Etude des incertitudes des modèles neuronaux sur la prévision hydrogéologique. Application à des bassins versants de typologies différentes

Nicolas Akil

▶ **To cite this version:**

Nicolas Akil. Etude des incertitudes des modèles neuronaux sur la prévision hydrogéologique. Application à des bassins versants de typologies différentes. Sciences de la Terre. IMT - MINES ALES - IMT - Mines Alès Ecole Mines - Télécom, 2021. Français. NNT : 2021EMAL0005 . tel-03767834

HAL Id: tel-03767834
https://theses.hal.science/tel-03767834

Submitted on 2 Sep 2022

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.

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'INSTITUT MINES-TELECOM (IMT) – ÉCOLE NATIONALE SUPÉRIEURE DES MINES D'ALÈS (IMT MINES ALÈS)

En Sciences de la Terre et de l'Eau

École doctorale [GAIA – Biodiversité, Agriculture, Alimentation, Environnement, Terre, Eau]
Portée par l'Université de Montpellier

Unité de recherche HydroSciences Montpellier

Etude des incertitudes des modèles neuronaux sur la prévision hydrogéologique - Application à plusieurs bassins de typologies différentes

Présentée par Nicolas AKIL

Le 15 septembre 2022

Sous la direction de Anne Johannet
et de Marc Vinches

Devant le jury composé de

Nathalie DORFLIGER, Danone	Rapporteure
Didier GRAILLOT, IMT Mines Saint-Etienne	Rapporteur
Marie-George TOURNOUD, Polytech Montpellier	Présidente du jury
Nicolas MASSEI, Université de Rouen	Examinateur
Guillaume ARTIGUE, IMT Mines Alès	Examinateur
Michaël SAVARY, aQuasys	Invité
Jérôme NICOLAS, BRGM	Invité
Anne JOHANNET, IMT Mines Alès	Directrice de thèse

IMT Mines Alès
École Mines-Télécom

COLLÈGE DOCTORAL
UNIVERSITÉ DE MONTPELLIER

Remerciements

Je tiens à remercier chaleureusement Madame Anne JOHANNET, ma directrice de thèse, qui m'a aidé dans ma recherche, m'a fait partager son expérience dans le domaine de l'intelligence artificielle appliquée à l'hydrologie, m'aidant ainsi à grandir, et qui a réalisé bon nombre de relectures. Elle a su encadrer cette thèse avec succès malgré les difficultés distancielles et sanitaires.

Je remercie également Monsieur Marc VINCHES, mon co-directeur de thèse qui m'a donné des outils statistiques pour mener à bien ce projet et qui a su poser les bonnes questions pour recadrer le projet.

Je souhaite également remercier l'intégralité des membres du Jury de thèse, composé de Mesdames Nathalie DORFLIGER et Marie-George TOURNOUD et Messieurs Nicolas MASSEI, Didier GRAILLOT, Séverin PISTRE, Jérôme NICOLAS et Michaël SAVARY. Cette gratitude est également valable pour les personnes ayant assisté aux différents comités de thèse, dont Messieurs Pierre ROUSSEL et Jérôme NICOLAS qui ont su poser des questions pertinentes sur le sujet et m'ont permis d'identifier certains verrous technologiques et Monsieur Luc MALHAUTIER pour avoir présidé un de ces comités.

Je remercie une nouvelle fois Monsieur Séverin PISTRE, référent de l'Ecole Doctorale, qui, outre ses questions lors des comités de thèse, est resté vigilant sur l'état psychologique des doctorants dont il avait la charge, suite aux confinements successifs causés par la pandémie du coronavirus.

Je désire exprimer toute ma sympathie envers le BRGM, qui a fourni les données dont j'avais besoin et qui s'est impliqué dans le projet, l'OUGC84 pour les données agricoles, GEONOSIS, pour avoir développé un outil de modélisation hydrogéologique performant, ainsi que l'ARNT, pour avoir participé au financement de cette thèse.

Bien évidemment, il me paraît primordial de remercier aQuasys, ainsi que son fondateur Monsieur Stéphane BARTHON et son directeur Monsieur Dominique THERIEZ, sans qui je n'aurais jamais pu vivre cette expérience et qui ont initié le projet. Je souhaite également remercier chacun des employés et alternants, actuels ou passés par cette entreprise, pour leur bonne humeur, leur soutien indéfectible et leur professionnalisme.

Bien entendu, je suis obligé de te remercier, Michael SAVARY, car tu as su être à la fois un collègue au sein de l'IMT Mines Alès, un collègue de travail à aQuasys, un encadrant de thèse pertinent et surtout un ami. Je n'ai pas assez de mots pour exprimer toute ma gratitude.

De même, je te remercie fort chaleureusement mon second encadrant, Guillaume ARTIGUE, pour tout le travail fourni, les nombreux conseils que tu as su me prodiguer, et aussi les discussions non-affiliées au travail. Sa sympathie et son professionnalisme sont pour moi des exemples à suivre.

J'ai une pensée particulière pour ceux qui m'ont hébergé toutes ces années lors de mes venues sur Alès. Je pense notamment à toi Laurence, et à toi David. Merci pour ces moments de partage.

Je remercie également tous ceux que j'ai pu côtoyer dans les laboratoires, ne serait-ce le temps d'un café, ou d'une partie de football, organisée par l'ATHEMA. Je n'oublie pas bien sûr mes fidèles acolytes de laboratoire, devenus ceux que j'appelle mes « Amis du Sud », Valentin, Aurélie, Michael, Bob, Cagri, Alienor, Maria, … et bien sur leurs amis et compagnons, Dylan, Delphine, Emeline, Solène, … Merci à tous d'avoir été présents, dans les moments difficiles, dans les moments de doutes ou encore juste pour les partages de bons moments (Prague restera un beau souvenir).

Je remercie aussi mes autres amis (ceux du Nord), Anthony, Mahir, Delphine, Chloé, Ester, Noémie, Carole… Merci à vous d'être ce que vous êtes, ne changez rien. Vous m'avez été d'un grand soutien et m'avez permis de souffler quand j'en avais besoin.

Je remercie aussi ma famille, sans qui je ne serais jamais devenu ce que je suis aujourd'hui. Je ne t'oublie pas ma sœur, et je suis très heureux que la famille s'agrandisse.

Pour finir, j'aimerais remercier de tout mon cœur ma compagne, Catherine, qui me rend meilleur chaque jour. Merci pour ton amour et ton soutien inconditionnels.

Sommaire

REMERCIEMENTS **3**

SOMMAIRE **5**

LISTE DES CARTES **9**

LISTE DES FIGURES **9**

LISTE DES TABLEAUX **12**

RESUME **15**

ABSTRACT **17**

PREAMBULE **19**

L'ENTREPRISE AQUASYS 19
LE CENTRE CREER D'IMT MINES ALES 19

Introduction

Chapitre I : La ressource en eau et sa gestion par la prévision hydrogéologique

I. INTRODUCTION **29**

II. LA RESSOURCE EN EAU, ENTRE ABONDANCE ET INSUFFISANCE **29**

II.1 ETAT DES LIEUX DE LA RESSOURCE EN EAU 29
II.2 CONTEXTE DEMOGRAPHIQUE : UNE CONSOMMATION ACCRUE 31
II.3 CONTEXTE CLIMATIQUE : DES EVENEMENTS EXTREMES EN AUGMENTATION 32

III. LA CONNAISSANCE POUR UNE MEILLEURE GESTION **34**

III.1 QUANTIFICATION DE LA RESSOURCE EN EAU 34
III.1.1. Eau de surface *34*
III.1.2. Eau souterraine *36*
III.1.3. Règlementations *36*
III.2 VULNERABILITE AUX EXTREMES CLIMATIQUES 37
III.2.1. Sécheresse *37*
III.2.2. Inondation *39*
III.3 PREVISION ET GESTION EN TEMPS REEL 41
III.3.1. Gestion en temps réel *41*
III.3.2. Approches de modélisation *42*

IV. DE LA MODELISATION A L'ACTION : LES IMPACTS DES INEXACTITUDES **44**

V. POSITION DE LA PROBLEMATIQUE **46**

Chapitre II : Méthodes : réseaux de neurones artificiels et autres méthodes statistiques

I. INTRODUCTION **51**

II. NEURONE FORMEL **51**

III. LE RESEAU DE NEURONES 52

III.1 DEFINITION 52
III.2 TYPOLOGIE 52
III.2.1. Caractéristiques principales 52
III.2.2. Modèles statique et dynamique 53
III.2.3. Principales architectures 54
III.3 APPRENTISSAGE DES RESEAUX DE NEURONES 58
III.3.1. Principe de l'apprentissage 58
III.3.2. Règles d'apprentissage 60
III.3.3. Dilemme biais-variance et sur-ajustement 63
III.3.4. Méthodes de régularisation 64
III.3.5. Choix des variables 66
III.4 UTILISATION ET CONCEPTION DES MODELES 69
III.5 EVALUATION DE LA PERFORMANCE D'UN MODELE 71
III.5.1. Evaluation de la qualité moyenne de la modélisation 71
III.5.2. Evaluation du décalage de la modélisation 73
III.5.3. Evaluation spécifique de la modélisation d'évènements 73
III.5.4. Evaluation de la surestimation de la modélisation 74
III.5.5. Evaluation de la justesse de la modélisation 74

IV. CONCLUSIONS DU CHAPITRE II 76

Chapitre III : De l'incertitude à une approche de l'exactitude dans la perspective d'une modélisation hydrogéologique

I. INTRODUCTION 79

II. DEFINITION DE L'INCERTITUDE 79

III. ORIGINE DES INCERTITUDES ET LEUR PRISE EN COMPTE 80

III.1 INCERTITUDES DE MESURE 81
III.1.1. Vision métrologique de l'incertitude de mesure 82
III.1.2. Prétraitements et qualité des données 85
III.1.3. Incertitudes rencontrées sur les grandeurs hydrologiques 85
III.2 INCERTITUDES ENVIRONNEMENTALES 91
III.3 INCERTITUDES DUES AUX PRETRAITEMENTS 91
III.4 INCERTITUDES DE DISTRIBUTION SPATIALE 92
III.4.1. Discrétisation spatiale 92
III.4.2. Polygones de Thiessen 92
III.4.3. Détermination des incertitudes par permutations d'entrées 93
III.5 INCERTITUDES DE REPRESENTATIVITE TEMPORELLE 94
III.6 INCERTITUDES DE MODELISATION 95
III.6.1. Approches d'ensemble 95
III.6.2. Incertitudes liées à l'initialisation et l'estimation des paramètres 95
III.6.3. Définitions et inférence statistique 99
III.6.4. Fonction de vraisemblance 99
III.6.5. Densités de probabilités 100

IV. PROPOSITION D'UNE METHODOLOGIE D'ESTIMATION DE L'INCERTITUDE DE MODELISATION 102

IV.1 INCERTITUDES IDENTIFIEES DANS CE TRAVAIL 102
IV.2 INCERTITUDES DE MODELISATION PAR DENSITE DE PROBABILITE 103
IV.2.1. Détermination du nombre d'initialisations optimal 103
IV.2.2. Application des densités de probabilités 105

V. CONCLUSIONS DU CHAPITRE III 112

Chapitre IV – Applications des méthodologies d'estimation des incertitudes

I. INTRODUCTION 117

II. LA NAPPE DU BASSIN DU BLAVET 117

II.1 CONTEXTE GENERAL 117
II.1.1. Généralités 117
II.1.2. Hydrographie et hydrogéologie 120
II.1.3. Climat 124
II.1.4. Géologie 126
II.1.5. Occupation des sols 128
II.2 RESEAU DE MESURES ET SITE D'ETUDE 131
II.3 RESULTATS DES EXPERIMENTATIONS 134
II.3.1. Au pas de temps journalier : modélisation à court-terme 134
II.3.2. Au pas de temps décadaire : modélisation à moyen-terme 160
II.3.3. Conclusion sur le bassin de la nappe du Blavet 175

III. LA NAPPE DE LA CRAIE DE CHAMPAGNE SUD ET CENTRE 177

III.1 CONTEXTE GENERAL 177
III.1.1. Généralités 177
III.1.2. Hydrographie et hydrogéologie 178
III.1.3. Climat 183
III.1.4. Géologie 184
III.1.5. Occupation des sols 186
III.1.6. Ecosystèmes 189
III.2 RESEAU DE MESURES ET SITE D'ETUDE 191
III.2.1. Présentation des données récoltées 191
III.2.2. Reconstitution des données de prélèvements pour l'irrigation 193
III.3 RESULTATS DES EXPERIMENTATIONS 194
III.3.1. Compléments à l'article 225
III.3.2. Principales conclusions de l'article 227

IV. DISCUSSIONS 229

IV.1 CONCLUSIONS GENERALES SUR LES MODELISATIONS REALISEES SUR LES DEUX BASSINS 229
IV.2 AVANTAGES ET INCONVENIENTS DE LA METHODOLOGIE DEVELOPPEE 229
IV.2.1. Avantages 229
IV.2.2. Inconvénients 231

Conclusion

BIBLIOGRAPHIE 238

A-D 238
E-H 242
I-L 244
M-P 248
Q-U 252
V-Z 254

Annexes

ANNEXE I : LISTE DES LOIS DE PROBABILITES TESTEES 258
ANNEXE II : PRESENTATION DES PRINCIPALES LOIS DE PROBABILITES UTILISEES 259
Loi Uniforme *259*
Loi Normale *259*
Loi de Gumbel *260*
Loi de Laplace *261*
Loi de Cauchy *261*
Loi du Cosinus surélevé *262*
Loi Logistique *262*
Loi de Kumaraswamy *263*
Loi de Slash *263*
Loi de Bhattacharjee *264*
Loi de Huber *264*
ANNEXE III : LE *MACHINE LEARNING* AU SERVICE DES PROBABILITES 266
Chaînes de Markov *266*
Algorithme de Metropolis-Hastings *266*
ANNEXE IV : OUTILS INFORMATIQUES UTILISES 268

Liste des cartes

CARTE IV-1 – PRESENTATION DE LA NAPPE DU BASSIN DU BLAVET. 119
CARTE IV-2 - OBSTACLES A L'ECOULEMENT SUR LE BLAVET ET SES AFFLUENTS. 121
CARTE IV-3 – ISOPIEZES DU BASSIN DU BLAVET LORS D'UN EVENEMENT SEC (SEPT.-DEC. 2017). 123
CARTE IV-4 – ISOPIEZES DU BASSIN DU BLAVET LORS D'UN EVENEMENT HUMIDE (JAN.-MARS 2014). 123
CARTE IV-5 – CLIMAT SUR LA NAPPE DU BASSIN DU BLAVET. 125
CARTE IV-6 - GEOLOGIE SIMPLIFIEE DE LA NAPPE DU BASSIN DU BLAVET. 126
CARTE IV-7 – INDICE DE DEVELOPPEMENT ET DE PERSISTANCE DES RESEAUX (IDPR) SUR LA NAPPE DU BASSIN DU BLAVET. 127
CARTE IV-8 - OCCUPATION DES SOLS SUR LA NAPPE DU BASSIN DU BLAVET (CORINE LAND COVER, 2018). 128
CARTE IV-9 - REGISTRE PARCELLAIRE GRAPHIQUE DU BASSIN DE LA NAPPE DU BLAVET (RPG, 2017). 130
CARTE IV-10 - LOCALISATION DES STATIONS DE MESURE DU BASSIN DU BLAVET 133
CARTE IV-11 - PRESENTATION DE LA NAPPE DE LA CRAIE DE CHAMPAGNE SUD ET CENTRE. 178
CARTE IV-12 - OBSTACLES A L'ECOULEMENT SUR LA SEINE AMONT, LA MARNE ET LEURS AFFLUENTS 180
CARTE IV-13 – ISOPIEZES DE LA NAPPE DE LA CRAIE LORS D'UN EVENEMENT SEC (SEPT.-DEC. 1990). 181
CARTE IV-14 – ISOPIEZES DE LA NAPPE DE LA CRAIE LORS D'UN EVENEMENT HUMIDE (JANV.-MARS 1982). 181
CARTE IV-15- CLIMAT SUR LE BASSIN DE LA NAPPE DE LA CRAIE DE CHAMPAGNE SUD ET CENTRE. 183
CARTE IV-16 - GEOLOGIE SIMPLIFIEE DE LA NAPPE DE LA CRAIE DE CHAMPAGNE SUD ET CENTRE. 184
CARTE IV-17 - INDICE DE DEVELOPPEMENT ET DE PERSISTANCE DES RESEAUX (IDPR) SUR LA NAPPE DE LA CRAIE. 186
CARTE IV-18 - OCCUPATION DES SOLS SUR LA NAPPE DE LA CRAIE DE CHAMPAGNE SUD ET CENTRE (CORINE LAND COVER, 2018). 187
CARTE IV-19 - REGISTRE PARCELLAIRE GRAPHIQUE DU BASSIN DE LA NAPPE DE LA CRAIE DE CHAMPAGNE (RPG, 2017). 189
CARTE IV-20 - RESEAU DE MESURES DU BASSIN DE LA NAPPE DE LA CRAIE DE CHAMPAGNE. 191

Liste des figures

FIGURE I-1 - LE CYCLE DE L'EAU. 30
FIGURE II-1 - SCHEMA DU FONCTIONNEMENT D'UN NEURONE. 52
FIGURE II-2 - SCHEMA D'UN PERCEPTRON MULTICOUCHE. 55
FIGURE II-3 - REPRESENTATION SCHEMATIQUE D'UN RESEAU LONG SHORT TERM MEMORY. 58
FIGURE II-4 - ILLUSTRATION DE L'EXISTENCE DE MINIMUMS LOCAUX. 61
FIGURE II-5 - DILEMME BIAIS-VARIANCE. 64
FIGURE II-6 - PRINCIPE DE LA VALIDATION CROISEE. 65

FIGURE II-7 - AUTOCORRELOGRAMMES SUR LA STATION DE HOLTZWIHR EN 2016 DES PLUIES (A.) ET DES DEBITS (B.). 67

FIGURE II-8 – EXEMPLE DE CORRELOGRAMME CROISE 68

FIGURE II-9 SCHEMA DE LA METHODOLOGIE DE CONCEPTION DE MODELES. 70

FIGURE III-1 - EVOLUTION THEORIQUE DES INCERTITUDES PAR RAPPORT A LA COMPLEXITE D'UN MODELE (SHRESTHA, 2009). 81

FIGURE III-2 - SCHEMATISATION DE L'INCERTITUDE DE MESURE. 82

FIGURE III-3 - EXEMPLE DE DIAGRAMME D'ISHIKAWA APPLIQUE A L'ERREUR DE MESURE (D'APRES TAILLADE, 2005). 83

FIGURE III-4 - DEFINITION DE LA JUSTESSE, DE LA FIDELITE ET DE L'EXACTITUDE. 84

FIGURE III-5 - SCHEMA DE L'INFLUENCE D'UN CONE DE RABATTEMENT CAUSE PAR UN POMPAGE. 89

FIGURE III-6 - SCHEMA DE FONCTIONNEMENT DE LA METHODOLOGIE UNEEC (D'APRES SOLOMATINE & SHRESTHA, 2009). 97

FIGURE III-7 - DISTRIBUTION DE LA DENSITE DE PROBABILITE. 100

FIGURE III-8 - SELECTION DE L'ARCHITECTURE PAR VALIDATION CROISEE. 104

FIGURE III-9 - ECART-TYPE DES SORTIES MODELISEES EN FONCTION DU NOMBRE D'INITIALISATIONS (SAVARY, 2018). 105

FIGURE III-10 - EXEMPLE DE REPRESENTATION DE MARGE D'ERREUR. 109

FIGURE III-11 - CALCUL DES NIVEAUX PIEZOMETRIQUES A PARTIR DES PROBABILITES DE JUSTESSE. 110

FIGURE III-12 SCHEMA DE LA METHODOLOGIE D'ESTIMATION DE L'INCERTITUDE D'INITIALISATION DES PARAMETRES DU MODELE. 111

FIGURE III-13 SCHEMA DE LA METHODOLOGIE D' L'ESTIMATION DE L'INCERTITUDE LIEE AUX RESULTATS DE MODELISATION. 112

FIGURE IV-1 - DISPONIBILITE DES DONNEES (NAPPE DU BLAVET). 132

FIGURE IV-2 - EVOLUTION MENSUELLE DU *PREDICTION CONFIDENCE CRITERION* EN FONCTION DU NOMBRE D'INITIALISATIONS. 145

FIGURE IV-3 - COMPARAISON ENTRE UNE DOUBLE ET UNE TRIPLE LOI NORMALE INFERANT LA DISTRIBUTION DES NIVEAUX PIEZOMETRIQUES OBSERVES. 147

FIGURE IV-4 - DENSITES DE PROBABILITE INFEREES A LA DISTRIBUTION DES NIVEAUX PIEZOMETRIQUES EN AUGMENTATION. 149

FIGURE IV-5 - MODELES DE JUSTESSE POUR LES NIVEAUX PIEZOMETRIQUES EN AUGMENTATION. 151

FIGURE IV-6 - DENSITES DE PROBABILITE INFEREES A LA DISTRIBUTION DES NIVEAUX PIEZOMETRIQUES EN DIMINUTION. 154

FIGURE IV-7 - MODELES DE JUSTESSE POUR LES NIVEAUX PIEZOMETRIQUES EN DIMINUTION. 156

FIGURE IV-8 - PROBABILITE DE JUSTESSE DES NIVEAUX PIEZOMETRIQUES MEDIANS SIMULES EN FONCTION DE LA PROBABILITE DES NIVEAUX PIEZOMETRIQUES OBSERVES. 158

FIGURE IV-9 - MODELISATION DES NIVEAUX PIEZOMETRIQUES AVEC INTERVALLE DE PREVISION OBTENU PAR UNE SELECTION DE LA PROBABILITE DE JUSTESSE A 0,9..... 159

FIGURE IV-10 - ARCHITECTURE SELECTIONNEE POUR UNE MODELISATION DES NIVEAUX PIEZOMETRIQUES A MOYEN-TERME. 160

FIGURE IV-11 - RESULTAT DE MODELISATION DE L'ENSEMBLE DE VALIDATION DES NIVEAUX D'EAU SOUTERRAINE A UN HORIZON DE PREVISION DE 30 JOURS. 161

FIGURE IV-12 - EVOLUTION DU CRITERE *PREDICTION CONFIDENCE CRITERION (C_{PC})* EN FONCTION DU NOMBRE D'INITIALISATIONS..... 162

FIGURE IV-13 - PREVISION MEDIANE A 30 JOURS DES NIVEAUX PIEZOMETRIQUES ET INTERVALLE DE PREVISION GENERE PAR 20 MODELISATIONS AVEC DES INITIALISATIONS DIFFERENTES. 162

FIGURE IV-14 - DENSITES DE PROBABILITE INFEREES A LA DISTRIBUTION DES NIVEAUX PIEZOMETRIQUES EN AUGMENTATION. 165

FIGURE IV-15 - MODELES DE JUSTESSE POUR LES NIVEAUX PIEZOMETRIQUES EN AUGMENTATION. 168

FIGURE IV-16 - DENSITES DE PROBABILITE INFEREES A LA DISTRIBUTION DES NIVEAUX PIEZOMETRIQUES EN DIMINUTION. 170

FIGURE IV-17 - MODELES DE JUSTESSE POUR LES NIVEAUX PIEZOMETRIQUES EN DIMINUTION. 172

FIGURE IV-18 - PROBABILITE DE JUSTESSE DES NIVEAUX PIEZOMETRIQUES MEDIANS SIMULES EN FONCTION DE LA PROBABILITE DES NIVEAUX PIEZOMETRIQUES OBSERVES..... 174

FIGURE IV-19 - MODELISATION DES NIVEAUX PIEZOMETRIQUES AVEC INTERVALLE DE PREVISION OBTENU PAR UNE SELECTION DE LA PROBABILITE DE JUSTESSE A 0,7..... 174

FIGURE IV-20 - EVOLUTION DES PRELEVEMENTS EN EAU AU PIEZOMETRE DE LA CRAIE A VAILLY. 182

FIGURE IV-21 - DISPONIBILITE DES DONNEES (NAPPE DE LA CRAIE). 192

FIGURE IV-22 - DOUBLES CUMULS POUR LE PLUVIOMETRE DE MAILLY..... 193

FIGURE IV-23 - COURBE D'IRRIGATION POTENTIELLE PAR UNITE DE SURFACE. 194

FIGURE IV-24 - PROBABILITE DE JUSTESSE DES NIVEAUX PIEZOMETRIQUES MEDIANS SIMULES EN FONCTION DE LA PROBABILITE DES NIVEAUX PIEZOMETRIQUES OBSERVES..... 227

FIGURE A1-1 - DENSITE DE PROBABILITE DE LA LOI UNIFORME. 259

FIGURE A1-2 - DENSITE DE PROBABILITE DE LA LOI NORMALE. 259

FIGURE A1-3 - DENSITE DE PROBABILITE DE LA LOI DE GUMBEL..... 260

FIGURE A1-4 - DENSITE DE PROBABILITE DE LA LOI DE LAPLACE. 261

FIGURE A1-5 - DENSITE DE PROBABILITE DE LA LOI DE CAUCHY. 261

FIGURE A1-6 - DENSITE DE PROBABILITE DE LA LOI DU COSINUS SURELEVE. 262

FIGURE A1-7 - DENSITE DE PROBABILITE DE LA LOI LOGISTIQUE. 262

FIGURE A1-8 - DENSITE DE PROBABILITE DE LA LOI DE KUMARASWAMY. 263

FIGURE A1-9 - DENSITE DE PROBABILITE DE LA LOI DE SLASH. 263

FIGURE A1-10 - DENSITE DE PROBABILITE DE LA LOI DE BHATTACHARJEE. 264

FIGURE A1-11 - DENSITE DE PROBABILITE DE LA LOI DE HUBER. 264

Liste des tableaux

TABLEAU I-1 - LISTE NON EXHAUSTIVE D'EVENEMENTS ASSOCIES AU RISQUE DE SECHERESSE EN EUROPE. BILAN ECONOMIQUE ET SOCIAL. 38

TABLEAU I-2 - LISTE NON EXHAUSTIVE D'EVENEMENTS ASSOCIES AU RISQUE INONDATIONS EN FRANCE. BILAN ECONOMIQUE ET SOCIAL. 40

TABLEAU III-1 - COMBINAISONS POSSIBLES POUR DES PERMUTATIONS DE DEUX ENTREES. 94

TABLEAU IV-1 - STATISTIQUES DES STATIONS SUR LE BASSIN DE LA NAPPE DU BLAVET. 132

TABLEAU IV-2 - INDICATEURS D'INFERENCE DES LOIS A DENSITE DE PROBABILITE. 148

TABLEAU IV-3 - CORRELATION DES MODELES DE JUSTESSE POUR LES NIVEAUX PIEZOMETRIQUES EN AUGMENTATION. 150

TABLEAU IV-4 - INDICATEURS D'INFERENCE DES LOIS A DENSITE DE PROBABILITE. 153

TABLEAU IV-5 - CORRELATION DES MODELES DE JUSTESSE POUR LES NIVEAUX PIEZOMETRIQUES EN DIMINUTION. 155

TABLEAU IV-6 - INDICATEURS D'INFERENCE DES LOIS A DENSITE DE PROBABILITE. 163

TABLEAU IV-7 - CORRELATION DES MODELES DE JUSTESSE POUR LES NIVEAUX PIEZOMETRIQUES EN AUGMENTATION. 166

TABLEAU IV-8 - INDICATEURS D'INFERENCE DES LOIS A DENSITE DE PROBABILITE. 169

TABLEAU IV-9 - CORRELATION DES MODELES DE JUSTESSE POUR LES NIVEAUX PIEZOMETRIQUES EN DIMINUTION. 171

TABLEAU IV-10 - STATISTIQUES DES STATIONS DE LA NAPPE DE LA CRAIE DE CHAMPAGNE. 192

TABLEAU IV-11 - INDICATEURS D'INFERENCE DES LOIS A DENSITE DE PROBABILITE. 225

TABLEAU IV-12 - INDICATEURS D'INFERENCE DES LOIS A DENSITE DE PROBABILITE. 226

« Je ne connais rien d'autre si propre à frapper l'imagination que cette merveilleuse forme d'ordre cosmique donnée par la Loi de Fréquence des Erreurs... Elle règne avec sérénité et en toute abnégation au milieu de la confusion sauvage. »

— Francis Galton, 1889, *Natural Inheritence*

Résumé

Etude des incertitudes des modèles neuronaux sur la prévision hydro(géo)logique. Application à des bassins versants de typologies différentes.

Les crues et les sécheresses sont deux des risques majeurs en France et seront accentués par le changement climatique. La modélisation et la prévision de ces phénomènes est donc un élément incontournable pour la gestion des nappes et des rivières.

Actuellement, les débits ou hauteurs d'eau sont principalement anticipés à partir de modèles à base physique ou conceptuelle. Bien qu'efficaces et nécessaires, la calibration et la mise en œuvre de ces modèles nécessitent souvent la réalisation d'études longues et coûteuses.

Concernant les nappes souterraines, le comportement des bassins d'étude est dans la plupart des cas difficile à appréhender, tant du fait de l'hétérogénéité des infiltrations et du sous-sol que des conditions d'observation coûteuses car nécessitant le plus souvent des forages. Les modèles hydrogéologiques prédictifs sont donc insuffisamment utilisés par les gestionnaires tant du fait de leur coût que du manque de confiance en leurs sorties.

Dans ce contexte, cette thèse, codirigée par l'IMT Mines Alès et conjointement financée par la société aQuasys et l'ANRT, a pour objectif de développer des modèles issus du paradigme systémique. En effet, ceux-ci ont l'avantage de pouvoir être conçus uniquement à partir des seules informations d'entrées et de sorties (pluies et débits/hauteurs), limitant ainsi les investigations et mesures complémentaires.

Les modèles choisis sont les réseaux neuronaux du fait de leurs capacités à identifier des phénomènes non-linéaires et dynamique. Cependant les incertitudes associées aux sorties de ces modèles sont rarement quantifiées et ceci ne permet pas aux gestionnaires de prendre des décisions de manière raisonnée. Cette thèse cherche donc à répondre à trois objectifs principaux :

1. Généralisation d'une méthode de conception de modèle adaptée à différentes variables (débits/hauteur des eaux de surface) et à des bassins de types très différents : bassins versants ou bassins hydrogéologiques (hauteur des eaux souterraines).

2. Évaluation des incertitudes liées à ces modèles, en particulier pour l'impact de l'initialisation des paramètres du modèle à réseaux de neurones, que celle de l'hétérogénéité des pluies de recharge.

3. Maitrise de ces incertitudes.

Plusieurs bassins exploités pour l'alimentation en eau potable ou à des fins agricoles sont utilisés pour répondre à ces problématiques : la nappe du bassin du Blavet en Bretagne et celle du bassin de la nappe de la Craie de Champagne sud et Centre. Ceux-ci ont permis de valider une méthodologie de conception de modèles appliquée au domaine hydrogéologique et d'estimer les incertitudes associées à partir de calculs statistiques.

Les principaux résultats de ce travail sont :

1. La validation de la méthode de conception de modèles sur différents cas d'étude.

2. La proposition d'une méthode générique de quantification des incertitudes dépendant de la situation hydrogéologique (recharge, récession), en particulier appliquée à l'initialisation des paramètres du modèle à réseaux de neurones, ainsi que d'une méthode originale d'estimation des incertitudes liées à l'hétérogénéité de la distribution spatiale des précipitations.

3. La définition d'un indice de confiance a priori permettant de contraindre l'estimation des incertitudes pour chaque situation hydrogéologique.

Mots-Clés : Réseaux neuronaux, Modélisation, Prévision, Incertitudes, Densités de probabilités, Bassins versants, Nappes souterraines

Abstract

Study of the uncertainties of neural models on hydro(geo)logical forecasting. Application to watersheds of different typologies.

Floods and droughts are two of the major risks in France. They will be accentuated by climate change. Modelling and forecasting of these phenomena is therefore an essential element for the management of water tables and rivers.

At present, water flows or heights are mainly anticipated from physical or conceptual models. Although efficient and necessary, the calibration and implementation of these models often require long and costly studies.

Concerning groundwater, the behaviour of the study basins is in most cases difficult to apprehend both because of the heterogeneity of the infiltrations and of the geology and because of the costly observation conditions as they often require drilling. Predictive hydrogeological models are therefore insufficiently used by managers due to both their cost and the lack of confidence in their outputs.

In this context, this thesis, co-directed by IMT Mines Alès and jointly financed by aQuasys and the ANRT, aims to develop models based on the systemic paradigm. Indeed, these models have the advantage of being able to be designed solely on the basis of input and output information (rainfall and flow/height), thus limiting the need for additional investigations and measurements.

The models chosen are neural networks because of their ability to identify non-linear and dynamic phenomena. However, the uncertainties associated with the outputs of these models are rarely quantified and this does not allow managers to take decisions in a reasoned manner. This thesis therefore seeks to address three main objectives:

1. Generalization of a model design method adapted to different variables (surface water flows/height) and to very different types of basins: watersheds or hydrogeological basins (groundwater height).

2. Evaluation of the uncertainties linked to these models, in particular for the impact of the initialization of the parameters of the neural network model, as well as that of the heterogeneity of the recharge rainfall.

3. Controlling these uncertainties.

Several basins exploited for drinking water supply or agricultural purposes are used to address these issues: the Blavet basin in Brittany and the Chalk basin in southern and central Champagne. These have enabled us to validate a model design methodology applied to the hydrogeological domain and to estimate the associated uncertainties based on statistical calculations.

The main outcomes of this work are :

1. Validation of the model design method on different study cases.

2. The proposal of a generic method for quantifying uncertainties depending on the hydrogeological situation (recharge, recession), in particular applied to the initialization of the parameters of the neural network model, as well as an original method for estimating the uncertainties linked to the heterogeneity of the spatial distribution of precipitation.

3. The definition of an a priori confidence index to constrain the estimation of uncertainties for each hydrogeological situation.

Keywords : Neural networks, Modelling, Forecasting, Uncertainties, Probability densities, Catchments, Groundwater.

Préambule

Les travaux qui suivent sont réalisés dans le cadre d'une thèse CIFRE. Ce type de thèse vise à renforcer l'insertion professionnelle des futurs docteurs. Pour cela, une collaboration entre la société aQuasys et l'équipe ERT d'IMT Mines Alès, rattachée à l'UMR HydroSciences a été initiée. Cette collaboration entre entreprise et école d'ingénieur est encadrée par ARMINES. Ces travaux sont également inscrits dans l'école doctorale GAÏA, qui encadre et s'assure de la faisabilité du projet. L'ANRT (Association Nationale de la Recherche et de la Technologie) est une association qui vise à promouvoir les collaborations de recherche public-privé en contribuant à financer des thèses.

L'entreprise aQuasys

L'entreprise aQuasys est une PME de onze salariés à temps plein, dont le siège social est localisé à Port-Saint-Père, à une vingtaine de kilomètres de Nantes. Cette entreprise est spécialisée dans la conception de progiciels d'aide à la gestion de la ressource en eau. La mission principale de cette structure est de permettre aux acteurs territoriaux et aux bureaux d'études environnementaux de valoriser leurs données environnementales pour qu'ils puissent anticiper au mieux les risques hydrologiques.

Fondée en 2006 par M. Stéphane BARTHON à la suite d'un projet collaboratif avec la DREAL Centre sur la gestion des données piézométriques, cette entreprise prend son essor en 2011 en effectuant son premier recrutement. L'année 2016 marque une nouvelle évolution au sein de l'entreprise, de par la labellisation SANDRE du progiciel SIeau v.6.3[1] et de l'évolution du personnel et de la hiérarchie, notamment de par l'arrivée de M. Dominique THERIEZ, en tant que directeur associé. En 2019, M. Stéphane GRENON intègre la société en tant que directeur commercial.

Aujourd'hui, aQuasys dispose d'un panel de quatre progiciels d'aide à la décision, de bancarisation et de valorisation des données environnementales, ainsi que des solutions web de récupération de données *open data* issues des banques de données nationales et d'un progiciel de modélisation hydrogéologique. Ses principaux clients sont issus de services publics (DREAL, Conseils départementaux, …), de syndicats mixtes, de bureaux d'études environnementaux, d'acteurs agricoles ou encore d'industries agro-alimentaires.

Le Centre CREER d'IMT Mines Alès

IMT Mines Alès est une école d'ingénieur appartenant à l'Institut Mines Télécom, basée à Alès dans le Gard. Fondée en 1843 par ordonnance du Roi Louis-Philippe, l'école des Mines d'Alès

[1] Voir le document officiel de la labellisation SANDRE.

(ancien nom d'IMT Mines Alès) se focalise initialement sur des travaux miniers, notamment après la découverte de mines de charbon à proximité d'Alès. Dans les années 1960, cette école se diversifie et s'oriente vers plusieurs secteurs d'activités, dont le secteur de l'environnement. Cette diversification fait suite à la baisse des activités minières en France et plus particulièrement autour de la ville d'Alès. Vingt ans plus tard, l'école d'ingénieur s'étend en implantant une annexe à Pau et une seconde à Nîmes (qui finira par être rapatriée à Alès en 2017). En 2012, l'école est rattachée à l'Institut Mines Télécom et intègre entièrement cet institut en 2017. L'école change alors de nom et devient IMT Mines Alès. Elle est co-accréditée au sein de l'école doctorale GAIA, et trois autres écoles doctorales pour délivrer le diplôme de docteur.

Aujourd'hui, IMT Mines Alès représente près de 1 400 élèves dont 1 150 élèves ingénieurs (Bac + 5), 150 en Master et mastère spécialisé (Bac + 5 et Bac + 6) et 100 élèves chercheurs (Bac + 8). Elle dispose également de trois centres de recherche spécialisés dans divers domaines d'activités (matériaux, environnement, informatique).

L'un d'entre eux est le CREER (Centre de Recherche et d'Enseignement en Environnement et en Risques). Ce Centre, créé en 1974, se focalise sur la gestion des ressources naturelles et des risques qui en découlent. Le Centre est dirigé par Mme Anne JOHANNET.

Le Centre CREER représente en 2021 24 enseignants-chercheurs, 16 personnes en soutien et développement, 5 post-doctorants et 27 doctorants, générant cette même année un total de 93 publications scientifiques.

Deux équipes de recherche contribuent aux travaux du Centre. La première équipe EUREQUA (EtUdes des RisquEs de la QUalité de l'Air) étudie les situations de risques et les problématiques liés à l'air. La seconde équipe, dans lequel s'inscrit le cadre de cette thèse, est l'équipe ERT (Eau, Ressources et Territoire). Cette équipe s'intéresse à l'eau continentale comme facteur fondamental de développement des sociétés. L'équipe aborde, comme thème principal, l'eau : ressource ou danger pour les sociétés humaines. Au 1er Janvier 2021, elle intègre l'UMR Hydrosciences Montpellier.

Les missions principales du laboratoire sont de faire progresser la science et ses applications, tout en répondant aux besoins communs de former des élèves à haut niveau, grâce à l'expertise des enseignants-chercheurs et encadrants, et de transférer au mieux l'expertise vers d'autres organismes et entreprises.

Introduction

L'eau est une ressource centrale pour les sociétés, si bien que sa disponibilité en quantité et en qualité suffisantes constitue une problématique majeure, dont la complexité est aggravée par le changement climatique. L'alimentation en eau potable, l'irrigation, et l'alimentation en eau des industries sont des enjeux particulièrement sensibles aux sécheresses. Si, en France, la consommation d'eau potable domestique par habitant a amorcé une baisse depuis 2004 après des années d'augmentation (SDES 2020), elle représentait encore 3,7 milliards de mètres cube en 2012, soit 145 litres par jour et par habitant, elle ne représente qu'une petite part de la consommation générale. Ce chiffre ne représente en effet que 17% des prélèvements sur le territoire français et 7,5% de l'empreinte incluant la consommation liée aux importations soit respectivement 519 et 650 mètres cube par an et par habitant (SDES 2020). L'eau potable représente une fraction importante de ces prélèvements. Ainsi, 32 milliards de m^3 d'eau ont été prélevés en 2017 pour l'eau potable, soit l'équivalent à 23 jours de pluies en France métropolitaine (SDES 2020). L'état quantitatif des ressources en eau de surface ou souterraines est très dépendant du climat à court et long terme, celui-ci étant lui-même associé à des variations de la consommation. Les ressources sont en effet très affectées par leur recharge issue des précipitations et par les prélèvements (naturels ou artificiels). Il s'agit là d'un équilibre à maintenir pour garantir la sécurité de l'approvisionnement.

A l'opposé, les inondations concernent un français sur quatre et causent en moyenne 650 à 800 millions d'euros de dégâts par an en France (Carré et Deutsch 2015). Plus de 300 millions d'euros par an sont dépensés par l'État et les collectivités locales pour le développement de programmes et plans d'actions pour la défense contre les inondations (Laimé 2015). Les exemples récents montrent que le sud-est du pays est particulièrement exposé (Alpes Maritimes, 2020 ; Gard et Lozère, 2020 ; Aude, 2018 ; Alpes Maritimes, 2015 ; etc.) mais le reste du territoire, avec une fréquence et une intensité plus faibles, peut aussi connaître des épisodes importants (Seine et affluents, 2016 ; Centre-ouest, 2010 ; Loire, 2008 ; Somme, 2001 ; etc.).

Ces deux phénomènes pourraient voir leur fréquence et leur intensité augmenter en contexte de changement climatique (Hartmann, et al. 2013), ce qui va accroître les besoins de gestion rapprochée des hydrosystèmes. Cette gestion passe par l'anticipation et nécessite donc le concours de modèles hydrologiques ou hydrogéologiques, bien souvent couplés à des modèles météorologiques. La démarche de modélisation n'est jamais aisée puisqu'elle nécessite une connaissance extensive des hydrosystèmes ainsi que des données variées et relativement nombreuses. Cela représente un coût financier et temporel qu'il est difficile d'assumer pour les gestionnaires. Le fonctionnement des bassins concernés n'est en effet souvent connu que partiellement et la qualité et la représentativité des mesures ne sont pas toujours satisfaisantes.

Pour effectuer une modélisation hydrologique ou hydrogéologique, il est possible d'utiliser des modèles à base physique. Si ces derniers ont l'avantage de pouvoir fournir des indications sur le fonctionnement du système en sus des sorties modélisées, ils nécessitent toutefois une bonne connaissance du système étudié, ce qui est parfois difficile compte tenu du caractère non linéaire, non stationnaire, hétérogène et anisotrope du fonctionnement des bassins versants. Par ailleurs, pour fournir des informations prédictives, ces modèles doivent être couplés avec des prévisions de leurs entrées, qui souffrent souvent d'un manque de fiabilité aux échelles de temps et d'espace nécessaires.

Une alternative intéressante, issue du paradigme systémique et récemment mise en lumière dans de nombreux travaux de recherche sous le terme « *machine learning* », utilise d'autres modèles, comme les réseaux de neurones, qui ne nécessitent pas de connaissances physiques approfondies sur le fonctionnement du système, et peuvent relier les entrées aux sorties par un modèle statistique. Ces modèles, de type boite noire, qui ont ainsi l'avantage de ne pas requérir de connaissance extensive du système sont donc tout particulièrement pertinents lorsqu'ils sont appliqués au champ des eaux souterraines, dont les processus sont particulièrement mal connus de par le manque de connaissances selon différentes échelles spatiales. De plus, ces modèles ne nécessitent pas de prévision de leurs entrées pour produire des prévisions, dans la limite d'un horizon proche du temps de réponse de l'hydrosystème. En revanche, du fait de leur caractère statistique, ils ont besoin de chroniques de données significatives, de qualité et de durée suffisante.

Quel que soit le type de modèle, les résultats fournis ne sont pas toujours assortis d'informations concernant leurs incertitudes, ne laissant que peu d'indications sur les erreurs pouvant être engendrées par le modèle, et limitant ainsi leur caractère opérationnel pour l'utilisateur final.

Dans ce contexte, cette thèse, co-financée par la société aQuasys et l'ANRT dans le cadre d'une convention CIFRE, en collaboration avec IMT Mines d'Alès, a pour objectif de développer des modèles à réseaux de neurones pour prévoir la disponibilité des eaux souterraines et avec pour objectif l'implémentation de méthodes d'évaluation et de représentation des incertitudes.

Les travaux menés dans le cadre de cette thèse s'intéressent aux eaux souterraines et ont pour but de répondre à trois objectifs principaux :

1. Déterminer l'origine potentielle des principales incertitudes, puis les réduire

2. Définir une méthode de conception de modèle adaptable aux différents types de bassins avec leurs différents types de variables

3. Évaluer et représenter les incertitudes liées à ces modèles en fonction des types de bassins visés.

Ce mémoire présente dans un premier temps le contexte hydrogéologique actuel. Dans un second temps, les outils et méthodes, incluant les méthodes de caractérisation d'un bassin ainsi que le fonctionnement des réseaux de neurones, sont présentés. La troisième partie est consacrée à la présentation des incertitudes et des densités de probabilités utilisées ainsi que de la méthodologie appliquée pour le calcul des incertitudes sur les sorties du modèle. Les deux bassins qui ont fait l'objet des travaux de modélisation et de calculs d'incertitudes : le bassin de la nappe du Blavet et le bassin de la nappe de la Craie de Champagne sud et centre sont ensuite présentés. Les résultats obtenus sur ces bassins ciblés sont ensuite décrits dans la dernière partie.

Chapitre I : La ressource en eau et sa gestion par la prévision hydrogéologique

I. Introduction

L'eau est essentielle et est à l'origine de toute vie. Cette petite molécule, composée seulement de deux atomes d'hydrogène et d'un atome d'oxygène, recouvre plus de 70% de la surface de notre planète et son volume total est estimé à 1,338 milliards de km^3. Cependant, seulement 2,5 % de cette eau est de l'eau douce, dont les deux-tiers sont contenus dans les glaciers (Antarctique et Groenland principalement). Ainsi, moins de 0,77 % des eaux sont douces et potentiellement utilisables pour l'homme (cours d'eau, plans d'eau et masses d'eau souterraines) (Eakins et Sharman 2010).

C'est au travers de ce contraste entre abondance et insuffisance que la ressource en eau est abordée dans cette partie, en établissant un état des lieux de la ressource en eau, puis en présentant les contextes démographiques et climatiques influant sa disponibilité et son usage.

Nous verrons ainsi que la démographie, et les activités anthropiques et enjeux qu'elle induit, et le changement climatique peuvent avoir une influence très variable sur la ressource en eau. Ces éléments nous permettrons de définir une problématique vis-à-vis de la gestion de la ressource en eau et de la modélisation hydrogéologique.

II. La ressource en eau, entre abondance et insuffisance

II.1 Etat des lieux de la ressource en eau

La ressource en eau est à la fois omniprésente et rare sur Terre. En effet, l'eau douce, *a priori* exploitable pour la consommation humaine, est peu abondante par rapport aux eaux salées des mers et océans. L'eau est présente sur Terre sous tous ses états (solide, liquide et gazeux) et chacun de ses états fait partie intégrante du cycle de l'eau, qui illustre comment circule et se renouvelle l'eau, si importante pour l'homme et les écosystèmes (Figure I-1).

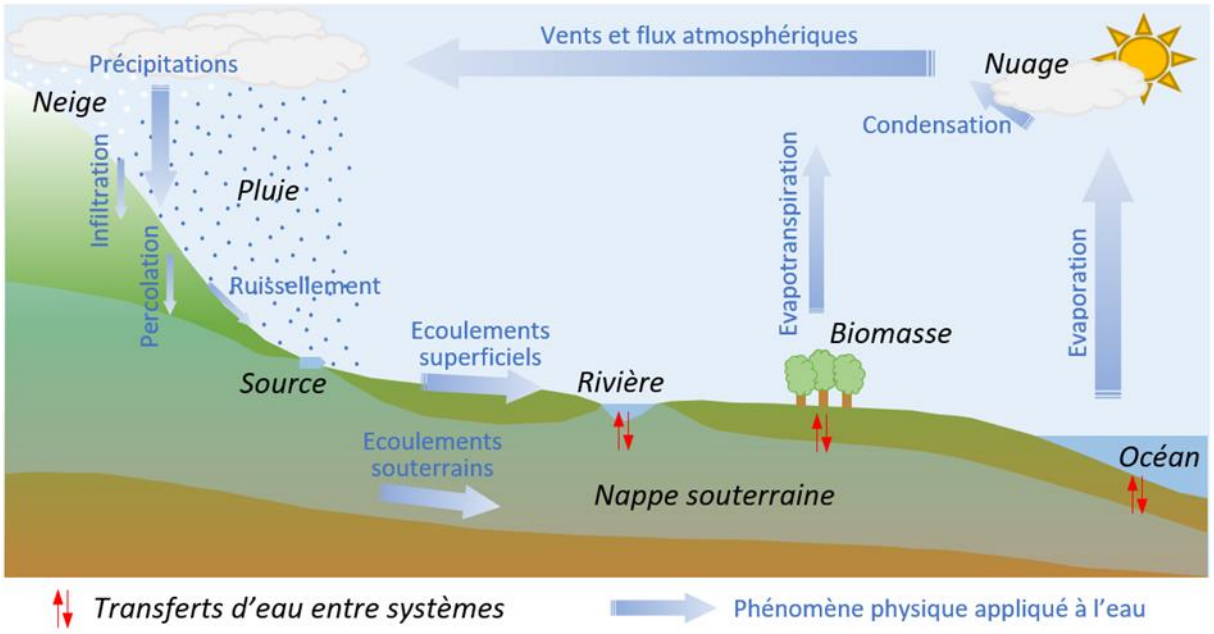


Figure I-1 - Le cycle de l'eau.

Il est ainsi possible de comprendre les phénomènes liés à la ressource en eau. Par exemple, des ruissellements dus à une forte pluie et une nappe souterraine en hautes eaux peuvent engendrer des inondations. A l'inverse, de fortes chaleurs couplées à l'absence de précipitations peuvent provoquer une forte évapotranspiration, asséchant ainsi les sols. Ce phénomène de sécheresse peut être aggravé par les prélèvements destinés à l'agriculture, aux industries ou à l'alimentation en eau potable.

En effet, le cycle de l'eau peut être perturbé par les activités anthropiques, qui fragilisent directement la ressource en eau, de par les prélèvements en eau ou les dérivations, l'expansion des zones artificialisées qui accroît l'imperméabilisation des sols, les pollutions, ou encore les ouvrages hydroélectriques. Toutefois, ces activités, même si elles pourraient être optimisées, sont la plupart du temps nécessaires pour l'homme.

Les eaux douces pour la consommation humaine sont majoritairement prélevées de deux manières : par forage pour un prélèvement souterrain (majoritaires en France avec 65 % des volumes prélevés (BNPE 2020)), ou directement en surface (rivière ou lac, naturel ou artificiel). Dans certains pays côtiers, il existe des usines de dessalement qui permettent de transformer les eaux de mer ou saumâtres en eau douce. Toutefois, cette technique s'avère couteuse et difficile à mettre en place et est surtout développée dans des pays où la ressource en eau est rare sur leur territoire, tel que Singapour, Israël ou encore le Qatar. En dépit de ces techniques, seulement 89 % de la population mondiale dispose d'un accès à l'eau potable (Observatoire des Inégalités, 2018).

Malgré l'existence de ces techniques de potabilisation de l'eau, la ressource en eau sera de moins en moins disponible, notamment à cause de l'explosion démographique, qui, d'une

part, accroît les risques de pollution, et d'autre part, appauvrit de plus en plus les nappes souterraines. Les enjeux climatiques sont aussi impliqués dans la moindre disponibilité de l'eau attendue dans le futur et il en sera question dans les prochaines sections.

II.2 Contexte démographique : une consommation accrue

La population mondiale ne cesse de croître, en passant de 2,5 milliards d'habitants dans les années 1950 à 7,7 milliards d'habitants en 2020. Chaque année, 83 millions d'humains supplémentaires sont dénombrés (De Plazaola, et al. 2019). Cette croissance démographique devrait se poursuivre, d'après les prévisions de l'ONU, atteignant ainsi 9,7 milliards d'habitants en 2050 et 10,9 milliards à l'orée du XXII^{ème} siècle (ONU 2019). Cependant, cette tendance est peut-être légèrement surévaluée. En effet, une étude du *Institute for Health Metrics and Evaluation* a utilisé un modèle prédictif ARIMA pour estimer la population mondiale et a montré que la population mondiale pourrait légèrement décroître à partir de 2064 pour atteindre au final 8,79 milliards d'humains en 2100 (Vollset, et al. 2020).

En France, la croissance démographique est plus faible que dans le reste du monde, compte-tenu du nombre de naissances qui diminue chaque année. Elle reste néanmoins une des plus élevées d'Europe, avec un solde naturel annuel supérieur à 200 000 depuis la fin de la seconde guerre mondiale et jusqu'en 2015. Au-delà de 2015, ce solde diminue, passant de 190 000 en 2016 à 140 000 en 2019 (INSEE 2020). Ainsi, 67,1 millions de Français recensés début 2020 se partagent la ressource en eau, contre 55,6 millions en 1982 (INSEE 2020).

Par ailleurs, la croissance démographique évolue conjointement à la proportion des surfaces artificialisées, passant de 8,3 % en 2006 à 9,4 % du territoire français en 2015 (De Plazaola, et al. 2019). Ces surfaces empêchent l'infiltration des eaux et tendent donc à limiter la recharge.

Pourtant, en France, les prélèvements d'eau potable sont en diminution, avec 5,4 milliards de mètres cube d'eau prélevés en 2011, soit en moyenne 400 millions de mètres cube de moins qu'au cours de la période 2000-2007 (Da Costa, et al. 2015). Ces prélèvements sont encore prévus à la baisse, avec une diminution estimée entre 7% et 12% en 2070, principalement causée par la diminution de la consommation en eau par foyer entre 25 et 29% et une diminution des fuites sur les réseaux de distribution d'eau de 11% (Da Costa, et al. 2015).

De plus, la consommation en eau par habitant est, elle aussi, en diminution. En effet, après être passée de 106 litres par habitant et par jour en 1975 à 161 litres par habitant et par jour en 1991 (Poquet et Maresca 2006), la consommation en eau journalière par habitant suit la même évolution que celle des prélèvements en eau potable, n'atteignant plus que 148 litres d'eau par habitant et par jour en 2015 (Da Costa, et al. 2015).

Pourtant, la consommation en eau journalière ne cesse d'augmenter, principalement à cause de la croissance démographique, avec une consommation en eau moyenne de 9,83 millions de mètres cube par jour en 2015 sur l'ensemble du territoire français contre 9,38 millions de mètres cube par jour en 1991[2].

Cette consommation en eau n'intervient pas uniquement sur l'eau potable mais aussi sur d'autres enjeux tels que l'agriculture ou l'industrie. En effet, l'accroissement de la population génère une demande en nourriture plus importante. Cela implique des pratiques agriculturales plus intensives et des surfaces agricoles plus importantes (cultures et élevages) nécessitant plus d'eau.

Ainsi, la disponibilité de la ressource en eau est réduite par la consommation liée à la pression démographique en France et surtout à échelle mondiale depuis 200 ans. Cette ressource est par ailleurs tout autant fragilisée par d'autres phénomènes.

II.3 Contexte climatique : des évènements extrêmes en augmentation

La croissance démographique va de pair avec l'avènement de l'ère industrielle au milieu du XIX^{ème} siècle. Cette dernière a engendré au fils des années la naissance de nombreux outils, machines, techniques, exploitations ou encore véhicules, améliorant certes la vie quotidienne des humains mais souvent au détriment de l'environnement. Les émissions de gaz à effet de serre et les autres pollutions générées par ces progrès techniques ont provoqué un dérèglement climatique, observable de nos jours, mais qui est la conséquence des activités anthropiques depuis maintenant près de 200 ans.

Pour illustrer ce propos, les rejets annuels mondiaux en CO_2, qui est le principal gaz à effet de serre lié aux activités anthropiques, ont augmenté de 17 gigatonnes (Gt), passant de 21 Gt en 1970 à 38 Gt en 2004 (Bernstein, et al. 2007).

Les rejets de gaz à effet de serre ont une incidence directe sur le climat. En effet, une étude menée par le Groupe d'experts Intergouvernemental sur l'Evolution du Climat (GIEC) a montré une hausse des températures visible à l'échelle mondiale, avec une augmentation moyenne de 0,85 °C sur la période 1880-2012, et une augmentation de 0,72°C entre 1951 et 2012 (Hartmann, et al. 2013). Il est donc possible de constater que la hausse des températures s'est accélérée au cours des soixante dernières années. Il est à noter également que cette hausse est inégalement répartie sur l'ensemble du globe, avec une hausse plus importante sur les massifs proches du pôle Nord (Bernstein, et al. 2007). En France, les températures suivent la

[2] Résultats obtenus à partir du produit entre la consommation en eau d'un habitant par jour et du nombre d'habitants en France pour les années 1991 et 2015.

même évolution, avec une augmentation estimée à plus de 1,0°C au cours du XX^{ème} siècle, supérieure à l'augmentation constatée à échelle mondiale (Météo-France 2015).

Cette hausse de températures moyennes engendre de fait une hausse des températures en été, augmentant ainsi le nombre de vagues de chaleurs et d'épisodes caniculaires (Hartmann, et al. 2013). De même, le nombre de jours et de nuits frais ont diminué tandis que le nombre de jours et de nuits chauds ont augmenté depuis 1950 (Hartmann, et al. 2013).

Cette hausse des températures intensifie le phénomène d'évapotranspiration, augmentant ainsi l'humidité spécifique de l'air depuis les années 1970 (Hartmann, et al. 2013).

Le réchauffement climatique engendre d'autres phénomènes extrêmes. En effet, il a été montré, avec toutefois un indice de confiance faible, que les épisodes de sécheresse ont augmenté depuis 1970, plus particulièrement sur le pourtour méditerranéen. De plus, l'intensité de ces épisodes de sécheresse s'est accrue (Hartmann, et al. 2013). Cependant, Peña-Angulo *et al.* ont observé que les tendances à la baisse des précipitations en Europe du Sud-Ouest (qui comprend le Portugal, l'Espagne, le sud de la France et de la Suisse, l'Italie, la Slovénie et la Croatie) sur la période 1850-2018 sont visibles à l'échelle interannuelle et interdécadaire. En revanche, il n'est pas possible de confirmer ces mêmes tendances sur des échelles de temps plus larges. Ainsi, il est difficilement possible d'associer les périodes de sécheresses et le forçage anthropogénique (Peña-Angulo, et al. 2020). Les auteurs de cet article nous conseillent d'étudier et de mieux comprendre les phénomènes thermodynamiques responsables de la variabilité des précipitations (Peña-Angulo, et al. 2020). L'étude de Vicente-Serrano *et al.* décrit des conclusions similaires sur l'Europe de l'Ouest pour la même période (Vicente-Serrano, et al. 2020). Ces constats s'accordent avec ceux réalisés sur les évolutions au XXe siècle sur l'humidité des sols en Europe (van der Schrier, et al. 2006), qui, notamment sur le pourtour méditerranéen, caractérisent les évolutions de sècheresse (Vicente-Serrano, et al. 2020).

A contrario, le nombre d'épisodes pluvieux importants et leurs intensités ont sensiblement augmenté dans de nombreuses régions depuis 1950, notamment en Europe et en Amérique du Nord (Hartmann, et al. 2013). Les intensités de pluies lors d'évènements extrêmes (notamment pour des précipitations d'au-delà de 200 mm/j) ont en moyenne augmentés de 22% sur le pourtour méditerranéen sur la période 1961-2015. L'hypothèse d'une cause anthropique à ce phénomène est privilégiée (Ribes, et al. 2019).

Rajczak et Schär ont modélisé les évènements pluvieux extrêmes à partir de *scenarii* d'émissions de gaz à effet de serre et de données historiques de précipitations sur le continent européen. Ils ont montré que ces phénomènes pluvieux extrêmes vont s'intensifier au cours des 50 prochaines années, particulièrement en hiver, avec toutefois des incertitudes non-négligeables (Rajczak et Schär 2017). Colmet-Daage et *al.* parviennent à des conclusions

similaires sur les bassins du Lez, de l'Aude (France) et de la Muga (Espagne) à partir des *scenarii* climatiques et des modèles SAFRAN (Colmet-Daage, et al. 2017).

Ainsi, le réchauffement climatique a un impact sur la ressource en eau en modifiant les équilibres du bilan hydrique. La tendance pour les prochaines années et décennies n'est pas à l'amélioration de ces conditions, avec une dégradation probable de la ressource en eau, d'un point de vue qualitatif comme quantitatif.

III. La connaissance pour une meilleure gestion

A une large échelle, la connaissance de la ressource en eau permet d'appréhender et de comprendre les différents processus hydrologiques et hydrogéologiques. Elle permet aussi d'observer les variations de chacun de ces processus dans le temps. De plus, une observation à grande échelle permet de mieux appréhender la dynamique et les non stationnarités de bassins versants dont le comportement est souvent considéré comme stationnaire (Kingston, et al. 2020).

III.1 Quantification de la ressource en eau

La connaissance de la ressource en eau repose principalement sur des observations, qu'elles soient climatiques, hydrologiques, géothermiques, pédologiques ou encore géologiques. Ces observations peuvent être quantitatives (mesures in situ, ponctuelles ou permanentes) ou qualitatives.

Parmi ces observations, la mise en place de dispositifs de mesures permet de récolter des données datées sur différents phénomènes. Ces réseaux de mesures sont un ensemble de stations de mesures, quantifiant ainsi un phénomène hydrogéologique, hydrologique ou climatique particulier. Ainsi, les historiques de données récoltés permettent d'avoir une première connaissance sur le fonctionnement d'un bassin. Cependant, la représentativité des mesures n'est pas nécessairement assurée, en particulier lorsque l'hydrosystème est hétérogène, lorsque l'historique disponible est bref ou encore lorsque la densité de stations est faible.

III.1.1. Eau de surface

La quantification de la ressource en eaux de surface consiste le plus souvent à évaluer le débit d'un cours d'eau. En effet, le débit permet de calculer un bilan hydrique lorsqu'il est mis en relation avec les conditions météorologiques et de caractériser un régime hydrologique.

Pour estimer le débit en une section d'un cours d'eau, celle-ci doit d'abord être mesurée. La vitesse de l'eau est ensuite estimée au travers de cette section, ce qui représente la difficulté la plus importante. La section peut parfois évoluer, notamment lors de crues importantes ou

à l'occasion de modifications anthropiques, rendant ainsi l'estimation de la vitesse difficile, notamment en hautes eaux.

Pour ce faire, dès 1910, Bureau appliqua la méthode du jaugeage par flotteur. Un fil, plus long que la largeur du cours d'eau, muni de flotteurs, est posé sur la rivière. Le fil, flottant sur l'eau, est entraîné par le courant. Les vitesses du courant en ces points sont ensuite estimées, puis la vitesse moyenne du cours d'eau est déterminée (Bureau 1910).

Une autre méthode consiste à utiliser un moulinet hydrométrique. Également appelé courantomètre, ce moulinet est composé d'un rotor, sur lequel sont disposées des hélices. Les forces engendrées par le courant du cours d'eau font tourner l'hélice. La vitesse de rotation de celle-ci détermine ainsi la vitesse du cours d'eau, permettant d'estimer le débit. Cette technique peut s'appliquer sur plusieurs portions de la section et à plusieurs profondeurs. Diconne met au point une méthodologie similaire où toutes les vitesses de chaque profondeur sont mesurées en une fois. Cette méthode de mesure, dite « par intégration », permet au moulinet de se déplacer de manière constante sur une verticale et de mesurer la vitesse du courant en tout point de celle-ci. Elle s'oppose à la méthodologie dite « par points », où la vitesse du courant est mesurée sur plusieurs points fixes d'une même verticale (Diconne 1972). L'avantage de la mesure de vitesse par intégration est qu'elle produit des résultats plus homogènes lorsque les vitesses d'une même verticale sont mesurées. De plus, cette technique est plus rapide à exécuter que celle « par points ». Néanmoins, elle n'est pas adaptée pour des rivières ayant des hauteurs d'eau faibles (Diconne 1972).

Des méthodes plus récentes viennent compléter le panel de méthodologies de mesure de la vitesse ou du débit. Les limnimètres sont des échelles permettant de mesurer la hauteur d'eau en surface. Bien qu'anciens, ils peuvent être couplés avec des caméras qui photographient ces limnimètres, permettant de lire et de mesurer le niveau de l'eau sans intervention manuelle (hors maintenance) (Gilmore, Birgand et Chapman 2013). Cette méthode génère toutefois des incertitudes sur la mesure du niveau de l'eau et de sa détection en raison de la qualité des images, les conditions environnementales et les interprétations informatiques des images (Gilmore, Birgand et Chapman 2013).

Certaines caméras sont également capables de mesurer la vitesse en surface d'un cours d'eau (Daigle, et al. 2013), (Tauro, et al. 2016). Cette technique nécessite des traceurs afin de les calculs soient possibles. Toutefois, les divers débris et autres végétaux peuvent aider à la détermination de la vitesse du cours d'eau. Cette méthodologie nécessite trois étapes, la première étant l'enregistrement des images pour laquelle un minimum de 30 images par seconde est requis. La deuxième étape consiste en un prétraitement des images (étalonnage des pixels, correction et correspondance des images). Cette étape permet de limiter toute distorsion des images enregistrées. La troisième étape correspond à l'analyse des images, par

le biais de calculs de corrélations croisées entre chaque image, permettant la détermination de la vitesse du cours d'eau (Daigle, et al. 2013), (Tauro, et al. 2016).

La mesure des écoulements en surface est donc relativement aisée dans la plupart des circonstances. Elle peut présenter des incertitudes que différentes méthodes tendent à réduire, en particulier en hautes eaux. Ceci étant, en termes de ressources, ces incertitudes ne sont pas de nature à fondamentalement modifier les équilibres du bilan hydrique.

III.1.2. Eau souterraine

Pour caractériser un aquifère, évaluer et démontrer son volume de ressource et ses conditions de recharge, on pratique une étude géologique et structurale, permettant notamment de déterminer les conditions de recharge potentielles et le lien avec les eaux de surface et les autres aquifères. L'étude des mesures piézométriques disponibles ainsi que la caractérisation hydraulique des forages existants permet alors de produire des cartes piézométriques qui renseignent sur la géométrie du toit de l'aquifère et sur son fonctionnement (isopièzes et lignes de courant). Des diagraphies, traçages et essais de pompage peuvent compléter ces éléments dans une phase de démonstration. Ils vont alors permettre de déterminer la plupart des caractéristiques hydrauliques de l'aquifère.

Au terme de ces études, la ressource en eau est la plupart du temps quantifiable et l'on peut y associer un volume de prélèvement potentiel qui permette d'assurer le renouvellement de la recharge. Ce volume doit tenir compte de la saisonnalité ainsi que d'éventuelles non stationnarités, notamment concernant l'occupation des sols, l'influence anthropique sur les écoulements de surface ou encore le climat.

III.1.3. Règlementations

Il existe diverses règlementations qui permettent la préservation de la ressource en eau. Par exemple, des arrêtés préfectoraux peuvent être énoncés, selon les bilans hydrogéologiques, interdisant ou limitant le prélèvement en eau pour les usagers. Chaque arrêté est pris au niveau d'une masse d'eau superficielle ou souterraine, avec quatre niveaux de vigilance. Ces arrêtés sont disponibles sur le site Propluvia. Ces arrêtés permettent ainsi de réagir localement et de manière réactive sur la gestion de la ressource en eau.

En France, chaque prélèvement en eau, ainsi que son ouvrage dédié, doivent *a minima* être déclarés, conformément à l'arrêté ministériel du 11 septembre 2003 et à l'article R214-1 du code de l'environnement. Cet arrêté permet ainsi de recenser au niveau national ces prélèvements en eau.

Au niveau européen, outre les aspects qualitatifs, la Directive Cadre sur l'Eau (DCE) fixe des objectifs sur les aspects quantitatifs par masse d'eau. Enoncée en 2000, cette directive permet

ainsi l'évaluation de l'état chimique, biologique, hydromorphologique et quantitatif et impose l'adoption de plans de gestion pour chaque masse d'eau afin d'atteindre un bon état des masses d'eau. L'atteinte du bon état général des masses d'eau a été demandé pour 2015, avec des reports possibles en 2021 et 2027. Ces directives permettent ainsi d'avoir une politique globale et commune et peuvent être plus strictes sur des échelles spatiales inférieures.

Toutes ces règlementations visent au maintien ou à l'atteinte d'un bon état hydrogéologique pour chaque masse d'eau. Néanmoins, elles ne tiennent pas compte de la vulnérabilité de ces masses d'eau face aux pressions anthropiques ou forçages climatiques.

III.2 Vulnérabilité aux extrêmes climatiques

La vulnérabilité d'un territoire correspond aux effets provoqués sur des enjeux (populations, industries, transports, agriculture, …) par un phénomène donné (Bourguignon 2014). L'aléa correspond à la survenance du phénomène en question qui peut se caractériser par une intensité et une probabilité d'occurrence ou une période de retour. Le risque résulte quant à lui d'un croisement spatial entre un aléa et des enjeux caractérisés par leur vulnérabilité. Le risque exprime alors le niveau d'exposition d'un enjeu donné à un aléa donné.

Les extrêmes climatiques influencent la ressource en eau. Ceux-ci peuvent générer des risques hydrologiques importants. Les deux principaux sont le risque d'inondation et le risque de sécheresse. Suivant la vulnérabilité des territoires et l'intensité des aléas, ces risques peuvent être élevés.

III.2.1. Sécheresse

La sécheresse a des conséquences directes sur l'agriculture, en limitant la production agricole voire en détruisant les cultures. En effet, les sécheresses causent en moyenne en France 110 millions d'euros de dégâts par an depuis 1989 pour le secteur agricole. A ce montant, il est possible d'ajouter 3 millions d'euros par an pour le secteur de la pêche (Spinoni, et al. 2016).

La sécheresse engendre également des difficultés d'approvisionnement en eau potable. Cela amène souvent à des restrictions d'eau pour les usagers. En France, comme énoncé dans la partie précédente, ces restrictions sont prises par le Préfet et sont accessibles sur le site Propluvia. En cas de rupture d'approvisionnement, les autorités peuvent être amenées à faire parvenir des rations d'eau aux habitants, comme en Finlande en 2003 (Spinoni, et al. 2016).

Outre les impacts directs de la sécheresse sur des enjeux tels que l'agriculture ou l'approvisionnement en eau potable, des feux de forêts peuvent également se déclarer, aggravant ainsi les dommages causés par la sècheresse, comme en Espagne entre 1994 et 1995 où plus de 28 800 ha de forêts ont brûlé (Spinoni, et al. 2016). En France, 18 800 ha de forêts se sont embrasés en juillet 2003 dans le Massif des Maures dans le Var (SOeS 2011).

Le tourisme peut également souffrir de la sécheresse. Les activités de baignade ou de montagne peuvent ne plus être accessibles à cause de la diminution des niveaux d'eau ou de l'absence de neige.

La sécheresse a aussi un impact important sur la production en électricité. En effet, un déficit en eau causé par la sécheresse engendre une baisse de production en électricité par les centrales hydroélectriques et des difficultés de refroidissement des installations nucléaires. D'autres secteurs économiques sont également touchés comme l'industrie ou la navigation commerciale. Des exemples pour les différents enjeux associés à la sécheresse sont détaillés sur le Tableau III-1.

Tableau III-1 - Liste non exhaustive d'évènements associés au risque de sécheresse en Europe. Bilan économique et social.

Lieu	Année	Enjeu	Coûts (millions d'€)	Victimes	Références
France	2003	Agriculture	550	-	(Spinoni, et al. 2016)
Finlande	2003	Eau potable	5,5	-	(Spinoni, et al. 2016)
Espagne	1990-95	Energie	210	-	(Spinoni, et al. 2016)
Mont Carmel, Israël	2010	Feux de forêts	*n.c.*	45	(Paz, et al. 2011)
Pedrógão Grande, Portugal	2017	Feux de forêts	*n.c.*	60	(Couto, et al. 2020)
Alpes, France	2006-07	Tourisme	144	-	(Spinoni, et al. 2016)
Barcelone, Espagne	2007-08	Tourisme	1 605	-	(Martin-Ortega, González-Eguino et Markandya 2012)

La sécheresse provoque également des dommages qui sont difficilement chiffrables, notamment en ce qui concerne la biodiversité et l'environnement. Nous avons vu auparavant que ce risque de sècheresse favorise l'apparition de feux de forêt. Il peut aussi provoquer une accélération de l'eutrophisation sur les lacs et cours d'eau, ainsi que la salinisation et l'acidification des eaux. Ce cas a été notamment recensé sur les lacs Alexandrina et Albert en Australie pour la période de sécheresse entre 2007 et 2010 (Li, et al. 2017). La sécheresse

engendre également une perte de la biodiversité floristique et faunistique, notamment concernant les espèces qui vivent dans un écosystème peu variant.

Nistor étudie la vulnérabilité croissante des eaux souterraines en 1990, 2020 et projetée en 2050 en Europe (Nistor 2020). Il utilise pour cela un indice qui prend en compte la disponibilité de la ressource en eau, un indice de charge de pollution, l'infiltration potentielle, la vulnérabilité de l'aquifère et la qualité de l'eau vis-à-vis des nitrates et de l'arsenic. Cette vulnérabilité de la ressource en eau est focalisée sur les risques de sécheresse et les risques de pollution.

Il constate que cette vulnérabilité est très inégalement répartie en Europe, avec des régions très vulnérables sur l'Europe Occidentale, et d'autres beaucoup moins vulnérables comme en Scandinavie. La vulnérabilité de la ressource en eau a légèrement augmenté entre 1990 et 2020 sur le Nord de l'Europe. Cette vulnérabilité s'accroit d'autant plus en 2050 sur l'ensemble du territoire européen (Nistor 2020).

Cette méthodologie a également été appliquée sur la région Grand Est pour estimer la vulnérabilité aux changements climatiques sur cette région. Il a été montré que le bassin parisien et les plaines rhénanes ont la vulnérabilité la plus forte de ce territoire. Cette vulnérabilité suit la même tendance à la hausse qu'en Europe (Haidu et Nistor 2020).

En parallèle, la disponibilité de la ressource en eau en Europe a, en moyenne, diminué entre 1990 et 2020 et cette diminution sera plus prononcée en 2050 (Nistor 2020). De même, l'humidité présente dans les sols tend à diminuer et augmente les risques de sécheresse sur l'ensemble de l'Europe (Grillakis 2019).

III.2.2. Inondation

Le réchauffement climatique n'agit pas que sur le risque de sécheresse mais aussi sur le risque d'inondation. En effet, l'accroissement des intensités des phénomènes pluvieux a une incidence sur les inondations. Toutefois, le GIEC ne dispose pas de données suffisantes pour attester d'une évolution quelconque sur ce risque. Le Groupe affirme d'ailleurs que les inondations sont aussi fortement influencées par le mode de gestion des cours d'eau, ce qui rend délicate la recherche des causes des évolutions parfois constatées (Hartmann, et al. 2013).

De nombreuses catastrophes naturelles liées aux inondations ont été observées. Le Tableau III-2 regroupe quelques évènements passés avec les coûts et les victimes causées par les inondations en France. Plusieurs types d'inondations sont répertoriées dans ce tableau.

Tableau III-2 - Liste non exhaustive d'évènements associés au risque inondations en France. Bilan économique et social.

Lieu - Cours d'eau	Date	Type d'inondations	Coûts (millions d'€)	Victimes	Références
Seine, France	Janvier 1910	Crues	1 600	*n.c.*	(DRIEE 2009)
Pyrénées Orientales, France	Octobre 1940	Episode méditerranéen Crues éclairs	*n.c.*	57*	(Boudou 2015)
Gard, France	Septembre 2002	Episode cévenole Crues éclairs	1 200	23	(Huet, et al. 2003)
Côte Atlantique, France	Février 2010	Submersion marine Tempête	1 147	59	(Migaud 2012)
Var, France	Juin 2010	Crues	816	25	(Migaud 2012, Esposito et Quemart 2012)
Loing et Seine, France	Mai-Juin 2016	Crues Remontées de nappes	1 120	5	(Delaunay, et al. 2018)
Loing, France	Janvier 2018	Crues Remontées de nappes	152	-	(Delaunay, et al. 2018)

** Environ 300 victimes supplémentaires sont à déplorer sur le territoire espagnol.*

A partir de *scenarii* climatiques inclus dans divers modèles, Alfieri *et al*. ont estimé que le risque inondation augmenterait d'ici 30 ans en Europe occidentale et centrale, que ce soit au niveau des coûts des sinistres ou de la population affectée (Alfieri, et al. 2018). Cette dernière étude ne tient pas compte de la croissance démographique, ni des enjeux associés.

Le réchauffement climatique a aussi un impact sur les inondations côtières. La hausse des températures accélère la fonte des glaciers continentaux, qui participent à l'élévation du niveau de la mer. En prenant en compte les *scenarii* climatiques les plus critiques, la fonte des glaces de l'Antarctique représenterait près de 2 mètres d'élévation de la mer en 2100 (Kopp, et al. 2017).

Face à ces projections alarmantes, la ressource en eau doit être gérée au mieux afin de limiter les impacts des aléas dont la fréquence et l'intensité sont probablement appelés à augmenter.

III.3 Prévision et gestion en temps réel

La gestion de la ressource en eau repose sur plusieurs grands axes :

- la connaissance des hydrosystèmes,
- la mesure, qui permet d'appréhender les phénomènes hydrologiques et les autres phénomènes naturels sur un bassin d'étude, et participe ainsi à la connaissance des hydrosystèmes,
- la communication, afin de sensibiliser les usagers et les décideurs sur les risques liés à cette ressource,
- la prévision, qui permet d'évaluer ces risques et leurs réalisations dans le futur,
- la prise de décision, afin de limiter ou d'anticiper ces risques lorsqu'ils se présentent.

III.3.1. Gestion en temps réel

Pour une compréhension plus fine du fonctionnement du bassin, des réseaux de mesures doivent y être implantés avec, si possible, une répartition spatiale judicieuse assurant une représentativité optimale de l'ensemble des données générées. Une représentativité satisfaisante des réseaux de mesure peut s'avérer compliquée à atteindre selon les bassins d'étude. La topographie du bassin d'étude, notamment sur des bassins montagneux, peut, par exemple, avoir des répercussions sur la représentativité spatiale des pluies. En effet, la quantité de précipitations peut différer fortement en fonction des vallées. Cela a été observé dans les Cévennes (Moussa, Chahinian et Bocquillon 2007), (Emmanuel, et al. 2017), ou encore dans les Pyrénées sur la côte basque (Lajournade et Lalanne-Berdouticq 1997).

Le réseau de mesure doit également respecter certaines contraintes. Par exemple, le couvert végétal peut perturber le bon fonctionnement de certains appareils de mesure. L'appareillage doit alors être, en règle générale, installé loin d'une végétation pouvant perturber la mesure. Il en va de même avec les points de prélèvements et ouvrages hydrauliques, modifiant localement et respectivement les niveaux d'eau souterrains et superficiels.

Afin de pouvoir prendre des décisions sur le bassin dont il a la charge de manière rapide et efficace, le gestionnaire a besoin d'un jeu de données représentatif, parfois avec une granularité temporelle très fine. On parle alors de gestion en temps réel. L'instrument de mesure est composé d'un capteur, qui quantifie le phénomène étudié, mais peut aussi contenir un afficheur numérique, un système de stockage des données ou un modem. Une fois mesurée par l'instrument de mesure, la donnée récoltée est envoyée directement par télétransmission via le modem jusqu'à un serveur, sous forme de fichiers texte le plus souvent. Ces fichiers sont ensuite déposés automatiquement (ou manuellement) sur une base de données. Il est à noter que l'envoi de ces fichiers représente une consommation énergétique

parfois significative pouvant conduire à des contraintes de mesure et de transmission de données supplémentaires.

De plus, le gestionnaire doit critiquer les données afin, par exemple d'ôter ou de corriger les valeurs qui apparaissent comme aberrantes, suite à un dysfonctionnement de l'appareil de mesure ou de la transmission des valeurs. Ces contrôles peuvent ainsi générer une latence pour la mise à disposition des données vers des banques nationales de données, mais ils sont surtout la plupart du temps impossibles à réaliser en temps réel. Cette problématique s'ajoute à celle du temps d'extraction et de mise en forme des données qui peut être important ainsi qu'à celle du coût de l'extraction pour certaines banques nationales.

L'enjeu de l'obtention de mesures en temps réel est de comprendre au plus vite l'état d'un hydrosystème et d'agir en conséquence. Cela peut être primordial dans certains cas, notamment lors de phénomènes rapides. Malgré tout, elle peut être compliquée à mettre en place car elle requiert du matériel (installations des instruments de mesure, serveurs et autres matériels informatiques) et de multiples compétences, que ce soit en environnement (étude des données récoltées, étude sur les dispositifs à mettre en place), en métrologie (mise en place et maintenance des instruments de mesure), en informatique (gestion des bases de données, administration réseaux informatiques) et parfois en génie civil (si présence d'ouvrages). Ces matériels et ce savoir-faire représentent un coût économique pour le gestionnaire.

Ces approches de mesures en temps réel ont un autre inconvénient. Si elles permettent au gestionnaire de réagir assez rapidement face à un évènement hydrologique ou climatique, elles ne permettent pas d'anticiper ces phénomènes. Pour passer du constat à l'anticipation, il est nécessaire de compléter la gestion en temps réel avec des systèmes prédictifs.

III.3.2. Approches de modélisation

Afin d'obtenir des informations prédictives sur l'hydrosystème considéré, il est courant d'utiliser des modèles hydrologiques ou hydrogéologiques. Les mesures récoltées en amont génèrent des séries temporelles qui alimentent directement ou indirectement la plupart des modèles, en temps réel comme avec l'historique.

Il existe plusieurs approches pour la modélisation hydrologique, comme hydrogéologique. La première d'entre elle est l'approche dite à base physique. Dans ce cas, le modèle utilise les relations physiques afin de reproduire un débit ou une hauteur d'eau. Ces relations reproduisent le comportement de phénomènes comme le ruissellement ou l'infiltration à partir de lois physiques connues comme la loi de Darcy pour les transferts en milieu saturé (Marchandise 2007). Dès 1969, Freeze et Harlan utilisent des équations différentielles afin de

représenter mathématiquement les processus hydrologiques et hydrogéologiques du bassin (Freeze et Harlan 1969).

L'avantage majeur de cette approche est qu'elle peut tenir compte de tous les phénomènes qui ont une influence sur l'hydrosystème, ce qui aide à la compréhension de son fonctionnement. Les principaux inconvénients de ces modèles sont la grande quantité d'informations qu'ils nécessitent sur le système, la représentativité imparfaite des relations physiques qui y sont implémentées et une certaine complexité de mise en œuvre associée à des temps de calcul importants. Ces temps de calculs sont utilisés pour estimer les paramètres du modèle, étape théoriquement non nécessaire, mais qui l'est en réalité, compte-tenu de la transcription mathématique imparfaite des phénomènes physiques et des variables d'entrées (Marchandise 2007), (Darras 2015). De plus, afin de leur conférer un caractère prédictif, il est fréquent que ces modèles doivent être couplés à des modèles météorologiques dont les erreurs et incertitudes peuvent se cumuler avec celles de la partie hydrologique.

Les modèles TOPKAPI (Liu et Todini 2002) et CEQUEAU (issu de l'INRS) (Morin et Paquet 2007) sont des exemples connus pour cette approche de modélisation. Le modèle hydrogéologique GAIA, qui est en cours de développement par le BRGM, sera également un modèle à base physique, qui s'appliquera sur le bassin sud-Aquitain (Wuilleumier, Cabaret et Saltel 2017), (André, et al. 2019).

On parle de modélisation conceptuelle lorsque l'on s'éloigne de l'objectif de représentation des processus physiques qui génèrent le débit ou la hauteur d'eau. Par exemple, le fonctionnement de l'aquifère est assimilé au fonctionnement d'un réservoir et le fonctionnement d'une rivière est assimilé à un second réservoir.

L'avantage de cette méthode est la compartimentation des phénomènes hydrologiques, qui permet une meilleure compréhension du fonctionnement de chaque réservoir. Malgré ce, une bonne connaissance du bassin d'étude reste nécessaire pour utiliser convenablement ces modèles. De même, ils doivent eux aussi disposer de prévisions de leurs entrées pour proposer une prévision de leurs sorties.

Le modèle SWAT (pour *Soil and Water Assessment Tool*) (Arnold, et al. 1998), le modèle TOPMODEL (Beven et Kirkby 1979), les modèles GR4J et GR6J (pour Génie Rural à 4 et 6 paramètres Journalier, issus de l'INRAE) (Perrin, Michel et Andréassian 2003, Pushpalatha 2013), le modèle HYDROL-DC (Wang, Chu et Zhang 2021) ou encore le modèle GARDENIA (pour modèle Global À Réservoirs pour la simulation de DÉbits et de NIveaux Aquifères, issu du BRGM) (Roche et Thiéry 1984) font partie de cette catégorie de modèles.

La modélisation empirique se focalise directement sur la relation entre les entrées d'un hydrosystème et ses sorties. L'objectif ici est d'utiliser des fonctions qui permettent de

retrouver cette relation. Elle est parfois associée à des « boîtes noires » car ces modèles n'expliquent pas comment l'hydrosystème se comporte. Contrairement aux autres approches, une connaissance partielle du bassin est donc suffisante pour réaliser ces modélisations. De plus, ces modèles ne nécessitent pas forcément de prévisions de leurs entrées pour fournir une prévision de leur sortie. Cependant, cette catégorie de modèles nécessite des données de qualité, sans lacune, avec parfois des séries temporelles assez longues pour avoir une modélisation hydrologique ou hydrogéologique acceptable.

Les modèles ARIMA (pour *AutoRegressive Integrated Moving Average*) et SARIMA (pour *Seasonal AutoRegressive Integrated Moving Average*) (Box et Jenkins 1970), NARMA (pour *Nonlinear AutoRegressive Moving Average*) (Connor, Atlas et Martin 1991) et les réseaux de neurones font partie de cette catégorie.

Tous ces modèles sont soumis aux incertitudes de leurs entrées ainsi qu'à celles liées à leurs éléments constitutifs et fournissent donc un résultat de modélisation entaché d'un certain niveau d'incertitude. L'évaluation de cette incertitude permet de circonscrire l'évaluation des erreurs possibles sur le résultat de modélisation. Ainsi, même avec une modélisation globalement correcte, un gestionnaire peut prendre de mauvaises décisions en l'absence d'évaluation de l'incertitude. Cette incertitude n'est encore que rarement fournie avec la sortie du modèle.

Ces types de modélisation sont mobilisés par le projet MétéEau des Nappes. Ce projet, mené par le Bureau de Recherche Géologique et Minière (BRGM), a abouti début 2021 au lancement d'un site WEB permettant d'accéder à un ensemble de services de visualisation et de prévision des niveaux de nappes en France (https://meteeaunappes.brgm.fr/fr). Les modélisations réalisées fournissent par exemple des sorties qui sont classées en fonction de l'Indice Piézométrique Standardisé (IPS), composant un intervalle de prévision (Mougin, et al. 2019).

Un autre projet, nommé AQUI-FR, permet de regrouper tous les travaux menés à une échelle régionale pour la prévision hydrogéologique. De nombreux organismes tels que le BRGM, le CNRS, Météo-France et des universités (par exemple Mines ParisTech, l'école nationale de météorologie CERFACS) collaborent pour regrouper ces travaux (Vergnes, et al. 2020).

IV. De la modélisation à l'action : les impacts des inexactitudes

La modélisation peut être utilisée comme outil d'aide à la décision. La qualité de la prise de décision dépendra donc de la qualité de la modélisation, en particulier en situation de prévision en temps réel. Cet outil d'aide à la décision doit donc être le plus exact possible et dans l'idéal, proposer un domaine de validité et une indication de sa justesse.

McMillian *et al*. citent trois exemples où l'inexactitude peut avoir un impact important sur la prise de décision. Le premier cas concerne les coûts engendrés par une mesure du débit sous-estimée en sortie d'une centrale hydroélectrique en Norvège. Dans ce cas, la centrale doit avoir un débit réservé supérieur à 11 $m^3.s^{-1}$ en période estivale. Le gouvernement norvégien applique des pénalités pour tout manquement à cette règle. Une erreur de 1 $m^3.s^{-1}$ sur la courbe de tarage des débits réservés peut engendrer des pertes de 5000 euros par jour compte-tenu des pertes dues aux pénalités et aux coûts en électricité, représentant jusqu'à 500 000 euros de pertes par été (comme ce qui s'est déroulé en 2010 pour la compagnie Statkraft), et 100 000 euros de pertes par été en cas de recalibrage des débits réservés (McMillan, et al. 2017).

Le second cas traite de la délimitation d'un bassin versant du lac Rotoroa en Nouvelle-Zélande. Les agriculteurs autour du lac ne peuvent mettre leur bétail à l'intérieur des limites du bassin versant pour limiter les risques de pollution. L'ajout d'une tolérance sur la limitation du bassin a permis d'augmenter la robustesse de la délimitation et de limiter les contestations (McMillan, et al. 2017).

Le troisième cas évoque les inconnues sur les tendances de la qualité de l'eau sur des bassins non jaugés en Nouvelle Zélande. La qualité de l'eau dépend souvent du débit. Cependant, pour les bassins non-jaugés, le débit n'est pas mesuré mais est approché par des modèles hydrologiques (cela concerne 547 des 785 bassins sur cette étude). Les tendances de qualité pour les bassins jaugés – avec des données de débits mesurés – ont aussi été étudiées par l'utilisation des modèles hydrologiques. Ces tendances ont été classées en moyenne comme étant correctes pour 90% d'entre elles, permettant de gagner en robustesse sur les bassins non jaugés (McMillan, et al. 2017).

Le dernier cas concerne le Decision Center for a Desert City (DCDC) qui est un organisme scientifique en Arizona. Cet organisme réalise de la recherche empirique et fournit des conseils sur la prise de décision pour la disponibilité de la ressource en eau en milieu urbain et pour l'adaptation au changement climatique. Cet organisme a pu identifier au cours de la dernière décennie plusieurs conséquences des inexactitudes sur la prise de décision. Cette prise de décision peut être prise par les politiques, qui se base dans certains cas sur des résultats scientifiques et leurs marges d'erreur (Larson, et al. 2015). Ces résultats peuvent parfois être mal retranscrits et/ou mal compris par les politiques, pouvant conduire à des prises de décision différentes (Larson, et al. 2015). Le DCDC met également en exergue une dimension sociale de ces divergences, les variations comportementales. Ils ont pu montrer par exemple que les personnes les plus sensibilisées aux phénomènes de sécheresses sont celles qui arrosent le plus leur jardin. Ainsi, la vision écologique et environnementale de la population, propagée par les recherches scientifiques, peut ne pas voir d'incidence sur leurs prises de décisions relatives à l'eau au quotidien (Larson, et al. 2015).

Le DCDC met l'accent également sur les différences entre les prises de décision politiques, qui doivent souvent répondre à des besoins à court-terme, et la recherche scientifique, qui répond à des besoins à des échelles de temps plus grandes. Ce décalage temporel génère ainsi des incohérences sur la prise de décision finale (Larson, et al. 2015).

A travers ces exemples, il est possible de constater que l'inexactitude de modélisation hydrologique peut avoir des impacts économiques (dégâts, amendes), socio-politiques (contestations, comportements, compréhension, temporalité), scientifiques (approches différentes), environnementaux (pollutions) et humains (santé, victimes).

V. Position de la problématique

Le réchauffement climatique est en cours et l'augmentation des émissions de CO_2 et autres gaz à effet de serre sont de nature à accentuer cette tendance. Ce dérèglement climatique a un fort impact sur notre environnement et plus particulièrement sur la ressource en eau. Les risques de sécheresse augmentent et sont aggravés par une démographie croissante qui réduit la ressource en eau (augmentation de la consommation de l'eau potable et des prélèvements en eau utilisés pour les exploitations agricoles et les industries). A l'opposé, le réchauffement climatique a aussi un impact sur les risques d'inondations, bien qu'il soit plus difficile à observer.

A l'échelle d'un bassin, les gestionnaires doivent prévenir ces risques et les gérer en cas de crise. Pour cela, une supervision du bassin au travers de la mesure est possible. Ces mesures sont un premier pas vers la connaissance du fonctionnement d'un bassin. Elles permettent de connaître son état hydrogéologique actuel et passé. Toutefois, ces bassins sont souvent mal connus (par manque de connaissances à différentes échelles spatiales), car leur fonctionnement hydrogéologique ou hydrologique est complexe. Les gestionnaires manquent généralement de moyens financiers et humains pour faire des études plus poussées. De plus, les réseaux de mesures sont parfois incomplets, mal dimensionnés, voire inexistants. Les données générées par ces réseaux peuvent comporter des lacunes, suite à des dysfonctionnements divers. La gestion en temps réel est compliquée à mettre en place et des latences entre la mesure et le stockage des données peuvent apparaître et compliquer la gestion des phénomènes.

La prévision des niveaux piézométriques et des débits semble alors être une solution afin d'anticiper les risques de sécheresse et d'inondation. Les approches de modélisation à base physique et conceptuelle nécessitent une grande connaissance du bassin. L'approche de modélisation empirique, que nous avons choisie, nécessite une connaissance moins importante. Ce type de modélisation, dont les réseaux de neurones font partie, implique une relation causale assez directe entre les entrées (pluie, débit, évapotranspiration, etc.) et les sorties du modèle (niveaux d'eau souterrains).

