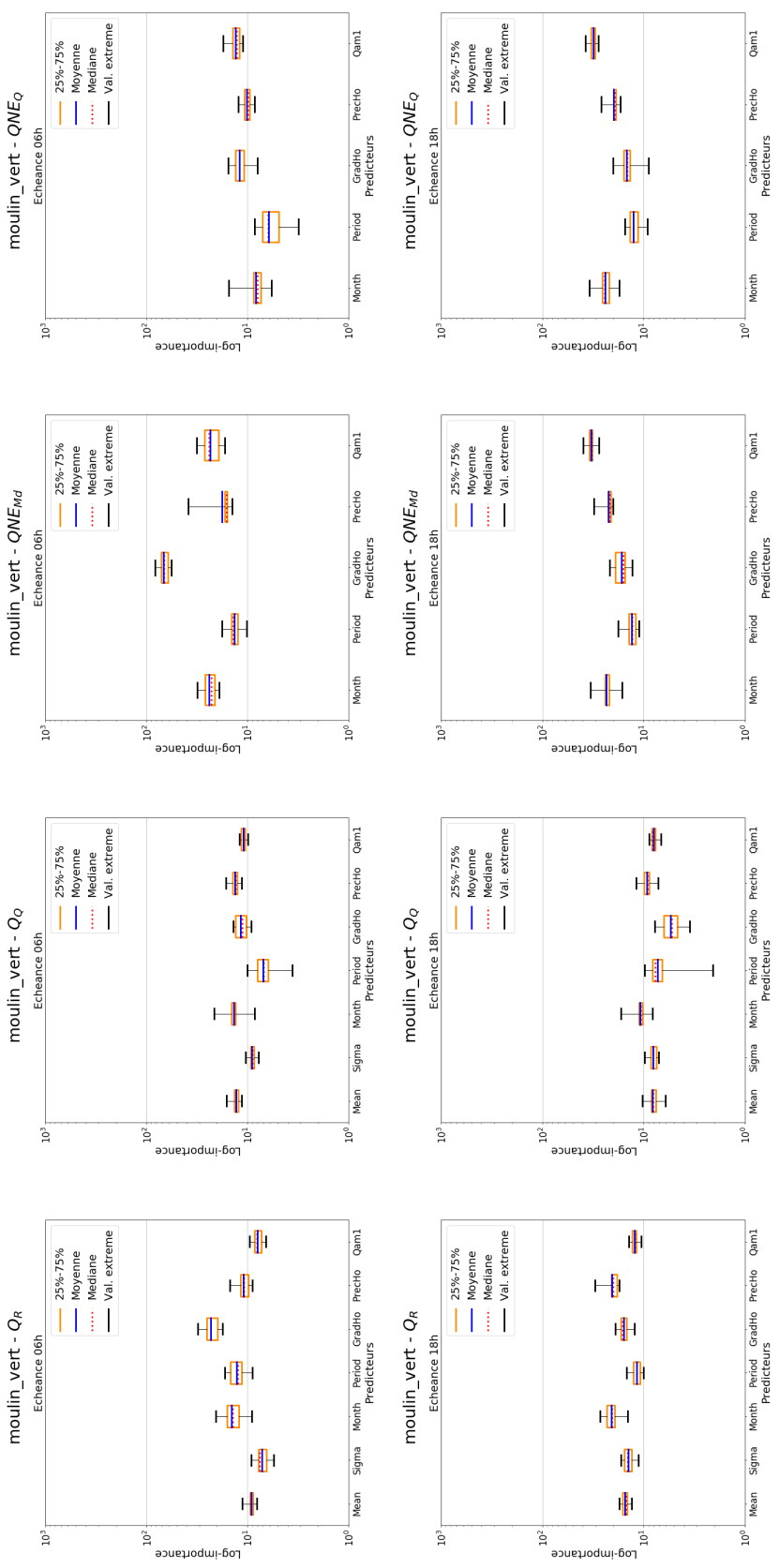


Figure 8.8: Importance des prédicteurs à Moulin-Vert pour les échéances 6 heures (ligne 1) et 18 heures (ligne 2) : Q_R (col 1), Q_Q (col 2), QNE_{Md} (col 3) et QNE_Q (col 4).

Pour chaque station et chaque échéance de calibration, il se dégage deux groupes pour lesquels la hiérarchisation de l'importance des prédicteurs est la même. En effet, on note les mêmes tendances concernant l'importance des prédicteurs pour les calibrations Q_R et Q_Q, et pour les calibrations QNE_{Md} et QNE_Q. Cela met en évidence que pour une échéance de prévision donnée, l'importance relative des prédicteurs est en premier lieu liée à l'ensemble de prédicteurs utilisés. Les différentes manières dont ils ont été construits importe peu. Cela signifie que l'importance relative des prédicteurs *Moyenne* et *Ecart-type* n'est pas liée au fait que l'ensemble hydraulique soit construit avec une hydrologie calibrée ou non, et l'importance relative du prédicteur Q_{am} n'est pas liée au fait que l'on utilise le modèle MORDOR-TS déterministe ou la prévision d'ensemble pour le construire.

D'une manière générale, les graphiques mettent également en évidence que le prédicteur *GradHo*, représentant le gradient de hauteur mesurée au temps de base, perd de l'importance avec l'augmentation de l'échéance de prévision. Ce prédicteur représente la dynamique de l'hydrogramme au moment du lancement de la prévision. Il est normal que celui-ci ait beaucoup d'importance aux courtes échéances, et qu'il soit beaucoup moins informatif aux longues échéances.

En-dehors de ces conclusions, il ne se dégage pas de tendance générale pour l'importance des prédicteurs selon les stations, les échéances et les types de calibrations.

8.2.4 Synthèse

Dans cette section, après avoir présenté les prédicteurs utilisés pour la construction des forêts aléatoires à Kervir et Moulin-Vert, nous avons discuté de la stratégie d'optimisation de ces forêts. Nous avons fait le choix d'utiliser un critère d'optimisation qui est un compromis entre la minimisation du CRPS et de la norme des histogrammes de rang des ensembles calibrés. Nous avons vu qu'en acceptant de dégrader légèrement les CRPS en période de validation, la fiabilité est plus nettement améliorée. Cette stratégie s'inscrit dans les préconisations de Wilks [2018a], qui proposent de ne pas utiliser uniquement la minimisation du CRPS pour l'optimisation des paramètres de post-traitement, mais d'y ajouter une pénalité en cas de manquement à la fiabilité.

Nous résumons ci-après les principaux résultats obtenus. Tout d'abord, les ensembles hydrauliques bruts présentent globalement de meilleures performances en termes de résolution et fiabilité si l'hydrologie a été calibrée au préalable. Ceci n'est cependant pas vrai pour la fiabilité aux courtes échéances en période de validation uniquement. Ce point est peut-être lié à une moins bonne fiabilité des ensembles bruts hydrologiques en période de validation à la station Kerjean. La calibration de l'hydrologie paraît donc indispensable pour les performances des ensembles hydrauliques bruts.

Nous avons également vu que les calibrations statistiques des ensembles hydrauliques bruts

mises en place permettent d'améliorer les performances des ensembles en termes de CRPS et de discrimination. En ce qui concerne la fiabilité des ensembles calibrés, les résultats ne sont pas consistants entre les périodes P_{Cal} et P_{Val} à Kervir. Il ne se dégage pas de stratégie de calibration quant à l'amélioration de la fiabilité des ensembles calibrés sur cette station. Par ailleurs, il a été mis en évidence que les performances de l'ensemble hydrologique choisi pour générer l'ensemble hydraulique et/ou pour les prédicteurs de la calibration QRF impactent directement la fiabilité des ensembles hydrauliques construits.

À Kervir, l'utilisation de prédicteurs issus du modèle MASCARET ne semble pas indispensable pour avoir de meilleurs CRPS. À Moulin-Vert, l'utilisation d'un ensemble hydraulique généré avec MASCARET n'est pas indispensable pour les CRPS jusqu'à l'échéance de prévision +18h. En revanche, pour les deux stations, l'utilisation de prédicteurs issus du modèle hydraulique est indispensable pour obtenir une meilleure discrimination des événements de dépassement de seuil. Il n'est donc pas pertinent d'envisager une stratégie fondée uniquement sur l'apprentissage statistique pour générer des ensembles calibrés ici. Il faut de plus garder à l'esprit qu'il n'est pas possible de reconstruire une chronique de prévision si l'on ne dispose pas des ensembles bruts hydrauliques, puisque c'est de celui-ci dont proviendra le motif permettant de faire la reconstruction.

Un des objectifs de cette étude est de définir quelle est la meilleure stratégie pour une application opérationnelle. Nous avons vu que la meilleure stratégie est celle pour laquelle une calibration est effectuée à chaque étape de la chaîne de prévision, c'est-à-dire que l'on calibre les ensembles hydrologiques et hydrauliques. Cela correspond à l'ensemble calibré Q_Q. Nous avons vu que le gain le plus important en ce qui concerne le CRPS est lié à la calibration de l'hydrologie.

Un test a été effectué pour appréhender l'importance du prédicteur Q_{am}. Nous avons vu que les performances des ensembles en termes de CRPS et discrimination des événements sont améliorées lorsque le prédicteur Q_{am} est construit avec le modèle MORDOR-TS déterministe lancé avec les mesures de pluie ANTILOPE. Cela signifie que l'amélioration des prévisions de pluie est un point clé pour une bonne prévision d'ensembles des hauteurs d'eau à Kervir et Moulin-Vert. Cependant, cela ne permet pas d'améliorer la fiabilité des ensembles en prévision.

L'étude de l'importance des prédicteurs a mis en évidence le fait que le gradient de la hauteur d'eau mesurée au temps de base perd de l'importance avec l'échéance de prévision. Nous n'avons pas mis en évidence d'autres tendances significatives.

L'ensemble des conclusions émises dans ce paragraphe sont naturellement à modérer. Les résultats ont été obtenus sur un seul bassin versant, et nous avons déjà mis en évidence des comportements légèrement différents pour les résultats aux deux stations Kervir et Moulin-Vert, sans que nous ayons pu identifier les origines de ces différences. Il serait pertinent d'appliquer cette méthodologie sur d'autres bassins versants afin de dégager éventuellement

des tendances générales pour la calibration statistique des ensembles hydrauliques.

8.3 Calibration QRF à Justice

Pour une bonne lecture du manuscrit, nous commençons ce chapitre par un rappel des principales caractéristiques de la calibration effectuée à la station Justice. Tout d'abord, ce sont les pics de marée haute qui sont calibrés. Nous rappelons que cela implique qu'aucune stratification n'est nécessaire. Une autre conséquence est que les prédicteurs ne sont donc pas les mêmes que pour les stations Kervir et Moulin-Vert. De plus, puisque nous n'avons pas de prévision d'ensemble de surcote à Plaisance qui permette de prendre en compte l'incertitude sur les niveaux à l'aval, trois expériences avec différentes prévisions à l'aval sont réalisées (voir paragraphe 8.1.1.1) : prévision de marée astronomique (PA), prévision de marée astronomique avec persistance de la surcote mesurée au temps de base ($PASC$), et pseudo-prévision d'ensemble construite par perturbation *a posteriori* des mesures (MP).

Le paragraphe 8.3.1 présente les prédicteurs utilisés. Dans le paragraphe 8.3.2, les deux critères d'optimisation des forêts aléatoires sont comparés. Les résultats comparés des trois expériences E_{PA}, E_{PASC} et E_{MP} réalisées avec le critère retenu sont présentées au paragraphe 8.3.3. Le paragraphe 8.3.4 présente la synthèse des résultats obtenus pour la calibration à Justice.

8.3.1 Prédicteurs

Le tableau 8.8 présente les prédicteurs utilisés pour la calibration des marées hautes à Justice. Les spécificités liées à l'une ou l'autre des trois expériences E_{PAS}, E_{PASC} ou E_{MP} sont indiquées en bleu. Pour mémoire, de même qu'aux stations Kervir et Moulin-Vert, chacune de ces trois expériences permet de générer six prévisions d'ensemble : deux prévisions d'ensemble brutes et quatre prévisions calibrées (voir paragraphe 8.1.3).

Nom	Unité	Description
Mean	m	Moyenne de l'ensemble brut
Sigma	m	Ecart-type de l'ensemble brut
Month		Mois de la prévision
Period		Période de la journée de la prévision (0h-6h, 6h-12h, 12h-18h, 18h-24h)
precPicJu	m	Hauteur de la marée haute précédant le temps de base à Justice
mareePicPl	m	Hauteur de la pleine mer astronomique prévue à Plaisance générant le pic à Justice (prévision SHOM) Pour E_{PASC}, on ajoute la surcote mesurée au temps de base
Q_{am}	m^3/s	Débit mesuré/prévu (MORDOR déterministe) à Tréodet 2 heures plus tôt
nivPicPl	m	Moyenne de l'ensemble des hauteurs de pleine mer prévues à Plaisance générant le pic à Justice

Tableau 8.8: Prédicteurs utilisés pour la calibration QRF de la station Justice sous influence maritime.

Dans le cas de l'expérience E_{PASC}, le prédicteur *mareePicPl* correspond à la prévision de la pleine mer astronomique à laquelle on ajoute la surcote mesurée au temps de base.

Concernant le prédicteur Q_{am}, nous avons choisi de prendre en compte le débit mesuré ou prévu à Tréodet 2 heures avant le moment calibré. En effet, l'étude de sensibilité a montré que pour les événements fluviaux ou fluviaux-maritimes, l'impact du débit amont se fait un peu sentir au moment des pleines mers à Justice. Il est donc pertinent de prendre en compte cette information dans les prédicteurs. De plus, il peut aussi y avoir un lien entre la valeur des débits amont et une éventuelle surcote/décote à la condition limite aval maritime. On fait le choix de ne prendre en compte que le débit de l'affluent le plus important, à un instant un peu arbitraire puisqu'on ne sait pas quel est le moment le plus pertinent pour la corrélation avec les surcotes/décotes. La durée de 2 heures considérée ici correspond approximativement au temps de propagation entre Tréodet et Kervir. Le choix pour le type de débit prévu est le même que pour la calibration à Kervir et Moulin-Vert : soit la prévision déterministe alimentée par RR3, soit la moyenne de la prévision calibrée *QRF-hydro-TS*.

Pour finir, le prédicteur *nivPicPl* correspondant à la moyenne des hauteurs de pleine mer prévues à Plaisance n'est disponible que lorsqu'une prévision d'ensemble est disponible à Plaisance. Il n'est utilisé ici que pour l'expérience E_{MP} pour laquelle une pseudo-prévision d'ensemble a été construite. Par construction, ce prédicteur a la même valeur quelle que soit l'échéance de prévision calibrée. Avec une vraie prévision d'ensemble, on s'attend au contraire à une diminution des performances de l'ensemble de surcotes avec l'échéance de prévision. Pour cette expérience, les prédicteurs *Month*, *Period*, *mareePicPl* et *nivPicPl* seront identiques pour toutes les échéances de calibration. Le prédicteur *precPicJu* sera le même pour des temps de base couvrant une période de 12 heures environ.

8.3.2 Comparaison des deux critères pour le choix de la meilleure forêt

L'étude de la comparaison des deux critères d'optimisation des forêts aléatoires est présentée en annexe Q.2.1. Pour mémoire, ces deux critères sont : optimisation du CRPS uniquement (Opt_{CRPS}) et optimisation de la fiabilité au préalable (Opt_{Rel}). Cette étude met en évidence les conclusions suivantes :

- La stratégie d'optimisation de la fiabilité Opt_{Rel} conduit, pour certaines échéances isolées, à des valeurs pour les CRPS ou pour les aires des courbes ROC complètement en-dehors de la tendance générale pour les autres échéances. Pour ces points particuliers, la stratégie n'est donc pas adaptée. Il serait possible de l'améliorer en conservant dans un premier temps davantage que seulement 3 forêts optimisés pour la fiabilité. Le CRPS serait ensuite optimisé dans cet ensemble.
- En période de calage, la stratégie Opt_{CRPS} fournit les meilleurs CRPS et des aires

pour les courbes ROC un peu meilleures, et Opt_{Rel} fournit les plus petites normes pour les histogrammes de rang des ensembles calibrés.

- En période de validation, la stratégie Opt_{CRPS} fournit toujours les meilleurs CRPS et les meilleures discriminations pour l'événement de dépassement de seuil. Cette stratégie fournit également des scores équivalents ou légèrement meilleurs pour la norme des histogrammes de rang.

La stratégie Opt_{CRPS} présentant des meilleures performances en validation, c'est celle-ci qui est retenue pour la calibration des pics de marée haute à Justice.

8.3.3 Résultats comparés des 3 expériences E_{PA}, E_{PASC} et E_{MP}

La figure 8.9 représente les valeurs des CRPS, la norme $\|.\|_2$ des histogrammes de rang et l'aire AUC de la courbe ROC pour le dépassement du seuil vert-jaune pour les six ensembles issus de l'expérience E_{PA} ($R_{R/PA}$, $R_{Q/PA}$, $Q_{R/PA}$, $Q_{Q/PA}$, $QNE_{Md/PA}$, $QNE_{Q/PA}$) pour les périodes de calage et de validation. Les figures 8.10 et 8.11 sont analogues pour les expériences E_{PASC} et E_{MP} respectivement. Les tableaux 8.9, 8.10 et 8.11 présentent la moyenne sur toutes les échéances des scores présentés sur les graphiques pour les trois expériences, pour les périodes de calage et de validation. Dans l'annexe Q.2.2, les histogrammes de rang pour les échéances 6, 12, 18 et 24 heures sont présentés pour les périodes de calage et de validation.

Score	$10^{-2}\|.\|_2$		100*CRPS		100 * AUC - seuil V/J	
Période	Cal	Val	Cal	Val	Cal	Val
$R_{R/PA}$	52.4	49.1	12.1	11.8	71.1	79.8
$R_{Q/PA}$	52.4	49.2	12.5	11.8	71.1	79.8
$Q_{R/PA}$	5.0	10.8	8.0	8.0	95.7	97.5
$Q_{Q/PA}$	4.7	10.9	8.2	8.0	96.0	97.8
$QNE_{Md/PA}$	6.6	11.3	8.1	7.9	96.4	98.0
$QNE_{Q/PA}$	6.8	11.3	8.0	7.9	96.9	98.0

Tableau 8.9: $\|.\|_2$, CRPS et aire ROC moyennés sur les 24 heures d'échéances à Justice. Expérience E_{PA}.

Score	$10^{-2}\|.\|_2$		100*CRPS		100 * AUC - seuil V/J	
Période	Cal	Val	Cal	Val	Cal	Val
$R_{R/PASC}$	39.3	45.7	6.7	6.3	89.5	92.6
$R_{Q/PASC}$	39.3	45.3	6.7	6.2	89.5	92.6
$Q_{R/PASC}$	5.8	11.5	5.6	4.9	96.8	99.1
$Q_{Q/PASC}$	5.6	11.4	5.6	4.9	96.7	99.0
$QNE_{Md/PASC}$	5.8	10.4	5.9	5.3	96.0	99.0
$QNE_{Q/PASC}$	5.9	10.3	5.9	5.3	96.2	99.2

Tableau 8.10: $\|.\|_2$, CRPS et aire ROC moyennés sur les 24 heures d'échéances à Justice. Expérience E_{PASC}.

Score	$10^{-2}\|.\|_2$		100*CRPS		100 * AUC - seuil V/J	
Période	Cal	Val	Cal	Val	Cal	Val
$R_{R/MP}$	30.2	32.3	4.7	4.6	99.9	99.9
$R_{Q/MP}$	30.1	32.5	4.8	4.6	99.9	99.9
$Q_{R/MP}$	7.4	12.2	2.6	2.1	99.8	99.9
$Q_{Q/MP}$	7.4	13.2	2.6	2.1	99.8	99.9
$QNE_{Md/MP}$	7.4	12.9	2.7	2.1	99.9	99.9
$QNE_{Q/MP}$	7.3	12.8	2.6	2.1	99.9	99.9

Tableau 8.11: $\|.\|_2$, CRPS et aire ROC moyennés sur les 24 heures d'échéances à Justice. Expérience E_{MP}.

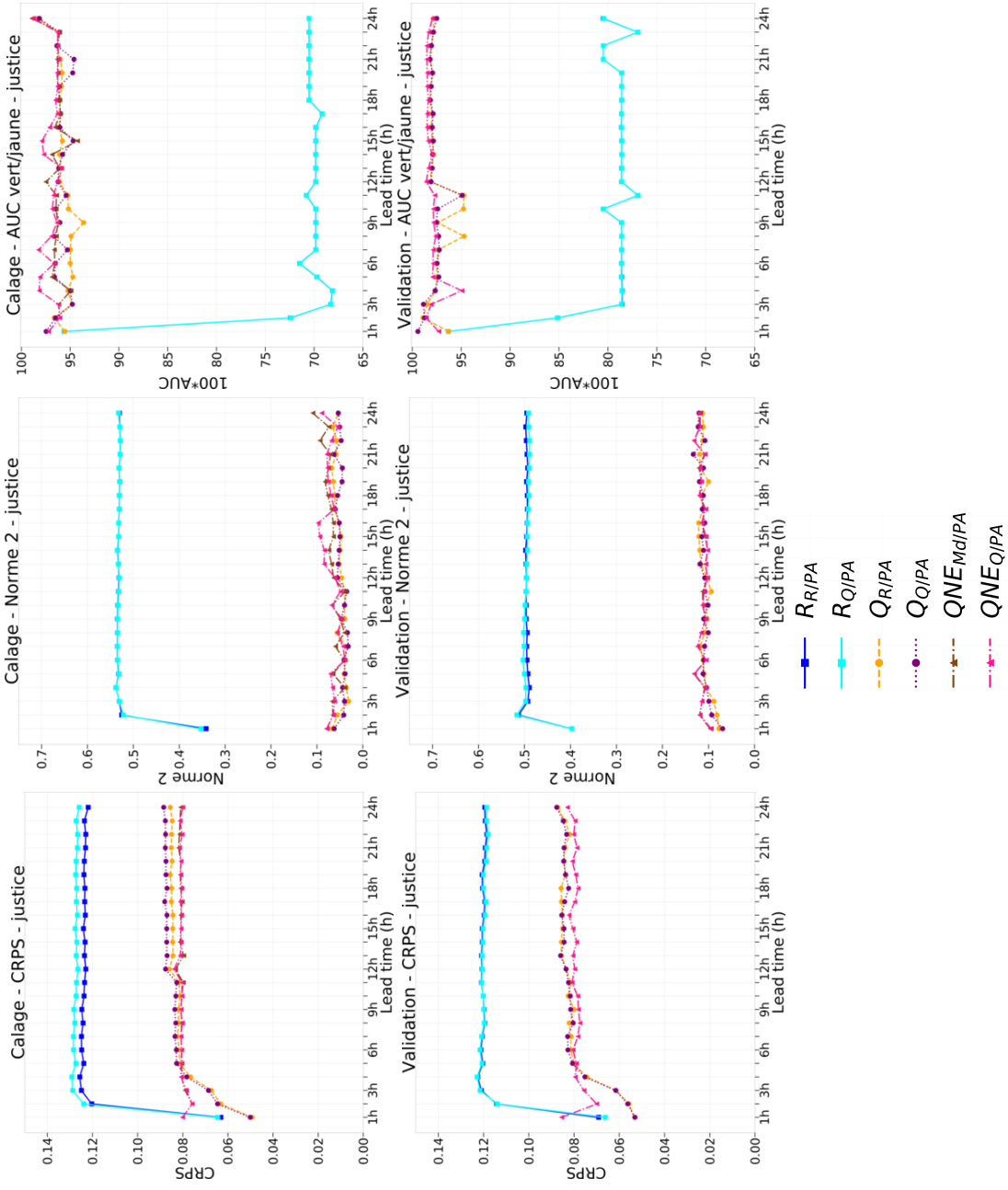


Figure 8.9: CRPS (col 1), $||.||_2$ (col 2), aire roc AUC pour le seuil passage vert-jaune (col 3) pour la calibration des pics de marée haute à Justice pour les périodes de calage (ligne 1) et de validation (ligne 2). Expérience E_{PA}.

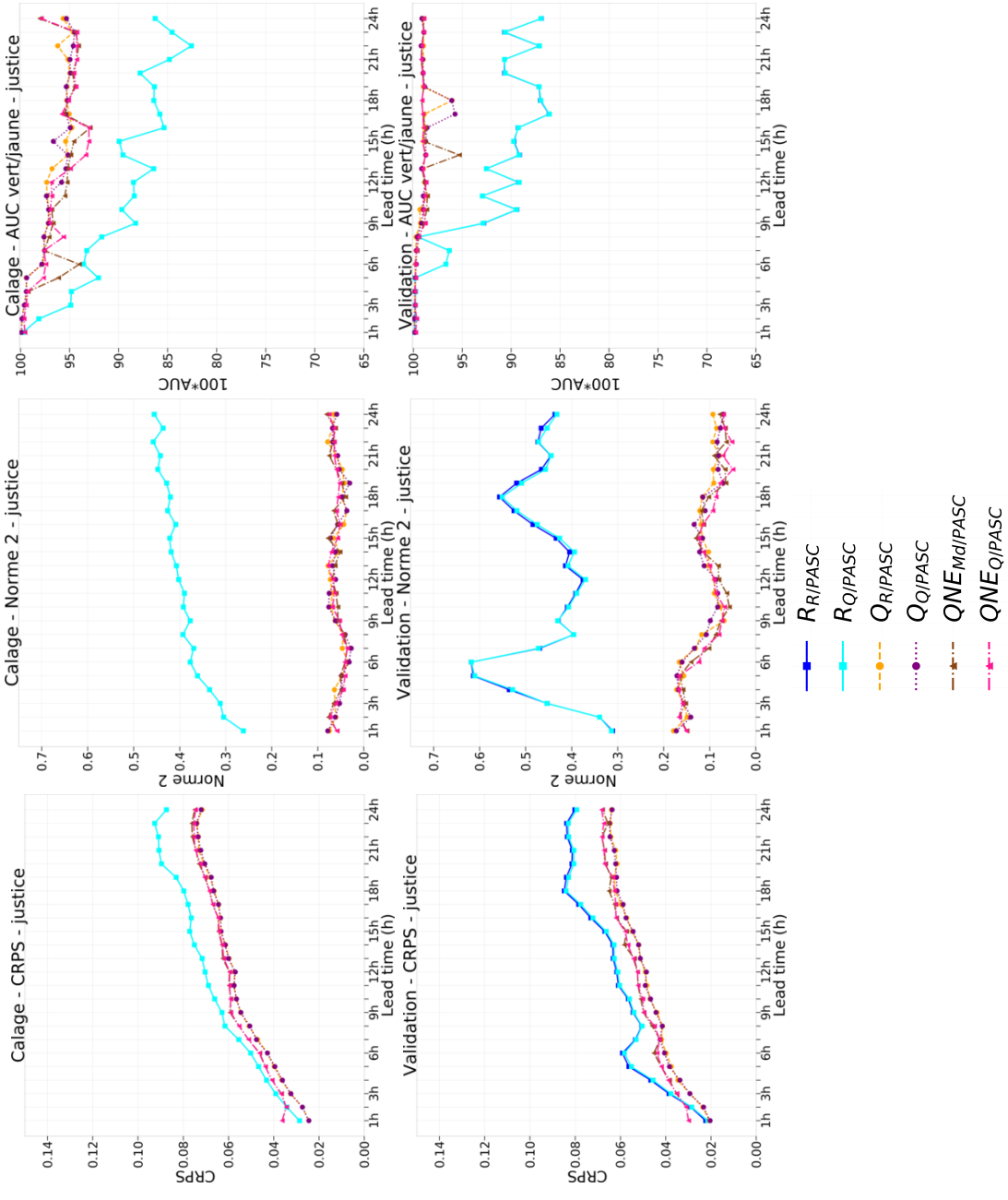


Figure 8.10: CRPS (col 1), n2 (col 2), aire roc pour le seuil passage vert-jaune (col 3) pour la calibration des pics de marée haute à Justice pour les périodes de calage (ligne 1) et de validation (ligne 2). Expérience E_{PASC}.

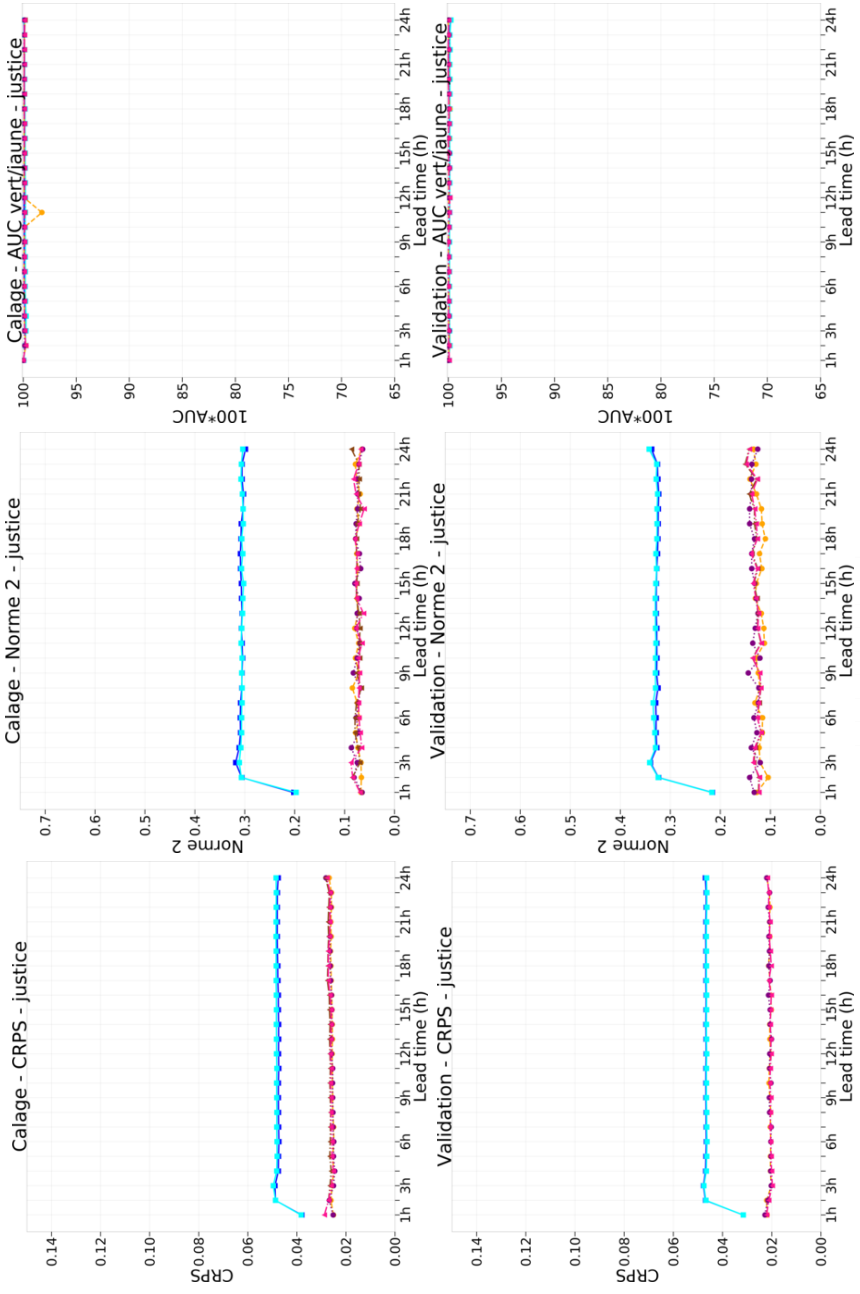


Figure 8.11: CRPS (col 1), $||\cdot||_2$ (col 2), aire roc AUC pour le seuil passage vert-jaune (col 3) pour la calibration des pics de marée haute à Justice pour les périodes de calage (ligne 1) et de validation (ligne 2). Expérience E_{MP}.

8.3.3.1 Analyse des ensembles bruts

Ensembles évalués dans ce paragraphe

⇒ les 2 ensembles bruts pour chaque expérience :

- E_{PA} : prévision astronomique
- E_{PASC} : prévision astronomique, surcote constante
- E_{MP} : mesure perturbée

R_R : ensemble hydraulique brut, condition limite amont *raw-HEF*.

R_Q : ensemble hydraulique brut, condition limite amont *QRF-hydro-TS*.

Nous analysons dans ce paragraphe les courbes R_R et R_Q associées aux trois expériences menées E_{PA}, E_{PASC} et E_{MP}. L'intérêt de la calibration préalable de l'ensemble hydrologique sera discuté, puis les performances relatives des différents ensembles bruts seront analysées au regard des trois scores étudiés (CRPS, norme de l'histogramme de rang et aire AUC sous la courbe ROC).

Calibration de l'hydrologie pour générer les ensembles bruts hydrauliques Pour les trois expériences, que l'on se place en période de calage ou de validation, les scores comparés des ensembles bruts R_R et R_Q sont vraiment très proches, voire identiques. Cela signifie que la calibration de l'hydrologie n'a pas de conséquence sur l'ensemble des marées hautes simulées à Justice. Ceci est cohérent avec l'étude de sensibilité qui a mis en évidence le fait que les hauteurs de pleine mer à Justice étaient peu influencées par le choix de la chronique pour la condition limite amont. Ainsi, le fait d'imposer un ensemble hydrologique calibré ou non à l'amont n'a pas d'importance notable pour les ensembles bruts à Justice. Il n'y a donc pas d'intérêt à calibrer l'hydrologie quand l'objectif est la calibration des hauteurs de pleine mer à Justice.

CRPS Dans un premier temps, nous analysons la tendance des courbes des CRPS des ensembles hydrauliques bruts R_R et R_Q pour les trois expériences. L'analyse des graphiques met en évidence que la courbe des valeurs des CRPS par échéance est plate à partir de l'échéance 2 heures pour les ensembles bruts des expériences E_{PA} et E_{MP} (figures 8.9 et 8.11). En revanche, elle est croissante pour les ensembles bruts issus de l'expérience E_{PASC} (figure 8.10). Ceci s'explique par le fait que pour les deux expériences E_{PA} et E_{MP}, la condition limite maritime présente les mêmes caractéristiques en termes d'erreur pour tous les pas de temps de prévision. Dit d'une autre manière, cela traduit le

fait que la condition limite imposée en prévision n'est pas modifiée par le choix du temps de base. En effet, pour l'expérience E_{PA}, le modèle est alimenté par les prévisions de marée astronomique à partir du temps de base. Pour l'expérience E_{MP}, le modèle est alimenté par les pseudo-prévisions d'ensemble, qui correspondent à des chroniques perturbées des niveaux à Plaisance. Ces chroniques sont générées de manière continue sur toute la période de calage P_{Cal} ou de validation P_{Val} et ne sont pas impactées par le choix d'un temps de base. Pour ces deux expériences, le premier pas de temps présente des meilleurs CRPS. Ceci s'explique par le fait que les prévisions de hauteur à échéance 1 heure à Justice correspondent à la remontée de la condition limite aval à Plaisance au cours de l'heure précédente, la durée de propagation des pics entre Plaisance et Justice étant comprise le plus souvent entre 30 et 45 minutes. Les mesures de hauteur à Plaisance imposées dans le modèle jusqu'au temps de base sont toujours présentes dans le modèle et apportent ainsi une information de qualité à Justice à l'échéance 1 heure, d'où des meilleurs CRPS à cette échéance. Pour l'expérience E_{PASC}, le comportement est différent. Les valeurs des CRPS des ensembles bruts sont croissantes avec les échéances de prévision. Cela traduit le fait que la qualité de la surcote ajoutée à la prévision de marée astronomique décroit avec l'échéance de prévision. Ceci est cohérent avec le fait que les phénomènes de surcote/décote possèdent une auto-corrélation temporelle.

Après avoir analysé la forme des courbes de CRPS avec les échéances, nous nous intéressons à présent aux valeurs comparées des CRPS des ensembles bruts R_R et R_Q pour les trois expériences. Nous pouvons noter ici que les CRPS des ensembles bruts sont légèrement meilleurs sur la période de validation que sur celle de calage. Les CRPS les plus faibles sont obtenus pour l'expérience E_{MP}, et les plus forts pour l'expérience E_{PA}. Ces résultats peuvent être expliqués par la qualité de la prévision à Plaisance, qui est bien meilleure pour E_{MP} puisqu'elle est basée sur une perturbation de la mesure. L'expérience E_{PASC} intègre la surcote mesurée au temps de base et présente des CRPS un peu dégradés par rapport à E_{MP}. L'expérience E_{PA} ne tient pas compte des surcotes/décotes. Cela ne paraît pas être une bonne stratégie pour la construction des ensembles bruts : nous notons que les CRPS de l'expérience E_{PASC} sont meilleurs que ceux de l'expérience E_{PA} à toutes les échéances de prévision. Cela signifie que l'ajout de la surcote mesurée au temps de base à la prévision de marée astronomique permet d'améliorer les CRPS pour toutes les échéances jusqu'à 24 heures. Jusqu'à l'échéance 4 heures environ, les CRPS des ensembles bruts issus de l'expérience E_{PASC} sont même meilleurs que ceux fournis par l'expérience E_{MP}. Sur la période de validation, les CRPS des ensembles bruts pour l'expérience E_{PASC} présentent un caractère cyclique, avec des pics de CRPS pour les échéances 5-6 heures et 18 heures. S'il n'est pas surprenant d'obtenir des caractéristiques cycliques liées au phasage de la marée, il est plus surprenant de ne pas les observer également sur la période de calage. Nous n'avons pas su expliquer cette différence de comportement.

$\|.\|_2$ des histogrammes de rang Les formes des courbes de la norme des histogrammes de rang pour les ensembles bruts R_R et R_Q présentent les mêmes caractéristiques que celle du CRPS : les expériences E_{PA} et E_{MP} présentent des courbes plates à partir de l'échéance 2 heures, contrairement à l'expérience E_{PASC} qui présente une tendance croissante. Les explications sont analogues également. Nous notons également comme pour le CRPS que la valeur des normes est cyclique en période de validation. Les pics se retrouvent aux mêmes échéances, à savoir 5-6 heures et 18 heures. Sur les deux périodes P_{Cal} et P_{Val}, la meilleure fiabilité des ensembles bruts est obtenue pour l'expérience E_{MP}. La fiabilité des ensembles bruts de l'expérience E_{PASC} est meilleure que celle de l'expérience E_{PA}.

Discrimination seuil vert-jaune La discrimination pour le passage du seuil vert-jaune est meilleure pour les ensembles bruts issus de l'expérience E_{PASC} que de l'expérience E_{PA}. La discrimination est excellente pour l'expérience E_{MP}.

Synthèse

Nous avons mis en évidence qu'il n'est pas utile d'utiliser un ensemble hydrologique calibré pour la génération de l'ensemble brut des pics de pleine mer à Justice, quelle que soit l'expérience considérée (E_{PA}, E_{PASC} ou E_{MP}).
Les performances des ensembles bruts issus de l'expérience E_{MP} sont meilleures pour tous les scores étudiés, mais cette expérience est uniquement réalisée à titre expérimental ici puisqu'elle n'est pas réalisable en opérationnel. En revanche, les deux expériences E_{PA} et E_{PASC} nous intéressent particulièrement car elles sont réalisables en opérationnel. Les résultats mettent en évidence un net avantage à ajouter la surcote mesurée au temps de base pour tous les critères étudiés pour les ensembles bruts R_R et R_Q. D'une manière générale, nous pouvons conclure que la qualité des ensembles bruts est d'autant meilleure que l'information sur la surcote prévue à Plaisance est de bonne qualité.

8.3.3.2 Caractéristiques des ensembles calibrés par rapport aux ensembles bruts correspondants

Dans ce paragraphe, nous comparons pour chaque expérience (E_{PA}, E_{PASC} ou E_{MP}) les performances des ensembles hydrauliques calibrés Q_R, Q_Q, QNE_{Md} et QNE_Q par rapport à celles des ensembles bruts correspondants. La comparaison des performances respectives des ensembles calibrés avec les trois expériences sera réalisée au paragraphe 8.3.3.3.

Ensembles évalués dans ce paragraphe

⇒ Comparaison des quantiles calibrés par rapport aux ensembles bruts correspondants pour chaque expérience prise indépendamment :

- E_{PA} : prévision astronomique
- E_{PASC} : prévision astronomique, surcote constante
- E_{MP} : mesure perturbée

Comparaison des trois ensembles issus de l'hydrologie non calibrée :

R_R : ensemble hydraulique brut, condition limite amont *raw-HEF*.

Q_R : quantiles calibrés à partir de l'ensemble hydraulique brut R_R.

QNE_{Md} : quantiles calibrés, mêmes prédicteurs que pour Q_R sauf les prédicteurs issus de statistiques de l'ensemble brut R_R.

Comparaison des trois ensembles issus de l'hydrologie calibrée *QRF-hydro-TS* :

R_Q : ensemble hydraulique brut, condition limite amont *QRF-hydro-TS*.

Q_Q : quantiles calibrés à partir de l'ensemble hydraulique brut R_Q.

QNE_Q : quantiles calibrés, mêmes prédicteurs que pour Q_Q sauf les prédicteurs issus de statistiques de l'ensemble brut R_Q.

Pour les trois expériences menées, toutes les calibrations statistiques mises en place génèrent des ensembles avec des meilleures performances en termes de CRPS, norme et discrimination du seuil de vigilance vert-jaune que les ensembles brut correspondants. Ceci est valable aussi bien sur la période de calage que de validation. Une calibration en bout de chaîne de prévision est donc indispensable pour la prévision d'ensemble des pics de marée haute à Justice. Par ailleurs, les résultats montrent que les ensembles Q_R et Q_Q sont équivalents. Ceci est évidemment lié au fait que les ensembles bruts R_R et R_Q sont quasiment identiques. De même, les ensembles QNE_{Md} et QNE_Q présentent des performance similaires. Tout ceci est cohérent avec le fait que les pics de marée haute à Justice sont très peu influencés par les débits amont.

CRPS Pour chaque expérience, les CRPS sont corrigés de manière semblable sur les périodes de calibration et de validation, ce qui montre une bonne robustesse des forêts aléatoires construites pour ce critère. Pour chaque expérience également, la forme de la courbe des CRPS calibrés garde la même tendance que celle des ensembles bruts : plate pour les expériences E_{PA} (après l'échéance 4 heures) et E_{MP}, croissante pour l'expérience E_{PASC}. D'une manière générale, pour une expérience donnée, les calibrations obtenues avec les quatre ensembles Q_R, Q_Q, QNE_{Md} et QNE_Q sont équivalentes en termes de valeur de CRPS. Pour l'expérience E_{PASC}, les CRPS calibrés sont très légèrement meilleurs avec l'utilisation du modèle hydraulique (ensembles $Q_{R/PASC}$ $Q_{Q/PASC}$) par rapport aux calibrations construites sans information issue du modèle hydraulique.

$\|.\|_2$ **des histogrammes de rang** Pour les trois expériences, les normes des histogrammes de rang des ensembles calibrés sont significativement supérieures en période de validation par rapport à la période de calage. Les forêts aléatoires présentent donc moins de robustesse pour la fiabilité que pour le CRPS. Cependant, il est important de noter que la fiabilité reste encore grandement améliorée par rapport aux ensembles bruts en période de validation. La fiabilité des ensembles calibrés sans utilisation d'information issue du modèle hydraulique (ensembles QNE_{Md} et QNE_Q) sont comparables à celles obtenues avec l'utilisation du modèle hydraulique.

Discrimination seuil vert-jaune La calibration statistique améliore de manière très nette la discrimination de l'événement *Dépassement du seuil vert-jaune* pour les deux expériences et E_{PA} et E_{PASC} sur les périodes de calage et de validation. Pour ces deux expériences, nous notons que les calibrations statistiques réalisées sans prédicteur issu du modèle hydraulique QNE_{Md} et QNE_Q ne présentent pas d'irrégularité de performance en validation, contrairement aux ensembles Q_R et Q_Q pour lesquelles les courbes AUC présentent des pics vers le bas pour certaines échéances. L'expérience E_{MP} présente des ensembles calibrés avec une excellente discrimination, de même que les ensembles bruts correspondants.

Synthèse

Nous avons vu ici que les calibrations statistiques mises en place présentent de meilleures caractéristiques que celles des ensembles bruts sur les périodes de calage et de validation. Une calibration de l'hydraulique, en bout de chaîne de prévision, s'avère donc indispensable.
D'une manière générale, les deux calibrations QNE_{Md} et QNE_Q, réalisées sans utiliser les sorties du modèle hydraulique, présentent des performances comparables à Q_R et Q_Q qui sont réalisées à partir des ensembles bruts. Pour la station Justice, il est pertinent de réaliser un modèle d'apprentissage statistique QRF sans construire

de modèle hydraulique MASCARET. Les performances des quantiles calibrés sont de qualité équivalente, et leur génération est ainsi beaucoup plus légère en termes de développement et de coût de calcul.

8.3.3.3 Meilleure stratégie à retenir concernant la condition limite aval en prévision

Ensembles évalués dans ce paragraphe

⇒ Comparaison des résultats en période de validation pour les trois expériences : E_{PA} (prévision astronomique), E_{PASC} (prévision astronomique, surcote constante et E_{MP} (mesure perturbée, expérience non réalisable en opérationnel).

Nous nous intéressons à présent aux performances respectives des ensembles calibrés dans les trois expériences afin d'évaluer la meilleure stratégie à adopter en opérationnel.

CRPS Les expériences qui fournissent les meilleurs CRPS calibrés sont dans l'ordre l'expérience E_{MP}, puis E_{PASC} et enfin E_{PA}. Ce classement est le même que pour les CRPS des ensembles bruts. Cela signifie que plus le CRPS de l'ensemble brut est faible, plus le CRPS de l'ensemble calibré est faible également. Il y a donc un intérêt certain à construire des ensembles bruts de bonne qualité. La calibration statistique ne permet pas de tout corriger *a posteriori*. Il faut aussi garder à l'esprit que certains prédicteurs utilisés pour la calibration statistique sont dépendants de la manière dont est généré l'ensemble brut. Par exemple, le prédicteur *mareePicPl* intègre la surcote au temps de base pour l'expérience E_{PASC}. Le prédicteur *nivPicPl* n'est pas disponible en l'absence de prévision d'ensemble de surcote à Plaisance. Nous verrons dans la partie 8.3.3.4 que ces prédicteurs revêtent une importance particulière pour la calibration.

$\|.\|_2$ des histogrammes de rang En période de calage, la fiabilité des ensembles calibrés pour l'expérience E_{MP} est légèrement moins bonne que pour les deux autres expériences, alors que l'ensemble brut correspondant présente les meilleures fiabilités parmi les trois expériences. En période de validation, cette différence et toujours visible bien que moins marquée. L'expérience E_{MP} présente ainsi la plus mauvaise fiabilité pour les ensembles calibrés. Les expériences E_{PA} et E_{PASC} présentent des ensembles calibrés dont la fiabilité est similaire.

Discrimination seuil vert-jaune L'expérience E_{PASC} présente une meilleure discrimination de l'événement de dépassement de seuil vert-jaune par rapport à l'expérience E_{PA}.

La prise en compte de la surcote au temps de base permet donc d'améliorer de manière significative la discrimination des ensembles bruts.

Synthèse

L'expérience E_{MP} génère des ensembles calibrés présentant les meilleurs CRPS et les meilleures discriminations en validation, mais les moins bonnes fiabilités. Elle n'est pas réalisable en opérationnel.
Si l'on se limite aux expériences réalisables en opérationnel, c'est la stratégie de l'expérience E_{PASC}, qui consiste à ajouter à la prévision de marée astronomique la surcote mesurée au temps de base, qui présente les meilleurs résultats en validation. C'est cette expérience qui sera conservée pour la suite du travail lors de la comparaison des résultats obtenus par calibration statistique et par assimilation de données.

8.3.3.4 Importance des prédicteurs

Calibrations étudiées dans ce paragraphe

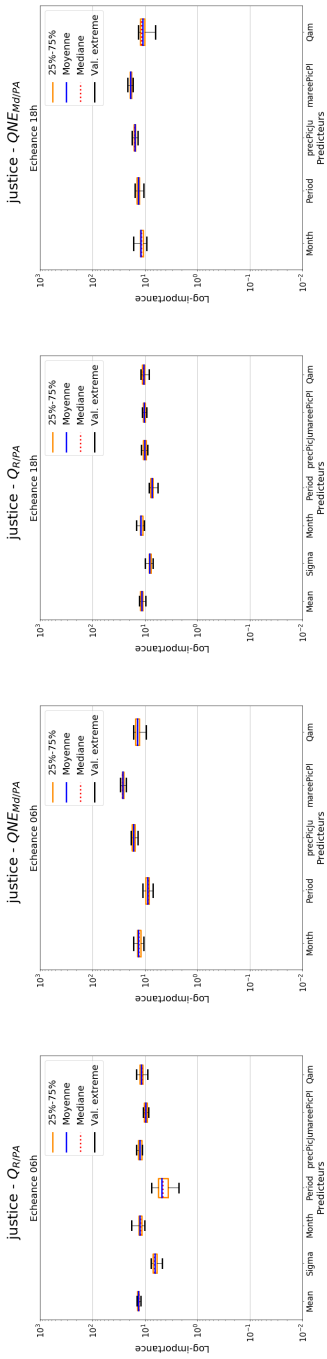
⇒ Importance des prédicteurs pour les trois expériences : E_{PA} (prévision astronomique), E_{PASC} (prévision astronomique, surcote constante et E_{MP} (mesure perturbée, expérience non réalisable en opérationnel).

Q_R : quantiles calibrés à partir de l'ensemble hydraulique brut R_R.

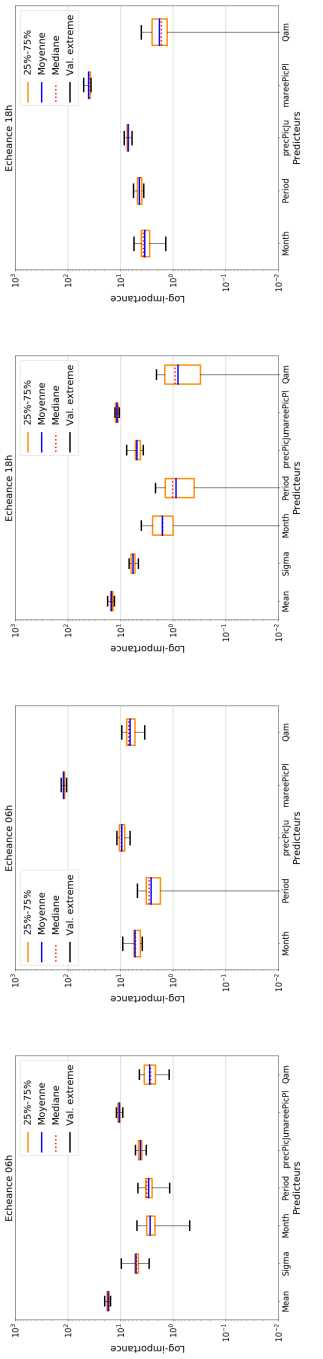
QNE_{Md} : quantiles calibrés, mêmes prédicteurs que pour Q_R sauf les prédicteurs issus de statistiques de l'ensemble brut R_R.

De même que pour la calibration à Kervir et Moulin-Vert, nous présentons dans ce paragraphe l'importance des prédicteurs utilisés pour les calibrations. Les figures 8.12a, 8.12b et 8.12c présentent l'importance des prédicteurs pour les échéances 6 heures et 18 heures pour les ensembles Q_R et QNE_{Md}, respectivement pour les trois expériences E_{PA}, E_{PASC} et E_{MP} à Justice. Nous ne montrons pas les résultats pour les calibrations Q_Q ni QNE_Q puisque nous avons vu qu'elles sont respectivement quasiment identiques aux calibrations Q_R et QNE_{Md}.

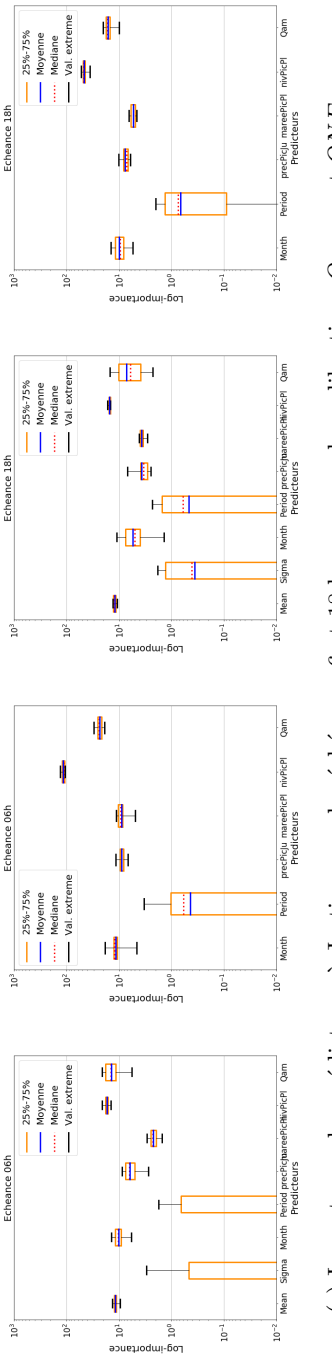
Pour tous les ensembles calibrés, le prédicteur *Period* caractérisant le moment de la prévision dans la journée est peu informatif. Il ne semble donc pas y avoir de corrélation entre la valeur du pic et le moment de la journée auquel il se produit. Ceci est normal si l'on considère la remontée de la marée astronomique. Il est possible que le phénomène de surcote/décote possède une variation quotidienne. En tout état de cause, même si c'est le



(a) Importance des prédicteurs à Justice pour les échéances 6 et 18 heures pour les calibrations $Q_{R/PA}$ et $QNE_{Md/PA}$.



(b) Importance des prédicteurs à Justice pour les échéances 6 et 18 heures pour les calibrations $Q_{R/PASC}$ et $QNE_{Md/PASC}$.



(c) Importance des prédicteurs à Justice pour les échéances 6 et 18 heures pour les calibrations $Q_{R/MP}$ et $QNE_{Md/MP}$.

cas, il n'est pas suffisamment important pour rendre ce prédicteur informatif. Le prédicteur *Sigma*, qui représente l'écart-type de la prévision, est également peu important.

Le prédicteur *precPicJu*, qui représente la valeur du pic de marée haute précédant le temps de base à Justice, est relativement important pour toutes les configurations, ce qui signifie qu'il existe une corrélation entre deux valeurs successives de marée haute à Justice. Ceci est cohérent puisque les valeurs des pics en périodes de mortes-eaux ou de vive-eaux sont d'amplitude similaire. Les valeurs de surcote et décote pour des pics à 12 heures d'intervalle sont également potentiellement corrélées.

Les prédicteurs les plus informatifs sont *Mean*[6] et/ou *mareePicPl*[7] pour les expériences E_{PA} et E_{PASC}, et *Mean* et/ou *nivPicPl*[8] pour l'expérience E_{MP}. Dans tous les cas, ces prédicteurs informent sur la valeur du pic de marée à Plaisance générant le pic de marée à Justice. Cette information est exactement le prédicteur *nivPicPl*, disponible dans l'expérience E_{MP}. Pour les deux autres expériences, en l'absence de ce prédicteur, c'est le prédicteur *mareePicPl* qui le remplace. Ce prédicteur représente la marée astronomique à Plaisance dans E_{PA}, il tient compte de la surcote au temps de base dans E_{PASC}. Il est moins performant que le prédicteur *nivPicPl* qui est en moyenne la valeur du pic mesuré à Plaisance. C'est pour cela que les CRPS et la discrimination des ensembles calibrés E_{PA} et E_{PASC} sont moins bons que pour l'ensemble E_{MP}. Le prédicteur *Mean* est important pour toutes les expériences. Ici, l'information sur la valeur du pic à Plaisance est remontée à la station Justice par l'intermédiaire du modèle hydraulique. L'analyse de l'importance des prédicteurs met donc en évidence que l'information la plus importante pour la calibration des pics de marée haute à Justice est la valeur de la hauteur du pic à Plaisance qui le génère. Ceci est cohérent avec la physique de l'écoulement. À Justice, c'est la courbe de remous qui remonte qui détermine la hauteur d'eau. Ces résultats sont également en accord avec les résultats de l'étude de sensibilité qui avait mis en lumière une importance prédominante de la condition limite aval pour la hauteur d'eau simulée à Justice au moment des pics.

Les prévisions opérationnelles à Justice tiennent compte de la prévision de surcote déterministe à Plaisance. Pour construire une prévision d'ensemble des pics de pleine mer à Justice, il faudrait *a minima* utiliser le prédicteur qui serait la somme de la prévision de marée astronomique et de la prévision de surcote déterministe. Pour aller plus loin, l'utilisation d'une prévision d'ensemble de surcotes permettrait de prendre également en compte les incertitudes sur la prévision de surcote.

[6] *Mean* : Moyenne de l'ensemble brut.

[7] *mareePicPl* : Hauteur de la pleine mer astronomique générant le pic à Justice. Pour l'expérience E_{PASC}, on ajoute la surcote mesurée au temps de base.

[8] *nivPicPl* : Moyenne de l'ensemble des hauteurs de pleine mer prévues à Plaisance générant le pic à Justice.

8.3.4 Synthèse

Dans cette section, nous avons d'abord présenté les prédicteurs utilisés pour la mise en place de la calibration QRF à Justice. Les prédicteurs sont différents de ceux utilisés à Kervir et Moulin-Vert car on y calibre les pics de marée haute. La stratégie retenue pour l'optimisation des forêts aléatoires est également différente : nous avons mis en évidence qu'une stratégie basée uniquement sur la minimisation du CRPS conduisait à des ensembles avec des meilleures performances globales en validation.

Une fois défini le cadre de la calibration statistique à Justice, la suite du chapitre s'est attachée à présenter les performances obtenues pour les ensembles bruts et calibrés. Trois expériences qui diffèrent par la condition limite aval utilisée en prévision ont été menées : E_{PA} (prévision de marée astronomique), E_{PASC} (prévision de marée astronomique avec ajout de la surcote mesurée au temps de base), et E_{MP} (perturbation de la surcote mesurée *a posteriori*). Nous résumons ci-après les principaux résultats obtenus.

Nous avons d'abord mis en évidence que la calibration des ensembles hydrologiques n'a pas d'impact sur les performances des ensembles bruts et calibrés des pics de pleine mer à Justice. Il n'y a donc pas d'intérêt à produire une hydrologie calibrée à l'amont. De plus, il a été montré que les performances des ensembles bruts sont d'autant meilleures que la surcote prise en compte à l'aval est de bonne qualité. En effet, l'expérience E_{MP} présente les meilleures performances, suivie de l'expérience E_{PASC}, puis de l'expérience E_{PA}. La prise en compte de la surcote au temps de base permet donc d'améliorer les performances des ensembles bruts sur les 24 heures de prévision.

En ce qui concerne les performances des ensembles calibrés, nous avons vu que pour chaque expérience, les calibrations statistiques améliorent les performances des ensembles bruts pour tous les critères étudiés. Une calibration de l'hydraulique au bout de la chaîne de prévision s'avère donc indispensable. Nous avons également vu que l'amélioration des CRPS et de la discrimination des ensembles est homogène sur les périodes de calage et de validation. En revanche, l'amélioration de la fiabilité est meilleure sur la période de calage, mais les calibrations statistiques mises en place améliorent quand même grandement la fiabilité par rapport aux ensembles bruts sur la période de validation. Les calibrations réalisées sans utiliser de prédicteur issu de statistiques d'un ensemble brut présentent des performances comparables à celles réalisées avec un ensemble hydraulique brut. Il est donc pertinent de construire un modèle d'apprentissage sans modèle hydraulique à Justice. Le coût de développement est alors nettement réduit, puisqu'il n'est pas nécessaire de construire le modèle, le caler, et générer des prévisions d'ensemble brutes. Pour finir, si l'on s'intéresse aux ensembles possibles d'obtenir en opérationnel, les résultats comparées des deux expérience E_{PA} et E_{PASC} mettent en évidence un net avantage à ajouter la surcote mesurée au temps de base pour les performances des ensembles bruts et calibrés. Dans la suite du travail, lors de la comparaison des performances des ensembles obtenus

par calibration statistique avec ceux obtenus par assimilation de données (chapitre 10), nous travaillerons uniquement avec l'expérience E_{PASC}. Puisque nous avons vu que la calibration de l'hydrologie n'a pas d'impact notable sur les résultats, nous travaillerons uniquement les ensembles $R_{R/PASC}$, $Q_{R/PASC}$ et $QNE_{Md/PASC}$, qui sont générés sans ensemble hydrologique calibré.

L'étude de l'importance des prédicteurs a mis en évidence que l'information essentielle pour une bonne calibration à Justice est la connaissance de la valeur du pic de marée haute à Plaisance qui le génère. Ce point est en accord avec les résultats de l'étude de sensibilité globale. C'est donc sur ce point là qu'il faut progresser pour améliorer la prévision à Justice. Pour une application opérationnelle, il serait pertinent de reprendre cette étude avec une prévision déterministe ou d'ensemble de surcote. Nous pouvons espérer que la pente visible sur les graphiques de CRPS en figure 8.10 soit plus faible. En effet, la qualité de l'information apportée par la surcote mesurée au temps de base est très bonne aux courtes échéances. À plus longue échéance, sa qualité diminue et le CRPS augmente. La connaissance d'une prévision de surcote permettrait de conserver une meilleure qualité de cette information tout au long de la prévision.

De même que pour les résultats émis à Kervir et Moulin-Vert, il est nécessaire de garder à l'esprit que les résultats obtenus ici sont à modérer du fait que le travail n'est réalisé que sur un seul bassin versant et sur une seule station. Les conclusions pourraient être différentes pour une station sous influence maritime située davantage en amont du réseau hydraulique. L'importance de l'ensemble hydrologique serait alors sûrement renforcé.

8.4 Conclusion générale

Ce chapitre a présenté la méthode de génération des ensembles hydrauliques bruts aux trois stations de vigilance et leur calibration statistique.

Les ensembles bruts sont générés dans un contexte opérationnel, en tenant compte des variables incertaines qui sont importantes pour les hauteurs d'eau prévues aux trois stations de prévision. Différentes expériences ont été menées, selon le choix pour les conditions limites amont et aval. Le calage des hyperparamètres des forêts aléatoires est réalisé sur une période de 3 ans, les mois d'été étant retirés. La validation des forêts est réalisée sur une période indépendante et permet d'estimer les performances des ensembles en opérationnel.

Les calibrations sont menées de manière différente aux stations fluviales Kervir et Moulin-Vert et à la station sous influence maritime Justice. À Kervir et Moulin-Vert, stations sous inflluence fluviale, la variable calibrée est la hauteur d'eau horaire. À Justice, la dynamique de l'écoulement est différente. Cette station est sous influence maritime. La variable calibrée est les hauteurs des pics de pleine mer.

Les prédicteurs utilisés pour les deux types de stations sont différents. Le choix des

prédicteurs est en partie guidé par les résultats de l'analyse de sensibilité. Les prédicteurs utilisés à Kervir et Moulin-Vert permettent de prendre en compte ce qui se passe au niveau des conditions limites amont, alors que les prédicteurs utilisés à Justice sont davantage focalisés sur ce qui se passe à l'aval. Les conclusions obtenues ne sont pas identiques sur ces deux types de station. Pour mémoire, les conclusions spécifiques à Kervir et Moulin-Vert sont présentées au paragraphe 8.2.4, et les conclusions à Justice sont présentées au paragraphe 8.3.4. Nous nous attachons à présenter à présent les similarités et les différences mises en évidence selon que la station est sous influence martime ou non.

Il y a assez peu de conclusions communes aux trois stations. En ce qui concerne la stratégie d'optimisation Opt_{Rel}, nous avons vu que dans tous les cas, elle conduisait à des courbes chaotiques pour les CRPS en période de calage. Un autre point commun réside dans le fait qu'une calibration de l'hydraulique en bout de la chaîne de prévision est indispensable. Par ailleurs, pour les trois stations, la fiabilité obtenue sur la période de validation est moins bonne que celle obtenue lors du calage. Ceci est surtout marqué à la station Kervir.

Les différences entre les deux types de stations résident tout d'abord dans le choix pour le critère d'optimisation des forêts. À Kervir et Moulin-Vert, nous avons vu qu'une stratégie d'optimisation combinée de la norme des histogramme de rang et du CRPS était préférable. À Justice en revanche, la calibration des pics de marée haute est plus performante en validation avec une optimisation du CRPS uniquement. Par ailleurs, une autre différence notable réside dans le besoin ou non de produire une hydrologie calibrée pour générer les ensembles hydrauliques. Pour les stations Kervir et Moulin-Vert, la meilleure stratégie consiste à calibrer hydrologie et hydraulique, donc il faut réaliser une calibration à chaque maillon de la chaîne de prévision. Nous avons vu que le gain le plus important en termes de CRPS est lié à la calibration de l'hydrologie. À Justice en revanche, la calibration de l'hydrologie n'a pas d'impact sur les performances des ensembles bruts et calibrés. Ces résultats sont cohérents avec les résultats de l'analyse de sensibilité globale. Pour finir, nous avons vu qu'à Justice, la prise en compte de prédicteurs issus de statistiques du modèle hydraulique n'est pas indispensable. En revanche, cette stratégie n'est pas pertinente pour la calibration à Kervir et Moulin-Vert, où l'utilisation des prédicteurs issus de l'hydraulique améliore la discrimination des événements de dépassement de seuil et les valeurs des CRPS pour les longues échéances. Cette différence est probablement liée au fait que les variables calibrées sont différentes. À Justice, ce sont les pics de pleine mer qui sont calibrés. Le moment précis auquel survient ce pic n'est pas pris en compte. Il n'est donc pas nécessaire de connaître la dynamique fine des écoulements. À Kervir et Moulin-Vert en revanche, ce sont les hauteurs d'eau horaires qui sont calibrés. La dynamique des écoulements intervient fortement, car si une hauteur d'eau est bien prévue mais décalée d'une heure, les résultats ne sont pas bons. La dynamique des écoulements est prise en compte dans la calibration grâce à l'utilisation de prédicteurs constuits à partir du modèle hydraulique MASCARET.

Chapitre 9

Filtre de Kalman d'ensemble (EnKF) dans le modèle MASCARET

Sommaire

9.1 Structure de l'assimilation de données mise en place 256

9.2 Réalisation des expériences jumelles 258

9.2.1 Description des expériences jumelles 259

9.2.2 Correction des Ks, des conditions limites amont ou aval séparément 262

9.2.2.1 Correction des Ks : $Ks@Z_{vigilance}$ 262

9.2.2.2 Correction des conditions aux limites amont : $CLa_{amont}@Z_{vigilance}$ 266

9.2.2.3 Correction de la condition limite aval : $CLb_{aval}@Z_{Justice}$ 274

9.2.3 Correction des Ks, des conditions limites amont et aval simultanément 276

9.2.3.1 Assimilation des hauteurs : $KsCL@Z_{vigilance}$ 277

9.2.3.2 Assimilation des hauteurs et des conditions limites : $KsCL@all_{zQ}$ 279

9.2.3.3 Assimilation des hauteurs et des conditions limites, test avec 279 membres : $KsCL@all_{zQ} - 279$ 280

9.2.4 Synthèse 281

9.3 Réalisation des expériences réelles 282

9.3.1 Description des expériences réelles 282

9.3.2 Résultats dans la phase d'analyse 285

9.3.2.1 Etude des paramètres analysés 285

9.3.2.2 Hauteurs simulées dans la phase d'analyse 286

9.3.3 Synthèse 288

Ce chapitre présente la mise en place d'assimilation de données au sein du modèle hydraulique MASCARET. De même que pour l'hydrologie, nous utilisons le filtre de Kalman d'ensemble de manière cyclée. Le paragraphe 9.1 présente la structure de l'assimilation de données mise en place pour les expériences jumelles et réelles. Le paragraphe 9.2 présente les résultats des expériences jumelles menées, et le paragraphe 9.3 présente les expériences réelles ainsi que les résultats obtenus dans la phase d'analyse. Les résultats en phase de prévision opérationnelle seront présentés au chapitre 10 dans lequel nous comparerons toutes les prévisions d'ensemble hydrauliques construites dans le cadre de ce travail.

9.1 Structure de l'assimilation de données mise en place

Nous présentons dans ce paragraphe les éléments clés de l'assimilation de données mise en place : choix du vecteur de contrôle, réalisation du cyclage des analyses, choix du nombre de membres et de l'erreur d'observation.

Vecteur de contrôle L'assimilation de données mise en place pour réaliser des prévisions d'ensemble avec le modèle hydraulique 1D MASCARET a pour objectifs la correction des coefficients de frottement du modèle et de ses conditions aux limites. On ne cherche pas à corriger l'état du modèle. En effet, la correction de l'état a un impact potentiellement limité dans le temps au moment d'un événement pluvieux, ce qui n'est pas le cas de la correction des paramètres. L'objectif étant de réaliser des prévisions jusqu'à une échéance de 24 heures, la correction des paramètres a été privilégiée. Par ailleurs, il est pertinent de considérer la persistance des coefficients de Strickler sur la phase de prévision opérationnelle puisque ces derniers ne varient pas très rapidement dans le temps.

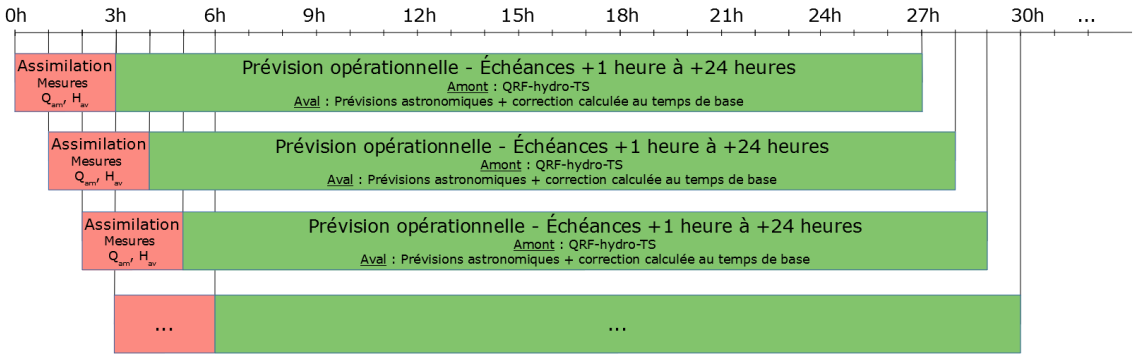
La correction d'une condition limite $f(t)$ variant temporellement est réalisée grâce à un paramétrage avec trois paramètres a, b et c tels que :

$$f(t) = a * f(t - c) + b.$$

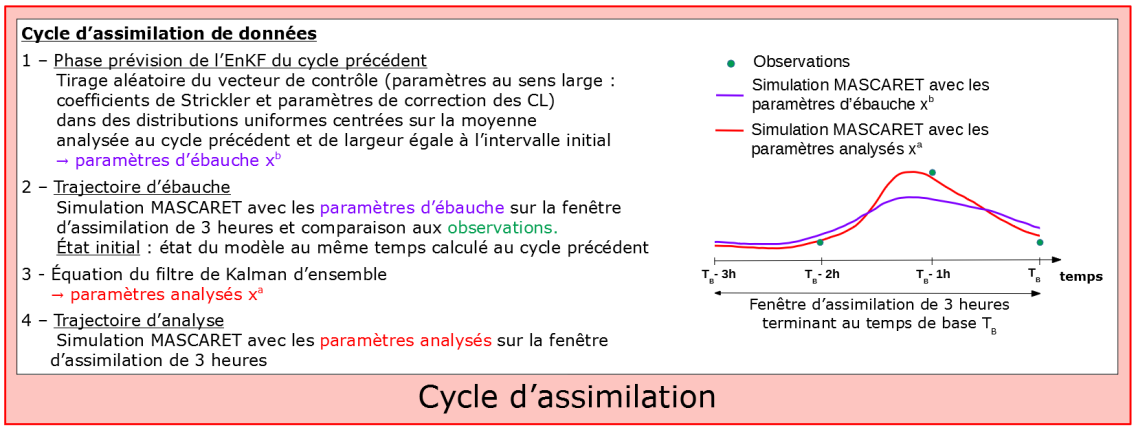
a représente une correction multiplicative, b une correction additive et c un déphasage de la chronique. Dans les expériences implémentées dans ce travail, seuls les paramètres de correction a et b ont été utilisés. f représente un débit pour une condition limite amont ou une hauteur d'eau pour la condition limite aval.

Le vecteur de contrôle est donc constitué des coefficients de frottement des zones retenues suite à l'étude d'incertitude ($Ks_{min}4$, $Ks_{min}7$, $Ks_{min}8$, $Ks_{min}11$ et $Ks_{maj}8$) et des coefficients a et/ou b pour chaque condition limite le cas échéant. Les éléments du vecteur de contrôle seront nommés *paramètres* même s'ils ne sont pas à strictement parler des paramètres du modèle (cas de la correction des conditions limites). La correction des éléments du vecteur de contrôle est réalisée par remplacement des valeurs.

Cyclage La figure 9.1 présente l'assimilation de données mise en place pour le modèle hydraulique MASCARET. La figure 9.1a présente le principe du cyclage des analyses et prévisions opérationnelles mis en place pour les expériences jumelles et réelles avec le modèle MASCARET. De même que lors de l'assimilation de données dans le modèle MORDOR-TS, les fenêtres d'assimilation durent 3 heures et elles sont cyclées de manière horaire. Les prévisions opérationnelles sont toujours réalisées sur une fenêtre de 24 heures. Nous rappelons que la fin de la fenêtre d'assimilation correspond au temps de base T_B, et qu'à ce moment, une prévision opérationnelle peut être lancée. Cette figure 9.1a correspond exactement au cas des expériences jumelles décrites au paragraphe 9.2 et à une des expériences réelles en ce qui concerne les forçages mentionnés. Dans le cas des expériences réelles, des expériences supplémentaires sont menées avec différents forçages sur les fenêtres d'assimilation et de prévision opérationnelle. Ce point sera détaillé au paragraphe 9.3.



(a) Cycles d'analyses et de prévisions mis en place pour l'assimilation de données dans le modèle MASCARET. Les forçages indiqués correspondent à ceux utilisés pour les expériences jumelles et pour une des expériences réelles menées.



(b) Description détaillée d'un cycle d'assimilation de données. À la fin de cette phase, l'état du modèle est en équilibre avec la valeur des paramètres analysés. Une prévision opérationnelle peut être lancée.

Figure 9.1: Structure de l'assimilation de données mise en place pour le modèle MASCARET. La fenêtre d'assimilation dure 3 heures. Les prévisions opérationnelles sont émises jusqu'à l'échéance 24 heures.

La figure 9.1b présente la réalisation d'un cycle d'assimilation de données en détail correspondant aux phases *Prévision* et *Analyse* de l'EnKF. Pour plus de détail, le lecteur pourra se reporter au paragraphe 3.4.6 qui décrit l'EnKF appliqué dans le cadre de ce travail. L'intervalle initial dont la largeur est utilisée à chaque cycle pour le tirage d'un nouveau vecteur de contrôle d'ébauche (étape 1 du schéma 9.1b) est le même que celui utilisé pour l'étude de sensibilité globale et pour la génération de l'ensemble brut (voir chapitre 8, tableau 8.1) pour les coefficients de Strickler. Pour le paramètre de correction multiplicative a, l'intervalle intial est [0.95, 1.05] et pour le paramètre de correction additive b, l'intervalle initial est [-10 cm, 10 cm]. Dans le cas présenté ici, les trajectoires d'ébauche et d'analyse (étapes 2 et 4) sont réalisées en alimentant le modèle MASCARET par les mesures à l'amont et à l'aval, avec correction des lois avec les paramètres a et/ou b le cas échéant. Par ailleurs, lors du calcul de la trajectoire d'analyse, l'état du modèle est sauvegardé au bout d'une heure. Il servira d'état initial pour la trajectoire d'ébauche du cycle suivant. À la fin du cycle d'assimilation (après l'étape 4), on se trouve au temps de base T_B et la prévision opérationnelle peut être lancée. Elle correspond à l'intégration déterministe des membres analysés sur la fenêtre de prévision opérationnelle. Dans le cas présenté ici, ils sont alors alimentés par les prévisions *QRF-hydro-TS* et par les prévisions de marée astronomique. Les paramètres de MASCARET (coefficients de Strickler et paramètres de correction des conditions aux limites a et/ou b) analysés sur la fenêtre d'assimilation sont alors persistés sur la fenêtre de prévision opérationnelle.

Nombre de membres et erreur d'observation Le nombre de membres est fixé à 49. Un test avec 279 membres est réalisé au paragraphe 9.2.3.3 pour étudier l'impact du nombre de membres sur les résultats. L'erreur d'observation est fixée à

- une valeur constante valant 5 cm pour les observations de hauteur d'eau ;
- un pourcentage de la valeur mesurée de 5% pour les observations de débit, comme pour les assimilations de données réalisées dans le modèle hydrologique MORDOR-TS.

9.2 Réalisation des expériences jumelles

Ce paragraphe présente les expériences jumelles réalisées pour évaluer l'algorithme d'assimilation de données dans le modèle MASCARET. L'objectif est l'amélioration de la modélisation des hauteurs d'eau au niveau des trois stations de vigilance Kervir, Moulin-Vert et Palais de Justice. La structure du modèle MASCARET de l'Odet ainsi que la localisation des stations sont rappelées sur la figure 9.2. Le principe des expériences jumelles a été décrit au paragraphe 5.2.1. Le lecteur pourra s'y reporter si besoin. Nous rappelons que le principe est de générer des pseudo-observations grâce au modèle utilisé avec des paramètres dits *de référence* qui sont différents des paramètres calés. Cette simulation est alors appelée simu-

lation *de référence*. Ensuite, en partant du modèle calé, on évalue comment l'assimilation de données permet de retrouver ou non les valeurs des paramètres de référence (espace de contrôle) et les valeurs simulées par la simulation de référence (espace des observations). La simulation dite *de contrôle*[1] correspond à la simulation sans assimilation de données, donc lancée avec les paramètres calés.

9.2.1 Description des expériences jumelles

Les expériences jumelles sont réalisées sur la période du 10 décembre 2013 au 31 décembre 2013. On travaille ainsi sur le même événement que lors des expériences jumelles avec le modèle hydrologique MORDOR-TS (voir paragraphe 2.1.1, tempête Dirk).

Création des observations *vraies* Comme expliqué précédemment, le vecteur de contrôle est constitué des coefficients de Strickler et des paramètres permettant de corriger les conditions limites. La simulation de référence, pour laquelle les paramètres sont modifiés par rapport au calage, est générée de la manière suivante :

- les coefficients de Strickler majorés de 15 par rapport à la valeur de calage ;
- une correction multiplicative est appliquée aux conditions limites amont $q(t)$, selon la formule $q_C(t) = a * q(t)$. a est fixé à 0.95.
- une correction additive est appliquée à la condition limite aval $h(t)$, selon la formule $h_C(t) = h(t) + b$. Le paramètre b est cyclique selon le moment de la journée.

Le tableau 9.1 résume les valeurs des paramètres utilisés pour les simulations de contrôle et de référence pour les expériences jumelles. Il est à noter que les choix réalisés pour les expériences jumelles ont pour conséquence une diminution des cotes d'eau dans la simulation de référence (augmentation des coefficients de Strickler et diminution des apports en eau à l'amont et à l'aval).

Simulation	$Ks4_{min}$	$Ks7_{min}$	$Ks8_{min}$	$Ks11_{min}$	$Ks8_{maj}$	a	b			
Contrôle	38.0	37.0	25.5	32.5	5	1.00	0.0			
Référence	53.0	52.0	40.5	47.5	20	0.95	0-6h 0	6h-12h -10 cm	12h-18h -20 cm	18h-24h -10 cm

Tableau 9.1: Valeur des paramètres pour les simulations de contrôle et de référence pour les expériences jumelles.

Pour les trois simulations de contrôle, de référence et d'assimilation, le modèle est alimenté par des mesures de débit à l'amont et de hauteur d'eau à l'aval sur la fenêtre d'assimilation. Sur la phase de prévision opérationnelle, le modèle est alimenté par les prévisions *QRF-hydro-TS* moyennées à l'amont. Pour mémoire, ce sont les prévisions hy-

[1] Attention à ne pas confondre la simulation de contrôle et le vecteur de contrôle, qui correspond à l'ensemble des paramètres corrigés par l'assimilation de données.

drologiques d'ensemble obtenues par calibration statistique sur lesquelles est appliqué un lissage des trajectoires. Pour les expériences dans lesquelles la condition limite amont est modifiée par un coefficient multiplicatif a, ce coefficient est également persisté en prévision. À l'aval, le modèle est alimenté par la prévision de marée astronomique du SHOM. Pour les expériences dans lesquelles la condition limite aval est modifiée par un coefficient additif b, la valeur de ce coefficient au temps de base est également persistée en prévision : elle est ajoutée à la prévision de marée astronomique.

Application de l'algorithme EnKF Dans toutes les expériences, des hauteurs d'eau sont assimilées aux stations de vigilance Kervir, Moulin-Vert et/ou Justice. Dans certaines expériences, des débits peuvent également être assimilés aux stations amont Tréodet, Kerjean et Ty-Planche. Plusieurs types d'expériences jumelles sont réalisées. Le nom de ces expérience est $el_1@el_2$. el_1 correspond aux variables corrigées, et el_2 correspond aux données assimilées.

Étude des performances de l'algorithme EnKF La performance de l'algorithme d'assimilation de données est analysée selon plusieurs critères :

- Étude temporelle des valeurs analysées pour les éléments du vecteur de contrôle (valeurs analysées pour tous les membres et moyenne de l'analyse) et comparaison avec la valeur imposée pour la simulation de référence.

- Calcul de l'erreur sur les hauteurs simulées au niveau de chaque station de vigilance pour la simulation de contrôle et la simulation avec assimilation de données par rapport à la simulation de référence. Pour la simulation avec assimilation de données, on calcule pour chaque pas de temps l'erreur ϵ_i entre la moyenne des membres et la simulation de référence.

- Calcul de la RMSE intégrée sur l'événement calculée le long des biefs du modèle MASCARET. Pour cette analyse, nous nous appuierons sur la localisation des 5 biefs du modèle MASCARET de l'Odet présentés sur la figure 9.2. Le bief b_1 est compris entre Tréodet et la confluence de l'Odet et du Jet. Le bief b_2 est compris entre l'aval du bief b_1 et la confluence entre l'Odet et le Steïr. Le bief b_3 est compris entre l'aval du bief b_2 et Plaisance. Le bief b_4 est compris entre Kerjean et la confluence du Jet et de l'Odet. Le bief b_5 est compris entre Ty-Planche et la confluence du Steïr et l'Odet. Ainsi, le bief b_{1-2-3} correspond à l'enchaînement des biefs b_1, b_2 et b_3, le bief b_{4-2-3} correspond à l'enchaînement des biefs b_4, b_2 et b_3 et le bief b_{5-3} correspond à l'enchaînement des biefs b_5 et b_3. Pour la simulation avec assimilation de données, pour chaque section de calcul de MASCARET, on calcule d'abord pour chaque pas de temps l'erreur ϵ_i entre la moyenne des membres et la simulation de référence, puis on calcule la norme 2 de ces erreurs sur tous les pas de temps considérés : $\sqrt{\frac{1}{N_{pas}} \sum_{i=1}^{N_{pas}} \epsilon_i^2}$.

• Calcul de la RMSE des erreurs de prévision en fonction de l'échéance pour les simulations de contrôle et l'assimilation de données. Pour une échéance de prévision donnée, on calcule d'abord pour chaque pas de temps l'erreur ϵ_i entre la moyenne des membres prévus et la simulation de référence, puis on calcule la norme 2 de ces erreurs sur tous les pas de temps considérés : $\sqrt{\frac{1}{N_{pas}} \sum_{i=1}^{N_{pas}} \epsilon_i^2}$.

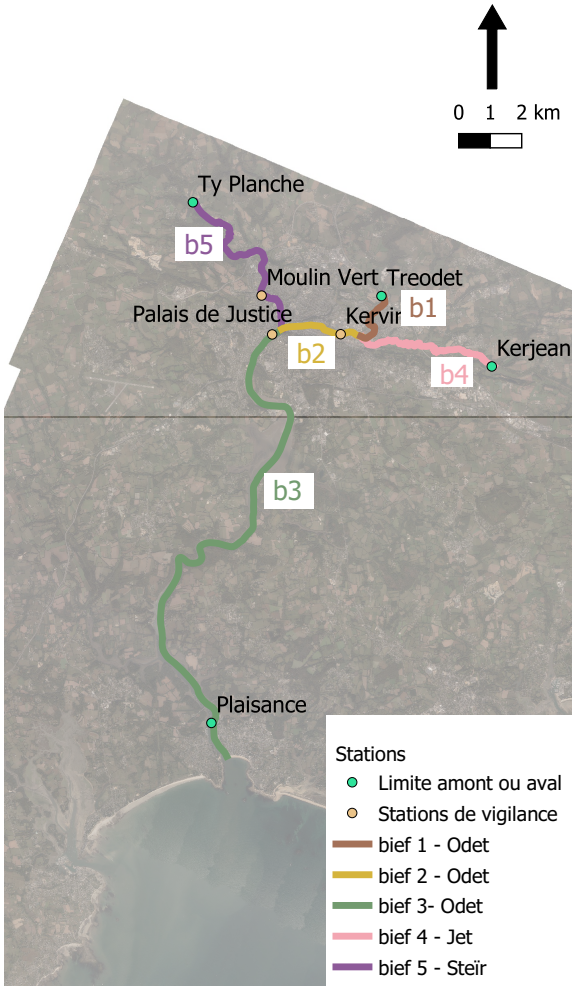


Figure 9.2: Modèle MASCARET de l'Odet. Visualisation des stations amont et aval, des stations de vigilance Kervir, Moulin-Vert et Justice et localisation des 5 biefs. Le bief b_{1-2-3} va de Tréodet vers Plaisance, le bief b_{4-2-3} va de Kerjean vers Plaisance et le bief b_{5-3} va de Ty-Planche vers Plaisance.

Pour tous les graphiques présentés, les valeurs des simulations analysées sont prises en fin de fenêtre d'assimilation (temps de base). De même que dans les chapitres concernant l'assimilation de données dans le modèle MORDOR-TS, sur tous les graphiques présentés :

- la légende *asm* en rouge est relative à l'analyse ;
- la légende *ctl* en bleu est relative à la simulation de contrôle ;
- la légende *ref* en noir est relative à la simulation de référence.

Selon les variables représentées sur les graphiques, ces termes seront relatifs à l'espace de contrôle (les paramètres) ou à l'espace des observations (les hauteurs d'eau aux stations de vigilance). Sur les graphiques relatifs aux résultats dans la phase de prévision opérationnelle, un nouveau élément de légende apparait :

- la légende *fcst* (pour *forecast*) en vert est relative à la prévision opérationnelle effectuée suite à la simulation de la trajectoire d'analyse.

Plan de la suite du chapitre Le paragraphe 9.2.2 présente les expériences jumelles dans lesquelles sont corrigés soit les coefficients de Strickler, soit les trois conditions limites amont de manière simultanée (correction multiplicative), soit la condition limite aval (correction additive). Le paragraphe 9.2.3 présente ensuite les expériences dans lesquelles sont corrigés simultanément les coefficients de Strickler, les trois conditions limites amont et la condition limite aval.

9.2.2 Correction des Ks, des conditions limites amont ou aval séparément

Dans ce paragraphe sont présentées les expériences jumelles construites pour la correction soit des coefficients de Strickler, soit des conditions limites amont, soit des conditions limites aval. Nous présentons les trois expériences suivantes :

- L'expérience $Ks@Z_{vigilance}$ vise à corriger les coefficients de Strickler des 5 zones retenues suite à l'étude d'incertitude en assimilant les cotes d'eau aux trois stations de vigilance. Les résultats de cette expérience sont présentés au paragraphe 9.2.2.1.
- L'expérience $CLa_{amont}@Z_{vigilance}$ vise à corriger simultanément les trois conditions limites amont par une correction multiplicative a, en assimilant les cotes d'eau aux trois stations de vigilance. Les résultats de cette expérience sont présentés au paragraphe 9.2.2.2.
- L'expérience $CLb_{aval}@Z_{Justice}$ vise à corriger la condition limite aval par une correction additive b en assimilant les cotes d'eau à Justice. Les résultats de cette expérience sont présentés au paragraphe 9.2.2.3.

9.2.2.1 Correction des Ks : $Ks@Z_{vigilance}$

Dans l'expérience $Ks@Z_{vigilance}$, les valeurs des coefficients de Strickler des 5 zones influentes retenues à l'issue de l'étude d'incertitude sont modifiées simultanément pour la simulation de référence, selon les valeurs indiquées dans le tableau 9.1.

Résultats dans la phase d'analyse La figure 9.3 présente les valeurs analysées des éléments du vecteur de contrôle au cours de l'événement. Ces graphiques mettent en

évidence une bonne correction des valeurs des Stricklers par l'algorithme d'assimilation de données. Nous notons que les coefficients de Strickler des zones 7 et 8 sont bien corrigées. Nous ne mettons donc pas en évidence de problème d'équifinalité notable entre la correction des coefficients de Strickler $Ks_{min}7$, $Ks_{min}8$ et $Ks_{maj}8$, qui déterminent majoritairement la hauteur d'eau simulée à Justice. Nous notons que les coefficients de Strickler $Ks_{min}11$ et $Ks_{maj}8$ mettent plusieurs cycles d'assimilation à atteindre la valeur de référence, alors que c'est beaucoup plus rapide pour les autres paramètres.

La figure 9.4 montre au niveau des trois stations de vigilance le résultat du calcul en fonction du temps de l'erreur sur la hauteur d'eau simulée par la simulation de contrôle et par la simulation avec assimilation de données (moyenne des membres) par rapport à la simulation de référence. Nous notons que les erreurs sur les hauteurs d'eau sont nettement réduites avec l'assimilation de données pour les trois stations. Les erreurs résiduelles sont inférieures aux erreurs d'observation (5 cm).

Il est intéressant de ne pas se limiter à étudier les erreurs aux stations de vigilance, mais d'analyser également les impacts de l'assimilation de données sur les valeurs simulées sur l'ensemble des biefs. Le nom des biefs a été décrit en figure 9.2. Les graphiques figure 9.5 montrent les erreurs intégrées sur toute la période d'étude (22 jours) le long des trois biefs b_{1-2-3}, b_{4-2-3} et b_{5-3} pour les simulations de contrôle et l'assimilation de données par rapport à la simulation de référence. Sur ces graphiques, l'indice du bief est indiqué en vert. Les limites des zones pour les Ks sont indiquées en rose. Ces graphiques mettent en évidence une bonne correction des niveaux simulés tout le long des biefs, et non pas uniquement au niveau des stations de vigilance au niveau desquelles des données sont assimilées.