Toutefois, tous les modèles hydrologiques et hydrogéologiques peuvent générer des erreurs de modélisation dont l'impact sur la décision peut être atténué par l'évaluation de leur qualité. Cette évaluation peut avoir de lourdes conséquences sur la prise de décision, notamment sur les domaines financiers, sociaux, politiques et environnementaux.

Dans ce contexte, les objectifs de cette thèse répondent aux problématiques vues précédemment. Le premier objectif vise à concevoir des modèles hydrogéologiques, à court et moyen termes, afin d'anticiper les risques hydrogéologiques potentiels de sécheresse. Le second objectif est de concevoir une méthodologie de détermination de la fiabilité de modélisation afin de minimiser les prises de décision erronées. Le dernier objectif est d'appliquer cette méthodologie sur les modèles conçus et de représenter efficacement leurs limites. Pour cela nous allons faire appel au concept d'incertitude et comment cette dernière peut être estimée de manière utile pour les gestionnaires.

Chapitre II : Méthodes : réseaux de neurones artificiels et autres méthodes statistiques

I. Introduction

Les réseaux de neurones sont des systèmes mathématiques capables d'apprendre à partir d'une base de données. Après avoir décrit le fonctionnement d'un neurone formel, qui est l'élément de base constituant les réseaux de neurones, nous présenterons les caractéristiques d'un réseau de neurones et leurs propriétés fondamentales. Nous présenterons ensuite différents types d'architectures, et les règles d'apprentissage qui permettent de calculer les paramètres des différentes architectures. Les limites de ce type d'approche comme le dilemme biais-variance seront ensuite abordées ainsi que les méthodes de régularisation qui permettent de s'en affranchir. Enfin, nous nous intéresserons aux différents critères d'évaluation de leur performance avant de décrire la méthodologie complète de conception de modèle à réseau de neurones

II. Neurone formel

Le premier modèle de neurone est proposé en 1943 par McCulloch et Pitts (McCulloch et Pitts 1943), pour modéliser le fonctionnement des neurones biologiques. Ses applications se sont ensuite diversifiées dans de nombreux domaines, y compris au-delà du domaine des neurosciences.

Le neurone formel est un opérateur mathématique qui applique une transformation généralement non linéaire de ses entrées **U** vers ses sorties **Y**.

Le calcul de la sortie d'un neurone s'effectue en deux étapes :

- La somme des entrées **U** pondérée par les paramètres **W** (aussi appelés poids ou coefficients synaptiques) est calculée pour délivrer le potentiel *V* du neurone (1).
- La fonction d'activation φ, généralement non linéaire, est appliquée au potentiel *V* pour générer la sortie du neurone $\hat{y}$ (2). Différentes fonctions d'activation peuvent être utilisées : des gaussiennes (Elsafi 2014), des tangentes hyperboliques ou des sigmoïdes (Elsafi 2014), (Kong-a-Siou 2011), qui sont souvent préférées pour l'identification de fonction en hydrologie.

$$V = \sum_{i=0}^{n} W_i u_i \quad (1)$$

$$\hat{y} = \varphi(V) = \varphi(\sum_{i=0}^{n} W_i u_i) \quad (2)$$

Le fonctionnement d'un neurone formel est schématisé sur la Figure II-1.

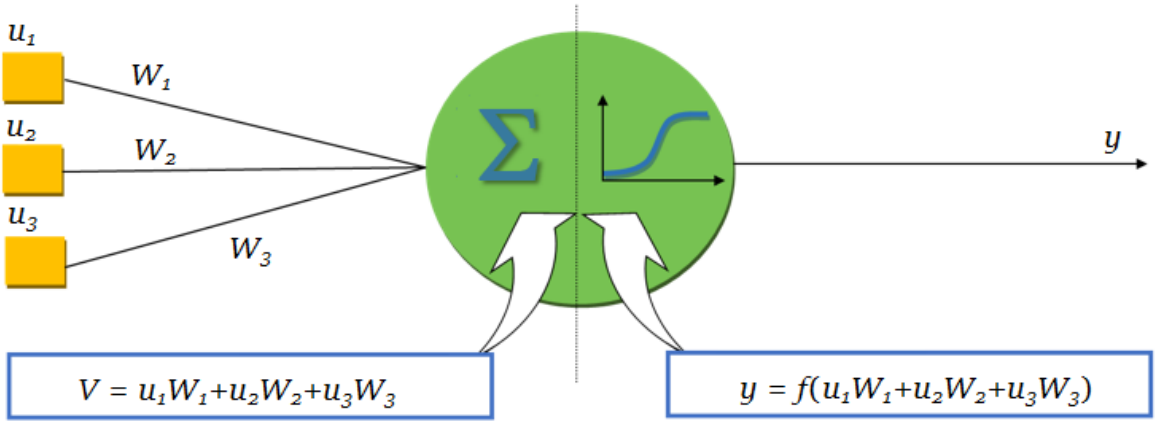


Figure II-1 - Schéma du fonctionnement d'un neurone.

III. Le réseau de neurones

III.1 Définition

Les réseaux de neurones constituent une famille de modèles statistiques non linéaires paramétrés utilisés dans de multiples domaines tels que la finance, l'imagerie, l'informatique statistique, la biologie ou encore l'hydrologie.

Ils sont une famille du *machine learning,* ou apprentissage automatique, qui s'intéresse à la conception, l'analyse et l'implémentation d'un ou de plusieurs algorithmes permettant l'apprentissage automatique d'une fonction à partir d'une base de données, de sorte de pouvoir reproduire cette fonction sur des exemples inconnus. Le *machine learning* est un embranchement de l'intelligence artificielle. On distingue deux types d'utilisation des réseaux de neurones : la régression, qui vise à identifier la fonction continue réalisé par un processus physique au moyen d'une régression, qui est le sujet de cette thèse, et la classification qui vise à attribuer un vecteur d'entrée à une classe de sortie. Comme les modèles et les méthodes peuvent varier assez grandement d'un de ces objectifs à l'autre, nous nous concentrerons dans cette partie sur les modèles et méthodes utilisés en régression.

III.2 Typologie

III.2.1. Caractéristiques principales

Les neurones sont assemblés selon différents schémas, ou graphes, appelés « architecture ».

Un réseau de neurones est composé de deux types de neurones (Dreyfus 2005) :

- des neurones de sortie, dont la sortie correspond également à une sortie du modèle,
- des neurones cachés, dont la sortie est une sortie intermédiaire du modèle, non mesurée sur le processus physique.

Selon la fonction à identifier Il est nécessaire de définir comment le temps intervient dans le modèle. Plusieurs types de modèles vont donc découler de ces choix.

III.2.2. Modèles statique et dynamique

III.2.2.a. Modèle statique

Avant de montrer les principales architectures utilisées, il est nécessaire de différencier les modèles dits « statiques », des modèles « dynamiques ».

Les modèles statiques, ou non-récurrents, sont des filtres à réponse impulsionnelle finie. Ils reçoivent des entrées qualifiées d'exogènes, c'est-à-dire qui ne sont pas issues des calculs précédents du modèle considéré (non récurrentes). Si ces entrées sont des séries temporelles, on les appliquera au réseau sur une certaine profondeur temporelle. Le temps n'a donc pas de rôle fonctionnel au sein du modèle. Le calcul effectué par les modèles statiques s'exprime de la façon suivante (3) :

$$\hat{y}(k) = \varphi(\mathbf{U}(k), \dots, \mathbf{U}(k - n_r + 1), \mathbf{W}) \quad (3)$$

Avec $\hat{y}$ étant la sortie du modèle au temps discret *k*, $\mathbf{U}$ le vecteur des entrées, $\mathbf{W}$ le vecteur des paramètres, n_r la taille des fenêtres temporelles et φ la fonction d'activation non-linéaire.

III.2.2.b. Modèle dynamique

Au contraire des modèles statiques, les modèles dynamiques sont des filtres à réponse impulsionnelle infinie. Outre les variables exogènes ils sont donc aussi alimentés par des variables récurrentes, ou bouclées.

Le bouclage permet au modèle de disposer d'une information sur ses sorties précédemment estimées. Dans certains cas ces sorties sont interprétées comme des sorties d'état et peuvent ainsi faciliter la représentation par le modèle d'un processus dynamique (4).

On écrit alors :

$$\hat{y}(k) = \varphi(\hat{\boldsymbol{y}}(k-1), \dots, \hat{\boldsymbol{y}}(k-r); \boldsymbol{u}(k), \dots, \boldsymbol{u}(k - n_r + 1), \mathbf{W}) \quad (4)$$

Où r représente l'ordre du modèle (nombre d'entrées récurrentes).

III.2.2.c. Modèle statique pour représenter un phénomène dynamique

Lorsqu'un bruit d'état trop important se rajoute sur la sortie estimée du modèle, Nerrand et *al*. ont montré qu'il peut être intéressant de remplacer l'entrée récurrente du modèle précédent par la variable équivalente mesurée sur le processus (Nerrand, Roussel-Ragot et

Personnaz, et al. 1993). On applique alors en entrée la variable que le modèle doit identifier (*y*), mais aux pas de temps précédents (5).

$$\hat{y}(k) = \varphi(y(k-1), \dots, y(k-r); \boldsymbol{u}(k), \dots, \boldsymbol{u}(k-n_r+1), \mathbf{W}) \quad (5)$$

Avec y la sortie mesurée sur le processus.

III.2.3. Principales architectures

De très nombreux types d'architectures de réseaux de neurones sont décrits dans la littérature. Nous limiterons la présentation ci-après à celles qui ont été utilisées durant cette thèse.

III.2.3.a. Perceptron multicouche

Le modèle le plus utilisé dans le domaine hydrogéologique est le perceptron multicouche (Figure II-2) (Kong-A-Siou, Johannet et Borrell Estupina, et al. 2015, Johannet, Taver, et al. 2012, Darras, Johannet, et al. 2014, Maca, Pech et Pavlasek 2014, Kong-A-Siou, Fleury, et al. 2014). Ce dernier est composé d'au moins deux couches de neurones, dont au moins une couche de neurones à fonction d'évaluation non linéaire, dits « cachés » et une couche de sortie. Le modèle est alimenté par un vecteur d'entrées **U** et délivre sa sortie $\hat{y}$. Le perceptron multicouche est une généralisation au cas non linéaire du perceptron simple, introduit par Rosenblatt en 1958 (Rosenblatt 1958). Ce dernier composé d'une couche de neurones de sortie notée $\hat{y}$, est limité à la résolution de problèmes linéaires.

Afin que le modèle puisse délivrer une sortie non nulle lorsque toutes les entrées exogènes sont nulles, il est nécessaire d'appliquer une entrée constante, appelé biais, dont le poids est noté W_0. Généralement cette entrée constante est égale à 1, le paramètre W_0 n'est donc pas distingué des autres en ajoutant cette entrée constante au vecteur des entrées **U**.

Lorsqu'il n'a qu'une sortie, la sortie $\hat{y}$ du perceptron multicouche est calculée ainsi (6) :

$$\hat{y}(k) = \varphi(W_{s,0}u_0 + \sum_{j=1}^{m}(W_{s,j}\varphi(\sum_{i=0}^{n} W_{j,i}u_i))) \quad (6)$$

Avec φ la fonction d'activation des neurones cachés, $W_{s,0}$ le biais de l'entrée constante u_0, $W_{s,j}$ le paramètre d'un neurone de la couche cachée relié à un neurone de la couche de sortie et $W_{j,i}$ le paramètre de la variable d'entrée u_i reliée à un neurone de la couche de sortie.

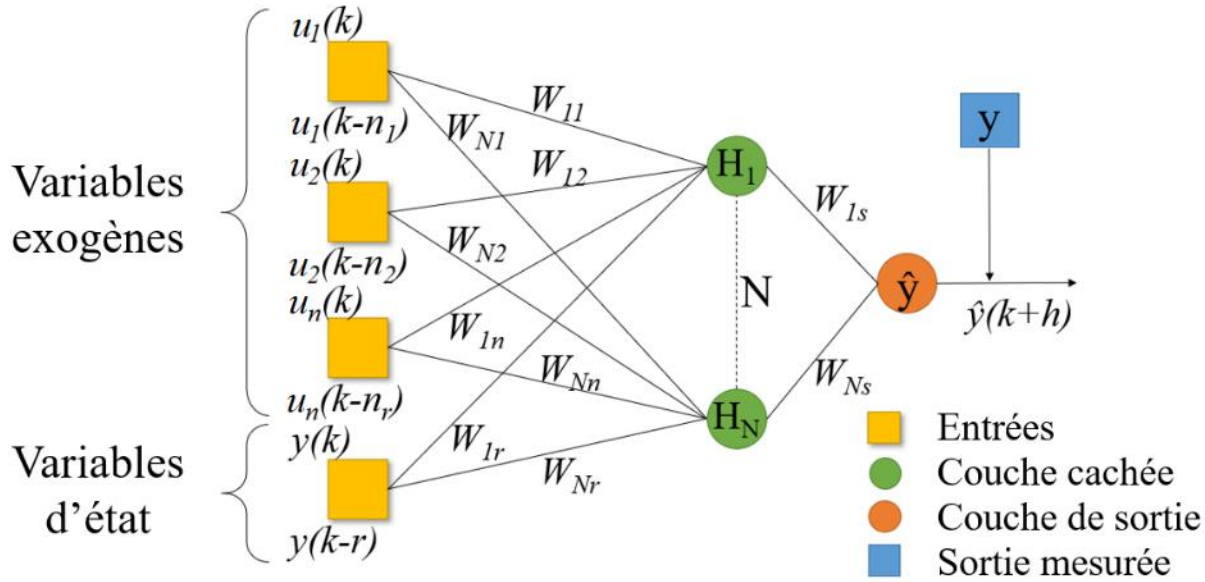


Figure II-2 - Schéma d'un perceptron multicouche.

III.2.3.b. Propriétés fondamentales du perceptron multicouches

III.2.3.b.i. Approximation universelle

Le perceptron multicouche possède la propriété d'approximation universelle (Hornik, Stinchombe et White 1989, Hornik 1991) (Funahashi 1989, Cybenkot 1989). Cette propriété exprime que : « Toute fonction bornée suffisamment régulière peut être approchée uniformément, avec une précision arbitraire, dans un domaine fini de l'espace de ses variables, par un réseau de neurones comportant une couche de neurones cachés en nombre fini, possédant tous la même fonction d'activation bornée, et un neurone de sortie linéaire. » (M. Toukourou 2009). La précision de l'approximation dépend du nombre de neurones cachés. Contrairement à son homologue à une couche, le perceptron multicouche est capable d'identifier des fonctions non-linéaires. Ainsi, il est très adapté quant à la modélisation hydro(géo)logique (Hornik, Stinchombe et White 1989, Hornik 1991).

III.2.3.b.ii. Parcimonie

Le perceptron multicouche est aussi parcimonieux, ce qui signifie que, par rapport à d'autres modèles statistiques, il utilise moins de paramètres pour identifier une fonction. Barron explique dans ses travaux que la parcimonie du modèle est due à la non-linéarité de la sortie vis à vis des paramètres, contrairement aux autres approximateurs non-linéaires dont la sortie est non linéaire seulement par rapport aux variables. Il a montré également que plus le nombre d'entrées augmente, plus le modèle est parcimonieux en comparaison d'un approximateur linéaire en ses paramètres (Barron 1991). L'avantage de la parcimonie est particulièrement sensible pour les bases de données réduites, comportant peu évènements rares.

Si l'on ne s'intéresse qu'au nombre de paramètres libres, les réseaux de neurones sont moins parcimonieux que la plupart des modèles hydrologiques à base physique. Cependant il faut bien avoir à l'esprit que, grâce à la propriété d'approximation universelle, le réseau de neurones, en plus de déterminer les « paramètres », doit aussi reconstituer la fonction du processus modélisé, à partir des fonctions non-linéaires que sont ses sigmoïdes.

III.2.3.c. Long-Short-Term Memory

Savary M. a soulevé la difficulté à intégrer efficacement des variables à des échelles de temps très différentes au sein d'une même modèle (Savary 2018). Cette difficulté peut être un frein à la modélisation hydrogéologique. En effet, les processus de long terme, comme le soutien d'étiage par des apports très différés, se superposent à des processus de court-terme, comme la réaction du système à une averse. Un autre type de modèle a donc été proposé, afin de pallier cette difficulté. Le modèle *Long-Short Term Memory*, ou LSTM, est un modèle récurrent composé de cellules, nommées « cellules LSTM ». L'avantage de cette architecture est qu'elle permet de conserver la mémoire des évènements sur le long terme, grâce aux termes récurrents, tout en leur garantissant de ne pas être atténués si leur rôle est important, grâce à des « portes » qui permettent de filtrer les entrées et les sorties (Figure II-3 - Représentation schématique d'un réseau Long Short Term Memory.). Ce réseau a été proposé par Hochreiter et Schmidhuber en 1997 (Hochreiter et Schmidhuber 1997), étudiée par Gers et son équipe (Gers, Schmidhuber et Cummins 2000) ; il fait actuellement partie des modèles les plus utilisés pour traiter des données temporelles (Glorot et Bengio 2010, Nérot 2018).

La particularité de ce réseau est la structure de ses cellules. Ces dernières sont organisées autour d'un perceptron à une couche comportant des entrées exogènes au temps discret k ($\mathbf{X}^k$) et des entrées récurrentes ($\mathbf{H}^{k-1}$). La sortie de ce perceptron à une couche apparait sous la forme de la tangente hyperbolique dans l'équation (10). Un objet spécifique, appelé « porte » est introduit, qui a pour objet de filtrer les variables d'entrée afin de les bloquer ou de leur permettre de modifier les vecteurs d'états, ces derniers étant décomposés en état de contexte (sortie calculée par la cellule : $\mathbf{C}^k$) et en état interne (sortie des neurones cachés $\mathbf{H}^k$). Les portes s'appliquent aux vecteurs en calculant leur produit de Hadamard (vecteur produit des composantes des vecteurs deux à deux) ; un des vecteurs a des composantes en {0,1} afin de bloquer la composante de l'autre vecteur si cette dernière est face à un « 0 », ou de la laisser passer si elle est face à un « 1 ».

Lorsque l'on a un réseau récurrent, sa sortie au temps discret k fait appel à son entrée, qui n'est autre que sa sortie au temps k-1. Elle même dépend à son tour de sa sortie au temps k-2, et ainsi de suite. Généralement on « déplie » la récurrence sur un certain horizon temporel fini en dupliquant le modèle, chaque « copie » étant constituée d'un modèle opérant avec un retard temporel donné. Le LSTM représente ainsi le « dépliement » temporel de la récurrence en assemblant les cellules en couches successives. Ce dernier se déploie selon deux axes : un

axe « temporel », qui transmet l'état interne ($\mathbf{H}^{k-1}$) de la cellule correspondant au temps discret *k*-1, vers la cellule correspondant au temps *k* via, entre autres, la porte de sortie ($\mathbf{O}^{k}$, eq. (9)). L'autre axe « entrée-sortie », permet de filtrer les informations entrant dans la cellule : $\mathbf{X}^{k}$, $\mathbf{H}^{k-1}$, via chacune des portes d'entrée ($\mathbf{I}^{k}$, eq. (8)), et d'oubli ($\mathbf{F}^{k}$, eq. (7)).

Les opérations réalisées au sein d'une cellule sont les suivantes :

1) Axe entrées-sorties.

$$\mathbf{F}^{k} = \sigma(\mathbf{W}_{F}\mathbf{X}^{k} + \mathbf{U}_{F}\mathbf{H}^{k-1}), \quad (7)$$

Avec $\mathbf{F}^{k}$ la porte d'oubli, avec $\mathbf{W}_{F}$ et $\mathbf{U}_{F}$, les matrices de paramètres associées, $\sigma(.)$, la fonction sigmoïde {0 ;1} utilisée par la porte, $\mathbf{X}^{k}$ et $\mathbf{H}^{k-1}$ les vecteurs d'entrées, respectivement exogènes et récurrentes.

$$\mathbf{I}^{k} = \sigma(\mathbf{W}_{I}\mathbf{X}^{k} + \mathbf{U}_{I}\mathbf{H}^{k-1}), \quad (8)$$

Avec $\mathbf{I}^{k}$ la porte d'entrée, avec $\mathbf{W}_{I}$ et $\mathbf{U}_{I}$, les matrices de paramètres associées, $\mathbf{X}^{k}$ et $\mathbf{H}^{k-1}$ et $\sigma(.)$, telles que définis précédemment.

2) Axe temporel : actualisations du contexte $\mathbf{C}^{k}$ et de $\mathbf{H}^{k-1}$ (état caché).

$$\mathbf{O}^{k} = \sigma(\mathbf{W}_{O}\mathbf{X}^{k} + \mathbf{U}_{O}\mathbf{H}^{k-1}), \quad (9)$$

Avec $\mathbf{O}^{k}$ la porte de sortie, avec $\mathbf{W}_{O}$ et $\mathbf{U}_{O}$, les matrices de paramètres associées, $\mathbf{X}^{k}$ et $\mathbf{H}^{k-1}$ et $\sigma(.)$, tels que définis précédemment.

$$\mathbf{C}^{k} = \mathbf{C}^{k-1} \otimes \mathbf{F}^{k} + \mathbf{I}^{k} \otimes \tanh(\mathbf{W}_{c}\mathbf{X}^{k} + \mathbf{U}_{c}\mathbf{H}^{k-1}), \quad (10)$$

Il s'agit du calcul du contexte interne de la cellule, avec $\otimes$, le produit de Hadamard, $\mathbf{W}_{c}$ et $\mathbf{U}_{c}$ les paramètres associés au perceptron à une couche déjà cité, $\mathbf{C}^{k}$, $\mathbf{F}^{k}$, $\mathbf{I}^{k}$, $\mathbf{X}^{k}$ et $\mathbf{H}^{k-1}$, tels que définis précédemment. On note que cet état de contexte est actualisé via son contexte précédent, bloqué, ou pas, par la porte d'oubli, à laquelle s'ajoute la sortie du perceptron à une couche, également bloquée, ou pas, par la porte d'entrée $\mathbf{I}^{k}$.

$$\mathbf{H}^k = \mathbf{O}^k \otimes \tanh(\mathbf{C}^k), \quad (11)$$

Il s'agit du calcul de l'état interne, avec les mêmes notations que précédemment. Cet état interne est filtré par la porte de sortie puis transmis à la cellule suivante.

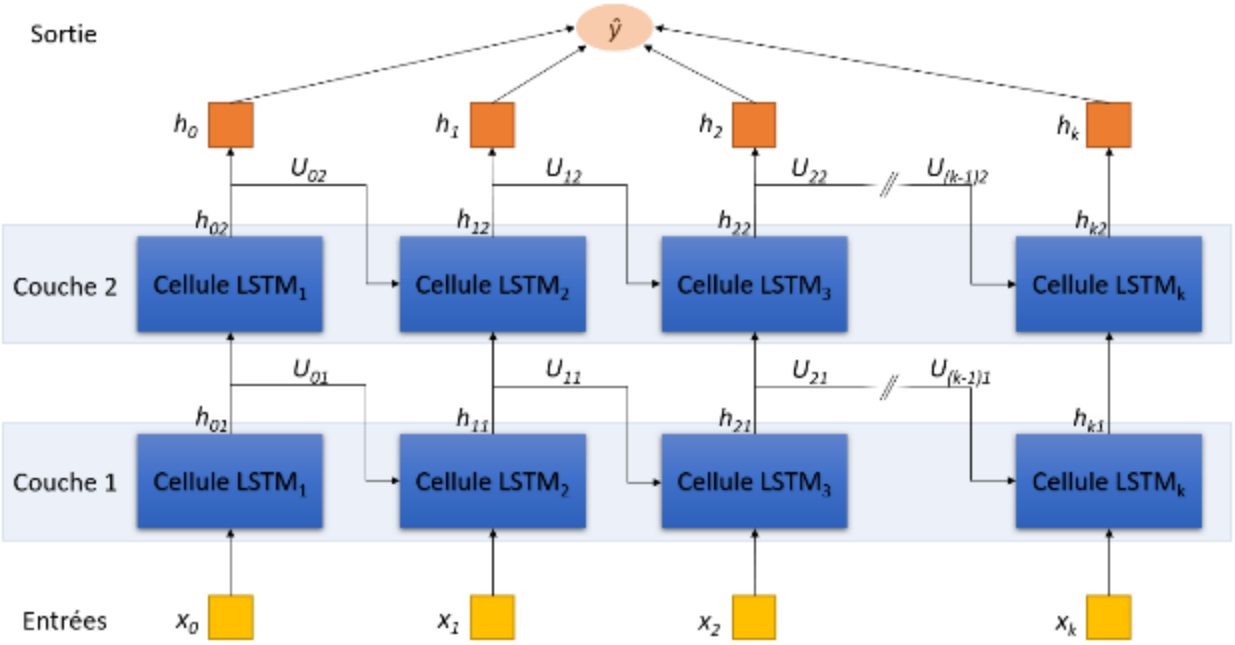


Figure II-3 - Représentation schématique d'un réseau Long Short Term Memory.

Ces architectures sont très utilisées pour le traitement des séquences, par exemple la reconnaissance de mots dans une phrase, ou la traduction automatique ; elles sont cependant moins utilisées que les perceptrons multicouches en hydrologie et hydrogéologie. Toutefois, ces architectures ont déjà été testées sur des modèles pluie-débit avec des résultats satisfaisants (Li, Kiaghadi et Dawson 2020), ainsi que pour la prévision des hauteurs d'eau souterraines en contexte agricole (Zhang, et al. 2018).

III.3 Apprentissage des réseaux de neurones

III.3.1. Principe de l'apprentissage

L'apprentissage correspond au calcul des paramètres du réseau de neurones de manière à minimiser, de manière itérative, une fonction de coût. La fonction de coût utilisée dans ce travail est l'erreur quadratique moyenne notée *J*. Comme la sortie du modèle dépend non linéairement de ses entrées et de ses paramètres, l'apprentissage est réalisé de manière itérative par une règle d'apprentissage. Pour cela il est nécessaire de disposer d'une base de données comportant un ensemble d'exemples représentant l'ensemble des comportements que le modèle doit apprendre. Un exemple est un couple de vecteurs connus : vecteur d'entrées $\mathbf{U}^p$ et vecteur de sortie associé : $\mathbf{Y}^p$, $p \in \mathbb{N}$.

L'apprentissage se déroule en cinq étapes :

1. Initialisation aléatoire des paramètres,
2. Propagation : choix d'un exemple et calcul des potentiels et des sorties des neurones depuis la couche des entrées jusqu'à la couche de sortie,
3. Calcul de la fonction de coût ***J***,
4. Correction des paramètres,
5. Retour à la deuxième étape tant que le critère d'arrêt n'est pas atteint (nombre de calculs ou minimum de la fonction de coût).

La fonction de coût la plus communément utilisée en hydrologie est l'erreur quadratique moyenne, MSE pour *Mean Squared Error* (12)*,* ou sa racine carrée (RMSE pour *Root Mean Squared Error*) (13).

$$J_{MSE} = \frac{1}{N} \sum_k (Y_{obs,k} - Y_{p,k})^2, \quad (12)$$

$$J_{RMSE} = \sqrt{\frac{1}{N} \sum_k (Y_{obs,k} - Y_{p,k})^2}, \quad (13)$$

Avec $Y_{obs,k}$ la sortie observée au temps *k*, $Y_{p,k}$ la sortie simulée ou estimée au temps *k* et N la longueur de la chronique.

Comme l'erreur quadratique est non-linéairement croissante avec l'augmentation de l'opérande, elle a tendance à apporter plus d'importance aux grandes valeurs. Pour diminuer cet effet on peut préférer la racine carrée de l'erreur quadratique, ou sa valeur absolue (éq. (13)).

Ainsi, pour s'adapter à l'objectif de modélisation, d'autres fonctions de coût sont également utilisées, telles que l'erreur absolue moyenne (MAE pour Mean Absolute Error (14)), la valeur moyenne de l'erreur (ME pour Mean Error (15)) ou encore l'erreur maximale absolue (AME pour Absolute Maximum Error (16)) (Kouki 2010, Bennett, et al. 2013) :

$$J_{MAE} = \frac{1}{N} \sum_k |(Y_{obs,k} - Y_{p,k})|, \quad (14)$$

$$J_{ME} = \frac{1}{N} \sum_k (Y_{obs,k} - Y_{p,k}), \quad (15)$$

$$J_{AME} = max\,|Y_{obs,k} - Y_{p,k}|, \quad (16)$$

Une itération est réalisée en présentant au réseau l'ensemble de la base de données dédiée à l'apprentissage. Les itérations sont répétées jusqu'à la satisfaction du critère d'arrêt. Il existe de très nombreuses variantes dans la façon de présenter les exemples au réseau : soit un par un dans l'ordre, soit dans un ordre aléatoire. De la même façon, les incréments aux paramètres peuvent être appliqués soit à chaque présentation d'un exemple comme indiqué

plus haut (gradient partiel, ou gradient stochastique), soit, cumulés à la fin de la présentation (gradient total, ou « batch » en anglais) soit avec des modes intermédiaires.

III.3.2. Règles d'apprentissage

La minimisation de la fonction de coût se fait grâce à une règle d'apprentissage qui opère sur un sous ensemble de la base de données : l'ensemble d'apprentissage. L'objectif de l'apprentissage est de calculer les paramètres du modèle sur l'ensemble d'apprentissage de manière à pouvoir généraliser le fonctionnement appris à d'autres sous-ensembles, jamais vus par le modèle. Cette fonctionnalité attendue est ce que l'on appelle la généralisation.

III.3.2.a. Normalisation préalable des variables

Les réseaux de neurones sont alimentés par de nombreuses données. Toutefois, celles-ci doivent être normalisées entre 0 et 1 (17), ou entre -1 et 1, dans le cas où les fonctions d'activation sont bornées telles que les sigmoïdes ou les tangentes hyperboliques.

$$X^{\mathrm{norm}} = \frac{(X - X_{\mathrm{min}})}{(X_{\mathrm{max}} - X_{\mathrm{min}})}, \quad (17)$$

X^{norm} étant la valeur normalisée, X la valeur à normaliser X_{min} la valeur minimale et X_{max} la valeur maximale du vecteur **X**.

Pour cela, il est souvent recommandé dans la littérature de réaliser une normalisation centrée-réduite, plutôt qu'une normalisation simple. Cela évite que la sortie calculée soit écrêtée par la fonction d'activation (Johannet 2011).

La normalisation centrée réduite s'écrit selon la formule suivante (18) :

$$\mathbf{\chi} = \frac{(X - \overline{X})}{\sigma_X} \quad (18)$$

Avec **X** le vecteur de variables d'entrée, $\overline{X}$ sa moyenne, σ_X son écart-type et $\mathbf{\chi}$ la centrée réduite de **X**.

Il est cependant préférable d'appliquer à ce calcul un facteur multiplicatif pour que les données soient normalisées entre 0 et 0,9 afin d'anticiper sur des données ultérieures qui dépasseraient le maximum historique enregistré.

III.3.2.b. Règles du premier ordre

La méthode la plus utilisée pour diminuer la valeur de la fonction de coût est la méthode du gradient. Etudiée par Cauchy au milieu du XIX^{e} siècle, puis généralisée par Curry en 1944 pour des optimisations non-linéaires (Curry 1944), la règle du gradient est un algorithme itératif permettant de trouver un minimum local d'une fonction. Partant d'un point de

fonctionnement donné, l'initialisation par exemple, le principe de ces règles est d'ajouter à la matrice des paramètres un incrément proportionnel à l'opposé de la dérivée, de manière à déplacer le point de fonctionnement du modèle vers le minimum de la fonction de coût.

La première étape est la propagation, qui consiste à calculer la sortie associée aux entrées appliquées, puis de déterminer le gradient ∇ de la fonction de coût J, c'est-à-dire de calculer la dérivée de la fonction par rapport à la matrice des paramètres $\mathbf{W}$. La règle du gradient la plus simple est celle dite de « gradient de plus grande pente » (19), elle s'exprime ainsi :

$$\mathbf{W}_{k+1} = \mathbf{W}_k - \alpha_k \nabla J(\mathbf{W}_k), \tag{19}$$

Avec α_k, le « pas du gradient ».

Le pas du gradient α_k, peut être choisi constant ou varier selon différentes stratégies en fonction de l'avancement de la minimisation de la fonction de coût. Comme la fonction de coût prend en général en compte des termes qui dépendent non linéairement des paramètres, elle est affectée de minima locaux (Figure II-4). Comme la règle du gradient de plus grande pente ne peut déplacer le point de fonctionnement qu'en minimisant sa fonction de coût, elle reste alors bloquée au fond de ces minima locaux.

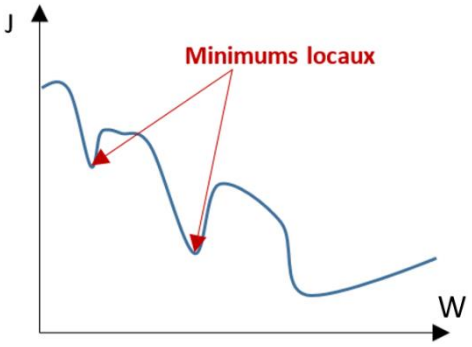


Figure II-4 - Illustration de l'existence de minimums locaux.

La règle de la rétropropagation (Rumelhart, Hinton et Williams 1986), très utilisée pour calculer le gradient de la fonction de coût dans les réseaux multicouches, est une règle de gradient de plus grande pente. En tant que règle d'apprentissage, elle souffre donc de cet inconvénient majeur. Elle est cependant très utilisée pour sa façon astucieuse de calculer le gradient de la fonction de coût au sein des réseaux multicouches, mais n'est plus que très peu utilisée dans sa forme simple de calcul des incréments.

Pour éviter le problème de sensibilité aux minima locaux, on peut, par exemple, ajouter un terme d'inertie qui dépend des incréments précédents et qui permet de « passer par-dessus » les minima locaux de faible ampleur.

Par exemple, la règle d'ADAM (pour *ADAptative Moment estimation*), conserve un historique des incréments précédents, pondérés de manière exponentiellement décroissante, afin d'obtenir un effet d'inertie qui permet au réseau de ne pas rester bloqué dans les minima locaux de la fonction de coût (Kingma et Ba 2014).

Cette règle utilise les estimations des moments d'ordre 1 et d'ordre 2 (moyenne et variance), des paramètres précédemment calculés, notés m_k pour la moyenne et v_k pour la variance.

L'incrément ici comprend ces deux termes pour le calcul de la matrice des paramètres **W**, selon l'équation (19), avec un incrément égal à $\Delta \mathbf{W}_k = -\alpha_k \frac{m_k}{\sqrt{v_k}+\hat{e}} \nabla J(\mathbf{W}_k)$ où $\hat{e}$ est un nombre très faible (10^{-8} en général), dont l'objet est que le dénominateur de l'incrément ne soit jamais nul.

$$\mathbf{W}_{k+1} = \mathbf{W}_k - \alpha_k \frac{m_k}{\sqrt{v_k}+\hat{e}} \nabla J(\mathbf{W}_k) \quad (19)$$

La moyenne m_k et la variance v_k sont calculées selon les équations (20) et (21), à partir d'hyper-paramètres β_i (Kingma et Ba 2014).

$$m_k = \beta_1 m_{k-1} + (1-\beta_1).\nabla J(\mathbf{W}_{k-1}) \quad (20)$$

$$v_k = \beta_2 v_{k-1} + (1-\beta_2).(\nabla J(\mathbf{W}_{k-1}))^2 \quad (21)$$

III.3.2.c. Règles du second ordre

Pour pallier les inconvénients précédemment cités, et diminuer le nombre d'itérations, la fonction de coût peut également être minimisée en utilisant les dérivées secondes. Le principe généralement adopté est que la fonction de coût est supposée de forme quadratique près du minimum global et que la règle utilisant les dérivées premières (matrice Jacobienne) peut être complétée par la matrice des dérivées secondes (matrice Hessienne). C'est le cas de la règle de Levenberg-Marquardt, qui simplifie le calcul des dérivées secondes, en négligeant les termes d'ordre 2 et en ne gardant que les produits des termes d'ordre 1 (les dérivées premières). Cette simplification permet de n'utiliser que les dérivées premières, ces dernières peuvent être calculées de manière simple par la rétropropagation. La règle de Levenberg-Marquardt utilise ainsi une matrice hessienne simplifiée de la fonction de coût pour mettre à jour les paramètres, selon l'équation (22) (Levenberg 1944, Marquardt 1963).

Par ailleurs, afin de n'appliquer le calcul du second ordre que proche du minimum global, les auteurs ont proposé de commencer la minimisation par une règle du premier ordre afin d'avancer rapidement vers le minimum global. Lorsque le modèle s'approche suffisamment du minimum global (ou supposé comme tel), on passe à la méthode mobilisant l'approximation de la Hessienne. L'arbitrage entre le fonctionnement du type premier ordre et du type second ordre est fait grâce à un hyper-paramètre appelé λ dans l'équation (23).

$$H(\mathbf{W}_k) = \sum_{ij,lm}^{N,N} \left(\frac{\partial J(\mathbf{W}_k)}{\partial W_{ij}}\right) . \left(\frac{\partial J(\mathbf{W}_k)}{\partial W_{lm}}\right)^T \quad (22)$$

$$\mathbf{W}_{k+1} = \mathbf{W}_k - [H(\mathbf{W}_k) + \lambda_{k+1}\mathbf{I}]^{-1} . \nabla J(\mathbf{W}_k) \quad (23)$$

Avec **I** la matrice identité, λ_{k+1} l'hyper-paramètre défini précédemment, $H(\mathbf{W}_k)$ l'approximation de la matrice hessienne au pas de temps *k*, W_{ij} et W_{lm} les paramètres par rapports auxquels sont calculés les gradients et *N* le nombre de paramètres.

Maca et al. ont étudiés sept règles d'apprentissage différentes et ont conclu que, pour leur cas d'étude, la règle de Levenberg-Marquardt était la plus performante (Maca, Pech et Pavlasek 2014). L'inconvénient de cette règle est qu'elle nécessite de calculer l'inverse d'une matrice de dimension (*N*, *N*) où *N* est le nombre de paramètres. Elle augmente donc très rapidement de dimension avec la complexité du modèle, ce qui la rend inopérante en pratique sur les problèmes de très grande dimension et est donc difficile à utiliser pour les réseaux comportant trop de paramètres.

Il existe d'autres règles qui approchent la matrice Hessienne, la plus connue étant l'algorithme de Broyden-Fletcher-Goldfarb-Shanno. Également connu sous l'acronyme BFGS, cet algorithme approche la valeur de la matrice hessienne à partir du gradient en recherchant le pas d'apprentissage optimal. L'inconvénient de cette technique est qu'elle nécessite de stocker de nombreux résultats intermédiaires en mémoire et est donc difficile à utiliser pour les réseaux comportant trop de paramètres (Broyden 1970, Fletcher 1970, Goldfarb 1970, Shanno 1970).

III.3.3. Dilemme biais-variance et sur-ajustement

L'objectif de la mise en œuvre de l'apprentissage des réseaux de neurones en régression est d'approcher au mieux la fonction de régression, grâce à la base de données disponible. Durant la phase d'apprentissage, à chaque itération, les paramètres du modèle sont calculés pour approcher la fonction de régression. Geman *et al.*, dès 1992, ont relevé le fait que l'erreur calculée sur l'ensemble d'apprentissage n'était pas un bon estimateur de l'erreur de généralisation : l'erreur calculée sur un autre ensemble, jamais vu par le modèle. Qui plus est, Geman *et al* ont montré que plus le modèle utilisé était complexe, c'est-à-dire plus il avait de paramètres libres, plus la différence entre les deux erreurs (apprentissage-généralisation) augmentait (Geman, Bienenstock et Doursat 1992). L'erreur peut être décomposée en trois termes :

- Le biais qui traduit comment l'espérance mathématique de la prévision est éloignée de la régression théorique (sans bruit),
- La variance qui rend compte de la variance des prédictions possibles,
- Et le bruit irréductible qui est compris dans les données.

Un biais trop élevé signifie généralement que le modèle n'est soit pas assez complexe (pas assez de paramètres libres), soit pas assez optimisé (nombre d'itérations trop faible), et qu'il n'est pas capable d'approcher suffisamment la fonction de régression.

Une variance trop élevée signifie généralement que le modèle est trop complexe ; soit du fait du trop grand nombre de paramètres libres, soit parce qu'il est trop optimisé sur l'ensemble d'apprentissage. Il est alors sur-ajusté à l'ensemble d'apprentissage et perd sa capacité de généralisation en se focalisant sur les détails de cet ensemble, incluant le bruit.

La recherche de cet équilibre entre biais et variance a été décrite par Geman *et al.* et connue sous le nom de « dilemme biais-variance » (Figure II-5) (Geman, Bienenstock et Doursat 1992).

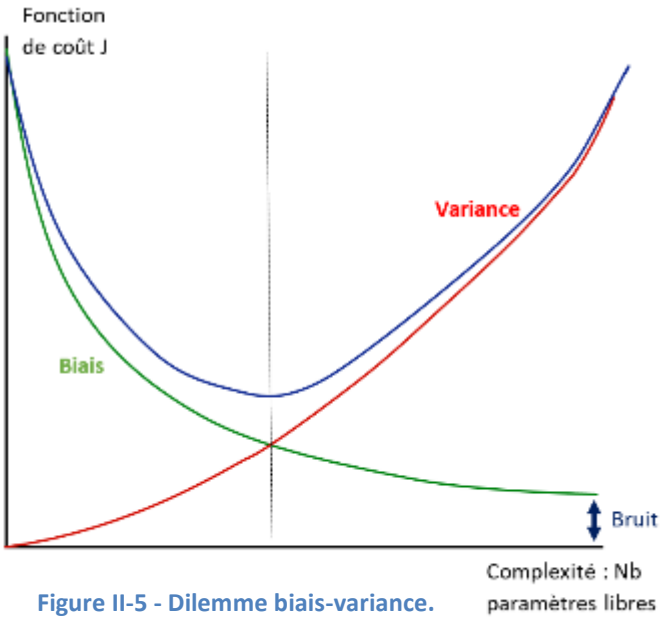


Figure II-5 - Dilemme biais-variance.

III.3.4. Méthodes de régularisation

Les méthodes de régularisation permettent de limiter la variance, tout en minimisant le biais. Plusieurs techniques existent ; trois d'entre elles, qui sont utilisées dans ces travaux, sont présentées ici.

III.3.4.a. Arrêt précoce

L'arrêt précoce vise à arrêter l'apprentissage lorsque la capacité de généralisation du modèle commence à se dégrader. Pour cela, un ensemble de données complémentaire est utilisé, appelé ensemble d'arrêt. L'apprentissage est effectué et l'erreur est mesurée sur l'ensemble d'arrêt au fur et à mesure que l'apprentissage se déroule. Lorsque le modèle se spécialise sur son ensemble d'apprentissage – on parle alors de sur-apprentissage si le modèle est trop spécialisé sur cet ensemble –, l'erreur sur l'ensemble d'arrêt augmente. On arrête donc l'apprentissage, et on conserve les paramètres du modèle correspondant à l'itération où la fonction de coût calculée sur l'ensemble d'arrêt est minimale.

Sjöberg et *al.* ont montré que l'arrêt précoce permet de diminuer le nombre de paramètres effectifs du modèle. Le raisonnement est le suivant : au début de l'apprentissage, les paramètres sont initialisés à des valeurs aléatoires faibles. Tant que les paramètres restent d'amplitude très faible, on peut considérer qu'ils ont un rôle négligeable dans la fonction

opérée par le modèle, et qu'on peut les négliger. Or, au fur et à mesure que l'apprentissage se déroule, les paramètres augmentent en amplitude entrainant ainsi une augmentation de la complexité du modèle. Ainsi, utiliser l'arrêt précoce permet d'arrêter l'apprentissage et a donc pour effet de réduire indirectement la complexité réelle du modèle (Sjöberg, et al. 1995).

Néanmoins, le choix de l'ensemble d'arrêt est essentiel. Toukourou et al. ont proposé une solution élégante pour ce choix : utiliser le sous ensemble de la base de donnée le mieux modélisé par l'ensemble d'apprentissage (Toukourou, et al. 2011). Ceci assure que l'apprentissage est auto cohérent avec l'ensemble d'arrêt qui lui est associé.

III.3.4.b. Validation croisée : un estimateur de l'erreur de généralisation

La validation croisée permet d'estimer l'erreur de généralisation sur l'ensemble de la base d'apprentissage, en situation de généralisation. Pour pratiquer la validation croisée (Stone 1974), l'ensemble d'apprentissage est sous-divisé en *N* sous-ensembles E_i de taille pertinente au regard de l'ensemble des données. Par un jeu de permutations successives, chacun des *N* ensembles est utilisé en validation, c'est-à-dire en généralisation, alors que tous les autres sont utilisés en apprentissage (Figure II-6). A chacune de ces permutations, on peut calculer la valeur de la fonction de coût $J(E_i)$, ou de tout autre score $S(E_i)$ calculé sur l'ensemble de validation E_i considéré. Après les *N* permutations, le score de validation croisée S_{VC} peut être calculé ainsi (24) :

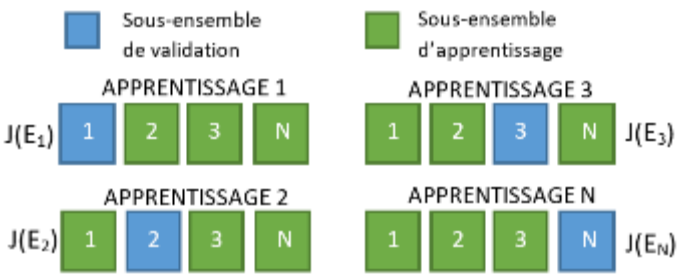


Figure II-6 - Principe de la validation croisée.

$$S_{VC} = \sqrt{\frac{1}{N}\sum_i^N S(E_i)^2} \qquad (24)$$

Il est donc ainsi possible d'obtenir une quantification de l'erreur de variance du modèle sur tout l'ensemble d'apprentissage, chaque sous-ensemble ayant été considéré successivement en situation de validation (*i.e.* généralisation).

III.3.4.c. Modèles d'ensemble

Les modèles d'ensemble sont un ensemble de modèles différents, appelés membres de l'ensemble, estimant la même grandeur. Ils peuvent différer de plusieurs façons : il est possible d'utiliser plusieurs méthodes d'apprentissage pour générer un ensemble de modèles (Džeroski et Ženko 2004), d'utiliser des techniques bayésiennes (Wagena, et al. 2019, He, et al. 2018, Liang, et al. 2013), d'utiliser des algorithmes générant aléatoirement des jeux de

paramètres (Lee et Kang 2016) ou d'utiliser différents modèles avec un apprentissage par validation croisée (Džeroski et Ženko 2004).

Par ailleurs il a été observé par Darras *et al*. que l'initialisation aléatoire des paramètres du modèle pouvait engendrer une forte variabilité en sortie (Darras, Johannet, et al. 2014). Ainsi il a été proposé par ces auteurs de réduire cette variabilité en réalisant un modèle d'ensemble dont les membres ne varient que par l'initialisation de leurs paramètres et dont la sortie à chaque pas de temps est la médiane des sorties de chaque membre, à chaque pas de temps. Cette méthode peut donc s'apparenter à une méthode de régularisation car elle réduit significativement la variance de la sortie du modèle d'ensemble et rend le modèle beaucoup plus robuste.

III.3.5. Choix des variables

La sélection de variables est une étape fondamentale dans la conception d'un modèle neuronal. Gardant en mémoire le dilemme biais-variance, il apparait essentiel de choisir le sous-ensemble de variables le plus réduit et le plus informatif, qui conduit au modèle de complexité minimale ; celui-ci induisant alors les meilleures performances de généralisation. C'est ce que l'on appelle la « sélection de variables ». Ce processus est une des étapes clé de toute modélisation statistique.

III.3.5.a. Sélection experte : choix des variables

Considérant que la variable de sortie est prédéterminée par l'objectif de la modélisation, dans certains types de problèmes il est possible de choisir les variables d'entrée du modèle de manière experte ; par exemple lorsque l'on dispose d'un modèle de type physique et que l'on connait donc les variables pertinentes à utiliser. Néanmoins comme les modèles statistiques réalisés par les réseaux de neurones offrent des fonctionnalités plus larges, il est possible, voire utile, d'élargir les possibilités de variables à appliquer au modèle. Dans ce contexte un premier moyen de sélection des variables d'entrées est l'analyse corrélatoire. Il s'agit alors de choisir les variables d'entrées qui sont le plus corrélées à la sortie du modèle.

III.3.5.b. Sélection statistique : analyses corrélatoires

La corrélation consiste à calculer le coefficient de corrélation linéaire entre deux signaux. On peut calculer la corrélation simple c'est-à-dire entre le signal et lui-même ou la corrélation croisée.

La corrélation simple, également appelée autocorrélation, permet le calcul du coefficient de corrélation entre un signal au temps discret *k* avec ce même signal décalé d'un pas de temps noté *l.* Le calcul de l'autocorrélation $C_{xx}(l)$ s'exprime de la façon suivante (25) :

$$C_{xx}(l) = \frac{1}{n} \sum_{i=1}^{n-l}(x_i - x_{moy})(x_{i+l} - x_{moy}) \cdot \frac{1}{\sigma_x \sigma_x} \quad (25)$$

Où *l* correspond au décalage temporel, *x* au signal étudié, σ_x à son écart-type et *n* au nombre d'éléments de la chronique *x* considérés pour ce calcul (Figure II-7).

Ce calcul permet ainsi de mieux connaitre les relations temporelles s'établissant au sein du phénomène étudié.

En particulier, Mangin a montré sur des bassins karstiques pyrénéens que l'on pouvait considérer qu'un débit était corrélé à lui-même jusqu'à un certain décalage temporel qu'il a appelé « l'effet mémoire » (Mangin 1984). Cet effet mémoire représente l'inertie du système. Le seuil de non corrélation a été défini par la valeur obtenue par l'amplitude de la corrélation au décalage considéré. Sur les bassins étudiés, Mangin a défini ce seuil à 0,2.

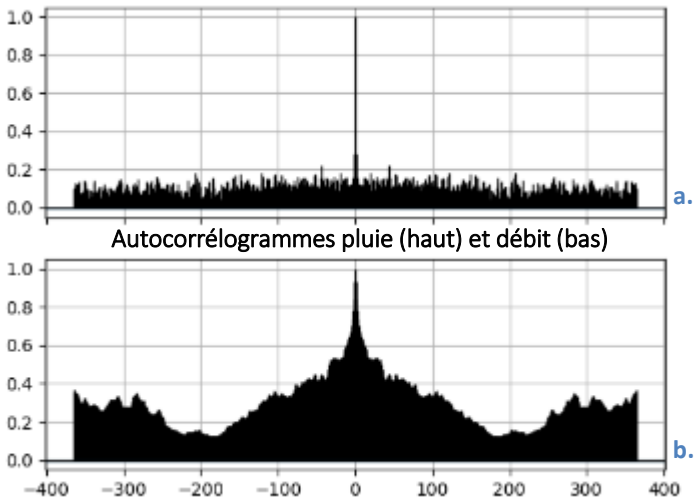


Figure II-7 - Autocorrélogrammes sur la station de Holtzwihr en 2016 des pluies (a.) et des débits (b.).

Ainsi, l'effet mémoire est obtenu en considérant le temps au-delà duquel la corrélation passe en dessous de ce seuil. Cette valeur reste empirique et dépend du nombre d'observations disponibles du bassin considéré mais nous l'avons conservé dans un premier temps par souci de simplicité.

Il est également possible, grâce à ce calcul, d'observer les cyclicités d'un signal.

Par ailleurs, la corrélation croisée détecte la coïncidence, c'est-à-dire la simultanéité, entre deux phénomènes. Elle permet ainsi d'étudier la relation entre deux signaux différents, en l'occurrence, la relation entre l'entrée d'un processus ou d'un modèle, telle que la pluie, et la sortie du processus ou du modèle, soit le débit ou la hauteur d'eau. Appliquée aux signaux temporels, la corrélation consiste à calculer le coefficient de corrélation entre deux signaux.

Grâce à cette technique corrélatoire, il est possible de définir le « temps de réponse » du bassin, c'est-à-dire le temps que met le processus, ou le système d'étude, à réagir face à une variation d'une de ses entrées. Celui-ci est défini lorsque la corrélation entre l'entrée *x* et la sortie *y* atteint son maximum (Mangin 1984). On peut aussi identifier l'« effet mémoire », tel que précédemment défini pour la corrélation simple. Le calcul de la corrélation croisée C_{xy}(l)

entre les signaux *x* et *y* est défini de la façon suivante (26), par exemple en utilisant le coefficient de corrélation linéaire de Pearson :

$$C_{xy}(l) = \frac{1}{n-l} \sum_{i=1}^{n-l} (x_i - x_{moy})(y_{i+l} - y_{moy}) \cdot \frac{1}{\sigma_x \sigma_y} \quad (26)$$

Où *l* correspond à un décalage temporel, *x* le signal de l'entrée, σ_x son écart-type, *y* le signal de la sortie, σ_y son écart-type et *n* le nombre d'éléments considérés de la chronique *x* ou *y*.

La Figure II-8 illustre comment sont déterminés le temps de réponse et l'effet mémoire à partir du corrélogramme croisé.

Ces informations sont utiles quant à la connaissance du fonctionnement d'un bassin.

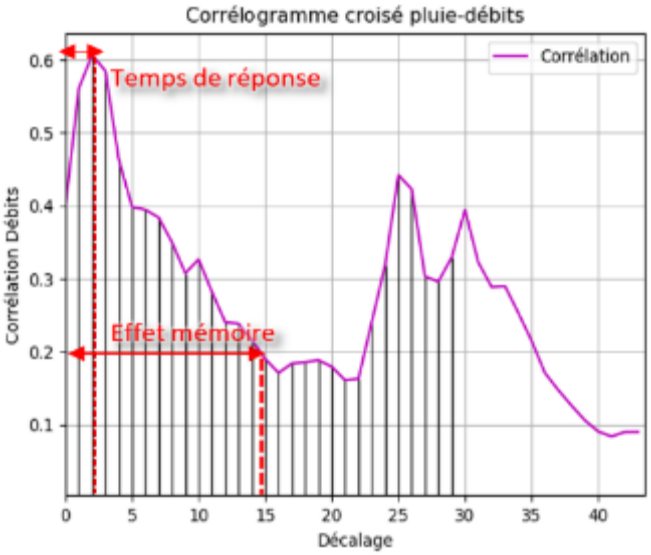


Figure II-8 – Exemple de corrélogramme croisé

III.3.5.c. Sélection des variables par corrélation croisée

Les analyses corrélatoires sont utilisées ici pour plusieurs objectifs dans le cadre de la modélisation hydrogéologique.

Dans un premier temps, elle permet de donner une indication sur les variables sélectionnables en entrée du modèle. La corrélation calculée permet ainsi de voir la relation potentielle entre deux phénomènes sur un bassin étudié.

Cependant, il est nécessaire que le modélisateur reste vigilant sur ces résultats car deux phénomènes peuvent avoir de fortes corrélations entre elles sans pour autant qu'il n'y ait de lien de causalité entre eux. Par exemple, la consommation en électricité est importante en hiver et les niveaux d'eau souterrains sont en règle générale plus importants en hiver. Il est donc possible que la corrélation entre ces deux phénomènes soit importante alors que la consommation en électricité n'engendre pas de hausse de niveaux d'eau. De même, certains phénomènes doivent être anti-corrélés. Par exemple, une forte évapotranspiration ou des températures en hausse engendrent une baisse des niveaux d'eau.

Dans un deuxième temps, une fois que les variables soient sélectionnées, il est possible d'avoir une pré-estimation de la taille des fenêtres temporelles de chaque variable d'entrée, grâce

aux temps de réponse et d'effet mémoire. Pour une estimation plus fine, il est absolument nécessaire de passer par les méthodes de régularisation décrites dans les sous-chapitres précédents. Cette pré-estimation de la taille des fenêtres temporelles permet de gagner en temps d'exécution de la validation croisée en limitant le champ des possibles.

III.4 Utilisation et conception des modèles

Lors des chapitres précédents nous avons présenté comment la mise en œuvre de réseaux de neurones nécessitait de séparer la base de données initiale en plusieurs sous-ensembles : l'ensemble d'apprentissage est utilisé lors de la phase d'apprentissage du modèle ; l'ensemble d'arrêt est utilisé pour réaliser un arrêt précoce, qui est une des méthodes de régularisation énoncée précédemment. Il est cependant nécessaire d'introduire un troisième sous-ensemble : l'ensemble de test qui est l'ensemble de données sur lequel la qualité du modèle sera, *in fine*, validée. Nous avons indiqué précédemment que l'apprentissage devait comporter une représentation de l'ensemble de l'espace des états cibles de la modélisation. L'ensemble d'arrêt est le sous-ensemble des données le plus compatible avec l'ensemble d'apprentissage ; et l'ensemble de test, que nous venons d'introduire, devrait permettre de tester le fonctionnement du modèle en relation avec l'objectif de modélisation. Si l'on s'intéresse aux crues, on prend alors comme ensemble de test l'évènement le plus représentatif, par exemple la crue la plus rapide, ou celle qui a le pic de crue le plus intense. De nombreuses études ont montré qu'en utilisant une méthodologie de conception de modèle robuste, et contrairement à ce que l'on trouve souvent écrit dans la littérature, le modèle était capable de généraliser le comportement appris à des ensemble de données extrêmes (Darras 2015, Artigue 2012, M. Toukourou 2009, Kong-a-Siou 2011).

Il n'existe pas de méthodologie universelle pour la définition de la taille des différents ensembles cités ci-dessus. Certains recommandent que l'ensemble de test représente au moins 60% de l'historique de données (Heddam 2017, Raschka 2018). Cela est expliqué par le principe de Pareto, qui explicite que seulement 20% des causes suffisent à expliquer 80% des effets. Bien que ces observations soient initialement réalisées sur des données fiscales, Juran généralise ce principe et le constate dans divers domaines (Juran 1954).

Toutefois, Guyon a démontré pour la modélisation que les résultats de modélisations sont meilleurs lorsque l'ensemble de test représente un pourcentage L_{test} égal à l'inverse de la racine carrée du nombre de paramètres libres du modèle (Guyon 1997) (27).

$$L_{test} = \frac{1}{\sqrt{N_W}} \tag{27}$$

Avec L_{test}, la taille de l'ensemble de test par rapport à l'ensemble des données en pourcentage et N_W le nombre de paramètres libres du modèle.

Or, contrairement aux domaines souvent cités pour l'utilisation de l'intelligence artificielle, l'hydrologie ne peut pas disposer de bases de données suffisamment fournies pour pouvoir appliquer ces considérations : si l'on peut augmenter à loisir les photos pour la reconnaissance d'image ou de caractères, on ne peut enrichir les bases de données temporelles qu'avec le temps. De plus, dans les applications critiques qui s'intéressent aux évènements extrêmes, il n'est pas possible d'appliquer le principe de Pareto. La durée temporelle des phénomènes peut aussi rajouter de la difficulté ; ainsi les remontées de nappe sont parfois des phénomènes très lents avec une constante de temps de plusieurs années, l'historique disponible est alors insuffisant. Il nous faut donc composer avec les bases de données disponibles et consacrer la majeure partie de cette dernière à l'apprentissage.

Ainsi après avoir affecté environ 60% des données à l'apprentissage, 20% à l'ensemble d'arrêt et 20% au test, nous avons, dans ce travail utilisé la méthodologie de conception de modèles, présentée sur le schéma II-9 suivant :

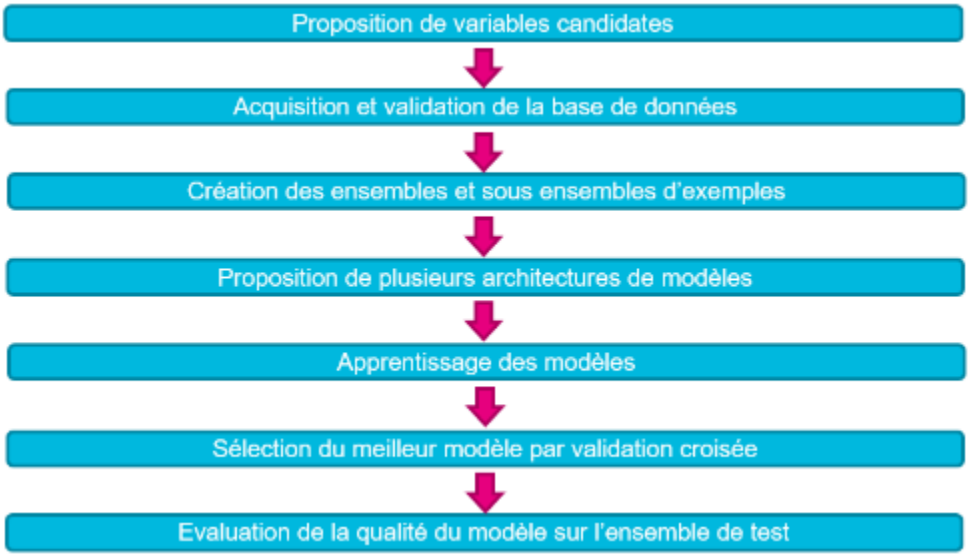


Figure II-9 Schéma de la méthodologie de conception de modèles.

Il est à noter que les étapes de proposition de plusieurs architectures de modèles et de sélection de la meilleure par la validation croisée peuvent contenir des architectures pouvant différer par :

a. Le type d'architecture, par exemple perceptron ou LSTM,
b. Les variables d'entrées utilisées
c. La dimension de l'ordre si le modèle est récurrent
d. Le nombre de couches cachées et de neurones par couches
e. Les fonctions d'évaluation des neurones

III.5 Evaluation de la performance d'un modèle

De manière générale la performance d'un modèle repose sur plusieurs critères, comme la synchronisation de pics, l'amplitude des pics, etc. L'évaluation de cette performance est rarement effectuée par un seul critère, car elle est multifactorielle.

Les principaux critères de performance sont définis ci-dessous. Beaucoup d'entre eux ont fait l'objet d'études afin de définir leurs champs d'application les plus pertinents (Waseem, et al. 2008, Moriasi, et al. 2007).

III.5.1. Evaluation de la qualité moyenne de la modélisation

III.5.1.a. Coefficient de corrélation de Pearson (*r*)

Le coefficient de corrélation de Pearson (*r*) est un indice utilisé régulièrement dans les mathématiques statistiques pour comparer deux courbes (Pearson 1896). Il peut toutefois s'appliquer, entre autres, en tant que critère de performance d'un modèle. Il s'écrit de la façon suivante (28) :

$$C_r = 1 - \frac{\sum_t^n (Y_{obs,k} - \overline{Y_{obs}})\,(Y_{p,k} - \overline{Y_p})}{\sqrt{\sum_t^n (Y_{obs,k} - \overline{Y_{obs}})^2 + \sum_t^n (Y_{p,k} - \overline{Y_p})^2}} \quad (28)$$

Cet indice est un critère de base pour attester de la qualité d'un modèle. Toutefois, l'usage de cet indice seul ne suffit pas pour attester de la qualité d'un modèle (Willmott 1981).

Avec $Y_{obs,k}$, la sortie observée au temps *k* et $Y_{p,k}$ la sortie prédite au temps *k* et $\overline{Y_{obs}}$ et $\overline{Y_p}$ leurs moyennes respectives.

III.5.1.b. Critère de Nash-Sutcliffe (C_{NSE})

Le critère de Nash-Sutcliffe (C_{NSE}), parfois raccourci en critère de Nash, est un indice permettant de quantifier la justesse d'une prévision. Ce critère est probablement le plus utilisé pour attester de la bonne qualité d'une simulation. Il varie entre -∞ et 1. Lorsqu'il vaut 0, le signal simulé par le modèle est en moyenne équivalent au signal correspondant à la moyenne de la sortie mesurée.

Un modèle est considéré comme bon lorsqu'il atteint 0,7 et très bon à partir de 0,9. Il est défini par l'équation suivante (Nash et Sutcliffe 1970) (29) :

$$C_{NSE} = 1 - \frac{\sum_t^n (Y_{obs,k} - Y_{p,k})^2}{\sum_t^n (Y_{obs,k} - \overline{Y_{obs}})^2} \quad (29)$$

Avec $Y_{obs,k}$, la sortie observée et $Y_{p,k}$ la sortie prédite.

Ce critère a beaucoup été étudié et certains ont créé d'autres critères à partir de celui-ci. C'est le cas de Mathevet qui a créé son indice à partir de celui de Nash-Sutcliffe pour minimiser l'impact des prévisions extrêmes parmi l'ensemble de la chronique (Mathevet, et al. 2006). Son indice se calcule ainsi (30) :

$$C_{2M} = \frac{C_{NSE}}{2 - C_{NSE}} \quad (30)$$

De même, Schaefli et Gupta ont soulevé le fait qu'un modèle bon en moyenne sur l'ensemble de la prévision peut être mauvais sur une plus petite échelle. Ils ont donc utilisé un critère analogue au critère de Nash, initialement proposé par Seibert, en utilisant des données de référence (31). Cela permet de ne pas tenir compte uniquement de la moyenne des sorties observées, permettant ainsi la prise en compte de phénomènes tels que la saisonnalité (Seibert 2001, Schaefli et Gupta 2007).

$$C_{BE} = 1 - \frac{\sum_t^n (Y_{obs,k} - Y_{p,k})^2}{\sum_t^n (Y_{obs,k} - Y_{ref,k})^2} \quad (31)$$

Où $Y_{ref,k}$ correspond à la sortie de référence au temps discret *k*.

De son côté, R. Moussa a montré qu'une prévision ayant un faible critère de Nash (0,3 par exemple) peut tout de même donner des informations intéressantes et être cohérente (exemple d'une simulation avec une mauvaise amplitude mais un bon décalage temporel). A l'inverse, une simulation avec un critère de Nash élevé n'est pas toujours révélatrice d'une bonne prévision (Moussa 2010).

Malgré ces remarques, le critère de Nash-Sutcliffe reste très usité pour la modélisation hydrogéologique. Cependant, il existe d'autres critères de performance attestant la qualité de la modélisation.

III.5.1.c. Critère de Kling-Gupta (KGE)

Le critère de Kling-Gupta est un critère basé sur la distance euclidienne. Celle-ci correspond à la distance entre le front de Pareto (point idéal) et la valeur de la fonction objectif (point optimal). Cette distance euclidienne **ED** peut être obtenue à partir de la décomposition du critère de Nash-Sutcliffe. En effet, Kling et Gupta ont montré que le critère de Nash peut être représenté en fonction de la corrélation α, du facteur de biais β (utilisant moyennes et variance) et la mesure de la variabilité entre l'observé et le simulé **r**. Le point idéal est obtenu lorsque $\alpha = 1$, $\beta = 0$ et $r = 1$ (Gupta, et al. 2009, Gerard 2010) (32) :

$$C_{NSE} = 2\alpha r - \alpha^2 - \beta^2{}_n \quad (32)$$

L'intérêt de ce critère est qu'il maximise la mesure de la variabilité entre les données observées et les données simulées ou prévues. Le critère varie lui aussi entre -∞ et 1, où plus il tend vers 1, plus le modèle est performant (36). Les équations ci-dessous permettent de le définir (Gupta, et al. 2009) :

$$\alpha = \frac{\sigma_p}{\sigma_{obs}} \tag{33}$$

$$\beta = \frac{\overline{Y_p} - \overline{Y_{obs}}}{\sigma_{obs}} \tag{34}$$

$$r = \frac{cov_{p,obs}}{\sigma_{obs}\sigma_p} \tag{35}$$

$$C_{KGE} = 1 - ED = 1 - \sqrt{(\alpha - 1)^2 + (\beta - 1)^2 + (r - 1)^2} \tag{36}$$

Avec σ_p et σ_{obs} respectivement les écart-types des données prédites et des données observées, Y_{moy} la moyenne et $cov_{p,obs}$ la covariance entre les chroniques observée et prédite.

III.5.2. Evaluation du décalage de la modélisation

III.5.2.a. Critère de persistance (C_P)

Contrairement aux précédents indices, le critère de persistance permet d'évaluer spécifiquement la prévision. Ce critère est fondé sur la notion de prévision naïve. Une prévision est dite naïve lorsque la valeur prévue est égale à la valeur observée à l'instant de prévision. Tout comme le critère de Nash-Sutcliffe, cet indice varie entre -∞ et 1. La prévision naïve a un critère de persistance égal à 0. Il se calcule de la façon suivante (Kitanidis et Bras 1980) (37) :

$$C_P = 1 - \frac{\sum_t^n (Y_{obs,k+h} - Y_{p,k+h})^2}{\sum_t^n (Y_{obs,k} - Y_{obs,k+h})^2} \tag{37}$$

Avec $Y_{obs,k}$, la sortie observée au temps k, $Y_{p,k}$ la sortie prévue au temps k et h l'horizon de prévision.

III.5.3. Evaluation spécifique de la modélisation d'évènements

III.5.3.a. Pourcentage de pic de crue synchrone (C_{SPPD})

Le critère de pic de crue synchrone C_{SPPD} (*Synchronous Percentage of the Peak Discharge*) s'exprime en pourcentage (Lin et Chen 2004) (38) ; il permet de quantifier si le modèle retranscrit bien l'amplitude d'une crue à l'instant où le maximum a été observé (Artigue 2012) (39) :

$$C_{SPPD} = 100 \cdot \frac{{Y_p}^{k,max} - {Y_{obs}}^{k,max}}{{Y_{obs}}^{k,max}} \tag{38}$$

$$C_{SPPD} = 100 \cdot \frac{{Y_p}^{k,max}}{{Y_{obs}}^{k,max}} \quad (39)$$

Avec ${Y_{obs}}^{max}$ la sortie observée à l'instant k_{max} où celle-ci est maximale, et ${Y_p}^{k,max}$ la sortie prédite à l'instant k_{max}.

III.5.3.b. Retard de pic de crue (C_{PD})

Le critère de retard de pic ou C_{PD} (*Peak Delay*) évalue, le retard entre le moment où la valeur maximale a été observée et le moment de la valeur maximale de la prévision. Une valeur négative indique que le signal de prévision est en avance par rapport au signal mesuré (Lin et Chen 2004) (40).

$$C_{PD} = {k_p}^{max} - {k_{obs}}^{max} \quad (40)$$

Avec ${k_p}^{max}$ le temps discret où la sortie prédite est maximale et ${k_{obs}}^{max}$ le temps discret où la sortie observée est maximale.

III.5.4. Evaluation de la surestimation de la modélisation

III.5.4.a. Critère du pourcentage de biais (PBIAS)

Le pourcentage de biais (PBIAS), introduit par Gupta et *al.* à la fin des années 1990, définit une tendance moyenne afin de savoir si les valeurs prédites sont surestimées ou sous-estimées par rapport aux valeurs mesurées. Ce pourcentage est nul lorsque qu'en moyenne il y a autant de surestimations que de sous-estimations. Si ce critère est négatif, cela signifie que le modèle a tendance à surestimer la sortie et *vice-versa* (Gupta, Sorooshian et Yapo 1999). L'équation de ce critère est la suivante (41) :

$$C_{PBIAS} = 100 \cdot \frac{\sum_t^n (Y_{obs,k} - Y_{p,k})}{\sum_t^n Y_{obs,k}} \quad (41)$$

Avec $Y_{obs,k}$, la sortie observée et $Y_{p,k}$ la sortie prédite.

III.5.5. Evaluation de la justesse de la modélisation

III.5.5.a. Critère Mean Prediction Interval (C_{MPI})

Le critère *Mean Prediction Interval* (C_{MPI}) représente la dispersion moyenne d'une prévision d'un modèle d'ensemble. Ce critère est calculé à partir des valeurs minimales et maximales pour chaque temps discret *k*. Il s'exprime avec une unité identique à celle de la sortie du modèle (Shrestha et Solomatine 2006) $C_{MPI} = \frac{1}{n}\sum_k^n (Y_{s_{max},k} - Y_{s_{min},k})$ (42)(42).

$$C_{MPI} = \frac{1}{n}\sum_{k}^{n}(Y_{s_{max},k} - Y_{s_{min},k}) \quad (42)$$

Avec $Y_{s_{min},k}$, la sortie minimale d'un modèle ensembliste au temps discret *k* et $Y_{s_{min},k}$, la sortie maximale du modèle ensembliste au temps *k*.

Il existe également une version normalisée de cet indice, le *Normalized Mean Prediction Interval Length* (C_{NMPIL}), qui est normalisé par les valeurs minimales et maximales souhaitées sur la série de données considérée (Khosravi, Nahavandi et Creighton 2010) (43).

$$C_{NMPIL} = \frac{1}{n}\sum_{k}^{n}\frac{Y_{s_{max},k} - Y_{s_{min},k}}{y_{max} - y_{min}} \quad (43)$$

Avec y_{max} et y_{min} correspondant respectivement aux valeurs maximales et minimales souhaitées sur la série de mesures considérées.

III.5.5.b. Critère Prediction Interval Coverage Probability (C_{PICP})

Le *Prediction Interval Coverage Probability Criterion* (C_{PICP}) est un indice qui dénombre le nombre moyen de valeurs de sortie incluses dans un intervalle donné (Shrestha et Solomatine 2006). Ce critère se calcule selon l'équation (44) :

$$C_{PICP} = \frac{1}{n}\sum_{k}^{n} f(Y_{p,k}) \quad (44)$$

avec $f(Y_{p,k}) = 1$ si $Y_{p,k} \in [Y_{s_{min},k}; Y_{s_{max},k}]$, sinon $f(Y_{p,k}) = 0$

Avec $Y_{s_{min},k}$, le minimum de l'intervalle au temps discret *k*, $Y_{s_{max},k}$, le maximum de l'intervalle au temps *k*, $Y_{p,k}$ la sortie calculée par le modèle au temps *k*, et *n* le nombre de valeurs considérées.

III.5.5.c. Critère Confidence Prediction Criterion (C_{PC})

Le *Confidence Prediction Criterion* (C_{PC}) est un indice quantifiant la « confiance » que l'on peut avoir dans la sortie d'un modèle. C'est le ratio entre le C_{PICP} *et le* C_{MPI} (Akil, et al. 2019) (45). Quand il est maximisé, il transcrit la capacité du modèle à délivrer le maximum de ses sorties dans l'intervalle le plus faible possible.

$$C_{PC} = \frac{C_{PICP}}{C_{MPI}} = \frac{\frac{1}{n}\sum_{k}^{n} f(Y_{p,k})}{\frac{1}{n}\sum_{k}^{n}(Y_{s_{max},k} - Y_{s_{min},k})} \quad (45)$$

avec $f(Y_{p,k}) = 1$ si $Y_{p,k} \in [Y_{s_{min},k}; Y_{s_{max},k}]$, sinon $f(Y_{p,k}) = 0$

Avec $Y_{s_{min},k}$, la sortie minimale d'un modèle ensembliste au temps discret *k*, $Y_{s_{min},k}$, la sortie maximale du modèle ensembliste au temps *k* et $Y_{p,k}$ la sortie prédite par le modèle ensembliste au temps *k*, et *n* le nombre de valeurs considérées.

IV. Conclusions du chapitre II

Ce chapitre a présenté les principes de base du fonctionnement d'un modèle à réseau de neurones. Les propriétés fondamentales du réseau de neurones, que sont l'approximation universelle et la parcimonie permettent de les utiliser pour la modélisation hydrogéologique.

Les différentes étapes de conception du modèle permettent ainsi de pallier à la difficulté du dilemme biais-variance, grâce à une sélection de variables judicieuse, à une architecture du modèle adaptée et à une application rigoureuse des méthodes de régularisation en réduisant les erreurs de variance.

Cependant, malgré une application stricte de la méthodologie décrite dans ce chapitre, le modèle génère des incertitudes sur le résultat de modélisation. L'ensemble des critères de performance mobilisés quantifient la qualité d'un résultat mais trouvent leurs limites pour estimer l'incertitude associée au résultat de modélisation.

Chapitre III : De l’incertitude à une approche de l’exactitude dans la perspective d’une modélisation hydrogéologique

I. Introduction

Ce chapitre vise à définir l'incertitude, selon différentes approches, afin de cerner les différentes composantes de cette notion. Sont détaillées les différentes sources d'incertitudes possibles affectant un résultat de modélisation, en distinguant les incertitudes inhérentes à la mesure et aux différentes variables d'entrée du modèle, les incertitudes de distribution spatiale des variables, les incertitudes liées à l'architecture du modèle et les incertitudes liées à l'initialisation et l'estimation des paramètres.

Nous observerons par la suite comment ces incertitudes sont estimées, en fonction de leurs origines. Nous détaillerons enfin une méthodologie d'estimation des incertitudes de modélisation que nous avons développée au cours de cette thèse, à partir d'éléments existants dans d'autres méthodologies et en utilisant les densités de probabilités.

II. Définition de l'incertitude

L'incertitude est l'antonyme de la « certitude ». Elle transcrit le fait qu'il existe une configuration, un cas de figure ou une possibilité, qui entraîne le non-déroulement d'une action, d'un résultat ou d'un phénomène attendu, ou de manière atténuée, la méconnaissance d'un phénomène, d'un évènement. D'après l'étude de Vold, le marqueur textuel d'incertitude le plus retrouvé dans les études scientifiques françaises est le verbe « sembler » (Vold 2006). Dans le contexte hydro(géo)logique, l'incertitude transcrit une variabilité liée à la probabilité qu'un évènement, par exemple une inondation, se produise. Cette variabilité peut être le résultat de la composition de trois facteurs qui sont : (i) le manque de connaissances ou de données, (ii) la variabilité propre des phénomènes naturels, (iii) l'intervention humaine, notamment sur la mesure (biais) et sur la prise de décision (Gouldby et Samuels 2009).

Plusieurs auteurs se sont intéressés aux liens entre incertitude et probabilité. La probabilité dite « physique », qui prend en compte la connaissance sur le phénomène, nécessite de décrire très précisément l'ensembles des causes (les entrées) et le résultat (la sortie). Elle aboutit, compte tenu de la complexité des phénomènes, à une description où chaque situation est quasiment unique, et la répétition de situations identiques est donc difficilement possible (Molina 2016). Ainsi, dans le contexte hydrologique, la fréquence de mesure d'un certain niveau d'eau diffère de sa probabilité, qui, en prenant en compte tous les facteurs qui ont conduit à ce niveau (état initial, météo), tendrait vers zéro. Il existe alors une infinité d'états possibles. En prenant en compte l'ensemble des causes, chaque situation devient alors strictement unique – même si les niveaux d'eau sont *in fine* égaux. L'approche « fréquentiste » permet de réconcilier ces deux approches en permettant de déterminer la probabilité « empirique » calculée en dénombrant l'apparition d'évènements identiques ; ainsi la

probabilité demeure un des outils qui permet le mieux de décrire l'incertitude (Willink 2009, Isukappali 1999).

L'incertitude est généralement définie à partir d'un concept générique décliné à domaine spécifique, tel que la modélisation, la mesure, les sondages, la microbiologie, etc. Dans ce manuscrit, il sera donc important de présenter les facettes générales de l'incertitude avant de la préciser à un niveau de détail plus fin. Dans notre étude, nous nous focaliserons sur les incertitudes de modélisation.

III. Origine des incertitudes et leur prise en compte

Les incertitudes de modélisation sont généralement de trois types (F. Bourgin 2014) :

- Les incertitudes sur les données, dues à l'appareillage et à la complexité des conditions environnementales, dont les erreurs dues à la résolution des mesures, en temporel ou de manière spatiale (Rodgers 1990).
- Les erreurs d'interprétation ou de traitement (exemple des courbes de tarages)
- Les incertitudes sur les modèles, soit du fait des paramètres, soit du fait de sa structure (fonction postulée plus ou moins en adéquation avec le phénomène représenté, …).

Shrestha s'est intéressé aux incertitudes affectant les modèles à apprentissage statistique. Ils distinguent, en sus des précédentes incertitudes, celles qui sont dues à :

- l'incomplétude de l'ensemble d'apprentissage ; en effet ce dernier ne peut pas représenter la totalité de l'espace des états comme cela serait nécessaire,
- au fait que les variables sont différentes en apprentissage (entrées et sortie) et en utilisation (seulement les entrées), les incertitudes sont donc différentes dans les deux configurations (Solomatine et Shrestha 2009),
- la structure du modèle, dues au fait que la fonction obtenue est *in fine* le résultat d'une méthode de minimisation itérative. Comme les paramètres ne peuvent identifier le minimum global de la fonction idéale (la fonction de régression idéale), une incertitude est générée. Cette incertitude intègre aussi celle due aux choix non optimaux des hyperparamètres (nombre de neurones cachés, pas de gradient …).
- au choix des différents ensembles d'apprentissage, de test et d'arrêt.

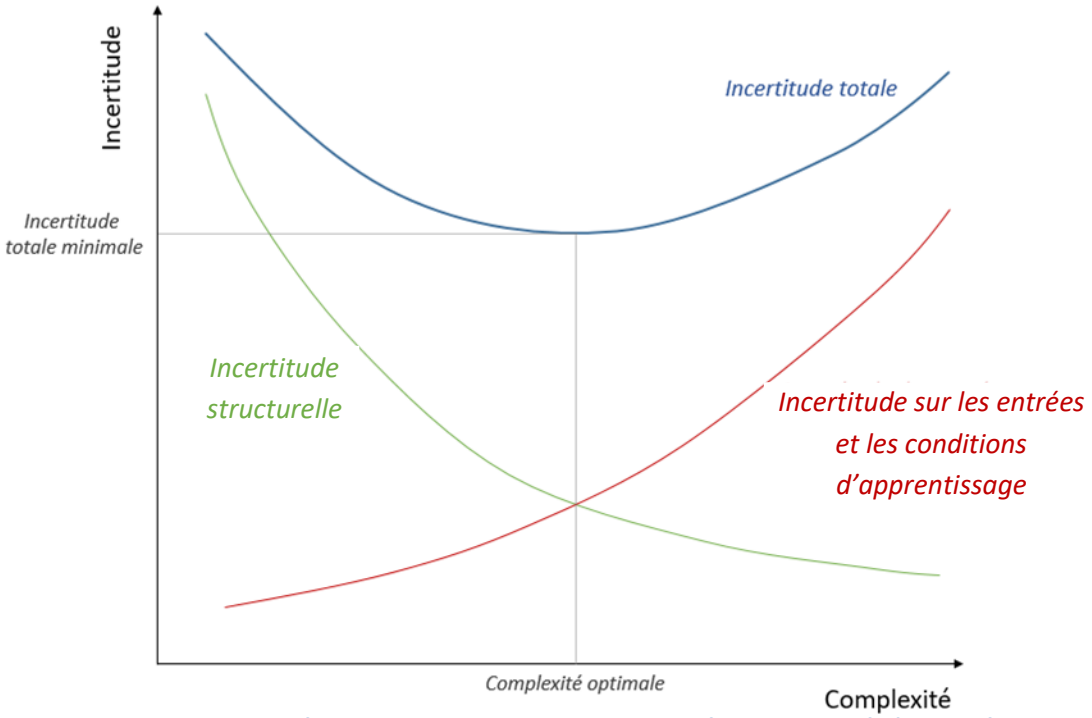


Figure III-1 - Evolution théorique des incertitudes par rapport à la complexité d'un modèle (Shrestha, 2009).

Shrestha reprend la formulation du dilemme biais variance et le représente de manière conceptuelle comme évolution de l'incertitude au fur et à mesure que la complexité du modèle augmente (Figure III-1). En partant d'une complexité faible, l'incertitude diminue principalement du fait que la structure du modèle s'adapte mieux à la fonction de régression sous-tendue par les données. Dans une deuxième partie du graphique, lorsque la complexité continue à croître, l'incertitude est alors due au bruit contenu dans les données et aux imperfections de l'apprentissage (détermination des sous-ensembles, choix des hyperparamètres et de la règle d'apprentissage, …).

Cette vision proposée par Shrestha est fondée sur l'analyse des incertitudes mais ne nous parait pas aussi simple à manipuler que celle proposée par Geman *et al.* qui nous a permis de proposer une méthode systématique de sélection de modèles (Geman, Bienenstock et Doursat 1992).

Dans la partie suivante nous allons étudier l'importance des incertitudes issues de différentes origines et comment elles peuvent être traitées.

III.1 Incertitudes de mesure

La métrologie (« science de la mesure ») est une des disciplines qui permet d'étudier les incertitudes de mesure. Celles-ci sont détaillées dans la partie suivante.

III.1.1. Vision métrologique de l'incertitude de mesure

L'incertitude correspond, selon la définition établie dans le Vocabulaire International de Métrologie (VIM), « à la dispersion des valeurs attribuées à un mesurande, à partir des informations utilisées », c'est-à-dire dans notre cas à la dispersion de la grandeur que l'on veut mesurer ou modéliser (JCGM 2008).

On distingue l'erreur de mesure, qui est à la différence entre la valeur vraie, c'est-à-dire la valeur réelle du phénomène, et la valeur mesurée (Willink 2009, JCGM 2008), de l'incertitude qui est la dispersion des mesures (Figure III-2). On distingue donc les incertitudes aléatoires des erreurs systématiques. Les erreurs systématiques génèrent ainsi toujours la même erreur et sont plus facilement quantifiables et souvent possibles à réduire ou à compenser.

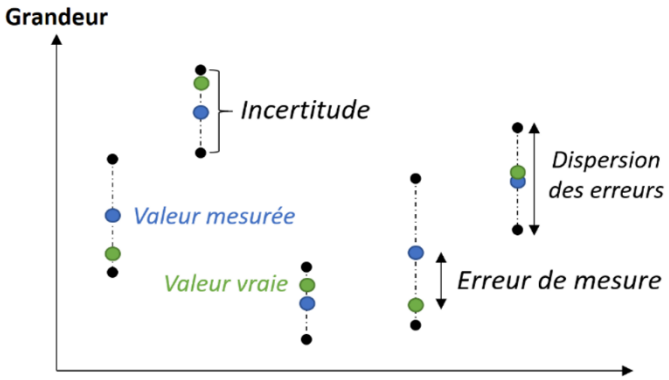


Figure III-2 - Schématisation de l'incertitude de mesure.

Considérant les incertitudes de mesures, on utilise la règle des « 5M ». Dérivée du diagramme d'Ishikawa – également connu sous le nom de diagramme de cause et effets –, cette règle explicite cinq origines de l'incertitude de mesure, représentées sur la Figure III-3 (Taillade 2005) :

- Le « Mesurande », correspondant à l'incertitude crée par les fluctuations de la grandeur physique mesurée,
- La « Méthode », correspondant aux incertitudes dues à la méthodologie et au protocole de mesure utilisés,
- Les « Moyens », où l'incertitude résulte des appareils de mesure, de leur calibration précision et de leur résolution.
- Le « Milieu », traduisant les incertitudes dues à la variabilité du milieu qui a un impact sur les appareils de mesure où la variable mesurée. Cela met en exergue l'interdépendance des phénomènes physiques ou naturels,
- La « Main d'œuvre », matérialisant que l'intervention humaine génère de l'incertitude.

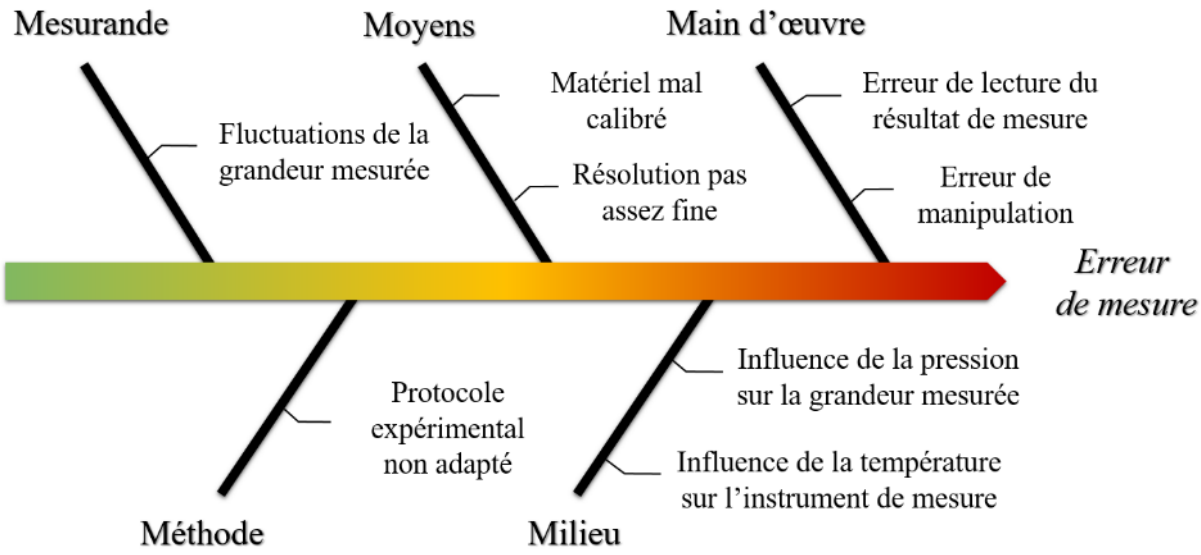


Figure III-3 - Exemple de diagramme d'Ishikawa appliqué à l'erreur de mesure (d'après Taillade, 2005).

Ces cinq sources d'incertitude affectent le résultat final. Si certaines erreurs sont difficiles à appréhender, il est possible d'estimer les erreurs de mesure selon une approche métrologique. En effet, Taillade en 2005 a expliqué dans « *Notions de Métrologie* », qui reprend les principes énoncés dans le VIM, que ces erreurs possèdent elles-mêmes plusieurs composantes, approchées selon la norme NF ISO 5725 de 1994 par deux méthodes distinctes (Taillade 2005). La méthode type A, qui évalue statistiquement les incertitudes suite à une série de mesures répétées. Elle évalue principalement deux composantes d'erreurs qui sont les erreurs de répétabilité et de reproductibilité.

En contexte hydrologique, les mesures de niveaux d'eau ou de débits sont instrumentées et font rarement intervenir plusieurs producteurs de données simultanément, ou des réplicas de mesures. Par conséquent, l'erreur de reproductibilité – qui correspond à l'erreur réalisée par plusieurs producteurs de données – et l'erreur de répétabilité – qui correspond à l'erreur d'une même mesure répétée – sont rarement estimables (Taillade 2005).

La seconde méthode, dite de type B quantifie directement des composantes d'erreurs intervenant dans l'opération de la mesure. Ces composantes comprennent les erreurs de résolution, les erreurs d'hystérésis, les erreurs de dérive, les composantes asymétriques (par exemple les erreurs générées par la forme d'un contenant), les erreurs de correction et les erreurs d'étalonnage (Taillade 2005).

La norme ISO 5725 renvoie à deux notions importantes : la justesse, qui correspond à la capacité d'un appareil ou d'une méthode à fournir en moyenne une valeur mesurée proche de la valeur vraie, et la fidélité, caractérisée par l'aptitude d'un appareil de mesure ou d'une méthode à fournir la même valeur à chaque mesure (dispersion des valeurs faible). La combinaison des deux (justesse et fidélité) engendre la notion d'exactitude (Taillade 2005) (Figure III-4).

Figure III-4 - Définition de la justesse, de la fidélité et de l'exactitude.

Les différentes erreurs composant l'incertitude liée aux données sont calculables ; on estime les erreurs suivantes (Taillade 2005) :

- l'erreur de résolution e_{res} : $e_{res} = \frac{res}{2\sqrt{3}}$, (46)
- l'erreur d'hystérésis e_{hyst} – caractérisant la dépendance d'un système à son état dans le passé : $e_{hyst} = \frac{x^{max} - x^{min}}{2\sqrt{3}}$, (47)
- l'erreur d'asymétrie e_{asym} – caractérisant la non-homogénéité de la morphologie d'un système : ($e_{asym} = \frac{d}{3\sqrt{2}}$, (48)
- l'erreur liée aux conditions environnementales e_{env} : $e_{env} = \frac{a}{\sqrt{2}}$, (49)
- l'erreur d'étalonnage e_{etal} – caractérisant les erreurs d'un appareil de mesure par rapport à un étalon ou à un appareil de référence : $e_{etal} = \frac{EMT}{\sqrt{3}}$, (50)
- les erreurs de corrections et les erreurs de dérive,

avec res la résolution de l'appareil de mesure, x la variable considérée, d la longueur de la base de l'asymétrie, a la variable environnementale influençant la mesure (par exemple la température dans le cas de la conductivité) et EMT l'Erreur Maximale Tolérée, indiquée sur les spécifications de l'appareillage.

Au total, l'erreur globale de mesure est définie comme étant la racine carrée de la somme du carré des erreurs, conformément au Vocabulaire International de Métrologie.

Les erreurs de mesure, du fait des variations des différents types d'erreurs en fonction des conditions de mesure contribuent fortement aux incertitudes de mesure

III.1.2. Prétraitements et qualité des données

Outre les erreurs de mesures présentées précédemment, plusieurs anomalies peuvent être identifiés. Considérons tout d'abord les valeurs aberrantes. Une valeur aberrante est selon F. Grubbs « une observation qui semble s'écarter nettement des autres membres de l'échantillon dans lequel elle se produit » (Grubbs 1969). Ces valeurs aberrantes peuvent être générées par la variabilité aléatoire inhérente au jeu de données ou être un résultat suite à une modification de procédure expérimentale, de calcul ou d'enregistrement de la valeur numérique (Grubbs 1969). Ces valeurs aberrantes nécessitent une correction (valeurs de pluie négatives par exemple) ou leur suppression si la correction n'est pas envisageable.