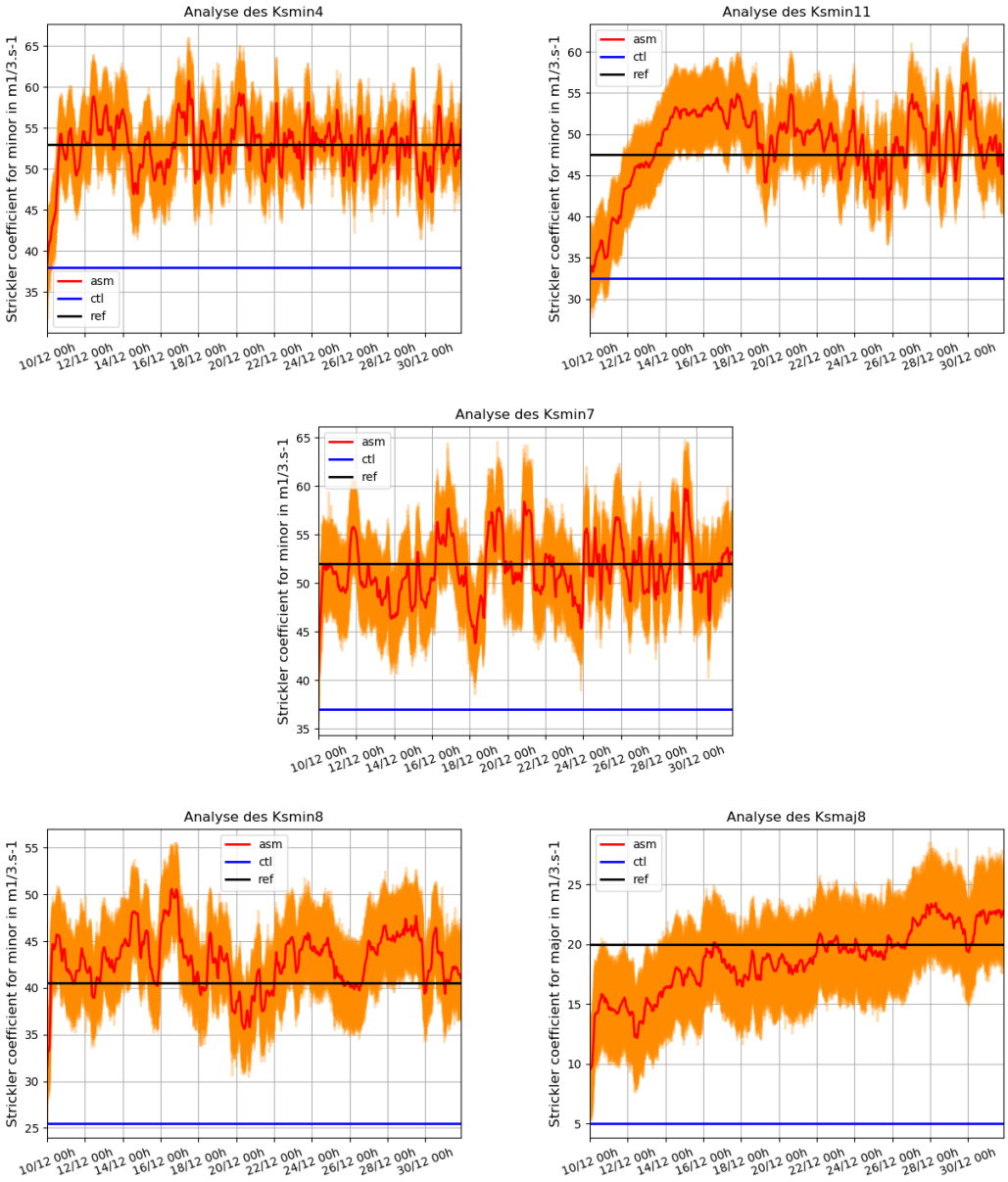


Figure 9.3: Expérience $Ks@Z_{vigilance}$ - Valeurs analysées pour le vecteur de contrôle (Ks) en fonction du temps. Les courbes orange représentent les valeurs pour chacun des 49 membres et la courbe rouge la valeur analysée moyenne.

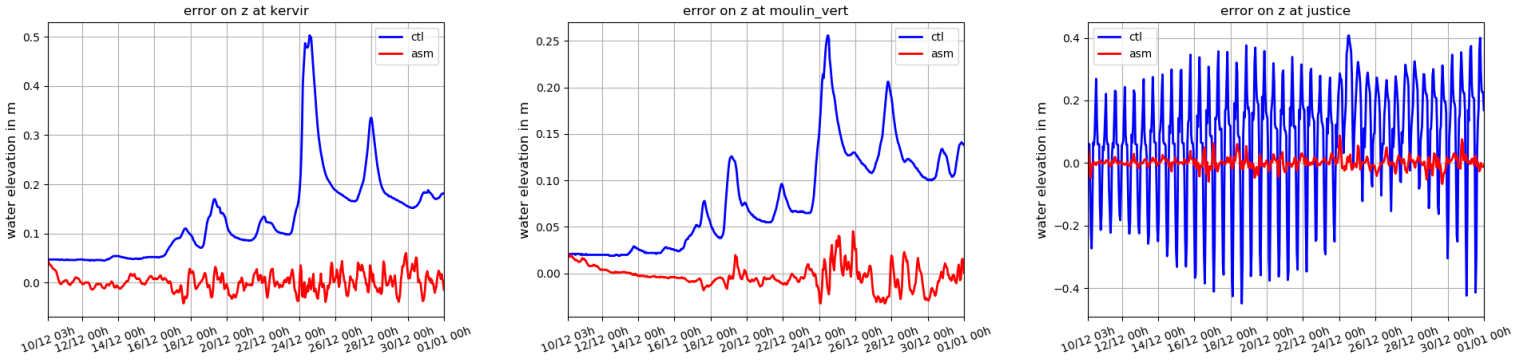


Figure 9.4: Expérience $Ks@Z_{vigilance}$ - Chroniques temporelles des erreurs aux stations de vigilance Kervir, Moulin-Vert et Justice pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (rouge) par rapport à la simulation de référence.

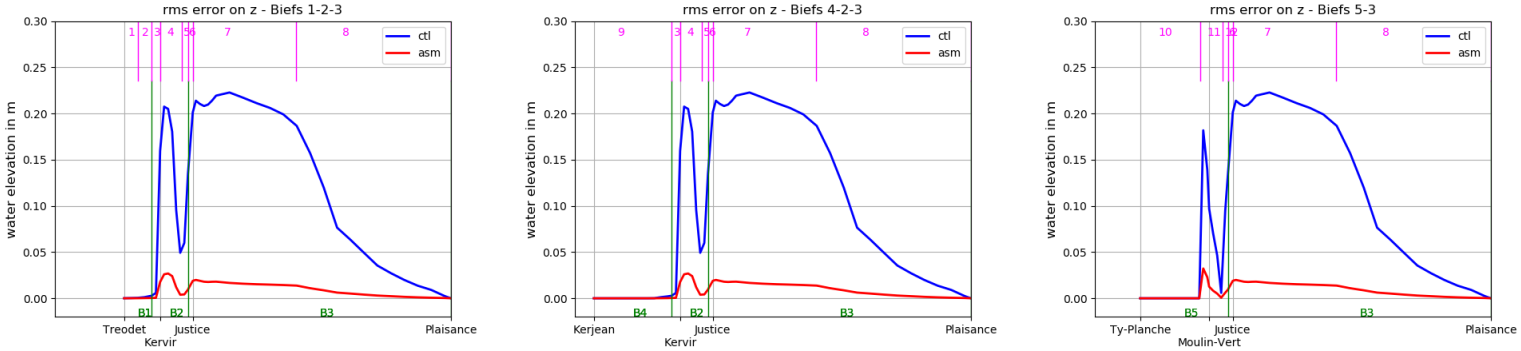


Figure 9.5: Expérience $Ks@Z_{vigilance}$ - RMSE le long des biefs du modèle MASCARET pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (rouge) par rapport à la simulation de référence.

Résultats dans la phase de prévision La figure 9.6 présente les résultats du calcul de la RMSE aux trois stations de vigilance pour les échéances de prévision variant de 1 heure à 24 heures. Nous notons que les prévisions sont nettement améliorées aux trois stations pour toutes les échéances présentées. Ici, on s'attendait à ne pas perdre de capacité prédictive du fait que les valeurs analysées du vecteur de contrôle convergent rapidement vers celles du vecteur de référence.

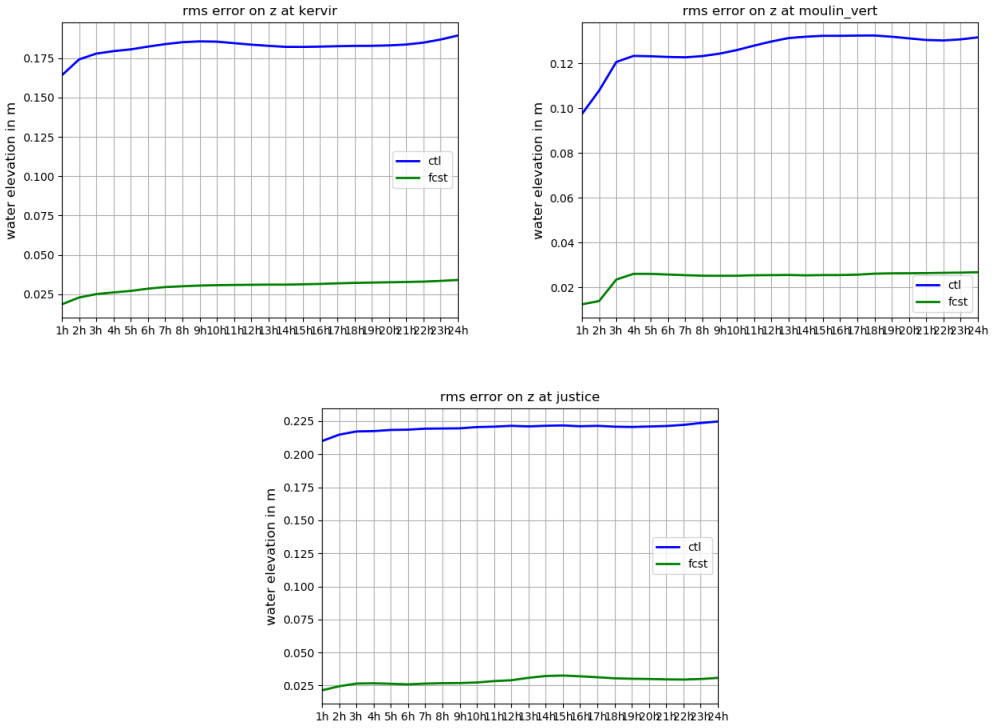


Figure 9.6: Expérience $Ks@Z_{vigilance}$ - RMSE des erreurs de prévision aux stations Kervir, Moulin-Vert et Justice en fonction de l'échéance pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (vert) par rapport à la simulation de référence.

9.2.2.2 Correction des conditions aux limites amont : $CLa_{amont}@Z_{vigilance}$

Dans l'expérience $CLA_{amont}@Z_{vigilance}$, la simulation de référence est générée en modifiant les lois des trois conditions limites amont avec un coefficient multiplicatif *a=0.95* constant dans le temps. Les hauteurs d'eau aux trois stations de vigilance Kervir, Moulin-Vert et Justice sont assimilées.

Résultats dans la phase d'analyse - Valeurs analysées La figure 9.7 présente les valeurs analysées du vecteur de contrôle au cours de l'événement. Ces graphiques mettent en évidence que l'assimilation de données permet de retrouver la valeur de référence à

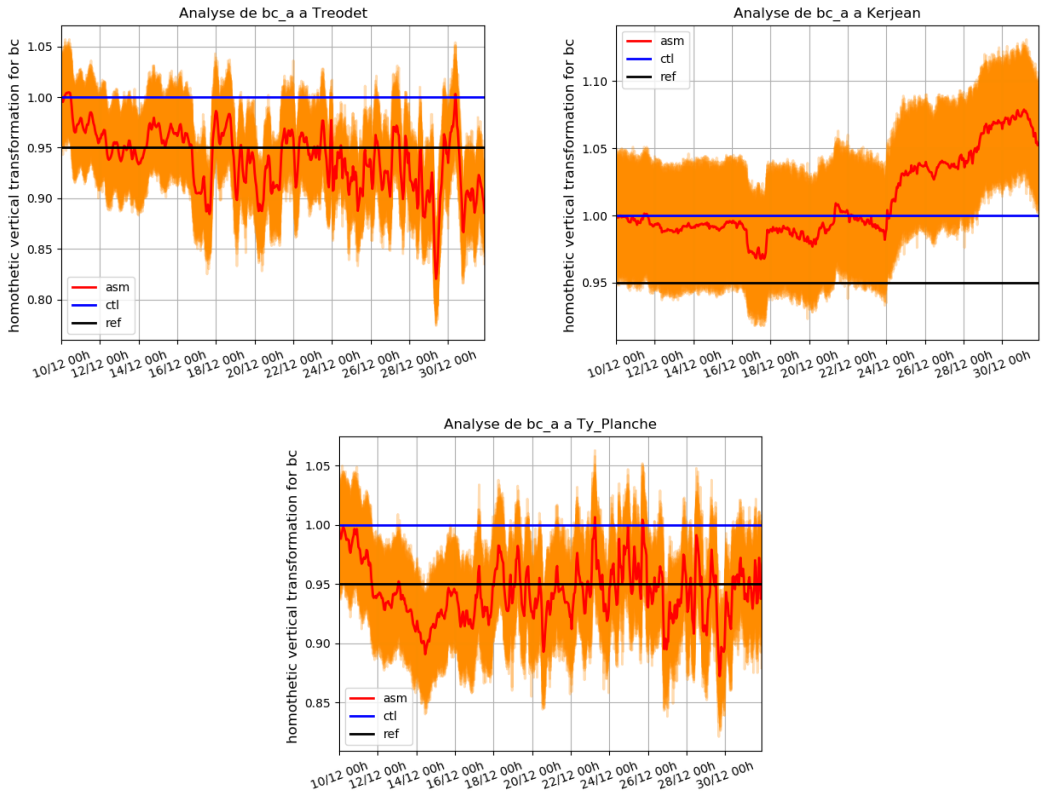


Figure 9.7: Expérience $CLA_{amont}@Z_{vigilance}$ - Valeurs analysées pour les éléments du vecteur de contrôle (paramètre de correction multiplicatif des conditions limites amont a) en fonction du temps. Les courbes orange représentent les valeurs pour chacun des 49 membres et la courbe rouge la valeur analysée moyenne.

Ty-Planche. À Tréodet, l'analyse sous-estime légèrement la référence, alors qu'elle la surestimne clairement à Kerjean. À Kerjean, l'algorithme d'assimilation de données n'a quasiment pas d'impact, voire un impact négatif (on s'éloigne de la valeur de référence). On note une tendance à l'augmentation du paramètre a à Kerjean à partir du 24 décembre, et cela correspond à une sous-estimation du paramètre a à Tréodet. La correction des paramètres à Tréodet et Kerjean est majoritairement déterminée par la valeur de la hauteur d'eau assimilée à Kervir, puisque c'est la station située immédiatement à l'aval. On note ici un problème d'équifinalité. Étant donné que le débit à Tréodet vaut environ 2.5 fois la valeur du débit à Kerjean, l'algorithme d'assimilation arrive à corriger le paramètre a à Tréodet de manière correcte, mais il ne fonctionne pas pour corriger la condition limite à Kerjean.

Afin d'investiguer la raison du non fonctionnement de l'algorithme pour corriger la condition limite à Kerjean, trois tests sont réalisés :

- Expérience $CLa@Z_{vigilance} - rea_{8h}$: cette expérience est analogue à la précédente, mais la durée de la fenêtre d'assimilation est de 8 heures à la place de 3 heures. Ce

test permet de voir si la longueur de la fenêtre d'assimilation peut être responsable d'une mauvaise correction à Kerjean. Les résultats des analyses pour les paramètres a à Tréodet et Kerjean sont présentées figure 9.8a.

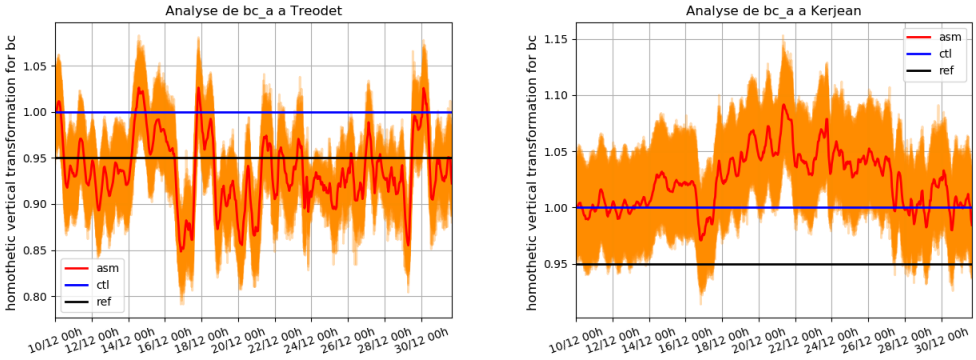
- Expérience $CLa@Z_{vigilance} - ech_{CL}$: cette expérience est également analogue à l'expérience $CLa_{amont}@Z_{vigilance}$, mais les chroniques de débits entre Tréodet et Kerjean sont échangées. Ce test permet de voir si c'est la différence d'ordre de grandeur entre les apports des deux biefs qui est responsable d'une mauvaise correction à Kerjean. Les résultats des analyses sont présentés en figure 9.8b.

- Expérience $CLa@Z_{vigilance+virtu-b4}$: dans cette expérience, on assimile en plus les hauteurs d'eau à une station virtuelle située à l'aval de Kerjean sur le bief b_4. Les résultats des analyses sont présentés en figure 9.8c.

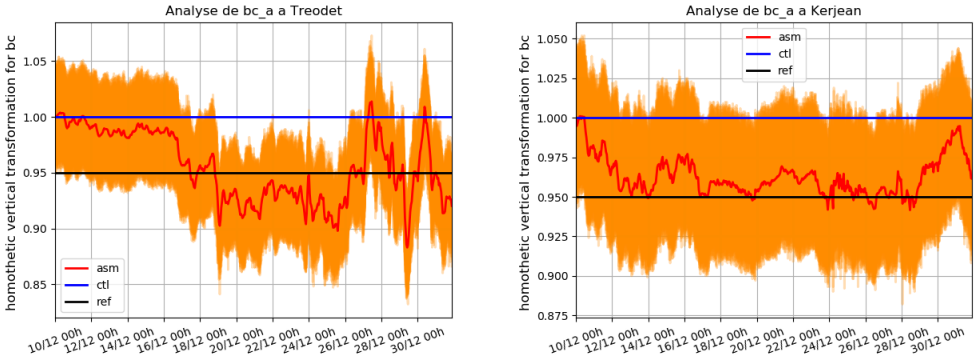
Nous notons d'abord que les trois expériences $CLa_{amont}@Z_{vigilance}$, $CLa@Z_{vigilance} - ech_{CL}$ et $CLa@Z_{vigilance+virtu-b4}$, pour lesquelles la taille de la fenêtre d'assimilation est la même (3 heures), fournissent exactement la même analyse à Ty-Planche (résultats non montrés ici). Ceci se comprend aisément car la hauteur d'eau à Moulin-Vert n'est pas influencée par les débits à Tréodet et Kerjean à cause de sa position sur le réseau hydraulique. Nous notons par ailleurs que l'augmentation de la taille de la fenêtre d'assimilation ne permet pas d'améliorer la correction de la condition limite à Kerjean (figure 9.8a). En revanche, l'échange des chroniques entre Tréodet et Kerjean permet de corriger la condition limite à Kerjean (figure 9.8b). La condition limite à Tréodet est alors moins bien corrigée que dans l'expérience de référence. Ainsi, le fait que les débits ne soient pas du même ordre de grandeur est en partie la cause de la mauvaise correction à Kerjean. L'ajout d'une station virtuelle sur le bief 4 (figure 9.8c) permet de retrouver les valeurs de référence pour les stations amont Tréodet et Kerjean. L'ajout de cette station permet de lever l'équifinalité. Nous pouvons conclure que la mauvaise correction des débits amont à Kerjean est liée essentiellement à l'équifinalité entre la correction à Tréodet et Kerjean, et une prédominance du forçage à Tréodet. La longueur de la fenêtre d'assimilation n'est pas à remettre en cause.

Résultats dans la phase d'analyse - Erreur temporelle et spatiale Pour la suite de l'étude des résultats, nous allons comparer les deux expériences $CLa_{amont}@Z_{vigilance}$ et $CLa@Z_{vigilance+virtu-b4}$ afin d'étudier l'impact d'un point de mesure supplémentaire sur le réseau.

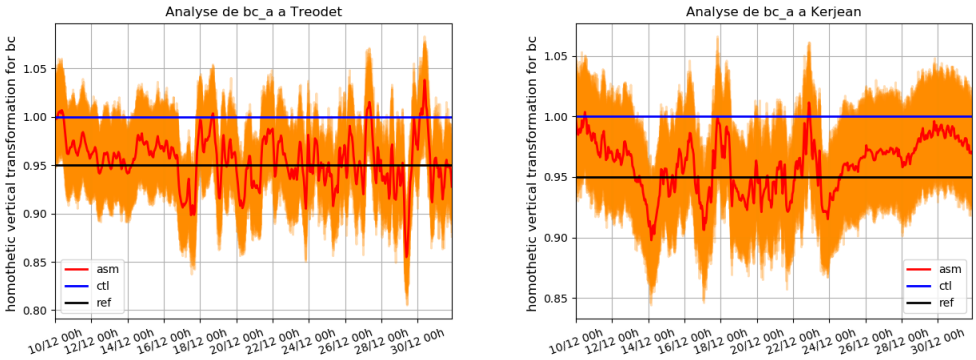
Les figures 9.9 et 9.10 montrent pour les trois stations de vigilance le résultat du calcul de l'erreur sur la hauteur d'eau simulée par la simulation de contrôle et par la simulation avec assimilation de données par rapport à la simulation de référence pour ces deux expériences. Nous notons que les erreurs sur les hauteurs d'eau sont réduites avec l'assimilation de données pour les trois stations dans les deux expériences. Les chroniques d'erreur sont



(a) Expérience $CLa@Z_{vigilance} - rea_{8h}$



(b) Expérience $CLa@Z_{vigilance} - ech_{CL}$



(c) Expérience $CLa@Z_{vigilance+virtu-b4}$

Figure 9.8: Valeurs analysées en fonction du temps pour la correction multiplicative des conditions amont.

quasiment identiques pour les deux expériences. Ainsi, l'erreur résiduelle à Kervir est très réduite dans l'expérience $CLa_{amont}@Z_{vigilance}$, bien que les paramètres analysés pour la correction des conditions aux limites à Tréodet et Kerjean soient éloignés de la valeur de référence (problème d'équifinalité évoqué plus haut). L'ajout de la station virtuelle permet de retrouver les valeurs de référence pour les paramètres analysés, mais n'améliore pas les résultats de l'assimilation de données aux stations de prévision.

Les graphiques figures 9.11 et 9.12 montrent les erreurs intégrées sur toute la période d'étude (22 jours) le long des biefs b_{1-2-3}, b_{4-2-3} et b_{5-3} pour les simulations de contrôle et l'assimilation de données par rapport à la simulation de référence pour les expériences $CLa_{amont}@Z_{vigilance}$ et $CLa@Z_{vigilance+virtu-b4}$. Nous notons que les graphiques sont analogues pour les deux expériences pour les biefs b_{1-2-3} et b_{5-3} : la correction des niveaux simulés est bonne sur ces biefs. Pour le bief b_{4-2-3} dont l'amont est la station Kerjean, on note une erreur résiduelle supérieure à celle de la simulation de contrôle à l'amont de la confluence avec le bief b_1 pour l'expérience $CLa_{amont}@Z_{vigilance}$. Celle-ci est due à l'équifinalité évoquée entraînant une mauvaise correction de condition limite à Kerjean. L'ajout de la station virtuelle sur le bief b_4 permet d'améliorer la simulation des débits sur ce bief.

Résultats dans la phase de prévision Les figures 9.13 et 9.14 montrent les résultats du calcul de la RMSE aux trois stations de vigilance en prévision pour les échéances variant de 1 heure à 24 heures pour les expériences $CLa_{amont}@Z_{vigilance}$ et $CLa@Z_{vigilance+virtu-b4}$. Nous notons que les erreurs de prévisions sont analogues pour les deux expériences. Les prévisions sont améliorées jusqu'à l'échéance 24 heures pour les trois stations. À Justice, nous mettons en évidence une erreur de prévision cyclique dont la période est la durée de la marée (environ 12 heures). Nous notons que cet aspect cyclique n'était pas présent pour l'expérience de correction des coefficients de Strickler $Ks@Z_{vigilance}$ (voir figure 9.6). On ne retrouve pas cet aspect cyclique sur les erreurs de prévision à Kervir et Moulin-Vert, ce qui signifie que le fait d'assimiler les hauteurs à Justice n'entraîne pas ici une correction cyclique sur tout le réseau. À Justice, l'erreur en prévision est plus faible pour les très courtes échéances, pour une échéance de l'ordre de 12 heures et pour une échéance de l'ordre de 24 heures. Cela correspond aux moments pour lesquels la phase de la marée est analogue à celle sur la fenêtre d'assimilation. On peut supposer que pour une certaine condition de marée, l'influence de la condition limite amont est la même 12 heures plus tard. Puisque la correction sur les conditions limites amont est la même pendant toute la prévision, l'erreur est du même ordre de grandeur.

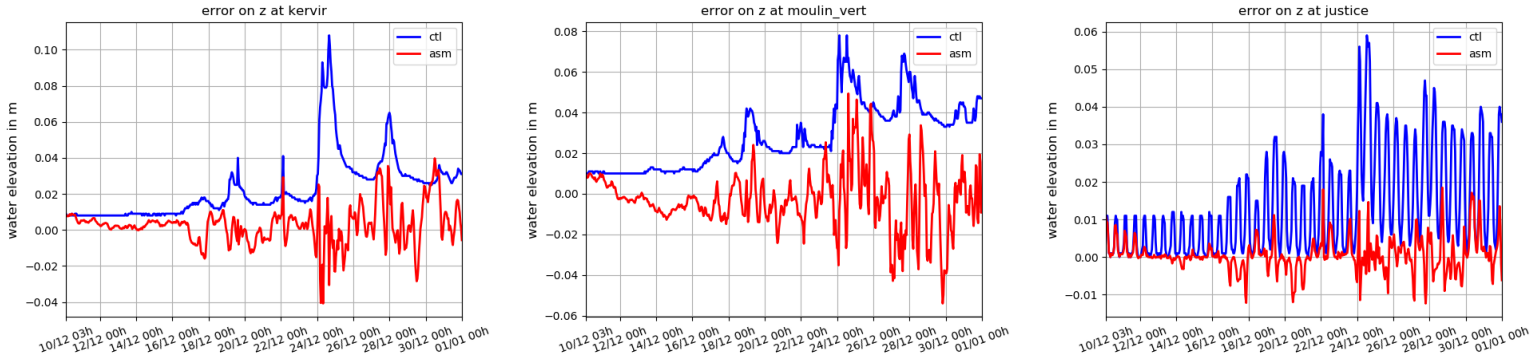


Figure 9.9: Expérience $CLA_{amont}@Z_{vigilance}$ - Chroniques temporelles des erreurs aux stations de vigilance Kervir, Moulin-Vert et Justice pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (rouge) par rapport à la simulation de référence.

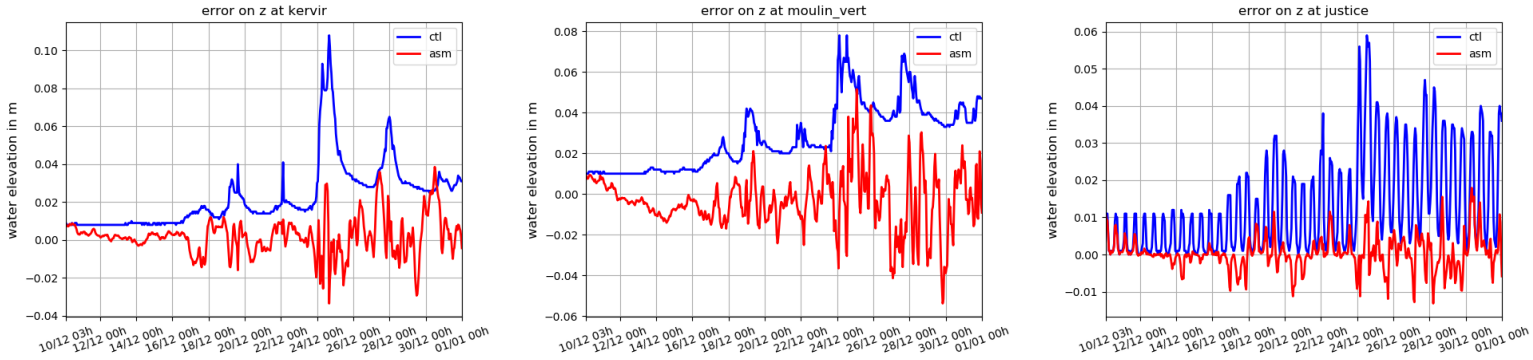


Figure 9.10: Expérience $CLa@Z_{vigilance+virtu-b4}$ - Chroniques temporelles des erreurs aux stations de vigilance Kervir, Moulin-Vert et Justice pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (rouge) par rapport à la simulation de référence.

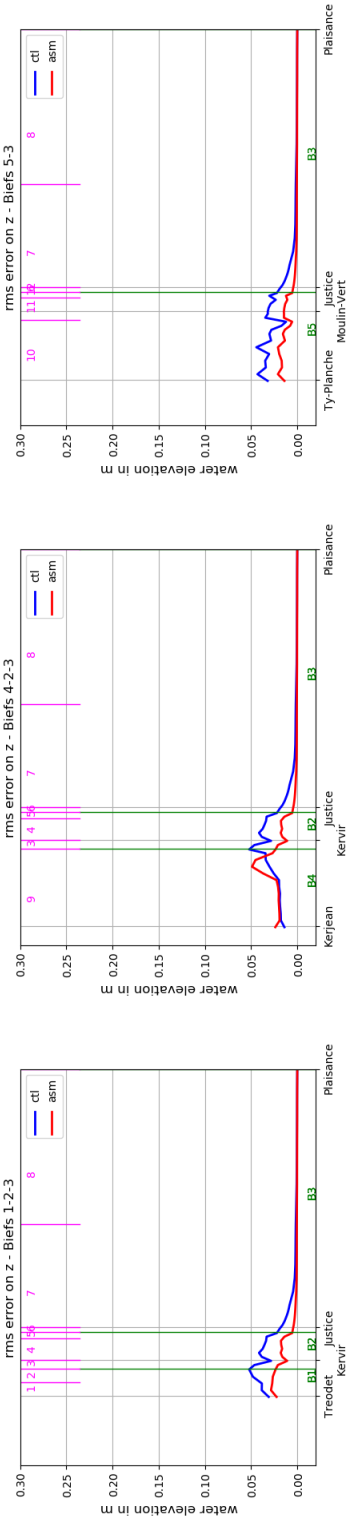


Figure 9.11: Expérience $CLA_{amont}@Z_{vigilance}$ - RMSE le long des biefs du modèle MASCARET pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (rouge) par rapport à la simulation de référence.

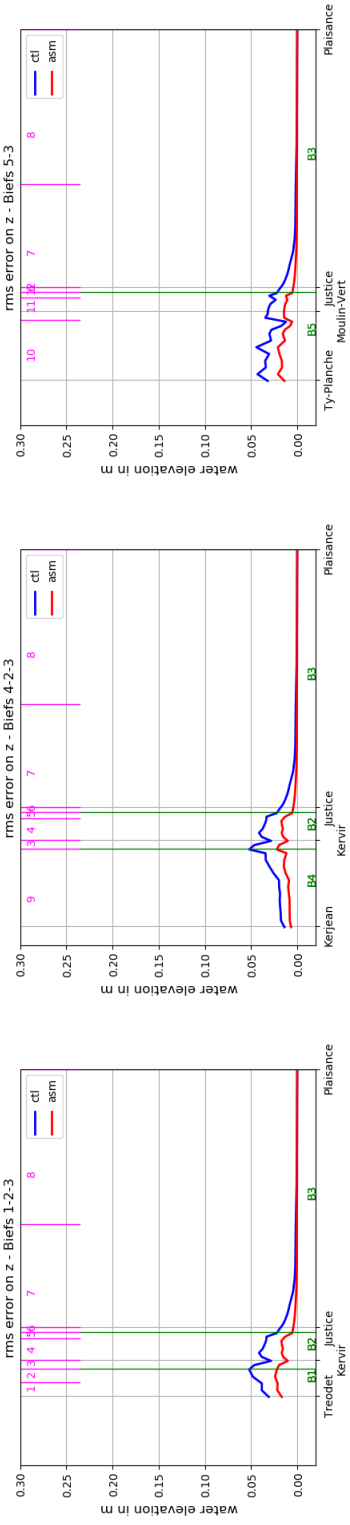


Figure 9.12: Expérience $CLa@Z_{vigilance+virtu-b4}$ - RMSE le long des biefs du modèle MASCARET pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (rouge) par rapport à la simulation de référence.

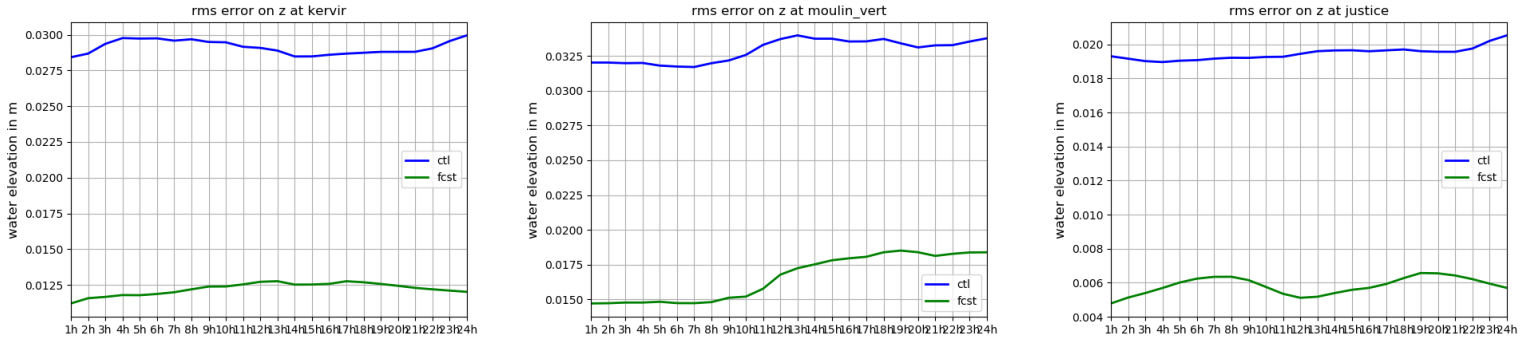


Figure 9.13: Expérience $CLA_{amont}@Z_{vigilance}$ - RMSE des erreurs de prévision aux stations Kervir, Moulin-Vert et Justice en fonction de l'échéance pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (vert) par rapport à la simulation de référence.

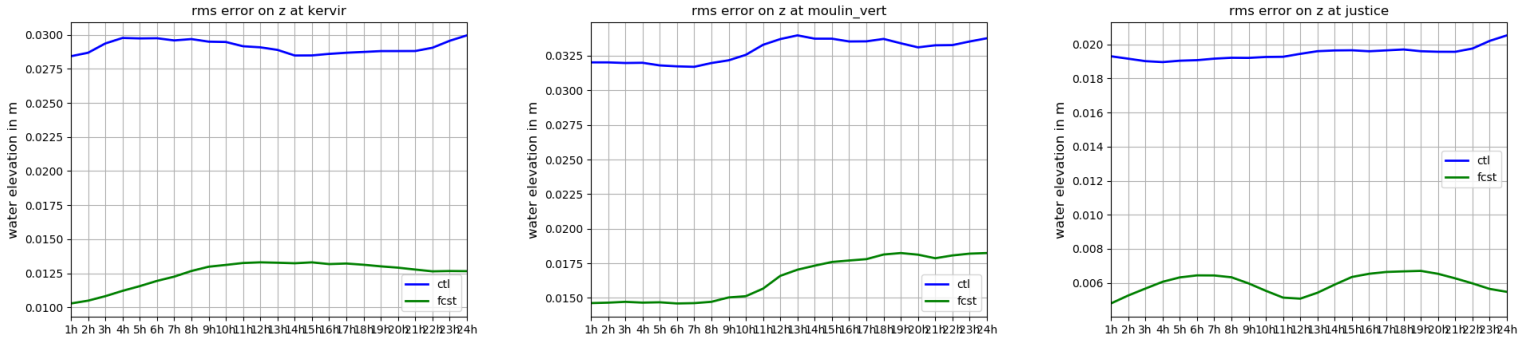


Figure 9.14: Expérience $CLa@Z_{vigilance+virtu-b4}$ - RMSE des erreurs de prévision aux stations Kervir, Moulin-Vert et Justice en fonction de l'échéance pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (vert) par rapport à la simulation de référence.

9.2.2.3 Correction de la condition limite aval : $CLb_{aval}@Z_{Justice}$

Dans l'expérience $CLb_{aval}@Z_{Justice}$, le coefficient additif b est ajouté à la condition limite aval pour la réalisation de la simulation de référence. Comme mentionné au paragraphe 9.2.1, celui-ci est variable en temps par quart de journée (voir tableau 9.1). Les hauteurs d'eau sont assimilées à la station Justice uniquement.

Résultats dans la phase d'analyse La figure 9.15 présente les valeurs analysées pour le vecteur de contrôle au cours de l'événement.

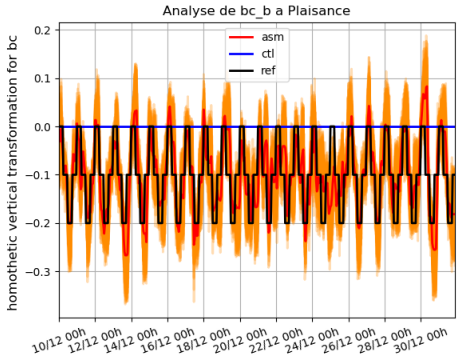


Figure 9.15: Expérience $CLb_{aval}@Z_{Justice}$ - Valeurs analysées pour l'unique élément du vecteur de contrôle (paramètre de correction additif de la condition limite aval b) en fonction du temps. Les courbes orange représentent les valeurs pour chacun des 49 membres et la courbe rouge la valeur analysée moyenne.

Ce graphique met en évidence que la valeur analysée du paramètre b est cyclique, comme la valeur imposée en référence. Les valeurs analysées suivent les valeurs imposées en référence. L'algorithme d'assimilation de données est ainsi capable de suivre une variation de la condition limite aval additive variable en temps. Il faut cependant garder à l'esprit qu'un temps d'adaptation est nécessaire à chaque nouvelle valeur imposée pour b, puisque tant que la fenêtre d'assimilation couvre des pas de temps pour lesquels la valeur de b n'est pas constante, la valeur analysée sera un compromis entre les deux valeurs différentes de la fenêtre.

La figure 9.16 montre pour la station Justice le résultat du calcul de l'erreur sur la hauteur d'eau simulée par la simulation de contrôle et par la simulation avec assimilation de données par rapport à la référence. Nous notons que l'erreur sur les hauteurs d'eau est réduite à Justice. Elle possède un caractère cyclique, avec des pics d'erreur négative à 0h et un pic d'erreur positive à 12h. On note aussi des pics d'erreur moins marqués entre 0h et 12h et 12h et 0h du jour suivant. Ceci est cohérent avec le paramètre b imposé pour la simulation de référence. À 0h par exemple, b passe de -10 cm à 0 cm. Sur le cycle qui assimile les hauteurs à 22h, 23h et 0h, l'assimilation doit faire un compris entre les données à 22h

et 23h pour lesquelles b vaut -10 cm, et la donnée à 0 h pour laquelle b vaut 0 cm. La tendance sera d'obtenir un b négatif, donc le niveau estimé en fin de fenêtre d'assimilation sera inférieur au niveau de la simulation de référence à 0h. On obtient alors une erreur négative. Le même raisonnement peut être fait pour les autres changements de la valeur de b. Le fait que les valeurs analysées mettent quelques cycles à retrouver la valeur de référence se répercute donc sur les erreurs sur les hauteurs d'eau simulées.

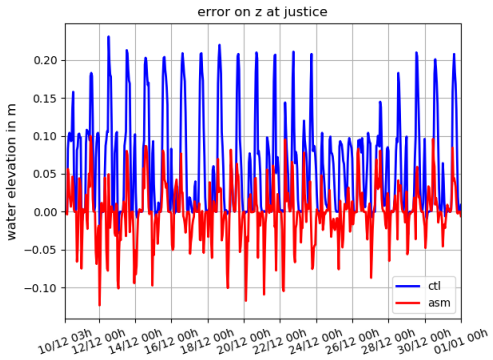


Figure 9.16: Expérience $CLb_{aval}@Z_{Justice}$ - Chroniques temporelles des erreurs à la station de vigilance Justice pour la simulation de contrôle (bleu) et la simulation avec assimilation de données en fin de fenêtre d'assimilation (rouge) par rapport à la simulation de référence.

Le graphique de la figure 9.17 montre les erreurs intégrées sur toute la période d'étude (22 jours) le long du bief b_{1-2-3} pour les simulations de contrôle et l'assimilation de données par rapport à la simulation de référence. Il n'est pas nécessaire de montrer ces graphiques sur les biefs b_{4-2-3} et b_{5-3} car l'erreur s'annule sur tous les biefs à partir de quelques centaines de mètres à l'amont de Justice. Ce graphique met en évidence une bonne correction des niveaux simulés sur la zone impactée par la cote imposée à Plaisance. Cette zone s'étend de Plaisance à un point situé sur le bief b_2 entre les stations Kervir et Justice.

Résultats dans la phase de prévision La figure 9.18 montre le résultat du calcul de la RMSE à Justice pour les échéances de prévision variant de 1 heure à 24 heures. Nous notons que les prévisions sont améliorées pour toutes les échéances. Nous notons également que la valeur de l'erreur sur la simulation de contrôle comme l'erreur résiduelle après assimilation sont cycliques. Comme lors de l'expérience précédente $CLa_{amont}@Z_{vigilance}$, la durée du cycle est celle d'un cycle de marée (environ 12 heures). Les explications données pour l'expérience précédente sont toujours valables. L'importance de l'amplitude du cycle est probablement amplifiée par le caractère cyclique du paramètre b imposé, qui varie toutes les 6 heures. Ici, l'erreur résiduelle est supérieure à l'erreur d'observation (5 cm).

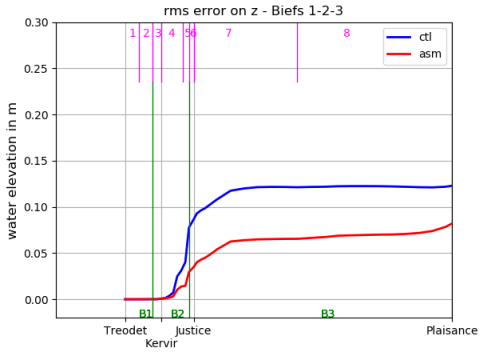


Figure 9.17: Expérience $CLb_{aval}@Z_{Justice}$ - RMSE le long du bief b_{1-2-3} du modèle MASCARET pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (rouge) par rapport à la simulation de référence.

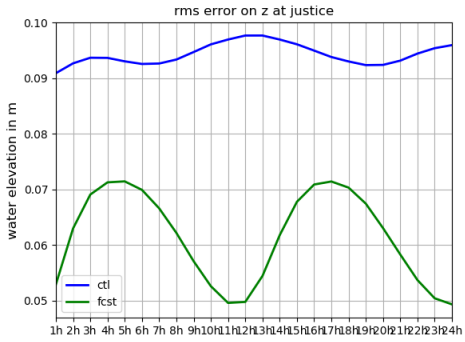


Figure 9.18: Expérience $CLb_{aval}@Z_{Justice}$ - RMSE des erreurs de prévision à la station Justice en fonction de l'échéance pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (vert) par rapport à la simulation de référence.

9.2.3 Correction des Ks, des conditions limites amont et aval simultanément

Le paragraphe 9.2.2 a été consacré aux expérience jumelles réalisées pour valider la correction des coefficients de Strickler, des conditions limites amont ou des conditions limites aval dans des expériences indépendantes. L'objet de ce présent paragraphe est l'étude de la correction simultanée des paramètres et des conditions aux limites. Le vecteur de contrôle est donc constitué des coefficients de Strickler des 5 zones, du paramètre multiplicatif a pour les trois conditions limites amont et du paramètre additif b pour la condition limite aval. Nous présentons les expériences suivantes :

- Dans l'expérience $KsCL@Z_{vigilance}$, les cotes d'eau aux trois stations de vigilance sont assimilées. Les résultats de cette expérience sont présentés au paragraphe 9.2.3.1.
- Dans l'expérience $KsCL@all_{zQ}$, les variables assimilées sont les cotes d'eau aux trois

stations de vigilance ainsi que les débits aux stations amont. Les résultats de cette expérience sont présentés au paragraphe 9.2.3.2. Cette expérience jumelle a de l'intérêt dans le cas où en expérience réelle, on intégrerait le modèle MASCARET sur la fenêtre d'assimilation en l'alimentant par des simulations MORDOR-TS. Cette piste de travail est intéressante car elle permet d'éviter le problème de changement de nature de l'erreur entre les phases d'assimilation et de prévision opérationnelle.

- L'expérience $KsCL@all_{zQ} - 279$ est similaire à l'expérience $KsCL@all_{zQ}$, mais l'ensemble possède 279 membres. Elle est présentée au paragraphe 9.2.3.3.

9.2.3.1 Assimilation des hauteurs : $KsCL@Z_{vigilance}$

Dans l'expérience $KsCL@Z_{vigilance}$, le vecteur de contrôle est composé des coefficients de Strickler des 5 zones, des correction multiplicatives avec le paramètre a pour les trois stations amont et de la correction additive avec le paramètre b à Plaisance. Les perturbations imposées pour la simulation de référence sont toujours celles définies dans le tableau 9.1. La majeure partie des résultats de cette expérience d'assimilation de données est présentée en annexe R.1.1 : vecteur de contrôle analysé (figures R.1 et R.2), chronique d'erreur aux stations de vigilance (figure R.3) et RMSE en prévision selon l'échéance (figure R.4). La RMSE le long des biefs est présentée figure 9.19.

Résultats dans la phase d'analyse La comparaison des valeurs analysées pour le vecteur de contrôle avec celles obtenues au paragraphe 9.2.2 (figures 9.3, 9.7 et 9.15) pour lesquels la correction des coefficients de frottement, des conditions limites amont et aval étaient réalisés de manière indépendante met en évidence que pour de nombreux éléments du vecteur de contrôle, les corrections sont proches. C'est le cas pour les paramètres $Ks4_{min}$, $Ks11_{min}$, $Ks7_{min}$, le paramètre a à Ty-Planche et à Kerjean et le paramètre b à Plaisance. Les tendances sont légèrement différentes pour les paramètres $Ks8_{min}$ et $Ks8_{maj}$. La tendance est significativement différente pour le paramètre a à Tréodet. Dans l'expérience présentée ici, la valeur analysée de a est significativement inférieure à la valeur imposée en référence, alors que la sous-estimation du paramètre n'était pas notable dans l'expérience $CLa_{amont}@Z_{vigilance}$.

Comme pour les expériences menées de manière indépendante, nous notons là aussi que l'erreur temporelle sur les hauteurs d'eau aux trois stations de vigilance sont nettement réduites avec l'assimilation de données pour les trois stations.

Les graphiques figure 9.19 montrent les erreurs intégrées sur toute la période d'étude (22 jours) le long des biefs b_{1-2-3}, b_{4-2-3} et b_{5-3} pour les simulations de contrôle et l'assimilation de données par rapport à la simulation de référence. Ces graphiques mettent en évidence une bonne correction des niveaux simulés le long des biefs à l'aval de Kervir et sur le Steïr. En revanche, à l'amont de Kervir, l'assimilation de données dégrade les

niveaux simulés par rapport à la simulation de contrôle. Cette tendance était déjà visible dans l'expérience $CLa_{amont}@Z_{vigilance}$ (voir figure 9.11 pour le bief b_{4-2-3}). Puisque la valeur analysée pour le paramètre a est du même ordre de grandeur à Kerjean et est dégradée à Tréodet par rapport à cette expérience, il est normal d'obtenir de mauvais résultats sur les biefs amont b_1 et b_4.

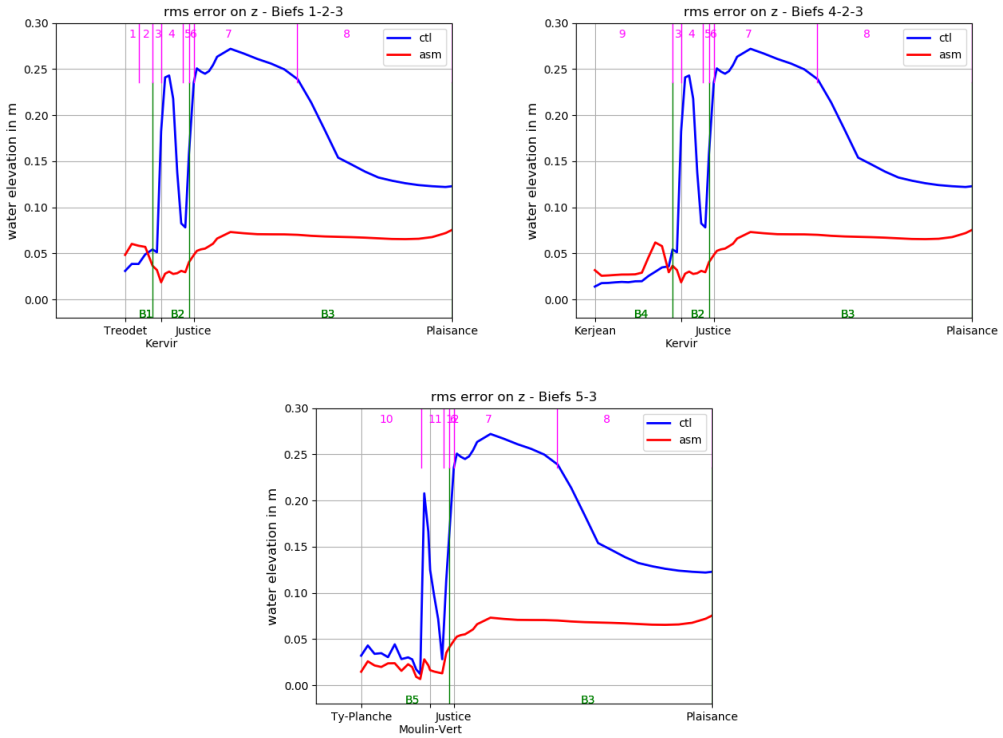


Figure 9.19: Expérience $KsCL@Z_{vigilance}$ - RMSE le long des biefs du modèle MASCARET pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (rouge) par rapport à la simulation de référence.

Résultats dans la phase de prévision De même que pour les expériences de correction des paramètres réalisées séparément, la correction simultanée permet d'améliorer les prévisions de manière notable jusqu'à l'échéance 24 heures. L'erreur résiduelle à Justice présente un caractère cyclique comme pour les expériences de correction des conditions limites amont $CLA_{amont}@Z_{vigilance}$ et aval $CLb_{aval}@Z_{Justice}$. Nous avons réalisé ici une expérience comparative d'assimilation de données avec une fenêtre d'assimilation de 12 heures (résultats non présentés dans ce manuscrit). De cette manière, à chaque cycle d'assimilation, les données assimilées couvrent toutes les phases de marée possibles. Dans cette expérience, le caractère cyclique des erreurs de prévision à Justice est toujours présent.

9.2.3.2 Assimilation des hauteurs et des conditions limites : $KsCL@all_{zQ}$

L'expérience $KsCL@all_{zQ}$ est analogue à la précédente en ce qui concerne les paramètres corrigés (coefficients de Stricklers, paramètres de correction multiplicative a et paramètres de correction additive b). En revanche, les variables assimilées sont toujours les hauteurs d'eau à Kervir, Moulin-Vert et Justice, auxquelles on ajoute les débits aux trois stations amont Tréodet, Kerjean et Ty-Planche.

Résultats dans la phase d'analyse Les figures présentant les valeurs analysées du vecteur de contrôle au cours de l'événement sont présentées en annexe R.1.2, figure R.5, pour les coefficients de Strickler et le paramètre de correction additive de la condition limite aval b. Les analyses réalisées sur ces différents paramètres sont analogues à celles obtenues avec l'expérience précédente $KsCL@Z_{vigilance}$. Les résultats sont en revanche différents pour les paramètres de correction multiplicative a aux trois stations amont, dont les valeurs analysées sont présentées sur la figure 9.20. Le fait d'assimiler les débits à l'amont permet cette fois-ci de très bien corriger les paramètres de correction multiplicative des lois à l'amont a.

Les autres résultats de cette expérience d'assimilation de données obtenus en phase d'analyse sont présentés en annexe R.1.2 : chronique d'erreurs aux stations de vigilance (figure R.6) et RMSE le long des biefs (figure R.7).

Nous notons que les erreurs aux trois stations de vigilance sont du même ordre de grandeur que lorsque l'on n'assimile pas les débits amont. Le fait de lever l'équifinalité n'améliore pas les résultats aux trois stations. De plus, nous mettons en évidence que le fait d'assimiler les débits à l'amont permet de corriger les lois à Tréodet et Kerjean de manière correcte. Nous ne retrouvons pas comme dans l'expérience $KsCL@Z_{vigilance}$ une dégradation des simulations sur les biefs b_1 et b_4. Ici, l'assimilation de données permet de réduire la RMSE le long de tous les biefs.

Résultats dans la phase de prévision La RMSE calculée en prévision est présentée en annexe R.1.2, figure R.8. Nous mettons en évidence le fait que les erreurs en prévision sont analogues à celles obtenues dans l'expérience $KsCL@Z_{vigilance}$ sans assimilation des débits amont.

Conclusion Le fait d'assimiler les débits aux trois stations amont permet de lever l'équifinalité pour la correction des paramètres a pour les stations Tréodet et Kerjean. Elle permet une très bonne analyse de ce paramètre aux trois stations amont. Cela permet d'améliorer les résultats de l'assimilation de données sur les biefs b_1 et b_4 qui sont directement à l'aval de ces stations. Ainsi, il est probable que des cartes d'inondation construites à partir de ces résultats seraient de meilleure qualité. En revanche, les résultats

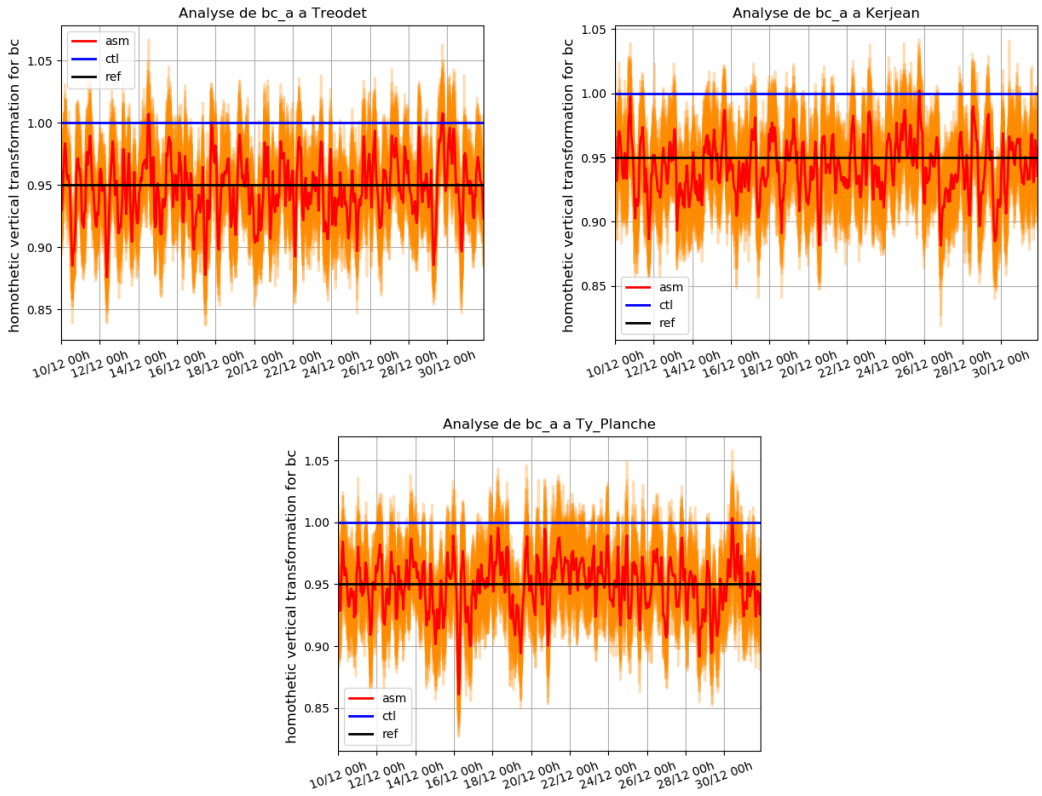


Figure 9.20: Expérience $KsCL@all_{zQ}$ - Valeurs analysées pour la correction des conditions limites amont avec le paramètre a en fonction du temps. Les courbes orange représentent les valeurs pour chacun des 49 membres et la courbe rouge la valeur analysée moyenne.

sur les trois stations de vigilance Kervir, Moulin-Vert et Justice sont quasiment identiques. Ils ne sont pas améliorés dans cette expérience.

9.2.3.3 Assimilation des hauteurs et des conditions limites, test avec 279 membres : $KsCL@all_{zQ} - 279$

L'expérience $KsCL@all_{zQ} - 279$ est identique à la précédente, sauf que le nombre de membres de l'ensemble est désormais 279 à la place de 49. Les résultats de cette expérience d'assimilation de données sont présentés en annexe R.1.3 : vecteur de contrôle analysé (figures R.9 et R.10), chronique d'erreur aux stations de vigilance (figure R.11), RMSE le long des biefs (figure R.12) et RMSE en prévision selon l'échéance (figure R.13).

Nous notons que les valeurs des paramètres analysés sont identiques à l'expérience avec 49 membres, sauf pour $Ks8_{min}$ et $Ks8_{maj}$. Pour ces deux paramètres, on note cependant que la tendance est la même pour les deux expériences. Les chroniques temporelles des erreurs sont les mêmes pour les trois stations dans les deux expériences, ainsi que les erreurs le

long des biefs. En prévision également, les RMSE sont identiques. Nous pouvons donc conclure que le passage de 49 à 279 membres n'améliore pas l'algorithme d'assimilation de données. Nous conservons donc pour la suite 49 membres.

9.2.4 Synthèse

Dans ce paragraphe, nous avons dans un premier temps présenté le cadre de la réalisation des expériences jumelles. Le vecteur de contrôle peut contenir les 5 coefficients de frottement retenus suite à l'étude d'incertitude et/ou des paramètres de correction multiplicative des conditions limites amont et/ou un paramètre de correction additive de la condition limite aval. Les données assimilées sont les hauteurs d'eau aux trois stations de vigilance Kervir, Moulin-Vert et Justice, et les débits aux trois stations amont Tréodet, Kerjean et Ty-Planche le cas échéant. La fenêtre d'assimilation dure 3 heures, elle contient donc 3 pas de temps d'observation. Le nombre de membres de l'algorithme EnKF est fixé à 49. Le cyclage est horaire et la fenêtre de prévision opérationnelle dure 24 heures.

Deux types d'expériences jumelles ont été réalisées. Dans un premier temps, nous avons cherché à corriger de manière indépendante les coefficients de frottement, les conditions limites amont et la condition limite aval. Ces expériences ont mis en évidence le fait que les coefficients de Strickler sont bien corrigés par l'algorithme. En revanche, un problème d'équifinalité est mis en évidence pour la correction des lois en condition limite amont entre Tréodet et Kerjean. Ce problème est levé par l'ajout d'une station virtuelle sur le bief b_4 situé à l'aval de Kerjean. La correction de la condition limite aval avec un paramètre additif variant dans le temps est de bonne qualité. Pour toutes ces expériences, les erreurs des hauteurs simulées aux trois stations de vigilance sont réduites par l'assimilation de données. En prévision, le modèle corrigé permet également d'améliorer les prévisions par rapport à la simulation de contrôle pour toutes les échéances étudiées, jusqu'à 24 heures.

Dans un second temps, tous les paramètres ont été corrigés simultanément. Nous avons mis en évidence que les valeurs des paramètres analysés sont souvent proches de celles obtenues avec les expériences indépendantes, sauf pour le paramètre de correction de la loi à Tréodet. Le fait d'assimiler les débits aux trois stations amont permet d'améliorer nettement les valeurs analysées pour les paramètres de correction des lois à l'amont. Cela n'a pas d'influence sur les résultats des simulations aux trois stations de vigilance, néanmoins, nous avons vu que la densification du réseau d'observation améliore les simulations sur l'ensemble du réseau hydraulique. Pour finir, nous avons montré que les résultats sont similaires lorsque l'on considère 279 membres dans la dernière expérience menée $KsCL@all_{zQ}$. La suite du travail est réalisée avec 49 membres.

Ces expériences jumelles permettent de valider l'algorithme d'assimilation de données développé dans le cadre de cette thèse pour le modèle MASCARET. Il est à présent possible de l'appliquer en conditions réelles dans un contexte opérationnel.

9.3 Réalisation des expériences réelles

Nous travaillons à présent en conditions réelles. Le paragraphe 9.3.1 présente les caractéristiques des expériences réelles menées. Ensuite, nous nous intéressons dans le paragraphe 9.3.2 aux résultats de l'assimilation de données dans la phase d'analyse. La présentation des résultats en phase prévision opérationnelle fera l'objet du chapitre suivant (chapitre 10), dans lequel seront comparées les méthodes de calibration statistique et d'assimilation de données pour la génération de prévisions d'ensemble hydrauliques.

9.3.1 Description des expériences réelles

Périodes pour les expériences réelles Les expériences réelles d'assimilation de données sont générées sur 3 hivers en dehors de la période utilisée pour la construction des forêts aléatoires (d'octobre 2011 à juin 2014, voir tableau 8.3 dans le chapitre 8). Cela permet de comparer les méthodes de calibration statistique et d'assimilation de données dans un contexte opérationnel. L'assimilation de données est lancée de manière cyclée sur chacun de ces hivers en démarrant avec les valeurs calées des coefficients de Strickler, les coefficients de correction multiplicative des lois amont $a = 1$ et le coefficient de correction additive de la loi aval $b = 0$. Les dates pour les trois hivers considérés sont choisies en fonction du début des événements pluvieux significatifs :

- $H_{2010/2011}$: du 26 octobre 2010 au 15 mars 2011. Durant cet hiver, la station de mesure à Justice n'était pas encore opérationnelle. Les cotes d'eau sont donc assimilées uniquement aux stations Kervir et Moulin-Vert.
- H_{2015} : du 1er janvier 2015 au 26 mars 2015.
- $H_{2015/2016}$: du 4 novembre 2015 au 15 avril 2016.

Quatre expériences Quatre configurations différentes sont testées pour les expériences réelles. Leurs caractéristiques sont résumées dans le tableau 9.2.

Comme pour les expériences jumelles, à chaque expérience est associé une simulation de contrôle qui correspond à la simulation du modèle MASCARET sans assimilation dont les forçages sont les mêmes que ceux de la fenêtre d'assimilation. Il y a donc deux simulations de contrôle. La première, appelée *SimCont-12*, correspond aux trois expériences *DA1-persist*, *DA1* et *DA2*. Le modèle MASCARET est alimenté par les mesures de débit à l'amont et de hauteur à l'aval. La seconde, appelée *SimCont-3*, correspond à l'expérience *DA3*. Le modèle MASCARET est alors alimenté par la simulation de MORDOR-TS déterministe alimenté par les mesures de pluie *ANTILOPE* à l'amont et par les mesures de hauteur à l'aval.

Pour toutes les expériences, lorsque les conditions limites amont sont corrigées, elles le

Expérience	Forçages fenêtre d'assimilation	Vecteur de contrôle	Observations assimilées	Forçage Q_{am} prévision	Forçage H_{av} prévision
DA1-persist	Q_{am} : mesures H_{aval} : mesures	Ks CL am av	Cotes	persistance	Marée astronomique + persistance surcote
DA1				*raw-HEF*	
DA2		Ks CL av		*QRF-hydro-TS*	
DA3	Q_{am} : MORDOR det H_{aval} : mesures	Ks CL am av	Cotes Débits	*raw-HEF*	

Tableau 9.2: Caractéristiques des 4 expériences d'assimilation de données. Le forçage MORDOR det correspond à une simulation MORDOR-TS déterministe forcée par les mesures de pluie ANTILOPE. Le vecteur de contrôle est constitué des coefficients de Strickler (Ks), du paramètre de correction additive b pour la condition limite aval (CL av) et des paramètres de correction multiplicative a pour les conditions limites amont (CL am) le cas échéant.

sont par un coefficient multiplicatif a. La valeur du paramètre analysée est alors persistée sur la fenêtre de prévision opérationnelle et appliquée sur le forçage utilisé. La condition limite aval est toujours corrigée par l'intermédiaire d'un paramètre additif b. En prévision, la surcote corrigée par le paramètre b est alors persistée. La valeur de la surcote corrigée est ajoutée à la prévision de marée astronomique.

Expérience *DA1-persist* : L'expérience *DA1-persist* est une expérience de référence, c'est la plus simple que l'on puisse mener. Les prévisions hydrologiques ne sont pas utilisées dans la phase de prévision, on persiste simplement les conditions limites amont au moment du temps de base. Les caractéristiques de cette expérience sont les suivantes :

- Les forçages sur la fenêtre d'assimilation sont les mesures pour le débit amont et la cote à l'aval.
- Les paramètres contrôlés sont les coefficients de Strickler et les conditions limites à l'amont et à l'aval.
- Les observations assimilées sont les cotes aux trois stations de vigilance Kervir, Moulin-Vert et Justice.
- En prévision opérationnelle, les conditions limites amont corrigées sont persistées. Ainsi, pour chaque station amont, on impose pour toute la phase de prévision une valeur constante égale à la valeur imposée au dernier pas de temps de la fenêtre d'assimilation. Cette valeur tient compte de la correction analysée pour le paramètre a.

Expérience *DA1* : Cette expérience est identique à l'expérience *DA1*-persist pour la partie analyse (forçages, variables corrigées et observations assimilées). Elle diffère de la précédente par les chroniques utilisées en prévision opérationnelle, puisqu'on utilise alors les prévisions hydrologiques brutes *raw-HEF*. Les paramètres multiplicatifs de correction des débits amont a sont donc appliqués aux chroniques de l'ensemble *raw-HEF*.

Expérience *DA2* : Cette expérience est identique à *DA1* en ce qui concerne les forçages sur la fenêtre d'assimilation et les observations assimilées. Les points pour lesquels elle

diffère de l'expérience *DA1* sont les suivants :

- Les paramètres contrôlés sont les coefficients de Strickler et la condition limite aval uniquement. On ne corrige pas les conditions limites amont.
- En prévision, les prévisions hydrologiques calibrées *QRF-hydro-TS* sont utilisées.

Ainsi, dans cette expérience, on ne corrige pas les débits amont en phase d'analyse. On considère que les débits hydrologiques calibrés sont bons et n'ont pas besoin d'être corrigés. La comparaison des expériences *DA1* et *DA2* permettra ainsi d'évaluer quelle est la meilleure approche pour corriger l'hydrologie : correction paramétrique de l'ensemble hydrologique brut *raw-HEF* par assimilation de données dans le modèle MASCARET (*DA1*) ou calibration statistique (*DA2*).

Expérience *DA3* : Dans cette expérience, on va essayer de diminuer le problème présent dans l'expérience *DA1* lié au changement de la nature de l'erreur entre les forçages utilisés sur la fenêtre d'assimilation et pour la prévision opérationnelle. Les caractéristiques sont les suivantes :

- Les forçages à l'amont sur la fenêtre d'assimilation sont les simulations réalisées avec le modèle MORDOR-TS déterministe forcé avec les pluies mesurées ANTILOPE. À l'aval, le modèle est toujours forcé par la cote mesurée à Plaisance.
- Les paramètres contrôlés sont les coefficients de Strickler et les conditions limites à l'amont et à l'aval.
- Les observations assimilées sont les cotes aux trois stations de vigilance Kervir, Moulin-Vert et Justice ainsi que les débits aux trois stations amont Tréodet, Kerjean et Ty-Planche.
- En prévision, le modèle est alimenté par la prévision hydrologique brute *raw-HEF*.

Dans l'expérience *DA1*, les débits amont sont corrigés en prévision avec une correction apprise sur la fenêtre d'assimilation sur laquelle le modèle est forcé par des débits mesurés. Nous n'avons donc pas la même nature de l'erreur sur les débits amont entre la phase d'analyse et de prévision. L'expérience *DA3* a pour but d'évaluer les performances des prévisions d'ensemble construites avec une assimilation de données pour laquelle les erreurs sur les fenêtres d'assimilation et de prévision opérationnelle sont davantage similaires. C'est pour cela que le modèle MASCARET est forcé dès la fenêtre d'assimilation par des débits simulés avec MORDOR-TS. Cependant, il est à noter que la nature de l'erreur est toujours un peu différente, puisque le modèle MORDOR-TS est alimenté sur la fenêtre d'assimilation par les mesures de pluie ANTILOPE, alors qu'il est alimenté par les prévisions de pluie RR3 en phase prévision.

L'exprience qui paraît la plus prometteuse pour l'opérationnel est l'expérience *DA2* dans laquelle on utilise un forçage amont corrigé par calibration statistique.

Nombre de membres et erreur d'observation Les expériences d'assimilation de données sont réalisées avec 49 membres. L'écart-type d'erreur d'observation est le même que pour les expériences jumelles, à savoir

- une valeur constante valant 5 cm pour les observations de hauteur d'eau ;
- un pourcentage de la valeur mesurée de 5% pour les observations de débit.

9.3.2 Résultats dans la phase d'analyse

Dans ce paragraphe, nous présentons les résultats de l'assimilation de données dans la phase d'analyse. Puisque les expériences *DA1-persist* et *DA1* sont identiques sur la période d'analyse, nous ne parlerons que des trois expériences *DA1*, *DA2* et *DA3*.

9.3.2.1 Etude des paramètres analysés

Les chroniques des paramètres analysés pour les trois expériences *DA1*, *DA2* et *DA3* sont présentées en annexe R.2 (figure R.14 pour l'hiver $H_{2010/2011}$, figure R.15 pour l'hiver H_{2015} et figure R.16 pour l'hiver $H_{2015/2016}$).

Nous notons tout d'abord une différence concernant les variations des paramètres analysés entre l'hiver $H_{2010/2011}$ pour lequel les cotes d'eau ne sont pas assimilées à Justice et les deux autres hivers. En effet, la variabilité des paramètres $Ks7_{min}$, $Ks8_{min}$, $Ks8_{maj}$, du paramètre b à Plaisance et dans une moindre mesure de $Ks11_{min}$ est beaucoup plus importante lorsque les hauteurs d'eau sont assimilées à Justice. Ceci est probablement lié au fait que les hauteurs à Justice sont très influencées par la marée, et potentiellement les erreurs de simulation possèdent également un caractère cyclique. Le paramètre $Ks11_{min}$ semble moins impacté par l'ajout de l'assimilation des hauteurs à Justice. Ceci est cohérent avec le fait que le bief sur lequel se trouve la zone 11 est quasiment coupé hydrauliquement du bief sur lequel se trouve Justice (voir chapitre 2, paragraphe 2.3.2.2). Les figures R.17 et R.18 placées en annexe R.2 permettent d'évaluer ces variations à l'échelle d'une journée : elles représentent les variations du vecteur de contrôle analysé sur une journée de l'hiver $H_{2010/2011}$ et de l'hiver H_{2015}, donc sans et avec assimilation des cotes d'eau à Justice. Lorsque des cotes d'eau sont assimilées à Justice (figure R.18), tous les paramètres autres que les paramètres pour la correction des débits amont a et $Ksmin_{11}$ présentent une variation cyclique. Lorsque les données à Justice ne sont pas disponibles (figure R.17), cette variation cyclique des valeurs analysées n'est pas présente. Ainsi, nous mettons en évidence que lorsque les hauteurs sont assimilées à Justice, les paramètres qui contrôlent cette hauteur ($Ksmin_7$, $Ksmin_8$, $Ksmaj_8$, et b à plaisance) dépendent eux aussi de la phase de la marée. Cette remarque servira en particulier à expliquer les performances des expériences d'assimilation de données pour la prévision des pics de marée haute à Justice.

Nous nous intéressons à présent aux valeurs analysées pour les paramètres a de correction

des conditions limites amont. Les conditions limites amont sont corrigées uniquement dans les expériences *DA1* et *DA3*. La différence entre ces deux expériences sur la fenêtre d'assimilation réside dans le choix du forçage amont pour le modèle MASCARET : dans l'expérience *DA1*, le modèle est forcé par les mesures de débit, alors que dans l'expérience *DA3*, il est forcé par une simulation hydrologique déterministe réalisée avec MORDOR-TS alimenté par les mesures de pluie ANTILOPE. Dans l'expérience *DA3*, l'erreur liée au modèle hydrologique est donc déjà présente dans la phase d'analyse. Le tableau 9.3 présente les valeurs minimales et maximales du paramètre analysé a aux trois stations amont pour ces deux expériences. Il met en évidence que globalement, la correction du paramètre a est beaucoup plus importante pour l'expérience *DA3*, bien que cela soit moins marqué pour l'hiver $H_{2015/2016}$. Puisque le modèle MASCARET est alimenté par des prévisions hydrologiques et non des mesures dans l'expérience *DA3*, les corrections à faire sur les conditions limites amont sont plus prononcées.

	DA1			*DA3*		
	$H_{2010/2011}$	H_{2015}	$H_{2015/2016}$	$H_{2010/2011}$	H_{2015}	$H_{2015/2016}$
a min Tréodet	0.98	0.92	1.00	0.47	0.68	1.00
a max Tréodet	1.07	1.24	1.65	1.54	1.74	1.66
a min Kerjean	0.96	0.90	0.74	0.45	0.79	1.00
a max Kerjean	1.04	1.07	1.17	1.35	1.41	1.40
a min Ty-Palnche	0.81	0.90	0.15	0.68	0.91	0.16
a max Ty-Planche	1.08	1.24	1.14	1.55	1.82	1.48

Tableau 9.3: Valeurs minimales et maximales du paramètre a analysé sur les chroniques des 3 hivers étudiés pour les ensembles *DA1* et *DA3*.

9.3.2.2 Hauteurs simulées dans la phase d'analyse

Les tableaux 9.4 et 9.5 présentent les résultats du calcul de la RMSE calculée entre les observations et les simulations analysées des expériences *DA1*, *DA2* et *DA3*. Les simulations analysées sont d'abord moyennées pour chaque pas de temps, puis la RMSE est calculée sur tous les pas de temps. On compare également ces RMSE à celles obtenues avec la simulation de contrôle correspondante. Les RMSE sont calculées sur les trois hivers de la période de validation simultanément (2 hivers pour la station Justice).

Ces tableaux montrent tout d'abord que la simulation de contôle présente une RMSE nettement plus petite lorsqu'elle est alimentée par les mesures de débit à l'amont (simulation *SimCont-12*) plutôt que par les débits issus des simulations avec MORDOR-TS (simulation *SimCont-3*). Ceci est normal car le calage du modèle MASCARET a été réalisé avec les mesures de débit aux stations amont. Les trois expériences d'assimilation de données permettent de diminuer de manière significative la RMSE à Kervir et Moulin-Vert. Sur ces deux stations, l'assimilation de données mise en place permet donc de corriger de manière satisfaisante les paramètres afin d'obtenir une bonne chronique analysée, même lorsque l'on part de forçages amont avec une erreur plus marquée (expérience *DA3*).

À la station Justice en revanche, les expériences *DA1* et *DA2* dégradent les performances par rapport à la simulation de contrôle correspondante *SimCont-12*. Seule l'expérience *DA3* permet d'améliorer l'erreur par rapport à la simulation de contrôle correspondante *SimCont-3*. On retrouve alors le même score que pour la simulation de contrôle *SimCont-12*. Ainsi, les expériences d'assimilation de données à Justice ne permettent pas d'améliorer les hauteurs simulées à Justice sur la fenêtre d'assimilation par rapport au modèle calé.

Pour les trois expériences, il est à noter que les meilleures RMSE en fin de fenêtre d'assimilation sont obtenues pour les trois stations pour l'expérience *DA3* pour les trois stations, alors que la simulation de contrôle correspondante donne les moins bons résultats.

Kervir			Moulin-Vert			Justice		
SimCont-12	*DA1*	*DA2*	*SimCont-12*	*DA1*	*DA2*	*SimCont-12*	*DA1*	*DA2*
9.5	1.7	1.3	4.7	2.8	2.3	5.7	9.1	7.9

Tableau 9.4: RMSE (cm) calculée par rapport aux mesures pour chaque station de vigilance pour les simulations de contrôle et avec assimilation de données pour les expériences *DA1* et *DA2*.

Kervir		Moulin-Vert		Justice	
SimCont-3	*DA3*	*SimCont-3*	*DA3*	*SimCont-3*	*DA3*
14.9	1.1	14.8	1.8	7.8	5.7

Tableau 9.5: RMSE (cm) calculée par rapport aux mesures pour chaque station de vigilance pour les simulations de contrôle et avec assimilation de données pour l'expérience *DA3*.

Ces tendances sont confirmées par l'examen des graphiques représentant les chroniques de hauteurs d'eau mesurées et simulées. Les figures 9.21a et 9.21b présentent les chroniques de hauteurs d'eau simulées par le modèle MASCARET pour les deux simulations de contrôle *SimCont-12* et *SimCont-3* et pour les trois expériences d'assimilation de données *DA1*, *DA2* et *DA3* au cours de deux événements d'importance moyenne : du 13 au 15 novembre 2010 et du 1er au 15 janvier 2016. Nous rappelons que les hauteurs d'eau ne sont pas assimilées à Justice en 2010. Pour les expériences d'assimilation de données, les hauteurs d'eau correspondent au dernier pas de temps de la fenêtre d'assimilation. La hauteur d'eau mesurée est également ajoutée aux graphiques.

Ces graphiques confirment que le modèle MASCARET calé génère de moins bons résultats lorsqu'il est alimenté par les simulations MORDOR-TS à l'amont : les chroniques *SimCont-3* sont plus éloignées des observations que les chroniques *SimCont-12*. Les trois expériences fournissent des simulations analysées qui semblent de qualité comparable sur ces événements. Cela rejoint les calculs des RMSE présentés aux tableaux 9.4 et 9.5 pour Kervir et Moulin-Vert. Pour Justice, le graphique présenté ne permet pas de mettre en évidence une dégradation des simulations analysées *DA1* et *DA2* par rapport à la simulation de contrôle *SimCont12* comme nous pouvons le voir avec le calcul des RMSE. Nous

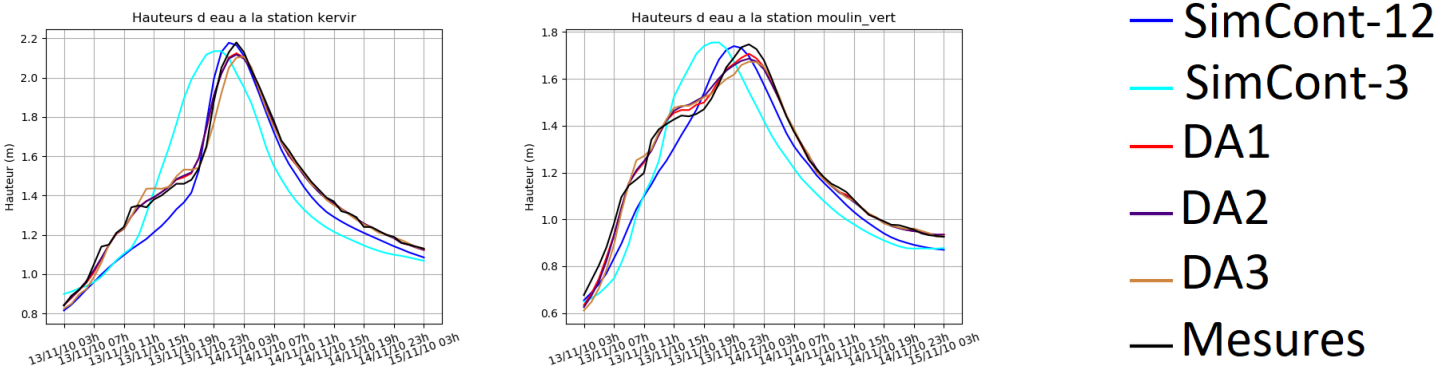
notons ponctuellement des erreurs d'analyse à Kervir et Moulin-Vert sur la période du 1er au 15 janvier 2016. Au cours de cet événement, on note une sous-estimation des hauteurs à Kervir et Moulin-Vert pour la simulation de contrôle *SimCont3*. Cette sous-estimation est probablement dûe à une sous-estimation des débits amont. À Justice, cela a des conséquences uniquement à marée basse. En effet, à marée haute, comme nous l'avons déjà mentionné, c'est la courbe de remous qui remonte de l'aval. La valeur du débit amont n'est pas impactante. Ces résultats ont déjà été mis en évidence au cours de l'étude de sensibilité.

9.3.3 Synthèse

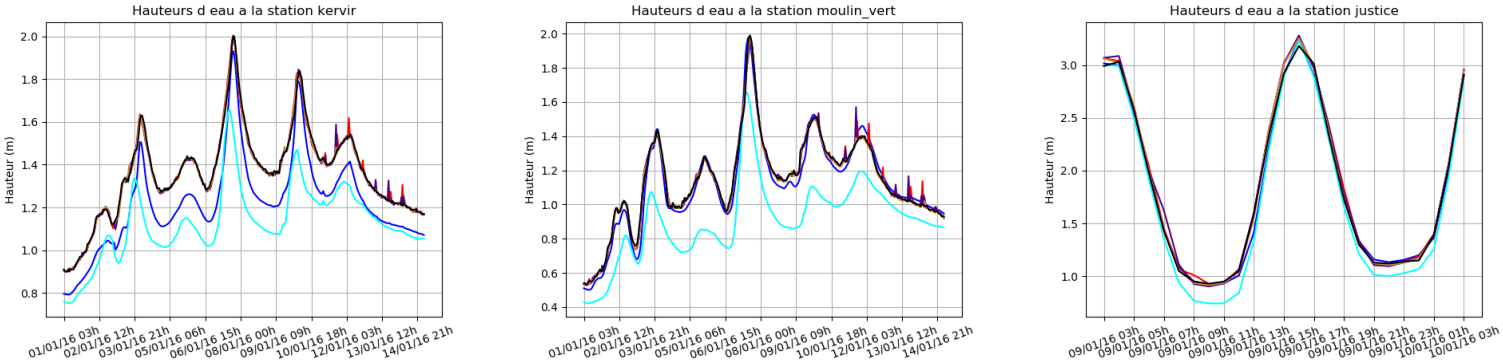
Dans ce chapitre, nous avons présenté les expériences jumelles réalisées dans le cadre de l'assimilation de données intégrée dans le modèle MASCARET. Elles ont permis de valider l'algorithme d'assimilation de données mis en place.

Ensuite, quatre expériences d'assimilation de données ont été décrites dans un cadre opérationnel. Si l'on s'intéresse uniquement à la phase d'analyse, deux d'entre elles sont identiques. Nous nous sommes donc attachés à examiner les performances de trois expériences sur la fenêtre d'assimilation. Pour ces trois expériences, l'assimilation de données permet d'améliorer significativement les hauteurs d'eau modélisées aux stations Kervir et Moulin-Vert. Les variations temporelles du vecteur de contrôle analysé sont la preuve que l'assimilation de données est pertinente par rapport à un calage du modèle.

À la station Justice en revanche, les performances obtenues ne dépassent jamais celles du modèle calé alimenté avec les mesures de débit à l'amont. On pourrait envisager un autre paramétrage de la correction de la condition limite aval. Ici, la correction appliquée est uniquement une correction additive de la surcote. Il serait intéressant de voir si les résultats à Justice pourraient être améliorés en ajoutant une correction multiplicative et/ou un déphasage.



(a) Stations Kervir et Moulin-Vert - Du 13 au 15 novembre 2010.



(b) Stations Kervir, Moulin-Vert et Justice - Du 1er au 15 janvier 2016 (Journée du 9 janvier 2016 pour Justice).

Figure 9.21: Chroniques de hauteurs d'eau : mesure, simulées par le modèle MASCARET calé alimenté par les mesures à l'amont, simulées par le modèle MASCARET calé alimenté par les simulations MORDOR-TS à l'amont et simulées avec assimilation de données pour les expériences *DA1*, *DA2* et *DA3*.

Chapitre 10

Comparaison des méthodes QRF et EnKF

Sommaire

10.1 Résultats à Kervir 293
10.2 Résultats à Moulin-Vert 296
10.3 Résultats à Justice 301
10.4 Synthèse 304

Ce chapitre a pour objet la comparaison des performances des prévisions d'ensemble hydrauliques générées dans le cadre de cette thèse. Nous rappelons ci-dessous les deux méthodologies développées ainsi que les ensembles associés :

- La calibration statistique des ensembles hydrauliques bruts a été présentée au chapitre 8. Nous avons conclu que la meilleure expérience à retenir pour générer des prévisions d'ensemble à la station Justice est l'expérience *PASC* (*Prévision Astronomique Surcote Constante*) dans laquelle la surcote mesurée au temps de base est persistée sur les 24 heures de prévision opérationnelle pour la génération des ensembles bruts et pour le calcul des prédicteurs associés. À Kervir et Moulin-Vert, le choix de la condition limite aval a un impact négligeable. Nous avons alors évalué 6 ensembles différents. Pour mémoire, le détail de leur génération est synthétisé figure 8.2. Nous résumons dans l'encart ci-dessous les caractéristiques de ces 6 ensembles. Pour la station Justice, nous avons en particulier montré que la calibration de l'ensemble hydrologique n'avait pas d'impact notable sur les ensembles construits. Pour cette station, pour plus de lisibilité, nous présentons uniquement les résultats obtenus avec les ensembles R_R, Q_R et QNE_{Md}.

Ensembles issus de la calibration statistique

R_R : ensemble hydraulique brut, condition limite amont *raw-HEF*.

Q_R : quantiles calibrés à partir de l'ensemble hydraulique brut R_R.

QNE_{Md} : quantiles calibrés, mêmes prédicteurs que pour Q_R sauf les prédicteurs issus de statistiques de l'ensemble brut R_R.

R_Q : ensemble hydraulique brut, condition limite amont *QRF-Hydro-TS*.

Q_Q : quantiles calibrés à partir de l'ensemble hydraulique brut R_Q.

QNE_Q : quantiles calibrés, mêmes prédicteurs que pour Q_Q sauf les prédicteurs issus de statistiques de l'ensemble brut R_Q.

- L'introduction d'assimilation de données ensembliste dans le modèle hydraulique MASCARET a fait l'objet du chapitre 9. Nous y avons décrit quatre expériences d'assimilation de données. Le tableau 10.1, déjà présenté au chapitre 9, synthétise leurs caractéristiques en termes de forçage sur la fenêtre d'assimilation, variables corrigées, observations assimilées et forçages en prévision. Il est à noter que les ensembles bruts R_R et R_Q générés dans

le cadre de la calibration statistique servent aussi de référence pour les ensembles obtenus par assimilation de données. Ils ne sont pas spécifiquement liés à la méthodologie QRF puisqu'elle est appliquée après leur génération.

<table>
<tr><th>Expérience</th><th>Forçages
fenêtre d'assimilation</th><th>Vecteur de
contrôle</th><th>Observations
assimilées</th><th>Forçage Q_{am}
prévision</th><th>Forçage H_{av}
prévision</th></tr>
<tr><td>DA1-persist</td><td rowspan="2">Q_{am} : mesures
H_{aval} : mesures</td><td rowspan="2">Ks
CL am av</td><td rowspan="3">Cotes</td><td>persistance</td><td rowspan="4">Marée
astrono-
mique
+
persistance
surcote</td></tr>
<tr><td>DA1</td><td>raw-HEF</td></tr>
<tr><td>DA2</td><td>Ks
CL av</td><td>QRF-hydro-TS</td></tr>
<tr><td>DA3</td><td>Q_{am} : MORDOR det
H_{aval} : mesures</td><td>Ks
CL am av</td><td>Cotes
Débits</td><td>raw-HEF</td></tr>
</table>

Tableau 10.1: Caractéristiques des 4 expériences d'assimilation de données.

Afin de comparer les méthodes dans un contexte opérationnel, les ensembles sont évalués sur des périodes indépendantes de celle qui a servi à la construction des forêts aléatoires. Nous travaillons sur les mêmes événements que ceux présentés lors de la réalisation des expériences réelles d'assimilation de données (voir paragraphe 9.3). Pour mémoire, il s'agit des hivers suivants :

- $H_{2010/2011}$: du 26 octobre 2010 au 15 mars 2011. Pendant cet hiver, les cotes d'eau ne sont pas assimilées à Justice du fait de leur indisponibilité.
- H_{2015} : du 1er janvier 2015 au 26 mars 2015.
- $H_{2015/2016}$: du 4 novembre 2015 au 15 avril 2016.

Le calcul des scores aux stations Kervir et Moulin-Vert est réalisé après stratification afin d'évaluer les performances des ensembles au cours des événements pluvieux. La stratification est réalisée à partir de la prévision issue de l'ensemble brut R_R : comme au chapitre 8 au moment de la calibration statistique des ensembles, nous retenons les pas de temps de prévision pour lesquels le maximum de l'ensemble dépasse le quantile q_{90} de hauteur mesurée à la station. Les résultats à Kervir sont présentés au paragraphe 10.1 et ceux à Moulin-Vert au paragraphe 10.2. À Justice, nous rappelons que l'objectif est de réaliser une prévision d'ensemble des pics de pleine mer. Les résultats à cette station sont présentés au paragraphe 10.3.

10.1 Résultats à Kervir

La figure 10.1 présente les caractéristiques des 10 ensembles construits en prévision à la station Kervir pour les échéances de prévision allant de 1 heure à 24 heures : CRPS, RMSE, norme $||.||_2$ des histogrammes de rang, et aire AUC sous la courbe ROC pour les événements de dépassement des seuils q_{90} et q_{99}. La RMSE est calculée à partir de la moyenne des ensembles. Elle est également calculée pour la prévision déterministe

chaînée, qui consiste en l'enchaînement du modèle MORDOR-TS déterministe et du modèle MASCARET déterministe. En annexe S.1.1 sont présentés les histogrammes de rang des 10 ensembles aux échéances 6, 12, 18 et 24 heures dont les normes sont synthétisées sur la figure 10.1.

Tout d'abord, nous notons que pour tous les scores présentés, les caractéristiques des ensembles générés par calibration statistique sont similaires à celles obtenues sur la période de validation des forêts aléatoire (voir graphique 8.4 en validation). Le fait d'avoir réduit la période d'analyse ne modifie pas les résultats.

Le graphique présentant les CRPS en prévision met en évidence qu'en-dehors de l'expérience d'assimilation *DA1-persist*, tous les ensembles construits par calibration statistique ou assimilation de données améliorent les CRPS par rapport à l'ensemble brut R_R pour toutes les échéances de prévision. Ces méthodes fournissent également de meilleurs CRPS que ceux de l'ensemble hydraulique brut alimenté par l'hydrologie calibrée R_Q, bien que les performances peuvent être équivalentes selon les ensembles considérés pour des échéances supérieures à 12 heures. En effet, comme nous l'avons vu au chapitre 8, les ensembles Q_R et QNE_{Md} présentent des CRPS analogues à l'ensemble R_Q pour les grandes échéances. L'ensemble obtenu par assimilation de données avec la persistance des conditions aux limites *DA1-persist* présente des CRPS qui augmentent au fil du temps, jusqu'à dépasser les CRPS de l'ensemble brut R_R. L'ensemble généré par assimilation de données *DA2* se dégage par ses performances en termes de valeurs de CRPS qui sont significativement inférieurs aux autres ensembles pour toutes les échéances de prévision[1]. L'expérience d'assimilation de données *DA3* fournit également des CRPS meilleurs que ceux obtenus par calibration statistique, sauf aux très courtes échéances. L'expérience d'assimilation de données *DA1* fournit des CRPS du même ordre de grandeur que ceux obtenus par calibration statistique, sauf aux longues échéances pour lesquelles ils sont de meilleure qualité et se rapprochent de ceux obtenus avec *DA3*.