Il existe aussi des « lacunes », c'est-à-dire l'absence de mesure. Dans certains cas, l'utilisateur peut utiliser des données sans combler ces lacunes. Mais généralement l'utilisateur va chercher à combler les manques. Ceci peut être fait par le biais d'interpolation, régressions linéaires et/ou autres calculs statistiques et géostatistiques.

Pour certains types de données, il est possible de combler les lacunes autrement que par des régressions linéaires. C'est notamment le cas pour les données pluviométriques, où il est possible d'appliquer la méthode des « doubles cumuls ». Cette méthode part de l'observation que les cumuls pluviométriques de deux séries de mesures, dans une unité régionale, sont proportionnelles. Ce constat permet de déterminer des données pluviométriques manquantes sur une station donnée, à partir des données pluviométriques d'une station voisine ayant des caractéristiques similaires (altitude de la station, qualité de l'instrument de mesure, absence de micro-climat, …), en comparant les cumuls des deux stations (Artigue 2012, Sebbar 2013, Brunet-Moret, et al. 1986) sommées selon les équations (51) et (52)$P_{i,j} = z_j P_i + e_{i,j}$ (52) :

$$P_{C_{i,j}} = \sum_{r=1}^{j} P_{i,r} \quad (51)$$

$$P_{i,j} = z_j P_i + e_{i,j} \quad (52)$$

Avec j l'instant où la donnée pluviométrique est souhaitée, i la station pluviométrique utilisée, $P_{C_{i,j}}$ la pluie cumulée à la station *i* jusqu'au temps *j*, z_j un indice régional de pluie pour l'instant j et $e_{i,j}$ des résidus.

III.1.3. Incertitudes rencontrées sur les grandeurs hydrologiques

Les grandeurs hydrologiques généralement utilisées pour la modélisation hydrogéologique et leurs incertitudes associées sont décrits dans les parties suivantes.

III.1.3.a. Précipitations

Les précipitations sont difficiles à observer. En effet, comme l'a énoncé Hiez, les précipitations sont en fait un champ, elles varient donc dans l'espace et le temps, comme par exemple un champ électrique. Il est donc nécessaire d'observer la pluie à la bonne échelle d'espace (Hiez 1977) et de temps. Usuellement, la pluie est mesurée de deux façons : à l'aide de pluviomètres au sol et de radars météorologiques (Darras 2015). Lorsque les observations des pluies sont faites avec des pluviomètres, ces derniers réalisent une mesure ponctuelle des précipitations tombées sur un territoire (Hiez 1977). Par exemple, dans les Cévennes, il a été montré que des pluviomètres pourtant distants entre eux d'une dizaine de kilomètres, peuvent mesurer des hauteurs de pluies très différentes, d'un facteur 3 à 5, notamment sur des cumuls observés lors d'épisodes pluvieux (Artigue 2012). Les radars météorologiques permettent d'avoir une meilleure vision du champ des précipitations sur un bassin, mais s'ils représentent convenablement les variations relatives d'un pixel à l'autre. En revanche, ils représentent mal la quantité tombée. Ainsi, Visserot a étudié la quantification des hauteurs d'eau tombées en Cévennes à Anduze et a comparé la valeur mesurée par un pluviomètre au sol et la valeur estimée par la réanalyse COMEPHORE de Météo France pour le pixel du radar situé à l'aplomb du pluviomètre. Sur une vingtaine d'évènements intenses une différence moyenne de 40% est notée sur le cumul de précipitations observées par le radar et celui mesuré par le pluviomètre (Visserot 2012).

Déployer une démarche d'observation des précipitations nécessite donc d'avoir un préalable une connaissance des caractéristiques régionales de ces dernières et une bonne connaissance des capacités de l'instrumentation.

III.1.3.a.i. Pluviométrie au sol

La pluviométrie au sol est le plus souvent mesurée à l'aide d'un pluviomètre à augets basculants. Ce dernier est composé d'un auget, qui, une fois rempli, bascule sous l'effet du poids de l'eau, représentant ainsi une hauteur d'eau bien définie (en général de 0,2 mm). Les précipitations sont ainsi obtenues ponctuellement au niveau de la station. Lorsque le réseau de pluviomètres n'est pas assez dense pour rendre compte de variations locales de pluies, cette appréciation ponctuelle de la pluie génère des incertitudes sur la distribution spatiale des pluies. Hors incertitude de distribution spatiale, les incertitudes de mesure par un pluviomètre, sont le plus souvent estimées à 10 %, et jusqu'à 20 % en cas de cumuls sur des évènements très pluvieux (Marchandise 2007).

III.1.3.a.ii. Radars météorologiques

Les radars météorologiques mesurent les précipitations quand elles sont encore dans l'atmosphère. Pour cela, ils envoient des ondes par impulsions et mesurent les échos, ou réflectivité. Il est ensuite difficile d'en déduire la quantité, la position et le déplacement des

précipitations du fait de la non-linéarité de la relation réflectivité-pluie, et des aberrations nombreuses qui peuvent se produire, par exemple s'il y a de la glace dans la masse d'eau. Les précipitations sont ici obtenues sur une large zone géographique et la résolution spatiale est usuellement le km², sur une étendue de 100 km² (Marchandise 2007).

Dans un contexte montagneux où les reliefs sont prononcés, ces reliefs peuvent générer des échos ou des masques, modifiant ainsi la mesure des précipitations. De plus, le faisceau radar peut être dévié à cause un changement d'indice de réfraction de l'air – causée par l'humidité de l'air par exemple –, d'atténuation des ondes électromagnétiques ou à cause du type de précipitations (Marchandise 2007). Toutes ces sources d'incertitudes peuvent être approchées par des méthodes géostatistiques (Cecinati, et al. 2017). Afin d'avoir des données de pluies plus proches de la réalité, il est nécessaire de coupler les données issues de radars avec celles issues de pluviomètres.

III.1.3.b. Débits et hauteur d'eau en rivière

Les débits et hauteurs d'eau dans les cours d'eau superficiels sont deux variables essentielles, souvent utilisées comme objectif de modélisation. En général, les débits ne sont pas mesurés directement car il est plus aisé de mesurer la hauteur d'eau. Celle-ci est alors convertie en débit à l'aide d'une courbe de tarage. Les courbes de tarage sont des abaques, générées par des jaugeages manuels sur une gamme de hauteurs d'eau, permettant de déterminer la valeur d'un débit à partir d'une hauteur d'eau mesurée.

Les hauteurs d'eau en rivière sont mesurées souvent à partir d'une échelle limnimétrique. Cette échelle peut être lue par un observateur ou par une caméra fixant l'échelle. Les hauteurs d'eau peuvent également être mesurées à l'aide d'un capteur de pression (Darras 2015) ou d'un laser. Dans les cours d'eau dont le lit n'est pas calibré de manière artificielle, la courbe de tarage est sujette à des évolutions au cours du temps du fait de la modification du lit de la rivière : atterrissement, érosion des berges, ou obstacles divers intervenant fortuitement. Il est donc nécessaire de recalibrer régulièrement cette courbe de tarage. En période de crue, lorsque des débordements ont lieu, il est alors extrêmement difficile de réaliser la courbe de tarage. Marchandise a estimé en 2007 que les incertitudes sur la détermination du débit sont autour de 20 % et peuvent atteindre 30 % pour les débits de crue des épisodes cévenols (Marchandise 2007).

III.1.3.c. Evapotranspiration et données climatiques

D'autres types de données météorologiques sont également utilisés pour la modélisation hydrogéologique, notamment ceux influençant l'évapotranspiration, comme la température, l'humidité relative, l'ensoleillement, la force du vent et la pression.

En cas d'indisponibilité d'accès aux données, l'évapotranspiration potentielle (ETP) peut être utilisée. Dans les modélisation statistiques, celle-ci est par exemple estimée selon une sinusoïde dite *a priori* en fonction de la position de l'année hydrologique dans l'année climatique et de l'amplitude de l'évapotranspiration potentielle (Johannet, Taver, et al. 2012). Cette dernière peut également être calculée à partir des données de températures (Thornthwaite 1948). Il existe d'autres expressions plus complexes, prenant en compte d'autres paramètres tels que la température de l'air, l'humidité relative maximale, la vitesse du vent, les radiations solaires, et des constantes physiques, permettent d'avoir un résultat plus précis, telle que la formule de Penman-Monteith (53) (Monteith 1965).

$$ETP = \frac{\Delta_{h_r} R_n + \rho_{air}\, c_{p_{air}} \delta P g_{air}}{(\Delta_{h_r} + \gamma\left(1 + \frac{g_{air}}{g_s}\right) L_v)} \quad (53)$$

Avec Δ_{h_r} la variation de saturation de l'humidité relative selon la température de l'air, R_n l'irradiance nette, $c_{p_{air}}$ la capacité thermique de l'air, ρ_{air} la densité de l'air sec, δP le déficit de pression de vapeur, g_{air} la conductivité hydraulique de l'air, g_s la conductivité des stomates et γ la constante psychrométrique.

Le calcul de l'évapotranspiration reste délicat comme en témoignent les résultats parfois contradictoires de certaines études (Djaman, Balde, et al. 2015, Djaman, Tabari, et al. 2016, Ndiaye, et al. 2017, Oudin 2004). Par ailleurs, il faut noter que le choix du modèle dépend de la localisation du site d'étude, l'évapotranspiration étant influencée par le couvert végétal, le climat, le relief ou encore la latitude. Cela illustre la complexité du calcul de l'évapotranspiration et suggère que les incertitudes qui lui sont liées sont difficiles à estimer.

III.1.3.d. Niveaux piézométriques

Le niveau piézométrique correspond, selon le dictionnaire français de l'hydrogéologie, au niveau d'eau supérieur d'une colonne liquide statique équilibrant la pression hydrostatique. Il est matérialisé par le niveau libre de l'eau au sein d'un tube ouvert disposé verticalement – appelé piézomètre (Castany et Margat 1977). Il matérialise ainsi localement le niveau de la nappe d'eau souterraine.

Plusieurs sources d'incertitudes existent sur la mesure de ce champ de niveau. Une des sources d'incertitudes est la représentativité des mesures piézométriques. Par exemple, si le piézomètre se trouve à proximité d'un cours d'eau, les niveaux d'eau souterraine peuvent être influencés par les hauteurs d'eau en rivière. Ces influences très locales peuvent être plus ou moins importantes, en fonction de la structure des sols, des contextes climatiques et hydrologiques, de la distance avec les cours d'eau superficiels ainsi que des obstacles et diverses infrastructures entre le cours d'eau et le piézomètre (Darras 2015).

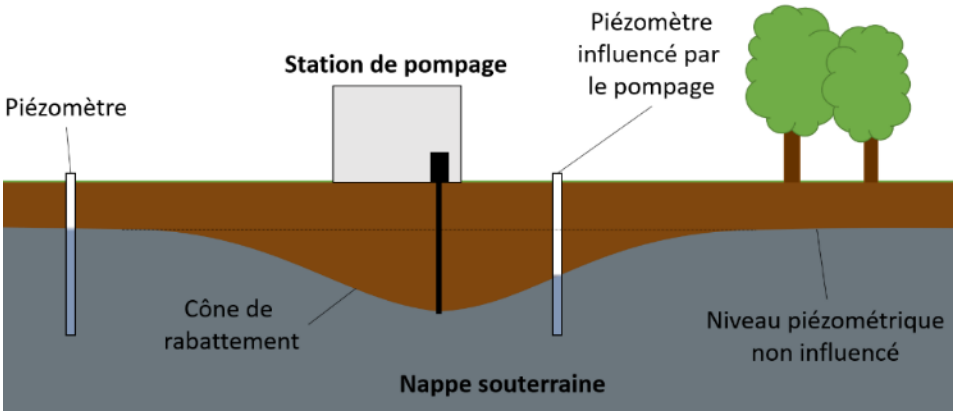


Figure III-5 - Schéma de l'influence d'un cône de rabattement causé par un pompage.

Les pompages souterrains – qu'ils soient pour un usage agricole, industriel ou pour l'adduction en eau potable – influent sur le niveau d'eau des piézomètres localisés à l'intérieur du périmètre du cône de rabattement (Figure III-5). Ce cône se forme lorsque les pompages sont plus importants que la recharge, créant une dépression hydrogéologique autour de la zone de pompage. Ce cône peut être plus ou moins déformé en fonction de l'homogénéité des caractéristiques physiques du sol (transmissivité, coefficient d'emmagasinement, …).

Une autre origine des incertitudes concerne directement la mesure des niveaux d'eau souterrains, selon l'appareillage utilisé. Les piézomètres sont le plus souvent équipés de capteur de pression, déterminant ainsi la hauteur d'eau, ou de capteurs avec un flotteur et un contrepoids reliés entre eux. Les incertitudes sur ces mesures restent relativement faibles, mais pas toujours négligeables. Pour une mesure à 25 mètres de profondeur, le capteur de pression a une précision de 1,25 cm (source : OTT) tandis que le capteur muni de flotteur est moins précis, avec une précision à 3,2 cm (source : OTT). Il existe également des capteurs à ultrasons, mesurant ainsi le niveau de l'eau par rapport au tube mais qui sont un peu moins fiables compte tenu des échos des ultrasons dans le piézomètre. Ils disposent d'une précision de 3,8 cm pour une mesure de 25 mètres de profondeur (source : OTT). Pour les données les plus anciennes, ainsi que pour les rares piézomètres n'étant pas supervisés de manière automatique, les niveaux piézométriques sont parfois mesurés manuellement à l'aide d'une sonde lumineuse. Les données mesurées ainsi sont alors moins précises que l'utilisation du capteur de pression et leurs qualités sont dépendantes de la bonne manipulation de l'utilisateur. Avec une erreur de manipulation réduite, l'utilisation d'une sonde lumineuse permet d'obtenir une précision d'environ 2,5 cm pour une mesure de 25 mètres de profondeur (source : OTT), soit 0,1%.

III.1.3.e. Prélèvements en eau et irrigation agricole

Les prélèvements en eau peuvent être judicieusement utilisées en modélisation. Toutefois, ces données sont difficilement accessibles, selon le bassin d'étude et la granularité temporelle souhaitée. La Banque Nationale des Prélèvements en Eau (BNPE) fournit des données

annuelles de prélèvements en eau et les classe en différents secteurs (énergie, adduction en eau potable, agriculture, industrie). Il est toutefois difficile d'acquérir des données avec une granularité temporelle plus fine et ceci nécessite un rééchantillonnage à des pas de temps plus petits, engendrant une incertitude sur les données.

Les prélèvements en eau comprennent les quantités d'eau utilisées pour l'irrigation agricole. Ces données sont déclarées auprès des DDT et DDTM (Direction Départementale des Territoires et de la Mer) pour des volumes prélevés supérieurs à 1 000 $m^3.an^{-1}$. Ces données sont difficiles à obtenir et peu adaptées à l'utilisation en modélisation hydrologique. Il est toutefois possible de les estimer en se basant sur les besoins en eau des cultures présentes sur un bassin et de la proportion de chaque type de culture présent sur ce même bassin. Chaque type de culture a des besoins en eau connus et nécessite des arrosages à certaines périodes en quantités bien définies. Il est alors possible de créer une courbe d'irrigation potentielle annuelle, qui peut être intégrée dans un modèle. Cette technique accentue toutefois les incertitudes liées aux entrées du modèle, puisqu'il s'agit de séries temporelles estimées et reconstituées à partir d'informations partielles.

III.1.3.f. Données complémentaires

D'autres types de données peuvent alimenter le modèle, telles que l'évapotranspiration réelle (ETR), la pluie efficace, ou encore les données d'éclusées de barrages hydroélectriques en rivière, de rejets de stations d'épurations dus à une densification de la population urbaine sur une partie du territoire d'étude, d'une perméabilité des sols favorisant le ruissellement ou non, ou d'échanges entre eaux superficielles et eaux souterraines. Cependant, ces données sensibles ne sont pas toujours disponibles.

La neige peut également être une entrée d'un modèle hydrogéologique selon le contexte du bassin (Riboust 2019). En hiver, les précipitations sont stockées sous forme de neige et peuvent générer des crues en surface à la fonte des neiges. Ce phénomène a pu être constaté sur le bassin du Baget, en Ariège dans les Pyrénées, pour la modélisation des débits (Johannet, Taver, et al. 2012).

Pour les études sur des sites côtiers, la marée peut être utilisée comme entrée ou sortie d'un modèle neuronal (Adib 2008).

Il est possible également d'utiliser la conductivité comme entrée d'un modèle car elle peut être caractéristique d'une arrivée d'eau (superficielle ou souterraine), tout comme la salinité pour les milieux halieutiques ou encore la turbidité, pour les échanges entre les eaux superficielles et souterraines (Savary 2018).

Quelles que soient les données utilisées, l'essentiel est de s'assurer de leur caractère explicatif vis-à-vis de la sortie que l'on souhaite modéliser. Pour cela, il est possible de se baser sur la

connaissance physique, hydrologique ou encore hydrogéologique du système ; on parlera alors de sélection experte de variables. Il est possible de recourir aux analyses statistiques et notamment les analyses corrélatoires afin de limiter la quantité de données au strict nécessaire.

III.2 Incertitudes environnementales

Les conditions physiques ont une influence sur la mesure. Par exemple, une pression atmosphérique élevée aura tendance à exercer une pression plus importante sur la surface du niveau d'eau dans le piézomètre et donc à abaisser le niveau de la nappe souterraine au niveau de la station, influençant ainsi le niveau de la nappe souterraine du fait de variations non purement hydrologiques (Rasmussen et Crawford 1997).

La température peut également influencer la mesure d'un phénomène en agissant directement sur le capteur, en dilatant plus ou moins celui-ci et modifiant ainsi légèrement la mesure. Cela concerne notamment les mesures de conductivité, de turbidité ou de salinité, qui utilisent des plaques métalliques où transitent des particules chargées. Plus ces plaques sont dilatées par la température, plus le nombre d'ions captés sera important, augmentant ainsi la valeur mesurée. La température agit également sur la taille des pluviomètres par le même phénomène de dilatation. Le réceptacle recevant les pluies, parfois en métal et donc susceptible d'être dilaté par la température, ne fait pas exactement la même taille en hiver et en été compte-tenu de l'écart de température et modifiant ainsi les mesures de pluies.

Le vent, quant à lui, créé des vagues en rivière ou en mer, modifiant ainsi la valeur des hauteurs d'eau, et par conséquent des débits (Touboul 2007). Le vent peut également avoir une incidence sur les précipitations, du fait de leur inclinaison. Cela modifie la quantité de précipitations captées par un pluviomètre au sol (Morell 1986).

Concernant les appareils de mesure, le niveau d'incertitude peut en partie être encadré par des valeurs fournies par le constructeur et par des calculs métrologiques complémentaires. Pour les phénomènes naturels (vagues, vent, rebond des gouttes sur le bord du pluviomètre, …), ces incertitudes peuvent être importantes, mais sont plus difficiles à quantifier.

III.3 Incertitudes dues aux prétraitements

Des incertitudes peuvent être générées en fonction de la manière dont les variables utilisées pour la modélisation hydrogéologique sont prétraitées. En effet, il peut s'agir de données brutes obtenues à la sortie d'un capteur, de données prétraitées en fonction des tendances précédentes ou d'abaques suite à des lacunes dans les séries temporelles, mais aussi de données diverses servant à calculer un autre type de variable, comme pour l'évapotranspiration par exemple. Ainsi, pour ce dernier point, le choix de la formule pour le calcul de ces données engendre des résultats différents car ne dépendent pas des mêmes

variables et génère ainsi de l'incertitude sur la donnée calculée. De même, la méthode utilisée pour combler les lacunes dans les chroniques joue un rôle sur l'incertitude liée à la donnée.

Un des prétraitements obligatoires pour l'utilisation d'un réseau de neurones est la normalisation des entrées et des sorties ; celle-ci consiste souvent à centrer et réduire la variable par son écart-type. Lorsque l'on simule les valeurs, cela n'a pas d'impact, mais dès lors que l'on travaille en prévision cela peut introduire un biais. En effet, une fois la prévision effectuée, si l'on effectue une dénormalisation centrée-réduite pour retrouver les bonnes grandeurs, il est nécessaire d'avoir la valeur moyenne et l'écart-type des valeurs que l'on veut prédire avant de réaliser ce calcul. Or, la valeur prédite est indisponible par définition, et elle peut potentiellement être une valeur extrémale et influencer la moyenne, comme l'écart type. Pour résoudre ce problème, des valeurs de références, correspondant le plus souvent aux valeurs observées dans le passé, sont utilisées, engendrant ainsi des biais dans le cas où la prédiction réalisée n'a pas les paramètres de dénormalisation identiques à ceux qu'aurait la série observée dans son intégralité.

III.4 Incertitudes de distribution spatiale

III.4.1. Discrétisation spatiale

La distribution spatiale des données en entrées d'un modèle est sources d'incertitudes. En particulier, beaucoup d'études se sont focalisées sur la distribution spatiale des données pluviométriques, mais avec une variabilité spatiale géologique, de relief, ou d'usage des sols également très importante. Ainsi, en fonction des sites d'étude les résultats de certaines études sur l'importance de cette source d'incertitude peuvent diverger (Troutman 1983, Le Lay et Saulnier 2007, Garçon 1999). Il est donc difficile de généraliser ces conclusions, ces incertitudes sont fortement dépendantes du site d'étude.

III.4.2. Polygones de Thiessen

Les incertitudes sur les données d'entrées et de sortie peuvent être causées par une mauvaise distribution spatiale des capteurs. Pour remédier à ce problème, Thiessen, en s'inspirant et appliquant les travaux de Voronoï portant sur la météorologie a proposé une méthode permettant de quantifier l'influence d'un pluviomètre sur un secteur. La méthodologie des polygones de Thiessen permet de définir une aire d'influence pour chaque station pluviométrique. Cette aire, rapportée en pourcentage en divisant par la superficie totale du bassin versant d'étude, permet ensuite le calcul de la pluviométrie moyenne, selon l'équation suivante (54) (Thiessen 1911) :

$$P_{moy} = \frac{\sum_i P_i A_i}{A} \quad (54)$$

Avec P_{moy} les précipitations moyennes sur l'ensemble du bassin versant, P_i les précipitations tombées sur l'aire A_i, A_i l'aire d'influence correspondant à un pluviomètre i et A la superficie du bassin versant.

Un des inconvénients principaux de cette technique est que les aires sont délimitées par des segments, sans gradient de l'influence des pluviomètres à la frontière des polygones. De plus, C. Damant *et al*. ont montré en 1983 que cette méthode peut engendrer de forts écarts entre la pluie calculée et la pluie observée par radar, pouvant dépasser 300 % d'erreurs. Ces fortes erreurs sont principalement dues au type d'évènement de précipitations – un fort orage local et bref engendre plus d'incertitudes que des précipitations réparties sur plusieurs heures (Damant, et al. 1983).

Afin de pallier ce problème, il est possible d'intégrer plusieurs pluviomètres comme entrées du modèle. Ainsi, pour un exemple d'un bassin ayant deux pluviomètres, il est possible d'intégrer en entrée :

- les données d'un des deux pluviomètres,
- les deux chroniques des deux pluviomètres,
- la moyenne (ou la somme) pondérée des deux pluviomètres,
- la différence des pluies date à date de ces deux pluviomètres,
- la variance de cette différence sur *n* pas de temps, …

L'ajout d'une ou plusieurs des possibilités apportées ci-dessus pour un bassin contenant deux pluviomètres permet ainsi de réduire les incertitudes liées à la distribution spatiale des données. Toutefois, il existe une méthodologie permettant d'estimer cette incertitude.

III.4.3. Détermination des incertitudes par permutations d'entrées

Dans ce travail, nous avons proposé d'étudier les incertitudes liées à la méconnaissance de la distribution spatiale des variables d'entrées en les perturbant à l'entrée du modèle. Appliquée de concert avec la validation croisée, présentée au chapitre précédent, il est alors possible de quantifier la dispersion des prévisions lorsque les entrées d'un même type de donnée sont interverties.

Afin d'illustrer cette méthodologie, supposons un modèle pluie-débit comportant deux entrées de pluie, mesurées sur deux sites différents. La configuration 1, indiquée dans le Tableau III-1, correspond au modèle pluie-débit avec les séries de données initiales, telles que mesurées. La validation croisée est effectuée pour déterminer la taille des fenêtres temporelles minimisant le score de validation croisée (Kong-A-Siou, Johannet et Borrell, et al. 2011). Une simulation est ensuite réalisée afin d'obtenir un premier résultat. Une permutation et duplication des entrées est ensuite effectuée, en conservant une architecture de modèle identique. Ainsi, pour la permutation/duplication 2, l'entrée A ayant sa propre fenêtre

temporelle, reçoit les données du pluviomètre B. A l'inverse, l'entrée B reçoit les données du pluviomètre A. Une nouvelle simulation est donc réalisée selon cette nouvelle configuration, donnant un nouveau résultat de modélisation. Cela est ensuite répété pour toutes les permutations/duplications listées sur le Tableau III-1. Pour *n* entrées de précipitations à perturber, on obtient donc n^n possibilités, cette méthode n'est donc applicable que pour peu d'entrées à permuter et dupliquer.

Tableau III-1 - Combinaisons possibles pour des permutations de deux entrées.

Permutations et duplications	Entrée A	Entrée B
1	A	B
2	B	A
3	A	A
4	B	B

L'ensemble des résultats de modélisation forment un ensemble de modèles. Il est donc possible de déterminer un intervalle correspondant à l'incertitude de modélisation par permutations des seules entrées. Par comparaison avec les incertitudes sans permutation des entrées de pluie, on peut donc en déduire l'augmentation des incertitudes dues à la variabilité spatiale des pluies.

Il a été observé que l'utilisation de plusieurs entrées de pluies issues de plusieurs stations augmente l'amplitude des valeurs des jeux de paramètres estimés par rapport à ceux obtenus suite à l'utilisation une à une de ces mêmes entrées (Chaubey, et al. 1999). Il est donc aussi possible d'étudier les incertitudes de distribution spatiale en permutant les paramètres reliés à ces entrées de pluies, sous réserve que le nombre de paramètres soit identique pour chacune des entrées.

III.5 Incertitudes de représentativité temporelle

La longueur des historiques de données utilisée pour l'apprentissage est une des sources d'incertitude. En effet, un apprentissage avec des données mesurées sur une courte période (2-3 années) est moins représentatif des comportements de l'hydrosystème d'étude qu'un apprentissage avec des historiques de données très grands, sur plusieurs décennies. Cela est dû aux caractéristiques de signaux hydrologiques : non stationnaires, non gaussiens et affectées de valeurs extrêmes. La durée minimale de mesures est dépendant du pas de temps utilisé. Par exemple sur l'on dispose de données au pas de temps mensuel, une dizaine d'années d'historique ne confère que 120 valeurs, ce qui est insuffisant pour prendre en compte les potentiels cycles pluriannuels des diverses variables d'entrées. À notre connaissance, il n'existe pas de travaux étudiant les impacts de la durée d'historique en fonction du pas de temps, de la variable hydrologique suivie et des caractéristiques du site d'étude. Ce travail reste à réaliser. De plus, le pas d'échantillonnage des données a une importance sur les incertitudes, car il ne faut pas que les données soient sous-échantillonnées car l'on perdrait alors une partie de l'information.

L'utilisation de la médiane s'avère un choix judicieux car cette statistique est moins sensible aux valeurs extrêmes que la moyenne (De Nijs et Klausen 2013). Pour les précipitations, les cumuls sont indispensables, réduisant par intégration l'incertitude possible.

Lorsque l'on utilise la validation croisée, le découpage des données en sous-ensembles joue un rôle sur l'incertitude car les scores de validation croisée sont calculés à partir de ces sous-ensembles. Ils ont alors une influence sur la sélection de l'architecture du modèle. Ce sont alors les choix arbitraires de découpage des évènements qui introduisent des incertitudes.

De même, la sélection des différents ensembles (apprentissage, arrêt et test) génère de l'incertitude sur le résultat de modélisation final.

III.6 Incertitudes de modélisation

III.6.1. Approches d'ensemble

L'approche d'ensemble, ou approche multi-modèles, utilise plusieurs modèles, qui peuvent être de nature ou de typologies différentes, et s'exécutent sur un même jeu de données, délivrant au modélisateur des sorties qui peuvent être exploitées en même temps. On parle alors de modèle d'ensemble et chaque modèle délivrant une sortie est un membre de cet ensemble. Grâce à cet ensemble les différentes sorties transcrivent autant de possibilités qui peuvent être utilisée pour traduire une incertitude (Arsenault 2015).

Cette technique peut également être utilisée sur un type de modèle, par exemple un réseau de neurones d'architecture définie mais qui permet de générer autant de modèles que l'on effectue d'initialisation différentes des paramètres de cette architecture (Darras 2015).

Il est possible également d'estimer l'incertitude de chaque modèle à partir d'un modèle de référence (*Skill Score* et score de Brier (Mylne, Evans et Clark 2002, Brier 1950)).

III.6.2. Incertitudes liées à l'initialisation et l'estimation des paramètres

III.6.2.a. Présentation de l'incertitude sur les paramètres

L'incertitude liée aux paramètres vient de leur initialisation et de leur calcul au cours du processus d'apprentissage (F. Bourgin 2014). Même si Kuzcera définit l'origine de ce type d'incertitude dans un contexte hydrogéologique, comme étant l'effet de la difficulté d'inférence exacte des différents paramètres lors de l'ajustement des modèles (Kuczera 1983), liée à l'équifinalité, il existe différentes méthodologies pour quantifier les incertitudes associées. Selon Bourgin (F. Bourgin 2014), on peut utiliser les approches bayésiennes, les approches post-traitement, les analyses de sensibilité.

III.6.2.b. Quantification des incertitudes

Parmi les approches utilisées dans la littérature pour quantifier les incertitudes, on trouve :

- les méthodes bayésiennes, qui permettent la détermination des erreurs de modélisation selon leurs origines.

Par exemple, la méthode BATEA, pour *Bayesian Total Error Analysis*, est une approche bayésienne s'appliquant aux modèle conceptuels pluie-débit et qui permet de mieux identifier les sources d'erreurs s'appliquant sur les paramètres (Thyer, et al. 2009). Les incertitudes de mesures peuvent également être approchées par cette méthode (Kavetski, Kuczera et Franks 2006), parfois en la couplant avec des algorithmes de Monte-Carlo (Motra, Dimmig-Osbürg et Hildebrand 2013). Cependant, ces approches bayésiennes sont limitées aux modèles conceptuels (Kavetski, Kuczera et Franks 2006).

- les analyses de sensibilité, qui permettent de faire varier les valeurs des paramètres du modèle.

Une des méthodes les plus communément connues, et s'appliquant à des modèles ayant peu de paramètres est la méthode LPSA pour *Local Parameter Sensitivity Analysis*. Cette technique effectue des variations de la valeur d'un seul paramètre à la fois et quantifie les variations de la sortie qui y sont associées. Ces variations sont alors interprétées comme reflétant l'incertitude liée à l'estimation des paramètres. Le calcul de la sensibilité, noté S^{Y}_{W}, est ensuite effectué (55) (Vanlier, et al. 2013) :

$$S^{Y}_{W_t} = \frac{Y(W_t + \Delta W_t)}{Y(W_t)} \cdot \frac{W_t}{\Delta W_t} \quad (55)$$

Avec Y la sortie calculée, W_t le paramètre considéré et ΔW_t les variations du paramètre considéré.

Ces méthodes restent longues à mettre en place, puisqu'elles font varier un à un les paramètres du modèle afin d'analyser les modifications induites sur les résultats de modélisation.

- les approches post-traitement, qui s'appuient sur des modèles dits « déterministes » préalablement calés, et analyse les différences entre les sorties des modèles et les chroniques mesurées (F. Bourgin 2014).

Il existe parmi ces approches post-traitement la méthode UNEEC, pour *Uncertainty Estimation based on local Errors and Clustering*. Cette méthode s'appuie sur la différence entre les valeurs de sortie simulées par le modèle et les données de sortie mesurées pour quantifier l'incertitude globale du modèle (Solomatine et Shrestha 2009). La méthode UNEEC est séquencée en trois étapes :

- *Clustering* : les valeurs des variables d'entrée sont ici regroupées en *clusters*, selon des caractéristiques définies au préalable (par exemple les débits de hautes eaux, les débits d'étiage, …),
- Estimation de la distribution de probabilité *P* d'erreur de chaque cluster, obtenue à partir d'un modèle déterministe **M**,
- Construction du modèle $\mathbf{U^p}$ à partir de la fonction de la distribution des probabilités d'erreur obtenue empiriquement (Solomatine et Shrestha 2009).

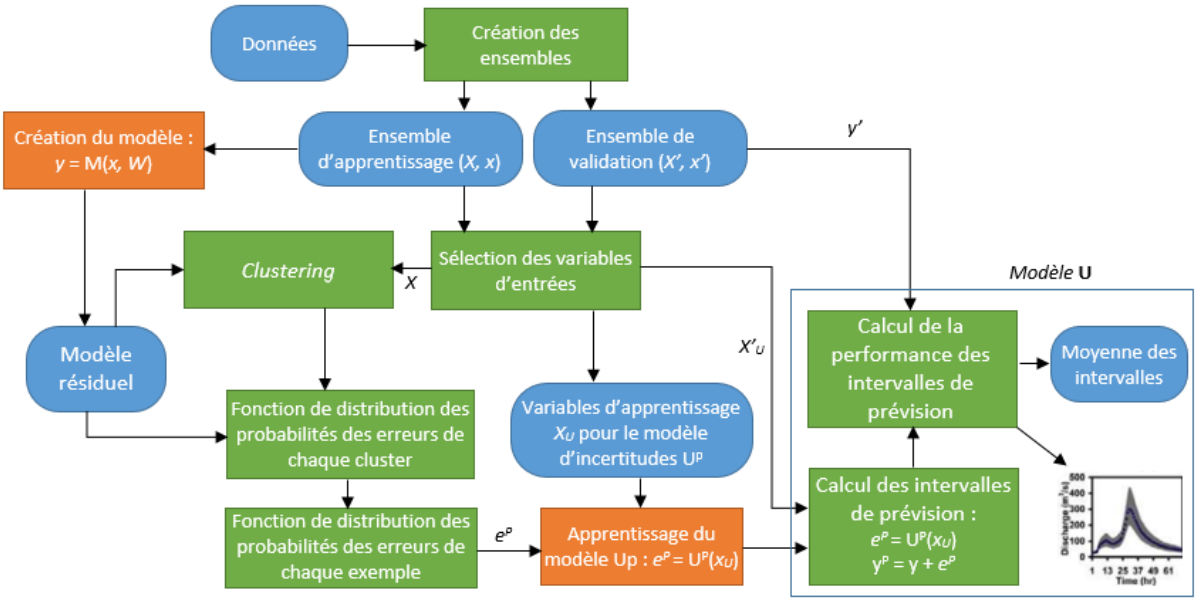


Figure III-6 - Schéma de fonctionnement de la méthodologie UNEEC (d'après Solomatine & Shrestha, 2009).

Plusieurs modèles sont ainsi utilisés en fonction de la probabilité *P* souhaitée pour construire les modèles $\mathbf{U^p}$ (Figure III-6) (Solomatine et Shrestha 2009, D. Shrestha 2009).

D'après Bourgin, ces approches soufrent de ne pas pouvoir définir l'origine des incertitudes (F. Bourgin 2014).

III.6.2.c. Méthodes d'échantillonnage aléatoire

Les méthodes d'échantillonnage aléatoire sont des approches d'estimation de l'incertitude régulièrement utilisées. Elles se basent sur des algorithmes de Monte-Carlo, décrits en Annexe III.

III.6.2.c.i. La méthode Generalized Likelihood Uncertainty Estimation (GLUE)

La méthode GLUE (pour *Generalized Likelihood Uncertainty Estimation*) part de l'observation de l'équifinalité : plusieurs jeux de paramètres d'un même modèle permettent d'obtenir des sorties semblables (Beven et Binley 1992).

La méthode GLUE se réalise en trois étapes. La première est une phase de caractérisation qui consiste à définir des modèles supposés être aptes à modéliser « convenablement » le fonctionnement du système d'étude. Ainsi, les variables d'entrées, la fonction des densités de probabilité et la structure des modèles sont choisies afin de construire des jeux de paramètres.

La seconde étape correspond à la phase de simulation. Ici, chaque jeu de paramètres sélectionné est utilisé dans le modèle, de façon aléatoire grâce à des algorithmes tels que la méthode de Monte-Carlo.

La troisième étape est la phase d'évaluation. Une estimation de la vraisemblance de chaque jeu de paramètres est réalisée à partir d'une fonction GLUE, dans le but de sélectionner les meilleurs d'entre eux à l'aide d'une valeur seuil. Cette valeur seuil, définie arbitrairement par l'utilisateur, dépend de la fonction GLUE utilisée et sert à définir le caractère « satisfaisant » d'un modèle. La prévision réalisée par chaque jeu retenu est ensuite pondérée pour pouvoir restituer une prédiction avec des intervalles, estimant ainsi l'incertitude.

La fonction GLUE, qui est à l'origine une fonction de vraisemblance, notée $\mathcal{L}$, peut s'écrire différemment selon les objectifs voulus. Elle se base le plus souvent sur la performance d'un modèle, sur la somme des carrés des résidus (Beven et Binley 1992, Beven et Binley 2014, F. Bourgin 2014).

P. Vallam et *al*. ont appliqué cette méthodologie sur des modèles hydrologiques pour quantifier les incertitudes liées à l'estimation des paramètres (Vallam, Qin et Yu 2014). Parmi trente paramètres, ils ont pu isoler ceux qui généraient le plus d'incertitude grâce à cette méthode. Ils ont toutefois évoqué l'absence de prise en compte des incertitudes liées à l'architecture du modèle sur leur méthode.

Autre inconvénient de cette méthode, elle est coûteuse en temps d'exécution car elle nécessite un grand nombre de jeux de paramètres pour fonctionner.

III.6.2.c.ii. La méthode Automatic Calibration and Uncertainty Assessment Using Response Surfaces (ACUARS)

La méthode ACUARS (pour *Automatic Calibration and Uncertainty Assessment Using Response Surfaces*) est une méthode similaire à la méthode GLUE. Elle se distingue en limitant le nombre de jeux de paramètres par des valeurs seuils et des poids d'importance. Pour cela, elle analyse les sorties du modèle, générées par des ensembles de paramètres qui satisfont les données d'apprentissage (Mugunthan et Shoemaker 2006).

Elle utilise le *cell-declustering*, une méthode géostatistique qui permet de découper une zone d'intérêt en un certain nombre de cellules équivalentes qui contiennent des *clusters* de données équiprobables (Deutsch 1989). Les paramètres sont ici ajustés à partir de facteurs de

corrections calculés en utilisant cet algorithme et de manière itérative, générant à chaque fois un modèle. A l'instar de la méthode GLUE, les modèles conservés sont ceux qui atteignent un score seuil. L'incertitude sur les paramètres du modèle est alors définie sur ces modèles.

III.6.3. Définitions et inférence statistique

Les probabilités peuvent aider à approcher les incertitudes. Classiquement on distingue l'approche Bayésienne qui utilise la connaissance des phénomènes pour établir des probabilités quantifiant la croyance que l'on peut avoir qu'un évènement se produise, de l'approche fréquentiste qui quantifie après l'expérience la fréquence « a posteriori » de l'apparition de l'évènement. Ainsi, la probabilité *a posteriori* peut se calculer de la façon suivante (56).

$$P(A) = \frac{n_{pos}}{nb_{expe}}, \tag{56}$$

avec A un évènement aléatoire quelconque, $P(A)$ sa probabilité calculée, n_{pos} le nombre de cas positifs ou réussis et nb_{expe} le nombre total d'expériences réalisés.

Dans notre étude, l'approche des probabilités « fréquentistes » sera préférée, car ces probabilités partagent une notion commune avec la modélisation hydrogéologique, qui est la mesure des expériences. Grace au calcul de probabilités on peut procéder à des inférences statistiques, qui consistent à inférer (induire) les propriétés statistiques d'une population à partir d'un échantillon de cette population.

III.6.4. Fonction de vraisemblance

La fonction de vraisemblance, notée $\mathcal{L}$, a été introduite par Fisher en 1922. Cette fonction évalue l'adéquation entre les paramètres d'un modèle de loi de probabilité donnée et les observations mesurées de réalisations d'une variable aléatoire. Par exemple si l'on suppose que nos expériences permettent de mesurer les réalisations indépendantes d'un phénomène qui suit une loi Normale $\mathcal{N}(\bar{x},\ {\sigma_x}^2)$, on a pour chaque expérience (57) :

$$f(x|\bar{x},\ {\sigma_x}^2) = \mathcal{N}(\bar{x},\ {\sigma_x}^2) = \frac{1}{\sigma_x\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2{\sigma_x}^2}}, \tag{57}$$

avec f la fréquence d'apparition des évènements, $\mathcal{N}$ à la loi normale de paramètres $\bar{x}$ et ${\sigma_x}^2$, respectivement l'espérance mathématique de x, et son écart type au carré.

Comme les évènements sont indépendants, les probabilités d'obtenir la suite des valeurs observées est le produit des probabilités de chacune des valeurs, soit (Fisher 1922) :

$$\mathcal{L}(x_1 \dots x_n; \theta) = \prod_{i=1}^{n} f(x_i\ |\bar{x}, {\sigma_x}^2) = \frac{1}{2\pi{\sigma_x}^2} e^{-\frac{\sum_i (x_i-\bar{x})^2}{2{\sigma_x}^2}} \tag{58}$$

On peut alors prendre le logarithme de cette expression, ce qui transforme les produits en somme et chercher son maximum en annulant sa dérivée. C'est ce qu'on appelle la méthode du maximum de vraisemblance (59). On résout donc :

$$\frac{\partial}{\partial \boldsymbol{\theta}} ln\, \mathcal{L}(x_1 \dots x_n; \theta) = 0 \qquad (59)$$

Avec $\boldsymbol{\theta}$ le vecteur des paramètres de la loi de probabilité, dans cet exemple, la loi normale.

Et l'on peut ainsi « inférer » la valeur du vecteur des paramètres $\boldsymbol{\theta}$.

III.6.5. Densités de probabilités

III.6.5.a. Définition

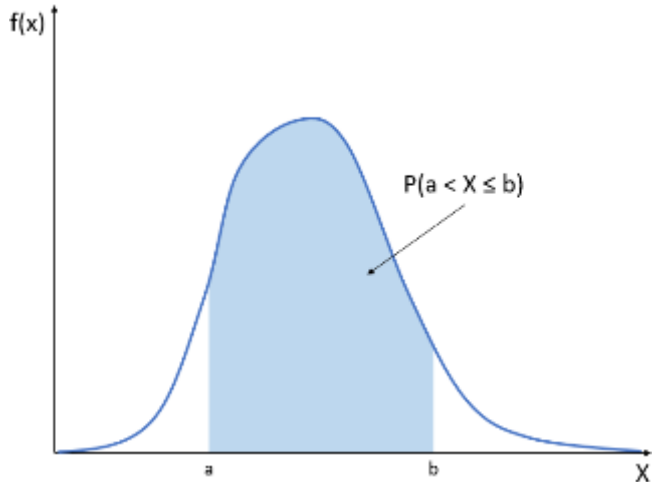


Figure III-7 - Distribution de la densité de probabilité.

La densité de probabilité, ou *pdf* – pour *Probability Density Function* –, correspond à une fonction définissant l'ensemble des valeurs de la distribution des probabilités.

On calcule la probabilité pour qu'une variable X soit comprise dans un intervalle $[a\,;\,b]$ en intégrant cette densité de probabilité (Figure III-7). La probabilité s'écrit alors de la façon suivante (60) :

$$P(a < X \leq b) = \int_a^b f(x)\, dx \qquad (60)$$

Avec a et b les limites de l'intervalle et $f(x)$ la fonction de densité de probabilité. Par définition des probabilités, ces fonctions sont toujours positives et leur intégrale vaut 1. Lorsque les densités de probabilités sont issues de mesures, nous utilisons le terme de densité de probabilité empirique, pour les différentier des densités de probabilité théoriques, par exemple celle de la loi normale (57). Il est possible également de déterminer les densités de probabilité cumulatives (*cdf* pour *Cumulative Density Function*), qui permettent de sommer la probabilité pour chaque valeur de X.

Les principales lois à densité de probabilité théoriques que nous avons utilisées dans ce travail sont explicitées en annexe, avec une présentation de chaque loi, une explication sur ses applications, la formule mathématique de la fonction et une représentation graphique de la loi standardisée.

III.6.5.b. Algorithme d'Espérance-Maximisation : comment déterminer les densités de probabilités mixtes

À l'issue d'un processus de mesure, il arrive que la densité de probabilité obtenue ne puisse être associée à une seule loi de probabilité connue. On peut donc chercher à représenter cette densité empirique sous forme de la superposition pondérée de plusieurs densités théoriques élémentaires, dont on va chercher à connaitre les vecteurs de paramètres respectifs et les coefficients de pondération qui les lient, pour représenter au mieux la densité de probabilité empirique.

L'algorithme d'espérance-maximisation (EM) est un algorithme itératif qui permet d'obtenir les différents paramètres des lois élémentaires de même type en utilisant les maximums de vraisemblance (Dempster, Laird et Rubin 1977). Cet algorithme est donc utilisé dans ce travail pour obtenir les paramètres des fonctions de densités de probabilité qui vont être composées, ainsi que leurs coefficients de pondération.

Cet algorithme se calcule en deux étapes distinctes. La première est l'étape d'espérance ou d'écriture du logarithme de la vraisemblance des réalisations, en suite à l'équation (61) :

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\theta}) = \sum_{i=1}^{n} ln\, f(x_i, \boldsymbol{\theta}) \qquad (61)$$

Avec $\boldsymbol{\theta}$ le vecteur des paramètres à déterminer, $\mathbf{x}$ le vecteur des observations x_i suivant la loi f à paramétrer et $\mathcal{L}$ la vraisemblance, *n* le nombre de mesures.

Dans un premier temps, les paramètres $\boldsymbol{\theta}$ de la loi supposée f et d'un vecteur $\mathbf{z}$ dont on ne sait à quelle distribution l'associer, sont initialisés. Tant que l'algorithme n'a pas convergé, l'espérance est évaluée (62):

$$E\left[\mathcal{L}(\mathbf{x}, \boldsymbol{\theta})\middle|\boldsymbol{\theta}^{(c)}\right] = E\left[\mathcal{L}((\mathbf{x}, \mathbf{z}), \boldsymbol{\theta})\middle|\boldsymbol{\theta}^{(c)}\right] - E\left[\sum_{i=1}^{n} ln\, f(z_i|(x_i, \boldsymbol{\theta}))\middle|\boldsymbol{\theta}^{(c)}\right] \qquad (62)$$

Avec c l'itération courante, $E\left[\mathcal{L}(\mathbf{x}, \boldsymbol{\theta})\middle|\boldsymbol{\theta}^{(c)}\right]$ l'espérance de la vraisemblance de **x**, $E\left[\sum_{i=1}^{n} \ln f(z_i|(x_i, \boldsymbol{\theta}))\middle|\boldsymbol{\theta}^{(c)}\right]$ l'espérance prise sur $\mathbf{z}$ et $E\left[\mathcal{L}\big((\mathbf{x}, \mathbf{z}), \boldsymbol{\theta}\big)\middle|\boldsymbol{\theta}^{(c)}\right]$ l'espérance de la vraisemblance sur **x** et **z** : l'ensemble des observations complété par **z**.

La deuxième étape est l'étape de maximisation. Pour cela, la vraisemblance est maximisée selon l'équation (63) suivante :

$$\boldsymbol{\theta}^{(c+1)} = arg\max_{\theta} E\left[\mathcal{L}\left((x,z),\boldsymbol{\theta}\right)\middle|\boldsymbol{\theta}^{(c)}\right] \quad (63)$$

Où $arg\max_{\theta}$ est le vecteur des paramètres $\boldsymbol{\theta}$ qui maximisent l'expression. L'itération c est ensuite incrémentée et les deux étapes sont répétées jusqu'à ce que les paramètres $\boldsymbol{\theta}$ soient optimaux (Dempster, Laird et Rubin 1977).

IV. Proposition d'une méthodologie d'estimation de l'incertitude de modélisation

IV.1 Incertitudes identifiées dans ce travail

Nous avons vu au début de ce chapitre que la notion d'incertitude sur une mesure correspond à la dispersion de cette mesure. Cette approche peut également être utilisée pour décrire l'incertitude sur les sorties d'un modèle. Un des moyens les plus directs pour accéder à cette dispersion est de créer un modèle d'ensemble dont les membres sont générés de manière à représenter la variabilité due au paramètre ou à l'hyperparamètre qui nous intéresse. Il est donc également possible, en utilisant un modèle d'ensemble, de cibler plus particulièrement les incertitudes qui dépendent d'un facteur spécifique. Dans les parties précédentes de ce chapitre nous avons identifié les sources d'incertitudes susceptibles d'avoir un impact sur la modélisation des niveaux de nappe. Parmi ces dernières, celles qui semblent avoir un impact majeur sont : (i) les prélèvements effectués qui sont très souvent non mesurés mais estimés de manière indirecte, (ii) l'évapotranspiration qui est aussi estimée de manière indirecte par un modèle, (iii) la mesure des variables hydrométéorologiques qui dépend du capteur considéré et dont la variabilité est souvent plus importante concernant les précipitations que les débits ou les hauteurs d'eau. A ces incertitudes difficiles à réduire, s'ajoutent des incertitudes dues à des prétraitements mal effectués ou à des échantillonnages, spatial ou temporel, inadéquats. Le volet concernant les prétraitements ne sera pas pris en compte Par la suite ; en revanche nous nous intéresserons à la variabilité spatiale des précipitations. S'il est clair que son rôle est majeur pour les eaux de surface, elle peut également générer une incertitude notable sur les eaux souterraines (Savelyeva, et al. 2010). Celle-ci sera estimée en permutant les entrées de mesure de pluies ponctuelles, comme énoncé précédemment. Cette technique peut être apparentée à une méthode de sensibilité particulière, car les entrées sont ainsi perturbées. L'avantage de cette méthode est que l'amplitude de la variation n'est pas arbitraire mais représente bien la variabilité du phénomène, telle qu'observée.

Le second facteur que nous prendrons en compte est directement lié à la technique de modélisation, il s'agit du calcul des paramètres du modèle.

Les incertitudes sur l'initialisation et l'estimation des paramètres du modèle sont les incertitudes les plus étudiées dans la littérature, que ce soit sur des modèles à base physique

(Arsenault 2015, Kuczera 1983, Mazzilli 2011) que sur des modèles de type réseaux neuronaux (Blundell, et al. 2015, Perreault Levasseur, Hezaveh et Wechsler 2017, Secchi, Zio et Di Maio 2008). Les méthodologies pour estimer les incertitudes liées aux paramètres sont souvent longues en temps d'exécution et en temps de traitement de données post-exécution.

Concernant les réseaux de neurones, on peut construire un intervalle de prévision en construisant un modèle d'ensemble dont les membres ne diffèrent que par l'initialisation de leurs paramètres avant l'apprentissage. Cette technique a été expérimentée par Darras et *al.* (Darras, Johannet, et al. 2014), par Kong-A-Siou *et al.* (Kong-A-Siou, Johannet et Borrell Estupina, et al. 2015), ainsi que par Kabir et *al.* qui utilisent cette méthodologie pour construire un intervalle de prévision, et estime l'incertitude à partir de différents critères de performance (Kabir, et al. 2018). L'avantage de cette technique est qu'elle est moins gourmande en données et par conséquent en temps d'exécution. Elle est donc privilégiée pour la suite des travaux. De plus, cette méthodologie nécessite peu de calculs supplémentaires et est facilement applicable et à mettre en œuvre dans un processus d'industrialisation.

Les deux autres sources majeures d'incertitudes : l'évapotranspiration et les prélèvements anthropiques n'ont pas été étudiés spécifiquement. Leur impact a été diminué en prenant en compte le modèle non récurrent, c'est-à-dire le modèle qui estime la hauteur d'eau future à partir de la hauteur d'eau actuelle (5)$\hat{y}(k) = \varphi(y(k-1), \dots, y(k-r); \boldsymbol{u}(k), \dots, \boldsymbol{u}(k-n_r+1), \mathbf{W})$ (5). En effet, Taver *et al.* ont montré que ce modèle était capable de prendre en compte les non stationnarités affectant le fonctionnement du bassin (Taver, et al. 2015).

IV.2 Incertitudes de modélisation par densité de probabilité

IV.2.1. Détermination du nombre d'initialisations optimal

IV.2.1.a. Sélection du modèle

Comme indiqué à la section II.4., il est nécessaire, préalablement à la sélection du modèle de connaitre suffisamment le bassin d'étude pour sélectionner les variables candidates et entrée et en sortie.

Suivant la méthodologie que nous avons explicitée à la section II.4, ces variables sont appliquées au modèle et le jeu de variables le plus pertinent est déterminé grâce à la validation croisée, qui nous permet de sélectionner le modèle ayant la plus faible variance.

Au préalable, nous avons effectué des analyses corrélatoires afin de pré-dimensionner la taille des fenêtres temporelles de chaque type d'entrée, puis des variations ont été appliquées autours de ces hyperparamètres pour ajuster plus finement la fenêtre temporelle conduisant au meilleur score de validation croisée (Kong-A-Siou, Johannet et Borrell, et al. 2011). Ce dernier a été calculé en utilisant le critère de persistance pour bien évaluer les capacités du modèle en prévision.

La sélection de l'architecture du modèle s'effectue en fonction des scores de validation croisée, lorsque ceux-ci correspondent au critère de persistance C_P et au critère de Nash-Sutcliffe C_{NSE} selon la Figure III-8. Sur cette figure, chaque point correspond aux critères de Nash et de persistance d'un modèle différent.

Le critère de persistance C_P a été privilégié au critère de Nash car la prévision hydrogéologique est plus souvent sensible à la persistance qu'à la qualité du modèle.

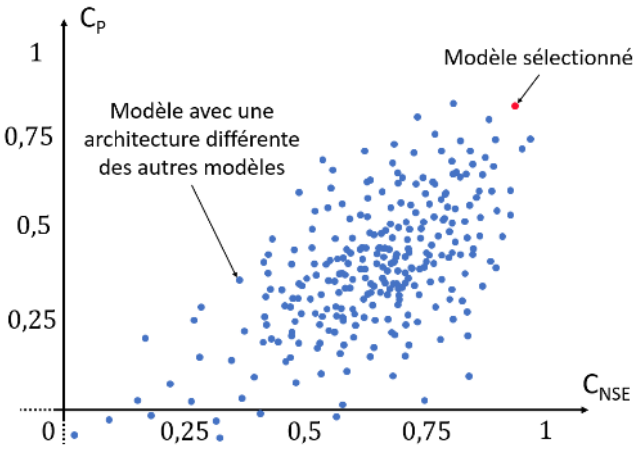


Figure III-8 - Sélection de l'architecture par validation croisée.

Afin d'augmenter la robustesse du processus de sélection de modèle, vis-à-vis de l'initialisation des paramètres, nous avons appliqué la validation croisée sur des modèles d'ensembles à chaque étape de la sélection. Pour cela, des ensembles de modèles sont constitués, où chaque membre de l'ensemble a la même architecture, mais une initialisation des paramètres différente. Pour réduire l'influence des membres « anormaux », la médiane des scores de validation croisée est utilisée.

IV.2.1.b. Nombre d'initialisations optimal

Une fois la sélection de l'architecture faite, nous revisitons le choix du nombre de membres de l'ensemble afin de représenter au mieux la dispersion des prévisions. Notons qu'un premier choix du nombre de membres a été fait pour réaliser la sélection du modèle ; ce qui était important à cette étape précédente était de représenter correctement la médiane de l'ensemble afin d'effectuer la sélection d'un modèle robuste ; dans cette seconde phase, il s'agit de représenter la dispersion des prévisions. Pour ce faire, une fois le modèle déterminé, une nouvelle validation croisée est réalisée avec des modèles d'ensemble constitués d'un nombre de membres croissant.

À partir de l'évolution du score de validation croisée en fonction du nombre de membres des ensembles, il est possible de définir le nombre de membres nécessaires. Des travaux similaires ont été réalisés par Darras et *al*. Sur la prévision du pic de crue par le C_{SPPD}, en fonction du nombre de membres (Darras, Johannet, et al. 2014). De même, Savary a réalisé des travaux similaires en se focalisant sur les écart-types de modélisation (Savary 2018) (Figure III-9).

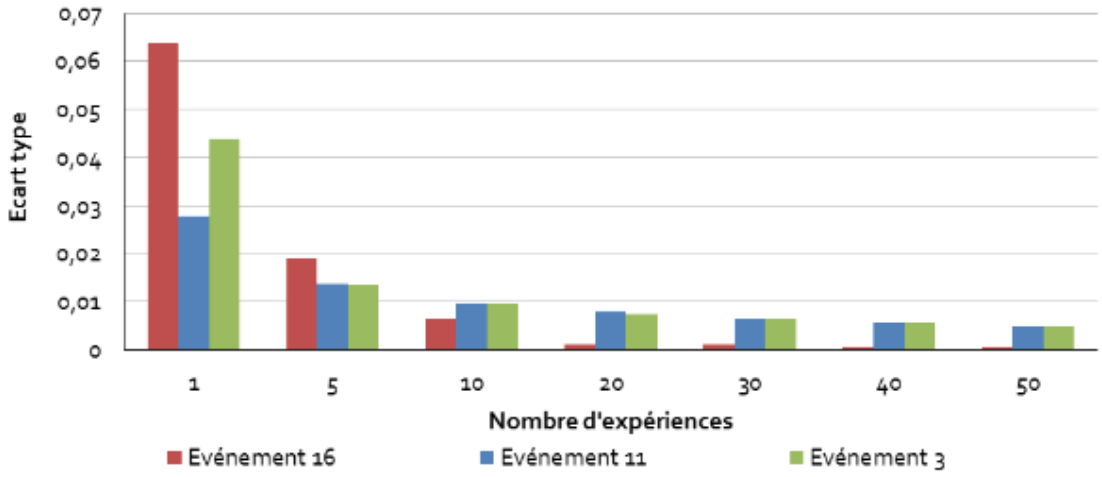


Figure III-9 - Ecart-type des sorties modélisées en fonction du nombre d'initialisations (Savary, 2018).

Logiquement, plus le nombre de membres augmente, plus l'intervalle de prévision augmente jusqu'à une certaine amplitude maximale. Le nombre de prévisions à l'intérieur de cet intervalle augmente puis se sature lorsque l'intervalle est si large qu'il englobe toutes les sorties possibles.

En pratique le score de validation croisée choisi est le *Confidence Prediction Criterion* (C_{PC}), présenté à la section II.5, et qui traduit la confiance que l'on peut avoir dans un modèle pour délivrer des sorties peu variables et dans le bon intervalle.

IV.2.2. Application des densités de probabilités

Pour déterminer des incertitudes de modélisation, nous avons adopté une démarche fréquentiste, qui permet de déterminer un intervalle de prévision avec un indice de confiance associé. Cette méthode fait appel aux densités de probabilité et se réalise en plusieurs étapes. Ces étapes sont décrites ici :

1) Choix de la taille des classes de hauteur d'eau de manière à assurer que chaque classe ait suffisamment d'échantillons. L'intérêt ici est d'avoir un bon compromis entre un nombre suffisant d'échantillons pour limiter les erreurs statistiques et la résolution des hauteurs d'eau prévues.

2) Calcul de la fréquence d'apparition de chaque classe. Cette approche « fréquentiste » permet ainsi de déterminer la densité de probabilité empirique f des données modélisées, l'équivalent d'une *pdf* (64) ; la normalisation suivante est réalisée :

$$P(X = c_i) = \frac{n_{x \in c_i}}{n_{total}} \quad (64)$$

Avec c_i la classe de hauteur d'eau, n le nombre d'apparitions dans cette classe, et x la valeur de la variable et n_{total} le nombre total d'échantillons.

3) Recherche d'une loi de distribution de probabilité, basée sur une superposition de lois connues, représentant convenablement la densité de probabilité empirique précédemment tracée.

Une dizaine de lois de probabilité ont pu être étudiées : la loi Normale, la loi de Gumbel, la loi de Laplace, la loi du Cosinus surélevé, la loi de Cauchy, la loi Logistique, la loi de Kumaraswamy, la loi de Slash, la loi de Bhattacharjee et la loi de Huber. Ces lois ont été présélectionnées parmi près d'une trentaine de lois, qui ont pu être étudiées. Les autres lois – dont la loi Uniforme – ayant une inférence très mauvaise avec la distribution des niveaux piézométriques observés ne sont pas présentées dans ce manuscrit. Toutes les lois à densités de probabilité précitées ont été testées et retenues ; elles sont présentées en Annexe I.

L'algorithme d'espérance-maximisation (EM) est utilisé pour effectuer du clustering, il permet de calculer la composition de lois théoriques élémentaires de manière à s'ajuster au mieux à la distribution empirique. Cet algorithme permet ici de retrouver les moments d'ordre 1 (moyenne) et d'ordre 2 (variance) de chaque composante de la distribution observée, ainsi que le pourcentage à attribuer à chacune des composantes. À partir de ces moyennes et variances, les paramètres de chaque loi à densité de probabilité de chaque composante sont déterminés et les lois à densité de probabilité sont calculées en sommant chacune des lois des composantes.

Afin de mesurer l'adéquation entre les lois à densités de probabilités et la distribution empirique des données, le coefficient de corrélation de Pearson (r^2) et le test de Kolmogorov-Smirnov ont été utilisés. Inventé en 1933, le test de Kolmogorov-Smirnov permet d'attester si la fonction représentant la distribution des valeurs d'un échantillon de données suit une loi à densité de probabilité. Ce test fonctionne en comparant les densités de probabilités cumulées (*cdf*) (Stephens 1992). Celui-ci renvoie la distance maximale entre les deux fonctions cumulées, selon l'équation suivante (65) :

$$D_n = sup_x |F_n(x) - F(x)| \quad (65)$$

Avec $F_n(x)$ la *cdf* empirique de l'échantillon étudié, $F(x)$ la *cdf* de la loi à densité de probabilité testée, D_n la distance.

Plus ce score tend vers zéro, plus la loi de probabilité utilisée, avec ses paramètres propres, peut être considérée comme satisfaisante pour décrire le phénomène observé, ou les observations effectuées.

Ces deux critères permettent ainsi de justifier ou non l'inférence d'une loi à densité de probabilité pour décrire la distribution des données hydrogéologiques observées. Ainsi, un premier tri peut être fait sur les lois pouvant être utilisées, à un niveau de confiance donné, parmi les lois étudiées. C'est à l'issue de ce premier « tri » que les densités présentées en Annexe II ont été retenues.

4) Conception de modèles de justesse.

Pour décrire la justesse du modèle nous nous sommes intéressés au fait que la valeur mesurée était inscrite dans l'intervalle de prévision. Soit la valeur mesurée appartient à cet intervalle, soit elle ne lui appartient pas. De plus, pour une classe, il est nécessaire de pouvoir prendre en compte le caractère dynamique de l'évolution des hauteurs d'eau. En effet une même hauteur peut correspondre à différentes situations hydrogéologiques. Pour approcher la notion d'état hydrogéologique, nous avons proposé de traiter différemment les deux situations suivantes : soit l'hydrosystème est en phase de recharge (pente du limnigramme positive ou constante), soit l'hydrosystème est en récession (pente négative). Nous avons donc quatre configurations qui représentent autant de lois à ajuster : pente positive ou négative et valeur mesurée appartenant ou n'appartenant pas à l'intervalle de prévision calculé. Ainsi, pour chaque temps discret k il est défini si la valeur de la variable d'étude mesurée est incluse dans l'intervalle de prévision. Puis pour chaque classe de hauteur d'eau, le nombre d'échantillons mesurés total et le nombre d'échantillons mesurés inclus dans l'intervalle de prévision sont comptabilisés. De même, le nombre d'échantillons mesurés étant en dehors de l'intervalle de prévision est aussi déterminé.

Pour chaque configuration, une loi à « densité de probabilité composite » est inférée : c'est une superposition de plusieurs lois identiques par leur nature, mais différentes par leurs paramètres (moyenne, variance par exemple pour la loi normale), qui doit représenter au mieux les données mesurées. Pour chaque classe de variables d'étude observée c_i, la justesse du modèle, c'est-à-dire son espérance mathématique : E_{c_i} est définie selon le calcul suivant (66).

$$\mathrm{E}_{c_i} = p_j(x) = \frac{f_{c_i}(x \in PI)}{f_{c_i}(x)} = \frac{f_{c_i}(x \in PI)}{f_{c_i}(x \in PI) + f_{c_i}(x \notin PI)} \quad (66)$$

Avec E_{c_i} la justesse du modèle pour la classe observée c_i, $p_j(x)$ la probabilité de justesse attribué à la variable x, PI l'intervalle de prévision, f_{c_i} la fonction de la loi à densité de probabilité, x la valeur hydrogéologique observée.

Enfin, ces calculs sont répétés en remplaçant les lois théoriques de densités de probabilité par la distribution réelle des données hydrogéologiques observées. Ces valeurs permettent de vérifier l'adéquation des courbes théoriques – lois à densité de probabilité – avec l'utilisation de données réelles – distribution réelle des valeurs de niveaux d'eau ou de débits.

La loi théorique de densité de probabilité la mieux inférée à la distribution hydrogéologique peut être donc définie à cette étape, grâce aux critères d'adéquation de courbes tels que le coefficient de corrélation de Pearson ou le test de Kolmogorov-Smirnoff.

5) Détermination de la marge d'erreur

Considérons une série de mesures supposée distribuée selon une loi normale $\mathcal{N}(\bar{x},\ {\sigma_x}^2)$, la probabilité que l'erreur d'une mesure soit comprise dans un intervalle donné est exprimée par la fonction « erreur » erf. Cependant, les paramètres moyenne et écart-type au sein de chaque classe sont inconnus et ne peuvent être calculés.

Pour pallier ce problème, la loi d'erreur de Gauss est utilisée afin d'estimer cette incertitude, sans avoir besoin de ces paramètres. Introduite par Glaisher en 1871 (Glaisher 1871), cette fonction est définie de la façon suivante (67) :

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-\xi^2}\, d\xi \tag{67}$$

Avec erf la fonction erreur de Gauss et x la variable.

La marge d'erreur peut être déterminée depuis la fonction de répartition de la loi normale, en utilisant cette loi d'erreur (68) :

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{\xi^2}{2}}\, d\xi = \frac{1}{2}\left(1 + \frac{\mathrm{erf}(x)}{\sqrt{2}}\right) \tag{68}$$

Avec $\Phi(x)$ la loi normale appliquée à la variable x.

La fonction d'erreur inverse permet de retrouver, à partir d'une probabilité connue, et en fonction de la loi normale et d'une variable complexe j. Ce calcul est alors réalisé pour chaque classe de valeurs hydrogéologiques. Le calcul de la fonction d'erreur inverse ne peut se réaliser que de façon approchée, à partir de la fonction suivante, selon la suite de Maclaurin (69) :

$$erf^{-1}(j) = \frac{\sqrt{\pi}}{2}\left(j + \frac{\pi}{12}\, j^3 + \frac{7\pi^2}{480}\, j^5 + \frac{127\pi^3}{40320}\, j^7 + \frac{4369\pi^4}{5806080}\, j^9 + \cdots\right) \tag{69}$$

Avec $erf^{-1}(j)$ étant la fonction erreur inverse et j le nombre complexe.

Pour notre étude, le nombre complexe j est remplacé à partir d'une probabilité de justesse p_j définie par l'utilisateur (70) :

$$j = \frac{p_j + 1}{2} \tag{70}$$

La marge d'erreur maximale ε peut alors se calculer ainsi (71) :

$$\varepsilon = \frac{erf^{-1}\left(\frac{p_j+1}{2}\right) \times \frac{1}{2}}{\sqrt{n}} \tag{71}$$

Le résultat de cette fonction représente alors l'erreur sur les calculs d'incertitude de modélisation. La moitié de cette marge d'erreur est soustraite au modèle de justesse pour chaque classe hydrogéologique afin de représenter la borne inférieure de l'erreur sur les calculs d'incertitude de modélisation. L'autre moitié de cette marge d'erreur est de la même manière additionnée à la justesse pour définir la borne supérieure de cette incertitude.

Il est possible d'utiliser cette marge d'erreur sous forme de critère afin de savoir le pourcentage de valeurs, issues des modèles de justesse utilisant des données hydrogéologiques observées, comprises dans cette marge d'erreur (Figure III-10).

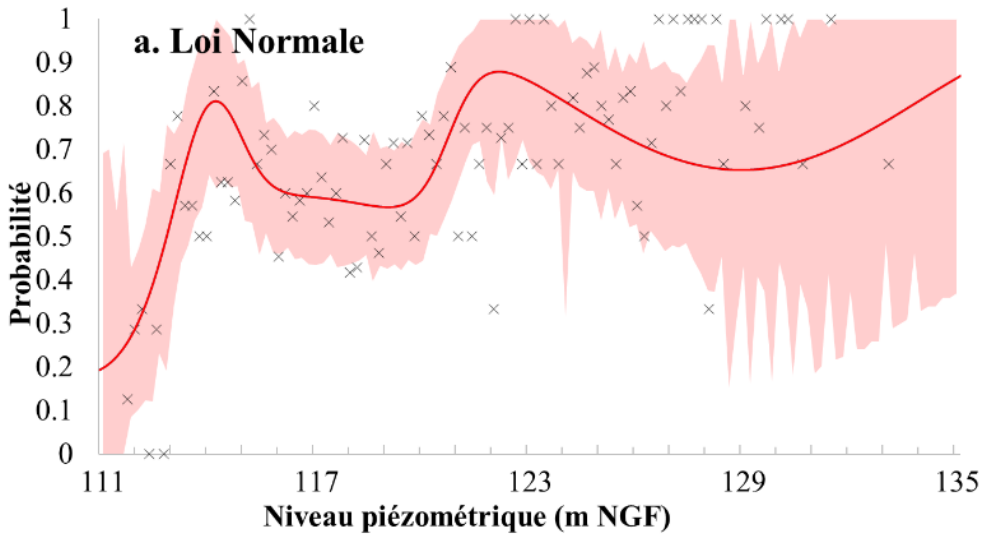


Figure III-10 - Exemple de représentation de marge d'erreur.

6) Détermination de l'intervalle de confiance à partir d'un indice de confiance.
Le second objectif de cette méthodologie est de définir l'intervalle de confiance à partir d'une probabilité de confiance p_{conf} fixée par l'utilisateur. Pour cela, il sera nécessaire d'obtenir un grand nombre de valeurs prédites pour le plus de classes

possibles ; un ensemble de modèles contenant de nombreux membres est donc nécessaire.

A chaque instant discret k, nous pouvons calculer les valeurs minimale et maximale modélisées sur cet ensemble de modèles, puis calculer le critère C_{PICP_n}(*Prediction Interval Probability Coverage*). La probabilité de justesse est retrouvée à partir des valeurs de niveaux piézométriques observés et modélisés médians et des modèles de justesse calculés précédemment.

Il est possible d'observer une relation affine entre la probabilité de justesse appliquée aux niveaux piézométriques observés $p_j\ (x_{obs})$ et celle appliquée aux niveaux piézométriques prévus médians $p_j\ (x_{médian})$. Les paramètres *a* et *b* de la fonction affine sont utilisés pour calculer la probabilité de justesse $p_{j,rec}\ (x_{obs})$ appliquée aux données observées (Figure III-11).

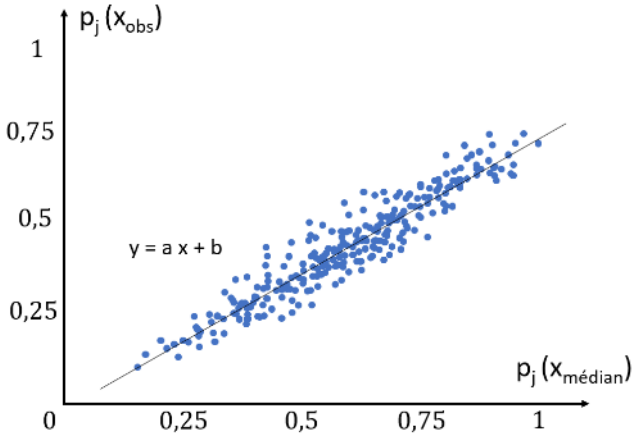


Figure III-11 - Calcul des niveaux piézométriques à partir des probabilités de justesse.

En parallèle, la probabilité de confiance p_{conf} imposée par l'utilisateur doit correspondre au résultat du critère C_{PICP_n} appliqué sur l'ensemble de modèles. Lorsque le nombre de modèles n est suffisamment grand, il est considéré que l'intervalle de prévision reste constant. Ainsi, une table de correspondance est alors construite entre la probabilité de confiance p_{conf} et le critère C_{PICP_n} en faisant varier la probabilité de confiance. Il est donc possible par la suite de retrouver les limites de l'intervalle de confiance, défini par la probabilité fixée par l'utilisateur selon les équations suivantes (72) (73) :

$$CI_{inf} = \frac{PI_{n_{inf}} \times p_{conf}}{p_{j,rec}\ (x_{obs})} = \frac{PI_{n_{inf}} \times f(C_{PICP_n})}{\left(a \times p_j\ (x_{médian})\right)+b} \quad (72)$$

$$CI_{sup} = \frac{PI_{n_{sup}} \times p_{conf}}{p_{j,rec}\ (x_{obs})} = \frac{PI_{n_{sup}} \times f(C_{PICP_n})}{\left(a \times p_j\ (x_{médian})\right)+b} \quad (73)$$

Avec n le nombre de membres de l'ensemble de modèles où n est grand, CI_{inf} et CI_{sup} respectivement les limites inférieures et supérieures de l'intervalle de confiance à calculer, $PI_{n_{inf}}$ et $PI_{n_{sup}}$ respectivement les limites inférieures et supérieures de l'intervalle de prévision du modèle d'ensemble à n modèles, C_{PICP_n} le critère *Prediction Interval Probability Coverage* appliqué au modèle d'ensemble à n modèles, $p_{j,rec}\ (x_{obs})$ la probabilité de justesse recalculée à partir des modèles de justesse et appliquée sur des données piézométriques observées, $p_j\ (x_{médian})$ la probabilité de justesse obtenue à partir des modèles de justesse et appliquée sur des valeurs

piézométriques médianes simulées et a et b les paramètres de la fonction affine présentée en Figure III-11.

Grâce à ces calculs, l'intervalle de confiance est établi ; il est ajustable en fonction d'une probabilité de confiance imposée par l'utilisateur. Ainsi, l'utilisateur peut imposer le taux de valeurs observées comprises dans l'intervalle de prévision. La théorie étant expliquée, cette méthodologie doit désormais être appliquée sur des cas concrets, sur plusieurs bassins versants.

Pour une meilleure compréhension, l'ensemble des étapes clés de la méthodologie élaborée sur les paragraphes précédents est résumée sur les schémas suivants, selon les sources d'incertitudes. Le premier schéma concerne l'estimation de l'incertitude d'initialisation des paramètres du modèle (Figure III-12).

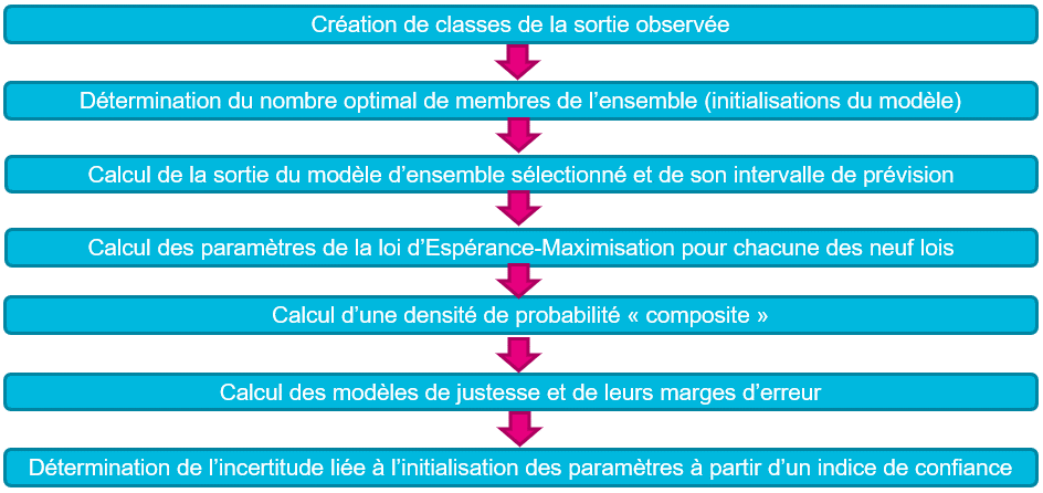


Figure III-12 Schéma de la méthodologie d'estimation de l'incertitude d'initialisation des paramètres du modèle.

Le second schéma concerne l'estimation de l'incertitude liée aux résultats de modélisation, incluant à la fois les incertitudes d'initialisation des paramètres du modèle et les incertitudes sur la distribution spatiales de variables d'entrée (Figure III-13).

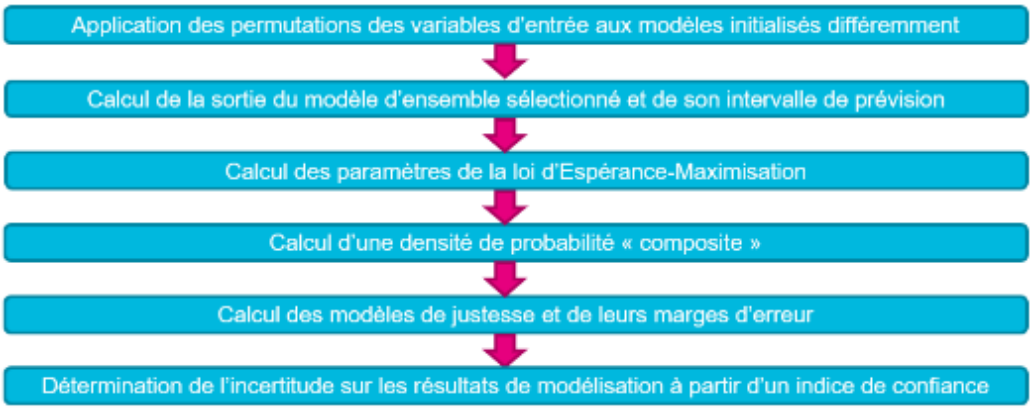


Figure III-13 Schéma de la méthodologie d'estimation de l'incertitude liée aux résultats de modélisation.

Il est à noter que cette méthodologie ne reprend pas les deux premières étapes de la méthodologie précédente. L'étape de création de classes de sortie observée a déjà été fixée. Il n'est alors pas nécessaire de la modifier. L'étape concernant l'obtention du nombre optimal de membres au sein d'un modèle d'ensemble est, quant à elle, spécifique aux incertitudes liées aux paramètres du modèle. L'ajout de l'incertitude de la distribution spatiale des données d'entrées n'a donc pas d'incidence sur cette étape.

V. Conclusions du Chapitre III

Nous avons vu que les incertitudes liées aux résultats de modélisation hydrogéologique dépendent des incertitudes sur les données, des incertitudes liées à l'architecture du modèle et les incertitudes liées à aux paramètres du modèle.

Les méthodes de régularisation nous permettent de réduire les incertitudes liées à l'architecture, et l'utilisation de données contrôlées nous autorise de faire le postulat que les incertitudes liées à la mesure sont faibles, comparées aux incertitudes liées à la distribution spatiale des pluies et à l'initialisation des paramètres.

Afin d'estimer les incertitudes sur la modélisation hydrogéologique, une première partie de la méthodologie consiste à évaluer l'impact de l'initialisation des paramètres du modèle sur l'incertitude globale dès la sélection du modèle en définissant un nombre optimal de modèles au sein d'un modèle ensembliste et d'observer la dispersion des résultats de modélisation de ces modèles.

La seconde partie de la méthodologie consiste à transcrire cette dispersion des résultats de modélisation en probabilité, en utilisant des lois à densités de probabilités. De plus, les incertitudes de distribution spatiale d'une variable est ajoutée en permutant les variables

d'entrée au sein du modèle sans modification de l'architecture. Cet ajout permet l'obtention d'un nouvel ensemble de modèles qui permet l'estimation des incertitudes d'initialisation des paramètres et des incertitudes de distribution spatiale.

Grâce à ces développements, il est désormais possible d'estimer les incertitudes liées aux résultats de modélisation, que nous appliquerons sur le chapitre suivant.

Chapitre IV – Applications des méthodologies d'estimation des incertitudes

I. Introduction

Ce chapitre applique la méthodologie de conception de modèles hydrogéologiques vue dans le chapitre II et la méthodologie d'estimation des incertitudes liées aux résultats de modélisation vue dans le chapitre III sur deux bassins hydrogéologiques de typologie différente, sur des horizons de prévision différents.

Le premier bassin d'étude est la nappe du bassin du Blavet, située en Bretagne, qui est une nappe côtière affleurante, en contexte rural, en amont de bassin, reposant sur un socle granitique.

Le second bassin d'étude est la nappe de la Craie de Champagne Sud et Centre, située dans la région Grand-Est. Il s'agit ici d'une nappe partiellement affleurante, en contexte rural malgré la présence de l'agglomération troyenne à proximité du point d'étude, et reposant sur un socle calcaire.

II. La nappe du bassin du Blavet

Dans cette première partie, nous présenterons le bassin d'étude ainsi que le réseau de mesure, puis les résultats obtenus pour une modélisation hydrogéologique à court-terme et à moyen-terme.

Ce bassin a été sélectionné compte-tenu de sa supposée simplicité de fonctionnement hydrogéologique, permettant ainsi une application de la méthodologie développée plus aisée, et de la disponibilité et de la qualité des données contenant peu de lacunes sur la période souhaitée.

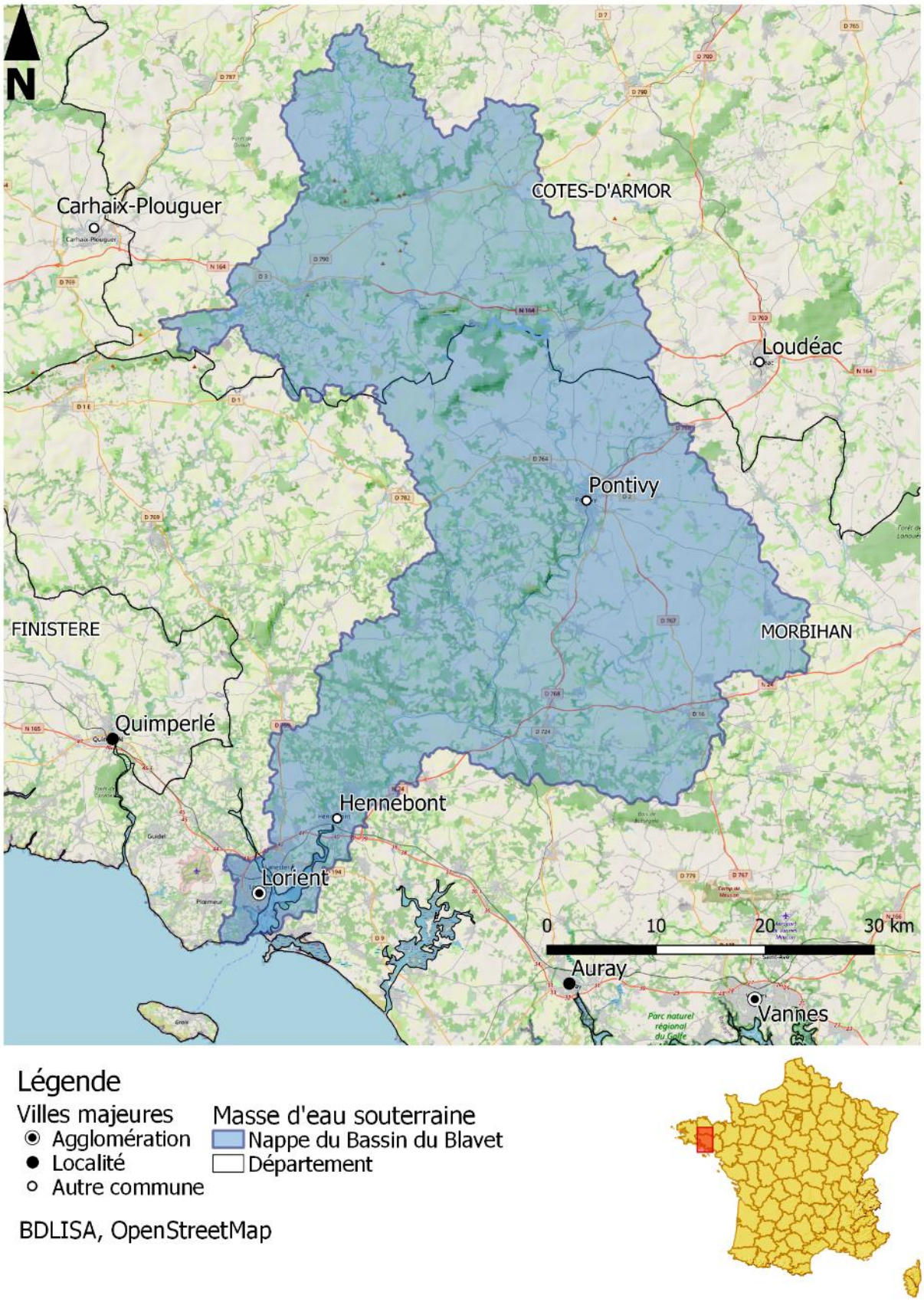
II.1 Contexte général

II.1.1. Généralités

Située au cœur de la Bretagne, la nappe du bassin versant du Blavet est gérée par le Syndicat Mixte du SAGE Blavet, structure porteuse et coordinatrice des dispositions du Schéma d'Aménagement et de Gestion des Eaux (SAGE) du bassin du Blavet.

Cette nappe de socle a été découpée en fonction du bassin versant superficiel du fleuve nommé le Blavet. En effet, elle siège dans un socle métamorphique avec des intrusions salines provenant de l'Océan Atlantique (BRGM 2009). Ces intrusions salines demeurent toutefois faibles à modérées et sont localisées le plus souvent près des côtes (Lucassou, Schroëtter, et al. 2019, Les Landes, et al. 2015).

Cette masse d'eau souterraine s'étend sur 2 128 km², à cheval sur deux départements : les Côtes d'Armor (22) en amont et le Morbihan (56) en aval. Elle chevauche notamment les villes de Lorient – et sa ville voisine Lanester – (80 000 habitants), Pontivy (15 000 habitants) et Hennebont (15 000 habitants) et s'étend sur 105 communes (Carte IV-1). De par la population impactée, de nombreux enjeux divers et variés, que nous verrons par la suite, dépendent de cet aquifère.



Carte IV-1 – Présentation de la nappe du bassin du Blavet.

II.1.2. Hydrographie et hydrogéologie

II.1.2.a. Hydrographie

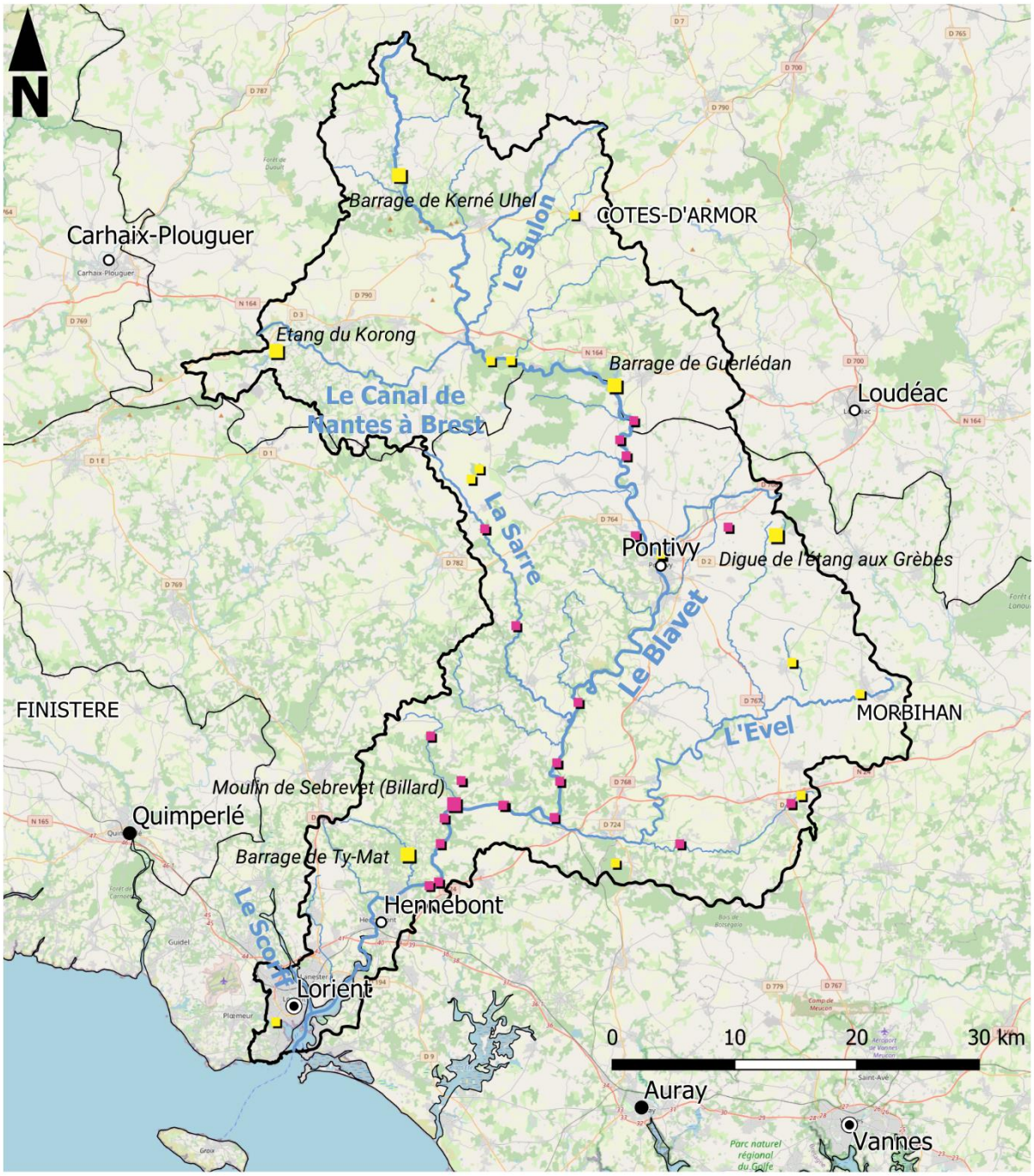
Le Blavet est un fleuve côtier breton, prenant sa source dans la commune costarmoricaine de Bourbriac située à 280 mètres d'altitude. Ce cours d'eau s'écoule du nord au sud et se jette dans l'Océan Atlantique après un trajet de 148,9 km. Son bassin versant est de la même taille que la nappe souterraine, soit 2 128 km² (SANDRE 2011). Il représente ainsi 7,7% de la superficie totale de la région Bretagne, faisant de lui le second bassin versant breton par la taille, derrière le bassin versant de la Vilaine (Houet 2006), (Lucassou et Mougin 2015).

Son débit moyen avant l'embouchure, est de 28 m^3/s, au niveau de la station Le Blavet à Languidic – Quellenec sur la période 1983-2016, avec des variations saisonnières importantes, allant de moins de 6 m^3/s au mois d'août en période d'étiage jusqu'à 67,7 m^3/s au mois de janvier en période de hautes eaux (Banque Hydro 2020). L'analyse spectrale des débits journaliers du Blavet entre 1939 et 1999 a montré une périodicité annuelle et une grande variabilité de ceux-ci (Hubert, et al. 2002).

Ses principaux affluents sont la Sarre (35 km pour un bassin de 132 km²) en rive droite, le Sulon (29 km pour un bassin de 100km²) et l'Evel (55 km pour un bassin de 478 km²) en rive gauche (Carte IV-2). Le Scorff rejoint le Blavet au niveau de l'embouchure du fleuve dans le Rade de Lorient.

Le réseau hydrométrique est assez dense, avec 9 stations encore en activité et 10 autres qui ont été fermées. Parmi les stations encore en activité, 6 stations se trouvent sur le Blavet, les 3 autres étant sur ses affluents.

Une des caractéristiques du Blavet est qu'il est fortement artificialisé, à l'exception de sa partie très en amont. En effet, la présence de barrages, utilisés pour l'alimentation en eau potable, du canal de Nantes à Brest ainsi que de ses aménagements et écluses réalisés en aval, au niveau de l'agglomération lorientaise (Lucassou et Mougin 2015), font de ce fleuve une masse d'eau fortement modifiée au regard de la Directive Cadre sur l'Eau (Carte IV-2).



Carte IV-2 - Obstacles à l'écoulement sur le Blavet et ses affluents.

II.1.2.b. Hydrogéologie

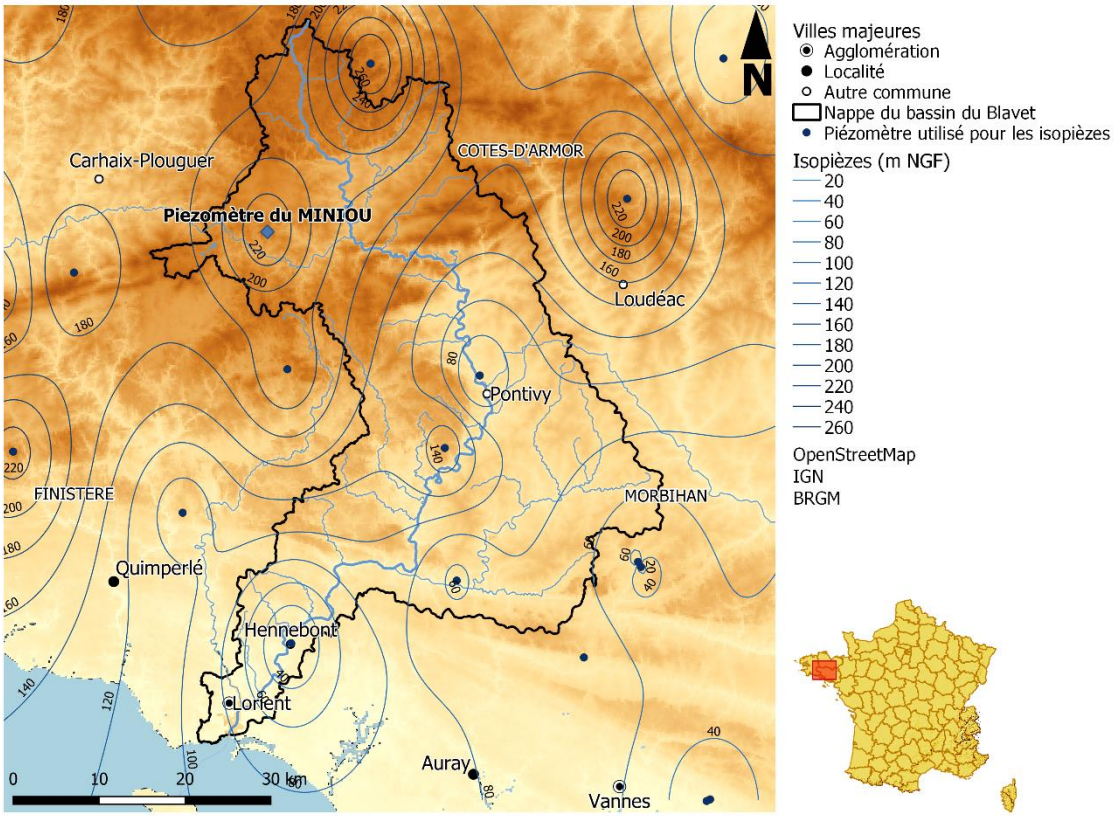
La nappe du bassin versant du Blavet est une nappe de socle. En 2006, un réseau de 52 piézomètres répartis sur l'ensemble de la Bretagne a permis de suivre l'évolution de cette nappe, ainsi que les nappes avoisinantes, dans le cadre du projet SILURES (Mougin, Carn, et al. 2004), (B. Mougin 2005), (B. Mougin 2006). Ce projet a eu pour but d'évaluer la ressource en eau souterraine, tant sur les aspects qualitatifs que quantitatifs par le biais de chroniques piézométriques, et de gérer le réseau de piézomètres permettant de faire cette évaluation (Mougin, Carn, et al. 2004), (B. Mougin 2005), (B. Mougin 2006).

A ce jour, il existe 57 piézomètres sur l'ensemble de la région dont 5 sur le territoire d'étude (ADES 2020). Tous les piézomètres dont leur mise en service est antérieure à 2006 (soit 52 des 57 piézomètres) sont équipés d'un capteur mono-voie flotteur Thalimède de la marque OTT, ainsi que d'un télétransmetteur (B. Mougin 2006).

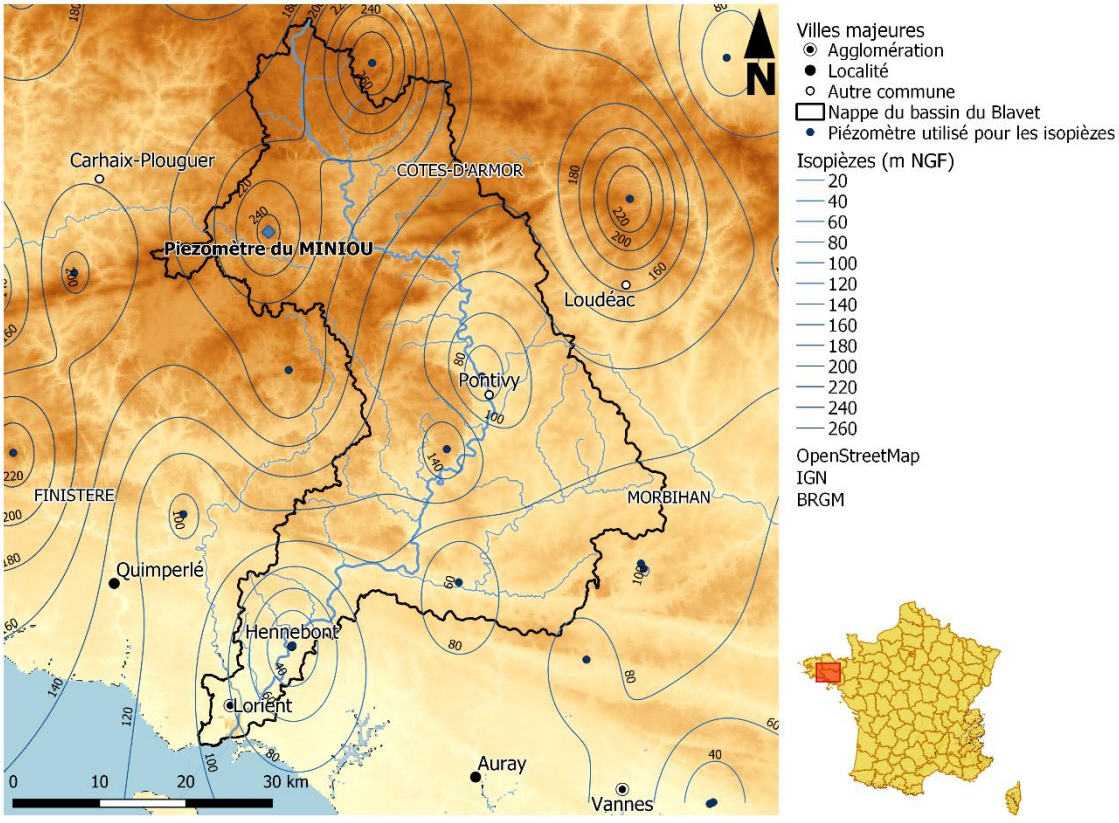
Les battements de nappe sont variables selon le positionnement des piézomètres. En effet, sur le piézomètre de Kerpert, l'amplitude maximale sur un historique piézométrique compris entre 2004 et 2006 est d'environ 3,5 mètres tandis que sur le piézomètre du Miniou à Rostrenen, il a été enregistré une variation de la nappe entre novembre et décembre 2005 de près de 9,5 mètres (B. Mougin 2006). Ces deux piézomètres se trouvent à moins de 15 km l'un de l'autre mais sont situés sur des terrains géologiques différents (l'un granitique, l'autre schisteux). Cette différence géologique peut potentiellement expliquer cette différence de variation piézométrique. Cela peut avoir aussi des répercussions sur les échanges entre eau superficielle et eau souterraine.

Les cartes IV-3 et IV-4 montrent les isopièzes de la nappe respectivement pour un évènement sec (début décembre 2017) et un évènement humide (début février 2014).

De manière cohérente, les isopièzes calculées sont plus basses pour un évènement sec que pour un évènement humide. De plus, ces isopièzes suivent le sens d'écoulement du Blavet, ainsi que le relief, avec des isopièzes plus hautes en amont du bassin (particulièrement sur le nord-ouest).



Carte IV-3 – Isopièzes du bassin du Blavet lors d'un évènement sec (sept.-déc. 2017).



Carte IV-4 – Isopièzes du bassin du Blavet lors d'un évènement humide (jan.-mars 2014).

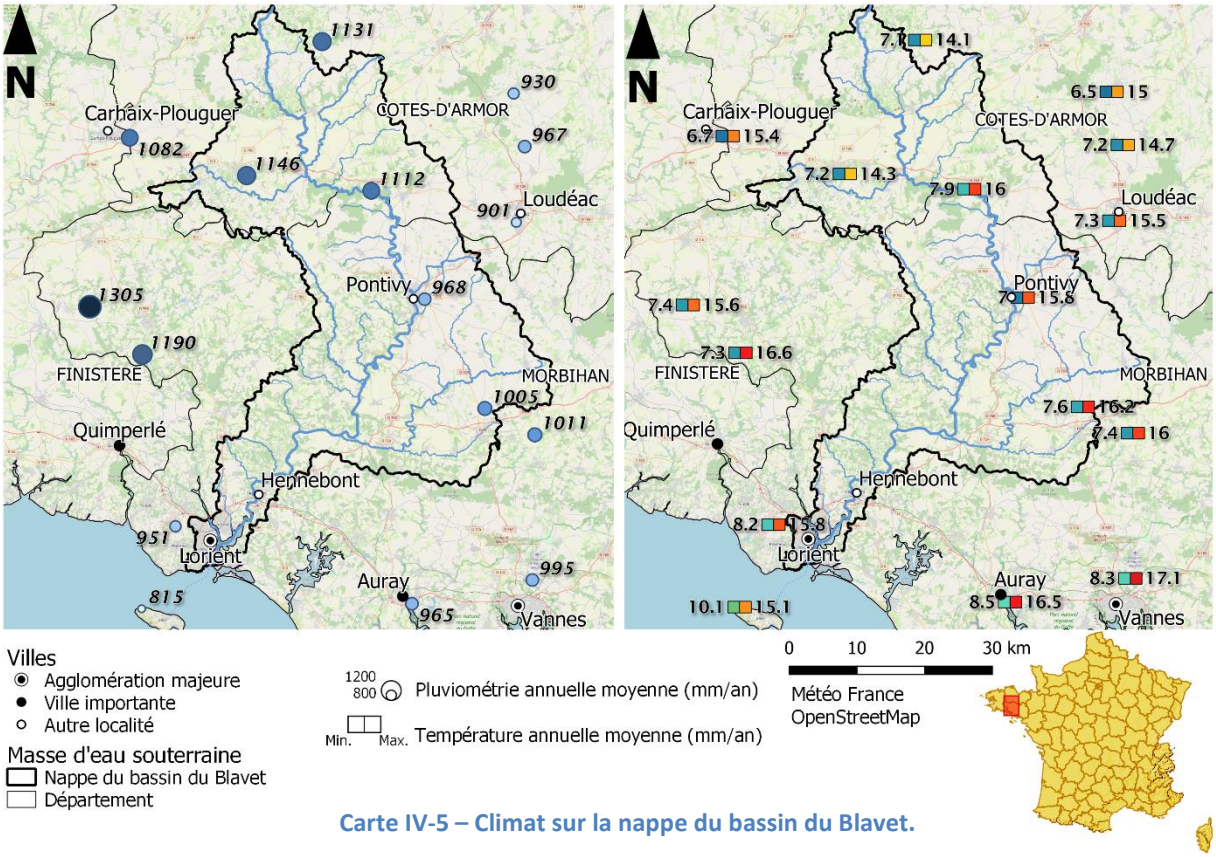
Le Blavet semble drainer la nappe éponyme et bénéficier ainsi d'abondantes contributions souterraines à son débit. En effet, Mougin et *al.* (2004) ont montré que les contributions souterraines sur les débits du Blavet au niveau de la station hydrométrique du Blavet à Kerien [Kerlouet] pouvaient être estimées à 57 %, à partir de modélisations d'écoulement réalisées avec l'outil du BRGM Gardénia, sur des historiques de données compris entre 1994 et 2000 (Mougin, Carn, et al. 2004).

Les prélèvements sont dans l'ensemble homogènes sur l'ensemble du bassin, hormis sur la partie la plus en amont du bassin, où les prélèvements sont plus faibles voire inexistants (BNPE 2020). Ceux-ci sont majoritairement utilisés pour l'alimentation en eau potable (45%), mais aussi pour l'industrie (20%), l'irrigation (23%) et les canaux (12%) (Syndicat de la Vallée du Blavet 2014). Il existe néanmoins quelques secteurs où les prélèvements souterrains sont plus importants, notamment en aval du bassin autour de Pontivy et d'Hennebont et de Lanrivain en amont du bassin de la nappe. En aval, le piézomètre du Miniou est par exemple utilisé pour l'alimentation en eau potable (AEP). D'après la Banque Nationale des Prélèvements en Eau, les prélèvements d'eau annuels effectués sur une moyenne faite entre 2008 et 2017 sont de 95 946 m^3 pour l'eau potable sur ce captage (BNPE 2020).

II.1.3. Climat

Le climat sur le bassin de la nappe du Blavet est océanique. Sur ce territoire, la pluviométrie annuelle moyenne est supérieure à celle de la France.

Le climat est assez homogène sur l'ensemble du bassin, à l'exception des côtes morbihannaises, où les températures minimales sont plus douces et la pluviométrie plus faible. Sur le nord du bassin, la pluviométrie annuelle moyenne oscille entre 900 et 1 150 mm/an. Sur les côtes, cette pluviométrie n'excède pas les 1 000 mm/an. Le gradient pluviométrique est aussi visible d'est en ouest (Houet 2006), où les pluies sont plus importantes annuellement sur la partie ouest de l'ordre de 100 à 200 mm/an. Les températures minimales annuelles moyennes sont entre 6,5 °C et 8,5 °C. Les écarts de températures maximales annuelles moyennes entre la côte Atlantique et les terres bretonnes sont moins restreints, avec des températures étant comprises entre 14,1 °C et 17,1 °C (Carte IV-5).

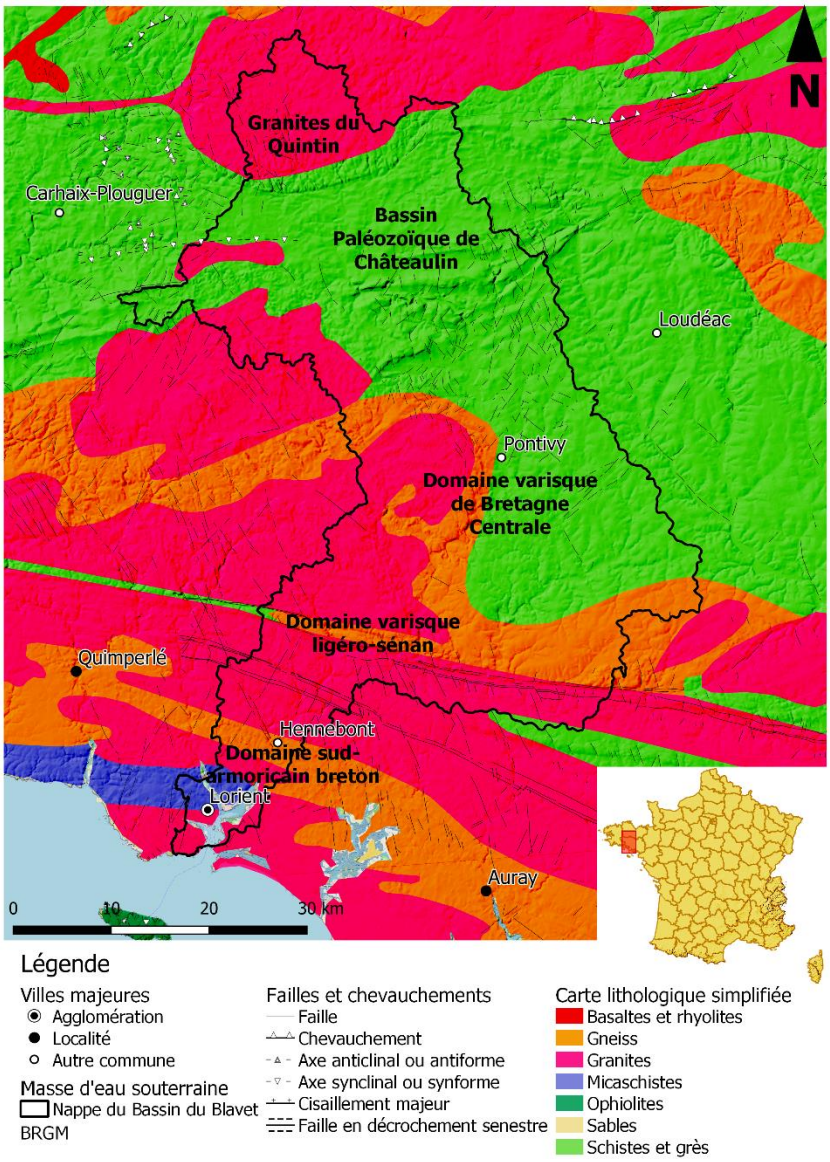


Carte IV-5 – Climat sur la nappe du bassin du Blavet.

Il est à noter que les stations climatologiques ayant servies à construire cette carte se basent initialement sur des historiques de données, issus de Météo-France, compris entre 1981 et 2010. Toutefois, de nombreuses stations n'existaient pas en 1981. Les chiffres présentés sur la Carte IV-5 ne sont montrés ici qu'à titre informatif et n'ont pour vocation que d'avoir une analyse globale du climat sur le territoire. Il n'est donc pas possible de comparer deux stations climatologiques entre elles.

II.1.4. Géologie

La nappe du bassin du Blavet siège dans des roches métamorphiques et magmatiques. Certaines roches sont d'origine sédimentaire – telles que les grès par exemple – qui se sont métamorphisées avec le temps. Toutes ces roches sont parfois profondément altérées sur la quasi-totalité de sa superficie grâce à un réseau de failles développé. Cela contrebalance la faible porosité de ces types de roches vis-à-vis de l'infiltration des eaux (Carte IV-6) (Lucassou et Mougin 2015), (Les Landes, et al. 2015).



Carte IV-6 - Géologie simplifiée de la nappe du bassin du Blavet.

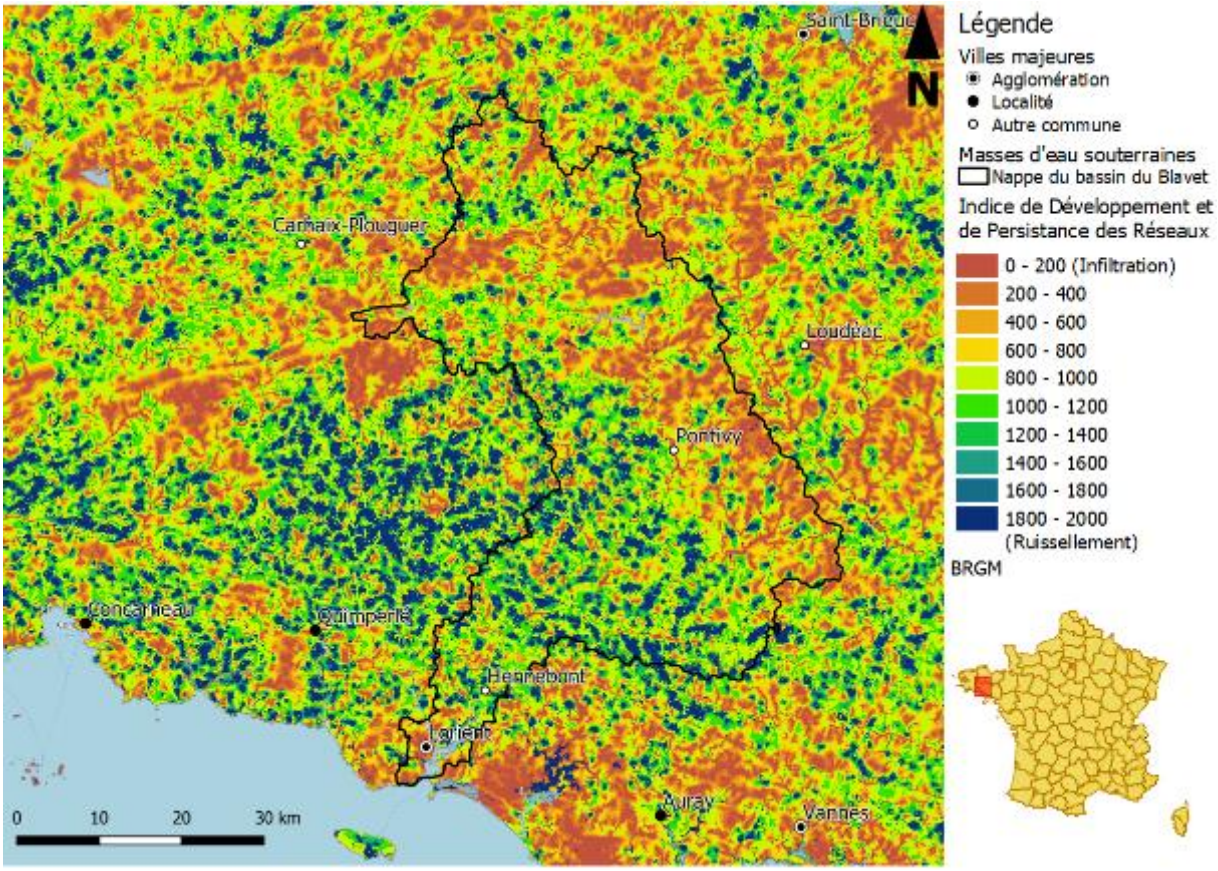
Il existe toutefois quelques disparités au sein de l'aquifère. La nappe du bassin du Blavet peut être divisée en cinq types d'aquifères différents. Le premier siège dans les granites de Quintin datant du Carbonifère sur la partie nord de la nappe, contenant de la hornblende et de la biotite (Lucassou et Mougin 2015), (Ballouard, et al. 2017).

Une seconde partie de la nappe se situe sur un empilement de roches métamorphiques et sédimentaires telles que des schistes et grès, nommé bassin paléozoïque de Châteaulin. Cette partie correspond à la partie aval du barrage de Kerné Uhel en surface (Lucassou et Mougin 2015).

Le troisième terrain est constitué de micashistes, de paragneiss ainsi que de leucogranites (granites ayant une faible teneur en micas noirs et sont appelés plus couramment les granites blanc) (Lucassou et Mougin 2015), (Ballouard, et al. 2017). Cet aquifère est connu sous le nom de domaine varisque de Bretagne Centrale (Lucassou et Mougin 2015).

Plus en aval, la nappe du bassin du Blavet repose sur le domaine varisque ligéro-sénan, composé de leucogranites d'Ergué-Languidic (Lucassou et Mougin 2015).

Pour finir, le dernier type d'aquifère correspond à des roches granitiques et autres roches métamorphiques telles que les gneiss et micaschistes, qui se trouvent en aval de la nappe sur le domaine sud armoricain breton (Lucassou et Mougin 2015).



Carte IV-7 – Indice de Développement et de Persistance des Réseaux (IDPR) sur la nappe du bassin du Blavet.

Il est possible de faire le rapprochement entre la perméabilité des sols et la géologie par l'utilisation de l'Indice de Développement et de Persistance des Réseaux (IDPR, Carte IV-7) (Mardhel et Allier 2010).

Sur ce bassin, il est constaté que les zones d'infiltration principales correspondent aux endroits où la présence de schistes et/ou de grès est majoritaire, c'est-à-dire sur l'ouest du bassin. *A contrario*, il est à noter que la région des granites du Quintin présente une infiltration relativement importante, laissant supposer un état avancé de son altération. Ces zones d'infiltration semblent correspondre aux zones ou la profondeur de nappe est la plus faible (Cartes IV-3 et IV-4).

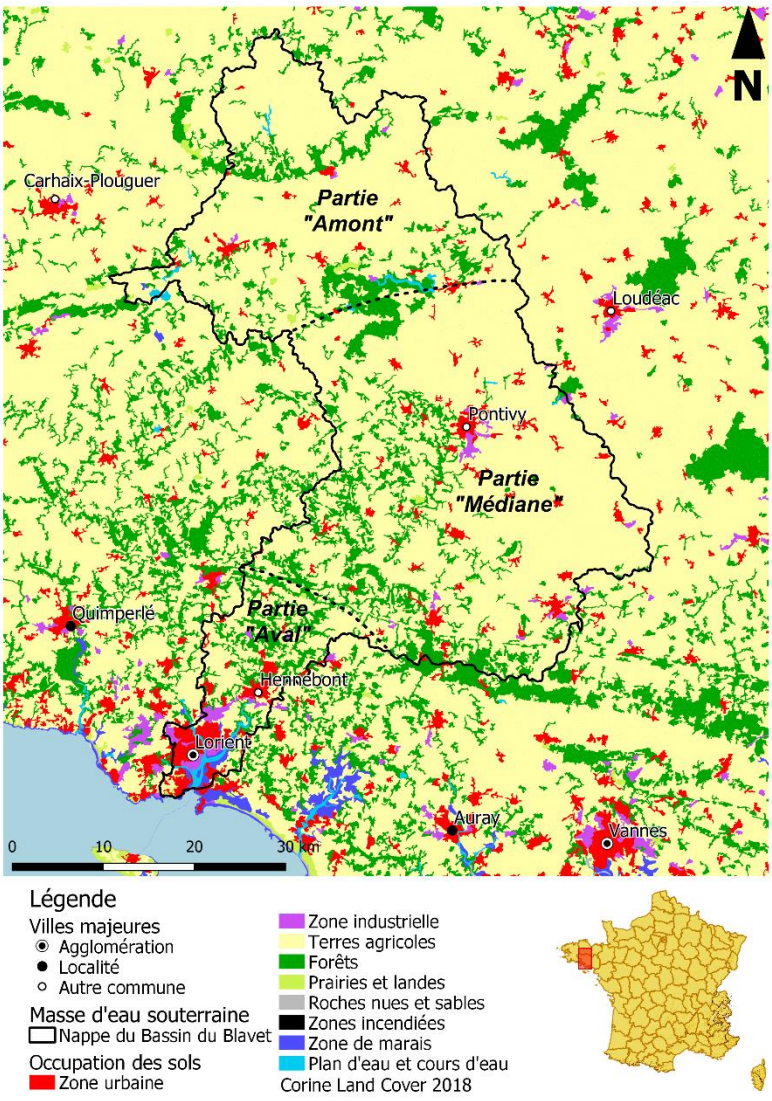
De plus, cet indice IDPR montre donc qu'en amont de la nappe, l'infiltration est plus forte que le ruissellement. Cela coïncide avec les études de Mougin et *al*. en 2004 vis-à-vis des

contributions souterraines sur le débit du Blavet, montrant ainsi la possibilité d'échanges entre les eaux superficielles et les eaux souterraines (Mougin, Carn, et al. 2004).

II.1.5. Occupation des sols

L'occupation des sols sur le bassin de Blavet est principalement marquée par les activités agricoles, représentant près de 80 % de la superficie totale du territoire, malgré la présence de forêts éparses (13 %), de quelques pôles urbains ou semi-urbains (4,5 %) et de zones humides et d'étendues d'eau (1,5 %) (Houet 2006).

Les activités anthropiques sur le territoire d'étude sont très hétérogènes. En 2006, Houet a scindé le territoire en trois parties distinctes : la partie « Amont », correspondant à la surface costarmoricaine, la partie « Médiane », correspondant au territoire compris entre le lac artificiel de Guerlédan et la confluence de l'Evel avec le Blavet, et la partie « Aval », qui regroupe les derniers kilomètres du Blavet et son embouchure, plus communément appelé « la Rade de Lorient » (Carte IV-8) (Houet 2006).



Carte IV-8 - Occupation des sols sur la nappe du bassin du Blavet (Corine Land Cover, 2018).

En tête de bassin, sur la partie « Amont », les activités agricoles sont majoritaires, avec la présence d'une agriculture d'élevage essentiellement bovine, pour la production de lait, et aviaire. De plus, les prélèvements pour l'alimentation en eau potable (AEP), ainsi que la production d'électricité, sont importants avec la présence de plusieurs captages souterrains mais surtout celle de deux barrages sur le Blavet : celui de Kerné Uhel en amont, situé à moins d'une quinzaine de kilomètres par rapport à sa source, et celui de Guerlédan, à la frontière entre les Côtes

d'Armor et le Morbihan (Carte IV-9). Ces deux barrages servent également à la régulation et à la gestion des crues ainsi qu'au soutien d'étiage en période estivale. Le territoire est ici essentiellement rural ; sa plus grande commune est Rostrenen, avec ses 3 000 habitants (Houet 2006).

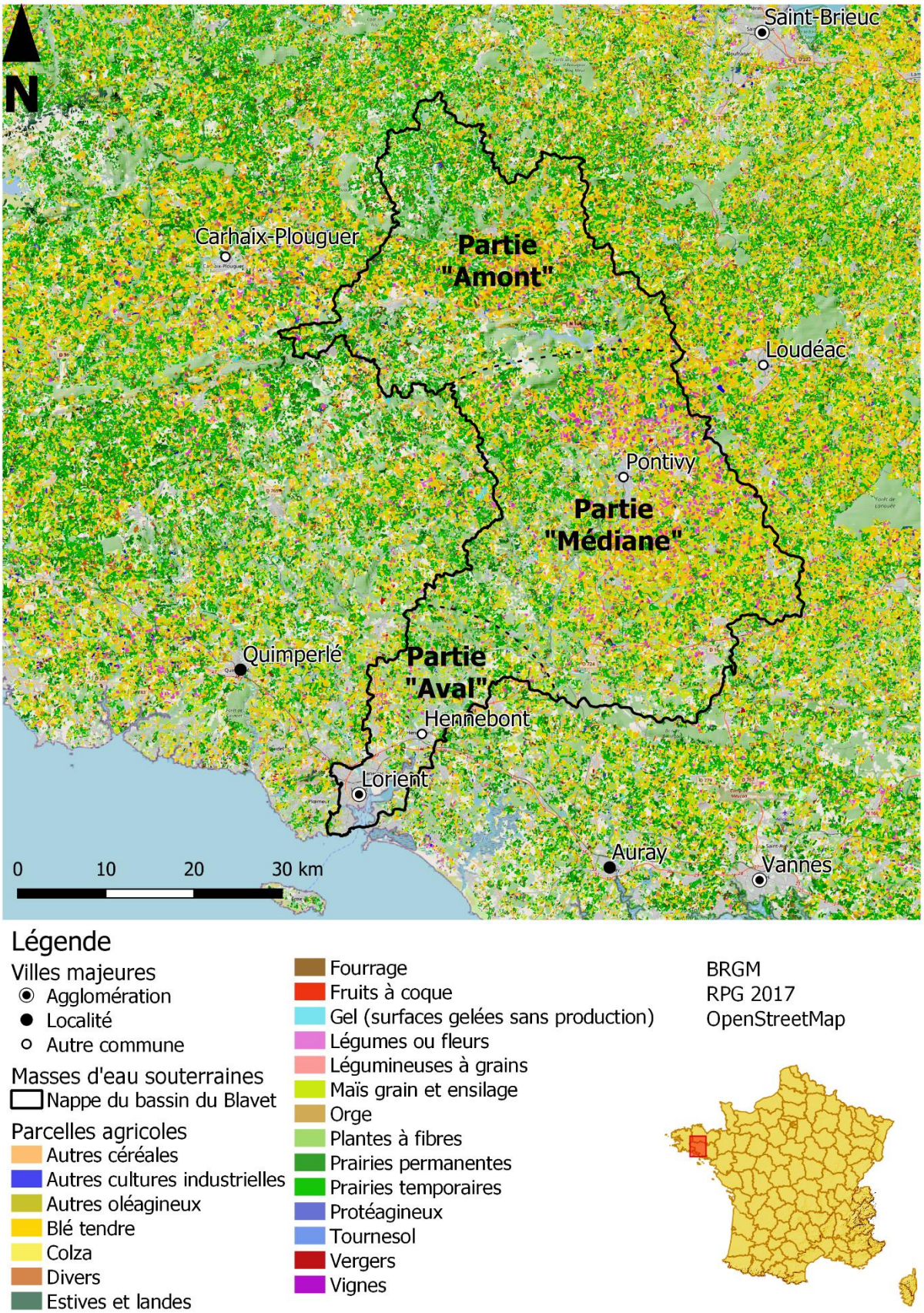
L'agriculture est très présente sur ce territoire, mais est essentiellement tournée vers l'élevage. En effet, seuls le maïs, le blé et l'orge sont cultivés sur la partie « Amont », au milieu des prairies pour l'élevage des bovins. Il est à noter qu'il n'existe principalement que des prairies au nord du barrage de Kerné Uhel (Carte IV-9).

La zone « Médiane » est très représentative du clivage observé sur le bassin. En effet, la dominance de cultures hors-sols sur la majeure partie de cette zone viennent contraster avec l'élevage intensif (élevages aviaires et porcins) et les industries agro-alimentaires (notamment les industries laitières) autour de Pontivy (Houet 2006).

En effet, à l'ouest, les prairies bovines dominent les cultures de blé et de maïs tandis qu'à l'est de cette partie, les cultures sont plus diversifiées et les prairies sont moins présentes. Parmi ces cultures, nous retrouvons les cultures de blé et de maïs, mais aussi d'orge, de colza et de légumineuses (Carte IV-9).

La partie « Aval » de la nappe du bassin du Blavet est l'opposé de sa partie « Amont », avec la présence d'un pôle urbain dense. En effet, les 11 communes les plus en aval, faisant partie de l'agglomération lorientaise, regroupent 50 % de la population présente sur le bassin, soit plus de 220 000 habitants. De plus, de nombreuses industries agroalimentaires et activités maritimes viennent renforcer cette urbanisation (Houet 2006).

La partie « Aval » contient moins de cultures que les précédentes zones de la nappe. Ici, les cultures de légumineuses et maraîchères (rose) se répartissent équitablement le territoire avec les cultures de maïs et de blé (Carte IV-9).



Carte IV-9 - Registre Parcellaire Graphique du bassin de la nappe du Blavet (RPG, 2017).

II.1.6. Ecosystèmes

Le bassin est également riche en écosystèmes. Pour rappel, un écosystème est un ensemble composé de la faune, de la flore et du milieu naturel dans lequel vivent ces espèces. Ces milieux naturels sont sensibles aux variations piézométriques. L'état de ces écosystèmes est un reflet potentiel de l'état de la ressource en eau, qu'elle soit superficielle ou souterraine. Une diminution de ces eaux peut engendrer un appauvrissement de la biodiversité locale. En effet, certaines espèces animales ou végétales attestent de la présence ou non d'eau. Il est donc intéressant de détailler les écosystèmes et la biodiversité présentes sur le territoire.

La présence de la réserve naturelle régionale des landes et marais de Glomel, situés à quelques kilomètres à l'ouest de Rostrenen, contribue à la biodiversité.

Cette réserve est scindée en deux sites, répartis sur près de 108 hectares (Barussaud et Blond 2015). Le premier est le marais de Magoar Penvern, où l'eau s'accumule en hiver et submerge les marais, landes tourbeuses et prairies humides. Le second est le site de Lann Bern, constitué de landes humides et tourbeuses. Ces habitats humides abritent ainsi de nombreuses espèces végétales et animales. En effet, près de 330 espèces entomologiques ont été recensées, dont 14 d'entre elles qui sont considérées comme étant déterminantes pour la désignation des ZNIEFF en Bretagne (Barussaud et Blond 2015).

Le bassin est également sous protection Natura 2000 sur sa partie amont. En effet, les têtes de bassin du Blavet et de l'Hyères sont composées de divers types d'habitants, tels que des landes sèches et humides, des zones humides, prairies tourbeuses et tourbières, ou encore diverses hêtraies. Grace à cette diversité d'habitats, cette zone Natura 2000 regorge d'espèces protégées, qu'elles soient animales, comme le *Rhinolophus ferrumequinum*, la *Lutra lutra* ou l'*Austropotamobius pallipes*, ou végétales, comme le *Vandenboschia speciosa* ou le *Luronium natans*. Au total, 25 espèces sont considérées comme importantes, dont 12 étant inscrites dans l'annexe II de la directive 92/43/CEE concernant la conservation des habitats naturels ainsi que de la faune et de la flore sauvages (Museum National d'Histoire Naturelle 1995).

Le Rade de Lorient a lui aussi été classé comme zone Natura 2000. Se trouvant en partie sur le bassin de la nappe du Blavet, localisé à son exutoire et le long de la côte Atlantique, 27 espèces d'oiseaux sont protégées par la directive 2009/147/CE, concernant la protection des oiseaux (Museum National d'Histoire Naturelle 1993).

II.2 Réseau de mesures et site d'étude

Le site d'étude de Rostrenen se situe sur la partie amont du bassin. Les stations présentées sont essentiellement celles qui ont une potentielle influence sur les niveaux de nappe sur cette partie amont. Ces stations ont été sélectionnées selon leur positionnement géographique,

leur représentativité vis-à-vis des phénomènes qu'elles mesurent à partir d'études existantes, et de la disponibilité des données. Ainsi, le site dispose de plusieurs stations :

- Deux stations météorologiques :
    - Kerpert (avec les données pluie et évapotranspiration) (22092001),
    - Rostrenen (ayant des données pluie) (22266001),
- Une station hydrométrique :
    - Le Blavet à Kerien [Kerlouet] (J5202110),
- Une station piézométrique :
    - Piézomètre du Miniou (BSS000XFPD),
- Une station marégraphique :
    - Port-Tudy (71).

Le Tableau IV-1 récapitule les statistiques des différents types de données récoltées par les différents types de stations :

Tableau IV-1 - Statistiques des stations sur le bassin de la nappe du Blavet.

Nom de station	Variable	Unité	Pas de temps	Valeur max.	Valeur min.	Médiane	Moyenne	Période
Port-Tudy	Marée	m	heure	6,140	0,157	3,241	3,173	1994-2019
Piézomètre du Miniou	Niveau d'eau	m NGF	jour	247,40	234,56	239,83	240,21	2005-2019
Le Blavet à Kerien	Débit	m^3/s	jour	3,200	0,015	0,208	0,325	1994-2019
Kerpert-Coldevennec	ETP	mm	jour	8,9	0	1,6	1,87	1994-2019
Kerpert-Coldevennec	Précipitations	mm	jour	78,7	0	0,4	2,81	1994-2019
Rostrenen	Précipitations	mm	jour	151,9	0	0,4	2,85	1975-1987

Le piézomètre du Miniou dispose de données entre 2005 et 2019. Les autres types de données commencent en 1994 hormis le pluviomètre de Rostrenen (illustré sur le graphique présenté en Figure IV-1). Ainsi, pour la modélisation hydrogéologique, seules les données postérieures à 2005 seront utilisées. De même, les courtes lacunes dans les chroniques piézométriques sont comblées par des régressions linéaires, étant entendu que la variabilité de court terme est suffisamment faible pour se permettre ce type de traitement.

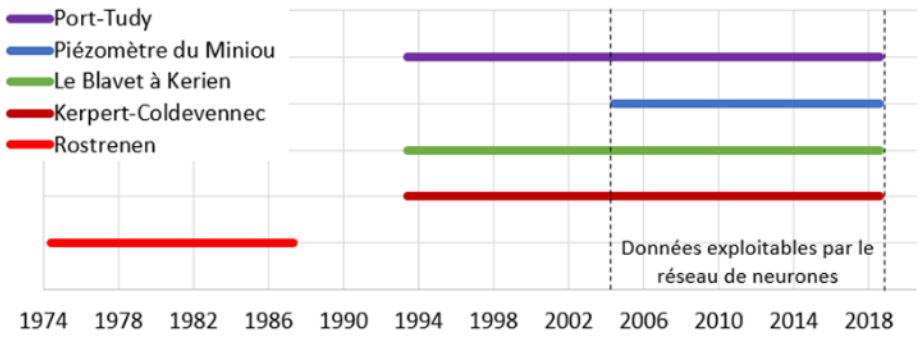
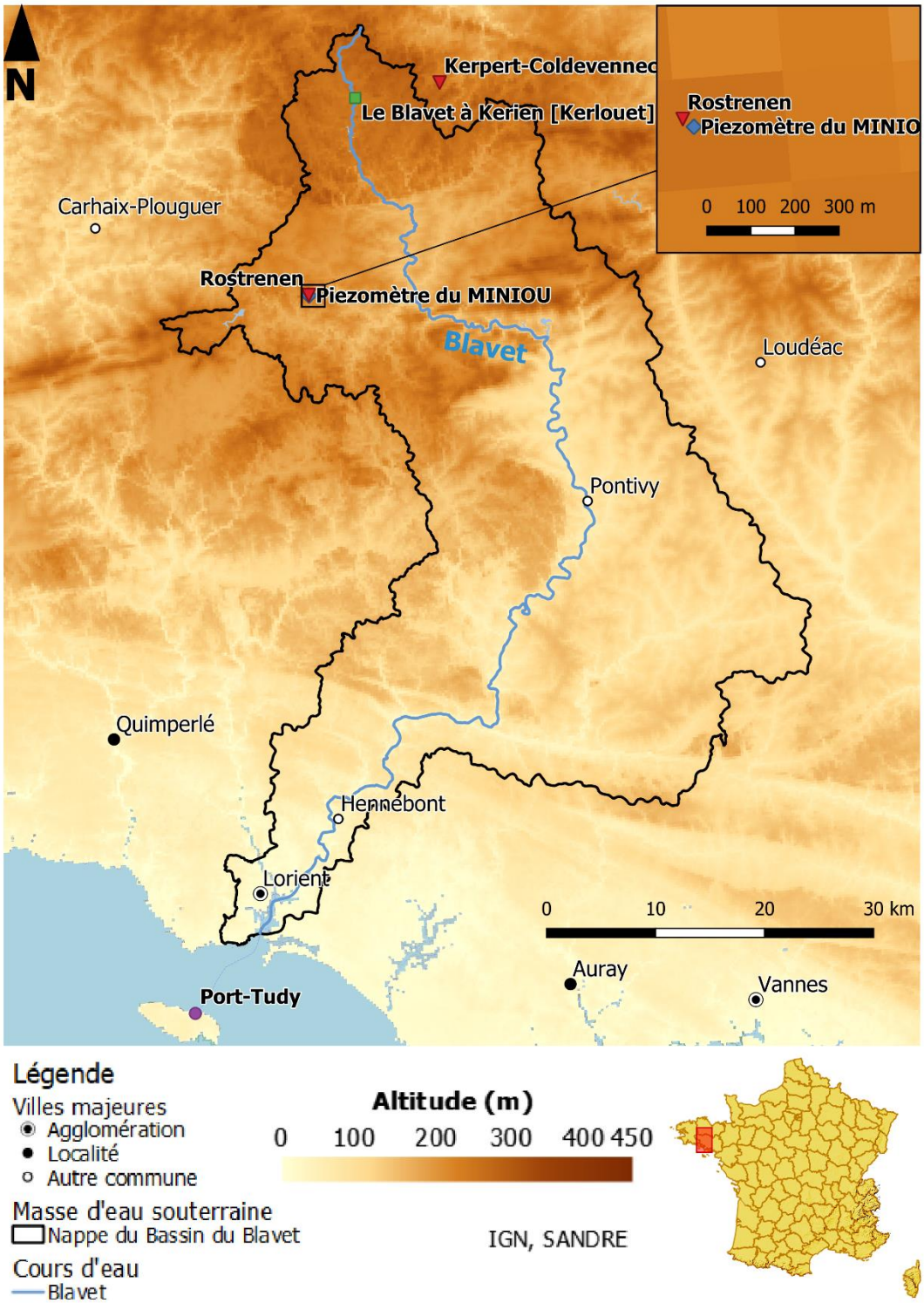


Figure IV-1 - Disponibilité des données (Nappe du Blavet).

La Carte IV-10 indique la position de chaque station sur le territoire d'étude.



Carte IV-10 - Localisation des stations de mesure du bassin du Blavet.

II.3 Résultats des expérimentations

II.3.1. Au pas de temps journalier : modélisation à court-terme

L'objectif est ici de concevoir un modèle capable de proposer une prévision à court-terme au pas de temps journalier afin de développer une méthodologie d'estimation de l'incertitude liée à la modélisation. Certains des résultats détaillés dans cette étude ont fait l'objet d'un acte de conférence (Akil, et al. 2019). Celui-ci est présenté à la suite de ce paragraphe et les approfondissements postérieurs à la publication sont présentés ensuite.

Quantification of Neural Network Uncertainties on the Hydrogeological Predictions by Probability Density Functions

Nicolas Akil[1,2], Guillaume Artigue[1], Michaël Savary[2], Anne Johannet[1], Marc Vinches[1]

[1]LGEI, IMT Mines Alès, Univ. Montpellier, 6 Avenue de Clavières. 30319 Alès Cedex FRANCE
[2]AQUASYS, 2 rue de Nantes, 44710 Port-Saint-Père FRANCE

guillaume.artigue@mines-ales.fr

Abstract. The risk of drought impacting the drinking water and agricultural production is worrying in the developed countries, especially in a changing climate context. To manage and prevent this phenomenon, real-time monitoring and predictive systems are emerging as the key solutions. In the field of artificial intelligence, neural networks are one of these predictive systems. This family of parameterized models is a composition of neuronal functions, which apply a non-linear transformation from their inputs to their outputs. These networks are able to learn a hydro(geo)logical system behaviour using a database composed of observed inputs (rainfall, evapotranspiration, etc.) and outputs (groundwater level, discharge, etc.), thanks to an algorithm minimizing a cost function between observed and simulated outputs. However, it remains difficult to assess the uncertainty generated by these models, possibly leading to misinterpretations by the end users. These uncertainties are mainly of three types. The first is related to the input data. Indeed, hydrosystems are surface elements whereas meteorological inputs are punctual elements. The interpolation error can, therefore, be significant because of the lack of knowledge between gauging stations. The second is the neural network model architecture itself. It is possible to deal with this source of uncertainty using regularization methods. Finally, the neural networks are submitted to uncertainties related to parameter initialization, before the training step. The initial parameters may have an important impact on the results. In this paper, we address the prediction of the *Blavet* groundwater level (*Bretagne*, France). In order to assess uncertainties, we will first focus on the parameters initialization of the model. Neuronal models are optimized using cross-validation and early stopping. Then, an ensemble model is realized, in which each member is the result of a unique set of parameters initialization. The purpose of the study is to define how many initializations are necessary to obtain a reasonable confidence interval for forecasts, with the smallest interval and the higher rate of observed points inside this interval. The best model will be determined using cross-validation scores thereby ensuring optimal robustness. We show that, in this case study, an ensemble model of 20 different initializations is sufficient to estimate uncertainty while preserving quality. In the second part, the resulting ensemble model will be used to estimate the global model uncertainty using probability density functions (pdf) applied to the distribution of groundwater level data and cross-validation scores of forecasts. It reveals that the groundwater level predictions are composed of two mixed distributions. Therefore, we will use the expectation-maximization algorithm (EM) to obtain parameters of mixed models. Mixed normal and mixed Gumbel laws, among five mixed distributions assessed, give the best groundwater distribution and are able to generate an abacus drawing uncertainty of model.

1. Introduction

Water related risks impact a large part of the population. On the one hand, floods frequently cause fatalities and damages. On the other hand, droughts affect drinking water and agricultural production. In a climate change context, with a rise of extreme phenomena frequency and duration [1], real-time monitoring and predictive systems are emerging as key solutions. To forecast the underground water level or river discharge, two main solutions emerge. The first relies on a deep knowledge of the considered system, dedicated to building physically based models. Unfortunately, this knowledge is most of the time difficult to acquire, leading to lower efficiency and high time and money consuming methods. The second one relies on the relation between system inputs and outputs, never making any strong hypothesis on the system operation, as long as the inputs are able to explain the outputs (such as rainfall for discharge). In the latter solution family, neural networks play an important role as they are known to be able to identify dynamical processes [2]. However, like any model would do, they generate uncertainties which are sometimes difficult both to quantify and to express, and can lead to misinterpretations by end users. These uncertainties mainly have three origins: input data (especially noise and spatial variability), model architecture and parameters initialization before the training step [3].

In the present paper, we propose a methodology to allow quantifying some of these uncertainties. The first part consists of building a reliable model to forecast groundwater level on a case study of north-western France. As [4] did with success on a southern France karst aquifer, the neural model will be optimized using cross-validation and early stopping. In the targeted region, we will state that the high quality of data and the low spatial variability are good enough reasons to overlook the uncertainties coming from data. Besides, the strict application of regularization methods should protect us from the uncertainties coming from the model architecture. We will thus focus on parameters initialization uncertainties, which are consequently considered as the main source of uncertainty in this case. The second part will consist of assessing this uncertainty, by generating ensemble models composed of members corresponding to different random sets of initial parameters. At last, the uncertainty will be expressed as a function of the state of the system in order to build an uncertainty pattern using statistical distributions. After having reminded the main characteristics of neural models and having presented the study area, we will describe more precisely the methodology used and discuss the results obtained.

2. Neural networks

o Background

A neuron is a mathematical operator applying a nonlinear transformation from its inputs to its outputs. In practice, it makes a weighted sum of its inputs transformed by an activation function thus giving an output. The weights of the sum are the parameters of the neural network.

Neurons can be combined in networks following architecture depending on: (i) the data describing the modelled system, and (ii) the purpose of the model. Neurons are then organised in layers of two types: output layers, whose outputs are the outputs of the model, and associated to measured values, and hidden layers, whose outputs are not associated to measured values [5].

o Neural networks models

Depending on how they deal with time, neural networks can be called “static” or “recurrent”. In the first case, time does not play any functional role and the input variables are exogenous (static model, (1)).

$$\hat{y}(k) = \varphi(\mathbf{x}(k), \dots, \mathbf{x}(k - n_r + 1), \mathbf{W}) \quad (1)$$

Where $\hat{y}(k)$ is the estimated output at the discrete time k; φ the nonlinear function implemented by the model; $\boldsymbol{x}$ the input vector; n_r the sliding time windows size that defines the length of the necessary exogenous data; $\mathbf{W}$ the vector of parameters.

In the second case, added to the exogenous variables, the result of the simulation at the previous time steps is used as an input variable (recurrent model, (2)).

$$\hat{y}(k) = \varphi(\hat{\boldsymbol{y}}(k-1), \dots, \hat{\boldsymbol{y}}(k-r); \boldsymbol{x}(k), \dots, \boldsymbol{x}(k-n_r+1), \mathbf{W}) \quad (2)$$

Where r is the order of the recurrent model.

A third case, called "feed-forward model" substitutes the recurrent input by the observations of the output at the previous time steps (3).

$$\hat{y}(k) = \varphi(\boldsymbol{y}(k-1), \dots, \boldsymbol{y}(k-r); \boldsymbol{x}(k), \dots, \boldsymbol{x}(k-n_r+1), \mathbf{W}) \quad (3)$$

Where $\boldsymbol{y}(k)$ is the observed value of the simulated variable at the discrete time k.

Recurrent models are used when the noise affecting outputs is known to be higher than the one affecting inputs. In contrast, feed-forward models, in which previously observed outputs are used as input variables, are used when the noise affecting input data is known to be the highest [6, 7].

One of the most common neural network structures is a feed-forward model called "multilayer perceptron" (MLP), for which the universal approximation property has been shown by [8]. This property mainly states that that kind of model (Figure 1) can approach any differentiable function as soon as it has enough hidden neurons and a database of sufficient quality.

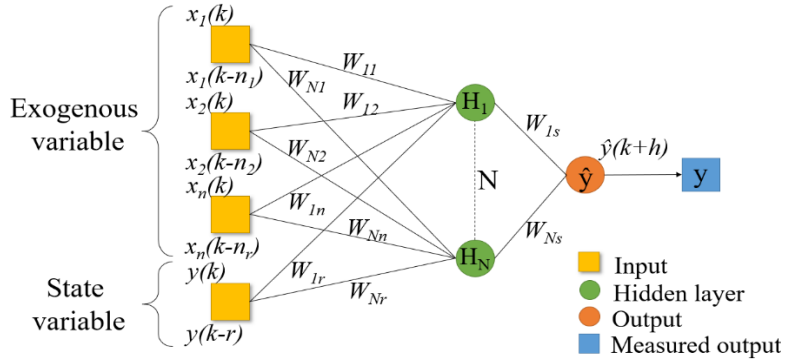


Figure 1. Multilayer Perceptron representation, with x_i, the exogenous variables, W_{ab}, the parameters, y, the measured output, $\hat{y}$, the simulated output, r, the order of the model, n_r, the input window width, H_N, hidden neuron, N, the number of hidden neurons, k, the discrete time and h, the lead time

Beside this property, a MLP also has the parsimony property, shown by [9], that states that if the function operated by the model depends on nonlinearly adjustable parameters, it is more parsimonious than if it depends linearly on its parameters. This property is even more valuable when the number of input variables increases.

- Neural networks training and overtraining: regularization methods

Training a neural network means finding the parameters vector such as a cost function is minimized. A training data set is dedicated to this task and a training algorithm is implemented. The purpose is to find the best approximation of the regression function to make the model adjust itself to the training set.

Consequently, a too complex model will adapt to the signal and to the random noise carried by this signal. On the other hand, a too simple model will not be able to adapt to the signal. This is called the "*bias/variance dilemma*" [10]. If the bias decreases too much, there is a risk that the model is adjusting to the noise, thus increasing the variance of its output, whereas a low variance and a low bias are desired. Therefore, the error on the training set is not a relevant estimator of the generalisation error.

We understand here that the result is highly depending on the data contained in the training set. To assess the generalization error of the model, one can use cross-validation [11]. Applied to the increasing model complexities, it provides a performance score that should decrease while the score the model obtains on the training set still increases, showing that the model performances in generalization are becoming too complex. Also, to stop training before the generalization capability decreases, one can use

early stopping [12], which consists of applying the model to a separate smaller data set and to stop training when the cost function reaches a minimum on this set.

In this paper, both these regularization methods will be applied, in order to ensure the robustness of the model [13]. Beside these issues, the neural models are sensitive to the initial random values of the parameters at the beginning of the training step. The modeller thus has to ensure that the result given by a model is robust and not depending on this random initialization. A certain number of these initializations may be used to create an ensemble that provides a range of possibilities [5, 14].

3. Study area: the *Blavet* groundwater basin

Located in north-western France, in the *Bretagne* region, the *Blavet* groundwater basin lies on more than 2,100 sq.km. It is mainly composed of metamorphic, magmatic and sandy rocks and its elevation ranges from -15 to 320 m.a.s.l. (Figure 2). Due to the low porosity of bedrock, this groundwater only exists thanks to regoliths and faults [15].

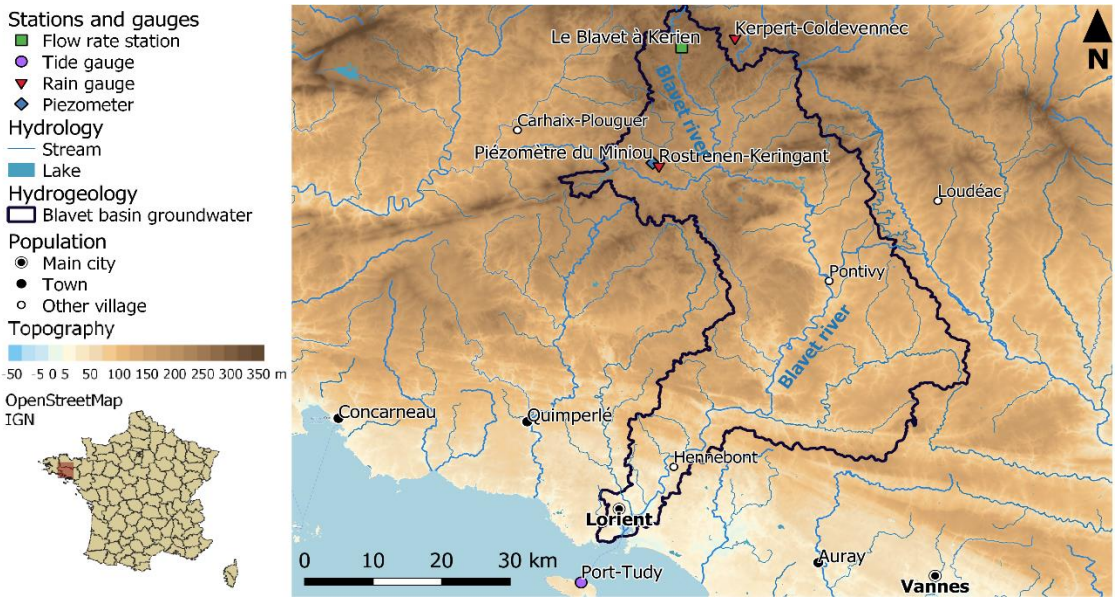


Figure 2. Map of the *Blavet* groundwater basin

Table 1. Statistics on *Blavet* time series

Station name	Variable	Unit	Time step	Max value	Min value	Median	Average
Miniou	Level	m.a.s.l	day	247.40	234.56	239.83	240.21
Blavet* in *Kerien	Discharge	m^3/s	day	3.200	0.015	0.208	0.325
Rostrenen-Keringant	Precipitation	mm	day	78.70	0.00	0.40	2.81
Rostrenen-Keringant	PET	mm	day	8.90	0.00	1.60	1.87

Data range from 2005 to 2019 at a daily time-step. They are issued from a meteorological station in *Rostrenen* (temperature, precipitation and evapotranspiration), a piezometer and a discharge station are also located in *Rostrenen* (Table 1). It has been shown that groundwater contributed at 57 % to river discharge at *Blavet* in *Kerien* during the 1994 to 2000 period [16]. Therefore, strong interactions between groundwater and surface water exist in this basin. From a climate point of view, the area is submitted to an oceanic climate, with an annual mean rainfall from 900 to 1,300mm and moderate temperatures, based on the 1981 to 2010 averages (*Météo France*).

4. Methodology

- Variable selection

Inputs are divided in a one-year-long-subsets, from January to December, except for first and last years that are not complete. These subsets are used for training, except for the one used for stop set (2017) and the one used for validation set that ranges from the end of 2018 to the beginning of 2019 (5 months). As the groundwater levels are available, a feed-forward model seems convenient. We also state that the water level measurement is accurate while rainfall and evapotranspiration measurements are not as much accurate, especially considering that there is only one meteorological station for a large surface and only one discharge station. In order to design the range of sliding input time windows (see section 2.2) that will be used for the model selection, cross correlations between available inputs and the output are calculated. They provide information on the explanatory power of the inputs on the outputs. They also show the response time, the memory effect of the system and the reasonable lead-time that could be reached. This lead-time is 3 days, which is long enough to make the output varying, and short enough to preserve forecast quality, even in the absence of rainfall forecast.

- Complexity selection

To ensure the robustness of the model, we explore different architectures and apply cross-validation. We thus select the depth of the sliding input time windows along with the complexity of the model. The cross-validation score S_{CV} is obtained by the median of the scores calculated for each combination of training / validation sets (Figure 3).

2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	
V	T	T	T	T	T	T	T	T	T	T	T	S	T	S_{2005}
T	V	T	T	T	T	T	T	T	T	T	T	S	T	S_{2006}
⋮														
T	T	T	T	T	T	T	T	T	T	T	V	S	T	S_{2016}
T	T	T	T	T	T	T	T	T	T	T	T	S	V	S_{2018}
													Median	S_{cv}

Figure 3. Schematic view of the cross-validation process; T is for training, V for validation, S for stop; S_{yyyy} is for the score calculated on the yyyy subset, S_{CV} is for the median of the scores

The score can be any relevant indicator of the performance. In the case of forecasting low variability variables, the persistence criterion [17] that measures the difference between a naïve forecast and the calculated forecast, is well suited (4).

$$C_P = 1 - \frac{\sum_k^n (Y_{p,k+h} - Y_{s,k+h})^2}{\sum_k^n (Y_{p,k} - Y_{p,k+h})^2} \quad (4)$$

Where $Y_{p,k}$ is the measured value at the discrete time k, $Y_{s,k}$ is the predicted value at the discrete time k and $Y_{p,k+h}$ the observed value at the discrete time $k+h$, with h, the lead time.

This selection process is applied for all the possible relevant architectures in conjunction with early stopping. Variables and complexity are thus appropriately selected, using a relevant score for the purpose of this study.

- Parameters initialization

After training, the parameters of the model possibly depend on their initial random value. Un uncertainty on the results is generated. The purpose of this study is to quantify this uncertainty, providing a range of possible outputs, called the prediction interval. This interval must be as small as possible (measured by the MPI criterion (Mean Prediction Interval), [18]) whereas it includes a maximum of observed values (which is measured by the PICP criterion (Prediction Interval Coverage Probability) [18]). The purpose of this step is to define how many random initializations are necessary to reach this optimum that will be calculated using SFPI (Spatial Frequency Prediction Interval) (5).

$$C_{SFPI} = \frac{C_{PICP}}{C_{MPI}} = \frac{\frac{1}{n}\sum_k^n f(Y_{p,k})}{\frac{1}{n}\sum_k^n (Y_{s_{max},k} - Y_{s_{min},k})}$$
$$\text{with } f(Y_{p,k}) = 1 \text{ if } Y_{p,k} \in [Y_{s_{min},k}; Y_{s_{max},k}], \text{ else } f(Y_{p,k}) = 0 \quad (5)$$

Where $Y_{p,k}$ is the measured value at the discrete time k and $Y_{s_{min},k}$ and $Y_{s_{max},k}$ the upper and lower bounds of the forecast interval.

Ensemble forecasts are thus calculated, with 3 to 120 members in each (respectively 3, 5, 10, 20, 30, 40, 50, 60, 80, 100 and 120 forecasts). For each of these intervals, the SFPI criterion is calculated, allowing defining an optimal number of initializations with a relevant benchmark.

- Representation of model uncertainty with probability density functions (pdf)

The uncertainty can be approached with *pdf* applied to output variable distribution. For this purpose, we fit 5 different *pdf* to the observed groundwater level distributions, which are Normal [19], Gumbel [20], Laplace [21], Cauchy [22] and Logistic laws [23], detailed respectively in (6), (7), (8), (9) and (10).

$$\mathcal{N}(\bar{x},\, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\bar{x})^2}{2\sigma^2}} \tag{6}$$

$$\mathcal{G}um(\bar{x},\beta) = \frac{e^{-\frac{(x-\bar{x})}{\beta}} e^{-e^{-\frac{(x-\bar{x})}{\beta}}}}{\beta} \tag{7}$$

$$\mathcal{L}ap(\bar{x},\, b) = \frac{1}{2b} e^{-\frac{(|x-\bar{x}|)}{b}} \tag{8}$$

$$\mathcal{C}au(x_0,\, a) = \frac{1}{\pi a\,(1+\left(\frac{x-x_0}{a}\right)^2)} \tag{9}$$

$$\mathcal{L}ogist(\bar{x},\, s) = \frac{e^{-\frac{(x-\bar{x})}{s}}}{s\left(1+e^{-\frac{(x-\bar{x})}{s}}\right)^2} \tag{10}$$

Where x is the variable, $\bar{x}$ its mean, x_0, its median, σ^2 its variance and a, b, β and s scale parameters.

In the case of *Blavet* groundwater level, a double distribution appears, corresponding to dry and wet seasons, leading to finding a mixed distribution. An expectation-maximization algorithm is applied to the 5 laws described, to obtain *pdf* parameters. The main purpose of this algorithm is to find a maximum likelihood set of parameters of a model when equations are not solvable. This algorithm is established in two steps. The first is the expectation step, where an expected value for log-likelihood parameters is defined, following assessed distribution or *pdf*. The second step is the maximization step, which consists of obtaining parameters that maximize the expected value, using an iterative process [24]. After finding parameters of each *pdf*, their inflection point is obtained, which allows separating mixed distribution in two *pdf*. Each *pdf* is multiplied by PICP, thus giving *pdf* of distribution included in prediction interval, by addition of the two *pdf*. The same process is used for water levels outside the prediction interval. It is lastly possible to rebuild mixed *pdf*, by addition of outside *pdf* and included *pdf*. The ratio between included prediction interval *pdf* and summoned *pdf* is calculated in order to obtain the probability to have a forecast inside the prediction interval for each groundwater level.

5. Results and discussions

- Model selected and deterministic results

The model selected by the combined variable and complexity selection processes, especially cross-validation with the persistence criterion has the architecture presented in Figure 4. This model is validated on the test set, which is totally unknown of the model, and gives the results shown in Figure 5, with Cp = 0.65. These results are qualitative enough to assess uncertainties without being dependent on the errors of the model.

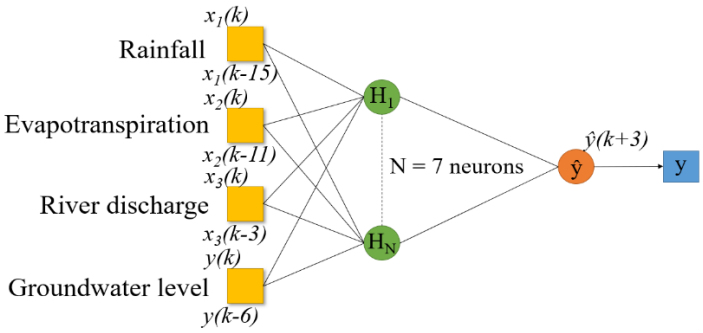


Figure 4. Selected architecture

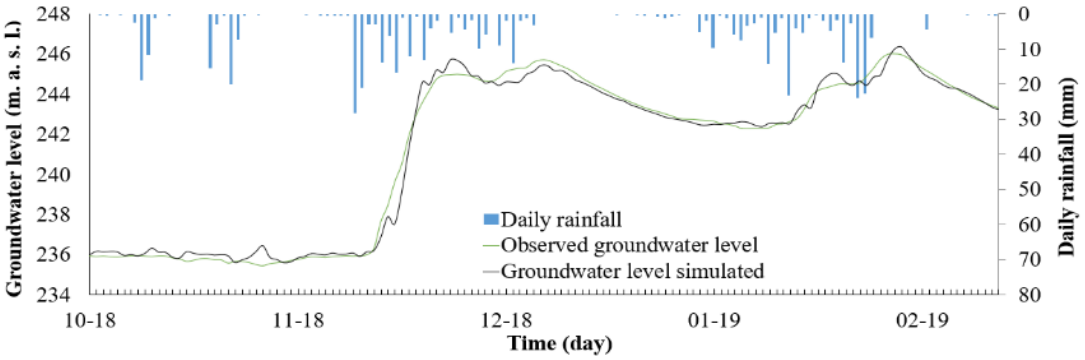


Figure 5. Validation of groundwater level forecasting at 3 days lead-time on the test set

- Optimal number of initializations and ensemble results

Once the architecture is selected and the model validated, ensemble forecasts are calculated. The number of members varies as described in section 4.3 while the performance is assessed using the SFPI criterion as a cross validation score.

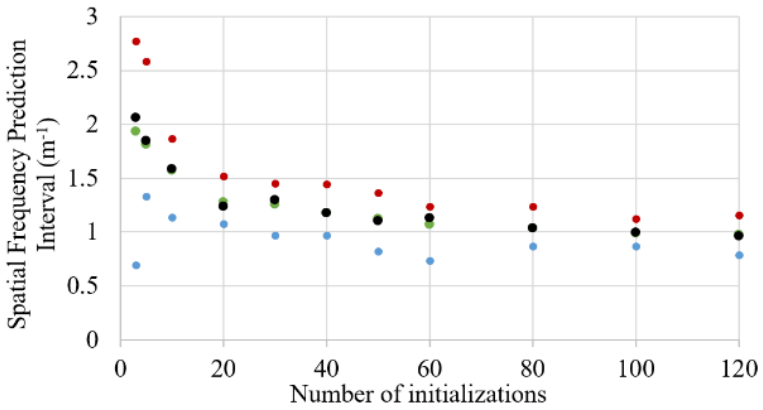


Figure 6. SFPI cross-validation score of ensemble models as a function of the number of members

As observed in Figure 6, a close to optimum performance is reached for 20 initializations. Even if the SFPI could be enhanced by using more members, the cost-benefit ratio (especially regarding calculation time) pleads in favour of the 20 members. Figure 7 shows the result for the test event, with a median PICP of 90% and an MPI of 0.41m, it reaches an SFPI of 1.25.

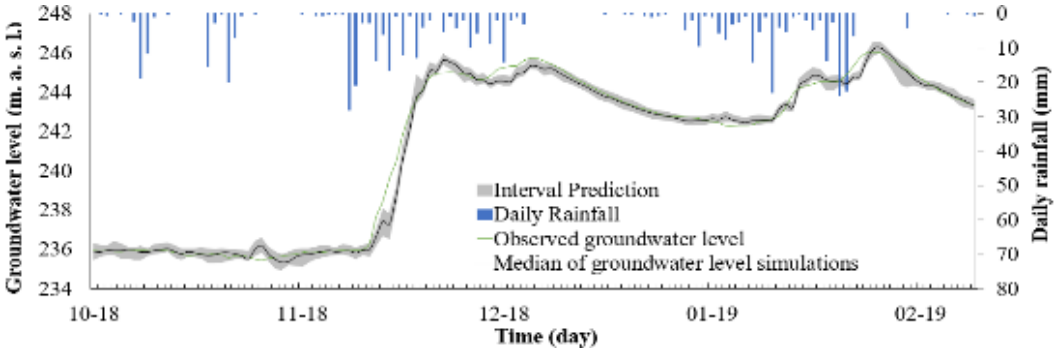


Figure 7. Representation of Figure 5 forecast with a prediction interval

- Representation of uncertainties

First, we need to focus on the ensemble forecast groundwater distribution. It shows higher occurrence frequencies for wet and dry periods (resp. >242 m.a.s.l. and <238 m.a.s.l.). Intermediate values are less frequent because of the quick evolution of levels occurring in them. Consequently, it appears that groundwater distribution is a double distribution and that an expectation-maximisation algorithm is needed to separate both *pdf*. We obtain inflexion points for each of the calculated laws that can be seen in Figure 8 and that range from 240.52 to 240.76 m.a.s.l. Using these values, mixed *pdf* are calculated and compared to groundwater distributions inside and outside the interval.

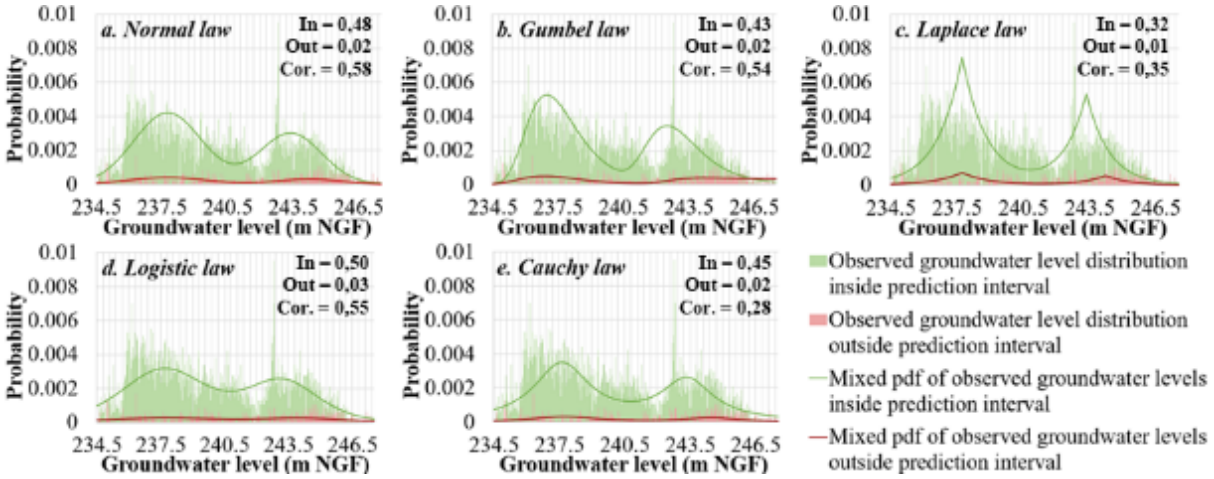


Figure 8. Mixed *pdf* of groundwater distributions. Determination coefficients are indicated in graphs ("In" means inside prediction interval, "out" means outside and "Cor." is model correctness)

The final step consists of calculating the abacus of model correctness, which is presented in Figure 9, comparing the calculated *pdf* with PICP scores observed for each piezometer level. PICP scores are ignored if they appear less than 3 times in the whole set. Except for Cauchy mixed *pdf*, each *pdf* is decreasing when approaching the extreme values. Normal, Gumbel and logistic mixed *pdf* seems to be the best of the 5 mixed laws, especially considering their determination coefficients. They also fit with observed groundwater level distributions. It remains a high uncertainty on the calculation of model correctness probability, determination coefficients rarely exceeding 0.5. Furthermore, PICP scores cannot be calculated for the highest groundwater levels, due to their low representation in the data set. This implies that abacus cannot ensure a reliable probability prediction for these levels. Lastly, this methodology requires large datasets, in order to have enough examples of high droughts or floods.

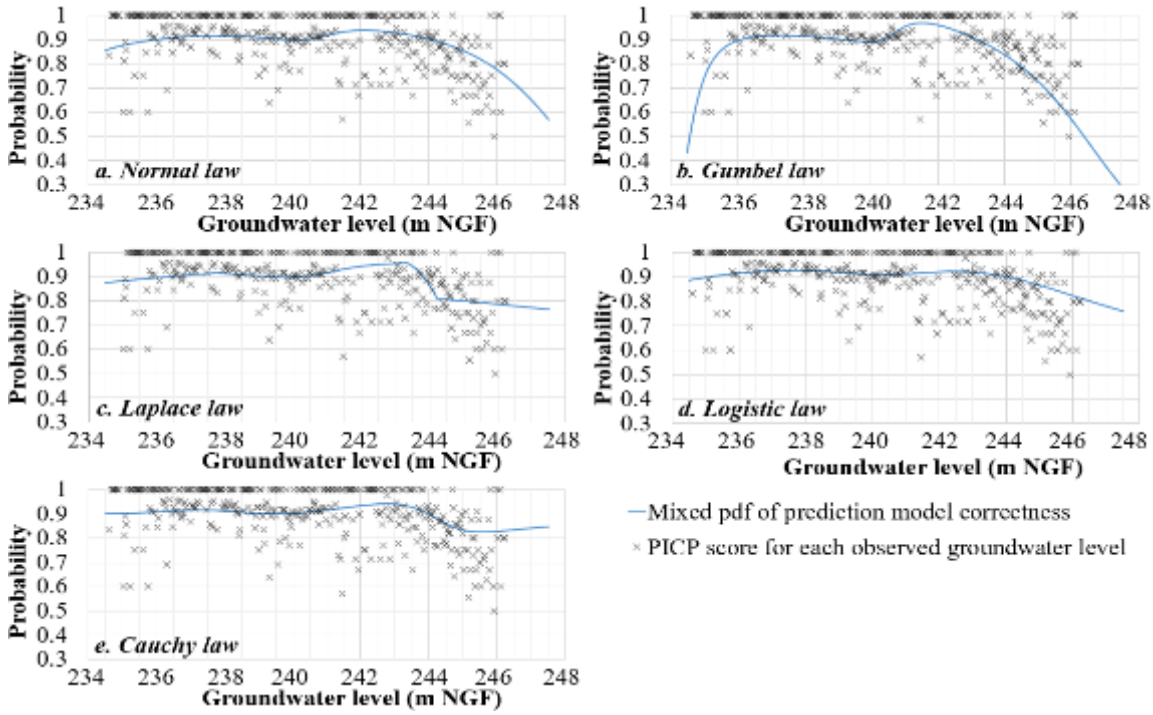


Figure 9. Abacus of mixed *pdf* of model correctness

6. Conclusions

This methodology allows forecasting neural network model uncertainty in the case of low uncertainty coming from the data and from the model architecture. This low uncertainty comes from a rigorous variable and complexity selection, applied to ensure the robustness of the model. The initialization of the parameters during the training step of the model is thus considered as the main source of uncertainty and it is first represented by ensemble forecasts, whose number of members must be defined as a function of the purpose of the model.

The use of the abacus created thanks to the law distribution fitting provides a significantly good representation of model uncertainty. With this abacus, it is possible to extrapolate uncertainties to unobserved value. The abacus is based on frequency distribution and as expected, it leads to: (i) a data set size sensibility and (ii) a low accuracy for an extreme event. Despite this limitation, this method is shown in our study, the possibility to use two different distribution laws; normal law and logistic law. Thanks to this final observation, the future work will focus on the application of multi-law distribution representation for the prediction uncertainty.

Acknowledgments

Authors would like to warmly thank the BRGM, which provided the dataset, J. Nicolas (BRGM) and S. Pistre (Univ. Montp.) for helpful discussions and the ANRT, which contributed to funding this work.

References

[1] D. L. Hartmann, "Observations: Atmosphere and Surface", *Climate Change 2013: The Phys. Sc. Basis*. Cont. of WG I to the 5th Ass. Report of the Inter. Panel on CC, 2, pp 159–254, 2013.

[2] L. Kong-A-Siou, "Neural networks for karst groundwater management: case of the Lez spring (Southern France)", *Env. Earth Sciences*, Volume: 74 Issue: 12, 7617-7632, 2015.

[3] F. Bourgin, "Comment quantifier l'incertitude prédictive en modélisation hydrologique? Travail exploratoire sur un grand échantillon de bassins versants", *PhD Thesis*, AgroParisTech, 230 p., HAL: tel-01130084, v. 2, 2014.

[4] L. Kong-A-Siou, A. Johannet, V. Estupina-Borrell, and S. Pistre, "Optimization of the generalization capability for rainfall–runoff modeling by neural networks: the case of the Lez aquifer (southern France)", *Environmental Earth Sciences*, 65(8), 2365-2375, 2012.

[5] G. Dreyfus, Neural networks, Methodology and applications, Springer, Berlin, 2005.
[6] O. Nerrand, P. Roussel-Ragot, L. Personnaz, G. Dreyfus, and S. Marcos, "Neural Networks and Nonlinear Adaptive Filtering: Unifying Concepts and New Algorithms.", Neural Comp 5(2), 165-199, 1993.
[7] V. Taver, A. Johannet, V. Borrel-Estupina, and S. Pistre, "Feed-forward vs recurrent neural network models for non-stationarity modelling using data assimilation and adaptivity", HSJ – J. des Sciences Hydrologiques, Vol. 60, Issue 7-8, Special Issue, 1242-1265, 2015.
[8] K. Hornik, M. Stinchombe and H. White, "Multilayer Feedforward Networks are Universal Approximators", *Neural Networks*, vol. 2, pp. 359–366, 1989.
[9] R. Barron, "Universal Approximation Bounds for Superpositions of a Sigmoidal Function." IEEE Trans. Inf. Theor. 39 (3): 930-45, 1993.
[10] S. Geman, E. Bienenstock and R. Doursat, "Neural Networks and the Bias/Variance dilemma", *Neural Computation*, vol. 4, pp. 1–58, 1992.
[11] M. Stone, "Cross-Validatory Choice and Assessment of Statistical Predictions", *Journal of the Royal Statistical Society*, Series B (Methodological), vol. 36, no. 2, pp. 111–147, 1974.
[12] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P. Y. Glorennec, H. Hjalmarsson and A. Juditsky, "Nonlinear Black-box Modeling in System Identification: a unified overview", *Automatica*, vol. 31, no. 12, pp. 1691–1724, 1995.
[13] L. Kong-A-Siou, A. Johannet, V. Estupina-Borrell, and S. Pistre, "Optimization of the generalization capability for rainfall–runoff modeling by neural networks: the case of the Lez aquifer (southern France)", Environmental Earth Sciences, 65(8), 2365-2375, 2012.
[14] T. Darras, A. Johannet, B. Vayssade, L. Kong-A-Siou and S. Pistre, "Influence of the Initialization of Multilayer Perceptron for Flash Flood Forecasting: Design of a Robust Model.", *International Work-Conference on Time Series Analysis*, 12 pp., 225–235, 2014.
[15] F. Lucassou and B. Mougin, "Essai d'élaboration d'indicateurs piézométriques pour la gestion quantitative AEP dans le département des Côtes d'Armor – Rapport final", *BRGM*, RP-53326-FR, 80 pp., 2015.
[16] B. Mougin, A. Carn, N. Debeglia, J. Perrin and E. Thomas, "SILURES Bretagne (Système d'Information pour la Localisation et l'Utilisation des Ressources en Eaux Souterraines)", *BRGM*, RP-52825-FR, 62 pp., 2004.
[17] P. K. Kitanadis and R. L. Bras, "Real-time forecasting with a conceptual hydrologic model: 2. Applications and results.", *Water Resources Research*, vol. 16, no. 6, pp. 1034–1044, 1980.
[18] D. L. Shrestha and D. P. Solomatine, "Machine learning approaches for estimation of prediction interval for the model output.", *Neural Networks*, vol. 19, pp. 225–235, 2006.
[19] R. A. Fisher, "On the mathematical foundations of theoretical statistics", *Philosophical Transactions of the Royal Society*, vol. 9, no. 222, pp. 309–368, 1922.
[20] E. J. Gumbel, "Méthodes graphiques pour l'analyse des débits de crues", *Revue de statistique appliquée*, vol. 5, no. 2, pp. 77–89, 1957.
[21] P. S. Laplace, "Mémoire sur la probabilité des causes par les évènements". *Mémoires de l'Academie Royale des Sciences*, Presentés par Divers Savan, vol. 6, pp. 621–656, 1774.
[22] L. Cauchy, "Sur les résultats moyens d'observations de même nature, et sur les résultats les plus probables". *Comptes Rendus Hebd. Seances Acad. Sci. Paris*, vol. 37, pp. 198–206, 1853.
[23] P. F. Verhulst, "Recherches mathématiques sur la loi d'accroissement de la population.". *Nouveaux mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles*, Académie de Bruxelles, vol. 18, pp. 1–38, 1845.
[24] P. Dempster, N. M. Laird and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm", *Journal of the Royal Statistical Society*. Series B (Methodological), vol. 39, no. 1, pp. 1–38, 1977.

II.3.1.a. Détermination du nombre de modèles optimal pour représenter l'incertitude de modélisation

La représentation de l'incertitude de modélisation est établie à partir de la dispersion de plusieurs modèles aux architectures identiques, différenciés uniquement par les paramètres obtenus à partir d'autant d'initialisations différentes de l'apprentissage. Le nombre de modèles optimal est ici déterminé en calculant le score de validation croisée médian lorsque ce score correspond au critère *Prediction Confidence Criterion* (C_{PC}) parmi les différents paramètres obtenus à partir d'initialisations différentes. Le critère C_{PC} a été nommé *Spatial Frequency Prediction Interval* (C_{SFPI}) dans l'acte de conférence (Akil, et al. 2019).

Le résultat obtenu dans la publication présentée précédemment détermine que 20 initialisations semblent être nécessaires pour obtenir un pourcentage de prévisions inclus dans l'intervalle de prévision suffisamment grand tout en ayant un intervalle de prévision le plus fin possible.

Afin d'observer le comportement du C_{PC} en fonction de la période de l'année, nous recalculons le résultat mensuellement par le biais de moyennes interannuelles appliquées aux différents sous-ensembles d'apprentissage pour 10, 20, 30 et 40 initialisations (Figure IV-2).

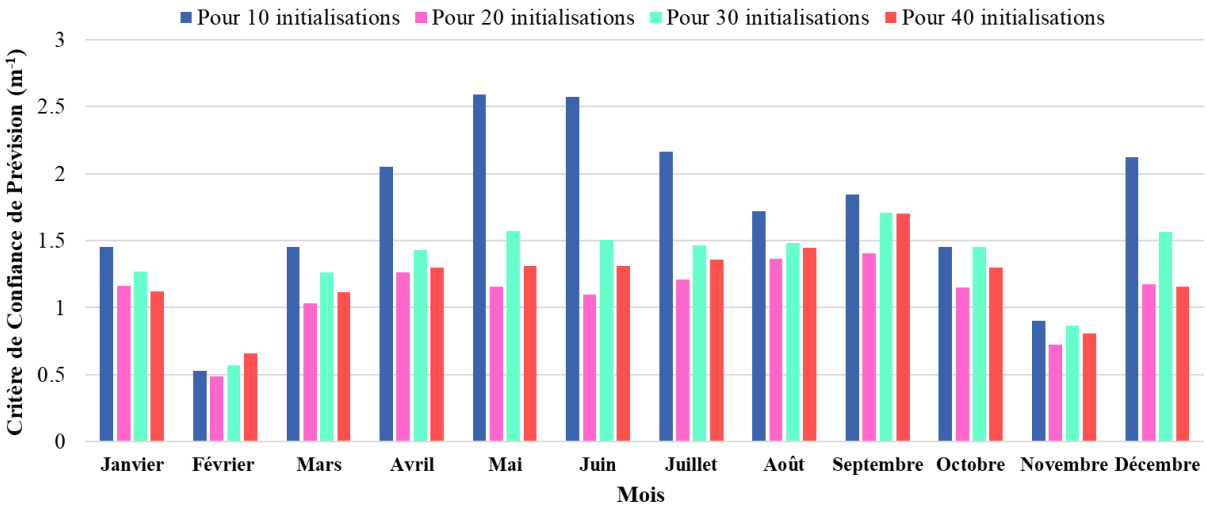


Figure IV-2 - Evolution mensuelle du *Prediction Confidence Criterion* en fonction du nombre d'initialisations.

Dans un premier temps, il est possible de voir que le critère *Prediction Confidence Criterion* (C_{PC}) est systématiquement plus important pour 10 initialisations (hormis pour le mois de février). Cela est d'autant plus visible pour les mois d'avril à août, ainsi que pour le mois de décembre. La période d'avril à août correspond ici au moment où la nappe du Blavet passe d'un état de hautes eaux à un état d'étiage.

Dans un second temps, les valeurs obtenues pour le critère C_{PC} avec 20, 30 et 40 initialisations sont globalement similaires pour l'ensemble des mois. Seul le mois de décembre fait

exception, où le critère C_{PC} est légèrement plus important lorsque 30 initialisations sont réalisées.

Ainsi, aucune variation mensuelle ne permet de favoriser un nombre d'initialisations au détriment d'un autre dans ce cas. Le choix initial peut être conservé.

Cette étude a toutefois pu mettre en évidence l'évolution du critère en fonction de l'état de recharge de la nappe.

II.3.1.b. Application des densités de probabilités

II.3.1.b.i. Modèles de justesse

Afin de construire les modèles de justesse, il est nécessaire d'observer la distribution des niveaux piézométriques. Dans l'acte de conférence, seules cinq lois à densités de probabilités (Normale, Gumbel, Laplace, Logistique et Cauchy) ont été inférées selon une double distribution des niveaux piézométriques. De plus, les travaux menés dans cette étude n'ont pas pris en compte l'évolution des niveaux piézométriques (Akil, et al. 2019). Pour rappel, ces niveaux d'eau représentent plus ou moins fidèlement l'état de la ressource en eau souterraine. Afin de compléter ces résultats, nous proposons ici de :

- tester l'inférence de quatre lois supplémentaires,
- utiliser une triple distribution pour améliorer sa représentativité vis-à-vis des données disponibles,
- discriminer les périodes de hausse et de baisse des niveaux piézométriques.

Ces éléments constituent l'aboutissement des premiers résultats présentés dans l'acte de conférence.

Afin de limiter les biais statistiques causés par les calculs de densités de probabilité qui vont suivre, les niveaux piézométriques ont été regroupés par classes de 3 cm.

La Figure IV-3 montre l'inférence d'une double loi Normale et d'une triple loi Normale sur la distribution des niveaux piézométriques inclus dans et en dehors des niveaux piézométriques observés. Les coefficients de corrélation obtenus sont plus élevés pour la triple distribution de la loi Normale que la double distribution (gain d'environ 0,1 en corrélation). L'inférence de la loi Normale est donc meilleure lorsqu'elle admet une triple distribution. C'est donc ce qui sera appliqué dans ces travaux.

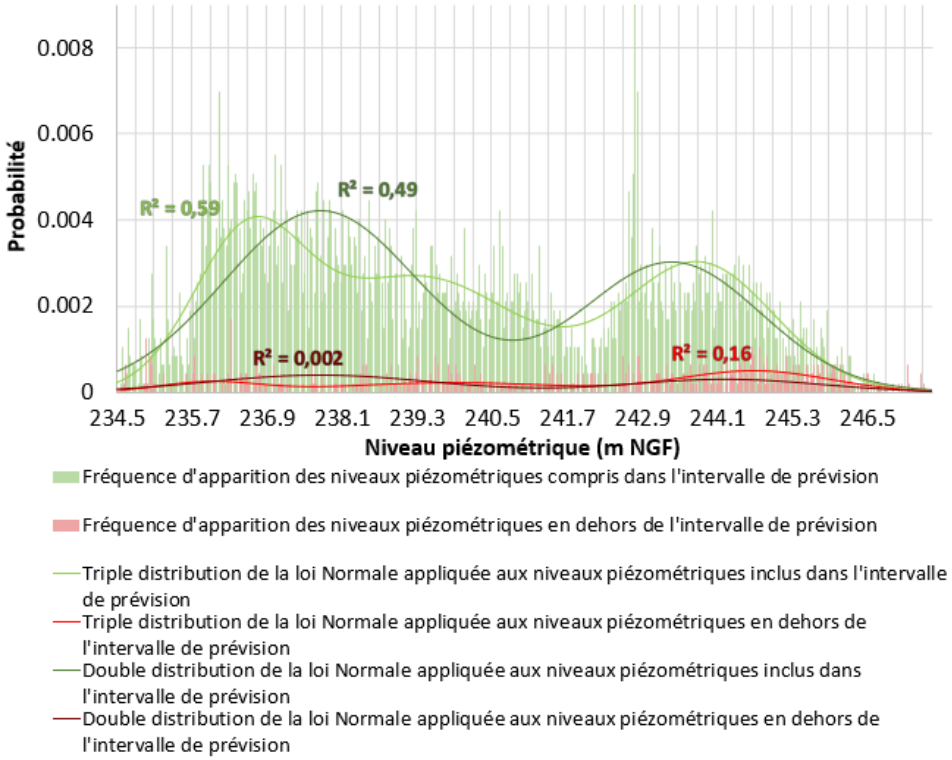


Figure IV-3 - Comparaison entre une double et une triple loi Normale inférant la distribution des niveaux piézométriques observés.

Il apparait par ailleurs opportun de distinguer deux grandes classes de comportements de l'hydrosystème. Les résultats des modèles de justesse seront donc présentés dans un premier temps lorsque les niveaux piézométriques sont en augmentation (période de recharge de la nappe), puis dans un second temps lorsque ces niveaux diminuent (période de vidange de la nappe).

Recharge de la nappe

Les niveaux piézométriques en augmentation sont classés en fonction de leur présence ou non au sein de l'intervalle de prévision (PI). Ces niveaux sont ensuite inférés par neuf lois à densités de probabilités, contre cinq auparavant. Les indicateurs d'inférence (coefficients de corrélation de Pearson r^2 et distances de Kolmogorov-Smirnov D_{KS}) sont présentés dans le Tableau IV-2 pour chaque loi à densité de probabilité inférée. Ces neuf lois sont présentées sur les graphiques présents dans la Figure IV-4.

Tableau IV-2 - Indicateurs d'inférence des lois à densité de probabilité.

Lois à densité de probabilité	r^2 (Inclus dans PI)	r^2 (Hors de PI)	D_{KS} (Inclus dans PI)	D_{KS} (Hors de PI)
Loi Normale	0,68	0,22	0,011	0,003
Loi de Gumbel	0,66	0,06	0,011	0,051
Loi de Laplace	0,66	0,18	0,012	0,004
Loi du Cosinus surélevé	0,69	0,25	0,006	0,002
Loi de Cauchy	0,68	0,18	0,049	0,009
Loi Logistique	0,70	0,21	0,033	0,005
Loi de Slash	0,63	0,01	0,056	0,034
Loi Bhattacharjee	0,68	0,13	0,007	0,026
Loi de Huber	0,68	0,12	0,009	0,026

La corrélation r^2 entre les lois inférées et à la distribution des niveaux piézométriques est proche de 0,65 pour les niveaux inclus dans l'intervalle de prévision. En revanche, cette corrélation est beaucoup plus faible (de l'ordre de 0,2) entre la distribution des hauteurs d'eau observées en dehors de l'intervalle de prévision et les lois inférées. Cette faible corrélation est à relier avec la faible fréquence d'apparition (6% des cas) d'une telle configuration. En considérant uniquement le coefficient de corrélation, les lois du Cosinus surélevé et Logistique semblent être les lois les mieux adaptées pour inférer la distribution des niveaux piézométriques.

Les distances de Kolmogorov-Smirnov pour chaque loi inférée tendent vers 0, signifiant qu'elles transcrivent assez correctement la distribution des hauteurs d'eau souterraines. La distance la plus faible, et par conséquent la meilleure loi, est obtenue par la loi du Cosinus surélevé, pour le cas où les niveaux sont dans et en dehors de l'intervalle de prévision.

Les différents indicateurs obtenus montrent ici que les lois les mieux inférées à la distribution des niveaux piézométriques observés semblent être la loi du Cosinus surélevé et la loi Normale. Les lois de Bhattacharjee et de Huber sont également bien corrélées avec la distribution des niveaux piézométriques observés compte tenu de leur dépendance avec la loi Normale.