Le graphique présentant le calcul de la RMSE met en évidence que l'ensemble brut R_R présente une erreur très légèrement inférieure à la prévision déterministe. La stratégie de génération de l'ensemble hydraulique brut permet ainsi de réduire l'erreur de prévision moyenne à la marge. Les erreurs de la moyenne des ensembles suivent les mêmes tendances que celles mises en évidence pour le CRPS. Là-encore, l'ensemble *DA2* présente des erreurs en prévision significativement réduites par rapport aux autres ensembles. Pour des échéances supérieures à 16 heures, les expériences d'assimilation de données *DA1* et *DA3* présentent elles aussi des erreurs inférieures à la calibration statistique.

En ce qui concerne la norme des histogrammes de rang, les ensembles générés par assimilation de données *DA1* et *DA3* présentent des normes du même ordre de grandeur que

[1]Nous rappelons que le CRPS est négativement orienté. Les plus petites valeurs pour ce score sont préférables.

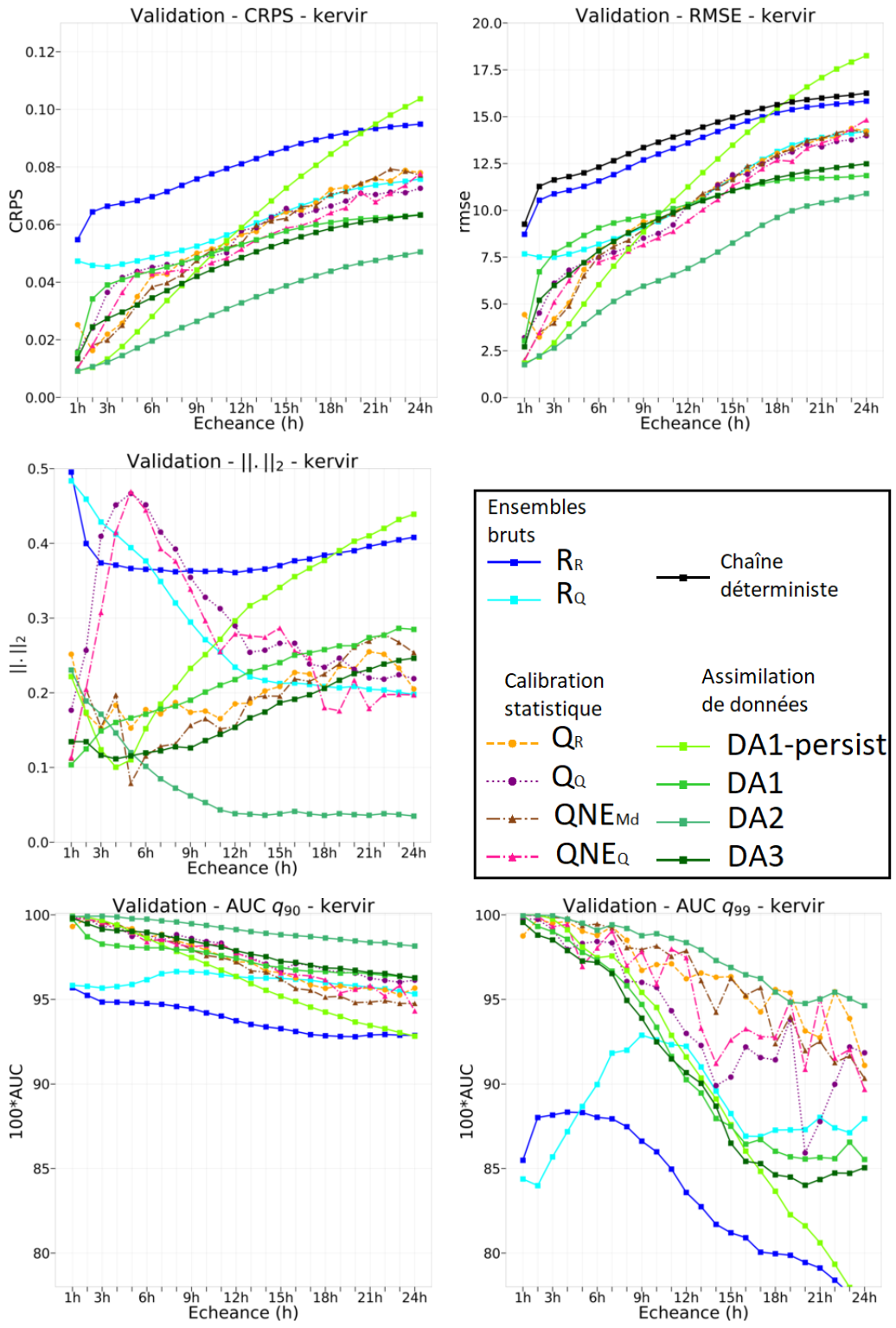


Figure 10.1: CRPS, RMSE, norme $||.||_2$ et aire ROC *AUC* pour les événements de dépassement des quantiles q_{90} et q_{99} pour les 10 prévisions d'ensemble hydrauliques générées pour les échéances de 1 heure à 24 heures à Kervir. Calculs effectués sur les 3 hivers de validation.

celles obtenus avec les calibrations Q_R et QNE_{Md}. L'expérience *DA3* génère cependant des ensembles plus fiables que l'expérience *DA1*. L'ensemble généré avec la persistance des conditions limites *DA1-persist* présente des normes d'histogrammes qui augmentent continuellement avec l'échéance de prévision à partir de l'échéance 4 heures. L'ensemble *DA2* présente là aussi des normes pour les histogrammes de rang significativement plus faibles que les autres ensembles à partir de l'échéance 6 heures. Pour les échéances plus courtes, la fiabilité est meilleure pour l'expérience *DA3*. Nous notons par ailleurs que les courbes de la norme $||.||_2$ ont la même tendance pour les ensembles R_Q et *DA2*, avec une décroissance de la norme jusqu'à l'échéance 12 heures environ puis une norme constante avec les échéances. Ces deux ensembles utilisent comme forçage l'ensemble hydrologique calibré *QRF-hydro-TS* en prévision. On peut faire la même remarque pour les ensembles R_R, *DA1* et *DA3*, qui utilisent tous les trois l'hydrologie brute *raw-HEF* en prévision. Ainsi, la fiabilité des ensembles hydrauliques semble liée au choix des ensembles hydrologiques utilisés en forçage.

Pour finir, l'analyse du score *AUC* met en évidence que l'ensemble *DA2* est le plus à même de prévoir le passage des seuils q_{90} et q_{99}. Les ensembles générés par assimilation de données *DA1* et *DA3* sont plutôt plus performants que ceux générés par calibration statistique pour le seuil q_{90}. Pour le seuil q_{99}, ces deux ensembles présentent une mauvaise capacité de prévision du dépassement du seuil. Dans tous les cas, l'ensemble *DA1-persist* présente de mauvaises performances.

À titre d'illustration, la figure 10.2 présente les prévisions cyclées à échéance 12 heures à Kervir réalisées avec les 10 ensembles construits dans le cadre de ce travail, pour l'événement du 13 au 17 novembre 2010. Les graphiques analogues aux échéances 6, 18 et 24 heures ainsi que les résultats sur l'événement du 2 au 15 janvier 2016 sont présentés en annexe S.1.2. Notons qu'il est difficile d'appréhender les performances respectives des différents ensembles à partir de ces graphiques. En particulier, si l'on n'y prend pas garde, il peut apparaître au premier abord que l'expérience *DA1-persist* est très performante puisque la prévision a exactement la même forme que la mesure, et l'écart-type est faible. Il faut bien garder à l'esprit que ces courbes doivent être analysées pour un temps donné t. On compare alors la prévision au temps t avec la mesure au même temps. Alors, la prévision *DA1-persist* devient nettement moins intéressante.

10.2 Résultats à Moulin-Vert

La figure 10.3 présente les résultats analogues de la figure 10.1 à la station Moulin-Vert. En annexe S.2 sont présentés les histogrammes de rang des 10 ensembles aux échéances 6, 12, 18 et 24 heures. On trouvera également dans cette annexe les graphiques des prévisions cyclées aux échéances 6, 12, 18 et 24 heures pour les 10 ensembles pour les événements du

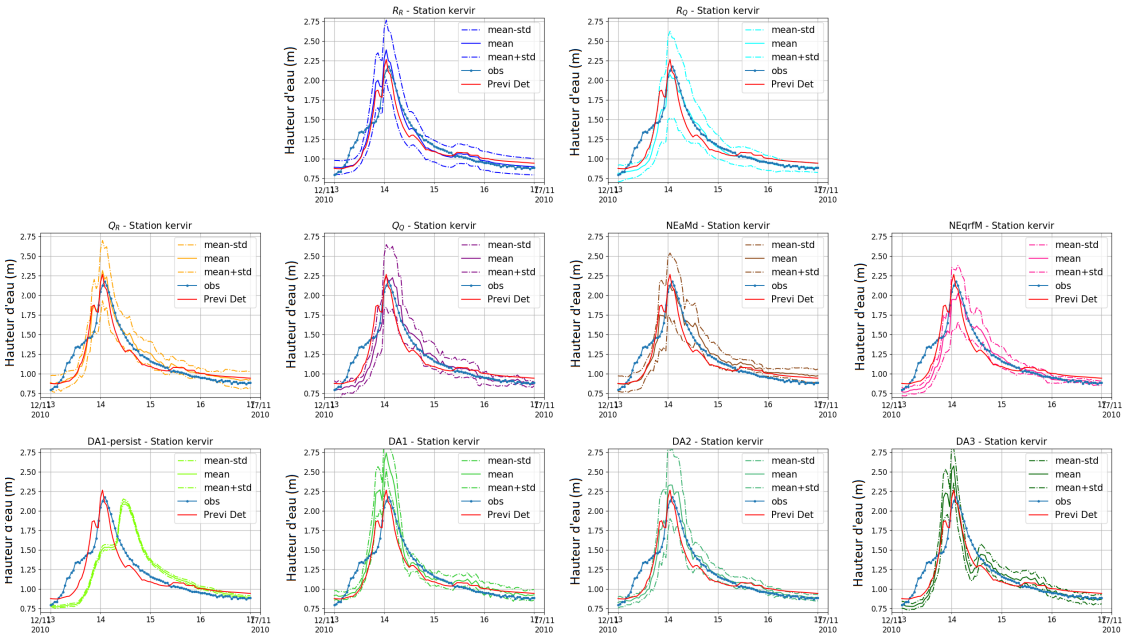


Figure 10.2: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 12 heures à Kervir. Événement du 13 au 17 novembre 2010.

13 au 17 novembre 2010 et du 2 au 15 janvier 2016.

De même que pour la station Kervir, les scores présentés pour les ensembles générés par calibration statistique sont similaires à ceux obtenus sur la période de validation des forêts aléatoires, qui est plus large que celle examinée ici. Par ailleurs, pour tous les scores calculés, on remarque là aussi comme à Kervir que l'ensemble *DA1-persist* présente des CRPS et RMSE croissants avec l'échéance de prévision, des normes pour les histogrammes de rang croissants à partir de l'échéance 3 heures et de mauvaises discriminations pour les événements de dépassement de seuil.

L'analyse des CRPS met en évidence deux groupes ayant des caractéristiques similaires. D'une part, les ensembles obtenus par calibration statistique et l'ensemble généré par assimilation de données *DA2* présentent des CRPS du même ordre de grandeur et la même tendance avec les échéances. L'ensemble brut R_Q présente des CRPS similaires à ceux de ce groupe. D'autre part, les ensembles *DA1* et *DA3* présentent des CRPS moins bons, et similaires à ceux obtenus avec l'ensemble brut R_R.

L'analyse des RMSE montre comme à Kervir que l'ensemble brut R_R présente une erreur plus petite que celle obtenue avec le chaînage déterministe des modèles MORDOR-TS et MASCARET. Comme pour les CRPS, les expériences d'assimilation de données *DA1* et *DA3* n'améliorent pas l'erreur de prévision. L'ensemble *DA2* présente les mêmes RMSE que l'ensemble R_Q, sauf pour les très courtes échéances pour lesquelles il est un peu moins performant. Les ensembles obtenus par calibration statistique présentent des erreurs du

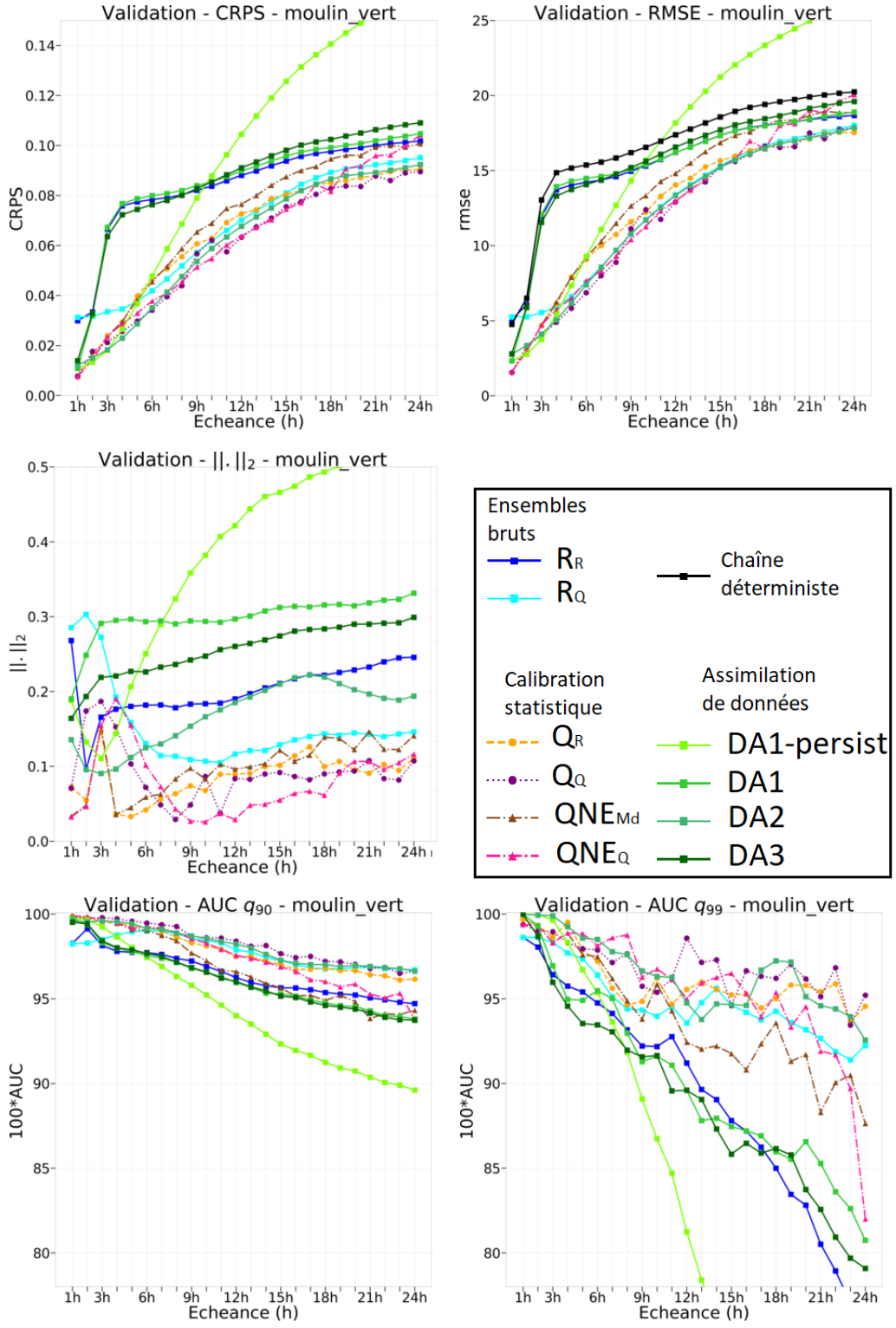


Figure 10.3: CRPS, RMSE, norme $||.||_2$ et aire ROC AUC pour les événements de dépassement des quantiles q_{90} et q_{99} pour les 10 prévisions d'ensemble hydrauliques générées pour les échéances de 1 heure à 24 heures à Moulin-Vert. Calculs effectués sur les 3 hivers de validation.

même ordre que celles de l'ensemble R_Q. Ainsi, pour la station Moulin-Vert, les méthodes mises en œuvre pour la génération des ensembles hydrauliques ne permettent pas de réduire l'erreur en prévision par rapport à l'ensemble brut R_Q alimenté par l'hydrologie calibrée *QRF-hydro-TS*.

L'analyse des normes $||.||_2$ des histogrammes de rang met en évidence que la calibration statistique permet d'obtenir les meilleures fiabilités. L'ensemble *DA2* ne fournit pas des ensembles très fiables comme à Kervir, bien que cet ensemble soit toujours celui qui possède les meilleures fiabilités parmi les expériences d'assimilation de données.

L'analyse de l'aire *AUC* sous les courbes ROC met en évidence que l'ensemble *DA2* génère les meilleurs ensembles pour ce score parmi les ensembles générés avec assimilation de données. Les performances de cet ensemble sont comparables à celles obtenues avec la calibration statistique Q_Q.

Ainsi, nous notons que les expériences d'assimilation de données fonctionnent significativement moins bien à Moulin-Vert qu'à Kervir. L'assimilation de données ne parvient pas à générer des ensembles ayant des performances supérieures à celles de l'ensemble brut généré avec la même condition limite amont. On peut se demander si cela est lié au problème de la manœuvre du seuil à l'aval de Moulin-Vert, pour lequel nous ne connaissons pas les dates ni les cotes *a posteriori* (voir explications au paragraphe 2.3.2.2, partie *Focus sur le calage de la station Moulin-Vert*). En effet, il est possible qu'à cause de la manœuvre du seuil qui peut ajouter un incrément de plusieurs dizaines de centimètres, le signe de l'écart entre le modèle et l'observation n'est pas le même à Justice et à Moulin-Vert. L'assimilation de données n'est alors pas capable de satisfaire les deux conditions. Pour infirmer ou confirmer cette hypothèse, nous représentons figure 10.4 la figure analogue à la figure 10.3, mais tracée uniquement sur l'hiver $H_{2010/2011}$ sur lequel aucune valeur n'est assimilée à Justice car la station n'était pas encore en service. Nous pouvons remarquer que pour l'ensemble *DA2*, les CRPS sont parmi les meilleurs. En revanche cela n'est plus vrai pour la RMSE. Ici, la fiabilité des ensembles obtenus par assimilation de données est globalement meilleure que celle obtenue par calibration statistique, si l'on met de côté l'ensemble DA1-persist. L'hypothèse formulée concernant la manœuvre du seuil est donc plausible. Il est à noter que cette manœuvre de seuil impacte potentiellement moins les résultats de la calibrations statistique. Les prédicteurs représentant le mois de la prévision et la hauteur d'eau mesurée au temps de base permettent en effet de discrimier les périodes pendant lesquelles le seuil est bas ou relevé (voir paragraphe 8.2.1 pour la définition des prédicteurs).

Par ailleurs, nous avons vu dans le paragraphe 8.2.3.5 que les résultats sont améliorés de manière notable à Moulin-Vert pour les CRPS et la discrimination des événements de dépassement de seuil lorsque le prédicteur représentant le débit simulé à Ty-Planche est issu d'une modélisation MORDOR-TS alimenté avec les mesures de pluie ANTILOPE.

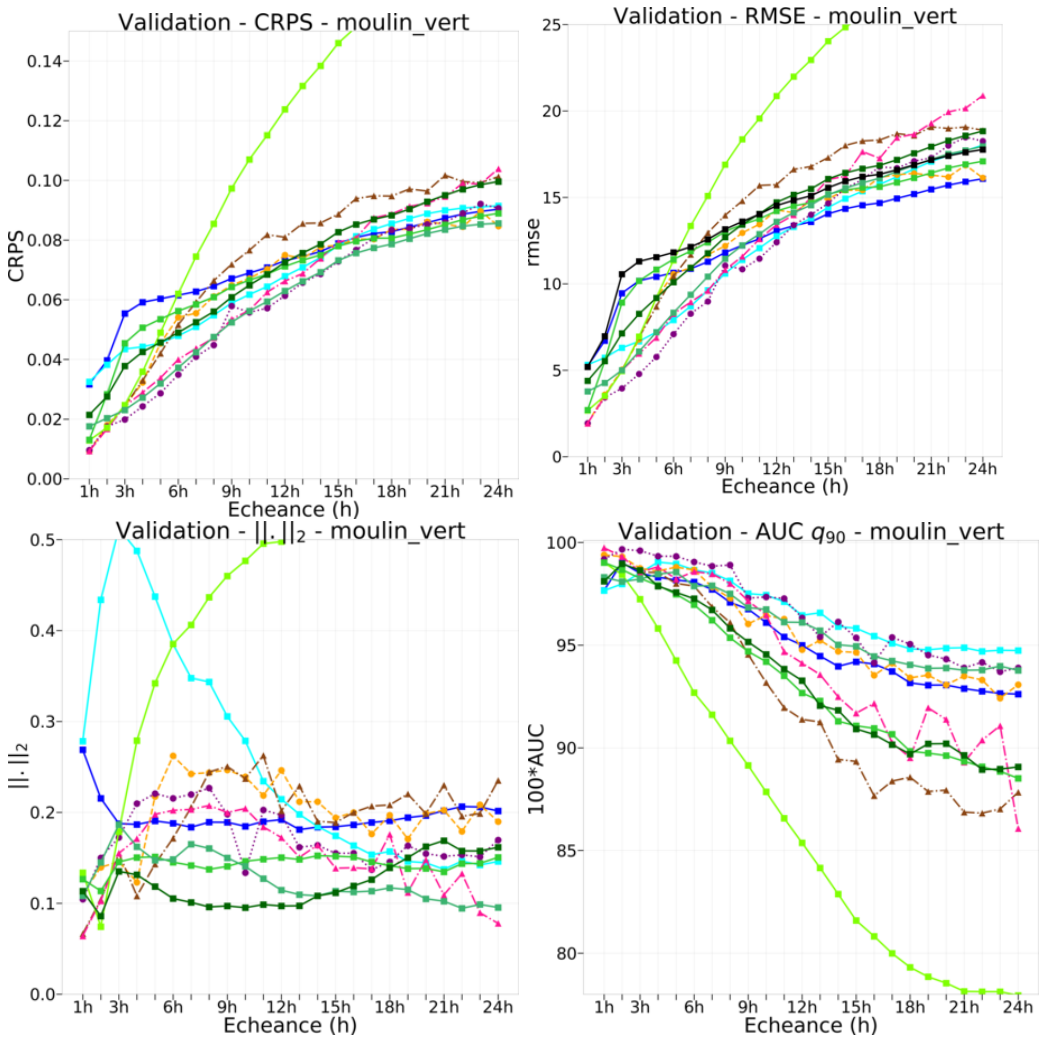


Figure 10.4: CRPS, RMSE, norme $||.||_2$ et aire ROC AUC pour les événements de dépassement des quantiles q_{90} et q_{99} pour les 10 prévisions d'ensemble hydrauliques générées pour les échéances de 1 heure à 24 heures à Moulin-Vert. Calculs effectués sur l'hiver $H_{2010/2011}$ sans assimilation de données à Justice. Se reporter à la figure 10.3 pour la légende.

Une piste d'amélioration des résultats obtenus à Moulin-Vert réside dans l'amélioration des pluies qui alimentent le modèle MORDOR-TS. Les ensembles obtenus par calibration statistique auraient alors potentiellement de meilleures performances. L'utilisation des prévisions de pluie PEAROME ou d'autres prévisions d'ensemble calibrées permettraient probablement d'améliorer les ensembles hydrauliques à Moulin-Vert.

10.3 Résultats à Justice

À la station Justice, nous rappelons que ce sont les pics de marée haute qui sont évalués. Aucune stratification n'est donc nécessaire. La figure 10.5 présente pour les ensembles obtenus par calibration statistique et les ensembles obtenus par assimilation de données les scores CRPS, RMSE, norme $||.||_2$ et aire AUC sous la courbe ROC. Pour mémoire, à Justice, le seuil retenu pour les courbes ROC est la hauteur confidentielle correspondant au passage de la vigilance verte à la vigilance jaune. La prévision déterministe est ici lancée avec persistance de la surcote mesurée au temps de base sur la période de prévision. En annexe S.3 sont présentés les histogrammes de rang des 10 ensembles aux échéances 6, 12, 18 et 24 heures. Nous rappelons que nous ne présentons ici que les résultats de l'ensemble brut R_R et des calibrations Q_R et QNE_{Md}, car nous avons montré au chapitre 8 que le forçage du modèle par l'hydrologie calibrée n'avait pas d'impact significatif sur les ensembles générés à Justice.

À Justice, nous notons comme à Kervir et Moulin-Vert que les scores présentés pour les ensembles générés par calibration statistique sont similaires à ceux obtenus sur la période de validation des forêts aléatoires (voir chapitre 8).

Pour les quatre expériences d'assimilation de données, les valeurs des CRPS dépendent de l'échéance calibrée de manière cyclique, un cycle correspondant à un cycle de marée (environ 12 heures). En effet, les CRPS sont faibles à très courte échéance, puis aux échéances entre +10h et +13h, puis à nouveau à partir de +22h. Ceci est lié à la correction cyclique des paramètres lors de la phase d'assimilation de données présentée au paragraphe 9.3.2.1. Ainsi, le modèle prévoit mieux les hauteurs d'eau à Justice lorsque la phase de la marée sur la fenêtre d'assimilation est la même que celle correspondant à l'échéance à prévoir. Une piste pour améliorer les résultats d'assimilation de données pourrait être de ne l'activer que pendant les phases de la marée correspondant aux pics de marée haute à Justice. Cependant, même lorsque l'on se trouve dans les conditions de prévision les plus favorables, on note que le CRPS des quatre ensembles d'assimilation de données sont toujours supérieurs à ceux obtenus avec la calibration statistique. Concernant les expériences d'assimilation de données, on note également que les trois expériences d'assimilation de données *DA1-persist*, *DA1* et *DA2* ont des performances comparables en termes de CRPS. L'expérience *DA3* présente des CRPS un peu plus faibles de manière

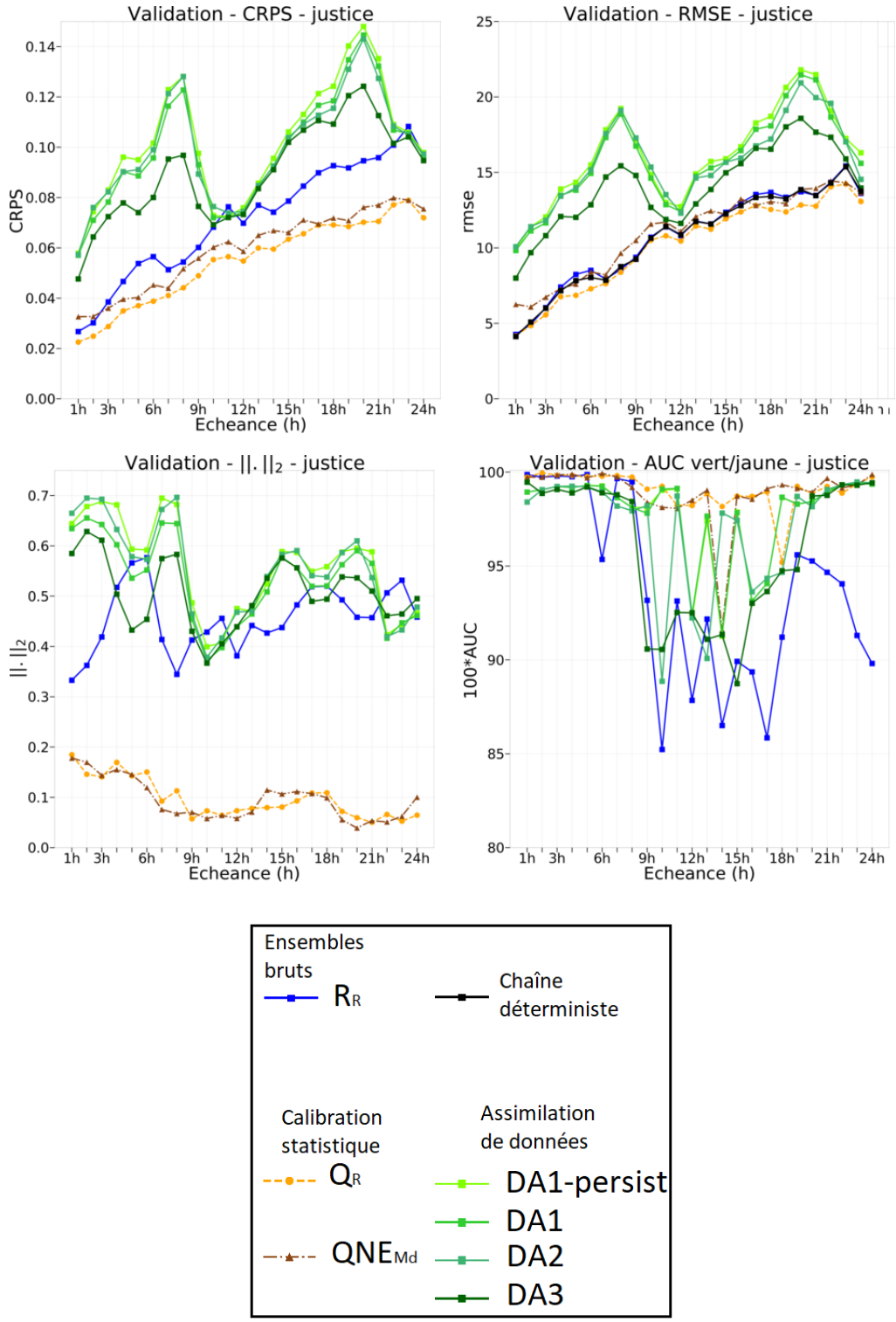


Figure 10.5: CRPS, RMSE, norme $||.||_2$2 et aire des courbes ROC pour le dépassement du seuil vert-jaune pour la prévision des pics de marée haute à Justice pour les 3 hivers confondus.

cyclique : les CRPS sont meilleurs surtout pour des échéances de prévision de 4 à 9 heures et de 19 à 21 heures.

En ce qui concerne la RMSE, les performances de l'ensemble brut R_R sont similaires à celles obtenues avec la simulation déterministe. Contrairement à la calibration à Kervir et Moulin-Vert, le fait d'intégrer les incertitudes de prévision n'a pas d'impact sur l'erreur obtenue avec l'ensemble R_R. Comme pour les CRPS, l'assimilation de données dégrade les erreurs de prévision par rapport à la prévision déterministe. Ce point avait déjà été mis en évidence lors de l'analyse des résultats sur la fenêtre d'assimilation (voir paragraphe 9.3.2.2). La calibration statistique génère des ensembles dont l'erreur est du même ordre de grandeur que la prévision déterministe. Ainsi, ni l'assimilation de données ni la calibration statistique ne permettent de diminuer la RMSE.

La calibration statistique s'avère de loin la plus efficace pour optimiser la fiabilité des ensembles. Les ensembles obtenus par calibration statistique présentent des normes d'histogrammes significativement inférieures à celles de l'ensemble brut. En revanche, l'assimilation de données a plutôt tendance à dégrader la fiabilité par rapport à l'ensemble brut.

Pour finir, la discrimination semble plus efficace avec la calibration statistique. Elle présente par ailleurs globalement des performances plus régulières selon les échéances que l'assimilation de données, surtout pour l'ensemble $NOTens@Mdet$ obtenu sans utilisation de modélisation hydraulique.

À titre d'illustration, la figure 10.6 présente les prévisions à échéance 12 heures des pics de marée haute obtenues avec les 10 ensembles construits dans le cadre de ce travail, pour les 2 hivers étudiés ici en validation. Les graphiques analogues aux échéances 6, 18 et 24 heures sont présentés en annexe S.3.

Ces courbes confirment que la prévision déterministe lancée avec persistance de la surcote mesurée au temps de base présente des résultats très proches à ceux de l'ensemble R_R moyenné. Ceci est cohérent avec l'analyse effectuée sur les RMSE. Elles mettent également en évidence que les prévisions moyennes issues des ensembles générés par assimilation de données présentent une plus grande dispersion que celles issues des ensembles générés par calibration statistique. Par ailleurs, si l'on compare les résultats des prévisions obtenues avec l'ensemble *DA2* par exemple aux échéances 6 heures et 12 heures, on note une surestimation des pics de marée haute à échéance 6 heures que l'on ne retrouve pas à échéance 12 heures. Ceci est cohérent avec les courbes obtenues pour les CRPS ou la RMSE aux échéances 6 et 12 heures.

Pour conclure, il semble préférable d'utiliser la calibration statistique à Justice pour prévoir les pics de marée haute. Avec l'ensemble QNE_{Md}, la calibration est faite sans faire tourner un modèle hydraulique au préalable. C'est relativement simple en opérationnel. Pour aller plus loin et envisager un passage vers l'opérationnel, il faudrait reprendre ce travail en utilisant des prévisions de surcote pour construire les prédicteurs pour la calibration statis-

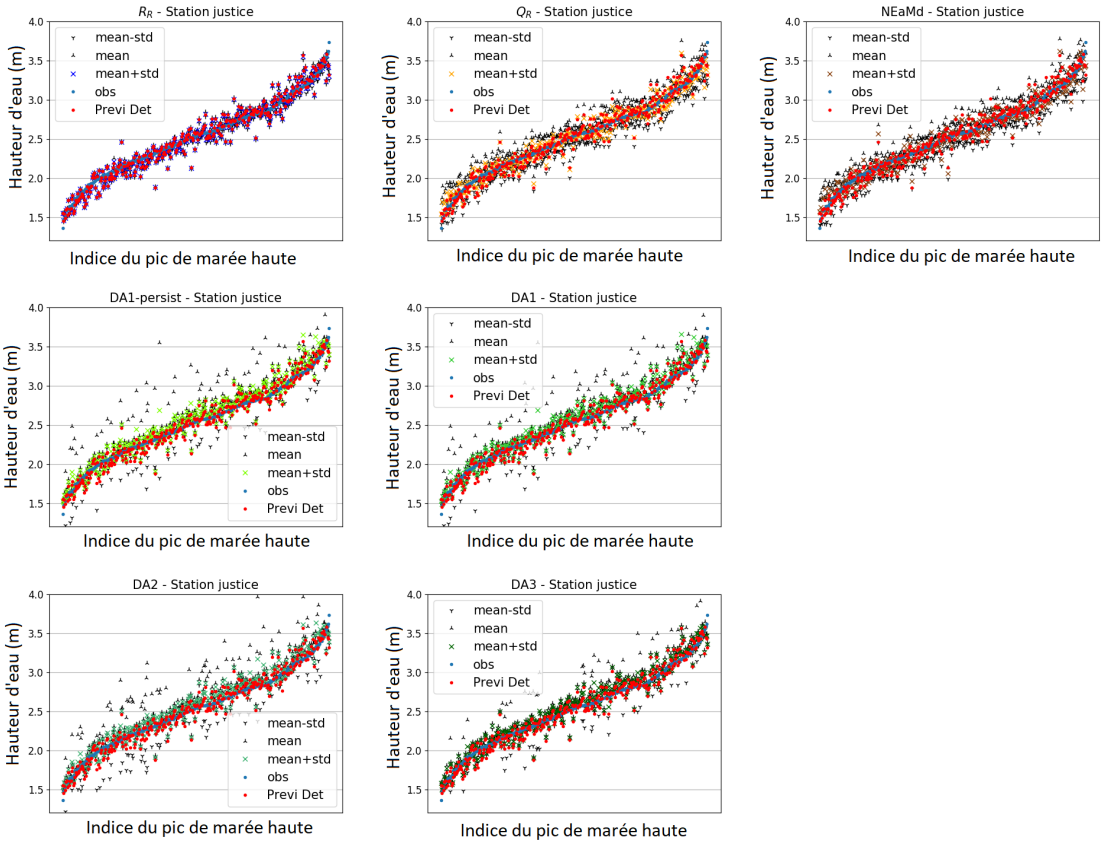


Figure 10.6: Moyenne, moyenne +/- écart-type pour les prévisions cyclées des pics de marée haute à échéance 12 heures à Justice. Pics de marée haute sur les 2 hivers de validation confondus classés par valeurs croissantes.

tique à Justice. En effet, nous avons montré au chapitre 8 que le prédicteur correspondant à la hauteur du pic à Plaisance est déterminant pour la calibration. Meilleur sera ce prédicteur, meilleurs seront les résultats de la calibration.

10.4 Synthèse

Dans ce chapitre, nous avons comparé les performances des prévisions d'ensemble hydrauliques construites dans le cadre de ce travail au niveau des trois stations de vigilance Kervir, Moulin-Vert et Justice. À Kervir et Moulin-Vert, nous nous intéressons aux prévisions horaires, tandis qu'à Justice, nous nous intéressons à la prévision des pics de marée haute.

Nous concluons tout d'abord sur les similitudes et les différences mises en évidence au niveau des 3 stations. Nous avons vu que pour les trois stations de vigilance, les performances des ensembles construits par calibration statistique sont analogues sur les 3 hivers considérés ici (environ 13 mois pour Kervir et Moulin-Vert, 8.5 mois à Justice) et sur la

période de validation des forêts aléatoires (31 mois à Kervir et Moulin-Vert, 16 mois à Justice). Les performances ne sont donc pas modifiées sur une période de validation plus courte. Pour les stations Kervir et Moulin-Vert, on note que l'ensemble hydraulique brut construit à partir des prévisions hydrologiques brutes R_R présente une erreur légèrement inférieure à la prévision déterministe. La prise en compte des erreurs de prévision hydrologiques et hydrauliques a permis d'améliorer à la marge les prévisions. Ceci n'est pas vrai pour la prévision à Justice. Enfin, pour Kervir et Moulin-Vert, l'ensemble créé par assimilation de données qui présente les meilleures performances est l'ensemble *DA2*. Ce n'est pas vrai non plus à Justice. Les différences entre d'une part Kervir et Moulin-Vert, et d'autre part Justice peuvent *a priori* être expliquée par un fonctionnement hydraulique différent à ces stations et au fait que ce sont les pics de marée haute qui sont calibrés à Justice.

Nous avons mis en évidence que la meilleure stratégie à retenir pour la réalisation de prévisions d'ensemble en opérationnel est différente pour chacune des stations. Nous présentons à présent les meilleures stratégies pour chacune de ces stations.

À Kervir, l'ensemble généré par assimilation de données *DA2* présente des caractéristiques significativement supérieures aux autres stratégies pour tous les scores présentés. La stratégie d'assimilation de données paraît donc la plus prometteuse, en ne corrigeant pas les conditions limites amont mais en utilisant les prévisions d'ensemble hydrologiques calibrées *QRF-hydro-TS* à l'amont.

À Moulin-Vert, les ensembles Q_Q et *DA2* présentent globalement les mêmes performances, sauf pour la fiabilité pour laquelle l'ensemble Q_Q est nettement supérieur. C'est donc l'ensemble obtenu par calibration statistique Q_Q qui semble être à privilégier à Moulin-Vert. Il est obtenu par calibration statistique de l'ensemble brut généré avec le modèle hydraulique MASCARET et alimenté par les prévisions d'ensemble hydrologiques calibrées *QRF-hydro-TS*. Nous retrouvons ici des conclusions analogues à ce qui a été mis en évidence pour les ensembles hydrologiques : l'assimilation de données produit des résultats analogues en termes de CRPS, RMSE et aire *AUC*, cependant, la calibration statistique produit des ensembles avec une meilleure fiabilité. Comme nous l'avons évoqué à propos de l'hydrologie, un travail spécifique sur les stratégies de réensemencement de l'ensemble à chaque cycle permettrait peut-être d'améliorer la fiabilité des ensembles obtenus pas assimilation de données. Par ailleurs, nous avons évoqué l'hypothèse que le comportement différent de l'assimilation de données à Kervir et à Moulin-Vert peut être lié au problème de manœuvre du seuil à l'aval de Moulin-Vert qui peut potentiellement dégrader l'assimilation de données. Grâce à des expériences menées en pluie parfaite, nous pouvons légitimement penser que la prise en compte de prévisions de pluie de meilleure qualité, telles que des prévisions d'ensemble calibrées, pourrait être une piste pour l'amélioration des prévisions d'ensemble hydrauliques.

À Justice, nous avons mis en évidence que l'assimilation de données générait des ensembles peu performants pour la prévision des pics de marée haute. La calibration statistique semble être une méthodologie bien plus intéressante. Cependant, les RMSE obtenues en moyennant les ensembles en prévision sont quasiment identiques pour les ensembles calibrés Q_R et QNE_{Md} que pour la prévision chaînée déterministe. Ainsi, le fait de construire des ensembles n'améliore pas la prévision. En revanche, l'intérêt de cette mise en place réside dans le fait d'accéder alors à une prévision probabiliste, et donc de pouvoir estimer des probabilités de dépassement de seuil. Par ailleurs, l'ensemble QNE_{Md}, qui présente des caractéristiques similaires à l'ensemble Q_R, est construit sans modélisation hydraulique. Pour l'obtenir, il suffit de construire les forêts aléatoires. D'un point de vue opérationnel, cela est intéressant puisqu'on peut se passer de la construction et du calage d'un modèle hydraulique.

Chapitre 11

Conclusion générale

Sommaire

11.1 Synthèse des résultats obtenus 309

11.1.1 Prévisions d'ensemble hydrologiques 309

11.1.2 Prévisions d'ensemble hydrauliques 311

11.2 Perspectives en termes de recherche 313

11.3 Perspectives d'application opérationnelle 315

11.3.1 Sur le bassin versant de l'Odet 315

11.3.2 Sur d'autres bassins versants 316

Le travail de thèse présenté dans ce manuscrit a été réalisé dans le cadre d'un partenariat entre le CEREMA[1], EDF R&D[2], le CERFACS[3] et le SCHAPI[4]. L'objectif est la mise en place et l'évaluation de prévisions d'ensemble hydrologiques et hydrauliques dans le cadre de la vigilance crues réalisée par les services de l'État afin de mieux appréhender et réduire les incertitudes dans un contexte de prévision à courte et moyenne échéance (24 heures). L'originalité de ce travail réside dans l'utilisation hybride de modèles à base physique et de modèles d'apprentissage sur un important volume de données. Dans cet objectif, les prévisions météorologiques forcent un modèle chaîné hydrologie-hydraulique afin de fournir des prévisions de débit et de hauteurs d'eau. Afin de prendre en compte les diverses sources d'incertitude liées aux modèles numériques, aux paramètres des modèles et aux données associées, l'approche déterministe est remplacée par une approche ensembliste ; on fournit ainsi un ensemble de prévisions de débits et hauteurs d'eau.

Le travail exploratoire a été mené sur le bassin versant de l'Odet dans le Finistère. La configuration de la prévision opérationnelle a été reprise. Une modélisation hydrologique couvre chacun des trois sous-bassins amont de l'Odet (modèles GRP ou MORDOR-TS), et une modélisation hydraulique couvre la partie aval du bassin jusqu'à l'exutoire maritime (modèle 1D MASCARET).

La même méthodologie a été suivie pour la génération des prévisions d'ensemble hydrologiques et hydrauliques. Dans un premier temps, des études de sensibilité globale des modèles hydrologiques GRP et MORDOR-TS des trois sous-bassins amont et du modèle hydraulique MASCARET du bassin aval sont réalisées. Cette étape est un préalable à la génération de prévisions d'ensemble. En effet, elle permet d'identifier les sources principales d'incertitude et ainsi de perturber les paramètres incertains significatifs pour la représentation des débits et des hauteurs d'eau prévus. La propagation de ces incertitudes avec les modèles MORDOR-TS et MASCARET aboutit à la création d'un ensemble brut pour l'hydrologie et pour l'hydraulique, les ensembles hydrologiques étant utilisés pour forcer les ensembles hydrauliques. Ensuite, deux méthodes de correction des ensembles sont investiguées : la calibration statistique de l'ensemble brut via la méthode des forêts aléatoires *Quantile Regression Forests* (QRF) et la calibration par assimilation de données via un filtre de Kalman d'ensemble (EnKF). Les ensembles générés dans le cadre de ce travail sont évalués en termes de fiabilité et de résolution. Les outils utilisés pour leur évaluation sont les histogrammes de rang, le CRPS et l'aire AUC sous les courbes ROC pour des événements de dépassement de seuil.

Le paragraphe 11.1 résume les résultats obtenus concernant la génération de prévisions d'ensemble hydrologiques et hydrauliques. Des perspectives de poursuite de ce travail

[1]Centre d'études et d'expertise sur les Risques, l'Environnement, la Mobilité et l'Aménagement
[2]Électricité de France, Recherche et Développement
[3]Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique
[4]Service Central d'Hydrométéorologie et d'Appui à la Prévision des Inondations

de recherche sont présentées au paragraphe 11.2. Pour finir, le paragraphe 11.3 présente des perspectives d'application pour la prévision opérationnelle réalisée par les Services de Prévision des Crues.

11.1 Synthèse des résultats obtenus

L'optimisation des forêts aléatoires pour la calibration des prévisions d'ensemble hydrologiques et hydrauliques brutes est réalisée sur une période de 3 ans, les mois d'été étant retirés. La comparaison des deux méthodologies de génération de prévisions d'ensemble est réalisée sur une période distincte de la période d'optimisation des forêts aléatoires afin d'évaluer les performances dans un contexte opérationnel. La période de validation a une durée de 7 ans pour les prévisions d'ensemble hydrologiques, les mois d'été étant toujours retirés. Pour les prévisions d'ensemble hydrauliques, la période de validation contient les événements pluvieux de trois hivers. Les performances des ensembles hydrologiques et hydrauliques sont analysées en réalisant une stratification. Cela permet de focaliser l'analyse des ensembles en période d'événements pluvieux. À la station Justice, la stratification n'est pas nécessaire puisqu'on étudie les performances de l'ensemble des pics de marée haute. Le paragraphe 11.1.1 résume les résultats obtenus pour la réalisation de prévisions d'ensemble hydrologiques tandis que le paragraphe 11.1.2 résume les résultats obtenus pour la réalisation de prévisions d'ensemble hydrauliques.

11.1.1 Prévisions d'ensemble hydrologiques

Des études de sensibilité globale ont été menées de manière comparée sur les deux modèles hydrologiques GRP (modèle global) et MORDOR-TS (modèle semi-distribué) implémentés sur les trois sous-bassins amont de l'Odet. Les sources d'incertitude prises en compte sont le forçage par la prise en compte des prévisions d'ensemble de précipitations PEAROME et les paramètres des modèles. Les prévisions PEAROME ont été préalablement calibrées par la méthode des forêts aléatoires pour améliorer leur fiabilité et leur résolution. Les résultats ont mis en évidence une grande similarité dans la réponse des deux modèles. La source d'incertitude prépondérante concerne le bilan en eau des modèles. Nous avons également mis en évidence une divergence des deux modèles en ce qui concerne le temps de réponse des sous-bassins versants. Une étude comparée entre des configurations tenant compte ou non de l'incertitude sur les pluies a été également menée avec le modèle MORDOR-TS pour valider la méthodologie de génération des ensembles bruts forcés avec une prévision déterministe. En effet, faute de disponibilité suffisante des prévisions PEAROME, le forçage utilisé pour la génération de l'ensemble brut hydrologique est les prévisions RR3 utilisées en opérationnel par les Services de Prévision des Crues. Pour le modèle MORDOR-TS, ces études de sensibilité ont ainsi mis en évidence deux paramètres

dont l'influence sur les débits simulés est négligeable. Les résultats des études de sensibilité globale permettent de valider la méthodologie utilisée pour la génération d'une prévision d'ensemble hydrologique brute avec le modèle MORDOR-TS forcé par une pluie déterministe.

Les prévisions d'ensemble hydrologiques brutes sont générées et calibrées statistiquement sur chacun des trois sous-bassins amont. La calibration est réalisée par échéance de prévision. Différentes approches sont comparées dans l'article *Strategies for hydrologic ensemble generation and calibration: on the merits of using model-based predictors* soumis au journal *Journal of Hydrology*. Dans l'approche *modèle uniquement*, des prévisions d'ensemble brutes *raw-HEF* sont générées par perturbation des 8 paramètres de MORDOR-TS significatifs pour les débits simulés selon les hypothèses de variabilité des paramètres utilisées dans l'étude de sensibilité globale. L'approche *modèle et données* consiste à calibrer statistiquement cet ensemble par la méthode QRF. Des prédicteurs issus de statistiques de l'ensemble brut sont utilisés. On obtient l'ensemble de quantiles *QRF-hydro*. L'approche *données uniquement* consiste en l'application d'une calibration QRF sans utiliser de prédicteur issu de statistiques de l'ensemble brut. Dans cette approche, aucun modèle hydrologique n'est nécessaire. Il s'agit en réalité d'un modèle d'apprentissage basé uniquement sur les données. On obtient l'ensemble de quantiles *QRF-nothydro*. Enfin, la méthode QRF étant appliquée par échéance de prévision, il est nécessaire de reconstruire des trajectoires réalistes pour une utilisation future comme forçage du modèle hydraulique MASCARET. L'application de la méthode de reconstruction *Ensemble Copula Coupling* (ECC) à l'ensemble *QRF-hydro* ainsi qu'un lissage des trajectoires *Trajectory Smoothing* permet de construire l'ensemble dénommé *QRF-hydro-TS*. Nous avons mis en évidence que les prévisions d'ensemble brutes *raw-HEF* sont sous-dispersives et sous-estiment les débits observés et que l'application de la calibration QRF permet d'obtenir des ensembles avec une meilleure fiabilité et des meilleurs CRPS sur les trois sous-bassins. Les meilleures performances sont obtenues avec l'approche *données et modèles* qui combine les résultats du modèle hydrologique avec des mesures disponibles avant le temps de base. Le lissage des trajectoires impacte légèrement les performances de l'ensemble *QRF-hydro* : l'ensemble *QRF-hydro-TS* présente une fiabilité légèrement moins bonne, mais des CRPS un peu meilleurs. Nous notons par ailleurs que l'étape de reconstruction n'est pas possible si l'on génère uniquement l'ensemble *QRF-nothydro*. En effet, la méthode ECC nécessite la disponibilité d'un motif qui sert de modèle à l'ensemble reconstruit.

La mise en place d'assimilation de données ensembliste grâce à un filtre de Kalman d'ensemble a pour objectifs la correction des paramètres du modèle MORDOR-TS. Ni les états du modèle ni ses paramètres ne sont corrigés. Ici aussi, on travaille sur chaque sous-bassin versant de manière indépendante. Les expériences jumelles réalisées visent à la correction soit d'un seul paramètre de MORDOR-TS, soit des 8 paramètres simultanément. Elles ont permis de valider l'algorithme d'assimilation de données mis en place. Les

expériences réelles sont aussi menées à l'échelle des sous-bassins, et le vecteur de contrôle est composé des 8 paramètres de MORDOR-TS. Les résultats en phase d'analyse montrent que les erreurs sur les débits simulés sont nettement diminuées. Dans l'espace du vecteur de contrôle, on note que selon les paramètres, les valeurs analysées peuvent être proches ou éloignées des valeurs de calage.

La comparaison de toutes les prévisions d'ensemble en prévision met en évidence que les deux stratégies, la calibration statistique et l'assimilation de données, permettent d'améliorer l'erreur RMSE par rapport à la prévision déterministe. Dans les configurations mises en place dans ce travail, la calibration statistique permet d'obtenir des ensembles avec des meilleures performances que l'assimilation de données. En effet, elle permet d'obtenir des CRPS meilleurs, de manière plus ou moins prononcée selon les stations, mais surtout elle permet d'obtenir une meilleure fiabilité et une meilleure discrimination des événements de dépassement de seuil. La suite de l'exploration de la chaîne de prévision est réalisée en prenant en compte l'ensemble brut *raw-HEF* et l'ensemble calibré *QRF-hydro-TS*.

11.1.2 Prévisions d'ensemble hydrauliques

L'étude de sensibilité globale du modèle MASCARET a été réalisée en considérant comme source d'incertitude les coefficients de frottement du modèle et les forçages en condition limite amont et aval. Elle a été menée de manière temporelle sur quatre types d'événements : fluvial, fluvio-maritime, maritime et pas d'événement particulier. La sensibilité du modèle aux différents paramètres d'entrée dépend du type d'événements considérés. Elle est également naturellement liée aux hypothèses faites concernant les plages de variation des paramètres du modèle MASCARET et l'amplitude de la perturbation des conditions aux limites considérée. Cette étude permet de mettre en évidence les paramètres influents pour la simulation des hauteurs d'eau aux trois stations de vigilance Kervir, Moulin-Vert et Justice situées dans l'agglomération quimpéroise. Ces résultats permettent de mettre en place la méthodologie pour la génération des prévisions d'ensemble brutes.

L'évaluation des performances des prévisions d'ensemble hydrauliques est différente selon les stations. L'ensemble des hauteurs d'eau horaires est évalué aux stations Kervir et Moulin-Vert, tandis que l'ensemble des pics de marée haute est évalué à Justice. Concernant la calibration statistique, ici aussi, nous testons des stratégies *données et modèle* et *modèle uniquement*. Après validation de l'algorithme d'assimilation de données EnKF grâce à des expériences jumelles, plusieurs expériences réelles sont testées dans un contexte opérationnel. Nous résumons ci-dessous les résultats obtenus pour les trois stations.

Kervir et Moulin-Vert Aux stations Kervir et Moulin-Vert, nous avons mis en évidence que les ensembles hydrauliques bruts présentent globalement de meilleures performances si l'hydrologie a été calibrée au préalable. La mise en place de calibration statistique permet d'améliorer les performances de ces ensembles bruts en périodes de calage et de validation. Cependant, on note que l'amélioration de la fiabilité est moins efficace sur la période de validation que sur la période de calage, surtout à Kervir. Nous avons également mis en évidence que les performances de l'ensemble hydrologique choisi pour générer l'ensemble hydraulique brut et/ou pour les prédicteurs de la calibration QRF impactent directement la fiabilité des ensembles hydrauliques construits (ensembles bruts ou calibrés). Pour ces stations, il n'est pas pertinent d'envisager une stratégie fondée uniquement sur l'apprentissage statistique pour générer des ensembles calibrés. L'utilisation de prédicteurs issus de statistiques de l'ensemble brut permet en effet d'améliorer la discrimination des événements de dépassement de seuil. La meilleure stratégie de calibration statistique est donc celle pour laquelle une calibration est effectuée à chaque étape de la chaîne de prévision, c'est-à-dire que l'on calibre les ensembles hydrologiques et hydrauliques. C'est l'ensemble appelé Q_Q dans le manuscrit. Le gain le plus important en termes de CRPS est lié à la calibration de l'hydrologie.

Les expériences réelles d'assimilation de données ont montré que sur la fenêtre d'assimilation, l'algorithme mis en place permet d'améliorer significativement les hauteurs d'eau modélisées aux stations Kervir et Moulin-Vert. En prévision, le meilleur ensemble créé par assimilation de données est *DA2*. Dans cette expérience, le modèle MASCARET est forcé par les mesures sur la fenêtre d'assimilation. Le vecteur de contrôle est constitué des coefficients de Strickler et du paramètre de correction additive de la condition limite aval b. Les cotes d'eau sont assimilées aux trois stations de vigilance. En prévision opérationnelle, le modèle est alimenté à l'amont par l'ensemble hydrologique calibré *QRF-hydro-TS*.

La comparaison des deux stratégies, calibration statistique et assimilation de données, ne permet pas de conclure sur une meilleure stratégie commune aux deux stations. À Kervir, l'ensemble *DA2* présente des caractéristiques significativement supérieures aux autres ensembles créés pour tous les scores présentés. À Moulin-Vert, les ensembles Q_Q et *DA2* présentent globalement les mêmes performances, sauf pour la fiabilité pour laquelle l'ensemble Q_Q est nettement supérieur. C'est donc l'ensemble obtenu par calibration statistique Q_Q qui semble être à privilégier à Moulin-Vert. Gardons à l'esprit que la modélisation à la station Moulin-Vert est délicate en raison du seuil situé quelques centaines de mètres à l'aval dont l'historique des manœuvres n'est pas connu. Par ailleurs, la modélisation avec MASCARET est délicate dans cette zone car les hypothèses de modélisation 1D ne sont pas vérifiées au niveau de ce seuil.

Justice À la station Justice, la calibration de l'ensemble hydrologique n'a pas d'impact significatif sur les performances des ensembles bruts et calibrés. En revanche, leur perfor-

mance est d'autant meilleure que la surcote prise en compte à l'aval est de bonne qualité. La calibration statistique améliore les performances des ensembles bruts pour tous les critères étudiés. Une calibration hydraulique en bout de chaîne est donc indispensable. Pour cette station, les calibrations réalisées sans utiliser de prédicteur issu de statistiques d'un ensemble brut présentent des performances comparables à celles réalisées avec un ensemble hydraulique brut. Il est donc pertinent de construire un modèle d'apprentissage sans modèle hydraulique à Justice. Le coût de développement est alors nettement réduit.

Les expériences réelles d'assimilation de données ont montré que sur la fenêtre d'assimilation, les performances obtenues ne dépassent jamais celles du modèle calé alimenté avec les mesures de débit à l'amont.

En prévision, l'assimilation de données génère des ensembles peu performants pour la prévision des pics de marée haute. La calibration statistique semble être une méthodologie bien plus intéressante, mais les RMSE obtenues en moyennant les ensembles en prévision sont quasiment identiques pour les ensembles calibrés et pour la prévision déterministe. Ce travail n'a donc pas permis d'améliorer la qualité de la prévision à Justice par rapport au chaînage déterministe des modèles hydrologiques et hydrauliques. Cependant, ce travail a permis de montrer qu'il est possible de générer des prévisions d'ensemble à Justice uniquement avec un modèle d'apprentissage statistique. Ce résultat est intéressant car cette mise en place est très légère pour une application opérationnelle. Par ailleurs, même à performance équivalente, une prévision d'ensemble apporte plus d'information qu'une prévision déterministe car elle permet d'appréhender l'incertitude de prévision. Des probabilités de dépassement de seuil peuvent être estimées.

11.2 Perspectives en termes de recherche

Des axes d'approfondissement ou de compléments sont envisageables suite à ce travail pour la génération de prévisions d'ensemble hydrologiques et hydrauliques.

Tout d'abord, comme nous l'avons expliqué, la chaîne de prévision d'ensemble ne tient pas compte des incertitudes concernant le forçage initial de la chaîne, à savoir les précipitations. Il serait donc pertinent de redéployer toute la méthodologie présentée en utilisant des prévisions d'ensemble météorologiques comme forçage initial. Les incertitudes de prévision seraient alors plus largement couvertes, et on pourrait espérer appréhender et réduire davantage les incertitudes le long de la chaîne. De la même manière, nous n'avons pas considéré de prévision de surcote à l'aval. Dans toutes les expériences opérationnelles, nous avons au mieux persisté la surcote mesurée au temps de base. Il serait pertinent de redéployer la chaîne de prévision hydraulique en prenant en compte des prévisions déterministes voire ensemblistes de surcotes à l'aval.

Par ailleurs, la démonstration n'a été réalisée ici que sur trois sous-bassins versants amont

de l'Odet pour les prévisions hydrologiques, et que sur trois stations de vigilance pour les prévisions hydrauliques. Afin d'apprécier la robustesse des résultats mis en évidence pour les prévisions d'ensemble hydrologiques, l'application de ces méthodologies à d'autres bassins versants paraît indispensable. Des éléments pourraient être adaptés pour les deux méthodologies présentées, comme les prédicteurs pour la calibration QRF ou encore la longueur de la fenêtre d'assimilation. En ce qui concerne les prévisions d'ensemble hydrauliques, nous avons déjà mis en évidence des comportements différenciés pour les trois stations de vigilance. Il est donc indispensable d'appliquer les méthodologies développées sur d'autres configurations.

Une perspective de complément concernant la génération de l'ensemble brut hydrologique et sa calibration statistique est de le générer en considérant l'ensemble issu du front de Pareto obtenu lors du calage du modèle MORDOR-TS. Ainsi, les modèles perturbés de MORDOR-TS présenteraient de meilleurs scores sur la période de calage qu'avec la méthodologie choisie ici. Cependant, nous avons montré que générer l'ensemble brut de cette façon n'est pas suffisant pour l'obtention de prévisions d'ensemble de bonne qualité. La calibration statistique reste alors indispensable.

Concernant l'assimilation de données dans le modèle hydrologique MORDOR-TS, il semble qu'il soit nécessaire d'approfondir les stratégies de réensemencement de l'ensemble à chaque cycle d'assimilation. Il est en effet possible qu'on puisse alors améliorer la fiabilité des ensembles en prévision. Enfin, une densification du réseau d'observations permettrait de lever en partie le problème d'équifinalité mis en évidence sur la fenêtre d'assimilation. Les prévisions pourraient alors être améliorées.

Il serait intéressant d'évaluer les performances d'un ensemble hydraulique à Kervir et Moulin-Vert qui serait construit en chaînant les prévisions hydrologiques calibrées et le modèle hydraulique déterministe. En effet, nous avons vu que le forçage est une des principales causes d'incertitude du modèle hydraulique. Cela permettrait d'évaluer l'importance de la perturbation des coefficients de Strickler pour la génération de l'ensemble.

Pour améliorer les résultats de l'assimilation de données à Justice, plusieurs pistes peuvent être envisagées. Tout d'abord, puisque l'objectif est la prévision des pics de marée haute et que nous avons mis en évidence que les performances de l'ensemble assimilé sont meilleures pour une échéance de prévision de 12 heures, il s'agirait de n'activer l'assimilation de données qu'autour des pics de marée haute. Pour améliorer les résultats de l'assimilation de données, on pourrait également envisager un paramétrage plus complexe de la correction de la surcote (ajout d'un coefficient de correction multiplicative et/ou d'un déphasage).

11.3 Perspectives d'application opérationnelle

11.3.1 Sur le bassin versant de l'Odet

En l'état actuel des résultats présentés, nous pouvons émettre les recommandations suivantes pour la mise en place d'un système de prévision d'ensemble sur le bassin de l'Odet.

Tout d'abord, il semble pertinent de réaliser la prévision d'ensemble hydrologique pour les trois stations amont Tréodet, Kerjean et Ty-Planche en générant un ensemble brut selon la méthodologie présentée, puis de lui appliquer une calibration statistique avec les forêts aléatoires construites dans le cadre de ce travail (ensemble *QRF-hydro-TS*). Les ressources informatiques nécessaires ne sont pas très importantes pour une application opérationnelle. Le coût de calcul principal de la méthode réside dans l'optimisation des forêts aléatoires et cette étape est réalisée au préalable. En opérationnel, le coût de calcul consistera essentiellement à lancer pour chaque temps de base 99 prévisions déterministes du modèle hydrologique pour chaque sous-bassin.

Pour la génération de prévisions d'ensemble à Kervir, il est conseillé de mettre en place la stratégie d'assimilation de données que nous avons appelée $DA2$[5]. Dans la configuration mise en place dans ce travail, pour chaque temps de base, cette stratégie nécessite deux fois l'intégration de 49 membres sur une fenêtre d'assimilation de 3 heures, puis l'intégration de ces 49 membres sur la fenêtre de prévision opérationnelle.

À Moulin-Vert, la meilleure stratégie est la génération d'un ensemble par calibration statistique selon la méthodologie Q_Q[6]. Il s'agit dans un premier temps de générer l'ensemble hydraulique brut R_Q alimenté par l'hydrologie calibrée *QRF-hydro-TS*. Pour cela, il faut générer 19 membres perturbés avec le modèle MASCARET, puis leur appliquer la calibration statistique. Nous avons vu que l'ensemble R_Q possède déjà de très bonnes performances, la calibration statistique permet essentiellement d'améliorer sa fiabilité et la discrimination pour le dépassement du seuil q_{99}.

À Justice, la mise en place d'un modèle d'apprentissage QRF paraît pertinent. Pour une application opérationnelle, il faudrait ré-optimiser une forêt aléatoire en prenant en compte dans les prédicteurs une prévision déterministe de surcote, ou ensembliste lorsque celle-ci sera disponible.

[5]Dans cette expérience, le modèle MASCARET est forcé par les mesures sur la fenêtre d'assimilation. Le vecteur de contrôle est constitué des coefficients de Strickler et du paramètre de correction additive de la condition limite aval b. Les cotes d'eau sont assimilées aux trois stations de vigilance. En prévision opérationnelle, le modèle est alimenté à l'amont par l'ensemble hydrologique calibré *QRF-hydro-TS*.

[6]La génération de cet ensemble hydraulique est réalisé par calibrations statistiques successives des ensembles bruts hydrologique et hydraulique, la prévision hydrologique calibrée servant de forçage à l'ensemble brut hydraulique.

11.3.2 Sur d'autres bassins versants

Dans l'optique d'une généralisation de la mise en place de prévisions d'ensemble hydrauliques sur d'autres bassins versants, il semble que la construction de prévisions hydrologiques de bonne qualité soit indispensable sauf dans le cas d'une influence maritime marquée. Ces influences respectives pourront être mises en évidence grâce à une étude de sensibilité globale telle que menée dans ce travail.

La bonne qualité des prévisions d'ensemble hydrologiques pourra être obtenue par calibration statistique, ou par assimilation de données dans l'hypothèse où l'on arrive à améliorer leur fiabilité grâce à différentes stratégies de réensemencement de l'ensemble à chaque cycle d'assimilation. La qualité des prévisions d'ensemble hydrologiques pourra potentiellement encore être améliorée avec la prise en compte de prévisions météorologiques de meilleure qualité, par exemple en utilisant des prévisions d'ensemble calibrées de précipitations.

En ce qui concerne la suite de la chaîne de prévision, nous n'avons pas mis en évidence de stratégie à privilégier pour la génération de prévisions d'ensemble hydrauliques pour des stations sous influence fluviale. Dans l'état actuel des connaissances, nous préconisons de développer en parallèle les deux méthodologies présentées afin d'analyser la meilleure stratégie pour chaque station de vigilance.

Pour une station sous influence maritime très marquée comme ici à Justice, nous avons mis en évidence que la stratégie est différente. La mise en place d'un modèle d'apprentissage statistique paraît de qualité comparable à un chaînage de modélisations hydrologiques et hydrauliques pour la prévision des pics de marée haute. Ce chaînage n'est donc a priori pas nécessaire et on pourra se concentrer sur l'application d'un modèle d'apprentissage type QRF uniquement.

Partie IV

ANNEXES

Annexe L

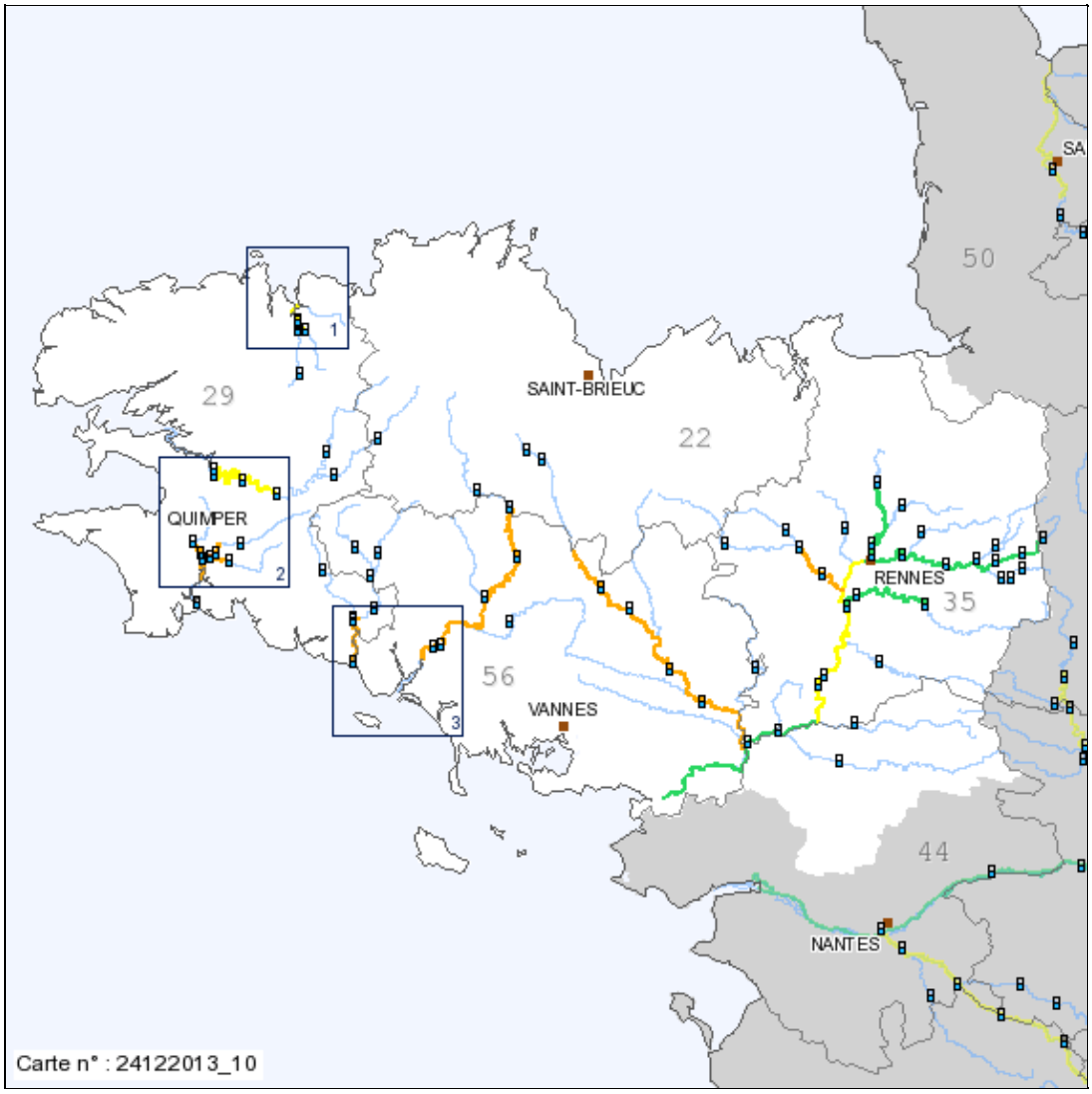
Compléments au chapitre 1

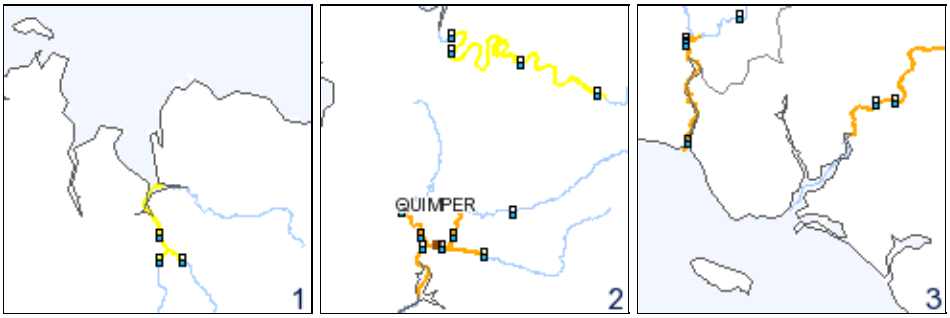
Sommaire

L.1 Bulletin de vigilance crues du SPC Vilaine et Côtiers Bretons du 23 décembre 2013 à 10h00 319

L.1 Bulletin de vigilance crues du SPC Vilaine et Côtiers Bretons du 23 décembre 2013 à 10h00

Bulletin du Service de Prévision des Crues Vilaine Côtiers Bretons





Rouge : Risque de crue majeure. Menace directe et généralisée de la sécurité des personnes et des biens.

Orange : Risque de crue génératrice de débordements importants susceptibles d'avoir un impact significatif sur la vie collective et la sécurité des biens et des personnes.

Jaune : Risque de crue ou de montée rapide des eaux n'entraînant pas de dommages significatifs, mais nécessitant une vigilance particulière dans le cas d'activités saisonnières et/ou exposées.

Vert : Pas de vigilance particulière requise.

BULLETIN D'INFORMATION
ORIGINE : Service de Prévision des Crues Vilaine Côtiers Bretons

Bulletin émis le : 24/12/2013 à 08:18
Prochain bulletin le : 24/12/2013 à 16:00

Etat maximal de vigilance sur le territoire du SPC : Orange

Département(s) en vigilance crue «orange» ou «rouge» :

Fait(s) nouveau(x) :
Passage en vigilance ORANGE des bassins du Blavet, Meu et Oust
Passage en vigilance JAUNE des bassins de Morlaix et de la Vilaine médiane

Qualification de la situation :
Vigilance ORANGE pour les bassins Laïta, Odet, Blavet, Meu et Oust
Vigilance JAUNE pour les bassins Morlaix, Aulne et Vilaine médiane

Situation actuelle et évolution prévue :
La tempête Dirk qui a balayé la Bretagne au cours de la nuit a été accompagnée de très importants cumuls de pluies. Les secteurs les plus arrosés se situent sur le Finistère sud et le centre Bretagne avec pontuellement près de 90mm en 24h. Le plus gros de l'épisode pluvieux aura quitté la région en cours de matinée laissant place à un régime d'averses.

Situation hydrologique par tronçon :

Nom	Vigilance	RSS
Morlaix	Jaune	☐
Aulne	Jaune	☐
Odet	Orange	☐
Laïta	Orange	☐
Blavet	Orange	☐
Vilaine aval	Vert	☐
Oust	Orange	☐
Vilaine médiane	Jaune	☐
Vilaine amont	Vert	☐
Meu	Orange	☐
Ille	Vert	☐
Seiche	Vert	☐

Tronçon : Morlaix

Les pluies enregistrées au cours du passage de la tempête Dirk ont été de 70mm en moyenne en 24h.
Les maximums ont été atteints aux stations en amont de Morlaix sur le Queffleuth (1,90m) et le Jarlot (1,80m) vers 07:00 ce mardi matin 24 décembre.
Les débordements actuellement en cours sur le centre ville de Morlaix perdureront jusqu'en milieu de journée.

Tronçon : Aulne

Les pluies enregistrées au cours du passage de la tempête Dirk ont été un peu plus importantes que prévues avec ponctuellement des valeurs proches de 80mm en 24h.
Cotes maximales prévues :
Pont Pol : 3,05 m +/− 20 cm mardi 24 décembre en cours d'après−midi
Pont Coblant : 3,95 m +/− 20 cm mardi 24 décembre en cours d'après−mid
Pont Routier : 2,15 m +/− 20 cm mardi 24 décembre au cours d'après−midi, début de soirée

---

Tronçon : Odet

Les pluies enregistrées au cours du passage de la tempête Dirk ont été de 60mm en moyenne sur le bassin en 24h.
Cotes maximales prévues :
Moulin Vert : 2,20 m +/− 15 cm mardi 24 décembre en fin de matinée
Kervir : 2,70 m +/− 15 cm mardi 24 décembre en milieu de journée

---

Tronçon : Laïta

Les pluies enregistrées au cours du passage de la tempête Dirk ont été un peu plus importantes que prévues avec ponctuellement des valeurs proches de 90mm en 24h.
Le niveau à la station Charles de Gaulle devrait atteindre 3,75 m +/− 15 cm mardi 24 décembre en matinée. La hausse devrait se poursuivre dans la journée pour une cote maximum de 3,95m +/−15cm en cours d'après−midi.

Tronçon : Blavet

Les pluies enregistrées au cours du passage de la tempête Dirk ont été de 60mm en moyenne en 24h
Cotes maximales prévues :
En matinée de ce mardi 24 décembre : 1,00m +/− 15 cm à Pontivy
En soirée de mardi 24 décembre : 1,60m +/− 15 cm à Languidic

Tronçon : Oust

Les pluies enregistrées au cours du passage de la tempête Dirk ont été de 55mm en moyenne en 24h.
Cotes prévues :
Cadoret : 3,85 m +/− 20 cm mardi 24 décembre en soirée, début de nuit du mardi au mercredi 25 décembre
Josselin : 1,10 m +/− 20 cm mardi 24 décembre en soirée, début de nuit du mardi au mercredi 25 décembre
Malestroit : 2,60 m +/− 20 cm dans la nuit du mardi au mercredi 25 décembre
Le Guélin : 7,40 m +/− 20 cm dans la nuit de mardi au mercredi 25 décembre

---

Tronçon : Vilaine médiane

Les pluies enregistrées au cours du passage de la tempête Dirk ont été de 40mm en moyenne en 24h sur les amonts du bassin de la Vilaine.
Cotes prévues :
Le Boel : 2,35m +/−20cm mercredi 25 décembre en milieu de journée
Guipry : 1,70m +/−20cm mercredi 25 décembre en fin de journée

---

Tronçon : Meu

Les pluies enregistrées au cours du passage de la tempête Dirk ont été de 50mm en 24h.
Cotes maximales prévues :
Monfort sur Meu : 2,65 m +/− 20 cm en fin de soirée du mardi 24 décembre
Bréal sous Monfort : 1,05 m +/− 20 cm en cours de nuit du mardi 24 décembre au mercredi 25 décembre.

Conséquences possibles :
Des inondations importantes sont possibles.
Les conditions de circulation peuvent être rendues difficiles sur l'ensemble du réseau et des perturbations peuvent affecter les transports ferroviaires.
Des coupures d'électricité peuvent se produire.
Les digues peuvent être fragilisées ou submergées.

Conseils de comportement :
Mettez−vous à l'abri.
Limitez tout déplacement sauf si absolument nécessaire et conformez−vous à la signalisation routière.
Tenez−vous informé de l'évolution de la situation (radio, etc...)
Veillez à la protection des biens susceptibles d'être inondés ou emportés (mobiliers, produits toxiques, appareil électriques, etc...).

Haut du bulletin

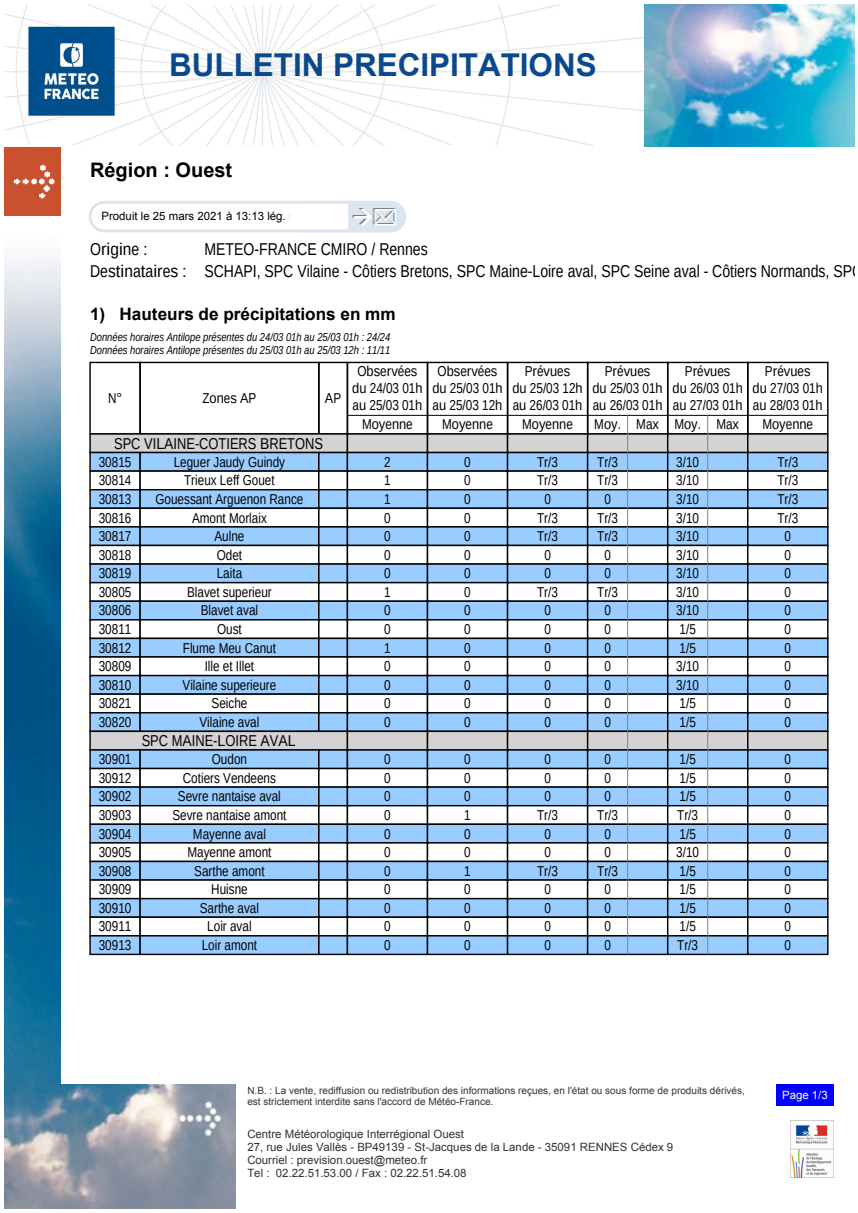
Annexe M

Compléments au chapitre 2

Sommaire

M.1 Bulletin de Précipitations émis par Météo-France 326
M.2 Description du modèle hydrologique GRP 327
M.3 Graphiques des débits mesurés et simulés avec GRP 329
M.4 Description du modèle hydrologique MORDOR-TS 333
M.5 Graphiques des débits mesurés et simulés avec MORDOR-TS . 338

M.1 Bulletin de Précipitations émis par Météo-France



METEO FRANCE

BULLETIN PRECIPITATIONS

Région : Ouest

Produit le 25 mars 2021 à 13:13 lég.

Origine : METEO-FRANCE CMIRO / Rennes

Destinataires : SCHAPI, SPC Vilaine - Côtiers Bretons, SPC Maine-Loire aval, SPC Seine aval - Côtiers Normands, SPC

1) Hauteurs de précipitations en mm

Données horaires Antilope présentes du 24/03 01h au 25/03 01h : 24/24
Données horaires Antilope présentes du 25/03 01h au 25/03 12h : 11/11

N°	Zones AP	AP	Observées du 24/03 01h au 25/03 01h	Observées du 25/03 01h au 25/03 12h	Prévues du 25/03 12h au 26/03 01h	Prévues du 25/03 01h au 26/03 01h		Prévues du 26/03 01h au 27/03 01h		Prévues du 27/03 01h au 28/03 01h
			Moyenne	Moyenne	Moyenne	Moy.	Max	Moy.	Max	Moyenne
SPC VILAINE-COTIERS BRETONS										
30815	Leguer Jaudy Guindy		2	0	Tr/3	Tr/3		3/10		Tr/3
30814	Trieux Leff Gouet		1	0	Tr/3	Tr/3		3/10		Tr/3
30813	Gouessant Arguenon Rance		1	0	0	0		3/10		Tr/3
30816	Amont Morlaix		0	0	Tr/3	Tr/3		3/10		Tr/3
30817	Aulne		0	0	Tr/3	Tr/3		3/10		0
30818	Odet		0	0	0	0		3/10		0
30819	Laita		0	0	0	0		3/10		0
30805	Blavet superieur		1	0	Tr/3	Tr/3		3/10		0
30806	Blavet aval		0	0	0	0		3/10		0
30811	Oust		0	0	0	0		1/5		0
30812	Flume Meu Canut		1	0	0	0		1/5		0
30809	Ille et Illet		0	0	0	0		3/10		0
30810	Vilaine superieure		0	0	0	0		3/10		0
30821	Seiche		0	0	0	0		1/5		0
30820	Vilaine aval		0	0	0	0		1/5		0
SPC MAINE-LOIRE AVAL										
30901	Oudon		0	0	0	0		1/5		0
30912	Cotiers Vendeens		0	0	0	0		1/5		0
30902	Sevre nantaise aval		0	0	0	0		1/5		0
30903	Sevre nantaise amont		0	1	Tr/3	Tr/3		Tr/3		0
30904	Mayenne aval		0	0	0	0		1/5		0
30905	Mayenne amont		0	0	0	0		3/10		0
30908	Sarthe amont		0	1	Tr/3	Tr/3		1/5		0
30909	Huisne		0	0	0	0		1/5		0
30910	Sarthe aval		0	0	0	0		1/5		0
30911	Loir aval		0	0	0	0		1/5		0
30913	Loir amont		0	0	0	0		Tr/3		0

N.B. : La vente, rediffusion ou redistribution des informations reçues, en l'état ou sous forme de produits dérivés, est strictement interdite sans l'accord de Météo-France.

Centre Météorologique Interrégional Ouest
27, rue Jules Vallès - BP49139 - St-Jacques de la Lande - 35091 RENNES Cédex 9
Courriel : prevision.ouest@meteo.fr
Tel : 02.22.51.53.00 / Fax : 02.22.51.54.08

Figure M.1: Bulletin de Précipitations (BP) émis le 25 mars 2021 à 13h00 par Météo-France.

M.2 Description du modèle hydrologique GRP

La figure M.2 rappelle le schéma fonctionnel du modèle hydrologique GRP.

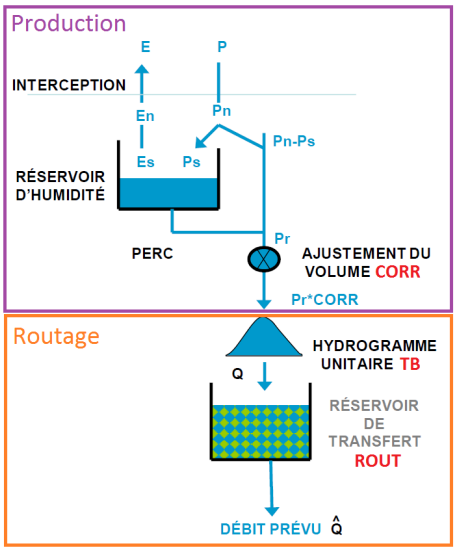


Figure M.2: Schéma fonctionnel du modèle hydrologique GRP sans assimilation de données - D'après IRSTEA IRSTEA [2017].

Les différentes étapes de calcul du modèle GRP sur un pas de temps t sont présentées ci-après.

Fonction de production

Interception de surface

Les précipitations P sont tout d'abord neutralisées par l'évapotranspiration E et forment ainsi la pluie nette P_n et le reliquat d'évapotranspiration potentielle E_n, selon les équations suivantes :

$$\begin{cases} Pn(t) = P(t) - E(t) & \text{et } En(t) = 0 & \text{si } P(t) <= E(t) \\ Pn(t) = 0 & \text{et } En(t) = E(t) - P(t) & \text{si } P(t) < E(t) \end{cases}$$

Réservoir de production

Si la pluie P n'a pas satisfait la demande évaporative E, la quantité Es est prélevée dans le réservoir de production et contribue ainsi à sa vidange selon l'équation :

$$\begin{cases} E_s(t) = \frac{s(t)(2-\frac{s(t)}{S_{max}})tanh(\frac{E_n(t)}{S_{max}})}{1+(1-\frac{s(t)}{S_{max}})tanh(\frac{E_n(t)}{S_{max}})} \\ P_s(t) = 0 \end{cases}$$

où s(t) représente le remplissage du réservoir S au pas de temps t. Dans le cas contraire, si la pluie P est supérieure à la demande évaporative E, le flux de pluie nette Pn est

partagé en deux : une partie va alimenter le réservoir de production dont la capacité S_{max} est fixée à 275 mm, l'autre partie va alimenter l'écoulement. Le partage du flux est réalisé selon l'équation :

$$\begin{cases} P_s(t) = \frac{S_{max}(1-(\frac{s(t)}{S_{max}})^2)tanh(\frac{P_n(t)}{S_{max}})}{1+\frac{s(t)}{S_{max}}tanh(\frac{P_n(t)}{S_{max}})} \\ E_s(t) = 0 \end{cases}$$

Le niveau du réservoir de production est alors actualisé :

$$s(t) = s(t) - E_s(t) + P_s(t)$$

Ce réservoir alimente l'écoulement par percolation, selon la formule :

$$Perc(t) = s(t)[1 - (1 + (\frac{s(t)}{5S_{max}})^4)]^{\frac{-1}{4}}$$

Le niveau du réservoir de production est à nouveau actualisé pour être à jour au pas de temps suivant :

$$s(t) = s(t) - Perc(t)$$

Ajustement du volume

La quantité d'eau totale allant alimenter l'écoulement est corrigée par le facteur d'ajustement de la pluie $Corr$. Cette correction vise à modifier le bilan en eau du modèle, et permet de prendre en compte d'éventuels échanges avec des nappes profondes ou des bassins voisins, et également de corriger la pluie qui entre dans le modèle qui est entachée d'incertitudes. Ce paramètre est déterminé lors du calage du modèle. On a ainsi :

$$\begin{cases} P_r(t) = P_n(t) - P_s(t) + Perc(t) \\ P'_r(t) = P_r(t) * Corr \end{cases}$$

Fonction de transfert

Hydrogramme unitaire

La pluie efficace P'_r est transférée au réservoir de routage par l'hydrogramme unitaire discrétisé au pas de temps horaire. Il permet de connaître à chaque pas de temps le pourcentage de volume de pluie transféré au réservoir de routage. Cet hydrogramme est caractérisé par le temps de base T_B. Ce paramètre est fixé lors de la phase de calage. Les coordonnées de l'hydrogramme unitaire sont données par

$$HU(k) = HS(k) - HS(k-1)$$

où la fonction HS représente l'hydrogramme unitaire cumulé (courbe en S) selon les équa-

tions suivantes :

$$\begin{cases} HS(k) = 0 & \text{si } k \leq 0 \\ HS(k) = \frac{k^1.25}{k^1.25+(T_B-k)^1.25} & \text{si } 0 \leq k \leq T_B \\ HS(k) = 1 & \text{si } k \geq T_B \end{cases}$$

Le débit transféré au réservoir de routage au pas de temps t est donné par le produit de convolution suivant :

$$q_{HU}(t) = \sum_{i=0}^{E(T_B)+1} P'_r(t-k+1)HU(k)$$

où $E(T_B)+1$ représente le nombre d'ordonnées non nulles de l'hydrogramme unitaire.

Réservoir de routage

Le réservoir de routage est alimenté par le flux issu de l'application de l'hydrogramme unitaire. Si on note $r(t)$ le niveau de remplissage du réservoir au pas de temps t, on a :

$$r(t) = r(t-1) + q_{HU}(t)$$

Le flux de vidange de ce réservoir au pas de temps t correspond au débit à l'exutoire du bassin versant au même pas de temps. Sa loi de vidange est caractérisée par la capacité *Rout* du réservoir dont la valeur est déterminée par le calage. La vidange est réalisée selon l'équation suivante :

$$Q(t) = \frac{r(t)^2}{r(t) + Rout}$$

où $r(t)$ représente le remplissage du réservoir au pas de temps t.