Les neuf lois inférées à la distribution des niveaux piézométriques en augmentation sont présentées sur la Figure IV-4. Conformément aux indicateurs d'inférence calculés précédemment, les courbes des densités de probabilités obtenues semblent suivre convenablement les valeurs piézométriques observées, à l'exception d'un pic autour 242,85 m NGF. Il est donc fréquent de retrouver un niveau de nappe entre 242,5 m et 243 m durant les remontées de nappe, malgré les variations hydro-climatiques interannuelles potentielles.

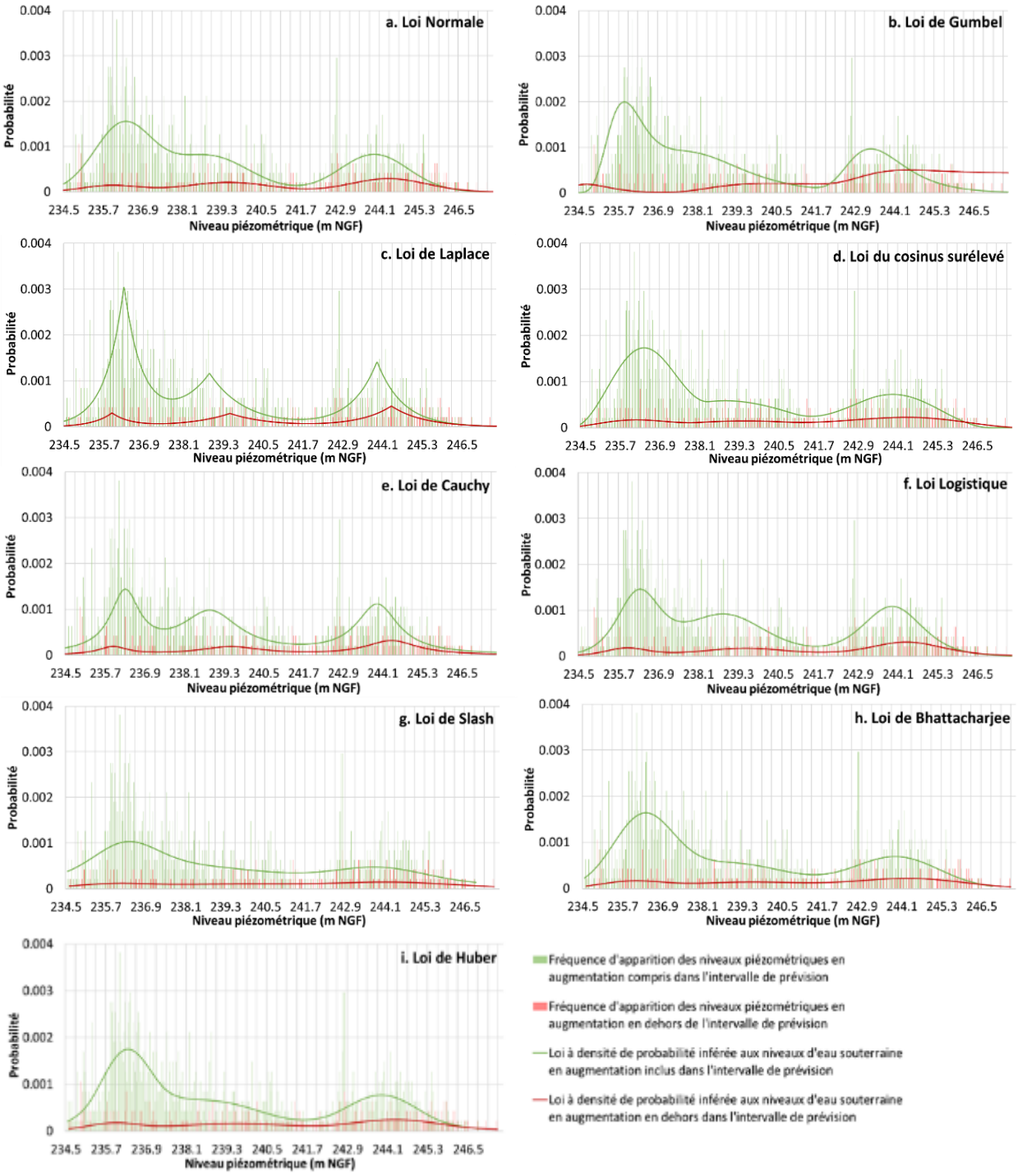


Figure IV-4 - Densités de probabilité inférées à la distribution des niveaux piézométriques en augmentation.

A partir des lois inférées, il est possible de construire des modèles de justesse, qui indiquent dans notre cas la probabilité que les niveaux piézométriques en augmentation soient inclus dans l'intervalle de prévision. Ces modèles peuvent être utilisés comme abaques pour évaluer la probabilité de justesse d'un résultat de modélisation du modèle d'étude pour la prévision des niveaux d'eau souterraine à un horizon de trois jours.

Les coefficients de corrélation entre les différents critères PICP obtenus à partir des valeurs piézométriques observées et le modèle de justesse obtenu par les lois à densités de

probabilité sont calculés pour chacune des neuf lois. De plus, les pourcentages de ces critères PICP inclus dans la marge d'erreur (M_E), établis grâce à la fonction erreur de Gauss, sont également calculés. Ces deux indicateurs sont présentés sur le Tableau IV-3 suivant :

Tableau IV-3 - Corrélation des modèles de justesse pour les niveaux piézométriques en augmentation.

Lois à densité de probabilité	***r²***	***M_E (%)***
Loi Normale	0,58	78,2
Loi de Gumbel	0,57	64,3
Loi de Laplace	0,54	76,9
Loi du Cosinus surélevé	0,58	79,0
Loi de Cauchy	0,47	77,3
Loi Logistique	0,58	77,7
Loi de Slash	0,50	78,2
Loi Bhattacharjee	0,51	79,4
Loi de Huber	0,50	80,3

Ici, les lois Normale, Logistique et Cosinus surélevé présentent les meilleurs coefficients de corrélation. Ces coefficients sont proches de 0,5 pour toutes les lois. Les lois de Bhattacharjee, Huber et Cosinus surélevé ont quant à elles les meilleurs pourcentages de critères inclus dans la marge d'erreur, avec près de 80% de points inclus dans cet intervalle.

Les modèles de justesse réalisés pour les niveaux piézométriques en augmentation présentent des similitudes selon les lois utilisées. En effet, la Figure IV-5 montre que ces densités de probabilités ont une même tendance globale, qui tend à diminuer la probabilité de justesse – correspondant à la probabilité que les valeurs mesurées soient incluses dans l'intervalle de prévision – lorsque les niveaux piézométriques augmentent. De plus, lorsque ces probabilités sont les plus importantes, le nombre de critères PICP obtenus à partir des valeurs observées étant à une probabilité de 1 est important également.

Il est toutefois possible de répartir ces neuf modèles de justesse en trois classes. La première classe regroupe les modèles utilisant la loi Normale, la loi de Gumbel et la loi du Cosinus surélevé. Dans cette classe, il est possible d'observer que les amplitudes des probabilités calculées par les densités de probabilité sont importantes. En effet, pour des valeurs piézométriques supérieures à 246 m NGF, les probabilités obtenues sont relativement faibles (inférieures à 0,4) et en forte diminution. Pour la loi du Cosinus surélevé et la loi de Gumbel, les probabilités obtenues sont également faibles lorsque les niveaux piézométriques sont faibles.

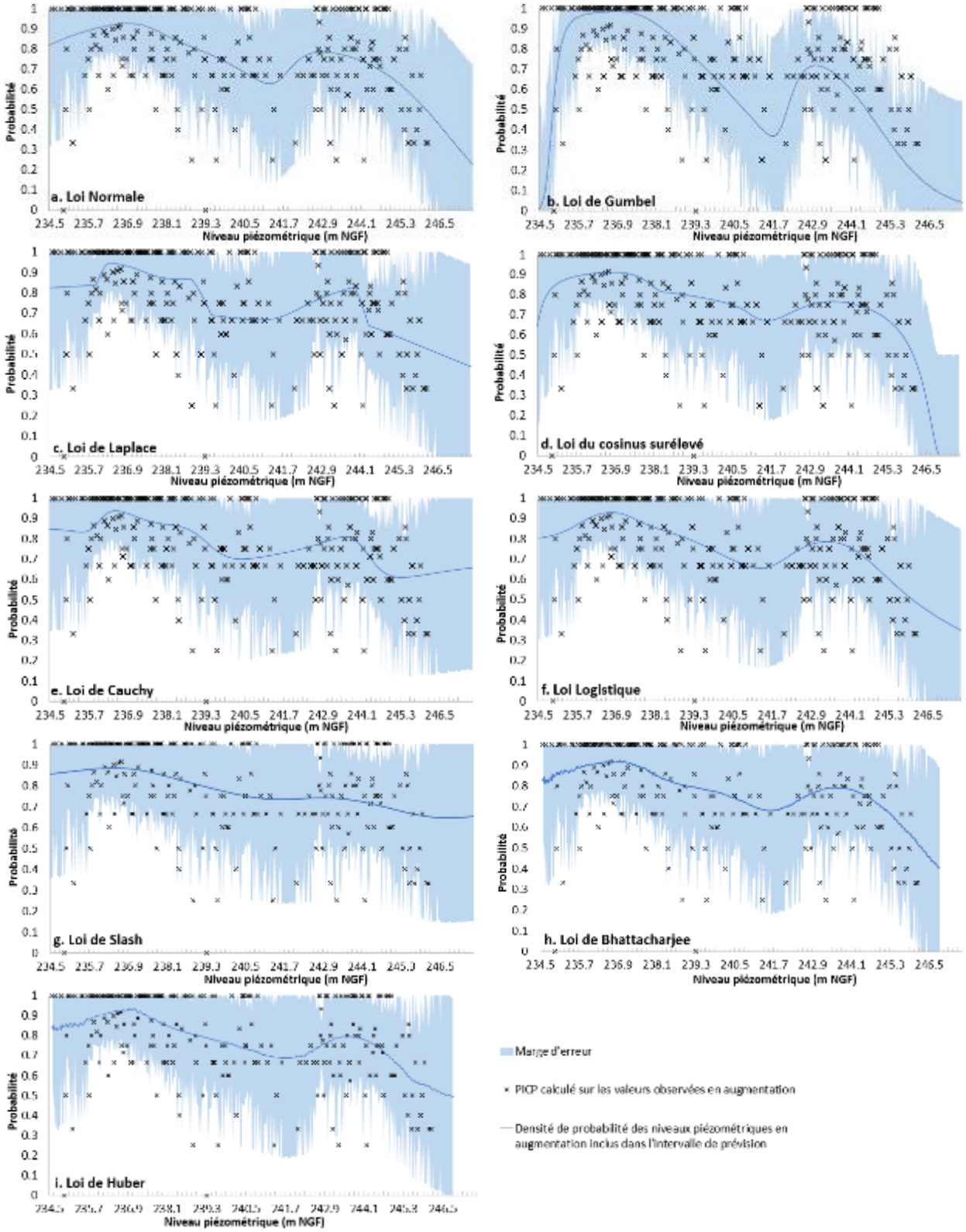


Figure IV-5 - Modèles de justesse pour les niveaux piézométriques en augmentation.

La seconde classe contient les modèles avec les lois de Laplace, de Cauchy et de Slash. Cette classe contient les modèles relativement plats, où les probabilités n'évoluent qu'entre 0,5 à 0,9. Pour les lois de Cauchy et de Slash, et contrairement aux sept autres lois, les probabilités

augmentent avec les niveaux piézométriques pour des fortes valeurs de hauteurs d'eau souterraine. Ce résultat pourrait apparaître peu cohérent avec l'ensemble d'apprentissage du modèle qui contient, par définition, peu de valeurs extrêmes. Pour autant, il témoigne probablement simplement du fait que la sensibilité de ces lois aux valeurs extrémales est plus forte (moins de variabilité de l'inférence). Il est probable que ces lois ne représentent donc pas correctement le comportement du modèle qui est censé être moins performant sur les valeurs extrêmes.

La troisième classe regroupe les modèles de justesse de la loi Logistique, la loi de Bhattacharjee et la loi de Huber. Cette classe présente des caractéristiques intermédiaires aux deux autres classes, avec des amplitudes de probabilité modérées et des baisses de probabilités assez prononcées pour des grandes valeurs de niveaux piézométriques.

Suite à ces expérimentations, la loi du Cosinus surélevé semble être la loi la mieux adaptée pour inférer la distribution des niveaux piézométriques en augmentation sur le bassin de la nappe du Blavet pour une prévision à trois jours. Parmi les neuf lois étudiées, cette loi possède les meilleurs indicateurs d'inférence vis-à-vis de la distribution des hauteurs d'eau souterraine croissantes, ainsi que des meilleurs indicateurs vis-à-vis des modèles de justesse élaborés. Il était donc pertinent d'expérimenter des lois supplémentaires.

Vidange de la nappe

Comme pour le cas précédent, les niveaux piézométriques en diminution sont classés en fonction de leur présence ou non au sein de l'intervalle de prévision (PI). Ces niveaux d'eau souterraine en diminution sont inférés par les mêmes lois à densités de probabilités utilisées pour le cas de recharge de la nappe. Les différents indicateurs d'inférence (coefficients de corrélation de Pearson r^2 et distances de Kolmogorov-Smirnov D_{KS}) sont présentés dans le Tableau IV-4.

Tableau IV-4 - Indicateurs d'inférence des lois à densité de probabilité.

Lois à densité de probabilité	r^2 (Inclus dans PI)	r^2 (Hors de PI)	D_{KS} (Inclus dans PI)	D_{KS} (Hors de PI)
Loi Normale	0,75	0,43	0,017	0,003
Loi de Gumbel	0,75	0,31	0,017	0,041
Loi de Laplace	0,64	0,47	0,026	0,003
Loi du Cosinus surélevé	0,77	0,49	0,011	0,002
Loi de Cauchy	0,67	0,48	0,077	0,005
Loi Logistique	0,75	0,49	0,016	0,003
Loi de Slash	0,73	0,36	0,143	0,026
Loi Bhattacharjee	0,73	0,43	0,014	0,023
Loi de Huber	0,73	0,44	0,019	0,032

La corrélation r^2 entre les lois inférées et à la distribution des niveaux piézométriques est proche de 0,7 pour les niveaux inclus dans l'intervalle de prévision contre 0,45 pour les niveaux d'eau souterraine observés en dehors de l'intervalle de prévision. Comme pour le cas précédent, la corrélation est plus faible pour les hauteurs d'eau en dehors de l'intervalle de prévision. En effet, moins de 4 % de l'historique de données concerne ces niveaux en diminution en dehors de l'intervalle de prévision. La loi du Cosinus surélevé, et dans une moindre mesure, les lois Normale et Logistique sont les lois qui disposent des meilleurs coefficients de corrélation.

Les distances de Kolmogorov-Smirnov pour chaque loi inférée sont proches de 0, ce qui indique que chacune des lois suit bien la distribution des niveaux piézométriques. Comme pour l'état de recharge de la nappe, la plus petite distance de Kolmogorov-Smirnov est obtenue par la loi du Cosinus surélevé, quand les niveaux sont dans et en dehors de l'intervalle de prévision. Les lois Normale et Logistique disposent également de faibles distances. En revanche, la loi de Slash obtient une distance de Kolmogorov-Smirnov relativement importante (environ dix fois supérieure vis-à-vis des distances obtenues par les autres lois).

Toutes les lois étudiées qui sont inférées à la distribution des hauteurs d'eau souterraine en diminution sont montrées dans la Figure IV-6 . La plupart des densités de probabilités (*pdf*) suivent bien les fréquences d'apparition de chaque niveau piézométrique observé, à l'exception de la loi de Slash pour les valeurs incluses dans l'intervalle de prévision. En effet, il est possible de constater que la *pdf* est légèrement inférieure aux valeurs observées, justifiant ainsi la grande distance de Kolmogorov-Smirnov pour cette loi.

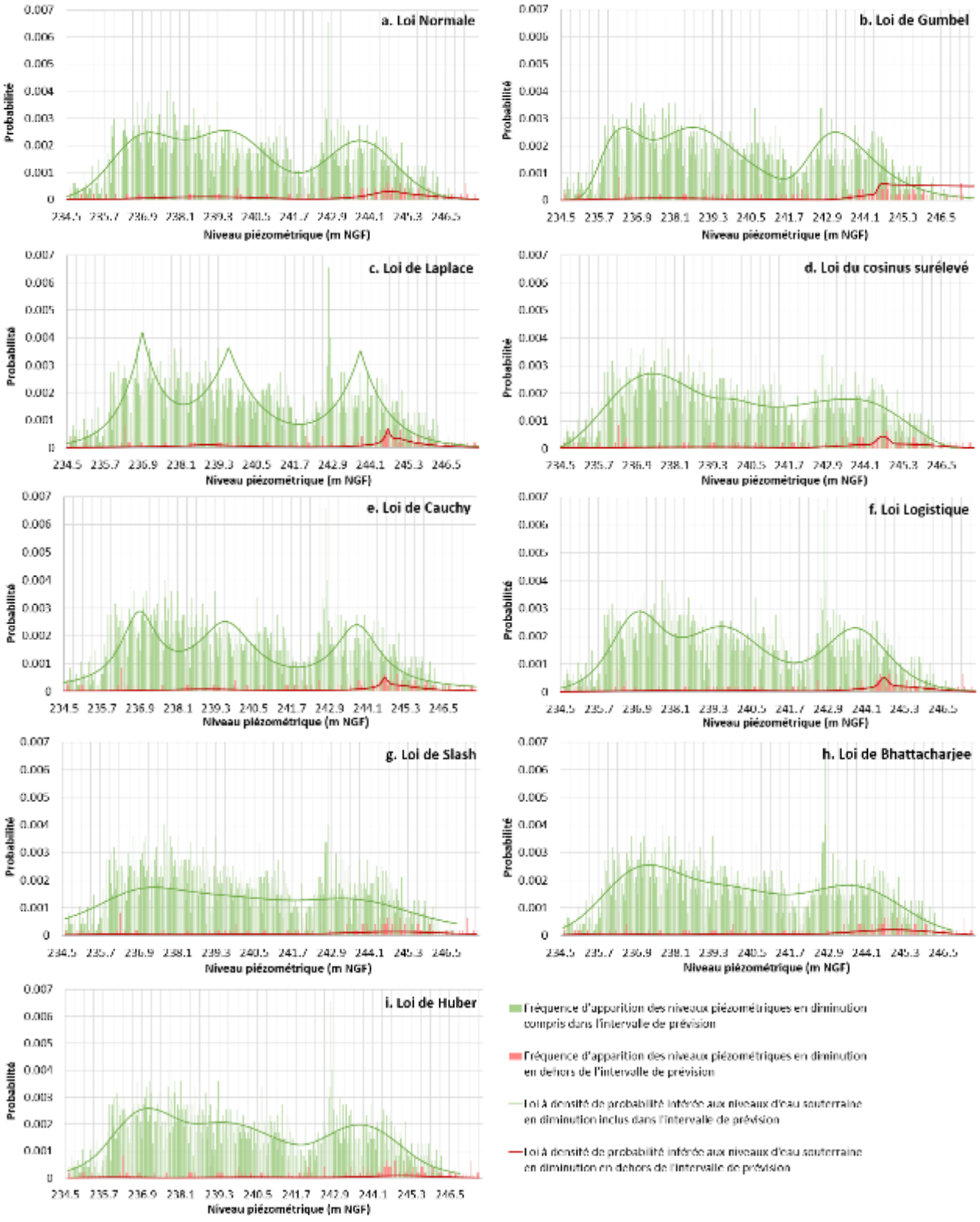


Figure IV-6 - Densités de probabilité inférées à la distribution des niveaux piézométriques en diminution.

Les modèles de justesse sont réalisés à partir des neuf lois inférées à la triple distribution des niveaux piézométriques en baisse inclus au sein de l'intervalle de prévision. Ces modèles sont présentés dans la Figure IV-7.

Pour chaque loi, la corrélation et le pourcentage de points inclus dans la marge d'erreur (M_E) entre le modèle réalisé et les critères PICP pour chaque niveau sont calculés. Les résultats de ces calculs sont présentés dans le Tableau IV-5.

Tableau IV-5 - Corrélation des modèles de justesse pour les niveaux piézométriques en diminution.

Lois à densité de probabilité	***r²***	***M_E (%)***
Loi Normale	0,76	94,0
Loi de Gumbel	0,72	88,3
Loi de Laplace	0,58	92,6
Loi du Cosinus surélevé	0,79	96,9
Loi de Cauchy	0,42	94,3
Loi Logistique	0,69	94,6
Loi de Slash	0,61	56,7
Loi Bhattacharjee	0,52	57,1
Loi de Huber	0,49	58,0

Les différents coefficients de corrélation obtenus diffèrent grandement en fonction de la loi étudiée, allant de 0,42 à 0,79. La loi du Cosinus surélevé obtient la corrélation la plus grande parmi les neuf lois étudiées, suivie par loi Normale. Le pourcentage de critères PICP inclus dans la marge d'erreur la plus élevée est également obtenu par la loi du Cosinus surélevé, avec près de 97% de ces points inclus dans cette marge d'erreur.

Ces scores attestent de la qualité de l'inférence des lois sur la distribution des niveaux piézométriques. Cela s'observe également sur les courbes obtenues par les modèles de justesse, montrés en Figure IV-7. Ces neuf modèles présentent une probabilité de justesse étant autour de 0,9 pour la plupart des niveaux piézométriques, montrant ainsi que le modèle utilisé est fidèle en générant des résultats de modélisation relativement proches les uns des autres, et/ou que l'intervalle de modélisation est suffisamment grand pour contenir les résultats de modélisation.

Deux classes de modèles peuvent être distinguées. La première rassemble les modèles avec la loi Normale, la loi de Gumbel et la loi du Cosinus surélevé. Ces trois modèles présentent une forte baisse des probabilités lorsque les niveaux piézométriques sont de plus en plus grands au-delà de 245 m NGF. De plus, plus spécifiquement pour la loi de Gumbel et la loi du Cosinus surélevé, ces probabilités décroissent avec les niveaux piézométriques quand ceux-ci sont bas (en dessous 235 m NGF).

La seconde classe regroupe les six autres lois. A l'inverse de la première classe de modèles, la probabilité diminue plus faiblement lorsque les hauteurs d'eau souterraine mesurées augmentent au-delà de 245 m NGF. Pour certaines lois, telles les lois de Laplace, de Cauchy et de Slash, ces probabilités augmentent pour des hauts niveaux piézométriques.

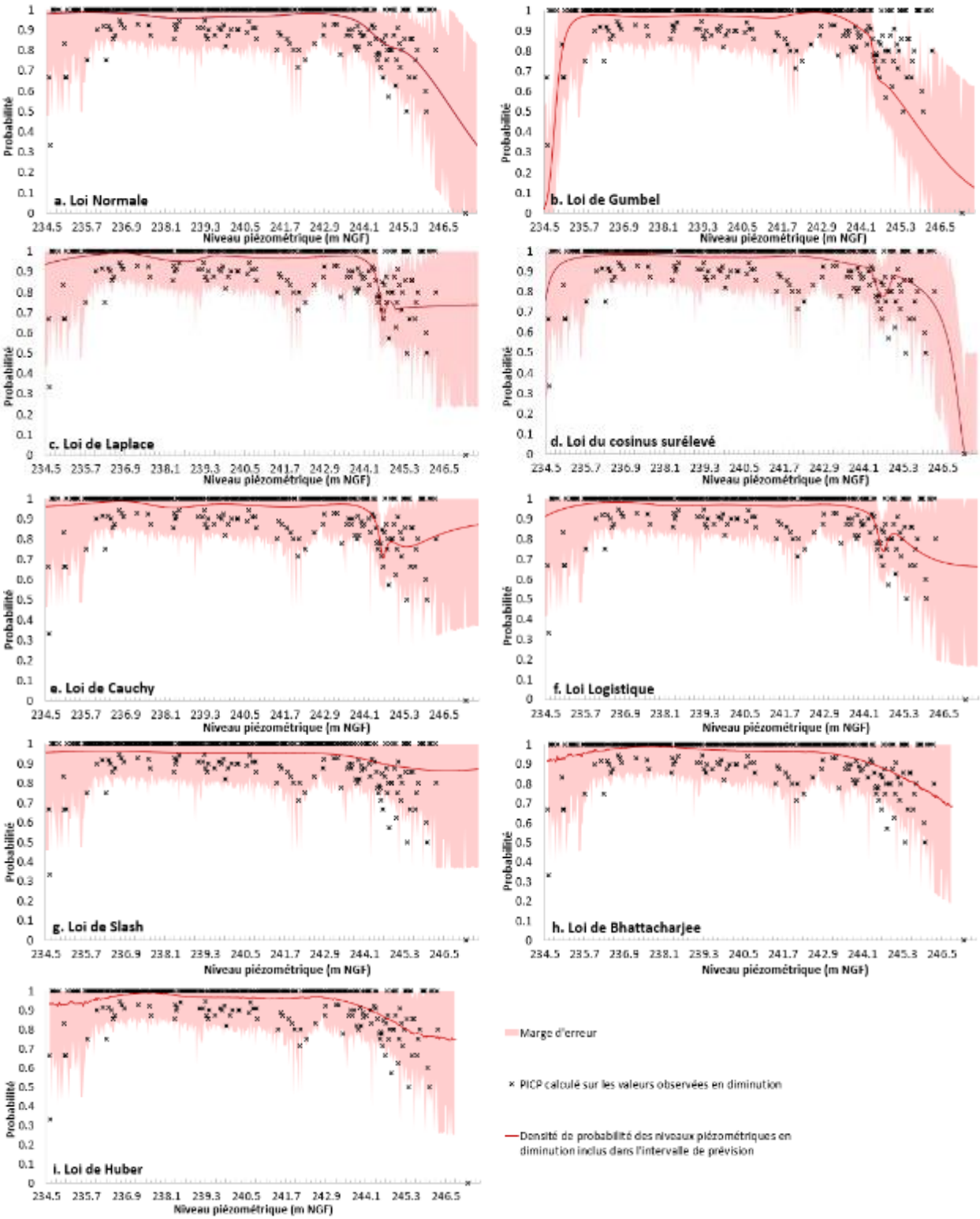


Figure IV-7 - Modèles de justesse pour les niveaux piézométriques en diminution.

Ainsi, l'utilisation de la loi du Cosinus surélevé est la meilleure des neuf lois étudiées pour inférer la distribution des niveaux piézométriques en diminution sur le bassin de la nappe du Blavet pour une prévision à 3 jours. Cette loi possède les meilleurs indicateurs d'inférence vis-à-vis de la distribution des hauteurs d'eau souterraine décroissants, ainsi que des meilleurs indicateurs vis-à-vis des modèles de justesse élaborés. Ici encore, il était pertinent d'essayer quatre lois supplémentaires.

Conclusion sur les modèles de justesse

L'utilisation de l'algorithme d'espérance-maximisation a montré qu'il est préférable d'admettre une triple distribution des niveaux d'eau souterraine plutôt qu'une double distribution. Les modèles de justesse établis avec neuf lois à densités de probabilités différentes ont montré qu'il est possible estimer la probabilité de justesse d'un résultat de modélisation et que certaines lois qui sont venues compléter la première étude étaient très performantes. En effet, la loi du Cosinus surélevé est la loi qui infère le mieux la distribution des niveaux piézométriques, qu'ils soient ascendants ou descendants. L'utilisation de la loi Normale, ou de la loi Logistique, semble être une alternative possible, bien qu'elles soient moins adaptées que la loi du Cosinus surélevé. Enfin, on a pu voir que la distinction des phases de recharge et de vidange conduisait à de meilleurs scores de performance.

Globalement, les modèles de justesse attestent d'une meilleure probabilité de justesse d'un résultat de modélisation dans un contexte de vidange de la nappe (de 0,05 à 0,25 selon les niveaux). Cela peut se constater sur les résultats de modélisation qui d'une part peuvent ne pas atteindre précisément les extrema, et d'autre part qui peuvent montrer des retards sur la remontée de nappe. Pour notre cas présent, ces phénomènes restent peu présents mais influent sur la justesse globale du modèle. De plus, la probabilité de justesse calculée dans les modèles est sensible aux évènements extrêmes. En effet, la marge d'erreur calculée dans ces modèles de justesse est maximale lorsque les niveaux d'eau souterraine sont au-delà de 246,5 m NGF.

L'étude de ce bassin présentée avec le même modèle utilisé mais une méthodologie légèrement différente (double inférence, non prise en compte des états précédents des niveaux d'eau souterraine et quatre lois non testées) avait montré que les lois Normale et Logistique étaient les mieux inférées à la distribution piézométrique (Akil, et al. 2019). La loi du Cosinus surélevé ne faisait pas partie de cette étude. Dans le cas présent, la loi Normale peut constituer une bonne alternative à la loi du Cosinus surélevé. En revanche, la loi Logistique présente quant à elle une inférence un peu moins bonne que la loi Normale.

Ainsi, afin de construire un modèle de justesse, il faut tenir compte du type de distribution inféré selon la distribution observée des données mais aussi de l'intérêt de la discrimination de certaines phases dans le comportement de l'hydrosystème. Ces éléments peuvent influencer le résultat obtenu de façon significative. L'expertise du modélisateur est donc

requise pour tester et sélectionner au mieux les lois qui apporteront l'information la plus pertinente au décisionnaire.

II.3.1.b.ii. Détermination de l'intervalle en fonction d'une justesse exigée

Les modèles de justesse calculés à partir de la loi du Cosinus surélevé sont utilisés pour déterminer un intervalle de prévision à partir d'une justesse exigée. Pour cela, il est nécessaire de mettre en relation les probabilités de justesse appliquées aux niveaux piézométriques observés, selon leur tendance croissante ou décroissante, avec ces mêmes probabilités de justesse appliquées aux résultats de modélisation médians. Le graphique présent en Figure IV-8 montre cette relation, qui se traduit par une fonction affine.

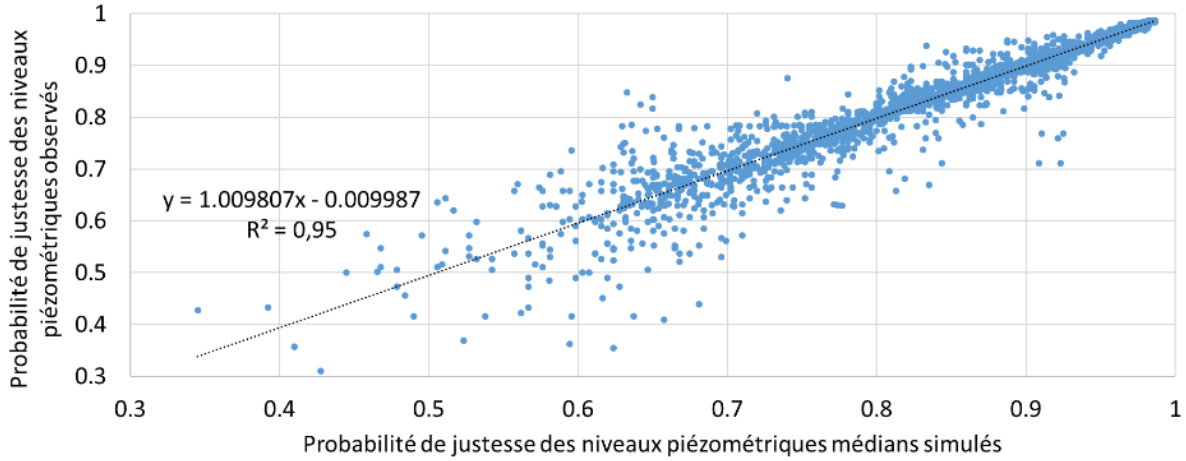


Figure IV-8 - Probabilité de justesse des niveaux piézométriques médians simulés en fonction de la probabilité des niveaux piézométriques observés.

Les coefficients de cette fonction affine sont utilisés afin de recalculer les limites de l'intervalle de prévision. Ainsi, l'amplitude de l'intervalle de prévision est modifiée à partir d'une justesse demandée par un utilisateur.

Dans l'exemple présenté en Figure IV-9, la probabilité de justesse exigée par l'utilisateur est de 0,9. Ici, à certains moments (mi-février 2019), l'intervalle de prévision obtenu est plus grand que l'intervalle de prévision obtenu par 120 initialisations différentes. Le modèle réalisé avec 120 initialisations correspond au plus grand modèle ensembliste réalisé au cours de cette étude. Sur l'ensemble de l'historique de données, le critère PICP obtenu avec cet intervalle de prévision est de 0,90, ce qui est conforme avec la justesse exigée par l'utilisateur.

Toutefois, sur l'ensemble de test, dont le graphique est observé Figure IV-9, ce critère n'est que de 0,80. Sur cet ensemble, la justesse exigée n'est pas respectée. Cela signifie que pour certaines années ou certains évènements spécifiques, la justesse ne peut être garantie avec cette méthodologie. La justesse n'est garantie que sur l'ensemble des données utilisées pour élaborer le modèle, mais pas pour le valider. Pour pouvoir garantir une justesse pour tout

évènement, l'utilisateur doit augmenter la justesse exigée afin d'élargir l'intervalle de prévision.

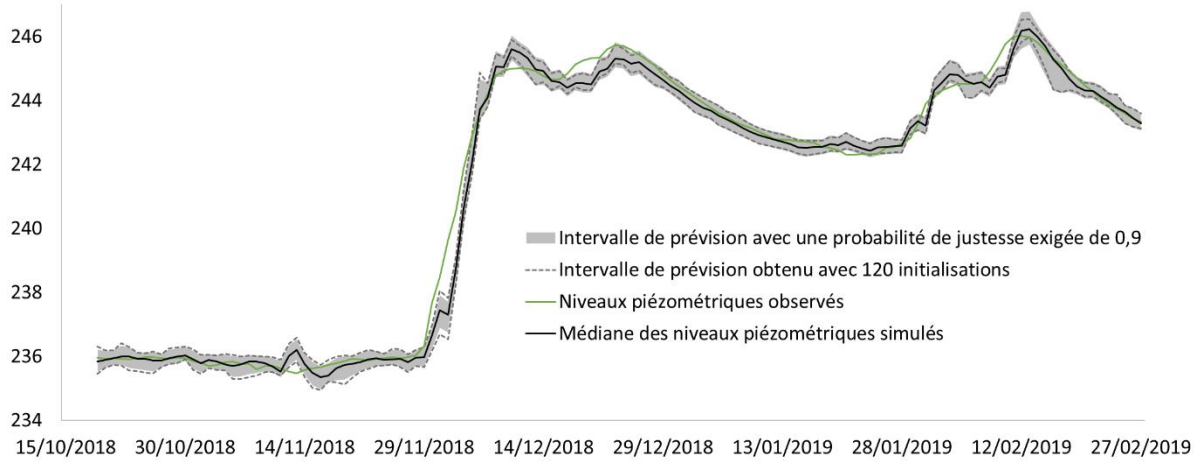


Figure IV-9 - Modélisation des niveaux piézométriques avec intervalle de prévision obtenu par une sélection de la probabilité de justesse à 0,9.

La méthodologie élaborée présente également un autre inconvénient. Il est impossible d'assurer une probabilité de justesse de 1. En effet, l'intervalle de prévision obtenu avec une probabilité de justesse exigée à 1 génère un critère PICP égal à 0,95. Ce décalage provient dans un premier temps que la modélisation n'est pas parfaite, car le modèle a quelques difficultés à représenter les valeurs extrêmales et les remontées de nappes. Dans un second temps, l'inférence des lois sur la distribution piézométrique génèrent des approximations, pouvant également expliquer ce décalage. Ainsi, au-delà de cette valeur de critère il est impossible de garantir la justesse du modèle, hormis en augmentant artificiellement l'intervalle de prévision au-delà de 1.

Ainsi, il est possible de modifier l'intervalle de prévision en fonction d'une probabilité de justesse exigée par un utilisateur en utilisant les résultats obtenus par les deux modèles de justesse avec la loi à densité de probabilité la mieux inférée (en l'occurrence la loi du Cosinus surélevé). Cette méthode tend à se spécialiser sur les données d'apprentissage, limitant sa capacité de généralisation à d'autres années et elle ne peut assurer une justesse demandée très élevée. Toutefois, dans des gammes de justesse exigibles raisonnables en l'état de l'art et avec des lois calculées sur des ensembles de données importants, la méthodologie est tout à fait exploitable en conditions opérationnelles.

Elle est par ailleurs appliquée dans son entièreté aux prévisions à moyen terme (30 jours) sur ce même bassin, au pas de temps décadaire, afin de voir si ces constats seront identiques ou non.

II.3.2. Au pas de temps décadaire : modélisation à moyen-terme

L'objectif est ici de concevoir un modèle capable de restituer une modélisation à moyen-terme. La méthodologie d'estimation de l'incertitude liée à la modélisation est alors ici testée pour des modélisations à plus long terme mais avec autant de pas de temps en horizon de prévision puisque les données deviennent décadaires. Comme pour le modèle à court-terme, l'historique de données a été séquencé par année civile pour définir les sous-ensembles. L'ensemble d'arrêt correspond à l'année civile de 2017. L'ensemble de test a, quant à lui, été élargi par rapport au modèle à court terme, compte-tenu du pas de temps décadaire. Cet ensemble de test inclut l'année civile de 2018 ainsi que les deux premiers mois de 2019. Les données, originellement journalières, ont été intégrées à un pas de temps de 10 jours. La fourniture de ces données décadaires en entrée du modèle permet à celui-ci de fournir un résultat de modélisation de 3 décades, soit 30 jours. Le modèle est tout d'abord sélectionné par validation croisée avec l'utilisation de méthodes de régularisation, comme au pas de temps journalier.

II.3.2.a. Sélection du modèle

Le modèle est ici sélectionné par plan d'expériences couplant arrêt précoce et validation croisée. La taille des fenêtres temporelles est présélectionnée par les analyses corrélatoires. La sélection du modèle par le score de validation croisée a été réalisée avec le critère de persistance (C_P). L'architecture du modèle sélectionnée est présentée dans la Figure IV-10.

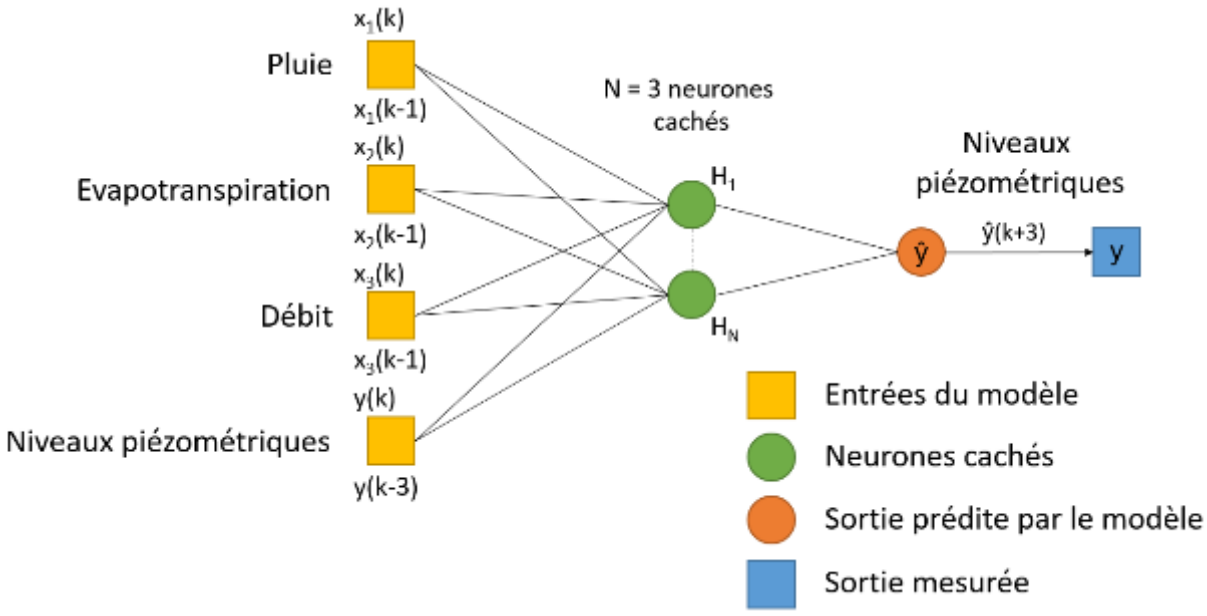


Figure IV-10 - Architecture sélectionnée pour une modélisation des niveaux piézométriques à moyen-terme.

Comme pour le modèle à court-terme, les entrées qui alimentent le modèle sont uniquement composées de données pluviométriques, d'évapotranspiration, de débit et de niveaux

piézométriques. Un perceptron multicouche classique a été préféré à d'autres architectures. L'algorithme d'ADAM a été ici utilisé comme règle d'apprentissage.

Le résultat de la modélisation sur l'ensemble de test est présenté en Figure IV-11.

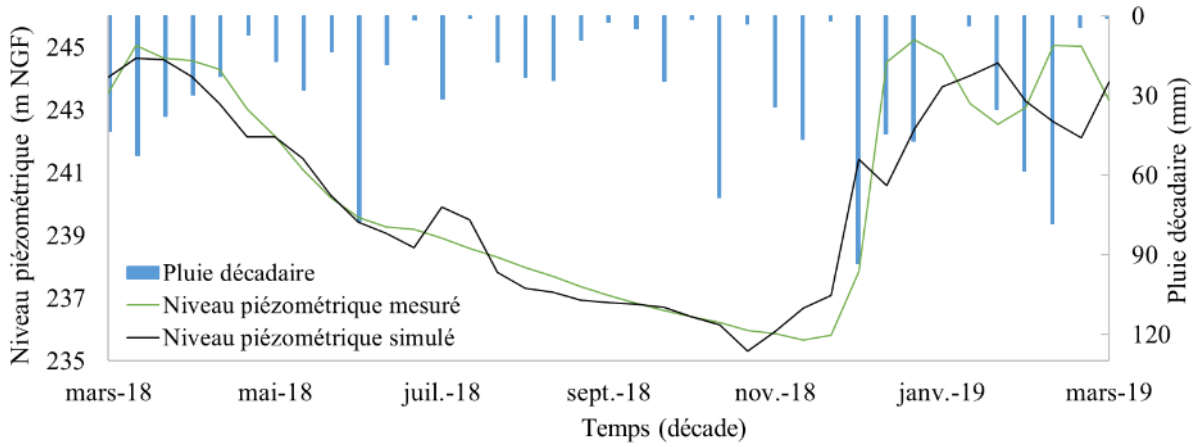


Figure IV-11 - Résultat de modélisation de l'ensemble de validation des niveaux d'eau souterraine à un horizon de prévision de 30 jours.

Le critère de Nash-Sutcliffe (C_{NSE}) est égal à 0,80. La persistance (C_P) est ici égale à 0,74, justifiant ainsi la bonne qualité de la modélisation. Le critère de biais (C_{PBIAS}) est ici de -0,94 %, indiquant ainsi que le modèle surestime légèrement les niveaux piézométriques.

Il est toutefois possible d'observer quelques imperfections, comme en juillet 2018, lorsque le modèle répond trop à une précipitation intense quelques semaines auparavant, ou sur la fin de la recharge, où le modèle peine à atteindre les niveaux les plus élevés.

II.3.2.b. Détermination du nombre de modèles optimal pour représenter l'incertitude de modélisation

Comme pour le premier modèle, le nombre de modèles optimal doit être déterminé en calculant le score de validation croisée médian lorsque celui-ci équivaut au critère *Prediction Confidence Criterion* (C_{PC}) en fonction du nombre d'initialisations de paramètres différentes.

La Figure IV-12 présente l'évolution de ce score en fonction du nombre d'initialisations comprises au sein d'un ensemble de modèles.

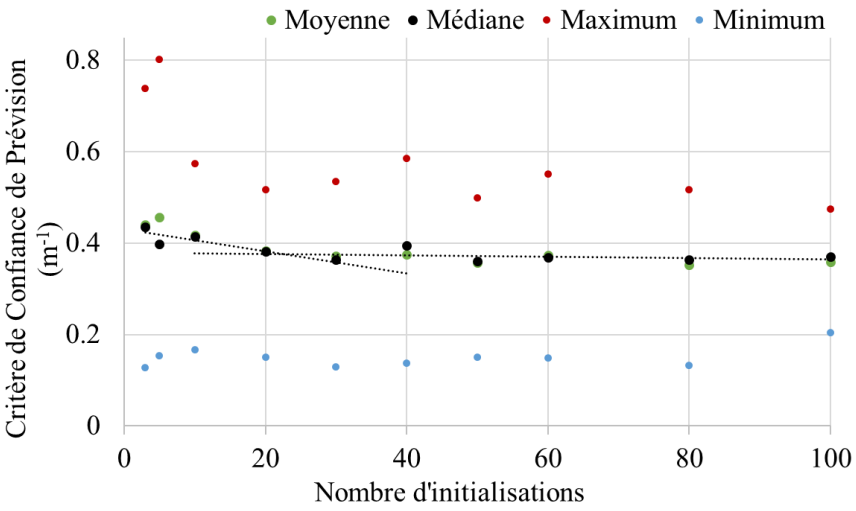


Figure IV-12 - Evolution du critère *Prediction Confidence Criterion* (C_{PC}) en fonction du nombre d'initialisations.

La Figure IV-12 montre que le critère C_{PC} médian diminue lorsque le nombre d'initialisations différentes augmente. Cependant, à partir de 20 initialisations, ce critère semble atteindre un seuil à 0,37 m^{-1}, indiquant que l'aggrandissement de intervalle de prévision ne justifie plus le gain en justesse du modèle, comme pour la prévision à court terme.

Le nombre d'initialisations optimal est donc établi à 20 initialisations pour la suite des travaux. Le résultat de modélisation sur l'ensemble de test est présenté sur la Figure IV-13, avec l'intervalle de prévision construit par les 20 modélisations issues de paramètres initiaux différents.

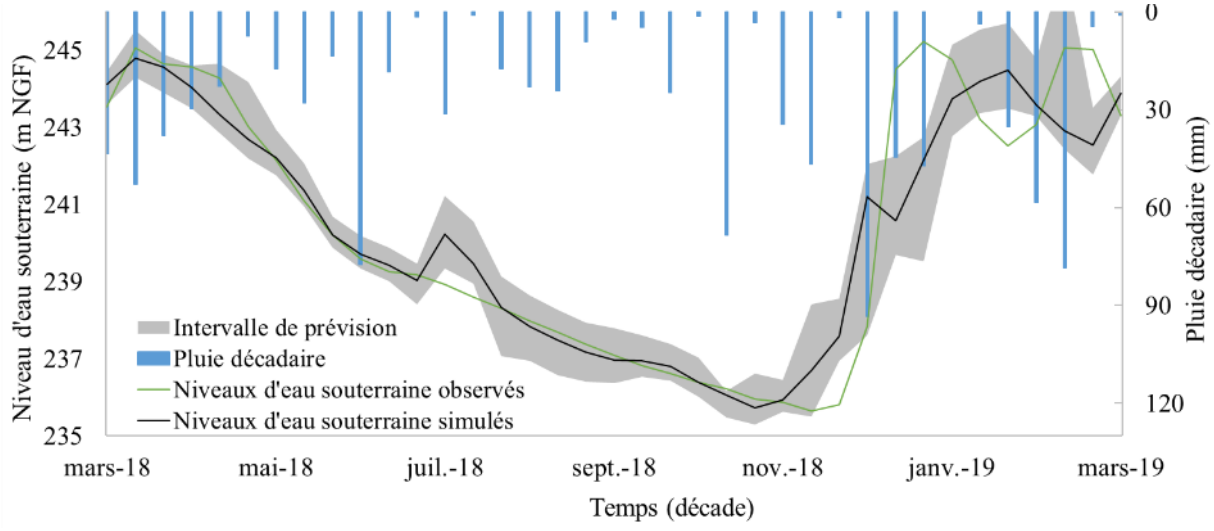


Figure IV-13 - Prévision médiane à 30 jours des niveaux piézométriques et intervalle de prévision généré par 20 modélisations avec des initialisations différentes.

Le score de validation croisée médian lorsqu'il correspond au critère de Nash-Sutcliffe (C_{NSE}) atteint 0,85. Lorsqu'il correspond à la persistance (C_P), ce score médian est égal à 0,79. Le score de validation croisée médian équivalant au critère de biais (C_{PBIAS}) est ici de -1,07 %. Ces scores montrent que le modèle est relativement bon, bien qu'il surestime légèrement les hauteurs d'eau souterraine. Il est toutefois moins bon que le modèle réalisé pour la simulation des niveaux piézométriques à court-terme.

L'intervalle moyen de prévision (C_{MPI}) est de 1,49 m et la probabilité que la médiane des prévisions soit comprise dans l'intervalle de prévision (C_{PICP}) est évaluée à 0,60 en validation croisée. Le score de validation croisée médian égal au critère de confiance de prévision (C_{PC}) est alors de 0,36 m^{-1}.

II.3.2.c. Application des densités de probabilités

II.3.2.c.i. Modèles de justesse

Les modèles de justesse sont calculés à partir de lois à densités de probabilité inférées à la distribution des niveaux piézométriques et selon leur évolution. Comme pour l'étude à court-terme, neuf lois ont été inférées selon une triple distribution. Pour limiter les biais statistiques engendrés par les calculs de densités de probabilité qui vont suivre sur les parties suivantes, les niveaux piézométriques ont été regroupés par classes de 20 cm.

Les résultats des modèles de justesse sont présentés dans un premier temps lorsque les niveaux piézométriques sont en augmentation (période de recharge de la nappe), puis dans un second temps lorsque ces niveaux diminuent (période de vidange de la nappe).

Recharge de la nappe

Les hauteurs d'eau de nappe en hausse sont définies en fonction de leur inclusion au sein de l'intervalle de prévision (PI). La distribution de ces hauteurs est ensuite inférée par des lois à densités de probabilité, listées dans le Tableau IV-6. Ce tableau contient également les indicateurs d'inférence (coefficients de corrélation de Pearson r^2 et distances de Kolmogorov-Smirnov D_{KS}) pour chacune des lois. Ces neuf inférences sont montrées sur les graphiques présents dans la Figure IV-14.

Tableau IV-6 - Indicateurs d'inférence des lois à densité de probabilité.

Lois à densité de probabilité	***r^2*** **(Inclus dans PI)**	***r^2*** **(Hors de PI)**	***D_{KS}*** **(Inclus dans PI)**	***D_{KS}*** **(Hors de PI)**
Loi Normale	0,20	0,17	0,012	0,005
Loi de Gumbel	0,26	0,08	0,010	0,009
Loi de Laplace	0,12	0,15	0,012	0,007
Loi du Cosinus surélevé	0,19	0,07	0,011	0,007
Loi de Cauchy	0,13	0,16	0,016	0,017

Loi Logistique	0,17	0,16	0,011	0,006
Loi de Slash	0,09	0,16	0,016	0,032
Loi Bhattacharjee	0,19	0,13	0,012	0,063
Loi de Huber	0,19	0,19	0,046	0,113

Il est constaté dans un premier temps que la corrélation entre les inférences et la distribution des niveaux piézométriques est faible (globalement entre 0,1 et 0,2), y compris pour les valeurs comprises dans l'intervalle de prévision. Les lois de Gumbel et Normale possèdent les meilleures corrélations pour les valeurs piézométriques incluses dans l'intervalle de prévision tandis que les lois Normale et de Huber possèdent de meilleurs résultats pour les valeurs en dehors de l'intervalle de prévision.

Parallèlement, les distances de Kolmogorov-Smirnov sont proches de zéro pour la plupart des lois, et dans l'ensemble équivalentes. Seule la loi de Huber présente des distances un peu plus éloignées de 0, indifféremment de l'appartenance ou non à l'intervalle de prévision. Les lois de Slash présentent quant à elles des distances un peu plus grandes pour les niveaux exclus de l'intervalle de prévision. Les lois de Gumbel, Cosinus Surélevé et Logistique possèdent quant à elles les distances de Kolmogorov-Smirnov les plus petites parmi les neuf lois.

Les graphiques présents sur la Figure IV-14 expliquent ces faibles valeurs vis-à-vis de la corrélation. En effet, la distribution des valeurs piézométriques observées (représentée en histogramme vert pour les valeurs incluses dans l'intervalle de prévision et en histogramme rouge pour les valeurs en dehors de l'intervalle de prévision) semble incomplète. Par exemple, aucune valeur n'est observée entre 242,0 m NGF et 242,4 m NGF, qu'elle soit incluse ou exclue de l'intervalle de prévision, malgré le regroupement des valeurs piézométriques par classes de 20 cm.

Le passage de données journalières en données décadaires induit que le nombre de données utilisées pour réaliser les calculs est divisé par dix. Ainsi, 429 valeurs décadaires composent notre historique de données compris entre 2005 et 2019, dont seulement 104 d'entre elles sont en augmentation. Ces 104 valeurs sont ensuite réparties sur l'ensemble des classes de 20 cm de hauteurs d'eau.

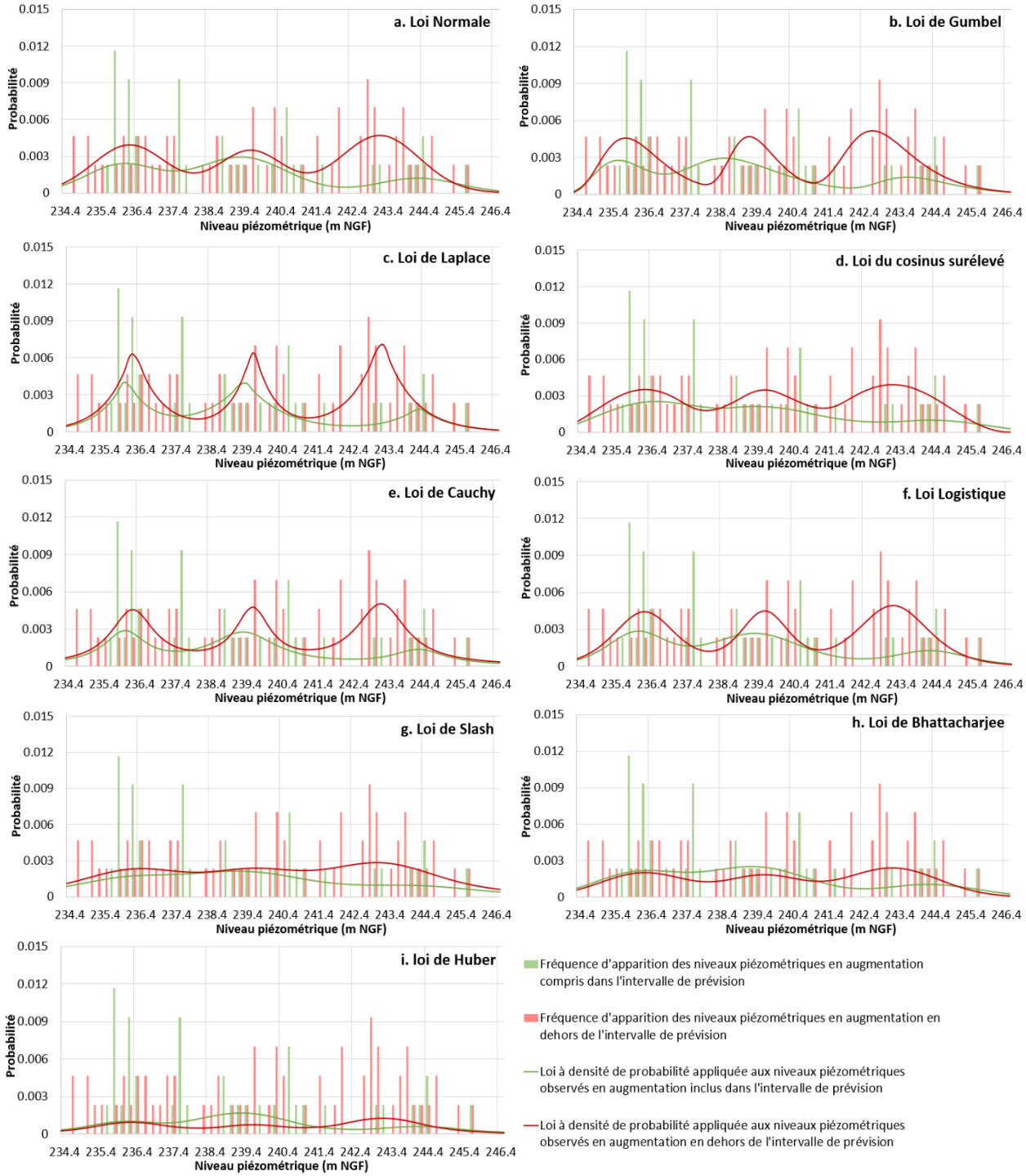


Figure IV-14 - Densités de probabilité inférées à la distribution des niveaux piézométriques en augmentation.

Ce manque de valeurs observées complexifie l'étude de l'incertitude de modélisation puisqu'il génère une incertitude statistique importante. L'ajout de données d'entrées au modèle est dans ce cas impossible, compte-tenu de la non disponibilité de celles-ci. Au premier abord, il serait possible d'envisager d'élargir les classes de valeurs piézométriques, afin qu'elles contiennent suffisamment de valeurs pour réduire et limiter les incertitudes statistiques. Cependant, augmenter la largeur de ces classes reviendrait à attribuer la même probabilité d'apparition à des valeurs de plus en plus éloignées entre elles. Cela tend alors à diminuer la

justesse des résultats présentés et par conséquent des modèles de justesse et de l'estimation de l'incertitude de modélisation. Le modélisateur doit donc trouver un compromis entre justesse des calculs et incertitude statistique.

Malgré cet inconvénient, il reste possible d'exploiter les graphiques présentés sur la Figure IV-14. Les lois inférées suivent globalement la distribution des niveaux piézométriques, hormis pour la loi de Huber, qui sous-évalue grandement cette distribution. Ce constat est cohérent avec les mauvais résultats obtenus vis-à-vis des indicateurs d'inférence.

Les différents modèles de justesse sont déterminés selon les neuf lois inférées à la distribution des hauteurs piézométriques en augmentation comprises dans l'intervalle de prévision. Ces modèles sont présentés sur la Figure IV-15.

La corrélation et le pourcentage de points inclus dans la marge d'erreur (M_E) entre le modèle réalisé et les critères PICP sont calculés pour chaque classe de niveaux piézométriques et pour chaque loi à densité de probabilité étudiée (Tableau IV-7) :

Tableau IV-7 - Corrélation des modèles de justesse pour les niveaux piézométriques en augmentation.

Lois à densité de probabilité	***r²***	***M_E (%)***
Loi Normale	0,15	57,1
Loi de Gumbel	0,20	50,0
Loi de Laplace	0,12	64,3
Loi du Cosinus surélevé	0,17	57,1
Loi de Cauchy	0,11	64,3
Loi Logistique	0,15	57,1
Loi de Slash	0,22	42,9
Loi Bhattacharjee	0,08	42,9
Loi de Huber	0,07	42,9

Les corrélations obtenues sont très proches de zéro pour chacune des lois. Le pourcentage de critères PICP inclus dans la marge d'erreur se situe autour de 50 % pour chaque loi. Ces mauvais résultats sont majoritairement causés par le manque de valeurs piézométriques en augmentation dans notre historique de données et ne sont pas représentatifs de la justesse des modèles générés. En effet, seulement treize critères PICP issus de valeurs piézométriques observées ont pu être calculées. L'incertitude des résultats présentés est assez grande, comme peut attester la grande amplitude de la marge d'erreur en tout point.

Néanmoins, l'incertitude de modélisation en fonction des niveaux piézométriques peut être approchée par la création des modèles de justesse. Les neuf lois génèrent des modèles de justesse avec des profils semblables. Toutefois, deux catégories de modèles peuvent être distingués. La première catégorie concerne les modèles générés par les lois Normale, Gumbel,

Cosinus surélevé, Logistique et Bhattacharjee. Ceux-ci présentent une forte amplitude des probabilités de justesse, qui oscillent entre 0,25 et 1, selon la valeur des niveaux piézométriques. De plus, ces modèles ont une particularité commune. La probabilité de justesse augmente fortement au-delà de 244 m NGF. Ce résultat apparaît comme être incohérent avec la théorie. A l'instar de la prévision à court-terme, ce résultat témoigne également de la plus forte sensibilité de ces lois aux valeurs extrémales, qui ne représentent pas correctement le comportement du modèle neuronal, qui a *a priori* plus de difficultés à modéliser les valeurs extrêmes compte-tenu de leur faible occurrence naturelle dans les chroniques.

La seconde catégorie regroupe les modèles issus des autres lois. Ici, l'amplitude des probabilités de justesse est légèrement plus restreinte, variant entre 0,2 et 0,75. Le profil des densités de probabilités est ici plus aplani. De plus, la probabilité de justesse pour les hautes valeurs de niveaux d'eau souterraine reste stable au-delà de 244 m NGF, dont notamment la loi de Huber.

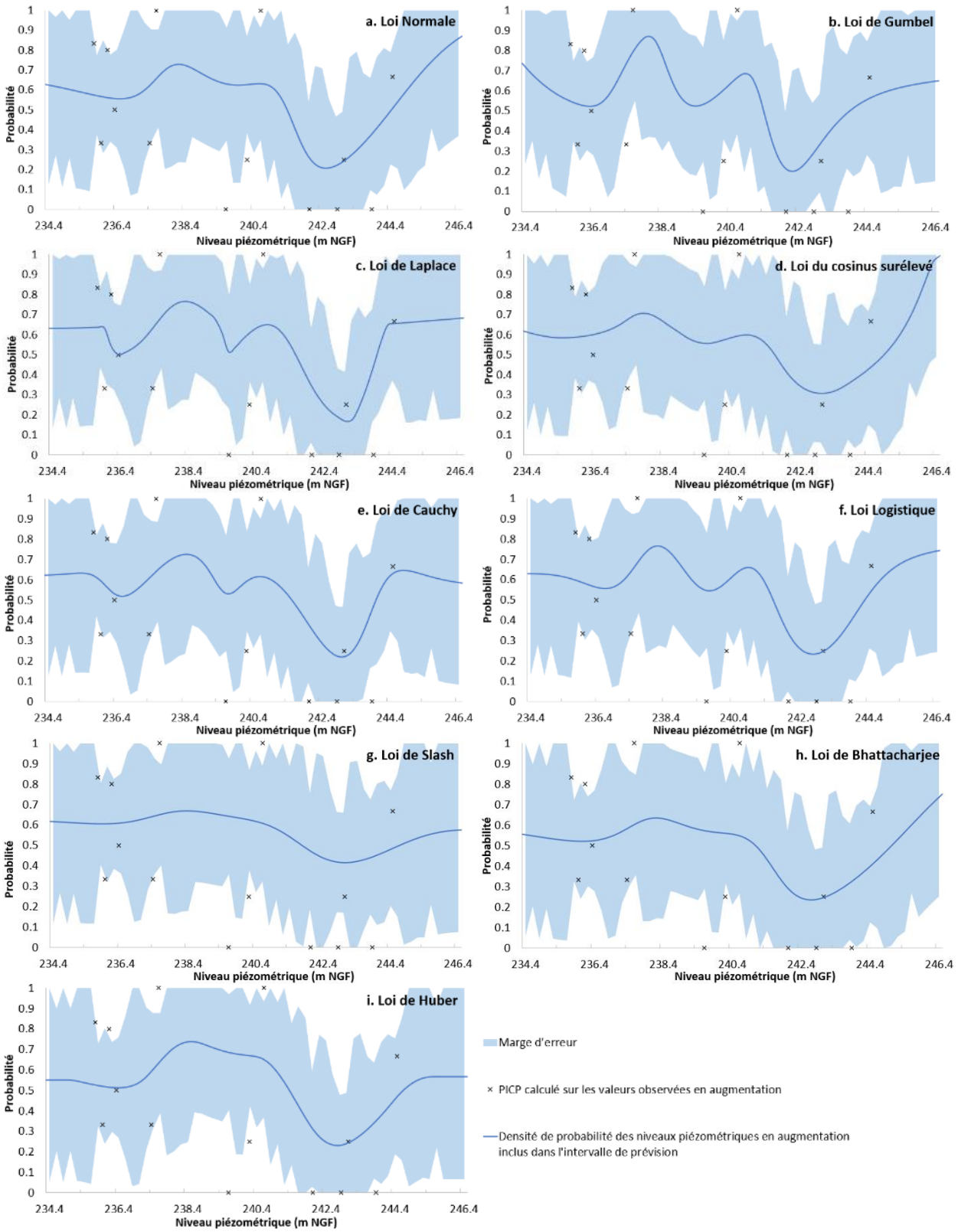


Figure IV-15 - Modèles de justesse pour les niveaux piézométriques en augmentation.

Compte-tenu de la forte incertitude sur les résultats émis par les calculs de probabilité de justesse, il est difficile de sélectionner une loi adaptée pour la détermination de l'incertitude

de modélisation des niveaux piézométriques en période de recharge de la nappe. La loi Logistique peut prétendre à être la loi la mieux inférée car elle présente de manière globale des meilleurs indicateurs d'inférence, malgré la forte incertitude présente sur ces indicateurs, représentée par des marges d'erreur importantes.

Vidange de la nappe

A l'instar de la recharge de la nappe, les niveaux piézométriques en période de vidange de la nappe sont triés selon leur présence ou non au sein de l'intervalle de prévision (PI). Les inférences des neuf lois à densités de probabilité avec la distribution des niveaux piézométriques en diminution sont étudiées en comparant la corrélation r^2 et la distance de Kolmogorov-Smirnov D_{KS} et présentées sur le Tableau IV-8 suivant.

Tableau IV-8 - Indicateurs d'inférence des lois à densité de probabilité.

Lois à densité de probabilité	***r^2*** **(Inclus dans PI)**	***r^2*** **(Hors de PI)**	***D_{KS}*** **(Inclus dans PI)**	***D_{KS}*** **(Hors de PI)**
Loi Normale	0,45	0,18	0,021	0,017
Loi de Gumbel	0,46	0,04	0,020	0,016
Loi de Laplace	0,41	0,18	0,025	0,020
Loi du Cosinus surélevé	0,40	0,17	0,016	0,013
Loi de Cauchy	0,43	0,18	0,046	0,036
Loi Logistique	0,47	0,18	0,017	0,015
Loi de Slash	0,48	0,26	0,078	0,016
Loi Bhattacharjee	0,48	0,23	0,015	0,096
Loi de Huber	0,46	0,18	0,240	0,105

La corrélation entre les densités de probabilités et la distribution piézométrique comprise dans l'intervalle de prévision est comprise entre 0,48 (loi de Slash et Bhattacharjee) et 0,40 (loi du Cosinus surélevé). Pour les niveaux en dehors de l'intervalle de prévision, les lois de Slash et de Bhattacharjee obtiennent la meilleure corrélation avec un résultat de 0,26 et 0,23. Les distances de Kolmogorov-Smirnov sont toutes proches de zéro, hormis pour la loi de Huber. La loi du Cosinus surélevé obtient les distances les plus courtes. En revanche, la loi de Bhattacharjee a une distance un peu plus grande par rapport aux autres lois (sauf loi de Huber) dans le cas où les niveaux d'eau souterraine sont en dehors de l'intervalle de prévision.

En observant la courbe présente dans la Figure IV-16, la loi de Bhattacharjee surestime légèrement la distribution des niveaux piézométriques en diminution non compris dans l'intervalle de prévision, expliquant ainsi une distance de Kolmogorov-Smirnov plus élevée. De même, la loi de Huber est nettement en dessous de la distribution des valeurs piézométriques comprises dans l'intervalle de prévision. Les autres densités de probabilité calculées suivent bien la distribution des valeurs de niveaux de nappe observées.

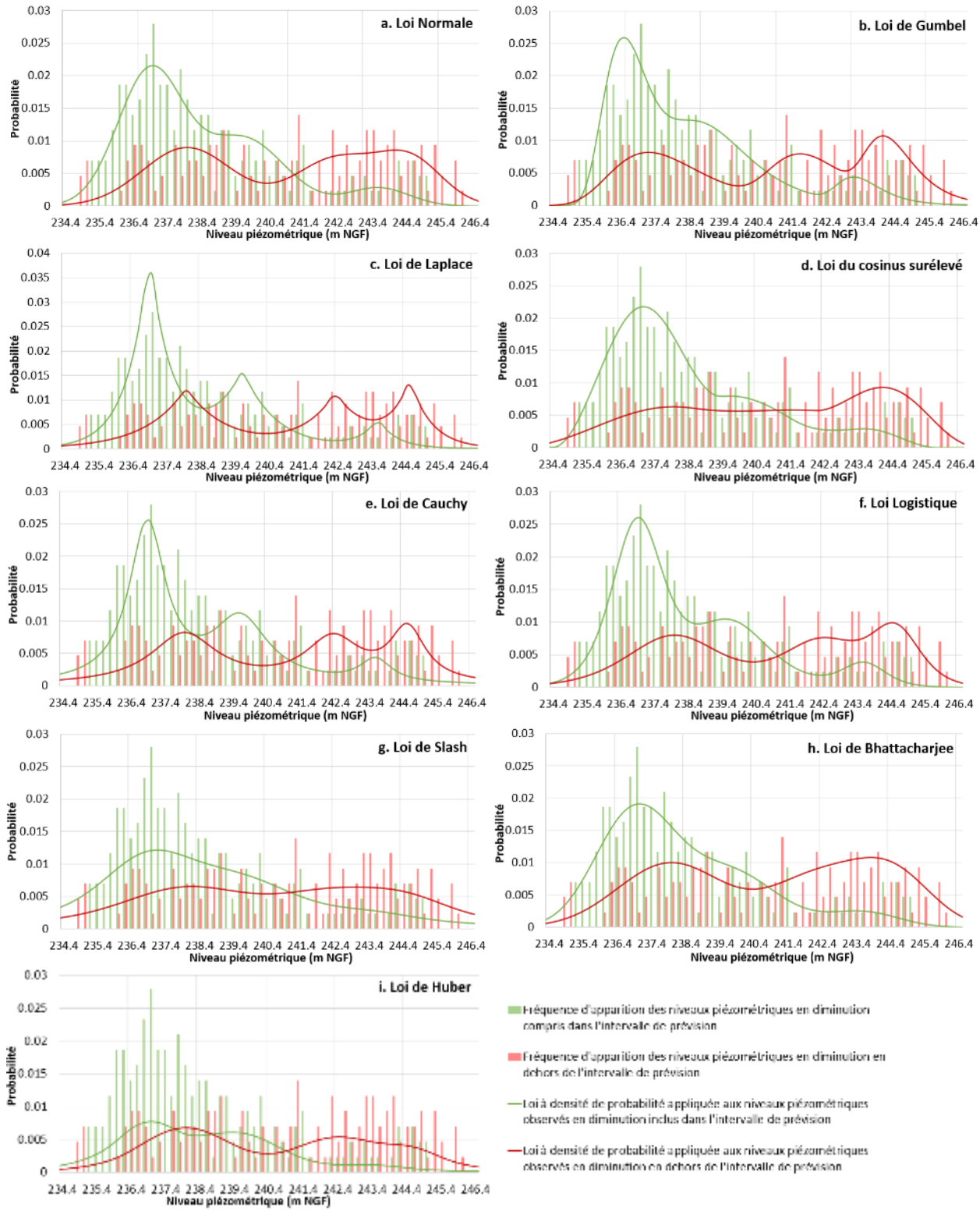


Figure IV-16 - Densités de probabilité inférées à la distribution des niveaux piézométriques en diminution.

La construction des modèles de justesse est réalisée à partir des neuf densités de probabilité inférées à la distribution des valeurs de niveaux de nappe en diminution comprises dans l'intervalle de prévision (Figure IV-17).

Le Tableau IV-9 indique les résultats obtenus de corrélation et de pourcentage de points inclus dans la marge d'erreur (M_E) entre le modèle réalisé et les critères PICP issus des niveaux piézométriques observés.

Tableau IV-9 - Corrélation des modèles de justesse pour les niveaux piézométriques en diminution.

Lois à densité de probabilité	***r²***	***M_E* (%)**
Loi Normale	0,40	68,0
Loi de Gumbel	0,45	68,0
Loi de Laplace	0,46	60,0
Loi du Cosinus surélevé	0,16	66,0
Loi de Cauchy	0,45	62,0
Loi Logistique	0,49	62,0
Loi de Slash	0,48	62,0
Loi Bhattacharjee	0,44	74,0
Loi de Huber	0,44	56,0

Les différents coefficients de corrélation obtenus varient entre 0,16 et 0,49 selon les lois. La Logistique et la loi de Slash obtiennent la meilleure corrélation parmi les neuf lois étudiées, suivies par loi Logistique. Le pourcentage de points dans la marge d'erreur le plus élevé est obtenu par la loi de Bhattacharjee. La mauvaise inférence de la loi de Huber se traduit ici par des indicateurs plus faibles que ceux obtenus par les autres lois.

La Figure IV-17 affiche les modèles de justesse issus des neuf densités de probabilités calculées. Pour chaque modèle, la probabilité de justesse est maximale autour de 236 m NGF, puis diminue progressivement au fur et à mesure que les niveaux piézométriques augmentent.

Les modèles ont en revanche des profils relativement hétérogènes bien qu'ils aient tous les mêmes tendances. Les profils des modèles issus de la loi du Cosinus surélevé et de Gumbel ont la particularité d'avoir une probabilité de justesse très basse pour les niveaux d'étiage les plus importants (inférieurs à 235 m NGF). Les lois de Gumbel, Laplace, Cauchy et Logistique (et à moindre mesure la loi Normale) génèrent des modèles avec de fortes fluctuations de probabilités de justesse lorsque les niveaux piézométriques sont compris entre 236 m NGF et 244 m NGF.

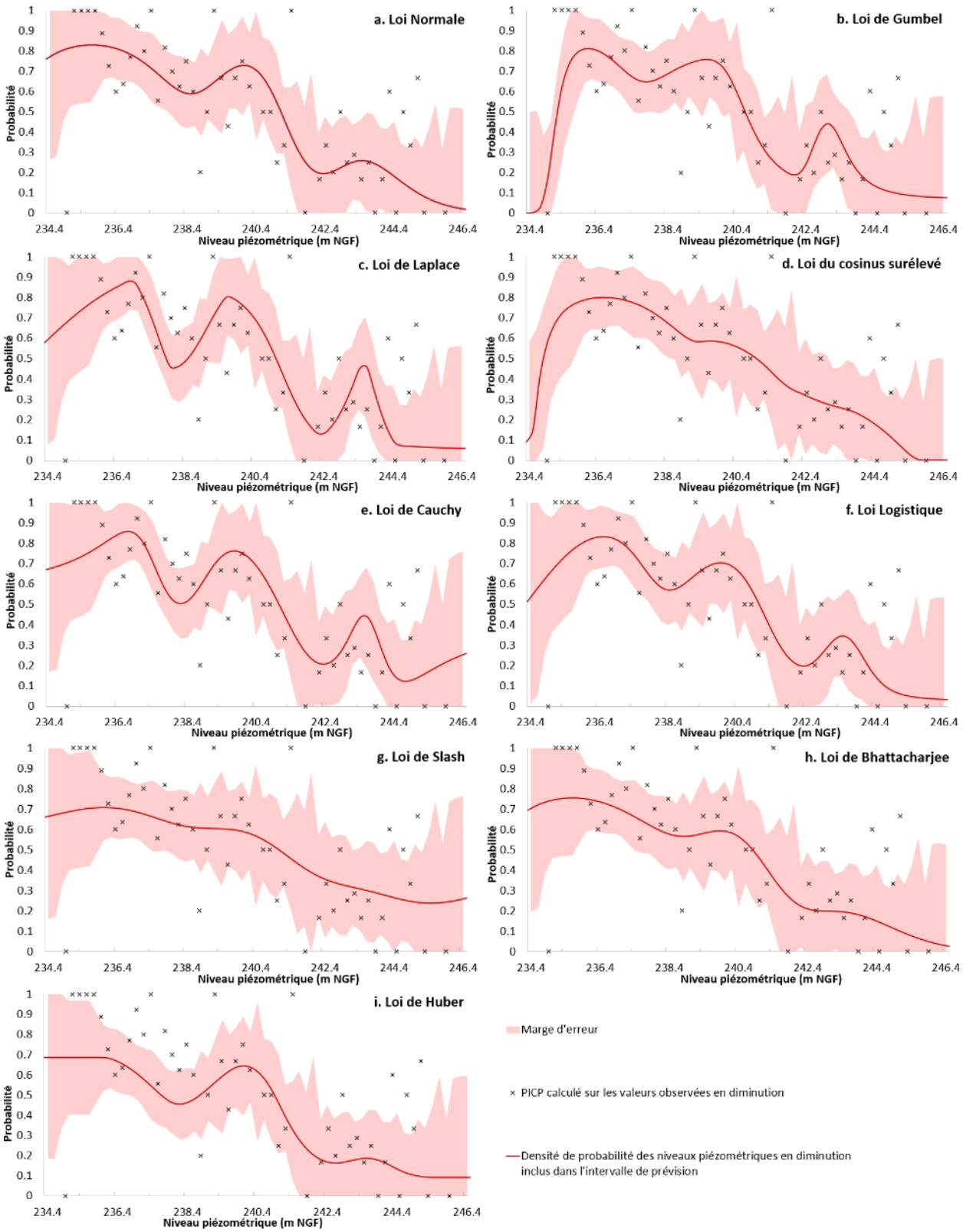


Figure IV-17 - Modèles de justesse pour les niveaux piézométriques en diminution.

Ainsi, l'utilisation de la loi Logistique est la meilleure des neuf lois étudiées pour inférer la distribution des niveaux piézométriques en diminution sur le bassin de la nappe du Blavet

pour une prévision à 30 jours. Cette loi possède un des meilleurs indicateurs d'inférence vis-à-vis de la distribution des hauteurs d'eau souterraine décroissants, ainsi que des meilleurs indicateurs vis-à-vis des modèles de justesse élaborés. Malgré des statistiques d'inférence presque équivalentes à la loi Logistique, la loi de Bhattacharjee ne peut être considérée comme étant une loi de référence pour la construction du modèle de justesse. En effet, la surestimation de cette loi vis-à-vis de la distribution des niveaux d'eau en dehors de l'intervalle de prévision peut avoir un impact sur le calcul des probabilités de justesse. L'utilisation de la loi Normale est ici préférable comme second choix, bien qu'elle possède des indicateurs d'inférence légèrement plus faibles que la loi de Bhattacharjee.

Conclusion sur les modèles de justesse

Les modèles de justesse établis avec neuf lois à densités de probabilité permettent d'évaluer l'incertitude de modélisation pour une prévision des niveaux d'eau souterraine de la nappe du bassin du Blavet à 30 jours. L'élaboration du modèle neuronal a nécessité l'utilisation de données décadaires en entrée. Cela a eu un fort impact sur la détermination de cette incertitude de modélisation et la conception des modèles de justesse puisque, comme nous l'avons vu pour le court terme, leur qualité dépend de la quantité de données disponibles. Ce constat est d'autant plus vrai pour les périodes de recharge de nappe puisqu'elles sont plus rapides que les vidanges et donc moins fréquentes dans la base.

Les modèles de justesse élaborés et proposés précédemment sont donc soumis à une plus grande incertitude que ceux réalisés pour la modélisation des niveaux d'eau souterraine à court-terme. Néanmoins, il a été possible de les concevoir. Les lois du Cosinus surélevé et Logistique semblent être les mieux adaptées et inférées à la distribution des valeurs piézométriques. L'avantage de la loi Logistique est qu'elle présente les meilleurs indicateurs d'inférence pour les niveaux piézométriques croissants et de bons résultats pour les valeurs piézométriques décroissantes. Ceci étant, on ne saurait recommander suffisamment de prudence dans l'interprétation de résultats statistiques basés sur des populations faibles.

II.3.2.c.ii. Détermination de l'intervalle en fonction d'une justesse exigée

Les modèles de justesse issus de la loi Logistique sont utilisés pour déterminer un intervalle de prévision à partir d'une justesse exigée. Les probabilités de justesse appliquées aux niveaux piézométriques observés, selon leurs évolutions précédentes, sont comparées avec ces mêmes probabilités de justesse appliquées aux résultats de modélisation médians. Le graphique présent en Figure IV-18 montre cette relation, qui se traduit par une fonction affine.

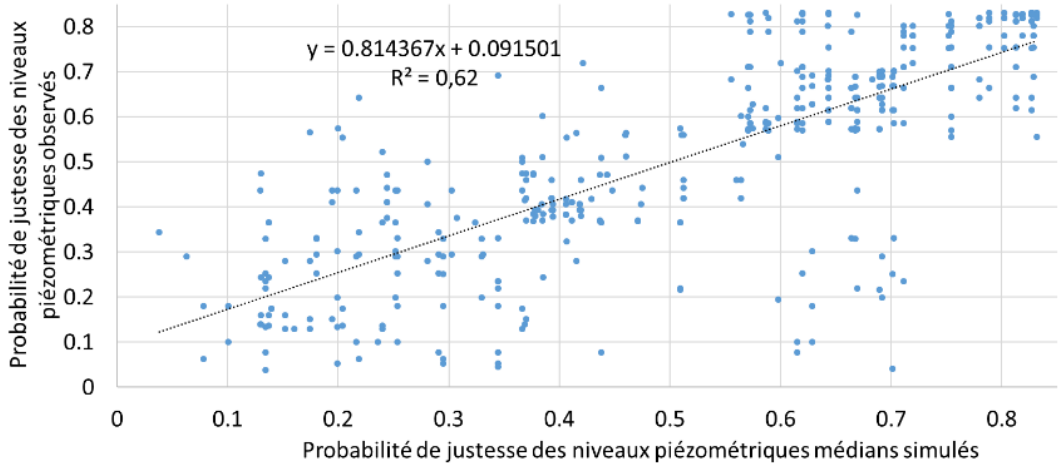


Figure IV-18 - Probabilité de justesse des niveaux piézométriques médians simulés en fonction de la probabilité des niveaux piézométriques observés.

Les coefficients de cette régression linéaire sont utilisés afin de recalculer les limites de l'intervalle de prévision. La corrélation de la courbe de tendance avec les probabilités de justesse est toutefois modérée, avec une corrélation de 0,62. Dans notre cas d'étude, il est à noter que la probabilité de justesse n'excède jamais 0,85.

La Figure IV-19 montre l'ensemble de test avec une probabilité de justesse exigée par l'utilisateur est de 0,7. L'intervalle de prévision exigé par l'utilisateur est ici plus grand que l'intervalle de prévision obtenu par 100 initialisations différentes. Le modèle réalisé avec 100 initialisations correspond au plus grand modèle ensembliste réalisé au cours de cette étude. L'intervalle de prévision exigé comporte plusieurs artéfacts (avril et décembre 2018) sur les valeurs minimales de cet intervalle. Ces artéfacts proviennent de la faible probabilité de justesse pour les niveaux piézométriques croissants observés proches de 244 m NGF (la probabilité est environ de 0,3 dans ce cas).

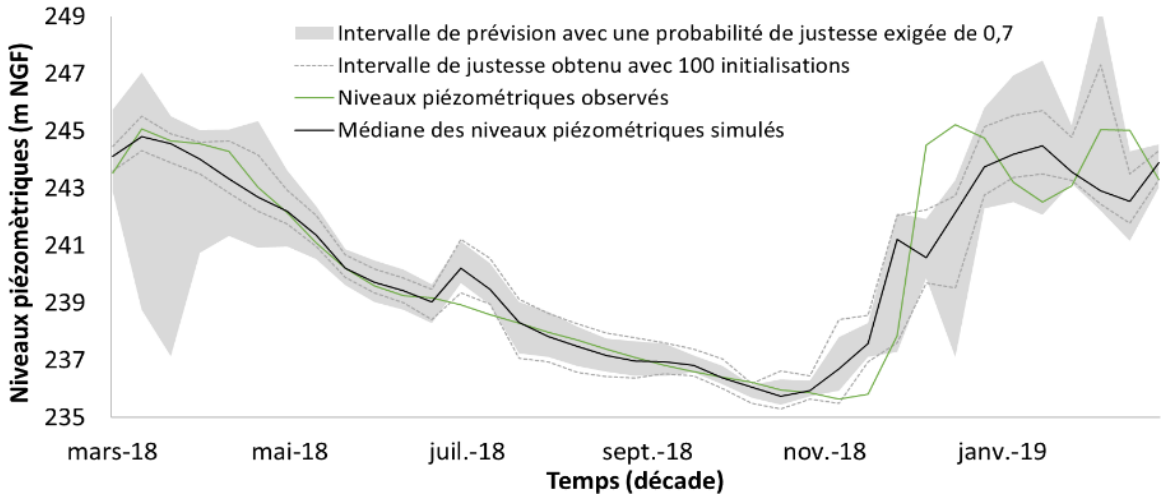


Figure IV-19 - Modélisation des niveaux piézométriques avec intervalle de prévision obtenu par une sélection de la probabilité de justesse à 0,7.

Sur l'ensemble de l'historique de données, le critère PICP obtenu avec cet intervalle de prévision est de 0,69, ce qui est proche avec la conformité vis-à-vis de la justesse exigée par l'utilisateur. Sur l'ensemble de test, dont le graphique est observé Figure IV-19, ce critère atteint 0,75. La justesse exigée est alors ici respectée pour ce test.

Malgré les problèmes causés par le manque de données, il a été possible d'élaborer la méthodologie d'estimation d'incertitude de modélisation et de conception d'un intervalle de prévision à partir d'une justesse exigée.

Toutefois, ce manque de données génère des incertitudes sur les modèles de justesse et produit des intervalles parfois très grands. Or la conception de l'intervalle de prévision avec une justesse exigée utilise les probabilités de justesse des meilleurs modèles de justesse. Ainsi, la qualité de l'intervalle de prévision généré par une probabilité de justesse exigée est dépendante à la quantité de données qui alimentent le réseau de neurones. Néanmoins, l'intervalle de prévision calculé respecte les exigences de l'utilisateur dans notre exemple, malgré ces problématiques. L'utilisateur doit donc être sensibilisé au fait qu'une exigence élevée peut se traduire par un intervalle large dans les périodes peu représentées de la distribution initiale des données.

II.3.3. Conclusion sur le bassin de la nappe du Blavet

C'est dans le contexte décrit au début de ce chapitre qu'il a été possible de modéliser et de prévoir les niveaux des eaux souterraines à court et moyen termes pour la nappe du Blavet. Dans les deux cas, un perceptron multicouche a été utilisé. La modélisation à court-terme a abouti à des prévisions de meilleure qualité par rapport celles obtenues pour une modélisation à moyen-terme. Les données d'entrées pour la modélisation à moyen-terme ont été dégradées à un pas de temps décadaire, réduisant ainsi la qualité de la prévision à un horizon toutefois trois fois plus élevé.

De plus, une méthodologie d'estimation de l'incertitude de la modélisation a été développée et appliquée. Cette méthodologie aboutit sur la conception de modèles de justesse et de génération d'intervalle de prévision en fonction d'une probabilité de justesse exigée par un utilisateur. La modélisation à moyen-terme a montré un inconvénient concernant la méthodologie de détermination de l'incertitude appliquée. Lorsque le modèle neuronal est alimenté par peu de données, le calcul des probabilités de justesse est influencé par des incertitudes statistiques, réduisant potentiellement la justesse de ces calculs. Ces incertitudes statistiques se transcrivent également sur la détermination de l'intervalle de prévision avec une exigence sur la justesse.

La modélisation à court terme, quant à elle, a permis de mettre en évidence la difficulté d'évaluer l'incertitude de modélisation pour les valeurs extrêmes. La faible occurrence de ces phénomènes engendre également des biais statistiques, rendant l'évaluation de l'incertitude de modélisation pour ces valeurs incertaine. Elle a aussi permis de mettre en exergue la

complexité de répondre aux besoins et exigences de l'utilisateur, causée par la variabilité des phénomènes hydrologiques et climatiques (et donc des données en entrées du réseau de neurones).

Les incertitudes de modélisation ont toutefois pu être approchées par la méthodologie développée. Les résultats obtenus ont été souvent satisfaisants, notamment en utilisant les lois à densités de probabilité qui se sont le mieux inférées à la distribution des valeurs piézométriques. Les lois du Cosinus surélevé (pour la modélisation à court-terme) et Logistique (pour la modélisation à moyen-terme) sont les lois ayant eu les meilleures inférences. De plus, ces incertitudes de modélisation ont été abordées en fonction de la dernière variation des niveaux piézométriques, qui est la représentation majeure de l'état hydrogéologique du système.

Cependant, ces premières conclusions sont élaborées sur l'étude d'un seul et unique bassin. Seront-elles identiques si un nouveau bassin est étudié ? De plus, les incertitudes liées à la dispersion spatiales des précipitations n'ont pas été abordée dans nos résultats. Ces dernières ont-elles une influence sur la détermination de l'incertitude de modélisation ? La partie suivante tentera de répondre à toutes ces interrogations, en étudiant les niveaux d'eau souterraine sur la nappe de la Craie de Champagne pour une prévision à moyen-terme.

III. La nappe de la Craie de Champagne sud et centre

La nappe de la Craie de Champagne Sud et Centre est le second bassin étudié au cours de nos travaux. Comme pour le premier bassin d'étude, nous présenterons dans un premier temps les caractéristiques de ce bassin. Puis, dans un second temps, nous présentons les résultats de modélisation obtenus ainsi que les incertitudes associées.

Ce bassin a été sélectionné compte-tenu de la profondeur des historiques de données disponibles et de la présence de plusieurs pluviomètres à proximité du point d'étude permettant ainsi de s'intéresser à la variabilité induite par la prise en compte de sources différentes de pluviométrie.

III.1 Contexte général

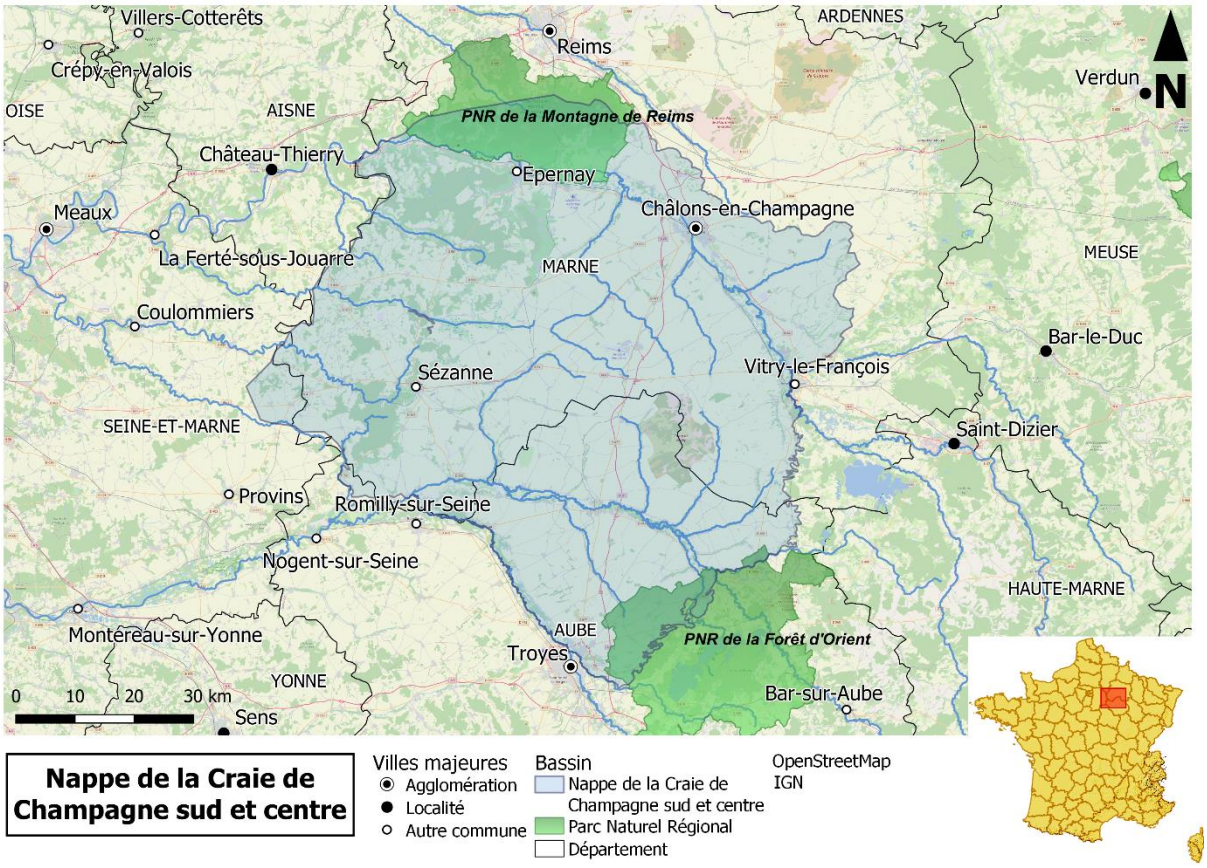
III.1.1. Généralités

Située sur la partie est du bassin parisien, la nappe de la Craie de Champagne sud et centre s'étend sur 5 928 km² principalement sur deux départements qui sont la Marne (73,6%) et l'Aube (25,0 %) (AESN 2015).

Elle est délimitée par la frontière entre le département de l'Aisne et de la Marne au nord-ouest, par la Seine au sud-ouest, les lacs et forêts du Parc Naturel Régional de la forêt d'Orient ainsi que de la Voire au sud-est et par le bassin de la Vesle au nord-est (Carte IV-11).

Cette nappe est libre et affleurante sur 71,5 % de sa surface et captive sur 28,5% (AESN 2015). Dans ses parties captives, elle est surmontée par les nappes alluvionnaires de l'Aube, de la Seine amont et de la Marne ainsi que des plaines entre la Marne et la Seine sur le nord du territoire d'étude.

La gestion de la nappe de la Craie de Champagne est conjointement réalisée par le BRGM et l'Agence de l'eau Seine-Normandie. Elle est principalement utilisée pour l'approvisionnement en eau potable.



Carte IV-11 - Présentation de la nappe de la Craie de Champagne sud et centre.

Elle alimente plusieurs grosses agglomérations telles que Troyes (62 000 habitants), Châlons-en-Champagne (45 000 habitants) et Epernay (23 000 habitants).

III.1.2. Hydrographie et hydrogéologie

III.1.2.a. Hydrographie

La nappe de la craie de Champagne est traversée par plusieurs grands cours d'eau. La Seine délimite la partie sud-est du bassin. Le premier de ses affluents majeurs est l'Aube. L'Aube est un cours d'eau, long de 248 km, qui prend sa source, tout comme la Seine, sur le plateau de Langres, plus précisément dans la commune de Praslay en Haute-Marne à 379 mètres d'altitude.

Son bassin versant s'étend sur 4 660 km² et occupe un tiers du territoire de la nappe souterraine. Son débit moyen à 40 km de sa confluence avec la Seine, sur la station hydrométrique de l'Aube à Arcis-sur-Aube est de 33,3 $m^3.s^{-1}$, avec des variations possibles entre hautes eaux et étiage comprises entre 15,5 et 58,5 $m^3.s^{-1}$ (valeurs moyennes mensuelles) sur une période comprise entre 1960 et 2019 (Banque Hydro 2020).

Ses principaux affluents sont l'Aujon (68 km pour un bassin de 481 km²), la Voire (56,1 km pour un bassin de 896 km²), la Superbe (39,5 km pour un bassin de 320 km²), l'Auzon (39,5 km pour un bassin de 166 km²) et la Barbuise (35,7 km pour un bassin de 200 km²) (Carte IV-12).

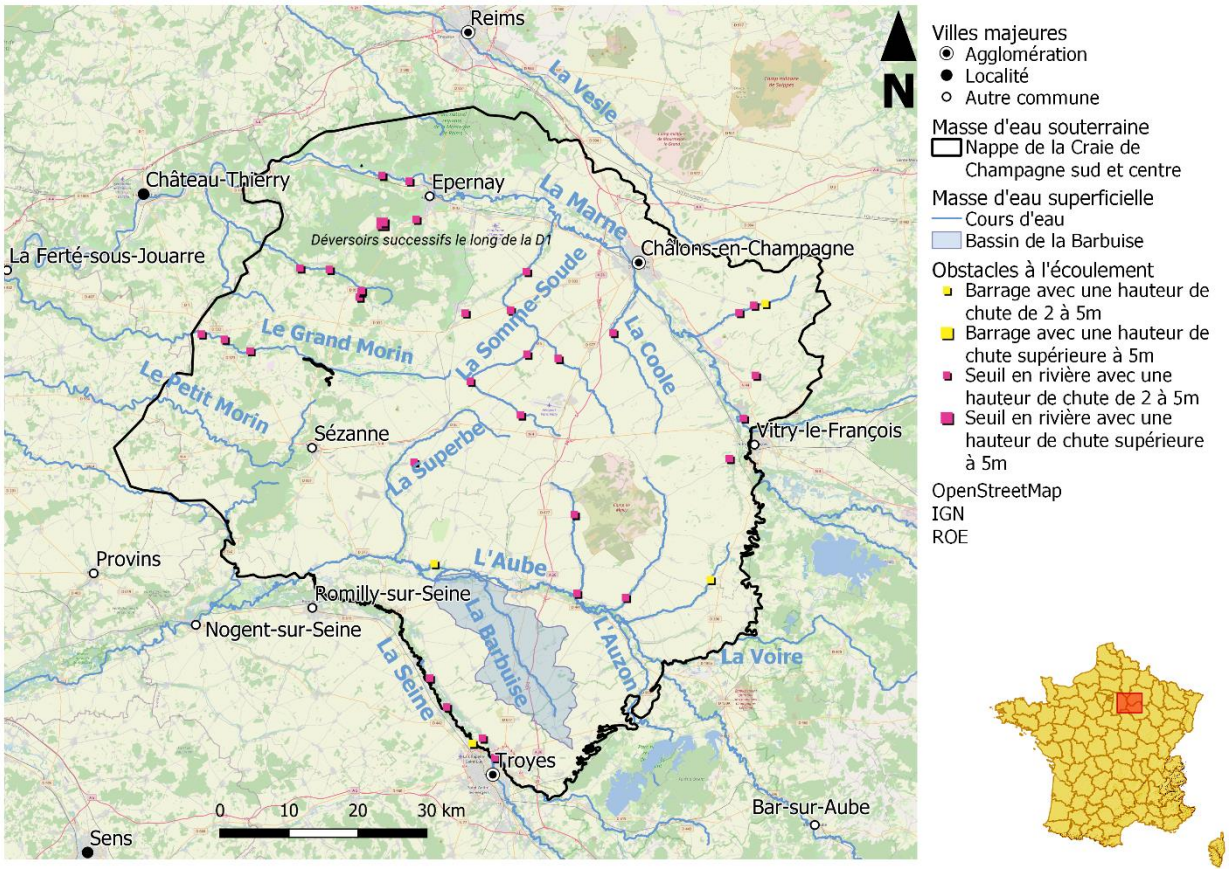
Pour ce dernier cours d'eau, il a été montré que les pompages dans la nappe de la Craie influencent les débits d'étiage, montrant que la Barbuise draine la nappe plus qu'elle ne l'alimente, en particulier en période d'étiage (Pinault, Allier et Chabart 2006).

La Marne traverse de part en part le nord de la nappe de la craie de Champagne. La Marne prend elle aussi sa source au pied du plateau de Langres sur la côte au Clos à 423 mètres d'altitude. Elle effectue ensuite un trajet de 514 km et conflue avec la Seine aux portes de Paris, à Alfortville.

Le bassin versant de la Marne est de 12 920 km², dont 3 458 km² surplombent le bassin de la nappe de la craie de Champagne, représentant ainsi 58,3 % de sa surface. Son débit moyen au niveau de Châlons-en-Champagne, sur la station hydrométrique de la Marne à Châlons-sur-Marne est de 72,1 $m^3.s^{-1}$, avec des variations hydrométriques comprises entre 35,1 et 128 $m^3.s^{-1}$ entre 1957 et 2020 (valeurs moyennes mensuelles) (Banque Hydro 2020).

Ses affluents principaux présents sur la nappe de la Craie de Champagne sud et centre sont le Grand Morin (118,2 km pour un bassin de 1 197 km²), le Petit Morin (86,3 km pour un bassin de 605 km²), la Somme-Soude (59,9 km pour un bassin de 485 km²) et la Coole (30,2 km pour un bassin de 171 km²) (Carte IV-12).

La Marne et l'Aube, traversant toutes deux la nappe de la Craie de Champagne sud et centre, confluent toutes les deux avec la Seine. Malgré la présence de ces grands cours d'eau, le réseau hydrographique est peu dense sur le territoire compte-tenu de la perméabilité importante des sols.



Carte IV-12 - Obstacles à l'écoulement sur la Seine amont, la Marne et leurs affluents.

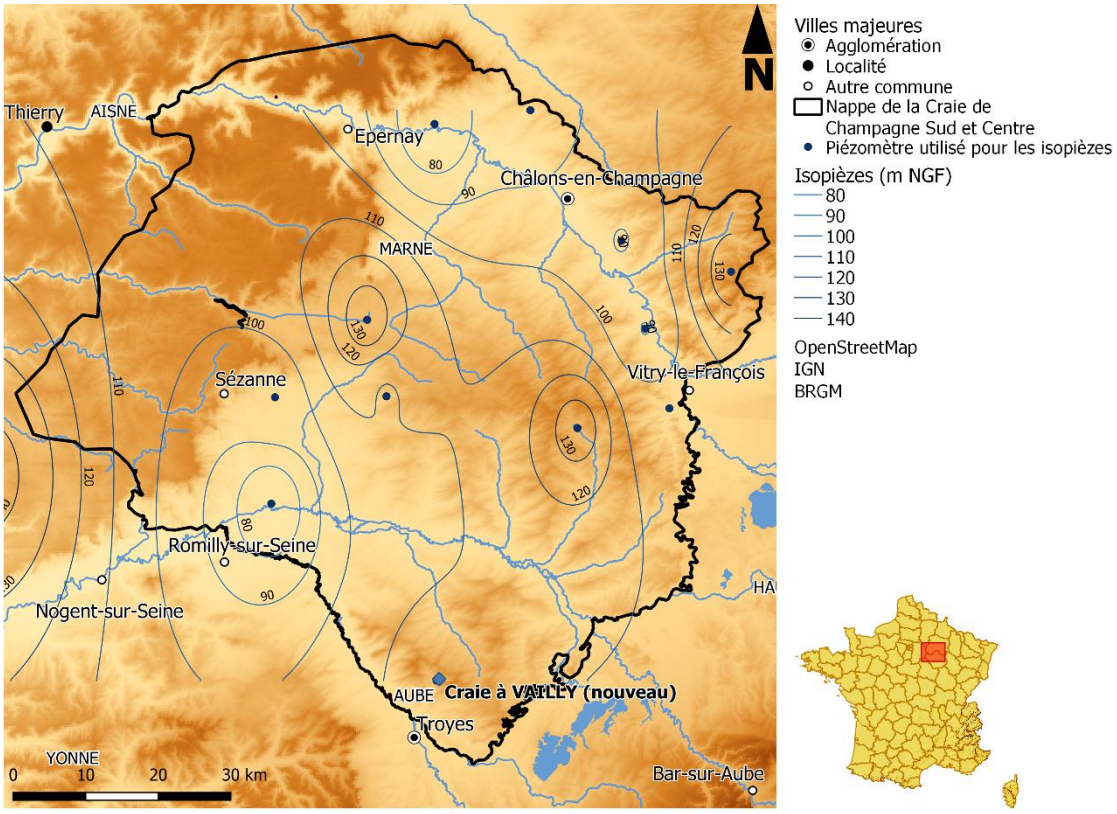
III.1.2.b. Hydrogéologie

La nappe de la Craie de Champagne sud et centre est une nappe sédimentaire (nappe crayeuse) partiellement captive sur sa partie occidentale, compte-tenu de sa superposition avec la nappe tertiaire de Champigny en Brie et Soissonnais, ainsi que sur sa partie septentrionale, avec la superposition des nappes de l'Eocène du bassin versant de l'Ourcq et du Lutétien – Yprésien du Soissonnais – Laonnois. Elle est en forte connexion avec les masses d'eau souterraines alluviales de la Marne, de l'Aube et de la Seine Amont.

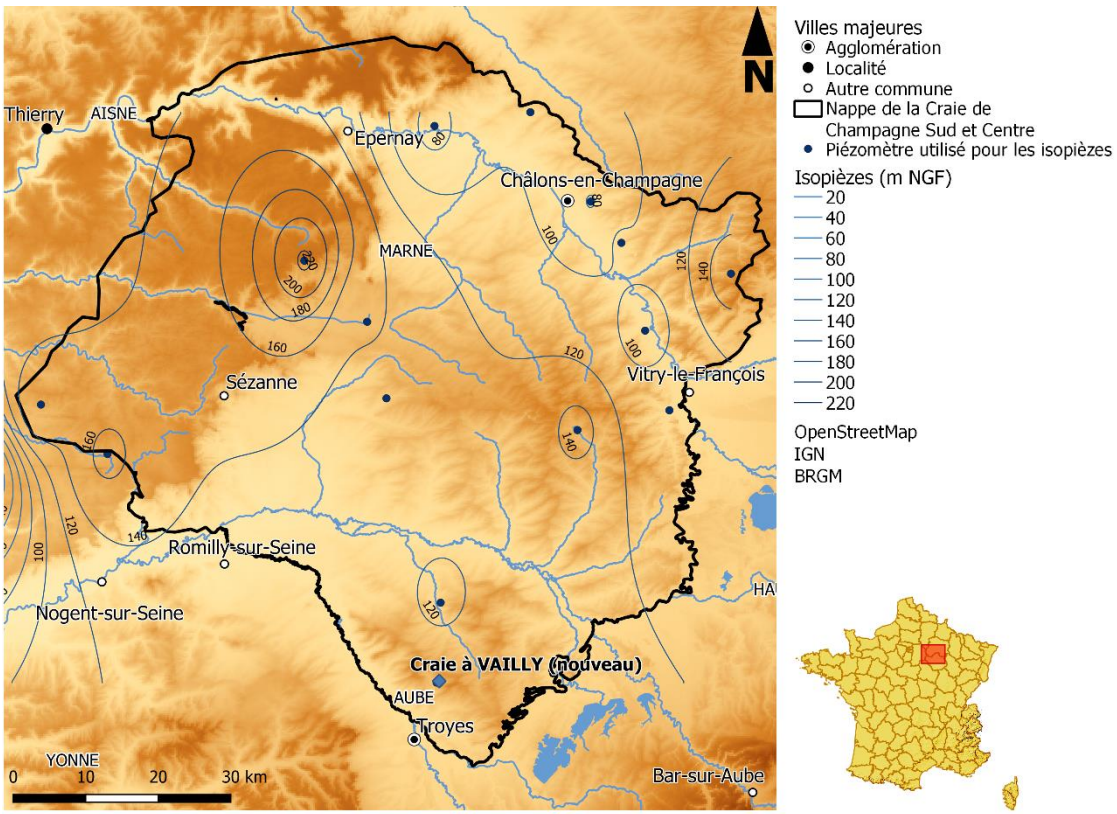
A l'inverse, la nappe de la Craie de Champagne sud et centre surplombe elle aussi la nappe de l'Albien-néocomien captif, qui s'étend sur l'ensemble du bassin parisien, soit près de 110 000 km², pour un aquifère estimé à 12 milliards de mètres cube d'eau (Lallahem 2002). Elle surplombe aussi le long de sa bordure sud-est la nappe des calcaires du Dogger entre Armançon et limite de district (AESN 2015).

Le temps de recharge de la nappe est estimée à 100 jours au niveau du piézomètre de Vailly (Pinault, Allier et Chabart 2006). La profondeur de la nappe est observée pour des évènements sec (mi-octobre 1990) et humide (février 1982), respectivement sur les cartes IV-13 et IV-14. À cause de la superposition des nappes il n'y avait pas de sens à prendre en compte des

piézomètres hors bassin et donc isolés de la nappe étudiée. Le calcul des isopièzes n'a donc pu être défini sur l'ensemble du bassin.



Carte IV-13 – Isopièzes de la nappe de la Craie lors d'un évènement sec (sept.-déc. 1990).



Carte IV-14 – Isopièzes de la nappe de la Craie lors d'un évènement humide (janv.-mars 1982).

Sur ces cartes, on note que les fluctuations piézométriques sont relativement faibles, excepté pour un piézomètre.

Une étude de variogrammes a été réalisé à partir de différents piézomètres répartis sur l'ensemble du territoire de la nappe de la Craie de Champagne. Cette étude a montré que quatre piézomètres sur les sept présents sur la nappe sont significativement représentatifs du fonctionnement de celle-ci (Putot, Verjus et Vernoux 2006). L'étude de ce variogramme a permis d'observer une variation interannuelle moyenne de la nappe de 6 mètres, dans le cas où elle est peu influencée par des cours d'eau (Pinault, Allier et Chabart 2006), (Putot, Verjus et Vernoux 2006). Cependant, sur certains points de mesures, tel que le piézomètre de Vailly, le battement de nappe peut atteindre 25 mètres (AESN 2015).

Le forage associé au piézomètre de la Craie de Vailly est entièrement utilisé pour l'alimentation en eau potable (AEP). D'après la Banque Nationale des Prélèvements en Eau, les prélèvements d'eau annuels effectués sur une moyenne faite entre 2012 et 2017 sont de 17 393 m^3, avec une tendance à la baisse, évoluant de plus de 22 000 m^3 prélevés en 2012 à un peu moins de 12 000 m^3 en 2016, bien que ces prélèvements annuels augmentent de 4 000 m^3 l'année suivante (BNPE 2020). Le graphique présent sur la Figure IV-20 indique l'évolution des prélèvements sur le piézomètre d'étude.

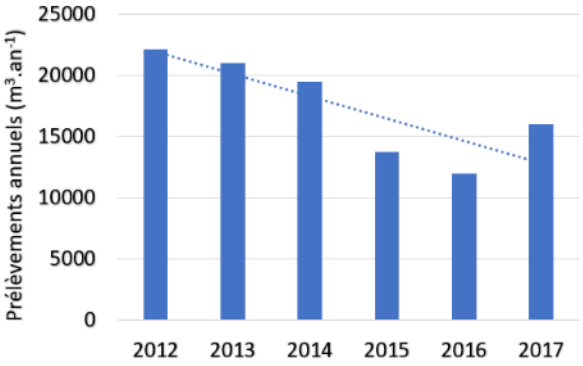


Figure IV-20 - Evolution des prélèvements en eau au piézomètre de la Craie à Vailly.

Les prélèvements d'eau souterraine sont homogènes sur l'ensemble du bassin, et en moyenne plus importants que ceux retrouvés sur le bassin de la nappe du Blavet. Ceux-ci sont majoritairement utilisés pour l'alimentation en eau potable, l'industrie, les canaux et dans une moindre mesure l'irrigation (Martin, Rouxel-David et Batkowski 2004).

Les prélèvements mensuels liés à l'agriculture sont considérés comme constants entre chaque année, avec des prélèvements répartis entre mai et septembre avec un pic de prélèvement sur les mois de juin (30 % des prélèvements) et juillet (40 % des prélèvements) (Stollsteiner 2013).

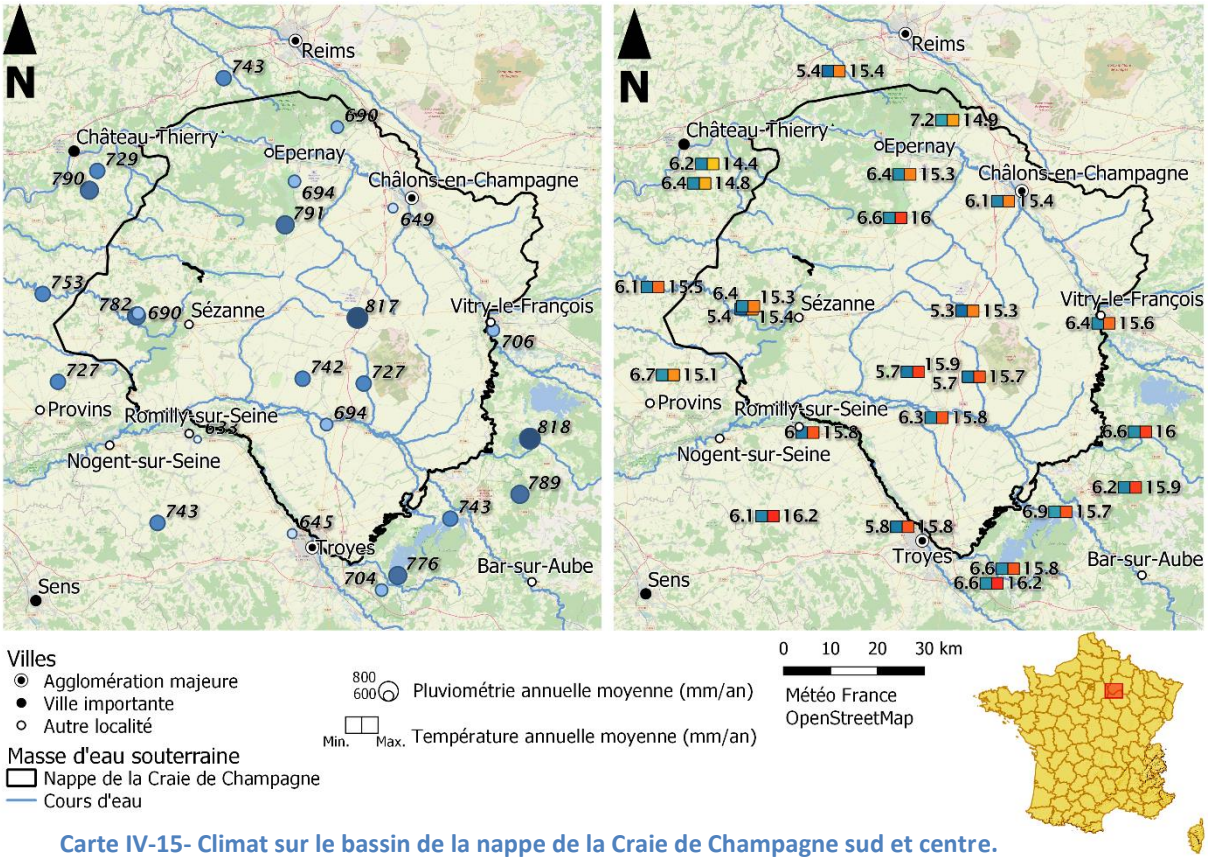
La nappe de la Craie de Champagne interagit fortement avec les cours d'eau superficiels, notamment avec la Barbuise (Pinault, Allier et Chabart 2006), (Stollsteiner 2013). Les prélèvements en eau pour l'irrigation accentuent l'étiage des cours d'eau avec un déphasage d'autant plus grand que le point de prélèvement et le cours d'eau ne soient éloignés (Stollsteiner 2013). Sur le bassin versant de la Barbuise, il est possible de prélever 1,1 millions

de mètres cube d'eau souterraine pour subvenir aux besoins d'irrigation sans que les niveaux d'eau de la Barbuise ne soient impactés, montrant l'importance de la contribution de la nappe à la rivière (Martin, Rouxel-David et Batkowski 2004).

III.1.3. Climat

Le climat sur le bassin de la nappe de la Craie de Champagne sud et centre est de type océanique dégradé, où l'impact climatologique des mers et océans est moins prononcé qu'un climat océanique.

Le climat champenois est homogène sur l'ensemble du bassin. La pluviométrie annuelle moyenne oscille entre 640 et 820 mm/an, comme indique la Carte IV-15. Les variations de cumuls observées sont faibles et très locales, mais non-négligeables. Les températures minimales annuelles moyennes varient entre 5,4°C et 7,2°C tandis que les températures maximales annuelles moyennes sont comprises entre 14,4°C et 16,2°C. Il est possible d'observer que le nord du bassin de la nappe de la Craie de Champagne est légèrement plus frais que le sud, avec un différentiel des minimales et des maximales de l'ordre de 1°C à 1,5°C.



Carte IV-15- Climat sur le bassin de la nappe de la Craie de Champagne sud et centre.

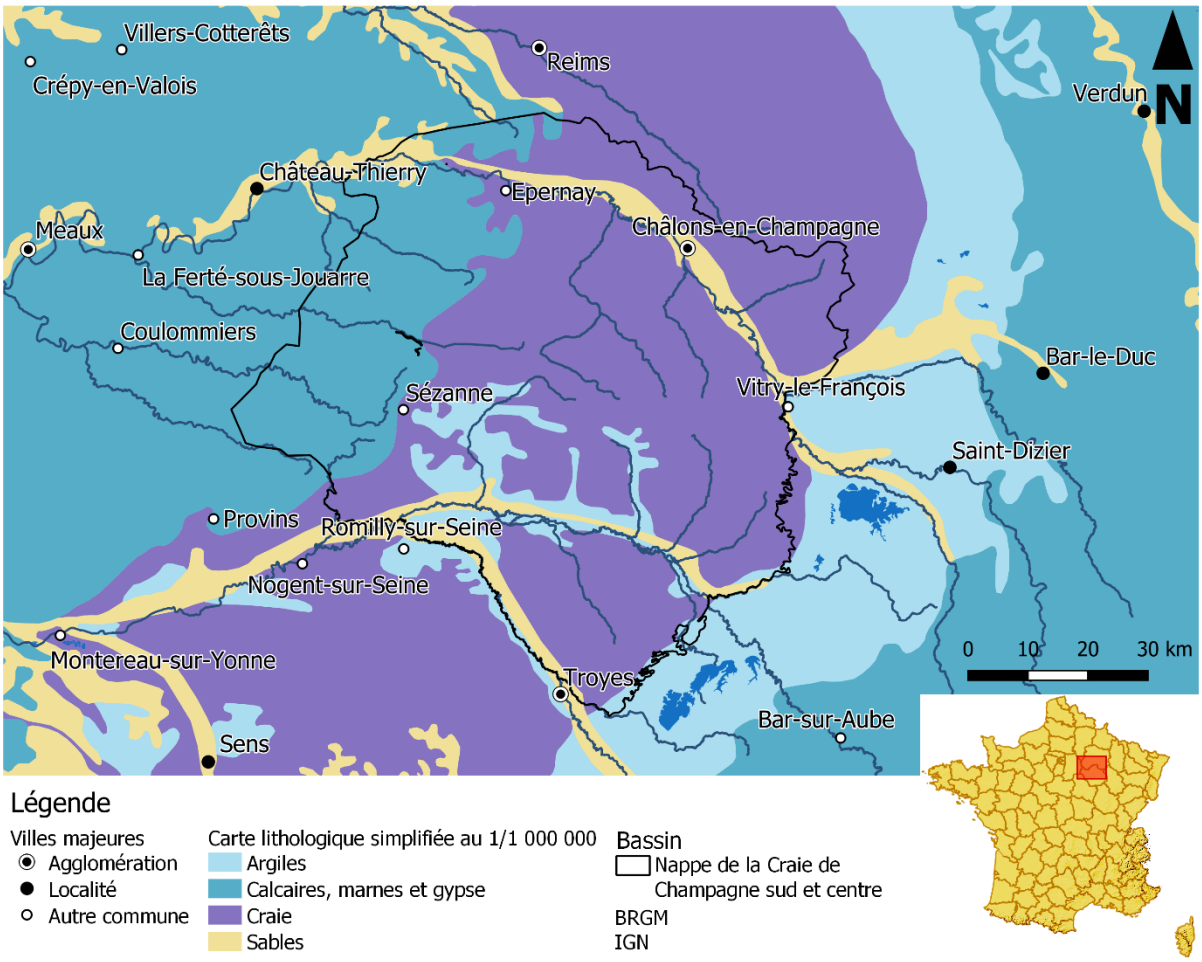
A l'instar de l'étude du climat sur le bassin de la nappe du Blavet, les stations climatologiques de Météo-France utilisées pour construire la carte ci-dessus se basent sur des historiques de données compris entre 1981 et 2010. Toutefois, quelques-unes de ces stations n'existaient

pas en 1981. Par conséquent, les statistiques indiquées sur la Carte IV-15 ne sont montrés ici qu'à titre informatif et n'ont pour but que d'avoir une vue d'ensemble du climat sur le bassin.

III.1.4. Géologie

La nappe de la Craie de Champagne sud et centre siège, comme son nom l'indique, dans des substrats de type sédimentaire. Plus précisément, les craies qui composent la majeure partie de l'aquifère sont des calcaires de type biomicritique avec des grains très fins, de l'ordre de 10 µm sur ce terrain d'étude (Stollsteiner 2013).

Une grande partie de sa surface se trouve donc sur des terrains crayeux datant de la période campanienne jusqu'à la période céno-turonienne (craie blanche et craie marneuse) (Thuon et Julio 2010), (Kloppmann, Dever et Edmunds 1998), bien qu'il soit possible de trouver des calcaires non crayeux et des marnes du Paléogène sur la partie nord-ouest, lorsque des nappes perchées se superposent à la nappe champenoise (AESN 2015), (Cao, et al. 2020). Pour finir, des sables et argiles du sparnacien sont présents le long des cours d'eau (Carte IV-).



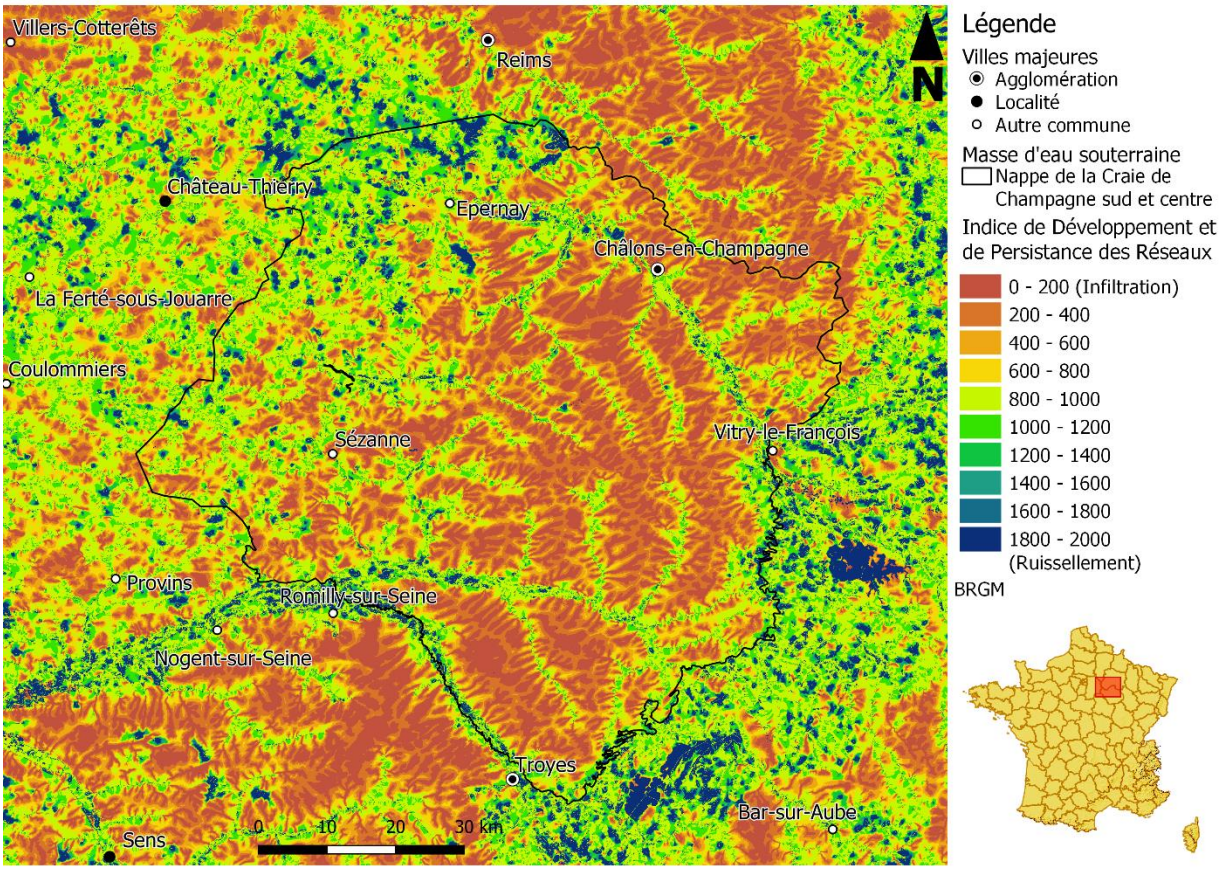
Carte IV-16 - Géologie simplifiée de la nappe de la Craie de Champagne sud et centre.

Ces craies ont une perméabilité et une porosité totale importantes (30 à 40%), laissant ainsi s'infiltrer les eaux superficielles et pluviales aisément (AESN 2015). Cette perméabilité, plus importante en surface, diminue au fur et à mesure que l'on s'enfonce dans l'aquifère, et la craie devient alors de plus en plus marneuse, jusqu'au substratum marneux imperméable (Stollsteiner 2013). La présence de karsts au nord du bassin facilite également l'infiltration des eaux (Lejeune et Devos 2001).

La perméabilité des sols crayeux sur le territoire d'étude se divise quant à elle en quatre catégories (Stollsteiner 2013) :

- La perméabilité « de matrice », correspondant à la structure de la roche et est peu perméable (10^{-6} à 10^{-7} $m.s^{-1}$),
- La perméabilité de diaclase, qui constitue le principal aquifère, avec une perméabilité plus élevée, laissant ainsi le mouvement des eaux souterraines (10^{-2} à 10^{-5} $m.s^{-1}$),
- La perméabilité d'origine tectonique située au niveau des forages d'exploitation à plus de 40 m de profondeur,
- La perméabilité de strate, située à plusieurs centaines de mètres de profondeur (Stollsteiner 2013).

La perméabilité importante sur ce bassin se constate également en analysant la Carte IV-17, montrant l'indice de persistance des réseaux (IDPR) sur le territoire d'étude (Mardhel et Allier 2010).



Carte IV-17 - Indice de Développement et de Persistance des Réseaux (IDPR) sur la nappe de la Craie.

En effet, les zones en rouge favorisent l'infiltration des eaux en surface, et correspondent à la quasi-totalité de la surface de la nappe de la Craie de Champagne, à l'exception de la partie nord-ouest, où la lithologie est plus marneuse que le reste du bassin, et des parties longeant les cours d'eau majeurs – la Seine, l'Aube, la Marne et dans une moindre mesure la Superbe – où le ruissellement est favorisé.

La partie nord-ouest, plus communément appelée « Montagne de Reims », est une région naturelle différente des autres d'un point de vue géologique. En effet, elle est constituée de différentes formations géologiques, le plus souvent calcaires, allant du quaternaire (-2,6 Ma) au Thanétien (-59,2 Ma), reposant sur les craies blanches du Campanien (-83,6 Ma) et du céno-turonien (-100 Ma). Cette différence géologique, par rapport au reste de la nappe, engendre des productions agriculturales et des écosystèmes différents (Lejeune et Devos 2001), (Damien 2019).

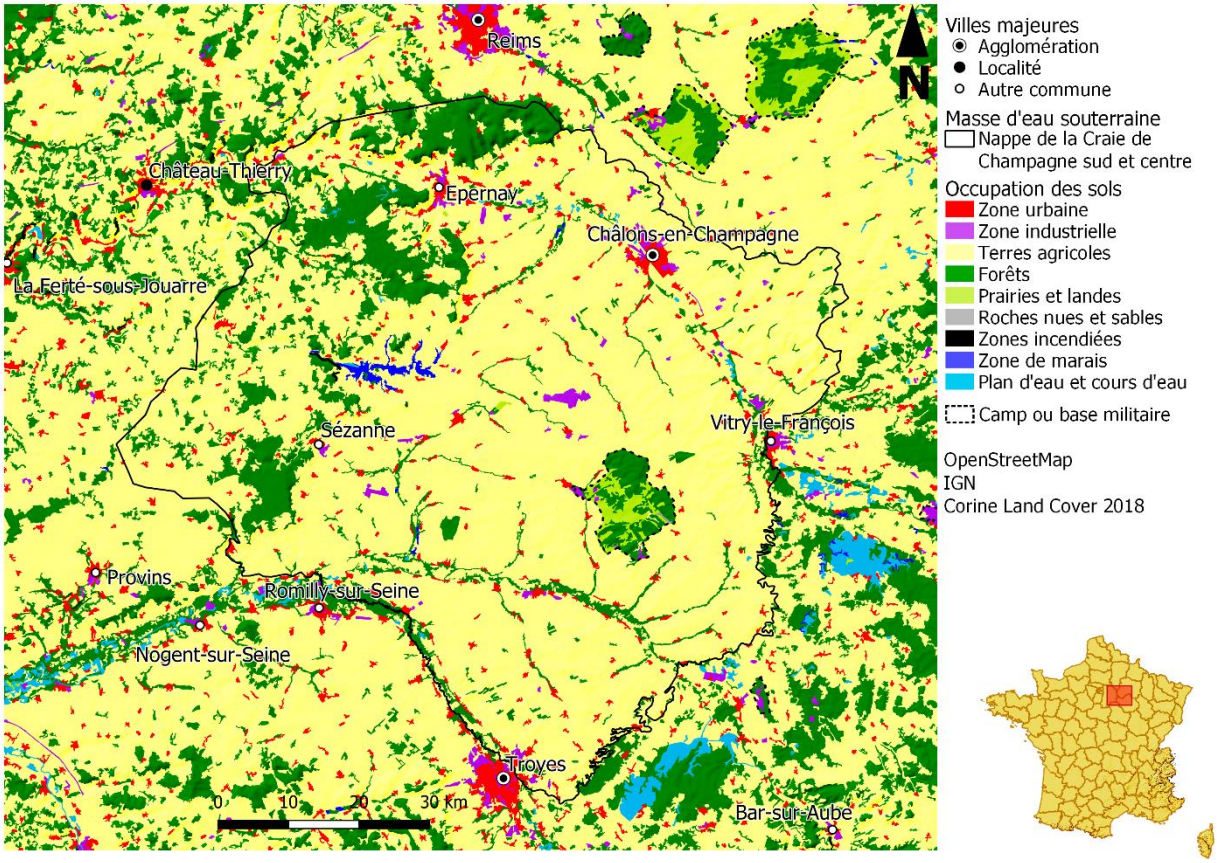
III.1.5. Occupation des sols

Le territoire de la nappe de la Craie champenoise est essentiellement rural, avec plus de 92 % de terres agricoles par rapport à la superficie totale du territoire. Il existe néanmoins quelques forêts représentant près de 6 % du bassin et les industries et centres urbains qui s'étendent

sur moins de 1,5 %. Il est également possible de noter la présence de quelques prairies (présentes principalement sur le camp militaire de Mailly) et marais sur le bassin de la nappe champenoise (Carte IV-18).

La craie peut être utilisée à des fins industrielles, notamment dans le papier, la porcelaine, le plastique, le verre, le caoutchouc, ou encore la peinture, car moins cher que le kaolin (Thuon et Julio 2010). Ainsi, il existe quelques industries sur le territoire utilisant la craie en matière première.

Le territoire est ici majoritairement rural. Les cultures dominent les quelques pôles urbains, répartis le long de la Seine (115 000 habitants sur 29 communes) et de la Marne (133 000 habitants sur 62 communes). Ainsi, en 2016, 492 communes se trouvent tout ou partie sur le territoire de la nappe champenoise, pour une population avoisinant 400 000 habitants (INSEE 2020). Il est à noter que la commune de Troyes a été intégrée dans ces statistiques bien que cette ville ne chevauche la nappe que sur moins de 1 % de sa superficie (Carte IV-18).



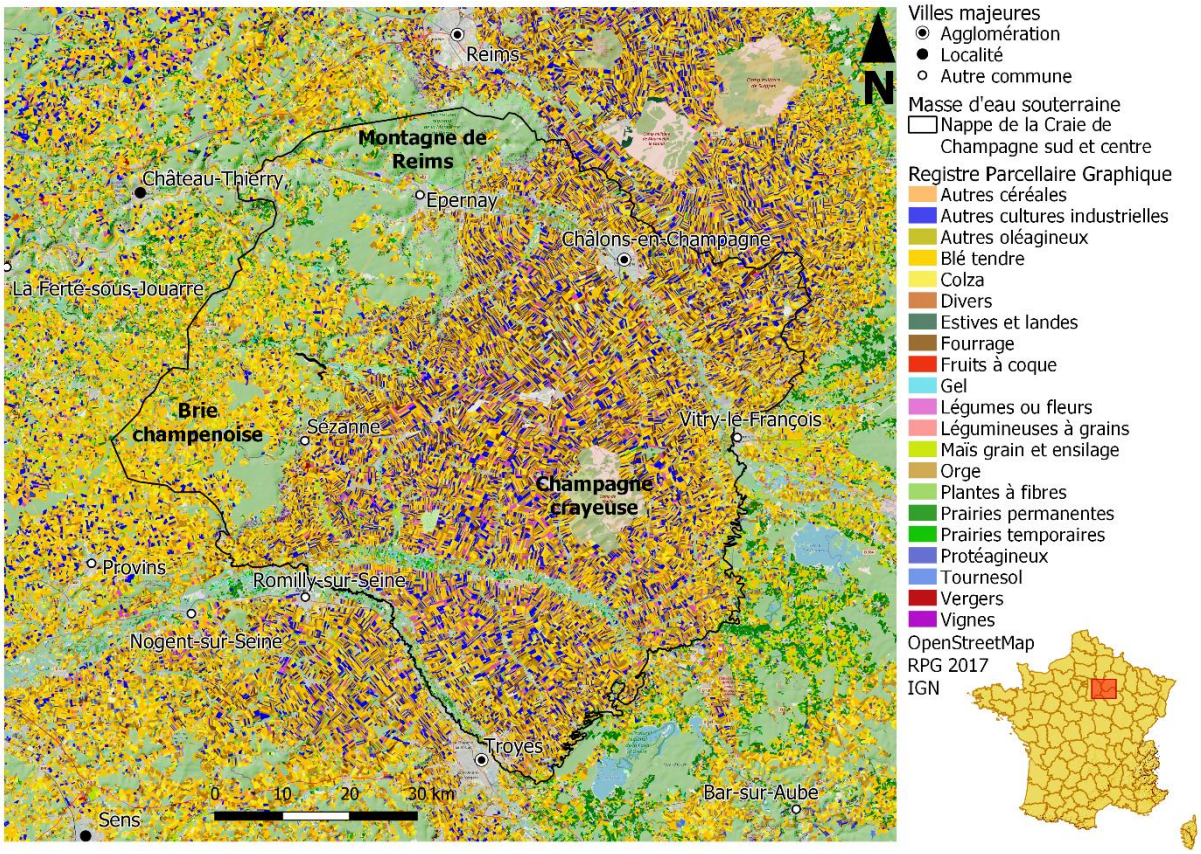
Carte IV-18 - Occupation des sols sur la nappe de la Craie de Champagne sud et centre (Corine Land Cover, 2018).

La nappe de Champagne s'étend également sur trois régions naturelles distinctes. La plus grande d'entre elles est la « Champagne crayeuse ». Également appelée « Champagne pouilleuse », elle occupe la majeure partie du territoire. Ses vastes plaines s'étendent de la vallée de l'Aisne, au pied des Ardennes, jusqu'aux confins du département de l'Aube

(Stollsteiner 2013). D'un point de vue hydrogéologique, elle correspond à la zone où la craie est perméable car directement affleurante. Elle contraste avec la « Champagne humide » qui borde la nappe champenoise sur sa partie orientale, où l'eau est abondante, de par la présence de forêts, de lacs et de zones humides. En revanche, sur notre territoire d'étude, l'infiltration des eaux pluviales étant élevée, les sols en surface sont secs, rendant l'agriculture difficile. La mécanisation et l'arrivée des amendements à la fin du XIX^{e} siècle a rendu l'exploitation des terres possible. Ainsi, à l'exception du camp militaire de Mailly et des ripisylves des cours d'eau, les prairies et forêts laissent place aux terres agricoles (Carte IV-18 et Carte IV-19).

Sur ces terres crayeuses poussent de nombreuses cultures industrielles, telles que la betterave et la pomme de terre. Des cultures de légumes, notamment les pois, viennent clairsemer le paysage champenois. Il est également possible de trouver en assez grande quantité des cultures d'orge, de blé tendre et de colza. Il existe en revanche peu d'élevage sur cette région, à l'exception de quelques troupeaux présents sur les prairies en bordures de cours d'eau. Enfin, il est possible d'observer de petits vignobles à la bordure ouest de la « Champagne crayeuse » (Carte IV-19) (Martin, Rouxel-David et Batkowski 2004), (Damien 2019).

La seconde région naturelle du bassin est la « Montagne de Reims ». Localisée entre Epernay et Reims, le point culminant de ce plateau est le mont Sinaï, atteignant 286 mètres d'altitude. Elle est constituée en majeure partie de forêts et de vignes à l'est et sur la vallée de la Marne. Quelques prairies viennent s'entremêler avec les vastes forêts. Il est toutefois possible de trouver quelques rares cultures de blé, de colza et de maïs. Ces forêts se trouvent principalement sur le plateau, tandis que les vignes se trouvent sur les flancs de collines. Les cultures, quant à elles, se situent au pied de la « Montagne de Reims » (Carte IV-19) (Lejeune et Devos 2001).



Carte IV-19 - Registre Parcellaire Graphique du bassin de la nappe de la Craie de Champagne (RPG, 2017).

La « Brie champenoise » est la troisième région naturelle, localisée sur la partie occidentale de la nappe. A l'instar de la région précédente, le relief de la « Brie champenoise » est plus vallonné que celui de la « Champagne crayeuse », avec des altitudes comprises entre 180 et 220 mètres. Bien que les forêts soient bien représentées sur cette région, les cultures agricoles restent omniprésentes sur ces terres humides (Damien 2019).

Le blé, le colza, le maïs et l'orge sont les fers de lance de l'économie agricole de la « Brie Champenoise ». Il est possible également de voir quelques côteaux de vignobles à la frontière de la « Champagne crayeuse ». Il subsiste également quelques prairies pour l'élevage, temporaires et permanentes, le long des cours d'eau – Petit Morin et Grand Morin – ainsi que quelques cultures industrielles (Carte IV-19) (Damien 2019).

La présence de toutes ces cultures sur ce bassin entraîne un fort besoin en eau. Les apports nécessaires en eau pour l'irrigation des cultures ont potentiellement un impact sur les niveaux d'eau souterraine. Ceux-ci seront estimés lorsque les données et réseaux de mesures seront présentés sur les paragraphes suivants.

III.1.6. Ecosystèmes

La nappe étudiée dispose d'un large panel d'écosystèmes, avec la présence d'une réserve naturelle et de nombreuses zones Natura 2000. La réserve naturelle régionale du marais de Reuves, est localisée au cœur de la nappe de la Craie de Champagne, dans les marais de Saint-

Gond. Riches en histoire, ces marais sont également riches en biodiversité et sont considérés comme étant des réserves d'eau potable. Ces marais, composés principalement de tourbes, regroupent 216 espèces floristiques et abritent de nombreux oiseaux, mais également libellules, papillons et batraciens (Damien 2019), (Museum National d'Histoire Naturelle, FR2100283 - Le Marais de Saint-Gond 1995).

De nombreux marais sont situés le long des cours d'eau tel que la Superbe et de la vallée de l'Aube. Ceux-ci sont des habitats idéaux pour l'avifaune, avec pas moins de 73 espèces d'oiseaux protégées, selon la directive 2009/147/CE vis-à-vis de la protection des oiseaux (Museum National d'Histoire Naturelle, FR2100285 - Marais de la Superbe 1995), (Museum National d'Histoire Natuelle 2005). La vallée de l'Aube abritent également de nombreuses forêts et prairies en ripisylve, abritant 51 espèces de plantes, d'oiseaux, de mammifères, d'invertébrés et de poissons dont 9 étant inscrites dans la directive 92/43/CEE (Museum National d'Histoire Naturelle, FR2100297 - Prairies et bois alluviaux de la basse vallée alluviale de l'Aube 1995).

Il existe également deux savarts sur le territoire de la nappe, au niveau des camps militaires de Suippes et de Mailly-le-Camp. Les savarts sont des steppes rases de rendosols calcaires laissés volontairement en jachère. Ces friches sont typiques des paysages champenois et sont un abri pour certains mammifères et autres espèces animales. Sept espèces différentes sont inscrites dans la directive 92/43/CEE (Museum National d'Histoire Naturelle 1995), (Museum National d'Histoire Naturelle 1995).

La Montagne de Reims est également un endroit où la biodiversité des habitats et des espèces est importante. Le massif forestier de la Montagne de Reims (versant sud) et étangs associés, ainsi que le massif forestier d'Epernay et étangs associés sont deux zones sylvicoles protégées, plus ou moins humides avec la présence d'étangs, de rivières et lacs entourés de chênaies, de hêtraies et de forêts alluviales d'aulnes et de frênes. Au total, pas moins de dix espèces sont suivies, selon la directive 92/43/CEE (Museum National d'Histoire Naturelle 1995), (Museum National d'Histoire Naturelle 1995).

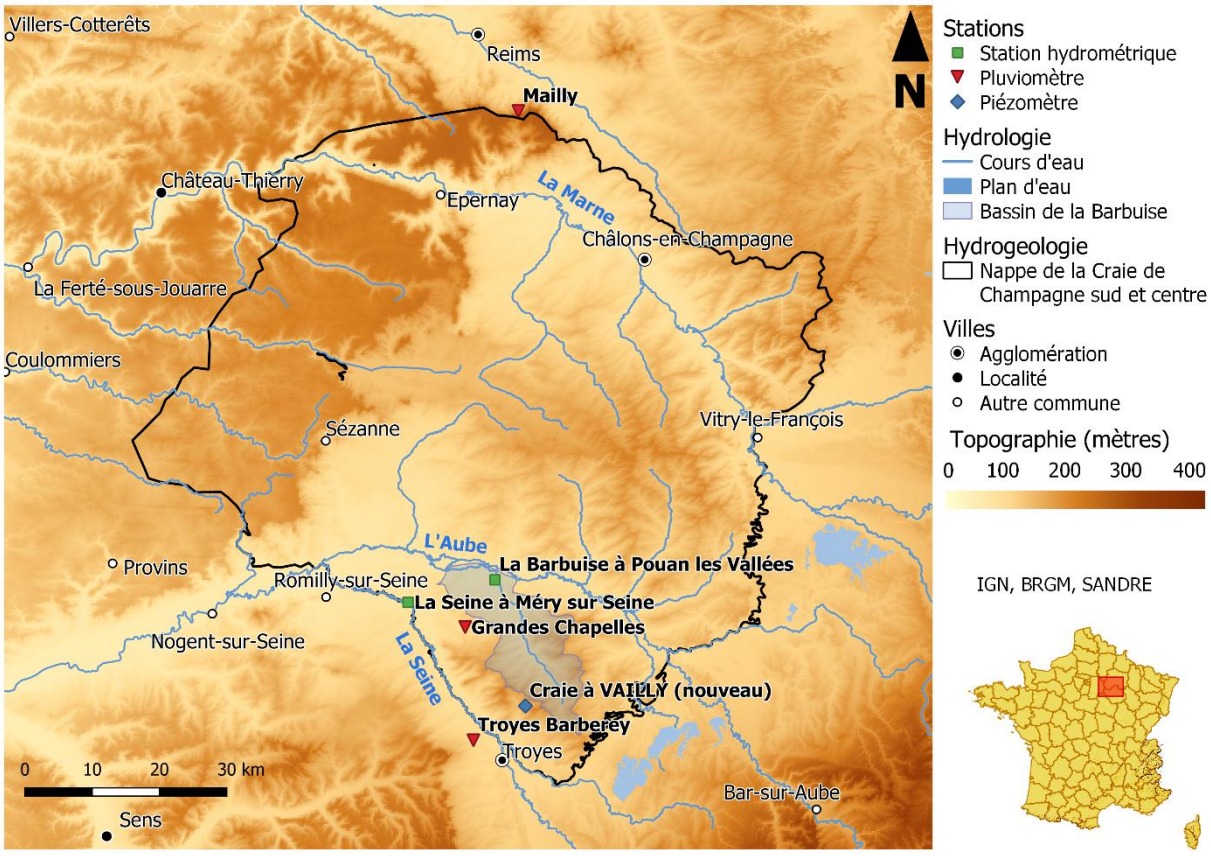
III.2 Réseau de mesures et site d'étude

III.2.1. Présentation des données récoltées

Le site d'étude de Vailly dispose d'un réseau de mesure de sept stations réparties sur le territoire d'étude (Carte IV-20) :

- Trois stations météorologiques :
  - Troyes-Barberey (pluie et évapotranspiration) (10030001),
  - Grandes-Chapelles (pluie) (10166002),
  - Mailly (pluie), station appartenant au Comité Interprofessionnel du Vin de Champagne (CIVC),
- Deux stations hydrométriques :
  - La Barbuise à Pouan-les-Vallées (débit) (H1513210),
  - La Seine à Méry-sur-Seine (débit) (H0810010),
- Une station piézométrique :
  - Piézomètre de la Craie à Vailly (BSS000WKVT).

Le site d'étude dispose également de données reconstituées de prélèvements en eau pour l'irrigation, réparties sur l'ensemble de territoire d'étude.



Carte IV-20 - Réseau de mesures du bassin de la nappe de la Craie de Champagne.

Le Tableau IV-10 récapitule les statistiques des différents types de données récoltées par les différents types de stations :

Tableau IV-10 - Statistiques des stations de la nappe de la Craie de Champagne.

Nom de station	Variable	Unité	Pas de temps	Valeur max.	Valeur min.	Médiane	Moyenne	Période
Craie à Vailly	Niveau d'eau	m NGF	Décade	134,75	109,75	119,95	120,56	1977-2018
La Barbuise à Pouan	Débit	$m^3.s^{-1}$	Décade	4,502	0	0,67	0,836	1972-2018
Seine à Méry-sur-Seine	Débit	$m^3.s^{-1}$	Décade	182,2	5,95	25,61	35,71	1977-2018
Mailly (MA)	Précipitations	mm	Décade	138,8	0	17	21,5	1977-2018
Grandes-Chapelles (GC)	Précipitations	mm	Décade	131,2	0	15,9	19,9	1972-2018
Troyes-Barberey (TB)	Précipitations	mm	Décade	86,4	0	13,2	17,3	1977-2018
Troyes-Barberey (TB)	ETP	mm	Décade	64,7	0	19	21,2	1977-2018
-	Irrigation	$m^3.ha^{-1}.mois^{-1}$	Mois	833,9	0,6	176	280,1	-

Il est à noter que le piézomètre de la Craie à Vailly dispose de données depuis son ouverture en 1969. Toutefois, entre 1969 et 1977, les données piézométriques de cette station ont un pas de temps mensuel. De plus, les chroniques piézométriques ont des lacunes, la plus grande étant entre novembre 1991 et janvier 1995. D'autres périodes plus courtes, d'une longueur d'environ un mois, sont également manquantes, notamment pour les mois d'avril 1985, de janvier et février 2003, de février 2005 et de mai 2008 (Figure IV-21). Ces périodes ont été enlevées pour toutes les stations.

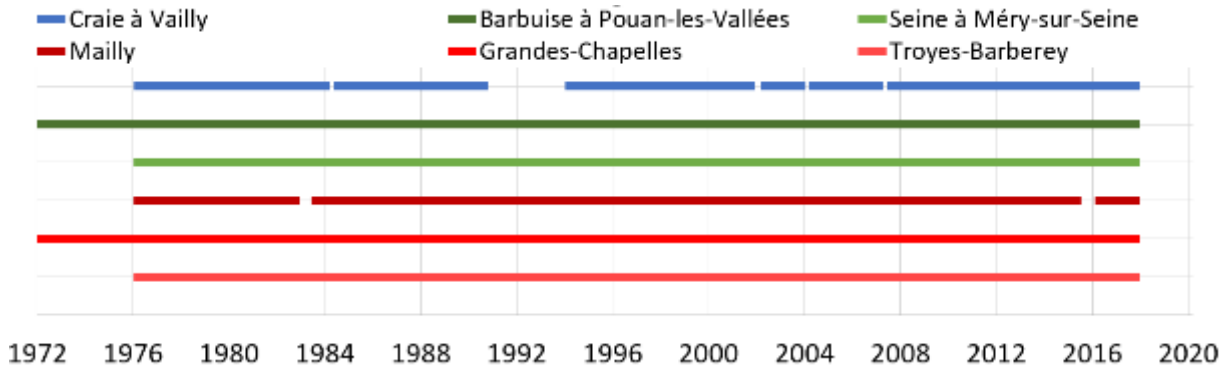


Figure IV-21 - Disponibilité des données (Nappe de la Craie).

De même, les données de pluie à Mailly contiennent deux périodes lacunaires, la première entre décembre 1983 et juin 1984 et la seconde entre août 2016 et février 2017. Celles-ci ont été comblées par la technique des doubles cumuls avec les stations de Reims-Courcy pour la première période et de Reims-Prunay pour la seconde période, compte-tenu de leur proximité avec le pluviomètre de Mailly. Cette technique permet de déterminer des données pluviométriques manquantes, à partir de données pluviométriques d'une station voisine ayant des caractéristiques similaires (altitude de la station, qualité de l'instrument de mesure,

absence de micro-climat, …), en comparant les données des deux stations, sommées selon l'équation (74).

$$P_{DC_{i,j}} = P_{i,0} + P_{i,1} + \cdots + P_{i,j-1} + P_{i,j} \quad (74)$$

Avec P_{DC}, la pluie cumulée,
i, la station pluviométrique utilisée
j, la donnée pluviométrique souhaitée.

Ceux-ci permettent d'observer la linéarité de la relation entre les cumuls des deux stations. Les données sont alors reconstruites à partir de l'équation de la courbe de tendance (Sebbar 2013), (Artigue 2012). Les deux « doubles-cumuls » calculés sont montrés sur la Figure IV-22.

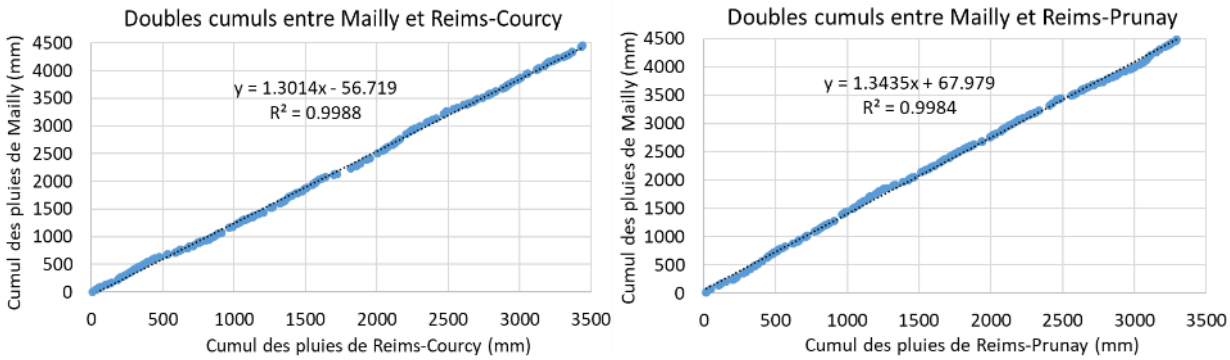


Figure IV-22 - Doubles cumuls pour le pluviomètre de Mailly.

Une fois la technique des « doubles cumuls » appliquée, les pluies du pluviomètre de Mailly peuvent être ainsi calculées avec Reims-Courcy pour 1983-1984 et avec Reims-Prunay pour 2016-2017.

III.2.2. Reconstitution des données de prélèvements pour l'irrigation

Les prélèvements en eau ont potentiellement une influence sur la modélisation. Les prélèvements en eau potable sont essentiellement pour un usage en eau potable, en irrigation et en industrie. Pour la commune de Vailly et les communes limitrophes, ces prélèvements sont utilisés à hauteur de 27,9 à 44 % pour l'adduction en eau potable contre 56 à 72,1 % pour l'irrigation sur une période de 2012 à 2018 (BNPE 2020). Cependant, ces données de prélèvements ne sont pas accessibles à des pas de temps décadaires, mais uniquement à des pas de temps annuels, sur la période entre 2008 et 2018.

Il est donc compliqué de reconstituer l'évolution de ces prélèvements en eau selon les périodes d'une année. Il est possible de s'en approcher pour un usage en irrigation, en utilisant des données agricoles. En combinant les données du Registre Parcellaire Graphique (RPG) de l'IGN et les besoins en eau mensuels de chaque type de culture (issus de l'OUGC 84), il est possible de reconstituer les irrigations potentielles sur le bassin. Ici, il est supposé que les besoins en eau d'une culture correspondent exactement au volume irrigué par parcelle.

Les principales cultures présentes sur le bassin avec leur besoin en eau par mois sont recensées sur le tableau présenté en Annexe II.

Une courbe d'irrigation moyenne par culture et par unité de surface est alors obtenue pour une année. En admettant que l'irrigation soit relativement constante au cours des années, il est possible d'appliquer cette courbe d'irrigation sur l'entièreté des chroniques. Le pas de temps de cette courbe étant mensuel, la courbe a été rééchantillonnée au pas de temps décadaire puis a été lissée. Ces irrigations moyennes potentielles par unité de surface sont ensuite appliquées en entrée du modèle (Figure IV-23).

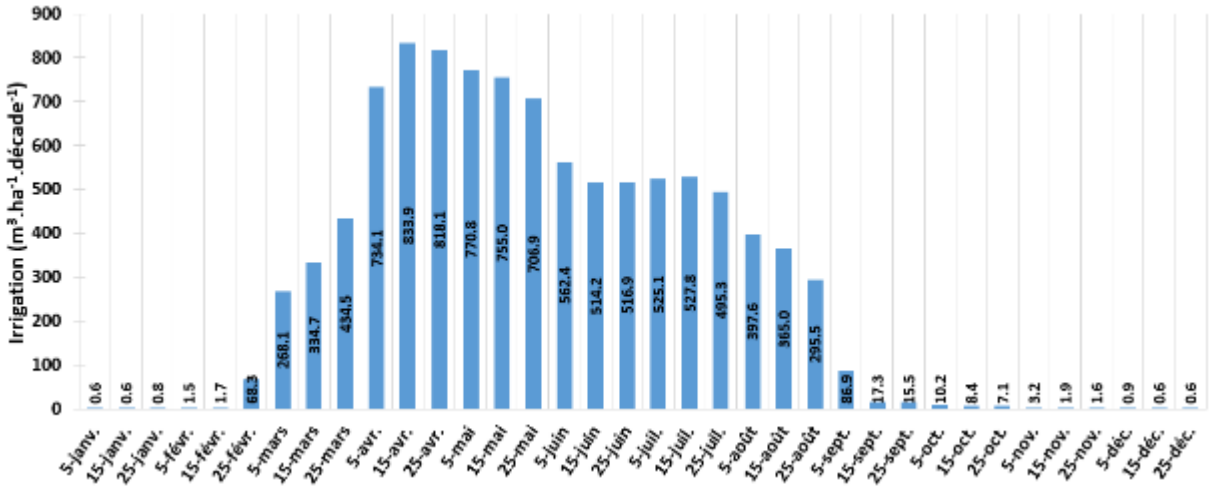


Figure IV-23 - Courbe d'irrigation potentielle par unité de surface.

Toutes ces données sont ensuite intégrées en entrées d'un modèle neuronal pour réaliser une prévision des niveaux d'eau souterraine avec un horizon de prévision de 20 jours, qui fait l'objet d'une publication présentée les paragraphes suivants.

III.3 Résultats des expérimentations

L'objectif est ici de concevoir un modèle capable fournir une prévision de qualité suffisante à moyen-terme (20 jours, soit deux pas de temps de 10 jours), destinée à la gestion de la sécheresse. En effet, si la problématique de la sécheresse est présente partout en France, cette région présente de nombreux enjeux sensibles, tels que la viticulture, l'agriculture et l'alimentation en eau potable.

Pour cela, les ensembles d'arrêt et de test choisis correspondent à des périodes présentant au moins un épisode sec. La période 2011-2013 a été choisie comme ensemble d'arrêt et la période 1988-1990 comme ensemble de test. Cette période est celle qui contient le plus de données piézométriques en dessous de la 2,5ennale sèche (correspondant au seuil Indicateur Piézométrique Standardisé (IPS) « Niveaux modérément bas ») obtenue mensuellement et qui franchit ce seuil le plus souvent. Pour cette étude, les données ont été divisées en sous-

ensembles de trois ans (hors lacunes) selon un cycle annuel (de janvier à décembre) afin d'observer au mieux les périodes de sécheresse. L'utilisation de données décadaires en entrée du modèle permet à celui-ci de fournir une prévision à un horizon de 20 jours. L'architecture du modèle est sélectionnée sur la base de plans d'expériences utilisant les différentes méthodes de régularisation et analyses corrélatoires détaillés dans l'article qui suit (N. Akil, G. Artigue et M. Savary, et al. 2021).

water

MDPI

Article

Uncertainty Estimation in Hydrogeological Forecasting with Neural Networks: Impact of Spatial Distribution of Rainfalls and Random Initialization of the Model

Nicolas Akil [1,2,*], Guillaume Artigue [1,*], Michaël Savary [2,*], Anne Johannet [1] and Marc Vinches [1]

[1] HydroSciences Montpellier, Univ. Montpellier, IMT Mines Ales, IRD, CNRS, 30100 Ales, France; anne.johannet@mines-ales.fr (A.J.); Marc.vinches@mines-ales.fr (M.V.)
[2] AQUASYS, 2 rue de Nantes, 44710 Port-Saint-Père, France
* Correspondence: nicolas.akil@aquasys.fr (N.A.); guillaume.artigue@mines-ales.fr (G.A.); michael.savary@aquasys.fr (M.S.)

Abstract: Neural networks are used to forecast hydrogeological risks, such as droughts and floods. However, uncertainties generated by these models are difficult to assess, possibly leading to a low use of these solutions by water managers. These uncertainties are the result of three sources: input data, model architecture and parameters and their initialization. The aim of the study is, first, to calibrate a model to predict Champagne chalk groundwater level at Vailly (Grand-Est, France), and, second, to estimate related uncertainties, linked both to the spatial distribution of rainfalls and to the parameter initialization. The parameter uncertainties are assessed following a previous methodology, using nine mixed probability density functions (*pdf*), thus creating models of correctness. Spatial distribution of rainfall uncertainty is generated by swapping three rainfall inputs and then observing dispersion of 27 model outputs. This uncertainty is incorporated into models of correctness. We show that, in this case study, an ensemble model of 40 different initializations is sufficient to estimate parameter uncertainty while preserving quality. Logistic, Gumbel and Raised Cosine laws fit the distribution of increasing and decreasing groundwater levels well, which then allows the establishment of models of correctness. These models of correctness provide a confidence interval associated with the forecasts, with an arbitrary degree of confidence chosen by the user. These methodologies have proved to have significant advantages: the rigorous design of the neural network model has allowed the realisation of models able to generalize outside of the range of the data used for training. Furthermore, it is possible to flexibly choose the confidence index according to the hydrological configuration (e.g., recession or rising water table).

Keywords: neural networks; uncertainty; hydrogeology; probability; probability density function; model; rainfall

Citation: Akil, N.; Artigue, G.; Savary, M.; Johannet, A.; Vinches, M. Uncertainty Estimation in Hydrogeological Forecasting with Neural Networks: Impact of Spatial Distribution of Rainfalls and Random Initialization of the Model. *Water* **2021**, *13*, 1690. https://doi.org/10.3390/w13121690

Academic Editor: Renato Morbidelli

Received: 28 April 2021
Accepted: 14 June 2021
Published: 18 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Water is an essential resource for life on Earth but also a hazard, through its scarcity during droughts or its abundance during floods. Water-related risks sometimes cause damage and fatalities and have a strong impact on water supply, agriculture and industries. The current climate change context has causes the rise of extreme phenomena frequency and duration [1]. Moreover, water demand is growing in developed countries due to change in water uses, thus becoming a major issue. Predictive systems can be used in order to manage and to prevent these hydro(geo)logical risks. Among the available solutions to forecast groundwater level or river discharge, two stand out. The first consists in using physically based models, which are supposed to represent a deep knowledge of the study basin. Unfortunately, this level of knowledge is often difficult to reach because of the heterogeneity and anisotropy of hydro-systems. Besides, these models require meteorological forecasts whose reliability, at the necessary space and time scales, can be insufficient. The second consists in statistical modeling, among which artificial neural networks (ANN) are widely represented. This family of models does not require a strong

knowledge about the system behavior, but does require a database representing the assumed relationship between input and output data. Moreover, ANN do not nec-essarily require forecasts of their inputs to provide output forecasts, as long as the lead-time remains below or close to the response time of the system. A specific model, the multilayer perceptron, is known to be able to identify any nonlinear but differenti-able function thanks to the universal approximation property [2]. The choice of ANN models thus relies on poor knowledge of the underground processes of the area and on very low operational calculation times. This avoids making an uncertain hypothesis about the hydrosystem. As any model will, ANN generate uncertainties that are diffi-cult to quantify and to communicate, the absence of which would optimize the deci-sion-making process for end users. Especially, as is often the case in hydrology and hy-drogeology, if the decision is based on threshold crossing, ambiguities in decisions are not acceptable. Thus, uncertainties can lead to mistakes and inconsistencies in deci-sion-making. For these reasons, [3] focus on the key issues for modelers, especially the issue of how far the model has been able to capture the catchment behavior and [4] focuses on the origins of the uncertainties. [4] described three main origins for these uncertainties: (i) input data, especially noise and non-measured spatial variability of these data, (ii) oversimplified structure and (iii) parameters determination. The origins of input data uncertainties are mainly related to data quality, input representativity in the basin related to spatial distribution, environmental conditions and errors from measurements (resolution, measuring instruments) [4,5]. The uncertainties in the model parameters are found in training performance, as well in initialization during the training step [4]. Bayesian Model Averaging was developed for parameter uncer-tainty estimation [6] and applied to the Pô streamflow in Italy. Uncertainty is repre-sented as a forecast interval with a certain probability of correctness [6]. As long as the chosen model generates multiple outputs with an ensemble model, an interval of un-certainty can be drawn. The uncertainty related to various input variables can thus be addressed, for example, the spatial distribution of rainfalls. The latter was approached with a Bayesian Forecasting System [7], coupled to a precipitation forecast, with inter-esting results [8], indicating that the uncertainty, estimated by the prediction interval delivered by the ensemble model, could be improved by post processing. Therefore, Bayesian models can be an alternate solution to estimate model uncertainties. This ap-proach is therefore an effective method for approximating the uncertainty of the var-ious hyperparameters of a model.

In the present paper, we propose a methodology to estimate the uncertainty gen-erated by both the neural network model itself and by the non-measured spatial het-erogeneity of rainfall. This work was carried out on the Champagne chalk aquifer (Northern France) as a case study, at a 10-day time-step. Predictions are achieved for up to 20 days (two time-steps). A reliable enough ANN model to forecast Champagne chalk groundwater level is built, thanks to a rigorous variable and complexity selection process [9,10] helped by the application of regularization methods [10], mainly cross-validation [11] and early stopping [12]. The uncertainties due to the model pa-rameters are then estimated [13]. Even though the noise in rainfall inputs is assumed to be limited due to the quality of rain gauges and to the 10 day time-step, the spatial variability of rainfall is significant, and the number of rain gauges may be insufficient to properly represent this variability. An original method to assess this uncertainty is to perform permutations and substitutions of the available rain gauges, simulating a spatial variation of rainfall and thus giving an ensemble of forecasts. These new fore-casts are finally embedded in the assessment and the representation of uncertainties method.

The article is organized as follows: the material and methods section presents the neural networks and the method used for model design. Then, the target basin used to implement this method, the Champagne chalk groundwater basin, the quality criteria used and, finally, the method used to estimate the uncertainties are described. A

correctness model implementation method is proposed, allowing the display of a confidence interval, choosing an a priori confidence index. Then the design method deployed to realize the neural network model is detailed in Section 3. Section 4 goes on to present the results of the prediction and an uncertainty estimation. Section 5 proposes a discussion and some paths of improvement, before the general conclusion.

2. Material and Methods

2.1. Neural Network Models

2.1.1. Definitions

An artificial neuron is a mathematical operator that first calculates its potential, i.e., the weighted sum of its inputs with its parameters, and secondly its output, applying a nonlinear transformation to its potential.

Neurons are combined inside a network following an architecture, which is built according to the targeted function: classification or regression. Neurons can be organized in layers of two types: (i) output layer, whose outputs are those of the model, and associated to measured values, and (ii) hidden layers, whose outputs are not associated to measured values [14].

One of the most common forms of neural network architecture in hydrology is the feed-forward model "multilayer perceptron" (MLP), for which the universal approximation property has been shown by [2,15]. This property states that the model (Figure 1) is able to approximate any differentiable function with an accuracy depending on the number of hidden neurons. The other property of this architecture is parsimony. This means that the model needs less parameters to describe phenomena, compared to other statistical nonlinear models [16]. This comes from the calculation of the output, which depends non-linearly on the inputs and the parameters. Parsimony is even more valuable when the number of input variables increases. For these reasons, this model is used in this study.

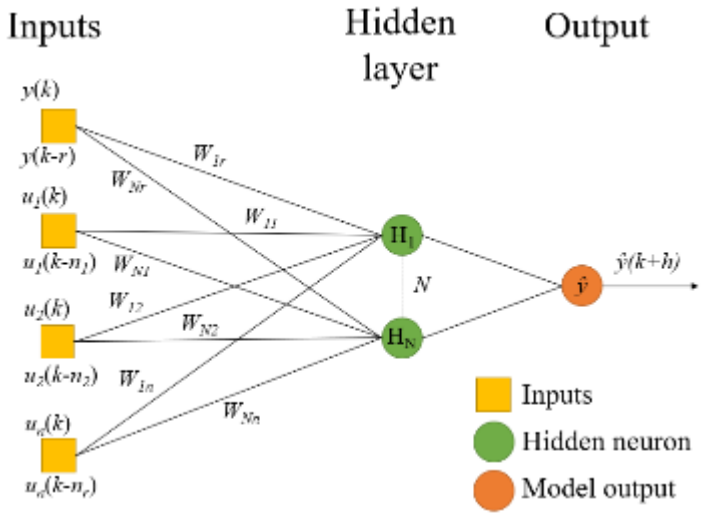


Figure 1. Multilayer Perceptron representation, with x_i, the exogenous variables; **W**, the matrix of parameters; y, the measured output; $\hat{y}$, the predicted output; r, the order of the model, n_r, the input window width; H_i (i = 1 to N) the hidden neurons; N, the number of hidden neurons; k, the discrete time and h, the lead time [13].

2.1.2. Role of Time in Neural Networks Models

Assuming the crucial role of time in forecasting, a neural network can have a dynamic behavior, according to its architecture [17].

In the first case, the static character of the model implies that time has no functional role and that input variables are all exogenous (1). The model is a finite impulse response filter.

$$\hat{y}(k) = \varphi(\mathbf{u}(k), \dots, \mathbf{u}(k - n_r + 1), \mathbf{W}) \quad (1)$$

where $\hat{y}(k)$ is the estimated output at the discrete time k; u the input vector; φ the nonlinear function implemented by the neural model; n_r the sliding time-window size that defines the length of the necessary history of exogenous data; and W the matrix of parameters.

In the second case, the recurrent model uses the result of the simulation at previous time-steps in addition to the exogenous variables (2). The model is thus an infinite impulse response filter and presents a dynamic behavior.

$$\hat{y}(k) = \varphi(\hat{\mathbf{y}}(k-1), \dots, \hat{\mathbf{y}}(k - r); \mathbf{u}(k), \dots, \mathbf{u}(k - n_r + 1), \mathbf{W}) \quad (2)$$

with the same notations as previously stated, and r, the order of the recurrent model; that is to say, the number of previous output values applied at the input of the model.

The recurrent model allows the simulation of a dynamic function: it is used when the noise added by output measurements is supposed to be higher than that affecting inputs. Practically, recurrent models must also be used when real-time data are unavailable [18].

In the third case, the feed-forward model, the recurrent inputs are substituted by the measurement of the output variables at previous time steps (3). The model is thus a finite impulse filter. It is static, rigorously speaking, but thanks to the addition of observed state variables as exogenous input variables, it is able to simulate dynamic behavior.

$$\hat{y}(k) = \varphi(\mathbf{y}(k-1), \dots, \mathbf{y}(k - r); \mathbf{u}(k), \dots, \mathbf{u}(k - n_r + 1), \mathbf{W}) \quad (3)$$

where $y(k)$ is the observed output of the simulated system at the discrete time k.

Feed-forward models are used if the noise due to the measurement of the output variables is low, or lower than the noise on inputs [17,19].

2.1.3. Training and Overfitting

Training a neural network consists of calculating the parameters set by minimizing a cost function, measuring the error between observed and simulated values. This stage uses a training rule applied on a subset of the database: the training set.

Afterward, the quality of the model is assessed on another subset: the test set. The test is used to assess the property of "generalization" of the model. It is never used during training, nor for optimization of the architecture.

Ref. [20] showed that the training error is not a relevant estimator of the test error. Indeed, the training error diminishes when the complexity (number of free parameters) increases, while the test error (the variance) increases. This phenomenon is called "bias/variance dilemma", and indicates that a too complex model perfectly fits the training data, also including the random noise contained by these data. This model is thus unable to correctly generalize to unknown data. Conversely, a too simple model will not be able to adapt to the signal. This leads to a high bias and a low variance.

2.1.4. Regularization Methods

Regularization methods can be used to prevent overfitting. The cross-validation method [11] is used in this study. It allows assessment of the generalization error and provides the cross-validation score that assesses the variance. It is used to select the variables and the complexity of the model (see Section 3).

The second regularization method used in this study is early stopping. Early stopping [12] considers that training the model too much increases the amplitude of the parameters and is equivalent to increasing the complexity. It thus stops training before the generalization capacity decreases. For this purpose, a subset of the database, called the "stop set" must be defined.

As shown by [21,22], the ensemble model can be used to reduce the uncertainty due to the initialization of parameters before training. At each time-step, the output is calculated as (4):

$$\hat{y}_M^k = Median_X\left(\hat{y}_i^k\right) \quad (4)$$

where $\hat{y}_M^k$ is the output of the ensemble model at time-step k; y_i^k is the output of one member of the ensemble at time-step k; and Median$_X$ represents the median calculated on the outputs of a set of X models. The choice of the number of models in the ensemble depends on the application.

2.1.5. Model Design

The first step in the model design is to split the database into several subsets: training set, stop set and test set. Various test sets can be chosen depending on the modelling purpose. For example, the most intense event can be chosen when floods are targeted.

The second step consists in choosing the relevant kind of model regarding the role of time (Section 2.1.2).

The third step is to realize cross-correlations between input data and between input and output data. These correlations allow the obtaining of a response time and a memory effect duration for each input variable. This gives a first overview of the hydro-system dynamics and allows preselecting the input windows width of the model and a relevant lead-time (Figure 1).

Then, in order to prevent overfitting, the architecture of the model is optimized using the cross-validation method. For this optimization, the output is that of the ensemble model. This architecture selection is carried out by adjusting the following hyperparameters using cross-validation:

- The window widths of the different (exogenous) input variables (n_r in Equations (1)–(3)).
- The "order" of the model, corresponding to the window width of the estimated (or observed, if the model is a feed-forward) targeted variable (output variable), for previous time-steps, applied at the input of the model (r in Equations (2) and (3)).
- The number of neurons in the hidden layers: N.

2.2. Study Area: The Champagne Chalk Groundwater Basin

2.2.1. Field Study Presentation

- Location

  Located in Northern France, in the Grand-Est region, the Champagne chalk groundwater basin area is estimated at 5927 sq.km. It corresponds mainly to the drainage of the rivers Marne and Aube, delimited by piezometric ridges characterized as follows: chalk limit on the eastern part, tertiary rocks on the western part, other hydrogeological basins on the northern limit, the Seine river for the southern part and, as a bedrock, marlstones [23]. Elevation varies from 40 to 286 m.a.s.l. (Figure 2).

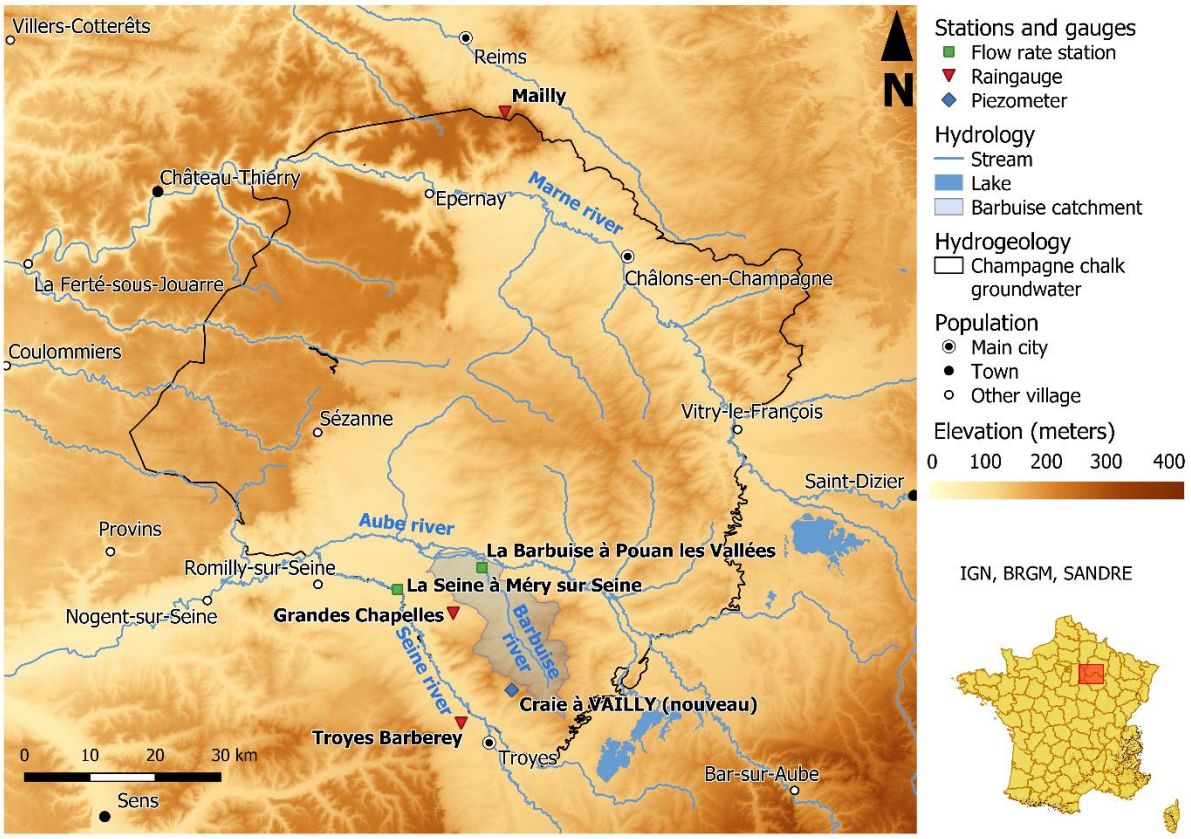


Figure 2. Map of the Champagne chalk groundwater basin (IGN, BRGM, BD Carthage).

- Water use

  Water is mainly used for tap water production and agriculture [23]. Annual water withdrawals via studied piezometer made on average between 2012 and 2017 are 17,393 m^3, however, showing a decreasing trend [24]. Water is also used for agriculture, with 61.5% of groundwater withdrawal for irrigation in 2017 (against 38.5% for tap water production) in Vailly (location of the studied piezometer) and neighboring towns [25].

- Climate

  The climate of the basin corresponds to a transition climate between oceanic and semi-continental climates. The mean annual rainfall varies from 640 to 820 mm, measured on 22 meteorological stations calculated on the 1981–2010 period [26]. The recharge is estimated at 160 mm/y [23].

- Geology and groundwater behavior

  This basin is mainly composed of chalk, and limestones to a lesser proportion, with sands and clay along the hydrographic network [27,28]. Intense shallow fracturing, mainly caused by climate action, has developed a significant permeability especially near the hydrographic network. Groundwater recharge time in the champagne basin is estimated at 100 days in our study piezometer (Craie à Vailly (nouveau)) [29], and the underground levels can increase from 6 m to 25 m [23,30]. Groundwater levels, especially in the Barbuise catchment area, which is close to the study piezometer, are influenced by the shallow water [27]. Consequently, the Barbuise river discharge is strongly correlated to piezometric levels at Craie à Vailly [27,29].

2.2.2. Database Presentation

Meteorological data are provided at three stations:

1. Troyes-Barberey (R_{TB}) (precipitation and potential evapotranspiration),
2. Grandes-Chapelles (R_{GC}) (precipitation),
3. Mailly (R_{MA}) (precipitation)

Groundwater levels are available at the Craie à Vailly piezometer (L_{CV}), and two discharge stations are located in Pouan-les-Vallées in Barbuise catchment (D_{BP}) and at Méry-sur-Seine in the Seine catchment (D_{SM}) (Table 1).

Irrigation (I) information was calculated from cultivated area and agricultural data. The cultivated area can be estimated thanks to (RPG2017 from IGN [31]). Agricultural data contains monthly water demand for each crop type (OUGC84). Therefore, it is possible to establish the monthly water demand by crop type for the total area of the basin. Monthly irrigation needs were then resampled at 10-days sampling rate.

Data ranges from 1977 to 2018 at a 10-day time-step, with a gap between August 1991 and January 1995 and another one from January 2014 to May 2014. Smaller gaps exist in data but never exceed two months, as in April 1985, December 1995, and early 2003 and 2005. Data from October to December 1990 have been set apart due to potential errors in groundwater level measurements, because the data are constant at 116.65 m.a.s.l. during this period. Even though short gaps (one or two time steps) were filled by simple interpolations, the other more important gaps were not filled due to the lack of information about water table variability during these gaps. This is particularly true during periods of extreme levels, such as during the test set, which has been shortened for this reason. Irrigation data are provided on Appendix A Figure A1.

Table 1. Simple statistics on Champagne chalk time series from 1977 to 2018.

Station Name	Measured Variable	Unit	Time Step	Max Value	Min Value	Median	Average
Craie at Vailly (L_{CV})	Level	m.a.s.l.	10 days	134.75	109.75	119.95	120.558
Barbuise at Pouan les Vallées (D_{BP})	Discharge	$m^3.s^{-1}$	10 days	4.50	0.00	0.67	0.836
Seine at Méry-sur-Seine (D_{SM})	Discharge	$m^3.s^{-1}$	10 days	182.2	5.95	25.61	35.71
Grandes-Chapelles (R_{GC})	Rain	mm	10 days	131.2	0.0	15.9	19.87
Troyes-Barberey (R_{TB})	Rain	mm	10 days	86.4	0.0	13.2	17.29
Mailly (R_{MA})	Rain	mm	10 days	138.8	0.0	17.0	21.50
Troyes-Barberey (PET)	Potential Evapo-transpiration	mm	10 days	64.7	0.0	19.0	21.20
Bassin (I)	Irrigation	$m^3.ha^{-1}.month^{-1}$	month	833.9	0.6	176.0	280.1

Cross correlations between inputs and output were calculated to better understand the behavior of the basin. These provide information on input-output relationships by showing the response time (mean delay between rain peak and discharge/groundwater level peak, in number of time steps) and the memory effect [32]. They help define the reasonable lead-time that can be reached. Correlation analysis is synthetized in Table 2.

Table 2. Correlation analysis. Diagonal shows the memory effect (in number of time-steps) when simple correlation is calculated (orange). When cross-correlation is calculated, blue cells show memory effect and green cells show response time. NC means that correlation score is always under 0.2, showing a very weak correlation, leading to possible misinterpretations of the memory effect.

	L_{CV}	ΔL_{CV}	D_{BP}	D_{SM}	R_{GC}	R_{TB}	R_{MA}	PET	I
L_{CV}	17	6	2	5	21	22	15	12	16
ΔL_{CV}	15	4	−4	0	1	1	1	4	8
D_{BP}	17	1	11	3	2	6	7	9	12
D_{SM}	19	4	15	5	1	1	1	3	7
R_{GC}	NC	2	NC	3	0	0	0	0	27
R_{TB}	NC	2	NC	3	1	1	0	0	27
R_{MA}	NC	3	NC	3	0	1	0	0	33
PET	19	11	16	9	NC	NC	NC	8	4
I	22	14	19	12	NC	NC	NC	11	7

Regarding the first line of Table 2 showing the response times of all variables over water levels at Craie at Vailly (L_{CV}) piezometer, it appears that the shortest response time is two time-steps for the discharge at Barbuise at Pouan les Vallées (D_{BP}). This indicates that, statistically, the discharges at Barbuise at Pouan les Vallées have a greater influence on the groundwater at Craie at Vailly after two time-steps delay. And that this response time is the shorter. This confirms the quick interaction between surface water and groundwater. Regarding the impact of surface water on both the water quality and the groundwater level, the two time-step lead-time was thus chosen. In this way, a lead-time of 20 days (two time-step) is considered as a good compromise between model accuracy and end users' needs. A shorter lead-time would reduce the interest of the forecast for the end users, while a longer lead-time would require the availability of the Barbuise at Pouan les Vallées discharge forecast. Thus, this lead time ensures that available inputs explain the output.

2.3. Quality Criteria

It is important to use impartial criteria to evaluate the quality of result. This study investigates two modeling goals: groundwater level prediction and uncertainty quantification. As a result, two kinds of criteria will be used, as presented below.

2.3.1. Quality of Fitting and Prediction

The Pearson's correlation coefficient allows quantifying of the linear relation between two variables. It varies between −1 to +1, and is the covariance divided by the product of the standard deviations of the two variables (5).

$$r^2 = \frac{cov(u,y)}{\sigma_u \sigma_y} \tag{5}$$

where $cov(u,y)$ is the covariance between variables u and y, σ_u and σ_y, respectively, are their standard deviation.

- The persistency criterion

This criterion was proposed by [33] for prediction (6). It must be close to 1. A 0 value represents the score of the naive forecasting (prediction value = the actual value), and a negative value means that the forecasting is even worse than the naive forecasting.

$$C_P = 1 - \frac{\sum_k^n (y_{k+h} - \hat{y}_{k+h})^2}{\sum_k^n (y_k - y_{k+h})^2} \tag{6}$$

where y_k is the measured value at the discrete time *k*, $\hat{y}_{k+h}$ is the simulated value at the discrete time *k*, y_{k+h} is the observed value at the discrete time *k+h*, *h*, the lead time, and *n* the number of samples of the considered dataset.

2.3.2. Uncertainties Quantification

The following criteria assume that several models are available for prediction, and that a prediction interval can therefore be defined between the largest and smallest of the forecast values, at each time step.

- Prediction Interval Coverage Probability

  This criterion expresses the empirical probability that the prediction interval contains the measured value (7). It represents a kind of accuracy of the predicted value. In the (7) equation, *f*(.) is the function of belonging to the prediction interval [34,35].

$$C_{PICP} = \frac{1}{n}\sum_{k}^{n} f(y_k, \hat{y}_k^{min}, \hat{y}_k^{max}) \tag{7}$$

with the same notations as before, and $f(y_k) = 1$ if $y_k \in [\hat{y}_k^{min}; \hat{y}_k^{max}]$, else $f(y_k) = 0$

- Mean Prediction Interval

  The Mean Prediction Interval, C_{MPI}, is the average of all the results set of the interval of prediction calculated at each time-step. It quantifies the mean scattering of the prediction [34], following (8).

$$C_{MPI} = \frac{1}{n}\sum_{k=1}^{n} (\hat{y}_k^{max} - \hat{y}_k^{min}) \tag{8}$$

with the same notations as before, y_k the measured value at the discrete time *k* and $\hat{y}_k^{max}$ and $\hat{y}_k^{min}$ the upper and lower bounds of the forecast interval.

- Prediction Confidence Criterion
  The Prediction Confidence Criterion, C_{PC}, is a ratio quantifying the performance of a predictor for providing a prediction having the highest empirical probability of lying within the smaller prediction interval (9). It is simply defined by the ratio between the two previous criteria [13].

$$C_{PC} = \frac{C_{PICP}}{C_{MPI}} = \frac{\frac{1}{n}\sum_{k}^{n} f(y_k, \hat{y}_k^{min}, \hat{y}_k^{max})}{\frac{1}{n}\sum_{k=1}^{n} (\hat{y}_k^{max} - \hat{y}_k^{min})} \tag{9}$$

  with the same notations as before, y_k is the measured value at the discrete time *k* and $\hat{y}_k^{max}$ and $\hat{y}_k^{min}$ the upper and lower bounds of the forecast interval at discrete time *k*.

2.4. Uncertainties Linked to the Initialization of the Parameters and to the Spatial Variability of the Rains

2.4.1. Variability Due to the Initialization of Parameters.

As presented in Section 2.1.4, the implementation of an ensemble model makes possible a reduction in the variability of the outputs caused by the random initialization of the model parameters during the training step. However, the number of members in the ensemble model needs to be determined in order to sufficiently reduce this variability. The purpose is to obtain the smallest prediction interval, measured by the MPI criterion (Mean Prediction Interval) [34], whereas this interval includes a maximum of the observed

values, measured by PICP criterion (Prediction Interval Coverage Probability) on the subset of interest, considered as a whole [34,35]. Therefore, the optimal number of random initializations (members of the ensemble) can be determined thanks to the calculation of the C_{PC} criterion (Prediction Confidence Criterion) [13] defined in Section 2.3.2, that synthetizes both criteria.

2.4.2. Spatial Rainfall Variability

In order to approach the uncertainty caused by the spatial variability of rainfalls, usually poorly described by a small number of rain gauges, once the training is over, we propose to run permutations and substitutions of rain gauges applied to the model inputs. This allows generation of another ensemble model of possible forecasts obtained in the test. As three rain gauges are available for the studied basin, 27 combinations, shown in Figure 3, composed the ensemble model. This new kind of ensemble model, whose members differ by the combination of used rain gauges, is denoted as ensemble-RG. All the permutations made give a range of outputs that can be considered as a prediction interval related to the spatial variability of rain.

Combination	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Raingauge 1	GC	GC	GC	GC	GC	GC	GC	GC	GC	TB	TB	TB	TB	TB	TB	TB	TB	TB	MA	MA	MA	MA	MA	MA	MA	MA	MA
Raingauge 2	GC	GC	GC	TB	TB	TB	MA	MA	MA	GC	GC	GC	TB	TB	TB	MA	MA	MA	GC	GC	GC	TB	TB	TB	MA	MA	MA
Raingauge 3	GC	TB	MA	GC	TB	MA	GC	TB	MA	GC	TB	MA	GC	TB	MA	GC	TB	MA	GC	TB	MA	GC	TB	MA	GC	TB	MA

Actual measurements GC : Grandes Chapelles raingauge TB : Troyes-Barberey raingauge MA : Mailly raingauge

Figure 3. Combinations of rainfall input data with three rain gauges.