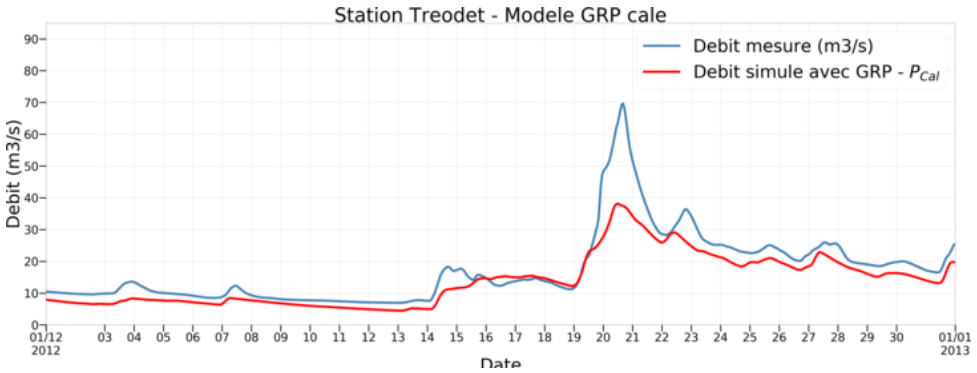
Assimilation de données

Le modèle opérationnel GRP exploite l'information fournie par le débit mesuré à l'instant de prévision. Elle se déroule en deux temps :

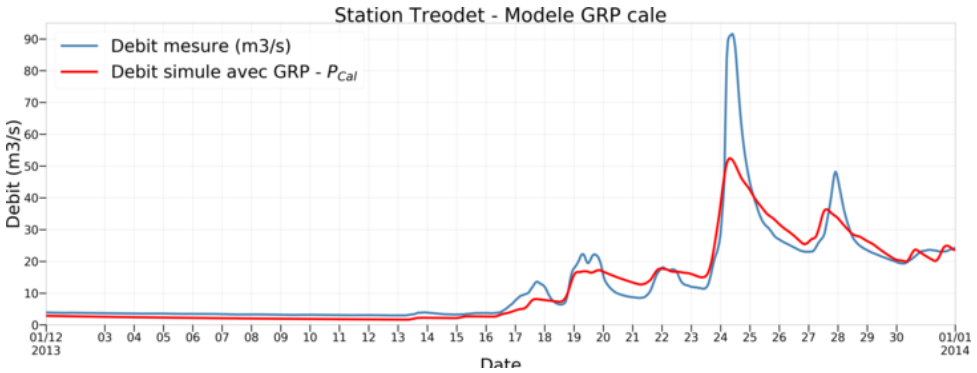
- une mise à jour du remplissage du réservoir de routage du modèle r(t) ;
- une exploitation de l'erreur du modèle faite au pas de temps précédent pour le pas de temps courant.

Nous rappelons que cette partie du code est inactive dans le travail présenté.

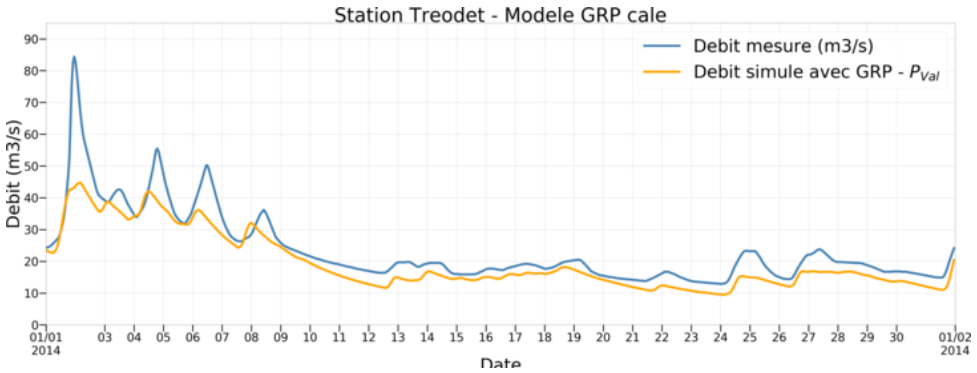
M.3 Graphiques des débits mesurés et simulés avec GRP



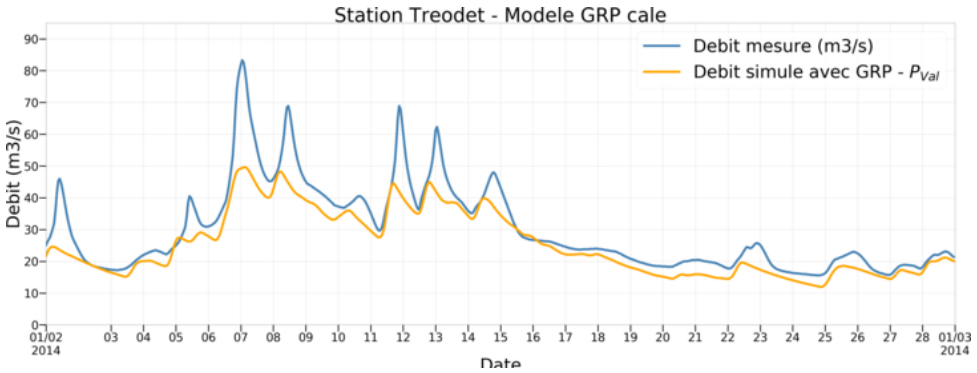
(a) Décembre 2012 - Période de calage



(b) Décembre 2013 - Période de calage

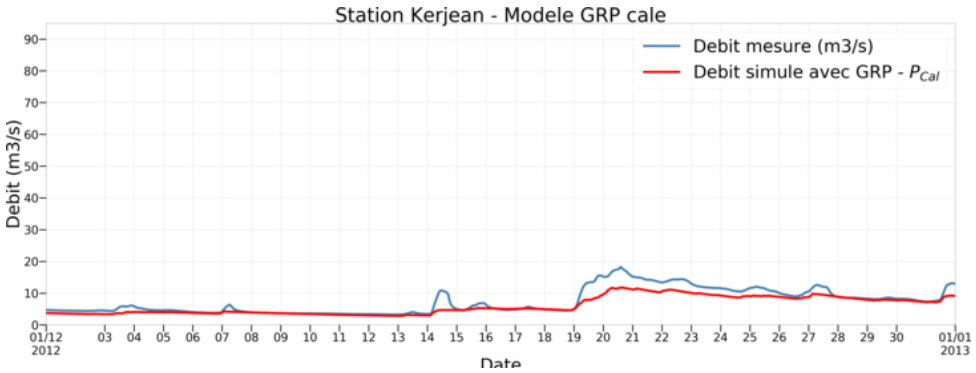


(c) Janvier 2014 - Période de validation

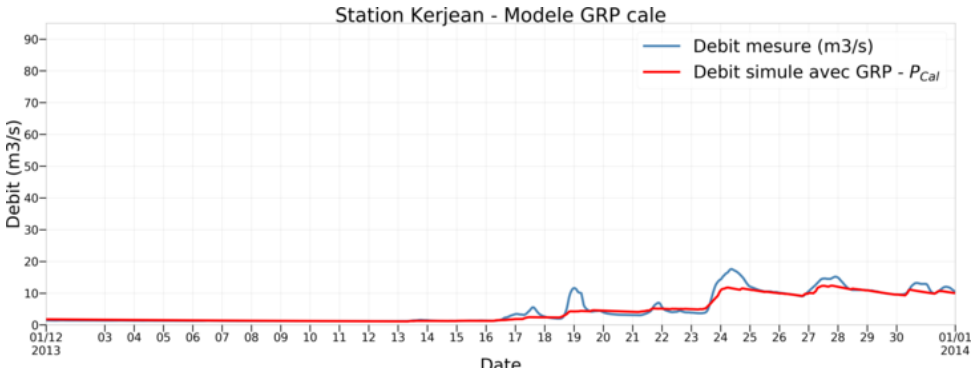


(d) Février 2014 - Période de validation

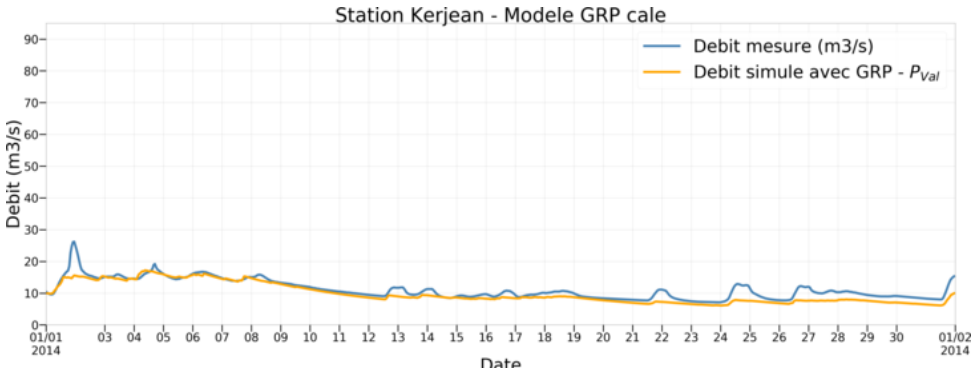
Figure M.3: Débits mesurés et simulés avec le modèle GRP à la station Tréodet.



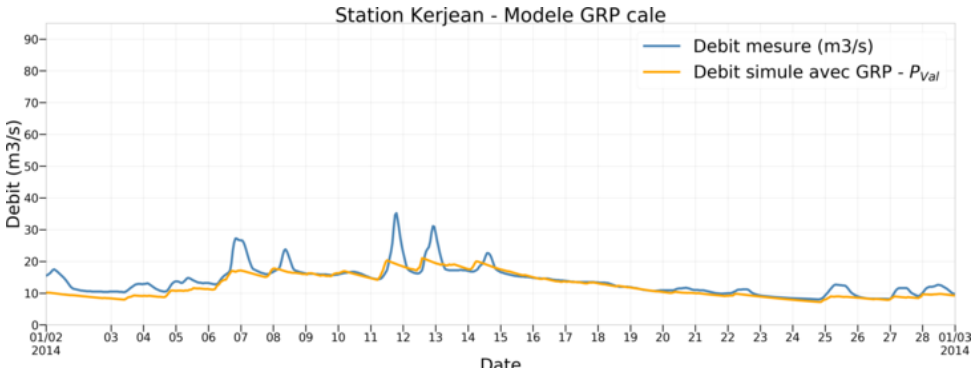
(a) Décembre 2012 - Période de calage



(b) Décembre 2013 - Période de calage

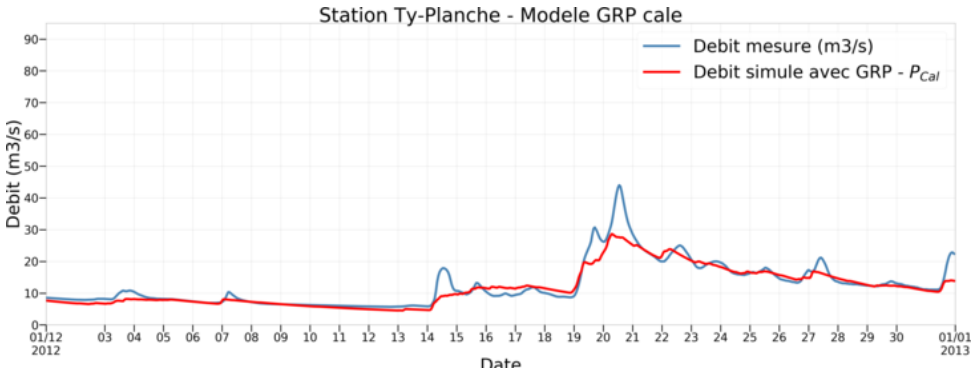


(c) Janvier 2014 - Période de validation

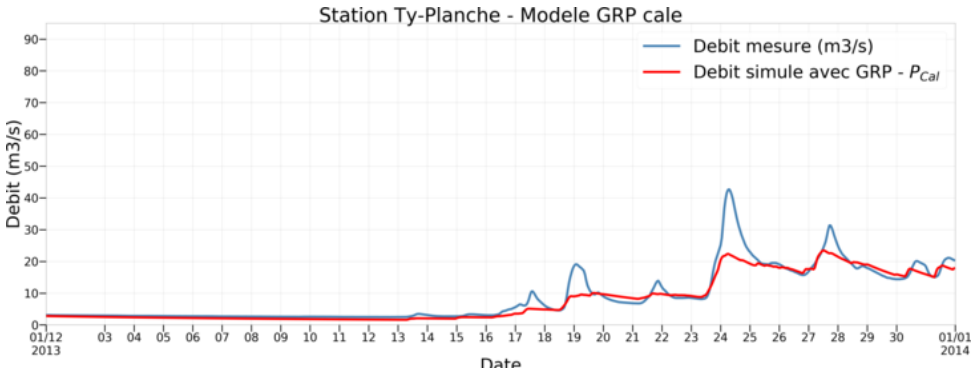


(d) Février 2014 - Période de validation

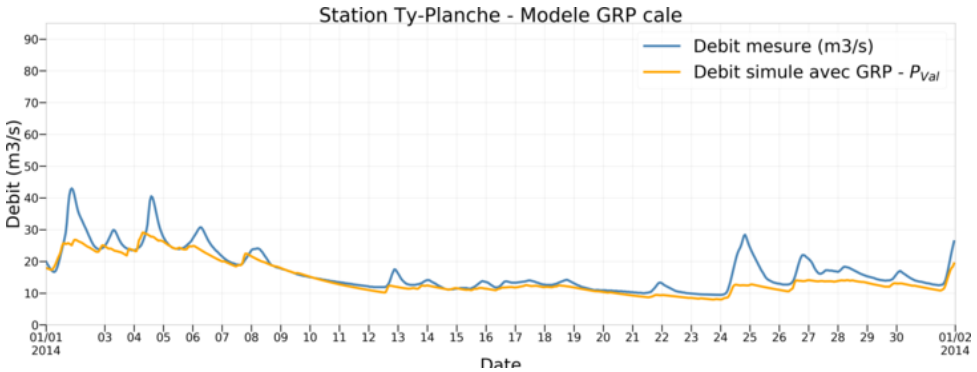
Figure M.4: Débits mesurés et simulés avec le modèle GRP à la station Kerjean.



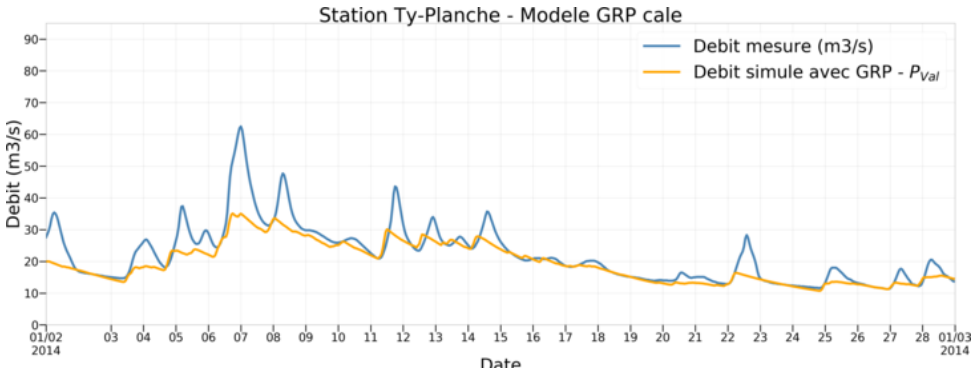
(a) Décembre 2012 - Période de calage



(b) Décembre 2013 - Période de calage



(c) Janvier 2014 - Période de validation



(d) Février 2014 - Période de validation

Figure M.5: Débits mesurés et simulés avec le modèle GRP à la station Ty-Planche.

M.4 Description du modèle hydrologique MORDOR-TS

Description du modèle

La figure M.6 rappelle le schéma fonctionnel du modèle hydrologique MORDOR-TS.

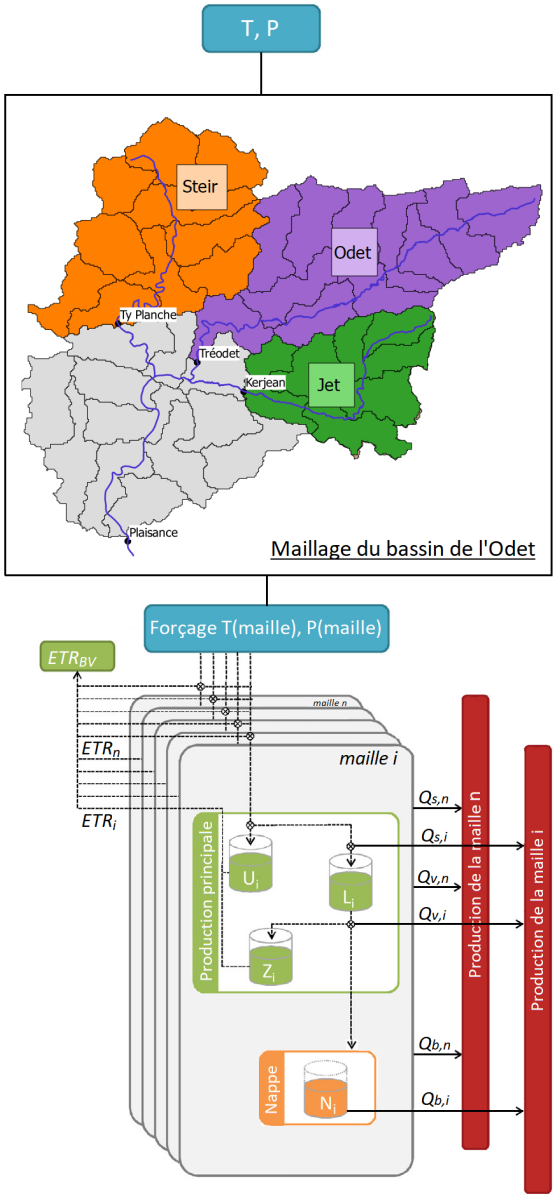


Figure M.6: Schéma fonctionnel du module de production du modèle hydrologique MORDOR-TS sur le bassin de l'Odet avec une seule bande d'altitude - D'après Rouhier et al. [2017].

À partir des forçages horaires de pluie et de température, le modèle fait évoluer dans chaque maille i les états des stocks U_i, L_i, Z_i et N_i. Ces réservoirs interagissent et alimentent l'écoulement à l'exutoire de la maille. Ce débit est ensuite transféré à l'exutoire du bassin

versant par une fonction de transfert. Pour plus de lisibilité, l'indice i relatif à la maille sera omis dans la suite de cette annexe.

Fonction de production

Le module neige n'est pas utilisé dans l'étude, car il n'y a pas d'événement neigeux sur le bassin de l'Odet susceptible d'avoir des conséquences hydrologiques. La figure M.7 présente les quatre réservoirs du module de production utilisés ainsi que les entrées et sorties de ces réservoirs.

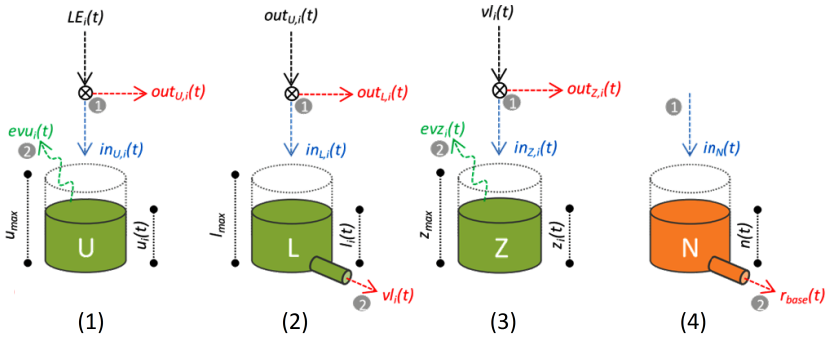


Figure M.7: Représentation schématique des 4 réservoirs U, L, Z et N de MORDOR-TS - D'après Garavaglia et al. [2017]

Correction des précipitation, évapotranspiration et interception de surface

Les forçages imposés à l'entrée du modèle ne sont pas parfaits. En ce qui concerne les précipitations, leur estimation a posteriori reste un exercice délicat. La pluie entrant dans le modèle connaît donc des incertitudes. Afin de pallier ce problème, le modèle corrige la pluie par le paramètre c_p, dont la valeur est déterminée lors du calage.

Le modèle utilise la notion d'évapotranspiration maximale e_{max}. Cette notion intègre l'évapotranspiration potentielle e_p classiquement utilisée (évapotranspiration maximale qui se produirait en cas d'approvisionnement en eau suffisant de la surface, valeur calculée principalement à partir de la température), et les caractéristiques de la végétation et son stade de développement. En effet, à conditions météorologiques équivalentes, la reprise évaporative sera différente entre le mois de mars et le mois de septembre car la végétation n'est pas au même stade de développement. Elle est également différente pour une forêt ou une garrigue. On introduit pour cela le coefficient cultural K_c, tel que:

$$e_{max}(t) = K_c(t) * e_p(t)$$

K_c évolue en fonction du rayonnement solaire extra-terrestre R_{pot} ($W.m^{-2}.jour^{-1}$) selon la

formule suivante :

$$K_c(t) = k_{min} + (1 - k_{min}) \frac{R_{pot} - min(R_{pot})}{max(R_{pot}) - min(R_{pot})}$$

Le paramètre k_{min} décrit l'amplitude de la variation saisonnière de K_c. Si k_{min} vaut 1, K_c vaut 1 et il n'y a pas de variation saisonnière de K_c. L'évapotranspiration maximale est égale à l'évapotranspiration potentielle. Si k_{min} est inférieur à 1, la variation saisonnière de K_c est proportionnelle au rayonnement. Si k_{min} est supérieur à 1, la variation saisonnière de K_c est inversement proportionnelle au rayonnement (cas de cultures avec plusieurs cycles annuels, ou déphasés). À noter que la valeur moyenne de K_c fixée à 1 pour limiter son impact sur le cumul évaporatif. K_c régit uniquement la saisonnalité de l'évapotranspiration. Le coefficient cultural est ainsi mis à jour à chaque pas de temps selon la formule

$$K_c(t) = \frac{K_c(t)}{\bar{K_c}}$$

Le paramètre k_{min} est fixé lors du calage. Les précipitations sont interceptées par le couvert végétal et réévaporés directement dans l'atmosphère : c'est l'interception.

Les deux paramètres c_p et k_{min} contrôlent ainsi la quantité d'eau participant à la production du modèle MORDOR-TS.

Le réservoir U

Le réservoir de surface U représente la capacité d'absorption des couches superficielles du sol. Sa capacité U_{max} est déterminée lors du calage. Son remplissage varie entre *0* et U_{max}. Ce réservoir voit en entrée la pluie après correction par le coefficient c_p. C'est la lame d'eau LE. Une partie out_U est directement refusée par le réservoir est dirigée vers le réservoir intermédiaire L. La saturation de ce réservoir est progressive. Le refus à l'infiltration varie linéairement avec le remplissage du réservoir. Plus il est rempli, plus le refus à l'infiltration est important. Si le réservoir est plein, toute la lame d'eau sera refusée. Le complément in_U entre dans le réservoir U. Les équations qui régissent ce réservoir sont les suivantes :

$$\begin{aligned} out_U(t) &= LE(t) - (U_{max} - u(t-1))(1 - exp(-\tfrac{LE(t)}{U_{max}})) \\ in_U(t) &= LE(t) - out_U(t) \end{aligned}$$

où $u(t)$ représente le remplissage du réservoir U au pas de temps t.

Si la demande évaporative n'a pas complètement été satisfaite lors de l'interception de surface, un complément est prélevé dans ce réservoir, de façon proportionnelle à son taux de remplissage :

$$ev_U(t) = [e_{max} - ev_0] \frac{u(t)}{U_{max}}$$

où e_{max} représente l'évapotranspiration maximale et ev_0 la fraction de la pluie évaporée

lors de l'interception de surface.

Le réservoir L

Le réservoir intermédiaire L représente les nappes de versant. Ce sont des nappes de faible profondeur, localisées directement dans le versant. Ce réservoir est caractérisé par sa capacité L_{max} et son exposant de vidage evl qui détermine sa loi de vidange. Ces deux paramètres sont déterminés lors du calage. Ce réservoir voit en entrée le refus du réservoir U out_U. Ce flux est décomposé en deux flux : une partie in_L entre dans le réservoir L. Le complément out_L n'entre pas dans le réservoir et va alimenter l'écoulement : c'est le ruissellement de surface r_{surf}. Les équations qui régissent ce réservoir sont les suivantes :

$$in_L(t) = out_U(t)[1 - (\frac{l(t-1)}{L_{max}})^2]$$
$$out_L(t) = out_U(t) - in_L(t)$$

Si le réservoir L se sature avant absorption complète de in_L, l'excédent est ajouté au flux out_L et alimente directement le ruissellement de surface r_{surf}. Ainsi, on a

$$\boxed{r_{surf}(t) = out_L(t) + max(0, in_L(t) - L_{max})}$$

La vidange du réservoir intermédiaire L est réalisée après le remplissage du réservoir selon la loi suivante :

$$v_L(t) = \frac{1}{evl * {L_{max}}^{evl-1}} * l(t)^{evl}$$

où l(t) représente le remplissage du réservoir L au pas de temps t.

Les paramètres L_{max} et evl définissent respectivement la capacité et la dynamique de ce réservoir.

Le réservoir Z

Si la teneur en eau dans le sol est en-dessous d'un seuil, l'eau ne s'y écoule plus gravitairement. Elle est alors uniquement disponible pour la transpiration des plantes. Le réservoir Z représente ce phénomène de piégeage capillaire du sol. Sa capacité Z_{max} est déterminée lors du calage. Ce réservoir Z voit en entrée la vidange de L v_L. Ce flux est décomposé en deux flux : une partie in_Z entre dans le réservoir Z. Le complément out_Z qui n'entre pas servira à alimenter l'écoulement de surface et les nappes profondes (voir partage des écoulements). Si le réservoir Z se sature avant absorption complète de v_L, l'excédent est ajouté au flux out_Z. Les équations qui régissent ce réservoir sont les suivantes :

$$in_Z(t) = v_L(t)(1 - \frac{z(t)}{Z_{max}})$$
$$out_Z(t) = v_L(t) - in_Z(t) + max(0, in_Z(t) - Z_{max})$$

Si la demande évaporative n'a pas complètement été satisfaite lors de l'interception de

surface, ni par le prélèvement dans le réservoir U, un complément est prélevé dans ce réservoir, de façon proportionnelle à son taux de remplissage :

$$ev_Z(t) = [e_{max} - ev_0 - ev_U]\frac{z(t)}{Z_{max}}$$

où e_{max} représente l'évapotranspiration maximale, ev_0 la fraction de la pluie évaporée lors de l'interception de surface et ev_u la fraction évaporée dans le réservoir U.

Partage des écoulements

L'eau non piégée par le réservoir Z out_Z est répartie entre un ruissellement retardé de versant r_{vers}, seconde composante du ruissellement, et l'alimentation d'un réservoir profond appelé N. Cette répartition est réalisée à l'aide d'un coefficient k_r, qui est déterminé lors de la phase de calage. La seconde composante du ruissellement est donc régit par l'équation suivante :

$$\boxed{r_{vers}(t) = k_r * out_Z(t)}$$
$$in_N(t) = (1 - k_r) * out_Z(t)$$

Le réservoir profond N

Le réservoir profond N conceptualise les nappes souterraines. Sa capacité est infinie. Sa vidange est conditionnée par le paramètre lkn. Elle représente le ruissellement de base r_{base}, la troisième et dernière composante du ruissellement. La valeur de lkn est déterminée lors du calage du modèle. Le calcul de la vidange du réservoir profond N se fait selon la loi suivante :

$$\boxed{r_{base}(t) = n(t) - (10^{lkn} * (evn - 1) + N(t)^{1-evn})^{\frac{1}{1-evn}}}$$

où n(t) représente le remplissage du réservoir N au pas de temps t.

Fonction de transfert

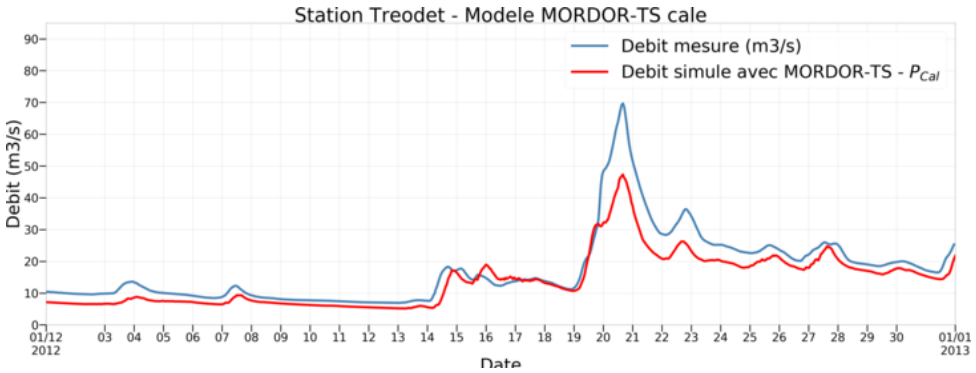
Pour chaque maille, les trois contributions à l'écoulement décrites précédemment r_{surf}, r_{vers} et r_{base} s'additionnent pour former l'écoulement total.

$$r_{total}(t) = r_{surf}(t) + r_{vers}(t) + r_{base}(t)$$

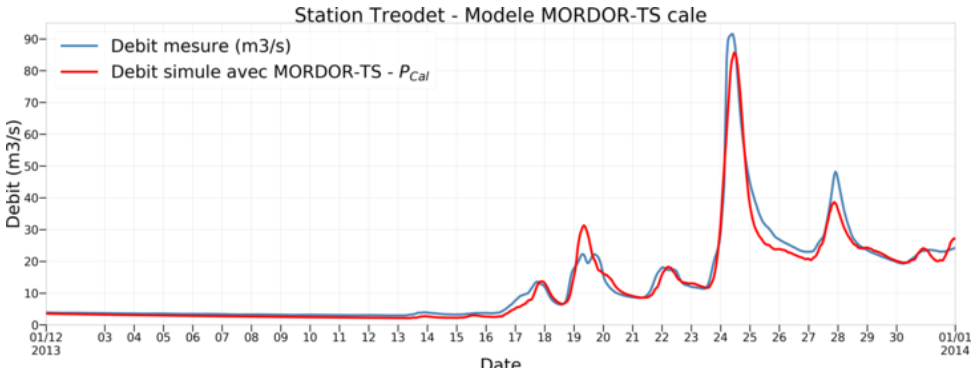
La fonction de transfert du modèle a pour rôle de transférer les contributions de chaque maille aux points d'intérêt du réseau hydrographique du bassin versant pour le calcul du débit. Chaque maille i est considérée équivalente à un bief décrit par deux distances [Monteil and Le Moine, 2014] : la distance hydraulique séparant l'exutoire de la maille (point aval du bief équivalent) et l'exutoire du domaine, et la longueur du bief équivalent. Sur chaque pas de temps de calcul, on suppose que la production de la maille i est apportée de manière uniforme sur toute la longueur du bief équivalent. Le routage intra-mailles

et inter-mailles est effectué à partir du modèle de l'onde diffusante, avec la célérité et la diffusivité indépendants du débit [Hayami, 1951]. On suppose que la célérité et la diffusivité sont les mêmes sur les versants et dans les cours d'eau. La linéarité de la formulation permet d'appliquer le principe de superposition. Les contributions de chaque maille sont propagées vers les différents points de contrôle de manière indépendante. Pour réduire le coût de calcul, le routage n'est en pratique effectué que pour les mailles du bassin dont l'exutoire est un point de contrôle.

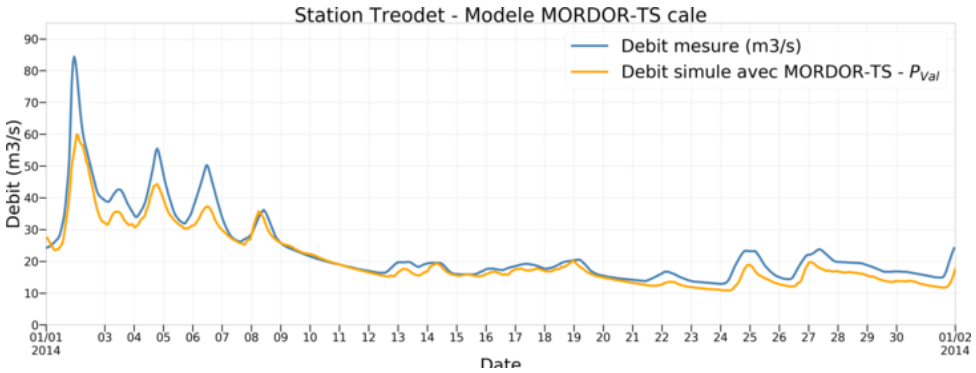
M.5 Graphiques des débits mesurés et simulés avec MORDOR-TS



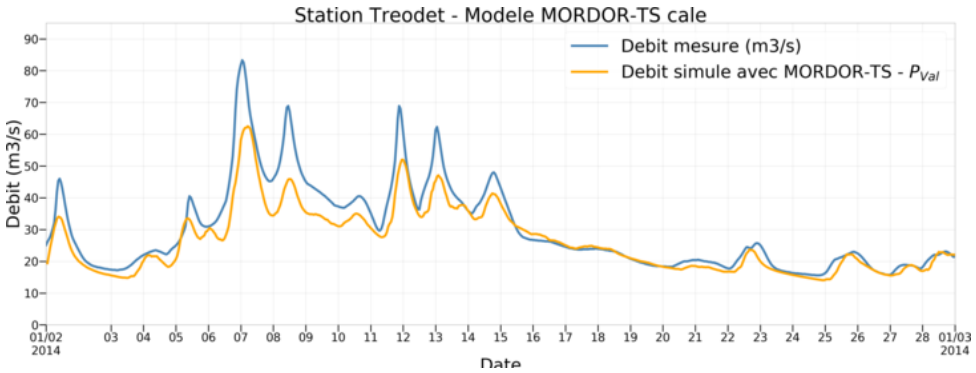
(a) Décembre 2012 - Période de calage



(b) Décembre 2013 - Période de calage

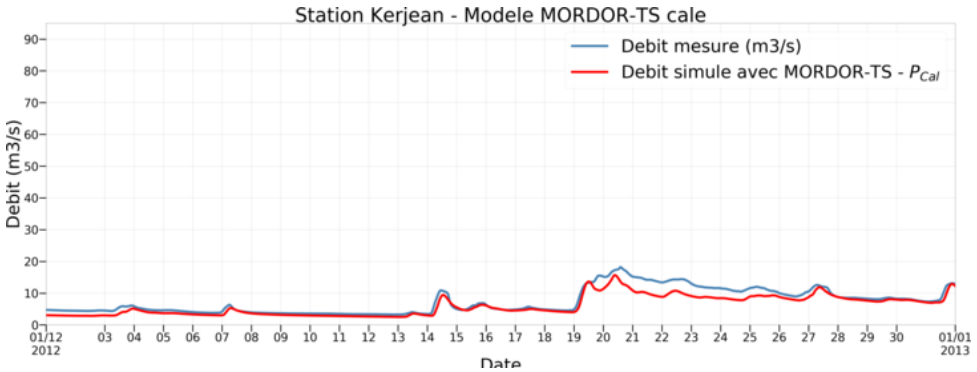


(c) Janvier 2014 - Période de validation

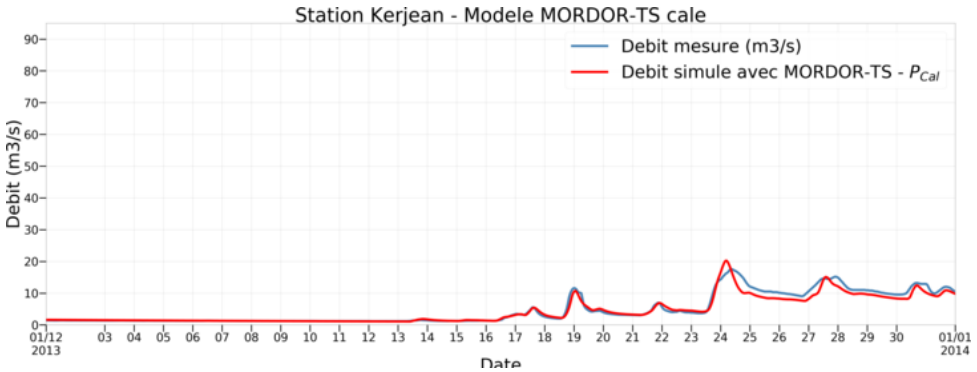


(d) Février 2014 - Période de validation

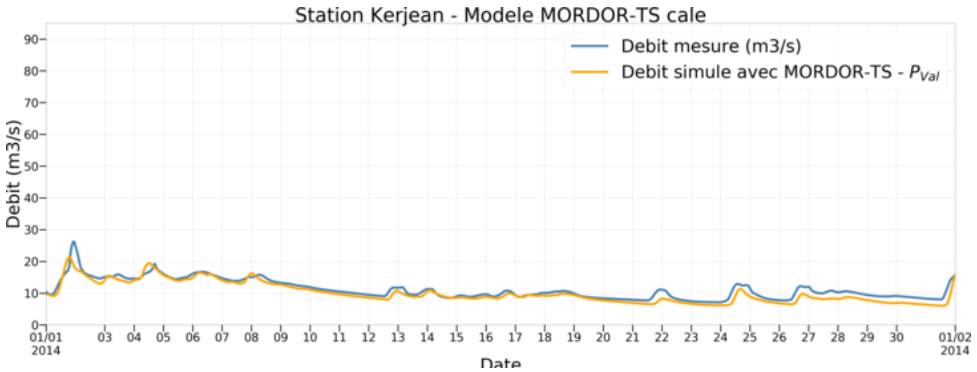
Figure M.8: Débits mesurés et simulés avec le modèle MORDOR-TS à la station Tréodet.



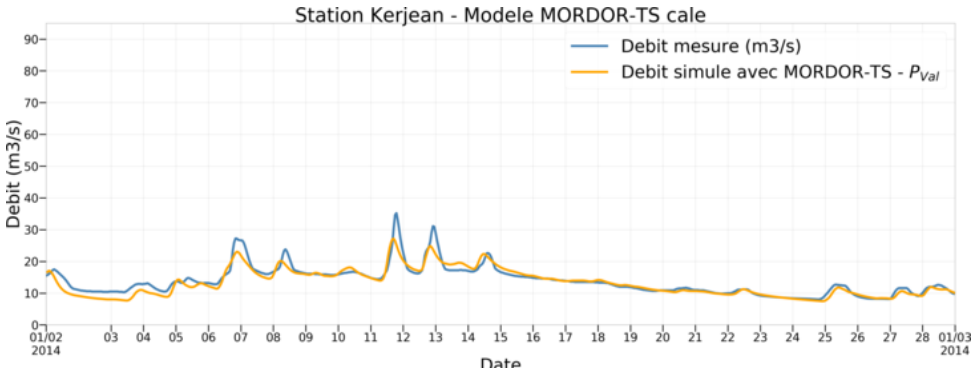
(a) Décembre 2012 - Période de calage



(b) Décembre 2013 - Période de calage

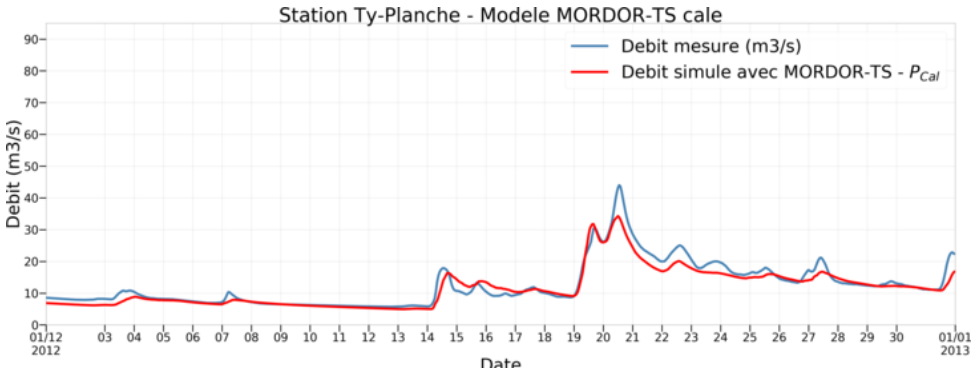


(c) Janvier 2014 - Période de validation

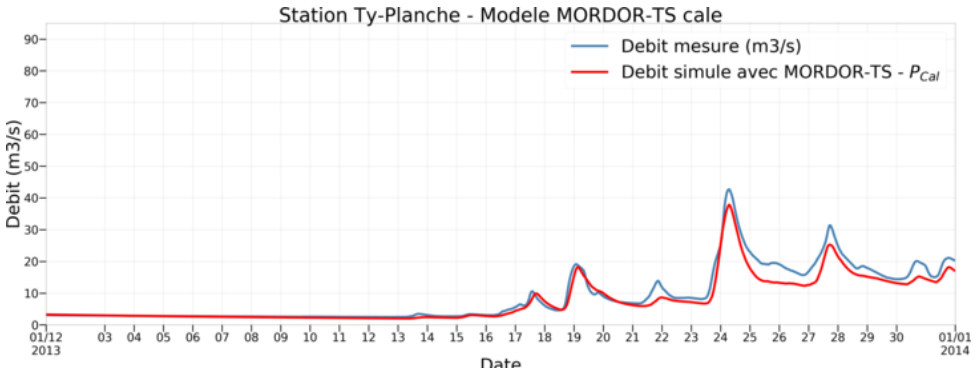


(d) Février 2014 - Période de validation

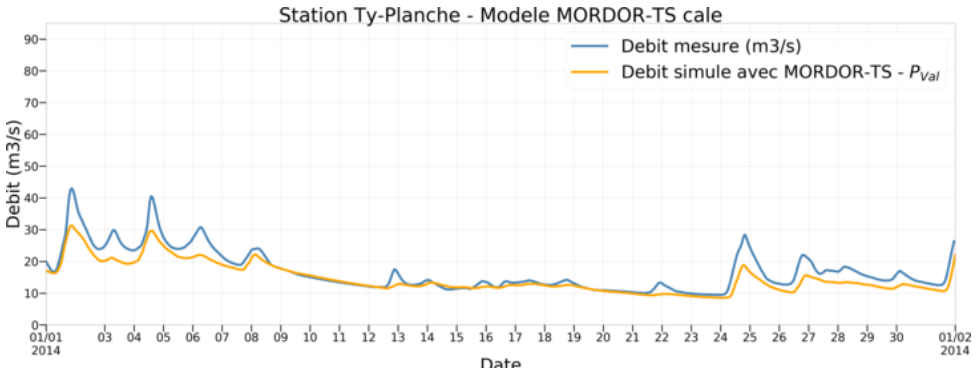
Figure M.9: Débits mesurés et simulés avec le modèle MORDOR-TS à la station Kerjean.



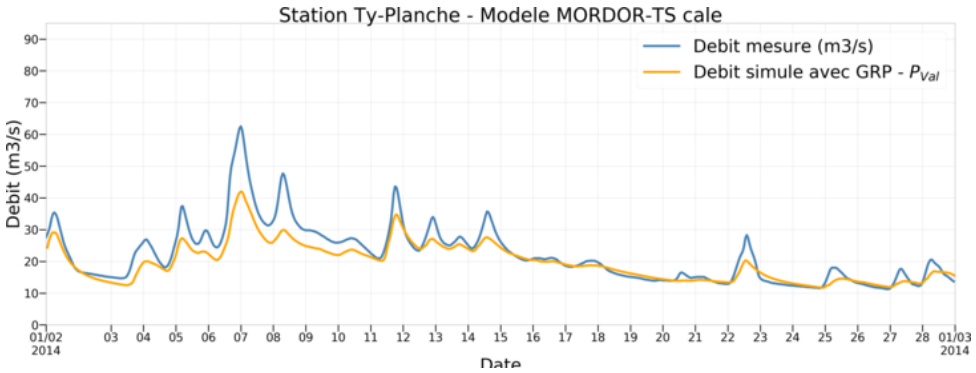
(a) Décembre 2012 - Période de calage



(b) Décembre 2013 - Période de calage



(c) Janvier 2014 - Période de validation



(d) Février 2014 - Période de validation

Figure M.10: Débits mesurés et simulés avec le modèle MORDOR-TS à la station Ty-Planche.

Annexe N

Compléments au chapitre 3

Sommaire

N.1 Méthodes d'analyse de sensibilité globale 344
N.1.1 Méthodes adaptées aux modèles linéaires 344
N.1.2 Calcul des indices de sensibilité dans le cadre ANOVA 345
N.1.3 Méthodes basées sur les moments 346
N.2 Calibration statistique 348
N.2.1 Méthodes paramétriques 348
N.2.1.1 Ensemble Model Output Statistics (EMOS) 348
N.2.1.2 Bayesian Model Averaging (BMA) 350
N.2.2 Méthodes non paramétriques 351
N.2.2.1 Recalibration d'histogrammes de rang 351
N.2.2.2 Régression quantile 352
N.3 Formulation de l'analyse avec deux estimateurs 353
N.3.1 Estimateur de variance minimal dans un cadre gaussien 353
N.3.2 Estimateur du maximum *a posteriori* dans un cadre linéaire gaussien 355

N.1 Méthodes d'analyse de sensibilité globale

Dans cette section, le paragraphe N.1.1 présente les méthodes d'analyse de sensibilité globale adaptées aux modèles linéaires. Le paragraphe N.1.2 présente les méthodes FAST et la méthode de Sobol qui permettent de calculer les indices de sensibilité des méthodes de décomposition de la variance ANOVA. Le paragraphe N.1.3 présente pour finir la méthode d'analyse de sensitivité globale basée sur les moments.

N.1.1 Méthodes adaptées aux modèles linéaires

Dans les méthodes de criblage (*screening* en anglais), l'amplitude des variations des valeurs de sortie du modèle est utilisée comme critère d'importance. Ces méthodes sont basées sur la discrétisation en plusieurs valeurs nommées niveaux. L'OAT (*One at A Time*, ou un-par-un) est une méthode classique de criblage, qui consiste à faire varier une variable incertaine à la fois de manière séquentielle, en utilisant deux ou trois niveaux par facteurs [Kleijnen, 2008]. Saltelli and Annoni [2010] synthétisent les limitations des méthodes OAT. La méthode de Morris [Morris, 1991] consiste à effectuer r fois un plan OAT, r étant classiquement compris entre 5 et 10. Le domaine de chaque entrée est discrétisé de manière uniforme en un même nombre de niveaux, qui dépend du nombre de répétitions r que l'on souhaite faire. Le point de départ de l'expérience OAT ainsi que la suite de directions pour lesquelles on évalue les nouvelles expériences sont tirés aléatoirement. À partir du calcul des effets élémentaires de la variable X_j pour le plan i notée E_j^i représentant l'accroissement du modèle entre deux points successifs, les indices de sensibilité $\mu_i = \frac{1}{r}\sum_{i=1}^{r} |E_j^i|$ (moyenne absolue des effets) et σ_i (écart-type des effets) sont calculés. Cette méthode permet de classer les entrées selon trois catégories : les entrées dont les effets sont négligeables, les entrées dont les effets impactent linéairement la sortie sans interaction avec les autres variables d'entrée et les entrées dont les effets sont non linéaires et/ou avec interactions.

D'autres méthodes, basées sur la régression linéaire, sont utilisables pour les modèles linéaires ou quasi-linéaires [Saltelli and Sobol', 1995]. Une régression linéaire entre la sortie du modèle et les entrées est réalisée. On vérifie que la relation est approximativement linéaire (analyse des résidus, coefficient de détermination R^2, [Saporta, 2006]). Si l'hypothèse de linéarité est validée, les indices de sensibilité peuvent être le coefficient de corrélation linéaire (coefficient de Pearson), le coefficient de régression standard (SRC) ou le coefficient de corrélation partielle (PCC) [Saporta, 2006]. L'échantillon doit être de taille $N >= k + 1$, k représentant le nombre de paramètres incertains. Le coefficient de Pearson vaut 1 si la sortie dépend linéairement des entrées, 0 si elle est indépendante. Si les variables d'entrée sont indépendantes, les coefficients SRC^2 forment une décomposition de la variance de réponse, et chaque SRC_i^2 donne la part de variance de la sortie expliquée par le facteur X_i. Le PCC indique la sensibilité de la sortie à une composante de l'entrée

lorsque les autres entrées sont inchangées.

N.1.2 Calcul des indices de sensibilité dans le cadre ANOVA

Il existe plusieurs méthodes pour calculer les indices de sensibilité des méthodes ANOVA. Nous décrivons ci-après la méthode FAST et la méthode de Sobol.

La méthode FAST

Dans la méthode FAST, pour *Fourier Amplitude Sensitivity Test* [Cukier et al., 1977, Saltelli et al., 1999], la décomposition de la variance fonctionnelle est basée sur la décomposition de Fourier. L'échantillonnage est réalisé en parcourant l'intervalle de variation des différents paramètres d'entrée avec des fréquences différentes. Une trajectoire déterministe, constituée d'une suite de N scénarios $X_j = (X_1, ..., X_k)_j$ est ainsi construite remplissant l'espace des entrées. Une fois les simulations effectuées, le calcul des indices de sensibilité est réalisé grâce à une analyse fréquentielle, en recherchant les entrées dont les variations sont en phase avec la réponse.

La méthode de Sobol

La méthode de Sobol [Sobol, 2001] est basée sur une procédure stochastique. Elle nécessite un très grand nombre de simulations, classiquement plusieurs dizaines de milliers. Saltelli et al. [2010] proposent une méthode d'estimation des indices du premier ordre et des indices totaux. Le principe est la réalisation de simulations appariées, pour lesquelles les entrées ne diffèrent que par un sous-ensemble. Les différences que l'on observe sur la sortie sont alors dues aux facteurs qui ont varié. Les principales étapes sont les suivantes :

(i) Tirage aléatoire de deux échantillons de taille (N, k) pour les k variables incertaines. Ces tirages sont stockés dans deux matrices appelées A et B. Il est recommandé d'utiliser une suite de Sobol plutôt qu'un tirage classique de Monte-Carlo.

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{Nk} \end{pmatrix} \quad B = \begin{pmatrix} b_{11} & \cdots & b_{1k} \\ \vdots & \ddots & \vdots \\ b_{N1} & \cdots & b_{Nk} \end{pmatrix}$$

(ii) Création de k matrices C_i formées par toutes les colonnes de A sauf la colonne i issue de B.

$$C_i = \begin{pmatrix} a_{11} & \cdots & a_{1,i-1} & b_{1i} & a_{1,i+1} & \cdots & a_{1k} \\ \vdots & \vdots & \ddots & \vdots & & & \\ a_{N1} & \cdots & a_{N,i-1} & b_{Ni} & a_{N,i+1} & \cdots & a_{Nk} \end{pmatrix}$$

(iii) Calcul des sorties du modèle pour toutes les vecteurs d'entrée correspondant aux

lignes dans A, B et C_i, , $i \in [1,k]$: $y_A^{(j)}$, $y_B^{(j)}$ et $y_{C_i}^{(j)}$, $j \in [1,N]$. Cela nécessite $N(k+2)$ intégrations du modèle : N intégrations pour la matrice A, N intégrations pour la matrice B, et Nk intégrations pour les k matrice C_i.

(iv) Calcul des indices du premier ordre et des indices totaux selon les formules suivantes :

$$S_i = \frac{V_{X_i}(E_{X_{\sim i}}(Y|X_i))}{Var(Y)} \tag{N.1}$$

$$S_{T_i} = \frac{E_{X_{\sim i}}(V_{X_i}(Y|X_{X_{\sim i}}))}{Var(Y)} \tag{N.2}$$

où V_{X_i}, E_{X_i} désignent la variance ou la moyenne calculées sur X_i, et $V_{X_{\sim i}}$, $E_{X_{\sim i}}$ désignent la variance ou la moyenne calculées sur l'ensemble des facteurs à l'exclusion de X_i. Les formules suivantes sont proposées dans Saltelli et al. [2010] :

$V_{X_i}(E_{X_{\sim i}}(Y|X_i)) = \frac{1}{N}\sum_{j=1}^{N} y_B^{(j)}(y_{C_i}^{(j)} - y_A^{(j)})$

$E_{X_{\sim i}}(V_{X_i}(Y|X_{X_{\sim i}})) = \frac{1}{2N}\sum_{j=1}^{N}(y_A^{(j)} - y_{C_i}^{(j)})^2$

$V(Y) = \frac{1}{N}\sum_{j=1}^{N}(y_A^{(j)})^2 - f_0^2$ avec $f_0^2 = \frac{1}{N}\sum_{j=1}^{N} y_A^{(j)} y_B^{(j)}$

Les indices obtenus par calcul numérique sont des estimations. Les estimateurs sont non-biaisés et convergent, mais il y a un risque d'imprécision selon la qualité et la taille de l'échantillon : les estimateurs de S_i et S_{T_i} peuvent être en-dehors de l'intervalle [0,1], ou l'estimation de l'indice total d'un facteur peut être inférieure à celle de son indice principal. Il est possible alors possible d'évaluer la précision des estimateurs par bootstrap [Archer et al., 1997]. On peut aussi vérifier la convergence des estimateurs en réalisant plusieurs estimations pour des échantillons de taille croissante.

Il est à noter que dans le module OpenTURNS utilisé dans la thèse, le calcul des indices de Sobol est réalisé selon la méthodologie décrite dans Saltelli [2002]. La différence avec l'algorithme plus récent décrit en annexe N.1.2 réside uniquement dans les formules pour le calcul des indices de sensibilité du premier ordre et totaux. Le principe pour le tirage des échantillons reste le même.

N.1.3 Méthodes basées sur les moments

Les méthodes ANOVA décrites au paragraphe 3.2.2.2 n'utilisent que les deux premiers moments (moyenne et variance) des distributions de Y conditionnées par les paramètres d'entrée. [Borgonovo, 2007, Borgonovo et al., 2011] ont développé une méthode pour pallier les limites des méthodes ANOVA (hypothèse de normalité et de paramètres indépendants) et ont proposé une mesure de sensibilité faisant intervenir l'intégralité de ces

distributions. Notons f_Y la densité de probabilité de la variable de sortie Y, et $f_{Y|X_i}$ la densité de probabilité conditionnelle de Y sachant X_i. Plus la densité conditionnelle $f_{Y|X_i}$ est différente de la densité non conditionnelle f_Y, plus le paramètre X_i a un impact sur la quantité d'intérêt. Des critères de distance entre les courbes f_Y et $f_{Y|X_i}$ peuvent être calculés et servir pour estimer des indices de sensibilité. [Borgonovo, 2007] propose la norme L_1. Pour cela, on définit une variable aléatoire $s(X_i)$, qui correspond à l'aire entre les deux courbes de densité de probabilité f_Y et $f_{Y|X_i}$:

$$s(X_i) = \int_R |f_Y(y) - f_{Y|X_i}| dy \tag{N.3}$$

L'intégration de cette distance suivant la loi de X_i permet de définir l'indice δ_i :

$$\delta_i = \frac{1}{2} E(s(X_i)) \tag{N.4}$$

δ_i est un indice normalisé qui représente le décalage provoqué par X_i sur la distribution de Y. Cet indice se généralise naturellement aux groupes de variables. Il vérifie la propriété $0 \leq \delta_i \leq 1$. Si Y est indépendante de X_i, alors $\delta_i = 0$. Cette méthode ne nécessite pas un échantillonnage particulier, et ne fait pas l'hypothèse que les variables sont indépendantes. Les limites qui peuvent être émises concernent la capacité à évaluer correctement les densités de probabilité. Notons également que cette analyse peut également être effectuée à partir des fonctions de répartition (*Cumulative Density Function CDF*) non conditionnelle et conditionnelles (méthode *PAWN*, Pianosi and Wagener [2015]).

N.2 Calibration statistique

Dans cette section, le paragraphe N.2.1 présente deux exemples de méthodes de post-traitement paramétriques : la méthode EMOS (*Ensemble Model Output Statistics*, Gneiting et al. [2005]) et la méthode BMA (*Bayesian Model Averaging*, Raftery et al. [2005]). Le paragraphe N.2.2 présente quant à lui deux exemples de méthodes de post-traitement non paramétriques : la recalibration d'histogrammes de rang et la régression quantile.

N.2.1 Méthodes paramétriques

Parmi les méthodes de post-traitement statistique paramétriques, deux approches sont largement utilisées : la méthode EMOS (*Ensemble Model Output Statistics*, Gneiting et al. [2005]) et la méthode BMA (*Bayesian Model Averaging*, Raftery et al. [2005]). Ces deux méthodes fournissent une densité de probabilité de la variable à prévoir. Le choix a priori de cette densité de probabilité dépend de la variable d'intérêt à calibrer. Les paramètres du modèle prédictif sont estimés sur un échantillon d'apprentissage glissant, qui est typiquement constitué des 20 à 40 jours précédents, en utilisant des techniques d'optimisation de scores [Gneiting and Katzfuss, 2014]. La méthode EMOS (paragraphe N.2.1.1) part du principe que tous les membres de l'ensemble apportent une information équivalente. La calibration consiste à extraire une information issue de la structure globale de l'ensemble (moyenne, variance, ou autre). Un modèle donnant la loi de la variable à prévoir conditionnellement aux informations résumées de l'ensemble est construit. Dans la méthode BMA (paragraphe N.2.1.2), on considère que chaque membre de l'ensemble est un scénario possible. La méthode de calibration consiste à pondérer les membres selon leur probabilité d'être le meilleur. Un modèle donnant la loi de la variable à prévoir conditionnellement à la valeur prise par chacun des membres est construit. Cette méthode fait partie des méthodes dites *par habillage*. L'introduction d'une pondération pour chacun des membres rend cette méthode attractive dans le cas où les membres ne sont plus échangeables, par exemple pour le post-traitement de prévisions d'ensemble multi-modèles.

N.2.1.1 Ensemble Model Output Statistics (EMOS)

La méthode EMOS (*Ensemble Model Output Statistic*, Gneiting et al. [2005]) fait partie de la famille des schémas de régression paramétrique. Dans cette méthode, on fait l'hypothèse que la densité prédictive de l'observation y sachant la prévision $(x_1, ..., x_{N_{ens}})$ est décrite par une unique loi de probabilité, dont les paramètres dépendent de la valeur de tous les membres de l'ensemble (moyenne et variance dans le cas gaussien) :

$$y|(x_1, ..., x_{N_{ens}}) \sim f(y|(x_1, ..., x_{N_{ens}}))$$

Le principe est donc l'ajustement d'une loi de probabilité à partir d'un échantillon d'apprentissage constitué de couples $(prévision, observation)$. En opérationnel, à chaque nouvelle prévision, la loi de probabilité ainsi définie est utilisée pour prévoir une densité de probabilité corrigée de l'observation. La mise en place et la qualité d'un modèle EMOS repose sur plusieurs points essentiels :

(i) Le choix de la loi de distribution prédictive de la variable à prévoir, qui doit être représentative de la variable post-traitée. Nous illustrons dans la suite plusieurs lois de probabilités utilisées classiquement selon la variable à prévoir.

(ii) Le choix des équations de régression, qui est déterminant pour le bon conditionnement des paramètres de la loi de distribution aux statistiques de l'ensemble brut.

(iii) Le choix du critère d'optimisation des paramètres. Une méthode classique est la maximisation de la vraisemblance. Gneiting et al. [2005] montrent cependant que cela mène à des prévisions sur-dispersives et proposent plutôt de minimiser le CRPS.

Le choix de la densité de probabilité de la prévision est guidé par la variable à prévoir. Gneiting et al. [2005], qui sont à l'origine de la première version d'EMOS, proposent de post-traiter des variables de climatologie gaussienne (pression et température) avec l'utilisation d'une loi normale. Il s'agit alors d'estimer la moyenne et l'écart-type de cette distribution à partir de l'ensemble. La moyenne est une correction affine de l'ensemble de la valeur des membres $(a + b_1 x_1 + ... + b_{N_{ens}})$ membres, et la variance est une fonction affine de la variance de l'ensemble $(c + dS^2$, si l'on note S^2 la variance de l'ensemble). Si les membres de l'ensemble sont interchangeables, alors $b_1 = ... = b_{N_{ens}}$. Pour calibrer des précipitations la loi normale n'est plus adaptée : la densité de probabilité est asymétrique, non négative, avec un point d'accumulation en zéro. Scheuerer and Hamill [2015] proposent de se baser sur la loi gamma, en décalant sa densité vers la gauche (*shifted*), puis en censurant cette densité à zéro (*censured*), ce qui consiste à remplacer toutes les valeurs négatives par zéro et permet de représenter le point d'accumulation. De cette manière, les probabilités de précipitation nulle et positive sont traitées simultanément. Cette loi est appelée *Censored Shifted Gamma*. Scheuerer [2014] ont proposé d'utiliser la loi des extrêmes censurée (*Censored Generalized Extreme Value*). Une loi normale tronquée en zéro est utilisée pour calibrer la vitesse du vent [Thorarinsdottir and Gneiting, 2010, Thorarinsdottir and Johnson, 2012] ou sa racine carrée [Hemri et al., 2014, Taillardat et al., 2016].

Des exemples de calibration avec la méthode EMOS dans le domaine météorologique sont disponibles dans Bellier [2018], Taillardat et al. [2016, 2019] pour la calibration de températures et de précipitations. Cette méthode, initialement développée dans le domaine de la météorologie, a également été appliquée avec succès à la calibration des débits hydrologiques [Bellier, 2018, Hemri et al., 2015].

N.2.1.2 Bayesian Model Averaging (BMA)

Dans la méthode BMA (*Bayesian Model Averaging*, Raftery et al. [2005]), on fait l'hypothèse que la densité prédictive de l'observation y sachant la prévision $(x_1, ..., x_{N_{ens}})$ est décrite par une combinaison linéaire de densités, chacune habillant un des membres de l'ensemble :

$$y|(x_1, ..., x_{N_{ens}}) \sim \sum_{i=1}^{x_{Nens}} w_i g(y|x_i)$$

Les poids w_i sont non négatifs et de somme égale à 1. Ils caractérisent la performance du membre correspondant sur l'échantillon d'apprentissage. Ils peuvent être interprétés comme la probabilité que le membre respectif fournisse la meilleure prévision [Raftery et al., 2005]. Dans le cas de membres interchangeables, $w_1 = ... = w_{ens} = 1/m$. Le terme $g(y|x_i)$ est appelé *noyau*, et représente la densité de probabilité *a posteriori* de y sachant le membre x_i sur les données de calage. Le noyau s'exprime à partir d'une loi paramétrique, qui, comme dans le cas de EMOS, doit être capable de représenter le comportement de l'observation. Très souvent, la loi normale est utilisée pour la densité prédictive g. Quand la variable à prévoir n'est pas gaussienne, les données peuvent subir une transformation de manière à les rendre approximativement gaussiennes. Bellier [2018], Duan et al. [2007], Hemri et al. [2013], Madadgar and Moradkhani [2014] ont ainsi utilisé une loi gaussienne pour le post-traitement de prévisions d'ensemble de débits après transformation. Pour la calibration de précipitations ou de vitesse de vent, la loi gamma peut être utilisée [Gneiting and Katzfuss, 2014]. L'ensemble brut doit au préalable être débiaisé de manière à corriger les biais systématiques. Ensuite, l'estimation des poids et de la variance des noyaux est généralement réalisée par maximisation du logarithme de la fonction de vraisemblance sur les données de calage. Contrairement à la méthode EMOS, la méthode BMA permet à la densité prédictive de prendre des formes très diverses et peut représenter des densités multi-modales. Cette stratégie permet donc plus de flexibilité. La figure N.1 est une illustration dans le cas de cinq membres échangeables, dont la variance des composantes gaussiennes des distribution de chaque membre est la même. Les cinq membres de la prévision préalablement débiaisés sont représentés par des ronds. Cette figure met en évidence que la distribution prédictive présente deux modes.

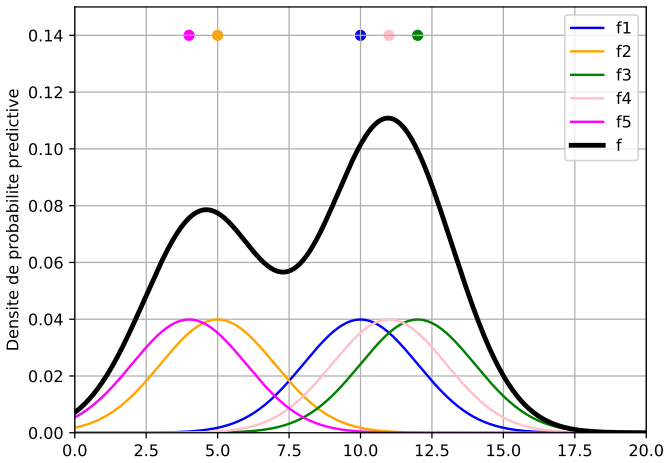


Figure N.1: Illustration de la distribution prédictive pour la méthode BMA (courbe noire), construite à partir d'un ensemble de membres échangeables de taille 5. Les points correspondent aux membres débiaisés de la prévision.

N.2.2 Méthodes non paramétriques

N.2.2.1 Recalibration d'histogrammes de rang

Il existe plusieurs méthodes non paramétriques. Nous décrivons dans ce paragraphe les méthodes de recalibration d'histogrammes de rang et la régression quantile.

Recalibration d'histogramme de rang La recalibration d'histogrammes de rang (en anglais *Rank histogram recalibration*, Hamill and Colucci [1997]) est la première méthode paramétrique à avoir été proposée. Elle était alors appliquée au post-traitement de prévisions d'ensemble de températures. Elle a pour objectif d'obtenir des ensembles post-traités avec un histogramme de rang plat. On note $(x_1, ..., x_{N_{ens}})$ les membres d'une prévision d'ensemble à un instant donné, y l'observation correspondante, et R l'histogramme de rang à $(N_{ens} + 1)$ rangs représentatif de l'observation comparée à l'ensemble, calculé sur une période historique. La probabilité que l'observation soit inférieure à la valeur du quantile d'ordre i de la prévision est $p(y < x_i) = \sum_{j=1}^{i} R_j$. On émet l'hypothèse que les probabilités sont uniformément distribuées entre la valeur 0 et le plus petit membre de l'ensemble, et également au sein d'un rang donné. Ainsi, pour un seuil T non égal à la valeur d'un quantile, une interpolation linéaire peut être réalisée :

$$p(x_i < y < T) = (\frac{T - x_i}{x_{i+1} - x_i}) R_{i+1}, \text{ avec } x_i < T < x_{i+1},$$
$$p(0 < y < T) = (\frac{T}{x_1}) R_1.$$

Pour un seuil supérieur à la valeur maximale de l'ensemble, une distribution de Gumbel est utilisée pour l'extrapolation.

N.2.2.2 Régression quantile

Une seconde méthode non paramétrique est la méthode de régression quantile (en anglais *quantile regression*). Son application au post-traitement des prévisions d'ensemble a été introduite par Bremnes [2004]. On note q_p le quantile tel que la probabilité que l'observation y soit inférieure à q_p soit p : $p(y < q_p) = p$. L'estimation du quantile q_p pour la prévision à un instant donné t est réalisée à partir du vecteur de prédicteurs à I éléments $x_t = (x_{t,1}, ...x_{t,I})$:

$$q_p(x_t) = a_q + \sum_{i=1}^{I} b_{q,i} x_{t,i}$$

Les coefficients de régression a_q et $b_{q,i}$ sont estimés pour chaque quantile par minimisation d'un score. Bremnes [2004] ont travaillé sur les cinq quantiles de prévision $q_{0.05}$, $q_{0.25}$, $q_{0.50}$, $q_{0.75}$ et $q_{0.95}$. Chacune des cinq régressions quantiles a été réalisée avec $I = 2$ prédicteurs que sont les quantiles $q_{0.25}$ et $q_{0.75}$ de l'ensemble.

N.3 Formulation de l'analyse avec deux estimateurs

Les deux paragraphes suivants N.3.1 et N.3.2 formulent l'analyse dans le cas linéaire gaussien en utilisant respectivement les estimateurs de maximum *a posteriori* et de minimum de variance. Toutes les quantités utilisées sont relatives au temps t_k. Pour alléger les notations, l'indice k sera omis.

N.3.1 Estimateur de variance minimal dans un cadre gaussien

Sous les hypothèses émises, l'estimateur BLUE (*Best Linear Unbiased Estimator*, meilleur estimateur linéaire non biaisé) permet de fournir une estimation du vecteur de contrôle analysé $\mathbf{x^a}$ connaissant les observations $\mathbf{y^o}$, ainsi qu'une estimation de l'incertitude de l'estimateur par le biais d'une matrice de covariance d'erreur $\mathbf{A}$. Le BLUE est un estimateur du minimum de variance.

Équation d'analyse du BLUE

Hypothèses

- Analyse linéaire (Linear de BLUE) : on cherche une analyse définie par une correction de l'ébauche qui soit linéaire avec les écarts entre observation et ébauche.
- Analyse optimale (Best de BLUE) : on cherche un vecteur de contrôle analysé aussi proche possible du vecteur de contrôle vrai dans le sens du minimum de variance.

alors

Le vecteur de contrôle analysé fourni par le BLUE est donné par les formules suivantes :

$$\mathbf{x^a} = \mathbf{x^b} + \mathbf{K}(\mathbf{y^o} - \mathbf{H}\mathbf{x^b}) \tag{N.5}$$

$$\mathbf{K} = \mathbf{B}\mathbf{H^T}(\mathbf{H}\mathbf{B}\mathbf{H^T} + \mathbf{R})^{-1}. \tag{N.6}$$

L'opérateur $\mathbf{K}$ est appelé *matrice de gain* de l'analyse et $(\mathbf{y^o} - \mathbf{H}\mathbf{x^b})$ est appelé *innovation*. La matrice de covariance d'erreur d'analyse est alors donnée par :

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}. \tag{N.7}$$

Nous cherchons à présent à démontrer les équations du BLUE. D'après les définitions des

erreurs données précédemment, on peut écrire :

$$\mathbf{x^a} = \mathbf{x^t} + \epsilon^{\mathbf{a}} \tag{N.8}$$

$$\mathbf{x^b} = \mathbf{x^t} + \epsilon^{\mathbf{b}} \tag{N.9}$$

$$\mathbf{y^t} = \mathbf{H}\mathbf{x^t} \tag{N.10}$$

$$\mathbf{y^o} = \mathbf{H}\mathbf{x^t} + \epsilon^{\mathbf{o}}. \tag{N.11}$$

L'équation N.5 découle directement du fait que l'on cherche une analyse définie par une correction de l'ébauche, linéaire avec les écarts entre observation et ébauche.

Comme indiqué dans les hypothèses, on cherche une analyse de variance d'erreur minimale. Cela revient à minimiser la trace de la matrice des covariances d'erreur d'analyse $Tr(\mathbf{A})$, c'est-à-dire la variance d'erreur totale. On cherche donc la matrice $\mathbf{K}$ qui minimise $Tr(\mathbf{A})$. Nous cherchons donc à écrire la matrice $\mathbf{A}$. D'après l'équation N.5 et les équations N.8 à N.11, on peut écrire :

$$\mathbf{x^t} + \epsilon^{\mathbf{a}} = \mathbf{x^t} + \epsilon^{\mathbf{b}} + \mathbf{K}(\mathbf{H}\mathbf{x^t} + \epsilon^{\mathbf{o}} - \mathbf{H}\mathbf{x^b})$$

donc

$$\epsilon^{\mathbf{a}} = \epsilon^{\mathbf{b}} + \mathbf{K}\mathbf{H}(\mathbf{x^t} - \mathbf{x^b}) + \mathbf{K}\epsilon^{o}$$

et finalement

$$\epsilon^{\mathbf{a}} = \epsilon^{\mathbf{b}}(\mathbf{I} - \mathbf{K}\mathbf{H}) + \mathbf{K}\epsilon^{\mathbf{o}}.$$

Puisque $\epsilon^{\mathbf{o}}$ et $\epsilon^{\mathbf{b}}$ sont non corrélées, leur covariance croisée est nulle. En développant l'expression $\epsilon^{\mathbf{a}}(\epsilon^{\mathbf{a}})^T$ et en prenant sa moyenne, nous pouvons exprimer la matrice de covariance d'erreur d'analyse A quel que soit l'opérateur K :

$$\mathbf{A} = (\mathbf{I} - \mathbf{K}\mathbf{H})\mathbf{B}(\mathbf{I} - \mathbf{K}\mathbf{H})^{\mathbf{T}} + \mathbf{K}\mathbf{R}\mathbf{K}^{\mathbf{T}}. \tag{N.12}$$

La trace de la matrice $\mathbf{A}$ s'exprime par

$$Tr(\mathbf{A}) = Tr(\mathbf{B}) + Tr(\mathbf{K}\mathbf{H}\mathbf{B}\mathbf{H}^{\mathbf{T}}\mathbf{K}^{\mathbf{T}}) - 2Tr(\mathbf{B}\mathbf{H}^{\mathbf{T}}\mathbf{K}^{\mathbf{T}}) + Tr(\mathbf{K}\mathbf{R}\mathbf{K}^{\mathbf{T}}).$$

C'est une fonction scalaire des coefficients de la matrice $\mathbf{K}$, continue et différentiable. Sa différentielle d_K est obtenue en prenant les termes au premier ordre par rapport à $\mathbf{K}$ de la différence $Tr(\mathbf{A})(\mathbf{K} + \mathbf{L}) - Tr(\mathbf{A})(\mathbf{K})$ où $\mathbf{L}$ est une matrice de perturbation arbitraire. On peut montrer que quelle que soit $\mathbf{L}$,

$$d_K[Tr(\mathbf{A})]\mathbf{L} = 2Tr([\mathbf{K}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathbf{T}} + \mathbf{R}) - \mathbf{B}\mathbf{H}^{\mathbf{T}}]\mathbf{L}^{\mathbf{T}}).$$

En annulant la dérivée d_K, on obtient la matrice de gain $\mathbf{K}$ qui minimise la variance du vecteur de contrôle analysé : c'est celle donnée par l'équation N.6. $\mathbf{K}$ est alors le gain optimal au sens des moindres carrés. La matrice $\mathbf{K}$ contient les coefficients de régression de la combinaison linéaire optimale au sens du minimum de variance entre l'ébauche $\mathbf{x^b}$ et les observations $\mathbf{y^o}$. La variance du vecteur de contrôle analysé $\mathbf{x^a}$ est minimale et non

biaisée. En remplaçant dans l'équation N.12 la valeur de $\mathbf{K}$ par celle du gain optimal défini par l'équation N.6, on en déduit la matrice de covariance d'erreurs d'analyse de l'équation N.7.

Nous notons que l'équation d'analyse N.5 du BLUE donne d'autant plus de poids aux observations que les variances de $\mathbf{R}$ sont petites par rapport à celles de $\mathbf{B}$. Avec l'hypothèse gaussienne, la loi de probabilité $p(\mathbf{x_k}|\mathbf{y_k^o})$ est normale de moyenne $\mathbf{x^a}$ et de matrice de covariance $\mathbf{A}$. Dans ce cas, l'approche du BLUE permet ainsi de déterminer complètement la loi de probabilité conditionnelle de l'approche bayésienne.

N.3.2 Estimateur du maximum *a posteriori* dans un cadre linéaire gaussien

Dans ce paragraphe, nous adoptons à présent une approche variationnelle en cherchant à exprimer l'analyse à partir de l'estimateur du maximum *a posteriori*. Nous cherchons ici à exprimer analytiquement la PDF $p^a(\mathbf{x})$. Puisqu'on suppose que le vecteur de contrôle d'ébauche $\mathbf{x^b}$ et les observations $\mathbf{y^o}$ suivent des lois gaussiennes et que l'opérateur d'observation est linéaire, on a les relations suivantes :

$$p^b(\mathbf{x}) \; \alpha \; exp(-\frac{1}{2}(\mathbf{x}-\mathbf{x^b})^{\mathbf{T}}\mathbf{B^{-1}}(\mathbf{x}-\mathbf{x^b}))$$
$$p(\mathbf{y^o}|\mathbf{x}) \; \alpha \; exp(-\frac{1}{2}(\mathbf{H}(\mathbf{x})-\mathbf{y^o})^{\mathbf{T}}\mathbf{R^{-1}}(\mathbf{H}(\mathbf{x})-\mathbf{y^o})).$$

Ainsi, la formule 3.13 issue du théorème de Bayes devient :

$$p^a(\mathbf{x}) \; \alpha \; exp(-\frac{1}{2}((\mathbf{x}-\mathbf{x^b})^{\mathbf{T}}\mathbf{B^{-1}}(\mathbf{x}-\mathbf{x^b}) + (\mathbf{H}(\mathbf{x})-\mathbf{y^o})^{\mathbf{T}}\mathbf{R^{-1}}(\mathbf{H}(\mathbf{x})-\mathbf{y^o}))).$$

Puisque l'on utilise l'estimateur du maximum *a posteriori*, la valeur d'analyse que l'on cherche correspond à la valeur de $\mathbf{x}$ qui maximise la densité de probabilité du vecteur de contrôle à posteriori $p^a(\mathbf{x})$. Cela est équivalent à la minimisation de l'opposé du logarithme de cette densité. Nous avons donc l'équivalence suivante :

$$max(p^a(\mathbf{x})) \iff min(-ln(p^a(\mathbf{x}))) = min(\mathcal{J}(\mathbf{x}))$$

où $\mathcal{J}$ représentant la fonction coût telle que $\mathcal{J}(\mathbf{x}) = -ln(p^a(\mathbf{x}))$, donc

$$\mathcal{J}(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x^b})^{\mathbf{T}}\mathbf{B^{-1}}(\mathbf{x}-\mathbf{x^b}) + \frac{1}{2}(\mathbf{H}(\mathbf{x})-\mathbf{y^o})^{\mathbf{T}}\mathbf{R^{-1}}(\mathbf{H}(\mathbf{x})-\mathbf{y^o}). \tag{N.13}$$

On note

$$\mathcal{J}(\mathbf{x}) = \mathcal{J}^b(\mathbf{x}) + \mathcal{J}^o(\mathbf{x})$$

avec
$$\mathcal{J}^b(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x^b})^{\mathbf{T}}\mathbf{B^{-1}}(\mathbf{x}-\mathbf{x^b})$$
$$\mathcal{J}^o(\mathbf{x}) = \frac{1}{2}(\mathbf{H}(\mathbf{x})-\mathbf{y^o})^{\mathbf{T}}\mathbf{R^{-1}}(\mathbf{H}(\mathbf{x})-\mathbf{y^o}).$$

Cette fonction coût est à la base d'une majorité d'algorithmes d'assimilation de données avec une approche variationnelle. $\mathcal{J}^b$ représente l'écart par rapport à l'ébauche, et $\mathcal{J}^o$ représente l'écart par rapport aux observations. Minimiser $\mathcal{J}$ représente un compromis entre être proche de l'ébauche et être proche des observations. Ces deux termes sont pondérés par l'inverse des matrices de covariances d'erreur. Ainsi, plus une donnée (modèle ou observée) est sûre, plus sa variance est faible et plus son poids dans la fonction coût est important.

Puisque $\mathbf{H}$ est linéaire, la fonction coût écrite en N.13 est quadratique en $\mathbf{x}$. Puisque les matrices $\mathbf{B}$ et $\mathbf{R}$ sont définies positives, cette fonction est convexe et possède un seul minimum que l'on peut estimer en annulant son gradient dont l'expression vaut :

$$\nabla\mathcal{J}(\mathbf{x}) = \mathbf{B^{-1}}(\mathbf{x}-\mathbf{x^b}) - \mathbf{H^T R^{-1}}(\mathbf{y^o}-\mathbf{Hx}).$$

Nous développons à présent que $\nabla\mathcal{J}(\mathbf{x^a}) = 0$ pour le vecteur de contrôle analysé :

$$\begin{aligned}
\mathbf{B^{-1}}(\mathbf{x^a}-\mathbf{x^b}) - \mathbf{H^T R^{-1}}(\mathbf{y^o}-\mathbf{Hx^a}) &= 0\\
\mathbf{B^{-1}}(\mathbf{x^a}-\mathbf{x^b}) + \mathbf{H^T R^{-1} Hx^a} - \mathbf{H^T R^{-1} y^o} &= 0\\
\mathbf{B^{-1}}(\mathbf{x^a}-\mathbf{x^b}) + \mathbf{H^T R^{-1} H}(\mathbf{x^a}-\mathbf{x^b}) &= \mathbf{H^T R^{-1} y^o} - \mathbf{H^T R^{-1} Hx^b}\\
(\mathbf{B^{-1}} + \mathbf{H^T R^{-1} H})(\mathbf{x^a}-\mathbf{x^b}) &= \mathbf{H^T R^{-1}}(\mathbf{y^o}-\mathbf{Hx^b}).
\end{aligned}$$

On obtient finalement l'expression du vecteur analysé :

$$\mathbf{x^a} = \mathbf{x^b} + \mathbf{K^*}(\mathbf{y^o}-\mathbf{Hx^b}) \tag{N.14}$$
$$\mathbf{K^*} = (\mathbf{B^{-1}} + \mathbf{H^T R^{-1} H})^{-1}\mathbf{H^T R^{-1}}. \tag{N.15}$$

Le gain $\mathbf{K}^*$ obtenu ici peut être réécrit de manière à retrouver l'équation du BLUE :

$$\begin{aligned}
\mathbf{K^*} &= (\mathbf{B^{-1}} + \mathbf{H^T R^{-1} H})^{-1}\mathbf{H^T R^{-1}}\\
&= (\mathbf{B^{-1}} + \mathbf{H^T R^{-1} H})^{-1}\mathbf{H^T R^{-1}}(\mathbf{R} + \mathbf{HBH^T})(\mathbf{HBH} + \mathbf{R^T})^{-1}\\
&= (\mathbf{B^{-1}} + \mathbf{H^T R^{-1} H})^{-1}(\mathbf{H^T} + \mathbf{H^T R^{-1} HBH^T})(\mathbf{HBH} + \mathbf{R^T})^{-1}\\
&= (\mathbf{B^{-1}} + \mathbf{H^T R^{-1} H})^{-1}(\mathbf{B^{-1}} + \mathbf{H^T R^{-1} H})\mathbf{BH^T}(\mathbf{HBH} + \mathbf{R^T})^{-1}.
\end{aligned}$$

On obtient finalement :

$$\mathbf{K^*} = \mathbf{BH^T}(\mathbf{HBH} + \mathbf{R^T})^{-1}. \tag{N.16}$$

Les équations N.14 et N.16 permettent de retrouver la formulation de l'analyse du BLUE.

Annexe O

Compléments au chapitre 5

Sommaire

O.1 Expérience *corr1par* - Résultats à Tréodet 358
O.2 Expérience *corr1par* - Résultats à Kerjean 363
O.3 Expérience *corr1par* - Résultats à Ty-Planche 368
O.4 Expérience *corr8par* - Résultats à Kerjean 373
O.5 Expérience *corr8par* - Résultats à Ty-Planche 375

O.1 Expérience *corr1par* - Résultats à Tréodet

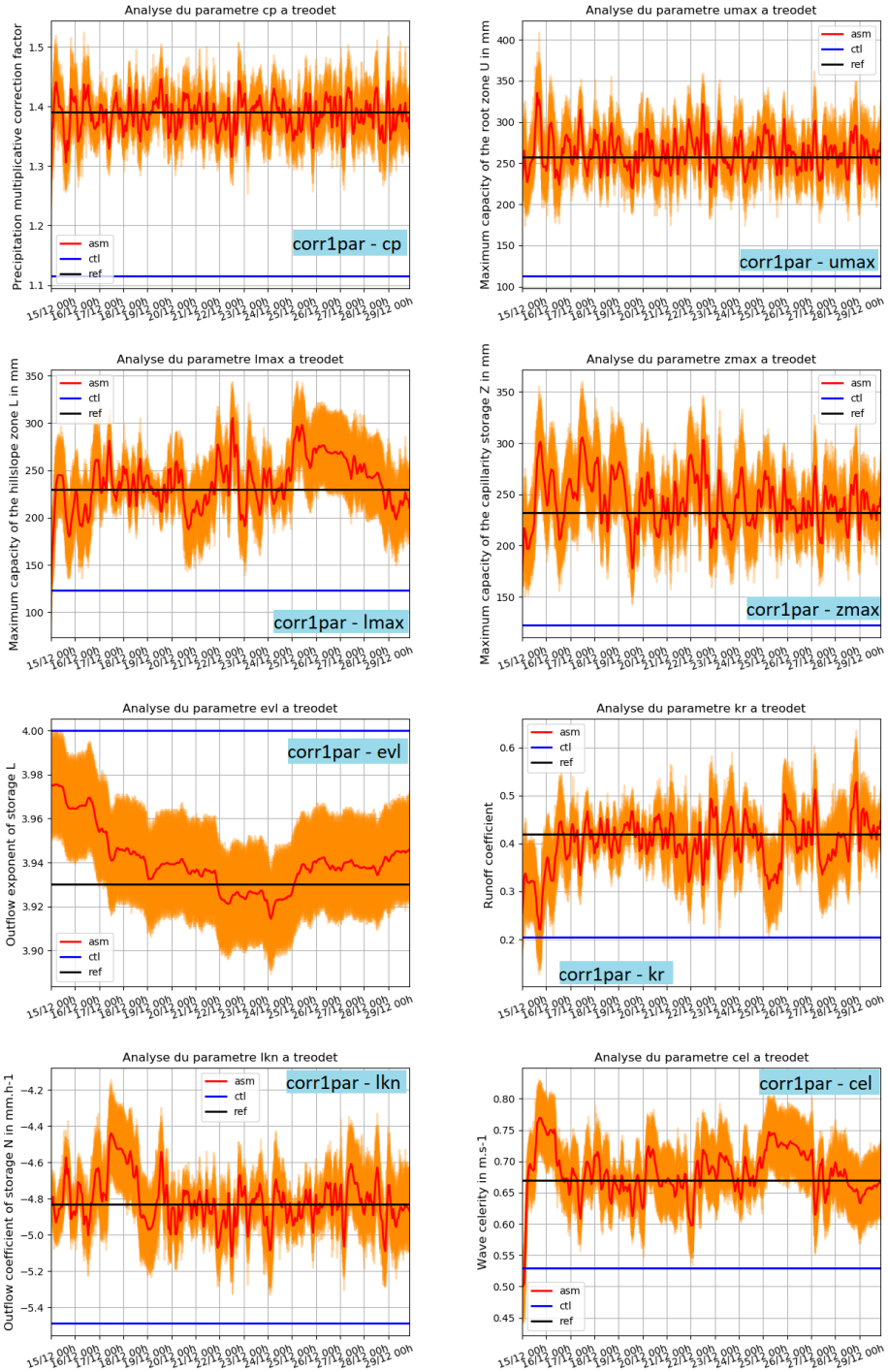


Figure O.1: Valeurs analysées pour le vecteur de contrôle (paramètres) en fonction du temps pour les 8 expériences *corr1par* menées à Tréodet. Les courbes orange représentent les valeurs pour chacun des 49 membres et la courbe rouge la valeur analysée moyenne.

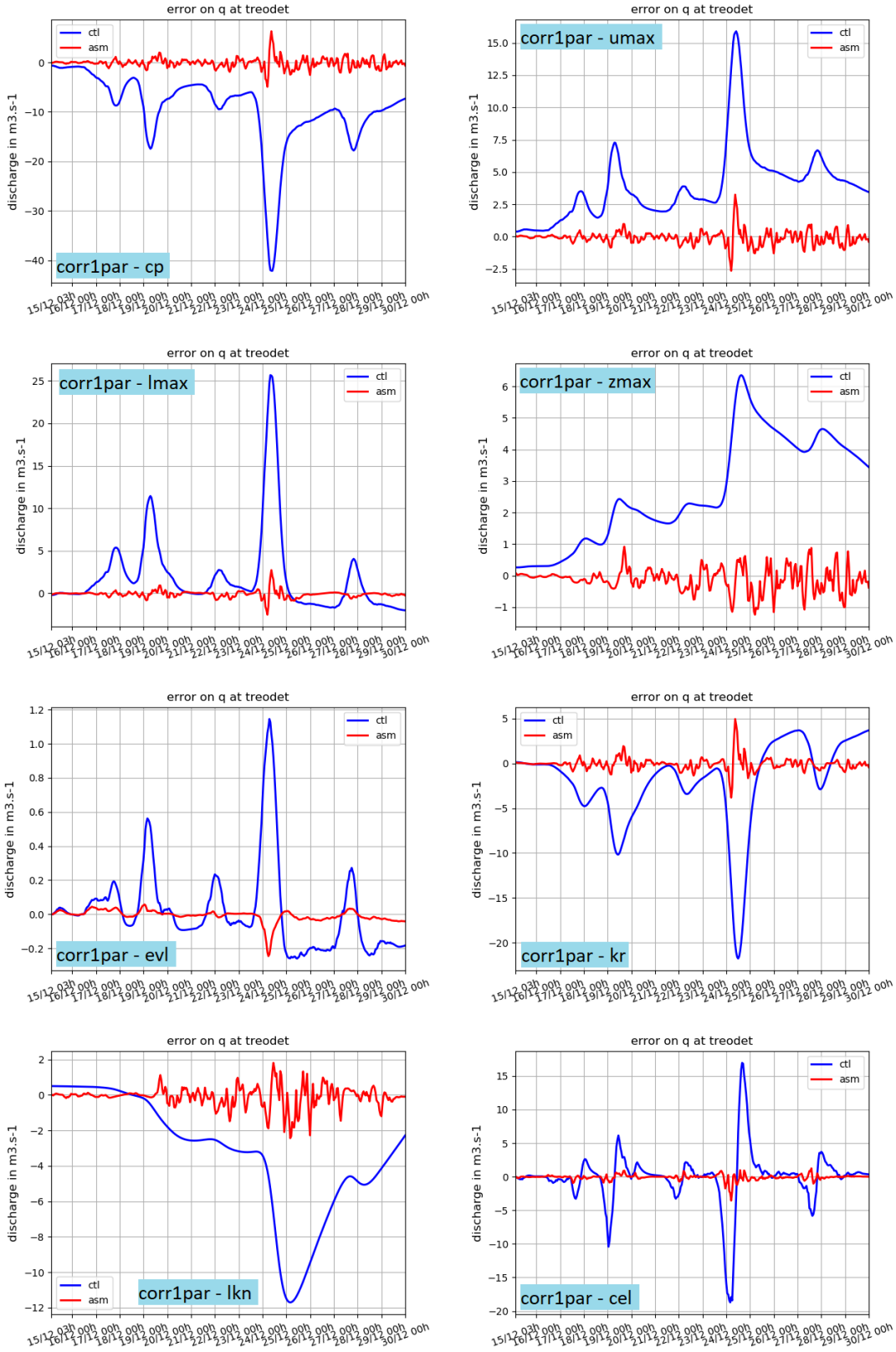


Figure O.2: Chroniques d'erreurs de la simulation de contrôle et de l'analyse par rapport à la simulation de référence pour les 8 expériences *corr1par* menées à Tréodet. La valeur de l'analyse est prise à la fin de la fenêtre d'assimilation.

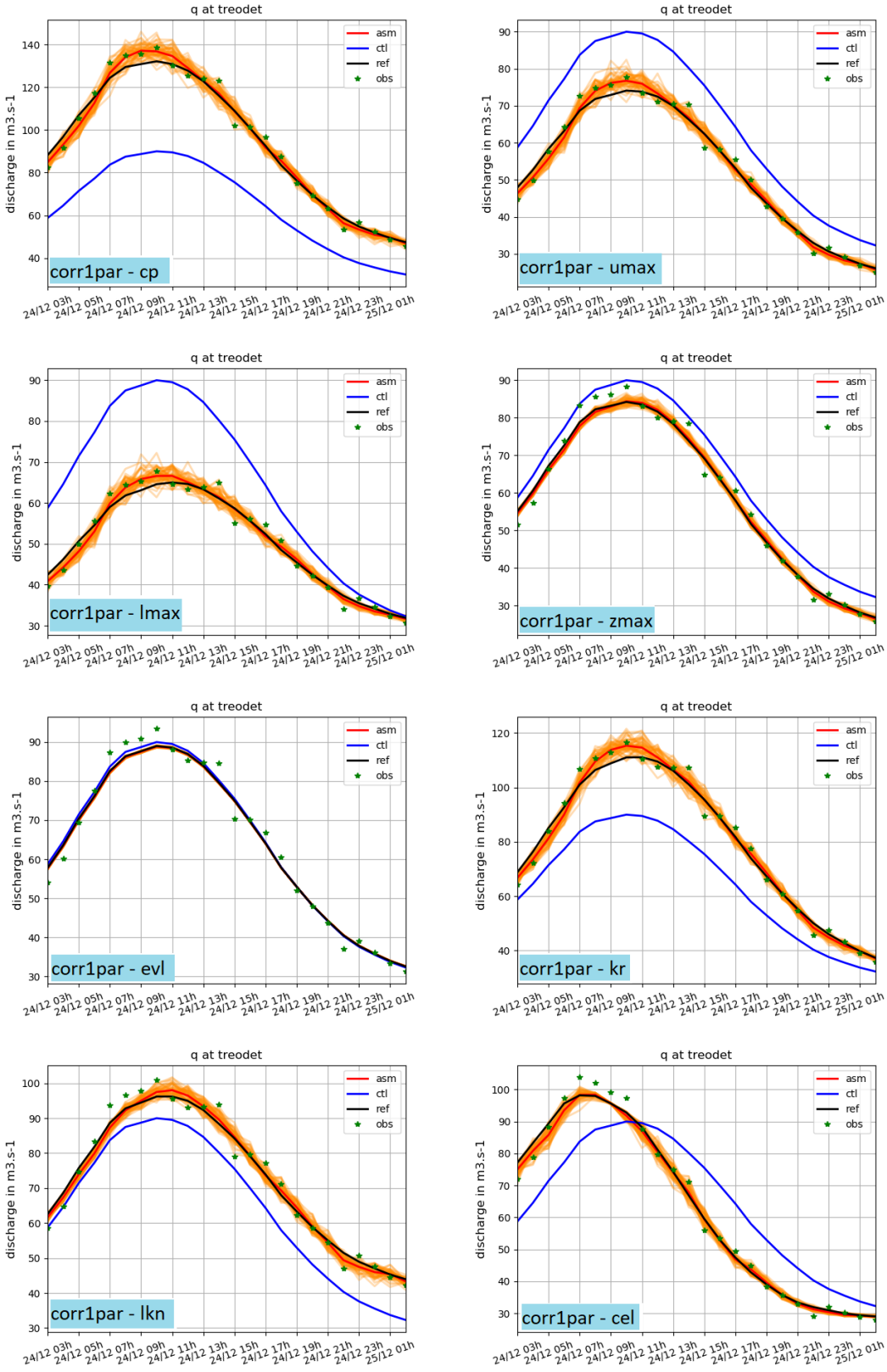


Figure O.3: Simulation de contrôle, simulation de référence, analyses et observations assimilées au moment du pic d'événement le 24 décembre 2013 pour les 8 expériences *corr1par* menées à Tréodet.

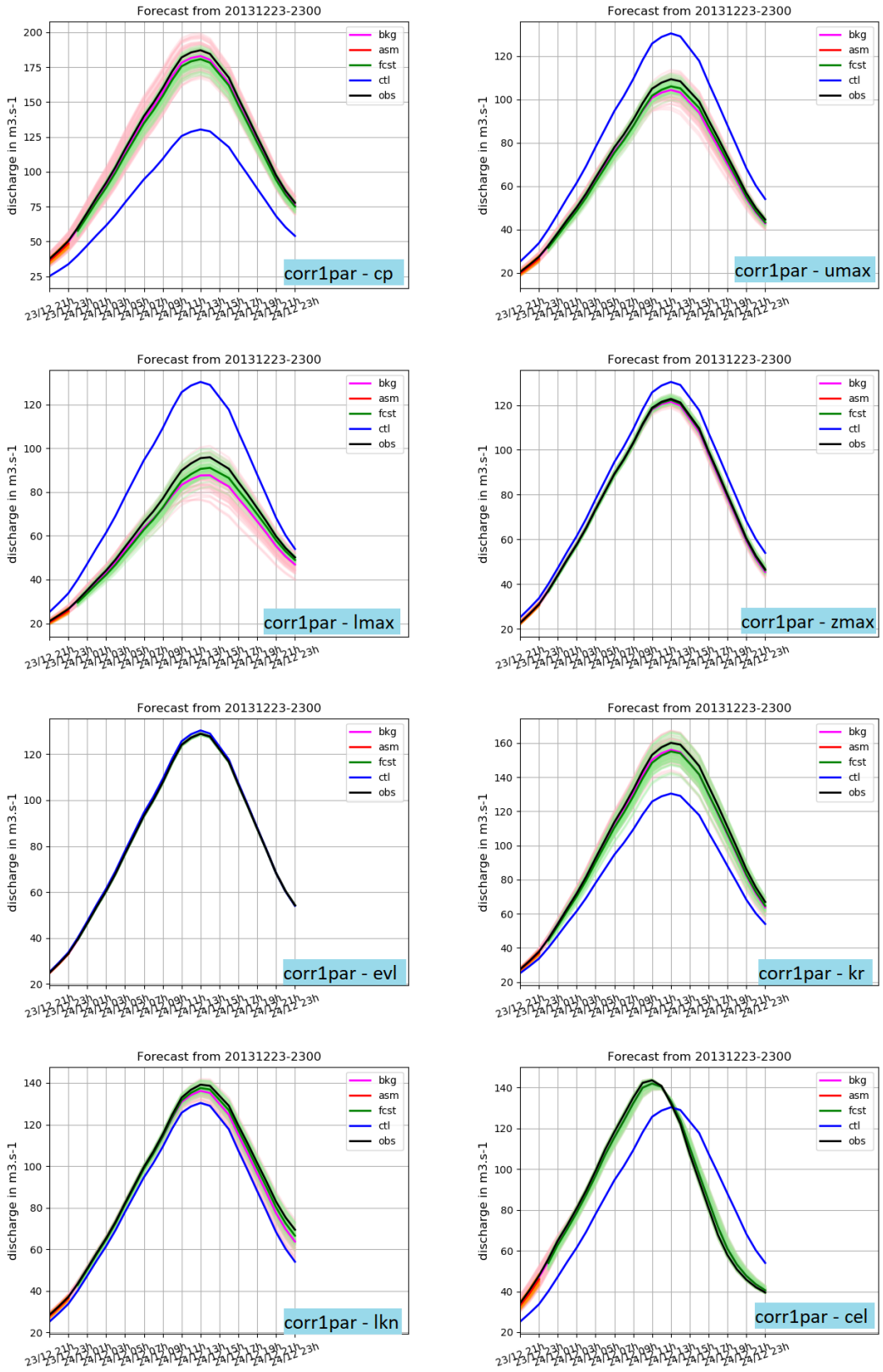


Figure O.4: Visualisation pour 8 expériences *corr1par* menées à Tréodet du cycle d'assimilation de données débutant le 23 décembre 2013 à 20 heures : ensemble d'ébauche (bkg), ensemble analysé (asm) et ensemble en prévision suite à l'analyse (fcst), simulation de contrôle (ctl) et simulation de référence (ref). À partir du temps de base (23 décembre à 23 heures), toutes les simulations sont alimentées par les prévisions de pluie RR3.

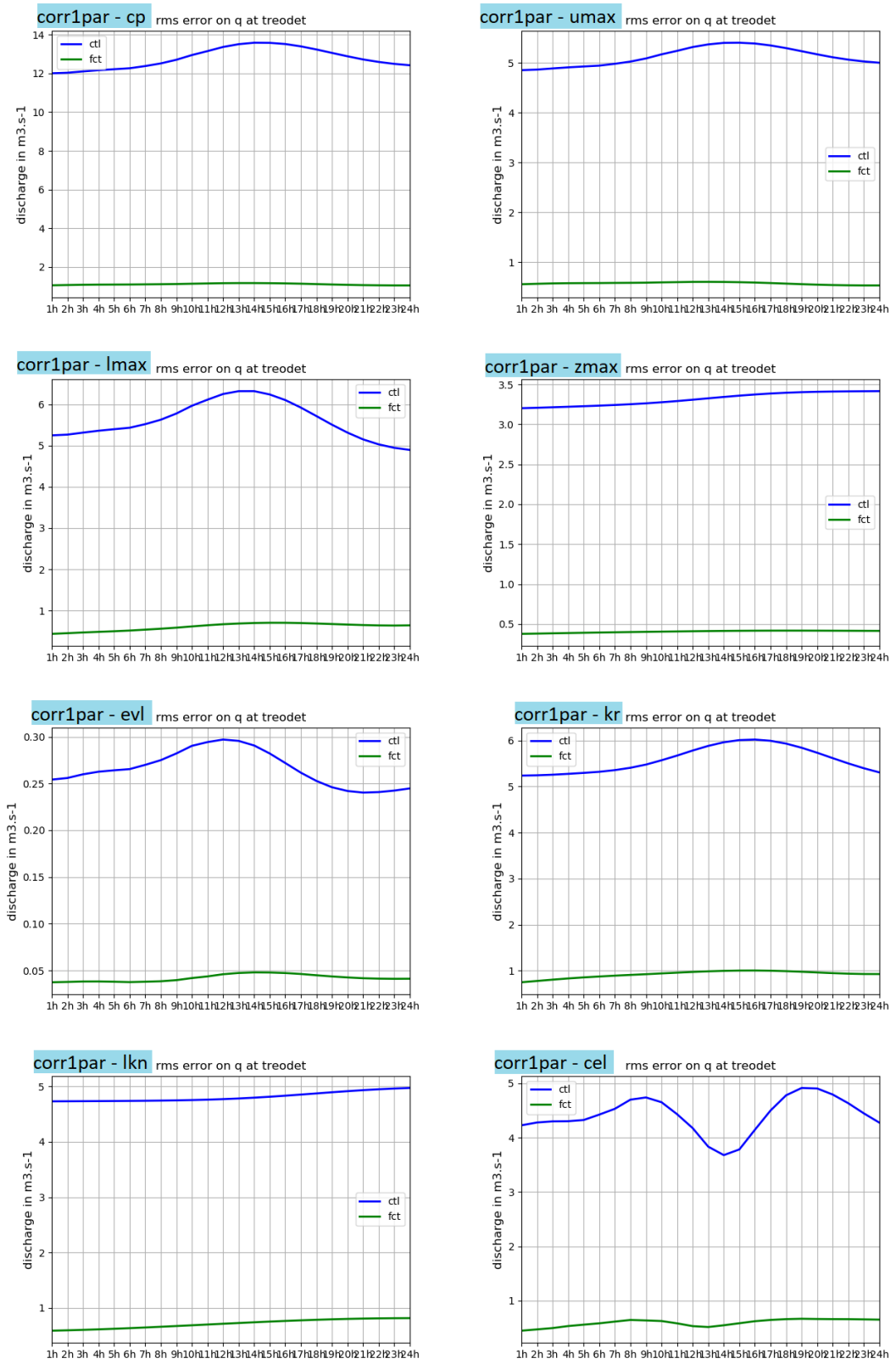


Figure O.5: RMSE des erreurs de prévision en fonction de l'échéance pour la simulation de contrôle (*ctl* en bleu) et le modèle analysé (*fcst* en vert) par rapport à la simulation de référence pour les 8 expériences *corr1par* menées à Tréodet. Toutes les simulations (de référence, de contrôle et le modèle corrigé) sont alimentées par les prévisions de pluie RR3 en prévision.

O.2 Expérience *corr1par* - Résultats à Kerjean

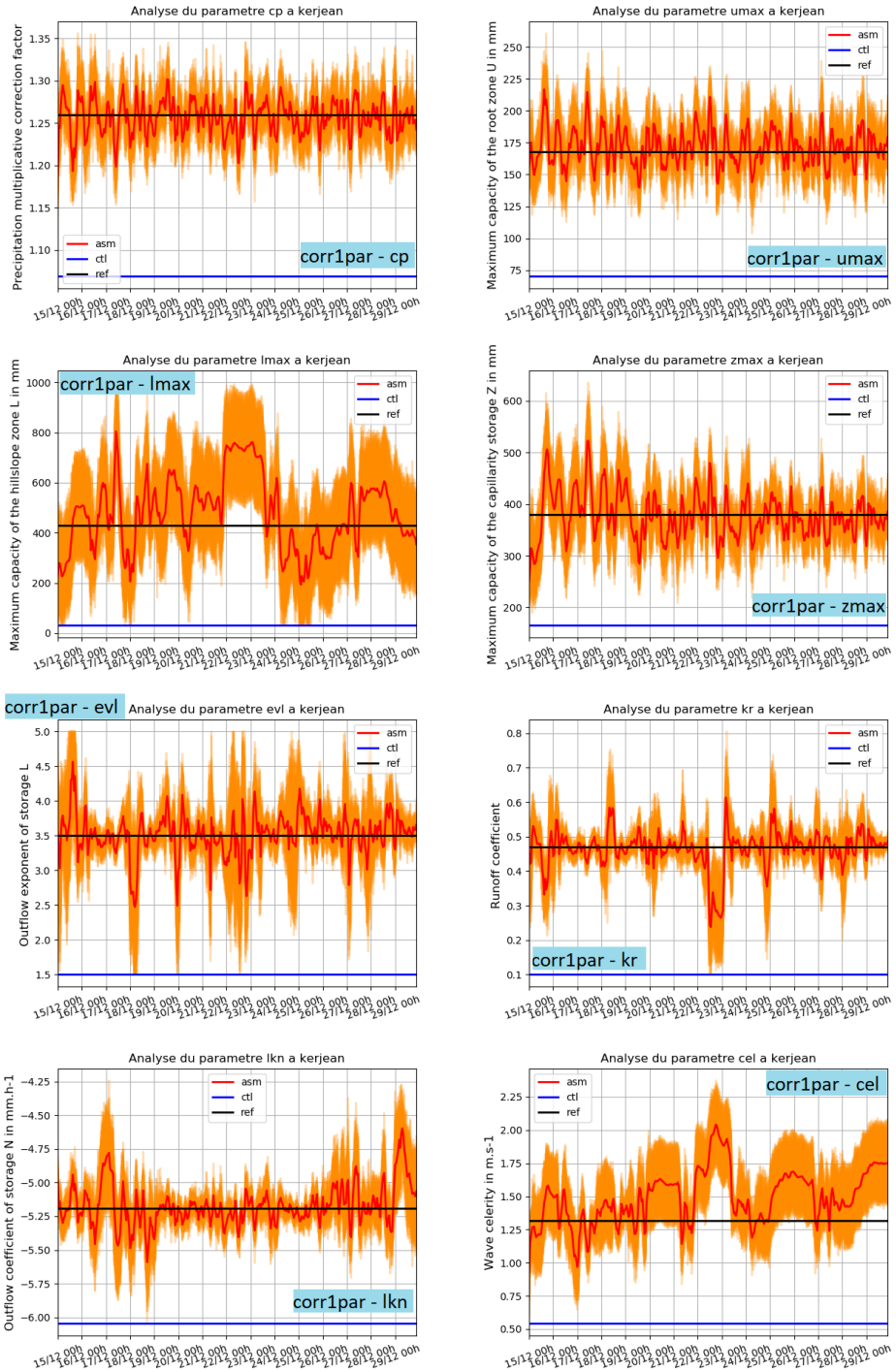


Figure O.6: Valeurs analysées pour le vecteur de contrôle (paramètres) en fonction du temps pour les 8 expériences *corr1par* menées à Kerjean. Les courbes orange représentent les valeurs pour chacun des 49 membres et la courbe rouge la valeur analysée moyenne.

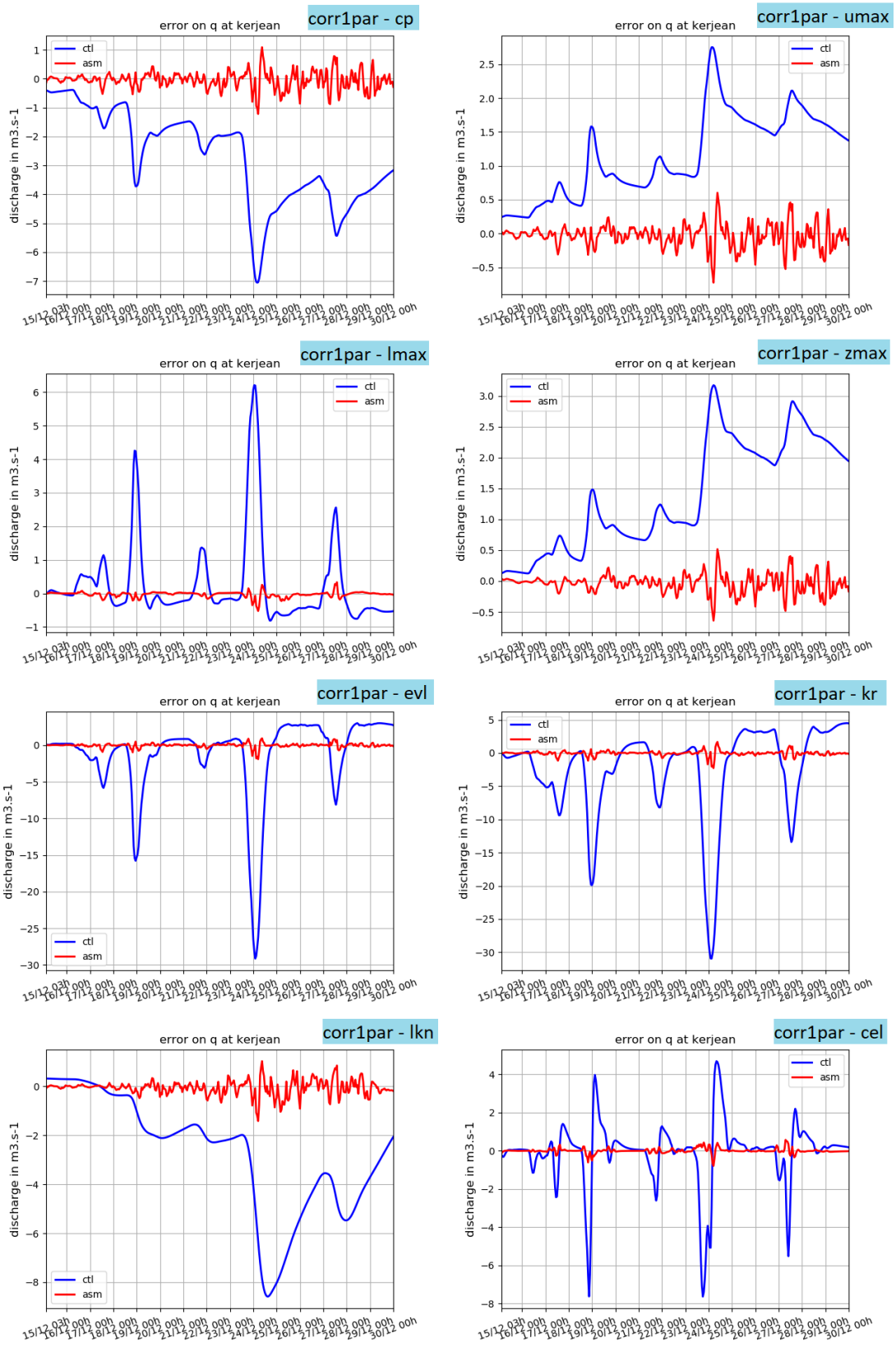


Figure O.7: Chroniques d'erreurs de la simulation de contrôle et de l'analyse par rapport à la simulation de référence pour les 8 expériences *corr1par* menées à Kerjean. La valeur de l'analyse est prise à la fin de la fenêtre d'assimilation.

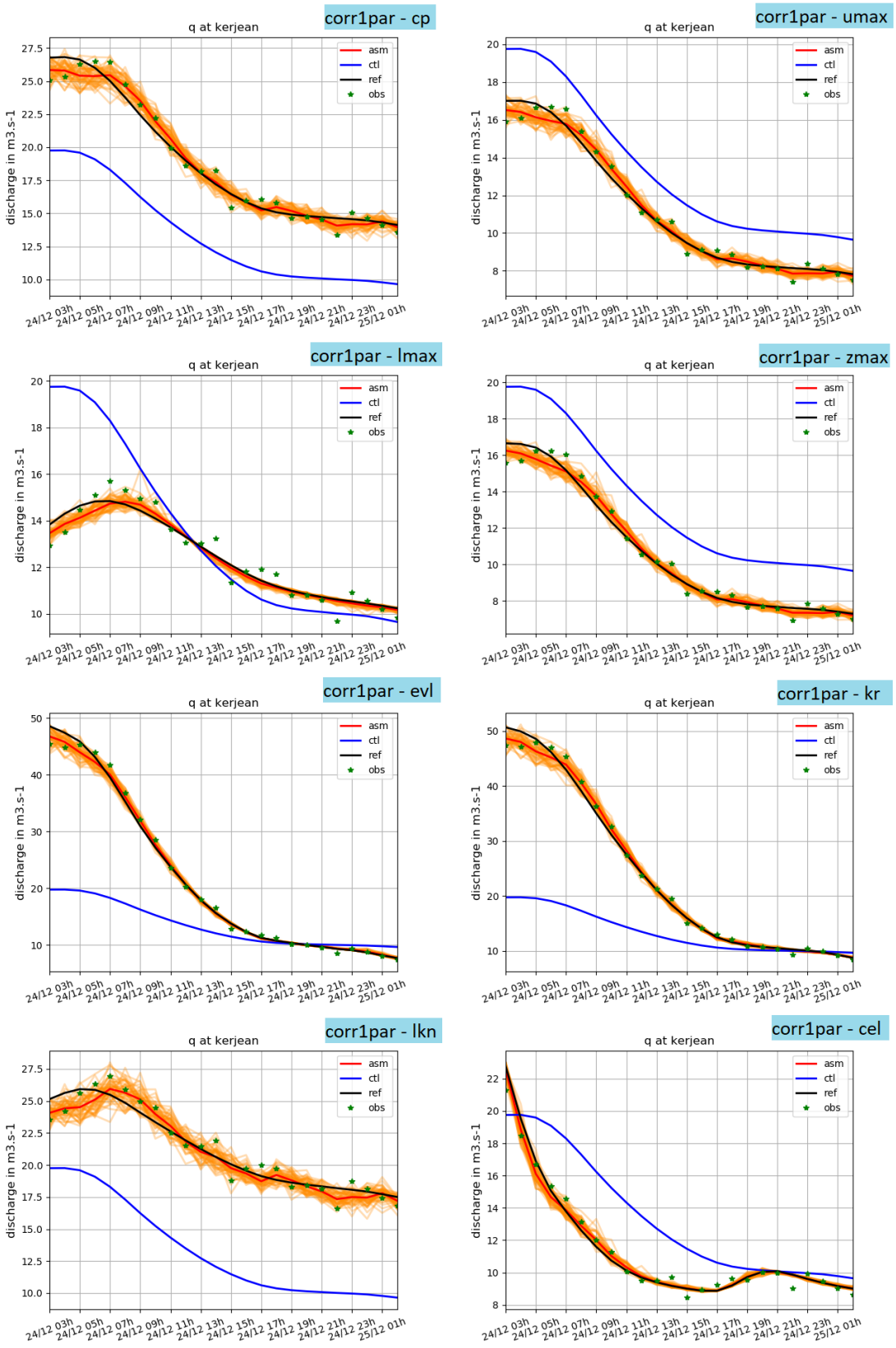


Figure O.8: Simulation de contrôle, simulation de référence, analyses et observations assimilées au moment du pic d'événement le 24 décembre 2013 pour les 8 expériences *corr1par* menées à Kerjean.

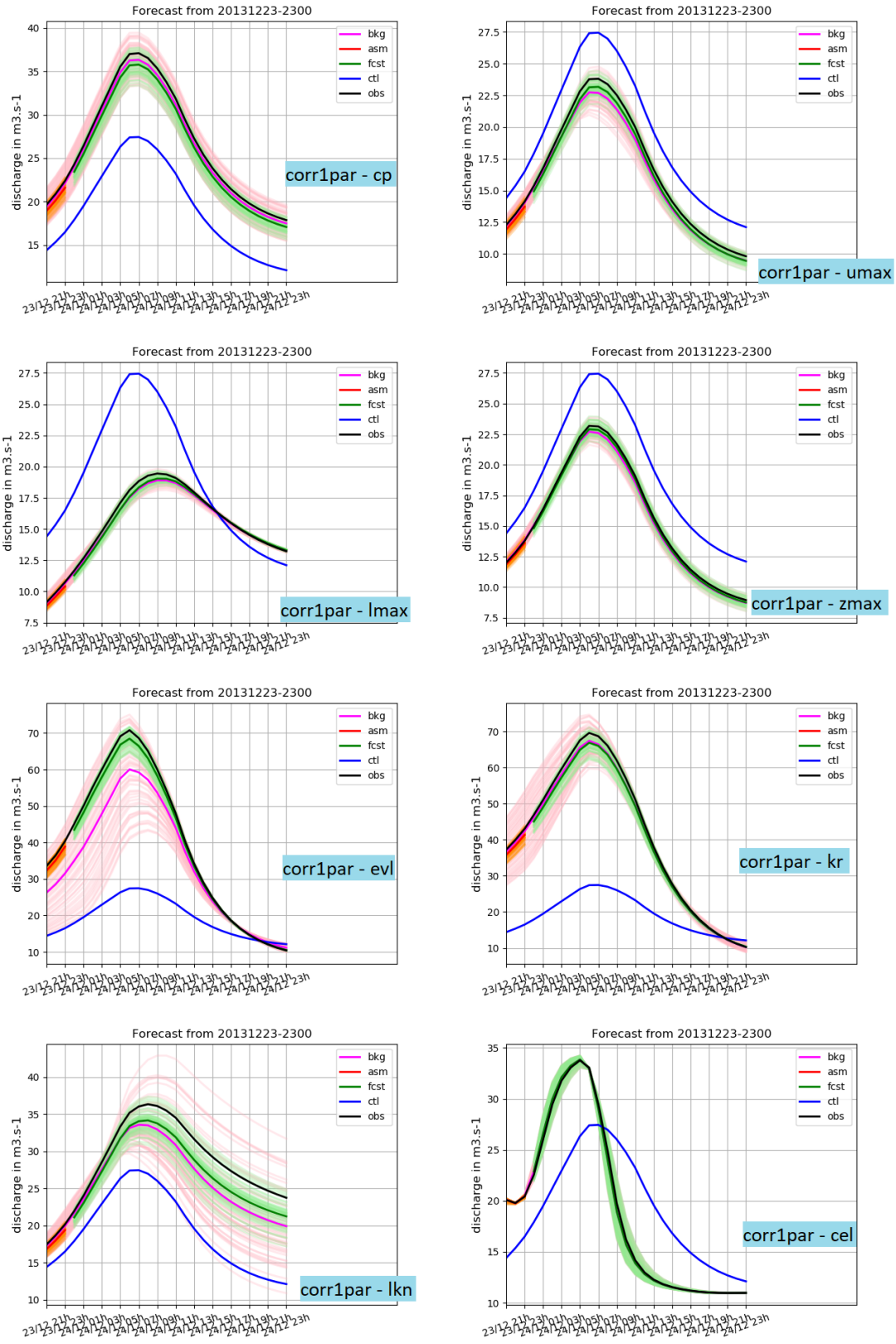


Figure O.9: Visualisation pour 8 expériences *corr1par* menées à Kerjean du cycle d'assimilation de données débutant le 23 décembre 2013 à 20 heures : ensemble d'ébauche (bkg), ensemble analysé (asm) et ensemble en prévision suite à l'analyse (fcst), simulation de contrôle (ctl) et simulation de référence (ref). À partir du temps de base (23 décembre à 23 heures), toutes les simulations sont alimentées par les prévisions de pluie RR3.

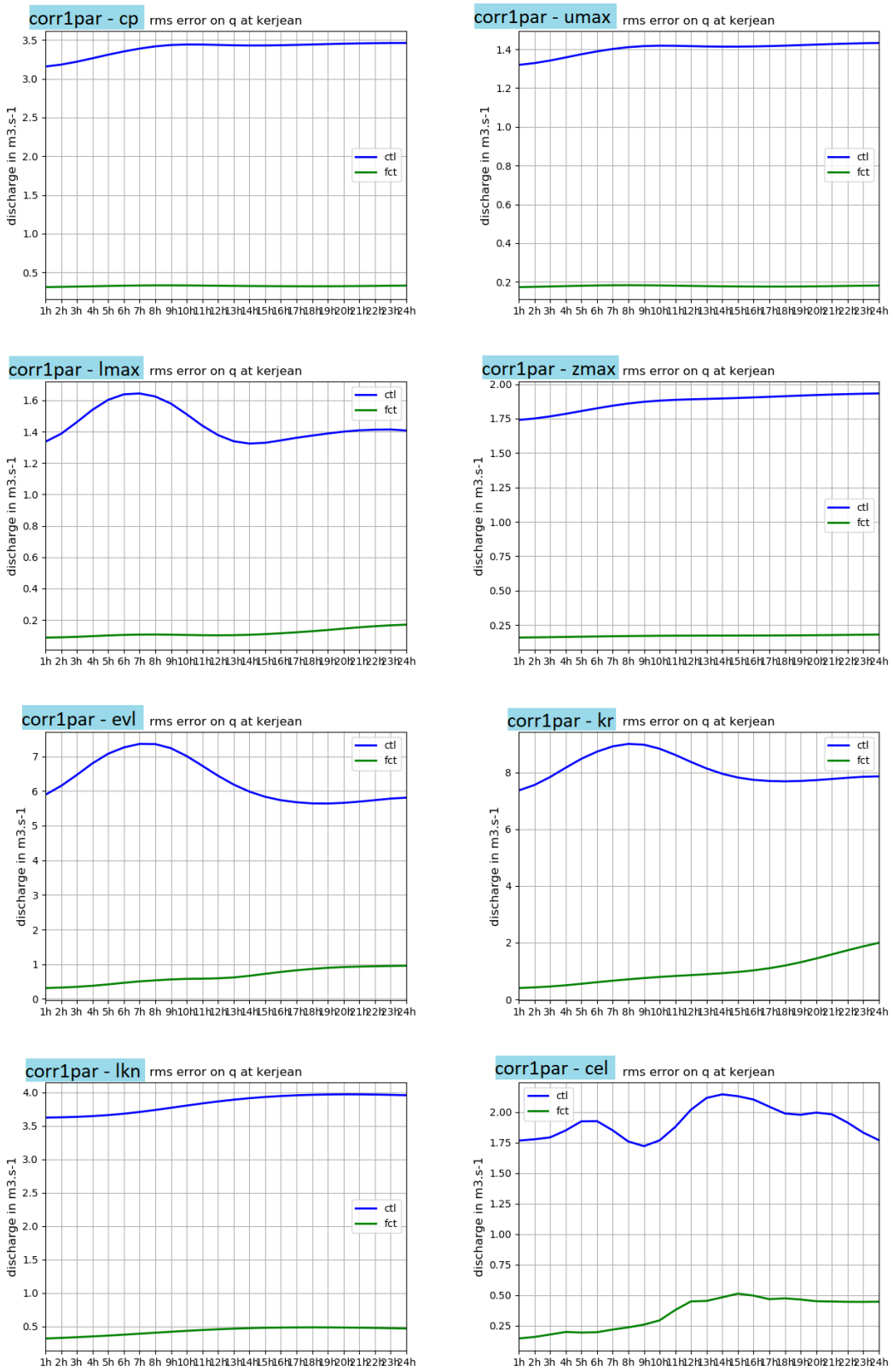


Figure O.10: RMSE des erreurs de prévision en fonction de l'échéance pour la simulation de contrôle (*ctl* en bleu) et le modèle analysé (*fcst* en vert) par rapport à la simulation de référence pour les 8 expériences *corr1par* menées à Kerjean. Toutes les simulations (de référence, de contrôle et le modèle corrigé) sont alimentées par les prévisions de pluie RR3 en prévision.

O.3 Expérience *corr1par* - Résultats à Ty-Planche

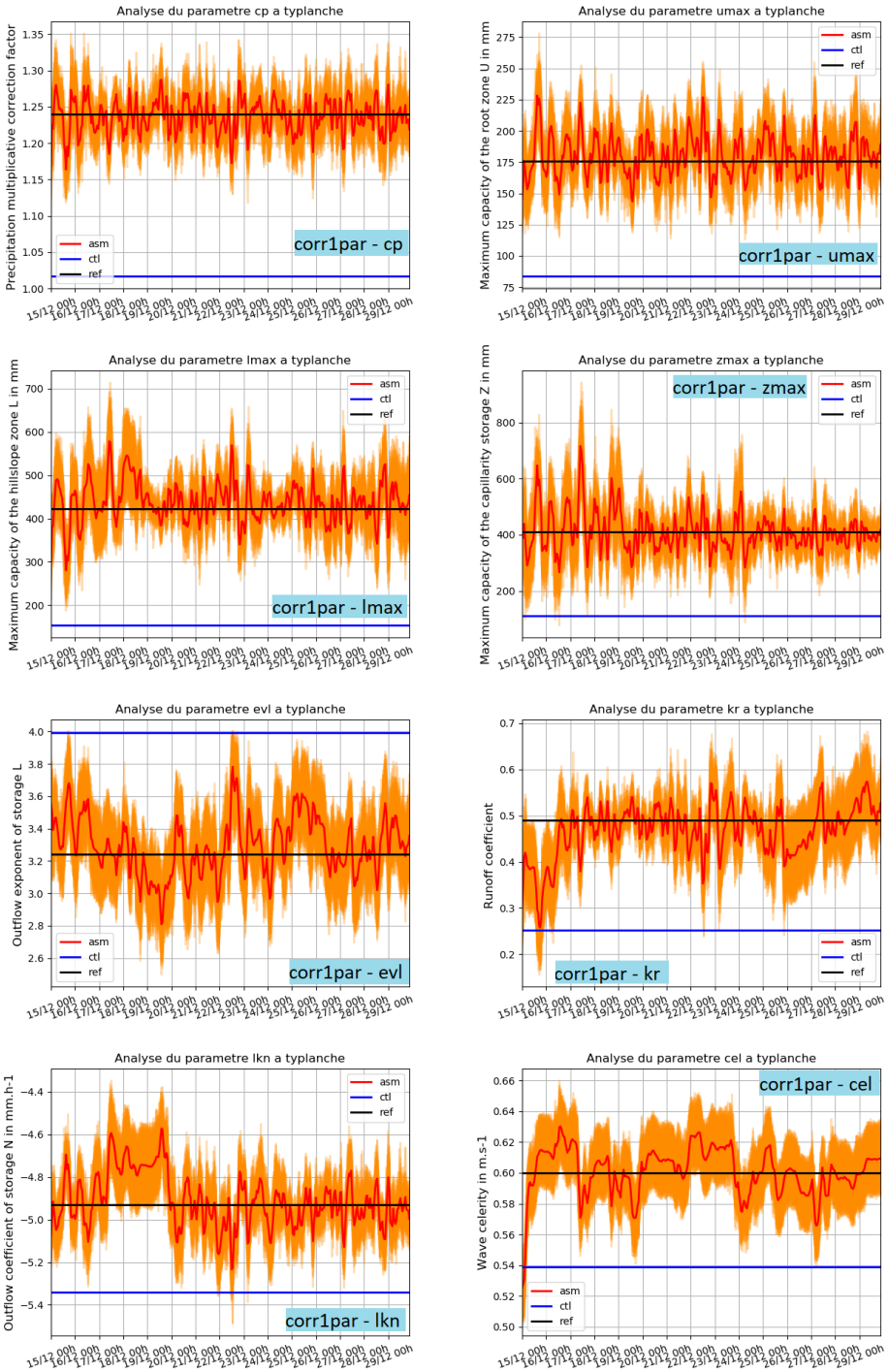


Figure O.11: Valeurs analysées pour le vecteur de contrôle (paramètres) en fonction du temps pour les 8 expériences *corr1par* menées à Ty-Planche. Les courbes orange représentent les valeurs pour chacun des 49 membres et la courbe rouge la valeur analysée moyenne.

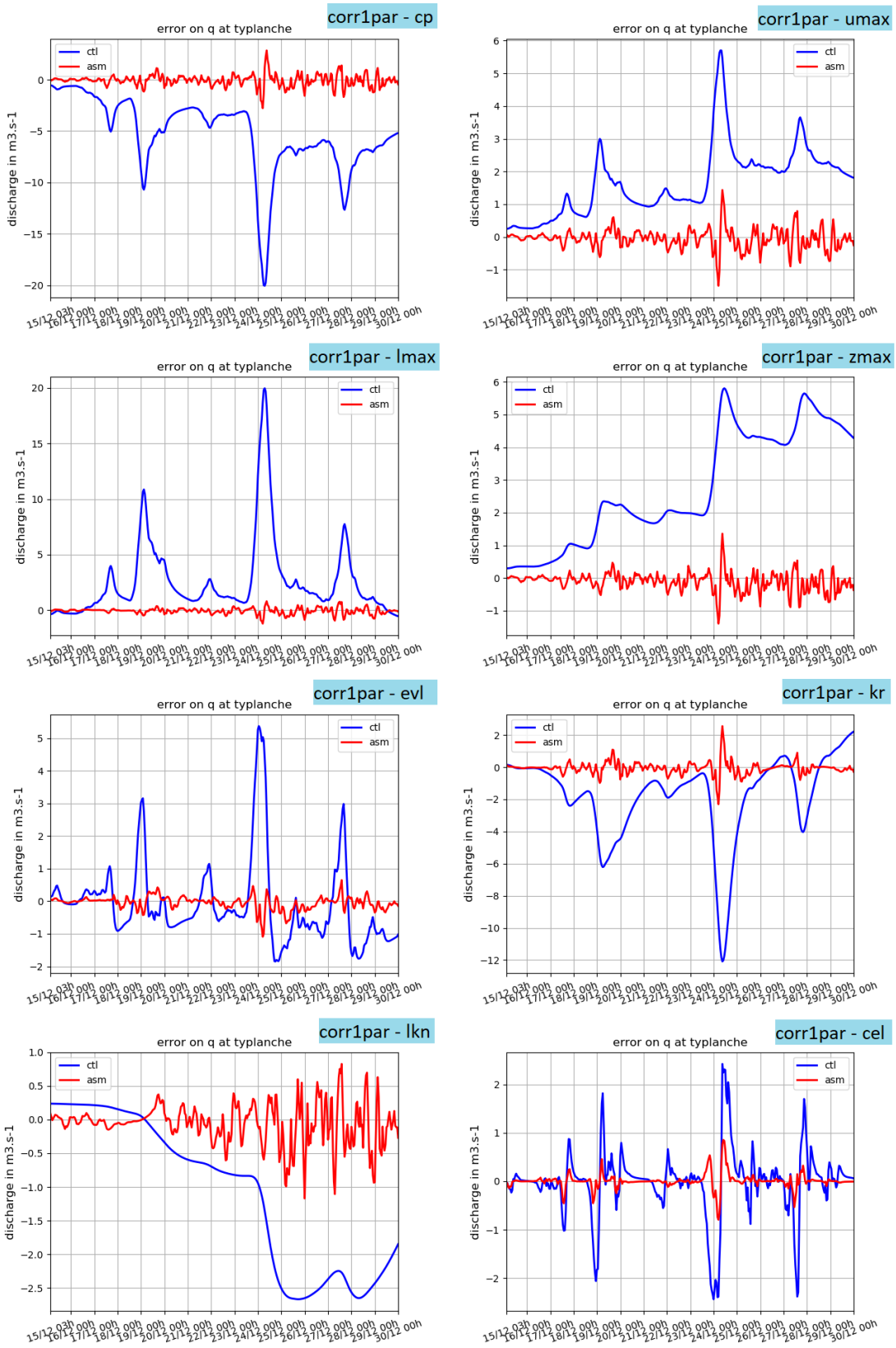


Figure O.12: Chroniques d'erreurs de la simulation de contrôle et de l'analyse par rapport à la simulation de référence pour les 8 expériences *corr1par* menées à Ty-Planche. La valeur de l'analyse est prise à la fin de la fenêtre d'assimilation.

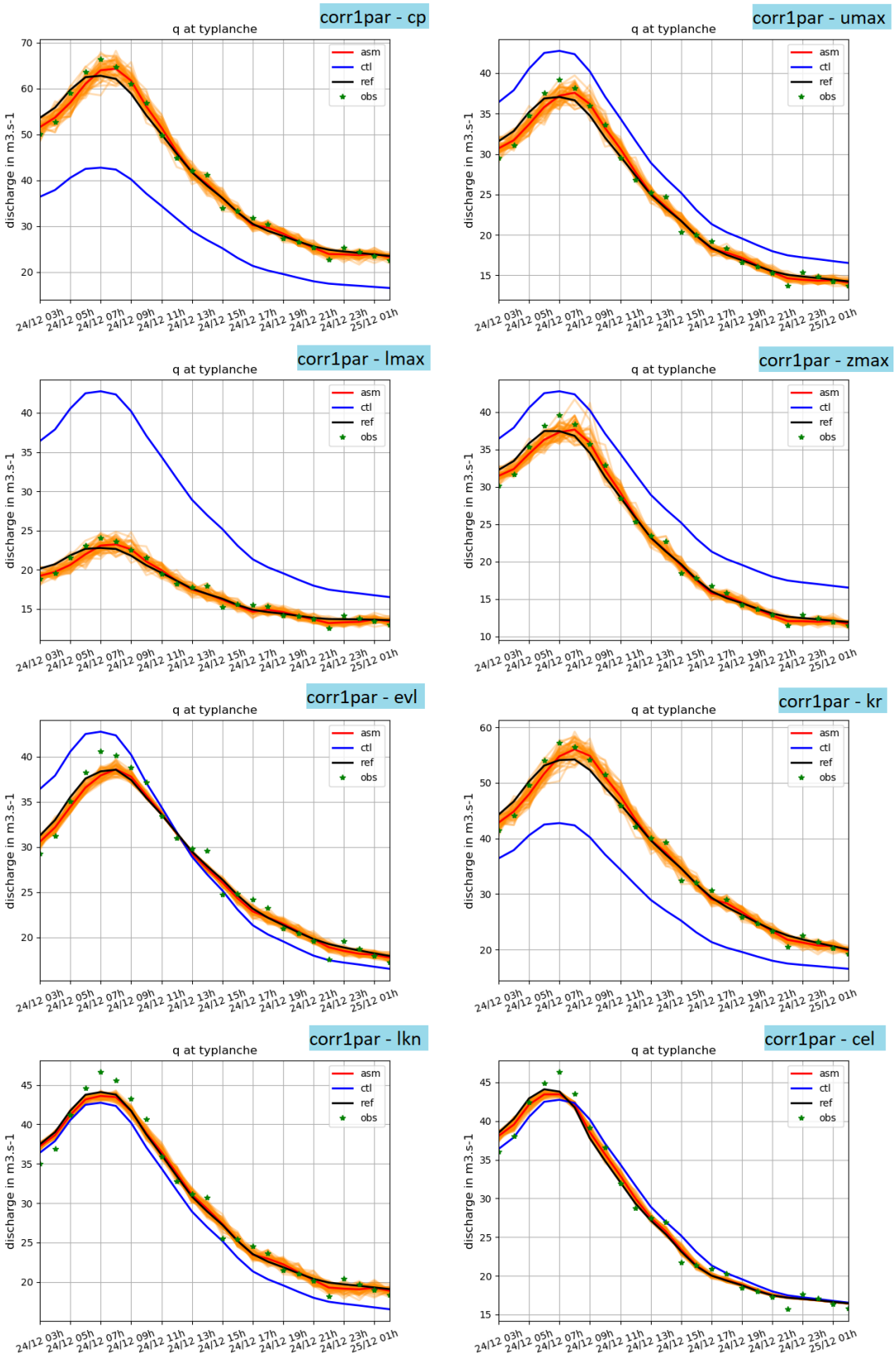


Figure O.13: Simulation de contrôle, simulation de référence, analyses et observations assimilées au moment du pic d'événement le 24 décembre 2013 pour les 8 expériences *corr1par* menées à Ty-Planche.

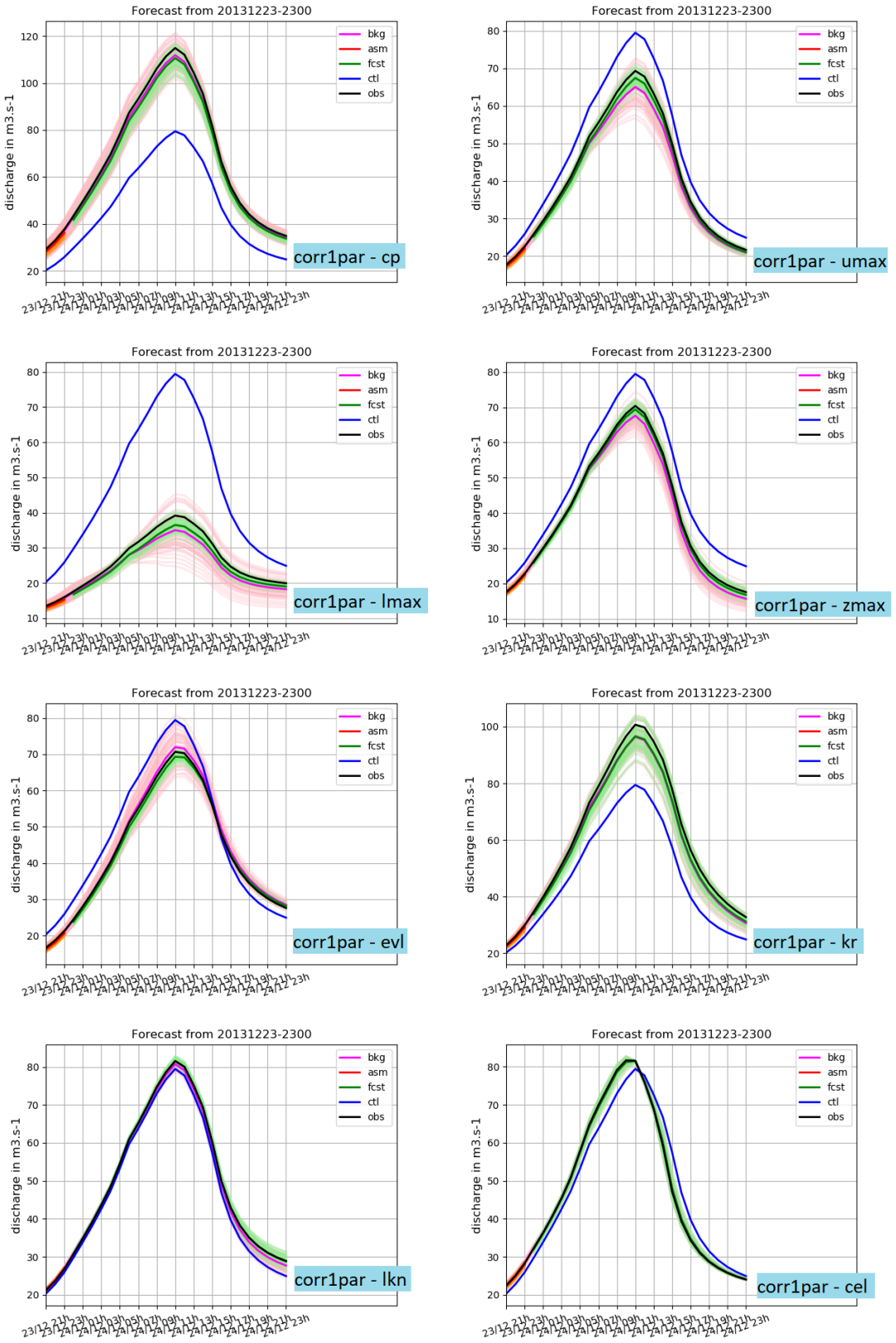


Figure O.14: Visualisation pour 8 expériences *corr1par* menées à Ty-Planche du cycle d'assimilation de données débutant le 23 décembre 2013 à 20 heures : ensemble d'ébauche (bkg), ensemble analysé (asm) et ensemble en prévision suite à l'analyse (fcst), simulation de contrôle (ctl) et simulation de référence (ref). À partir du temps de base (23 décembre à 23 heures), toutes les simulations sont alimentées par les prévisions de pluie RR3.

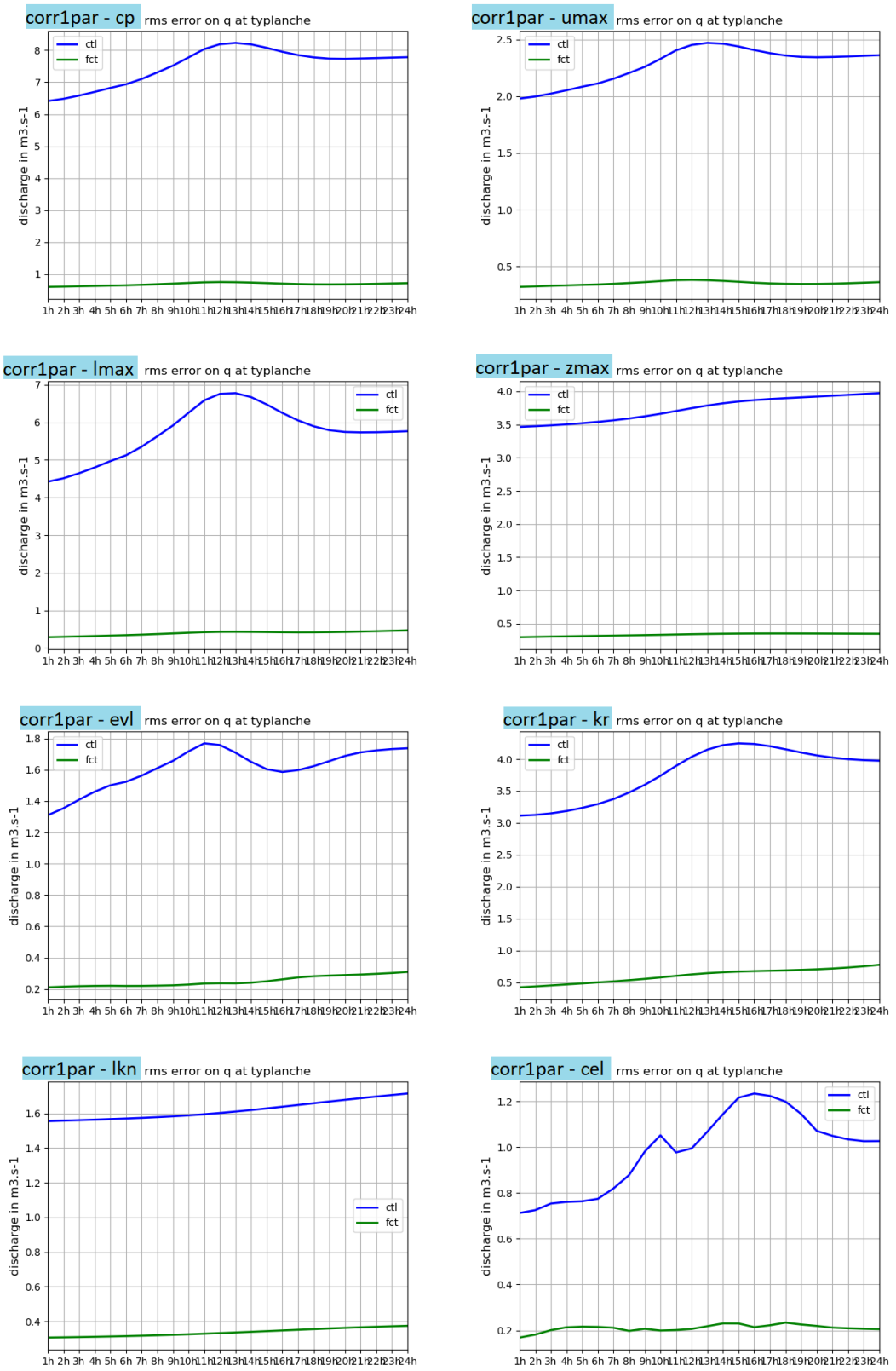


Figure O.15: RMSE des erreurs de prévision en fonction de l'échéance pour la simulation de contrôle (*ctl* en bleu) et le modèle analysé (*fcst* en vert) par rapport à la simulation de référence pour les 8 expériences *corr1par* menées à Ty-Planche. Toutes les simulations (de référence, de contrôle et le modèle corrigé) sont alimentées par les prévisions de pluie RR3 en prévision.

O.4 Expérience *corr8par* - Résultats à Kerjean

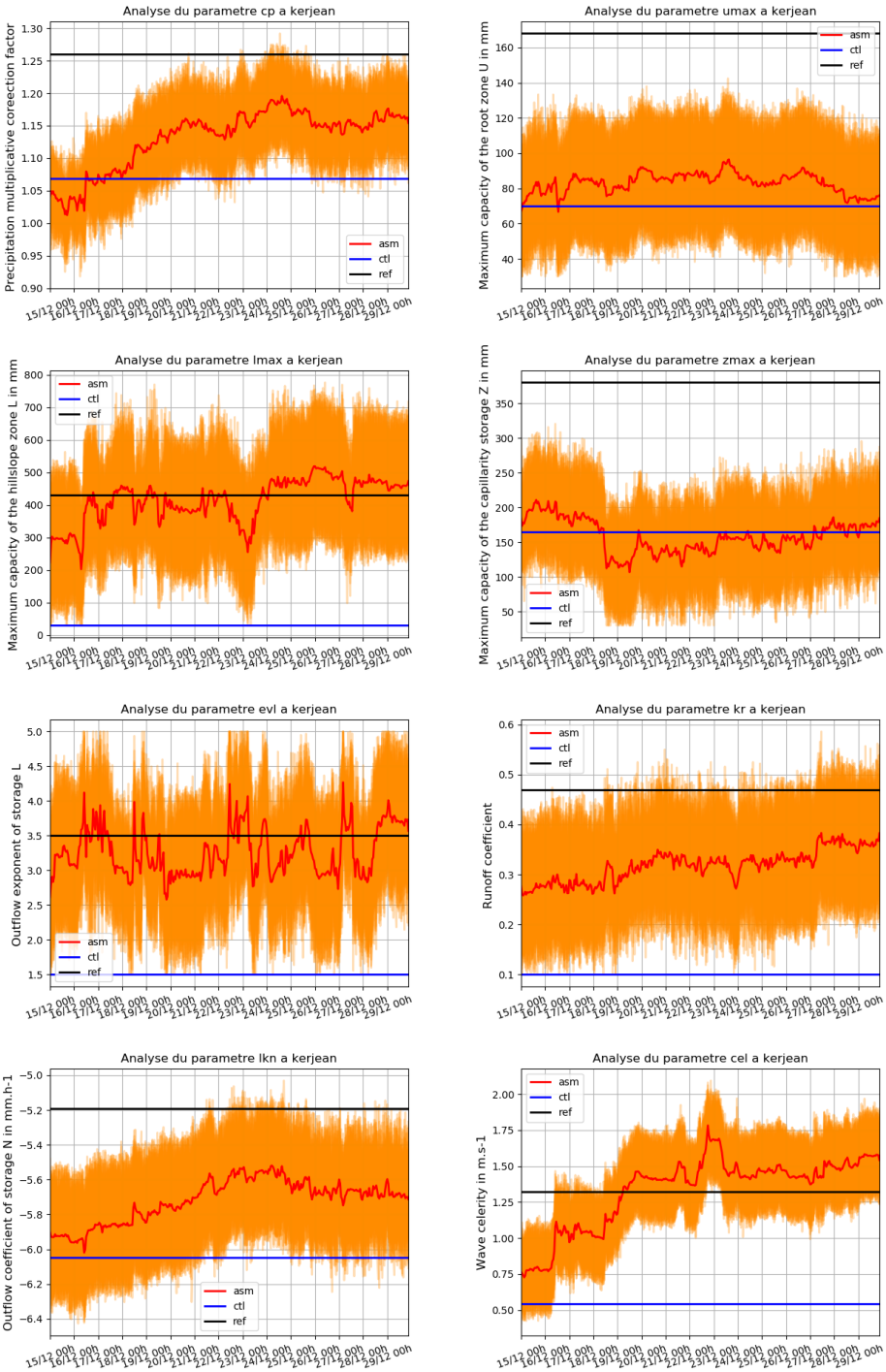


Figure O.16: Valeurs analysées pour le vecteur de contrôle (paramètres) en fonction du temps pour l'expérience *corr8par* menée à Kerjean.

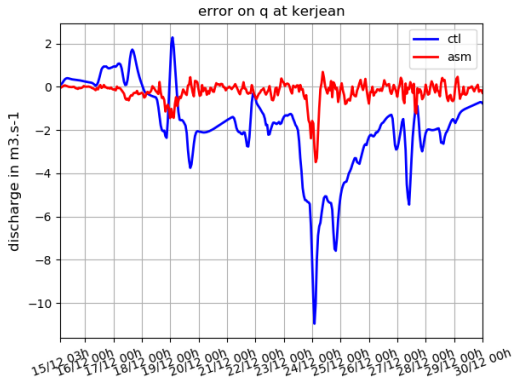


Figure O.17: Chronique temporelle des erreurs à Kerjean pour la simulation de contrôle (bleu) et pour l'analyse par rapport à la simulation de référence - Expérience *corr8par*.

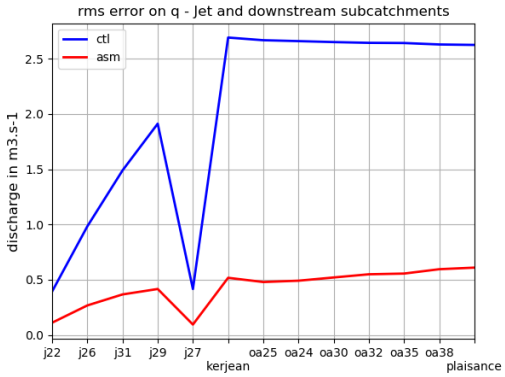


Figure O.18: RMSE sur les mailles du Jet de l'amont vers l'aval pour le contrôle (bleu) et de l'analyse pour les trois expériences par rapport à la simulation de référence.

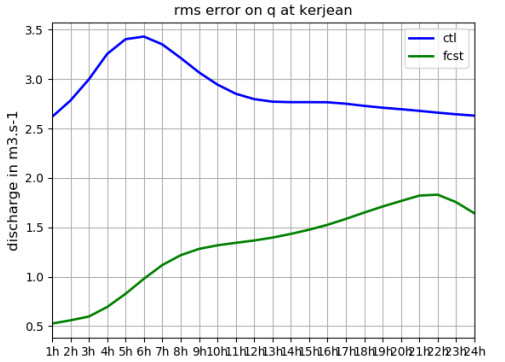


Figure O.19: RMSE des erreurs de prévision à Kerjean en fonction de l'échéance pour la simulation de contrôle (bleu) et le modèle analysé par rapport à la simulation de référence.

O.5 Expérience *corr8par* - Résultats à Ty-Planche

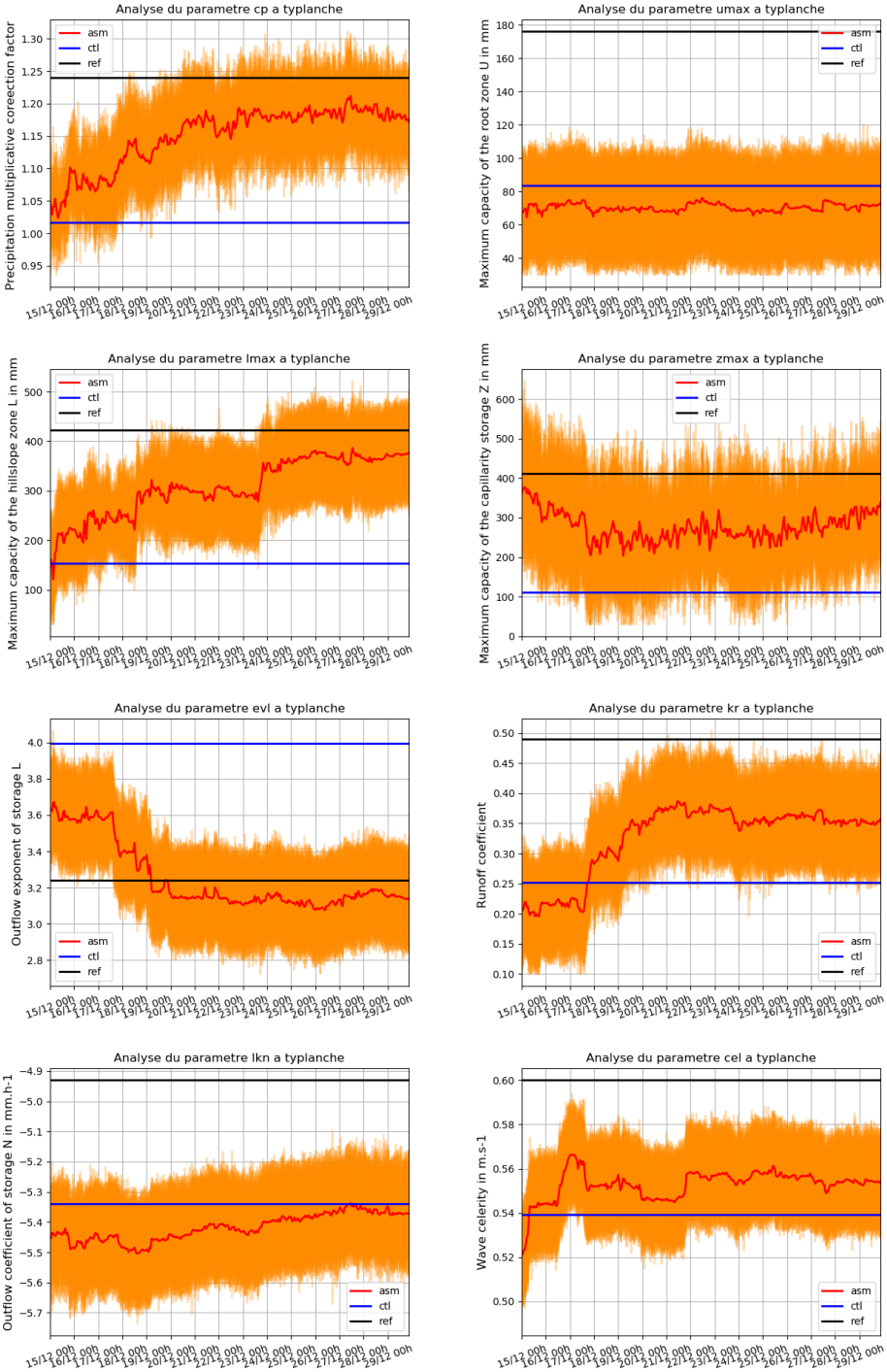


Figure O.20: Valeurs analysées pour le vecteur de contrôle (paramètres) en fonction du temps pour l'expérience *corr8par* menée à Ty-Planche.

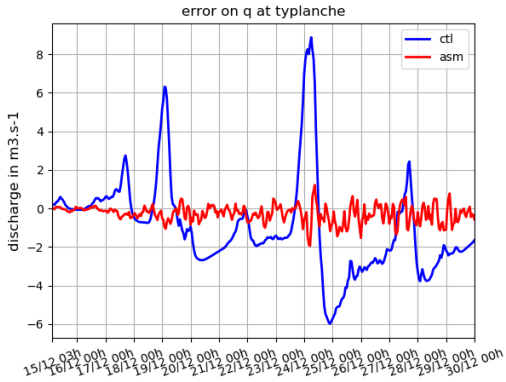


Figure O.21: Chronique temporelle des erreurs à Ty-Planche pour la simulation de contrôle (bleu) et pour l'analyse par rapport à la simulation de référence - Expérience *corr8par*.

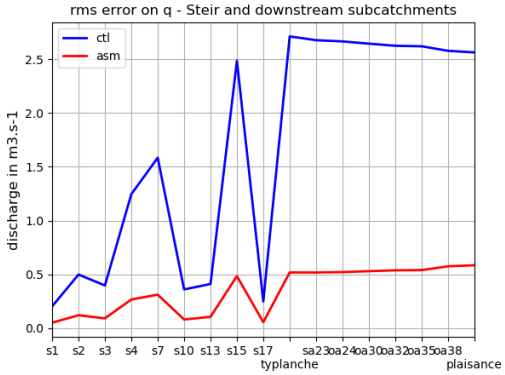


Figure O.22: RMSE sur les mailles du Steïr de l'amont vers l'aval pour le contrôle (bleu) et de l'analyse pour les trois expériences par rapport à la simulation de référence.

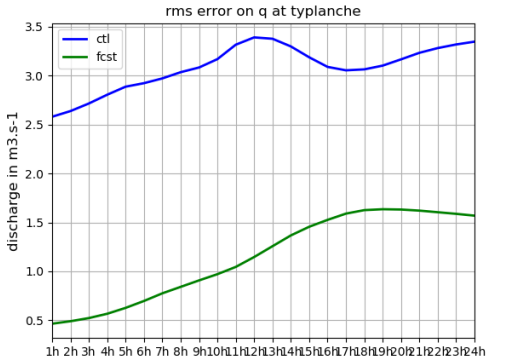


Figure O.23: RMSE des erreurs de prévision à Ty-Planche en fonction de l'échéance pour la simulation de contrôle (bleu) et le modèle analysé par rapport à la simulation de référence.

Annexe P

Compléments au chapitre 6

Sommaire

P.1 Ensembles à Tréodet 378
P.1.1 Graphiques sur les événements 378
P.2 Ensembles à Kerjean 380
P.2.1 Scores par échéance 380
P.2.2 Histogrammes de rang 381
P.2.3 Graphiques sur les événements 383
P.3 Ensembles à Ty-Planche 385
P.3.1 Scores par échéance 385
P.3.2 Histogrammes de rang 386
P.3.3 Graphiques sur les événements 388

P.1 Ensembles à Tréodet

P.1.1 Graphiques sur les événements

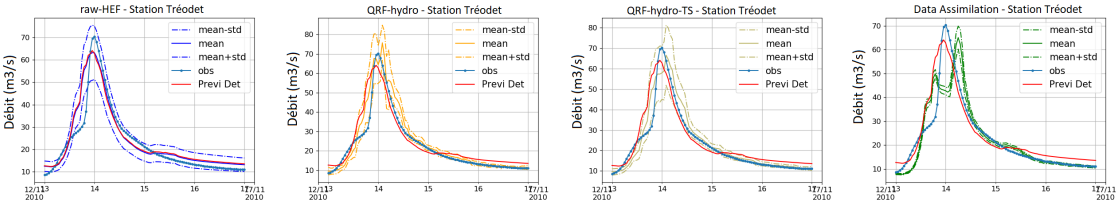


Figure P.1: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 6 heures à Tréodet. Événement du 13 au 17 novembre 2010.

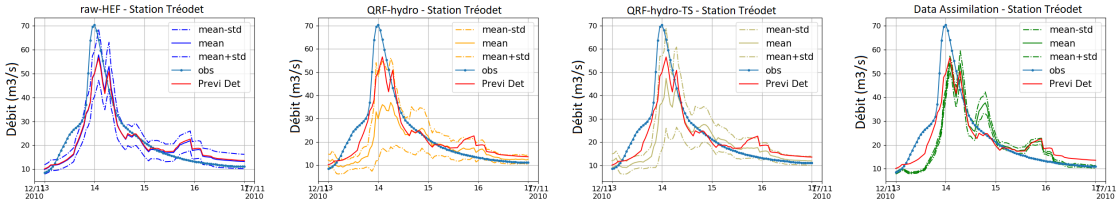


Figure P.2: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 18 heures à Tréodet. Événement du 13 au 17 novembre 2010.

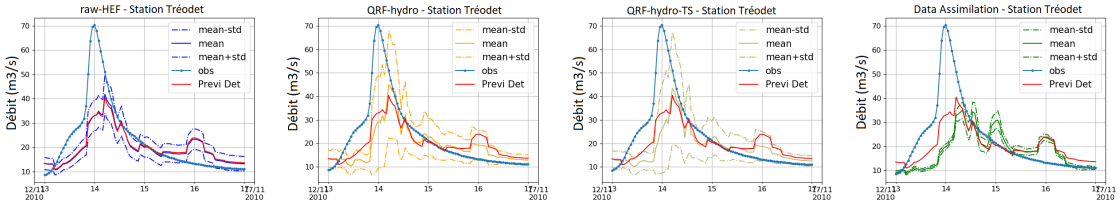


Figure P.3: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 24 heures à Tréodet. Événement du 13 au 17 novembre 2010.

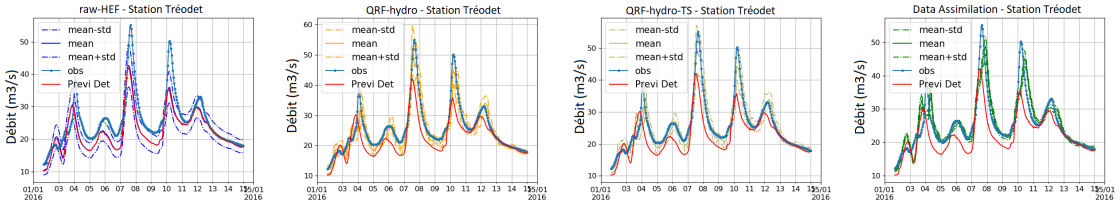


Figure P.4: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 6 heures à Tréodet. Événement du 2 au 15 janvier 2016.

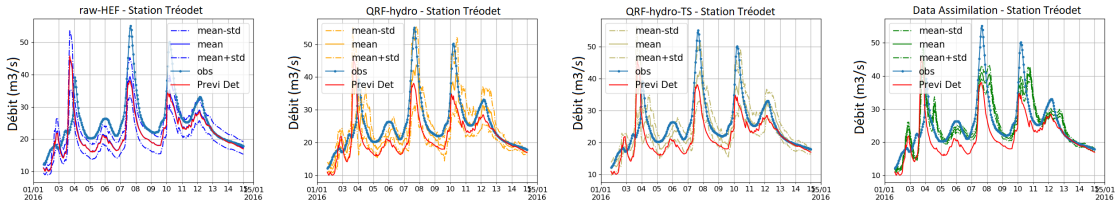


Figure P.5: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 12 heures à Tréodet. Événement du 2 au 15 janvier 2016.

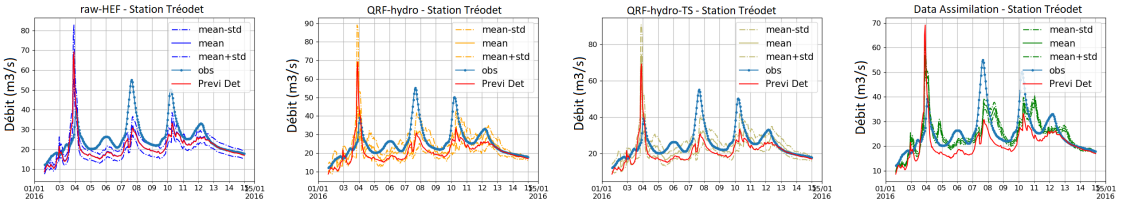


Figure P.6: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 18 heures à Tréodet. Événement du 2 au 15 janvier 2016.

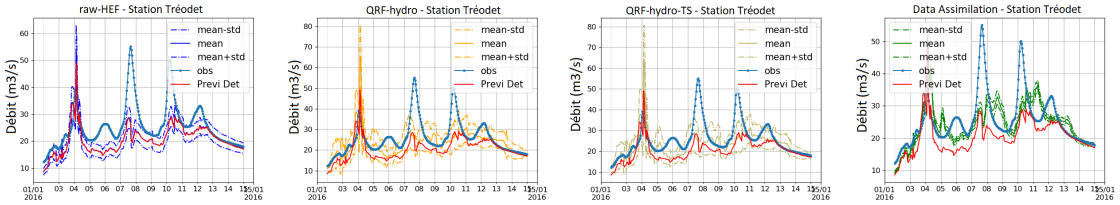


Figure P.7: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 24 heures à Tréodet. Événement du 2 au 15 janvier 2016.

P.2 Ensembles à Kerjean

P.2.1 Scores par échéance

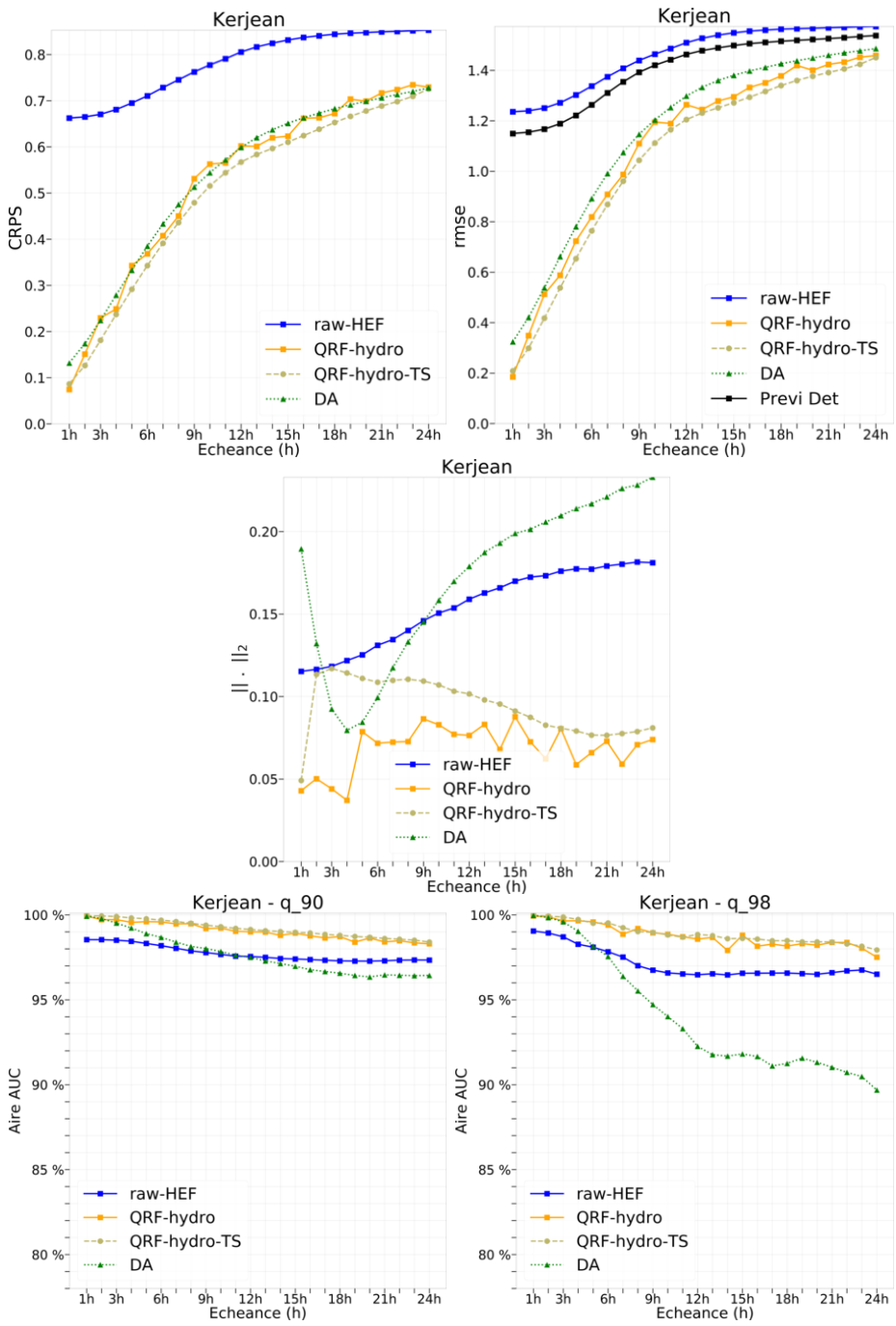
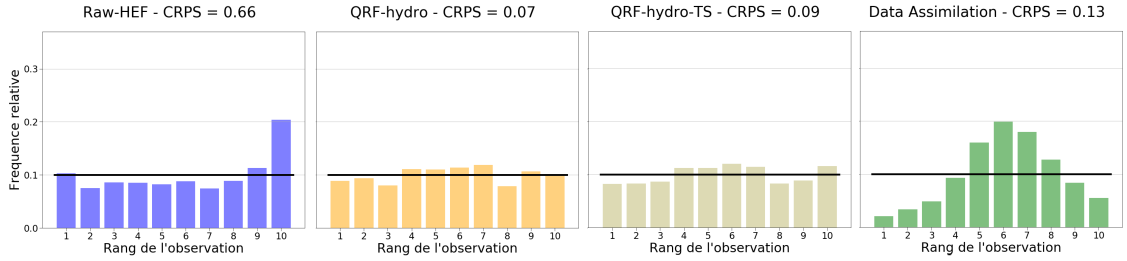
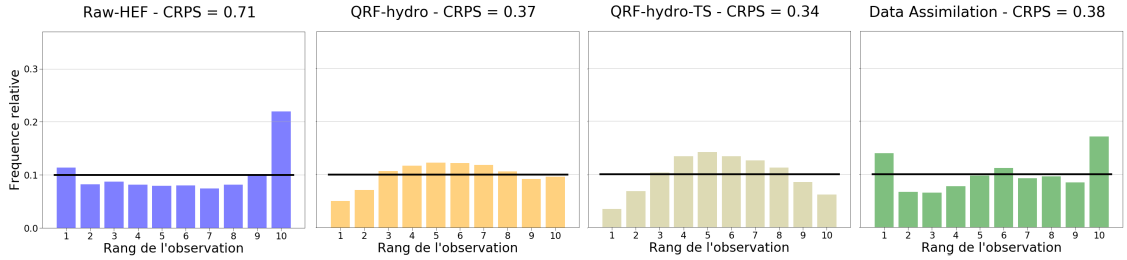


Figure P.8: CRPS, RMSE, $||.||_2$ et aire ROC *AUC* pour les événements de dépassement des quantiles q_{90} et q_{98} pour les prévisions d'ensemble *raw-HEF*, *QRF-hydro*, *QRF-hydro-TS* et *Data assimilation* pour les échéances de 1 heure à 24 heures à Kerjean.

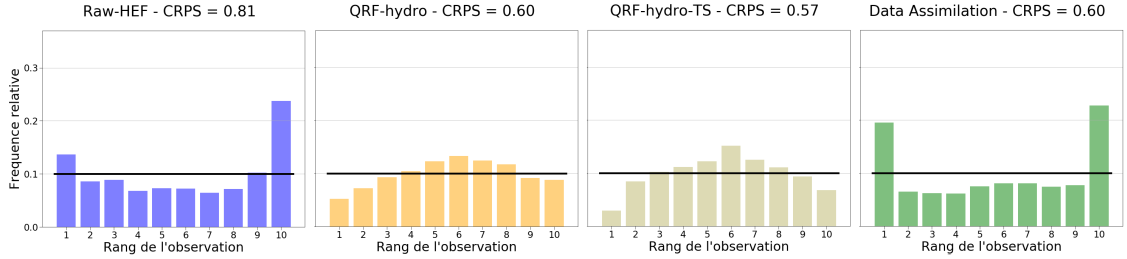
P.2.2 Histogrammes de rang



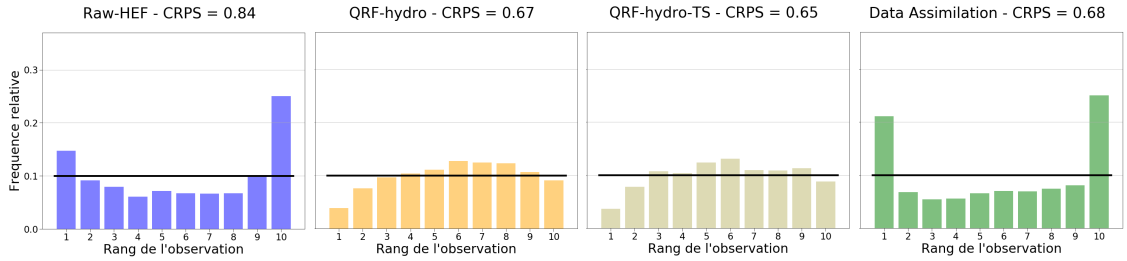
(a) Echéance 1 heure.



(b) Echéance 6 heures.

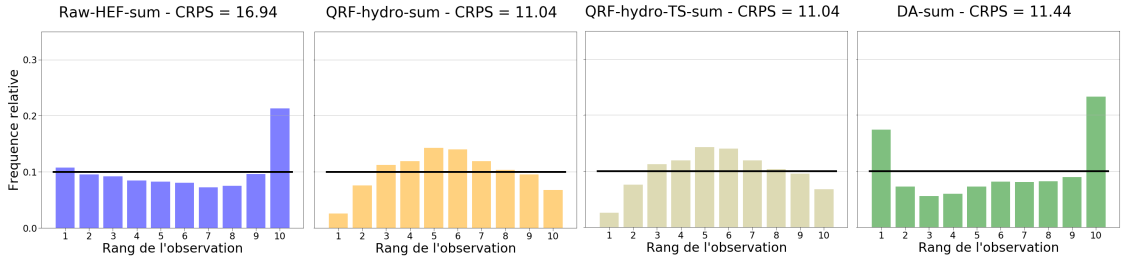


(c) Echéance 12 heures.

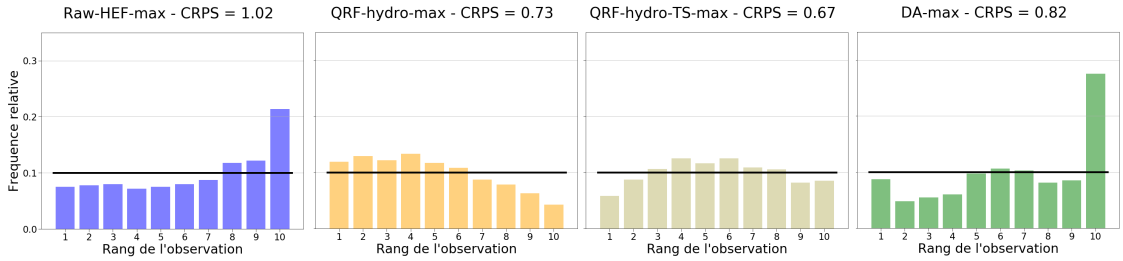


(d) Echéance 18 heures.

Figure P.9: Histogrammes de rang des prévisions pour les ensembles *raw*, *QRF-hydro*, *QRF-hydro-TS* et *Data assimilation*, pour les échéances 1 heure, 6 heures, 12 heures et 18 heures à Kerjean.



(a) Somme de la prévision



(b) Maximum de la prévision

Figure P.10: Histogrammes de rang pour la somme et le maximum des prévisions pour les ensembles *raw*, *QRF-hydro*, *QRF-hydro-TS* et *Data Assimilation* à Kerjean.

P.2.3 Graphiques sur les événements

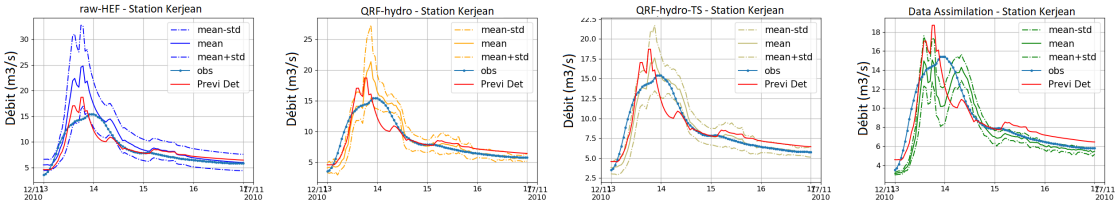


Figure P.11: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 6 heures à Kerjean. Événement du 13 au 17 novembre 2010.

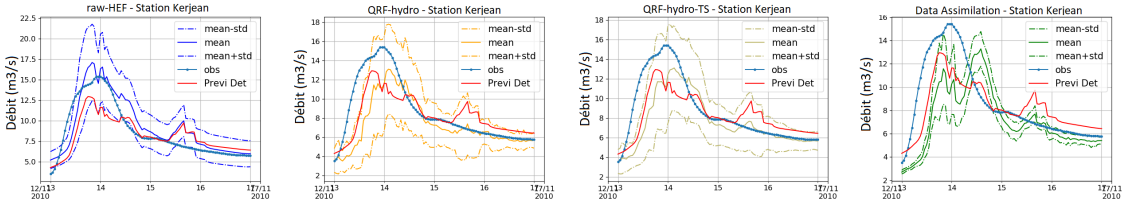


Figure P.12: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 12 heures à Kerjean. Événement du 13 au 17 novembre 2010.

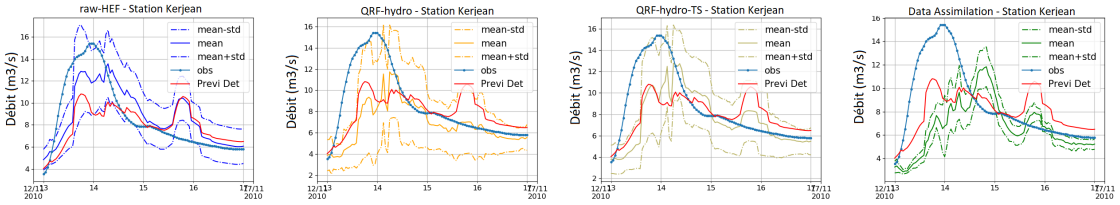


Figure P.13: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 18 heures à Kerjean. Événement du 13 au 17 novembre 2010.

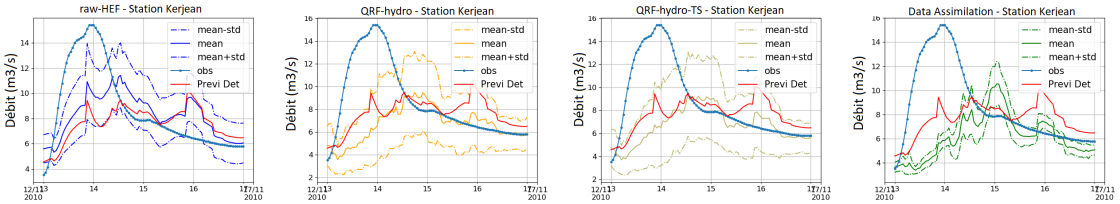


Figure P.14: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 24 heures à Kerjean. Événement du 13 au 17 novembre 2010.

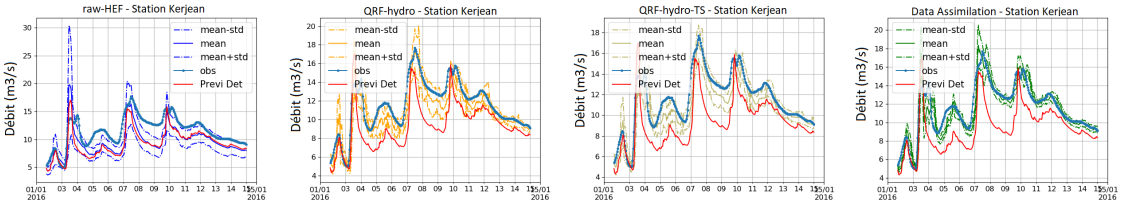


Figure P.15: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 6 heures à Kerjean. Événement du 2 au 15 janvier 2016.

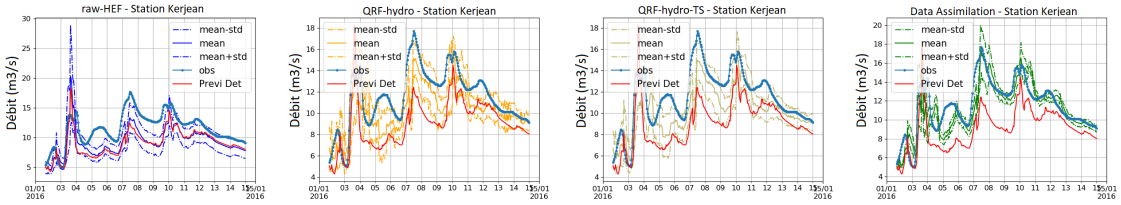


Figure P.16: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 12 heures à Kerjean. Événement du 2 au 15 janvier 2016.

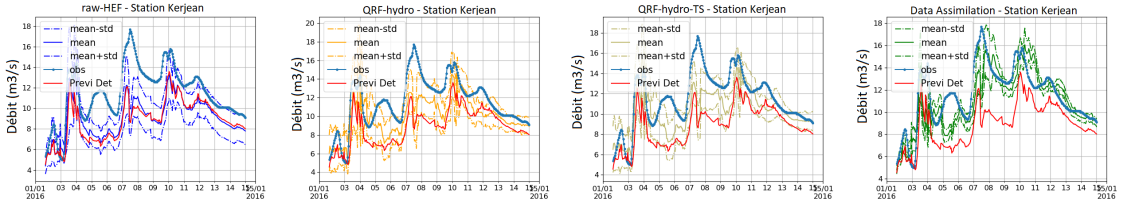


Figure P.17: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 18 heures à Kerjean. Événement du 2 au 15 janvier 2016.

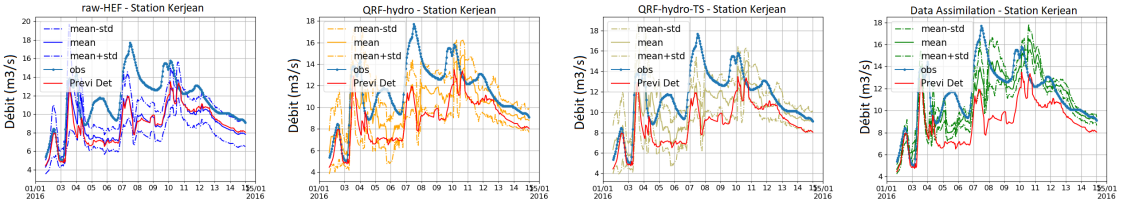


Figure P.18: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 24 heures à Kerjean. Événement du 2 au 15 janvier 2016.

P.3 Ensembles à Ty-Planche

P.3.1 Scores par échéance

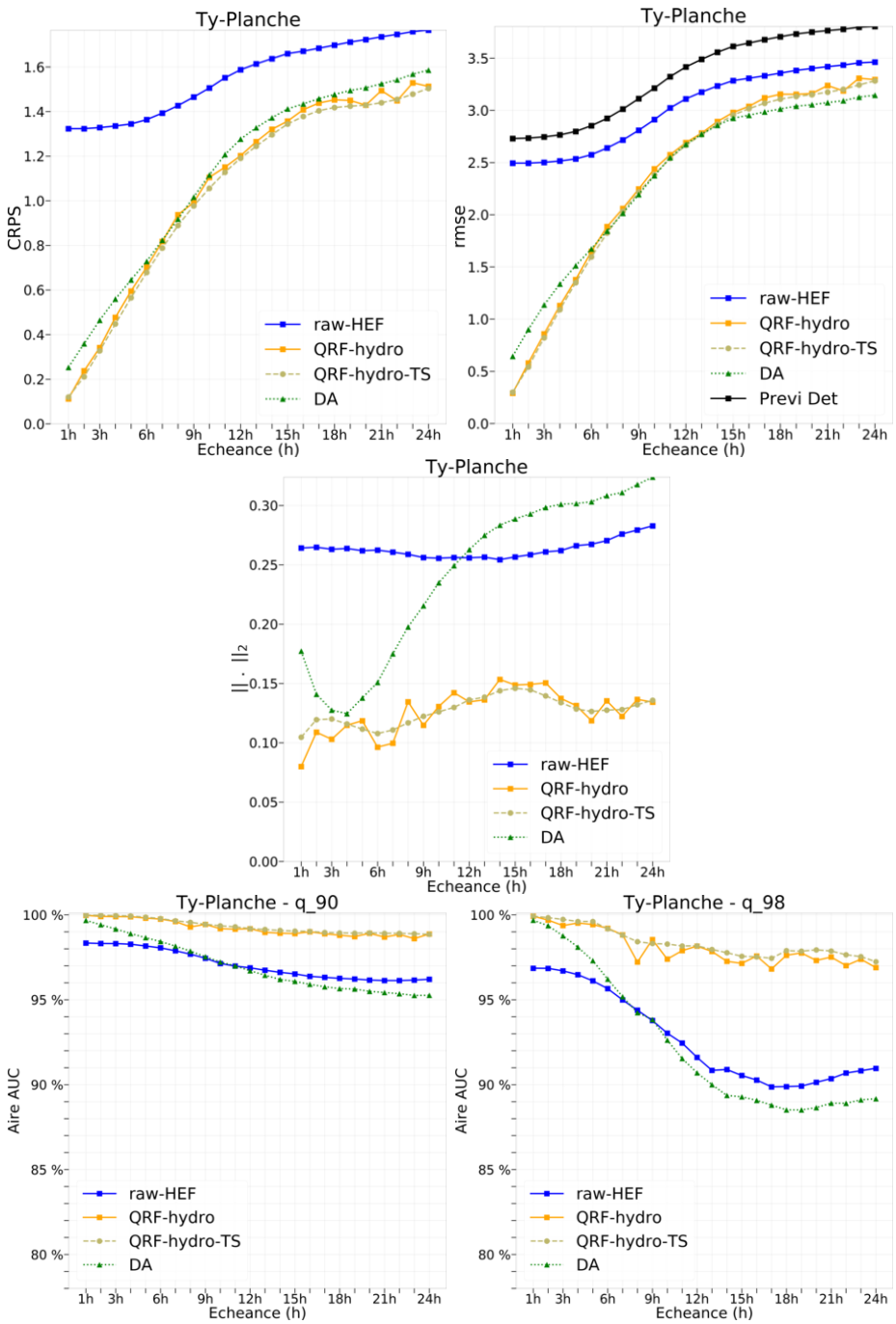
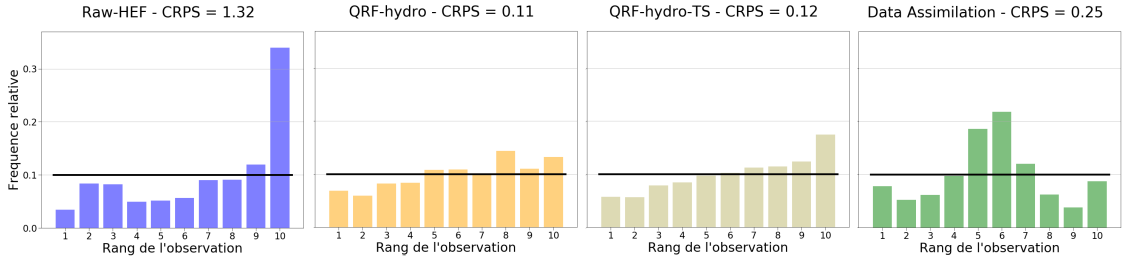
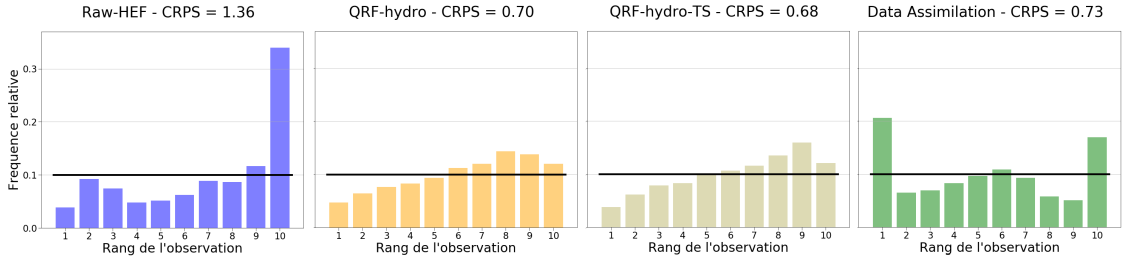


Figure P.19: CRPS, RMSE, $||.||_2$ et aire ROC *AUC* pour les événements de dépassement des quantiles q_{90} et q_{98} pour les prévisions d'ensemble *raw-HEF*, *QRF-hydro*, *QRF-hydro-TS* et *Data assimilation* pour les échéances de 1 heure à 24 heures à Ty-Planche.