2.5. Estimation of Empirical Confidence Intervals Using Probability Density Functions

2.5.1. Method

The purpose of this section is to present the method used to estimate a confidence interval for predictions achieved by the model. The original process consists in:

- Establishing the frequencies of appearance of the water level classes histogram; this is then considered as an empirical probability density function (pdf) of the data;
- Fitting a theoretical well-known pdf, for example the normal one, to the empirical pdf by adjusting its parameters. If necessary, thanks to the Expectation-Maximization algorithm (EM) [36,37], the theoretical pdf can be a composition of several pdfs of the same type, each one having different parameters; this composite pdf is called the target pdf. The algorithm provides the constituent parameters of the theoretical elementary theoretical pdfs as well as the weights that enable them to be assembled to fit the target pdf;
- Starting from target pdf, determining a probability of occurrence of the measured value inside the predicted interval for each class;
- For a given confidence index (for example 95%), and for each class, supposing the data verify the constraints of a normal law and establishing a model of "correctness" using the erf (error function). This provides the estimated error associated to each class;
- Finally, drawing the possible errors on the water chart.

The Expectation-Maximization algorithm follows two different steps: the expectation step and the maximization step. The expectation step consists in defining an expected value for log-likelihood parameters of the target pdf. The maximization step consists in obtaining parameters that maximize the expected value, using an iterative process [36].

The Expectation-Maximization process is applied four times to four categories of samples: (i) samples with a negative slope ($y_i^k - y_i^{k+1} > 0$), (ii) samples with a positive slope ($y_i^{k+1} - y_i^k > 0$), (iii) samples included inside the MPI and (iv) samples not included inside the MPI. Discriminating between negative and positive slopes is useful in considering the current conditions in the system: decreasing water levels are the sign of a

draining of groundwater whereas increasing water levels are the sign of a refill of groundwater stock. This corresponds to two different physical behaviors that should not be mixed to capture the guiding factors of the system.

2.5.2. Chosen Probability Density Functions

Nine probability density functions were chosen (Table 3) (Equations (10)–(18)).

Table 3. Equations of nine probability density functions.

Name of *pdf* Law	Formula	Eq.	References
Normal	$\mathcal{N}(\bar{x},\sigma^2)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}$	(10)	[38,39]
Gumbel	$\mathcal{G}um(\bar{x},\beta)=\frac{e^{-\frac{(x-\bar{x})}{\beta}}e^{-e^{-\frac{(x-\bar{x})}{\beta}}}}{\beta}$	(11)	[40]
Laplace	$\mathcal{L}ap(\bar{x},b)=\frac{1}{2b}e^{-\frac{(\lvert x-\bar{x}\rvert)}{b}}$	(12)	[41]
Raised Cosine	$\mathcal{C}osr(\bar{x},s)=\frac{1}{2s}\,(1+\cos\left(\frac{(x-\bar{x})}{s}\,\pi\right))$	(13)	[42,43]
Cauchy	$\mathcal{C}au(x_0,a)=\frac{1}{\pi a\,(1+\left(\frac{x\,-\,x_0}{a}\right)^2)}$	(14)	[44,45]
Logistic	$\mathcal{L}ogist(\bar{x},s)=\frac{e^{-\frac{(x-\bar{x})}{s}}}{s\left(1+e^{-\frac{(x-\bar{x})}{s}}\right)^2}$	(15)	[46]
Slash	$\mathcal{S}\ell a(x)=\frac{\varphi(0)-\varphi(x)}{x^2}$	(16)	[47]
Bhattacharjee	$\mathcal{B}hat(\bar{x},\sigma_x,a)=\frac{1}{2a}\left(\Phi\left(\frac{x-\mu+a}{\sigma}\right)-\Phi\left(\frac{x-\mu-a}{\sigma}\right)\right)$	(17)	[48]
Huber	$\mathcal{H}ub(z)=\frac{1}{2\sqrt{2\pi}\left(\Phi(z)\,-\,\frac{\phi(z)}{z\,-\,\frac{1}{2}}\right)}\,e^{-\rho_z(x)}$	(18)	[49,50]

where x is the variable, $\bar{x}$ its average, x_0, its median, σ its standard deviation, σ^2 its variance, a, b, β and s scale parameters, φ the normalized normal distribution, ϕ the normal law, Φ the cumulative normal law, z the degree of robustness and ϱ_z the Huber loss. The Huber loss depends on the degree of robustness and can be written following Equation (19) [49].

$$\rho_z(x)=\begin{cases}\frac{1}{2}x^2 & ,|x|\le \mathrm{z}\\ z|x|\,-\frac{1}{2}z^2 & ,|x|>\mathrm{z}\end{cases}' \tag{19}$$

3. Model Design

3.1. Definition of Subsets for Training Testing, Stop and Cross-Validation

As presented previously in Section 2.1.4, it is necessary to define the training, stop and test sets. For this purpose, it is usual to split the database into three subsets. To be consistent with the needs of the end users, the test set is chosen as the driest period of the database: the 1988–1990 period. It will be used to assess the quality of the model generalization. Requesting that the model will be able to predict correctly the extreme dry period of the database is a strong requirement to ensure generalization capacity. The 2011–2013 period is used for the stop set (12th subset). The remainder is the training set. Cross validation is used to select input variables and complexity, as presented in Section 2.1.5. To run cross-validation, we also have to define the cross-validation subsets inside the training set. Each one must contain a sufficient amount of data. A length of three years (108 samples based on the 10-day time step) is thus chosen for each cross-validation subset. Thus, 13 cross-validation subsets are defined as shown in Table 4. They are used in validation of each one its turn, and for each turn all other cross-validation subsets are used

in training (T on Table 4). Therefore, the training set is divided in two kinds of subset: 12 training subsets and a cross-validation subset. At the end of the 13 sequences of training, 13 cross-validation scores are calculated and the resulting global cross-validation score is calculated as follows (20):

$$S_{CV} = \frac{1}{D} Median(C_{P_q}) \quad (20)$$

with D=13 the number of cross validation subsets, C_P the score of persistency, and q the number of the considered subset.

Table 4. Split subsets; T is for training, V for validation, S for stop, T_e for test; C_{P_q} is for the score calculated on the q subset.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Scores
V	T	T	T	T	T	T	T	T	T	T	S	T	T	T_e	C_{P_1}
T	V	T	T	T	T	T	T	T	T	T	S	T	T	T_e	C_{P_2}
T	T	V	T	T	T	T	T	T	T	T	S	T	T	T_e	C_{P_3}
T	T	T	V	T	T	T	T	T	T	T	S	T	T	T_e	C_{P_4}
															…
T	T	T	T	T	T	T	T	T	T	V	S	T	T	T_e	$C_{P_{11}}$
T	T	T	T	T	T	T	T	T	T	T	S	V	T	T_e	$C_{P_{13}}$
T	T	T	T	T	T	T	T	T	T	T	S	T	V	T_e	$C_{P_{14}}$
														Median	S_{CV}

3.2. Choice of the Model and Complexity Selection

A single hidden layer model is used because of its lower complexity whereas its performance is sufficient for the modelling objectives.

According to the recommendations made in Section 2.1.2, the chosen predictor is the feed-forward model because the measurement of water level is accurate, which is not the case for the rainfall-field estimation. Indeed, not only are the rain gauges inaccurate, but the rain also has a spatial variability not sufficiently described by the three available rain gauges. Therefore, it can be assumed that the noise affecting the output of the process (the water table) is lower than the noise affecting the inputs (the rain).

Hyperparameters window-width ranges for rains, irrigation, and shallow water discharge, were chosen using correlation analysis as shown in Table 5, suggested by [10].

For training, the Levenberg-Marquardt algorithm, which is a second order learning method [51,52], has been used with 100 epochs for each experience.

Several architectures were tried with different complexities, and for each one the cross-validation score was calculated in order to select the best. Table 3 synthetizes the investigated architectures and the selected model for the two time-steps' lead-time (justified in Section 2.2), in order to simulate groundwater levels at a middle term in a drought context. The selected model is also shown in Figure 4. Table 5 presents the selected hyperparameters during the design stage.

Table 5. Selected architecture using the design procedure synthetized in 2.1.5.

Model Element	Selected Hyperparameters		Tested Range Values	
Order	r (LCV)	3	(3–6)	(8–14)
Exogenous input window-widths	n_1 (I)	8	(7–10)	(9–12)
	n_2 (PET)	12	(9–12)	
	n_3 (DSM)	5	(2–5)	(2–5)
	n_4 (DBP)	5	(2–5)	(7–12)
	n_5 (RGC)	2	(1–4)	(7–12)
	n_6 (RTB)	2	(1–4)	(7–12)
	n_7 (RMA)	3	(1–4)	(7–12)
Number of hidden neurons	N	3	(2–10)	(2–10)

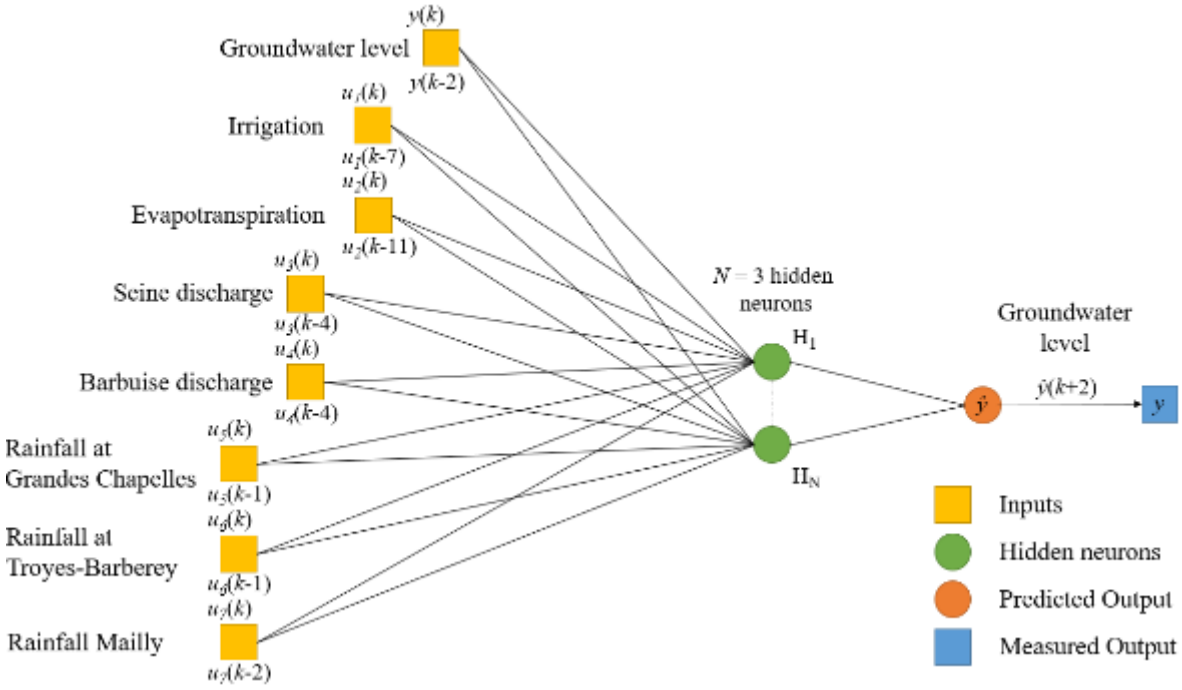


Figure 4. Selected architecture.

4. Results

4.1. Optimal Number of Members in Ensemble Models

Once the architecture is selected as presented in 2.1.5 and in 3.1.2, ensemble forecasts are calculated between 3 to 120 members in each (respectively 3, 5, 10, 20, 30, 40, 50, 60, 80, 100 and 120 forecasts). CPC, Prediction Confidence Criterion, is calculated for each ensemble, allowing definition of the optimal number of members, i.e., the number of members whose parameters are randomly initialized (X in equation 4).

Figure 5 presents the evolution of the CPC versus the number of members in the ensemble models. Schematically, the curve can be approximated by two straight lines whose intersection is at around 40 members. The first line decreases when the number of members in the ensemble increases, corresponding to a stage where the MPI increases. The second line corresponds to a plateau that indicates the stability of the two criteria that make up the CPC. The intersection of two lines corresponds to the minimal number of initializations for which the gain of ensemble starts to become stationary. We thus propose this value (40 members in Figure 5) as the number X of members. Although the CPC could possibly be enhanced by using more members, the cost-benefit ratio (especially regarding calculation time) pleads in favor of this choice.

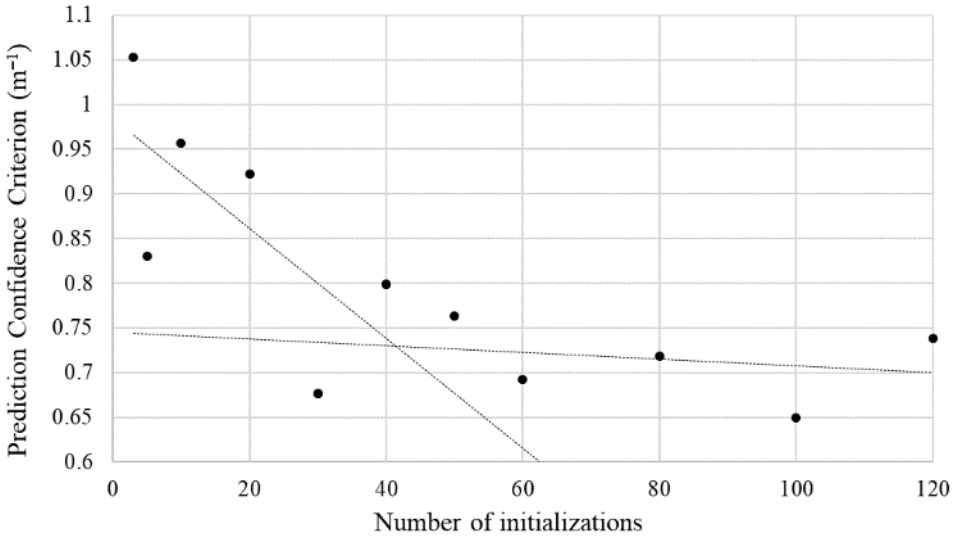


Figure 5. C_{PC} cross-validation score of ensemble models as a function of the number of members (X).

4.2. Prediction Results

The cross-validation persistency score for the architecture reaches C_p = 0.65 and the test score is C_p = 0.40. The performances are thus lower on the test set than on the other sets during cross validation. This is consistent with the choice of the test set, which corresponds to the drier period of the database. Nevertheless, the quality of the forecast presented in Figure 6, made for the year with the driest summer in the entire database, shows that: (1) the model is capable of generalizing to periods of extreme behavior, (2) confirms the interest in being able to visualize uncertainty, so that the manager can analyze the most uncertain parts of the limnigram. As a reminder, the three last months of 1990 were not considered due to possible errors in piezometric levels.

Figure 6 also shows the prediction interval for the test event. In this case, P_{ICP} = 0.39, M_{PI} = 0.62 m, and C_{PC} = 0.63 m^{-1}. It appears that the prediction is fairly close to the measurement except for the early spring of 1990, for which the forecast level is not low enough. On the other hand, the grey band showing the uncertainty is very thin and does not contain enough observed values (P_{ICP} = 0.39) to be able to inspire confidence in the end users.

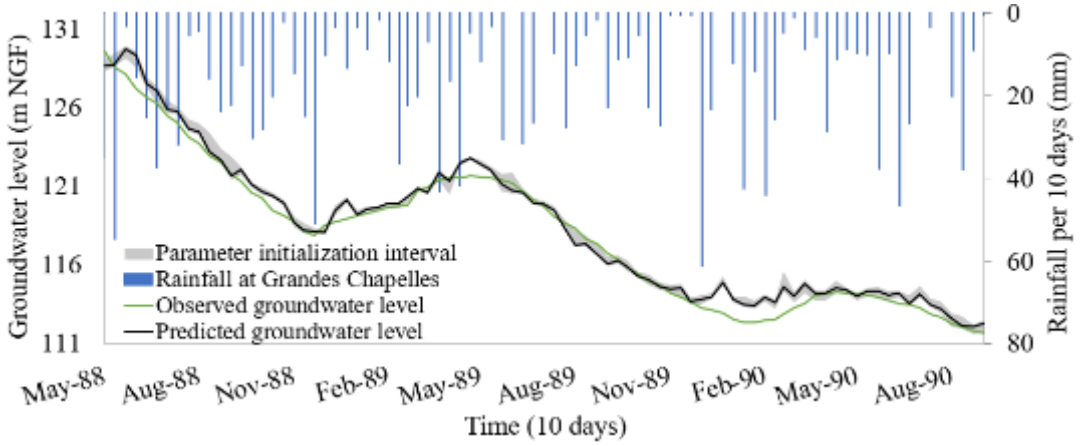


Figure 6. Prediction of groundwater levels at 20 days lead-time on the test set. The three last months of 1990 were not taken into account due to possible errors in piezometric levels. C_p = 0.40; C_{PICP} = 0.39; C_{MPI} = 0.62 m.

4.3. Representation of Uncertainties Caused by the Initialization Parameters

4.3.1. Theoretical Composite pdf for Four Distributions

The empirical *pdf*s of training and stop sets are represented in one *pdf* using classes of observed water levels of 20-cm-wide. This interval is the result of a compromise between the class width and the number of samples in each class.

Remember that we have chosen to represent four distributions: two for the sign of the slope of the groundwater evolution, and two for the belonging or non-belonging of the observed sample to the prediction interval. For each distribution, a theoretical composite *pdf* is built as presented in Section 2.5.2.

To illustrate the procedure, let us consider the specific distribution of water levels with a decreasing slope, using a Normal law as theoretical law. The obtained theoretical composite *pdf* is presented in (Figure 7).

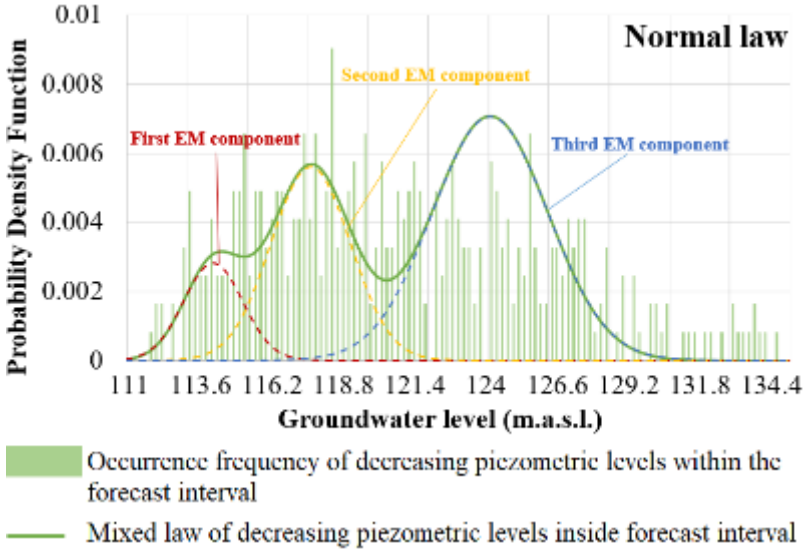


Figure 7. Theoretical composite *pdf* of piezometric levels, obtained using Esperance Maximization (EM) algorithm using a Normal law and applied to observed values having a decreasing slope and being inside the prediction interval. Each elementary normal law is represented with a different color and denoted as "EM component."

One can note that the water levels from 115 to 125 m.a.s.l show the highest frequencies whereas above 132 m.a.s.l. or under 115 m.a.s.l. observed groundwater values are underrepresented.

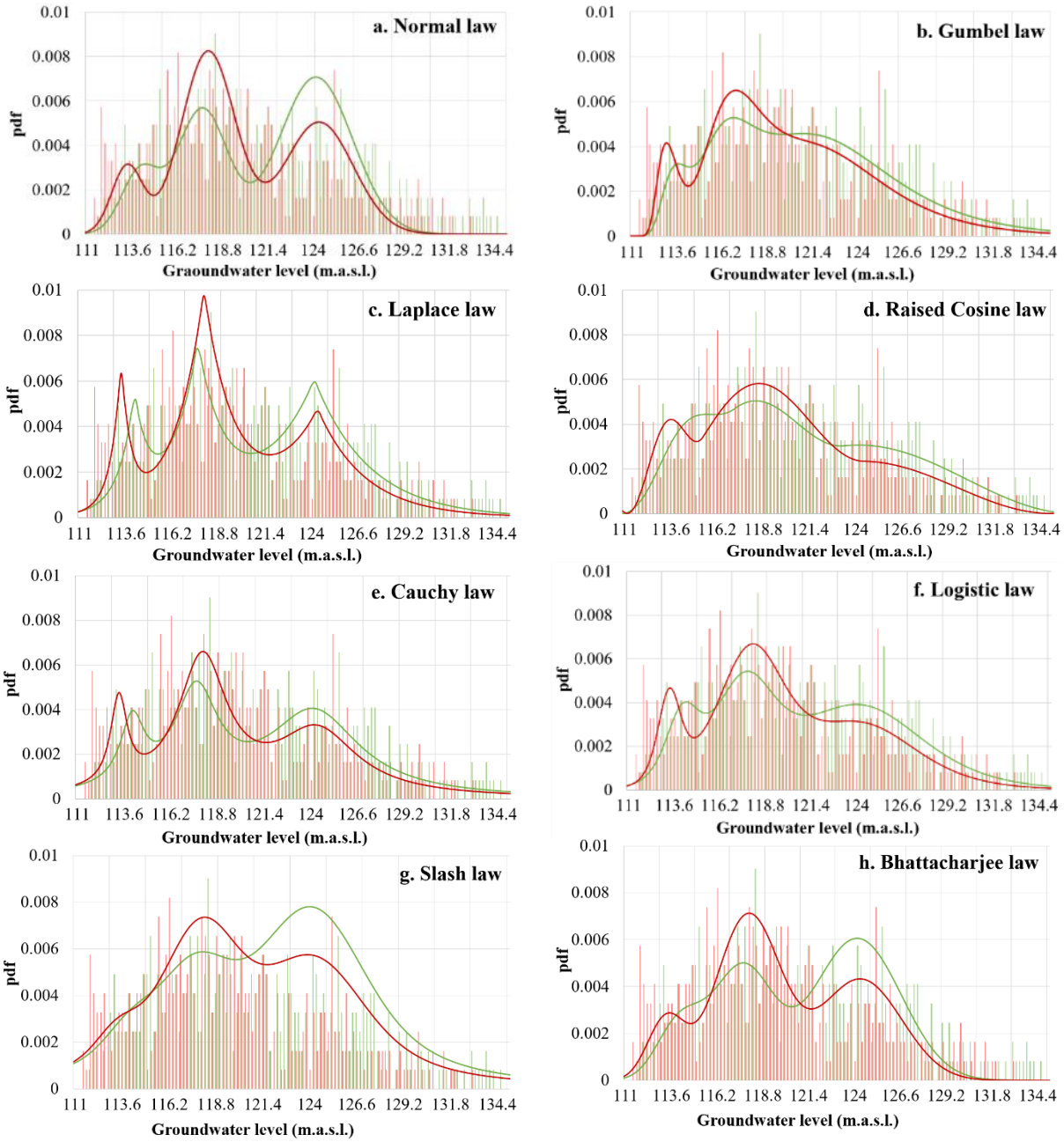
According to the Figure 7, the Esperance Maximization (EM) algorithm provides the parameters for each elementary distribution. For the Normal law, these parameters are the mean, the variance and the lambda, which is the maximum amplitude of each elementary distribution. The theoretical mixed *pdf*, fitted by measured groundwater level distribution classes, is obtained by the sum of the three components, as shown in Figure 7. This process is repeated for each of the three other groundwater levels configurations, and for the nine different studied laws (presented in Table 3).

Let us now examine which of the theoretical laws presented in Table 3 provides the best C_{PICP} on the Train+Stop dataset. To this end Figure 8 shows the representation, explained in Figure 7, for each one of the theoretical laws, regarding the two distributions of values inside the prediction interval (green) or outside this interval (red). Table 6 shows the correlations between the empirical distribution and the theoretical laws, for the measured groundwater levels inside the prediction interval (r_{in}^2) and outside the prediction interval (r_{out}^2). Best correlations are shown in green and worst in red. It

appears in Table 6 that the best adjustment is achieved by the Raised Cosine theoretical law.

Table 6. Pearson's correlation coefficients between the distributions with negative slope of groundwater evolution, and the theoretical composite *pdf*; r_{in}^2 applies to measurements inside prediction interval, r_{out}^2 applies to measurements outside the prediction interval.

Law	Normal	Gumbel	Laplace	Raised Cosine	Cauchy	Logistic	Slash	Bhatta-Charjee	Huber
r_{in}^2	0.62	0.74	0.66	0.76	0.69	0.75	0.64	0.64	0.50
r_{out}^2	0.68	0.74	0.67	0.77	0.70	0.75	0.72	0.75	0.65



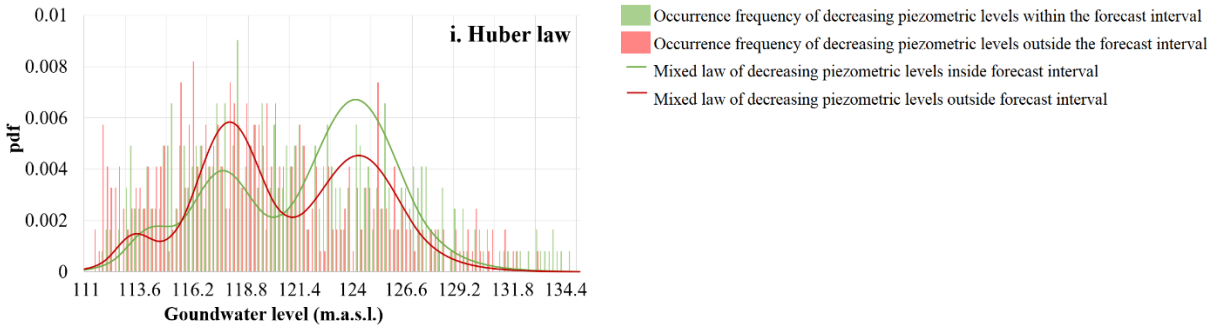


Figure 8. Composite *pdf* for decreasing measured groundwater distributions with (**a**) Normal law, (**b**) Gumbel law, (**c**) Laplace law, (**d**) Raised Cosine law, (**e**) Cauchy law, (**f**) Logistic law, (**g**) Slash law, (**h**) Bhattacharjee law and (**i**) Huber law.

For measured water levels having negative slopes (Figure 8), C_{PICP} = 0.51, meaning that the probability of the interval of prediction containing the observed value is similar to the probability of it not containing the observed value, whatever the groundwater level and theoretical composite *pdf* law selected. We can also notice that Normal, Bhattacharjee and Huber laws have the same kind of pattern whereas Logistic, Gumbel and Raised Cosine laws have similar shapes. This can be explained by the fact that Slash, Bhattacharjee and Huber are derived from the Normal law. Logistic, Gumbel and Raised Cosine laws seem to fit well with observed groundwater level distribution inside the prediction interval, having a Pearson's correlation coefficient over 0.74 for measured water levels inside the prediction interval, whereas other laws provide correlations ranging from 0.50 to 0.69 (Table 6 and Figure 8).

Raised Cosine, Gumbel, Logistic and Bhattacharjee laws have the best linear correlation with groundwater level distribution when they are outside the prediction interval, with a correlation from 0.74 to 0.77.

For measured water levels having positive slopes (Figure 9), C_{PICP} = 0.24, meaning that the observed groundwater levels outside the prediction interval are more numerous than groundwater levels inside the prediction interval. Pearson's correlation coefficients between the distribution of groundwater levels and the composite theoretical laws are shown in Table 7.

Table 7. Pearson's correlation coefficients between the distributions with positive slope of groundwater evolution and the theoretical composite *pdf*; r_{in}^2 applies to measurements inside prediction interval, r_{out}^2 applies to measurements outside the prediction interval.

Law	Normal	Gumbel	Laplace	Raised Cosine	Cauchy	Logistic	Slash	Bhatta-Charjee	Huber
r_{in}^2	0.43	0.41	0.47	0.44	0.45	0.44	0.43	0.43	0.42
r_{out}^2	0.52	0.54	0.54	0.66	0.55	0.61	0.63	0.57	0.52

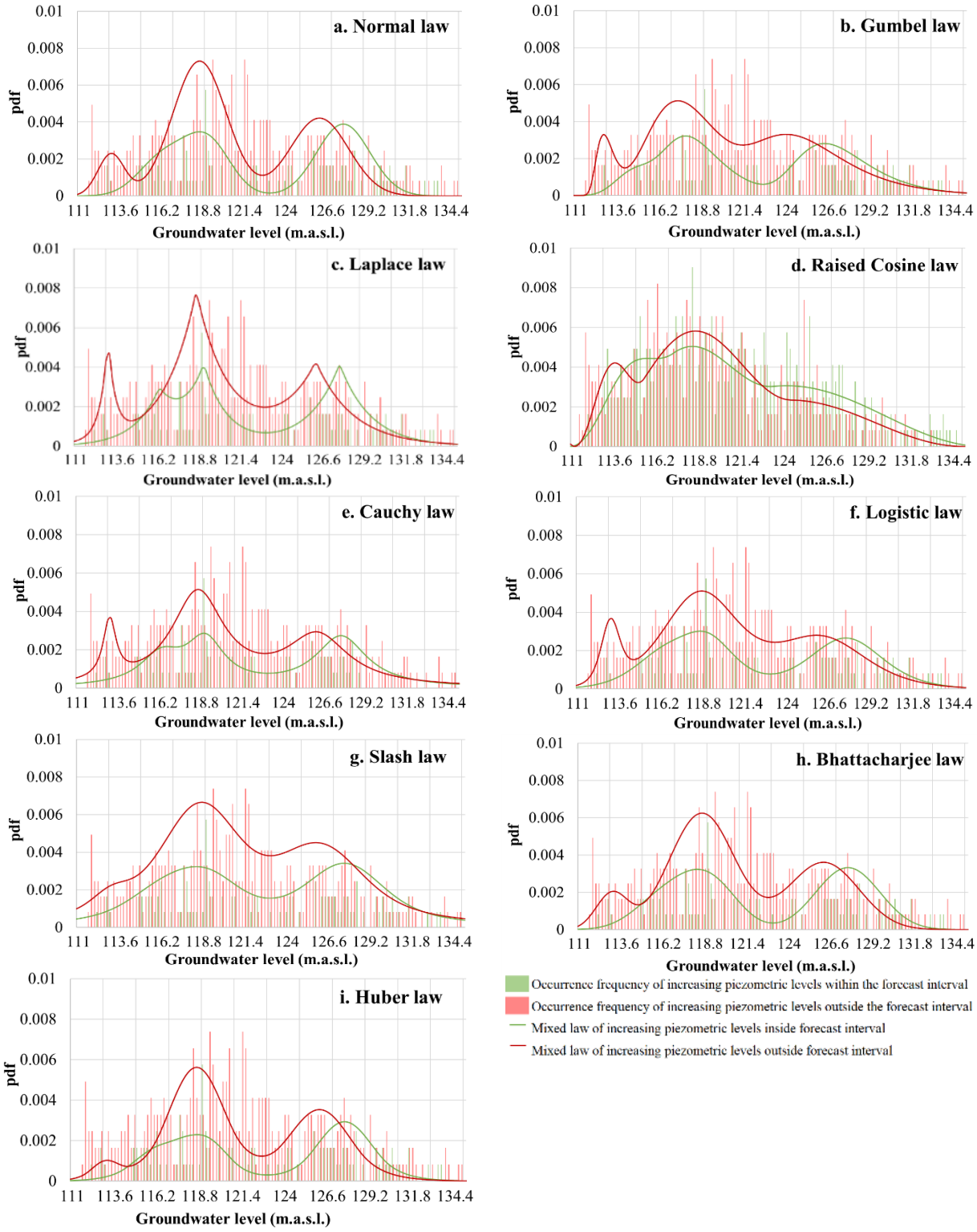


Figure 9. Mixed *pdf* for increasing groundwater distributions. with (**a**) Normal law, (**b**) Gumbel law, (**c**) Laplace law, (**d**) Raised Cosine law, (**e**) Cauchy law, (**f**) Logistic law, (**g**) Slash law, (**h**) Bhattacharjee law and (**i**) Huber law.

As seen previously, Normal, Slash, Bhattacharjee and Huber laws have a similar theoretical composite *pdf*. Curves representing observed values outside the prediction interval present a large bell around 119 m.a.s.l. for all of these four laws. This produces a lower correlation than for other laws (except for Slash law, which is smoother), with r^2 varying between 0.52 and 0.57. Slash, Raised Cosine and Logistic laws seem to provide the best correlated composite *pdf*, with a correlation between 0.61 and 0.66. On the other

hand, the composite *pdf* representing observed values inside the prediction interval has two flared "peaks" at 117 m.a.s.l. and 127 m.a.s.l. for the nine laws. However, correlations are low due to the small frequencies of increasing groundwater levels inside prediction interval. Laplace and Cauchy laws appear to be the laws with the best fit, with a correlation above 0.45. Raised Cosine and Logistic correlations reach only 0.44.

This illustrates the model's difficulties obtaining relevant prediction intervals when the slope of groundwater levels is positive.

4.3.2. Error Margins

The last step consists in calculating the models of correctness. This consists in calculating the probability that the measured value lies within the prediction interval, and in adding errors around this probability for each class, following the procedure presented in 2.5.1. This is done for both increasing and decreasing measured groundwater level distributions. For each distribution, the probability that the measured value belongs to the interval of prediction is calculated by numerical integration, and the possible associated error is deduced using the number of samples inside the considered class using the *erf* (error function, the inverse of the Normal law), supposing that the data follows a normal distribution, and for a predefined confidence index. Then, for each class, a model of "correctness" is established and provides a confidence interval around the probability (shadow band around the probability in Figures 10 and 11). To provide the charts in Figures 10 and 11, a predefined confidence index of 0.95 is chosen, and several specific considerations are adopted:

- It is supposed that the distribution of samples inside a class follows a Normal Distribution,
- When a class contains no sample, for example, the class around 135 m.a.s.l., the error is maximum and is divided into two parts: 50% above 50% underneath the probability.
- When a class contains very few samples (less than three), this class is not considered for r_C^2 and M_E calculations.

Considering for example the distributions of increasing measured groundwater levels (Figure 10), all models of correctness show a poor correlation with the C_{PICP} (each cross is a C_{PICP} calculated thanks to the ensemble model), always having a Pearson's correlation under 0.3. This is consistent with the high dispersion of C_{PICP}. However, the percentage of these C_{PICP} included inside the calculated error margin seems to be a better indicator of the quality of the model of correctness. In this case, six laws have more than 75% of C_{PICP} inside the error margin. Pearson's correlation coefficients and the error margin indicator for Cauchy and Slash law's models of correctness are the highest, with, respectively, 0.25 and 75% and 0.22 and 80.3% values (Table 8).

Except for the Cauchy and Slash models of correctness, the probability increases for highest groundwater values (over than 130 m.a.s.l). For all models of correctness, the probability of a correct prediction varies from 0.2 to 0.4. The low probability of correctness still shows the difficulty of the model in forecasting increasing groundwater levels, with a low prediction interval.

Table 8. Error margin (E_M), and Pearson's correlation coefficients (r_C^2) between the model of correctness and the empirical C_{PICP} calculated for each 20-cm groundwater levels having a positive slope.

Law	Normal	Gumbel	Laplace	Raised Cosine	Cauchy	Logistic	Slash	Bhatta-charjee	Huber
r_C^2	0.15	0.04	0.24	0.19	0.25	0.24	0.22	0.16	0.20
E_M	76.3%	73.7%	72.4%	76.3%	75.0%	73.7%	80.3%	76.3%	77.6%

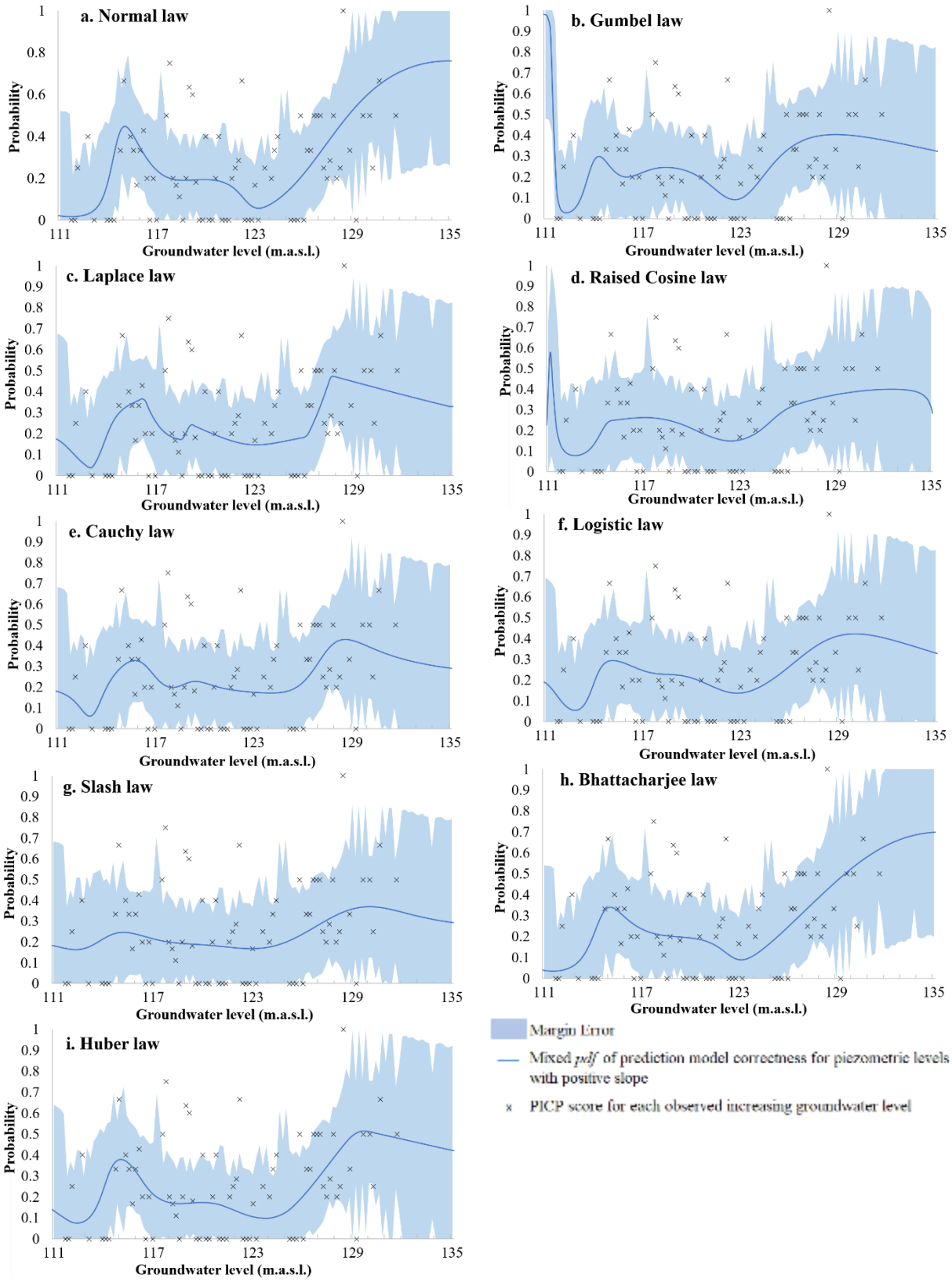


Figure 10. Models of correctness and Error margin (M_E), calculated from composite *pdf* of increasing levels and using a confidence index of 95% for (**a**) Normal law, (**b**) Gumbel law, (**c**) Laplace law, (**d**) Raised Cosine law, (**e**) Cauchy law, (**f**) Logistic law, (**g**) Slash law, (**h**) Bhattacharjee law and (**i**) Huber law.

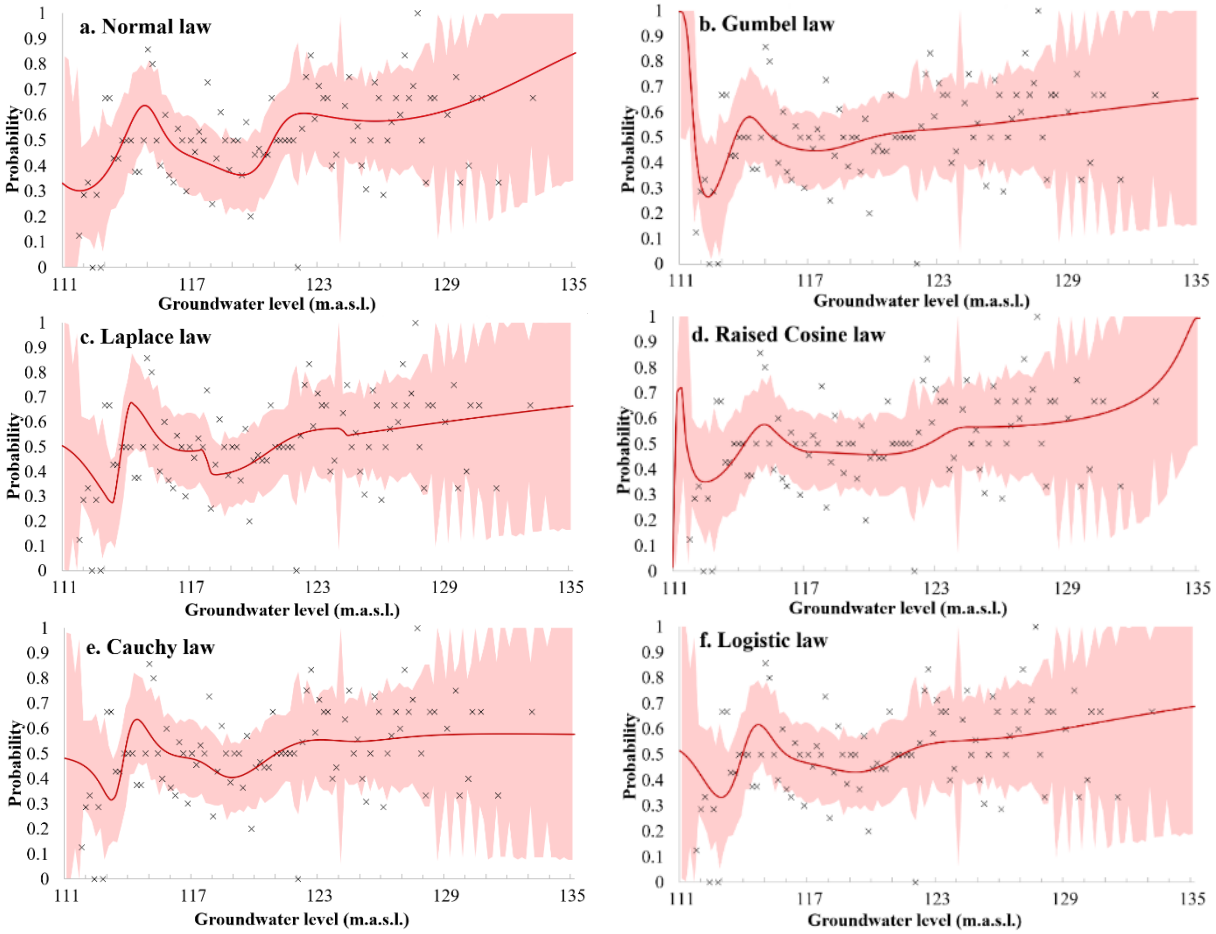
The same kind of models of correctness are made for decreasing groundwater levels, shown in Figure 11.

Correlations between the C_{PICP} and the model of correctness are still low, with an average value around 0.2 for all laws. The highest correlation comes from the model of correctness provided by the Slash mixed *pdf*. However, the crosses representing the C_{PICP} inside the error margins, reaching more than 79% for Logistic law (Table 9), are slightly higher than the ones obtained for decreasing groundwater levels (Table 8). Figure 11 shows that the models of correctness of the nine laws have a similar shape, with a stagnation of probability for observed piezometric levels above 125 m.a.s.l. The probability of correctness, for each law and each groundwater level, is above or equal to 0.5.

However, these models of correctness only consider the uncertainty linked to the parameter's initialization before training. The uncertainty linked to the spatial representativeness of rain measurements is the next step, in order to consider and draw the two major origins of uncertainty.

Table 9. Error margin (M_E), and Pearson's correlation coefficients (r_C^2) between the model of correctness and the empirical C_{PICP} calculated for each class of 20-cm groundwater levels having a negative slope.

Law	Normal	Gumbel	Laplace	Raised Cosine	Cauchy	Logistic	Slash	Bhatta-charjee	Huber
r_C^2	0.02	0.21	0.16	0.21	0.23	0.23	0.30	0.28	0.27
M_E	74.4%	70.3%	70.3%	72.5%	70.3%	79.1%	71.4%	73.6%	70.3%



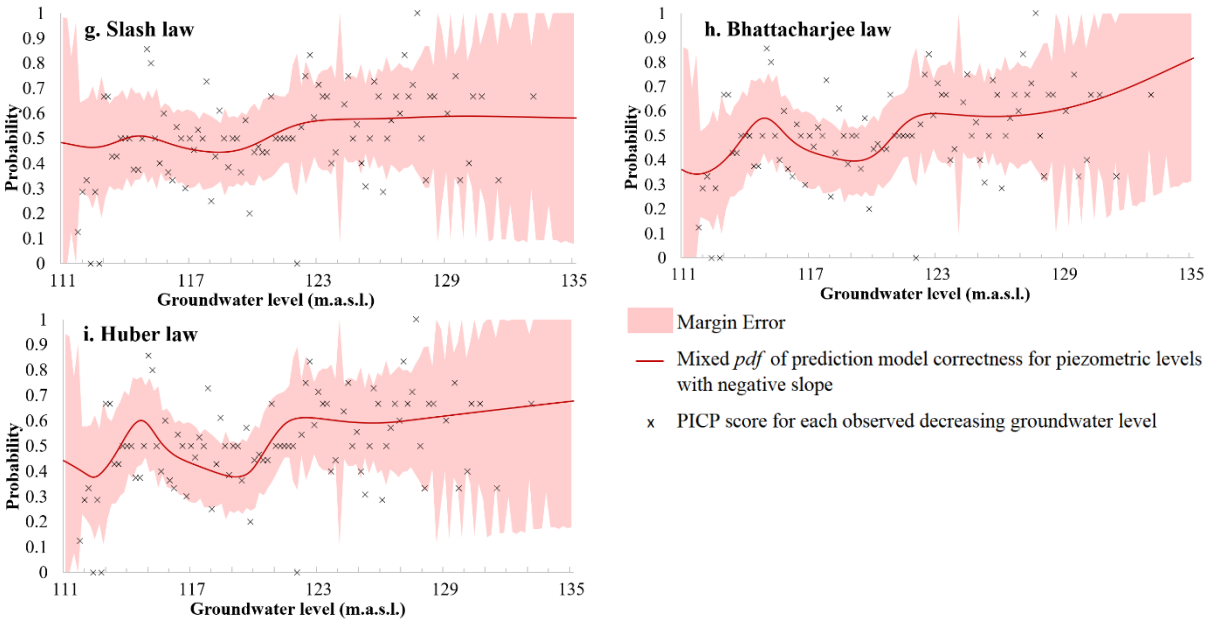


Figure 11. Models of correctness and Error margin (M_E), calculated from composite *pdf* of decreasing levels and using a confidence index of 95% for (**a**) Normal law, (**b**) Gumbel law, (**c**) Laplace law, (**d**) Raised Cosine law, (**e**) Cauchy law, (**f**) Logistic law, (**g**) Slash law, (**h**) Bhattacharjee law and (**i**) Huber law.

4.4. Determination of Spatial Distribution of Rainfall Uncertainty

In order to assess the uncertainty linked to the spatial representativeness of rain measurements, we propose the operation of permutations and substitutions of data between rain gauge inputs. Doing this, following Figure 3, we obtain 27 different datasets. As the method previously applied requires a subset devoted to its assessment, independent of training, test and stop sets, we reused the cross-validation process in order to estimate the uncertainty in the 13 subsets used in cross-validation, which will then be used in cross-uncertainty assessment. This method also has the advantage of producing a sufficient number of values in a validation situation, since 27 permutations are performed on the rain gauges for each one of the 13 cross-uncertainty assessment sets. Added to the variability due to the initialization of the parameters, which is also considered, this method generates an ensemble model integrating the two types of variability: that due to rainfall and that due to the initialization of the parameters (27*40 members for each validation set).

Based on this ensemble, the prediction values are calculated: the median, higher and lower values, allowing the definition of a prediction interval for each time step and for each configuration. Thus, we find that the C_{PICP} equals 0.49 for the test set, shown in Figure 12. This was 0.39, considering only variability due to the parameter's initialization. Regarding the prediction interval, it logically became wider when including the rainfall variability, with C_{MPI} criterion of 0.73 m and a C_{PC} of 0.68 m^{-1}.

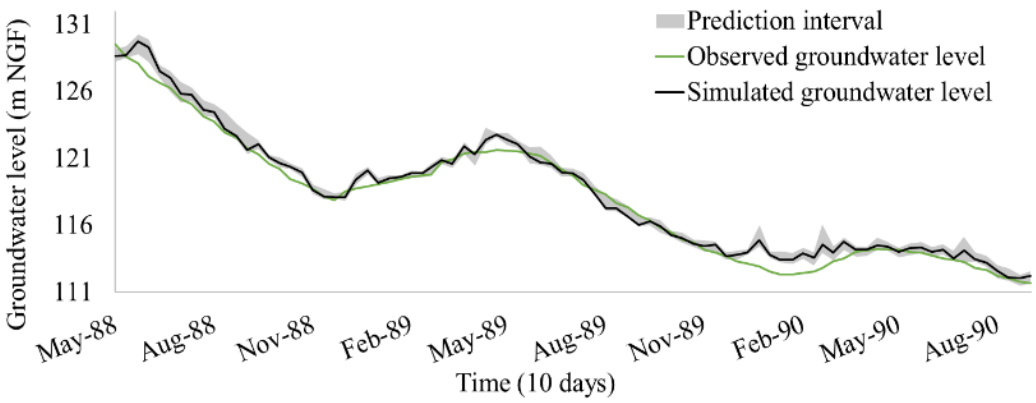


Figure 12. Groundwater level forecasting with 20 days' lead-time. Grey band shows the variability due to the parameter's initialization and to the rain spatial variability. C_P = 0.40; C_{PICP} = 0.49; C_{MPI} = 0.73 m.

4.5. Impact of the Spatial Distribution of Rainfall Uncertainty on the Model of Correctness

A new model of correctness is built as before, using the new ensemble model including the uncertainty due to the parameter's initialization, and the variability of rainfalls.

In order not to lengthen the presentation of the work, only the results of the laws with the set of best fits are presented, i.e., Gumbel, Raised Cosine, Logistic and Slash laws. Table 10 presents the correlations and indexes in a similar way to the presentation of Tables 6–9.

Table 10. Pearson's coefficients of correlation (r^2) between the measured water level distribution and the composite pdf; correlations between the model correctness and the empirical C_{PICP} calculated for each class of 20-cm groundwater levels (r_C^2); and Error Margins (EM).

Groundwater Level Class	Criteria	Laws			
		Gumbel	Raised Cosine	Logistic	Slash
Positive Slope	r_{in}^2	0.47	0.59	0.55	0.55
	r_{out}^2	0.52	0.56	0.53	0.54
	r_C^2	0.19	0.33	0.20	0.15
	$\%EM$	60.5%	61.8%	60.5%	63.2%
Negative Slope	r_{in}^2	0.78	0.79	0.77	0.72
	r_{out}^2	0.74	0.79	0.77	0.71
	r_C^2	0.52	0.41	0.49	0.47
	$\%EM$	70.3%	72.5%	71.4%	67.0%

As outlined above, the permutations of rainfall logically increase the forecast interval. Consequently, more observed values of water level belong to this interval. The Pearson correlation coefficients presented in Table 10 are therefore significantly higher than those of Tables 8 and 9. The Raised cosine law is that which clearly generates the highest correlations among the different composite pdfs. It will be thus used hereafter

4.6. Definition of a Confidence Interval

The model of correctness is calculated using a confidence index. This defines a confidence interval for each measured water level regardless of its configuration (positive or negative slope, inside or outside the prediction interval). A predefined confidence index thus provides a confidence interval.

Using the Raised Cosine's models of correctness, which presented the best fit, we have varied the confidence index from 0.60 to 0.95 by steps of 0.5. The corresponding

C_{PICP} are gathered in Table 11, and Figure 13 shows the confidence interval obtained for the confidence index of 0.90.

Table 11. Evolution of PICP and MPI scores following confidence index.

Confidence Index	C_{PICP} (Train + Test Datasets)	C_{PICP} (Test Set)	C_{MPI} (m)	C_{MPI} (m) (Without Extreme Values)
0.60	0.60	0.42	2.51	2.20
0.65	0.65	0.45	2.80	2.45
0.70	0.70	0.52	3.18	2.78
0.75	0.75	0.58	3.75	3.28
0.80	0.81	0.62	4.66	4.08
0.85	0.86	0.68	6.14	5.37
0.90	0.91	0.81	8.47	7.42
0.95	0.95	0.94	17.44	15.27

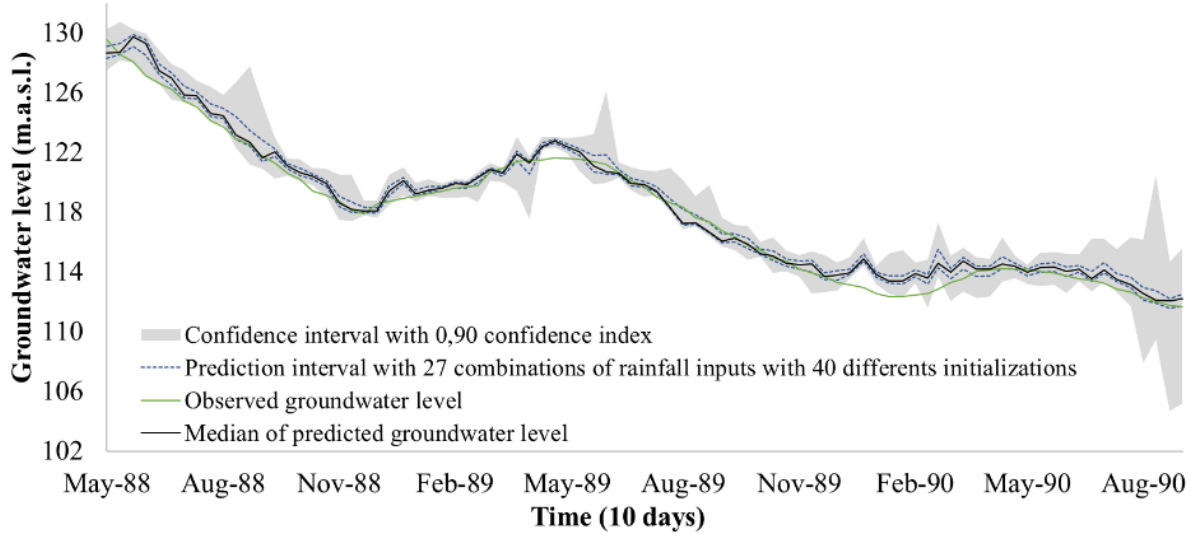


Figure 13. Groundwater level forecasting for 20 days' lead-time and confidence interval calculated with 0.90 confidence index. C_P = 0.40; C_{PICP} = 0.81; C_{MPI} = 8.47 m.

As expected, one can note in Table 11 that the confidence interval is wider when the confidence index is higher. The confidence interval associated with the highest confidence index (0.95) is very high (17.44 m) and probably not so useful to an end user. Nevertheless, the C_{MPI} decreases more quickly than the confidence index, which allows the manager to choose a compromise according to his requirements.

5. Discussion

5.1. Role of Rain in the Forecast Interval

Considering the two visualisations of the prediction intervals presented in Figure 6 and Figure 12, it appears first of all that the latter are rather weak, in particular if expressed as a percentage of the maximum water table beat (25 m; cf. Table 1), C_{MPI} = 0.62 m (2.5%) for the first and, respectively, C_{MPI} = 0.73 m (3%) for the second. These small intervals seem to be very accurate, but they do not provide any real added value for the user, since the measured water level does not always fall within this interval. On average, the C_{PICP} provides the probability that the prediction interval from the model contains the measured value; this is 39% for the former and 49% for the latter. Thus, even considering

the uncertainty caused by the measurement of rainfall variability (Figure 12), the model is only correct, on average, one time out of two.

Looking more specifically, in the test set in Figure 6, the model is better at predicting recessions. Both predicted values, water levels and confidence interval, are low during recessions. This is also shown by the correlations calculated on the composite pdf when the slopes are negative (Tables 6–9). For this reason, two different configurations have been separated to calculate the pdf: the case where the slope is positive, and the case where the slope is negative.

On further refinement, the measured and predicted water level curves deviate from each other mainly as a result of a strong rainfall pulse ($p > 40$ mm/day). This suggests that a specific phenomenon occurs in this configuration, which could be related to the influence of the Barbuise River. Indeed, if we refer to Table 2, we note that the response time linking the Discharge of the Barbuise and Pouan les Vallées (D_{BP}) is 2 decades; however, in Figure 6, we can see that the most important rainfall episode of February 1990 ($p > 60$ mm) influences the water table in less than one decade, its maximum effect appearing at 2 decades. It thus appears that infiltration with faster dynamics occurs during heavy rainfall episodes and that the model has difficulties in representing these fast and rare infiltrations. Moreover, the prediction interval is rather smaller for the responses to these episodes than for the other configurations, both for Figure 6 and Figure 12, suggesting that, during very wet episodes, the spatial variability of the rainfall events, at the decadal step, does not have a great impact on the response. Given the objective of the modelling, which is to predict low water, this double property, errors in prediction and low uncertainty during high rain pulses, can be considered as not being prohibitive.

Even so, in order to improve the representation of an uncertainty that would be more useful to managers, for example that would allow manager to choose the confidence interval that suits him, we have introduced the correctness models, delivering an "error margin" reported for the forecasts in the form of a confidence interval. This error margin is itself controllable by a user-defined confidence index. This addition has the same shortcomings with respect to major rainfall events but allows the manager to adapt the visualisation of the uncertainty to his needs. It could be used for managing pollutant intrusion during floods and for low levels anticipation, by choosing different confidence indices for these different uses.

5.2. *Role of the Amount of Data*

An important limitation to note is the amount of extreme data. While the range defining the water level classes of the training set has been chosen to contain at least 10 measurements per class, it can be noted that the data for the extreme class is only observed once. There is a lack of extreme data to be able to build reliable *pdf*. In particular, the chosen test set contains the lowest data over the entire history of the database (1997–2018), and the number of samples in the lowest groundwater levels classes is therefore very low; 0 for the lowest, and a few units for the others. Remember that the calculation of the correctness model assigns the maximum uncertainty when there has never been any measured value. This is the case for the dry period of the test set and this explains why the visualized confidence intervals for the summer 1990 are so large while the measured and predicted values are very close.

An improvement could be obtained by choosing different laws of probability in order to minimize the uncertainties in this specific condition. For example, in Figure 11, it can be seen that Gumbel's law gives the smallest error margin. This result is consistent with the application domain of Gumbel's law, which is aimed at distributions with extreme events. The confidence interval thus could be improved by using this specific law for very low and very high water level values.

6. Conclusions

The goal of this paper was to define a generic method able to estimate the uncertainty generated by both the neural network model itself and by the non-measured spatial heterogeneity of rainfalls. This work was carried out on the Champagne chalk aquifer (Northern France), at a 10-day time-step up to 20 days lead time and could be used for the purpose of tap water production and agriculture groundwater management.

A reliable enough ANN model to forecast groundwater level at "Craie à Vailly" piezometer was built, and tested on the driest summer of the entire 40 years database. The uncertainties due to the model parameters were then estimated following a simple method that takes into account the variability caused by the initialization of parameters. Added to this variability, the paper investigated the impact of the non-measured spatial variation on rainfall.

For this purpose, a set of permutations and substitutions of measurements at rainfall stations, combined in an original way with the implementation of the cross-validation process, was proposed and allowed the calculation of forecasts and uncertainties on a test set, never used in training, stopping, nor in cross-uncertainty assessment subsets. This test set contains the data for the most severe drought in the database. From the uncertainties found on this test set, a correctness model was proposed which provides, for the requirement of a global confidence index, a confidence interval to be applied to each forecast value.

Several limitations were identified. The main one is related to the amount of data: when there is no historical data in the range of values considered by the prediction, the uncertainty is maximal. It would be possible to improve the correctness model by choosing a more appropriate statistical law.

This methodology is original and can be deployed on other hydro-systems having other types of surface or subsurface features and different climate contexts. Applications may either make forecasts or propose a confidence interval associated with these forecasts, with the degree of confidence chosen by the user. These methodologies have proved significant advantages: the rigorous design of the neural network model has allowed the realization of a model capable of generalizing to a range of data that exceeded the range of the training set. Furthermore, it is possible to flexibly choose the confidence index according to the hydrological configuration (e.g., recession or rising water table).

Thanks to this methodology, a mid-term groundwater forecast with its own uncertainty can be provided. Our model, specialized for droughts showing as driest events on test dataset, allows for prevention of dry events, which can be anticipated nearly three weeks before, allowing agricultural and water supply end users to anticipate this risk, by issuing, for example, water withdrawal restriction orders or by carrying out water transfers from dams and reservoirs.

Author Contributions: Conceptualization, N.A., G.A. and M.S.; methodology, N.A., G.A., M.S. and A.J.; software, N.A. and A.J.; validation, G.A. and M.S.; formal analysis, N.A.; investigation, N.A.; resources, all authors; data curation, N.A.; writing—original draft preparation, N.A.; writing—review and editing, G.A., M.S., A.J. and M.V.; visualization, N.A.; supervision, A.J. and M.V.; project administration, A.J. and M.V. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the ARNT (Association Nationale de la Recherche et de la Technologie), grant number 1700720.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data sharing not applicable. No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: Authors would like to warmly thank the BRGM, which provided the dataset, J. Nicolas (BRGM) and S. Pistre (Univ. Montpellier) for helpful discussions, IT services of IMT Mines Ales and AQUASYS, for technical support and infrastructure deployed during coronavirus crisis, OUGC84, for sharing agricultural data, GEONOSIS, for modelling software RNF Pro and the ANRT, which contributed to funding this PhD thesis.

Conflicts of Interest: The authors declare no conflict of interest, except for a commercial use. Readers must have to keep in mind that presented results are not generalizable for others basins or others outputs and need additional studies. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix A

	Agricultural Area (ha)	Agricultural Area (%)	Water requirement by crop type ($m^3.ha^{-1}.y^{-1}$)	*January*		*February*		*March*		*April*		*May*	
				Water requirement (%)	Water requirement ($m^3.ha^{-1}$)	Water requirement (%)	Water requirement ($m^3.ha^{-1}$)	Water requirement (%)	Water requirement ($m^3.ha^{-1}$)	Water requirement (%)	Water requirement ($m^3.ha^{-1}$)	Water requirement (%)	Water requirement ($m^3.ha^{-1}$)
Soft wheat	130492.7	30.32	2000	0.00	0.00	0.00	0.00	0.20	121.28	0.30	181.92	0.40	242.56
Barley	94802.7	22.03	3000	0.00	0.00	0.00	0.00	0.15	99.14	0.35	231.32	0.45	297.41
Beet	57120.8	13.27	2750	0.00	0.00	0.00	0.00	0.25	91.23	0.25	91.23	0.25	91.23
Rapeseed	55213.6	12.83	2500	0.00	0.00	0.00	0.00	0.05	16.04	0.90	288.68	0.05	16.04
Alfalfa	27524.5	6.40	3000	0.00	0.00	0.00	0.00	0.00	0.00	0.10	19.20	0.30	57.60
Potato	12015.2	2.79	2500	0.00	0.00	0.00	0.00	0.05	3.49	0.05	3.49	0.10	6.98
Corn	9364.3	2.18	4500	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.08	7.85
Permanent grasslands	8492.9	1.97	4000	0.00	0.00	0.00	0.00	0.00	0.00	0.05	3.94	0.10	7.88
Peas	5087.8	1.18	2000	0.00	0.00	0.00	0.00	0.00	0.00	0.20	4.72	0.50	11.80
Medicinal plants	2121.4	0.49	2000	0.00	0.00	0.00	0.00	0.00	0.00	0.10	0.98	0.10	0.98
Temporary grasslands	1972.0	0.46	4000	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.92	0.10	1.84
Hemp	1838.3	0.43	5000	0.00	0.00	0.05	1.08	0.05	1.08	0.05	1.08	0.15	3.23
Lentils	1821.7	0.42	4000	0.00	0.00	0.00	0.00	0.05	0.84	0.05	0.84	0.05	0.84
Vineyards (wine grapes)	1772.0	0.41	1500	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Poppy seeds	1319.6	0.31	4000	0.05	0.62	0.05	0.62	0.05	0.62	0.05	0.62	0.10	1.24
Oats	1267.6	0.29	4000	0.00	0.00	0.00	0.00	0.05	0.58	0.05	0.58	0.10	1.16
Fava bean	1112.0	0.26	2000	0.00	0.00	0.00	0.00	0.00	0.00	0.20	1.04	0.50	2.60
Sunflower	1040.5	0.24	3500	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.42
Oignon	959.2	0.22	3000	0.00	0.00	0.00	0.00	0.00	0.00	0.45	2.97	0.45	2.97
Fescue	854.6	0.20	4000	0.00	0.00	0.00	0.00	0.05	0.40	0.05	0.40	0.05	0.40
Monthly water requirement ($m^3.ha^{-1}.month^{-1}$)				**0.62**		**1.70**		**334.69**		**833.91**		**755.01**	

June		*July*		*August*		*September*		*October*		*November*		*December*	
Water requirement (%)	Water requirement ($m^3.ha^{-1}$)	Water requirement (%)	Water requirement ($m^3.ha^{-1}$)	Water requirement (%)	Water requirement ($m^3.ha^{-1}$)	Water requirement (%)	Water requirement ($m^3.ha^{-1}$)	Water requirement (%)	Water requirement ($m^3.ha^{-1}$)	Water requirement (%)	Water requirement ($m^3.ha^{-1}$)	Water requirement (%)	Water requirement ($m^3.ha^{-1}$)
0.10	60.64	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.05	33.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.75	273.69	1.00	364.93	0.75	273.69	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.30	57.60	0.30	57.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.25	17.44	0.25	17.44	0.30	20.93	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.25	24.53	0.35	34.34	0.27	26.49	0.05	4.91	0.00	0.00	0.00	0.00	0.00	0.00
0.20	15.76	0.30	23.64	0.25	19.70	0.05	3.94	0.05	3.94	0.00	0.00	0.00	0.00
0.30	7.08	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.20	1.96	0.25	2.45	0.25	2.45	0.10	0.98	0.00	0.00	0.00	0.00	0.00	0.00
0.20	3.68	0.30	5.52	0.25	4.60	0.05	0.92	0.05	0.92	0.00	0.00	0.00	0.00
0.20	4.30	0.20	4.30	0.20	4.30	0.05	1.08	0.05	1.08	0.00	0.00	0.00	0.00
0.15	2.52	0.25	4.20	0.25	4.20	0.10	1.68	0.05	0.84	0.05	0.84	0.00	0.00
0.30	1.85	0.40	2.46	0.20	1.23	0.10	0.62	0.00	0.00	0.00	0.00	0.00	0.00
0.15	1.86	0.15	1.86	0.15	1.86	0.10	1.24	0.05	0.62	0.05	0.62	0.05	0.62
0.20	2.32	0.25	2.90	0.20	2.32	0.10	1.16	0.05	0.58	0.00	0.00	0.00	0.00
0.30	1.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.30	2.52	0.50	4.20	0.15	1.26	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.10	0.66	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.15	1.20	0.25	2.00	0.25	2.00	0.10	0.80	0.05	0.40	0.05	0.40	0.00	0.00
514.21		**527.83**		**365.03**		**17.32**		**8.38**		**1.86**		**0.62**	

Figure A1. Details of the construction and of the values of the irrigation inputs for the different types of crops as a function of the surface, the water requirement and the season (by months).

References

1. Hartmann, D.L.; Klein Tank, A.M.G.; Rusticucci, M.; Alexander, L.V.; Brönnimann, S.; Charabi, Y.A.-R.; Dentener, F.J.; Dlugokencky, E.J.; Easterling, D.R.; Kaplan, A.; et al. Observations: Atmosphere and Surface. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*; Cambridge University Press: Cambridge, UK, 2013; Chapter 2, pp. 159–254.

2. Hornik, K.; Stinchombe, M.; White, H. Multilayer Feedforward Networks are Universal Approximators. *Neural Netw.* **1989**, *2*, 359–366.
3. Wagener, T.; Montanari, A. Convergence of approaches toward reducing uncertainty in predictions in ungauged basins. *Water Resour. Res.* **2011**, *47*, doi:10.1029/2010WR009469.
4. Bourgin, F. *Comment Quantifier l'incertitude Prédictive en Modélisation Hydrologique? Travail Exploratoire sur un Grand Échantillon de Bassins Versants*; AgroParisTech: Paris, France, 2014.
5. Solomatine, D.P.; Shrestha, D.L. A novel method to estimate model uncertainty using machine learning techniques. *Water Resour. Res.* **2009**, *45*, 1–16.
6. Biondi, D.; Todini, E. Comparing Hydrological Postprocessors Including Ensemble Predictions into Full Predictive Probability Distribution of Streamflow. *Water Resour. Res.* **2018**, *54*, 9860–9882.
7. Krzysztofowicz, R. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resour. Res.* **1999**, *35*, 2739–2750.
8. Biondi, D.; Luca, D.L.D. Performance assessment of a Bayesian Forecasting System (BFS) for real-time flood forecasting. *J. Hydrol.* **2013**, *479*, 51–63.
9. Wu, W.; Dandy, G.C.; Maier, H.R. Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling. *Environ. Model. Softw.* **2014**, *54*, 108–127.
10. Kong-A-Siou, L.; Johannet, A.; Borrell, V.; Pistre, S. Complexity selection of a neural network model for karst flood forecasting: The case of the Lez Basin (southern France). *J. Hydrol.* **2011**, *403*, 367–380.
11. Stone, M. Cross-Validatory Choice and Assessment of Statistical Predictions (With Discussion). *J. R. Stat. Soc. Ser. B* **1974**, *38*, 111–147.
12. Sjöberg, J.; Zhang, Q.; Ljung, L.; Benveniste, A.; Delyon, B.; Glorennec, P.-Y.; Hjalmarsson, H.; Juditskys, A. Nonlinear Black-box Modeling in System Identification: A Unified Overview. *Automatica* **1995**, *31*, 1691–1724.
13. Akil, N.; Artigue, G.; Savary, M.; Johannet, A.; Vinches, M. Quantification of Neural Network Uncertainties on the Hydrogeological Predictions by Probability Density Functions. In Proceedings of the IOP Conference Series: Earth and Environmental Science, Prague, Czech Republic, 9–13 September 2019; p. 10.
14. Dreyfus, G. *Neural Networks, Methodology and Applications*; Springer: Berlin, Germany, 2005; p. 509.
15. Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Netw.* **1991**, *4*, 251–257.
16. Barron, A.R. Approximation bounds for superpositions of a sigmoidal function. In Proceedings of the IEEE International Symposium on Information Theory—Proceedings, San Antonio, TX, USA, 17–22 January 1993; pp. 930–945.
17. Nerrand, O.; Roussel-Ragot, P.; Personnaz, L.; Dreyfus, G.; Marcos, S. Neural Networks and Nonlinear Adaptive Filtering: Unifying Concepts and New Algorithms. *Neural Comput.* **1993**, *5*, 165–199.
18. Artigue, G.; Johannet, A.; Borrell, V.; Pistre, S. Flash flood forecasting in poorly gauged basins using neural networks: Case study of the Gardon de Mialet basin (southern France). *Nat. Hazards Earth Syst. Sci.* **2012**, *12*, 3307–3324.
19. Taver, V.; Johannet, A.; Borrell-Estupina, V.; Pistre, S. Feed-forward vs recurrent neural network models for non-stationarity modelling using data assimilation and adaptivity. *Hydrol. Sci. J.* **2015**, *60*, 1242–1265.
20. Geman, S.; Bienenstock, E.; Doursat, R. Neural Networks and the Bias/Variance dilemma. *Neural Comput.* **1992**, *4*, 1–58.
21. Darras, T.; Johannet, A.; Vayssade, B.; Long-a-Siou, L.; Pistre, S. Influence of the Initialization of Multilayer Perceptron for Flash Floods Forecasting: How Designing a Robust Model. In *International Work-Conference on Time Series 2014*; Springer: Granada, Spain, 2014; p. 13.
22. Kong-A-Siou, L.; Johannet, A.; Borrell Estupina, V.; Pistre, S. Optimization of the generalization capability for rainfall-runoff modeling by neural networks: The case of the Lez aquifer (southern France). *Environ. Earth Sci.* **2012**, *65*, 2365–2375.
23. AESN. *Fiche de Caractérisation de la ME HG20—Masse d'eau souterraine HG208 "Craie de Champagne Sud et Centre*; AESN: Nanterre, France, 2015.
24. Météo France. Données Publiques—Observations In Situ. Available online: https://donneespubliques.meteofrance.fr/?fond=rubrique&id_rubrique=26 (accessed on 8 September 2020).
25. BNPE. Base de données des prélèvements en eau, OPR0000035289. Available online: https://bnpe.eaufrance.fr/acces-donnees/codeOuvrage/OPR0000035289/annee/2014 (accessed on 15 January 2021).
26. BNPE. Base de Données des Prélèvements en eau, Vailly (10). Available online: https://bnpe.eaufrance.fr/?q=acces-donnees/codeCommune/10391/annee/2017/etCommunesAdjacentes (accessed on 31 May 2021).
27. Stollsteiner, P. *Connaissance des Ressources Réellement Disponibles sur L'ensemble des Bassins Versants Crayeux*; BRGM: Reims, France, 2013.
28. Kloppmann, W.; Dever, L.; Edmunds, W.M. Residence time of Chalk groundwaters in the Paris Basin and the North German Basin: A geochemical approach. *Appl. Geochem.* **1998**, *13*, 593–606.
29. Pinault, J.-L.; Allier, D.; Chabart, M. *Prévision des Volumes d'eau Exploitables de 10 Bassins versants en Champagne Crayeuse*; BRGM: Reims, France, 2006.
30. Putot, E.; Verjus, P.; Vernoux, J.F. *Qualification des Piézomètres du Réseau de Bassin Seine Normandie en 2005*; BRGM: Massy, France, 2006.
31. IGN. Registre Parcellaire Graphique (RPG): Contours Des Parcelles et Îlots Culturaux et Leur Groupe de Cultures Majoritaire. Available online: https://www.data.gouv.fr/fr/datasets/registre-parcellaire-graphique-rpg-contours-des-parcelles-et-ilots-culturaux-et-leur-groupe-de-cultures-majoritaire/#_ (accessed on 14 September 2020).

32. Mangin, A. Pour une meilleure connaissance des systèmes hydrologiques à partir des analyses corrélatoire et spectrale. *J. Hydrol.* **1984**, *67*, 25–43.
33. Kitanidis, P.K.; Bras, R.L. Real-time forecasting with a conceptual hydrologic model: 2. Applications and results. *Water Resour. Res.* **1980**, *16*, 1034–1044.
34. Shrestha, D.L.; Solomatine, D.P. Machine learning approaches for estimation of prediction interval for the model output. *Neural Netw.* **2006**, *19*, 225–235.
35. Khosravi, A.; Nahavandi, S.; Creighton, D. A prediction interval-based approach to determine optimal structures of neural network metamodels. *Expert Syst. Appl.* **2010**, *37*, 2377–2387.
36. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum Likelihood from Incomplete Data via the EM Algorithm. *J. R. Stat. Soc. Ser. B* **1977**, *39*, 1–22.
37. Benaglia, T.; Chauveau, D.; Hunter, D.R.; Young, D.S. Mixtools: An R Package for Analyzing Finite Mixture Models. *J. Stat. Softw.* **2009**, *32*, 29.
38. De Moivre, A. *The Doctrine of Chances, or, A Method of Calculating the Probability of Events in Play*, III ed.; Chelsea Publishing Company: Londres, UK, 1756; p. 368.
39. Fisher, R.A. On the mathematical foundations of theoretical statistics. *Philos. Trans. R. Soc.* **1922**, *9*, 309–368.
40. Gumbel, E.J. Méthodes graphiques pour l'analyse de débits de crues. *Rev. De Stat. Appl.* **1957**, *5*, 77–89.
41. Laplace, P.-S. Mémoire sur la probabilité des causes par les évènements. In *Œuvres complètes de Laplace*; Mémoires de l'Academie Royale des Sciences, Ed.; Divers Savan: Paris, France, 1774; Chapter 6, pp. 621–656.
42. Rinne, H. *Location-Scale Distributions Linear Estimation and Probability Plotting Using MATLAB*; Justus Liebig University: Giessen, Allemagne, 2010.
43. Sinha, R.K. A Thought on Exotic Statistical Distributions. *Int. J. Math. Comput. Sci.* **2012**, *6*, 49–52.
44. Cauchy, A.-L. CALCUL DES PROBABILITÉS—Sur les résultats moyens d'observations de même nature, et sur les résultats les plus probables. In *Oeuvres Complètes*; Cambridge University Press: Cambridge, UK, 1853; pp. 94–104.
45. Stigler, S.M. *An Historical Note on the Cauchy Distribution*; Biometrika Trust: Oxford, UK, 1974, pp. 375–380.
46. Verhulst, P.F. Recherches mathématiques sur la loi d'accroissement de la population. In *Nouveaux mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles*; Nabu Press: Bruxelles, Belgique, 1845; Chapter 18, pp. 1–38.
47. Rogers, W.H.; Tukey, J.W. Understanding some long-tailed symmetrical distributions. *Stat. Neerl.* **1972**, *26*, 211–226.
48. Bhattacharjee, G.P.; Pandit, S.N.N.; Mohan, R. Dimensional Chains Involving Rectangular and Normal Error-Distributions. *Technometrics* **1963**, *5*, 404–406.
49. Huber, P.J. Robust estimation of a location parameter. *Ann. Math. Stat.* **1964**, *35*, 73–101.
50. Huber, P.J. *Robust Statistics*; Wiley: Hoboken, NJ, USA, 1981; p. 308.
51. Levenberg, K. A method for the solution of certain non-linear problems in least squares. *Q. Appl. Math.* **1944**, *2*, 164–168.
52. Marquardt, D.W. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *J. Soc. Ind. Appl. Math.* **1963**, *11*, 431–441.

III.3.1. Compléments à l'article

Les différents compléments seront séquencés par paragraphes. Ceux-ci concernent :

- la sélection du modèle :

Plusieurs architectures ont été testées sur ce bassin. Un réseau de type perceptron multicouche avec trois couches de neurones sans les entrées d'irrigation a donné des résultats très similaires à ceux obtenus sur l'architecture retenue (C_P = 0,65 en validation croisée et C_P = 0,38 en test). Cependant, ce modèle n'a pas été retenu compte-tenu de la forte complexité du modèle, du au nombre important de paramètres libres.

De même, un réseau *Long-Short Term Memory* a donné des résultats un peu moins bons (avec un C_P = 0,35 en test). Celui-ci n'a pas été retenu car le perceptron multicouche semble être légèrement plus performant sur ce bassin.

Il est à noter également que le critère de biais (C_{PBIAS}) est ici de -0,27% pour le test sur le modèle retenu, attestant de la légère surestimation du modèle, en particulier constatée lors de la remontée de nappe en début d'année 1989 et surtout en période d'étiage entre décembre 1989 et mars 1990.

- les modèles de justesse :

Les distances de Kolmogorov-Smirnov D_{KS} ont été calculées pour justifier l'inférence des lois sur les différentes distributions. Celles-ci ont été ajoutées dans le Tableau IV-11 dans le cas où les niveaux de nappe augmentent (période de recharge), accompagnés des coefficients de corrélation de Pearson r^2.

Tableau IV-11 - Indicateurs d'inférence des lois à densité de probabilité.

Lois à densité de probabilité	***r^2*** **(Inclus dans PI)**	***r^2*** **(Hors de PI)**	***D_{KS}*** **(Inclus dans PI)**	***D_{KS}*** **(Hors de PI)**
Loi Normale	0,43	0,52	0,017	0,055
Loi de Gumbel	0,41	0,52	0,016	0,044
Loi de Laplace	0,47	0,54	0,013	0,041
Loi du Cosinus surélevé	0,44	0,66	0,014	0,032
Loi de Cauchy	0,45	0,55	0,008	0,019
Loi Logistique	0,44	0,61	0,012	0,037
Loi de Slash	0,43	0,63	0,043	0,116
Loi Bhattacharjee	0,43	0,57	0,016	0,049
Loi de Huber	0,42	0,52	0,012	0,052

Les distances de Kolmogorov-Smirnov sont proches de zéro pour la plupart des lois. Seule la loi de Slash surévalue les valeurs observées et obtient des distances un peu plus élevées que pour les autres lois. La loi de Cauchy obtient les distances les plus courtes, malgré une corrélation plus faible que la plupart des lois. Les lois de Gumbel, Cosinus surélevé et Logistique obtiennent également des distances relativement courtes, ce qui semble être en accord avec les conclusions faites avec le coefficient de corrélation dans l'article.

À la vue des résultats évoqués dans le paragraphe précédent et dans l'article, les lois de Logistique, Cosinus Surélevé, Gumbel, Cauchy et Slash semblent être les lois les mieux adaptées à l'inférence des niveaux piézométriques en augmentation.

Ce même travail a été réalisé pour les niveaux d'eau souterraine en diminution (vidange de la nappe), où les distances D_{KS} sont présentées sur le Tableau IV-12.

Tableau IV-12 - Indicateurs d'inférence des lois à densité de probabilité.

Lois à densité de probabilité	***r²*** **(Inclus dans PI)**	***r²*** **(Hors de PI)**	***D_{KS}*** **(Inclus dans PI)**	***D_{KS}*** **(Hors de PI)**
Loi Normale	0,66	0,68	0,048	0,037
Loi de Gumbel	0,74	0,74	0,030	0,021
Loi de Laplace	0,66	0,67	0,029	0,028
Loi du Cosinus surélevé	0,76	0,77	0,030	0,022
Loi de Cauchy	0,69	0,70	0,022	0,037
Loi Logistique	0,75	0,75	0,027	0,022
Loi de Slash	0,64	0,72	0,188	0,170
Loi Bhattacharjee	0,64	0,75	0,042	0,032
Loi de Huber	0,50	0,65	0,064	0,073

Les distances de Kolmogorov-Smirnov les plus courtes sont obtenues par les lois du Cosinus Surélevé, Logistiques et Cauchy. La loi de Slash obtient ici des distances éloignées de zéro, laissant supposer une mauvaise inférence de cette loi sur les niveaux de nappe en diminution.

Ainsi, les lois de Gumbel, Logistique et Cosinus Surélevé semblent être les mieux adaptées pour inférer les niveaux d'eau souterraine en diminution. Cette conclusion est sensiblement identique pour les niveaux de nappe en augmentation, en ajoutant la loi de Slash et correspond aux conclusions faites dans l'article.

- la détermination de l'intervalle en fonction d'une justesse exigée :

Les modèles de justesse issus de la loi du Cosinus surélevé sont utilisés pour déterminer un intervalle de prévision à partir d'une justesse exigée. Les incertitudes liées à la distribution

spatiale des pluies sont incluses dans cette étude. Les probabilités de justesse appliquées aux niveaux piézométriques observés, en fonction de s'ils sont en augmentation ou en diminution, sont comparées avec ces mêmes probabilités de justesse appliquées aux résultats de modélisation médians. La Figure IV-24 montre cette relation, qui se transcrit par une fonction linéaire.

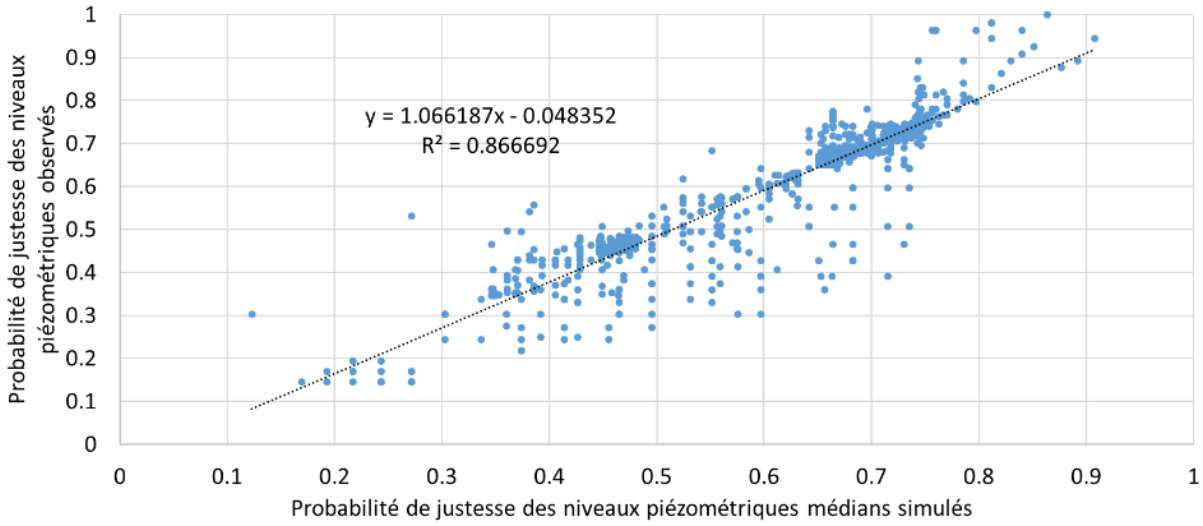


Figure IV-24 - Probabilité de justesse des niveaux piézométriques médians simulés en fonction de la probabilité des niveaux piézométriques observés.

Les limites de l'intervalle de prévision avec une justesse exigée sont obtenues à partir des coefficients de la droite de régression linéaire. La corrélation de la courbe de tendance avec les probabilités de justesse est bonne, égale à 0,87.

Pour ce cas d'étude avec le modèle utilisé, la probabilité de justesse maximale atteinte est de 0,96, lorsque cette justesse est imposée à 1.

III.3.2. Principales conclusions de l'article

La modélisation des niveaux piézométriques à moyen terme (horizon de prévision de 20 jours) s'est réalisée par l'utilisation d'un perceptron multicouche. Les différentes méthodes de régularisation ont permis de gagner en robustesse et de limiter les incertitudes liées à l'architecture du modèle.

Les incertitudes de modélisation ont été estimées à partir d'une méthodologie conçue lors de cet article, dont les prémices ont été visibles sur l'étude sur le Blavet (Akil, et al. 2019). Un des avantages de la méthodologie est qu'elle peut calculer les incertitudes de modélisation en fonction de leurs origines, et ce, de manière généralisable. Dans notre cas, ces incertitudes se sont principalement portées sur l'initialisation des paramètres, en utilisation des modèles d'ensembles avec une initialisation différentes pour chaque membre. Ces incertitudes de

modélisation comprennent également les incertitudes issues de la variabilité spatiale des pluies, en permutant les entrées de pluies du modèle.

La conception des modèles de justesse a vu la bonne inférence de la loi du Cosinus Surélevée vis-à-vis de la distribution des niveaux piézométriques sur les quatre catégories de hauteurs d'eau. Ces catégories dépendent à la fois de l'état précédent de l'hydrosystème (recharge ou vidange de la nappe) et de l'appartenance ou non de la modélisation à l'intervalle de prévision.

A partir de ces modèles, l'intervalle de prévision a pu être reconstitué à partir d'une justesse exigée par un utilisateur. Ce point renforce l'aspect commercialisable de la méthodologie.

Malgré tous ces atouts de cette méthodologie, cette dernière contient quelques inconvénients. Le premier est que cette méthode nécessite un nombre de données en entrée suffisant, ce qui est le cas dans notre étude, hormis pour les valeurs extrêmes, dont leurs incertitudes sont plus difficiles à estimer. Cela peut influencer le choix final de la loi utilisée pour les modèles de justesse et créer des intervalles de prévision grands pour ces valeurs.

Un autre inconvénient de cette étude concerne le bassin. L'agriculture, la viticulture et l'industrie textile sont les fers de lance de cette région. Ces secteurs d'activité ont des besoins en eau et sont sensibles à toute sécheresse. La prévision à moyen-terme des niveaux de nappe est une des solutions possibles afin d'anticiper ce risque. Cette modélisation demeure toutefois ardue, de par le secteur agricole omniprésent, qui nécessite de nombreux prélèvements en eau pour ces secteurs et l'alimentation en eau potable. Ces prélèvements en eau induisent une modification potentielle des niveaux de nappe, ce qui est problématique pour modéliser les périodes d'étiage.

La modélisation hydrogéologique reste toutefois possible avec la méthodologie élaborée, engendrant une qualité du modèle relativement correcte. La méthodologie d'estimation des incertitudes de modélisation ont mis en évidence le manque de justesse du perceptron multicouche utilisé pour prévoir les hauteurs d'eau souterraine en période de recharge de la nappe. Les multiples stations pluviométriques autour de notre point d'étude a permis de compléter la méthodologie, en ajoutant les incertitudes liées à la dispersion spatiale des pluies. Cet ajout a permis d'élargir l'intervalle de prévision et d'augmenter le nombre de prévisions considérées comme étant bonnes, sans modifier la qualité du modèle neuronal.

IV. Discussions

IV.1 Conclusions générales sur les modélisations réalisées sur les deux bassins

Plusieurs conclusions peuvent être tirées des expériences menées. Du point de vue de la performance absolue des modèles, les travaux sur le bassin du Blavet ont conduit à de meilleurs résultats que ceux portant sur le bassin de la nappe de la Craie. Les critères de performance y étaient en effet plus élevés, de même que les probabilités de justesse. Par ailleurs, dans les deux cas, on a pu observer des scores inférieurs lorsque les niveaux piézométriques étaient ascendants. Enfin, l'utilisation de données décadaires permet de limiter le bruit dans les données ainsi que de proposer des prévisions plus lointaines, mais on a pu remarquer que la baisse du nombre d'exemples liée à l'utilisation de grands pas de temps était défavorable aux performances de modèles et inférences.

Indépendamment de la performance des modèles, les lois qui ont montré les meilleures performances pour représenter la distribution des niveaux piézométriques sont les lois du Cosinus Surélevé et dans une moindre mesure, celles de Gumbel et Logistique.

La méthodologie d'estimation de l'incertitude de modélisation par densités de probabilité a pu être appliqué pour les trois modèles réalisés. La conception de modèles de justesse et la réalisation d'un intervalle de prévision fixé à partir d'une justesse exigée ont permis d'apporter des éléments pour répondre à la question de l'estimation de l'incertitude de modélisation. Néanmoins, cette méthodologie présente plusieurs avantages et inconvénients.

IV.2 Avantages et inconvénients de la méthodologie développée

IV.2.1. Avantages

D'abord, la méthodologie développée tient compte de l'état du système et du modèle, représentés par la valeur de leurs sorties respectives (hypothèse de bruit d'état). Cette approche est cohérente avec la non linéarité de la réponse de ce type de système face à une stimulation, non linéarité en grande partie liée à l'état du système au moment de la stimulation. Par ailleurs, la méthodologie tient aussi compte de l'évolution courante de cet état afin de représenter les comportements distincts de recharge et de vidange du système.

La méthodologie élaborée repose sur deux étapes, toutes deux utiles à un modélisateur, et commercialisables par une entreprise. La première est l'estimation de l'incertitude de modélisation par l'utilisation de modèles de justesse. Une fois le modèle de justesse conçu pour un ensemble de modèles neuronaux, le modélisateur est capable de fournir un résultat de modélisation avec une incertitude associée à un client. Le modèle de justesse sert ici d'abaque pour estimer cette incertitude. La seconde étape consiste à générer un intervalle de

confiance à partir d'une justesse exigée par un client. Cela permet au modélisateur de garantir la justesse de son modèle au client.

La méthodologie d'estimation de l'incertitude de modélisation présentée ici est applicable pour tout type de sortie de modèle. Elle est certes optimisée pour les sorties hydrogéologiques, telles que les niveaux piézométriques, mais est aussi applicable pour des sorties de modèles hydro-climatiques, tels les débits et hauteurs d'eau en rivière, l'évapotranspiration ou encore la pluie. Il est potentiellement envisageable d'utiliser tout type de sortie temporelle pour pouvoir estimer l'incertitude de modélisation par cette technique.

Un autre atout associé à cette méthodologie est qu'elle est applicable pour tout type de modèle, tant que ceux-ci soient aptes à fournir plusieurs résultats de modélisations. Les exemples montrés sur les deux bassins étudiés se sont réalisés avec des modèles de type réseaux de neurones. Toutefois, la méthodologie développée peut être appliquée à un modèle à réservoir par exemple. En effet, si ce modèle peut générer plusieurs simulations en faisant varier ses conditions initiales ou ses paramètres, la méthodologie d'estimation de l'incertitude de modélisation par densités de probabilité sera alors adaptable.