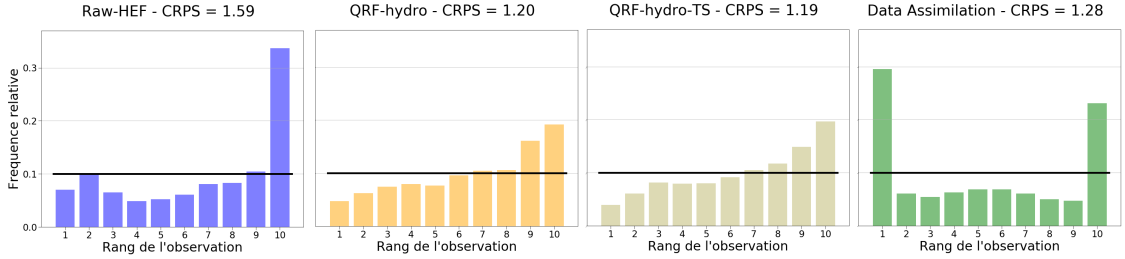
P.3.2 Histogrammes de rang



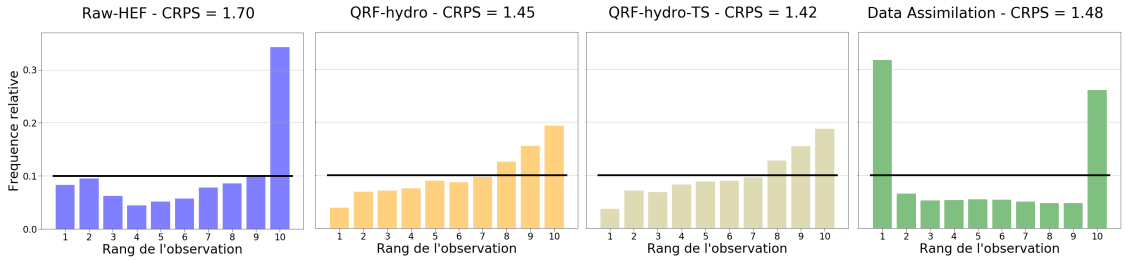
(a) Echéance 1 heure.



(b) Echéance 6 heures.

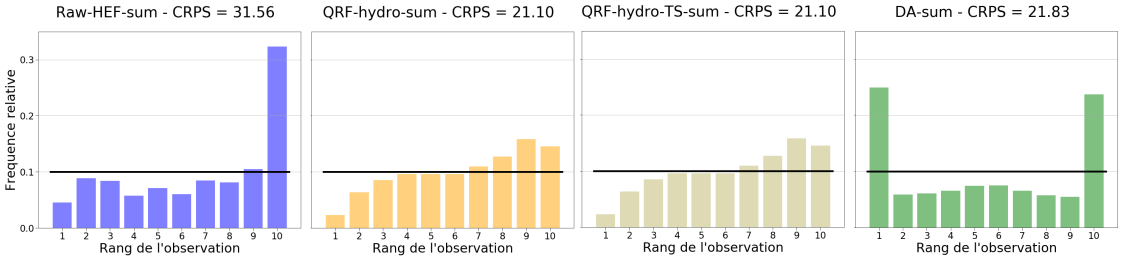


(c) Echéance 12 heures.

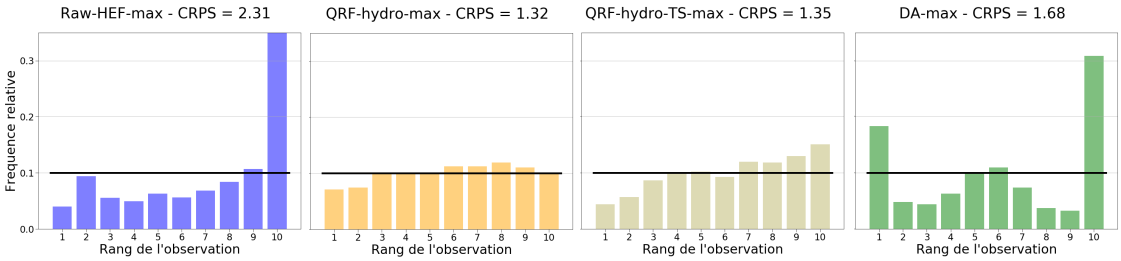


(d) Echéance 18 heures.

Figure P.20: Histogrammes de rang des prévisions pour les ensembles *raw*, *QRF-hydro*, *QRF-hydro-TS* et *Data assimilation*, pour les échéances 1 heure, 6 heures, 12 heures et 18 heures à Ty-Planche.



(a) Somme de la prévision



(b) Maximum de la prévision

Figure P.21: Histogrammes de rang pour la somme et le maximum des prévisions pour les ensembles *raw*, *QRF-hydro*, *QRF-hydro-TS* et *Data Assimilation* à Ty-Planche.

P.3.3 Graphiques sur les événements

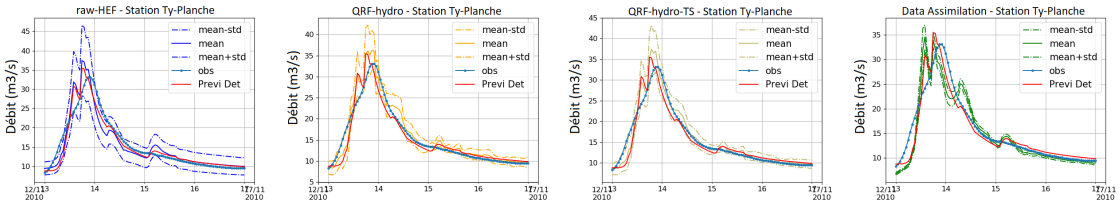


Figure P.22: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 6 heures à Ty-Planche. Événement du 13 au 17 novembre 2010.

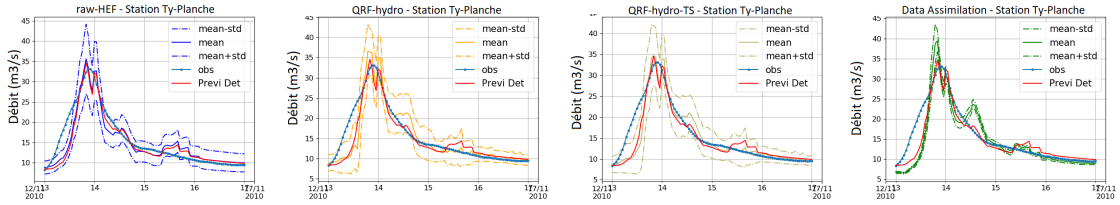


Figure P.23: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 12 heures à Ty-Planche. Événement du 13 au 17 novembre 2010.

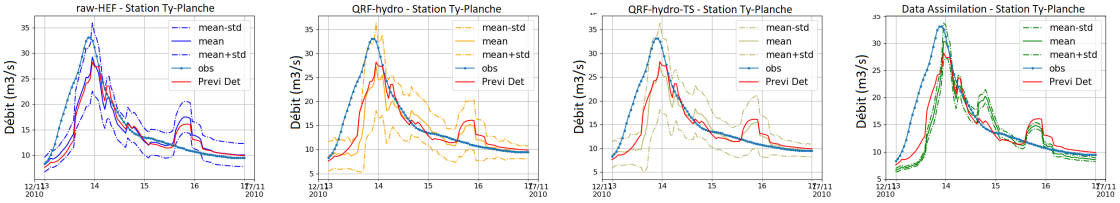


Figure P.24: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 18 heures à Ty-Planche. Événement du 13 au 17 novembre 2010.

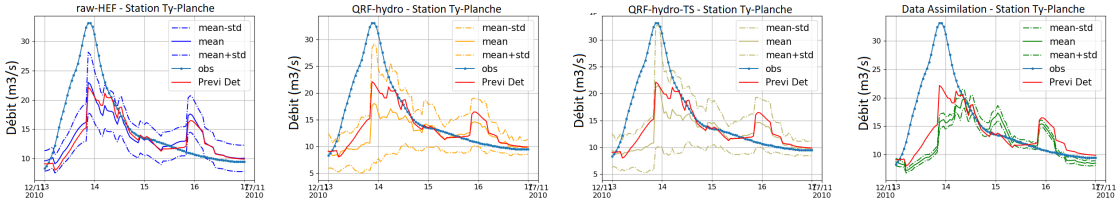


Figure P.25: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 24 heures à Ty-Planche. Événement du 13 au 17 novembre 2010.

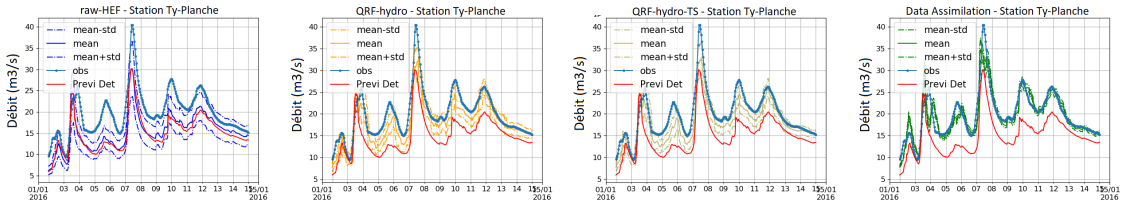


Figure P.26: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 6 heures à Ty-Planche. Événement du 2 au 15 janvier 2016.

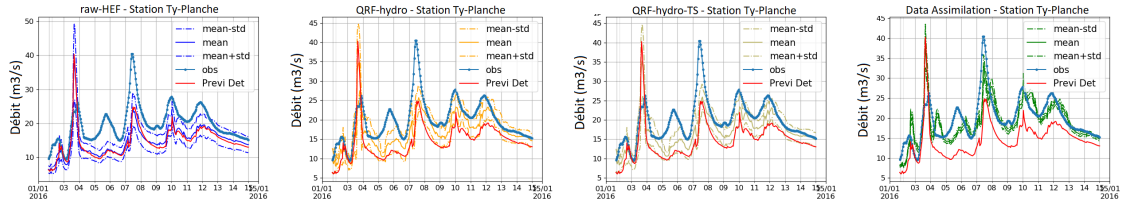


Figure P.27: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 12 heures à Ty-Planche. Événement du 2 au 15 janvier 2016.

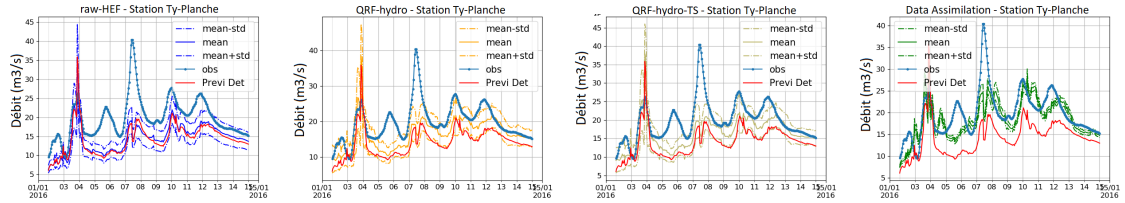


Figure P.28: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 18 heures à Ty-Planche. Événement du 2 au 15 janvier 2016.

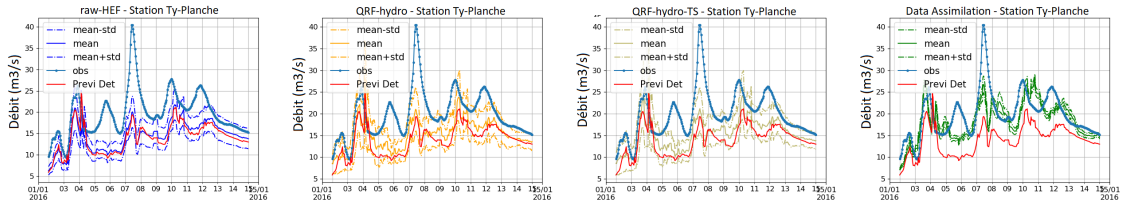


Figure P.29: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 24 heures à Ty-Planche. Événement du 2 au 15 janvier 2016.

Annexe Q

Compléments au chapitre 8

Sommaire

Q.1 Calibration à Kervir et Moulin-Vert 392

Q.1.1 Comparaison des critères Opt_{Rel} et Opt_{CRPS} pour le choix de la meilleure forêt 392

Q.1.2 Histogrammes de rang 395

Q.2 Calibration à Justice 400

Q.2.1 Comparaison des critères Opt_{Rel} et Opt_{CRPS} pour le choix de la meilleure forêt 400

Q.2.2 Histogrammes de rang 403

Q.1 Calibration à Kervir et Moulin-Vert

Q.1.1 Comparaison des critères Opt_{Rel} et Opt_{CRPS} pour le choix de la meilleure forêt

Ensembles évalués dans ce paragraphe

⇒ les 6 ensembles de l'expérience E_{PA} : ensembles bruts construits avec le modèle MASCARET alimenté par les prévisions de marée astronomique.

R_R : ensemble hydraulique brut, condition limite amont *raw-HEF*.

Q_R : quantiles calibrés à partir de l'ensemble hydraulique brut R_R.

QNE_{Md} : quantiles calibrés, mêmes prédicteurs que pour Q_R sauf les prédicteurs issus de statistiques de l'ensemble brut R_R.

R_Q : ensemble hydraulique brut, condition limite amont *QRF-hydro-TS*.

Q_Q : quantiles calibrés à partir de l'ensemble hydraulique brut R_Q.

QNE_Q : quantiles calibrés, mêmes prédicteurs que pour Q_Q sauf les prédicteurs issus de statistiques de l'ensemble brut R_Q.

La figure Q.1 présente les scores des six ensembles R_R, R_Q, Q_R, Q_Q, QNE_{Md}, et QNE_Q sur la période de calage P_{Cal} pour les échéances allant de 1 heure à 24 heures, pour les deux façons d'optimiser les forêts aléatoires : optimisation du CRPS uniquement (Opt_{CRPS}) et optimisation de la fiabilité au préalable (Opt_{Rel}). Les scores présentés sont les valeurs des CRPS, de la norme $\|.\|_2$ des histogrammes de rang et des aires des courbes ROC pour le dépassement des quantiles q_{90} et q_{99}. Dans le tableau Q.1, les moyennes des scores sur les 24 heures d'échéances pour les deux stratégies d'optimisation des forêts sont présentées. La figure Q.2 et le tableau Q.2 présentent les résultats analogues en période de validation P_{Val}. La comparaison n'est pas montrée ici pour Moulin-Vert, mais les conclusions émises dans ce paragraphe y sont également valables.

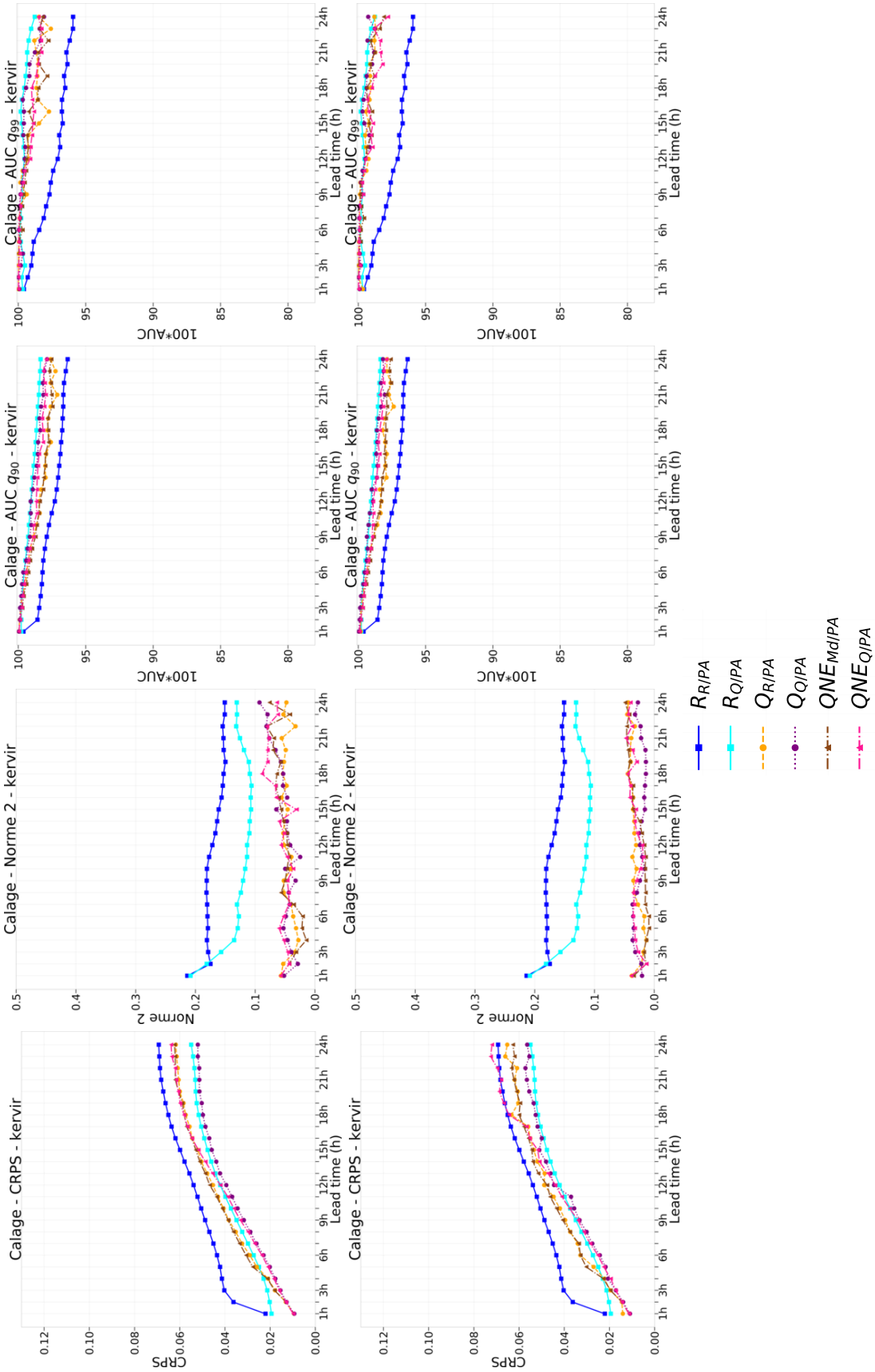


Figure Q.1: Comparaison des deux types d'optimisation des forêts aléatoires à Kervir pour l'expérience E_{PA} : Opt_{CRPS} (ligne 1) et Opt_{Rel} (ligne 2). Période P_{Cal}.

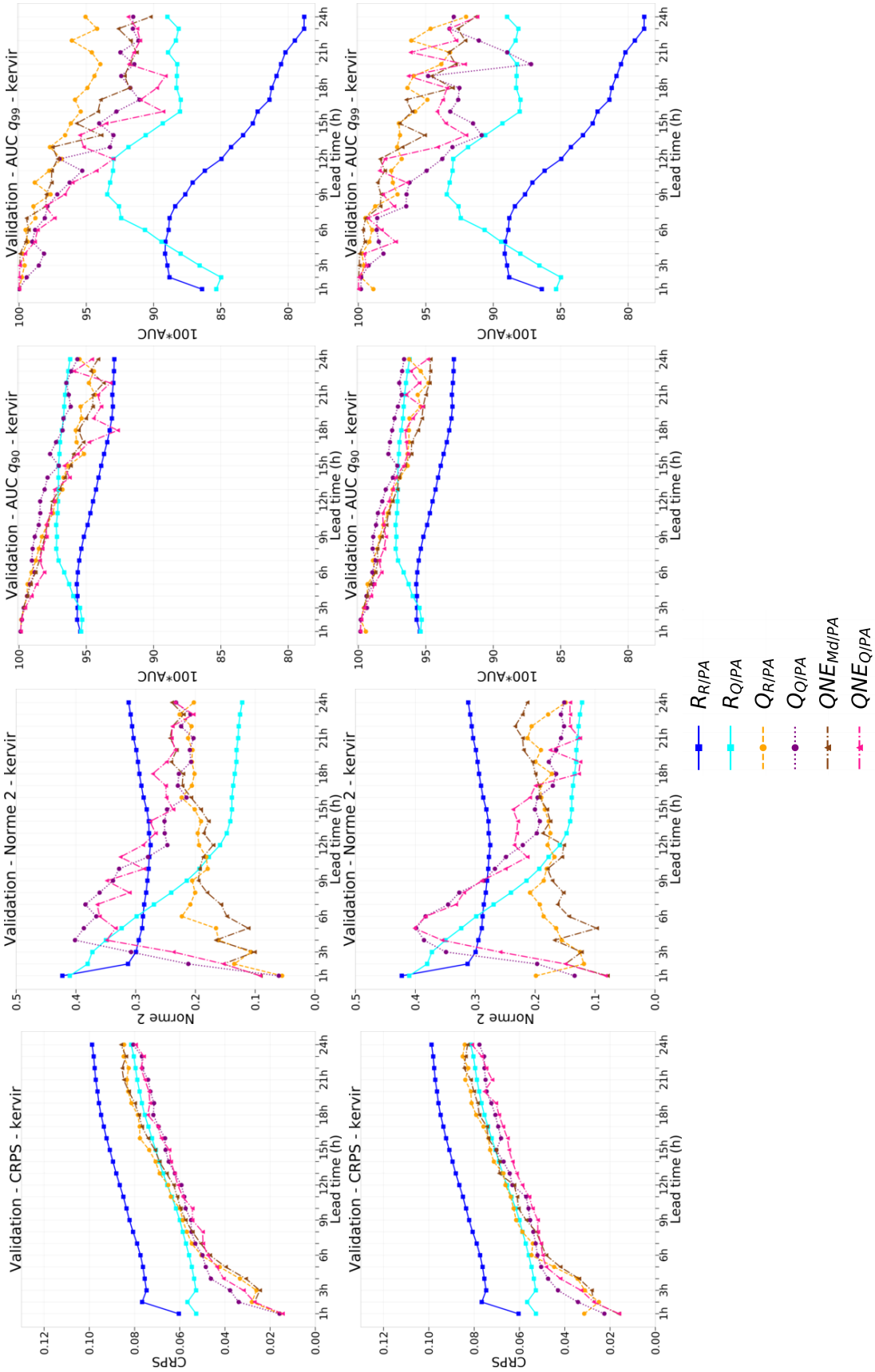


Figure Q.2: Comparaison des deux types d'optimisation des forêts aléatoires à Kervir pour l'expérience E_{PA} : Opt_{CRPS} (ligne 1) et Opt_{Rel} (ligne 2). Période P_{Val}.

Ensemble	$10^{-2}\|.\|_2$		100*CRPS		100 * AUC - q_{90}		100 * AUC - q_{99}	
R_R	16.9		5.4		97.5		97.4	
R_Q	12.7		4.0		99.1		99.6	
	Opt_{Rel}	Opt_{CRPS}	Opt_{Rel}	Opt_{CRPS}	Opt_{Rel}	Opt_{CRPS}	Opt_{Rel}	Opt_{CRPS}
Q_R	3.2	4.6	4.5	4.3	98.6	98.5	99.4	99.2
Q_Q	2.4	5.3	4.0	3.7	99.0	98.9	99.5	99.5
QNE_{Md}	2.7	5.0	4.5	4.4	98.6	98.5	99.4	99.2
QNE_Q	3.3	5.8	4.5	4.1	98.8	98.8	99.3	99.3

Tableau Q.1: $\|.\|_2$, CRPS et aire ROC moyennés sur les 24 heures d'échéances à Kervir. Expérience E_{PA}. Période P_{Cal}.

Ensemble	$10^{-2}\|.\|_2$		100*CRPS		100 * AUC - q_{90}		100 * AUC - q_{99}	
R_R	29.7		8.6		94.4		84.5	
R_Q	20.8		6.6		96.6		89.6	
	Opt_{Rel}	Opt_{CRPS}	Opt_{Rel}	Opt_{CRPS}	Opt_{Rel}	Opt_{CRPS}	Opt_{Rel}	Opt_{CRPS}
Q_R	17.9	18.7	6.4	6.3	97.5	97.2	97.0	97.1
Q_Q	23.4	26.6	6.0	5.9	98.1	98.0	94.7	95.1
QNE_{Md}	17.4	18.6	6.1	6.1	97.2	97.0	96.7	96.0
QNE_Q	22.3	26.6	5.7	5.8	97.4	96.8	96.0	94.7

Tableau Q.2: $\|.\|_2$, CRPS et aire ROC moyennés sur les 24 heures d'échéances à Kervir. Expérience E_{PA}. Période P_{Val}.

Sur la période de calage, la stratégie Opt_{Rel} entraîne une courbe de CRPS moins lisse selon les échéances. La stratégie Opt_{CRPS} permet d'obtenir des performances en termes de CRPS légèrement meilleures que la stratégie Opt_{Rel}. En revanche, la stratégie Opt_{Rel} améliore plus nettement les normes des histogrammes de rang. Ainsi, en calage, si l'on accepte de dégrader légèrement le CRPS, on améliore significativement la fiabilité des forêts aléatoires.

Sur la période de validation, la tendance des résultats est un peu différente. Les valeurs des CRPS pour les deux types d'optimisation sont quasiment équivalentes. En revanche, la fiabilité reste toujours un peu meilleure avec les forêts construites avec la stratégie $OptimRel$.

Les courbes ROC ne sont pas significativement impactées par le choix de l'optimisation des forêts sur les deux périodes, même si l'aire ROC est très légèrement meilleure avec la stratégie Opt_{Rel}.

Pour la suite de l'étude, nous faisons le choix de conserver le critère Opt_{Rel} pour l'optimisation des forêts aléatoires à Kervir et Moulin-Vert.

Q.1.2 Histogrammes de rang

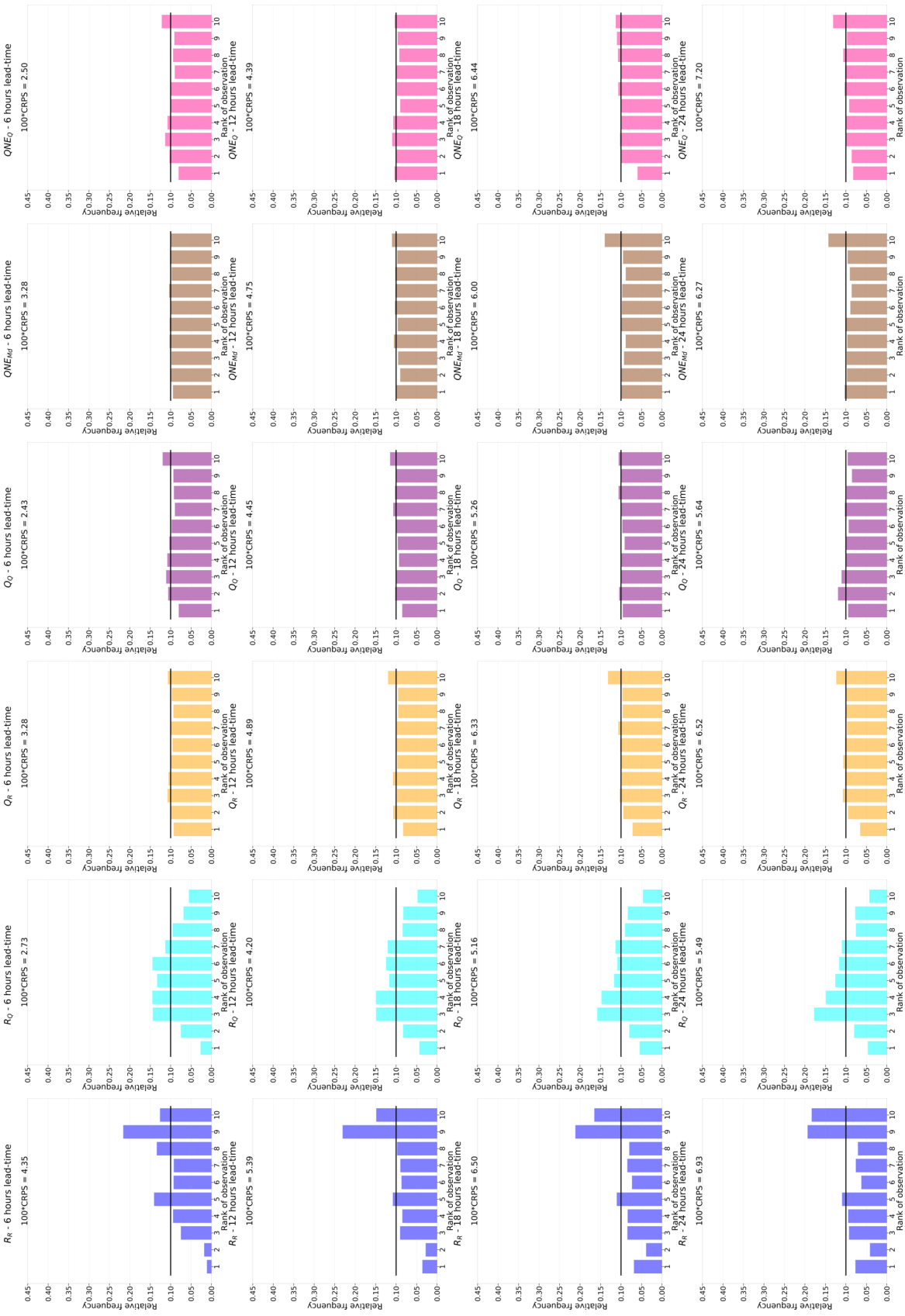


Figure Q.3: Histogrammes de rang pour la stratification M_{90} à Kervir pour les échéance 6, 12, 18 et 24 heures - Ensembles $R_{R/PA}$, $R_{Q/PA}$, $Q_{Q/PA}$, $QNE_{Md/PA}$ et $QNE_{Q/PA}$. Période de calage P_{Cal}.

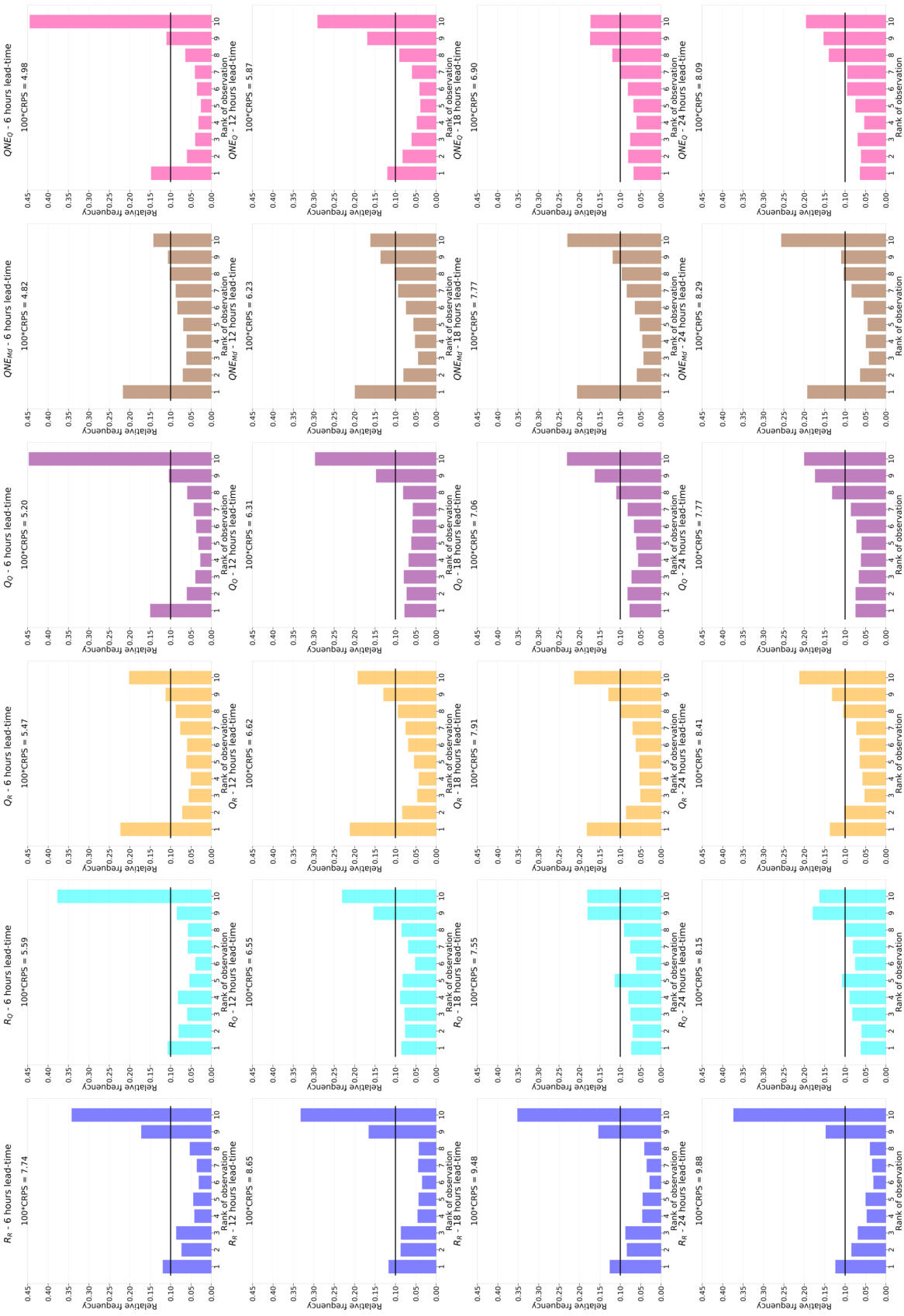


Figure Q.4: Histogrammes de rang pour la stratification M_{90} à Kervir pour les échéance 6, 12, 18 et 24 heures - Ensembles $R_{R/PA}$, $R_{Q/PA}$, $Q_{R/PA}$, $Q_{Q/PA}$, $QNE_{Md/PA}$ et $QNE_{Q/PA}$. Période de validation P_{Val}.

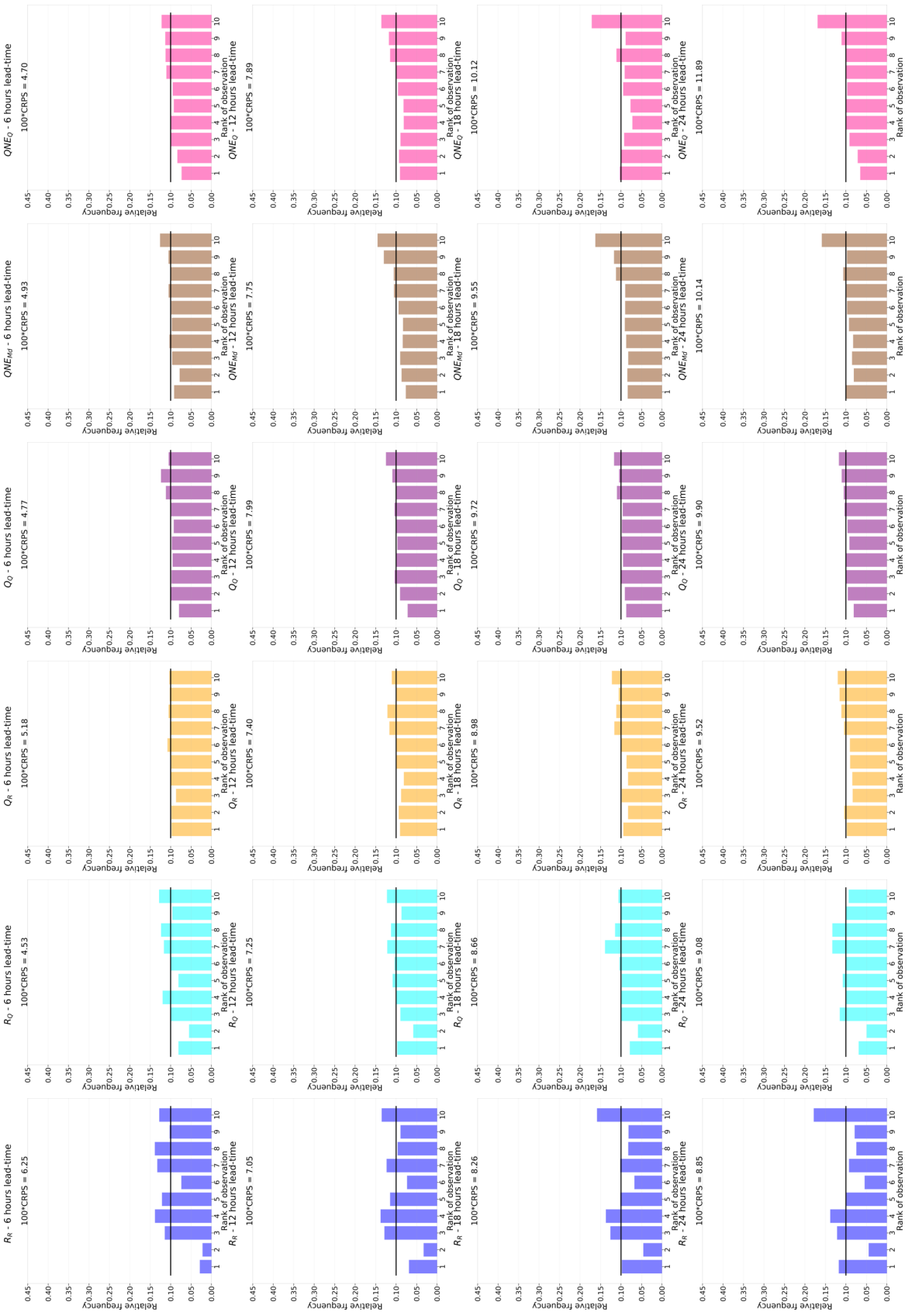


Figure Q.5: Histogrammes de rang pour la stratification M_{90} à Moulin-Vert pour les échéance 6, 12, 18 et 24 heures - Ensembles $R_{R/PA}$, $R_{Q/PA}$, $Q_{R/PA}$, $Q_{Q/PA}$, $QNE_{Md/PA}$ et $QNE_{Q/PA}$. Période de calage P_{Cal}.

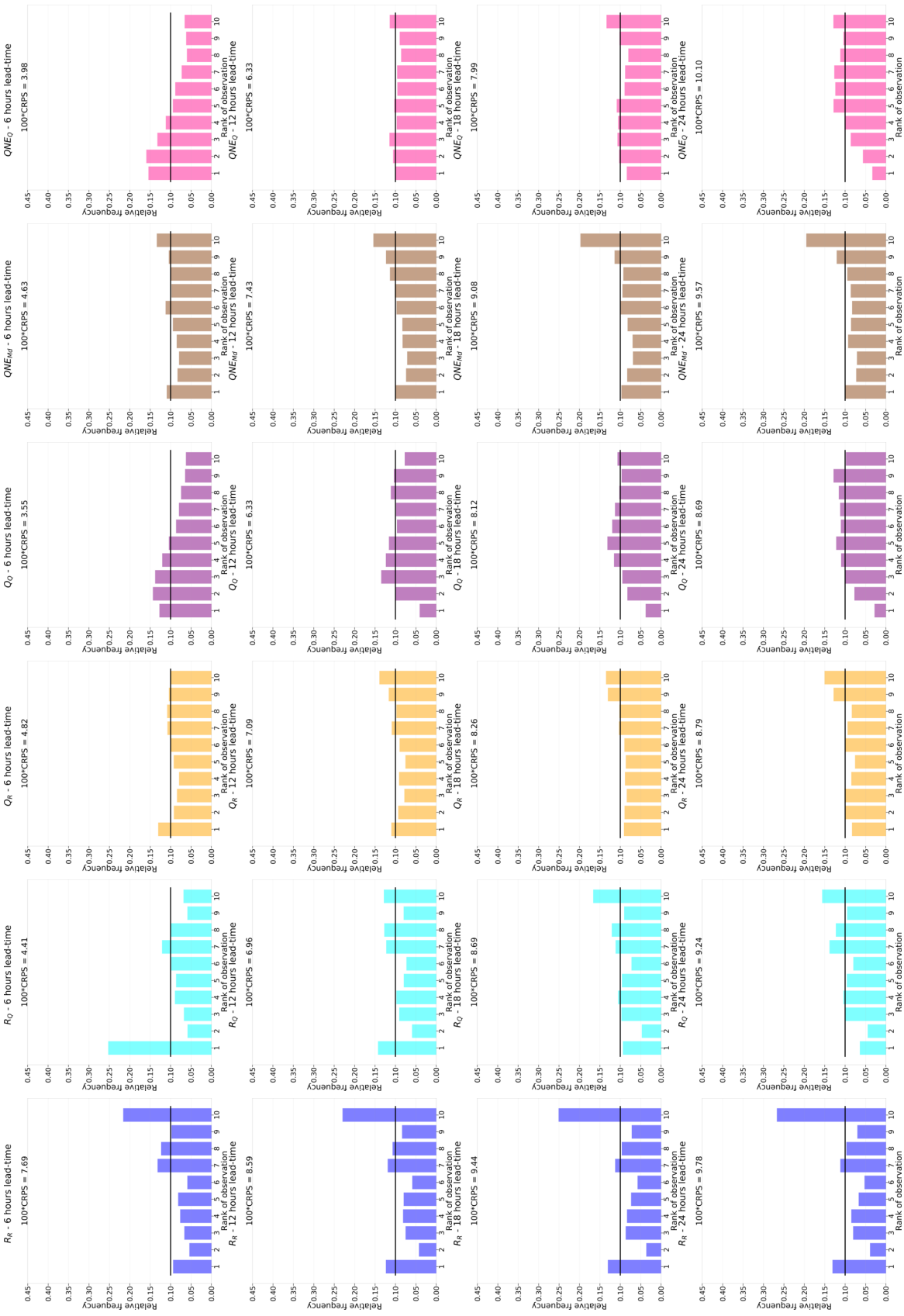


Figure Q.6: Histogrammes de rang pour la stratification M_{90} à Moulin-Vert pour les échéance 6, 12, 18 et 24 heures - Ensembles $R_{R/PA}$, $R_{Q/PA}$, $Q_{R/PA}$, $Q_{Q/PA}$, $QNE_{Md/PA}$ et $QNE_{Q/PA}$. Période de validation P_{Val}.

Q.2 Calibration à Justice

Q.2.1 Comparaison des critères Opt_{Rel} et Opt_{CRPS} pour le choix de la meilleure forêt

Ensembles évalués dans ce paragraphe

⇒ les 6 ensembles de l'expérience E_{PASC} : la condition limite maritime est la prévision de marée astronomique à laquelle on ajoute la surcote mesurée au temps de base.

R_R : ensemble hydraulique brut, condition limite amont *raw-HEF*.

Q_R : quantiles calibrés à partir de l'ensemble hydraulique brut R_R.

QNE_{Md} : quantiles calibrés, mêmes prédicteurs que pour Q_R sauf les prédicteurs issus de statistiques de l'ensemble brut R_R.

R_Q : ensemble hydraulique brut, condition limite amont *QRF-hydro-TS*.

Q_Q : quantiles calibrés à partir de l'ensemble hydraulique brut R_Q.

QNE_Q : quantiles calibrés, mêmes prédicteurs que pour Q_Q sauf les prédicteurs issus de statistiques de l'ensemble brut R_Q.

La figure Q.7 présente les scores des six ensembles $R_{R/PASC}$, $R_{Q/PASC}$, $Q_{R/PASC}$, $Q_{Q/PASC}$, $QNE_{Md/PASC}$, et $QNE_{Q/PASC}$ sur la période de calage P_{Cal} pour les échéances allant de 1 heure à 24 heures, pour les deux façons d'optimiser les forêts aléatoires : optimisation du CRPS uniquement (Opt_{CRPS}) et optimisation de la fiabilité au préalable (Opt_{Rel}). Les scores présentés sont les mêmes que pour la calibration à Kervir et Moulin-Vert. Les aires AUC sont celles de courbes ROC tracées pour le dépassement du seuil de vigilance vert-jaune. Dans le tableau Q.3, les moyennes des scores sur les 24 heures d'échéances pour les deux stratégies d'optimisation des forêts sont présentées. La figure Q.8 et le tableau Q.4 présentent les résultats analogues en période de validation P_{Val}.

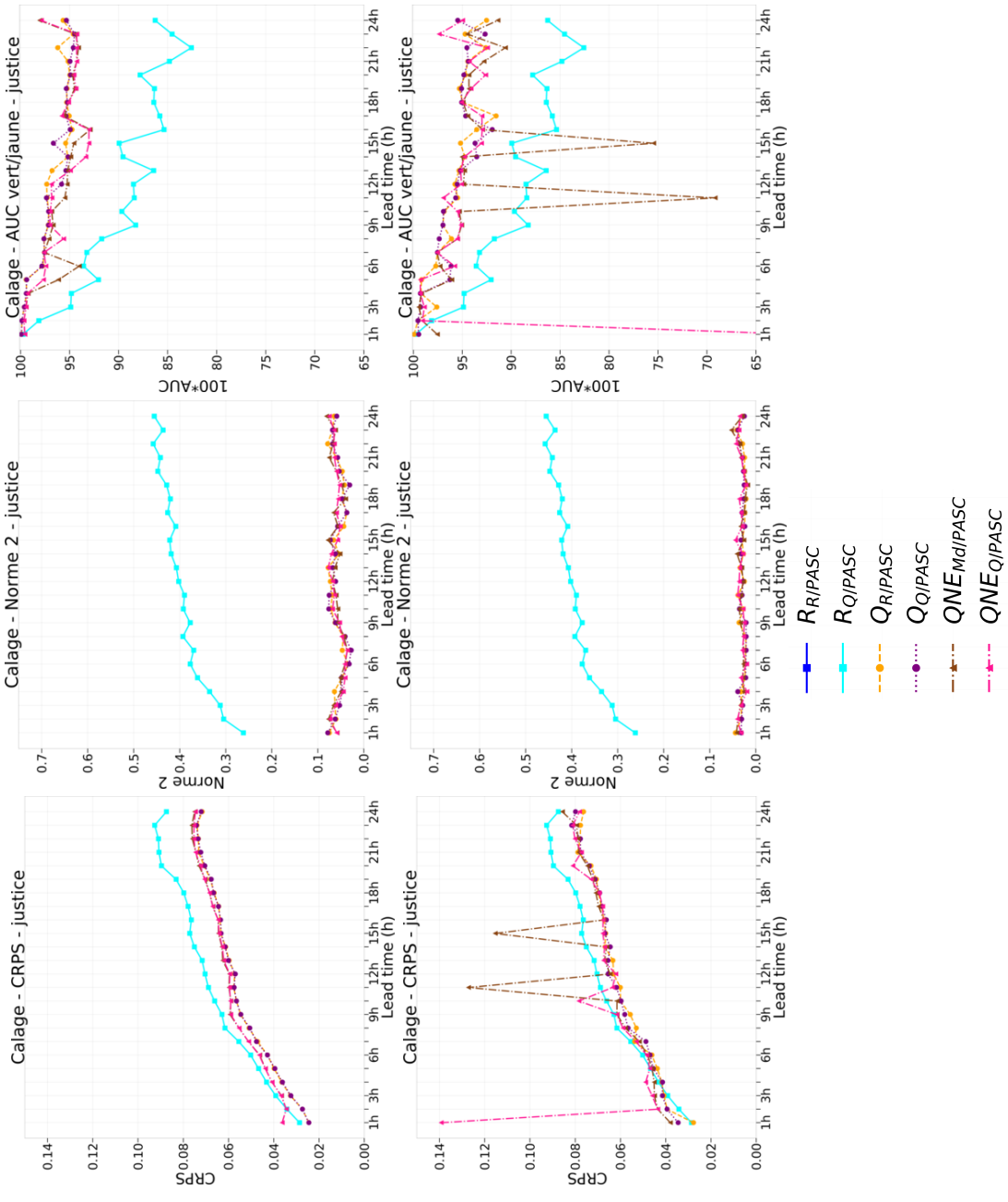


Figure Q.7: Comparaison des deux types d'optimisation des forêts aléatoires à Justice pour l'expérience E_{PASC} : Opt_{CRPS} (ligne 1) et Opt_{Rel} (ligne 2). Période P_{Cal}.

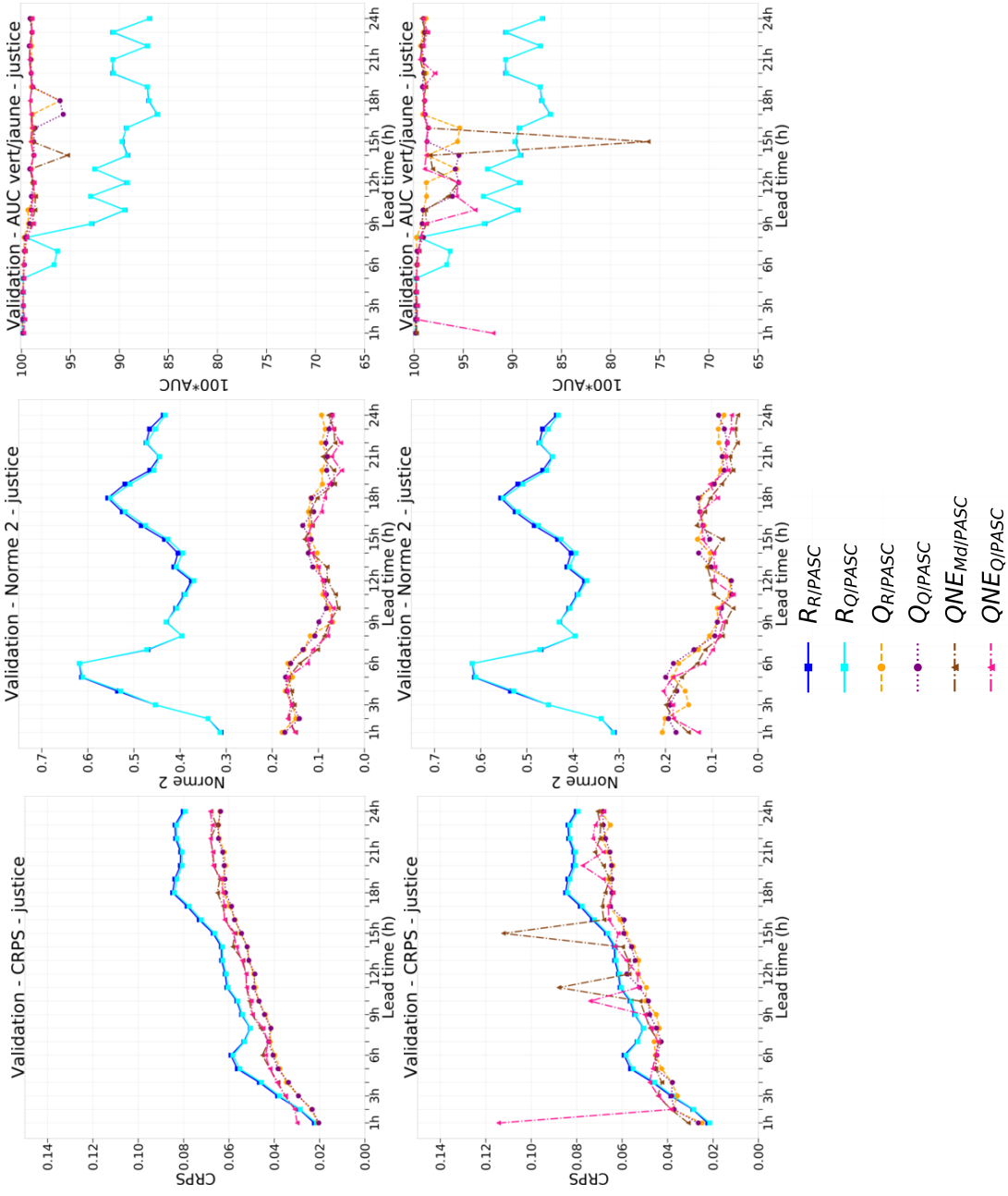


Figure Q.8: Comparaison des deux types d'optimisation des forêts aléatoires à Justice pour l'expérience E_{PASC} : Opt_{CRPS} (ligne 1) et Opt_{Rel} (ligne 2). Période P_{Val}.

Ensemble	$10^{-2}\|.\|_2$		100*CRPS		100 * AUC - seuil V/J	
$R_{R/PASC}$	39.3		6.7		89.5	
$R_{Q/PASC}$	39.3		6.7		89.5	
	Opt_{Rel}	Opt_{CRPS}	Opt_{Rel}	Opt_{CRPS}	Opt_{Rel}	Opt_{CRPS}
$Q_{R/PASC}$	2.8	5.8	6.0	5.6	95.9	96.8
$Q_{Q/PASC}$	2.8	5.6	6.1	5.6	95.9	96.7
$QNE_{Md/PASC}$	3.1	5.8	6.7	5.9	93.4	96.0
$QNE_{Q/PASC}$	3.1	5.9	6.8	5.9	94.1	96.2

Tableau Q.3: $\|.\|_2$, CRPS et aire ROC moyennés sur les 24 heures d'échéances à Justice. Expérience E_{PASC}. Période P_{Cal}.

Ensemble	$10^{-2}\|.\|_2$		100*CRPS		100 * AUC - seuil V/J	
$R_{R/PASC}$	45.7		6.3		92.6	
$R_{Q/PASC}$	45.3		6.2		92.6	
	Opt_{Rel}	Opt_{CRPS}	Opt_{Rel}	Opt_{CRPS}	Opt_{Rel}	Opt_{CRPS}
$Q_{R/PASC}$	11.7	11.5	5.3	4.9	98.8	99.1
$Q_{Q/PASC}$	11.7	11.4	5.3	4.9	98.7	99.0
$QNE_{Md/PASC}$	10.4	10.4	6.0	5.3	98.0	99.0
$QNE_{Q/PASC}$	10.6	10.3	6.1	5.3	98.3	99.2

Tableau Q.4: $\|.\|_2$, CRPS et aire ROC moyennés sur les 24 heures d'échéances à Justice. Expérience E_{PASC}. Période P_{Val}.

La stratégie d'optimisation de la fiabilité Opt_{Rel} conduit, pour certaines échéances isolées, à des valeurs pour les CRPS ou pour les aires des courbes ROC complètement en-dehors de la tendance générale pour les autres échéances. Pour ces points particuliers, la stratégie n'est donc pas adaptée. Il serait possible de l'améliorer en conservant dans un premier temps davantage que seulement 3 forêts optimisés pour la fiabilité. Le CRPS serait ensuite optimisé dans cet ensemble.

En période de calage, la stratégie Opt_{CRPS} fournit les meilleurs CRPS et des aires pour les courbes ROC un peu meilleures, et Opt_{Rel} fournit les plus petites normes pour les histogrammes de rang des ensembles calibrés. En période de validation, la stratégie Opt_{CRPS} fournit toujours les meilleurs CRPS et les meilleures discriminations pour l'événement de dépassement de seuil. Cette stratégie fournit également des scores équivalents ou légèrement meilleurs pour la norme des histogrammes de rang. La stratégie Opt_{CRPS} présentant des meilleures performances en validation, c'est celle-ci qui est retenue pour la calibration des pics de marée haute à Justice.

Q.2.2 Histogrammes de rang

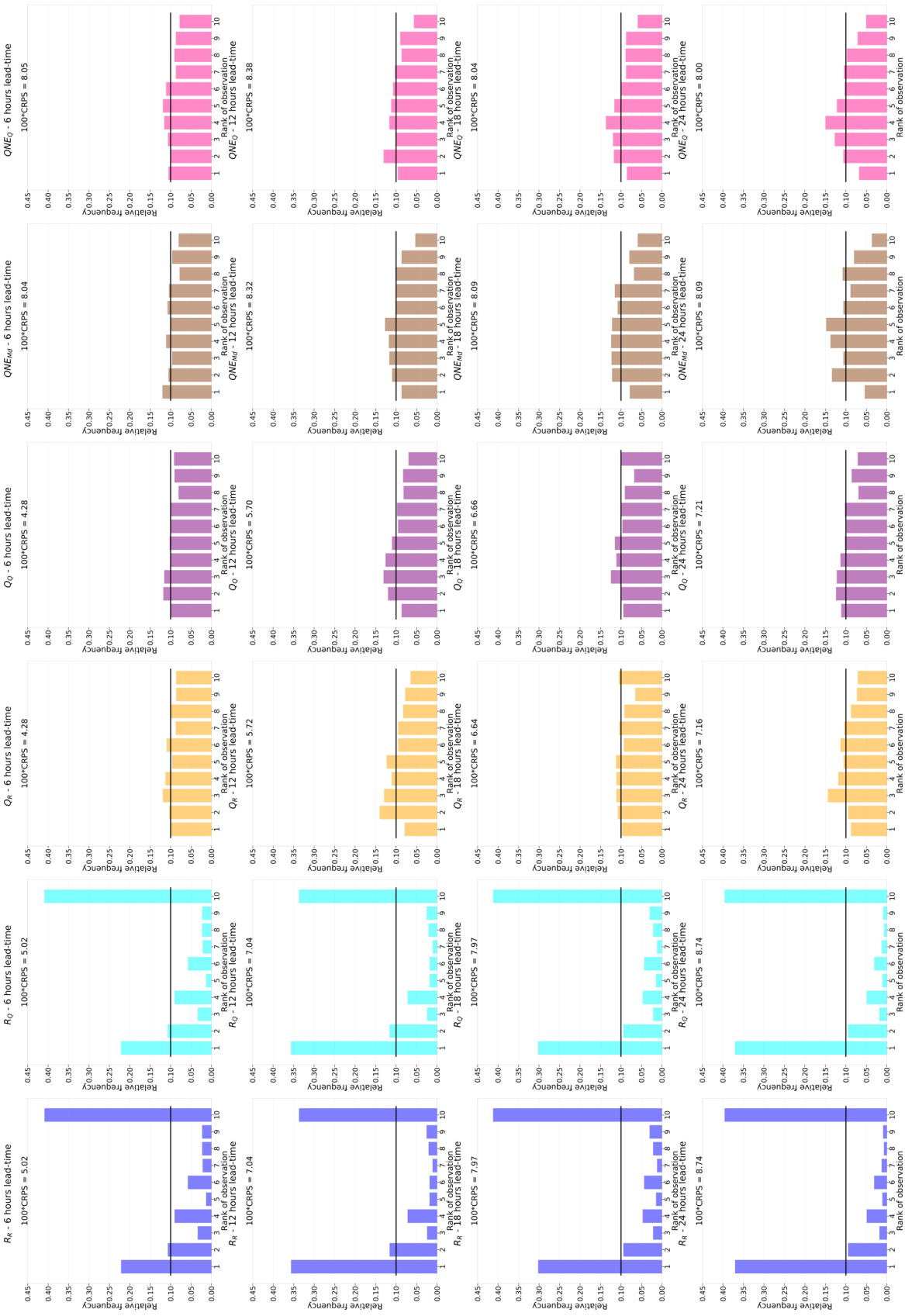


Figure Q.9: Histogrammes de rang à Justice pour les échéance 6, 12, 18 et 24 heures - Ensembles $R_{R/PASC}$, $R_{Q/PASC}$, $Q_{R/PASC}$, $Q_{Q/PASC}$, $QNE_{Md/PASC}$ et $QNE_{Q/PASC}$. Période de calage P_{Cal}.

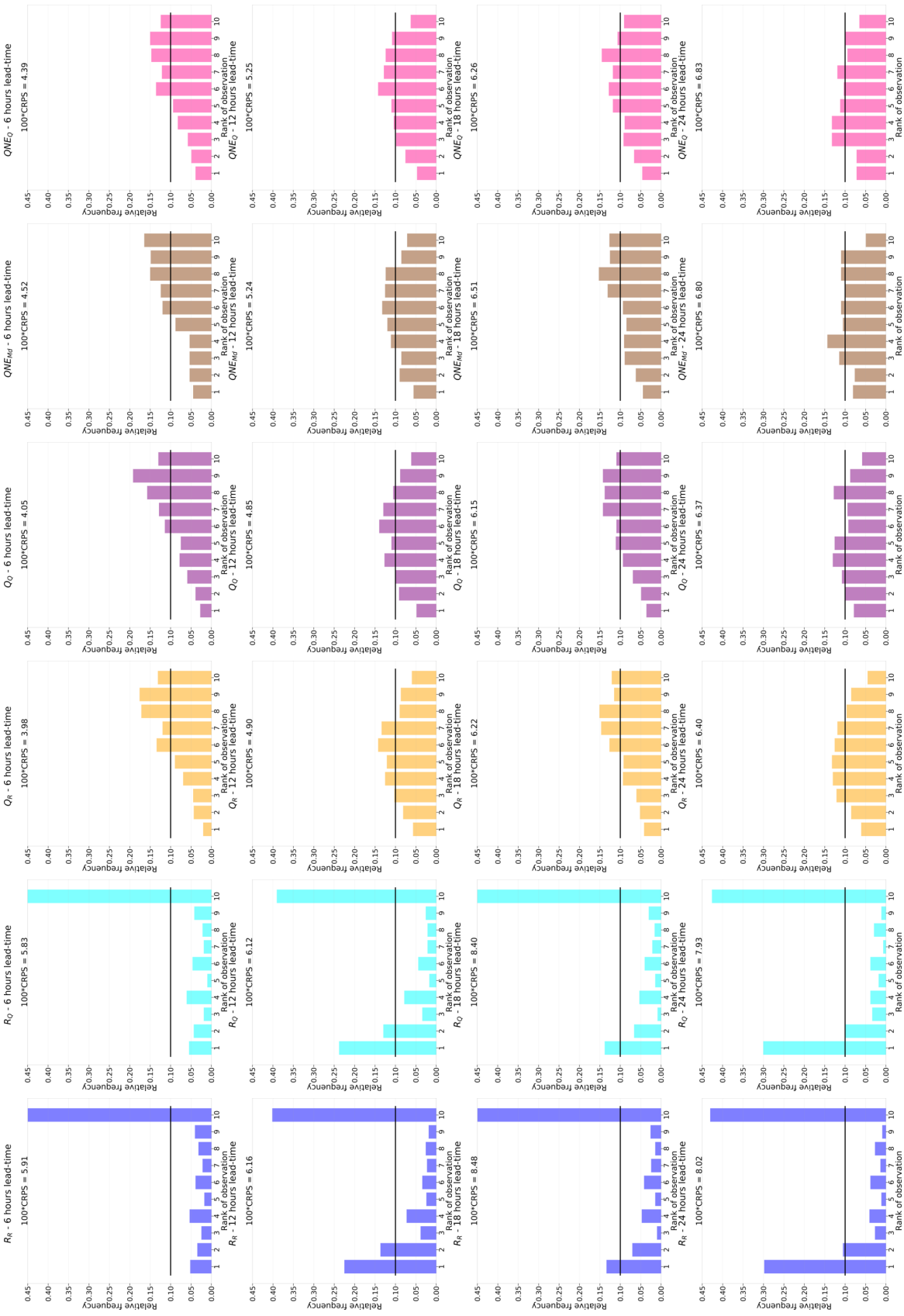


Figure Q.10: Histogrammes de rang à Justice pour les échéance 6, 12, 18 et 24 heures - Ensembles $R_{R/PASC}$, $R_{Q/PASC}$, $Q_{R/PASC}$, $Q_{Q/PASC}$, $QNE_{Md/PASC}$ et $QNE_{Q/PASC}$. Période de validation P_{Val}.

Annexe R

Compléments au chapitre 9

Sommaire

R.1 Expériences jumelles : correction simultanée des Ks et des conditions limites amont et aval 408

R.1.1 Assimilation des hauteurs : $KsCL@Z_{vigilance}$ 408

R.1.2 Assimilation des hauteurs et des débits amont : $KsCL@all_{zQ}$ . . 411

R.1.3 Assimilation des hauteurs et des conditions limites, test avec 279 membres : $KsCL@all_{zQ} - 279$ 414

R.2 Expériences réelles 418

R.1 Expériences jumelles : correction simultanée des Ks et des conditions limites amont et aval

R.1.1 Assimilation des hauteurs : $KsCL@Z_{vigilance}$

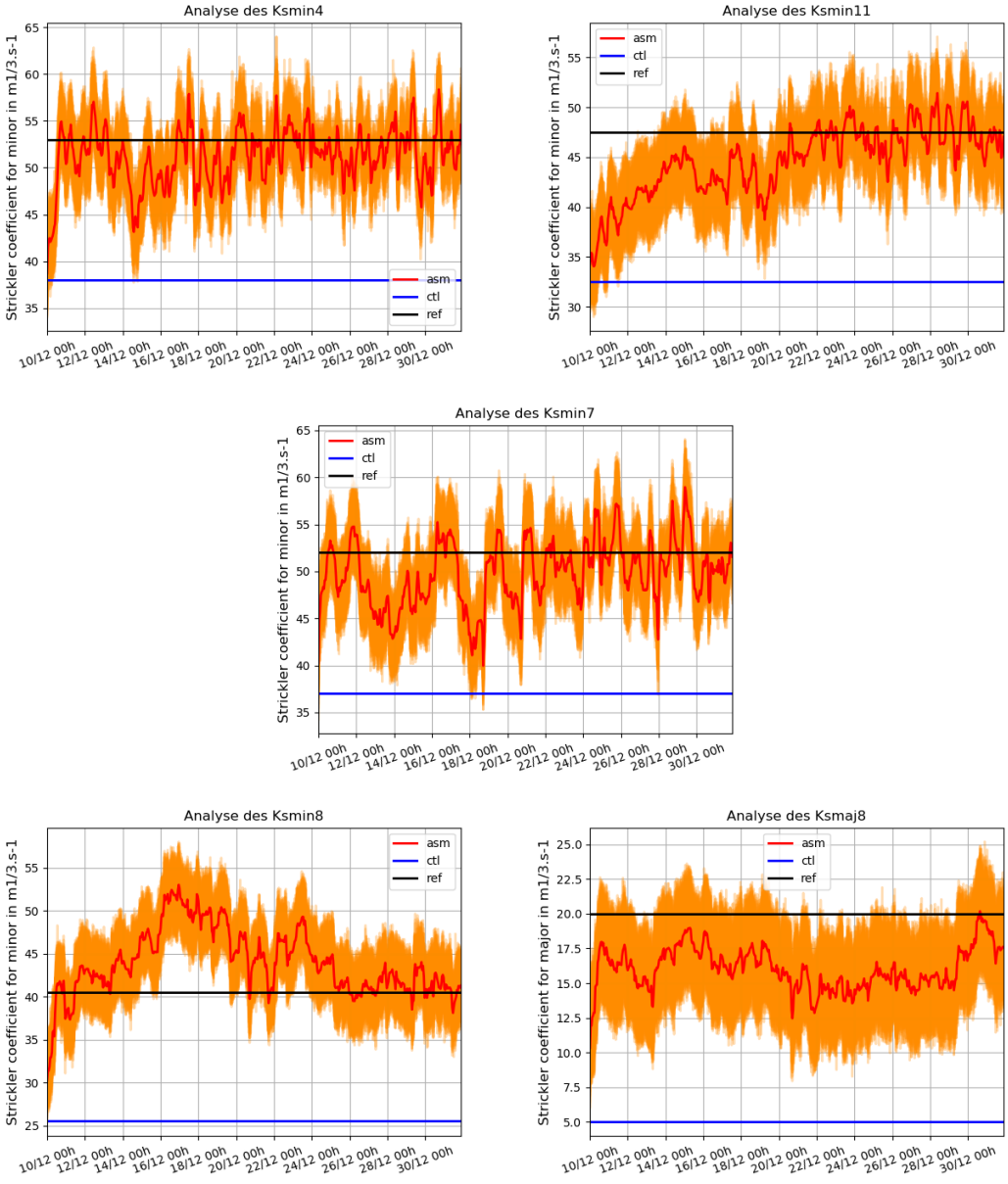


Figure R.1: Expérience $KsCL@Z_{vigilance}$ - Valeurs analysées pour les coefficients de Strickler en fonction du temps. Les courbes orange représentent les valeurs pour chacun des 49 membres et la courbe rouge la valeur analysée moyenne

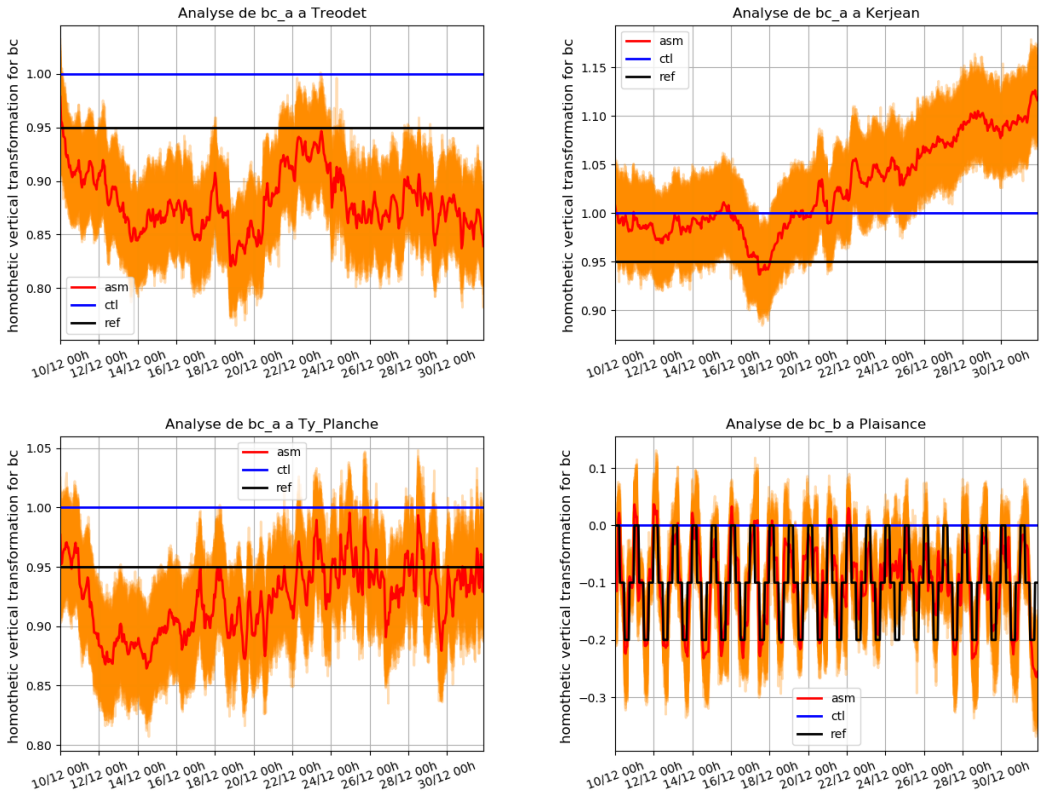


Figure R.2: Expérience $KsCL@Z_{vigilance}$ - Valeurs analysées pour la correction des conditions aux limites en fonction du temps. Les courbes orange représentent les valeurs pour chacun des 49 membres et la courbe rouge la valeur analysée moyenne

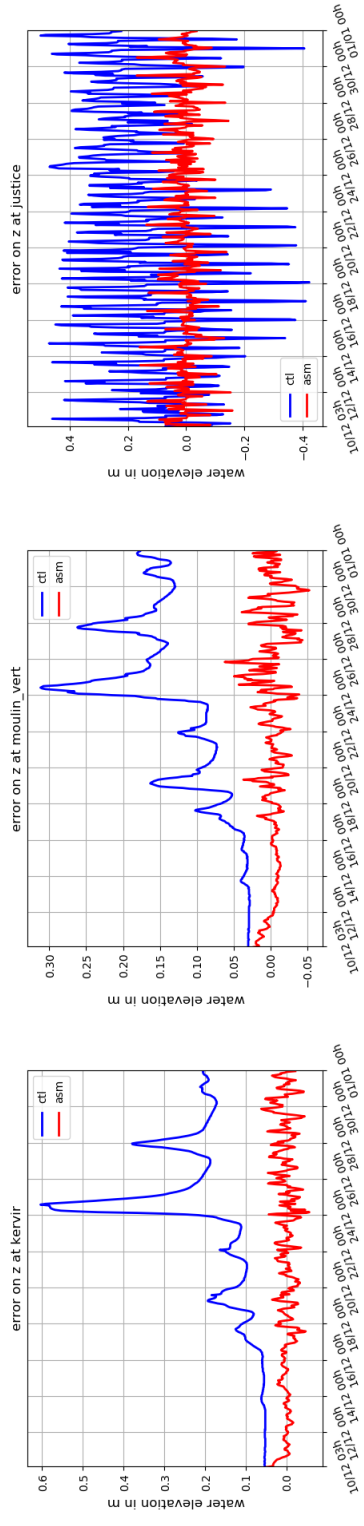


Figure R.3: Expérience $KsCL@Z_{vigilance}$ - Chroniques temporelles des erreurs aux stations de vigilance Kervir, Moulin-Vert et Justice pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (rouge) par rapport à la simulation de référence.

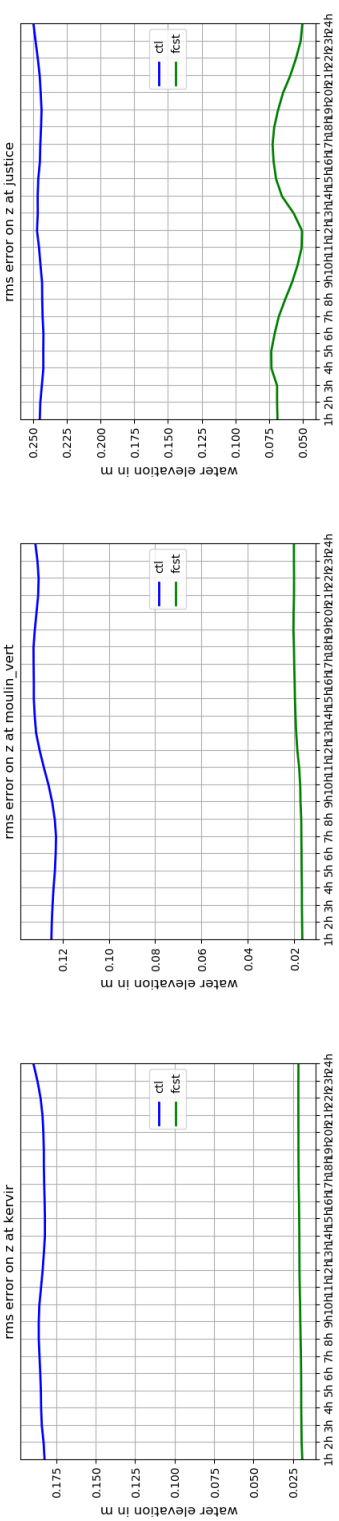


Figure R.4: Expérience $KsCL@Z_{vigilance}$ - RMSE des erreurs de prévision aux stations Kervir, Moulin-Vert et Justice en fonction de l'échéance pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (vert) par rapport à la simulation de référence.

R.1.2 Assimilation des hauteurs et des débits amont : $KsCL@all_{zQ}$

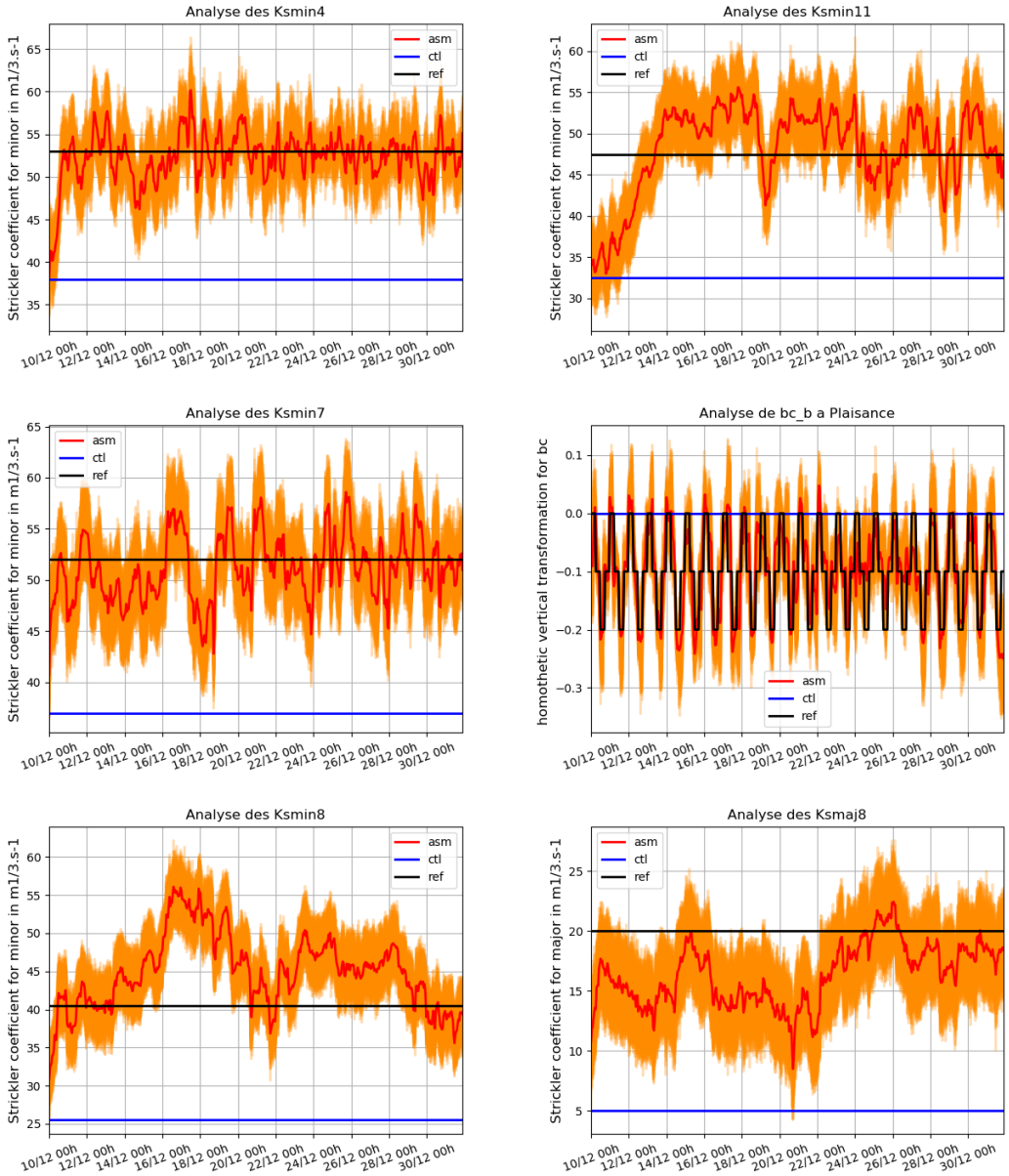


Figure R.5: Expérience $KsCL@all_{zQ}$ - Valeurs analysées en fonction du temps pour les coefficients de Strickler. Les courbes orange représentent les valeurs pour chacun des 49 membres et la courbe rouge la valeur analysée moyenne.

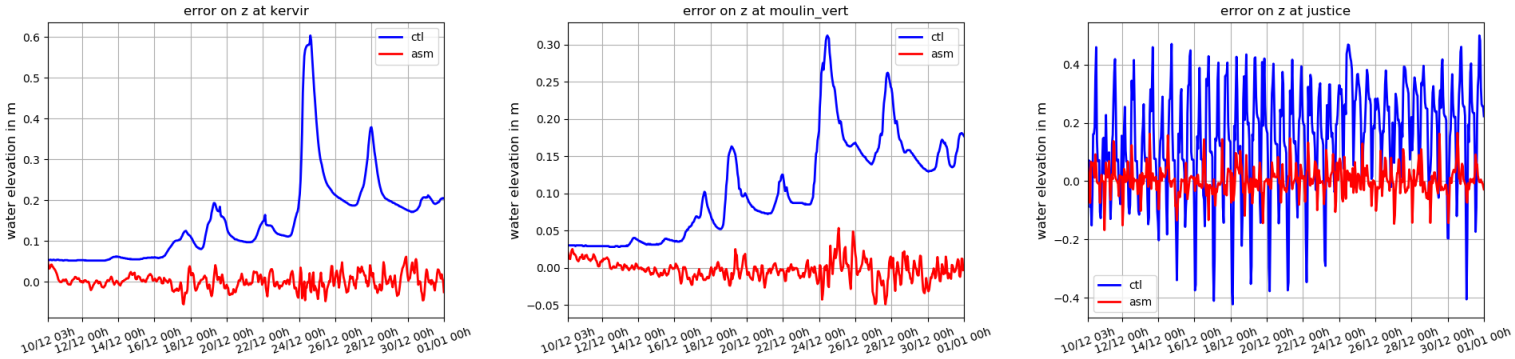


Figure R.6: Expérience $KsCL@all_{zQ}$ - Chroniques temporelles des erreurs aux stations de vigilance Kervir, Moulin-Vert et Justice pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (rouge) par rapport à la simulation de référence.

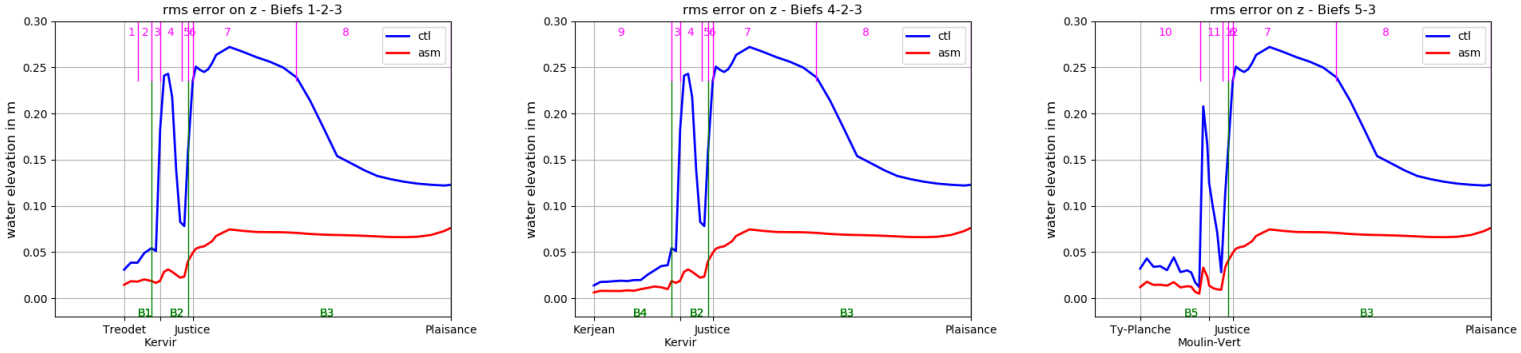


Figure R.7: Expérience $KsCL@all_{zQ}$ - RMSE le long des biefs du modèle MASCARET pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (rouge) par rapport à la simulation de référence.

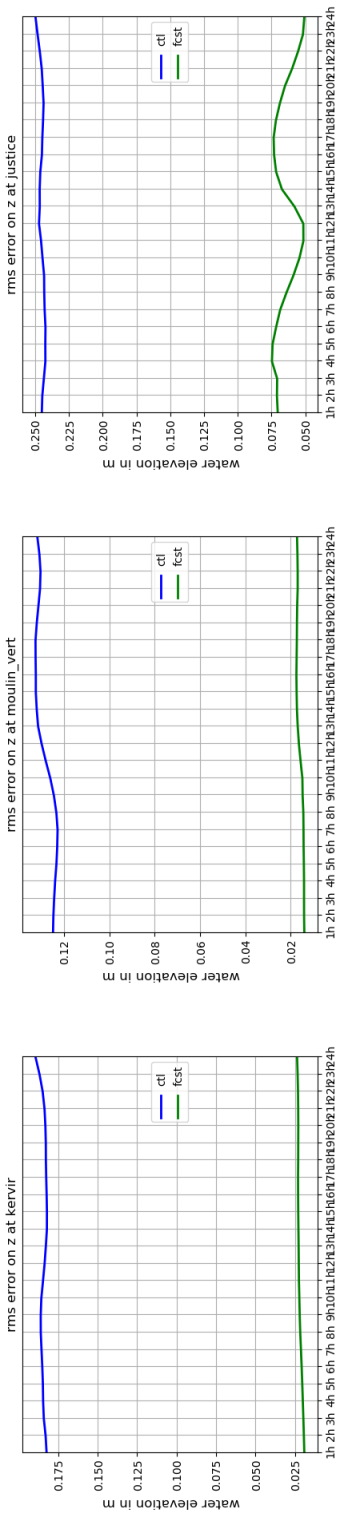


Figure R.8: Expérience $KsCL@all_zQ$ - RMSE des erreurs de prévision aux stations Kervir, Moulin-Vert et Justice en fonction de l'échéance pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (vert) par rapport à la simulation de référence.

R.1.3 Assimilation des hauteurs et des conditions limites, test avec 279 membres : $KsCL@all_{zQ} - 279$

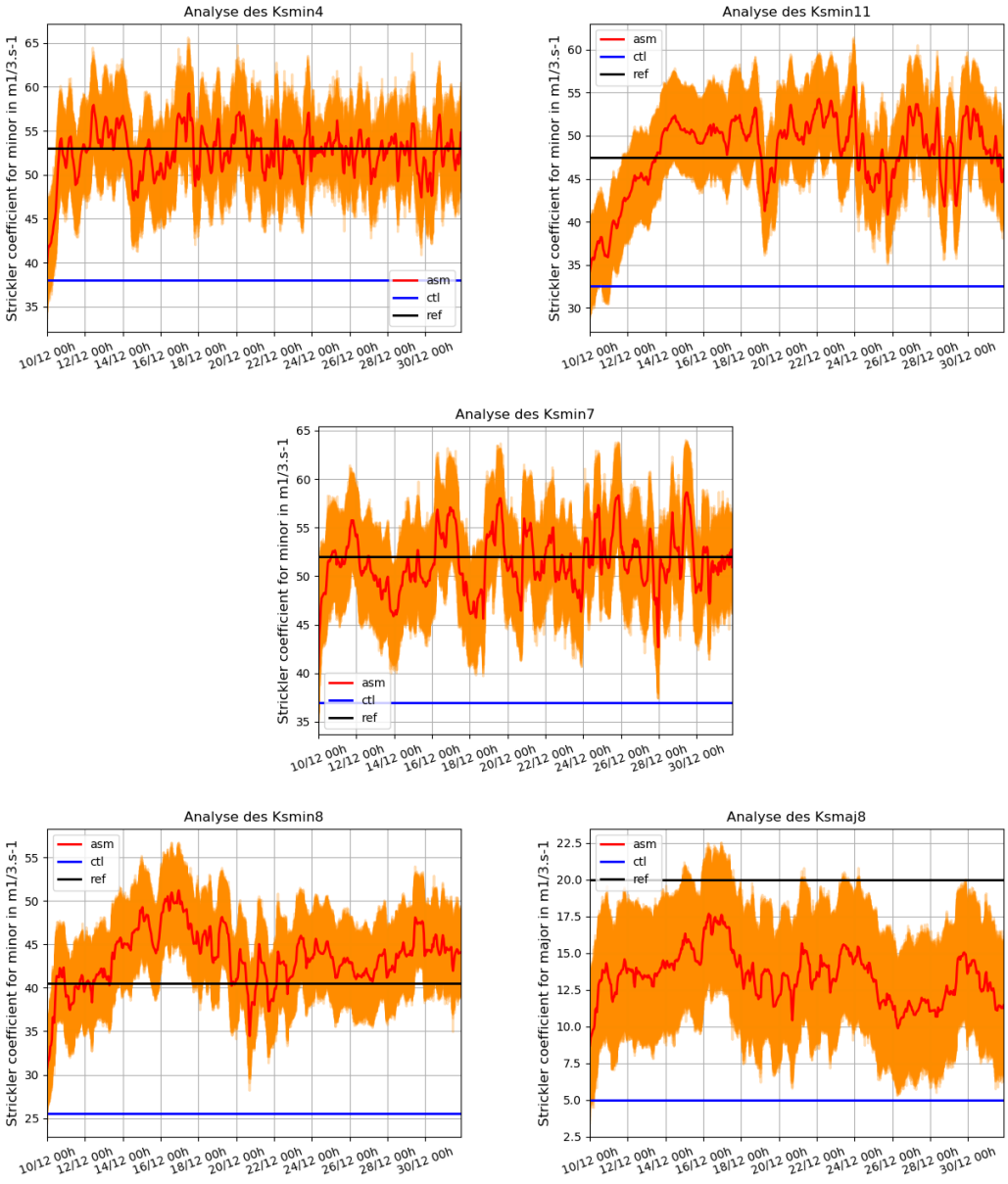


Figure R.9: Expérience $KsCL@all_{zQ} - 279$ - Valeurs analysées en fonction du temps pour les coefficients de Strickler. Les courbes orange représentent les valeurs pour chacun des 49 membres et la courbe rouge la valeur analysée moyenne.