De même, l'avantage de cette méthode est qu'elle peut incorporer plusieurs sources d'incertitudes. Dans nos études, l'incertitude de modélisation calculée incorpore l'incertitude liée à l'initialisation des paramètres et l'incertitude liée à la représentativité spatiale des précipitations. Il est toutefois possible d'imaginer d'ajouter l'incertitude des données d'entrées. Pour cela, les données d'entrées sont composées des valeurs avec ajout et retrait des erreurs de mesure possibles. Les résultats de modélisation engendrent ainsi la création d'un nouvel intervalle de prévision, qui complète l'intervalle de prévision existant.

L'application de cette méthodologie ne requiert pas de connaissances spécifiques sur le bassin d'étude. Les connaissances basiques sur le territoire suffisent pour estimer les incertitudes de modélisation. L'approche de cette technique est ici majoritairement statistique.

Cette technique est dite discrète. L'estimation de l'incertitude de modélisation n'influe pas sur les résultats de modélisation, ni sur ses incertitudes associées. Cependant, l'estimation de l'incertitude fournie par cette méthode sera elle-même soumise à une incertitude, plus ou moins grande selon les cas de figure.

Cette méthodologie nécessite peu de maintenance. Seule la mise à jour de données intégrées en entrées d'un modèle hydrogéologique engendre le recalcul des incertitudes de modélisation. Parallèlement, le coût financier et matériel pour appliquer cette méthode est faible. L'utilisation d'un logiciel de modélisation couplé à des programmes libres de droits et d'une licence Excel suffisent pour déterminer l'incertitude de modélisation. Le coût financier

se résume principalement au salaire du modélisateur et du faible coût de maintenance des infrastructures bureautiques et informatiques.

IV.2.2. Inconvénients

Cependant, cette méthodologie n'est pas parfaite et peut ne pas convenir dans certains cas. En effet, l'application de cette méthodologie requiert l'utilisation de beaucoup de données afin d'avoir potentiellement un modèle de bonne qualité, un biais statistique faible induisant une estimation convenable des incertitudes de modélisation. En effet, la qualité de l'estimation des incertitudes dépend fortement de la quantité de données en entrée du modèle.

Ce problème induit un second problème. Le biais statistique évoqué au paragraphe précédent provient de manque de données. Toutefois, il est possible que ce manque de données soit présent uniquement sur une plage de valeurs piézométriques. Cela est régulièrement le cas pour les valeurs dites extrêmes (fort étiage ou période de hautes eaux). Cela est d'autant plus contraignant si le modèle hydrologique utilisé est dédié à ces évènements, comme l'a été le réseau de neurones utilisé pour la modélisation des niveaux d'eau souterraine sur le bassin de la nappe de la Craie de Champagne. Cela engendre donc une forte incertitude sur l'estimation des incertitudes de modélisation pour ces valeurs.

Pour solutionner ces deux problèmes, il existe toutefois des possibilités. Les classes de valeurs de sortie du modèle peuvent être raisonnablement élargies afin qu'elles comprennent suffisamment de valeurs. Il est aussi possible de ne pas prendre en compte la variation de l'état de la sortie observée pour réduire encore plus le biais statistique. Cette troisième solution engendre toutefois la perte d'un des avantages de cette méthode.

Les études précédentes ont également montré que pour des niveaux piézométriques observés ayant une faible probabilité de justesse, l'intervalle de prévision ayant une exigence de justesse imposée sera très grand. Ainsi, les modèles avec des performances moindres seront sujet à des intervalles de prévision larges.

De même, la justesse exigée ne peut dépasser la justesse maximale. Il est par exemple impossible d'exiger une justesse à 0,99 si la justesse maximale (correspondant à la justesse calculée lorsque celle qui est exigée est de 1) est égale à 0,95. De plus, l'exigence de la justesse n'est valable qu'en moyenne sur l'ensemble de des données d'entrées du modèle. Cette exigence peut ne pas être respectée sur une période donnée. Il est toutefois envisageable de grossir artificiellement la justesse exigée afin d'anticiper ce problème.

Le dernier inconvénient est la mise en place de cette technique. En effet, cette méthode requiert des compétences en informatique, environnement et en statistiques. De plus, sa mise en place peut être longue, de par les temps de calculs nécessaires lors des plans d'expériences

avec validations croisées. Ce temps de calcul dépend néanmoins fortement du bassin étudié, de l'architecture sélectionnée et des données utilisées. L'augmentation constante des capacités de calcul devrait être de nature à limiter cet inconvénient dans les prochaines années.

Conclusion

La conception d'un modèle opérationnel permettant de prévoir la disponibilité de la ressource en eau a été la trame directrice de ce mémoire. Parfois abondante, parfois rare, la ressource en eau reste un sujet complexe à appréhender. Nous avons tout d'abord présenté la méthodologie de conception de modèles basés sur les réseaux de neurones, méthodologie généralisable à d'autres hydrosystèmes, mais aussi robuste et fiable. Cette dernière se déroule en trois étapes : l'étude du territoire (par le biais d'analyses de corrélation et de bibliographie), le traitement des données d'entrée, l'élaboration et la sélection d'un modèle en utilisant les méthodes de régularisation. Cette méthodologie permet en outre de minimiser l'impact des incertitudes liées à l'architecture du modèle. Ainsi, l'utilisation des réseaux de neurones, combinée à cette méthodologie de conception de modèles, a permis de mener à bien la modélisation des niveaux d'eau souterraine à court et moyen terme, pour répondre aux trois problématiques définies dans l'introduction de ce manuscrit, sur deux bassins différents, avec des contextes hydrologiques, climatiques, environnementaux et anthropiques variés.

Si les réseaux de neurones délivrent une prévision, ils ne fournissent pas une estimation de l'incertitude associée à ce résultat. Aussi, l'utilisation d'outils complémentaires a été nécessaire pour mener à bien nos objectifs, tels que les calculs de probabilités et calculs statistiques. Ainsi, une méthodologie innovante utilisant ces outils statistiques a été développée au cours de cette thèse afin d'estimer ces incertitudes. Cela a conduit à la conception de modèles de justesse à partir de lois à densité de probabilité, neuf lois à densité de probabilité ont été utilisées pour estimer les incertitudes, en suivant une méthodologie originale développée au cours de cette thèse. La loi du Cosinus surélevé et la loi Logistique ont le plus souvent été montrées comme les plus pertinentes à utiliser pour répondre aux trois problématiques abordées.

La méthodologie développée au cours de ces travaux s'est révélée être cohérente et a plusieurs avantages quant à l'estimation de ces incertitudes : elle permet de prendre en compte simultanément plusieurs sources d'incertitudes, comme l'incertitude relative à l'initialisation du modèle ou à la variabilité spatiale des précipitations ; elle aboutit à la création d'un intervalle de prévision créé à partir d'un indice de confiance exigé par l'utilisateur final, l'intervalle de prévision s'élargit ou se rétrécit en fonction de la justesse exigée ; elle est potentiellement applicable à tout type de sortie, pas seulement aux niveaux d'eau souterraine, pour des modèles avec des horizons plus ou moins longs et pour tout type de bassin.

Néanmoins, une des limitations de la méthodologie proposée est, comme pour tout système statistique, de nécessiter beaucoup de données de qualité pour pouvoir estimer les lois statistiques nécessaires. Cette limitation s'est notamment révélée lors de l'étude des niveaux piézométriques à moyen terme sur le bassin du Blavet. Le manque de données causé par le rééchantillonnage de ces données du pas de temps journalier au pas de temps décadaire a eu

un impact sensible sur l'incertitude qui lui est associée. Ce type d'impact a été moins observé sur le bassin de la nappe de la Craie de Champagne, où le modèle neuronal utilise un historique de données plus grand en entrée. Les meilleurs résultats, tant sur le résultat de modélisation que sur leurs incertitudes associées, ont été obtenus sur le bassin du Blavet pour une prévision à trois jours des hauteurs d'eau souterraine.

Ce travail a donc fourni les résultats attendus : une méthodologie de conception de modèle fiable et robuste, assortie d'une méthode de visualisation des incertitudes, contrôlée par l'utilisateur. Ce travail ouvre également de nombreuses pistes d'améliorations possibles ; il est possible de raffiner encore les méthodologies utilisées pour obtenir des modèles plus justes, plus robustes et moins incertains. En effet, l'utilisation de prévisions de pluies en entrée du modèle aurait pu générer de meilleurs résultats. Enfin, l'utilisation de *scenarii* climatiques aurait permis de modéliser les niveaux d'eau souterraine sur du long terme.

La méthodologie de détermination des incertitudes liées au résultats de modélisation a été testée uniquement sur des modèles de type réseau de neurones, mais les calculs d'incertitudes sont transposables sur d'autres modèles. Une étude similaire sur des modèles de type à réservoir est prévue après ces travaux.

Enfin, la question de la qualité des données d'entrées a été abordée mais pas incluse dans nos exemples d'application de la méthodologie de détermination de l'incertitude de modélisation. Il serait utile d'approfondir ce point en intégrant les calculs métrologiques définis dans ce manuscrit pour chaque entrée. Pour ce faire, il faudrait dans un premier temps calculer pour chaque entrée les différentes composantes des erreurs de mesure par le biais des outils métrologiques (erreurs de répétabilité, d'étalonnage, environnementales, …). On effectue ensuite une somme quadratique de ces erreurs. La racine de cette somme correspond ainsi à l'erreur de mesure qui peut être retranchée ou ajoutée à la valeur mesurée. Ainsi, pour chaque entrée d'un modèle, il est possible d'alimenter le modèle avec des données auxquelles serait retranchée ou additionnée cette erreur de mesure générant ainsi une variabilité qui pourra être considérée dans l'incertitude calculée.

Ce travail de thèse se termine ainsi de manière très positive avec d'une part des résultats que l'entreprise pourra industrialiser de manière opérationnelle immédiatement, et d'autre part, toute une gamme de potentialités à étudier pour améliorer la fiabilité de la méthode. La collaboration entre l'entreprise AQUASYS et IMT Mines Alès, rendue possible par le financement de l'ANRT qui est ici encore remerciée, a pu déboucher sur des résultats scientifiques et opérationnels et permet d'envisager nombre de collaborations futures.

Le projet MétéEau des Nappes fait partie des pistes de collaborations futures. Pour rappel, ce projet, mené par le BRGM, propose un ensemble de services permettant le suivi et la prévision des niveaux aquifères en France.

Actuellement, six bassins tests sont à l'étude, dont font partie le bassin de la nappe du Blavet et le bassin de la nappe de la Craie de Champagne. Les travaux réalisés au cours de cette thèse sont ici une ébauche pour voir si l'estimation des incertitudes sur les résultats de modélisation est possible. Le travail post-thèse consistera à adapter les méthodologies développées sur le projet « MétéEau des Nappes ».

Bibliographie

A-D

ADES. «ADES : Recherche piézométrique sur région Bretagne.» *ADES : Recherche piézométrique sur région Bretagne.* 2020.

Adib A. «Determining water surface elevation in tidal rivers by ANN.» *Proceedings of the Institution of Civil Engineers: Water Management* 161, n° 2 (4 2008): 83-88.

AESN. «Fiche de caractérisation de la ME HG208 - Masse d'eau souterraine HG208 "Craie de Champagne Sud et Centre".» AESN, 2015, 127.

Akil N., Artigue G., Savary M., Johannet A. et Vinches M. «Quantification of Neural Network Uncertainties on the Hydrogeological Predictions by Probability Density Functions.» *IOP Conference Series: Earth and Environmental Science.* Institute of Physics Publishing, 2019. 10.

Akil N., Artigue G., Savary M., Johannet J. et Vinches M. «Uncertainty Estimation in Hydrogeological Forecasting with Neural Networks: Impact of Spatial Distribution of Rainfalls and Random Initialization of the Model.» *Water* 13, n° 1690 (2021): 29.

Alfieri L., Dottori F., Betts R., Salamon P. et Feyen L. «Multi-model projections of river flood risk in Europe under global warming.» *Climate* (MDPI AG) 6, n° 1 (2018).

André L., Cabaret O., Douez O., Saplairoles M. et Wuilleumier A. «Projet GAIA. Année 4. Avancement des volets hydrogéologique et hydrogéochimique. Rapport d'étape.» *Projet GAIA. Année 4. Avancement des volets hydrogéologique et hydrogéochimique. Rapport d'étape.* n° RP-69126-FR. 7 2019. 1-209.

Arnold J. G., Srinivasan R., Muttiah R. S. et Williams J. R. «Large Area Hydrologic Modeling and Assessment Part I: Model Development.» *Journal of the American Water Resources Association* 34, n° 1 (2 1998): 73-89.

Arsenault R. «Équifinalité, incertitude et procédures multi-modèle en prévision hydrologique aux sites non-jaugés.» École de technologie supérieure, Université du Québec, 2015, 377.

Artigue G. «Prévision des Crues Éclair par Réseaux de Neurones : Généralisation aux Bassins non Jaugés.» Université Montpellier II Sciences et Techniques du Languedoc, 2012, 318.

Ballouard C.,. Poujol M, Boulvais P. et Zeh A. «Crustal recycling and juvenile addition during lithospheric wrenching: The Pontivy-Rostrenen magmatic complex, Armorican Massif (France), Variscan belt.» *Gondwana Research* (Elsevier Inc.) 49 (9 2017): 222-247.

Banque Hydro. «La Marne à Châlons-sur-Marne - SYNTHESE.» *La Marne à Châlons-sur-Marne - SYNTHESE.* 2020.

Banque Hydro. «L'Aube à Arcis-sur-Aube - SYNTHESE.» *L'Aube à Arcis-sur-Aube - SYNTHESE.* 2020.

Banque Hydro. «Le Blavet à Languidic [Quellenec] - SYNTHESE.» *Le Blavet à Languidic [Quellenec] - SYNTHESE.* 2020. 1-2.

Barron A. R. «Approximation bounds for superpositions of a sigmoidal function.» *IEEE International Symposium on Information Theory - Proceedings.* Institute of Electrical and Electronics Engineers Inc., 1991. 85.

Barussaud E. et Blond C.. «Etude d'impact faune-flore dans le cadre du projet de parc éolien de Botsay-Glomel (Côtes-d'Armor).» Barussaud Expertise Territoriale, Vannes, France, 2015, 93.

Benaglia T. , Chauveau D., Hunter D. R. et Young D. S.. «mixtools: An R Package for Analyzing Finite Mixture Models.» 2009.

Bennett N. D., Croke B., Guariso G., Guillaume J., Hamilton S., Jakeman A. J. et Andréassian V.. «Characterising performance of environmental models.» *Environmental Modelling and Software* 40 (2 2013): 1-20.

Bernstein L., et *al.* «Changements climatiques 2007 : rapport de synthèse : un rapport du groupe d'experts intergouvernemental sur l'évolution du climat.» *Changements climatiques 2007 : rapport de synthèse : un rapport du groupe d'experts intergouvernemental sur l'évolution du climat.* Contribution des Groupes de travail I, II et III au quatrième Rapport d'évaluation du Groupe d'experts intergouvernemental sur l'évolution du climat, 2007. 103.

Beven K. J. et Kirkby M. J.. «A physically based, variable contributing area model of basin hydrology.» *Hydrological Sciences Bulletin* 24, n° 1 (1979): 43-69.

Beven K. J. et Binley A.. «GLUE: 20 years on.» *Hydrological Processes* (John Wiley and Sons Ltd) 28, n° 24 (11 2014): 5897-5918.

Beven K. J. et Binley A.. «The future of distributed models: Model calibration and uncertainty prediction.» *Hydrological Processes* 6, n° 3 (1992): 279-298.

Bhattacharjee G. P., Pandit S. N. N. et Mohan R.. «Dimensional Chains Involving Rectangular and Normal Error-Distributions.» *Technometrics* 5, n° 3 (1963): 404-406.

Blundell C., Cornebise J., Kavukcuoglu K. et Wierstra D.. «Weight Uncertainty in Neural Networks.» Édité par Google DeepMind. *Proceedings of the 32nd International Conference on Machine Learning.* Lille, France, 2015. 10.

BNPE. «Base de données des prélèvements en eau.» *Base de données des prélèvements en eau.* 2020.

Boudou M.. «Monographie des inondations d'octobre 1940 dans les Pyrénées-Orientales.» *Monographie des inondations d'octobre 1940 dans les Pyrénées-Orientales.* 11 2015. 1-35.

Bourgin F.. «Comment quantifier l'incertitude prédictive en modélisation hydrologique? Travail exploratoire sur un grand échantillon de bassins versants.» AgroParisTech, 2014, 230.

Bourguignon D.. «Événements et territoires - le coût des inondations en France : analyses spatio-temporelles des dommages assurés.» Université Paul Valéry - Montpellier III, Montpellier, France, 2014, 312.

Box G. E. P. et Jenkins G. M.. *Time Series Analysis : forecasting and control.* Holden Day. 1970.

BRGM. «Fiche Masse d'eau : FRGG010.» *Fiche Masse d'eau : FRGG010.* 2009. 1-.

Brier G. W. «Verification of forecasts expressed in terms of probability.» Édité par J Caskey. *Monthly Weather Review* 78, n° 1 (1950): 3.

Broyden C. G. «The Convergence of a Class of Double-rank Minimization Algorithms 2. The New Algorithm.» *Journal of the Institute of Mathematics and Its Applications* 6, n° 3 (1970): 222-231.

Brunet-Moret Y., Chaperon P., Lamagat J. P. et Molinier M. *Monographie hydrologique du fleuve Niger.* ORSTOM. Vol. 1. MONOGRAPHIES HYDROLOGIQUES No 8, 1986.

Bureau H. «Nouvelle méthode de jaugeage par flotteurs.» *Houille Blanche*, n° 5 (1910): 117-122.

Cao F., Jaunat, J., Vergnaud-Ayraud V., Devau N., Labasque T., Guillou A., Guillaneuf A., Hubert J., Aquilina L. et Ollivier P. «Heterogeneous behaviour of unconfined Chalk aquifers infer from combination of groundwater residence time, hydrochemistry and hydrodynamic tools.» *Journal of Hydrology*, 2020.

Carré C. etDeutsch J. C. *L'eau dans la ville: Une amie qui nous fait la guerre.* Editions de l'Aube. Electre, 2015.

Castany G. et Margat J. *Dictionnaire Français d'hydrogéologie.* Editions du BRGM. Édité par Bureau de Recherche Géologique et Minière. 1977.

Cauchy A.-L. «CALCUL DES PROBABILITÉS. — Sur les résultats moyens d'observations de même nature, et sur les résultats les plus probables.» Dans *Oeuvres complètes*, de Augustin-Louis Cauchy, 94-104. Cambridge University Press, 1853.

Cecinati F., Rico-Ramirez M. A., Heuvelink G. B. M. et Han D. «Representing radar rainfall uncertainty with ensembles basedon a time-variant geostatistical error modelling approach.» *Journal of Hydrology* 548, 391-405, 2017.

Chaubey I., Haan C. T., Grunwald S. et Salisbury J. M. «Uncertainty in the model parameters due to spatial variability of rainfall.» *Journal of Hydrology* 220 (1999): 48-61.

Colmet-Daage A., Sanchez-Gomez E., Ricci S., Llovel C., Borrell Estupina V., Quintana-Seguí P., Llasat M. C. et Servat E. «Evaluation of uncertainties in mean and extreme precipitation under climate changes for northwestern Mediterranean watersheds from high-resolution Med and Euro-CORDEX ensembles.» *Hydrology and Earth System Sciences Discussions* (Copernicus GmbH), 3 2017: 1-27.

Connor J., Atlas L. E. et Martin D. R. «Recurrent Networks and N ARMA Modeling.» 1991. 301-308.

Couto F. T., Iakunin M., Salgado R., Pinto P., Viegas T., et Pinty J.-P. «Lightning modelling for the research of forest fire ignition in Portugal.» *Atmospheric Research* (Elsevier Ltd) 242 (9 2020).

Curry H. B. «The Method of Steepest Descent for Non-linear Minimization Problems.» Édité par American Mathematical Society. *Quarterly of Applied Mathematics* 2, n° 3 (1944): 258-261.

Cybenkot G. «Mathematics of Control, Signals, and Systems Approximation by Superpositions of a Sigmoidal Function*.» 1989, 303-314.

Da Costa V., Jobard E., Marquay J., Ollagnon M., Plat B., et Radureau S. «Les services publics d'eau et d'assainissement en France Données économiques, sociales et environnementales - Sixième édition.» *Les services publics d'eau et d'assainissement en France Données économiques, sociales et environnementales - Sixième édition.* 10 2015. 108.

Daigle A., Bérubé F., Bergeron N., et Matte P. «A methodology based on Particle image velocimetry for river ice velocity measurement.» *Cold Regions Science and Technology* 89 (5 2013): 36-47.

Damant C., Austin G. L., Bellon A., et Broughton R. S. «Errors in the Thiessen technique for estimating areal rain amounts using weather radar data.» *Journal of Hydrology* (Elsevier Scientific Publishing Company) 62 (1983): 81-94.

Damien T.. «Comportement hydrodynamique des Marais de Saint-Gond (Marne, France).» Université de Reims Champagne-Ardennes, 2019, 351.

Darras T. «Prévision des crues rapides par apprentissage statistique.» Université de Montpellier, 2015, 200.

Darras T., Johannet A., Vayssade B., Kong-a-Siou L., et Pistre S. «Influence of the Initialization of Multilayer Perceptron for Flash Floods Forecasting: How Designing a Robust Model.» *International work-conference on Time Series 2014.* 2014. 13.

De Moivre A. *The Doctrine of Chances, Or, A Method of Calculating the Probability of Events in Play.* IIIe édition. Londres: Chelsea Publishing Company, 1756.

De Nijs R., et Klausen T. L. «On the expected difference between mean and median.» *Electronic Journal of Applied Statistical Analysis* 6, n° 1 (2013): 110-117.

De Plazaola J.-P., Rignols É., Tavernier J.-L., Mayo-Simbsler S., Besbes M., Betti C., Caste F., Fidani G., Gabriel C., Leduc C., Michel V., Nabet M., Naudy-Fesquet I., Olivier R., Pauly J.-N., Poulos D., Rey-Witz L., Roosz P., Tronyo J *Tableaux de l'économie française.* INSEE. 2019.

Delaunay A., Ménoret B., Debard M.-H., et Fringant S. «Crue de la Seine et de ses affluents de janvier-février 2018 Retour d'expérience.» *Crue de la Seine et de ses affluents de janvier-février 2018 Retour d'expérience.* n° 012268-01. 12 2018. 1-129.

Dempster A. P., Laird N. M., et Rubin D. B. « Maximum Likelihood from Incomplete Data Via the EM Algorithm .» *Journal of the Royal Statistical Society: Series B (Methodological)* (Wiley) 39, n° 1 (9 1977): 1-22.

Deutsch C. «DECLUS: a fortran 77 program for determining optimum spatial declustering weights.» *Computers and Geosciences* 15, n° 3 (1989): 325-332.

Diconne M.-A. «Les jaugeages des débits des cours d'eau au moulinet hydrométrique par la méthode d'intégration.» *La Houille Blanche*, n° 1 (2 1972): 35-44.

Djaman K., Baldea A. B., Sow A., Muller B., Irmak S., N'Diaye M. K., Manneh B., Moukoumbi Y. D., Futakuchi K.,et Saito K.c «Evaluation of sixteen reference evapotranspiration methods under sahelian conditions in the Senegal River Valley.» *Journal of Hydrology: Regional Studies* (Elsevier) 3 (3 2015): 139-159.

Djaman K., Tabari H., Balde A. B., Diop L., Futakuchi K., et Irmak S. «Analyses, calibration and validation of evapotranspiration models to predict grass-reference evapotranspiration in the Senegal river delta.» *Journal of Hydrology: Regional Studies* (Elsevier) 8 (12 2016): 82-94.

Dreyfus G. *Neural networks, methodology and applications.* Berlin: Springer, 2005.

DRIEE. «Le risque inondation dans le bassin de la Seine - 1910, Et demain ?» *Le risque inondation dans le bassin de la Seine - 1910, Et demain ?* 2009.

Džeroski S., et Ženko B. «Is Combining Classifiers Better than Selecting the Best One?» Ljubljana, 2004, 8.

E-H

Eakins B. W., et Sharman G. F. «Volumes of the World's Oceans from ETOPO1.» *Volumes of the World's Oceans from ETOPO1.* NOAA National Geophysical Data Center, 2010. 1.

Elsafi S. H. «Artificial Neural Networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan.» *Alexandria Engineering Journal.* Vol. 53. n° 3. Elsevier, 2014. 655-662.

Emmanuel I., Payrastre O., Andrieu H., et Zuber F. «A method for assessing the influence of rainfall spatial variability on hydrograph modeling. First case study in the Cevennes Region, southern France.» *Journal of Hydrology*, 2017.

Esposito C., et Quemart P. «Crue de juin 2010 dans le Var : Comparaison de la Zone Inondée et des Zones Inondables Connues.» Congrès SHF : «Évènements extrêmes fluviaux et maritimes», 2012. 1-8.

Fisher R. A. «On the mathematical foundations of theoretical statistics.» *Philosophical Transactions of the Royal Society* 9, n° 222 (1922): 309-368.

Fletcher R. «A New Approach to Variable Metric Algorithms.» *Computer Journal* 13, n° 3 (1970): 317-322.

Freeze R. A., et Harlan R. L. «Blueprint for a Physically-Based, Digitally-Simulated Hydrologic Response Model.» *Blueprint for a Physically-Based, Digitally-Simulated Hydrologic Response Model.* Vol. 9. © North-Holland Publishing Co, 1969. 237-258.

Funahashi K. I. «On the approximate realization of continuous mappings by neural networks.» *Neural Networks* 2, n° 3 (1989): 183-192.

Garçon R. «Modèle global pluie-débit pour la prévision et la prédétermination des crues.» *La Houille Blanche* (EDP Sciences), n° 7-8 (12 1999): 88-95.

Geman S., Bienenstock E., et Doursat R. «Neural Networks and the Bias/Variance dilemma.» *Neural Computation* 4 (1992): 1-58.

Gerard L. «Sensibilité des performances d'un modèle de prévision des crues au critère de calage.» CEMAGREF, Antony, France, 2010, 18.

Gers F. A., Jürgen Schmidhuber, et Cummins F. «Learning to forget: Continual prediction with LSTM.» *Neural Computation* (MIT Press Journals) 12, n° 10 (2000): 2451-2471.

Gilmore T. E., Birgand F., et Chapman K. W. «Source and magnitude of error in an inexpensive image-based water level measurement system.» *Journal of Hydrology* 496 (2013): 178-186.

Glaisher J. W. L. « XXXII. On a class of definite integrals .» *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* (Informa UK Limited) 42, n° 280 (10 1871): 294-302.

Glorot X., et Bengio Y. «Understanding the difficulty of training deep feedforward neural networks.» *Journal of Machine Learning Research.* 2010. 249-256.

Goldfarb D. «A Family of Variable-Metric Methods Derived by Variational Means.» *Mathematics of Computation* 24 (1970): 23-26.

Gouldby B., et Samuels P. «Language of Risk - Project définitions (Second edition).» FLOODsite - Integrated Flood Risk Analysis and Management Methodologies, 2009, 55.

Grillakis M. G. «Increase in severe and extreme soil moisture droughts for Europe under climate change.» *Science of the Total Environment* (Elsevier B.V.) 660 (4 2019): 1245-1255.

Grubbs F. E. «American Society for Quality Procedures for Detecting Outlying Observations in Samples.» *American Society for Quality Procedures for Detecting Outlying Observations in Samples.* Vol. 11. n° 1. 1969. 1-21.

Gumbel E. J. «Méthodes graphiques pour l'analyse de débits de crues.» *Revue de statistique appliquée* 5, n° 2 (1957): 77-89.

Gupta H. V., Kling H., Yilmaz K. K., et Martinez G. F.. «Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling.» *Journal of Hydrology* 377, n° 1-2 (10 2009): 80-91.

Gupta H. V., Sorooshian S., et Yapo P. O. «Status of automatic calibration for hydrologic models: Comparison with multilevel expert calibration.» *Journal of Hydrologic Engineering* (ASCE) 4, n° 2 (4 1999): 135-143.

Guyon I. «A scaling law for the validation-set training-set size ratio.» AT&T Bell Laboratories, Berkeley, 1997, 11.

Haidu I., et Nistor M. M. «Groundwater vulnerability assessment in the Grand Est region, France.» *Quaternary International* (Elsevier Ltd) 547 (5 2020): 86-100.

Hartmann D. L., et al. «Observations: Atmosphere and Surface.» Chap. 2 dans *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.*, de Dennis L. Hartmann, et al., édité par Cambridge University Press, 159-254. 2013.

He S., Guo S., Liu E., Yin J., Chen K., et Wu X. «Uncertainty analysis of hydrological multi-model ensembles based on CBP-BMA method.» *Hydrology Research* (IWA Publishing) 49, n° 5 (10 2018): 1636-1651.

Heddam S. *Fuzzy neural network (EFuNN) for modelling dissolved oxygen concentration (DO).* Vol. 113, chez *Intelligent Systems Reference Library*, de Salim Heddam, 231-253. Springer Science and Business Media Deutschland GmbH, 2017.

Hiez G. «L'homogénéité des données pluviométriques.» *Ser. Hydrol.* 14, n° 2 (1977): 129-172.

Hochreiter S., et Schmidhuber J. «Long Short-Term Memory.» *Neural Computation* 9, n° 8 (1997): 32.

Hornik K. «Approximation capabilities of multilayer feedforward networks.» *Neural Networks* 4, n° 2 (1991): 251-257.

Hornik K, Stinchombe M., et White H. «Multilayer Feedforward Networks are Universal Approximators.» *Neural Networks* 2 (1989): 359-366.

Houet T. «Occupation des sols et gestion de l'eau : Modélisation prospective en paysage agricole fragmenté (Application au SAGE du Blavet).» Université de Rennes II, 2006, 368.

Huber P. J. «Robust estimation of a location parameter.» *The Annals of Mathematical Statistics* 35, n° 1 (1964): 73-101.

Huber P. J. *Robust statistics.* Wiley, 1981. 317.

Hubert P. , Tchiguirinskaia I., Bendjoudi H., Schertzer D., et Lovejoy S. «Multifractal modeling of the Blavet river discharge at Guerledan.» *Proceedings of the Third Inter-Celtic Colloquium on Hydrology and Management of Water Resources.* Galway, 2002. 131-138.

Huet P., Martin X., Prime J.-L., Foin P., Laurain C., et Cannard P.. «Retour d'expériences des crues de septembre 2002 dans les départements du Gard, de l'Hérault, du Vaucluse, des Bouches-du-Rhône, de l'Ardèche et de la Drôme.» *Retour d'expériences des crues de septembre 2002 dans les départements du Gard, de l'Hérault, du Vaucluse, des Bouches-du-Rhône, de l'Ardèche et de la Drôme.* 6 2003. 1-133.

I-L

INSEE. «Bilan démographique 2019.» *Bilan démographique 2019.* 2020. 20.

INSEE. *Les résultats des recensements de la population.* 28 12 2020. https://www.insee.fr/fr/information/2008354 (accès le 01 2021).

Isukappali S. S. «Uncertainty analysis of transport-transformation models.» State University of New Jersey, New Brunswick, Etats-Unis, 1999, 200.

JCGM. «Évaluation des données de mesure-Guide pour l'expression de l'incertitude de mesure Evaluation of measurement data-Guide to the expression of uncertainty in measurement.» 2008, 121.

Johannet A. «Modélisation par apprentissage statistique des systèmes naturels, ou en interaction avec un environnement naturel. Applications aux karsts, crues éclair et en robotique.» Université Pierre et Marie Curie, 2011, 82.

Johannet A., Taver V., Kong-A-Siou L., Borrell Estupina V., Pistre S., Mangin A., Vayssade B., Vinches M., et Bertin D. «Les réseaux de neurones artificiels pour la modélisation hydrodynamique des aquifères karstiques : présentation générale et illustrations.» *Karstologia : revue de karstologie et de spéléologie physique* (PERSEE Program) 60, n° 1 (2012): 45-59.

Juran J. M. «Universals in Management Planning and Controlling Two Examples from World War II.» 1954, 13.

Kabir H. M. D., Khosravi A., Hosen M. A., et Nahavandi S. «Neural Network-Based Uncertainty Quantification: A Survey of Methodologies and Applications.» *IEEE Access.* Vol. 6. Institute of Electrical and Electronics Engineers Inc., 3 6 2018. 36218-36234.

Kavetski D., Kuczera G., et Franks S. W. «Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory.» *Water Resources Research* 42, n° 3 (3 2006).

Khosravi A., Nahavandi S., et Creighton D. «A prediction interval-based approach to determine optimal structures of neural network metamodels.» *Expert Systems with Applications* 37, n° 3 (3 2010): 2377-2387.

Kingma D. P., et Lei Ba J. «Adam: A Method for Stochastic Optimization.» *International Conference on Learning Representations 2015.* 2014. 15.

Kingston D. G., Massei N., Dieppois B., Hannah D., Hartmann A., Lavers D., Vidal J.-P. «Moving beyond the catchment scale: Value and opportunities in large-scale hydrology to understand our changing world.» *Hydrological Processes.* John Wiley and Sons Ltd, 2020.

Kitanidis P. K., et Bras R. L. «Real-time forecasting with a conceptual hydrologic model: 2. Applications and results.» *Water Resources Research* 16, n° 6 (1980): 1034-1044.

Kloppmann W., Dever L., et Edmunds W. M. «Residence time of Chalk groundwaters in the Paris Basin and the North German Basin: A geochemical approach.» *Applied Geochemistry*, 1998.

Kong-A-Siou L. «Modélisation des crues de bassins karstiques par réseaux de neurones. Cas du bassin du Lez (France).» Université de Montpellier II, 2011, 232.

Kong-A-Siou L., Johannet A., Borrell Estupina V., et Pistre S. «Neural networks for karst groundwater management: case of the Lez spring (Southern France).» *Environmental Earth Sciences* (Springer Verlag) 74, n° 12 (12 2015): 7617-7632.

Kong-A-Siou L., Johannet A., Borrell V., et Pistre S. «Complexity selection of a neural network model for karst flood forecasting: The case of the Lez Basin (southern France).» *Journal of Hydrology* 403, n° 3-4 (6 2011): 367-380.

Kong-A-Siou L., Fleury P., Johannet A., Borrell Estupina V., Pistre S., et Dörfliger N. «Performance and complementarity of two systemic models (reservoir and neural networks) used to simulate spring discharge and piezometry for a karst aquifer.» *Journal of Hydrology* (Elsevier) 519, n° PD (11 2014): 3178-3192.

Kopp R. E., DeConto R. M. , Bader D. A., Hay C. C. , Horton R. M., Kulp S., Oppenheimer M., Pollard D., et Strauss D. H. «Evolving Understanding of Antarctic Ice-Sheet Physics and Ambiguity in Probabilistic Sea-Level Projections.» *Earth's Future* (John Wiley and Sons Inc) 5, n° 12 (12 2017): 1217-1233.

Kouki S. «Apprentissage multi-objectifs de réseaux de neurones pour la prévision des crues.» Université de Laval, Laval, Canada, 2010, 59.

Kovalev V., Kalinovsky A., et Kovalev S. «Deep Learning with Theano, Torch, Caffe, TensorFlow, and Deeplearning4J: Which One Is the Best in Speed and Accuracy?» *XIII International Conference on Pattern Recognition and Information Processing .* Minsk, Belarus, 2016. 99-103.

Kuczera G. «Improved Parameter Inference in Catchment Models 1. Evaluating Parameter Uncertainty.» Monash University, 1983, 1151-1162.

Kuczera G., et Parent E. «Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm.» *Journal of Hydrology* (Elsevier Sci B.V.) 211, n° 1-4 (11 1998): 69-85.

Kumaraswamy P. «A generalized probability density function for double-bounded random processes.» *Journal of Hydrology* 46, n° 1-2 (1980): 79-88.

Laimé M.. *Inondations : une faillite française.* 2015. http://www.eauxglacees.com/Inondations-une-faillite-francaise, 15 10 2019.

Lajournade C., et Lalanne-Berdouticq G. «Les crues torrentielles dans les Pyrénées occidentales : contexte pluviométrique, géomorphologique et anthropique des événements extrêmes/Torrential floods in the western Pyrenees: rainfall, geomorphic and anthropic conditions for extrem events.» *Géomorphologie : relief, processus, environnement* (PERSEE Program) 3, n° 2 (1997): 111-120.

Lallahem S. «STRUCTURE ET MODELISATION HYDRODYNAMIQUE DES EAUX SOUTERRAINES : Application à l'aquifère crayeux de la bordure nord du bassin de Paris.» Polytech Lille, 2002, 242.

Laplace P-S. «Mémoire sur la probabilité des causes par les évènements.» Chap. 6 dans *Œuvres complètes de Laplace*, de Pierre-Simon Laplace, édité par Mémoires de l'Academie Royale des Sciences., 621–656. Paris, France: Divers Savan, 1774.

Larson K. L., White D. D., Gober P., et Wutich A. «Decision-making under uncertainty for water sustainability and urban climate change adaptation.» *Sustainability (Switzerland)* (MDPI AG) 7, n° 11 (2015): 14761-14784.

Le Lay M., et Saulnier G. M. «Exploring the signature of climate and landscape spatial variabilities in flash flood events: Case of the 8-9 September 2002 Cévennes-Vivarais catastrophic event.» *Geophysical Research Letters* 34, n° 13 (7 2007): 5.

Lee D. H., et Kang D. S. «The Application of the Artificial Neural Network Ensemble Model for Simulating Streamflow.» *Procedia Engineering.* Elsevier Ltd, 2016. 1217-1224.

Lejeune O., et Devos A. «Caractéristiques et spécificités du karst de la Montagne de Reims ; comparaison avec les autres karsts de contact lithostratigraphique de l'est du Bassin parisien.» Édité par Association des géographes de l'Est. *Revue Géographique de l'Est* 41, n° 4 (2001): 15.

Les Landes A. A., Aquilina L., Davy P., Vergnaud-Ayraud V., et Le Carlier C. «Timescales of regional circulation of saline fluids in continental crystalline rock aquifers (Armorican Massif, western France).» *Hydrology and Earth System Sciences* (Copernicus GmbH) 19, n° 3 (3 2015): 1413-1426.

Levenberg K. «A method for the solution of certain non-linear problems in least squares.» *Quarterly of Applied Mathematics* (American Mathematical Society (AMS)) 2, n° 2 (7 1944): 164-168.

Li S., Bush R. T., Mao R., Xiong L., et Ye C. «Extreme drought causes distinct water acidification and eutrophication in the Lower Lakes (Lakes Alexandrina and Albert), Australia.» *Journal of Hydrology* (Elsevier B.V.) 544 (1 2017): 133-146.

Li W., Kiaghadi A., et Dawson C. N. «High Temporal Resolution Rainfall Runoff Modelling Using Long-Short-Term-Memory (LSTM) Networks.» 2 2020.

Liang Z., Wang D., Guo Y., Zhang Y., et Dai R. «Application of bayesian model averaging approach to multimodel ensemble hydrologic forecasting.» *Journal of Hydrologic Engineering* 18, n° 11 (2013): 1425-1436.

Lin G. F., et Chen L. H. «A non-linear rainfall-runoff model using radial basis function network.» *Journal of Hydrology* 289, n° 1-4 (4 2004): 1-8.

Liu Z., et Todini E. «Towards a comprehensive physically-based rainfall-runoff model.» *Towards a comprehensive physically-based rainfall-runoff model.* Vol. 6. n° 5. 2002. 859-881.

Lucassou F., Schroëtter J.-M., Baptiste J., et Coppo N. «Sensibilité des aquifères côtiers bretons aux intrusions salines - Rapport final.» *Sensibilité des aquifères côtiers bretons aux intrusions salines - Rapport final.* 7 2019. 1-230.

Lucassou F., et Mougin B. «Essai d'élaboration d'indicateurs piézométriques pour la gestion quantitative AEP dans le département des Côtes d'Armor - Rapport final.» BRGM - BRGM/RP-64123-FR, Rennes, France, 2015, 157.

M-P

Maca P., Pech P;, et Pavlasek J. «Comparing the selected transfer functions and local optimization methods for neural network flood runoff forecast.» *Mathematical Problems in Engineering* (Hindawi Publishing Corporation) 2014 (2014).

Mangin A. «Pour une meilleure connaissance des systèmes hydrologiques à partir des analyses corrélatoire et spectrale.» *Journal of Hydrology* 67 (1984): 25-43.

Marchandise A. «Modélisation hydrologique distribuée sur le Gardon d'Anduze; étude comparative de différents modèles pluie-débit, extrapolation de la normale à l'extrême et tests d'hypothèses sur les processus hydrologiques.» Université de Montpellier II - Sciences et Techniques du Languedoc, 2007, 214.

Mardhel V, et Allier D. *Indice de développement et de persistance des réseaux (IDPR).* 2010. https://www.esrifrance.fr/iso_album/p30_brgm.pdf.

Marquardt D. W. «An Algorithm for Least-Squares Estimation of Nonlinear Parameters.» *Journal of the Society for Industrial and Applied Mathematics* (Society for Industrial & Applied Mathematics (SIAM)) 11, n° 2 (6 1963): 431-441.

Martin J.-C., Rouxel-David E., et Batkowski D. «Elaboration d'une règle de gestion volumique de la ressource en eau du bassin de la Barbuise (Aube) - Rapport final.» BRGM - BRGM/RP-53178-FR, Reims, France, 2004, 93.

Martin-Ortega J., González-Eguino M., et Markandya A. «The costs of drought: The 2007/2008 case of Barcelona.» *Water Policy* 14, n° 3 (2012): 539-560.

Mathevet T., Michel C., Andréassian V., et Perrin C. «A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins.» *IAHS-AISH Publication*, n° 307 (2006): 211-219.

Mazzilli N. «Sensibilité et incertitude de modélisation sur les bassins versants à forte composante karstique.» Université de Montpellier, Montpellier, France, 2011, 234.

McCulloch W. S., et Pitts W H. «A logical calculus of the ideas immanent in nervous activity.» *Bulletin of Mathematical Biophysics* (Taylor and Francis) 5 (1943): 115-133.

McKay M. D., Beckman R. J., et Conover W. J. «A comparison of three methods for selecting values of input variables in the analysis of output from a computer code.» *Technometrics* 42, n° 1 (2000): 55-61.

McMillan H., Seibert J., Petersen Overleir A., Lang M., White P., Snelder T., Rutherford K., Krueger T., Mason R., et Kiang J. «How uncertainty analysis of streamflow data can reduce costs and promote robust decisions in water management applications.» *Water Resources Research.* Vol. 53. n° 7. Blackwell Publishing Ltd, 1 7 2017. 5220-5228.

Météo-France. «Le réchauffement climatique observé à l'échelle du globe et en France.» *Le réchauffement climatique observé à l'échelle du globe et en France.* 2015.

Metropolis N., et Ulam S. «The Monte Carlo Method.» *Journal of the American Statistical Association* 44, n° 247 (1949): 335-341.

Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., et Teller E.. «Equation of state calculations by fast computing machines.» *The Journal of Chemical Physics* 21, n° 6 (1953): 1087-1092.

Migaud D. «Les enseignements des inondations de 2010 sur le littoral atlantique (Xynthia) et dans le Var, discours de Didier Migaud, Premier président de la cours des Comptes.» *Les enseignements des inondations de 2010 sur le littoral atlantique (Xynthia) et dans le Var, discours de Didier Migaud, Premier président de la cours des Comptes.* Cours des Comptes, 7 2012. 1-8.

Molina M. G. «Probabilidad e incertidumbre, una aproximación desde la historia del pensamiento.» *Revista de Economia Institucional*, 2016: 101-122.

Monteith J. L. «Evaporation and environment.» *Symposia of the Society for Experimental Biology* 19 (1965): 205-234.

Morell M. «Influence de la topographie et de l'exposition aux vents dominants sur les mesures de précipitation.» *Influence de la topographie et de l'exposition aux vents dominants sur les mesures de précipitation.* 9 1986. 1-22.

Moriasi D. N., Arnold J. G., Van Liew M. W., Bingner R. L., Daren Harmel R., et Veith T. L. «Model evaluation guidelines for systematic quantification of accuracy in watershed simulations.» *Transactions of the ASABE* 50, n° 3 (5 2007): 885-900.

Morin G., et Paquet P. «Modèle hydrologique CEQUEAU.» *Modèle hydrologique CEQUEAU.* n° R000926. INRS-ETE, 2007. 1-458.

Motra H. B., Dimmig-Osbürg A., et Hildebrand J. «Influence of measurement uncertainties on results of creep prediction of concrete under cyclic loading.» Édité par J.G.M. Van Mier, G. Ruiz, C. Andrade, R.C. Yu et X.X. Zhang. *VIII International Conference on Fracture Mechanics of Concrete and Concrete Structures (FraMCoS-8).* New York, 2013. 805-814.

Mougin B. «Atlas du réseau piézométrique de Bretagne SILURES Suivi (52 ouvrages).» BRGM - BRGM/RP-54576-FR, Rennes, France, 2006, 59.

Mougin B. «SILURES Suivi ,Gestion du réseau de surveillance piézométrique régional - Annuaire 2005.» BRGM - BRGM/RP-54288-FR, Rennes, France, 2005, 27.

Mougin B., Carn A., Debeglia N., Perrin J., et Thomas E.. «SILURES Bretagne (Système d'Information pour la Localisation et l'Utilisation des Ressources en Eaux Souterraines). Rapport d'avancement de l'année 2.» BRGM - BRGM/RP-52825-FR, Rennes, France, 2004, 62.

Mougin B., Nicolas J., Vigier Y., Bessière H. ET Loigerot S. «" MétéEau Nappes " : un site Internet contenant des services utiles à la gestion des étiages.» , UNESCO-SHF “ Sécheresses, étiages et déficits en eau ”, 2019. 1-2.

Moussa R. «When monstrosity can be beautiful while normality can be ugly: assessing the performance of event-based flood models.» *Hydrological Sciences Journal* (Informa UK Limited) 55, n° 6 (8 2010): 1074-1084.

Moussa R., Chahinian N., et Bocquillon C. «Distributed hydrological modelling of a Mediterranean mountainous catchment - Model construction and multi-site validation.» *Journal of Hydrology* (Elsevier) 337, n° 1-2 (4 2007): 35-51.

Mugunthan P., et Shoemaker C. A. «Assessing the impacts of parameter uncertainty for computationally expensive groundwater models.» *Water Resources Research* 42, n° 10 (10 2006).

Museum National d'Histoire Natuelle. «FR2112012 - Marigny, Superbe, vallée de l'Aube.» Commission Européenne, 2005, 14.

Museum National d'Histoire Naturelle. «FR2100257 - Savart du camp militaire de Mailly-le-Camp.» Commission Européenne, 1995, 8.

Museum National d'Histoire Naturelle. «FR2100259 - Savart du camp militaire de Suippes.» Commission Européenne, 1995, 10.

Museum National d'Histoire Naturelle. «FR2100283 - Le Marais de Saint-Gond.» Commission Européenne, 1995, 11.

Museum National d'Histoire Naturelle. «FR2100285 - Marais de la Superbe.» Commission Européenne, 1995, 9.

Museum National d'Histoire Naturelle. «FR2100297 - Prairies et bois alluviaux de la basse vallée alluviale de l'Aube.» Commission Européenne, 1995, 11.

Museum National d'Histoire Naturelle. «FR2100312 - Massif forestier de la Montagne de Reims (versant sud) et étangs associés.» Commission Européenne, 1995, 10.

Museum National d'Histoire Naturelle. «FR2100314 - Massif forestier d'Epernay et étangs associés.» Commission Européenne, 1995, 7.

Museum National d'Histoire Naturelle. «FR5300007 - Têtes de bassin du Blavet et de l'Hyères.» Commission Européenne, 1995, 10.

Museum National d'Histoire Naturelle. «FR5310094 - Rade de Lorient.» Commission Européenne, 1993, 8.

Mylne K. R., Evans R. E., et Clark R. T. «Multi-model multi-analysis ensembles in quasi-operational medium-range forecasting.» *Meteorol. Soc* 128 (2002): 361-384.

Nash J. E., et Sutcliffe J. V. «River flow forecasting through, Part I. A conceptual models discussion of principles.» *Journal of Hydrology* 10 (1970): 282-290.

Ndiaye P. M., Bodian A., Diop L., et Djaman K. «Évaluation de vingt méthodes d'estimation de l'évapotranspiration journalière de référence au Burkina Faso.» *Physio-Géo* (OpenEdition) 11 (1 2017): 129-146.

Nérot B. «Réseaux de neurones : présentation de la méthode LSTM - Mise en œuvre sur un cas de prévision hydrologique.» IMT Mines Alès, Alès, France, 2018, 59.

Nerrand O., Roussel-Ragot P., Personnaz L., Dreyfus G., et Marcos S. «Neural Networks and Nonlinear Adaptive Filtering: Unifying Concepts and New Algorithms.» *Neural Computation* (MIT Press - Journals) 5, n° 2 (3 1993): 165-199.

Nistor M. M. «Groundwater vulnerability in Europe under climate change.» *Quaternary International* (Elsevier Ltd) 547 (5 2020): 185-196.

Observatoire des Inégalités, . «L'accès à l'eau potable dans le monde.» *L'accès à l'eau potable dans le monde.* 28 06 2018.

ONU. «How certain are the United Nations global population projections?» *How certain are the United Nations global population projections?* 12 2019. 4.

Oudin L. «Recherche d'un modèle d'évapotranspiration potentielle pertinent comme entrée d'un modèle pluie-débit global.» ENGREF, 2004, 495.

Paz S., Carmel Y., Jahshan F., et Shoshany M. «Post-fire analysis of pre-fire mapping of fire-risk: A recent case study from Mt. Carmel (Israel).» *Forest Ecology and Management* 262, n° 7 (10 2011): 1184-1188.

Pearson K. *Mathematical Contributions to the Theory of Evolution. - III. Regression, Heredity, and Panmixia.* Vol. 187, chap. 7 dans *On the Matheatical Theory of Evolution*, de Karl Pearson, 253-318. Londres: Philosophical Transactions of the Royal Society of London, 1896.

Peña-Angulo D., Vicente-Serrano S. M., Domínguez-Castro F., Murphy C., Reig F., Tramblay Y., Trigo R. M., Luna M. Y., Turco M., Noguera I., Aznárez-Balta M., García-Herrera R., Tomas-Burguera M., et El Kenawy A. «Long-term precipitation in Southwestern Europe reveals no clear trend attributable to anthropogenic forcing.» *Environmental Research Letters* (IOP Publishing Ltd) 15, n° 9 (9 2020).

Perreault Levasseur L., Yashar D. Hezaveh, et Risa H. Wechsler. «Uncertainties in Parameters Estimated with Neural Networks: Application to Strong Gravitational Lensing.» *The Astrophysical Journal* (American Astronomical Society) 850, n° 1 (11 2017): L7.

Perrin C., Michel C., et Andréassian V. «Improvement of a parsimonious model for streamflow simulation.» *Journal of Hydrology* (Elsevier) 279, n° 1-4 (8 2003): 275-289.

Pinault J.-L., Allier D., et Chabart M. «Prévision des volumes d'eau exploitables de 10 bassins versants en champagne crayeuse.» BRGM, Reims, 2006, 110.

Poquet G., et Maresca B. «La consommation d'eau baisse dans les grandes villes européennes.» *La consommation d'eau baisse dans les grandes villes européennes.* 4 2006. 4.

Pushpalatha R.. «Low-flow simulation and forecasting on French river basins: a hydrological modelling approach.» AgroParisTech, 2013, 230.

Putot E., Verjus P., et Vernoux J.-F. «Qualification des piézomètres du réseau de bassin Seine Normandie en 2005.» BRGM, Massy, 2006, 96.

Q-U

Rajczak J., et Schär C. «Projections of Future Precipitation Extremes Over Europe: A Multimodel Assessment of Climate Simulations.» *Journal of Geophysical Research: Atmospheres* (Blackwell Publishing Ltd) 122, n° 20 (10 2017): 10,773-10,800.

Raschka S. «Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning.» University of Wisconsin, Madison, 2018, 49.

Rasmussen T. C., et Crawford L. A. «Identifying and removing barometric pressure effects in confined and unconfined aquifers.» *Ground Water* (Ground Water Publ Co) 35, n° 3 (1997): 502-511.

Ribes A., Thao S., Vautard R., Dubuisson B., Somot S., Colin J., Planton S., et Soubeyroux J.-M. «Observed increase in extreme daily rainfall in the French Mediterranean.» *Climate Dynamics* (Springer Verlag) 52, n° 1-2 (1 2019): 1095-1114.

Riboust P. «De la neige au débit : de l'intérêt d'une meilleure contrainte et représentation de la neige dans les modèles.» Université de la Sorbonne, Paris, 2019, 323.

Rinne H. «Location-Scale Distributions Linear Estimation and Probability Plotting Using MATLAB.» Justus Liebig University, Giessen, Allemagne, 2010, 260.

Ritzema H. P. «Frequency and regression analysis of hydrologic data.» Chap. 6 dans *Drainage Principles and Applications*, de H. P. Ritzema, 47. Wageningen, Netherlands: International Institute for Land Reclamation and Improvement, 1994.

Roche P. A., et Thiéry D. «Simulation Globale de Bassins Hydrologiques - Introduction à la Modélisation et Description du Modèle GARDENIA.» *Simulation Globale de Bassins Hydrologiques - Introduction à la Modélisation et Description du Modèle GARDENIA.* n° 84 SGN 337. 1984.

Rodgers C. D. «Characterization and Error Analysis of Profiles Retrieved From Remote Sounding Measurements.» *Journal of Geophysical Research* 95, n° D5 (1990): 5587–5595.

Rogers W. H., et Tukey J. W. «Understanding some long-tailed symmetrical distributions.» *Statistica Neerlandica* 26, n° 3 (1972): 211-226.

Rosenblatt F. «The perceptron: A probabilistic model for information storage and organization in the brain.» *Psychological Review* 65, n° 6 (11 1958): 386-408.

Rumelhart D., Hinton G., et Williams R. J. «Learning representations by back-propagating errors.» *Nature* 323, n° 6088 (10 1986): 533-536.

SANDRE. «Le Blavet.» *Le Blavet.* 7 2011.

Savary M. «Prévision de la turbidité par apprentissage statistique : application au captage AEP d'Yport (Normandie).» Université de Rouen Normandie, La Havre, France, 2018, 276.

Savelyeva E., Utkin S., Kazakov S., et Demyanov V. «Modeling Spatial Uncertainty for Locally Uncertain Data.» De Elena Savelyeva, Sergey Utkin, Sergey Kazakov et Vasyliy Demyanov, 295-306. 2010.

Schaefli B., et Gupta H. V. «Do Nash values have value?» *Hydrological Processes.* Vol. 21. n° 15. 15 7 2007. 2075-2080.

SDES. «DATALAB - Eau et milieux aquatiques - Les chiffres clés.» *DATALAB - Eau et milieux aquatiques - Les chiffres clés.* 12 2020. 1-128.

Sebbar A. «Etude de la variabilité et de l'évolution de la pluviométrie au Maroc (1935-2005): Réactualisation de la carte des précipitations.» Université Hassan II Mohammedia, Casablanca, 2013, 188.

Secchi P, Zio E., et Di Maio F. «Quantifying uncertainties in the estimation of safety parameters by using bootstrapped artificial neural networks.» *Annals of Nuclear Energy* 35, n° 12 (12 2008): 2338-2350.

Seibert J. «On the need for benchmarks in hydrological modelling.» *Hydrological Processes* (Wiley) 15, n° 6 (4 2001): 1063-1064.

Shanno D. F. «Conditioning of Quasi-Newton Methods for Function Minimization.» *Source: Mathematics of Computation* 24, n° 111 (1970): 647-656.

Shrestha D. L. «Uncertainty Analysis in Rainfall-Runoff Modelling: Application of Machine Learning Techniques.» Delft University of Technology, 2009, 222.

Shrestha D. L., et Solomatine D. P. «Machine learning approaches for estimation of prediction interval for the model output.» *Neural Networks* 19, n° 2 (3 2006): 225-235.

Sinha R. K. «A Thought on Exotic Statistical Distributions.» *International Journal of Mathematical and Computational Sciences* 6, n° 1 (2012): 49-52.

Sjöberg Jonas, Zhang Q., Ljung L., Benveniste A., Delyon B., Glorennec P.-Y., Hjalmarsson H., et Juditsky A. «Nonlinear Black-box Modeling in System Identification: a Unified Overview.» 1995, 1691-1724.

SOeS. «Commissariat général au développement durable.» *Commissariat général au développement durable.* 2011.

Solomatine D. P., et Shrestha D. L. «A novel method to estimate model uncertainty using machine learning techniques.» *Water Resources Research* 45, n° 1 (2009).

Spinoni J., Naumann G., Vogt J., et Barbosa P. *Meteorological droughts in Europe events and impacts : past trends and future projections.* Publications Office of the European Union, 2016.

Stephens M. A. *An appreciation of Kolmogorov's 1933 paper.* Stanford University, Stanford, California: Office of Naval Research, 1992, 19.

Stigler S. M. «An Historical Note on the Cauchy Distribution.» 1974, 375.

Stollsteiner P. «Connaissance des ressources réellement disponibles sur l'ensemble des bassins versants crayeux.» BRGM, Reims, 2013, 158.

Stone M. «Cross-Validatory Choice and Assessment of Statistical Predictions (With Discussion).» *Journal of the Royal Statistical Society: Series B (Methodological)* (Wiley) 38, n° 1 (9 1974): 111-147.

Syndicat de la Vallée du Blavet. «Usage de la ressource en eau.» *Usage de la ressource en eau.* 2014.

Taillade F. «Notions de métrologie.» 2005, 50.

Tauro F., Olivieri G., Petroselli A., Porfiri M., et Grimaldi S. «Flow monitoring with a camera: a case study on a flood event in the Tiber River.» *Environmental Monitoring and Assessment* (Springer International Publishing) 188, n° 2 (2 2016): 1-11.

Taver V., Johannet A., Borrell Estupina V., et Pistre S. « Feed-forward vs recurrent neural network models for non-stationarity modelling using data assimilation and adaptivity .» *Hydrological Sciences Journal* (Informa UK Limited) 60, n° 7-8 (8 2015): 1242-1265.

Thiessen A. H. «Climatological data for July 1911, Precipitations averages for large areas.» *Monthly Weather Review*, 1911: 1082-1084.

Thornthwaite C. W. «An Approach toward a Rational Classification of Climate.» *Geographical Review* (JSTOR) 38, n° 1 (1 1948): 55.

Thuon Y., et Julio C. «Evaluation de la ressource dans le cadre de la révision du schéma départemental des carrières : département de l'Aube.» BRGM, Reims, 2010, 73.

Thyer M., Renard B., Kavetski D., Kuczera G., Franks S. W., et Srikanthan S. «Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis.» *Water Resources Research* 45, n° 12 (12 2009).

Touboul J. «Etude de l'interaction entre le vent et les vagues scélérates.» *Etude de l'interaction entre le vent et les vagues scélérates.* 11 2007. 1-187.

Toukourou M. S. «Application de l'apprentissage artificiel à la prévision des crues éclair.» MinesParisTech, 2009, 177.

Toukourou M. S., Johannet A., Dreyfus G, et Ayral P. A. «Rainfall-runoff modeling of flash floods in the absence of rainfall forecasts: The case of "cévenol flash floods".» *Applied Intelligence* 35, n° 2 (10 2011): 178-189.

Troutman B. M. «Runoff Prediction Errors and Bias in Parameter Estimation Induced by Spatial Variability of Precipitation.» 1983, 791-810.

V-Z

Vallam P., Qin X. S., et Yu J. J. «Uncertainty Quantification of Hydrologic Model.» *APCBEE Procedia* (Elsevier BV) 10 (2014): 219-223.

van der Schrier G., Briffa K. R., Jones P. D., et Osborn T. J. «Summer Moisture Variability across Europe.» *Journal of Climate* 19 (6 2006): 2818-2834.

Vanlier J., Tiemann C. A., Hilbers P. A. J., et van Riel N. A. W. «Parameter uncertainty in biochemical models described by ordinary differential equations.» *Mathematical Biosciences* 246, n° 2 (12 2013): 305-314.

Vergnes J.-P., Roux N., Habets F., Ackerer P., Amraoui N., Besson F., Caballero Y., Courtois Q., De Dreuzy J.-R., Etchervers P., Gallois N., Leroux D. J., Longuevergne L., Le Moigne P., Morel T., Munier S., Regimbeau F., Thiéry D., Viennot P. «The AquiFR hydrometeorological modelling platform as a tool for improving groundwater resource monitoring over France: evaluation over a 60-year period.» *Hydrology and Earth System Sciences* 24, 2020, 633-654.

Verhulst P. F. «Recherches mathématiques sur la loi d'accroissement de la population.» Chap. 18 dans *Nouveaux mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles*, de P. F. Verhulst, 1-38. Bruxelles, Belgique, 1845.

Vicente-Serrano S. M., Domínguez-Castro F. , Murphy C., Hannaford J., Reig F. , Peña-Angulo D., Tramblay Y., Trigo R. M., Mac Donald N., Luna M. Y., Mc Carthy M., Van der Schrier G., Turco M., Camuffo D., Noguera I., García-Herrera R. , Becherini F., Della Valle A., Tomas-Burguera M., et El Kenawy A. «Long-term variability and trends in meteorological droughts in Western Europe (1851–2018).» *International Journal of Climatology* (John Wiley and Sons Ltd) 41, n° S1 (1 2020): E690-E717.

Visserot Y. «Intégration de l'information hydrométéorologique spatialisée pour la modélisation et la prévision des crues éclair à l'aide des réseaux de neurones.» Mémoire de stage de Master Professionnel Gestion des Catastrophes et des Risques Naturels, Département de Géographie UFR, Université Paul Valéry - Montpellier III, Montpellier, 2012, 108.

Vold E. T. «Epistemic modality markers in research articles: a cross-linguistic and cross-disciplinary study.» *International Journal of Applied Linguistics* 16, n° 1 (2006): 61-82.

Vollset S. E., Goren E., Yuan C.-W., Cao J., Smith A. E., Hsiao T., Bisignano C., Azhar G. S., Castro E., Chalek J., Dolgert A. J., Frank T., Fukutaki K., Hay S. I., Lozano R., Mokdad A. H., Nandakumar V., Pierce M., Pletcher M., Robalik T., Steuben K. M., Wunrow H. Y., Zlavog B. S., et Murray C. J. L. «Fertility, mortality, migration, and population scenarios for 195 countries and territories from 2017 to 2100: a forecasting analysis for the Global Burden of Disease Study.» *The Lancet*, 2020.

Wagena M. B., Bhatt G., Buell E., Sommerlot A. R., Fuka D. R., et Easton Z. M. «Quantifying model uncertainty using Bayesian multi-model ensembles.» *Environmental Modelling and Software* (Elsevier Ltd), 7 2019: 89-99.

Wang N., Chu X., et Zhang X. «Functionalities of surface depressions in runoff routing and hydrologic connectivity modeling.» *Journal of Hydrology* (Elsevier B.V.) 593 (2 2021).

Waseem M., Mani N., Andiego G., et Usman M. «A review of criteria of fit for hydrological models.» *International Research Journal of Engineering and Technology* 4, n° 11 (2008): 1765-1772.

Willink R. *Measurement uncertainty and probability.* Vol. 9781107021938. Cambridge University Press, 2009.

Willmott C. J. «On the validation of models.» *Physical Geography* 2, n° 2 (1981): 184-194.

Wuilleumier A., Cabaret O., et Saltel M. «Projet GAIA. Année 3. Avancement du volet hydrogéologique. Rapport final.» *Projet GAIA. Année 3. Avancement du volet hydrogéologique. Rapport final.* n° RP-66849-FR. 11 2017. 1-192.

Zhang J., Zhu Y., Zhang X., Ye M., et Yang J. «Developing a Long Short-Term Memory (LSTM) based model for predicting water table depth in agricultural areas.» *Journal of Hydrology* 561 (6 2018): 918-929.

Annexes

Annexe I : Liste des lois de probabilités testées

#	Nom de la loi	Statut	Commentaires
1	Loi Normale	*Validé*	-
2	Loi de Gumbel	*Validé*	-
3	Loi de Laplace	*Validé*	-
4	Loi du Cosinus surélevée	*Validé*	-
5	Loi de Cauchy	*Validé*	-
6	Loi Logistique	*Validé*	-
7	Loi de Slash	*Validé*	-
8	Loi de Bhattacharjee	*Validé*	-
9	Loi de Huber	*Validé*	-
10	Loi de Kumaraswamy	*En partie validé*	Ne fonctionne que sur certains bassins.
11	Loi Uniforme	*En partie validé*	Fonctionne sur des bassins avec peu de fluctuations de niveaux.
12	Loi Sécante Hyperbolique	*En partie validé*	Ne fonctionne que sur certains bassins.
13	Loi de Gumbel Inverse	*Testé*	La loi de Gumbel est mieux adaptée, celle-ci a donc été écartée.
14	Loi de Pareto	*Testé*	Pas adapté aux données hydrogéologiques.
15	Loi Gamma	*Testé*	Pas adapté aux données hydrogéologiques.
16	Loi de Weibull	*Testé*	Pas adapté aux données hydrogéologiques.
17	Loi du Khi²	*Testé*	Pas adapté aux données hydrogéologiques.
18	Loi Exponentielle	*Testé*	Pas adapté aux données hydrogéologiques.
19	Loi de Rayleigh	*Testé*	Pas adapté aux données hydrogéologiques.
20	Loi de Burr	*Testé*	Pas adapté aux données hydrogéologiques.
21	Loi Double-Normale	*Testé*	Pas adapté aux données hydrogéologiques.
22	Loi Logit-Normale	*Testé*	Pas adapté aux données hydrogéologiques.
23	Loi Demi-Cercle	*Testé*	Pas adapté aux données hydrogéologiques.
24	Loi triangulaire	*Testé*	Pas adapté aux données hydrogéologiques.

Annexe II : Présentation des principales lois de probabilités utilisées

Loi Uniforme

La loi uniforme ($\mathcal{U}$) est la loi à densité de probabilité la plus simple. La probabilité est équiprobable entre les valeurs minimales et maximales d'une variable.

Elle s'écrit selon l'équation suivante (75):

$$\mathcal{U} = f(x) = \frac{1}{x_{max} - x_{min}} \qquad (75)$$

Avec x la variable, x_{min} la valeur minimale de la variable x et x_{max} sa valeur maximale.

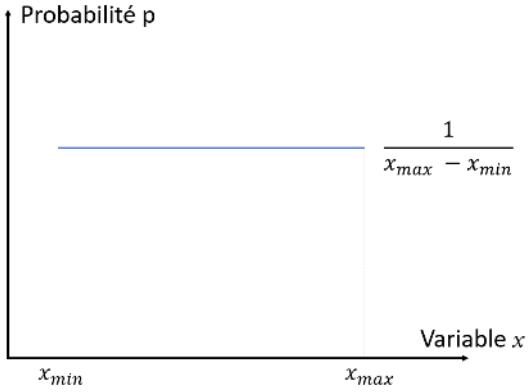


Figure A1-1 - Densité de probabilité de la loi uniforme.

En dehors de l'intervalle [x_{min} ; x_{max}], la valeur de la probabilité vaut zéro. La représentation graphique de cette densité de probabilité est présentée sur la Figure A1-1.

Loi Normale

La loi normale ($\mathcal{N}$), également appelée « loi de Laplace-Gauss », a été étudiée par de nombreux mathématiciens.

La première version de la loi normale fut explicitée par A De Moivre en 1756 (De Moivre 1756). Elle fut ensuite étudiée par P. S. Laplace, qui énonce la première table de la loi de probabilité en 1781 et énonce le théorème central limite en 1810 et Gauss, qui utilise la loi normale à des erreurs d'observations en astronomie en 1809.

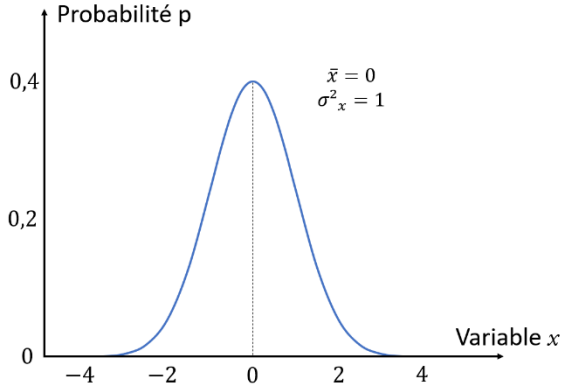


Figure A1-2 - Densité de probabilité de la loi normale.

R. Fisher a introduit également plusieurs concepts statistiques en se basant sur cette loi (Fisher 1922).

Elle est le plus souvent utilisée pour modéliser des phénomènes naturels provenant d'évènements aléatoires. Sa distribution est expliquée sur l'équation (76) et sa représentation graphique sur la Figure A1-2.

$$\mathcal{N}(\bar{x}, {\sigma_x}^2) = f(x|\bar{x}, {\sigma_x}^2) = \frac{1}{\sigma_x\sqrt{2\pi}}\, e^{-\frac{(x-\bar{x})^2}{2{\sigma_x}^2}} \qquad (76)$$

Avec x la variable, $\bar{x}$ la valeur moyenne de la variable x et σ_x son écart-type et ${\sigma_x}^2$ sa variance.

Contrairement à la loi uniforme, la loi normale est une loi paramétrique du second ordre, ayant pour paramètres la moyenne de la variable et sa variance. Lorsque la moyenne est nulle et la variance vaut 1, la loi normale est alors centrée-réduite, car centrée sur la moyenne et réduite de par sa variance.

Loi de Gumbel

La loi de Gumbel ($\mathcal{Gum}$) est une loi à densité de probabilité introduite par E. J. Gumbel en 1957 pour étudier les phénomènes de crues (Gumbel 1957).

Issue de la loi d'extremum généralisée, elle est plus généralement utilisée pour observer des phénomènes extrêmes.

Sa densité de probabilité est représentée sur la Figure A1-3, dont le tracé correspond à une application de l'équation (77).

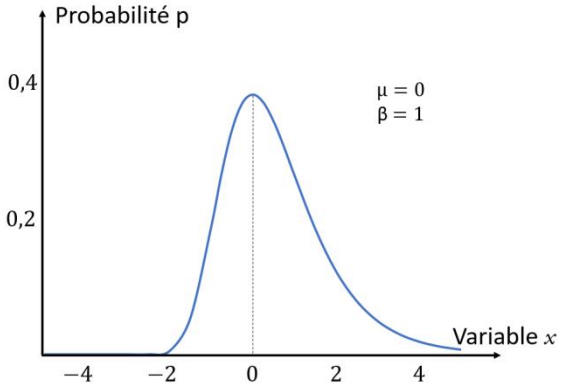


Figure A1-3 - Densité de probabilité de la loi de Gumbel.

$$\mathcal{Gum}(\mu, \beta) = f(x|\mu,\ \beta) = \frac{e^{-\frac{(x-\mu)}{\beta}} e^{-e^{-\frac{(x-\mu)}{\beta}}}}{\beta} \qquad (77)$$

Avec β le paramètre d'échelle, μ le paramètre de position, où $\mu = \bar{x} + \beta\gamma$, γ la constante d'Euler-Mascheroni, x la variable et $\bar{x}$ la valeur moyenne de la variable x.

La constante d'Euler-Mascheroni se calcule de la façon suivante (78) :

$$\gamma = \lim_{n\to\infty}(1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} - \ln(n)) \cong 0{,}577215664\ldots \qquad (78)$$

Le paramètre d'échelle, quant à lui, peut s'optimiser en appliquant la relation (79) ci-après :

$$\beta = \frac{\sigma_x\sqrt{6}}{\pi} \qquad (79)$$

Avec σ_x étant l'écart-type de x.

La particularité de cette loi est qu'elle est asymétrique. Il est possible d'inverser l'asymétrie en rendant négatif le terme $x - \mu$.

Loi de Laplace

La loi double exponentielle, aussi appelée « loi de Laplace » ($\mathcal{Lap}$), est une loi à densité de probabilité à deux paramètres s'apparentant à une association de deux lois exponentielles opposées et accolées entre elles (Laplace 1774). Cette loi est utilisée pour caractériser des phénomènes extrêmes.

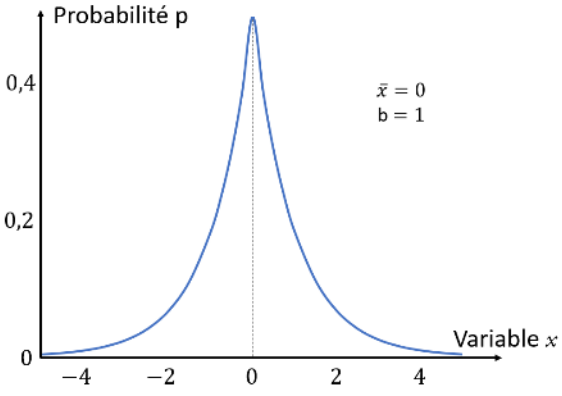


Figure A1-4 - Densité de probabilité de la loi de Laplace.

L'équation de la densité de probabilité se trouve sur l'équation (80) et sa représentation graphique sur la Figure A1-4.

$$\mathcal{Lap}(\bar{x},\, b) = f(x|\bar{x},\, b) = \frac{1}{2b} e^{-\frac{(|x-\bar{x}|)}{b}} \tag{80}$$

où $\quad b = \sqrt{\frac{{\sigma_x}^2}{2}}$ (81)

Avec x la variable, b le paramètre d'échelle, $\bar{x}$ la valeur moyenne de x et ${\sigma_x}^2$ sa variance.

Loi de Cauchy

La loi de Cauchy ($\mathcal{Cau}$), aussi connue sous le nom de « loi de Lorentz », est une loi à densité de probabilité bi-paramétrique, principalement utilisée en spectroscopie, en physique, ainsi qu'en hydrologie lorsqu'elle est appliqué à des évènements extrêmes (Cauchy 1853).

Cette loi fut attribuée à Cauchy en 1853, bien qu'elle ait été étudiée par Poisson quelques années plus tôt en 1824 (Stigler 1974).

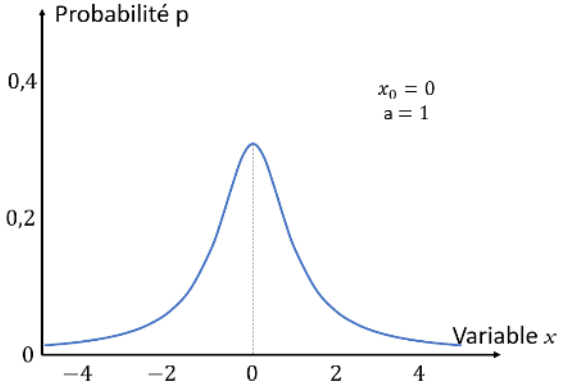


Figure A1-5 - Densité de probabilité de la loi de Cauchy.

Son équation s'écrit de la façon suivante (82) (*cf*. Figure A1-5) :

$$\mathcal{Cau}(x_0,\, a) = f(x|x_0,\, a) = \frac{1}{\pi a\,(1+\left(\frac{x-x_0}{a}\right)^2)} \tag{82}$$

Avec x la variable, x_0 la valeur médiane de la variable x et a le paramètre d'échelle. La particularité de cette loi est qu'elle est divergente pour des valeurs tendant vers l'infini. Il est donc impossible de calculer la moyenne à partir de cette loi.

Loi du Cosinus surélevé

La loi du cosinus surélevé ($\mathcal{Cosr}$) est une loi à densité de probabilité avec deux paramètres (Rinne 2010, Sinha 2012). Elle est utilisée dans le domaine des statistiques et de la théorie des probabilités.

Elle est aussi utilisée en tant que filtre en télécommunication pour limiter les interférences inter-symbole (distorsion de signal).

Elle s'écrit selon l'équation suivante (83) (*cf.* Figure A1-6) :

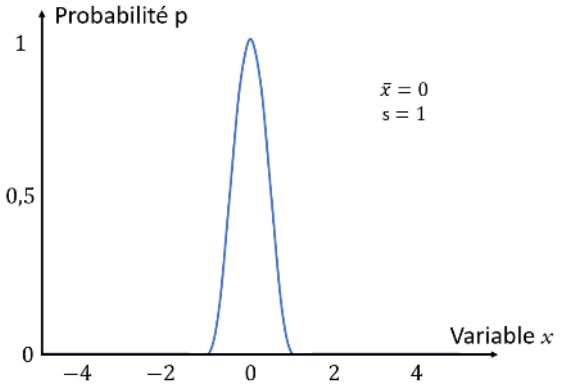


Figure A1-6 - Densité de probabilité de la loi du cosinus surélevé.

$$\mathcal{Cosr}(\bar{x}, s) = f(x|\bar{x},\ s) = \frac{1}{2s}\ (1 + \cos\left(\frac{(x-\bar{x})}{s}\ \pi\right)) \tag{83}$$

Avec x la variable, $\bar{x}$ la valeur moyenne de la variable x et s le paramètre d'échelle. Ce paramètre d'échelle peut être optimisé en utilisant l'écart-type σ_x de la variable x (84).

$$\mathrm{s} = \frac{\sqrt{3}}{\pi}\ .\sigma_x \tag{84}$$

A l'instar de la loi uniforme, cette loi n'est valable que dans un intervalle, qui est celui-ci compris entre $\bar{x} - s$ et $\bar{x} + s$.

Loi Logistique

La loi logistique ($\mathcal{Logist}$) est une loi à densité de probabilité bi-paramétrique (Ritzema 1994, Verhulst 1845).

Elle est utilisée dans le domaine de la modélisation, notamment avec les réseaux de neurones à propagation avant (*i.e.* perceptrons multicouches) ainsi que pour la caractérisation des phénomènes hydrologiques. Elle est également utilisée pour les classements de joueurs et équipes sportives (85) (*cf.* Figure A1-7).

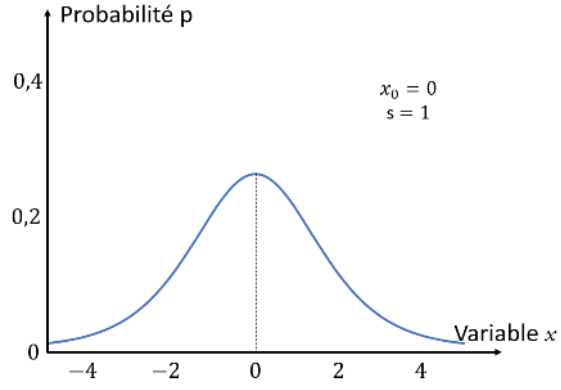


Figure A1-7 - Densité de probabilité de la loi logistique.

$$\mathcal{Logist}(\bar{x},\ s) = f(x|\bar{x},\ s) = \frac{e^{-\frac{(x-\bar{x})}{s}}\ e^{-e^{-\frac{(x-\bar{x})}{s}}}}{s} \tag{85}$$

Avec x la variable, $\bar{x}$ la valeur moyenne de la variable x et s le paramètre d'échelle. De la même manière que pour celui de la loi du cosinus surélevé, le paramètre d'échelle peut être optimisé en utilisant l'écart-type σ_x de la variable x.

Loi de Kumaraswamy

La loi de Kumaraswamy ($\mathcal{K}uma$) est une loi à densité de probabilité paramétrique ayant deux paramètres a et b, avec une variable évoluant sur un intervalle [0 ; 1]. La variable doit donc être normalisée entre 0 et 1 pour qu'elle puisse être appliquée sur un intervalle plus grand. La Figure A1-8 montre un exemple d'application sur un intervalle [-4 ; 4].

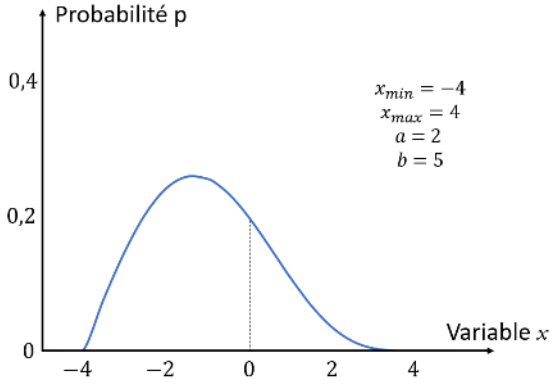


Figure A1-8 - Densité de probabilité de la loi de Kumaraswamy.

P. Kumaraswamy a proposé cette loi, visible sur l'équation (86), afin de prendre en compte les processus ayant des limites minimales et maximales. La méthodologie a ensuite été appliquée sur pluies annuelles (Kumaraswamy 1980).

$$\mathcal{K}uma(a,\ b) = f(x|a\ ,\ b) =\ abx^{a-1}(1-x^a)^{b-1} \qquad (86)$$

Avec x la variable et a et b, les paramètres de la loi.

Loi de Slash

La loi de Slash ($\mathcal{S}\ell a$) est une loi de probabilité issue des lois normale ($\mathcal{N}$) et uniforme ($\mathcal{U}$). Proposée par W. Rogers et J. Tukey en 1972, cette loi permet d'avoir une densité plus plane que la loi normale, en particulier pour des variables éloignées de sa moyenne (*cf*. Figure A1-9) (Rogers et Tukey 1972). La loi de Slash s'effectue selon l'équation suivante (87) :

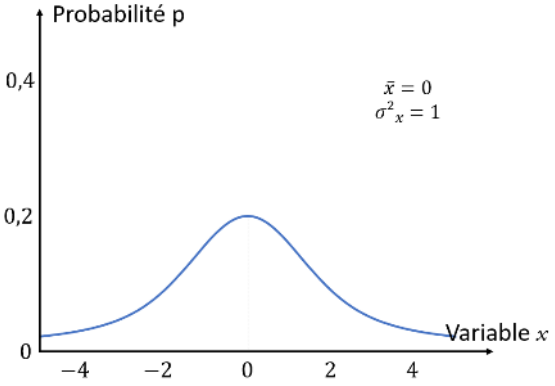


Figure A1-9 - Densité de probabilité de la loi de Slash.

$$\mathcal{S}\ell a(x) =\ f(x|\bar{x},\ {\sigma_x}^2) = \frac{\varphi(0)-\varphi(x)}{x^2} \qquad (87)$$

Avec x la variable, $\bar{x}$ la valeur moyenne de la variable x, ${\sigma_x}^2$ sa variance, $\varphi(0)$ la loi normale centrée-réduite pour $x = 0$ et $\varphi(x)$ la loi normale centrée-réduite en fonction de x. Cette loi est discontinue en zéro. La loi prend alors la valeur suivante (88) :

$$\lim_{x\to 0} f(x) = \frac{1}{2\sqrt{2\pi}} \qquad (88)$$

Loi de Bhattacharjee

La loi de Bhattacharjee ($\mathcal{Bhat}$) est une loi à densité de probabilité à trois paramètres. Cette loi utilise les densités de probabilité cumulatives de la loi normale ($\mathcal{N}$), dont sa moyenne provient d'une loi uniforme ($\mathcal{U}$) (Bhattacharjee, Pandit et Mohan 1963).

La loi de Bhattacharjee est utilisée pour l'étude de tolérance en production industrielle.

Sa représentation graphique est présentée sur la Figure A1-10 et son équation est la suivante (89) :

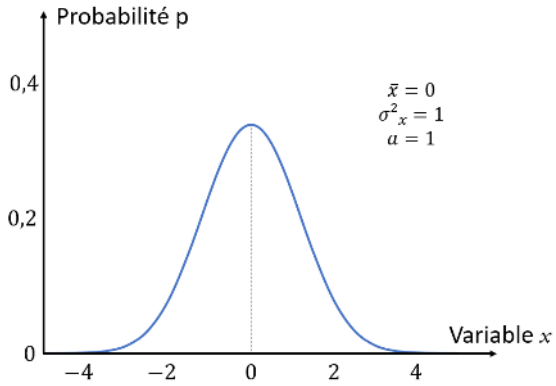


Figure A1-10 - Densité de probabilité de la loi de Bhattacharjee.

$$\mathcal{Bhat}(\bar{x}, \sigma_x, a) = f(x|\bar{x}, {\sigma_x}^2, a) = \frac{1}{2a}\left(\Phi\left(\frac{x-\mu+a}{\sigma}\right) - \Phi\left(\frac{x-\mu-a}{\sigma}\right)\right) \quad (89)$$

Avec x la variable, $\bar{x}$ la valeur moyenne de la variable x, σ_x l'écart-type, a le paramètre de forme et $\Phi(x)$ la loi normale cumulative appliquée à la variable x.

Loi de Huber

La loi de Huber ($\mathcal{Hub}$) est une loi à densité de probabilité paramétrique avec trois paramètres différents.

Elle est proposée par Huber en 1964 en tant que fonction de coût dans une procédure d'estimation, qui peut se retrouver dans la modélisation par réseau de neurones par exemple (90) (P. Huber 1964). La densité de probabilité a été proposée quelques années plus tard, visible sur la Figure A1-11 (91) (P. Huber 1981).

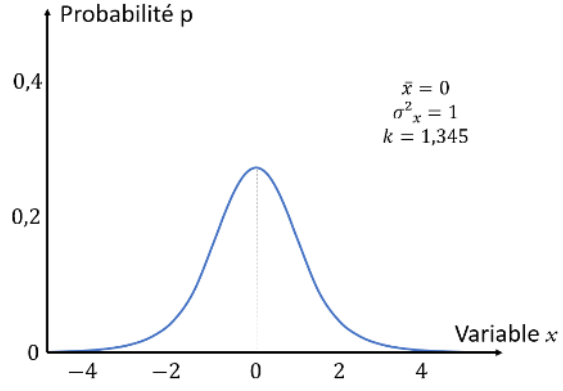


Figure A1-11 - Densité de probabilité de la loi de Huber.

$$\rho_k(x) = \begin{cases} \frac{1}{2}x^2 & ,|x| \le k \\ k|x| - \frac{1}{2}k^2 & ,|x| > k \end{cases} \quad (90)$$

$$\mathcal{Hub}(k) = f(x|\bar{x}, {\sigma_x}^2, k) = \frac{1}{2\sqrt{2\pi}\left(\Phi(k) - \frac{\phi(k)}{k-\frac{1}{2}}\right)} e^{-\rho_k(x)} \quad (91)$$

Avec x la variable, $\bar{x}$ la valeur moyenne de la variable x, ${\sigma_x}^2$ sa variance, k le degré de robustesse, $\rho_k(x)$ la fonction de coût de Huber exprimée en fonction de la variable x, $\phi(k)$ la loi normale pour k et $\Phi(k)$ la loi normale cumulative en fonction de k.

Annexe III : Le *machine learning* au service des probabilités

Plusieurs méthodologies brièvement présentées dans ce manuscrit qui n'ont pas été utilisées pour l'estimation des incertitudes de modélisation utilisent des algorithmes identiques. Ceux-ci sont décrits sur les paragraphes suivants.

Chaînes de Markov

Une chaîne de Markov est un processus stochastique qui conditionne les évènements futurs en fonction des évènements présents uniquement. De tels processus sont ainsi capables de simuler des évènements sans mémoire. Ils sont parfois considérés comme étant une architecture de réseau de neurones (92) :

$$P(X_{n+1} = i | X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = P(X_{n+1} = i | X_n = i_n) \quad (92)$$

Ces processus sont utilisés dans divers domaines tels que les statistiques, la sécurité informatique ou encore la bio-informatique.

Algorithme de Metropolis-Hastings

Les méthodes de Monte-Carlo sont des méthodes algorithmiques qui cherchent à calculer une valeur en utilisant des procédés randomisés. Leur nom provient d'un quartier de la principauté de Monaco, qui regroupe les casinos et introduit la notion de chance – et donc de probabilité. Elles permettent pour la plupart de rééchantillonner les données d'entrées aléatoirement. Ces méthodes peuvent être utilisées pour définir une incertitude sur les paramètres d'un modèle.

Parmi elles se trouve l'algorithme de Metropolis-Hastings, qui permet de générer des valeurs aléatoirement suivant une fonction de densité de probabilité.

Cet algorithme se réalise en deux étapes. La première étape concerne l'initialisation d'un vecteur X. Ensuite, il est nécessaire de choisir une fonction de densité de probabilité f symétrique, c'est-à-dire qu'elle satisfasse la condition $f(x_t|\overline{x_t}) = f(\overline{x_t}|x_t)$, qui soit proportionnelle à la fonction de distribution P souhaitée (Metropolis et Ulam 1949, Metropolis, et al. 1953). Il est commun, surtout en hydro(géo)logie, de choisir une répartition gaussienne (loi normale).

La seconde étape consiste à définir pour chaque temps t un échantillon x' qui provienne de la fonction de distribution f. Une fois l'échantillon choisi, un ratio d'acceptation α est calculé (Metropolis et Ulam 1949, Metropolis, et al. 1953) (93) :

$$\alpha = \frac{f(x')}{f(x_t)} = \frac{P(x')}{P(x_t)} \quad (93)$$

Un nombre aléatoire u compris entre 0 et 1 est généré. Ce nombre u est comparé au ratio d'acceptation α :

- Si $u \leq \alpha, x_{t+1} = x'$,
- Sinon si $u > \alpha, x_{t+1} = x_t$

L'avantage des méthodes de Monte-Carlo est qu'elles sont des algorithmes robustes capables de résoudre des problèmes complexes. Cependant, pour des modèles complexes ou fortement non-linéaires, les calculs réalisés sont eux-aussi complexes et donc coûteux en temps (D. Shrestha 2009). De plus, les variables x dites incertaines sont difficiles à échantillonner à partir leur distribution de densités de probabilité sauf si celle-ci est bien approximée par une fonction de distribution multi-normale (Kuczera et Parent 1998).

Il est fréquent de coupler des chaînes de Markov avec des méthodes de Monte-Carlo. En effet, la méthode de Metropolis-Hastings est, de fait, une chaîne de Markov compte-tenu que l'échantillon suivant dépend uniquement de l'échantillon présent.

Kuzcera et Parent ont utilisé les méthodes **MCMC** (*Monte-Carlo Markov Chain*), en utilisant l'algorithme de Metropolis-Hastings (Kuczera et Parent 1998). Cette méthode est utilisée pour quantifier les incertitudes liées à l'estimation des paramètres W en regardant la dispersion des différents échantillons obtenus. Il a été observé que cette technique génère des incertitudes suffisamment importantes pour inclure les données observées. Elle est en revanche moins performante, avec une exactitude inversement proportionnelle à la racine carrée du nombre de modèles utilisés (Kuczera et Parent 1998).

Il existe également d'autres méthodes d'échantillonnage similaires aux méthodes de Monte-Carlo, telle que la méthode d'échantillonnage par hypercube latin, notées **LHS** (pour *Latin Hypercube Sampling*) (McKay, Beckman et Conover 2000).

Annexe IV : Outils informatiques utilisés

Toutes ces techniques requièrent l'utilisation d'outils, tels que des logiciels, ou encore des bibliothèques logicielles, souvent nommées librairies, qui sont des collections de portions de code, prêtes à être utilisées dans un programme.

RNF Pro™ (Java)

RNF Pro™ est un progiciel de modélisation développé par la société GEONOSIS en Java. Cet outil dispose d'une interface graphique permettant de concevoir une architecture et de paramétrer un réseau de neurones. Il dispose également des techniques corrélatoires et des méthodes de régularisation. Différents algorithmes d'apprentissage sont également inclus, tel que l'algorithme de Levenberg-Marquardt. Cet outil est largement utilisé au sein de l'IMT Mines Alès pour une utilisation à but pédagogique et de recherche scientifique.

Tensorflow (Python)

Certains modèles ont été créés à partir de la bibliothèque logicielle TensorFlow (https://www.tensorflow.org) en Python. Cette bibliothèque *open-source* a été développée par Google en novembre 2015. Elle est également disponible dans d'autres langages de programmation tels que le C, le Go!, le Java et le JavaScript. Cependant, la librairie en langage Python est la plus complète et a donc été préférée.

D'autres bibliothèques sont également utilisées, notamment pour le traitement des données en amont et la réalisation de certains calculs tels que Numpy (www.numpy.org), Scipy (https://www.scipy.org) et Pandas (https://pandas.pydata.org). Les graphiques sont obtenus grâce à la bibliothèque Matplotlib (https://matplotlib.org).

La bibliothèque TensorFlow a été préférée à d'autres (telles que Theano, Torch (et son dérivé en Python PyTorch) ou encore Caffe) car elle fait partie, avec Theano, des plus simples à implémenter et des plus performantes (Kovalev, Kalinovsky et Kovalev 2016).

De plus, TensorFlow dispose de la plus grande communauté, ce qui est utile en cas de blocage lors des développements. Elle dispose également de nombreuses règles d'apprentissage, bien qu'elle ne dispose pas de règles utilisant la dérivée seconde telles que l'algorithme de Levenberg-Marquardt ou l'algorithme de Broyden-Fletcher-Goldfarb-Shanno (BFGS). Parmi elles, se retrouvent les règles « classiques » telles que la règle du gradient et de la rétropropagation. D'autres règles moins connues sont également présentes, dont notamment la règle d'ADAM.

mixTools (R)

MixTools est un package R, sous licence publique GPL2 et GPL3 (*General Public Licence 2 & 3*), créé par Benaglia et *al*. en 2009 (Benaglia, et al. 2009). Cet outil permet de retrouver les paramètres d'une fonction paramétrée telle que les lois à densité de probabilité mixte, en utilisant l'algorithme espérance-maximisation.