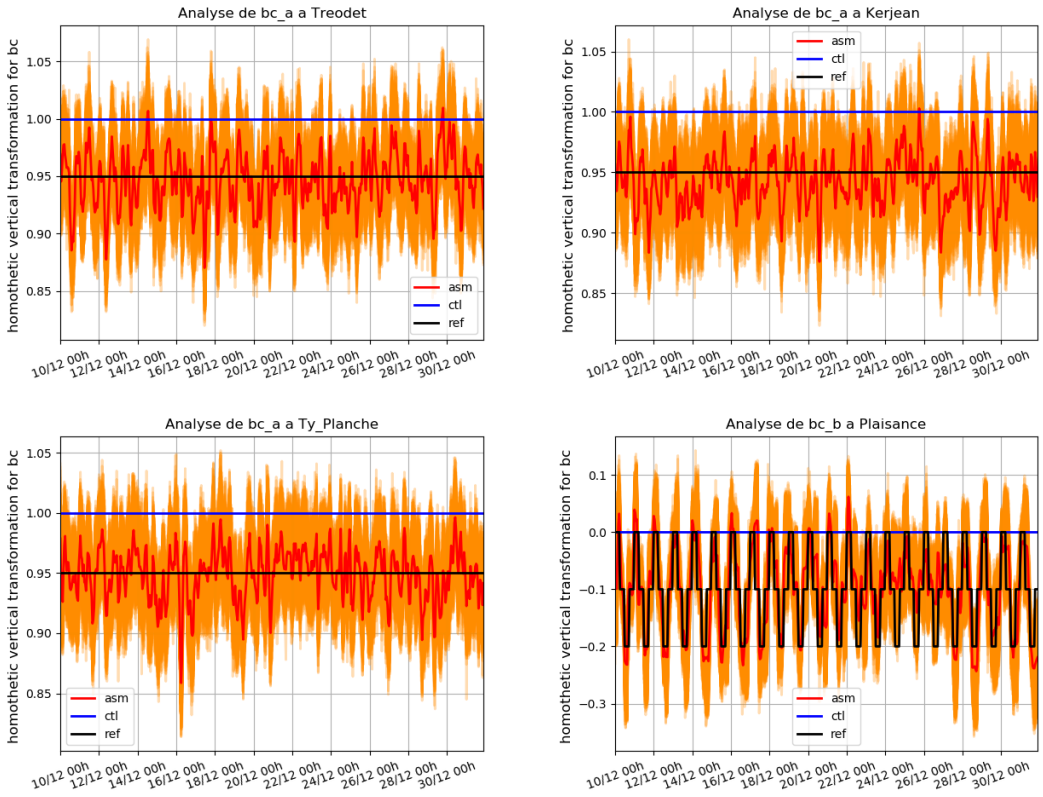


Figure R.10: Expérience $KsCL@all_{zQ} - 279$ - Valeurs analysées pour la correction des conditions aux limites en fonction du temps. Les courbes orange représentent les valeurs pour chacun des 49 membres et la courbe rouge la valeur analysée moyenne

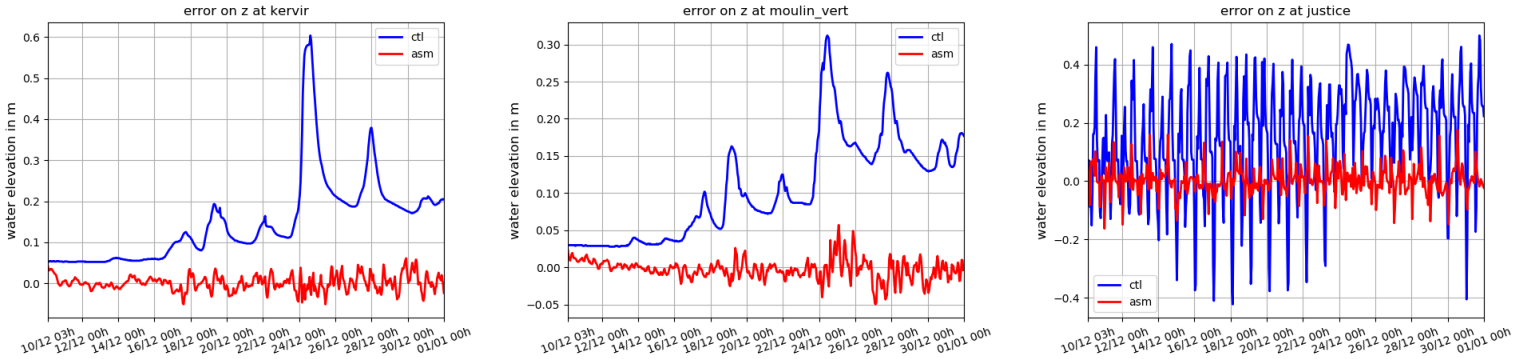


Figure R.11: Expérience $KsCL@all_{zQ} - 279$ - Chroniques temporelles des erreurs aux stations de vigilance Kervir, Moulin-Vert et Justice pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (rouge) par rapport à la simulation de référence.

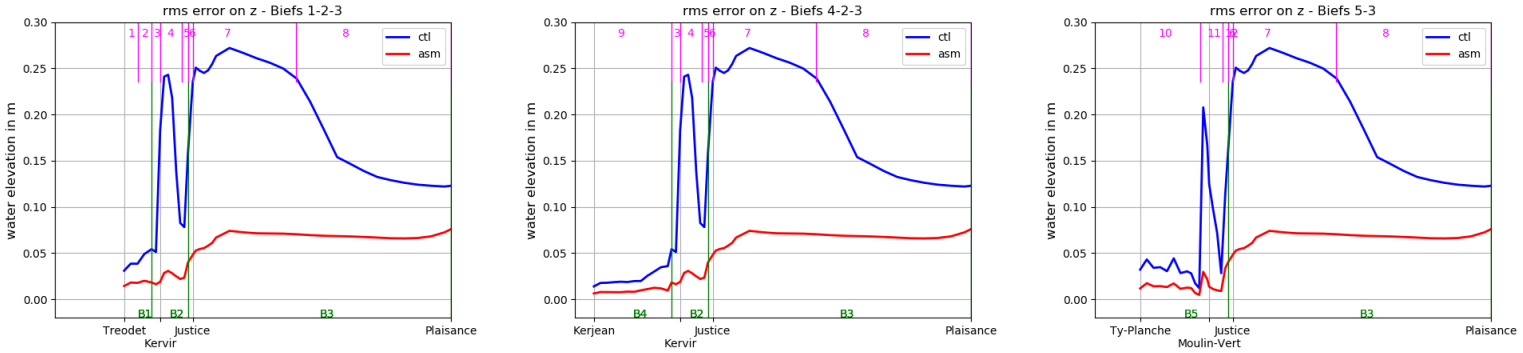


Figure R.12: Expérience $KsCL@all_{zQ} - 279$ - RMSE le long des biefs du modèle MASCARET pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (rouge) par rapport à la simulation de référence.

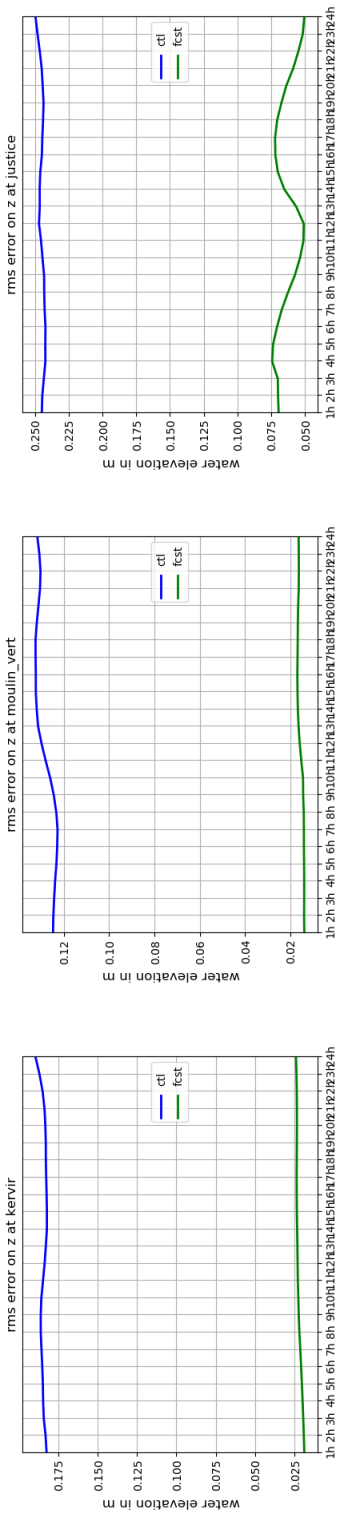


Figure R.13: Expérience $KsCL@all_zQ - 279$ - RMSE des erreurs de prévision aux stations Kervir, Moulin-Vert et Justice en fonction de l'échéance pour la simulation de contrôle (bleu) et la simulation avec assimilation de données (vert) par rapport à la simulation de référence.

R.2 Expériences réelles

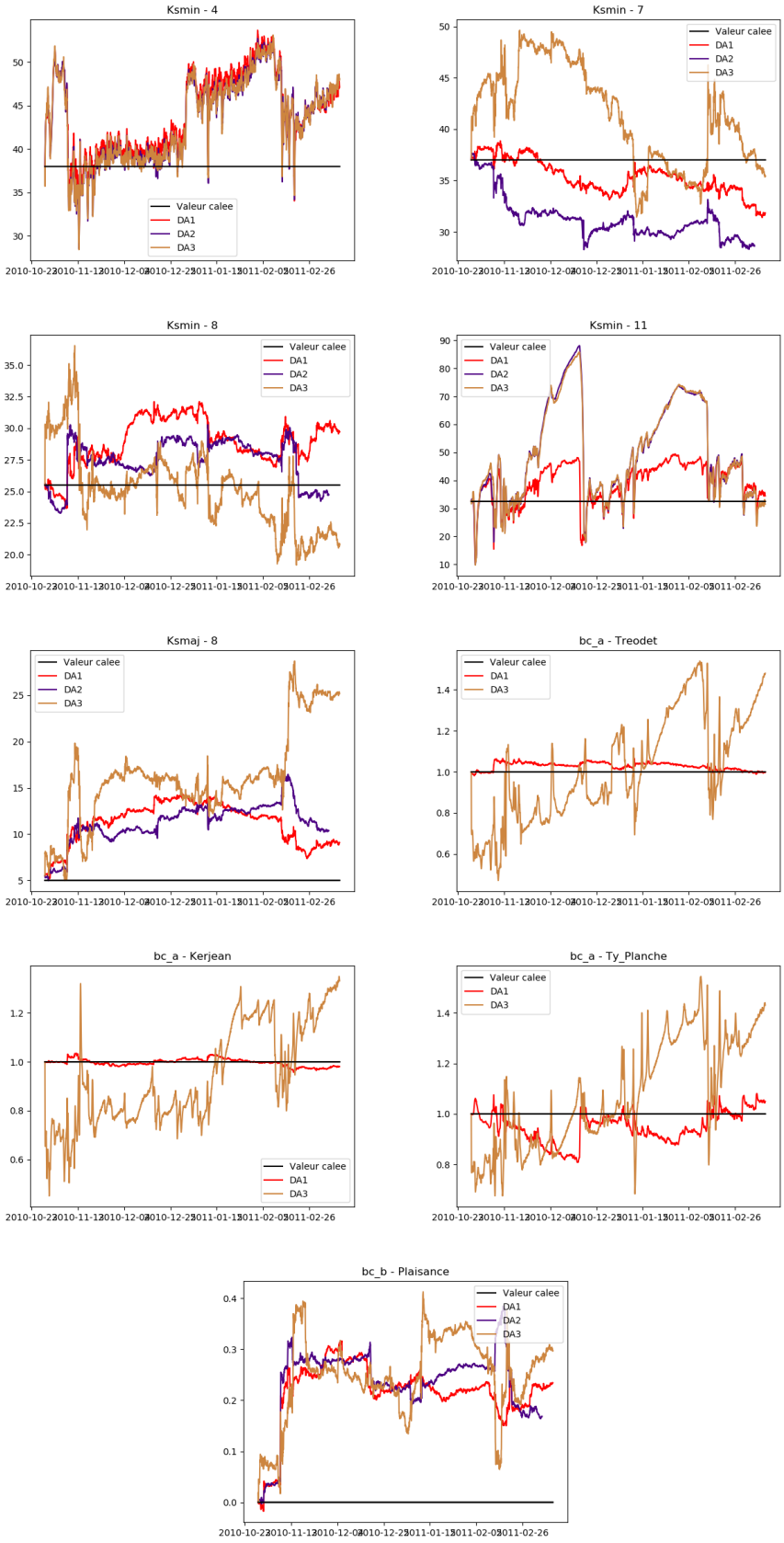


Figure R.14: Valeurs analysées pour les éléments du vecteur de contrôle pour les expériences *DA1*, *DA2* et *DA3* pour la période du 29 octobre 2010 au 15 mars 2011.

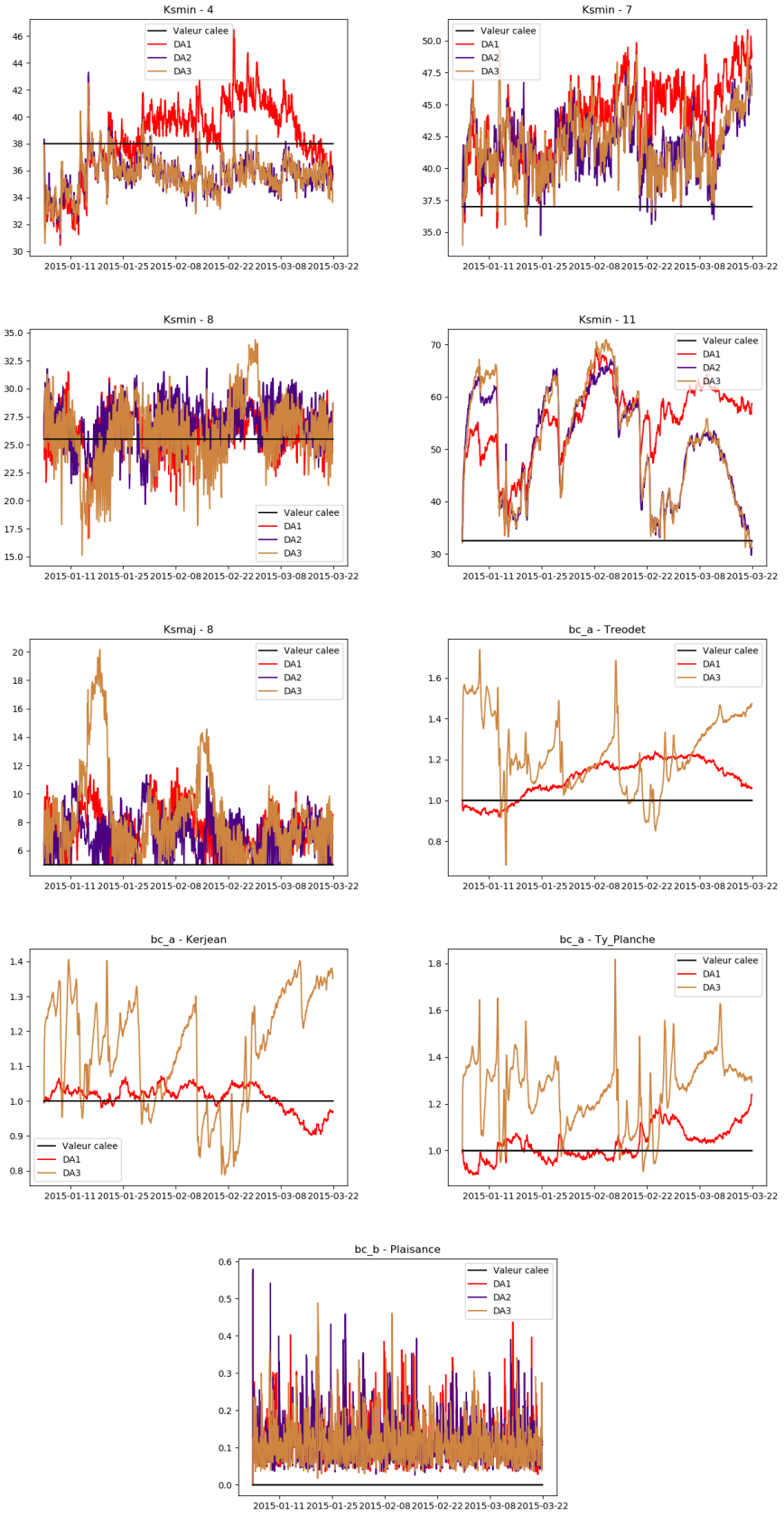


Figure R.15: Valeurs analysées pour les éléments du vecteur de contrôle pour les expériences *DA1*, *DA2* et *DA3* pour la période du 4 janvier 2015 au 26 mars 2015.

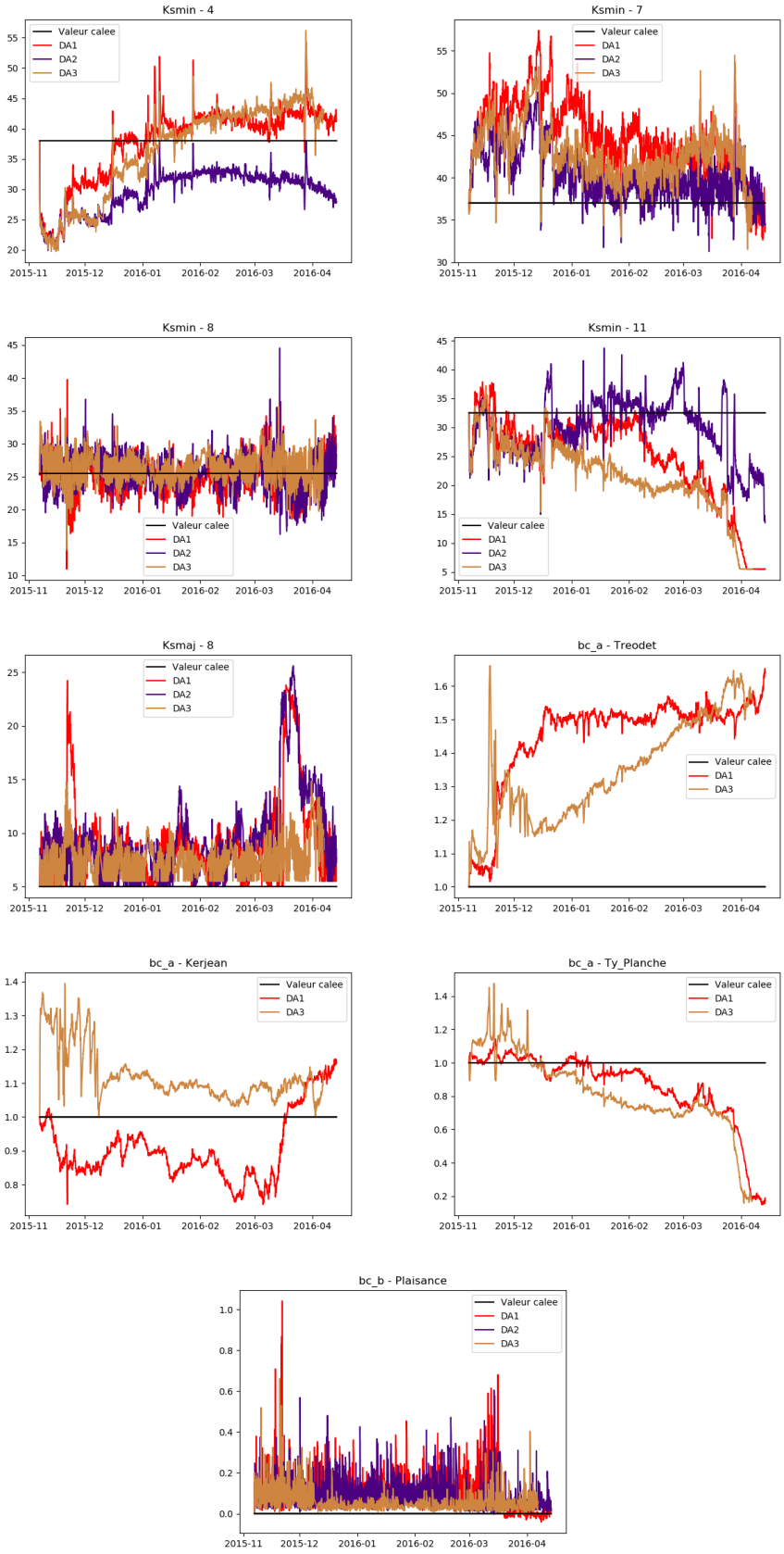


Figure R.16: Valeurs analysées pour les éléments du vecteur de contrôle pour les expériences *DA1*, *DA2* et *DA3* pour la période du 7 novembre 2015 au 15 avril 2016.

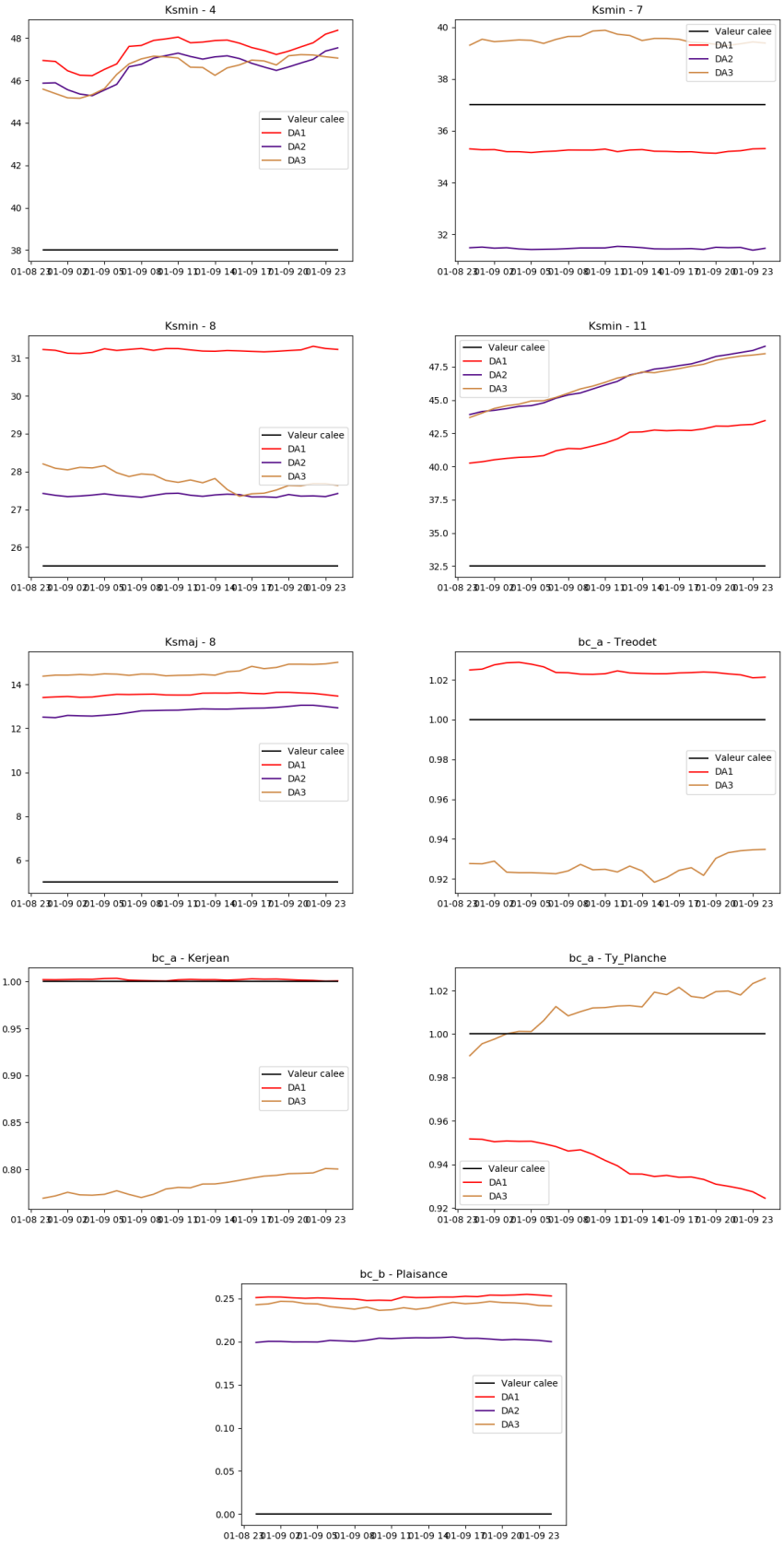


Figure R.17: Valeurs analysées pour les expériences *DA1*, *DA2* et *DA3* pour la journée du 9 janvier 2011

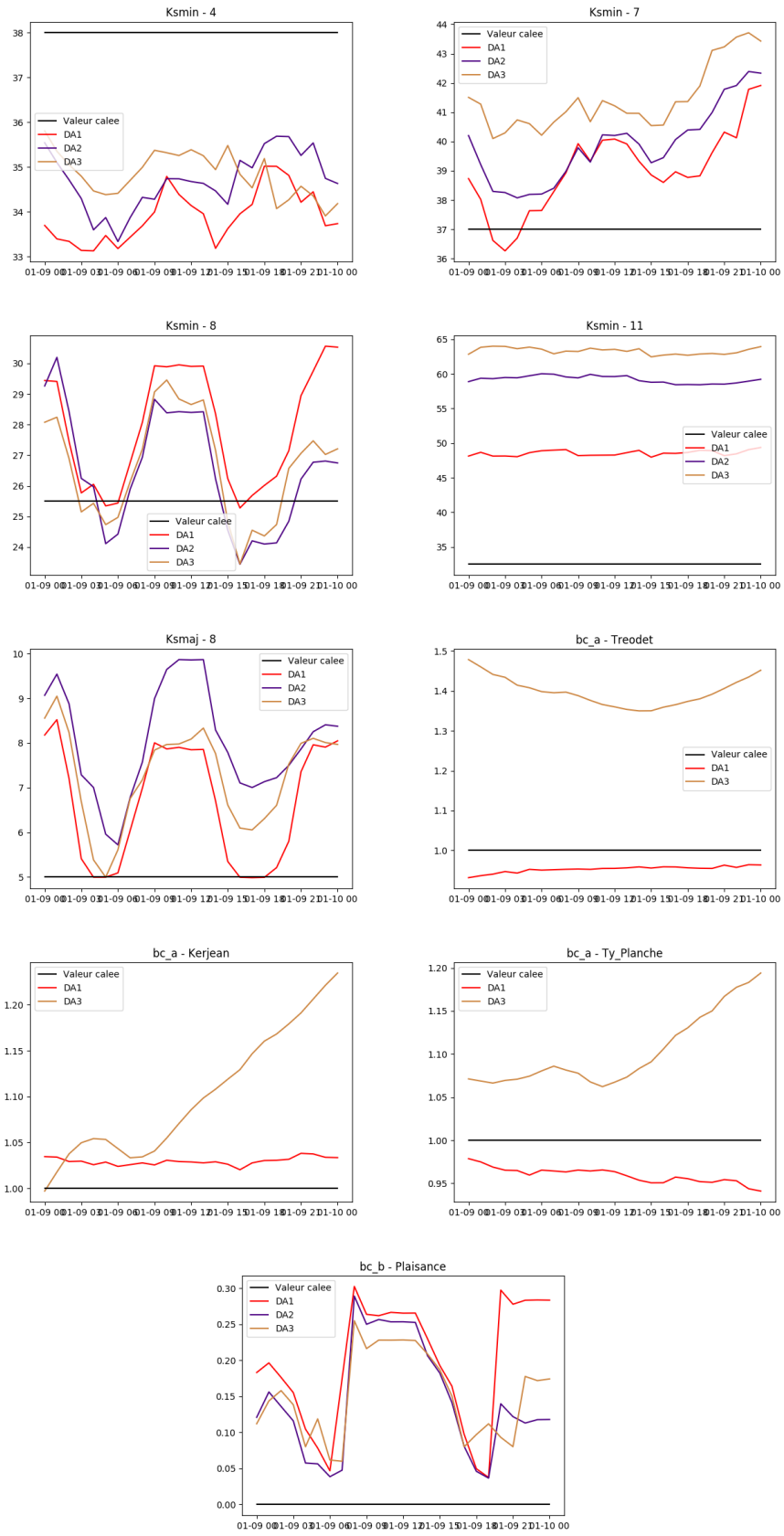


Figure R.18: Valeurs analysées pour les expériences *DA1*, *DA2* et *DA3* pour la journée du 9 janvier 2015

Annexe S

Compléments au chapitre 10

Sommaire

S.1 Ensembles à Kervir 426

S.1.1 Histogrammes de rang 426

S.1.2 Graphiques sur des événements 428

S.2 Ensembles à Moulin-Vert 432

S.2.1 Histogrammes de rang 432

S.2.2 Graphiques sur des événements 434

S.3 Ensembles à Justice 439

S.3.1 Histogrammes de rang 439

S.3.2 Graphiques des pics de pleine mer 440

S.1 Ensembles à Kervir

S.1.1 Histogrammes de rang

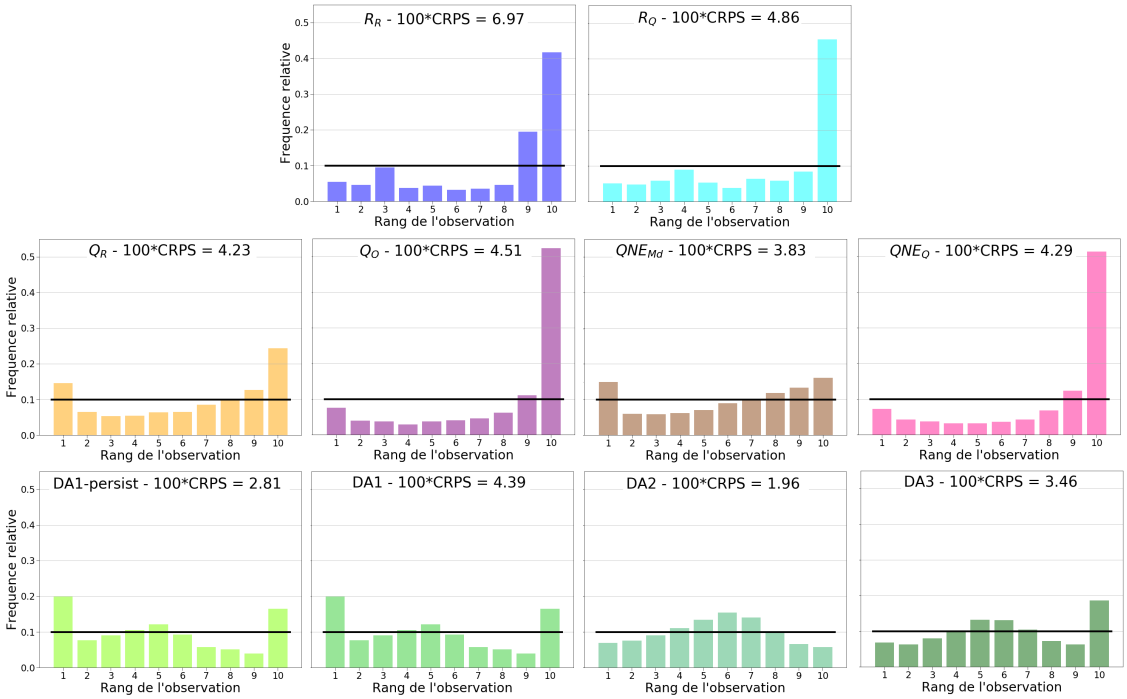


Figure S.1: Histogrammes de rang pour les 3 hivers confondus à l'échéance 6 heures à Kervir.

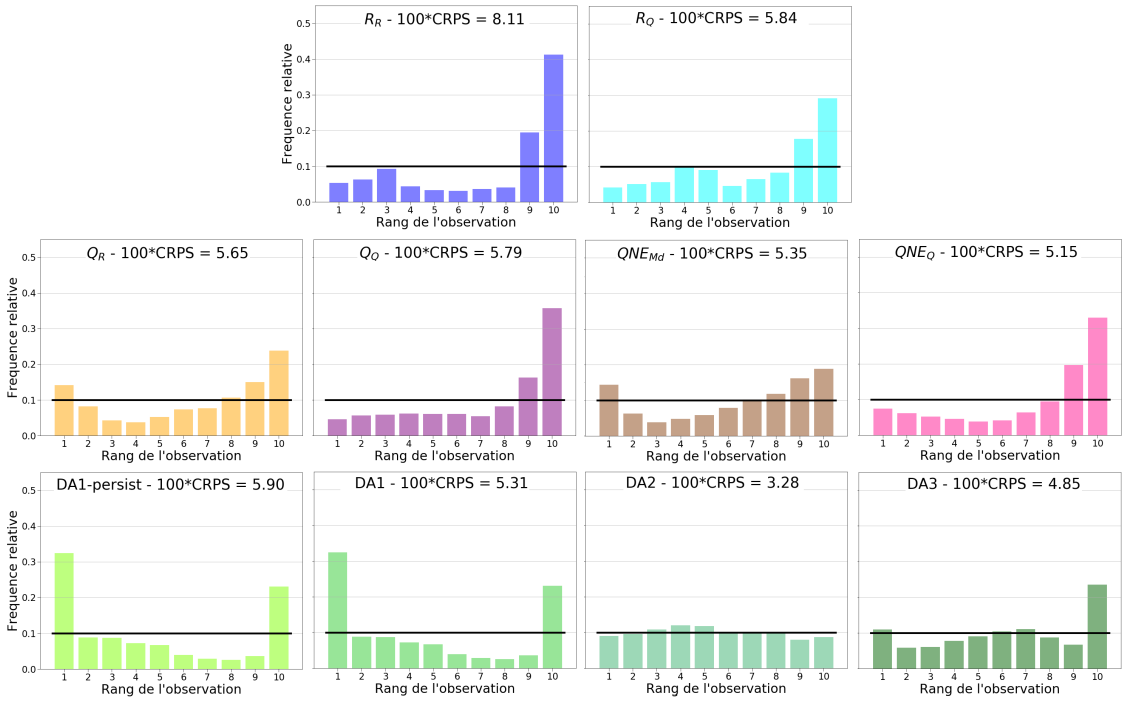


Figure S.2: Histogrammes de rang pour les 3 hivers confondus à l'échéance 12 heures à Kervir.

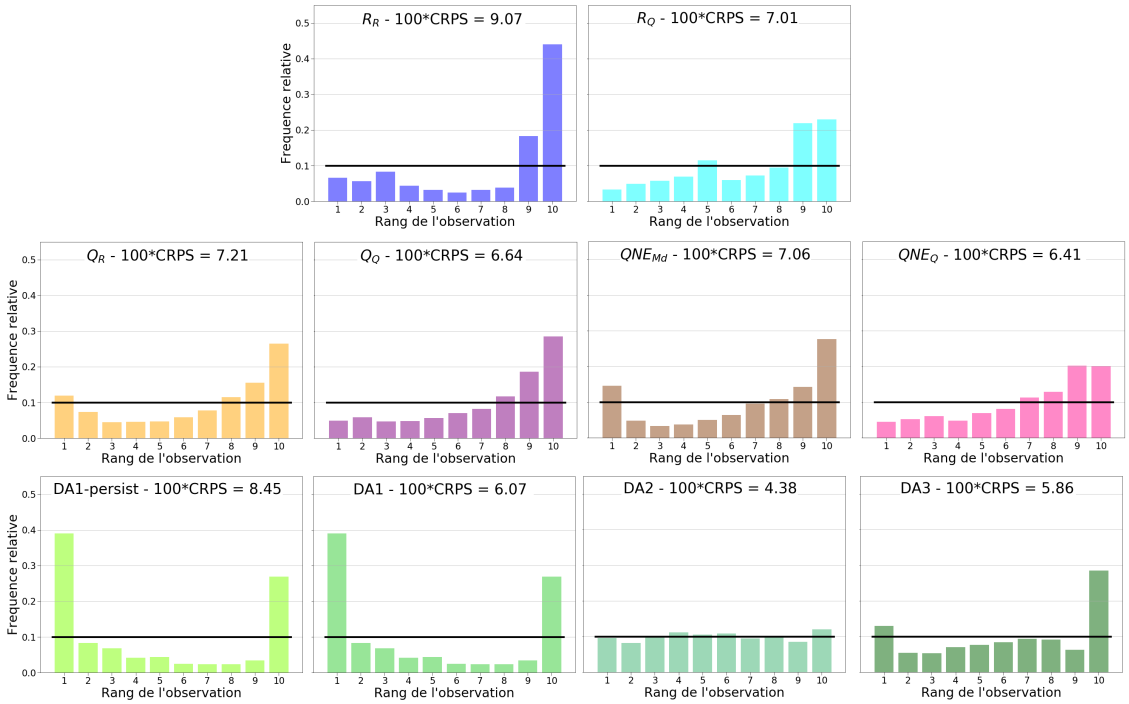


Figure S.3: Histogrammes de rang pour les 3 hivers confondus à l'échéance 18 heures à Kervir.

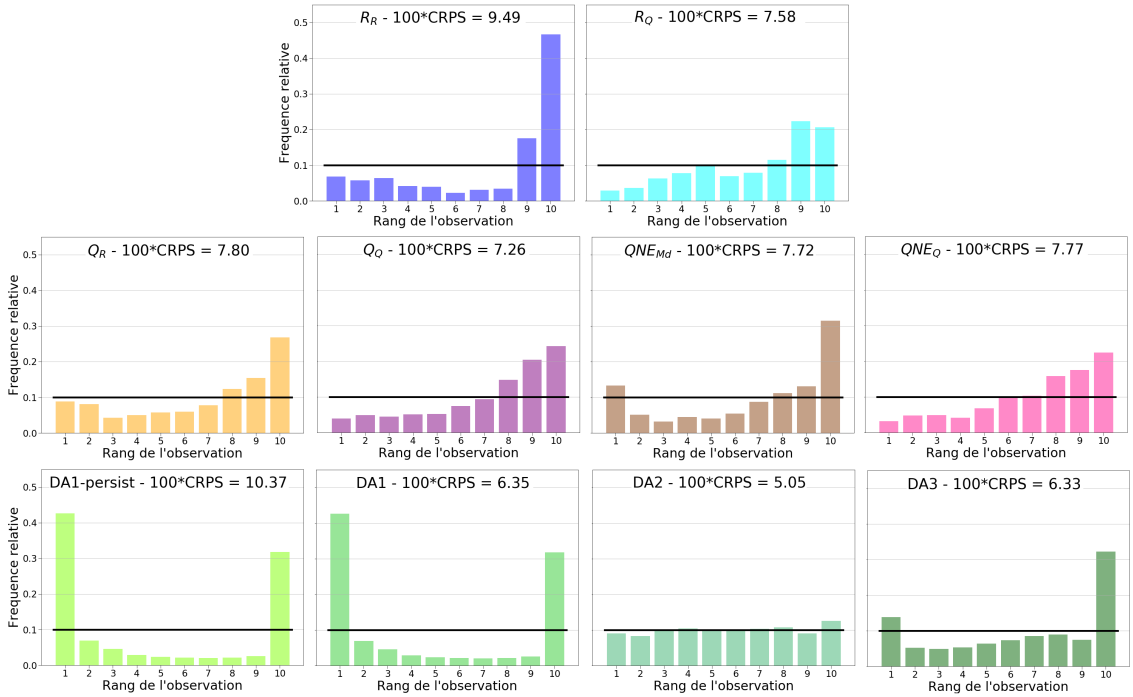


Figure S.4: Histogrammes de rang pour les 3 hivers confondus à l'échéance 24 heures à Kervir.

S.1.2 Graphiques sur des événements

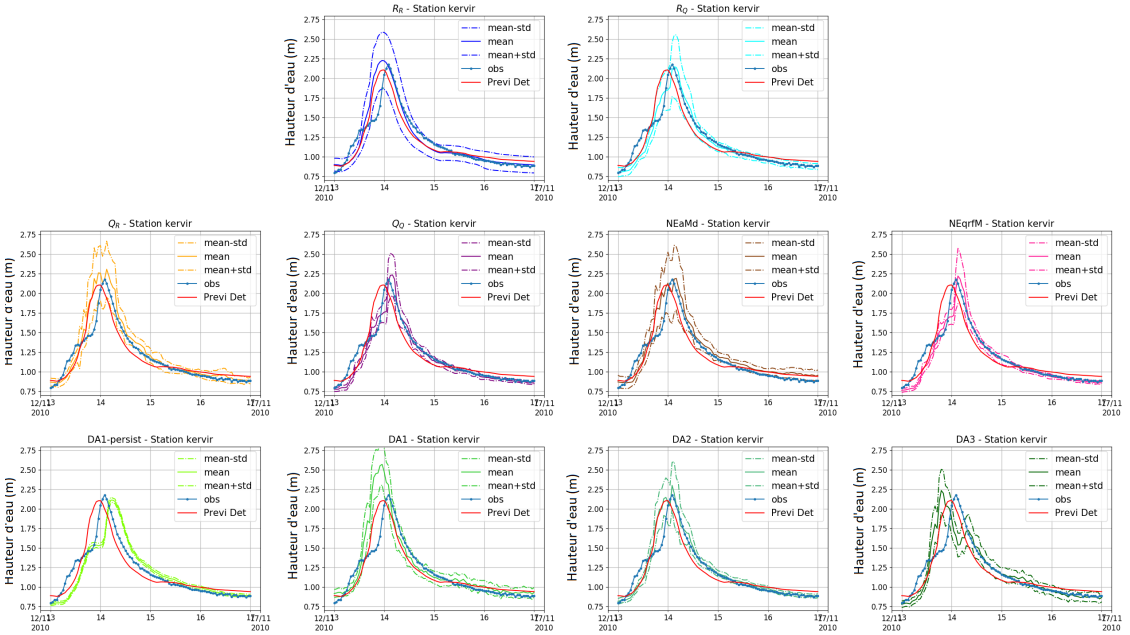


Figure S.5: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 6 heures à Kervir. Événement du 13 au 17 novembre 2010.

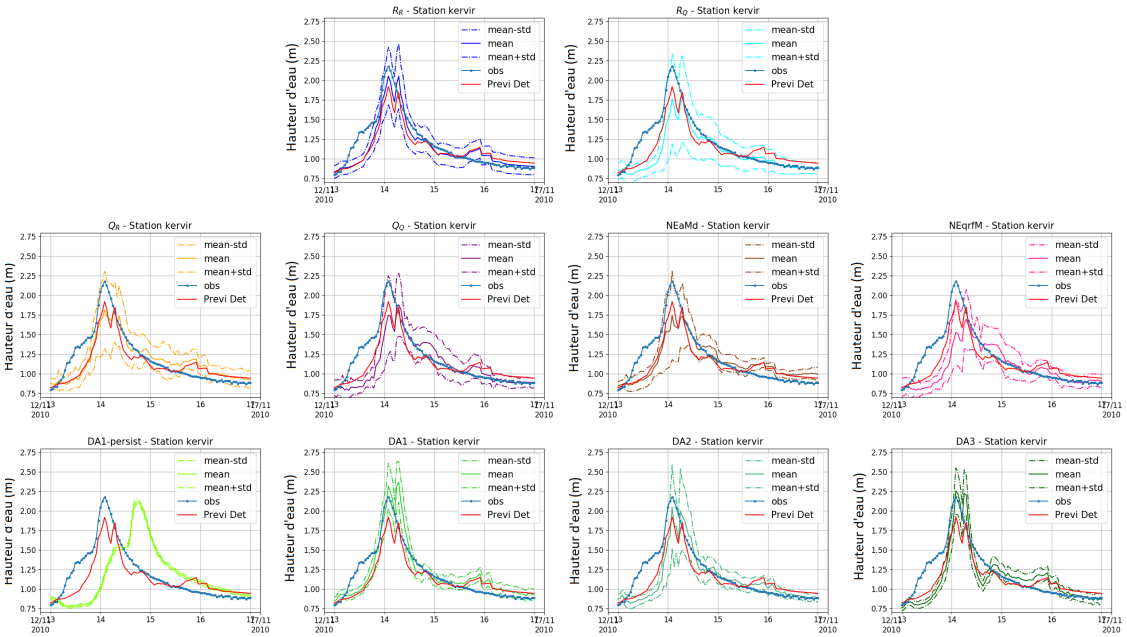


Figure S.6: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 18 heures à Kervir. Événement du 13 au 17 novembre 2010.

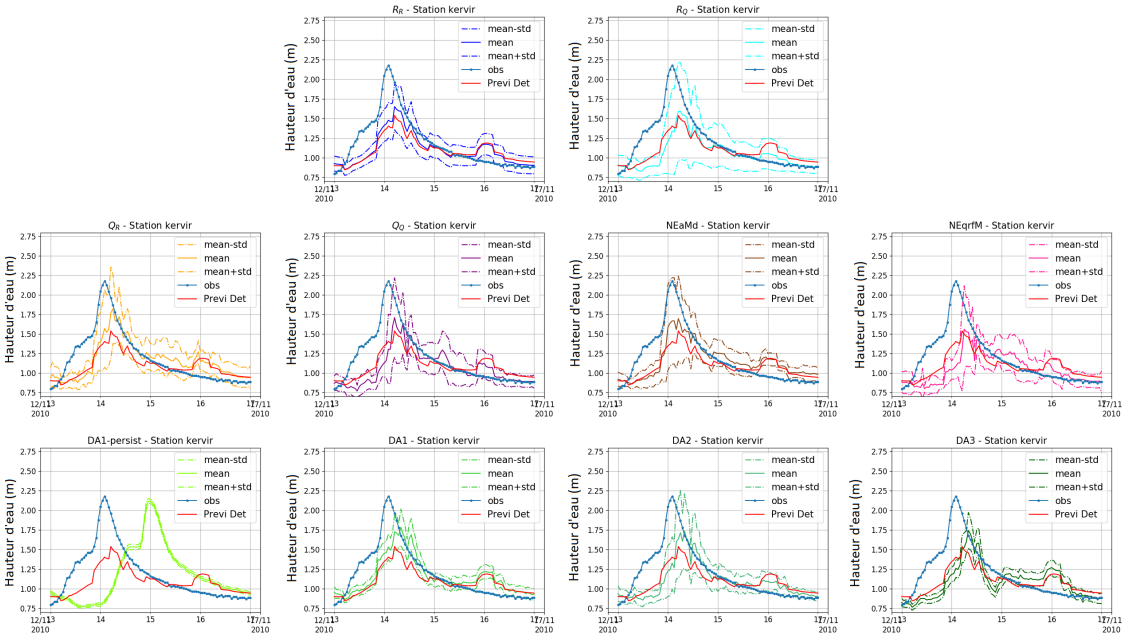


Figure S.7: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 24 heures à Kervir. Événement du 13 au 17 novembre 2010.

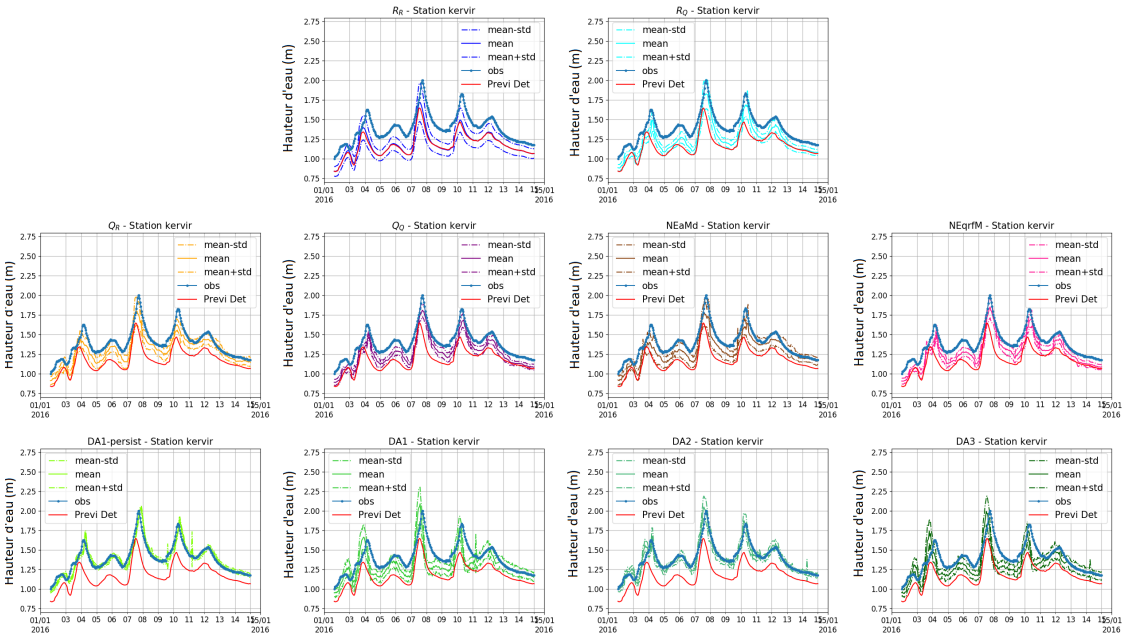


Figure S.8: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 6 heures à Kervir. Événement du 2 au 15 janvier 2016.

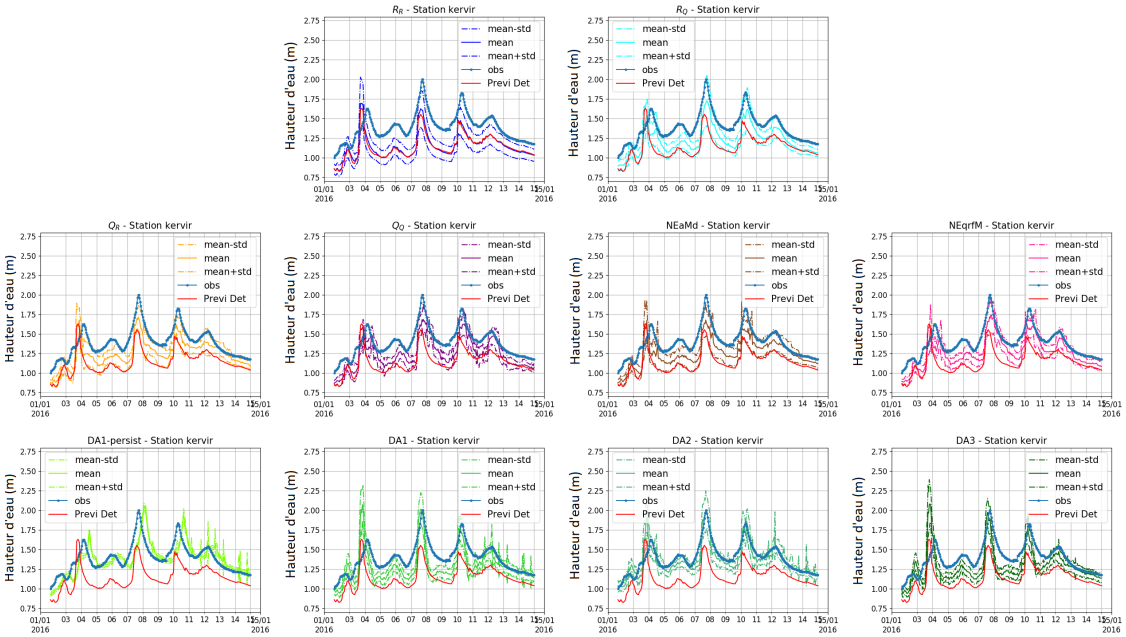


Figure S.9: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 12 heures à Kervir. Événement du 2 au 15 janvier 2016.

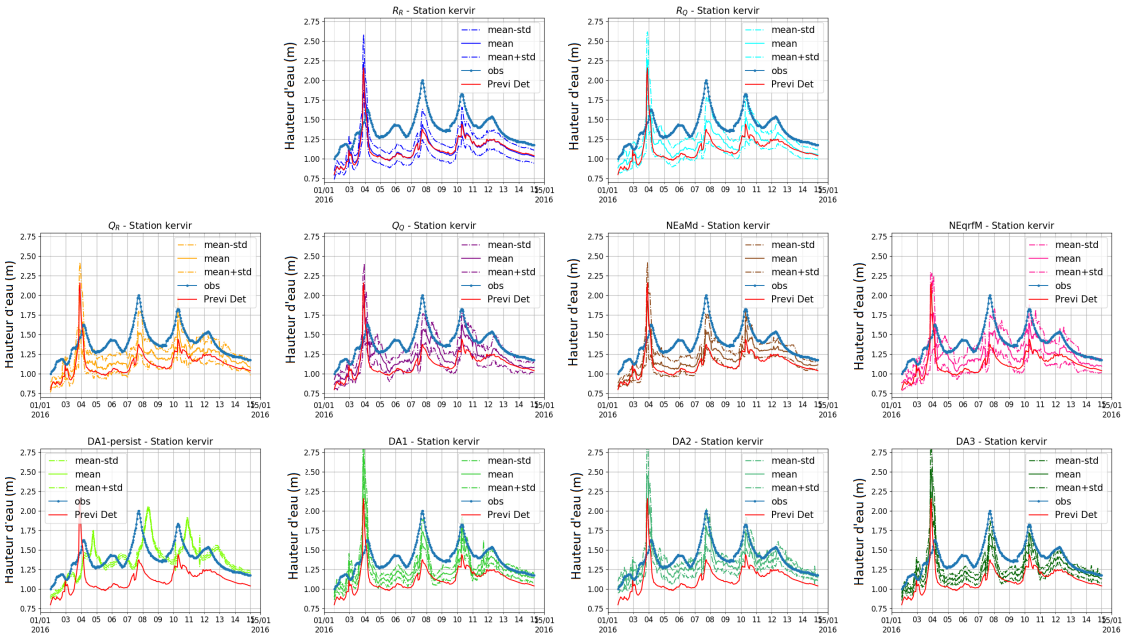


Figure S.10: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 18 heures à Kervir. Événement du 2 au 15 janvier 2016.

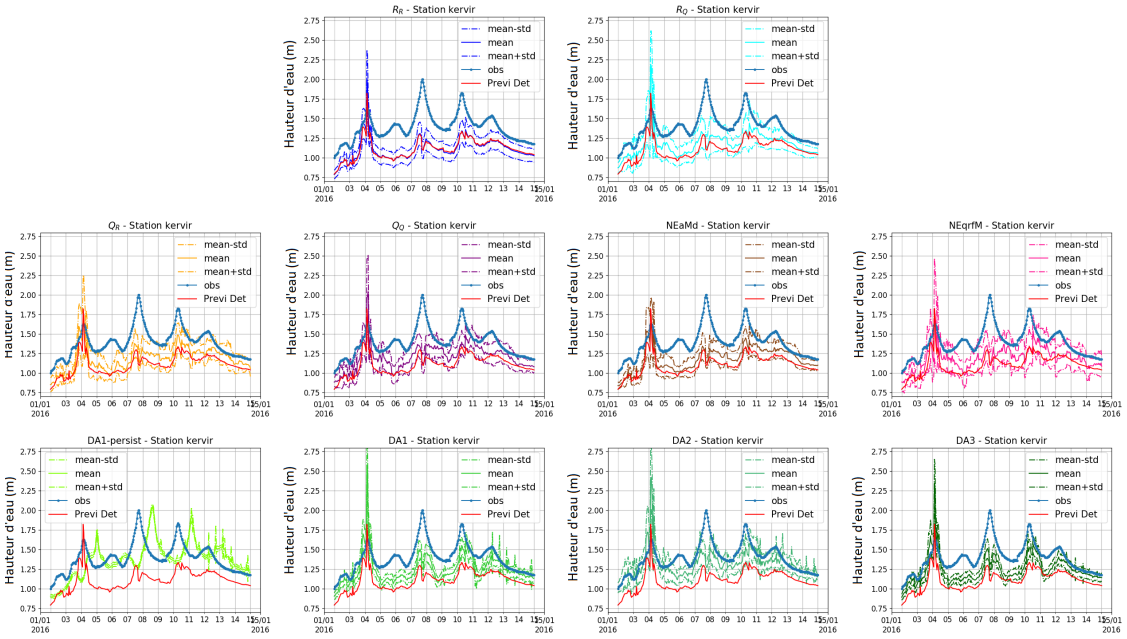


Figure S.11: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 24 heures à Kervir. Événement du 2 au 15 janvier 2016.

S.2 Ensembles à Moulin-Vert

S.2.1 Histogrammes de rang

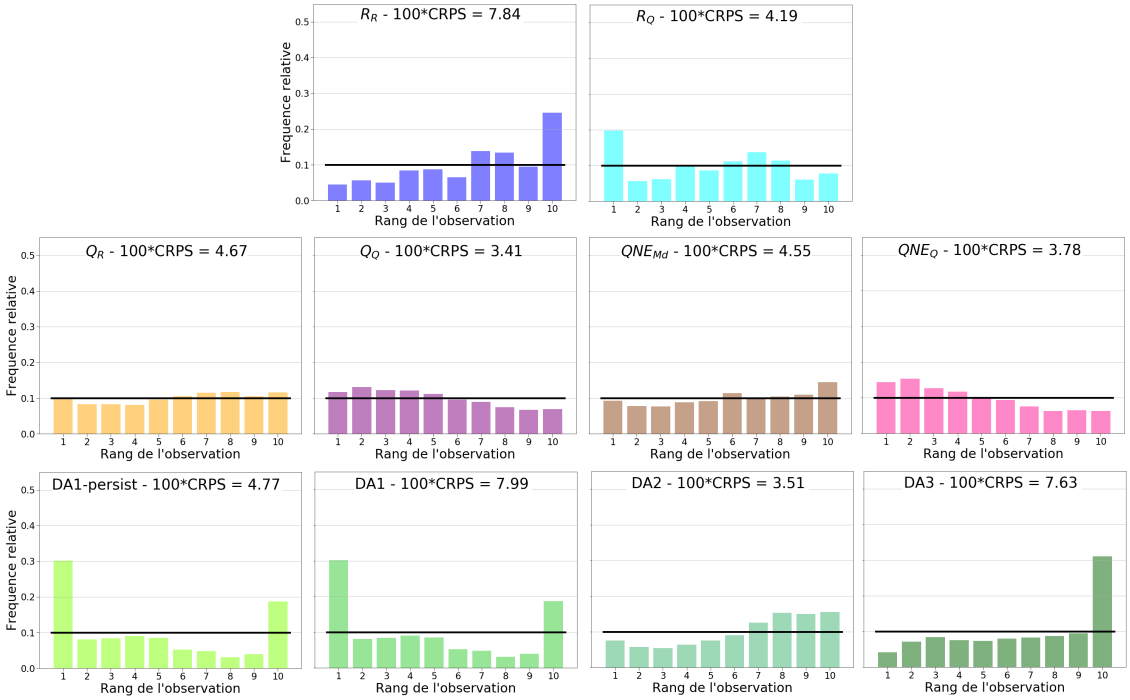


Figure S.12: Histogrammes de rang pour les 3 hivers confondus à l'échéance 6 heures à Moulin-Vert.

Figure S.13: Histogrammes de rang pour les 3 hivers confondus à l'échéance 12 heures à Moulin-Vert.

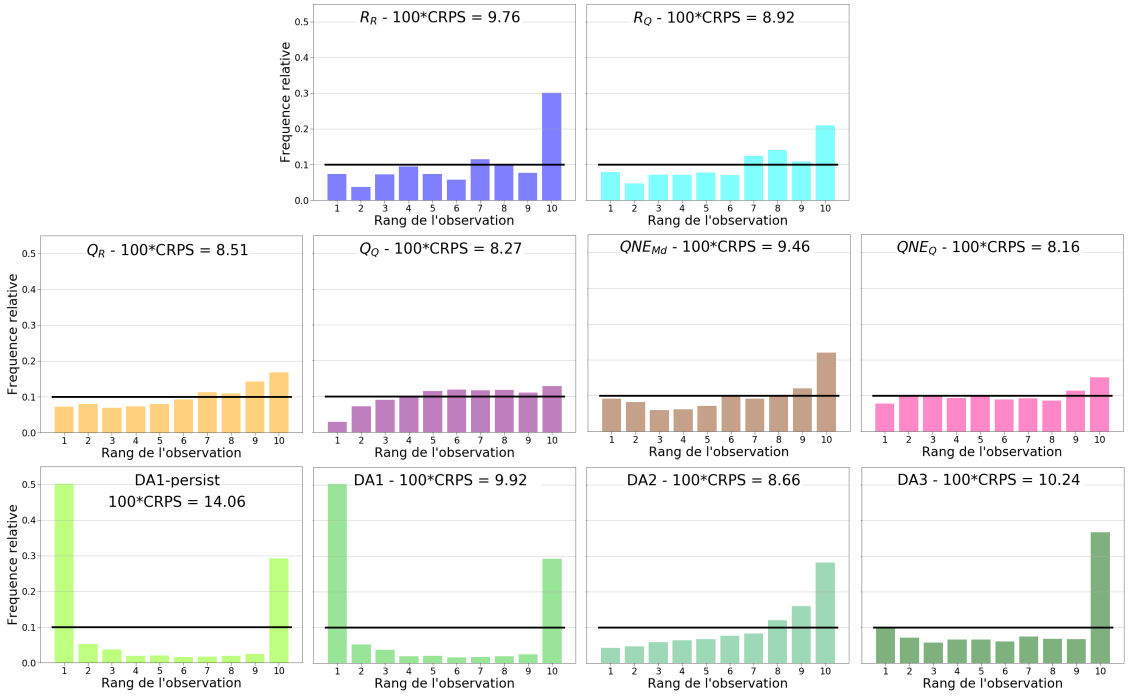


Figure S.14: Histogrammes de rang pour les 3 hivers confondus à l'échéance 18 heures à Moulin-Vert.

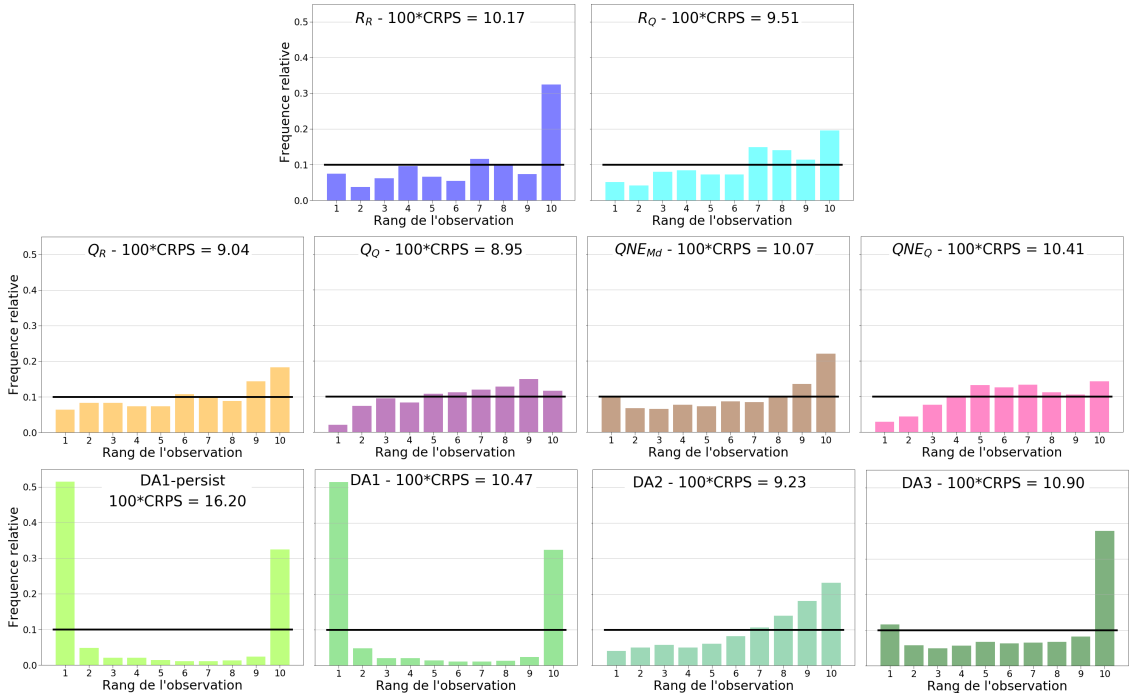


Figure S.15: Histogrammes de rang pour les 3 hivers confondus à l'échéance 24 heures à Moulin-Vert.

S.2.2 Graphiques sur des événements

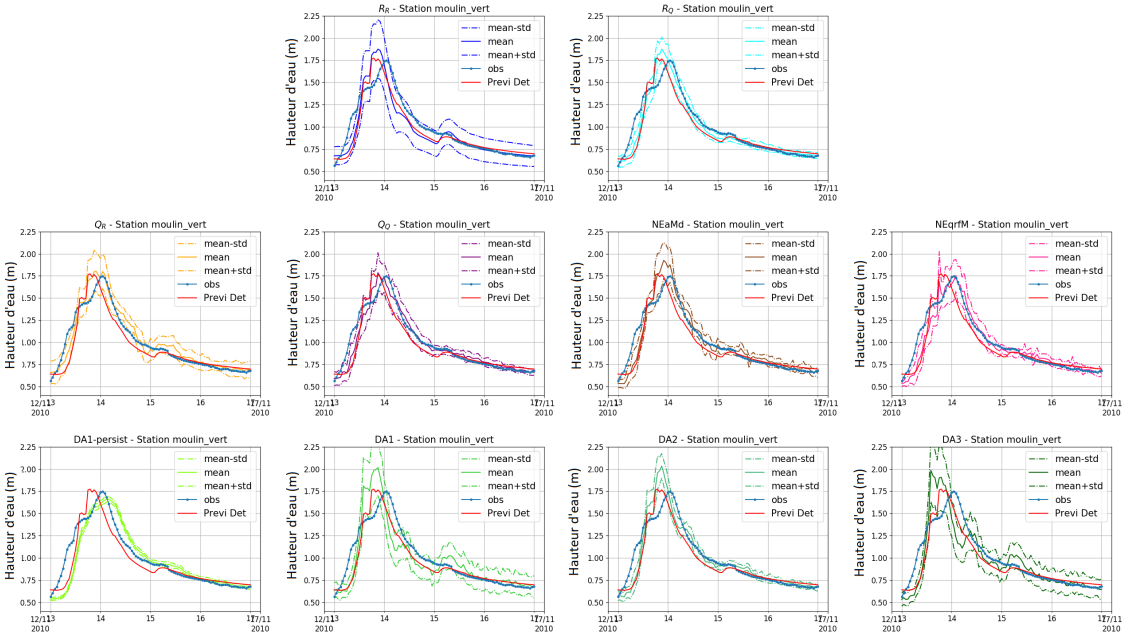


Figure S.16: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 6 heures à Moulin-vert. Événement du 13 au 17 novembre 2010.

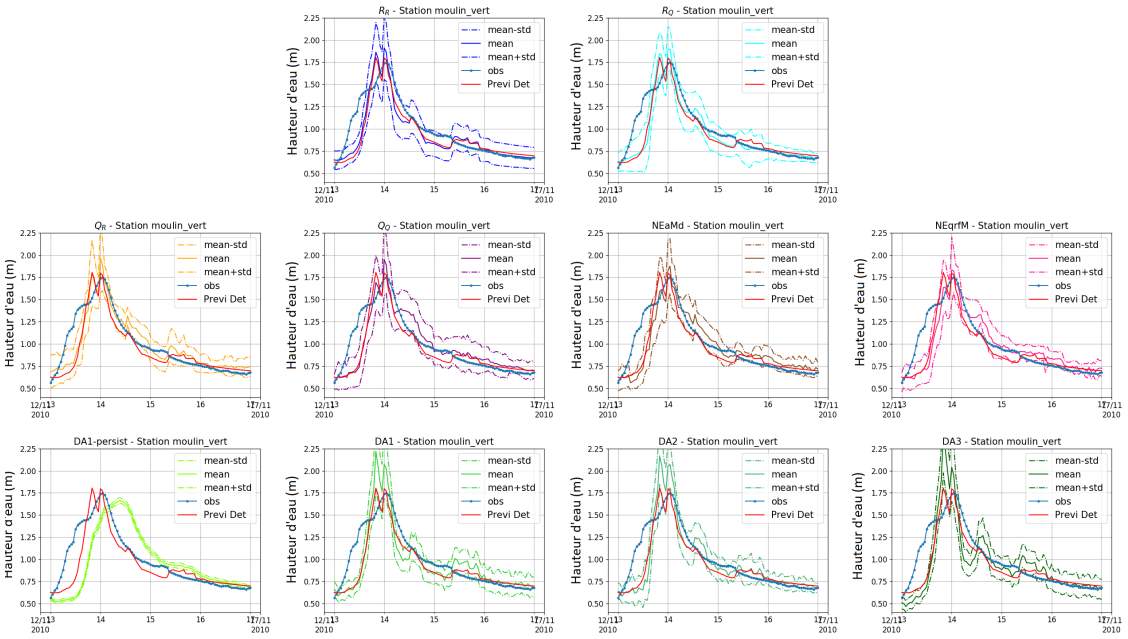


Figure S.17: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 12 heures à Moulin-vert. Événement du 13 au 17 novembre 2010.

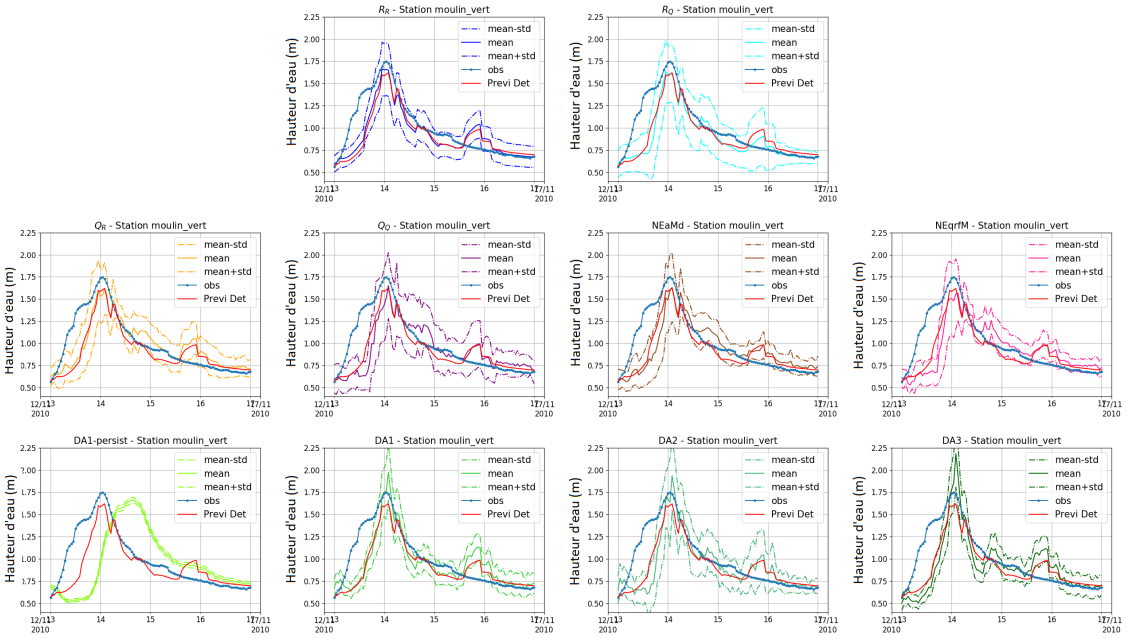


Figure S.18: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 18 heures à Moulin-vert. Événement du 13 au 17 novembre 2010.

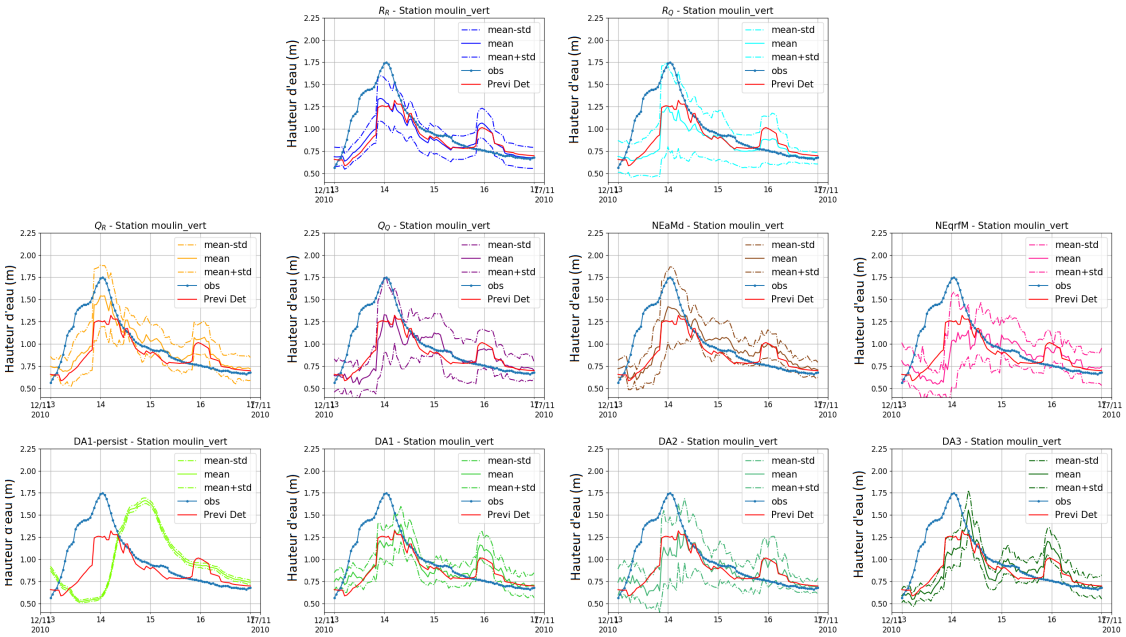


Figure S.19: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 24 heures à Moulin-vert. Événement du 13 au 17 novembre 2010.

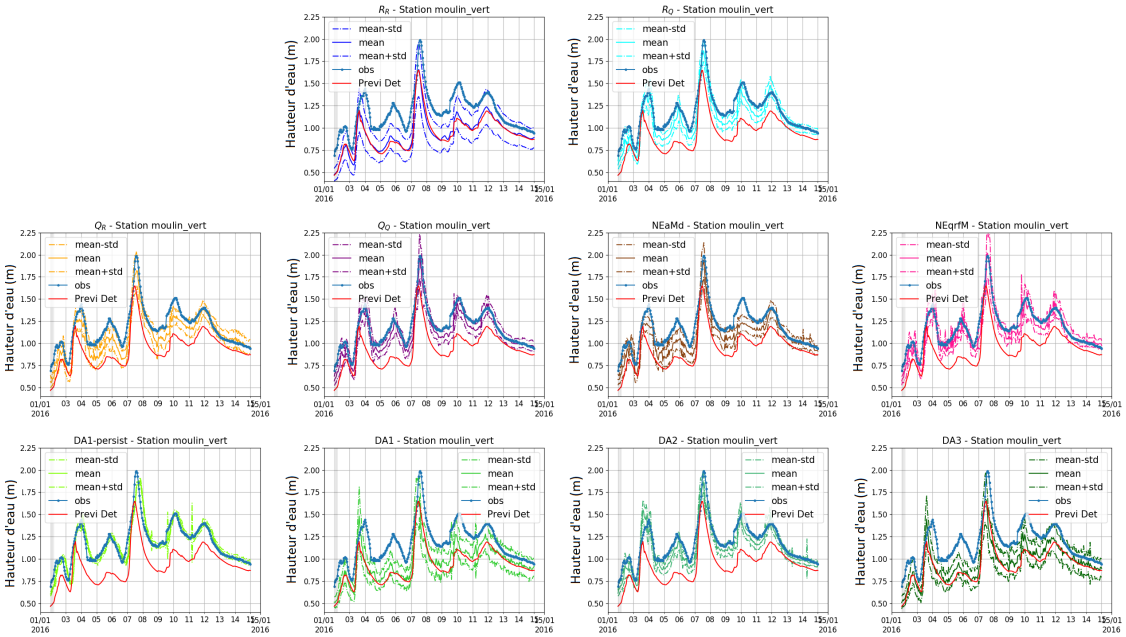


Figure S.20: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 6 heures à Moulin-vert. Événement du 2 au 15 janvier 2016.

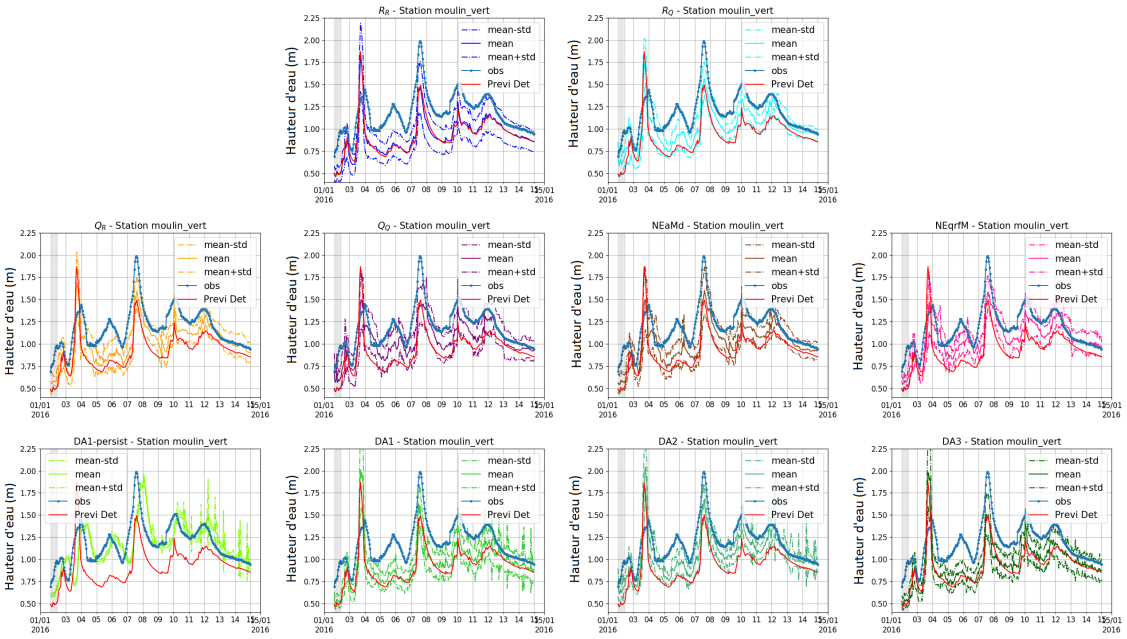


Figure S.21: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 12 heures à Moulin-vert. Événement du 2 au 15 janvier 2016.

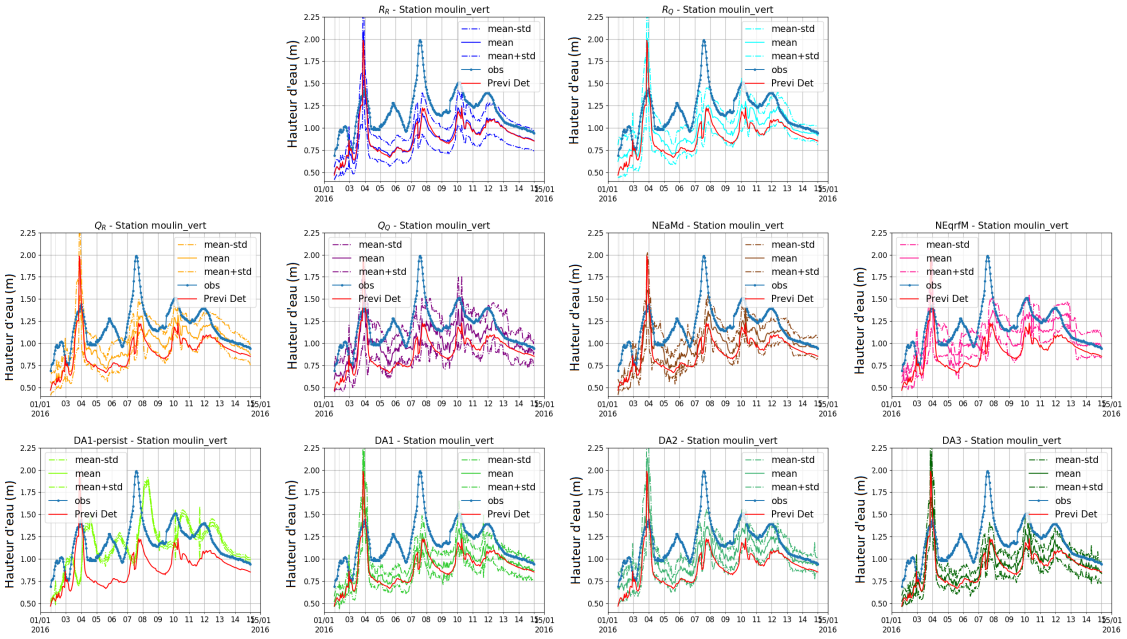


Figure S.22: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 18 heures à Moulin-vert. Événement du 2 au 15 janvier 2016.

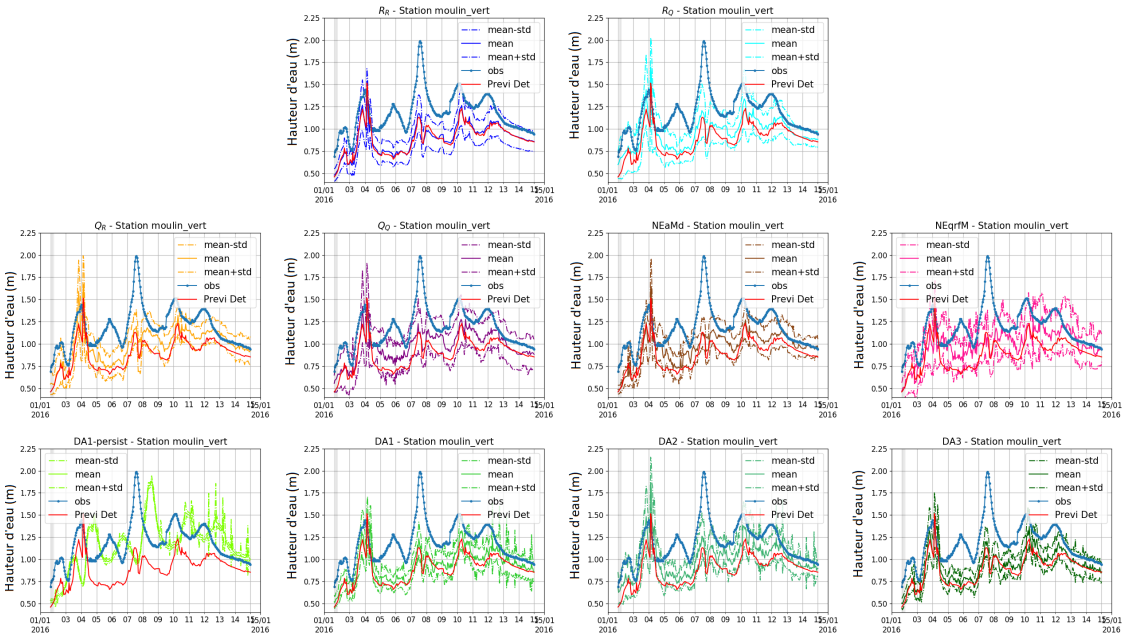


Figure S.23: Moyenne, moyenne +/- écart-type pour les prévisions cyclées à échéance 24 heures à Moulin-Vert. Événement du 2 au 15 janvier 2016.

S.3 Ensembles à Justice

S.3.1 Histogrammes de rang

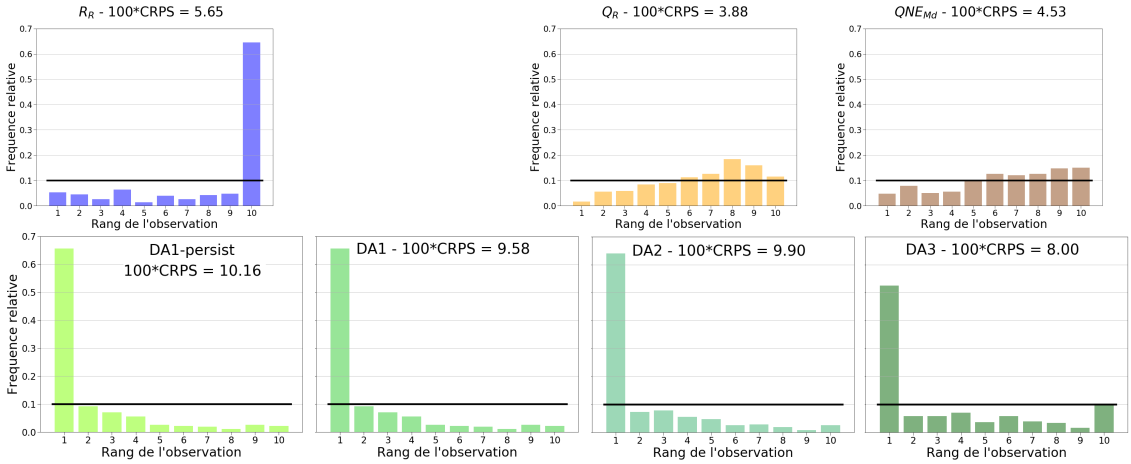


Figure S.24: Histogrammes de rang pour les 3 hivers confondus à l'échéance 6 heures à Justice.

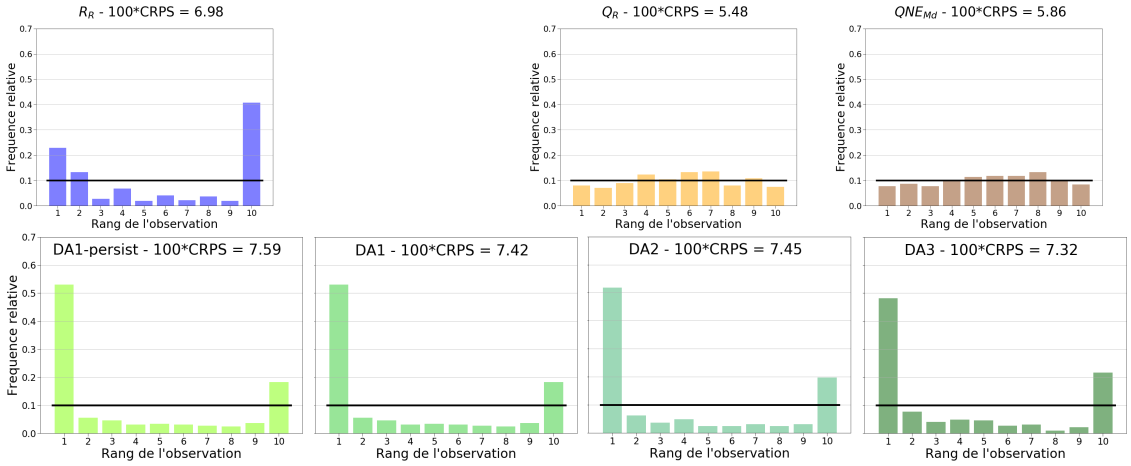


Figure S.25: Histogrammes de rang pour les 3 hivers confondus à l'échéance 12 heures à Justice.

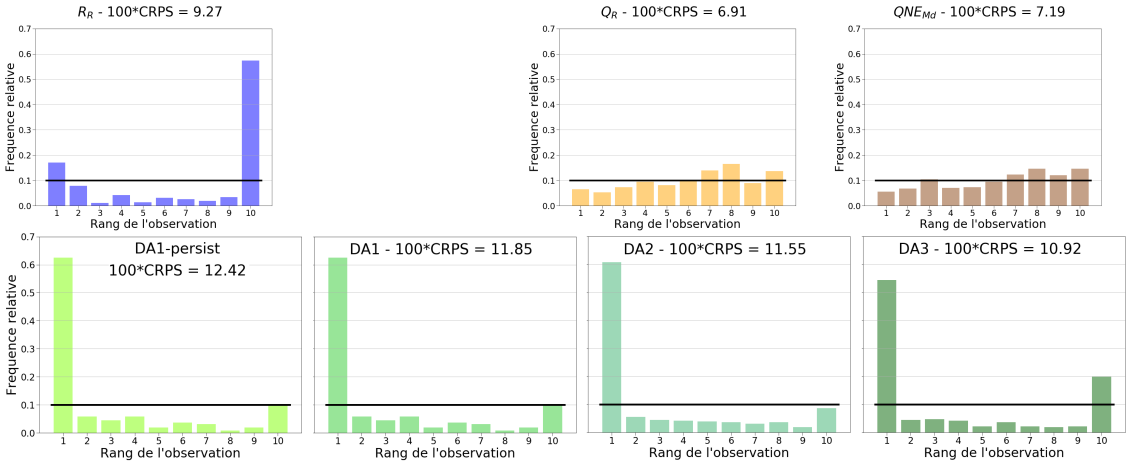


Figure S.26: Histogrammes de rang pour les 3 hivers confondus à l'échéance 18 heures à Justice.

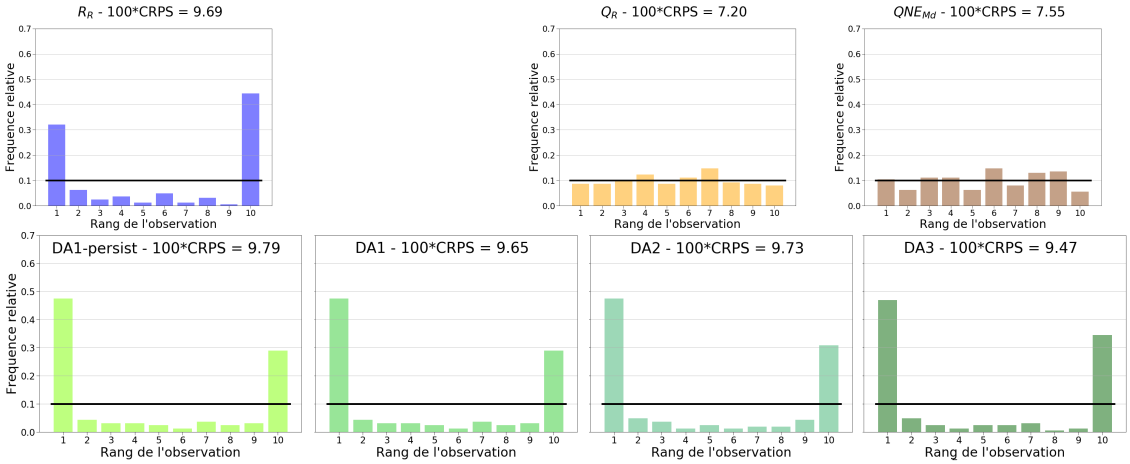


Figure S.27: Histogrammes de rang pour les 3 hivers confondus à l'échéance 24 heures à Justice.

S.3.2 Graphiques des pics de pleine mer

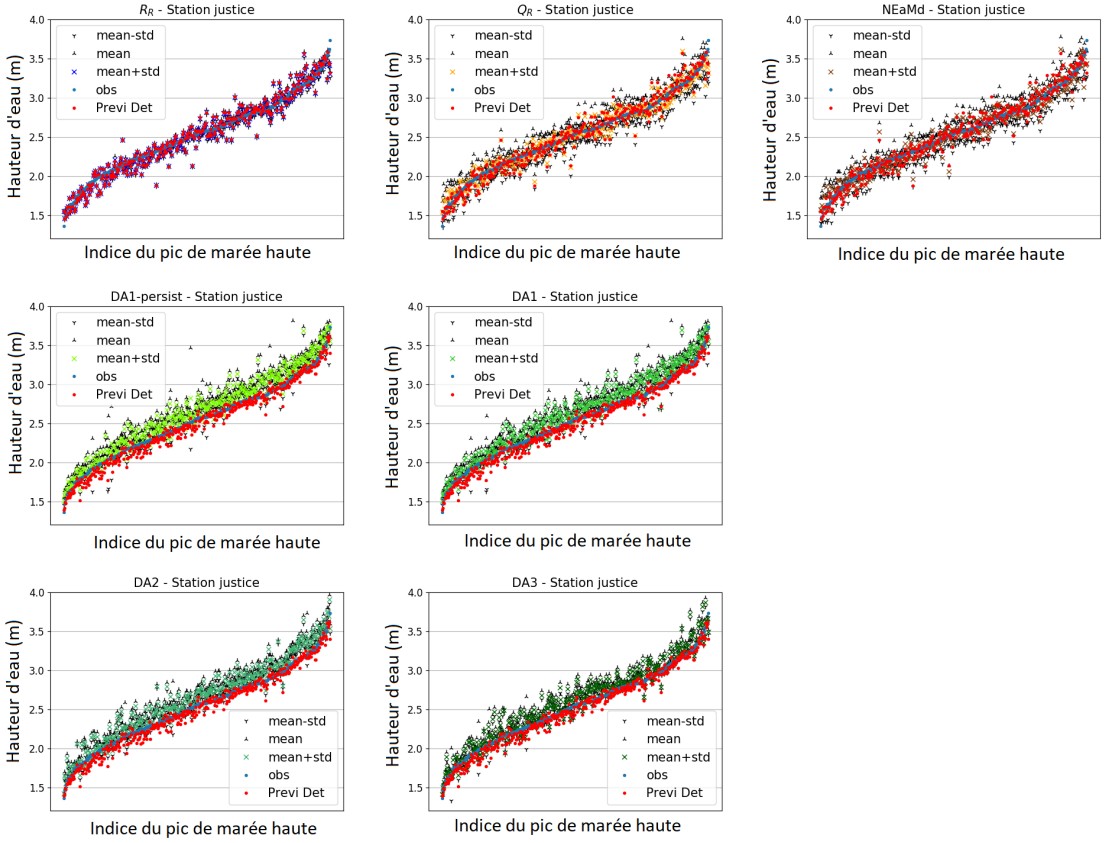


Figure S.28: Moyenne, moyenne +/- écart-type pour les prévisions cyclées des pics de marée haute à échéance 6 heures à Justice. Pics de marée haute sur les 2 hivers de validation confondus classés par valeurs croissantes.

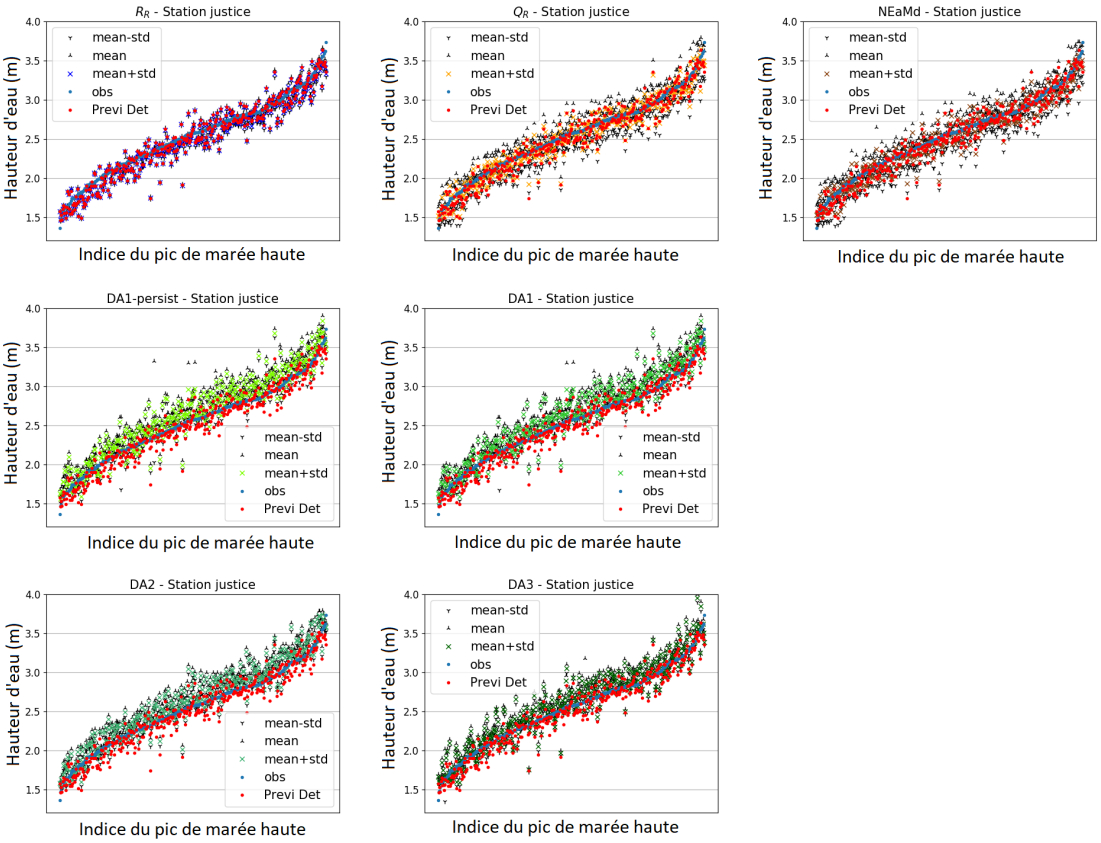


Figure S.29: Moyenne, moyenne +/- écart-type pour les prévisions cyclées des pics de marée haute à échéance 18 heures à Justice. Pics de marée haute sur les 2 hivers de validation confondus classés par valeurs croissantes.

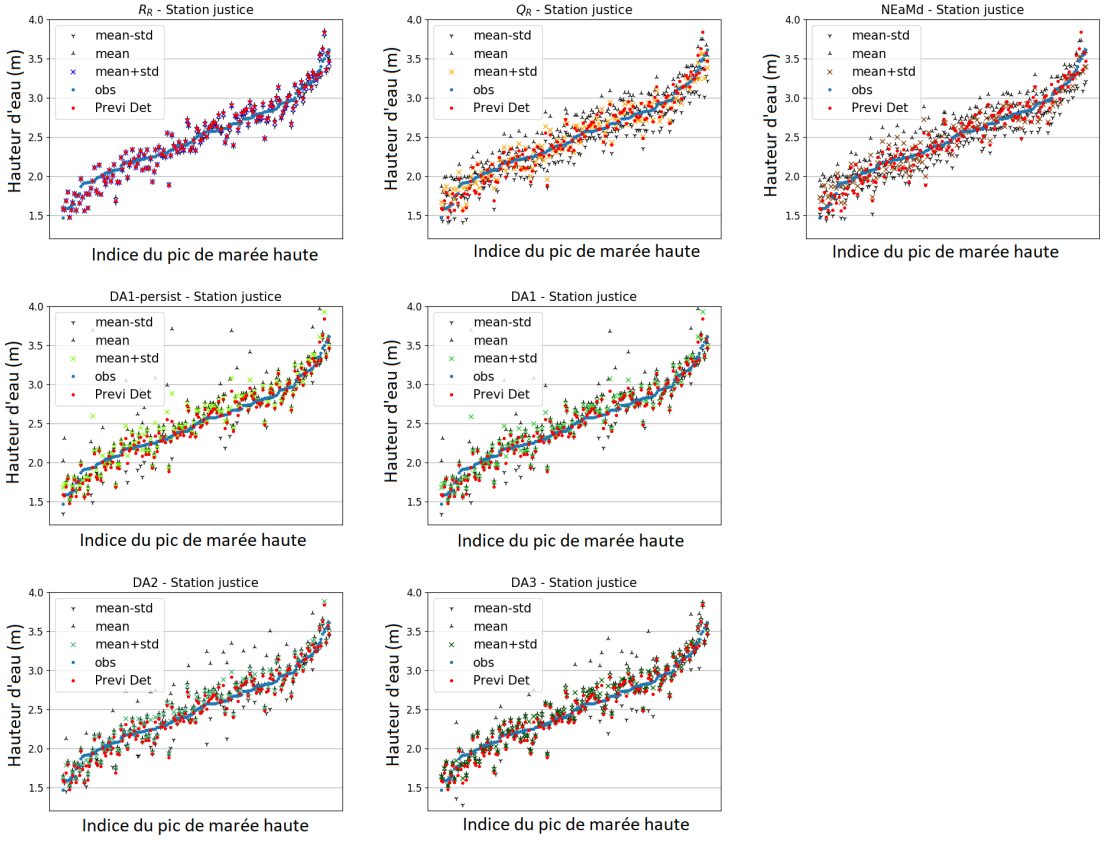


Figure S.30: Moyenne, moyenne +/- écart-type pour les prévisions cyclées des pics de marée haute à échéance 24 heures à Justice. Pics de marée haute sur les 2 hivers de validation confondus classés par valeurs croissantes.

Références

L. Alfieri, F. Pappenberger, F. Wetterhall, T. Haiden, D. Richardson, and P. Salamon. Evaluation of ensemble streamflow predictions in Europe. *Journal of Hydrology*, 517: 913–922, 2014. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2014.06.035. URL https://www.sciencedirect.com/science/article/pii/S0022169414004958.

E. A. Anderson. National Weather Service river forecast system: snow accumulation and ablation model. 1973. URL https://repository.library.noaa.gov/view/noaa/13507.

J. L. Anderson. A Method for Producing and Evaluating Probabilistic Forecasts from Ensemble Model Integrations. *Journal of Climate*, 9(7):1518–1530, 1996. doi: 10.1175/1520-0442(1996)009<1518:AMFPAE>2.0.CO;2. URL https://doi.org/10.1175/1520-0442(1996)009<1518:AMFPAE>2.0.CO;2. Number: 7.

J. L. Anderson and S. L. Anderson. A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts. *MONTHLY WEATHER REVIEW*, 127:18, 1999.

K. M. Andreadis and D. P. Lettenmaier. Assimilating remotely sensed snow observations into a macroscale hydrology model. *Advances in Water Resources*, 29(6): 872–886, June 2006. ISSN 03091708. doi: 10.1016/j.advwatres.2005.08.004. URL https://linkinghub.elsevier.com/retrieve/pii/S0309170805002058.

K. M. Andreadis, E. A. Clark, D. P. Lettenmaier, and D. E. Alsdorf. Prospects for river discharge and depth estimation through assimilation of swath-altimetry into a raster-based hydrodynamics model. *Geophysical Research Letters*, 34(10):L10403, May 2007. ISSN 0094-8276. doi: 10.1029/2007GL029721. URL http://doi.wiley.com/10.1029/2007GL029721.

G. E. B. Archer, A. Saltelli, and I. M. Sobol. Sensitivity measures,anova-like Techniques and the use of bootstrap. *Journal of Statistical Computation and Simulation*, 58(2): 99–120, May 1997. ISSN 0094-9655, 1563-5163. doi: 10.1080/00949659708811825. URL http://www.tandfonline.com/doi/abs/10.1080/00949659708811825.

J. G. Arnold, R. Srinivasan, R. S. Muttiah, and J. R. Williams. LARGE AREA HYDROLOGIC MODELING AND ASSESSMENT PART I: MODEL DEVELOPMENT1. *JAWRA Journal of the American Water Resources Association*, 34(1):73–89, 1998. doi: https://doi.org/10.1111/j.1752-1688.1998.tb05961.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1752-1688.1998.tb05961.x.

M. Asch, M. Bocquet, and M. Nodet. *Data assimilation: methods, algorithms, and applications*. Fundamentals of Algorithms. SIAM, 2016. URL https://hal.inria.fr/hal-01402885.

L. Auger, O. Dupont, S. Hagelin, P. Brousseau, and P. Brovelli. AROME–NWC: a new nowcasting tool based on an operational mesoscale forecasting system. *Quarterly Journal of the Royal Meteorological Society*, 141(690):1603–1611, July 2015. ISSN 0035-9009, 1477-870X. doi: 10.1002/qj.2463. URL https://onlinelibrary.wiley.com/doi/10.1002/qj.2463.

G. Balsamo, A. Beljaars, K. Scipal, P. Viterbo, B. van den Hurk, M. Hirschi, and A. K. Betts. A Revised Hydrology for the ECMWF Model: Verification from Field Site to Terrestrial Water Storage and Impact in the Integrated Forecast System. *Journal of Hydrometeorology*, 10(3):623–643, June 2009. ISSN 1525-7541, 1525-755X. doi: 10.1175/2008JHM1068.1. URL http://journals.ametsoc.org/doi/10.1175/2008JHM1068.1.

C. Bandaragoda, D. G. Tarboton, and R. Woods. Application of TOPNET in the distributed model intercomparison project. *Journal of Hydrology*, 298(1-4):178–201, Oct. 2004. ISSN 00221694. doi: 10.1016/j.jhydrol.2004.03.038. URL https://linkinghub.elsevier.com/retrieve/pii/S0022169404002446.

S. Barthelemy. *Assimilation de données ensembliste et couplage de modèles hydrauliques 1D-2D pour la prévision des crues en temps réel. Application au réseau hydraulique "Adour maritime"*. PhD Thesis, 2015.

S. Barthélémy, S. Ricci, M. Rochoux, E. Le Pape, and O. Thual. Ensemble-based data assimilation for operational flood forecasting – On the merits of state estimation for 1D hydrodynamic forecasting through the example of the "Adour Maritime" river. *Journal of Hydrology*, 552:210–224, Sept. 2017. ISSN 00221694. doi: 10.1016/j.jhydrol.2017.06.017. URL https://linkinghub.elsevier.com/retrieve/pii/S0022169417304201.

P. Bates and A. De Roo. A simple raster-based model for flood inundation simulation. *Journal of Hydrology*, 236(1-2):54–77, Sept. 2000. ISSN 00221694. doi: 10.1016/S0022-1694(00)00278-X. URL https://linkinghub.elsevier.com/retrieve/pii/S002216940000278X.

P. D. Bates, M. S. Horritt, and T. J. Fewtrell. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of*

Hydrology, 387(1-2):33–45, June 2010. ISSN 00221694. doi: 10.1016/j.jhydrol.2010.03.027. URL https://linkinghub.elsevier.com/retrieve/pii/S0022169410001538.

J. Bellier. *Prévisions hydrologiques probabilistes dans un cadre multivarié: quels outils pour assurer fiabilité et cohérence spatio-temporelle?* PhD thesis, Université Grenoble Alpes, 2018. URL https://tel.archives-ouvertes.fr/tel-01950725.

J. Bellier, I. Zin, and G. Bontron. Sample Stratification in Verification of Ensemble Forecasts of Continuous Scalar Variables: Potential Benefits and Pitfalls. *Monthly Weather Review*, 145(9):3529–3544, Sept. 2017. ISSN 0027-0644, 1520-0493. doi: 10.1175/MWR-D-16-0487.1. URL http://journals.ametsoc.org/doi/10.1175/MWR-D-16-0487.1. Number: 9.

L. Berthet. *Prévision des crues au pas de temps horaire : pour une meilleure assimilation de l'information de débit dans un modèle hydrologique*. PhD Thesis, AgroParisTech, 2010.

L. Berthet, A. Valéry, R. Garçon, R. Marty, L. Moulin, D. Puygrenier, O. Piotte, M. Le Lay, B. Janet, and F. Duquesne. Cohérence des prévisions et place de l'expertise : les nouveaux défis pour la prévision des crues. *La Houille Blanche*, (1):5–12, Feb. 2019. ISSN 0018-6368, 1958-5551. doi: 10.1051/lhb/2019001. URL https://www.shf-lhb.org/10.1051/lhb/2019001. Number: 1.

Berthet, Lionel, Gaume, Éric, Piotte, Olivier, les groupes de travail du Conseil scientifique et technique du SCHAPI, and du chantier d\'intérêt commun SCHAPI - SPC Estimation des incertitudes de prévision. Evaluer et communiquer les incertitudes associées aux prévisions hydrologiques pour mieux partager l'information. *La Houille Blanche*, (4):18–24, 2016. doi: 10.1051/lhb/2016035. URL https://doi.org/10.1051/lhb/2016035. Number: 4.

K. Beven. A manifesto for the equifinality thesis. *Journal of Hydrology*, 320(1-2):18–36, Mar. 2006. ISSN 00221694. doi: 10.1016/j.jhydrol.2005.07.007. URL https://linkinghub.elsevier.com/retrieve/pii/S002216940500332X.

K. Beven and A. Binley. The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, 6(3):279–298, 1992. doi: 10.1002/hyp.3360060305. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.3360060305. Number: 3.

K. Beven and J. Freer. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, page 19, 2001.

S. Biancamaria, M. Durand, K. Andreadis, P. Bates, A. Boone, N. Mognard, E. Rodríguez, D. Alsdorf, D. Lettenmaier, and E. Clark. Assimilation of virtual wide swath altimetry to improve Arctic river modeling. *Remote Sensing of Environment*, 115(2):373–381, Feb. 2011. ISSN 00344257. doi: 10.1016/j.rse.2010.09.008. URL https://linkinghub.elsevier.com/retrieve/pii/S0034425710002816.

B. Bonan. *Assimilation de données pour l'initialisation et l'estimation de paramètres d'un modèle d'évolution de calotte polaire*. PhD thesis, Université de Grenoble, 2013. URL https://tel.archives-ouvertes.fr/tel-00930097v2.

E. Borgonovo. A new uncertainty importance measure. *Reliability Engineering & System Safety*, 92(6):771–784, June 2007. ISSN 09518320. doi: 10.1016/j.ress.2006.04.015. URL https://linkinghub.elsevier.com/retrieve/pii/S0951832006000883.

E. Borgonovo, W. Castaings, and S. Tarantola. Moment Independent Importance Measures: New Results and Analytical Test Cases: Moment Independent Importance Measures. *Risk Analysis*, 31(3):404–428, Mar. 2011. ISSN 02724332. doi: 10.1111/j.1539-6924.2010.01519.x. URL http://doi.wiley.com/10.1111/j.1539-6924.2010.01519.x.

L. Bouilloud, K. Chancibault, B. Vincendon, V. Ducrocq, F. Habets, G.-M. Saulnier, S. Anquetin, E. Martin, and J. Noilhan. Coupling the ISBA Land Surface Model and the TOPMODEL Hydrological Model for Mediterranean Flash-Flood Forecasting: Description, Calibration, and Validation. *Journal of Hydrometeorology*, 11(2):315–333, Apr. 2010. ISSN 1525-7541, 1525-755X. doi: 10.1175/2009JHM1163.1. URL http://journals.ametsoc.org/doi/10.1175/2009JHM1163.1.

F. Bourgin. *Comment quantifier l'incertitude prédictive en modélisation hydrologique ? : Travail exploratoire sur un grand échantillon de bassins versants*. PhD Thesis, AgroParisTech, 2014. URL http://www.theses.fr/2014AGPT0016/document.

F. Bourgin, M. Ramos, G. Thirel, and V. Andréassian. Investigating the interactions between data assimilation and post-processing in hydrological ensemble forecasting. *Journal of Hydrology*, 519:2775–2784, Nov. 2014. ISSN 00221694. doi: 10.1016/j.jhydrol.2014.07.054. URL http://linkinghub.elsevier.com/retrieve/pii/S0022169414005885.

F. Bouttier. La prévision d'ensemble, 2021. URL https://www.encyclopedie-environnement.org/air/prevision-densemble/.

F. Bouttier and R. Buizza. La prévision d'ensemble. *La Météorologie*, 100(Spécial anniversaire 25 ans):29–36, 2018. doi: 10.4267/2042/65140. URL https://lameteorologie.fr/issues/2018/100/meteo_2018_100_29. Number: Spécial anniversaire 25 ans.

L. Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, Aug. 1996. ISSN 0885-6125, 1573-0565. doi: 10.1007/BF00058655. URL http://link.springer.com/10.1007/BF00058655. Number: 2.

L. Breiman. Random Forests. *Machine Learning*, 45(4):5–32, 2001. doi: 10.1023/A:1010933404324. Number: 4.

J. B. Bremnes. Probabilistic wind power forecasts using local quantile regression. *Wind Energy*, 7(1):47–54, 2004. doi: 10.1002/we.107. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/we.107. Number: 1.

P. Brisset, J. Monnier, P.-A. Garambois, and H. Roux. On the assimilation of altimetric data in 1D Saint-Venant river flow models. page 27, 2018.

P. Brousseau, L. Berre, F. Bouttier, and G. Desroziers. Background-error covariances for a convective-scale data-assimilation system: AROME-France 3D-Var. *Quarterly Journal of the Royal Meteorological Society*, 137(655):409–422, Jan. 2011. ISSN 00359009. doi: 10.1002/qj.750. URL http://doi.wiley.com/10.1002/qj.750. Number: 655.

G. Brunner. HEC-RAS, River Analysis System Hydraulic Reference Manual. Version 5.0. US Army Corps of Engineers – Hydrologic Engineering Center (HEC), 2016.

C. A. Buahin, N. Sangwan, C. Fagan, D. R. Maidment, J. S. Horsburgh, E. J. Nelson, V. Merwade, and C. Rae. Probabilistic Flood Inundation Forecasting Using Rating Curve Libraries. *JAWRA Journal of the American Water Resources Association*, 53(2):300–315, Apr. 2017. ISSN 1093474X. doi: 10.1111/1752-1688.12500. URL http://doi.wiley.com/10.1111/1752-1688.12500.

R. Buizza, M. Milleer, and T. N. Palmer. Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society*, 125(560):2887–2908, 1999. doi: 10.1002/qj.49712556006. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49712556006.

G. Burgers. Analysis Scheme in the Ensemble Kalman Filter. *MONTHLY WEATHER REVIEW*, 126:6, 1998.

G. Candille and O. Talagrand. Evaluation of probabilistic prediction systems for a scalar variable. *Quarterly Journal of the Royal Meteorological Society*, 131(609):2131–2150, July 2005. ISSN 00359009, 1477870X. doi: 10.1256/qj.04.71. URL http://doi.wiley.com/10.1256/qj.04.71. Number: 609.

A. Carrassi, M. Bocquet, L. Bertino, and G. Evensen. Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdisciplinary Reviews: Climate Change*, 9(5):e535, Sept. 2018. ISSN 17577780. doi: 10.1002/wcc.535. URL http://doi.wiley.com/10.1002/wcc.535.

J.-L. Champeaux, P. Dupuy, O. Laurantin, I. Soulan, P. Tabary, and J.-M. Soubeyroux. Les mesures de précipitations et l'estimation des lames d'eau à Météo-France : état de l'art et perspectives. *La Houille Blanche*, (5):28–34, Oct. 2009. ISSN 0018-6368, 1958-5551. doi: 10.1051/lhb/2009052. URL http://www.shf-lhb.org/10.1051/lhb/2009052. Number: 5.

S. Cheng, J.-P. Argaud, B. Iooss, D. Lucor, and A. Ponçot. Error covariance tuning in variational data assimilation: application to an operating hydrological model. *Stochastic Environmental Research and Risk Assessment*, 35(5):1019–1038, May 2021. ISSN 1436-3240, 1436-3259. doi: 10.1007/s00477-020-01933-7. URL https://link.springer.com/10.1007/s00477-020-01933-7.

M. Clark, S. Gangopadhyay, L. Hay, B. Rajagopalan, and R. Wilby. The Schaake Shuffle: A Method for Reconstructing Space–Time Variability in Forecasted Precipitation and Temperature Fields. *Journal of Hydrometeorology*, 5(1):243–262, 2004. doi: 10.1175/1525-7541(2004)005<0243:TSSAMF>2.0.CO;2. URL https://doi.org/10.1175/1525-7541(2004)005<0243:TSSAMF>2.0.CO;2. Number: 1.

M. P. Clark, D. E. Rupp, R. A. Woods, X. Zheng, R. P. Ibbitt, A. G. Slater, J. Schmidt, and M. J. Uddstrom. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. *Advances in Water Resources*, 31(10):1309–1324, Oct. 2008a. ISSN 03091708. doi: 10.1016/j.advwatres.2008.06.005. URL https://linkinghub.elsevier.com/retrieve/pii/S0309170808001012.

M. P. Clark, A. G. Slater, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay. Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models: DIFFERENCES BETWEEN HYDROLOGICAL MODELS. *Water Resources Research*, 44(12), Dec. 2008b. ISSN 00431397. doi: 10.1029/2007WR006735. URL http://doi.wiley.com/10.1029/2007WR006735.

H. Cloke and F. Pappenberger. Ensemble flood forecasting: A review. *Journal of Hydrology*, 375(3-4):613–626, Sept. 2009. ISSN 00221694. doi: 10.1016/j.jhydrol.2009.06.005. URL https://linkinghub.elsevier.com/retrieve/pii/S0022169409003291. Number: 3-4.

M. Courbariaux. *Contributions statistiques aux prévisions hydrométéorologiques par méthodes d'ensemble*. PhD Thesis, 2017.

R. I. Cukier, H. B. Levine, and K. E. Shuler. Nonlinear sensitivity analysis of multiparameter model systems. page 2, 1977.

G. J. M. De Lannoy, R. H. Reichle, P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest. Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter: CORRECTING FOR FORECAST BIAS IN SOIL MOISTURE ASSIMILATION. *Water Resources Research*, 43(9), Sept. 2007. ISSN 00431397. doi: 10.1029/2006WR005449. URL http://doi.wiley.com/10.1029/2006WR005449.

G. J. M. De Lannoy, R. H. Reichle, K. R. Arsenault, P. R. Houser, S. Kumar, N. E. C. Verhoest, and V. R. N. Pauwels. Multiscale assimilation of Advanced Microwave Scanning Radiometer-EOS snow water equivalent and Moderate Resolution Imaging Spectroradiometer snow cover fraction observations in northern Colorado: SATELLITE-OBSERVED SNOW DATA ASSIMILATION. *Water Resources Research*, 48(1), Jan. 2012. ISSN 00431397. doi: 10.1029/2011WR010588. URL http://doi.wiley.com/10.1029/2011WR010588.

A. De Roo, M. Odijk, G. Schmuck, E. Koster, and A. Lucieer. Assessing the effects of land use changes on floods in the meuse and oder catchment. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere*, 26(7-8):593–599, Jan. 2001. ISSN 14641909. doi: 10.1016/S1464-1909(01)00054-5. URL https://linkinghub.elsevier.com/retrieve/pii/S1464190901000545.

A. P. J. De Roo, C. G. Wesseling, and W. P. A. Van Deursen. Physically based river basin modelling within a GIS: the LISFLOOD model. *Hydrological Processes*, 14(11-12):1981–1992, 2000. doi: https://doi.org/10.1002/1099-1085(20000815/30)14:11/12<1981::AID-HYP49>3.0.CO;2-F. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/1099-1085%2820000815/30%2914%3A11/12%3C1981%3A%3AAID-HYP49%3E3.0.CO%3B2-F.

C. de Saint-Aubin, L. Garandeau, O. Roulle, G. Noël, Y. Laborda, F. Mannessiez, M. Sorel, A. Marchandise, B. Vincendon, C. Sorbet, and V. Pourret. LE PROJET CHROME, UNE COLLABORATION VIGICRUES – MÉTÉO-FRANCE POUR L'EXPÉRIMENTATION D'UNE PRÉVISION D'ENSEMBLE HYDROLOGIQUE SUR LES BASSINS MÉDITERRANÉENS. page 10, 2018.

C. Dechant and H. Moradkhani. Radiance data assimilation for operational snow and streamflow forecasting. *Advances in Water Resources*, 34(3):351–364, Mar. 2011. ISSN 03091708. doi: 10.1016/j.advwatres.2010.12.009. URL https://linkinghub.elsevier.com/retrieve/pii/S0309170810002381.

C. M. DeChant and H. Moradkhani. Improving the characterization of initial condition for ensemble streamflow prediction using data assimilation. *Hydrology and Earth System Sciences*, 15(11):3399–3410, Nov. 2011. ISSN 1607-7938. doi: 10.5194/hess-15-3399-2011. URL https://hess.copernicus.org/articles/15/3399/2011/.

L. Delle Monache, J. P. Hacker, Y. Zhou, X. Deng, and R. B. Stull. Probabilistic aspects of meteorological and ozone regional ensemble forecasts. *Journal of Geophysical Research: Atmospheres*, 111(D24), 2006. doi: 10.1029/2005JD006917. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2005JD006917. Number: D24.

L. Delle Monache, F. A. Eckel, D. L. Rife, B. Nagarajan, and K. Searight. Probabilistic Weather Prediction with an Analog Ensemble. *Monthly Weather Review*, 141(10):3498–3516, Oct. 2013. ISSN 0027-0644, 1520-0493. doi: 10.1175/MWR-D-12-00281.1. URL http://journals.ametsoc.org/doi/abs/10.1175/MWR-D-12-00281.1. Number: 10.

J. Demargne, P. Javelle, D. Organde, L. Garandeau, and B. Janet. Intégration des prévisions immédiates de pluie à haute-résolution pour une meilleure anticipation des crues soudaines. *La Houille Blanche*, (3-4):13–21, Oct. 2019. ISSN 0018-6368, 1958-5551. doi: 10.1051/lhb/2019023. URL https://www.shf-lhb.org/10.1051/lhb/2019023.

D. Demeritt, S. Nobert, H. L. Cloke, and F. Pappenberger. The European Flood Alert System and the communication, perception, and use of ensemble predictions for operational flood risk management: EFAS FOR OPERATIONAL FLOOD MANAGEMENT. *Hydrological Processes*, 27(1):147–157, Jan. 2013. ISSN 08856087. doi: 10.1002/hyp.9419. URL http://doi.wiley.com/10.1002/hyp.9419.

DGPR-SRNH. Guide méthodologique - Conception et mise en oeuvre d'un système d'avertissement local aux crues. URL http://www.auvergne-rhone-alpes.developpement-durable.gouv.fr/IMG/pdf/guide_methodo_sdal_vf_cle6a1141_1_.pdf.

G. Di Baldassarre, G. Schumann, P. D. Bates, J. E. Freer, and K. J. Beven. Flood-plain mapping: a critical discussion of deterministic and probabilistic approaches. *Hydrological Sciences Journal*, 55(3):364–376, Apr. 2010. ISSN 0262-6667, 2150-3435. doi: 10.1080/02626661003683389. URL https://www.tandfonline.com/doi/full/10.1080/02626661003683389.

DREALCentre. Schéma Directeur de Prévision des Crues - Bassin Loire-Bretagne, Dec. 2012. URL http://www.auvergne-rhone-alpes.developpement-durable.gouv.fr/IMG/pdf/sdpc-lb-approb_2012_cle03e2ba.pdf.

Q. Duan, N. K. Ajami, X. Gao, and S. Sorooshian. Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Advances in Water Resources*, 30(5):1371–1386, May 2007. ISSN 03091708. doi: 10.1016/j.advwatres.2006.11.014. URL https://linkinghub.elsevier.com/retrieve/pii/S030917080600220X. Number: 5.

V. Ducrocq, B. Boudevillain, C. Bouvier, I. Braud, N. Fourrie, C. Lebeaupin-Brossier, P. Javelle, O. Nuissier, O. Payrastre, H. Roux, I. Ruin, and B. Vincendon. LE PRO-

GRAMME HYMEX - CONNAISSANCES ET PREVISION DES PLUIES INTENSES ET CRUES RAPIDES EN REGION MEDITERRANENNE. page 10, 2018.

M. Durand, K. M. Andreadis, D. E. Alsdorf, D. P. Lettenmaier, D. Moller, and M. Wilson. Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model. *Geophysical Research Letters*, 35(20):L20401, Oct. 2008. ISSN 0094-8276. doi: 10.1029/2008GL034150. URL http://doi.wiley.com/10.1029/2008GL034150.

B. Efron and C. Stein. The Jackknife Estimate of Variance. *The Annals of Statistics*, 9(3): 586–596, 1981. doi: 10.1214/aos/1176345462. URL https://doi.org/10.1214/aos/1176345462. Number: 3.

M. B. Ek, K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley. Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model. *Journal of Geophysical Research: Atmospheres*, 108(D22), 2003. doi: https://doi.org/10.1029/2002JD003296. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002JD003296.

C. M. Emery. *Contribution de la future mission altimétrique à large fauchée SWOT pour la modélisation hydrologique à grande échelle*. PhD thesis, Toulouse 3 Paul Sabatier, 2017.

C. M. Emery, S. Biancamaria, A. Boone, P.-A. Garambois, S. Ricci, M. C. Rochoux, and B. Decharme. Temporal Variance-Based Sensitivity Analysis of the River-Routing Component of the Large-Scale Hydrological Model ISBA–TRIP: Application on the Amazon Basin. *Journal of Hydrometeorology*, 17(12):3007–3027, Dec. 2016. ISSN 1525-755X, 1525-7541. doi: 10.1175/JHM-D-16-0050.1. URL http://journals.ametsoc.org/doi/10.1175/JHM-D-16-0050.1. Number: 12.

C. M. Emery, A. Paris, S. Biancamaria, A. Boone, S. Calmant, P.-A. Garambois, and J. Santos da Silva. Large-scale hydrological model river storage and discharge correction using a satellite altimetry-based discharge product. *Hydrology and Earth System Sciences*, 22(4):2135–2162, Apr. 2018. ISSN 1607-7938. doi: 10.5194/hess-22-2135-2018. URL https://hess.copernicus.org/articles/22/2135/2018/.

C. M. Emery, S. Biancamaria, A. Boone, S. Ricci, M. C. Rochoux, V. Pedinotti, and C. H. David. Assimilation of wide-swath altimetry observations to correct large-scale river routing model parameters. preprint, Rivers and Lakes/Remote Sensing and GIS, June 2019. URL https://hess.copernicus.org/preprints/hess-2019-242/hess-2019-242.pdf.

G. Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99: 10143–10162, 1994.

G. Evensen. The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4):343–367, Nov. 2003. ISSN 1616-7341, 1616-7228. doi: 10.1007/s10236-003-0036-9. URL http://link.springer.com/10.1007/s10236-003-0036-9.

R. Faivre, B. Iooss, S. Mahévas, D. Makowski, and H. Monod. *Analyse de sensibilité et exploration de modèles*. Collection Savoir-Faire. Editions Quae, 2013. URL https://hal.archives-ouvertes.fr/hal-01173750.

I. Farouk, E. Cosme, S. Metref, J. Gailhard, and M. Le-Lay. Hydrological data assimilation using the particle filter in a semi-distributed model MORDOR-SD. other, oral, Mar. 2020. URL https://meetingorganizer.copernicus.org/EGU2020/EGU2020-18010.html.

F. Ferretti, A. Saltelli, and S. Tarantola. Trends in sensitivity analysis practice in the last decade. *Science of The Total Environment*, 568:666 – 670, 2016. ISSN 0048-9697. doi: https://doi.org/10.1016/j.scitotenv.2016.02.133. URL http://www.sciencedirect.com/science/article/pii/S0048969716303448.

J. Freer, K. Beven, and B. Ambroise. Bayesian Estimation of Uncertainty in Runoff Prediction and the Value of Data: An Application of the GLUE Approach. *Water Resources Research*, 32(7):2161–2173, July 1996. ISSN 00431397. doi: 10.1029/95WR03723. URL http://doi.wiley.com/10.1029/95WR03723.

P. A. Garambois, H. Roux, K. Larnier, W. Castaings, and D. Dartus. Characterization of process-oriented hydrologic model behavior with temporal sensitivity analysis for flash floods in Mediterranean catchments. *Hydrology and Earth System Sciences*, 17 (6):2305–2322, June 2013. ISSN 1607-7938. doi: 10.5194/hess-17-2305-2013. URL https://www.hydrol-earth-syst-sci.net/17/2305/2013/. Number: 6.

L. Garandeau, A. Belleudy, P. Javelle, D. Organde, J. Demargne, C. D. Saint-Aubin, and C. Fouchier. VIGICRUES FLASH, UN SERVICE AUTOMATIQUE D'AVERTISSEMENT POUR LES CRUES RAPIDES. page 11, 2018.

F. Garavaglia, M. Le Lay, F. Gottardi, R. Garçon, J. Gailhard, E. Paquet, and T. Mathevet. Impact of model structure on flow simulation and hydrological realism: from a lumped to a semi-distributed approach. *Hydrology and Earth System Sciences*, 21 (8):3937–3952, Aug. 2017. ISSN 1607-7938. doi: 10.5194/hess-21-3937-2017. URL https://www.hydrol-earth-syst-sci.net/21/3937/2017/. Number: 8.

R. Garçon. Prévision opérationnelle des apports de la Durance à Serre-Ponçon à l'aide du modèle MORDOR. Bilan de l'année 1994-1995. *La Houille Blanche*, (5):71–76, Aug. 1996. ISSN 0018-6368, 1958-5551. doi: 10.1051/lhb/1996056. URL http://www.shf-lhb.org/10.1051/lhb/1996056. Number: 5.

R. Garçon, B. houdant, F. Garavaglia, T. Mathevet, E. Paquet, and J. Gailhard. Expertise humaine des prévisions hydrométéorologiques et communication de leurs incertitudes dans un contexte décisionnel. *La Houille Blanche*, (5):71–80, 2009. doi: 10.1051/lhb/2009058. URL https://doi.org/10.1051/lhb/2009058. Number: 5.

N. Georgas, A. Blumberg, T. Herrington, T. Wakeman, F. Saleh, D. Runnels, A. Jordi, K. Ying, L. Yin, V. Ramaswamy, A. Yakubovskiy, O. Lopez, J. McNally, J. Schulte, and Y. Wang. The stevens flood advisory system: operational H3E flood forecasts for the greater New York / New Jersey Metropolitan Region. *International Journal of Safety and Security Engineering*, 6(3):648–662, Aug. 2016. ISSN 2041-9031, 2041-904X. doi: 10.2495/SAFE-V6-N3-648-662. URL http://www.witpress.com/doi/journals/SAFE-V6-N3-648-662.

T. Gneiting and M. Katzfuss. Probabilistic Forecasting. *Annual Review of Statistics and Its Application*, 1(1):125–151, Jan. 2014. ISSN 2326-8298, 2326-831X. doi: 10.1146/annurev-statistics-062713-085831. URL http://www.annualreviews.org/doi/10.1146/annurev-statistics-062713-085831. Number: 1.

T. Gneiting and A. E. Raftery. Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, 102(477):359–378, Mar. 2007. ISSN 0162-1459, 1537-274X. doi: 10.1198/016214506000001437. URL http://www.tandfonline.com/doi/abs/10.1198/016214506000001437. Number: 477.

T. Gneiting, A. E. Raftery, A. H. Westveld, and T. Goldman. Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation. *Monthly Weather Review*, 133(5):1098–1118, 2005. doi: 10.1175/MWR2904.1. URL https://doi.org/10.1175/MWR2904.1. Number: 5.

T. Gneiting, F. Balabdaoui, and A. E. Raftery. Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2):243–268, Apr. 2007. ISSN 1369-7412, 1467-9868. doi: 10.1111/j.1467-9868.2007.00587.x. URL http://doi.wiley.com/10.1111/j.1467-9868.2007.00587.x. Number: 2.

M. Gomez, S. Sharma, S. Reed, and A. Mejia. Skill of ensemble flood inundation forecasts at short- to medium-range timescales. *Journal of Hydrology*, 568:207–220, Jan. 2019. ISSN 00221694. doi: 10.1016/j.jhydrol.2018.10.063. URL https://linkinghub.elsevier.com/retrieve/pii/S0022169418308345.

N. Goutal. MASCARET : a 1-D open-source software for flow hydrodynamic and water quality in open channel networks. page 8, 2012.

N. Goutal and F. Zaoui. MASCARET theory guide - Version v8p2. Technical report, 2020.

N. Goutal, C. Goeury, R. Ata, S. Ricci, N. El Mocayd, M. Rochoux, H. Oubanas, I. Gejadze, and P. Malaterre. Uncertainty quantification for river flow simulation applied to a real test case: the Garonne valley. In *SimHydro 2017*, Advances in Hydroinformatics: Simhydro 2017. Choosing the Right Model in Applied Hydraulics, pages 169–187, Nice, France, June 2017. doi: 10.1007/978-981-10-7218-5_12. URL https://hal.inrae.fr/hal-02609249.

P. Grönquist, T. Ben-Nun, N. Dryden, P. Dueben, L. Lavarini, S. Li, and T. Hoefler. Predicting Weather Uncertainty with Deep Convnets. *arXiv:1911.00630 [physics, stat]*, Dec. 2019. URL http://arxiv.org/abs/1911.00630. arXiv: 1911.00630.

H. V. Gupta, H. Kling, K. K. Yilmaz, and G. F. Martinez. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1):80 – 91, 2009. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2009.08.003. URL http://www.sciencedirect.com/science/article/pii/S0022169409004843. Number: 1.

J. Habert, S. Ricci, E. Le Pape, O. Thual, A. Piacentini, N. Goutal, G. Jonville, and M. Rochoux. Reduction of the uncertainties in the water level-discharge relation of a 1D hydraulic model in the context of operational flood forecasting. *Journal of Hydrology*, 532:52–64, Jan. 2016. ISSN 00221694. doi: 10.1016/j.jhydrol.2015.11.023. URL https://linkinghub.elsevier.com/retrieve/pii/S002216941500904X.

J. W. Hall, S. Tarantola, P. D. Bates, and M. S. Horritt. Distributed Sensitivity Analysis of Flood Inundation Model Calibration. *Journal of Hydraulic Engineering*, 131(2):117–126, Feb. 2005. ISSN 0733-9429, 1943-7900. doi: 10.1061/(ASCE)0733-9429(2005)131:2(117). URL http://ascelibrary.org/doi/10.1061/%28ASCE%290733-9429%282005%29131%3A2%28117%29.

J. W. Hall, S. A. Boyce, Y. Wang, R. J. Dawson, S. Tarantola, and A. Saltelli. Sensitivity Analysis for Hydraulic Models. *Journal of Hydraulic Engineering*, 135(11):959–969, Nov. 2009. ISSN 0733-9429, 1943-7900. doi: 10.1061/(ASCE)HY.1943-7900.0000098. URL http://ascelibrary.org/doi/10.1061/%28ASCE%29HY.1943-7900.0000098.

T. M. Hamill. Interpretation of Rank Histograms for Verifying Ensemble Forecasts. *Monthly Weather Review*, 129(3):550–560, Mar. 2001. ISSN 0027-0644, 1520-0493. doi: 10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2.

URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%282001%29129%3C0550%3AIORHFV%3E2.0.CO%3B2. Number: 3.

T. M. Hamill and S. J. Colucci. Verification of Eta–RSM Short-Range Ensemble Forecasts. *Monthly Weather Review*, 125(6):1312–1327, June 1997. ISSN 0027-0644, 1520-0493. doi: 10.1175/1520-0493(1997)125<1312:VOERSR>2.0.CO;2. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0493%281997%29125%3C1312%3AVOERSR%3E2.0.CO%3B2. Number: 6.

T. M. Hamill and J. S. Whitaker. Probabilistic Quantitative Precipitation Forecasts Based on Reforecast Analogs: Theory and Application. *Monthly Weather Review*, 134(11): 3209–3229, Nov. 2006. ISSN 0027-0644, 1520-0493. doi: 10.1175/MWR3237.1. URL http://journals.ametsoc.org/doi/abs/10.1175/MWR3237.1. Number: 11.

T. M. Hamill, J. S. Whitaker, and C. Snyder. Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter. *MONTHLY WEATHER REVIEW*, 129:15, 2001.

T. M. Hamill, G. T. Bates, J. S. Whitaker, D. R. Murray, M. Fiorino, T. J. G. Jr, Y. Zhu, and W. Lapenta. NOAA'S SECOND-GENERATION GLOBAL MEDIUM-RANGE ENSEMBLE REFORECAST DATASET. page 14, 2013.

T. Hastie, R. Tibshirani, and J. Friedman. *The elements of statistical learning*. Springer. 2009. URL https://web.stanford.edu/~hastie/Papers/ESLII.pdf.

S. Hayami. On The Propagation Of Flood Waves. Technical report, Disaster prevention research institute, Dec. 1951. URL https://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/123641/1/b01p0n000p01.pdf.

S. Hemri, F. Fundel, and M. Zappa. Simultaneous calibration of ensemble river flow predictions over an entire range of lead times: Calibration of Ensemble River Flow Predictions. *Water Resources Research*, 49(10):6744–6755, Oct. 2013. ISSN 00431397. doi: 10.1002/wrcr.20542. URL http://doi.wiley.com/10.1002/wrcr.20542.

S. Hemri, M. Scheuerer, F. Pappenberger, K. Bogner, and T. Haiden. Trends in the predictive performance of raw ensemble weather forecasts. *Geophysical Research Letters*, 41(24):9197–9205, 2014. doi: https://doi.org/10.1002/2014GL062472. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014GL062472.

S. Hemri, D. Lisniak, and B. Klein. Multivariate postprocessing techniques for probabilistic hydrological forecasting. *Water Resources Research*, 51(9):7436–7451, 2015. doi: 10.1002/2014WR016473. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014WR016473. Number: 9.

H. Hersbach. Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems. *Weather and Forecasting*, 15(5):559–570, Oct. 2000. ISSN 0882-8156, 1520-0434. doi: 10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0434%282000%29015%3C0559%3ADOTCRP%3E2.0.CO%3B2. Number: 5.

V. Häfliger, E. Martin, A. Boone, S. Ricci, and S. Biancamaria. Assimilation of Synthetic SWOT River Depths in a Regional Hydrometeorological Model. *Water*, 11(1):78, Jan. 2019. ISSN 2073-4441. doi: 10.3390/w11010078. URL https://www.mdpi.com/2073-4441/11/1/78.

W. Hoeffding. A Class of Statistics with Asymptotically Normal Distribution. *The Annals of Mathematical Statistics*, 19(3):293–325, 1948. doi: 10.1214/aoms/1177730196. URL https://doi.org/10.1214/aoms/1177730196. Number: 3.

T. M. Hopson and P. J. Webster. A 1–10-Day Ensemble Forecasting Scheme for the Major River Basins of Bangladesh: Forecasting Severe Floods of 2003–07*. *Journal of Hydrometeorology*, 11(3):618–641, June 2010. ISSN 1525-755X, 1525-7541. doi: 10.1175/2009JHM1006.1. URL http://journals.ametsoc.org/doi/abs/10.1175/2009JHM1006.1. Number: 3.

R. Hostache, X. Lai, J. Monnier, and C. Puech. Assimilation of spatially distributed water levels into a shallow-water flood model. Part II: Use of a remote sensing image of Mosel River. *Journal of Hydrology*, 390(3-4):257–268, Sept. 2010. ISSN 00221694. doi: 10.1016/j.jhydrol.2010.07.003. URL https://linkinghub.elsevier.com/retrieve/pii/S0022169410004166.

P. L. Houtekamer, L. Lefaivre, J. Derome, H. Ritchie, and H. L. Mitchell. A System Simulation Approach to Ensemble Prediction. *Monthly Weather Review*, 124(6):1225–1242, June 1996. ISSN 0027-0644. doi: 10.1175/1520-0493(1996)124<1225:ASSATE>2.0.CO;2. URL https://doi.org/10.1175/1520-0493(1996)124<1225:ASSATE>2.0.CO;2. Number: 6.

B. Iooss and P. Lemaître. A Review on Global Sensitivity Analysis Methods. In G. Dellino and C. Meloni, editors, *Uncertainty Management in Simulation-Optimization of Complex Systems*, volume 59, pages 101–122. Springer US, Boston, MA, 2015. ISBN 978-1-4899-7546-1 978-1-4899-7547-8. doi: 10.1007/978-1-4899-7547-8_5. URL http://link.springer.com/10.1007/978-1-4899-7547-8_5.

IRSTEA. *Calage et application opérationnelle du modèle de prévision de crue GRP - Manuel d'utilisation*. IRSTEA, Mar. 2017.

B. Janet, A. Escudier, J. Laffargue, O. Piotte, and L. Garandeau. LES DERNIÈRES ÉVOLUTIONS DU RÉSEAU DE LA PRÉVISION DES CRUES. page 11, 2018.

P. Javelle. Evaluating flash-flood warnings at ungauged locations using post-event surveys: a case study with th. page 14, 2014.

P. Javelle, D. Organde, J. Demargne, C. Saint-Martin, C. de Saint-Aubin, L. Garandeau, and B. Janet. Setting up a French national flash flood warning system for ungauged catchments based on the AIGA method. *E3S Web of Conferences*, 7:18010, 2016. ISSN 2267-1242. doi: 10.1051/e3sconf/20160718010. URL http://www.e3s-conferences.org/10.1051/e3sconf/20160718010.

I. T. Jolliffe and D. B. Stephenson. *Forecast verification: a practitioner's guide in atmospheric science*. Wiley, Chichester, 2003. ISBN 978-0-471-49759-2.

R. E. Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.

R. E. Kalman and R. S. Bucy. New Results in Linear Filtering and Prediction Theory. *Journal of Basic Engineering*, 83(1):95–108, Mar. 1961. ISSN 0021-9223. doi: 10.1115/1.3658902. URL https://asmedigitalcollection.asme.org/fluidsengineering/article/83/1/95/426820/New-Results-in-Linear-Filtering-and-Prediction.

F. Khorashadi Zadeh, J. Nossent, F. Sarrazin, F. Pianosi, A. van Griensven, T. Wagener, and W. Bauwens. Comparison of variance-based and moment-independent global sensitivity analysis approaches by application to the SWAT model. *Environmental Modelling & Software*, 91:210–222, May 2017. ISSN 13648152. doi: 10.1016/j.envsoft.2017.02.001. URL https://linkinghub.elsevier.com/retrieve/pii/S1364815217301159.

J. P. C. Kleijnen. *Design and analysis of simulation experiments*. Number 111 in International series in operations research & management science. Springer, New York, 2008. ISBN 978-0-387-71812-5. OCLC: ocn141380110.

V. Koren, S. Reed, M. Smith, Z. Zhang, and D.-J. Seo. Hydrology laboratory research modeling system (HL-RMS) of the US national weather service. *Journal of Hydrology*, page 22, 2004.

R. Krzysztofowicz. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resources Research*, 35(9):2739–2750, Sept. 1999. ISSN 00431397. doi: 10.1029/1999WR900099. URL http://doi.wiley.com/10.1029/1999WR900099. Number: 9.

R. Krzysztofowicz. The case for probabilistic forecasting in hydrology. *Journal of Hydrology*, 249(1-4):2–9, Aug. 2001. ISSN 00221694. doi: 10.1016/S0022-1694(01)00420-6. URL https://linkinghub.elsevier.com/retrieve/pii/S0022169401004206.

S. V. Kumar, C. D. Peters-Lidard, D. Mocko, R. Reichle, Y. Liu, K. R. Arsenault, Y. Xia, M. Ek, G. Riggs, B. Livneh, and M. Cosh. Assimilation of Remotely Sensed Soil Moisture and Snow Depth Retrievals for Drought Estimation. *Journal of Hydrometeorology*, 15 (6):2446–2469, Dec. 2014. ISSN 1525-755X, 1525-7541. doi: 10.1175/JHM-D-13-0132.1. URL http://journals.ametsoc.org/doi/10.1175/JHM-D-13-0132.1.

H. Lü, Z. Yu, Y. Zhu, S. Drake, Z. Hao, and E. A. Sudicky. Dual state-parameter estimation of root zone soil moisture by optimal parameter estimation and extended Kalman filter data assimilation. *Advances in Water Resources*, 34(3):395–406, Mar. 2011. ISSN 03091708. doi: 10.1016/j.advwatres.2010.12.005. URL https://linkinghub.elsevier.com/retrieve/pii/S0309170810002344.

V. Laborie. *Quantification d'incertitudes et assimilation de données pour la modélisation hydrodynamique bidimensionnelle : application au modèle de prévision des hautes eaux de l'estuaire de la Gironde*. PhD Thesis, 2020. URL http://www.theses.fr/2020PESC1006/document.

V. Laborie, S. Ricci, M. De Lozzo, N. Goutal, Y. Audouin, and P. Sergent. Quantifying forcing uncertainties in the hydrodynamics of the Gironde estuary. *Computational Geosciences*, 24(1):181–202, Feb. 2020. ISSN 1420-0597, 1573-1499. doi: 10.1007/s10596-019-09907-7. URL http://link.springer.com/10.1007/s10596-019-09907-7.

M. Le Lay and G. M. Saulnier. Exploring the signature of climate and landscape spatial variabilities in flash flood events: Case of the 8-9 September 2002 Cévennes-Vivarais catastrophic event: HYDROLOGICAL INTERACTIONS ON FLASH FLOOD. *Geophysical Research Letters*, 34(13):n/a–n/a, July 2007. ISSN 00948276. doi: 10.1029/2007GL029746. URL http://doi.wiley.com/10.1029/2007GL029746.

N. Le Moine, F. Hendrickx, J. Gailhard, R. Garçon, and F. Gottardi. Hydrologically Aided Interpolation of Daily Precipitation and Temperature Fields in a Mesoscale Alpine Catchment. *Journal of Hydrometeorology*, 16(6):2595–2618, Dec. 2015. ISSN 1525-755X, 1525-7541. doi: 10.1175/JHM-D-14-0162.1. URL http://journals.ametsoc.org/doi/10.1175/JHM-D-14-0162.1. Number: 6.

M. Leisenring and H. Moradkhani. Snow water equivalent prediction using Bayesian data assimilation methods. *Stochastic Environmental Research and Risk Assessment*, 25(2): 253–270, Feb. 2011. ISSN 1436-3240, 1436-3259. doi: 10.1007/s00477-010-0445-5. URL http://link.springer.com/10.1007/s00477-010-0445-5.

S. Lerch, T. L. Thorarinsdottir, F. Ravazzolo, and T. Gneiting. Forecaster's Dilemma: Extreme Events and Forecast Evaluation. *Statistical Science*, 32(1):106–127, Feb. 2017.

ISSN 0883-4237. doi: 10.1214/16-STS588. URL http://projecteuclid.org/euclid.ss/1491465630. Number: 1.

W. Li, Q. Duan, C. Miao, A. Ye, W. Gong, and Z. Di. A review on statistical postprocessing methods for hydrometeorological ensemble forecasting. *WIREs Water*, 4(6):e1246, 2017. doi: 10.1002/wat2.1246. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/wat2.1246.

X. Liang, D. P. Lettenmaier, E. F. Wood, and S. J. Burges. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres*, 99(D7):14415–14428, 1994. doi: https://doi.org/10.1029/94JD00483. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/94JD00483.

Y. Liu, A. H. Weerts, M. Clark, H.-J. Hendricks Franssen, S. Kumar, H. Moradkhani, D.-J. Seo, D. Schwanenberg, P. Smith, A. I. J. M. van Dijk, N. van Velzen, M. He, H. Lee, S. J. Noh, O. Rakovec, and P. Restrepo. Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. *Hydrology and Earth System Sciences*, 16(10):3863–3887, Oct. 2012. ISSN 1607-7938. doi: 10.5194/hess-16-3863-2012. URL https://hess.copernicus.org/articles/16/3863/2012/.

E. N. Lorenz. Deterministic Nonperiodic Flow. *Journal of the Atmospheric Sciences*, 20(2):130–141, Mar. 1963. ISSN 0022-4928. doi: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2. URL https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2. Number: 2.

E. N. Lorenz. Atmospheric predictability as revealed naturaally occuring analogues. *Journal of the Atmospheric Sciences*, 26:636–646, 1969. doi: https://doi.org/10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2. URL https://journals.ametsoc.org/view/journals/atsc/26/4/1520-0469_1969_26_636_aparbn_2_0_co_2.xml.

S. Madadgar and H. Moradkhani. Improved Bayesian multimodeling: Integration of copulas and Bayesian model averaging. *Water Resources Research*, 50(12):9586–9603, Dec. 2014. ISSN 00431397. doi: 10.1002/2014WR015965. URL http://doi.wiley.com/10.1002/2014WR015965.

H. Madsen and C. Skotner. Adaptive state updating in real-time river flow forecasting—a combined filtering and error forecasting procedure. *Journal of Hydrology*, 308(1-4):302–312, July 2005. ISSN 00221694. doi: 10.1016/j.jhydrol.2004.10.030. URL https://linkinghub.elsevier.com/retrieve/pii/S0022169404005372.

J. E. Matheson and R. L. Winkler. Scoring Rules for Continuous Probability Distributions. *Management Science*, 22(10):1087–1096, June 1976. ISSN 0025-1909, 1526-5501. doi:

10.1287/mnsc.22.10.1087. URL http://pubsonline.informs.org/doi/abs/10.1287/mnsc.22.10.1087. Number: 10.

H. McMillan, T. Krueger, and J. Freer. Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality: BENCHMARKING OBSERVATIONAL UNCERTAINTIES FOR HYDROLOGY. *Hydrological Processes*, 26 (26):4078–4111, Dec. 2012. ISSN 08856087. doi: 10.1002/hyp.9384. URL http://doi.wiley.com/10.1002/hyp.9384.

N. Meinshausen. Quantile Regression Forests. *Journal of Machine Learning Research*, 7: 17, 2006.

T. Michon and W. Castaings. Stratégie de calage du modèle hydrologique semi-distribué MORDOR-SD. Technical report, TENEVIA, 2017.

I. Mirouze. *Régularisation de problèmes inverses à l'aide de l'équation de diffusion, avec application à l'assimilation variationnelle de données océaniques*. PhD thesis, Université de Toulouse, Toulouse, Sept. 2010.

I. Mirouze and S. Ricci. Smurf: System for Modelling with Uncertainty Reduction, and Forecasting. *Journal of Open Research Software*, 9:2, Feb. 2021. ISSN 2049-9647. doi: 10.5334/jors.312. URL http://openresearchsoftware.metajnl.com/articles/10.5334/jors.312/.

H. L. Mitchell and P. L. Houtekamer. An Adaptive Ensemble Kalman Filter. *MONTHLY WEATHER REVIEW*, 128:18, 2000.

N. E. Moçayd, S. Ricci, N. Goutal, M. C. Rochoux, S. Boyaval, C. Goeury, D. Lucor, and O. Thual. Polynomial Surrogates for Open-Channel Flows in Random Steady State. page 24, 2017.

C. Monteil and N. Le Moine. Cequeau version EDF R&D 5.1.1 : Note de Principe. Technical report, 2014.

C. Monteil, F. Zaoui, N. Le Moine, and F. Hendrickx. Technical note: the caRamel R package for Automatic Calibration by Evolutionary Multi Objective Algorithm. *Hydrology and Earth System Sciences Discussions*, 2019:1–16, 2019. doi: 10.5194/hess-2019-259. URL https://www.hydrol-earth-syst-sci-discuss.net/hess-2019-259/.

H. Moradkhani, S. Sorooshian, H. V. Gupta, and P. R. Houser. Dual state–parameter estimation of hydrological models using ensemble Kalman filter. *Advances in Water Resources*, 28(2):135–147, Feb. 2005. ISSN 03091708. doi: 10.1016/j.advwatres.2004.09.002. URL https://linkinghub.elsevier.com/retrieve/pii/S0309170804001605.

H. Moradkhani, G. S. Nearing, P. Abbaszadeh, and S. Pathiraja. Fundamentals of Data Assimilation and Theoretical Advances. In Q. Duan, F. Pappenberger, A. Wood, H. L. Cloke, and J. C. Schaake, editors, *Handbook of Hydrometeorological Ensemble Forecasting*, pages 675–699. Springer Berlin Heidelberg, Berlin, Heidelberg, 2019. ISBN 978-3-642-39925-1. doi: 10.1007/978-3-642-39925-1_30. URL https://doi.org/10.1007/978-3-642-39925-1_30.

M. D. Morris. Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics*, 33(2):161–174, 1991. ISSN 00401706. URL http://www.jstor.org/stable/1269043. Publisher: [Taylor & Francis, Ltd., American Statistical Association, American Society for Quality].

L. Moulin, A. Abonnel, D. Puygrenier, A. Valéry, and R. Garçon. Prévision hydrométéorologique opérationnelle à EDF-DTG - Progrès récents et état des lieux en 2018. *LHB*, (2):44–54, 2019. doi: 10.1051/lhb/2019014. URL https://doi.org/10.1051/lhb/2019014. Number: 2.

A. Mueller, C. Baugh, P. Bates, and F. Pappenberger. Probabilistic Inundation Forecasting. In Q. Duan, F. Pappenberger, J. Thielen, A. Wood, H. L. Cloke, and J. C. Schaake, editors, *Handbook of Hydrometeorological Ensemble Forecasting*, pages 1–14. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016. ISBN 978-3-642-40457-3. doi: 10.1007/978-3-642-40457-3_59-1. URL https://doi.org/10.1007/978-3-642-40457-3_59-1.

A. H. Murphy. A New Vector Partition of the Probability Score. *Journal of Applied Meteorology and Climatology*, 12(4):595 – 600, 1973. doi: 10.1175/1520-0450(1973)012<0595:ANVPOT>2.0.CO;2. URL https://journals.ametsoc.org/view/journals/apme/12/4/1520-0450_1973_012_0595_anvpot_2_0_co_2.xml. Place: Boston MA, USA Publisher: American Meteorological Society.

A. H. Murphy. What Is a Good Forecast? An Essay on the Nature of Goodness in Weather Forecasting. *Weather and Forecasting*, 8(2): 281 – 293, 1993. doi: 10.1175/1520-0434(1993)008<0281:WIAGFA>2.0.CO;2. URL https://journals.ametsoc.org/view/journals/wefo/8/2/1520-0434_1993_008_0281_wiagfa_2_0_co_2.xml. Place: Boston MA, USA Publisher: American Meteorological Society.

E. M. Nabil. *La décomposition en polynôme du chaos pour l'amélioration de l'assimilation de données ensembliste en hydraulique fluviale*. PhD Thesis, 2017.

J. Nash and J. Sutcliffe. River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, 10(3):282–290, Apr. 1970. ISSN

00221694. doi: 10.1016/0022-1694(70)90255-6. URL https://linkinghub.elsevier.com/retrieve/pii/0022169470902556. Number: 3.

J. C. Neal, P. M. Atkinson, and C. W. Hutton. Flood inundation model updating using an ensemble Kalman filter and spatially distributed measurements. *Journal of Hydrology*, 336(3-4):401–415, Apr. 2007. ISSN 00221694. doi: 10.1016/j.jhydrol.2007.01.012. URL https://linkinghub.elsevier.com/retrieve/pii/S0022169407000388.

T. Ngo-Duc, T. Oki, and S. Kanae. A variable streamflow velocity method for global river routing model: model description and preliminary results. *Hydrology and Earth System Sciences Discussions*, 4(6):4389–4414, Dec. 2007. ISSN 1812-2116. doi: 10.5194/hessd-4-4389-2007. URL https://hess.copernicus.org/preprints/4/4389/2007/.

J. Nicolau. Short-range ensemble forecasting. page 6, 2002.

S. Nie, J. Zhu, and Y. Luo. Simultaneous estimation of land surface scheme states and parameters using the ensemble Kalman filter: identical twin experiments. *Hydrology and Earth System Sciences*, 15(8):2437–2457, Aug. 2011. ISSN 1607-7938. doi: 10.5194/hess-15-2437-2011. URL https://hess.copernicus.org/articles/15/2437/2011/.

S. J. Noh, A. H. Weerts, O. Rakovec, H. Lee, and D.-J. Seo. Assimilation of Streamflow Observations. In Q. Duan, F. Pappenberger, J. Thielen, A. Wood, H. L. Cloke, and J. C. Schaake, editors, *Handbook of Hydrometeorological Ensemble Forecasting*, pages 1–36. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018. ISBN 978-3-642-40457-3. doi: 10.1007/978-3-642-40457-3_33-2. URL https://doi.org/10.1007/978-3-642-40457-3_33-2.

J. Noilhan and J.-F. Mahfouf. The ISBA land surface parameterisation scheme. *Global and Planetary Change*, 13(1-4):145–159, June 1996. ISSN 09218181. doi: 10.1016/0921-8181(95)00043-7. URL https://linkinghub.elsevier.com/retrieve/pii/0921818195000437.

H. Oubanas. River discharge estimation from synthetic SWOT-type observations using variational data assimilation and the full Saint-Venant hydraulic model. *Journal of Hydrology*, page 10, 2018.

L. Oudin. *Recherche d'un modèle d'évapotranspiration potentielle pertinent comme entrée d'un modèle pluie-débit global*. PhD Thesis, 2004.

L. Oudin, F. Hervieu, C. Michel, C. Perrin, V. Andréassian, F. Anctil, and C. Loumagne. Which potential evapotranspiration input for a lumped rainfall–runoff model? *Journal of Hydrology*, 303(1-4):290–306, Mar. 2005. ISSN 00221694. doi: 10.1016/j.jhydrol.2004.08.026. URL https://linkinghub.elsevier.com/retrieve/pii/S0022169404004056. Number: 1-4.

T. Palmer, J. Barkmeijer, R. Buizza, E. Klinker, and D. Richardson. L'avenir de la prévision d'ensemble. *La Météorologie*, 8(36):22, 2002. ISSN 0026-1181. doi: 10.4267/2042/36203. URL https://lameteorologie.fr/issues/2002/36/meteo_2002_36_22. Number: 36.

F. Pappenberger, K. J. Beven, N. M. Hunter, P. D. Bates, B. T. Gouweleeuw, and J. Thielen. Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). page 13, 2005.

F. Pappenberger, K. J. Beven, M. Ratto, and P. Matgen. Multi-method global sensitivity analysis of flood inundation models. *Advances in Water Resources*, 31(1):1–14, Jan. 2008. ISSN 03091708. doi: 10.1016/j.advwatres.2007.04.009. URL https://linkinghub.elsevier.com/retrieve/pii/S0309170807000747.

F. Pappenberger, H. L. Cloke, G. Balsamo, T. Ngo-Duc, and T. Oki. Global runoff routing with the hydrological component of the ECMWF NWP system. *International Journal of Climatology*, 30(14):2155–2174, 2010. doi: 10.1002/joc.2028. URL https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.2028. Number: 14.

F. Pappenberger, J. Thielen, and M. Del Medico. The impact of weather forecast improvements on large scale hydrology: analysing a decade of forecasts of the European Flood Alert System. *Hydrological Processes*, 25(7):1091–1113, Mar. 2011. ISSN 08856087. doi: 10.1002/hyp.7772. URL http://doi.wiley.com/10.1002/hyp.7772.

F. Pappenberger, T. C. Pagano, J. D. Brown, L. Alfieri, D. A. Lavers, L. Berthet, F. Bressand, H. L. Cloke, M. Cranston, J. Danhelka, J. Demargne, N. Demuth, C. de Saint-Aubin, P. M. Feikema, M. A. Fresch, R. Garçon, A. Gelfan, Y. He, Y. Z. Hu, B. Janet, N. Jurdy, P. Javelle, L. Kuchment, Y. Laborda, E. Langsholt, M. Le Lay, Z. J. Li, F. Mannessiez, A. Marchandise, R. Marty, D. Meißner, D. Manful, D. Organde, V. Pourret, S. Rademacher, M. H. Ramos, D. Reinbold, S. Tibaldi, P. Silvano, P. Salamon, D. Shin, C. Sorbet, E. Sprokkereef, V. Thiemig, N. K. Tuteja, S. J. van Andel, J. S. Verkade, B. Vehviläinen, A. Vogelbacher, F. Wetterhall, M. Zappa, R. E. Van der Zwan, and J. Thielen-del Pozo. Hydrological Ensemble Prediction Systems Around the Globe. In Q. Duan, F. Pappenberger, J. Thielen, A. Wood, H. L. Cloke, and J. C. Schaake, editors, *Handbook of Hydrometeorological Ensemble Forecasting*, pages 1–35. Springer Berlin Heidelberg, Berlin, Heidelberg, 2016. ISBN 978-3-642-40457-3. doi: 10.1007/978-3-642-40457-3_47-1. URL https://doi.org/10.1007/978-3-642-40457-3_47-1.

Y.-Y. Park, R. Buizza, and M. Leutbecher. TIGGE: Preliminary results on comparing and combining ensembles. *Quarterly Journal of the Royal Meteorological Society*, 134

(637):2029–2050, Oct. 2008. ISSN 00359009, 1477870X. doi: 10.1002/qj.334. URL http://doi.wiley.com/10.1002/qj.334.

O. Payrastre, V. Andreassian, S. Anquetin, X. Beaufils, L. Bonnifait, B. Boudevillain, O. Caumont, J.-D. Creutin, P. Davy, V. Ducrocq, A. Escudier, C. Fouchier, E. Gaume, B. Janet, P. Javelle, D. Lague, L. Lebouc, A. Lovat, C. Lutoff, D. Moncoulon, P. Naulin, T. Onfroy, C. Perrin, F. Pons, I. Ruin, G. Terti, and B. Vincendon. AMELIORER LA PREVISION IMMEDIATE DES CRUES SOUDAINES ET DE LEURS IMPACTS : LES OBJECTIFS DU PROJET DE RECHERCHE ANR PICS (2018-2021). page 6, 2018.

C. Perrin, C. Michel, and V. Andréassian. Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, 279(1):275 – 289, 2003. ISSN 0022-1694. doi: https://doi.org/10.1016/S0022-1694(03)00225-7. URL http://www.sciencedirect.com/science/article/pii/S0022169403002257. Number: 1.

D. T. Pham. Stochastic Methods for Sequential Data Assimilation in Strongly Nonlinear Systems. *MONTHLY WEATHER REVIEW*, 129:14, 2001.

F. Pianosi and T. Wagener. A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environmental Modelling & Software*, 67: 1–11, May 2015. ISSN 13648152. doi: 10.1016/j.envsoft.2015.01.004. URL https://linkinghub.elsevier.com/retrieve/pii/S1364815215000237.

A. E. Raftery, T. Gneiting, F. Balabdaoui, and M. Polakowski. Using Bayesian Model Averaging to Calibrate Forecast Ensembles. *Monthly Weather Review*, 133(5):1155–1174, 2005. doi: 10.1175/MWR2906.1. URL https://doi.org/10.1175/MWR2906.1. Number: 5.

M. Ramos. *Qualité et valeur des prévisions hydrologiques d'ensemble*. Habilitation à diriger des recherches, HDR, Sorbonne Université, 2018. URL https://hal.inrae.fr/tel-02608754.

R. H. Reichle, D. B. McLaughlin, and D. Entekhabi. Variational data assimilation of microwave radiobrightness observations for land surface hydrology applications. *IEEE Transactions on Geoscience and Remote Sensing*, 39(8):1708–1718, Aug. 2001. ISSN 1558-0644. doi: 10.1109/36.942549.

R. H. Reichle, D. B. Mclaughlin, and D. Entekhabi. Hydrologic Data Assimilation with the Ensemble Kalman Filter. *MONTHLY WEATHER REVIEW*, 130:12, 2002.

S. Ricci, A. Piacentini, O. Thual, E. Le Pape, and G. Jonville. Correction of upstream flow and hydraulic state with data assimilation in the context of flood forecasting. preprint,

Rivers and Lakes/Modelling approaches, 2011. URL https://hess.copernicus.org/preprints/7/9067/2010/hessd-7-9067-2010.pdf.

R. Romanowicz and K. Beven. Dynamic real-time prediction of flood inundation probabilities. *Hydrological Sciences Journal*, 43(2):181–196, Apr. 1998. ISSN 0262-6667, 2150-3435. doi: 10.1080/02626669809492117. URL https://www.tandfonline.com/doi/full/10.1080/02626669809492117.

L. Rouhier. *Régionalisation d'un modèle hydrologique distribué pour la modélisation de bassins non jaugés. Application aux vallées de la Loire et de la Durance*. Theses, Sorbonne Université, Oct. 2018. URL https://tel.archives-ouvertes.fr/tel-02409965. Issue: 2018SORUS316.

L. Rouhier, M. Le Lay, F. Garavaglia, N. Le Moine, F. Hendrickx, C. Monteil, and P. Ribstein. Impact of mesoscale spatial variability of climatic inputs and parameters on the hydrological response. *Journal of Hydrology*, 553:13–25, Oct. 2017. ISSN 00221694. doi: 10.1016/j.jhydrol.2017.07.037. URL http://linkinghub.elsevier.com/retrieve/pii/S0022169417305012.

H. Roux, D. Labat, P.-A. Garambois, M.-M. Maubourguet, J. Chorda, and D. Dartus. A physically-based parsimonious hydrological model for flash floods in Mediterranean catchments. *Natural Hazards and Earth System Sciences*, 11(9):2567–2582, Sept. 2011. ISSN 1684-9981. doi: 10.5194/nhess-11-2567-2011. URL https://nhess.copernicus.org/articles/11/2567/2011/.

P. T. Roy, N. E. Moçayd, S. Ricci, J.-C. Jouhaud, N. Goutal, M. De Lozzo, and M. C. Rochoux. Comparison of Polynomial Chaos and Gaussian Process surrogates for uncertainty quantification and correlation estimation of spatially distributed open-channel steady flows. *arXiv:1705.01440 [physics, stat]*, Oct. 2017. doi: 10.1007/s00477-017-1470-4. URL http://arxiv.org/abs/1705.01440. arXiv: 1705.01440.

J. Sacks, W. J. Welch, T. J. Mitchell, and H. P. Wynn. Design and Analysis of Computer Experiments. *Statistical Science*, 4(4):409–423, 1989. doi: 10.1214/ss/1177012413. URL https://doi.org/10.1214/ss/1177012413. Publisher: The Institute of Mathematical Statistics.

A. Saltelli. Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, 145(2):280–297, May 2002. ISSN 00104655. doi: 10.1016/S0010-4655(02)00280-1. URL https://linkinghub.elsevier.com/retrieve/pii/S0010465502002801.

A. Saltelli, editor. *Global sensitivity analysis: the primer*. John Wiley, Chichester, England ; Hoboken, NJ, 2008. ISBN 978-0-470-05997-5. OCLC: ocn180852094.

A. Saltelli and P. Annoni. How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software*, 25(12):1508–1517, Dec. 2010. ISSN 13648152. doi: 10.1016/j.envsoft.2010.04.012. URL https://linkinghub.elsevier.com/retrieve/pii/S1364815210001180. Number: 12.

A. Saltelli and R. Bolado. An alternative way to compute Fourier amplitude sensitivity test (FAST). *Computational Statistics & Data Analysis*, 26(4):445–460, Feb. 1998. ISSN 01679473. doi: 10.1016/S0167-9473(97)00043-1. URL https://linkinghub.elsevier.com/retrieve/pii/S0167947397000431. Number: 4.

A. Saltelli and I. M. Sobol'. About the use of rank transformation in sensitivity analysis of model output. *Reliability Engineering & System Safety*, 50(3):225 – 239, 1995. ISSN 0951-8320. doi: https://doi.org/10.1016/0951-8320(95)00099-2. URL http://www.sciencedirect.com/science/article/pii/0951832095000992. Number: 3.

A. Saltelli, S. Tarantola, and K. P.-S. Chan. A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. *Technometrics*, 41(1):39–56, 1999. ISSN 00401706. doi: 10.2307/1270993. URL http://www.jstor.org.extranet.enpc.fr/stable/1270993. Number: 1.

A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259–270, Feb. 2010. ISSN 00104655. doi: 10.1016/j.cpc.2009.09.018. URL https://linkinghub.elsevier.com/retrieve/pii/S0010465509003087.

J. Samuel, P. Coulibaly, G. Dumedah, and H. Moradkhani. Assessing model state and forecasts variation in hydrologic data assimilation. *Journal of Hydrology*, 513:127–141, May 2014. ISSN 00221694. doi: 10.1016/j.jhydrol.2014.03.048. URL https://linkinghub.elsevier.com/retrieve/pii/S0022169414002339.

G. Saporta. *Probabilités, Analyse des données et statistique*. Jan. 2006. URL https://hal.archives-ouvertes.fr/hal-01125195.

J. Schaake, J. Pailleux, J. Thielen, R. Arritt, T. Hamill, L. Luo, E. Martin, D. McCollor, and F. Pappenberger. Summary of recommendations of the first workshop on Postprocessing and Downscaling Atmospheric Forecasts for Hydrologic Applications held at Météo-France, Toulouse, France, 15-18 June 2009. *Atmospheric Science Letters*, 11(2):59–63, June 2010. ISSN 1530261X, 1530261X. doi: 10.1002/asl.267. URL http://doi.wiley.com/10.1002/asl.267. Number: 2.

R. Schefzik. A similarity-based implementation of the Schaake shuffle. *Monthly Weather Review*, 144(5):1909–1921, May 2016. ISSN 0027-0644, 1520-0493. doi: 10.1175/MWR-D-15-0227.1. URL http://arxiv.org/abs/1507.02079. arXiv: 1507.02079.

R. Schefzik, T. L. Thorarinsdottir, and T. Gneiting. Uncertainty Quantification in Complex Simulation Models Using Ensemble Copula Coupling. *Statistical Science*, 28(4):616–640, Nov. 2013. ISSN 0883-4237. doi: 10.1214/13-STS443. URL http://projecteuclid.org/euclid.ss/1386078881. Number: 4.

M. Scheuerer. Probabilistic Quantitative Precipitation Forecasting Using Ensemble Model Output Statistics. *Quarterly Journal of the Royal Meteorological Society*, 140(680):1086–1096, Apr. 2014. ISSN 00359009. doi: 10.1002/qj.2183. URL http://arxiv.org/abs/1302.0893. Number: 680.

M. Scheuerer and T. M. Hamill. Statistical Postprocessing of Ensemble Precipitation Forecasts by Fitting Censored, Shifted Gamma Distributions*. *Monthly Weather Review*, 143 (11):4578–4596, Nov. 2015. ISSN 0027-0644, 1520-0493. doi: 10.1175/MWR-D-15-0061.1. URL http://journals.ametsoc.org/doi/10.1175/MWR-D-15-0061.1.

G. J.-P. Schumann, J. C. Neal, N. Voisin, K. M. Andreadis, F. Pappenberger, N. Phanthuwongpakdee, A. C. Hall, and P. D. Bates. A first large-scale flood inundation forecasting model: Large-Scale Flood Inundation Forecasting. *Water Resources Research*, 49(10):6248–6257, Oct. 2013. ISSN 00431397. doi: 10.1002/wrcr.20521. URL http://doi.wiley.com/10.1002/wrcr.20521.

Y. Seity, P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson. The AROME-France Convective-Scale Operational Model. *Monthly Weather Review*, 139(3):976–991, Mar. 2011. ISSN 0027-0644, 1520-0493. doi: 10.1175/2010MWR3425.1. URL http://journals.ametsoc.org/doi/abs/10.1175/2010MWR3425.1. Number: 3.

D.-J. Seo, V. Koren, and N. Cajina. Real-Time Variational Assimilation of Hydrologic and Hydrometeorological Data into Operational Hydrologic Forecasting. *JOURNAL OF HYDROMETEOROLOGY*, 4:15, 2003.

D.-J. Seo, L. Cajina, R. Corby, and T. Howieson. Automatic state updating for operational streamflow forecasting via variational data assimilation. *Journal of Hydrology*, 367 (3-4):255–275, Apr. 2009. ISSN 00221694. doi: 10.1016/j.jhydrol.2009.01.019. URL https://linkinghub.elsevier.com/retrieve/pii/S0022169409000390.

A. G. Slater and M. P. Clark. Snow Data Assimilation via an Ensemble Kalman Filter. *Journal of Hydrometeorology*, 7(3):478–493, June 2006. ISSN 1525-7541, 1525-755X. doi: 10.1175/JHM505.1. URL http://journals.ametsoc.org/doi/10.1175/JHM505.1.

I. M. Sobol. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation*, 55(1):271 – 280, 2001. ISSN 0378-4754. doi: https://doi.org/10.1016/S0378-4754(00)00270-6. URL http://www.sciencedirect.com/science/article/pii/S0378475400002706. Number: 1.

M. Taillardat, O. Mestre, M. Zamo, and P. Naveau. Calibrated Ensemble Forecasts Using Quantile Regression Forests and Ensemble Model Output Statistics. *Monthly Weather Review*, 144(6):2375–2393, June 2016. ISSN 0027-0644, 1520-0493. doi: 10.1175/MWR-D-15-0260.1. URL http://journals.ametsoc.org/doi/10.1175/MWR-D-15-0260.1. Number: 6.

M. Taillardat, A.-L. Fougères, P. Naveau, and O. Mestre. Forest-based methods and ensemble model output statistics for rainfall ensemble forecasting. *arXiv preprint arXiv:1711.10937*, 2017.

M. Taillardat, A.-L. Fougères, P. Naveau, and O. Mestre. Forest-Based and Semiparametric Methods for the Postprocessing of Rainfall Ensemble Forecasting. *Weather and Forecasting*, 34(3):617–634, June 2019. ISSN 0882-8156, 1520-0434. doi: 10.1175/WAF-D-18-0149.1. URL http://journals.ametsoc.org/doi/10.1175/WAF-D-18-0149.1. Number: 3.

O. Talagrand, R. Vautard, and B. Strauss. Evaluation of probabilistic prediction systems. In *Workshop on Predictability, 20-22 October 1997*, pages 1–26, Shinfield Park, Reading, 1997. ECMWF. URL https://www.ecmwf.int/node/12555.

M. Tangara. *Nouvelle méthode de prévision de crue utilisant un modèle pluie-débit global*. PhD Thesis, 2005. URL http://www.theses.fr/2005EPHEA002.

A. Thiboult, F. Anctil, and M.-A. Boucher. Accounting for three sources of uncertainty in ensemble hydrological forecasting. *Hydrology and Earth System Sciences*, 20(5):1809–1825, May 2016. ISSN 1607-7938. doi: 10.5194/hess-20-1809-2016. URL https://www.hydrol-earth-syst-sci.net/20/1809/2016/. Number: 5.

J. Thielen, J. Schaake, R. Hartman, and R. Buizza. Aims, challenges and progress of the Hydrological Ensemble Prediction Experiment (HEPEX) following the third HEPEX workshop held in Stresa 27 to 29 June 2007. *Atmospheric Science Letters*, 9(2):29–35, Apr. 2008. ISSN 1530261X, 1530261X. doi: 10.1002/asl.168. URL http://doi.wiley.com/10.1002/asl.168. Number: 2.

J. Thielen, J. Bartholmes, P. de Tourvoie, and A. Cedex. The European Flood Alert System – Part 1: Concept and development. *Hydrol. Earth Syst. Sci.*, page 16, 2009.

G. Thirel, E. Martin, J.-F. Mahfouf, S. Massart, S. Ricci, and F. Habets. A past discharges assimilation system for ensemble streamflow forecasts over France – Part

1: Description and validation of the assimilation system. *Hydrology and Earth System Sciences*, 14(8):1623–1637, 2010a. doi: 10.5194/hess-14-1623-2010. URL https://www.hydrol-earth-syst-sci.net/14/1623/2010/. Number: 8.

G. Thirel, E. Martin, J.-F. Mahfouf, S. Massart, S. Ricci, F. Regimbeau, and F. Habets. A past discharge assimilation system for ensemble streamflow forecasts over France – Part 2: Impact on the ensemble streamflow forecasts. *Hydrology and Earth System Sciences*, 14(8):1639–1653, 2010b. doi: 10.5194/hess-14-1639-2010. URL https://www.hydrol-earth-syst-sci.net/14/1639/2010/. Number: 8.

J. Thomas Steven Savage, F. Pianosi, P. Bates, J. Freer, and T. Wagener. Quantifying the importance of spatial resolution and other factors through global sensitivity analysis of a flood inundation model. *Water Resources Research*, 52(11):9146–9163, 2016. doi: https://doi.org/10.1002/2015WR018198. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015WR018198.

T. L. Thorarinsdottir and T. Gneiting. Probabilistic forecasts of wind speed: ensemble model output statistics by using heteroscedastic censored regression. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 173(2):371–388, Apr. 2010. ISSN 09641998, 1467985X. doi: 10.1111/j.1467-985X.2009.00616.x. URL http://doi.wiley.com/10.1111/j.1467-985X.2009.00616.x.

T. L. Thorarinsdottir and M. S. Johnson. Probabilistic Wind Gust Forecasting Using Nonhomogeneous Gaussian Regression. *Monthly Weather Review*, 140(3):889–897, Feb. 2012. ISSN 0027-0644, 1520-0493. doi: 10.1175/MWR-D-11-00075.1. URL http://journals.ametsoc.org/doi/10.1175/MWR-D-11-00075.1.

A.-L. Tiberi-Wadier, M. Taillardat, N. Goutal, S. Ricci, P. Sergent, F. Bouttier, and E. Le Pape. Vers une approche ensembliste de la prévision des crues. In *De la prévision des crues à la gestion de crise*, Avignon, France, Nov. 2018. SHF. URL https://hal.archives-ouvertes.fr/hal-03250002.

A.-L. Tiberi-Wadier, N. Goutal, S. Ricci, P. Sergent, and C. Monteil. Sensitivity Analysis of the Mascaret model on the Odet River. In *Telemac-Mascaret User Club*, Toulouse, France, Oct. 2019. URL https://hal.archives-ouvertes.fr/hal-03250128.

A.-L. Tiberi-Wadier, N. Goutal, S. Ricci, P. Sergent, M. Taillardat, F. Bouttier, and C. Monteil. Strategies for hydrologic ensemble generation and calibration: on the merits of using model-based predictors. *Journal of Hydrology*, page 126233, 2021. ISSN 0022-1694. doi: https://doi.org/10.1016/j.jhydrol.2021.126233. URL https://www.sciencedirect.com/science/article/pii/S0022169421002808.

F. Tonelli, L. Casicci, A. Montani, S. Pecora, and M. Vergnani. Pre-operational use of a meteorological and hydrological/hydraulic ensemble approach on the Po River. page 1, 2011.

P. T.Roy, S. Ricci, R. Dupuis, R. Campet, J.-C. Jouhaud, and C. Fournier. BATMAN: Statistical analysis for expensive computer codes made easy. *The Journal of Open Source Software*, 3(21):493, 2018. ISSN 2475-9066. doi: 10.21105/joss.00493. URL http://dx.doi.org/10.21105/joss.00493. Number: 21.

D. Tuan Pham, J. Verron, and M. Christine Roubaud. A singular evolutive extended Kalman filter for data assimilation in oceanography. *Journal of Marine Systems*, 16 (3-4):323–340, Oct. 1998. ISSN 09247963. doi: 10.1016/S0924-7963(97)00109-7. URL https://linkinghub.elsevier.com/retrieve/pii/S0924796397001097.

J. M. Van Der Knijff, J. Younis, and A. P. J. De Roo. LISFLOOD: a GIS-based distributed model for river basin scale water balance and flood simulation. *International Journal of Geographical Information Science*, 24(2):189–212, Feb. 2010. ISSN 1365-8816, 1362-3087. doi: 10.1080/13658810802549154. URL http://www.tandfonline.com/doi/abs/10.1080/13658810802549154.

S. Vannitsem, D. Wilks, and J. Messner. *Statistical Postprocessing of Ensemble Forecasts*. 2018. ISBN 978-0-12-812372-0.

J. A. Velázquez, F. Anctil, M. H. Ramos, and C. Perrin. 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, Feb. 2011. ISSN 1680-7359. doi: 10.5194/adgeo-29-33-2011. URL https://www.adv-geosci.net/29/33/2011/.

M. Verlaan and A. W. Heemink. Tidal flow forecasting using reduced rank square root filters. *Stochastic Hydrology and Hydraulics*, 11(5):349–368, Oct. 1997. ISSN 0931-1955, 1436-3259. doi: 10.1007/BF02427924. URL http://link.springer.com/10.1007/BF02427924.

Viatgé, Julie, Berthet, Lionel, Marty, Renaud, Bourgin, François, Piotte, Olivier, Ramos, Maria-Helena, and Perrin, Charles. Vers une production en temps réel d'intervalles prédictifs associés aux prévisions de crue dans Vigicrues en France. *LHB*, (2):63–71, 2019a. doi: 10.1051/lhb/2019016. URL https://doi.org/10.1051/lhb/2019016. Number: 2.

Viatgé, Julie, Pinna, Thomas, Ficchì, Andrea, Perrin, Charles, Dorchies, David, Garandeau, Léa, and Tilmant, François. Vers une plus grande flexibilité temporelle du modèle opérationnel de prévision des crues GRP. *LHB*, (2):72–80, 2019b. doi: 10.1051/lhb/2019017. URL https://doi.org/10.1051/lhb/2019017. Number: 2.

J.-P. Vidal, E. Martin, L. Franchistéguy, M. Baillon, and J.-M. Soubeyroux. A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *International Journal of Climatology*, 30(11):1627–1644, Sept. 2010. ISSN 08998418. doi: 10.1002/joc.2003. URL http://doi.wiley.com/10.1002/joc.2003. Number: 11.

B. Vincendon, V. Ducrocq, O. Nuissier, and B. Vié. Perturbation of convection-permitting NWP forecasts for flash-flood ensemble forecasting. *Natural Hazards and Earth System Science*, 11(5):1529–1544, May 2011. ISSN 1684-9981. doi: 10.5194/nhess-11-1529-2011. URL http://www.nat-hazards-earth-syst-sci.net/11/1529/2011/. Number: 5.

B. Vincendon, S. Edouard, and V. Ducrocq. De l'incertitude dans un système de prévision d'ensemble des crues rapides méditerranéennes. *La Houille Blanche*, (3-4):22–30, Oct. 2019. ISSN 0018-6368, 1958-5551. doi: 10.1051/lhb/2019051. URL https://www.shf-lhb.org/10.1051/lhb/2019051.

T. Wagener, D. P. Boyle, M. J. Lees, H. S. Wheater, H. V. Gupta, and S. Sorooshian. A framework for development and application of hydrological models. *Hydrology and Earth System Sciences*, 5(1):13–26, Mar. 2001. ISSN 1607-7938. doi: 10.5194/hess-5-13-2001. URL https://hess.copernicus.org/articles/5/13/2001/.

A. H. Weerts, G. Y. El Serafy, S. Hummel, J. Dhondia, and H. Gerritsen. Application of generic data assimilation tools (DATools) for flood forecasting purposes. *Computers & Geosciences*, 36(4):453–463, Apr. 2010. ISSN 00983004. doi: 10.1016/j.cageo.2009.07.009. URL https://linkinghub.elsevier.com/retrieve/pii/S009830041000021X.

A. P. Weigel. Ensemble Forecasts. In *Forecast Verification*, pages 141–166. John Wiley & Sons, Ltd, 2012. ISBN 978-1-119-96000-3. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9781119960003.ch8.

D. S. Wilks. *Statistical methods in the atmospheric science, Third Edition*. 2011.

D. S. Wilks. Enforcing calibration in ensemble postprocessing: Enforcing Calibration in Ensemble Postprocessing. *Quarterly Journal of the Royal Meteorological Society*, 144 (710):76–84, Jan. 2018a. ISSN 00359009. doi: 10.1002/qj.3185. URL http://doi.wiley.com/10.1002/qj.3185.

D. S. Wilks. Univariate ensemble post-processing. In S. Vannitsem, D. S. Wilks, and J. W. Messner, editors, *Statistical postprocessing of ensemble forecasts*, Elsevier, pages 49–89. Amsterdam, Netherlands, 2018b.

Y. Yoon, M. Durand, C. J. Merry, E. A. Clark, K. M. Andreadis, and D. E. Alsdorf. Estimating river bathymetry from data assimilation of synthetic SWOT measurements. *Journal of Hydrology*, 464-465:363–375, Sept. 2012. ISSN 00221694. doi:

10.1016/j.jhydrol.2012.07.028. URL https://linkinghub.elsevier.com/retrieve/pii/S0022169412006294.

Y. Yoon, P. Garambois, R. C. Paiva, M. Durand, H. Roux, and E. Beighley. Improved error estimates of a discharge algorithm for remotely sensed river measurements: Test cases on S acramento and G aronne R ivers. *Water Resources Research*, 52(1):278–294, Jan. 2016. ISSN 0043-1397, 1944-7973. doi: 10.1002/2015WR017319. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/2015WR017319.

I. Zalachori. *Prévisions hydrologiques d'ensemble: développements pour améliorer la qualité des prévisions et estimer leur utilité*. PhD Thesis, AgroParisTech, 2013.

I. Zalachori, M.-H. Ramos, R. Garçon, T. Mathevet, and J. Gailhard. Statistical processing of forecasts for hydrological ensemble prediction: a comparative study of different bias correction strategies. *Advances in Science and Research*, 8:135–141, July 2012. ISSN 1992-0636. doi: 10.5194/asr-8-135-2012. URL http://www.adv-sci-res.net/8/135/2012/.

M. Zamo, O. Mestre, P. Arbogast, and O. Pannekoucke. A benchmark of statistical regression methods for short-term forecasting of photovoltaic electricity production. Part II: Probabilistic forecast of daily production. *Solar Energy*, 105:804 – 816, 2014. ISSN 0038-092X. doi: https://doi.org/10.1016/j.solener.2014.03.026. URL http://www.sciencedirect.com/science/article/pii/S0038092X14001601.